Radar based high resolution ensemble precipitation analyses over the French Alps

Matthieu Vernay[1], Matthieu Lafaysse[1], and Clotilde Augros[2]

[1]Univ. Grenoble Alpes, Université de Toulouse, Météo-France, CNRS, CNRM, Centre d'Études de la Neige, Grenoble, France

[2]CNRM, Université de Toulouse, Météo-France, CNRS, Toulouse, France

Correspondence: matthieu.vernay@meteo.fr

Abstract.

Reliable estimation of precipitation fields at high resolution is a key issue for snow cover modelling in mountainous areas, where the density of precipitation networks is far too low to capture ~~their complex variability~~ the complex variability of these fields with topography. Adequate quantification of the remaining uncertainty in precipitation estimates is also necessary for further assimilation of complementary snow observations in snow models. Radar observations provide spatialised estimates of precipitation with high spatial and temporal resolution, and are often combined with rain gauge observations to improve the accuracy of the estimate. However, radar measurements suffer from significant shortcomings in mountainous areas (in particular, unrealistic spatial patterns due to ground clutter leading to local systematic biases). Precipitation fields simulated by high-resolution numerical weather prediction (NWP) models provide an alternative estimate, but suffer from widespread systematic biases and positioning errors. Even though these uncertainties can be partially described by ensemble NWP systems and systematic errors can be reduced by statistical post-processing, NWP precipitation estimates are still not reliable enough for the requirements of high resolution snow cover modelling.

In this study, better precipitation estimates are obtained through a specific analysis based on a combination of all these available products. First, a pre-processing step is proposed to mitigate the main deficiencies of precipitation estimates by radar and gauges~~precipitation estimation products~~, focusing on reducing unrealistic spatial patterns. This method also provides a spatialised estimate of the associated error in mountainous areas, based on a climatological analysis of both radar and NWP-estimated precipitation. Three ensemble daily precipitation analysis methods are then proposed, first using only the modified precipitation estimates and associated errors, then combining them with ensemble NWP simulations based on the Particle Filter and Ensemble Kalman Filter data assimilation algorithms. The performance of the different precipitation analysis methods is evaluated at a local scale using independent ski resort precipitation observations. The evaluation of the pre-processing step shows its ability to remove the main spatial artefacts coming from the radar measurements and to improve the precipitation estimates at the local scale. The local scale evaluations of the ensemble analyses do not demonstrate an additional benefit of ensemble NWP forecasts, but their contrasted spatial patterns are challenging to evaluate with the available data.

1 Introduction

Monitoring snow cover in mountainous areas is essential for a wide range of ~~human activities~~ practical applications and scientific applications (IPCC, 2022). The complex topography of these areas leads to a very high spatial variability of meteorological and snow conditions (e.g. Clark et al., 2011), which is not fully sampled by any existing in situ observing network, especially at high altitudes (Thornton et al., 2022). Operational applications such as water resource management or avalanche forecasting, which require detailed monitoring of meteorological conditions and snow cover over large mountainous areas, suffer from this lack of observational information. The use of numerical snow models provides a more continuous spatial and temporal coverage than observations. The complexity of such models varies widely depending on their application (Krinner et al., 2018). However, all seasonal snow modelling systems are affected by the strong dependence of the snow cover state at any time on its past evolution since the first snowfall. This long-term dependence means that any simulation error at any time can affect all subsequent simulations, resulting in an accumulation of errors throughout the winter.

Satellite observations of some snow properties provide a great opportunity to identify and reduce these errors (Awasthi and Varade, 2021; Largeron et al., 2020). Methods based on ensemble data assimilation algorithms have been developed to process these observations ~~,~~ (Magnusson et al., 2017; Cluzet et al., 2021). According to Cluzet et al. (2022) and Deschamps-Berger et al. (2022), these methods primarily use snow observations to compensate for errors in the precipitation forcing of the snow cover model. Quantifying the uncertainties in the precipitation fields is therefore essential to fully benefit from the assimilation of snow observations. However, all the papers cited above rely only on stochastic perturbations of the precipitation dataset, obtained with homogeneous and rather arbitrary error estimates in the absence of more advanced quantification of precipitation uncertainty.

Existing snow cover modelling systems mostly use precipitation inputs provided by numerical weather prediction (NWP) output, surface observations or a combination of these two sources of information (Morin et al., 2020). Surface observations provide reliable local estimates of precipitation~~. However~~, but are affected by systematic undercatch in case of solid precipitation or in windy conditions (Rasmussen et al., 2012; Kochendorfer et al., 2020). In addition, the under-sampling of higher elevations (Thornton et al., 2022) means that there is a lack of information on the spatial distribution of elevation-dependent variables, such as precipitation (Mott et al., 2023). On the contrary, high-resolution NWP models produce spatialised estimates of precipitation fields at different spatio-temporal resolutions. Lundquist et al. (2019) argue that such models can simulate annual precipitation accumulation in mountainous areas better than estimates from gauge or radar-based observations. However, they suffer from biases and positioning errors in individual events and in seasonal accumulations. These errors are problematic for snow cover modelling ~~(Vionnet et al., 2016, 2019; Haddjeri et al., 2023)~~(Vionnet et al., 2016, 2019; Haddjeri et al., 2023) . A combination of surface observations and NWP output is used in some operational snow modelling systems (SAFRAN, Durand et al., 1993; Lespinas et al., 2015) to provide precipitation estimates at scales of a few hundred square kilometres. However, scarce observations may be insufficient to constrain high-resolution precipitation analyses in mountainous areas (Soci et al.,

2016), even when specifically designed for this purpose (Schirmer and Jamieson, 2015).

Radar measurements provide high resolution spatial estimates of precipitation. However, they are subject to uncertainties in mountainous regions, mainly due to the interaction between the radar beam and the terrain ~~(ground clutter and partial masks, Germann et al.~~ (ground clutter and partial masks) : Germann et al. (2022); Foresti et al. (2018); Faure et al. (2019); Yu et al. (2018); Foehn et al. (2018); G . Methods have been developed to correct radar-based precipitation fields (Vogl et al., 2012) and to assess the associated uncertainty (Kirstetter et al., 2010; Villarini et al., 2014; Kirstetter et al., 2015). ~~However the low quality of precipitation estimates based on radar measurements in complex terrain, even when combined with~~ The combination of radar measurements and in-situ observations of precipitation (Sideris et al., 2014; Sivasubramaniam et al., 2019; Champeaux et al., 2009) ~~, currently prevents their direct use to successfully force a snowpack model (Haddjeri et al., 2023)~~does not fully mitigate these uncertainties, as in-situ measurements generally do not sample areas where these uncertainties are the most significant. More sophisticated products combining NWP outputs, surface observations and precipitation estimates from radar measurements (CaPA, Fortin et al., 2015, 2018; Khedhaouiria et al., 2022) suffer from significant biases in winter (Lespinas et al., 2015). The potential of using radar observations for detailed snowpack modelling has only been investigated on a relatively large scale over the French Alps (Birman et al., 2017). Over the Austrian Alps, the SNOWGRID system (Olefs et al., 2013) uses radar observations via the INCA now-casting system (Haiden et al., 2011) to force a simple snowpack model designed for hydrological applications. The low quality of precipitation estimates based on radar measurements in complex terrain currently prevents their direct use to successfully force a detailed snowpack model at high spatial resolution (Haddjeri et al., 2023). In particular, the spatial structure of the error associated with such products in mountainous areas and its overall magnitude have not been investigated in depth.

As noted above, Cluzet et al. (2022) and Deschamps-Berger et al. (2022) showed that any precipitation analysis designed for a snow cover modelling system with assimilation of snow observations must include an estimate of the precipitation analysis errors. ~~This can be done through ensemble analysis and benefit from ensemble versions of NWP models developed specifically for this purpose~~In particular, snow data assimilation is effective when uncertainties in the precipitation forcing are correctly identified and accounted for, which can be achieved through accurate and reliable ensemble precipitation analysis (Cluzet et al., 2021). In a context of high resolution modelling, radar-based precipitation analyses present the advantage to provide already spatialised precipitation estimates. A variety of methods have been developed to produce ensembles of estimated precipitation from radar and gauge measurements with varying degrees of complexity (Clark and Slater, 2005; Germann et al., 2009; Mandapaka and Germann, 2010; Dai et al., 2014; Kirstetter et al., 2015; Frei and Isotta, 2019)~~, but they do not use NWP outputs~~. However, they were not designed to meet the requirements of snow data assimilation in a high-resolution snowpack modelling system. Ensemble methods combining radar-based precipitation estimates and NWP output are more common in the now-casting context (Foresti et al., 2012; Nerini et al., 2019; Atencia et al., 2020a; Sideris et al., 2020). In this case, NWP output are used to propagate the precipitation estimation in time but not to mitigate the inherent flaws of the radar product itself.

~~The~~ To address this gap, the aim of this study is to explore the combination of different products based on radar, gauges and NWP data (section 2) to produce ensemble precipitation analyses (section 3.2) over mountainous areas. ~~This preliminary study focuses specifically on the French Alps, but the proposed methodology can be applied to any mountainous area with at least one year of radar-based and NWP model precipitation estimates. These analyses are expected to appropriately quantify precipitation uncertainties in order to guarantee the efficiency of assimilation of snow observations in snow cover simulations.~~ An evaluation of the quality of several available precipitation estimation products is first performed (section 4.1). Then, a pre-processing step of the best product is proposed to remove spatial artefacts. Finally, this study develops three different methods for ensemble analysis of daily precipitation to investigate the benefits of combining observational precipitation estimates and NWP outputs. These methods are then applied to produce ensemble analyses of daily precipitation at a $1\mathrm{km}$ resolution (see section 3.2). Section 4.2 evaluates their performance and section 5.3 discusses their respective advantages and disadvantages. This study focuses specifically on the French Alps, but the proposed methodology can be applied to any mountainous area with at least one year of radar-based and NWP model precipitation estimates. These analyses are expected to appropriately quantify precipitation uncertainties in order to guarantee the efficiency of assimilation of snow observations in snow cover simulations.

2 Precipitation dataset

This study focuses on the French Alps region (Figure 1) for the period from 1 August 2021 to 1 August 2022. ~~Evaluation data is only available for the period from 1 December 2021 to 30 April 2022.~~

Three different observational precipitation estimates and one from a high-resolution NWP model were considered. To avoid confusion between snow variables, we follow the international classification for seasonal snow on the ground (Fierz et al., 2009) and express all precipitation in $\mathrm{kg\,m^{-2}}$. 24 h accumulation of all precipitation products up to 08:00 CEWT are analysed in this study.

2.1 Radar product (PANTHERE)

PANTHERE (Tabary, 2007; Figueras i Ventura and Tabary, 2013) is an operational Quantitative Precipitation Estimation (QPE) product based on the combination of most of the French metropolitan radar data. It has a ~~horizontal resolution of 1x1 and~~ 1 km horizontal resolutionand a temporal resolution of 5 minutes, aggregated in this study into 24 h precipitation accumulations at 8:00 ~~CET~~ CEWT each day. Each radar measures the reflectivity and dual polarisation variables with a resolution of $240\ \mathrm{m} \times 0.5$, up to a maximum range of 255 km, and at several elevation angles (see grey radar beams in Figure 2a). After a correction step to account for measurement problems, the reflectivity Z is converted to an instantaneous precipitation rate R ($\mathrm{kg\,m^{-2}\,h^{-1}}$) using a Z-R relationship (Marshall and Palmer, 1948) that is constant in space and time. This precipitation rate is then corrected by applying a Vertical Profile of Reflectivity (VPR) correction factor. The purpose of this step is to correct for the expected variation in reflectivity with altitude due to different hydrometeor types, and in particular to correct for the increase in reflectivity in the melt region, known as the "bright band". Although this method is effective in most cases, it has

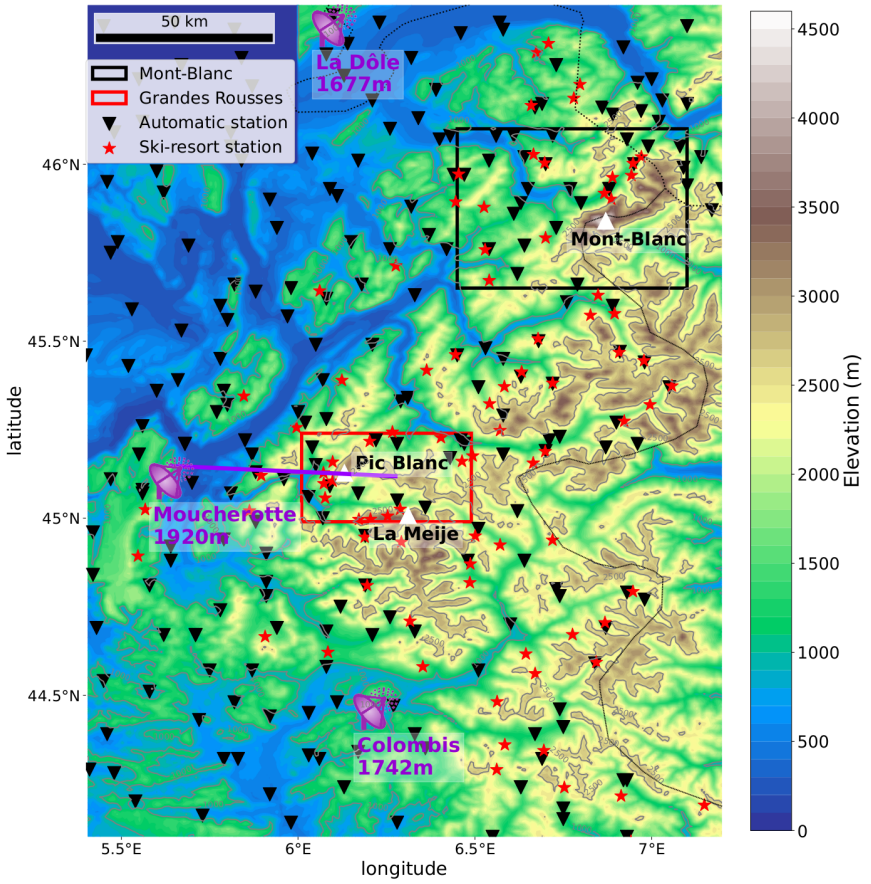

Figure 1. Relief at 250 m resolution of the French Alps domain used in this study, showing the three radars used in radar-based precipitation estimation products, the automatic observation stations and the reference ski-resort observation stations used for verification. The Grandes-Rousses and Mont-Blanc areas, on which parts of this study are focused, are framed and the cross section of Figure 2 is marked with a grey line.

some limitations: it assumes a constant precipitation rate below the bright band, so processes such as evaporation or low-level enhancement of precipitation must be considered (Le Bastard et al., 2019).

After this processing of the volumetric radar data, the precipitation rate at ground level is estimated from a combination of colocated precipitation rates ~~measured~~ estimated at all heights, weighted by their quality index, which depends on, among other factors, the height of the measurement. The final 5-minute QPE at ground level is then obtained by accumulating the precipitation rates over time. In this study, precipitation accumulations over 24 h are considered. Despite all these steps to calculate the QPE, its quality in space is variable. In general, the uncertainty of the estimate increases with the height of the radar beam above the ground, implying that the most valuable radar information comes from the lowest radar heights and

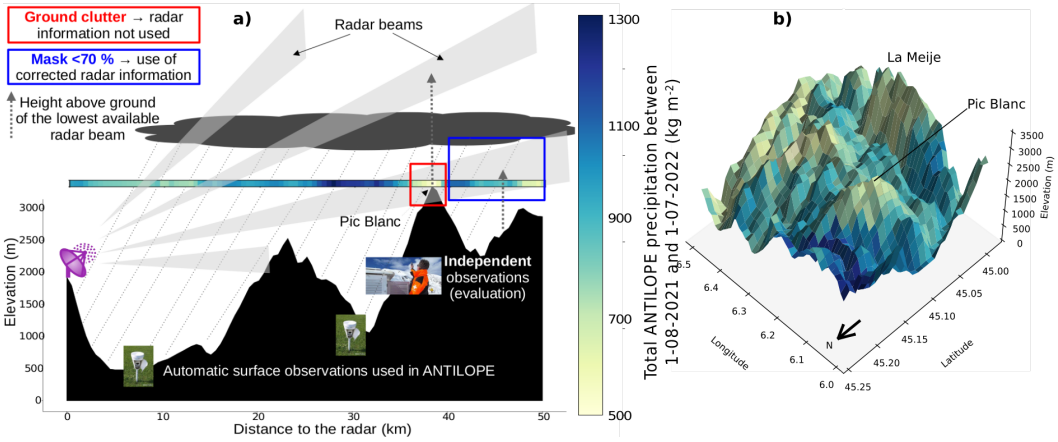

Figure 2. a) Illustration of radar measurement issues from the Moucherotte radar towards the Pic Blanc in the Grandes Rousses massif (grey line in Figure 1). Radar beams from four elevation angles are shown to illustrate the vertical sampling of the atmosphere and its limitations in complex terrain. The effect of ground clutter management in the PANTHERE algorithm and the link with the underestimation of precipitation over mountain ridges is illustrated over the Pic Blanc (framed in red). The horizontal coloured band in panel a) is the ANTILOPE precipitation accumulation on the ground between 1 August 2021 and 1 July 2022 along the transect between the Moucherotte radar, which is the primary contributor to the PANTHERE precipitation estimation in this area, and the Pic Blanc. The position of two automatic gauges and a ski-resort observation site are also shown to illustrate the altitudes of typical in-situ observations used in the ANTILOPE product and of the reference observations used in this study, as well as the lack of in-situ observations at altitudes above 2000 m. b) ANTILOPE precipitation accumulation between 1 August 2021 and 1 July 2022 over the relief of the Grandes Rousses domain (framed in red in Figure 1).

that the quality of the precipitation estimate tends to decrease far away from the radar. In particular, Figure 2a illustrates that stratiform precipitation systems are not detected by radar beams above the top of precipitation clouds.

In mountainous areas, the radar beam will also often intercept the ground. Radar elevations affected by ground clutter are rejected by the algorithm (red frame on Figure 2) and the lowest information comes from the next beam above, which affects the quality of the precipitation estimate. In addition, the conical shape of the radar beam means that the beam width increases with distance from the radar (up to about 1000 m at a distance of 50 km). It can also be partially affected by the presence of a mountain (blue frame in Figure 2). In this case, the affected beam information is considered unusable behind the mountain if the mask blocks more than 70% of the total energy, otherwise the signal is corrected for attenuation.

Faure et al. (2017) evaluated the quality of PANTHERE precipitation estimation over the French Alps, and showed an increasing underestimation of precipitation towards the east due to radar beam blockage and increasing distance from radars. They also identified specific areas affected by significant underestimations related to ground clutter handling (as shown in Figure 2) and concluded that the clutter correction is ineffective in a high mountain context. Similarly, Faure et al. (2019) studied the vertical distribution of PANTHERE precipitation estimates and highlighted a general overestimation of the radar QPE at the bottom of the valleys and an underestimation at the highest altitudes.

2.2 ~~Radar and gauges fusion~~ Radar-gauge combination product: ANTILOPE

ANTILOPE (Champeaux et al., 2009) is an operational composite analysis combining radar precipitation estimates from PANTHERE and precipitation observations from automatic gauges (see Figure 1). It is available at a ~~resolution of 1x1~~ 1 km resolution and 24 h accumulations at 8:00 ~~CET~~ CEWT have been used in this study. The fusion of these two sources of information is based on a scale separation between small scale convective and large scale stratiform precipitation. Precipitation associated with radar-detected convective cells is corrected by a spatialisation of the local differences between radar and gauge precipitation estimates using inverse distance interpolation. Large scale precipitation is estimated by a spatialisation of the gauge values by an ordinary kriging with external drift method, using either a correlogram computed from radar images or an exponential variogram model if no radar image is available. In contrast to PANTHERE, the quality of ANTILOPE precipitation estimation in mountainous areas is not well documented. The strong dependence of the ANTILOPE product on the radar-based precipitation estimate suggests that the main drawbacks of radar measurements described in section 2.1 also affect the ANTILOPE quality. However, the use of rain-gauge observations may reduce the ~~errors' magnitude~~ magnitude of the errors, with possible exceptions in case of assimilation of non-heated gauges not detected by the control steps.

2.3 Gauge kriging

In order to document the added value of the radar information used in ANTILOPE, a precipitation estimation based on the same kriging method (using an exponential variogram) of the same set of gauges (automatic stations, Figure 1) as those used in the ANTILOPE algorithm, but without any radar information was set up and evaluated for this study. This represents 512 stations across the study area with elevations ranging from 0 m a.s.l to 2730 m a.s.l and a mean elevation of 809 m a.s.l. The resulting precipitation fields have the same 1 km resolution as the ANTILOPE and PANTHERE products and are used similarly. The temporal resolution used in this study is 24 h.

2.4 ~~Evaluation data~~

~~The evaluation dataset comes from the ski-resort observation network of the ski resorts of the French Alps (Fig. 1). These observations are not used by the ANTILOPE product (and therefore not used in gauge kriging). This network provides a set of daily man-made meteorological and snowpack observations specifically designed for avalanche forecasting during the winter season (generally from mid-December to mid-April, depending on the opening and closing dates of the resorts). For this study, observations of 24 precipitation accumulation in a bucket weighted at 8:00 (CET) are used as the reference for all evaluations. Consequently, all precipitation products are evaluated in terms of 24 water equivalent accumulations starting at 8:00 CET. A human estimate of the highest altitude reached by the rain-snow limit during the same period is also available and used in this work. When focusing on solid precipitation, only days where the rain-snow limit altitude was below the station altitude were considered.~~

~~Ski-resort observations have limitations for the evaluation of gridded precipitation estimates : Human measurement time may vary slightly between stations and days the mountainous environment is known to affect the measurements from the gauges,~~

~~with possible snow accumulation in the gauges or under-catchment~~ As already mentioned, gauges measurements are known to be affected by significant undercatch in case of solid precipitation and in windy conditions ~~(these measurement errors are sometimes detected and corrected or removed in the monitoring process of the observations and also affect the automatic gauges used by the ANTILOPE product) local gauge observations do not have the same representativeness as a gridded estimate over a pixel of about 1×1~~ (Rasmussen et al., 2012; Kochendorfer et al., 2020).

2.4 Numerical weather prediction model

In this study, data from the deterministic NWP model AROME (Seity et al., 2011; Brousseau et al., 2016) and its ensemble version have been used in a complementary way.

The French operational high-resolution NWP model AROME provides hourly precipitation forecasts with a ~~resolution of 1×~~1 km resolution. Daily precipitation accumulations (liquid and solid) are directly derived from the 24 h forecast at validity time ~~6~~7:00 UTC (8:00 CEWT). Yearly AROME precipitation accumulation is then calculated as the accumulation of daily precipitation between 1 August 2021 and 1 August 2022.

A 16-member ensemble version of the AROME NWP model is also operational at a ~~resolution of 2.5×~~2.5 km resolution (Bouttier et al., 2016). Its precipitation forecasts are statistically post-processed with the method developed by Taillardat et al. (2019), based on quantile regression forests, to provide an unbiased and well-distributed ensemble of hourly precipitation forecasts, hereafter referred to as PEAROME. The calibration uses ANTILOPE precipitation estimates as reference and is performed at two AROME EPS initialisation times (9:00 UTC and 21:00 UTC) with lead times up to 45 h. (Taillardat and Mestre, 2020). Physically realistic precipitation patterns are then reconstructed using the ensemble copula coupling method of Schefzik et al. (2013). However, the training method has been evaluated over ~~metropolitan~~ mainland France (Taillardat and Mestre, 2020) and its performance over the French Alps is not well known. The daily precipitation used in this study is the precipitation accumulation between lead times ~~9~~10 h and ~~33~~34 h from the 21:00 UTC initialisation time. The raw precipitation fields are downscaled to the ANTILOPE 1~~×1~~ km grid by a bi-linear interpolation so that model and observation can be compared at the same spatial and temporal resolution.

2.5 Evaluation data

The evaluation dataset comes from the observation network of the ski resorts of the French Alps (Figure 1). These observations are not used by the ANTILOPE product (and therefore not used in gauge kriging). This network provides a set of daily man-made meteorological and snowpack observations specifically designed for avalanche forecasting during the winter season (generally from mid-December to mid-April, depending on the opening and closing dates of the resorts). For this study, observations of 24 h precipitation accumulation in a bucket weighted at 8:00 CET are used as the reference for all evaluations. Consequently, all precipitation products are evaluated in terms of 24 h water equivalent accumulations starting at 8:00 CET (7:00 UTC in winter, 6:00 UTC in summer). Evaluation data are only available for the period from 1 December 2021 to 30 April 2022. This means that there is a 1 h gap between the precipitation accumulation of the different products (up to 7:00 UTC) and the reference observations (up to 8:00 CET) for the month of April after the time change. However, this gap is expected to

have very limited impact on the results, as fewer reference observations are available in April (due to the closure of ski resorts) and only two significant precipitation events occur during this period. A human estimate of the highest altitude reached by the rain-snow limit during the same period is also available and used in this work. When focusing on solid precipitation, only days where the rain-snow limit altitude was below the station altitude were considered.

Ski-resort observations have limitations for the evaluation of gridded precipitation estimates :

- Human measurement time may vary slightly between stations and days
- the mountainous environment is known to affect the measurements from the gauges, with possible snow accumulation in the gauges or undercatch in windy conditions (these measurement errors are sometimes detected and corrected or removed in the monitoring process of the observations and also affect the automatic gauges used by the ANTILOPE product)
- local gauge observations do not have the same representativeness as a gridded estimate over a pixel of about 1 km^2

Nevertheless, these are the only available independent data for the evaluation of the various precipitation products.

3 Method

3.1 ANTILOPE observation error

A precipitation analysis requires the specification of the error of the various products involved. As mentioned earlier, radar-based observations are known to suffer from important shortcoming in mountainous areas (Germann et al., 2022). In particular, unrealistic spatial patterns of ANTILOPE yearly precipitation accumulation estimates can be visually identified over some mountain ridges. Figure 2b shows that ANTILOPE yearly precipitation accumulations over the ridges of Pic-Blanc and La Meije (around 600 $\mathrm{kg\,m^{-2}}$) are ~~only about half those over the lower elevations areas in between~~ unrealistically low. In comparison, the same accumulation at lower altitudes between the two peaks benefiting from direct gauge measurements is more than twice as large (up to 1300 $\mathrm{kg\,m^{-2}}$). Figure 3a shows the same pattern over the Mont-Blanc (circle in red) where yearly precipitation accumulation (less than 200 $\mathrm{kg\,m^{-2}}$ per year, five times less than in the Chamonix Valley circle in green) is unrealistically low. These patterns are also visible in the daily precipitation fields (Figure 5) and are probably due to the presence of ground clutter (see section 2.1).

The most common approach to overcome the limitations of radar-based precipitation estimates in mountainous areas is to combine them with other sources of information or to apply a calibration step (Germann et al., 2022). However, these methods suffer from the lack of observations at high elevations (Figure 1). McRoberts and Nielsen-Gammon (2017) proposed a method to detect and correct pixels affected by partial beam blockage based on an analysis of the radar precipitation estimation climatology, but it does not specifically focus on ground clutter, which is prevalent in mountainous areas. Haiden et al. (2011) uses radar precipitation estimates where ground clutters have been statistically filtered, and other perturbations are dealt with in a preprocessing step bases on a climatological scaling of the radar data. Here, we propose a method to mitigate the unrealistic

spatial patterns in ANTILOPE precipitation fields. A first method to estimate the climatological errors of ANTILOPE is described in section 3.1.1 and an evaluation of the method is provided in section 4.2. This estimated climatological error is then used to remove the unrealistic spatial structures in ANTILOPE 24 h precipitation fields (Section 3.1.4).

3.1.1 ANTILOPE climatological error estimation

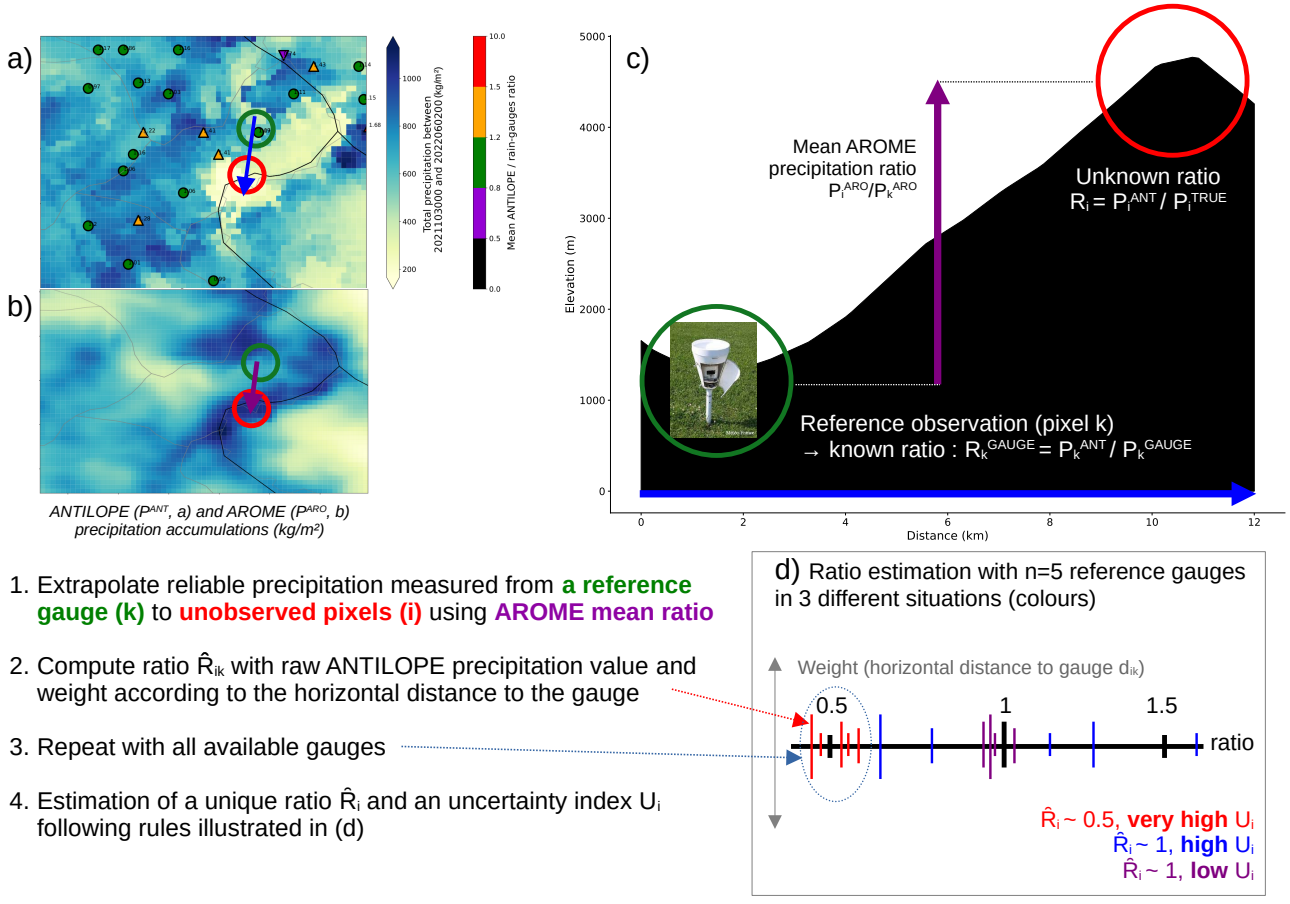

Figure 3. Illustration of the climatological ANTILOPE error estimation over the Mont-Blanc (circled in red) based on (a) the Yearly ANTILOPE precipitation accumulation (~~RR~~^{ANT}P^{ANT}) from 30 October 2021 to 2 June 2022 and the mean ANTILOPE / gauge ratio (R^{GAUGE}) observed at automatic stations and (b) the AROME precipitation accumulation. The marker associated with each evaluation station on panel a is a circle if the ratio between ANTILOPE and the reference gauge ~~pixel~~ precipitation estimates is not considered significant (between 0.8 and ~~the Mont-Blanc pixel~~ 1.2) and a triangle facing up (~~RR~~$^{ARO}_{i}$/~~RR~~$^{ARO}_{k}$, ~~purple arrow~~resp. down) if this ratio is significantly higher (resp. lower) than 1.

We propose a method based on a comparison between ANTILOPE, the automatic gauge measurements (whether used in ANTILOPE or not) and the AROME yearly precipitation accumulations to provide :

- an estimate $\hat{R}$ of the ~~real~~ true climatological ANTILOPE ratio (R^{TRUE}) defined as the ratio between the ANTILOPE precipitation accumulation (~~$\mathrm{RR}^{\mathrm{ANT}}$~~ $\mathrm{P}^{\mathrm{ANT}}$) and the ~~real~~ true precipitation accumulation (~~$\mathrm{RR}^{\mathrm{TRUE}}$~~ $\mathrm{P}^{\mathrm{TRUE}}$, unknown) over the same period :

$$R^{\mathrm{TRUE}} = \frac{\mathrm{RR}^{\mathrm{ANT}}}{\mathrm{RR}^{\mathrm{TRUE}}} \frac{\mathrm{P}^{\mathrm{ANT}}}{\mathrm{P}^{\mathrm{TRUE}}} \tag{1}$$

  Where a gauge is available, we assume ~~$\mathrm{RR}^{\mathrm{GAUGE}} = \mathrm{RR}^{\mathrm{TRUE}}$~~ $\mathrm{P}^{\mathrm{GAUGE}} = \mathrm{P}^{\mathrm{TRUE}}$ so that:

$$\hat{R} = R^{\mathrm{GAUGE}} = \frac{\mathrm{RR}^{\mathrm{ANT}}}{\mathrm{RR}^{\mathrm{GAUGE}}} \frac{\mathrm{P}^{\mathrm{ANT}}}{\mathrm{P}^{\mathrm{GAUGE}}} \tag{2}$$

  If $\hat{R}$ can be spatialized, it can be used as a correction factor for daily precipitation fields in order to remove systematic errors

- a spatialized ANTILOPE uncertainty index, which measures the relative confidence between the different pixels of the domain

The general idea of the method is to extrapolate the ratios measured at ~~gauges locations~~ location of gauges to any location without reference observation considering precipitation accumulations simulated by the AROME NWP model, which are considered to better represent mean vertical gradients. The estimation method is illustrated on Figure 3 over the Mont-Blanc pixel (hereafter referred to with the subscript i and circled in red on Figure 3a,b,c).

3.1.2 Ratio estimation

The estimation of the ratio $\hat{R}_i$ for any pixel i (e.g. Mont-Blanc point in Figure 3) is based on :

- the ratio ~~between the~~ bewteen ANTILOPE yearly precipitation accumulation ~~at pixel i $\mathrm{RR}_i^{\mathrm{ANT}}$ and the ANTILOPE accumulation $\mathrm{RR}_k^{\mathrm{ANT}}$~~ and gauge measurements R_k^{GAUGE} at nearby gauges k (circled in green in Figure 3a,b,c)~~, far from~~ ;
- the ratio between the ANTILOPE yearly precipitation accumulation at pixel i ~~of~~ $\mathrm{P}_i^{\mathrm{ANT}}$ and the ANTILOPE accumulation $\mathrm{P}_k^{\mathrm{ANT}}$, with a distance d_{ik} separating pixels i and k;
- the ratio between the AROME yearly precipitation accumulations ~~$\mathrm{RR}_i^{\mathrm{ARO}}$ and $\mathrm{RR}_k^{\mathrm{ARO}}$~~ $\mathrm{P}_i^{\mathrm{ARO}}$ and $\mathrm{P}_k^{\mathrm{ARO}}$ over pixels i and k (indicated by the purple arrow on Figure 3b,c)

$$\hat{R_{ik}} = \frac{\mathrm{RR}_i^{\mathrm{ANT}}}{\mathrm{RR}_k^{\mathrm{ANT}}} R_k^{\mathrm{GAUGE}} \cdot \frac{\mathrm{RR}_k^{\mathrm{ARO}}}{\mathrm{RR}_i^{\mathrm{ARO}}} \frac{\mathrm{P}_i^{\mathrm{ANT}}}{\mathrm{P}_k^{\mathrm{ANT}}} \cdot \frac{\mathrm{P}_k^{\mathrm{ARO}}}{\mathrm{P}_i^{\mathrm{ARO}}} \tag{3}$$

This estimated ratio conveys the hypothesis that the expected precipitation accumulation over an unobserved pixel can be retrieved from the precipitation accumulation observed at a nearby gauge by applying the AROME precipitation accumulation

ratio between these two locations. The underlying hypothesis is that AROME simulates realistic vertical gradients of precipitation even if it may be biased.

Considering all $k \in [1, n]$ nearby gauges (Figure 3a) and giving them a weight proportional to the distance d_{ik} from pixel i (e.g. Mont-Blanc) :

$$w_{ik} = max(0, 1 - \frac{d_{ik}}{d_0}) \tag{4}$$

(where $d_0 = 20$ km is the distance range parameter, arbitrarily chosen based on the density of the surface observation network), a weighted ensemble of estimated ratios is obtained (bars of the same colour in Figure 3d).

The weighted mean (M_i) and the spread (S) of this weighted ensemble of estimated ratios for pixel i are :

$$M_i = \frac{1}{W_i} \sum_{k=1}^{n} w_{ik} \hat{R_{ik}} \tag{5}$$

and

$$S_i = \sqrt{\frac{1}{W_i} \sum_{k=1}^{n} w_{ik} (\hat{R_{ik}} - M_i)^2} \quad \text{where} \quad W_i = \sum_{k=1}^{n} w_{ik} \tag{6}$$

The conversion of this weighted ensemble into a single ratio value R_i is illustrated in Figure 3d and follows the idea that :

- The lower the weighted spread S_i (i.e. the ratios $\hat{R_{ik}}$ estimated with the n nearby gauges are similar) the closer the estimated ratio $\hat{R_i}$ is to the weighted ensemble mean M_i.
- The higher the weighted spread S_i (i.e. the ratios $\hat{R_{ik}}$ estimated with the n nearby gauges give divergent information), the closer the estimated ratio $\hat{R_i}$ is to 1 (no relevant information can be derived, so no correction is applied). This situation is illustrated by the blue bars in Figure 3d.

Thus, the estimated ratio over the pixel i (the Mont-Blanc in the example) is given by :

$$\hat{R_i} = K_i \times M_i + 1 \times (1 - K_i) \quad \text{where} \quad K_i = e^{-\frac{S_i \times \ln(10)}{|M_i - 1|}} \; (K_i = 1 \quad \text{if} \quad M_i = 0) \tag{7}$$

K_i ensures that the estimated ratio lies between 1 and the weighted mean (M_i) depending on the spread S_i of the estimated ratio values. The constant exponential decay is fixed to ensure that if the weighted spread of the ensemble of estimated ratios among the n gauges (S_i) is high, the confidence that M_i is a good estimator of R_i^{TRUE} is low.

3.1.3 Uncertainty index estimation

The associated climatological uncertainty index U_i is defined so that it increases with:

- $|\hat{R_i} - 1|$, with a magnitude depending on the uncertainty associated with the estimation method, measured by K_i.
- the discrepancies between the individual ratio estimates associated with each gauge, which increase with S_i and $|M_i - 1|$.

The relationship between ANTILOPE errors and R^{TRUE} was estimated over rain gauge stations by computing a linear regression between ANTILOPE error compared to reference gauge measurements and the ratio R^{GAUGE} between ANTILOPE and reference gauges (see figure D1). This statistical relationship is used with the other estimators $\hat{R}_i$ and M_i to estimate an uncertainty U_i at each point of the domain.

U_i (in kg m^{-2}) is thus defined as a sum of two terms :

$$U_i = (1 + K_i) \times (A_i + 1) + (1 + S_i) \times (B_{ii} + 1) \tag{8}$$

where

$$A_i = \begin{cases} -20.137 \times (\hat{R}_i - 1) & \text{if } \hat{R}_i < 1 \\ 16.787 \times (\hat{R}_i - 1) & \text{if } \hat{R}_i \geq 1 \end{cases} \quad \text{and} \quad B_i = \begin{cases} -20.137 \times (M_i - 1) & \text{if } M_i < 1 \\ 16.787 \times (M_i - 1) & \text{if } M_i \geq 1 \end{cases} \tag{9}$$

~~In~~ By construction, $U_i \geq 2$ kg m^{-2}, and in extreme cases:

- ~~when~~ When all estimated $\hat{R_{ik}}$ are identical, $S_i = 0$ and $K_i = 1$, so U_i is obtained from A_i and proportional to $|\hat{R}_i - 1|$. This value increases as the estimated ratio deviates from 1, thereby indicating a likely systematic bias.
- ~~on~~ On the contrary, when the estimated $\hat{R_{ik}}$ are very contrasted (large S_i and K_i close to 0), U_i is mainly obtained from B_i and proportional to ~~$|M_i - 1|$~~$|M_i - 1| \times (1 + S_i)$. This value increases with S_i, increasing the estimated uncertainty in case diverging ratio estimates are obtained when applying the method to different reference gauges.

Figure 4 shows the ratio (Figure 4a) and uncertainty (Figure 4b) fields estimated with this method using data from winter 2021/2022. The yearly accumulation field after correction with the estimated ratio is shown in Figure 7b.

3.1.4 Dynamic correction of ANTILOPE daily precipitation fields

To deal with the unrealistic spatial patterns described in Section 3.1, a preprocessing step is applied to the daily ANTILOPE precipitation estimation fields and is illustrated in Figure 5. A first correction is applied to mitigate climatological errors using the estimated climatological ratio field obtained in Section 3.1.1 (Figures 4a and 5b). For each pixel i of the domain with a precipitation value ~~RR_i^{ANT}~~ P_i^{ANT} estimated by ANTILOPE (Figure 5a), the precipitation value after climatological correction (Figure 5c) is given by ~~$\text{RR}_i^d = \frac{\text{RR}_i^{\text{ANT}}}{R_i}$~~ $\text{P}_i^d = \frac{\text{P}_i^{\text{ANT}}}{R_i}$.

A modified Weighted Moving Average (WMA) filter is then applied using the uncertainty index (Figures 4b and 5c) obtained with the method described in Section 3.1.1. The WMA window includes all pixels within a $d_0 = 20$ km radius from pixel i (red circle in Figure 5c). To estimate the filtered precipitation of pixel i, each pixel k within this window is assigned a weight $c_{ik} = \frac{1 - d_{ik}/d_0}{U_k/U_0}$, where d_{ik} is the distance to the pixel at the centre of the moving window, U_k is the uncertainty index of pixel k (see Section 3.1) and $U_0 = 1$ kg m^{-2} a normalisation factor. Considering all n pixels of the window, the weighted mean

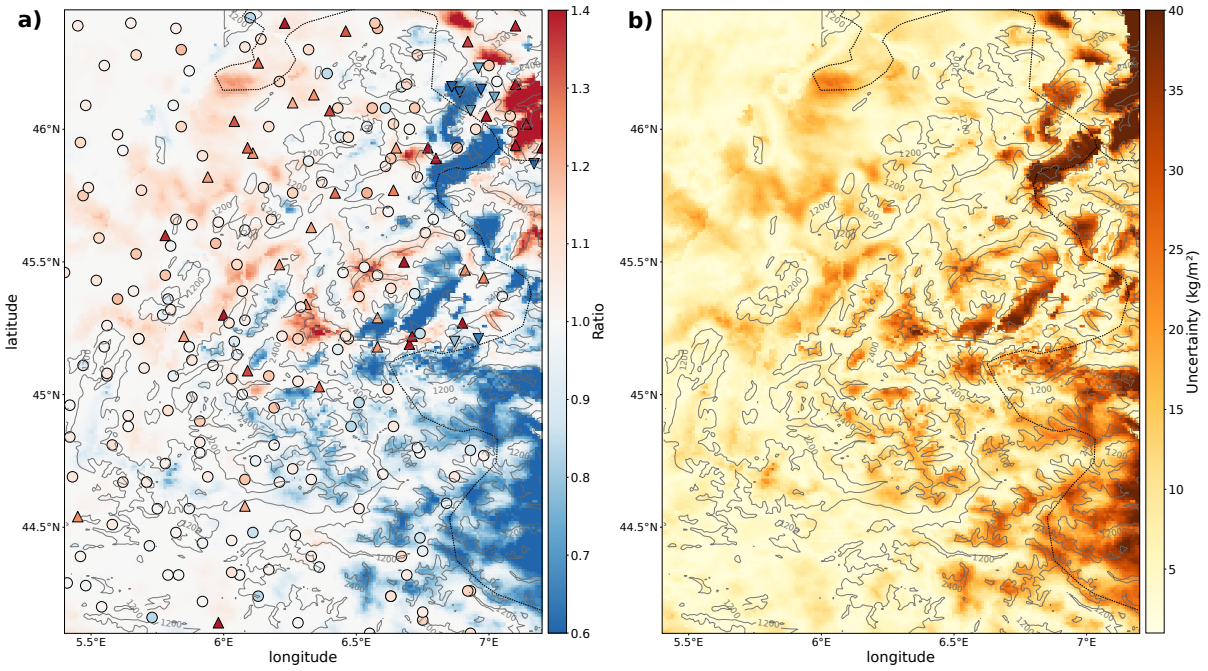

Figure 4. (a) Estimated climatological ratio ($\hat{R}$) based on the mean ratios between ANTILOPE and automatic gauge precipitation measurements and (b) uncertainty index (U) obtained using the method described in section 3.1.1. The grey lines indicate the 1200 m, 2400 m and 3600 m isolevels to illustrate the relationship between the estimated ratio ~~/error~~ and the relief.

precipitation ($\widetilde{\overline{\mathrm{RR}_i^d}\overline{\mathrm{P}_i^d}}$) and the variance ($V_i$) of all weighted precipitation values over the window associated to pixel $_i$ are :

$$\begin{cases} \overline{\mathrm{P}_i^d} = \frac{1}{C_i} \sum_{k=1}^{n} c_{ik} \mathrm{P}_{ik}^d \\ V_i = \frac{1}{C_i} \sum_{k=1}^{n} c_{ik} (\mathrm{P}_{ik}^d - \overline{\mathrm{P}_i^d})^2 \end{cases} \quad \text{where} \quad C_i = \sum_{k=1}^{n} c_{ik} \tag{10}$$

While a standard WMA filter would have considered ~~$\overline{\mathrm{RR}_i^d}$~~ $\overline{\mathrm{P}_i^d}$ as the new precipitation estimate for pixel i, in this work the filtered value is combined with the local value with respective weights depending on the estimated spatial error (Figure 5e):

$$\widetilde{\mathrm{RR}_i} = \frac{\mathrm{RR}_i^d c_{ii} + \overline{\mathrm{RR}_i^d} \frac{C_i}{n}}{c_{ii} + \frac{C_i}{n}} \frac{\mathrm{P}_i^d c_{ii} + \overline{\mathrm{P}_i^d} \frac{C_i}{n}}{c_{ii} + \frac{C_i}{n}} \tag{11}$$

This weighted correction ensures that a pixel with a low estimated spatial error ($c_{ii} >> \frac{C_i}{n}$) is not affected by the WMA filter while a pixel with a high estimated spatial error ($c_{ii} << \frac{C_i}{n}$) is closer to the average precipitation value in the localisation area.

The observation error (Figure 5f) is finally defined as the weighted spatial standard deviation of the daily precipitation over this window :

$$E_i = \sqrt{V_i} \tag{12}$$

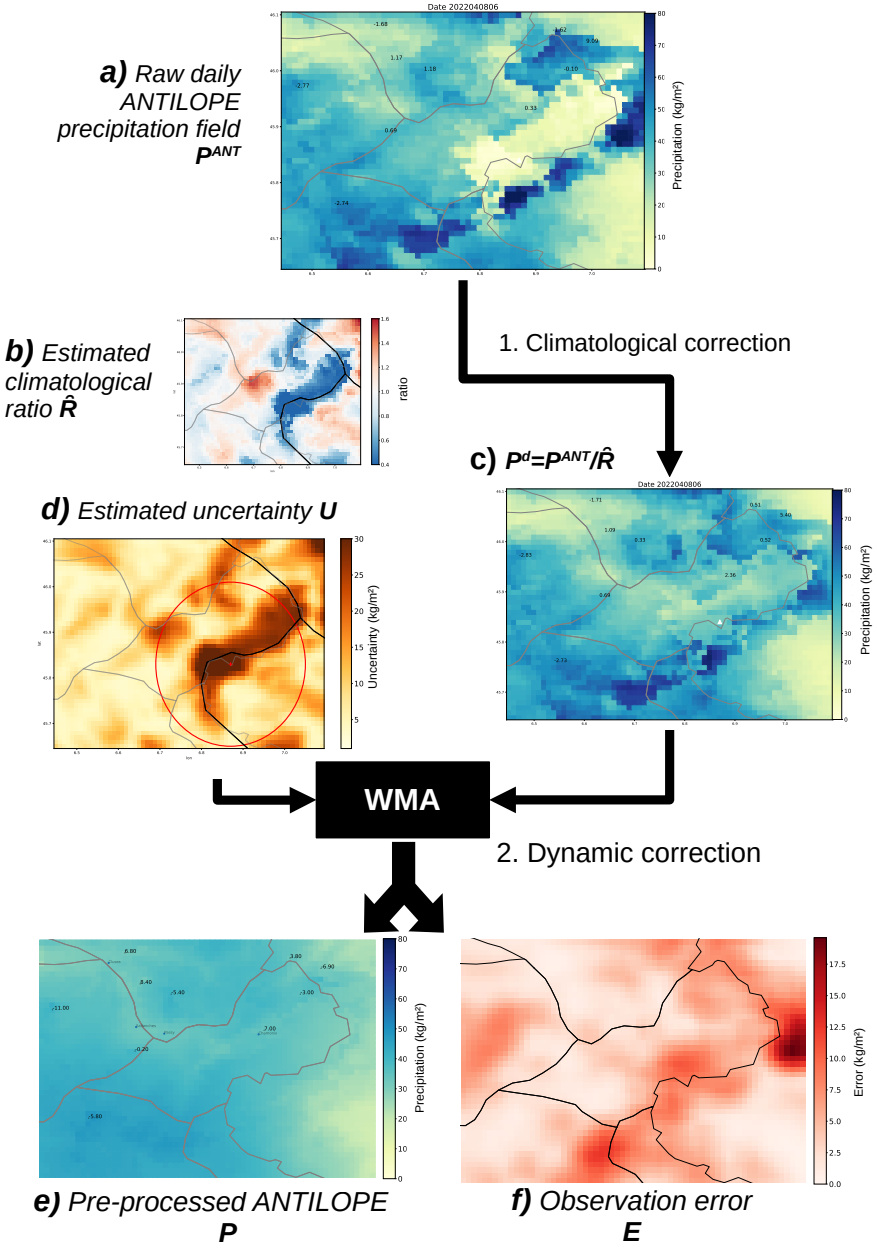

Figure 5. Illustration of the ANTILOPE pre-processing step applied for daily precipitation fields. First the climatological ratio ($\hat{R}$) estimated with the method described in Section 3.1.1 is used as a climatological correction factor to mitigate systematic errors. The resulting precipitation field ($\cancel{RR}^{d}\underset{\sim}{P^{d}}$) is then filtered with a modified Weighted Moving Average (WMA) method, which produces a smoother output daily precipitation field ($\cancel{RR}\underset{\sim}{P}$) as well as an associated error field (E). The weights used in the WMA step are a combination of an inverse distance weighting and the uncertainty index (U) estimated using the method described in Section 3.1.1.

3.2 Ensemble analysis

Three different methods are proposed to produce ensemble precipitation analyses of daily precipitation fields based on the pre-processed ANTILOPE precipitation fields obtained with the algorithm described in Section 3.1.4. Two of these methods are based on ensemble data assimilation algorithms that combine an observation (pre-processed ANTILOPE precipitation field) and an ensemble of NWP model outputs (post-processed PE-AROME precipitation fields).

3.2.1 Random Sampling

The first ensemble analysis method uses output fields of the ANTILOPE preprocessing step described in Section 3.1.4 ~~.~~ to construct 16 precipitation fields ~~are constructed~~ by random perturbations around the corrected precipitation field $\widetilde{P^i}$ (Figure 5e).

We have chosen to separate the spatial and dynamical components of the observation error in a similar way to Villarini et al. (2014). The ~~magnitude of the perturbation is thus determined by two terms : a~~ spatial observation error E_i ~~resulting from the preprocessing step (Section 3.1.4), which~~ accounts for the ~~spatial variability of the ANTILOPE quality described in Section 3.1 . a dynamic error~~ intrinsic spatial structure of the error associated with measurement issues in the ANTILOPE product described in section 3.1 and comes from the preprocessing step (Section 3.1.4). The dynamic error is associated to the uncertainty of each individual event and is expressed as a fraction of the precipitation intensity. The magnitude of this error has been estimated from the linear regression between the ANTILOPE precipitation intensity and the associated error (not shown) and is set at 30% of the precipitation intensity.

The random perturbation for a given ensemble member m is then determined by two different values :

- N^m is sampled from a normal distribution centered on 0 and with a variance of 1 and applied as a multiplicative factor to the spatial component of the error
- G^m is sampled from a gamma distribution (with a shape parameter $k = 2$ and a scale parameter $\theta = \sqrt{1/k}$ ensuring a variance of 1) shifted by $k\theta$ so that the mean of the distribution is 0

For each pixel i of the domain with a corrected precipitation ~~RR$_i$~~ P$_i$ (equation 11) and an associated error E_i (Equation 12), the precipitation in member m of the output ensemble is thus :

$$\text{~~RR~~P}_i^m = \max(\text{~~RR~~P}_i(1 + 0.3G^m) + N^m.E_i, 0) \tag{13}$$

Drawing random values from distributions with a mean centred on 0 ensures that the resulting ensemble mean is expected to be ~~RR$_i$~~P$_i$, unless the magnitude of the spatial error is greater than the precipitation value itself (in which case the 0 $\mathrm{kg\,m^{-2}}$ lower bound may shift the ensemble mean upwards).

This Random Sampling (RS) analysis is the direct translation of the ANTILOPE pre-processing step, which produces a corrected precipitation field and associated error, into an ensemble analysis. It can be considered as a benchmark product for ensemble analysis.

3.2.2 Particle Filter

The second ensemble analysis is based on particle filter theory. Particle Filters (PF) are sequential Monte Carlo algorithms commonly used for data assimilation in non-linear systems (van Leeuwen, 2009). These methods are based on the approximation of the model probability density function (PDF) by an ensemble of model states (called ensemble members and denoted X in the following equations). When an observation ~~RR_i~~ P_i is available, this ensemble of model states is updated in two steps. First, the different members are weighted according to their likelihood (distance to the observation). Then, the model state PDF is updated by resampling the different members according to their weights: members with high weights are replicated, while those with lower weights are dropped. To deal with the known problems of the PF algorithm for large numbers of observations, Snyder et al. (2008) suggests to reduce the dimension of the problem by splitting a large set of observation into smaller subsets. Here we chose to apply the particle filter algorithm for each pixel independently, reducing the problem dimension to 1 observation for a 16-member ensemble~~, as suggested by Snyder et al. (2008)~~. Thus, the likelihood of the PE-AROME precipitation forecast X_i^m of member m over a pixel i can be simply formulated as a function of the corresponding ANTILOPE pre-processed precipitation ~~RR_i~~ P_i (Equation 11) and its associated error E_i:

$$L_i = e^{-\left(\frac{(X_i^m - \mathrm{RR}_i)}{E_i}\right)^2 - \left(\frac{(X_i^m - \mathrm{P}_i)}{E_i}\right)^2} \tag{14}$$

The main disadvantage of this local approach is the loss of spatial consistency. In fact, the analysed precipitation accumulation of neighbouring pixels can be obtained from combinations of different members. To obtain consistent precipitation fields, a field reconstruction step based on an ensemble copula coupling method (Schefzik et al., 2013) is applied during the re-sampling step. This re-sampling procedure is based on preserving the rank structure of the ensemble members from the original ensemble to the analysed one, ensuring, for example, that the same member has the highest precipitation value on a pixel before and after the analysis. In this way, large-scale structures are preserved in the analysis, although local discrepancies may persist.

3.2.3 Ensemble Kalman Filter

The last ensemble analysis is based on the Ensemble Kalman Filter algorithm (EnKF, Evensen, 2003). EnKF are Monte Carlo implementations of the Kalman Filter (Kalman, 1960), where an ensemble of simulations is used to sample the model state (background) distribution. The main advantage over the PF is that it preserves the spatial consistency of the analysed fields. We have chosen an approach inspired by that of Atencia et al. (2020b), using only the analysis step of the EnKF, which consists in updating each ensemble member with the information coming from an observation. It is assumed that the cross-correlations between the background and the observation are negligible, so that each pixel of the domain is processed independently. For a given pixel i, the precipitation estimated by the ensemble member m (X_i^m) is pulled towards the corresponding observation (ANTILOPE pre-processed precipitation ~~RR_i~~ P_i from equation 11):

$$\widetilde{X_i^m} = X_i^m + K_i(\mathrm{RR}_i - X_i^m) \tag{15}$$

where the so called gain factor K_i defines the relative influence of the original background value and the observation, based on the observation error E_i defined in equation 12 and the background error represented by the 16-member ($N = 16$) ensemble spread $B_i = \frac{1}{N-1}\sum_{m=1}^{N}(X_i^m - \overline{X_i^m})(X_i^m - \overline{X_i^m})$:

$$K_i = B_i(B_i + E_i)^{-1} \quad (16)$$

3.3 Evaluation Method

The aim of this work is to produce an ensemble precipitation analysis with a resolution of 1 km over the French Alps, which should ideally fulfil the standard requirements expected from ensemble modelling systems (reliability and resolution). This ensemble analysis should ~~also meet the following specific conditions:~~

~~reduced systematic bias~~ reduce the systematic biases of the ensemble mean and improve precipitation estimation as compared to other existing products~~improved predictive added value of the analysis as compared to other existing products ensemble members~~. Finally, spatial artefacts affecting the radar-based precipitation fields must not be propagated to the analysis, meaning that each individual ensemble member should feature realistic spatial structures.

While the first three requirements can be evaluated at the point scale with the available ~~local~~ independent observations of the ski-resorts network, the lack of a spatialised reference observation of the precipitation fields implies that an objective evaluation of the spatial structure of the analysed fields is not directly possible at this stage.

The different products were evaluated during the 2021/2022 winter season. Each available ski-resort observation of 24 h precipitation at ~~7~~8:00 CET was compared with the corresponding 24 h precipitation estimate at ~~6~~7:00 UTC on the corresponding 1 km pixel for each product. First, a comprehensive evaluation of existing precipitation estimation products is provided. ~~For~~ In the case of deterministic products, this evaluation focuses on ~~systematic errors and biases~~ the magnitude of systematic biases and their spatial distribution as well as ~~spatial structure issues~~ on the consistency of the spatial structure of the precipitation ~~field, while for the PEAROME product ensemble characteristics are also considered~~fields. For ensemble products, ensemble characteristics such as the spread skill are also taken into account. As all proposed ensemble analysis methods rely on a pre-processing of the ANTILOPE precipitation estimation product, an objective evaluation of the pre-processing step and its performance compared to existing products is provided. A subjective comparison of the spatial structures of the precipitation fields is also discussed. Finally, the three ensemble analysis methods are evaluated, and their performances are compared in terms of mean bias, spread skill and overall predictive value added.

Biases, ratios and root mean square deviations are computed for each reference station by comparing the estimated and observed time series of 24 h precipitation ~~values time series~~records. For ensemble products, the estimated value is taken as the ensemble mean. The predictive added value of the different products is assessed using two probabilistic scores commonly used for ensemble prediction systems : the Brier Score (BS, Brier, 1950) and the Continuous Rank Probability Score (CRPS, Matheson and Winkler, 1976; Candille and Talagrand, 2005). More details on these scores can be found in Appendix B. Another important aspect of ensemble forecasting is the reliability of the ensemble spread : an ensemble forecasting system is reliable if the magnitude of its spread for different events and locations matches the associated spread of the ensemble mean error. This can be inferred from a scatter plot of the ensemble spread against the ensemble mean error (Hopson, 2014).

4 Results

4.1 Evaluation of existing precipitation products

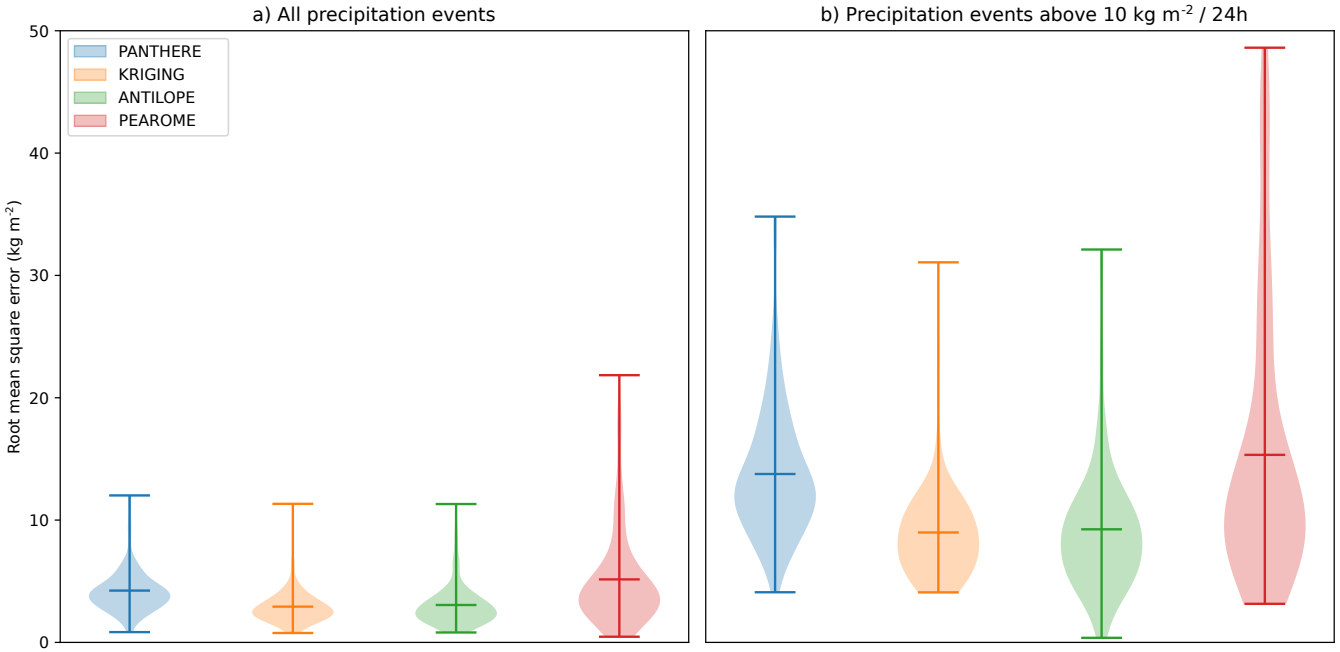

Figure 6. Distribution of the root mean square deviation of the existing precipitation estimation products from all ski-resort reference observations of daily precipitation (a) and daily precipitation above 10 $\mathrm{kg\,m^{-2}}$ (b) over all evaluation stations. Precipitation estimates come from radar-only observations (PANTHERE), a kriging of gauge observations (KRIGING), a fusion of radar and gauge observations (ANTILOPE), and the ensemble mean of the post-processed high-resolution ensemble NWP model (PEAROME).

The main results of the evaluation of the existing products (PANTHERE, gauge-kriging, ANTILOPE and post-processed PEAROME) are summarised in Figure 6 showing the distributions of the root mean square deviation from the ski-resorts reference observations for all data (Figure 6a) and for observed precipitation events above 10 $\mathrm{kg\,m^{-2}}$ in 24 h only (Figure 6b). This ~~Figure~~ figure shows that the precipitation estimation from the network of automatic observation stations (KRIGING) performs better than the estimation from radar measurements only (PANTHERE) in mountainous areas. The RMSE of the combination of these two sources of information (ANTILOPE) is in the same order of magnitude as that of the KRIGING product, but is sometimes improved for precipitation events above 10~~kg/m2/24h~~ $\mathrm{kg\,m^{-2}\,/\,24h}$ (b). Similar results are obtained when only snowfall events are considered (Figure E1). Other results (frequency of relative errors less than 20%, not shown) also support the choice of ANTILOPE as the best precipitation product available for an ensemble analysis system, especially since ANTILOPE performs equally well when only solid precipitation is considered ~~when only solid precipitation is considered~~ (see Figure E1 in Appendix ~~B). However, the spatial distribution of ANTILOPE's performance shows that even if there is~~

no widespread systematic bias, there is a high spatial variability in its performance. C). Figure 7a shows that the mean ratio between ANTILOPE estimated precipitation and the corresponding ski-resort reference observations can reach both very high (above 1.5, in red) and very low (below 0.5, in black) values, even for stations separated by only a few kilometres. However, the spatial distribution of ANTILOPE's performance shows that even if there is no widespread systematic bias, there is a high spatial variability in its performance. Moreover, the lack of evaluation stations at high altitudes or near mountain ridges does not allow a good documentation of the spatial artefacts already discussed in section 3.1.

Finally, the RMSE of the PEAROME ensemble mean is much higher than that of all observation-based products (see Figure effig:rmse6). This indicates that even after statistical post-processing, PEAROME forecasts provide a less reliable estimate of precipitation than observation-based products. An evaluation using metrics specifically adapted to ensemble forecasting, such as the Brier Score (BS) or the Continuous Ranked Probability Score (CRPS), shows that the probabilistic nature of PEAROME does not compensate for deficiencies in deterministic estimation (Figure 9).

Yearly ANTILOPE precipitation accumulation over the entire French Alps domain and mean ratio with the available evaluation data from the ski-resorts network before (a) and after (b) application of the climatological correction factor obtained by the observation error estimation method.

4.2 Evaluation of precipitation analyses

4.2.1 Impact of the pre-processing procedure

The ensemble precipitation analyses in this study are based on the ANTILOPE preprocessing step , as explained in Section 5. This step described in Section 3.1.4, which relies on the error estimation method described in Section 3.1.1. Figure 7b demonstrates the ability of this method to identify ANTILOPE biases, showing the same information as Figure 7a but with unbiased precipitation estimates using the ratio from Figure 4a. Spatial artefacts that are visible in the yearly precipitation accumulation of Figure 7a appear to be reduced in Figure 7b. In addition, the mean ratio with reference ski-resorts observations after climatological correction is often closer to 1 (green dots) than for raw ANTILOPE precipitation estimates.

Figures 8 and 9 show the performance of the full pre-processing step. The error estimation method reduces biases with ratios closer to 1 after the ANTILOPE pre-processing step (light blue) than before (brown), as confirmed by Figure 8. In addition, Figure 9 shows that this pre-processing step improves both CRPS and BS, regardless of the precipitation threshold.

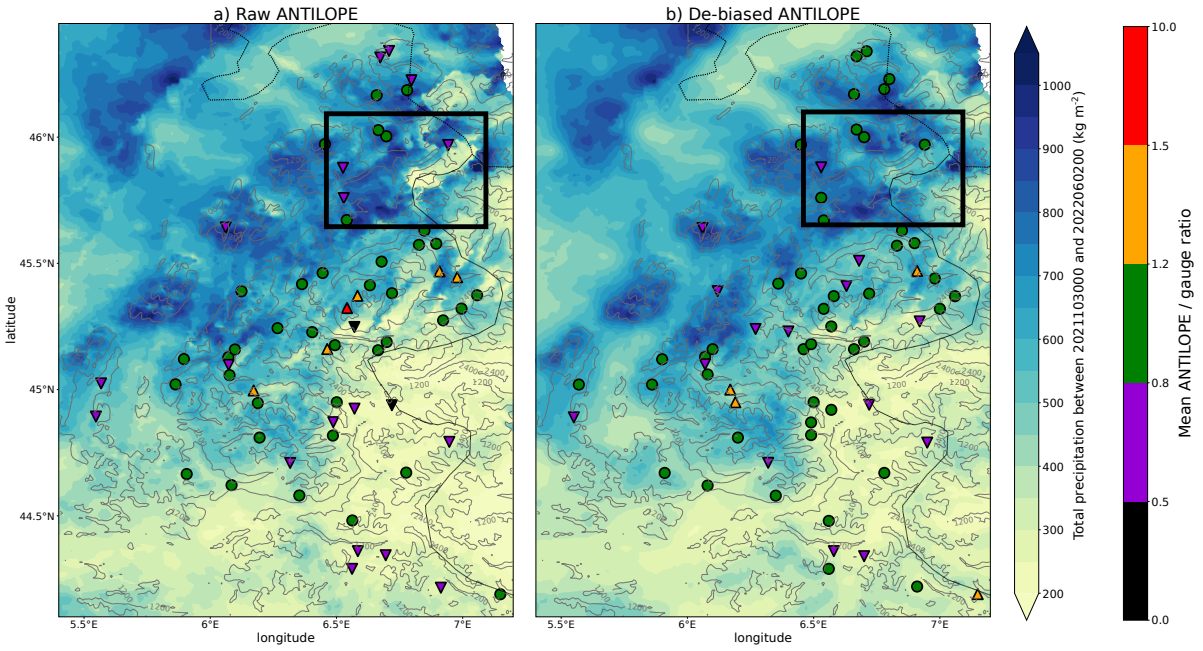

Figure 7. Yearly ANTILOPE precipitation accumulation over the entire French Alps domain and mean ratio with the available evaluation data from the ski-resorts network before (a) and after (b) application of the climatological correction factor obtained by the observation error estimation method. The marker associated with each evaluation station is a circle if the ratio between ANTILOPE and the reference gauge precipitation estimates is not considered significant (between 0.8 and 1.2) and a triangle facing up (resp. down) if this ratio is significantly higher (resp. lower) than 1. The black rectangle indicates the Mont Blanc area from which the precipitation fields shown in Figures 3 and 5 are derived.

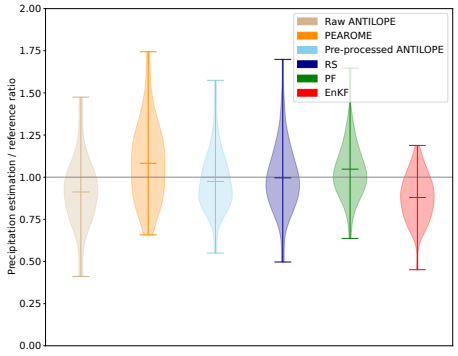

Figure 8. Ratio between estimated and observed accumulated precipitation over all ski-resorts reference stations for the different methods assessed in this study.

4.2.2 Skill of ensemble analyses

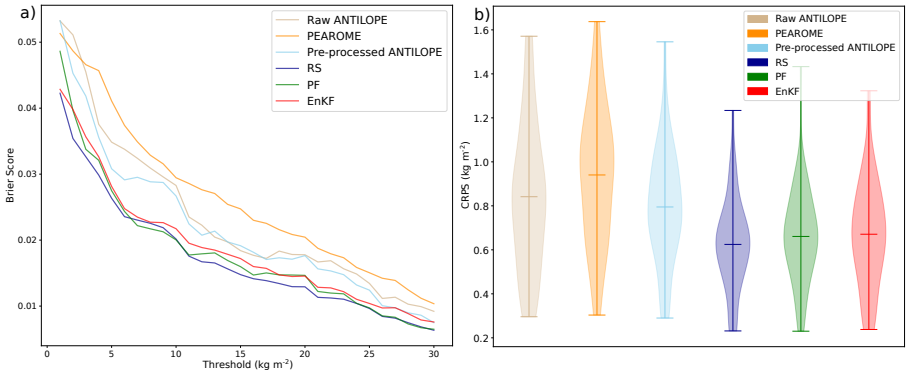

Figure 9. Evolution of the mean Brier score over all evaluation stations and dates for different daily precipitation thresholds ranging from 1 $\mathrm{kg\,m^{-2}}$ to 30 $\mathrm{kg\,m^{-2}}$ (a). Distribution of the mean continuous rank probability score over the reference stations (b). Six precipitation estimation products are shown: existing products in orange shades, deterministic products in light colors and ensemble products in dark colors.

All ensemble analyses (RS, PF and EnKF) evaluated in Figures 8 and 9 show better performance than the pre-processed ANTILOPE and PE-AROME products respectively. Figure 9a shows a clear improvement of the BS for the three ensemble analysis methods, which all have similar performance for low to moderate precipitation events (up to 10 $\mathrm{kg\,m^{-2}}$ in 24 h). However, for precipitation above 10 $\mathrm{kg\,m^{-2}}$, the basic RS method outperforms the two methods based on ensemble data analysis. For precipitation events above 25 $\mathrm{kg\,m^{-2}}$, the PF and EnKF analyses do not perform significantly better than the deterministic ANTILOPE pre-processing. This result is confirmed by the CRPS values (Figure 9b), which are significantly lower for the RS analysis than for the PF and EnKF analyses.

~~Distribution of the ratio between the precipitation accumulation estimated by the different products and the ski-resorts reference observations.~~

~~Evolution of the mean Brier score over all evaluation stations and dates for different daily precipitation thresholds ranging from 1 to 30 (a). Distribution of the mean continuous rank probability score over the reference stations (b). Six precipitation estimation products are shown: existing products in orange shades, deterministic products in light colors and ensemble products in dark colors.~~

4.2.3 Spread-skill consistency

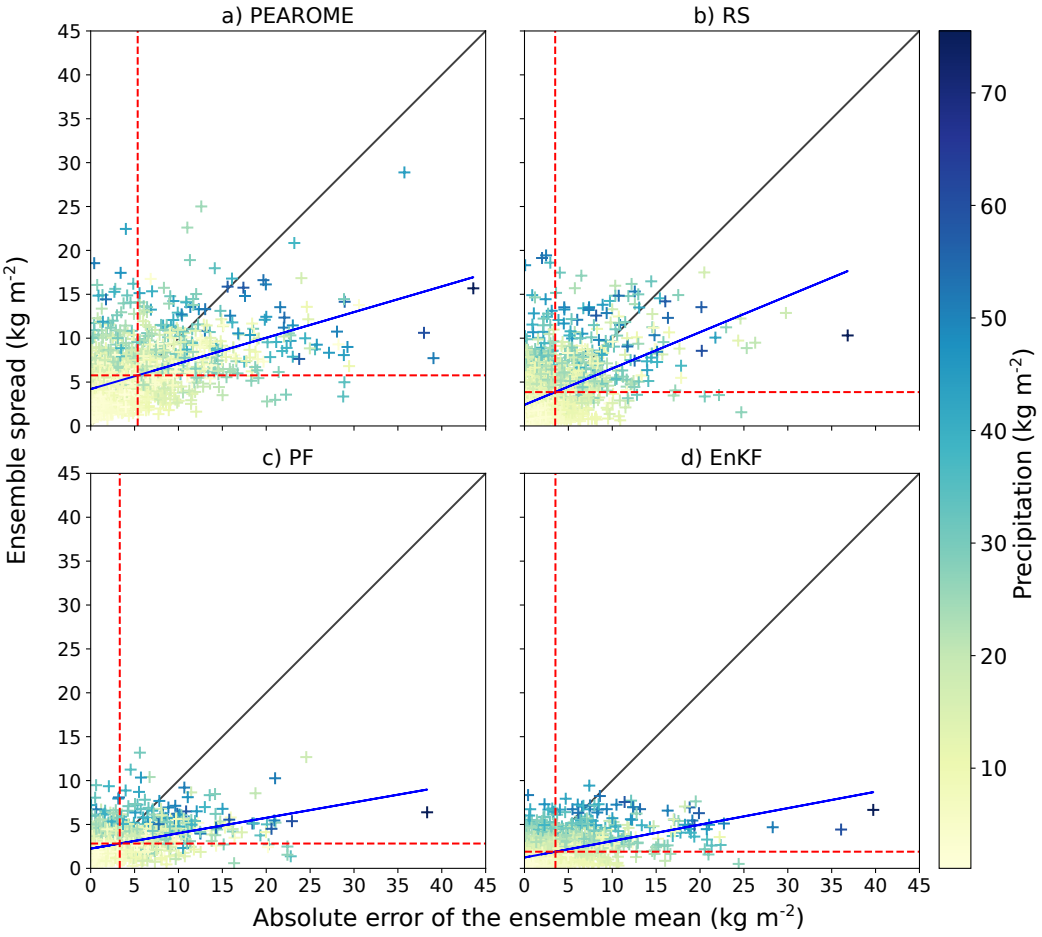

Figure 10. Spread skill of the different ensemble products. Each cross represents a specific evaluation point and date, its color indicates the corresponding 24 h precipitation. The blue line is the regression line of the scatter plot. The horizontal (resp. vertical) red line shows the mean ensemble error (resp. ensemble spread).

The spread of the raw PEAROME ensemble and the three ensemble analyses is shown in Figure 10 as a function of the ensemble mean. Ideally, the spread should be of the same magnitude as the quadratic error of the ensemble mean ~~(along the black line Hopson,~~ (along the one-to-one line) as described by Hopson (2014). The post-processing of the PEAROME ensemble (Figure 10a) is specifically designed to optimise statistical ensemble properties such as the mean spread. Consistently, the magnitude of the ensemble spread is comparable to the magnitude of the ensemble mean error, with approximately the same mean value (indicated by the red dashed lines).

The spread skill of the RS analysis (Figure 10b) shows that the magnitude of the ensemble spread ~~is also quite comparable to~~ match the magnitude of the ensemble mean error, with the spread on average slightly higher than the error (red dashed lines).

The blue regression line shows that high errors also tend to be underestimated by the ensemble spread, but ~~to a lesser extent~~not as much as for the raw PEAROME product. The RS ensemble spread is more often close to 0 $\mathrm{kg\,m^{-2}}$, indicating a higher confidence in the precipitation estimate than the PEAROME ensemble.

Figures 10c and d show that the spread of the two ensemble analyses based on ensemble data assimilation methods clearly underestimate the magnitude of the ensemble mean error. The spread of both analyses barely exceeds 10 $\mathrm{kg\,m^{-2}}$, while the ensemble mean error is often above 15 $\mathrm{kg\,m^{-2}}$.

It is also worth noting that the ensemble mean error of the three ensemble analyses is on average significantly lower than that of the PEAROME ensemble (vertical red dashed lines indicate an error almost twice lower) and does not reach the same extreme values (in particular, the RS analysis mean error exceeds 25 $\mathrm{kg\,m^{-2}}$ only once).

~~Spread skill of the different ensemble products. Each cross represents a specific evaluation point and date, its color indicates the corresponding 24 precipitation. The blue line is the regression line of the scatter plot and the horizontal (resp. vertical) red line shows the mean ensemble error (resp. ensemble spread).~~

4.2.4 Spatial patterns

~~Finally,~~ Figure 11 shows an example of four precipitation fields for each of the four ensemble products over the Mont-Blanc area (see Figure 1) for the same date as the fields in Figure 5. This illustrates the spatial variability of precipitation produced by the different analysis methods. The RS analysis produces very smooth fields, similar to the corresponding preprocessed ANTILOPE field (Figure 5e), but with different precipitation intensities. The fields resulting from the PF and EnKF analyses are a combination of this pre-processed ANTILOPE field and the corresponding PEAROME fields, also visible in Figure 11e and f. Despite the field reconstruction step described in section 3.2.2, the application of the particle filter algorithm at the pixel scale results in rather noisy fields. EnKF precipitation fields better preserve the spatial variability of the corresponding PEAROME fields, with a convergence of intensity towards the pre-processed ANTILOPE field. As explained in section 3.3, the lack of spatialised reference observations prevents any objective evaluation of the spatial structures of the different analyses, except that noisy PF precipitation fields are a serious shortcoming for further exploitation of this precipitation analysis.

~~Example of 24 precipitation fields of 4 December 2021 over the Mont-Blanc area (see Figure 1). a) Raw ANTILOPE observation. b) Pre-processed ANTILOPE field. c) Members 1 to 4 of the post-processed PEAROME ensemble. d) Members 1 to 4 of the Random Sampling analysis around the pre-processed ANTILOPE field (b). e) Members 1 to 4 of the Particle Filter analysis. f) Members 1 to 4 of the Ensemble Kalman Filter analysis. The full 16-member ensembles for each ensemble product are shown in Appendix C.~~

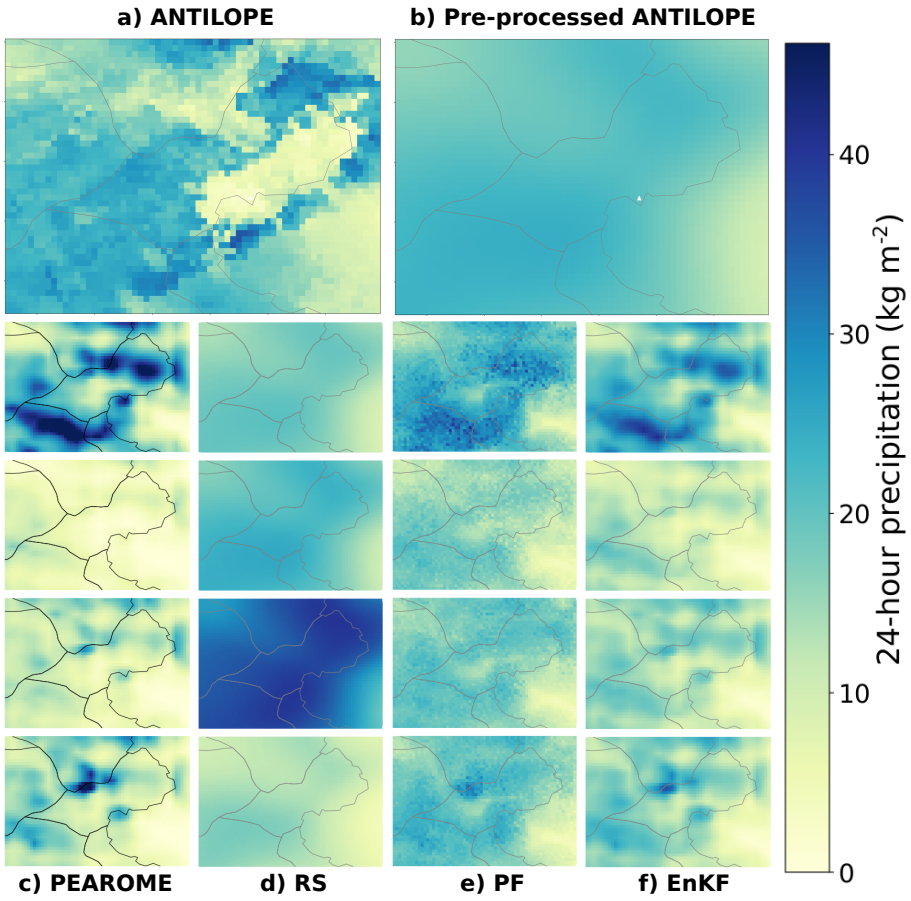

Figure 11. Example of 24 h precipitation fields of 4 December 2021 over the Mont-Blanc area (see Figure 1). a) Raw ANTILOPE observation. b) Pre-processed ANTILOPE field. c) Members 1 to 4 of the post-processed PEAROME ensemble. d) Members 1 to 4 of the Random Sampling analysis around the pre-processed ANTILOPE field (b). e) Members 1 to 4 of the Particle Filter analysis. f) Members 1 to 4 of the Ensemble Kalman Filter analysis. The full 16-member ensembles for each ensemble product are shown in Appendix C.

5 Discussion

5.1 Precipitation estimation products in mountainous areas

The evaluation of different precipitation ~~estimation~~ products with an independent observation network in ski resorts (section 4.1) shows that observation-based estimates ~~are more competitive~~ provide better results than forecasts from the PE-AROME ensemble, even after a post-processing step designed to remove biases and improve reliability. This contradicts, at least for the French Alps, the conclusion of Lundquist et al. (2019) that high-resolution precipitation models outperform observations in mountainous areas. However, these products have complementary qualities and shortcomings.

Although radar observation provides spatialised information, its quality is very heterogeneous in space and suffers from important limitations in mountainous areas. Ground clutter and partial beam blocking are common due to the complex topography, and the management of these disturbances in the radar processing algorithm described in section 2.1 does not compensate for the loss of information.

The most obvious effect of the interception of a radar beam by a mountain is a significant underestimation of the mean precipitation above the pixels affected by ground clutter, which produces spatial artefacts in the precipitation fields. This degradation of precipitation estimation due to increased height of the lowest usable radar measurement above ground is particularly pronounced for stratiform precipitation systems with relatively small vertical extent ~~(?)~~(Germann et al., 2022). This is the case for the majority of the winter precipitation events in the French Alps on which this study focuses, thus limiting the extrapolation of its conclusions to convective situations.

Radar-based precipitation estimates are also less reliable over valleys, where the height difference between the radar beams and the ground is more important (Figure 2) and precipitation variations below the beams are neglected. However, this problem is compensated in the ANTILOPE product by the use of many in-situ measurements at these altitudes (Figure 1). On the contrary, higher elevation areas are under-sampled (Figure 1 shows that ~~very few~~ only 31 out of 512 available automatic gauges are ~~available~~ above 2000 m a.s.l.), so radar measurement problems at these altitudes cannot be well corrected. The lack of reference observations over ridges (Figure 1) also makes it impossible to assess the magnitude of the error at these altitudes. This led us to develop a preprocessing step, described in section 3.1.4, to mitigate the resulting spatial artefacts and provide a spatialised assessment of the highly variable uncertainty of this product.

5.2 Observation error

Most of the results of this study are based on the spatialisation of the error associated with the ANTILOPE precipitation product. Despite the positive results at the local scale shown in section 4.2, this method ~~suffers from~~ has some limitations and possible improvements have been identified :

- The WMA filter described in section 3.1.4 tends to smooth precipitation fields. This is a necessary compromise to remove unrealistic spatial patterns that would be very detrimental to snow modelling (with unrealistic snow amounts near ridges, as illustrated by Haddjeri et al., 2023, with the ANTILOPE raw product). However, it can potentially erase realistic spatial structures well detected by the radar, even if the formulation of the equation 11 and the parameters in the error estimation method have been chosen to ensure minimal modification over pixels with no clear systematic errors. High intensity precipitation kernels are therefore often smoothed by the preprocessing step, which may affect the detection of extreme events. The spatial structures of the analysed precipitation fields could not be evaluated with the reference data set used in this study. An indirect evaluation of simulated snow depths based on the precipitation analyses presented here is planned in a future study, following the workflow already proposed by Haddjeri et al. (2023). ~~It~~ In such a study, we will compare simulated snow depths with snow depth maps derived from high resolution Pleiades satellite images (Deschamps-Berger et al., 2020) over the Grandes Rousses domain.

- The climatological approach of this method makes it highly dependent on ANTILOPE evolutions (changes in the algorithm, in the available radar or in the in-situ measurements used), so that regular recalibration would probably be required.

- Additional available radar information, not considered in this work, could also provide relevant information for the estimation of the observation error (height above ground of the lowest available radar beam, position of ground clutter,...).

- As most gauges are located in valleys or at low altitudes, the uncertainty index is generally higher at high altitudes (Figure 4b). To mitigate the over-representation of low elevation pixels in the WMA filter, the weights computed in the preprocessing step (section 3.1.4) could take into account the elevation difference between pixels.

- The uncertainty information generated by the kriging with external drift algorithm producing the raw ANTILOPE product could be used to improve the estimation of the associated observation error.

- The error estimation method described in Section 3.1.1 could be applied to daily precipitation fields in order to improve the handling of instances where the ANTILOPE error is significantly different from its climatological value.

5.3 Choice of the data assimilation method

The random sampling ensemble analysis (RS) combined with the preprocessing method described in section 3.1.4 proved to fulfil most of the requirements for a suitable ensemble precipitation analysis in complex terrain, even if the accuracy of the spatial structures of the individual fields remains to be evaluated. This result is particularly promising for the future goal of assimilating satellite snow observations into snowpack simulations. It accurately accounts for precipitation uncertainties and thus addresses the main limitation raised by Cluzet et al. (2022); Deschamps-Berger et al. (2022) to take full advantage of the assimilation of snow observations. It is based on a very simple ensemble generator and could probably be significantly improved using methods inspired by those reviewed in Mandapaka and Germann (2010). The RS analysis is also the most computationally efficient and requires less data handling.

On the contrary, the information provided by high-resolution NWP models may suffer from larger errors on average, but is more physically consistent and can therefore be expected to have more spatially homogeneous errors than precipitation estimation derived from radar measurements. Ensemble forecasts also provide uncertainty information that is lacking in deterministic observational products. This motivated the use of the AROME climatology for the ANTILOPE pre-processing step, as well as the PEAROME daily precipitation fields through data assimilation methods. However, the evaluation presented in section 4.2 shows that most of the NWP added value comes from the use of the AROME climatology, at least at the local scale. In particular, the EnKF and PF analyses proved to be very under-dispersive. This deficiency could be very detrimental in snow simulation systems based on ensemble assimilation of snow observations with e.g. the particle filter ~~Cluzet et al. (2022); Deschamps-Berger et al. (2022)~~(Cluzet et al., 2022; Deschamps-Berger et al., 2022), as in many cases no optimal scenario would be available among the backgrounds. The assimilation of snow observations would then fail to provide better estimates of snow conditions. The under-dispersion of both PF and EnKF analyses is an inherent aspect of these

ensemble data assimilation algorithms: the main concept of these algorithms is to modify an ensemble of initial simulations (the background state) with the knowledge of an observation and its associated error. Both PF and EnKF methods ensure that the resulting analysed ensemble has a smaller dispersion than the background ensemble: the PF analysis is a sub-ensemble of the background ensemble and the EnKF equations described in section 3.2.3 imply a convergence of all background members to the observation as long as the observation error is comparable to the background one (Hotta and Ota, 2021)). Thus, a well-dispersed background ensemble, such as the post-processed PEAROME ensemble used in this study, can only lead to under-disperse analysed ensembles when PF or EnKF ensemble data assimilation algorithms are applied using observations with lower error than the background~~, such as~~ . This is typically the case with the ANTILOPE product, at least where reference data is available. However, in instances where the ANTILOPE uncertainty is significantly higher than the PEAROME error, the impact of these algorithms on the background ensemble is minimal.

The consistency of the spatial structures produced by these analyses could not be objectively assessed. However, the PF analysis fields suffer from obvious spatial artefacts (even after a field reconstruction step). This PF method is also numerically more expensive when applied at the pixel scale, which can be a severe limitation for application over large areas. Applying the PF algorithm simultaneously to the whole domain would lead to degeneracy due to the very high number of available observations compared to the ensemble size. A compromise solution would be to apply the PF algorithm over well-chosen subdomains, as suggested by Cluzet et al. (2021), but this would require further strong hypotheses for the definition of the subdomains. Considering all these elements, the analysis method based on the EnKF algorithm is clearly more suitable than the one based on the PF, but it does not outperform the RS analysis based on observations only and relying on a simplistic ensemble generation method, at least at the local scale used for the evaluation in this study. This result implies that the information provided by direct radar and in-situ precipitation measurements cannot be significantly improved by daily high-resolution NWP forecasts up to 33 h lead time~~, despite major shortcomings~~. However, only a spatialised evaluation could determine whether the spatial structures produced by a high-resolution NWP model and transferred to the EnKF analysis (see Figure 11) can prove to be complementary information with added value.

6 Conclusions

This study first provides an evaluation of ~~the advanced~~ stat-of-the-art precipitation estimation products available over the French Alps. It focuses in particular on radar-based observational products, the quality of which is not well documented, but is often considered to be poor in complex terrain. Our evaluation against an independent dataset of ~~local~~ 24 h precipitation accumulation from ski resorts showed that radar measurements combined with precipitation estimation from an automatic gauge network provide the best precipitation estimates among all products considered. In particular, the performance of such a product was shown to be significantly better than the output of the most advanced high-resolution numerical weather prediction model available in this study's area. This is noteworthy because previous studies have suggested that these models have outperformed observation-based precipitation estimates in some areas (Lundquist et al., 2019). However, significant shortcomings related to expected radar measurement problems were also identified. ~~The authors~~ In this study, we developed a method to estimate the

spatial distribution of climatological errors associated with unrealistic spatial patterns, such as unrealistically low accumulation over ridges, in annual and daily precipitation fields. This method has been developed and evaluated for the French Alps, but can be applied to any mountainous area covered by similar precipitation estimation products. ~~They also applied a correction algorithm~~ A correction algorithm was applied to daily precipitation estimates using the mean precipitation gradient from the AROME NWP model. Although no objective evaluation of the spatial structure of the daily precipitation fields produced by this method could be carried out, this correction seems to be able to mitigate spatial artefacts visible in the annual precipitation accumulations and proved to improve precipitation estimates ~~at the local scale~~over a set of reference stations. Some possible refinements of the method have also been outlined for future improvements, such as taking advantage of additional available radar information or the inclusion of topographic criteria in the weighting step.

Three different ensemble analysis methods were then implemented based on the corrected precipitation fields obtained by the previous method. A first reference analysis randomly samples values around these observed precipitation fields from their estimated error. Analyses were then conducted using two ensemble data assimilation algorithms, namely the particle filter and the ensemble Kalman filter. These algorithms were used to combine the corrected precipitation fields with those produced by a high-resolution numerical weather prediction model ensemble. The aim was to explore the complementarity of these two sources of information. The method based on the particle filter turned out to be both more numerically expensive and to suffer from major drawbacks, especially regarding the excessive noise in the produced precipitation fields. The method based on the ensemble Kalman filter proved to be more suitable, but did not outperform the ensemble analysis based only on the reference observation dataset used in this study. However, the spatial structures obtained by both methods are significantly different and a spatialised evaluation will be necessary to determine which is the most appropriate.

Therefore, the next step will be to use the ensemble precipitation analyses described in this article to force a snow model. This will allow the resulting simulated snow properties to be compared with corresponding satellite observations thanks to the memory effect of the snowpack, which ensures a strong relationship between the state of the snowpack and past snowfall (Haddjeri et al., 2023). The main challenge of this indirect evaluation of precipitation analyses will be the introduction of additional uncertainties such as the height of the rain-snow limit (Vionnet et al., 2022) and the errors of the snowpack model (Essery et al., 2013; Lafaysse et al., 2017).

Code and data availability. All data used in this study are open access and can be downloaded online : https://portail-api.meteofrance.fr/web/fr/. The various codes used to process the data, perform evaluations, apply the correction method and perfrom ensemble analyses are available on request at https://github.com/vernaym/These.

Author contributions. M.V collected all the data, carried out the various simulations and evaluations, developed the method and wrote the paper. M.L and C.A supervised every stage of this work, directed various scientific decisions and contributed to the writing of this paper.

Competing interests. No competing interests where involved in this paper.

Acknowledgements. This work was partly funded by the TRISHNA-Cryosphere project of the Centre National d'Etudes Spatiales (CNES) and the SENSASS project funded by the Region Auvergne Rhône-Alpes (France). The CNRM-CEN is part of LabEx OSUG@2020. The authors thank Dominique Faure and the Meteo France radar team for their help in understanding radar measurement issues in mountainous areas.

Appendix A: ANTILOPE climatological error estimation method illustration

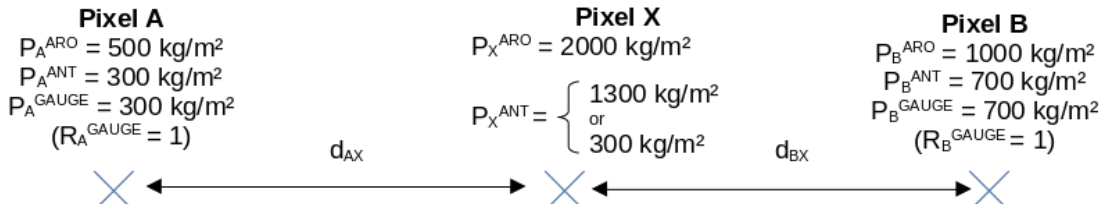

Figure A1. Illustration of the ANTILOPE climatological error estimation method on a pixel X with two available reference gauges (on pixels A and B)

This section is a numerical illustration of the estimation of the ANTILOPE climatological error on a specific pixel, designated X, based on two nearby available gauges on pixels A and B, which are situated at the same distance from pixel X ($d_{AX} = d_{BX}$ and $w_{AX} = w_{BX}$ as illustrated in Figure A1. It is assumed that the ANTILOPE product provides an unbiased estimation of precipitation accumulation for pixels A and B. This implies that the ANTILOPE yearly precipitation accumulation is the same as the one measured at the gauge : $P_A^{ANT} = P_A^{GAUGE} = 300\ \mathrm{kg\,m^{-2}}$ and $P_B^{ANT} = P_B^{GAUGE} = 700\ \mathrm{kg\,m^{-2}}$ (thus, $R_A^{GAUGE} = R_B^{GAUGE} = 1$). Furthermore, it is assumed that the AROME climatology exhibits four times more precipitation precipitation on pixel X ($P_X^{ARO} = 2000\ \mathrm{kg\,m^{-2}}$) than on pixel A ($P_A^{ARO} = 500\ \mathrm{kg\,m^{-2}}$) and twice the precipitation than on pixel B ($P_B^{ARO} = 1000\ \mathrm{kg\,m^{-2}}$).

Numerical results of the error estimation method are given for two different values of ANTILOPE precipitation accumulation on pixel X (P_X^{ANT}) and summarised in table A1. The first case ($P_X^{ANT} = 1300\ \mathrm{kg\,m^{-2}}$) illustrates a situation where the ANTILOPE is free of artifacts on pixel X. This means that the ANTILOPE yearly precipitation accumulation is consistent with the values on pixels A and B, as well as with the AROME precipitation ratios. In this situation, estimated ratios from reference pixels A and B ($\hat{R_{XA}}$ and $\hat{R_{XB}}$) are both relatively close to 1 (1.08 and 0.93 respectively). Therefore, the estimated ratio on pixel X ($\hat{R_X}$) is very close to 1 and the associated uncertainty ($U_X = 2.2\ \mathrm{kg\,m^{-2}}$) is close to the minimum value possible (2 $\mathrm{kg\,m^{-2}}$, see section 3.1.3). The second case illustrates a situation where the ANTILOPE is affected by ground clutter on pixel X. This has been identified as leading to unrealistically low precipitation accumulations (in this example, $P_X^{ANT} = 300\ \mathrm{kg\,m^{-2}}$). In this case, ratios estimated from the reference pixels A and B are both significantly lower than 1 and consistent with each other, leading to an estimated ratio $\hat{R_X} = 0.23\ \mathrm{kg\,m^{-2}}$. The associated uncertainty is consequently very high ($U_X = 48.6\ \mathrm{kg\,m^{-2}}$). The contribution of the first term of equation 8 is $(1 + K_X) \times (A_X + 1) = 31.8$, twice the contribution of the second term $(1 + K_X) \times (A_X + 1) = 16.8$. The contribution of the first term for approximately two-thirds of the ultimate uncertainty can be interpreted as a predominance of an identified systematic bias in the uncertainty estimation over diverging information brought about by the presence of gauges in pixels A and B.

In order to complete the illustration, tables A2 and A3 give the values of the different steps of the error estimation method applied to pixels A and B benefiting from gauge measurements. It is assumed that pixels A and B are 10 km apart ($d_{AB} = 10$ km and $w_{AB} = w_{BA} = 0.5$). Since the precipitation ratios between pixels A and B are different for AROME and ANTILOPE, the estimated ratios $\hat{R_{BA}}$ and $\hat{R_{AB}}$ are significantly different from 1 (1.17 and 0.86 respectively). However, the inverse distance

Table A1. Values of the different steps of the error estimation method depending on the value of the yearly ANTILOPE precipitation accumulation on pixel X (P_X^{ANT}).

P_X^{ANT} [kg m^{-2}]	$\hat{R_{XA}}$	$\hat{R_{XB}}$	M_x	S_X	K_X	$\hat{R_X}$	A_X [kg m^{-2}]	B_X [kg m^{-2}]	U_X [kg m^{-2}]
1300	1.08	0.93	1.00	0.08	0.00	1.00	0.00	0.10	2.19
300	0.25	0.21	0.23	0.02	0.98	0.25	15.11	15.46	48.60

weighting ensures that the ultimate value of the estimated ratios $\hat{R_A}$ and $\hat{R_B}$ does not significantly differ from the actual ANTILOPE/GAUGE ratios R_A^{GAUGE} and R_B^{GAUGE} (1.01 and 0.99 respectively). Similarly, the estimated uncertainty remains minimal for each pixel (3.61 kg m^{-2} and 3.62 kg m^{-2} respectively) thanks to the presence of a reference gauge with an identical climatology.

Table A2. Values of the different steps of the error estimation method applied to pixel A when $d_{AB} = 10$ km.

$\hat{R_{BA}}$	M_A	S_A	K_A	$\hat{R_A}$	A_A [kg m^{-2}]	B_A [kg m^{-2}]	U_A [kg m^{-2}]
1.17	1.06	0.08	0.24	1.01	0.23	0.96	3.61

Table A3. Values of the different steps of the error estimation method applied to pixel B when $d_{BA} = 10$ km.

$\hat{R_{AB}}$	M_B	S_B	K_B	$\hat{R_B}$	A_B [kg m^{-2}]	B_B [kg m^{-2}]	U_B [kg m^{-2}]
0.86	0.95	0.07	0.24	0.99	0.23	0.96	3.62

Appendix B: Evaluation scores

The Brier Score (BS) assesses the ability of an ensemble to forecast a threshold being exceeded. For each given event k, the predicted probability y_k of a given threshold being exceeded (the number of members forecasting the event divided by the size of the ensemble) is compared with the corresponding binary observation o_k of the threshold being exceeded, and for N events the BS is given by :

$$BS = \frac{1}{N}\sum_{k=1}^{N}(y_k - o_k)^2 \tag{B1}$$

Although the BS ~~has~~ was originally designed for ensemble systems, it is also commonly used in a deterministic context DeMaria et al. (2009); Vernay et al. (2015), where the predicted probability y_k only takes values of 0 and 1. The Brier Score ranges from 0 to 1, with 0 corresponding to a perfect score. It is computed for different precipitation thresholds from 1 kg m^{-2} to 30 kg m^{-2} to check the ability of each product to forecast a variety of precipitation events.

The Continuous Rank Probability Score (CRPS, Matheson and Winkler (1976)) is a measure of the difference between the ensemble cumulative distribution function F_y predicted for a given event and the Heaviside function centered on the associated

observation H_o :

$$CRPS = \int_{\mathbb{R}} (F_y(z) - H_o(z))^2 dz \tag{B2}$$

Unlike to the ensemble mean bias, the CRPS takes into account the error of each member of the ensemble: an unbiased ensemble with a large spread can have a higher CRPS than a slightly biased ensemble with a small spread.

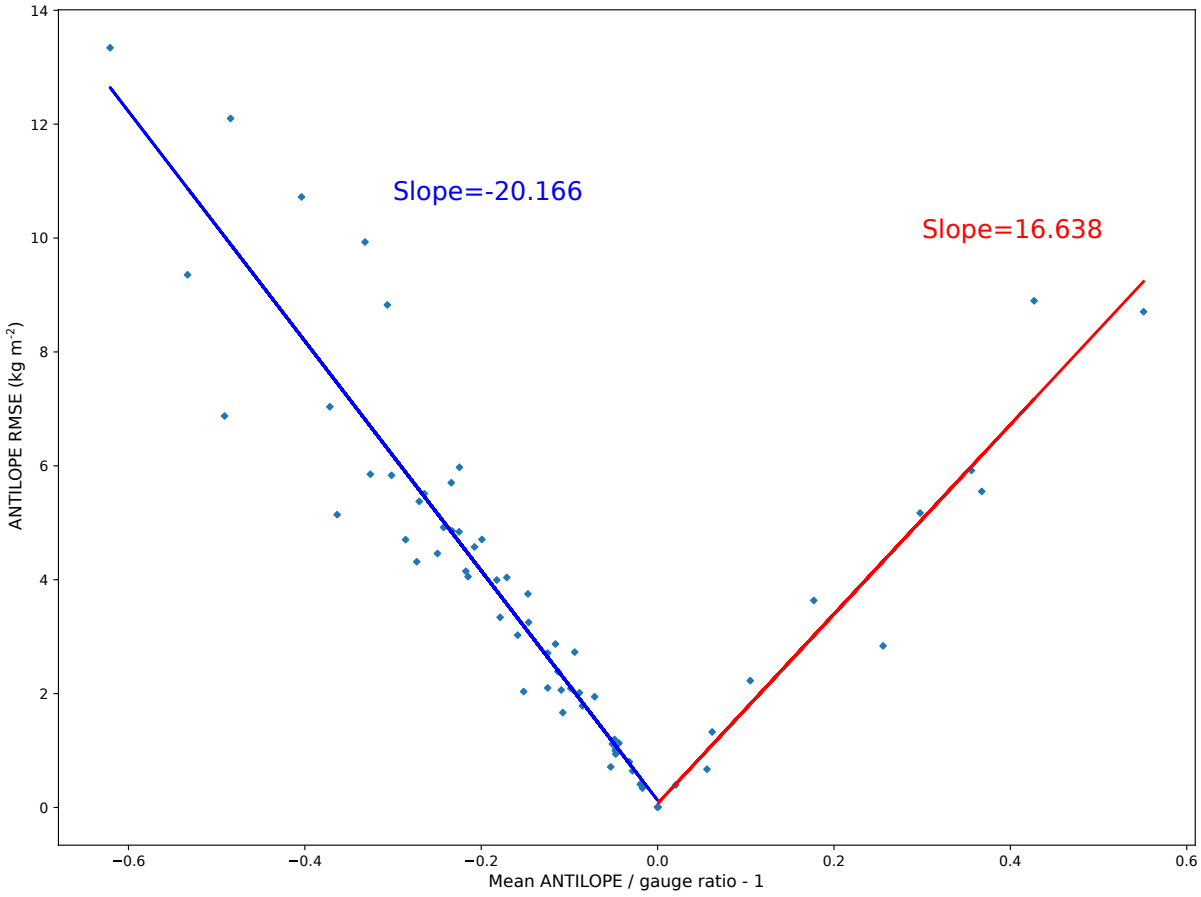

Figure D1. Linear regression between ANTILOPE / gauge mean ratio ($R^{\text{GAUGE}} - 1$) and the corresponding rmse over the 68 reference stations. The slope coefficients are used in the method described in Section 3.1.1

Appendix C: Additional results

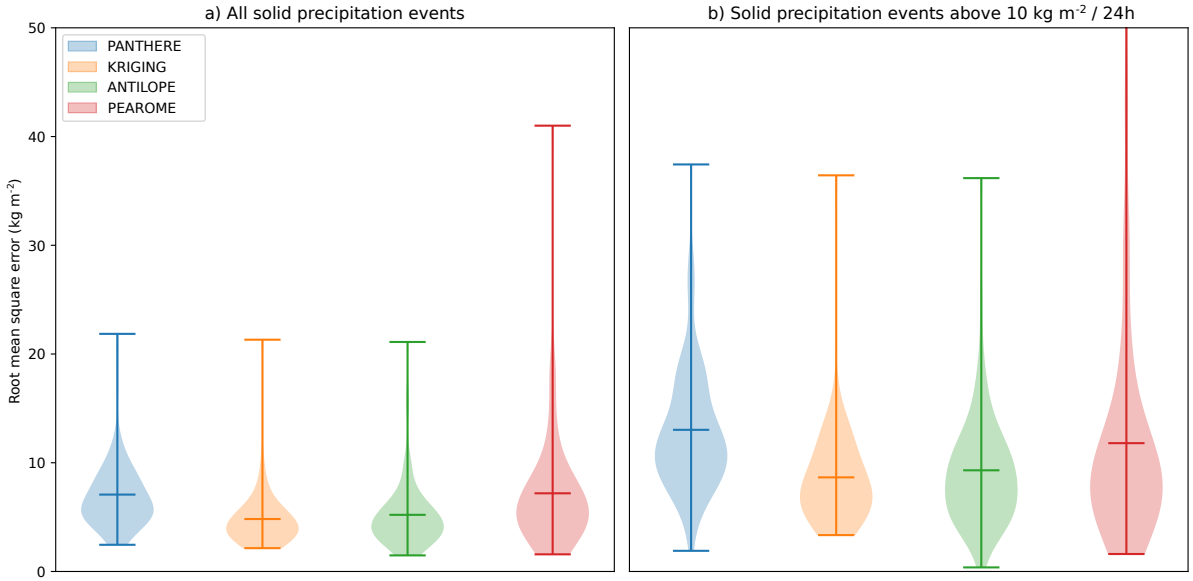

Figure E1. Same as Figure 6 but considering only snowfall events. The precipitation phase is determined by the additional ski-resort observation of the maximum altitude reached by rain during the observation period. Only situations where this maximum altitude is below the station altitude are considered here.

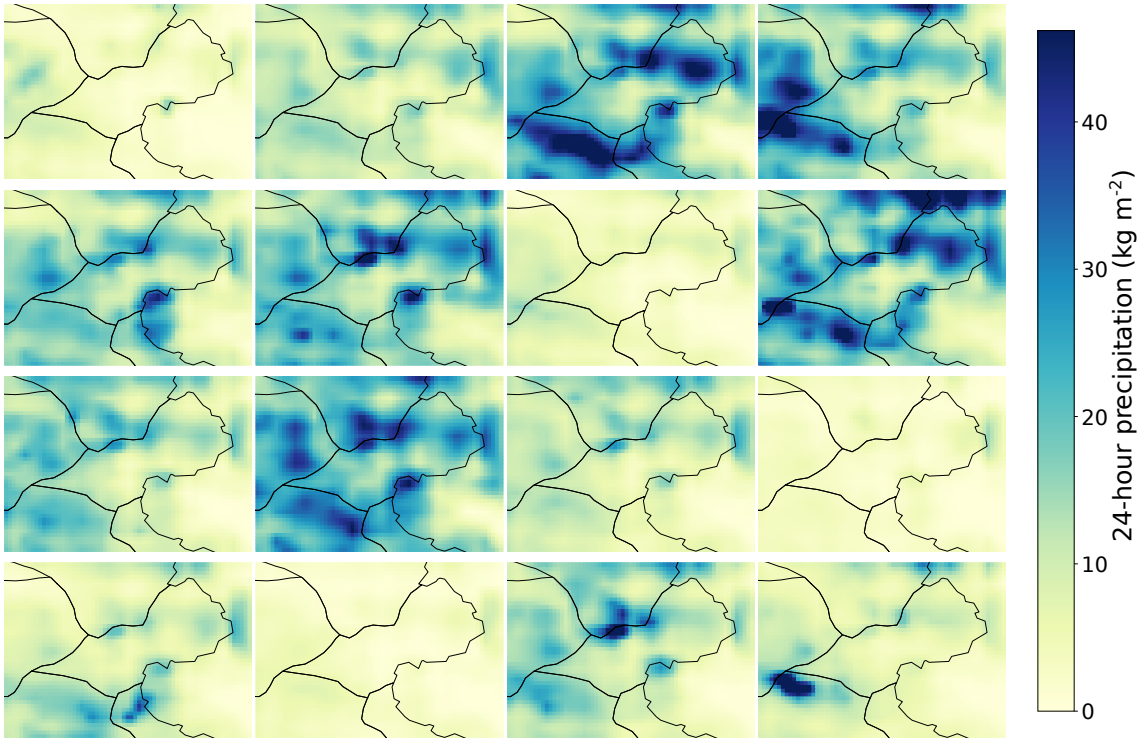

Figure G1. Raw post-processed PEAROME ensemble 24 h precipitation fields of 4 December 2021, downscaled to the ANTILOPE 1 km resolution over the Mont-Blanc area (see Figure 1). The corresponding raw ANTILOPE observation and the associated pre-processed fields are shown in Figure 11.

Appendix F: Ensemble precipitation analyses

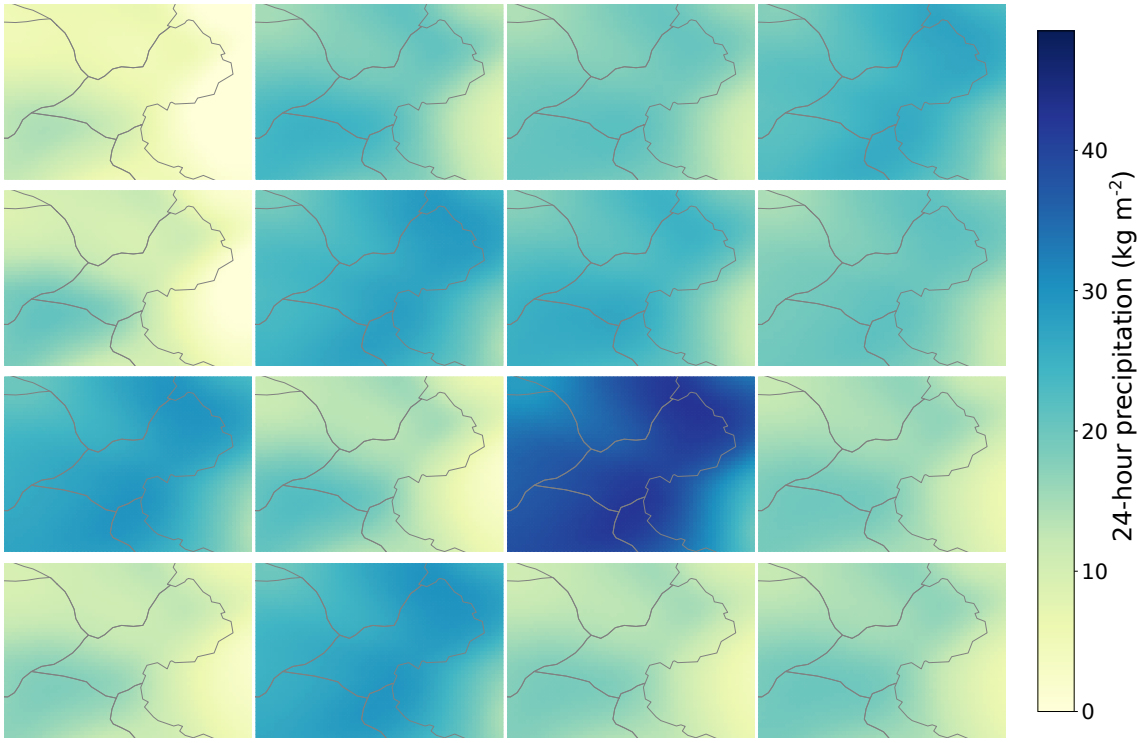

Figure H1. 24 h ensemble precipitation analysis of 4 December 2021 over the Mont-Blanc area (see Figure 1) obtained by a random sampling around the pre-processed ANTILOPE field (Section 3.2.1). The corresponding raw ANTILOPE observation and the associated pre-processed fields are shown in Figure 11.

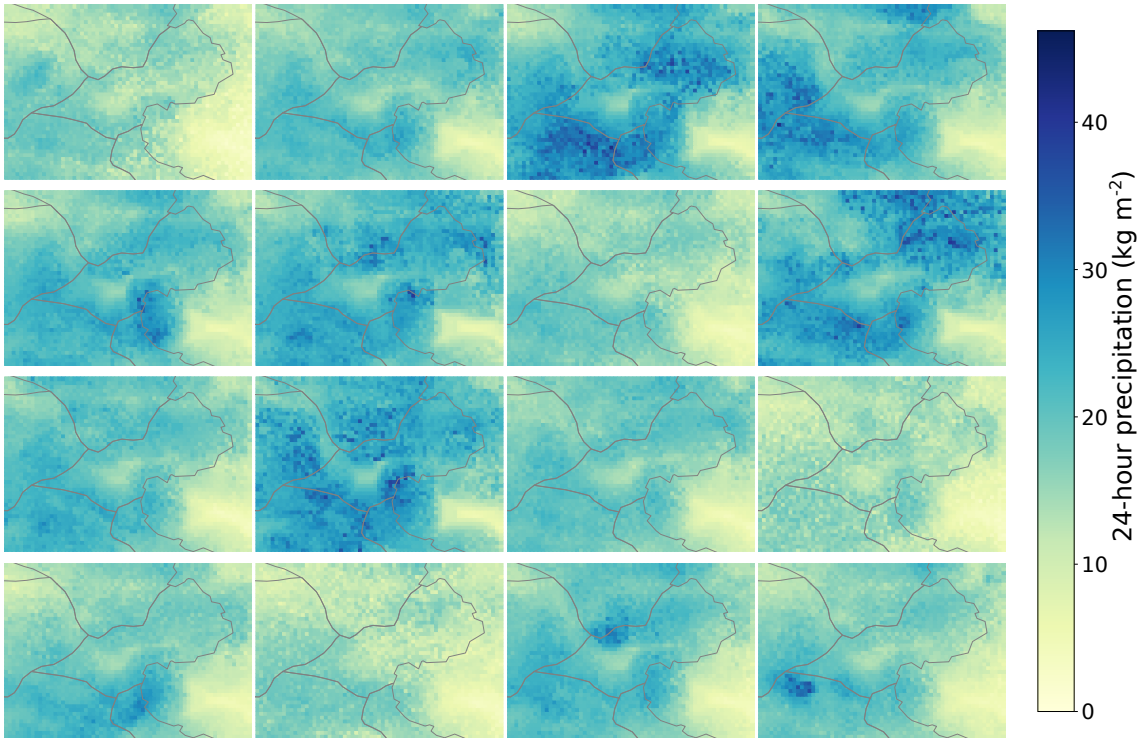

Figure I1. 24 h ensemble precipitation analysis of 4 December 2021 over the Mont-Blanc area (see Figure 1) obtained with the Particle Filter method (Section 3.2.2). The corresponding raw ANTILOPE observation and the associated pre-processed fields are shown in Figure 11 and the post-processed PEAROME ensemble is shown in Figure G1.

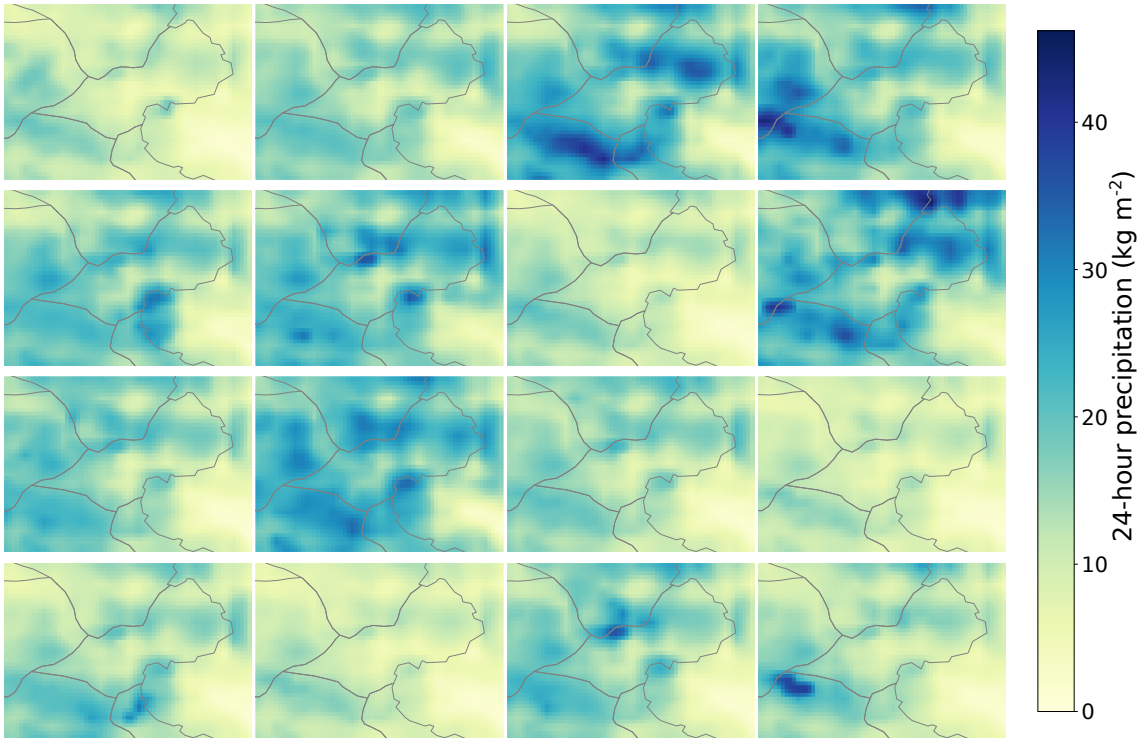

Figure J1. 24 h ensemble precipitation analysis of 4 December 2021 over the Mont-Blanc area (see Figure 1) obtained with the Ensemble Kalman Filter method (Section 3.2.3). The corresponding raw ANTILOPE observation and the associated pre-processed fields are shown in Figure 11 and the post-processed PEAROME ensemble is shown in Figure G1.

References

Atencia, A., Wang, Y., Kann, A., and Meier, F.: Localization and flow-dependency on blending techniques, Meteorologische Zeitschrift, 29, 231–246, https://doi.org/10.1127/metz/2019/0987, 2020a.

Atencia, A., Wang, Y., Kann, A., and Wastl, C.: A probabilistic precipitation nowcasting system constrained by an Ensemble Prediction System., Meteorologische Zeitschrift, 29, 183–202, https://doi.org/10.1127/metz/2020/1030, 2020b.

Awasthi, S. and Varade, D.: Recent advances in the remote sensing of alpine snow: a review, GIScience & Remote Sensing, 58, 852–888, https://doi.org/10.1080/15481603.2021.1946938, 2021.

Birman, C., Karbou, F., Mahfouf, J.-F., Lafaysse, M., Durand, Y., Giraud, G., Mérindol, L., and Hermozo, L.: Precipitation Analysis over the French Alps Using a Variational Approach and Study of Potential Added Value of Ground-Based Radar Observations, Journal of Hydrometeorology, 18, 1425–1451, https://doi.org/10.1175/JHM-D-16-0144.1, 2017.

Bouttier, F., Raynaud, L., Nuissier, O., and Ménétrier, B.: Sensitivity of the AROME ensemble to initial and surface perturbations during HyMeX, Quarterly Journal of the Royal Meteorological Society, 142, 390–403, https://doi.org/https://doi.org/10.1002/qj.2622, 2016.

Brier, G. W.: Verification of forecasts expressed in terms of probability, Monthly Weather Review, 78, 1 – 3, https://doi.org/https://doi.org/10.1175/1520-0493(1950)078<0001:VOFEIT>2.0.CO;2, 1950.

Brousseau, P., Seity, Y., Ricard, D., and Léger, J.: Improvement of the forecast of convective activity from the AROME-France system, Quarterly Journal of the Royal Meteorological Society, 142, 2231–2243, https://doi.org/https://doi.org/10.1002/qj.2822, 2016.

Candille, G. and Talagrand, O.: Evaluation of probabilistic prediction systems for a scalar variable, Quarterly Journal of the Royal Meteorological Society, 131, 2131–2150, https://doi.org/https://doi.org/10.1256/qj.04.71, 2005.

Champeaux, J.-L., Dupuy, P., Laurantin, O., Soulan, I., Tabary, P., and Soubeyroux, J.-M.: Les mesures de précipitations et l'estimation des lames d'eau à Météo-France : état de l'art et perspectives, La Houille Blanche, Numero 5, 28–34, https://doi.org/10.1051/lhb/2009052, 2009.

Clark, M. P. and Slater, A.: Probabilistic Quantitative Precipitation Estimation in Complex Terrain, Journal of Hydrometeorology, 7, 3–22, 2005.

Clark, M. P., Hendrikx, J., Slater, A. G., Kavetski, D., Anderson, B., Cullen, N. J., Kerr, T., Örn Hreinsson, E., and Woods, R. A.: Representing spatial variability of snow water equivalent in hydrologic and land-surface models: A review, Water Resources Research, 47, https://doi.org/https://doi.org/10.1029/2011WR010745, 2011.

Cluzet, B., Lafaysse, M., Cosme, E., Albergel, C., Meunier, L.-F., and Dumont, M.: CrocO_v1.0: a particle filter to assimilate snowpack observations in a spatialised framework, Geoscientific Model Development, 14, 1595–1614, https://doi.org/10.5194/gmd-14-1595-2021, 2021.

Cluzet, B., Lafaysse, M., Deschamps-Berger, C., Vernay, M., and Dumont, M.: Propagating information from snow observations with CrocO ensemble data assimilation system: a 10-years case study over a snow depth observation network, The Cryosphere, 16, 1281–1298, https://doi.org/10.5194/tc-16-1281-2022, 2022.

Dai, Q., Han, D., Rico-Ramirez, M., and Srivastava, P. K.: Multivariate distributed ensemble generator: A new scheme for ensemble radar precipitation estimation over temperate maritime climate, Journal of Hydrology, 511, 17–27, https://doi.org/https://doi.org/10.1016/j.jhydrol.2014.01.016, 2014.

DeMaria, M., Knaff, J. A., Knabb, R., Lauer, C., Sampson, C. R., and DeMaria, R. T.: A New Method for Estimating Tropical Cyclone Wind Speed Probabilities, Weather and Forecasting, 24, 1573 – 1591, https://doi.org/https://doi.org/10.1175/2009WAF2222286.1, 2009.

Deschamps-Berger, C., Gascoin, S., Berthier, E., Deems, J., Gutmann, E., Dehecq, A., Shean, D., and Dumont, M.: Snow depth mapping from stereo satellite imagery in mountainous terrain: evaluation using airborne laser-scanning data, The Cryosphere, 14, 2925–2940, https://doi.org/10.5194/tc-14-2925-2020, 2020.

Deschamps-Berger, C., Cluzet, B., Dumont, M., Lafaysse, M., Berthier, E., Fanise, P., and Gascoin, S.: Improving the Spatial Distribution of Snow Cover Simulations by Assimilation of Satellite Stereoscopic Imagery, Water Resources Research, 58, e2021WR030271, https://doi.org/https://doi.org/10.1029/2021WR030271, e2021WR030271 2021WR030271, 2022.

Durand, Y., Brun, E., Mérindol, L., Guyomarc'h, G., Lesaffre, B., and Martin, E.: A meteorological estimation of relevant parameters for snow models, Ann. Glaciol., 18, 65–71, 1993.

Essery, R., Morin, S., Lejeune, Y., and B Ménard, C.: A comparison of 1701 snow models using observations from an alpine site, Advances in Water Resources, 55, 131–148, https://doi.org/https://doi.org/10.1016/j.advwatres.2012.07.013, snow–Atmosphere Interactions and Hydrological Consequences, 2013.

Evensen, G.: The Ensemble Kalman Filter: Theoretical Formulation and Practical Implementation, Ocean Dynamics, 53, 343–367, https://doi.org/10.1007/s10236-003-0036-9, 2003.

Faure, D., Gaussiat, N., Dupuy, P., Delrieu, G., Yu, N., and Sarter, F.: Quality analysis of the 2016 quantitative precipitation estimates in the French Alps, 2017.

Faure, D., Delrieu, G., and Gaussiat, N.: Impact of the Altitudinal Gradients of Precipitation on the Radar QPE Bias in the French Alps, Atmosphere, 10, https://doi.org/10.3390/atmos10060306, 2019.

Fierz, C., R.L, A., Y, D., P, E., Greene, E., Mcclung, D., Nishimura, K., Satyawali, P., and Sokratov, S.: The international classification for seasonal snow on the ground (UNESCO, IHP (International Hydrological Programme)–VII, Technical Documents in Hydrology, No 83; IACS (International Association of Cryospheric Sciences) contribution No 1), 2009.

Figueras i Ventura, J. and Tabary, P.: The New French Operational Polarimetric Radar Rainfall Rate Product, Journal of Applied Meteorology and Climatology, 52, 1817–1835, https://doi.org/10.1175/JAMC-D-12-0179.1, 2013.

Foehn, A., Hernandez, J. G., Schaefli, B., and De Cesare, G.: Spatial interpolation of precipitation from multiple rain gauge networks and weather radar data for operational applications in Alpine catchments, JOURNAL OF HYDROLOGY, 563, 1092–1110, https://doi.org/10.1016/j.jhydrol.2018.05.027, 2018.

Foresti, L., Kanevski, M. F., and Pozdnoukhov, A.: Kernel-Based Mapping of Orographic Rainfall Enhancement in the Swiss Alps as Detected by Weather Radar, IEEE Transactions on Geoscience and Remote Sensing, 50, 2954–2967, 2012.

Foresti, L., Sideris, I. V., Panziera, L., Nerini, D., and Germann, U.: A 10-year radar-based analysis of orographic precipitation growth and decay patterns over the Swiss Alpine region, Quarterly Journal of the Royal Meteorological Society, 144, 2277–2301, https://doi.org/https://doi.org/10.1002/qj.3364, 2018.

Fortin, V., Roy, G., Donaldson, N., and Mahidjiba, A.: Assimilation of radar quantitative precipitation estimations in the Canadian Precipitation Analysis (CaPA), Journal of Hydrology, 531, 296–307, https://doi.org/https://doi.org/10.1016/j.jhydrol.2015.08.003, hydrologic Applications of Weather Radar, 2015.

Fortin, V., Roy, G., Stadnyk, T., Koenig, K., Gasset, N., and Mahidjiba, A.: Ten years of science based on the Canadian Precipitation Analysis: A CaPA system overview and literature review, Atmosphere-Ocean, 56, 178–196, https://doi.org/10.1080/07055900.2018.1474728, 2018.

Frei, C. and Isotta, F. A.: Ensemble Spatial Precipitation Analysis From Rain Gauge Data: Methodology and Application in the European Alps, Journal of Geophysical Research: Atmospheres, 124, 5757 – 5778, 2019.

Germann, U., Berenguer, M., Sempere-Torres, D., and Zappa, M.: REAL—Ensemble radar precipitation estimation for hydrology in a mountainous region, Quarterly Journal of the Royal Meteorological Society, 135, 445–456, https://doi.org/https://doi.org/10.1002/qj.375, 2009.

Germann, U., Boscacci, M., Clementi, L., Gabella, M., Hering, A., Sartori, M., Sideris, I. V., and Calpini, B.: Weather Radar in Complex Orography, Remote Sensing, 14, https://doi.org/10.3390/rs14030503, 2022.

Ghaemi, E., Gabella, M., Foelsche, U., Sideris, I., and Nerini, D.: The effect of altitude on the uncertainty of radar-based precipitation estimates over Switzerland, International journal of remote sensing, 44, 2495–2517, https://doi.org/10.1080/01431161.2023.2203339, 2023.

Haddjeri, A., Baron, M., Lafaysse, M., Le Toumelin, L., Deschamp-Berger, C., Vionnet, V., Gascoin, S., Vernay, M., and Dumont, M.: Exploring the sensitivity to precipitation, blowing snow, and horizontal resolution of the spatial distribution of simulated snow cover, EGUsphere, 2023, https://doi.org/10.5194/egusphere-2023-2604, 2023.

Haiden, T., Kann, A., Wittmann, C., Pistotnik, G., Bica, B., and Gruber, C.: The Integrated Nowcasting through Comprehensive Analysis (INCA) System and Its Validation over the Eastern Alpine Region, Weather and Forecasting, 26, 166–183, https://doi.org/10.1175/2010WAF2222451.1, 2011.

Hopson, T. M.: Assessing the Ensemble Spread–Error Relationship, Monthly Weather Review, 142, 1125 – 1142, https://doi.org/https://doi.org/10.1175/MWR-D-12-00111.1, 2014.

Hotta, D. and Ota, Y.: Why does EnKF suffer from analysis overconfidence? An insight into exploiting the ever-increasing volume of observations, Quarterly Journal of the Royal Meteorological Society, 147, 1258–1277, https://doi.org/https://doi.org/10.1002/qj.3970, 2021.

IPCC: High Mountain Areas, p. 131–202, Cambridge University Press, https://doi.org/10.1017/9781009157964.004, 2022.

Kalman, R. E.: A New Approach to Linear Filtering and Prediction Problems, Journal of Basic Engineering, 82, 35–45, https://doi.org/10.1115/1.3662552, 1960.

Khedhaouiria, D., Bélair, S., Fortin, V., Roy, G., and Lespinas, F.: Using a hybrid optimal interpolation–ensemble Kalman filter for the Canadian Precipitation Analysis, Nonlinear Processes in Geophysics, 29, 329–344, https://doi.org/10.5194/npg-29-329-2022, 2022.

Kirstetter, P.-E., Delrieu, G., Boudevillain, B., and Obled, C.: Toward an error model for radar quantitative precipitation estimation in the Cévennes–Vivarais region, France, Journal of Hydrology, 394, 28–41, https://doi.org/https://doi.org/10.1016/j.jhydrol.2010.01.009, flash Floods: Observations and Analysis of Hydrometeorological Controls, 2010.

Kirstetter, P.-E., Gourley, J. J., Hong, Y., Zhang, J., Moazamigoodarzi, S., Langston, C., and Arthur, A.: Probabilistic precipitation rate estimates with ground-based radar networks, Water Resources Research, 51, 1422–1442, https://doi.org/https://doi.org/10.1002/2014WR015672, 2015.

Kochendorfer, J., Earle, M. E., Hodyss, D., Reverdin, A., Roulet, Y.-A., Nitu, R., Rasmussen, R., Landolt, S., Buisán, S., and Laine, T.: Undercatch Adjustments for Tipping-Bucket Gauge Measurements of Solid Precipitation, Journal of Hydrometeorology, 21, 1193 – 1205, https://doi.org/10.1175/JHM-D-19-0256.1, https://journals.ametsoc.org/view/journals/hydr/21/6/jhm-d-19-0256.1.xml, 2020.

Krinner, G., Derksen, C., Essery, R., Flanner, M., Hagemann, S., Clark, M., Hall, A., Rott, H., Brutel-Vuilmet, C., Kim, H., Ménard, C. B., Mudryk, L., Thackeray, C., Wang, L., Arduini, G., Balsamo, G., Bartlett, P., Boike, J., Boone, A., Chéruy, F., Colin, J., Cuntz, M., Dai, Y., Decharme, B., Derry, J., Ducharne, A., Dutra, E., Fang, X., Fierz, C., Ghattas, J., Gusev, Y., Haverd, V., Kontu, A., Lafaysse, M., Law, R., Lawrence, D., Li, W., Marke, T., Marks, D., Ménégoz, M., Nasonova, O., Nitta, T., Niwano, M., Pomeroy, J., Raleigh, M. S., Schaedler, G., Semenov, V., Smirnova, T. G., Stacke, T., Strasser, U., Svenson, S., Turkov, D., Wang, T., Wever, N., Yuan, H., Zhou, W.,

and Zhu, D.: ESM-SnowMIP: assessing snow models and quantifying snow-related climate feedbacks, Geoscientific Model Development, 11, 5027–5049, https://doi.org/10.5194/gmd-11-5027-2018, 2018.

Lafaysse, M., Cluzet, B., Dumont, M., Lejeune, Y., Vionnet, V., and Morin, S.: A multiphysical ensemble system of numerical snow modelling, The Cryosphere, 11, 1173–1198, https://doi.org/10.5194/tc-11-1173-2017, 2017.

Largeron, C., Dumont, M., Morin, S., Boone, A., Lafaysse, M., Metref, S., Cosme, E., Jonas, T., Winstral, A., and Margulis, S. A.: Toward Snow Cover Estimation in Mountainous Areas Using Modern Data Assimilation Methods: A Review, Frontiers in Earth Science, 8, 325, https://doi.org/10.3389/feart.2020.00325, 2020.

Le Bastard, T., Caumont, O., Gaussiat, N., and Karbou, F.: Combined use of volume radar observations and high-resolution numerical weather predictions to estimate precipitation at the ground: methodology and proof of concept, ATMOSPHERIC MEASUREMENT TECHNIQUES, 12, 5669–5684, https://doi.org/10.5194/amt-12-5669-2019, 2019.

Lespinas, F., Fortin, V., Roy, G., Rasmussen, P., and Stadnyk, T.: Performance Evaluation of the Canadian Precipitation Analysis (CaPA), Journal of Hydrometeorology, 16, 2045 – 2064, https://doi.org/10.1175/JHM-D-14-0191.1, 2015.

Lundquist, J., Hughes, M., Gutmann, E., and Kapnick, S.: Our Skill in Modeling Mountain Rain and Snow is Bypassing the Skill of Our Observational Networks, BULLETIN OF THE AMERICAN METEOROLOGICAL SOCIETY, 100, 2473–2490, https://doi.org/10.1175/BAMS-D-19-0001.1, 2019.

Magnusson, J., Winstral, A., Stordal, A. S., Essery, R., and Jonas, T.: Improving physically based snow simulations by assimilating snow depths using the particle filter, Water Resources Research, 53, 1125–1143, https://doi.org/https://doi.org/10.1002/2016WR019092, 2017.

Mandapaka, P. and Germann, U.: Radar-Rainfall Error Models and Ensemble Generators, vol. 191, pp. 247–264, https://doi.org/10.1029/2010GM001003, 2010.

Marshall, J. S. and Palmer, W. M. K.: THE DISTRIBUTION OF RAINDROPS WITH SIZE, Journal of Atmospheric Sciences, 5, 165 – 166, https://doi.org/https://doi.org/10.1175/1520-0469(1948)005<0165:TDORWS>2.0.CO;2, 1948.

Matheson, J. E. and Winkler, R. L.: Scoring Rules for Continuous Probability Distributions, Management Science, 22, 1087–1096, https://doi.org/10.1287/mnsc.22.10.1087, 1976.

McRoberts, D. B. and Nielsen-Gammon, J. W.: Detecting Beam Blockage in Radar-Based Precipitation Estimates, Journal of Atmospheric and Oceanic Technology, 34, 1407 – 1422, https://doi.org/10.1175/JTECH-D-16-0174.1, 2017.

Morin, S., Horton, S., Techel, F., Bavay, M., Coléou, C., Fierz, C., Gobiet, A., Hagenmuller, P., Lafaysse, M., Ližar, M., Mitterer, C., Monti, F., Müller, K., Olefs, M., Snook, J. S., van Herwijnen, A., and Vionnet, V.: Application of physical snowpack models in support of operational avalanche hazard forecasting: A status report on current implementations and prospects for the future, Cold Regions Science and Technology, 170, 102 910, https://doi.org/https://doi.org/10.1016/j.coldregions.2019.102910, 2020.

Mott, R., Winstral, A., Cluzet, B., Helbig, N., Magnusson, J., Mazzotti, G., Quéno, L., Schirmer, M., Webster, C., and Jonas, T.: Operational snow-hydrological modeling for Switzerland, Frontiers in Earth Science, 11, https://doi.org/10.3389/feart.2023.1228158, 2023.

Nerini, D., Foresti, L., Leuenberger, D., Robert, S., and Germann, U.: A Reduced-Space Ensemble Kalman Filter Approach for Flow-Dependent Integration of Radar Extrapolation Nowcasts and NWP Precipitation Ensembles, Monthly Weather Review, 147, 987 – 1006, https://doi.org/10.1175/MWR-D-18-0258.1, 2019.

Olefs, M., Schöner, W., Suklitsch, M., Wittmann, C., Niedermoser, B., Neururer, A., and Wurzer, A.: SNOWGRID – A New Operational Snow Cover Model in Austria, 2013.

Rasmussen, R., 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, Bulletin of the American Meteorological Society, 93, 811 – 829, https://doi.org/10.1175/BAMS-D-11-00052.1, https://journals.ametsoc.org/view/journals/bams/93/6/bams-d-11-00052.1.xml, 2012.

Schefzik, R., Thorarinsdottir, T. L., and Gneiting, T.: Uncertainty Quantification in Complex Simulation Models Using Ensemble Copula Coupling, Statistical Science, 28, 616 – 640, https://doi.org/10.1214/13-STS443, 2013.

Schirmer, M. and Jamieson, B.: Verification of analysed and forecasted winter precipitation in complex terrain, The Cryosphere, 9, 587–601, https://doi.org/10.5194/tc-9-587-2015, 2015.

Seity, Y., Brousseau, P., Malardel, S., Hello, G., Bénard, P., Bouttier, F., Lac, C., and Masson, V.: The AROME-France Convective-Scale Operational Model, Monthly Weather Review, 139, 976 – 991, https://doi.org/10.1175/2010MWR3425.1, 2011.

Sideris, I. V., Gabella, M., Erdin, R., and Germann, U.: Real-time radar–rain-gauge merging using spatio-temporal co-kriging with external drift in the alpine terrain of Switzerland, Quarterly Journal of the Royal Meteorological Society, 140, 1097–1111, https://doi.org/https://doi.org/10.1002/qj.2188, 2014.

Sideris, I. V., Foresti, L., Nerini, D., and Germann, U.: NowPrecip: localized precipitation nowcasting in the complex terrain of Switzerland, Quarterly Journal of the Royal Meteorological Society, 146, 1768–1800, https://doi.org/https://doi.org/10.1002/qj.3766, 2020.

Sivasubramaniam, K., Sharma, A., and Alfredsen, K.: Merging radar and gauge information within a dynamical model combination framework for precipitation estimation in cold climates, Environmental Modelling & Software, 119, 99–110, https://doi.org/https://doi.org/10.1016/j.envsoft.2019.05.013, 2019.

Snyder, C., Bengtsson, T., Bickel, P., and Anderson, J.: Obstacles to High-Dimensional Particle Filtering, Monthly Weather Review, 136, 4629 – 4640, https://doi.org/https://doi.org/10.1175/2008MWR2529.1, 2008.

Soci, C., Bazile, E., Besson, F., and Landelius, T.: High-resolution precipitation re-analysis system for climatological purposes, Tellus A: Dynamic Meteorology and Oceanography, 68, 29 879, https://doi.org/10.3402/tellusa.v68.29879, 2016.

Tabary, P.: The New French Operational Radar Rainfall Product. Part I: Methodology, Weather and Forecasting, 22, 393 – 408, https://doi.org/10.1175/WAF1004.1, 2007.

Taillardat, M. and Mestre, O.: From research to applications – examples of operational ensemble post-processing in France using machine learning, Nonlinear Processes in Geophysics, 27, 329–347, https://doi.org/10.5194/npg-27-329-2020, 2020.

Taillardat, M., Fougères, A.-L., Naveau, P., and Mestre, O.: Forest-Based and Semiparametric Methods for the Postprocessing of Rainfall Ensemble Forecasting, Weather and Forecasting, 34, 617 – 634, https://doi.org/10.1175/WAF-D-18-0149.1, 2019.

Thornton, J. M., Pepin, N., Shahgedanova, M., and Adler, C.: Coverage of In Situ Climatological Observations in the World's Mountains, Frontiers in Climate, 4, https://doi.org/10.3389/fclim.2022.814181, 2022.

van Leeuwen, P. J.: Particle Filtering in Geophysical Systems, Monthly Weather Review, 137, 4089 – 4114, https://doi.org/https://doi.org/10.1175/2009MWR2835.1, 2009.

Vernay, M., Lafaysse, M., Mérindol, L., Giraud, G., and Morin, S.: Ensemble forecasting of snowpack conditions and avalanche hazard, Cold Regions Science and Technology, 120, https://doi.org/10.1016/j.coldregions.2015.04.010, 2015.

Villarini, G., Seo, B.-C., Serinaldi, F., and Krajewski, W. F.: Spatial and temporal modeling of radar rainfall uncertainties, Atmospheric Research, 135-136, 91–101, https://doi.org/https://doi.org/10.1016/j.atmosres.2013.09.007, 2014.

Vionnet, V., Dombrowski-Etchevers, I., Lafaysse, M., Quéno, L., Seity, Y., and Bazile, E.: Numerical Weather Forecasts at Kilometer Scale in the French Alps: Evaluation and Application for Snowpack Modeling, Journal of Hydrometeorology, 17, https://doi.org/10.1175/JHM-D-15-0241.1, 2016.

Vionnet, V., Six, D., Auger, L., Dumont, M., Lafaysse, M., Quéno, L., Réveillet, M., Dombrowski-Etchevers, I., Thibert, E., and Vincent, C.: Sub-kilometer Precipitation Datasets for Snowpack and Glacier Modeling in Alpine Terrain, Frontiers in Earth Science, 7, https://doi.org/10.3389/feart.2019.00182, 2019.

Vionnet, V., Verville, M., Fortin, V., Brugman, M., Abrahamowicz, M., Lemay, F., Thériault, J. M., Lafaysse, M., and Milbrandt, J. A.: Snow level from post-processing of atmospheric model improves snowfall estimate and snowpack prediction in mountains, Water Resources Research, 58, e2021WR031 778, https://doi.org/https://doi.org/10.1029/2021WR031778, e2021WR031778 2021WR031778, 2022.

Vogl, S., Laux, P., Qiu, W., Mao, G., and Kunstmann, H.: Copula-based assimilation of radar and gauge information to derive bias-corrected precipitation fields, Hydrology and Earth System Sciences, 16, 2311–2328, https://doi.org/10.5194/hess-16-2311-2012, 2012.

Yu, N., Gaussiat, N., and Tabary, P.: Polarimetric X-band weather radars for quantitative precipitation estimation in mountainous regions, QUARTERLY JOURNAL OF THE ROYAL METEOROLOGICAL SOCIETY, 144, 2603–2619, https://doi.org/10.1002/qj.3366, 2018.