Response to Reviewer

Journal: Earth System Science Data

Manuscript: essd-2026-467, "A 30-Year 1-km Daily Precipitation and Air Temperature Dataset for the Po River District (Italy)"

Authors: H. Salehi et al.

Dear Editor, dear Reviewer,

Thank you for the detailed and constructive review. We address each point below. The three main concerns were well founded. Rather than only adding new analyses, we rebuilt the dataset and released it as version 2.0. We refer to the submitted dataset as version 1.0.

The revised dataset uses the climatology-aided (anomaly) approach recommended by the reviewer for both variables:

- for temperature, daily anomalies are differences from PRISM-type normals;
- for precipitation, the anomaly approach is applied separately to wet-day occurrence and to wet-day amount (a two-part scheme).

We explain the two-part scheme under Comment 2.1. We first tested an anomaly method based on a single ratio of the daily total to its normal, as in APGD. Like the submitted version, this approach spreads precipitation over too many days and underestimates wet-day intensity. Separating occurrence and amount removes both biases.

We reproduce the reviewer's comments in full, in grey italics, and number them for reference (Comment 1.1, 1.2, …). Our response follows each comment. Section, figure and table numbers refer to the revised manuscript. Figures and tables numbered R1, R2, … are included in this letter; their counterparts in the revised manuscript are given in the captions. A summary at the end of the letter lists each change and where it appears in the revised manuscript.

General comments

Comment 0

The paper "A 30-Year 1-km Daily Precipitation and Air Temperature Dataset for the Po River District (Italy)" by H. Salehi et al. presents new high-resolution daily fields for precipitation and temperature for the Po River District for the 30-year period 1991-2020.

The topic addressed by the paper is very important because there is strong needing of high-resolution fields of meteorological variables for many applications. However, the paper has three fundamental deficits that, in my opinion, make it not acceptable for publication. The first concerns the actual spatial resolution of the temperature and precipitation fields, the second concerns the adopted methodology that does not allow sufficient robustness to missing data, the third concerns the lack of comparisons with the best data sets available for the study area.

Response. Thank you for identifying the three main problems so clearly. We agree with all three points. We addressed them as follows:

1. **Anomaly (climatology-aided) interpolation.** Daily fields are obtained by combining 1991–2020 monthly normals with interpolated daily anomalies (Sect. 4.1–4.3).

- **Normals.** They are derived with a PRISM-type local weighted regression on the station normals. For precipitation there are two monthly normals: total precipitation and wet-day frequency (≥ 1 mm). The normal intensity of wet days follows from the two.
- **Temperature.** Anomalies are differences from the normal and are added back to it.
- **Precipitation.** The anomaly of wet-day occurrence and the ratio of the wet-day amount to its normal intensity are interpolated separately and then recombined (two-part scheme; Comment 2.1).

2. **Local instead of domain-wide relations.** A spherical variogram is fitted for every grid cell and day from the 50 nearest active stations, and ordinary kriging uses the 7 nearest stations. The temperature lapse rate is now estimated locally and daily. The single domain-wide lapse rate of v1.0 has been abandoned.
3. **Weight-based spatial support**, used as a proxy for effective resolution, is quantified from the kriging weights as suggested. We computed the distances at which the weights fall to 2/3, 1/2 and 1/3 of the largest weight, by region and by decade (Sect. 5.2 and 6.2). Per-cell maps of these support distances are distributed with the dataset.
4. **Robustness to missing data** is assessed in two ways (Sect. 5.3 and 6.3):
   - two complementary 50 % deletion experiments (30 realisations each): temporal missingness (station-days removed, normals kept), with a four-year check for precipitation, and network sparsity (stations removed, normals rebuilt);
   - the Aosta–Gran San Bernardo point test proposed by the reviewer.
5. **Comparison with the regional reference datasets** named by the reviewer (Sect. 5.4 and 6.4): APGD (Isotta et al., 2014; version extended to 2019), ARCIS precipitation and temperature (Pavan et al., 2019, 2026), Manara et al. (2026; IH-GAR), Crespi et al. (2018) and Brunetti et al. (2014). The comparison now also decomposes precipitation into wet-day frequency and intensity.

While preparing the revision we also corrected problems in the station data:

6. **Day-definition harmonisation.** A large part of the station records were labelled with the day on which the 24-h accumulation ends rather than the day it covers. The accumulation windows themselves are 00–24 UTC for the Italian regional networks and EEAR-Clim. The labels were harmonised so that each value is stored under the calendar day it covers (Sect. 3.2).
7. **Quality control and network.** The station-day quality control was redone as a fully automatic and reproducible procedure (Sect. 3.3). The official MeteoSwiss stations within 20 km of the District, including Col du Grand St-Bernard, were added (Sect. 2.2).
8. **Calibrated uncertainty.** Each daily value is now distributed with a calibrated 1σ uncertainty (Sect. 4.4).

Compared with the submitted version, v2.0 better represents precipitation climatology, wet-day occurrence, wet-day intensity, and temperature. It also provides uncertainty estimates and documents the spatial support of each value. We also state the remaining trade-offs clearly, including the lower probability of detecting wet days and the greater sensitivity of the daily wet/dry pattern to station loss (Comments 2.1 and 2.3). The remaining limitations are described in the abstract, Sect. 7 (Usage notes and limitations), and the dataset README.

1. Actual spatial resolution of the temperature and precipitation fields

Comment 1.1

The nominal resolution of the produced dataset is 1 km, but the actual resolution of it is not clear from the paper. The number of available stations within the considered area is between 1100 and 1200, but considering

that about half of the data are missing, the interpolation methods works, on average, with fewer than 600 stations. Considering that the area has an extension of 90000 km2, this means that there is about 1 station per 150 km2, which corresponds to an actual resolution greater than 10 km. However, in my opinion this is an optimistic assumption because for precipitation the actual resolution is probably further reduced by the smoothing introduced by the interpolation method. For temperature, the use of elevation helps better capturing the local values, even if deriving only one lapse rate for all points and elevations is probably not the best possible choice.

Response. The reviewer is right that the effective resolution was not documented clearly, and the estimate is reasonable. We added a section on effective resolution: Sect. 5.2 describes the method and Sect. 6.2 gives the results (see Comment 1.3 below).

We addressed both points:

- **Network density.** The revised network comprises 1636 precipitation and 1570 temperature stations. This includes stations in a 20 km buffer around the District,56 precipitation and 15 temperature MeteoSwiss stations.
    - The median record length of the precipitation stations is 16.7 years of valid data, and 388 stations have fewer than 10 years.
    - The mean number of stations reporting per day (all network stations) grows from 369 in 1991 to 1203 in 2020.
    - All network stations enter the daily interpolation. 1375 precipitation and 1221 temperature stations have a complete set of station normals; the others use the gridded normal at their location in the missing months (Comment 2.1).
    - Leave-one-out, station-wise scores (Table R5): leave-one-out values are computed at all 1636 stations. The station-wise statistics use the 1627 stations with at least 365 days with observations; the other nine stations have shorter records (104–348 days).
    - Leave-one-out, comparison of versions (Table R4): the 1580 stations shared with the network of the submitted version, for which leave-one-out values exist in all three versions (v1.0, the single-ratio anomaly version tested during the revision, and v2.0). The 56 MeteoSwiss stations were added in this revision, so v1.0 has no values there. All three versions are evaluated on the same 9 249 370 station-days.
    - Data availability is shown in Fig. 2 and by elevation in Fig. 3, and a new panel shows the distribution of record lengths.
- **Domain-wide lapse rate.** The single lapse rate of v1.0 has been replaced by a lapse rate estimated locally for every cell and day, together with PRISM-type normals (see our response to Comment 1.7).

Comment 1.2

The rather low actual resolution of the dataset is clearly documented by the first panel of figure 8, reporting the mean annual precipitation over the considered area. This panel shows that areas characterized by very strong spatial gradients in rainfall climatologies (for example the Aosta Valley region or the Valtellina area - see Crespi et al., 2018) present spatial distributions strongly smoothed by the interpolation methodology. Regarding this panel, I am also very surprised by the average annual rainfall indicated, which is 559.7 mm. This is about half the value suggested in the scientific literature for the Po River basin! A so low value of the average annual rainfall could not sustain the observed discharge of the Po river (about 1500 m3/s at the Pontelagoscuro station).

Response.

Mean annual precipitation. The reviewer is right: 559.7 mm is implausible. The error came from the script used to make Fig. 8, not from the dataset itself.

- The district-mean annual precipitation is 1036 mm in the submitted dataset (v1.0) and 1051 mm in the revised dataset (v2.0).
- For comparison, Manara et al. (2026) give 1048 mm on the same cells, and Crespi et al. (2018) give 1044 mm.
- A mean discharge of about 1500 $m^3 s^{-1}$ at Pontelagoscuro corresponds to about 680 mm yr^{-1} of runoff over the ~70 000 km^2 basin upstream of the gauge. This cannot be sustained by 560 mm yr^{-1} of precipitation, but it is consistent with the values above. The District (about 86 700 km^2; 88 074 cells of the 1-km analysis grid) is larger than this basin, so the two numbers are not used to compute a runoff ratio.

The figure has been corrected (Fig. 13; Fig. R1).

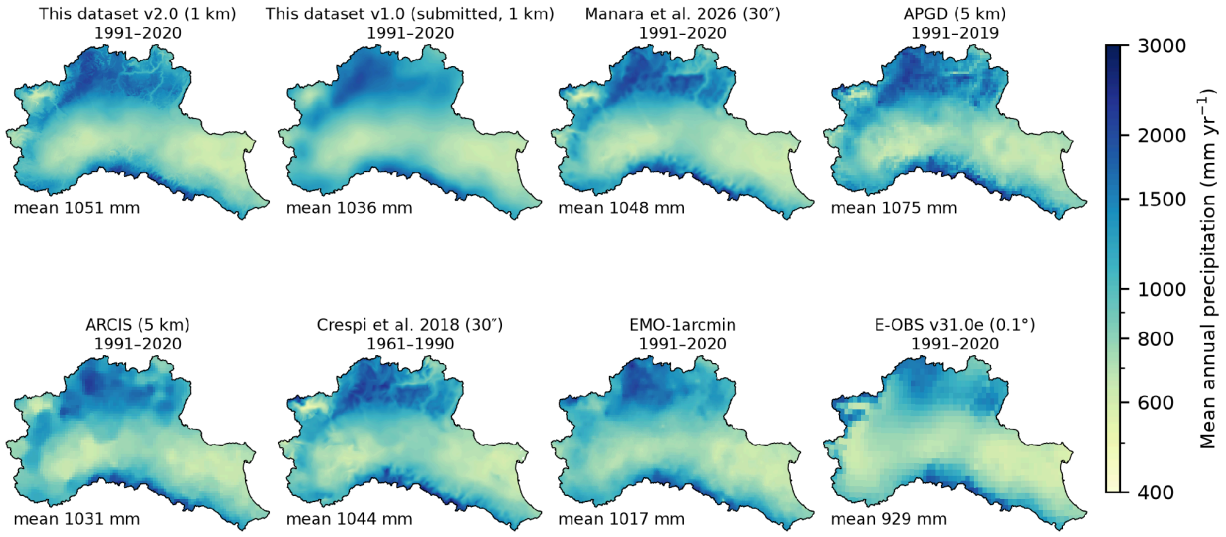


Figure R1. Mean annual precipitation of this dataset (v2.0 and the submitted v1.0) and of the reference datasets, each shown on its native grid (1991–2020; APGD 1991–2019; Crespi et al. 2018: 1961–1990). The District mean (1-km cells inside the District) is given in each panel. (Corresponds to Fig. 13 of the revised manuscript.)

Smoothing of the climatological gradients. The revised climatology reproduces the spatial variability of the reference climatologies better than the submitted version (Table R1, Fig. R1):

Table R1. Spatial standard deviation of mean annual precipitation across 1-km cells (mm) and pattern correlation with Crespi et al. (2018).

Region	v1.0 (submitted)	v2.0 (revised)	Manara et al. (2026)
Whole district, SD	320 mm	353 mm	369 mm
Aosta Valley, SD	222 mm	238 mm	268 mm
Valtellina, SD	190 mm	233 mm	318 mm
Aosta Valley, correlation with Crespi et al. (2018)	0.52	0.70	—

Spatial variability is still lower than in Manara et al. (2026), especially in Valtellina. We state this limitation in Sects. 6.4 and 7.

Comment 1.3

In my opinion, a dataset like the one proposed is of little use without a clear and documented discussion of the actual resolution of the proposed fields. I understand that addressing this topic extensively is not easy, but I believe that an in-depth analysis of the kriging weights could be very helpful in better understanding this important aspect. An analysis of this type could, for example, identify—for each variable, day, and location—the

distances at which the weights decrease to 2/3, 1/2, and 1/3, and subsequently present and discuss a series of distributions to show the values obtained in different areas and—given the significant increase in available data from the early years to the later ones—across different time periods.

Response. We ran the analysis suggested by the reviewer (Sect. 5.2 and 6.2, Fig. 10 and Table 3 of the revised manuscript; Fig. R2 and Table R2 here). For each variable, day, and 1-km cell we recorded the distance to the farthest kriging station whose weight is at least 2/3, 1/2 and 1/3 of the largest weight assigned on that cell-day.

Ordinary-kriging weights can be negative and are not necessarily monotonic with distance. The measure is therefore defined on positive weights only: R(f) is the distance to the farthest station whose positive weight is at least the fraction f of the largest positive weight. The distributions are presented by region (whole District, Aosta Valley, Valtellina, rest of the District) and by decade.

Table R2. Median distance (km) to the farthest kriging station whose weight is at least the given fraction of the largest weight (whole District, all cell-days of the decade).

Weight ≥ … of the maximum	Precipitation 1991–2000	Precipitation 2011–2020	Temperature 1991–2000	Temperature 2011–2020
2/3	13.8 km	8.2 km	20.8 km	11.2 km
1/2	16.8 km	10.8 km	23.2 km	13.2 km
1/3	19.8 km	12.2 km	25.2 km	14.2 km

In 2011–2020 the 1/2-weight distance for precipitation is 8.2 km in the Aosta Valley and 13.8 km in Valtellina.

These distances describe the weight-based spatial support of each estimate and serve as a proxy for effective resolution. They are consistent with the reviewer's estimate. The characteristic weight-based support of the daily anomaly fields is about 8–20 km for precipitation and 11–25 km for temperature, and it is coarser in the 1990s than in recent years, reflecting the growth of the network. Because kriging uses seven stations, these distances describe support within that neighbourhood. They are limited by station spacing and should not be read as an independent, for example spectral, estimate of resolution.

For comparison:

- the MeteoSwiss documentation of APGD, which is based on a much denser pan-Alpine network, gives an effective resolution of approximately 10–20 km;
- Isotta et al. (2014) give 10–25 km for daily precipitation.

We now state this clearly in the abstract, Sect. 6.2, and the Usage notes (Sect. 7). The climatological layer, i.e. the 1991–2020 normals that model the local dependence on elevation (see our response to Comment 1.7), is represented on the 1-km grid and contains terrain-dependent spatial structure at that scale. The daily anomaly component superimposed on it has the coarser weight-based support quantified above, and we now describe the product in these terms.

We also measured daily precipitation smoothing directly with leave-one-out (LOO) cross-validation:

- the station-median ratio of the predicted to the observed standard deviation of daily precipitation is 0.96;
- heavy precipitation is underestimated: the station-median ratio of the 99th percentile is 0.95, and that of the mean annual maximum daily precipitation is 0.93 (Sect. 6.1).

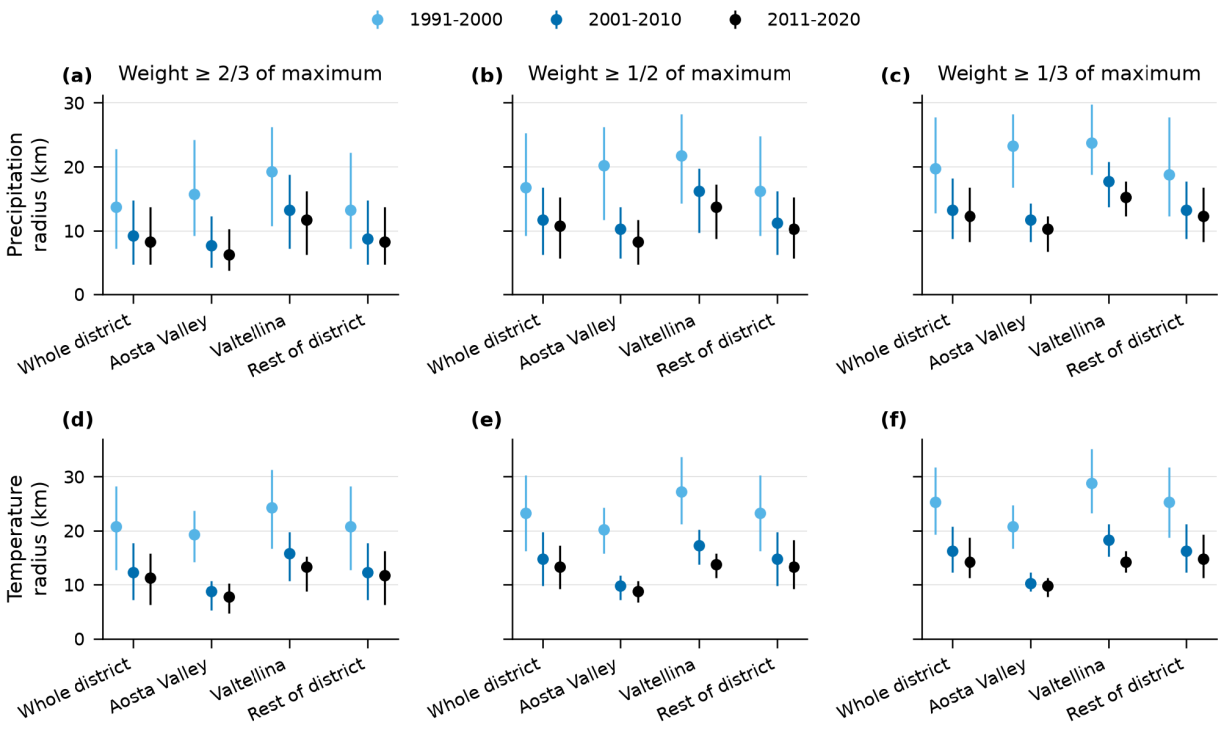


Figure R2. Effective resolution of the daily fields: distance (km) to the farthest kriging station whose weight is at least 2/3 (a, d), 1/2 (b, e) and 1/3 (c, f) of the largest weight on that cell-day, for precipitation (a–c) and temperature (d–f), by region and decade. Points: median over all cell-days; bars: interquartile range. (Corresponds to Fig. 10 of the revised manuscript.)

Comment 1.4

Moreover, other analyses could concern case studies (events) for which high-resolution fields could be acquired. For precipitation these fields can e.g. be acquired from radar measurements. The variances of the radar spatial fields can then be compared with those of the files obtained by kriging and this can permit to understand how much the radar fields must be smoothed to get spatial variability comparable to those of the kriging fields.

Response. We explored the radar option. A quality-controlled radar composite that is independent of the rain gauges and covers the whole District, including its Swiss and French parts, is not available to us for the study period. Radar composites adjusted to the same rain gauges would not provide an independent measure of spatial variance.

Instead, we assessed effective resolution in three ways:

- the kriging weights (Comment 1.3);
- the leave-one-out variance ratio (Comment 1.3);
- the comparison with the 30-arc-second fields of Manara et al. (2026) (Comment 3.1).

As a case study, we compared this dataset with the stations and with all reference products for Storm Alex (2–5 October 2020; Comment 3.1, Fig. R12 and Table R13). The radar-variance approach proposed by the reviewer is mentioned in Sect. 7 as a direction for future work.

Comment 1.5

The topic of the actual resolution of the fields is not covered at all in the paper, but from i) the method used (ordinary kriging for precipitation), ii) the number of stations, iii) the high fraction of missing data and iv) some of the figures presented in the paper, I believe I can deduce that, at least for precipitation, the effective resolution is well above 10 km. In my opinion, the actual resolution of the precipitation fields is not higher than

that of the ARCIS data set (Pavan et al., 2019), which has the relevant benefit of covering a longer period (it starts in 1961) and being regularly updated.

Response. We agree that the daily precipitation fields have an effective resolution comparable to ARCIS, and we now say this in Sect. 6.2. The dataset still adds value for hydrological modelling of the Po River District because:

- it provides precipitation and mean temperature on the same grid, from the same station network and with the same processing;
- it covers the whole District, including its trans-national headwaters, with Swiss stations in the network;
- its 1-km climatological layer resolves elevation-dependent normals;
- it reproduces both the wet-day frequency and the wet-day intensity of the station records (Comment 3.1), which matters for hydrological applications;
- each daily value carries a calibrated uncertainty.

We narrowed the novelty statement in the Introduction accordingly. ARCIS and Manara et al. (2026) are now discussed there, and both are part of the comparison (Comment 3.1). The longer period and regular updating of ARCIS are acknowledged in the Introduction and in Sect. 7.

Comment 1.6

I believe it is nonetheless useful for many applications to provide datasets at a nominal resolution higher than the actual resolution; however, it is essential to address the issue of actual resolution so that users understand whether the value provided for a specific point is truly representative of that point or, instead, represents an average value for an area of hundreds of square kilometres.

Response. We agree. Version 2.0 now gives users this information at every location:

- **Weight-based spatial-support maps.** Per-cell maps of the 2/3-, 1/2- and 1/3-weight distances, by decade, are distributed with the dataset (Fig. 5c, d of the revised manuscript).
- **Uncertainty.** Each daily value carries a calibrated 1σ uncertainty (Sect. 4.4).
- **Usage guidance.** The Usage notes (Sect. 7) and the dataset README state how the two components should be interpreted.
    - The climatological component is represented on the 1-km grid and contains terrain-dependent spatial structure at that scale.
    - The daily anomaly component is observationally supported over characteristic distances of roughly 11–17 km for precipitation and 13–23 km for temperature (the 1/2-weight distance in 2011–2020 and 1991–2000, respectively).
    - Point-scale extremes are smoothed.

Comment 1.7

Evaluating the effective resolution of the fields produced is, of course, only part of the problem. Alongside this assessment, it would be very important to identify methods to improve this resolution. This can be done by modelling the relationship between meteorological variables and the physiographic characteristics of the territory by means of the PRISM (Parameter-elevation Regressions on Independent Slopes Model) approach. In this paper this approach is used only for temperature and also for this variable it is used in a rather trivial way as the link between temperature and elevation is estimated considering all stations in the study area without taking into account that the temperature-elevation relation can have relevant spatial differences and can be different for different atmospheric layers. Moreover, it can depend on the aspect, on the slope, on the presence of valleys, etc. All these

. features can of course be considered also in the kriging following the elevation effect. This step depends, however, on the availability of stations. If a point is in a valley and the closest stations are on the surrounding slopes the estimation can be strongly biased in case of significant temperature inversions that could be better captured by means of more distant stations but with more similar physiographic features.

Response. We revised the method in this direction for both variables.

Normals. The 1991–2020 monthly normals are estimated with a PRISM-type local weighted regression (Daly et al., 2008) on the nearest stations with a valid normal.

- For precipitation, the regression is ln P against elevation (15 stations). The monthly wet-day frequency is regressed locally against elevation on the logit scale (15 stations).
- For temperature, the regression is T against elevation (25 stations).

Station weights combine distance and elevation difference. Through the elevation-difference weight, a valley-floor cell is driven by valley stations rather than by stations on the surrounding slopes. This is the situation described by the reviewer, in which more distant stations with similar physiographic features are preferred.

We also tested two additional physiographic weights:

- topographic facet (eight aspect classes on a smoothed DEM);
- for temperature, topographic position (valley floor versus slope or ridge).

All weighting parameters, including the number of stations, were selected by leave-one-out cross-validation of the station normals.

- The elevation-difference weighting improved the leave-one-out skill and was retained for both variables.
- Facet weighting and the topographic-position predictor did not improve the leave-one-out skill and were not retained.
- Slope steepness was not tested as a separate predictor. Its effect on the normals is partly represented by the elevation term, and the facet test suggests that the station network is too sparse to constrain additional terrain predictors.

These tests are reported in Sect. 4.2.

Two further rules were selected by leave-one-out:

- **Precipitation is not extrapolated upward** above the highest of the local stations. Without this rule, the normals of cells lying more than 500 m above their local network were biased high (median +36 % in leave-one-out at such stations). With it, the annual leave-one-out MAE of the normals decreases from 91.7 to 90.0 mm, which corresponds to a relative MAE of 7.9 %.
- **Temperature above the highest local station** uses a fixed lapse rate instead of the local lapse rate, which is flattened by valley inversions. We tested −5.5, −6.5, −7.5, −8.5 and −9.8 °C km^{-1} and the local lapse rate. The value −7.5 °C km^{-1} gave the lowest monthly leave-one-out MAE and reduced the warm bias at such stations from +0.59 °C (local lapse rate) to +0.11 °C. The monthly leave-one-out MAE of the temperature normals is 0.61 °C.

The leave-one-out error of the annual wet-day frequency normal is 1.4 percentage points (bias −0.2 points; median relative error 3.9 %).

Daily anomalies. The daily temperature lapse rate is estimated locally from the 50 nearest active stations of each cell, rather than domain-wide. It therefore varies in space and from day to day, and it can be weak or positive during inversion episodes. Extrapolation of the anomaly lapse rate beyond the elevation range of the local stations is limited to 300 m.

One limitation remains, and we state it in Sect. 7: inversions confined to narrow valleys smaller than the station neighbourhood are only partly captured when no station is located on the valley floor.

Agreement with Brunetti et al. (2014). The resulting temperature normals closely match the elevation dependence of this climatology (Table R3). The spatial correlation of the annual normals is 0.987 (Fig. R3).

Table R3. Elevation gradient of the annual temperature normal (°C km^{-1}).

Region	This dataset (1991–2020)	Brunetti et al. (2014) (1961–1990)
Whole District	−5.30	−5.33
Aosta Valley	−5.74	−5.52
Valtellina	−5.66	−5.50

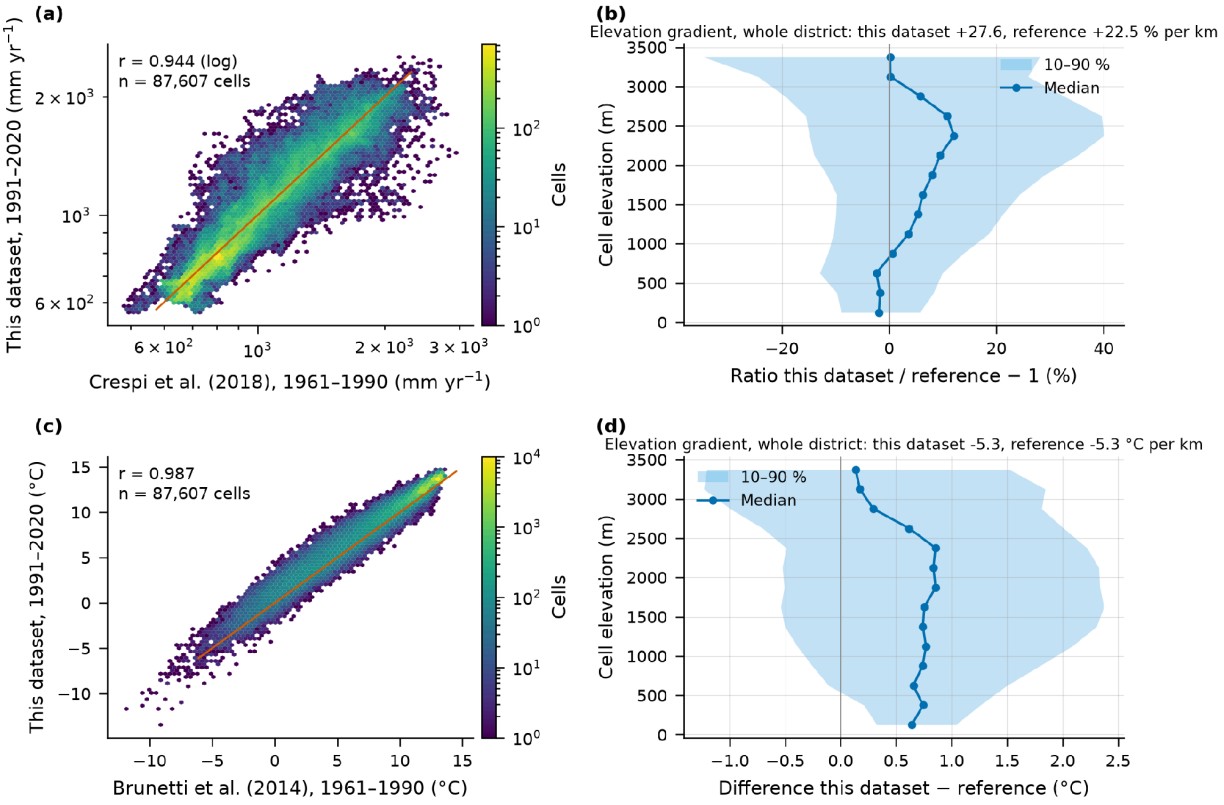


Figure R3. 1991–2020 normals of this dataset compared with the 1961–1990 climatologies of Crespi et al. (2018; precipitation; a, b) and Brunetti et al. (2014; temperature; c, d): cell-by-cell density scatter (a, c) and ratio or difference by elevation band (b, d), with the district-wide elevation gradients. (Corresponds to Fig. 15 of the revised manuscript.)

2. Robustness to missing data of the adopted methodology

Comment 2.1

High-resolution fields of meteorological variables are often produced by means of the anomaly method, which is based on the combination of fields of climate normals (usually defined over standard 30-year periods) and anomalies from them. The former depend only on space, whereas the latter depend on space and time. The benefit of this combination is that the definition of the climate normals and the anomalies can be performed with different methods and, when useful, also with different data sets. A well-known example of this approach is

the daily Alpine Precipitation Grid Data Set (Isotta et al., 2014), which is based on fields of monthly normals from Schwarb (2000) and anomalies from them. In this case, the fields of the monthly normals were derived with a PRISM approach allowing to model (at the local level) the link between the physiographic features of the territory and the precipitation normals. The same approach to define spatial fields of monthly precipitation normals has been adopted by Crespi et al. (2018) for Italy and more recently by Manara et al. (2026) for northern Italy and the northern part of central Italy. For the Italian territory, such spatial fields of climate normals are available also for temperature (Brunetti et al., 2014). Once the spatial fields of the climate normals (e.g. over the 30-year period 1991-2020) are defined, any time series of observations can be converted into an anomaly record and these anomaly records can be interpolated by means of many methods. The most important benefit of this approach is that it can be applied also to data sets with a relevant fraction of missing data, provided that the issue of missing data is correctly managed when the station climate normals are defined. This because the anomalies have much greater spatial coherence than the corresponding absolute values, which makes the interpolation of the anomalies robust to missing data. However, the same robustness is not guaranteed when the decomposition into normal values and anomalies is not used. For example, if we consider a point in the Aosta Valley between the city of Aosta and the Great S. Bernard Pass, it is located near a very dry station (Aosta) (about 500 mm of rainfall per year) and at the same time near a very wet station (Gran San Bernardo) (over 2000 mm of rainfall per year). It is quite clear that obtaining with kriging the rainfall for that point produces a record that is more affected by the missing data at one of the two stations than by any other factor linked to climate variability and change.

Response. We adopted the anomaly approach described by the reviewer (Sect. 4.1–4.3). For precipitation we apply it separately to wet-day occurrence and to wet-day amount, for the reasons given below.

Station normals and missing data. For each station and calendar month, the 1991–2020 normal is computed from months with at least 80 % valid days. Missing years are reconstructed from up to three highly correlated neighbouring stations: precipitation by ratio and temperature by difference.

- The neighbours must have $r \geq 0.70$ (precipitation) or $r \geq 0.85$ (temperature) on at least 36 common months.
- A normal is accepted when at least 5 years are observed and at least 20 years are available after reconstruction.

Using these rules, 1375 precipitation stations and 1221 temperature stations have a complete set of twelve station normals. For the remaining stations, in the months without a station normal, the gridded normal at the station location is used (these station-months are flagged). For precipitation, the monthly wet-day frequency (≥ 1 mm) of each station is computed from the same months.

Daily interpolation, temperature. Each daily observation is converted into a difference from its station normal. The differences are interpolated by ordinary kriging with the local lapse rate, and the result is added to the gridded normal.

Daily interpolation, precipitation (two-part scheme). The gridded normals give, for every cell and month, the mean precipitation M, the wet-day frequency p and the mean wet-day intensity $i = M / (p \cdot \text{days})$. We build the daily field in three steps:

1. **Occurrence.** The station wet-day indicator (1 if P ≥ 1 mm, 0 otherwise) minus the station frequency normal is interpolated by ordinary kriging. The result is added to the gridded p, which gives a daily wet-day index π for every cell.
2. **Amount.** At the wet stations, the ratio of the observed amount to the normal intensity of the station is interpolated by ordinary kriging. Multiplying by the gridded i gives the wet-day amount A for every cell.

3. **Recombination.** A cell is wet when π exceeds a threshold. The threshold is set for each cell and calendar month so that the number of wet days over 1991–2020 equals the number implied by the frequency normal. Daily precipitation is A on wet cells and 0 elsewhere.

Days with 0–1 mm are dry in the occurrence step. At the stations they make up 10.8 % of all days but only 1.3 % of the total precipitation. Because M includes these amounts, they are carried into the wet-day intensity normal i, which is therefore about 1.3 % higher than the mean intensity of days with at least 1 mm. As a result, the total is preserved, but values between 0 and 1 mm are almost absent from the gridded fields (0.04 % of the predicted station-days). This is stated in Sect. 4.3 and in the README.

Separating occurrence from amount follows the approach of E-OBS, in which precipitation occurrence is interpolated separately from the amounts (Haylock et al., 2008; Cornes et al., 2018). Related schemes estimate the probability of precipitation separately (Newman et al., 2015).

Why we use two parts. We first implemented the anomaly method with a single ratio of the daily total to its normal. That version improved the climatological normals and leave-one-out MAE compared with the submitted dataset. However, it kept one problem of interpolating daily totals: between stations the wet area spreads, so wet days are too frequent and too weak.

Table R4 compares the three versions on identical station-days. The submitted version (v1.0) overestimates wet-day frequency by 4.0 percentage points (16 %) and underestimates intensity by 14 %. The single-ratio anomaly version reduces neither bias substantially. The two-part version (v2.0) reproduces both. Fig. R4 and Fig. R5 show that this holds at all elevations and across the domain.

Table R4. Leave-one-out evaluation of three precipitation versions on identical station-days (1991–2020; 9 249 370 station-days at the 1580 stations with leave-one-out values in all three versions). All scores are pooled over these station-days. Wet day: ≥ 1 mm. POD and FAR refer to wet-day detection.

	Stations	v1.0 (submitted)	Single-ratio anomaly (tested)	v2.0 (revised, two-part)
Wet-day frequency	25.4 %	29.4 %	28.6 %	25.2 %
Mean wet-day intensity	11.71 mm	10.06 mm	10.25 mm	11.97 mm
99th percentile	44.4 mm	41.1 mm	41.3 mm	41.8 mm
KGE	—	0.868	0.879	0.889
Correlation	—	0.900	0.914	0.909
MAE	—	1.267 mm	1.181 mm	1.221 mm
RMSE	—	4.11 mm	3.83 mm	3.95 mm
Bias	—	−0.06 %	−0.59 %	+0.15 %
POD	—	0.913	0.915	0.866
FAR	—	0.212	0.188	0.127

The single-ratio version has the lowest MAE and RMSE and the highest correlation. These squared and absolute error measures favour fields that spread precipitation over more days. The two-part version accepts an MAE that is 3 % higher in exchange for:

- correct wet-day frequency and intensity;
- a higher KGE;
- a false-alarm rate that is lower by a third.

Its probability of detection is lower (0.87 against 0.91). For hydrological forcing, wet-day occurrence and event intensity matter because they affect infiltration and runoff generation. For this use, we prefer this trade-off and state it in Sect. 6.1. We do not claim that it necessarily improves every hydrological application.

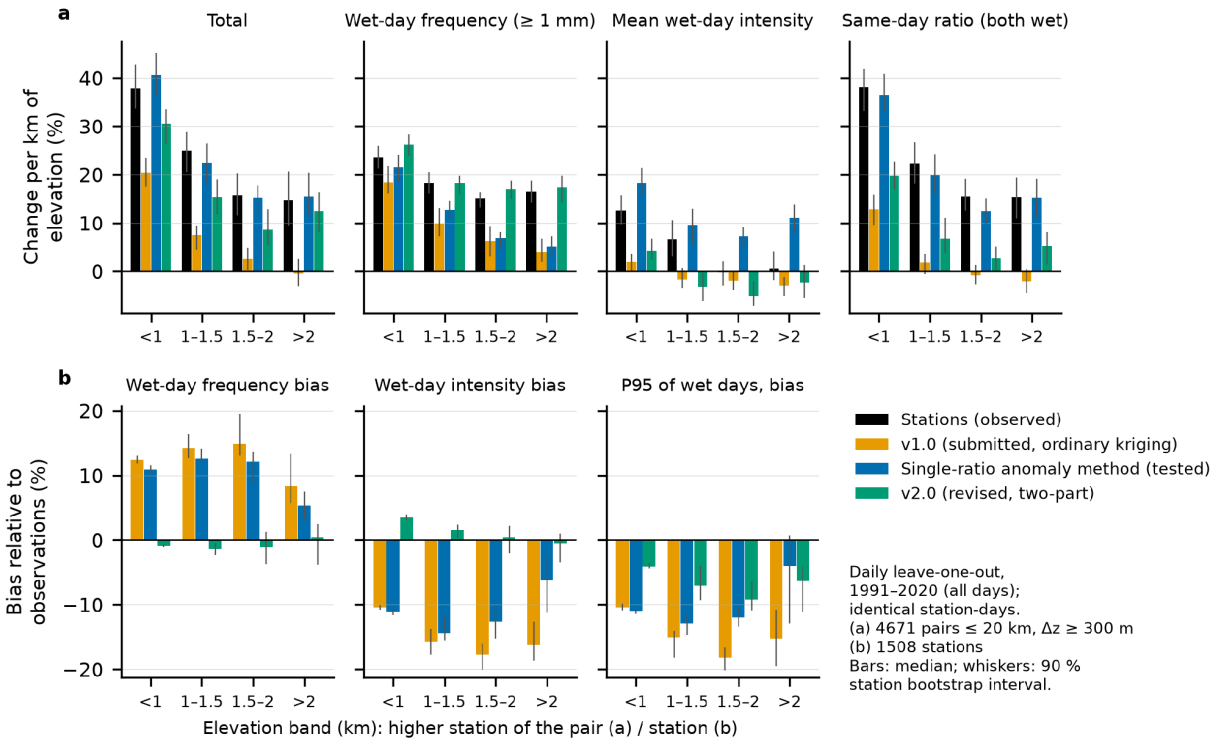


Figure R4. Leave-one-out evaluation of the three precipitation versions on identical station-days (1991–2020). (a) Change per km of elevation of the total, the wet-day frequency, the wet-day intensity and the same-day ratio, for station pairs at most 20 km apart with an elevation difference of at least 300 m, by elevation band of the higher station. (b) Bias of wet-day frequency, wet-day intensity and the 95th percentile of wet days relative to the observations, by station elevation. Bars: median; whiskers: 90 % station-bootstrap interval.

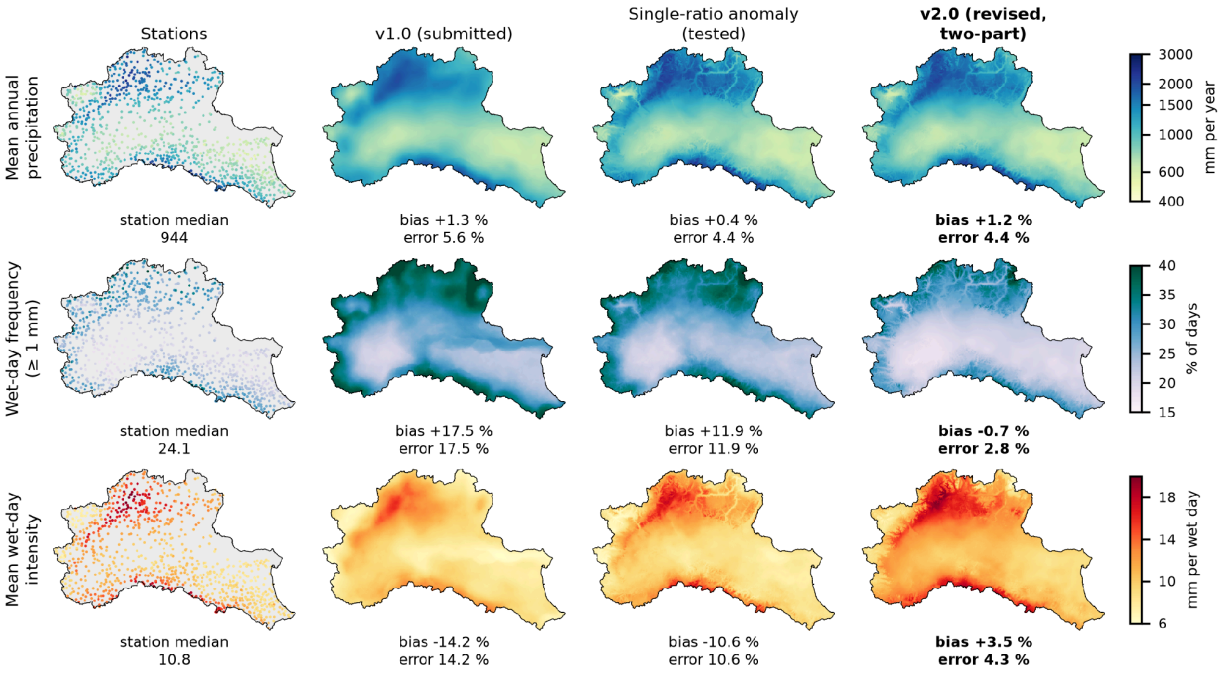


Figure R5. Mean annual precipitation, wet-day frequency (≥ 1 mm) and mean wet-day intensity, 1991–2020: station climatologies and the gridded fields of the submitted version (v1.0), the single-ratio anomaly version (tested) and the revised

version (v2.0, two-part). Bias and error: median and median absolute value of the relative difference between the gridded value and the station climatology at the 933 in-domain stations with at least 10 years of data.

Leave-one-out without leakage. In leave-one-out cross-validation, the withheld station is also excluded from the estimation of the gridded normals (precipitation, wet-day frequency, temperature) at its location. This avoids an optimistic bias of the anomaly method.

Skill. The station-wise leave-one-out skill of the revised dataset is given in Table R5. Metrics are means over stations with at least 365 days, and KGE follows Gupta et al. (2009).

Table R5. Station-wise leave-one-out skill of the revised dataset (v2.0), with the values reported for the submitted dataset (v1.0) in brackets. Scores are means of the station-wise values over the 1627 stations with at least 365 days. Wet days: observed precipitation ≥ 0.1 mm, the definition used in the submitted manuscript, kept here so that the bracketed values remain comparable (elsewhere in this letter a wet day is ≥ 1 mm).

	KGE	CC	MAE	RMSE	Bias
Precipitation, all days	0.83 (0.84)	0.90 (0.90)	1.25 mm (1.28)	3.75 mm (3.79)	+1.3 % (+1.1 %)
Precipitation, wet days	0.82 (0.82)	0.89 (0.88)	3.14 mm (3.05)	5.98 mm (6.02)	−1.8 % (−4.3 %)
Temperature	0.91 (0.88)	0.99 (0.98)	0.89 °C (1.14)	1.13 °C (1.50)	−0.02 °C (+0.02)

The values in brackets come from the submitted manuscript. They were computed on a different network and without the day-definition harmonisation, so they are not strictly comparable. The like-for-like comparison is Table R4, in which v1.0 was re-evaluated on the same station-days.

For precipitation, the station-wise means of v2.0 are similar to those reported for v1.0. On identical station-days, v2.0 has a higher KGE (0.889 against 0.868) and a lower MAE (1.221 against 1.267 mm), and it removes the frequency and intensity biases. We also corrected the definition of the KGE in the manuscript.

The Aosta–Gran San Bernardo situation described by the reviewer was tested explicitly (see Comment 2.4).

Comment 2.2

Not all high-resolution fields of meteorological variables are naturally produced by means of the anomaly method. Many other methods can be used. However, if the goal of the fields is to highlight climatic variability and change, in this case the station records should not present a relevant fraction of missing data.

Response. We agree. The revised Usage notes (Sect. 7) state two things: the station records were not homogenised, and the dataset is primarily intended for hydrological applications rather than for the detection of climate trends. The statement in the submitted manuscript that the dataset is suitable for trend analysis has been removed.

To check temporal consistency, we compared district-mean annual series with the reference datasets (Table R6; Table 7 and Sect. 6.5 of the revised manuscript). Trends are Theil–Sen slopes, and significance is tested with the Mann–Kendall test.

- None of the precipitation trends is significant ($p \geq 0.4$).
- Their sign and magnitude agree with APGD, Manara et al. (2026) and ARCIS.
- The temperature trend lies between those of the reference products.

Table R6. District-mean annual series: correlation with the reference datasets and Theil–Sen trends.

Variable	Reference (period)	Annual correlation	Trend, this dataset	Trend, reference
Precipitation	APGD (1991–2019)	0.996	+1.5 % per decade	+2.3 % per decade
Precipitation	Manara et al. 2026 (1991–2020)	0.999	+0.8 % per decade	+1.1 % per decade
Precipitation	ARCIS (1991–2020)	0.998	+0.8 % per decade	+0.7 % per decade
Temperature	ARCIS (Tmax+Tmin)/2 (1991–2020)	0.997	+0.51 °C per decade	+0.55 °C per decade
Temperature	EMO-1arcmin (1991–2020)	0.989	+0.51 °C per decade	+0.56 °C per decade

Temperature	E-OBS (1991–2020)	0.988	+0.51 °C per decade	+0.37 °C per decade

We also checked whether the growth of the network creates artificial trends. At the 171 stations with at least 28 complete years, we compared the trend of each observed series with that of its leave-one-out prediction, which depends only on the surrounding (changing) network.

- The median difference in annual precipitation trend is −0.02 % per decade; that in wet-day frequency is +0.15 percentage points per decade.
- At individual stations the trends are only moderately reproduced (correlation 0.50 for annual precipitation). Sect. 7 therefore advises against deriving local trends from the dataset.

Comment 2.3

Using a methodology robust to missing data is, of course, only part of the problem. Alongside this assessment, it would be very important to assess the effect of missing data on the data set. This can e.g. be investigated by focus on a period of full data availability (e.g. the most recent years) and by obtaining a high number of different data sets by randomly deleting 50 % of the data. Several analyses on these data sets can then permit to estimate the uncertainty introduced by this large fraction of missing data.

Response. We performed two versions of this experiment (Sect. 5.3 and 6.3; Fig. 11 and Table 4 of the revised manuscript; Fig. R6 and Tables R7–R9 here), and repeated the temporal-missingness experiment for precipitation on a four-year window. The two main experiments use the same design elements:

- **Days.** 120 days of the data-rich decade 2011–2020, as suggested by the reviewer (12 days per year, 10 per calendar month, drawn from the stratified day sample used throughout the evaluation).
- **Realisations.** Thirty in each experiment.
- **Methods.** Daily fields were produced with the method of the dataset (two-part scheme for precipitation, anomaly method for temperature) and with direct (non-climatological) interpolation.

The two versions differ in what is removed:

1. **Temporal missingness**, which is the design proposed by the reviewer. In each realisation, every valid station-day is withheld with probability 0.5. No station is removed from the network, and the normals are those of the full network. The withheld station-days are predicted from the remaining observations of the same day. As a conservative choice, the target normal at a withheld station is its leave-one-out grid normal. As a baseline, the same station-days were also predicted by ordinary leave-one-out with the complete network, so that the effect of the missing data can be isolated.
2. **Network sparsity.** In each realisation, a random 50 % of the stations is withheld for the whole period, and the normals are rebuilt from the retained stations only. This is a stricter test, because the climatological layer also loses half of its information.

In both versions, the fields were evaluated against the withheld observations and against the fields produced with the full network, on a fixed sample of 5000 grid cells.

Table R7. Temporal-missingness experiment (50 % of the station-days withheld, normals of the full network): errors at the withheld station-days, and their increase relative to the full-network leave-one-out on the same station-days (median over 30 realisations).

	Method of the dataset	Direct interpolation
Precipitation, daily MAE, complete network	1.148 mm	1.097 mm
Precipitation, daily MAE, 50 % of station-days missing	1.308 mm	1.242 mm
Precipitation, increase of MAE	+0.159 mm	+0.146 mm (smaller in 29 of 30 realisations)
Precipitation, station-mean absolute bias[1], complete network → 50 % missing	14.6 → 17.1 %	13.4 → 15.5 %

Precipitation, wet-day frequency (observed 26.7 %), complete network → 50 % missing	25.6 → 25.5 %	30.8 → 31.7 %
Precipitation, wet-day intensity (observed 9.18 mm), complete network → 50 % missing	9.56 → 9.58 mm	7.92 → 7.72 mm
Temperature, daily MAE, complete network → 50 % missing	0.811 → 0.845 °C	0.825 → 0.875 °C
Temperature, increase of MAE	+0.034 °C (smaller in 30 of 30)	+0.051 °C
Temperature, station-mean absolute bias, complete network → 50 % missing	0.46 → 0.45 °C	0.49 → 0.50 °C

[1] Median over stations of the absolute relative difference between predicted and observed mean precipitation at the withheld station-days. With 120 days, about 60 days per station are withheld, so this value is dominated by day-to-day sampling noise rather than by climatological bias (for comparison, the median absolute leave-one-out error of the 30-year mean precipitation is 5.8 %). The four-year check below gives a stable estimate.

Four-year check (precipitation). We repeated the temporal-missingness experiment on all 1461 days of 2017–2020, with 10 realisations and evaluation at the stations only. This window is long enough for the occurrence calibration of the dataset (per station and calendar month) and for a stable station-mean bias. Values are means over the 10 realisations, complete network → 50 % missing:

- daily MAE: 1.271 → 1.445 mm (+0.174) with the two-part scheme, 1.195 → 1.352 mm (+0.157) with direct interpolation; the increase is smaller for direct interpolation in 10 of 10 realisations;
- station-mean absolute bias[1]: 7.6 → 8.1 % (+0.5) with the two-part scheme, 7.6 → 8.4 % (+0.8) with direct interpolation; the increase is smaller for the two-part scheme in 10 of 10 realisations;
- mean station bias: +1.4 → +1.4 % with the two-part scheme, +2.4 → +3.3 % with direct interpolation;
- wet-day frequency (observed 25.7 %): 25.6 → 25.6 % against 29.0 → 29.8 %;
- wet-day intensity (observed 12.27 mm): 12.43 → 12.40 mm against 10.87 → 10.61 mm.

Table R8. Network-sparsity experiment (50 % of the stations withheld, normals rebuilt): errors at the withheld stations (median over 30 realisations).

	Method of the dataset	Direct interpolation	Better in
Precipitation, daily MAE	1.288 mm (two-part)	1.235 mm	direct, 30 of 30 realisations
Precipitation, station-mean absolute bias	13.8 % (two-part)	12.7 %	direct, 29 of 30
Temperature, daily MAE	0.864 °C (anomaly)	0.873 °C	anomaly, 27 of 30
Temperature, station-mean absolute bias	0.48 °C (anomaly)	0.51 °C	anomaly, 29 of 30

Table R9 quantifies how much the gridded fields themselves change in the network-sparsity experiment. In the temporal-missingness experiment the changes are smaller:

- precipitation: 0.59 mm day^{-1} (27 %) with the two-part scheme and 0.55 mm day^{-1} (25 %) with direct interpolation;
- temperature: 0.25 °C and 0.32 °C.

Table R9. Network-sparsity experiment: difference between the fields produced with half of the network and with the full network, on 5000 grid cells × 120 days (median over 30 realisations).

	Method of the dataset	Direct interpolation
Precipitation, mean absolute difference (relative value[1])	0.61 mm day^{-1} (28 %)	0.55 mm day^{-1} (25 %)
Precipitation, 95th percentile of absolute difference	3.7 mm day^{-1}	2.8 mm day^{-1}
Precipitation, change of the spatial mean	0.00 mm day^{-1}	+0.01 mm day^{-1}
Temperature, mean absolute difference	0.31 °C	0.32 °C
Temperature, 95th percentile of absolute difference	0.90 °C	0.97 °C

Temperature, change of the spatial mean	−0.01 °C	+0.01 °C

[1] Sum of the absolute differences divided by the sum of the full-network values over the same cell-days (all days, including dry days).

The results differ between the two variables, and we report them as they are.

- **Temperature.** In both experiments the anomaly method is better than direct interpolation, and it is also less sensitive to the missing data. With temporal missingness, its MAE increases by 0.034 °C against 0.051 °C, in all 30 realisations, and the station-mean bias does not change.
- **Precipitation, daily errors.** The expected robustness gain of the climatological layer does not appear in the daily errors. With temporal missingness, the MAE of the two-part scheme increases by about as much as that of direct interpolation (+0.16 against +0.15 mm), and its MAE is higher already with the complete network. With network sparsity, it is slightly higher still. The reason is that the two-part scheme reproduces the intermittency of precipitation: each cell is either wet or dry. Misclassified wet and dry days produce larger absolute errors than the smooth, widespread light precipitation of direct interpolation. The same property explains its higher MAE in Table R4.
- **Precipitation, station-mean bias.** On 120 days, the station-mean bias is larger for the two-part scheme and grows more (Table R7). On four years, where this measure reflects the climatological error rather than sampling noise, both methods start from the same value (7.6 %), and the two-part scheme degrades less with missing data (+0.5 against +0.8 points) and does not drift in the mean (+1.4 % against +2.4 → +3.3 %). This is the climatological robustness the reviewer refers to; it appears in the station climatology, not in the daily MAE.
- **Precipitation, frequency and intensity.** With the two-part scheme, wet-day frequency and intensity are essentially unaffected by the missing data (25.6 → 25.5 % and 9.56 → 9.58 mm, against 26.7 % and 9.18 mm observed). With direct interpolation, both are biased already with the complete network, and the bias grows when data are missing (30.8 → 31.7 % and 7.92 → 7.72 mm).
- **Local-gradient robustness.** The Aosta–Gran San Bernardo test (Comment 2.4) shows that the climatological layer substantially reduces the sensitivity to station loss at this particularly strong local precipitation gradient.
- **Configuration.** In all experiments, the occurrence threshold of the two-part scheme is calibrated to the wet-day frequency of the leave-one-out grid normal at the withheld station, which is computed without that station. The observed frequency at the withheld station-days is never used, so no information leaks from the evaluation data. With 120 days, the number of wet days is matched per station over all evaluated days (pooled), instead of per station and calendar month as in the dataset; the four-year check uses the per-month calibration of the dataset. The absolute values are therefore not identical to those of the full leave-one-out (Table R4).

We report these figures in Sect. 6.3 as an estimate of the uncertainty introduced by missing data.

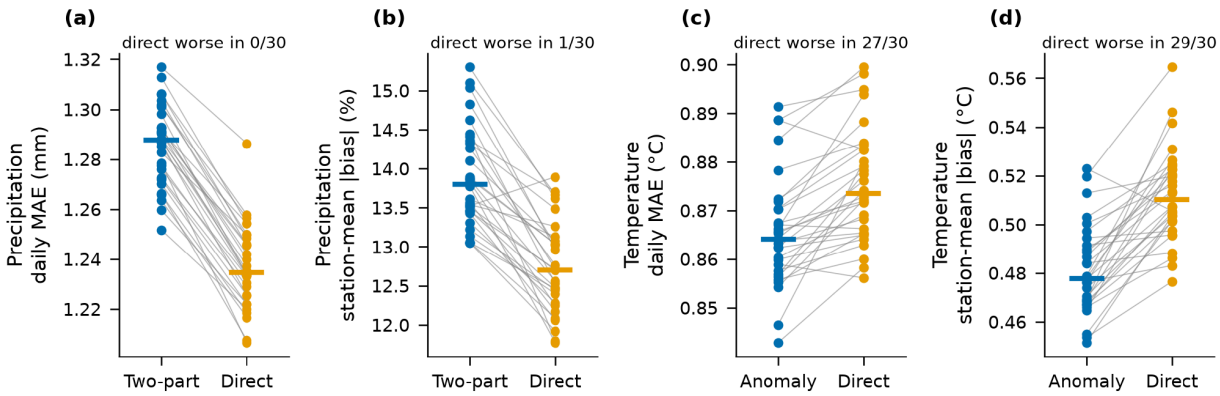


Figure R6. Network-sparsity experiment: 50 % random station deletion with normals rebuilt (30 realisations, 120 days of 2011–2020). Errors at the withheld stations with the method of the dataset (two-part scheme for precipitation, anomaly method for temperature) and with direct interpolation. Grey lines join the two methods for the same realisation; horizontal bars are medians. (Corresponds to Fig. 11 of the revised manuscript.)

Comment 2.4

Another option can be to focus on selected points (e.g. the point previously discussed between Aosta and Great S. Bernard) and by trying to understand the effect of missing data on the corresponding temperature and precipitation records. The point between Aosta and Great S. Bernard will e.g. show clearly biased minimum and maximum precipitation values when Great S. Bernard and Aosta data are, respectively, missing.

Response. We carried out this test (Sect. 5.3 and 6.3; Fig. 12 and Table 5 of the revised manuscript; Fig. R7 and Table R10 here).

An important finding was that the Gran San Bernardo station was not part of the submitted network. In the revision we added the official MeteoSwiss stations (SwissMetNet and manual precipitation network) within 20 km of the District, including Col du Grand St-Bernard (2472 m). Before inclusion, they passed the same quality control and a duplicate screening against the existing network.

We computed precipitation and temperature at the pass and at 1/4, 1/2 and 3/4 of the Aosta–Gran San Bernardo transect, under four scenarios: all stations, without Aosta, without Gran San Bernardo, and without both. The normals were rebuilt for each scenario.

Table R10. Aosta–Gran San Bernardo point test, precipitation 1991–2020: mean annual total (mm), 99th percentile of daily precipitation (mm) and mean annual maximum daily precipitation (mm), with all stations and without Gran San Bernardo (GSB). Two-part: method of the dataset; direct: interpolation without normals. Values in brackets: change relative to all stations.

Point	Method	Annual total, all	Annual total, without GSB	P99, all	P99, without GSB	Annual max, all	Annual max, without GSB
1/4	Two-part	655	642 (−2 %)	22.8	22.2 (−2 %)	42.3	42.2 (0 %)
1/4	Direct	758	668 (−12 %)	25.5	23.0 (−10 %)	46.6	46.9 (+1 %)
1/2	Two-part	719	683 (−5 %)	24.2	22.9 (−5 %)	43.4	42.9 (−1 %)
1/2	Direct	858	688 (−20 %)	26.9	22.9 (−15 %)	46.6	43.3 (−7 %)
3/4	Two-part	982	865 (−12 %)	27.9	25.2 (−10 %)	46.6	46.3 (−1 %)
3/4	Direct	1118	793 (−29 %)	31.3	23.7 (−24 %)	50.5	42.7 (−15 %)
GSB	Two-part	1590	987 (−38 %)	42.0	24.6 (−41 %)	65.8	44.5 (−32 %)
GSB	Direct	1603	928 (−42 %)	45.6	25.2 (−45 %)	69.9	44.0 (−37 %)

The test supports the reviewer's point (Table R10).

- **Direct interpolation.** The loss of GSB biases the transect low: annual totals decrease by 12–29 % and the heavy-precipitation statistics by up to 24 % (P99) and 15 % (annual maximum).
- **Two-part scheme.** The annual totals decrease by 2–12 % and P99 by 2–10 %, while the annual maximum changes by 1 % or less along the transect.
- **Loss of Aosta.** It has the opposite effect, as the reviewer anticipated, but it is small. At 1/4 of the transect, the annual total increases by 3.9 % with direct interpolation (from 758 to 787 mm) and by 1.7 % with the two-part scheme (from 655 to 666 mm).

The daily minimum is zero at all points in all scenarios and is therefore not informative for precipitation; the full distribution is represented by the annual totals, P99 and the annual maximum. At the pass, the mean annual temperature changes from 0.34 °C (with the station) to 0.67 °C (without it).

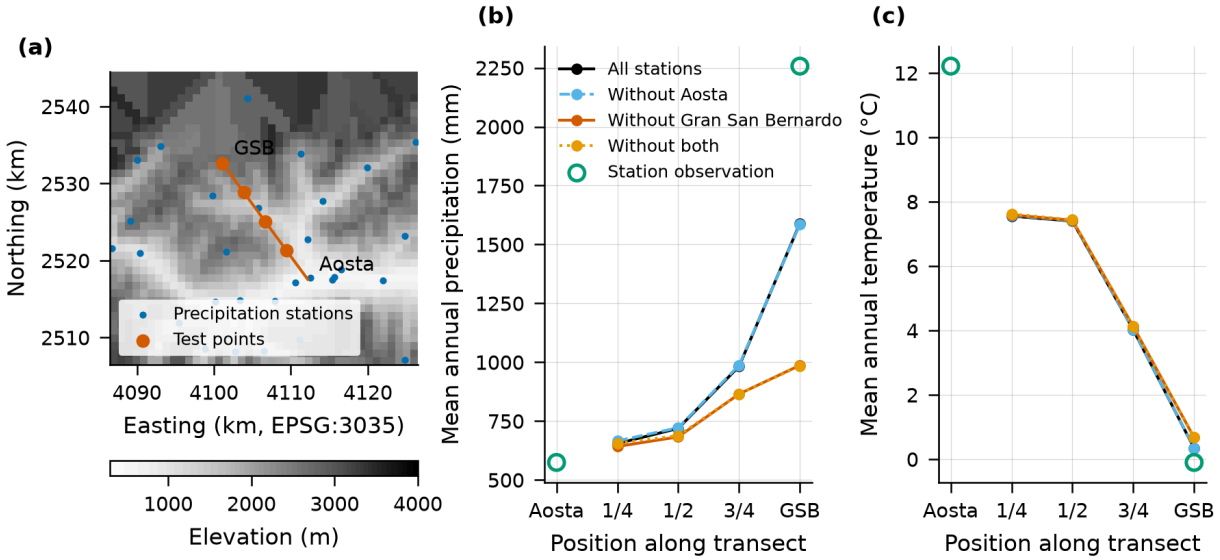


Figure R7. Aosta–Gran San Bernardo point test. (a) Terrain, precipitation stations and test points; (b) mean annual precipitation and (c) mean annual temperature along the transect with all stations, without Aosta, without Gran San Bernardo and without both (method of the dataset; normals rebuilt for each scenario). Open circles: station observations. (Corresponds to Fig. 12 of the revised manuscript.)

Limitation at the pass (Fig. R8). Gran San Bernardo records 2257 mm yr^{-1} (1991–2020). The estimates of this dataset at the pass are:

- **Station withheld (leave-one-out):** 56 % below the station value, with daily r = 0.71.
- **Station included:** the 1-km cell containing the station is 31 % below the station value.

This difference remains even when the station is included. The reason is that the station is only one part of the final gridded estimate. The daily value of a cell is the kriged anomaly multiplied by the gridded normal of that cell. The gridded normal comes from the local PRISM-type regression on the 15 nearest stations, which is not constrained to pass through any single station. At the cell of Gran San Bernardo it is 1677 mm, i.e. 26 % below the station normal of 2259 mm (sum of the twelve monthly normals; the 2257 mm given above is the mean of the observed annual totals). The remaining difference arises because the kriged anomaly at the cell centre, with a nugget in the daily variogram, is not identical to the anomaly of the station itself.

The neighbouring gauges within 4–10 km record 600–1400 mm yr^{-1} (for example 961 mm at 2084 m, 4.4 km from the pass). The pass-top excess is therefore a point-scale signal that the surrounding network does not support.

Other gridded products show the same behaviour at this cell:

- ARCIS: −67 %;
- Manara et al. (2026): −45 %;
- E-OBS: −28 %;
- APGD: −17 %;
- EMO-1arcmin, which assimilates the station: −1.3 %.

We also tested a short-range residual correction designed to reproduce such station values. It degraded the leave-one-out skill of the normals at every range tested and was therefore not adopted. This limitation is discussed explicitly in Sect. 6.3, Sect. 7 and the dataset README.

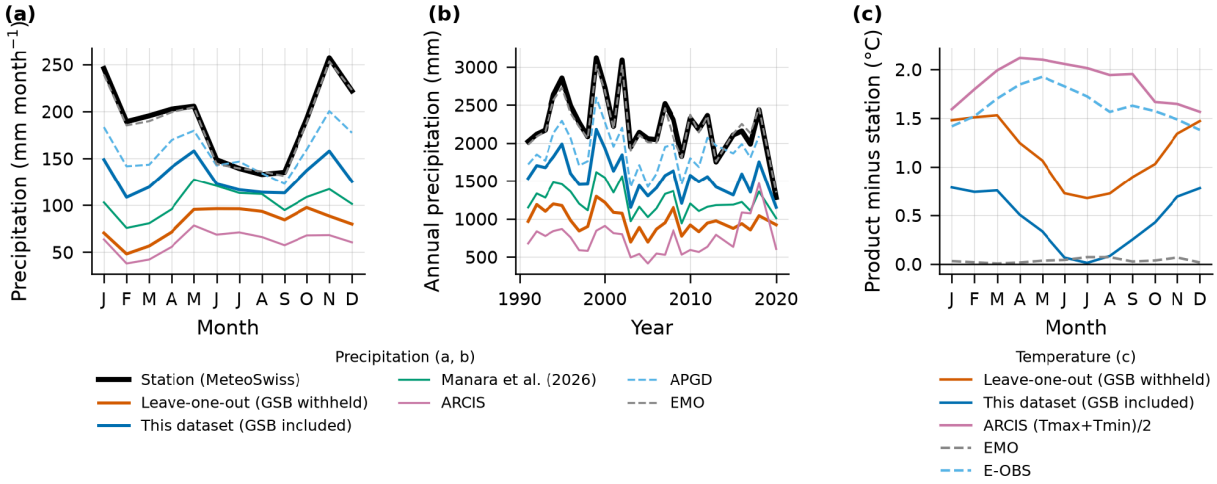


Figure R8. Col du Grand St-Bernard (MeteoSwiss, 2472 m), 1991–2020: (a) mean monthly and (b) annual precipitation, and (c) mean monthly temperature difference to the station, for this dataset with the station withheld (leave-one-out) and included, and for the reference datasets (dashed: products that include the station). APGD: MeteoSwiss release RapdD v2.0. (Corresponds to Supplement Fig. S1 of the revised manuscript.)

3. Comparison with the best datasets available

Comment 3.1

In addition to the two limitations previously discussed, the paper suffers from a fundamental shortcoming: it fails to compare its results with the best datasets available for the study area. Indeed, the only datasets used for comparison are E-OBS and EMO-1arcmin, both of which lack homogeneous data coverage for the region in question. Conversely, other datasets—which are far more robust for this study area—are ignored simply because they do not provide both temperature and precipitation data. This is e.g. the case of the Alpine Precipitation Grid Data Set (Isotta et al., 2014), the 5 km resolution ARCIS precipitation data set (Pavan et al., 2019) (now this data set is available also for temperature (Pavan et al., 2026), the Manara et al. (2026) 30-arc-second resolution precipitation data set, the Crespi et al. (2018) data set of monthly precipitation normals, the Brunetti et al. (2014), data set of monthly temperature normal, etc. Although they differ in characteristics from the dataset presented in this paper (in terms of spatial resolution, time period covered, and data aggregation), these datasets represent a valuable resource for evaluating its quality. An example is again the spatial distribution of mean annual precipitation presented in the first panel of figure 8. This distribution could be compared with the corresponding distribution from Isotta et al. (2014) (the data are now updated to 2019), Manara et al. (2026), Pavan et al. (2019) (the ARCIS data are regularly updated) and also (even though the investigated period is different) with those from Crespi et al. (2018). These comparisons would be very useful for better understanding the quality of the presented dataset.

Response. We agree, and a new section compares the dataset with all of the products named by the reviewer (Sect. 5.4 and 6.4; Figs. 13–15 and Table 6 of the revised manuscript; Fig. R1, Fig. R3, Fig. R9, Fig. R10, Fig. R11 and Fig. R12 and Tables R11–R13 here):

- **APGD (Isotta et al., 2014):** the current MeteoSwiss release (dataset doi:10.18751/Climate/Griddata/APGD/1.0; product code RapdD, version 2.0, as given in the MeteoSwiss product documentation; 5 km), which extends the dataset to 2019 as noted by the reviewer. The comparison covers 1991–2019.
- **ARCIS precipitation (Pavan et al., 2019) and temperature (Pavan et al., 2026):** ARCIS temperature provides Tmax and Tmin, which we compared as (Tmax+Tmin)/2. This is not strictly like-for-like with a 24-h mean. At Gran San Bernardo the difference is only +0.05 °C on average, but it can reach a few tenths of a degree in valleys with strong diurnal cycles, and part of the mean difference with ARCIS may come from this difference in definition.
- **Manara et al. (2026; IH-GAR):** daily precipitation at 30 arc-seconds.
- **Crespi et al. (2018) and Brunetti et al. (2014):** annual climatologies for 1961–1990.
- **EMO-1arcmin (Salamon et al., 2026) and E-OBS v31.0e (Cornes et al., 2018):** retained from the submitted manuscript (1 arc-minute and 0.1°).

The 1-km fields were aggregated onto each reference grid. These products rely partly on the same station networks. We therefore present the comparison as a consistency assessment rather than an independent validation. The independent evaluation is provided by leave-one-out cross-validation (Comment 2.1) and the deletion experiment (Comment 2.3).

Mean annual precipitation fields. Fig. R1 (Fig. 13 of the revised manuscript) compares the mean annual precipitation fields, as suggested by the reviewer, including the submitted version (v1.0) and Crespi et al. (2018). The District means are:

- this dataset: 1051 mm;
- APGD (1991–2019): 1075 mm;
- Manara et al. (2026): 1048 mm;
- ARCIS: 1031 mm;
- EMO-1arcmin: 1017 mm;
- E-OBS: 929 mm.

Table R11. Precipitation, 1991–2020, this dataset compared with each reference on the reference grid (bias = this dataset minus reference).

Reference	Mean annual bias	Bias above 2000 m	Spatial correlation of climatology	Median daily correlation
APGD	−2.0 %	−8.5 %	0.95	0.83
Manara et al. (2026)	+0.3 %	−1.3 %	0.96	0.80
ARCIS	+2.0 %	+8.6 %	0.94	0.83
EMO-1arcmin	+3.3 %	+1.3 %	0.92	0.81
E-OBS	+13.1 %	−4.2 %	0.78	0.81

Crespi et al. (2018). The reference period is different. Nevertheless, the spatial correlation with this climatology is 0.93 and the mean difference is +0.5 % (+7.1 % above 2000 m). The elevation gradients are +28 % km^{-1} in this dataset and +22 % km^{-1} in Crespi et al. (2018) (Fig. R3).

High elevations. Above 2000 m the reference products differ from each other by up to 18 %, from 1097 mm (ARCIS) to 1296 mm (APGD) on their own cells. Our values on the same cells (1186–1210 mm) lie within this range: below APGD, which has the densest high-Alpine network, and Manara et al. (2026), and above ARCIS.

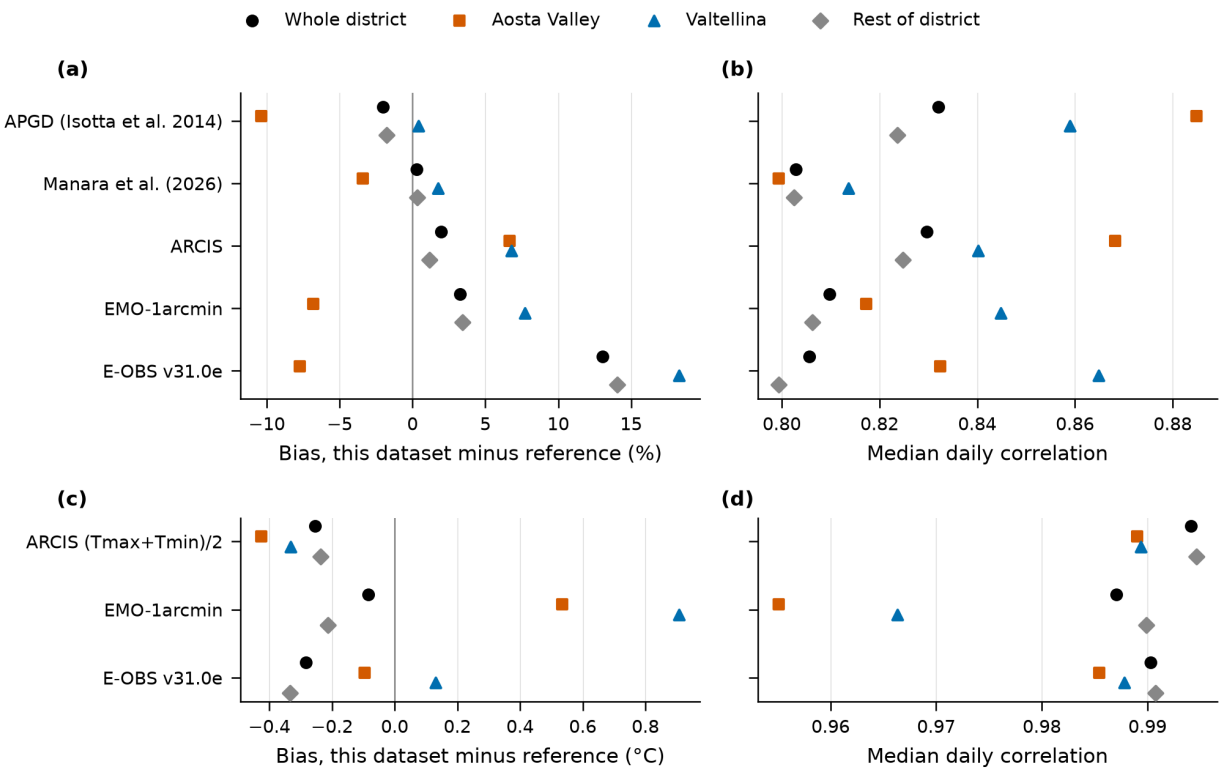


Figure R9. Consistency with the regional and European gridded datasets, 1991–2020: area-mean bias (a, c) and median daily correlation (b, d) for precipitation (a, b) and temperature (c, d), for the whole District and three sub-regions. Biases are this dataset minus reference, computed on the reference grid. (Corresponds to Fig. 14 of the revised manuscript.)

Wet-day frequency and intensity. Following the analysis of Comment 2.1, we compared all products with the station climatologies at the 933 in-domain stations with at least 10 years of data (Table R12, Fig. R10 and Fig. R11). Each product was taken on its native grid. For this dataset we also give the leave-one-out values, in which each station is withheld.

Table R12. Station-median bias / median absolute error (%) of the 1991–2020 station climatologies (933 stations).

Product	Mean annual precipitation	Wet-day frequency (≥ 1 mm)	Wet-day intensity
This dataset, leave-one-out	+0.3 / 5.8	−0.8 / 3.7	+2.9 / 5.6
This dataset, gridded	+1.2 / 4.4	−0.7 / 2.8	+3.5 / 4.3
Manara et al. (2026)	+0.5 / 4.8	+1.3 / 3.4	−0.2 / 3.7
ARCIS	+1.4 / 6.1	+5.6 / 6.5	−4.0 / 5.4
APGD	+3.4 / 7.1	+7.1 / 7.9	−3.5 / 5.8
EMO-1arcmin	+0.6 / 7.1	+14.0 / 14.0	−12.6 / 12.9
E-OBS	−5.9 / 11.1	+2.6 / 5.8	−7.5 / 9.4

Several points emerge from Table R12:

- ARCIS, APGD, EMO-1arcmin and E-OBS overestimate wet-day frequency and underestimate intensity. In EMO-1arcmin the two errors largely cancel in the annual total.
- Manara et al. (2026) is nearly unbiased overall. Above 2000 m, however, it underestimates frequency (−6 %) and overestimates intensity (+6 %).
- The frequency bias of this dataset is within ±2.2 % in all elevation bands. Its intensity is slightly overestimated (+2 to +4 %).

- The leave-one-out error of this dataset in mean annual precipitation is 5.0–6.7 % below 1000 m, against 4.2–4.6 % for Manara et al. (2026). Above 2000 m it is 11.4 %, against 14.0–14.9 % for Manara et al. (2026), ARCIS, APGD and EMO-1arcmin and 30.1 % for E-OBS.

The stations are, at least in part, also inputs of the reference products, so their errors are likely optimistic.

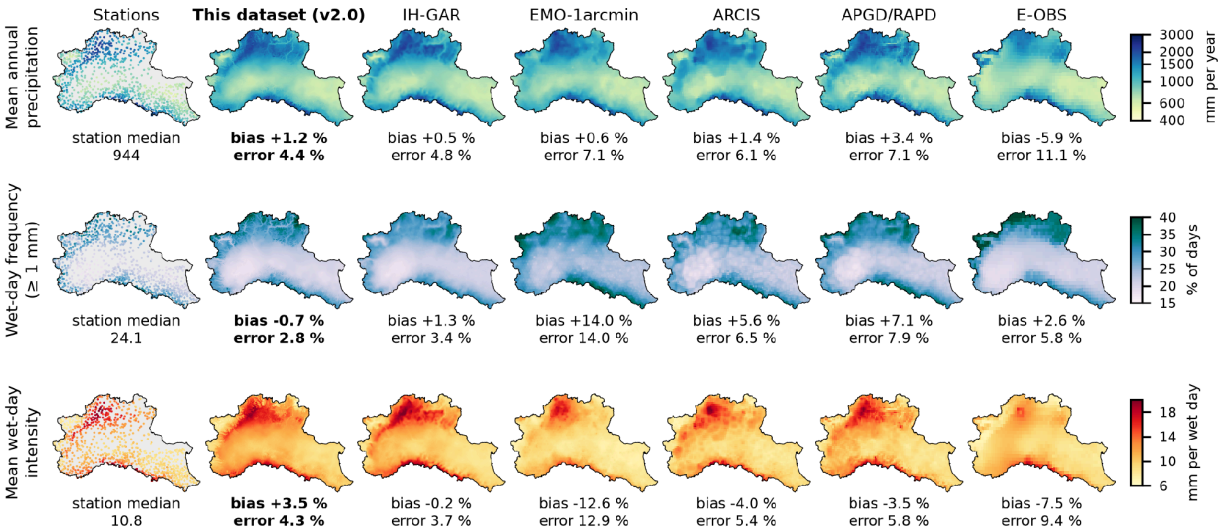


Figure R10. Mean annual precipitation, wet-day frequency (≥ 1 mm) and mean wet-day intensity, 1991–2020, of the station climatologies, this dataset and five reference products, each on its native grid (APGD: 1991–2019). Bias and error: median and median absolute value of the relative difference between the gridded value and the station climatology at the 933 in-domain stations with at least 10 years of data. Product labels: IH-GAR = Manara et al. (2026); APGD/RAPD = APGD, MeteoSwiss release RapdD.

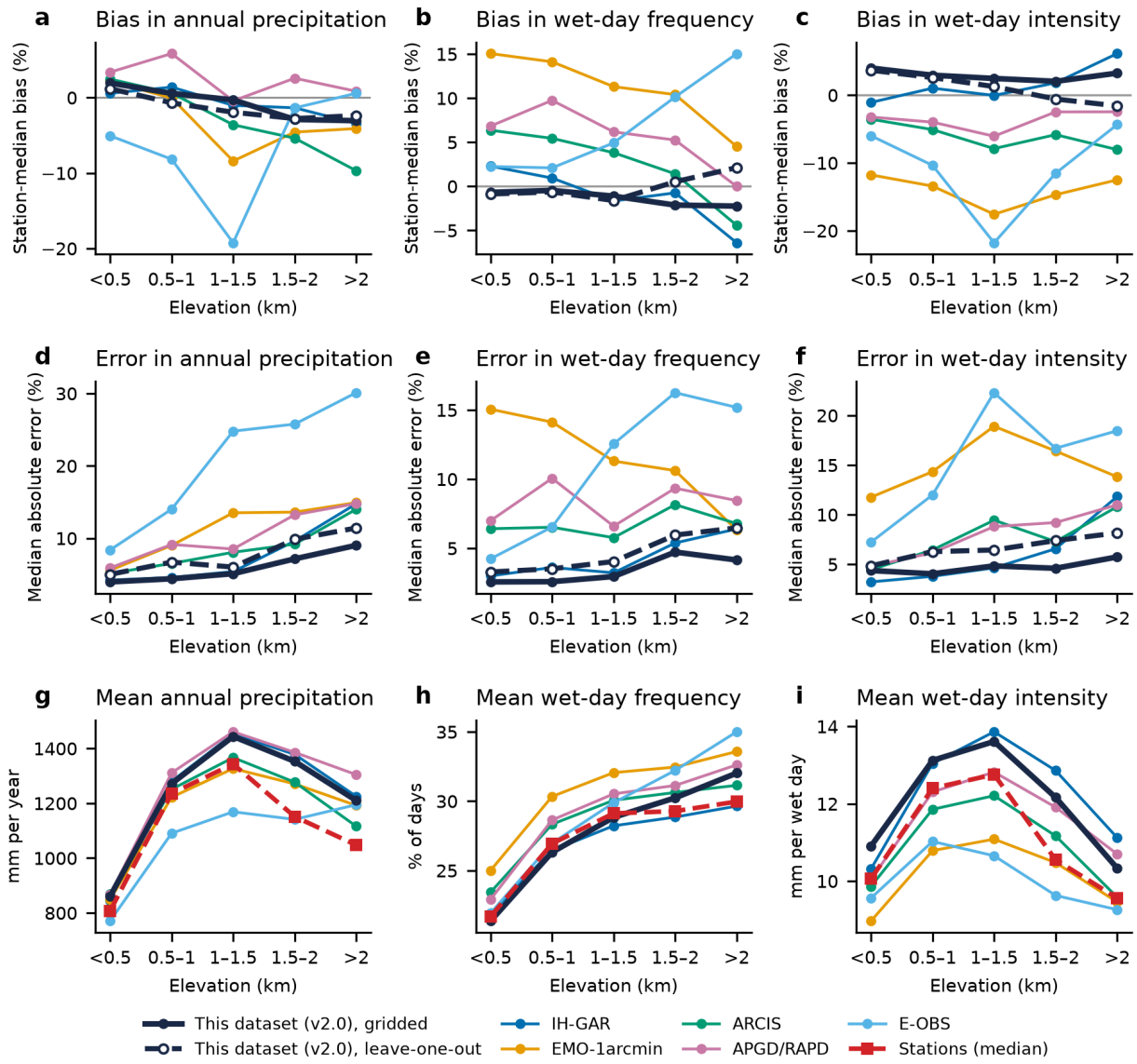


Figure R11. Station-median bias (a–c), median absolute error (d–f) and mean values (g–i) of annual precipitation, wet-day frequency and wet-day intensity by elevation band, for this dataset (gridded and leave-one-out) and five reference products, at the 933 in-domain stations (514, 229, 95, 64 and 30 stations per band). Panels g–i show area means of each gridded product over all cells of the band; the red line is the median of the station climatologies, which do not represent the band area at high elevation. Product labels: IH-GAR = Manara et al. (2026); APGD/RAPD = APGD, MeteoSwiss release RapdD.

Temperature.

- Area-mean differences: −0.25 °C with ARCIS (Tmax+Tmin)/2, −0.08 °C with EMO-1arcmin and −0.28 °C with E-OBS.
- Median daily correlation: 0.99 for all three.
- Brunetti et al. (2014): the mean difference of +0.72 °C reflects the warming between 1961–1990 and 1991–2020. The spatial correlation (0.987) and the lapse rates (Table R3) agree closely.

Event: Storm Alex (2–5 October 2020). Fig. R12 shows the four-day accumulation at the stations and in all products, and Table R13 compares the products with the station totals.

- District-mean accumulation on each reference grid (this dataset against the reference): 96 against 94 mm for ARCIS (spatial correlation 0.93), 95 against 98 mm for Manara et al. (2026) (0.97), 95 against 72 mm for EMO-1arcmin (0.83) and 96 against 90 mm for E-OBS (0.75).
- At the stations, the leave-one-out estimate of this dataset, which is independent of the station it is compared with, has a total bias of −1.8 % and underestimates the 147 stations with at least 200 mm by 11.6 %. The reference products are probably not independent of many of these stations, because they draw on the same regional networks; their station lists are not available to us, so this cannot be checked station by station. Even so, the leave-one-out underestimation of the largest totals lies between that of Manara et al. (2026) (−9.0 %) and ARCIS (−16.3 %), and is much smaller than that of EMO-1arcmin (−40.3 %) and E-OBS (−47.1 %).
- No product reproduces the local maxima. The largest station total, 647 mm at Sambughetto (Valle Strona, ARPA Piemonte), is 450 mm in this dataset (359 mm in leave-one-out), 432 mm in Manara et al. (2026) and 220 mm in ARCIS. The neighbouring stations within 10 km recorded 236–375 mm, so this maximum is very local. Both this value and the second-largest total (614 mm at Candoglia Toce, 11 km away) passed the station-day quality control, including the test for isolated high values (Sect. 3.3). The two gauges are independent of each other, and ARPA Piemonte reports more than 650 mm in the municipality of Valstrona and more than 600 mm in that of Mergozzo for this event (Arpa Piemonte, 2020). We therefore consider them a real, very local orographic maximum. Around the two stations, the ARCIS cell containing each station and its eight neighbours give 190–322 mm. The station list of ARCIS is not available to us, so we cannot determine whether these two stations enter that product. The peak value on the ARCIS grid is 473 mm in this dataset, against 550 mm in ARCIS. This is consistent with the smoothing of extremes discussed under Comment 1.3.
- The submitted version (v1.0) has a similar point bias, but its field is smoother and spreads the core of the event over a larger area (Fig. R12c). Its daily series also places the main storm day on 3 instead of 2 October, a direct consequence of the uncorrected end-day labels (see the day-definition harmonisation under Further changes); the four-day window removes this effect from Table R13.
- **Sensitivity to the event window and day definition.** After the day-label alignment applied in all comparisons, the stations and all references have their maximum on 2 October. With the wider window 1–6 October, the bias at the stations with at least 200 mm changes only slightly: −10.6 % for the leave-one-out estimate of this dataset, −10.6 % for Manara et al. (2026), −17.4 % for ARCIS, −35.8 % for EMO-1arcmin and −43.7 % for E-OBS. The underestimation in EMO-1arcmin and E-OBS is therefore not an effect of the window or of the day definition.

Table R13. Storm Alex, accumulated precipitation 2–5 October 2020, compared with the station totals at the 847 in-domain stations with complete records on the four days (mean station total 115 mm). Each product is sampled at the cell containing the station. Only the leave-one-out values are independent of the stations.

Product	Bias, all stations (%)	Bias, stations ≥ 200 mm (%)	MAE (mm)	Correlation
This dataset, v2.0 (stations included)	−1.0	−6.0	12.5	0.98
This dataset, v2.0, leave-one-out	−1.8	−11.6	23.3	0.92
Submitted version, v1.0	+1.8	−6.7	20.8	0.94
Manara et al. (2026)	−0.5	−9.0	17.0	0.95
ARCIS	−5.3	−16.3	23.3	0.89
EMO-1arcmin	−25.1	−40.3	38.7	0.76
E-OBS	−11.6	−47.1	44.1	0.68

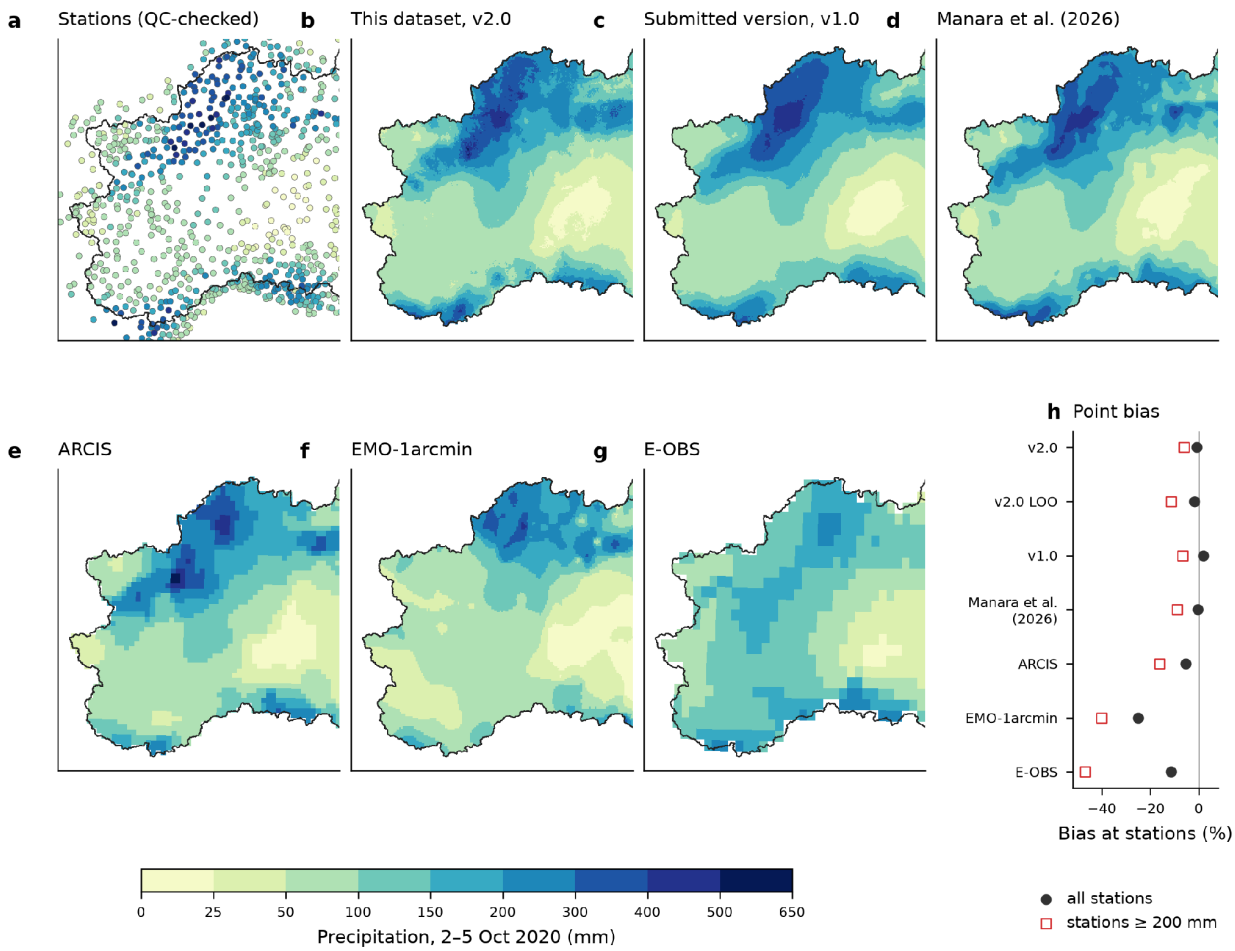


Figure R12. Storm Alex, accumulated precipitation 2–5 October 2020: (a) stations with complete records (QC-checked), (b–g) this dataset (v2.0), the submitted version (v1.0) and four reference products, each on its native grid and masked to the Po River District, and (h) bias of each product at the in-domain stations, for all stations and for the stations with at least 200 mm. v2.0 LOO: leave-one-out estimate, the only one independent of the stations.

4. Final comment

Comment 4.1

In light of the described limitations, I believe the paper cannot be accepted for publication in ESSD.

Response. We understand the reviewer's concern about the submitted version. We believe the rebuilt dataset and revised manuscript now address the three main problems:

- the weight-based spatial support, used as a proxy for effective resolution, is quantified, documented and distributed as maps;
- the interpolation now uses the anomaly method with PRISM-type normals, for precipitation separately for occurrence and amount. Its robustness to missing data has been tested, and the results are reported as they are. The method is robust for temperature and at the strong local precipitation gradient between Aosta and Gran San Bernardo. In the 50 % deletion experiments, the daily precipitation errors of the two-part scheme increase about as much as, or slightly more than, those of direct interpolation, while its wet-day frequency and intensity remain unaffected;

- the dataset is compared with all of the reference products the reviewer recommended, including a decomposition into wet-day frequency and intensity.

We also state the remaining limitations clearly. We thank the reviewer; these comments led to a much stronger dataset and manuscript.

Further changes not requested by the reviewer

- **Day-definition harmonisation (Sect. 3.2, Fig. 4).**
  - **Problem.** A large fraction of the station records were labelled with the end day of the 24-h accumulation.
  - **Detection.** Lag correlation (−1, 0, +1 day) of each station-year with EMO-1arcmin (precipitation) and with both EMO-1arcmin and E-OBS (temperature, day-to-day changes). Stations without a clear reference classification were tested against their nearest classified neighbours.
  - **Independent check.** The result was verified on neighbouring station pairs alone, without any gridded reference (see Effect).
  - **MeteoSwiss.** SwissMetNet totals are 00–00 UTC. The manual network is 06–06 UTC; it was kept under its start day, the calendar day of largest overlap, and was not shifted. SwissMetNet days without the 00–00 UTC total were filled with the 06–06 UTC total. The 33 manual stations and these filled days (together 4.4 % of the station-days 1991–2020) therefore keep a 6-h offset relative to the other networks; this is stated as a limitation in Sect. 7.
  - **Correction.** 934 precipitation and 382 temperature records were relabelled to the calendar day they cover. In 16 precipitation records the convention changed during the period.
  - **Effect.** Among neighbouring station pairs, the share still offset by one day decreased from 5.0 % to 0.45 % for precipitation and from 6.6 % to 0.62 % for temperature.
- **Quality control (Sect. 3.3).** Quality control is applied in two stages.
  - **Stage 1: station and series screening** (as in the submitted version). It removed 61 precipitation and 25 temperature stations, and 0.21 % (precipitation) and 1.84 % (temperature) of the values at the retained stations.
  - **Stage 2: new, fully automatic station-day QC.** An iterative procedure based on leave-one-out errors checks for false zeros, isolated extremes, accumulations after gaps, stuck sensors, range errors and spatial temperature outliers. It flagged 0.031 % of the precipitation and 0.057 % of the temperature station-days.
- **Calibrated uncertainty (Sect. 4.4).** Within ±1.96σ, the raw kriging standard deviation covered 96 % (precipitation) and 83 % (temperature) of the leave-one-out errors. It is now scaled by a factor estimated per season and class (for precipitation, the class of the predicted value: 0, 0–1, 1–5, 5–20, 20–50, > 50 mm). Cross-validated on independent halves of the station network, the calibrated coverage is 95 % for both variables and 94.8–95.0 % in every precipitation class, including > 50 mm, where the raw coverage was 89 %.

  For precipitation, σ is the scaled kriging standard deviation of the daily interpolation. It does not represent the uncertainty of the wet/dry decision explicitly. On 47 % of the station-days predicted dry, σ is zero, because all neighbouring stations are dry. This is stated in Sect. 4.4 and in the README.
- **Coordinate reference system.** The grid is now defined in ETRS89-LAEA (EPSG:3035).
- **Corrections to the text.**
  - We corrected the definition of the KGE (Gupta et al., 2009) and the mean annual precipitation reported in Fig. 8.

- The statement that no comparable product exists was narrowed (see Comment 1.5).
- The statement that the dataset is suitable for trend analysis was removed (see Comment 2.2), and so was the reference to network sensitivity experiments in Sect. 5.1 of the submitted manuscript, which were not shown.
- The discussion of Fig. 10 of the submitted manuscript stated that coarser grids have fewer wet days and higher intensities. The opposite is true, and our own comparison shows it: spatial smoothing increases the number of wet days and reduces their intensity (Comment 3.1). The paragraph was rewritten.
- TerraClimate is a monthly, not a daily, product; the Introduction was corrected.
- The variogram fitting criterion is now stated: weighted least squares between the empirical and the model semivariogram, with weights proportional to the number of station pairs in each lag class, minimised by particle swarm optimisation.
- The grid is defined in ETRS89-LAEA; the Usage notes now refer to projected coordinates (EPSG:3035) instead of latitude and longitude.

- **Data and code availability (Sect. 8).** Version 2.0 is published on Zenodo as a new version of the submitted record. It contains:
  - the daily fields with uncertainty;
  - the 1991–2020 monthly normals of precipitation, wet-day frequency and temperature;
  - the weight-based spatial-support maps;
  - aggregate validation tables;
  - a README describing the limitations.

  The processing code is released openly with its own DOI. The interpolation, including the two-part scheme and its monthly occurrence thresholds, is also implemented in the open-source GEOframe framework (Krigings component, run from the OMS console). For 1991–2020 it reproduces the daily precipitation grids of the dataset in at least 99.9998 % of the cell-days of every year.

Summary of changes

Table R14. Summary of the responses and their location in the revised manuscript.

Comment	Main change	Location in the revised manuscript
0	Dataset rebuilt (v2.0): anomaly method with PRISM-type normals; two-part (occurrence × amount) scheme for precipitation; local variograms and lapse rates	Abstract; Sect. 3–8
1.1	Network extended with MeteoSwiss stations; record lengths reported; domain-wide lapse rate replaced	Sect. 2.2, 4.2; Figs. 2–3
1.2	Mean annual precipitation corrected (1051 mm); climatological gradients improved	Sect. 6.4; Fig. 13
1.3	Kriging-weight support distances (2/3, 1/2, 1/3) by region and decade, defined on positive weights; smoothing of daily variability and extremes quantified	Sect. 5.2, 6.1, 6.2; Fig. 10; Table 3
1.4	Radar comparison not possible (no independent radar); event comparison (Storm Alex); radar as future work	Sect. 6.4, 7
1.5	Effective resolution comparable to ARCIS acknowledged; novelty statement narrowed	Introduction; Sect. 6.2, 7
1.6	Weight-based spatial-support maps and calibrated uncertainty distributed; usage guidance	Sect. 4.4, 7, 8; Fig. 5
1.7	PRISM-type normals for precipitation, wet-day frequency and temperature; local daily lapse rate; tests of physiographic predictors (elevation difference, facet, topographic position)	Sect. 4.2; Fig. 15

2.1	Anomaly method with missing-data handling of the normals; two-part precipitation scheme and its evaluation against the submitted and single-ratio versions; leave-one-out without leakage	Sect. 4.1–4.3, 6.1; Figs. 6–9
2.2	Trend-analysis claim removed; statement on homogenisation and intended use; temporal consistency with references and at long-record stations	Sect. 6.5, 7; Table 7
2.3	50 % deletion experiments (30 realisations each, 2011–2020): temporal missingness (station-days removed, normals kept) and network sparsity (stations removed, normals rebuilt); errors and change of the gridded fields for both methods; four-year temporal-missingness check for precipitation (2017–2020, 10 realisations)	Sect. 5.3, 6.3; Fig. 11; Table 4
2.4	Aosta–Gran San Bernardo point test (annual totals, P99, annual maxima); GSB station added	Sect. 5.3, 6.3; Fig. 12; Table 5; Fig. S1
3.1	Comparison with APGD, ARCIS P and T, Manara et al. (2026), Crespi et al. (2018), Brunetti et al. (2014), EMO-1arcmin and E-OBS, including wet-day frequency and intensity and the Storm Alex event against stations	Sect. 5.4, 6.4; Figs. 13–15; Table 6
4.1	All three deficits addressed	—

References

Arpa Piemonte: Eventi alluvionali in Piemonte. Evento del 2–3 ottobre 2020, Arpa Piemonte, Torino, https://www.arpa.piemonte.it/sites/default/files/media/2024-07/rapporto_evento_02_ottobre_2020_.pdf, 2020.

Brunetti, M., Maugeri, M., Nanni, T., Simolo, C., and Spinoni, J.: High-resolution temperature climatology for Italy: interpolation method intercomparison, Int. J. Climatol., 34, 1278–1296, https://doi.org/10.1002/joc.3764, 2014.

Cornes, R. C., van der Schrier, G., van den Besselaar, E. J. M., and Jones, P. D.: An ensemble version of the E-OBS temperature and precipitation data sets, J. Geophys. Res.-Atmos., 123, 9391–9409, https://doi.org/10.1029/2017JD028200, 2018.

Crespi, A., Brunetti, M., Lentini, G., and Maugeri, M.: 1961–1990 high-resolution monthly precipitation climatologies for Italy, Int. J. Climatol., 38, 878–895, https://doi.org/10.1002/joc.5217, 2018.

Daly, C., Halbleib, M., Smith, J. I., Gibson, W. P., Doggett, M. K., Taylor, G. H., Curtis, J., and Pasteris, P. P.: Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States, Int. J. Climatol., 28, 2031–2064, https://doi.org/10.1002/joc.1688, 2008.

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

Haylock, M. R., Hofstra, N., Klein Tank, A. M. G., Klok, E. J., Jones, P. D., and New, M.: A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006, J. Geophys. Res., 113, D20119, https://doi.org/10.1029/2008JD010201, 2008.

Isotta, F. A., Frei, C., Weilguni, V., et al.: The climate of daily precipitation in the Alps: development and analysis of a high-resolution grid dataset from pan-Alpine rain-gauge data, Int. J. Climatol., 34, 1657–1675, https://doi.org/10.1002/joc.3794, 2014.

Manara, V., et al.: A new daily high-resolution gridded precipitation dataset for the Italian Greater Alpine Region (1951–2023): spatio-temporal analysis and hydrological applications, J. Hydrol. Reg. Stud., 67, 103776, https://doi.org/10.1016/j.ejrh.2026.103776, 2026.

Newman, A. J., Clark, M. P., Craig, J., Nijssen, B., Wood, A., Gutmann, E., Mizukami, N., Brekke, L., and Arnold, J. R.: Gridded ensemble precipitation and temperature estimates for the contiguous United States, J. Hydrometeorol., 16, 2481–2500, https://doi.org/10.1175/JHM-D-15-0026.1, 2015.

Pavan, V., Antolini, G., Barbiero, R., et al.: High resolution climate precipitation analysis for north-central Italy, 1961–2015, Clim. Dynam., 52, 3435–3453, https://doi.org/10.1007/s00382-018-4337-6, 2019.

Pavan, V., et al.: A new operational dataset of gridded minimum and maximum temperature over North-Central Italy 1991 to present, Clim. Serv., 43, 100689, https://doi.org/10.1016/j.cliser.2026.100689, 2026.

Schwarb, M.: The Alpine precipitation climate: evaluation of a high-resolution analysis scheme using comprehensive rain-gauge data, PhD thesis, ETH Zürich, 2000.

Salamon, P., Sperzel, T., Gomes, G. R., Radke-Fretz, M., Lemke, C.-D., Russo, C., Schweim, C., Zsoter, E., Dosio, A., Vomero, M., Ziese, M., and Grimaldi, S.: EMO-1: an improved version of the high-resolution multi-variable gridded meteorological dataset for Europe, Earth Syst. Sci. Data Discuss. [preprint], https://doi.org/10.5194/essd-2025-723, in review, 2026.