






Article

Adjustment Methods Applied to Precipitation Series with Different Starting Times of the Observation Day

Francesca Becherini ¹, Claudio Stefanini ^{2,*}, Antonio della Valle ³, Francesco Rech ⁴, Fabio Zecchini ⁴
and Dario Camuffo ³

- ¹ National Research Council-Institute of Polar Sciences, Via Torino 155, 30172 Venice Mestre, Italy; francesca.becherini@cnr.it
- ² Department of Environmental Sciences, Informatics and Statistics, Ca' Foscari University of Venice, 30170 Venice Mestre, Italy
- ³ National Research Council-Institute of Atmospheric Sciences and Climate, Corso Stati Uniti 4, 35127 Padua, Italy; a.dellavalle@isac.cnr.it (A.d.V.); d.camuffo@isac.cnr.it (D.C.)
- ⁴ Regional Agency for Environmental Protection and Prevention of Veneto, Via Ospedale Civile 24, 35121 Padua, Italy; francesco.rech@arpa.veneto.it (F.R.); fabio.zecchini@arpa.veneto.it (F.Z.)
- * Correspondence: claudio.stefanini@unive.it

Abstract: The study of long precipitation series constitutes an important issue in climate research and risk assessment. However, long datasets are affected by inhomogeneities that can lead to biased results. A frequent but sometimes underestimated problem is the definition of the climatological day. The choice of different starting times may lead to inhomogeneity within the same station and misalignment with other stations. In this work, the problem of temporal misalignment between precipitation datasets characterized by different starting times of the observation day is analyzed. The most widely used adjustment methods (1 day and uniform shift) and two new methods based on reanalysis (NOAA and ERA5) are evaluated in terms of temporal alignment, precipitation statistics, and percentile distributions. As test series, the hourly precipitation series of Padua and nearby stations in the period of 1993–2022 are selected. The results show that the reanalysis-based methods, in particular ERA5, outperform the others in temporal alignment, regardless of the station. But, for the periods in which reanalysis data are not available, 1-day and uniform shift methods can be considered viable alternatives. On the other hand, the reanalysis-based methods are not always the best option in terms of precipitation statistics, as they increase the precipitation frequency and reduce the mean value over wet days, NOAA much more than ERA5. The use of the series of a station near the target one, which is mandatory in case of missing data, can sometimes give comparable or even better results than any adjustment method. For the Padua series, the analysis is repeated at monthly and seasonal resolutions. In the tested series, the adjustment methods do not provide good results in summer and autumn, the two seasons mainly affected by heavy rains in Padua. Finally, the percentile distribution indicates that any adjustment method underestimates the percentile values, except ERA5, and that only the nearby station most correlated with Padua gives results comparable to ERA5.

Keywords: daily precipitation; time of observation; time series; aggregation method; adjustment method; hourly dataset; daily extremes



Citation: Becherini, F.; Stefanini, C.; della Valle, A.; Rech, F.; Zecchini, F.; Camuffo, D. Adjustment Methods Applied to Precipitation Series with Different Starting Times of the Observation Day. *Atmosphere* **2024**, *15*, 412. <https://doi.org/10.3390/atmos15040412>

Academic Editor: Agnieszka Krzyżewska

Received: 18 February 2024

Revised: 15 March 2024

Accepted: 19 March 2024

Published: 26 March 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Long-term precipitation records are of great importance in climate research and risk assessment. Therefore, quality-controlled and homogenized data are needed for improved climate-related decision-making processes [1]. In addition, due to the spatial variability in precipitation, the availability of ground-based observations and high spatial station density are basic requirements to provide reliable results [2].

Recently, there has been growing interest in series of precipitation at daily [3,4] and sub-daily [5] resolutions, as these are the typical timings of extreme events. The study

of weather extremes is crucial for the scope of climatological analyses, climate change impact assessment, and future climate projections. The introduction of automatic rain gauges in the mid-20th century has allowed continuous observations, providing sub-daily and sub-hourly values. Before the standardization operated by WMO, founded in 1951, the number and times of daily observations were not standardized and depended on the station, period, and even the observer. Data were collected manually, so the observation time depended on several factors, some subjective, e.g., the health and other commitments of the observer, the weather, and the accessibility to the instrument [6]. Therefore, in the early instrumental period, the investigation of climate change effects on short-duration events, including the extreme ones, is often difficult [7], as only daily values are available.

Automatic weather stations introduce more flexibility in the definition of the climatological day, and in the methods used to calculate climatological parameters over time, e.g., daily average, as more values are available. In fact, even if the duration of the day is consistent between different stations, i.e., 24 h, the starting time of the climatological day may not be. For instance, if the climatological day starts at 08 LT (i.e., 8:00 am Local Time), the daily precipitation total is the amount collected from 08 LT on the previous day to 08 LT on the reporting day. For long-term analysis, especially in periods in which data are scarce, it may be necessary to use composite datasets derived from a combination of stations close to the target one. The misalignment in daily precipitation totals due to a different definition of the climatological day affects the compatibility between different stations, but also within the same station, thus limiting the possibility of using long-term observation-based data for spatial analyses and model development [8]. WMO [9] states the following: (i) The definition of the climatological day has to be clearly stated in the metadata of the station. The scarcity or lack of metadata in the early instrumental period makes it difficult to identify a bias due to data misalignment, if any, and, consequently, to apply the most appropriate adjustment. (ii) Any change in the definition of the climatological day should be avoided, as it can lead to potential inhomogeneity. Such a change is particularly crucial when extreme events are considered and sub-daily values are not available. For identifying extreme daily precipitation events, WMO [10] recommends the comparison of the daily totals with certain fixed thresholds or percentiles. An extreme event may be lost if it had been broken by the change in the climatological day.

The issue of the time of observation adjustment emerged during the reconstruction of the precipitation series of Padua, one of the longest precipitation series in Italy (for details, see [6] and references herein). That work required facing several problems that often affect early precipitation series, e.g., missing readings, cumulative amounts [11], and gaps [12]. One of the most used filling-gaps method consists of the use of contemporary records from one or more stations in the same climatic area [12]. The definition of the climatological day of these datasets is very likely absent in the early instrumental period, and often different in the modern period. For the 20th century, three main datasets are available: (1) the Meteorological Observatory of the Water Magistrate (WM) from 1920 to the 1990s; (2) the Meteorological Service of the Italian Air Force (AF) at Padua Airport from 1951 to 1990; and (3) the Regional Agency for the Prevention and Protection of the Environment in the Veneto Region (ARPA), from 1980 up to now. These three precipitation datasets are characterized by different definitions of the climatological day: WM set the start time at 09 LT; AF at 00 UTC, i.e., 01 LT; and ARPAV at 00 LT of the target day. Therefore, the reconstruction of the precipitation series of Padua since 1920 requires the adjustment of datasets characterized by different climatological days.

The problem of the time of observation misalignment is well known, and several adjustment methods have been proposed. One of the simplest methods was to shift morning observations back to one calendar day [13,14]. Other methods foresaw disaggregating daily precipitation amounts to hourly and then aggregating back following a different definition of the climatological day [13–16]. Disaggregation can be performed using actual hourly observations [13] or under the assumption that a daily total is distributed uniformly across all hours in a 24 h period [14,16]. These methods improved interstation temporal

correlations in the following scenarios: adjustment of daily precipitation from a morning to midnight-to-midnight observation time [13,14,16]; adjustment of daily precipitation from an afternoon to midnight-to-midnight observation time [14]; and adjustment of daily precipitation from an afternoon to a morning time [14]. Nevertheless, they showed some drawbacks, e.g., the increase in precipitation frequency and temporal autocorrelation, and the decrease in average intensity and extremes [13–16]. In addition, the absolute optimal method cannot be established, because it depends on the dataset, its specific application, and the local climatology. Therefore, starting from the results of previous studies, in this work, the problem of the inconsistencies in the observing times between stations was explored more in depth, to find a reliable method to overcome the issue of the temporal misalignment in the three available precipitation datasets of Padua, thus providing a unique aligned series from the early-20th century up to the present day. To achieve this main aim, several modern (1993–2022) datasets at an hourly resolution were used as testing sets, with the following specific objectives:

- (i) To apply adjustment methods already tested in the literature to modern datasets;
- (ii) To test for the first time two further adjustment methods, based on reanalysis;
- (iii) To compare the alignment of the series adjusted to the original series of stations located near the target one;
- (iv) To determine the impact of all the methods considered on the identification of extreme days;
- (v) To explore the feasibility of the application of the adjustment methods considered to WM and AF series.

The identification of the best method to adjust the three available precipitation datasets of Padua is extremely important for long-term climatic studies, as the Meteorological Observatory of the Water Magistrate constitutes a precious source of precipitation data at a daily resolution for several Italian locations, from 1920 to the 1990s. Moreover, the described methodology is general, not case-specific, and it can be applied to precipitation datasets of other countries and periods. In addition, the two new methods proposed, based on reanalysis, have a large potential of application to modern series.

2. Materials and Methods

2.1. Datasets

In 1980, the Department of Biology of the Padua University installed a weather station in the historical Botanical Garden, in the city center, that in May 2000 passed under the control of the Regional Agency for the Prevention and Protection of the Environment in the Veneto Region (ARPAV). This constitutes the main source of precipitation amount in Padua at an hourly resolution until present. The reference station for this study is the one named “Orto Botanico” (OB, the Botanical Garden), located in the city center (Figure 1), which can be considered the continuation of the former “Specola” station [6]. Five other stations (Table 1) were selected considering their proximity to OB (Figure 1). In March 2019, the OB station was closed and the meteorological instruments were moved to another place, called “Padova CUS”, about 2 km away (Table 1). The stations “Orto Botanico” and “Padova CUS” were considered as only one station because of their proximity and are named simply “Padua” (Pd) throughout the text. The datasets from the considered stations cover the 1993–2022 period, except for the Padua series that starts in October 1993, and Tribano in January 1996. The percentage of data available during the working period of each station is indicated in Table 1. Further information regarding the stations, such as yearly average precipitation amount and number of rainy days, is reported in Tables S1 and S2 of Supplementary Material, respectively.

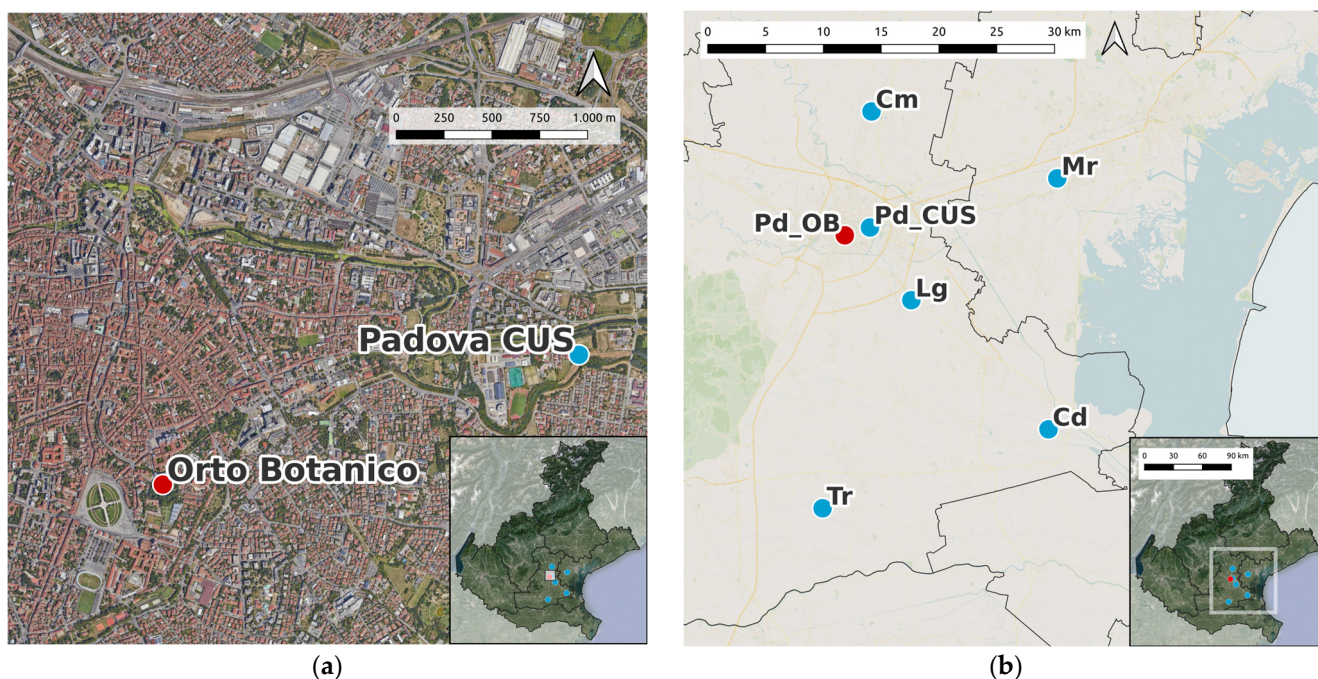


Figure 1. Location of the ARPAV meteorological stations listed in Table 1: (a) Veneto region; (b) zoom on Padua center.

Table 1. ARPAV meteorological stations in proximity to Padua, and data availability with respect to the 1993–2022 period.

Name	Acronym	Elevation (m a.g.l.)	Lat	Long	Distance from OB (km)	Data Availability
Orto Botanico		12	45.40	11.88	0	October 1993–December 2022 (97.1%)
Padova CUS	Pd	12	45.40	11.91	2.3	January 1993–December 2022 (99.5%)
Legnaro	Lg	7	45.35	11.95	8.0	January 1993–December 2022 (99.1%)
Campodarsego	Cm	16	45.49	11.91	11.0	January 1993–December 2022 (99.4%)
Codevigo	Cd	0	45.24	12.10	24.4	January 1993–December 2022 (99.4%)
Mira	Mr	3	45.44	12.12	19.0	January 1996–December 2022 (99.0%)
Tribano	Tr	3	45.19	11.85	23.8	January 1996–December 2022 (99.0%)

2.2. Methodology

The overall methodology applied to solve the issue of the temporal misalignment of a precipitation series visualized in Figure 2 includes the main steps described hereunder. The hourly precipitation measurements of the target station were taken as input of the process.

- (i) Application of the selected adjustment methods to the series of the target station and calculation of performance indicators. Three out of five adjustment methods are derived from the literature while two further methods, based on reanalysis, are tested for the first time;
- (ii) If there is at least an adjustment method that is better performing than the misaligned series, go to the next step; otherwise, return to the misaligned series (output 1, end of process);
- (iii) If there are contemporary precipitation data of a nearby station available, go to the next step; otherwise, select the series adjusted with the best performing method (output 2, end of process);

- (iv) Calculate performance indicators of the series of nearby stations;
- (v) If there is at least one nearby series that is better performing than the misaligned series, go to the next step; otherwise, return to the series adjusted with the best performing method (output 2, end of process);
- (vi) If the series of the nearby station is better performing than the adjusted series, select the series of the nearby station (output 3, end of process); otherwise, select the series adjusted with the best performing method (output 2, end of process).

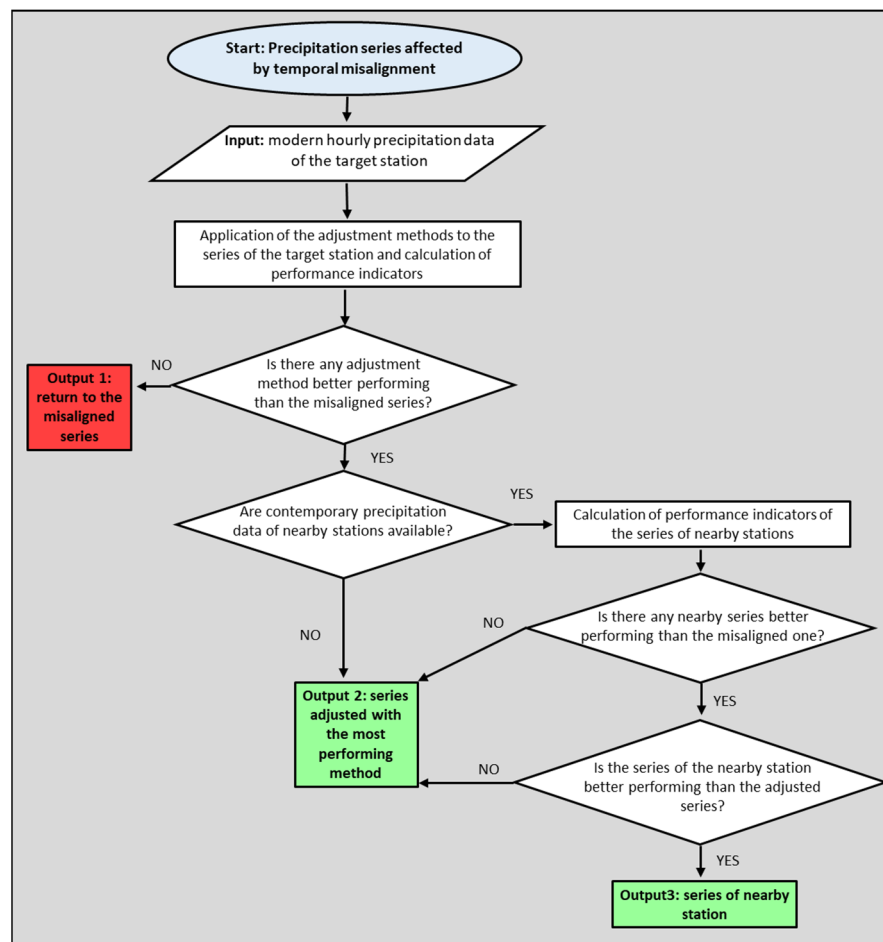


Figure 2. Flowchart of the methodology used to evaluate the adjustment methods of precipitation series with different definition of climatological day.

2.3. Homogeneity Tests

Firstly, several homogeneity tests were applied to the precipitation datasets to detect discontinuities and regime shifts. The most used absolute tests were selected to identify change points based on shifts in the mean: Buishand [17], Pettitt [18], and von Neumann ratio [19]. The yearly amounts and the monthly anomalies were set as testing variables. Relative tests are generally favored over absolute ones as they use the difference time series of the target station with neighboring stations to identify breaks or change points [20,21]. These reference series are supposed to have the same climate as the target station and thus can be used to detect inhomogeneities [22]. In modern homogenization tests, reference series themselves do not need to be homogeneous but encompass the same climatic signal as the target [1]. The relatively new software package Climatol, version 4.0.7, developed by the Spanish State Meteorological Agency (AEMET) under the R programming language [23], was used as the relative test. This package holds functions for quality control, homogenization, and missing data infilling of climatological series, and it has already been applied to precipitation series [1,24]. The homogenization is based on the Standard

Normal Homogeneity Test (SNHT) [25], considering other series as reference to detect inhomogeneities in the test series: when the SNHT statistics are greater than a prescribed threshold, the series is split at the point of maximum SNH, moving all data before the break to a new series that is incorporated into the data. This procedure is performed iteratively, splitting only the series with the higher SNHT values at every cycle, until no series is found inhomogeneous. As the SNHT was originally designed to find a single breakpoint in a series, it was first applied to stepped overlapping temporal windows, and then to the complete series. In the final stage, the method infills missing data in all homogeneous series and sub-series. As reference series to be used by the algorithm, data of stations listed in Table 1 were used. To infill the missing data and compute the homogeneity tests, the algorithm does not use the proximity criterion but evaluates the correlation between datasets. The Buishand and Pettitt tests are usually more performing when a break appears in the middle of the series, whereas the ability of SNHT is in favor of identifying inhomogeneities at the beginning/end [26]. Climatol can detect multiple change-points, as the process is iterative, and the procedure is applied to all the sub-series in which the test series is decomposed by the breakpoints detected at each step. Overall, absolute and relative tests were employed to obtain more reliable results and to take advantage of the specific features of each method.

The instrumental threshold of a rain gauge has a significant influence on the distribution of precipitation, more enhanced in frequency and less in amount [27]. Therefore, the results of the analyses depend on the choice of the threshold to define the wet days. In recent years, two ARPAV stations, i.e., Lg since 11 October 2005 and Cm since 4 May 2009, have been equipped with heated funnels, and this change had an impact on minor accumulations, as the false amounts up to 0.6 mm, mostly caused by dew or fog, were reduced [28]. To avoid bias due to this change, in the following statistical analysis, only data above the threshold of 1 mm/day were considered, following the WMO definition of a wet day [9].

2.4. Adjustment Methods

In the dataset with the climatological day starting at 09 LT (herewith named 9–9 dataset), the precipitation total of the target day, d_j , is the sum of the quantities collected from 09 LT on the previous day, d_{j-1} , to 09 LT on the target day. In modern series, the climatological day generally coincides with the civil local day, i.e., the 24 h interval from one midnight to the following midnight, subsequently shortened to 0–24 series. Starting from hourly observations, five different daily aggregation methods were considered and compared to the 0–24 series, which was thus considered as reference:

(1) 9–9

9–9 daily series is considered as is, i.e., daily precipitation total is the sum of the hourly amounts collected from 9 LT of d_{j-1} to 9 LT of d_j .

(2) 9–9 1-day shift (named simply “1 day” in the following) [13]

This method shifts the daily amounts of the 9–9 series back one calendar day, because most of the daily amount of the 9–9 series is collected in the previous day. Therefore, the precipitation amount of the target day, d_j , is simply associated with the previous day, d_{j-1} .

(3) 9–9 shift uniform (named simply “unif”) [15]

This method reapportions 9–9 daily totals from a 2-day moving window surrounding the target date, $P_{adj,j} = (P_j \cdot F_j) + (P_{j+1} \cdot F_{j+1})$, where $P_{adj,j}$ is the adjusted amount for the target day j ; P_j and P_{j+1} are the original 9–9 reported daily totals for the target and next days, respectively; and F_j and F_{j+1} are the fractions of P_j and P_{j+1} , respectively, to be included in the estimate of $P_{adj,j}$. Because the uniform method assumes that a reported daily total is distributed uniformly across all hours within its respective 24 h period, F_j and F_{j+1} are determined directly by the number of hours of overlap

between the 24 h periods, represented by P_j and P_{j+1} , and the new $P_{adj,j}$, i.e., $F_j = 9$ and $F_{j+1} = 15$ (Figure 3).

	DAY BEFORE (j-1)																								TARGET DAY (j)																								DAY AFTER (j+1)																							
hours of the day	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	1	2	3	4	5	6	7	8	9																								
9-9 series	9-9 TARGET DAY (j)																								9-9 DAY AFTER (j+1)																																															
adjusted series																									9 h of the 9-9 target day j									15 h of the 9-9 day after j+1																																						

Figure 3. Overlapping between the hours of observation of 9–9 series and the adjusted series, with the latter composed of $F_j = 9$ h of the target day j in the 9–9 series, and $F_{j+1} = 15$ h of the day after $(j + 1)$ in the 9–9 series.

(4) 9–9 shift ERA5 (named “ERA5”) [29]

Like method (3) but F_j and F_{j+1} are determined by means of the reanalysis (0.25° resolution, 1940–today). The simulated 9–9 amount of the target day and of the day after is determined using hourly reconstructed data, and the fractions of precipitation that occurred on those days are calculated. Then, the fractions F_j and F_{j+1} , are multiplied by the 9–9 amount of day j and day $j + 1$, respectively, and the results are added to obtain the total amount of the target day j .

(5) 9–9 shift NOAA (named “NOAA”) [30]

Like method (4) but using the NOAA 20CRv3 reanalysis to determine the fractions F_j and F_{j+1} . Unlike ERA5, this dataset uses only pressure observations as input and monthly sea surface temperatures as boundary conditions, covers the period 1836–2015 (experimentally extended to 1806), has a coarser resolution ($\sim 0.75^\circ$), and provides 3-hourly data.

2.5. Performance Indicators

The adjustment methods presented in Section 2.4 were validated in two main aspects: (i) temporal alignment between the original and adjusted series, i.e., if and how much the adjustment methods applied to the 9–9 series of the target station improve the alignment with the 0–24 series of the same station; and (ii) precipitation statistics.

The indicators used to evaluate temporal alignment and precipitation statistics are listed in Tables 2 and 3, respectively, including formulas to calculate them and possible values assumed. The symbols used in the formulas and their interpretation are described at the bottom of the tables.

Table 2. Indicators used to evaluate the temporal alignment of the test and reference series.

Name	Short Name	Formula	Range Values
Root-Mean-Square Error	RMSE	$\sqrt{\frac{\sum_{i=1}^N (y_i - o_i)^2}{N}}$	≥ 0 (ideal)
Mean Absolute Error	MAE	$\frac{1}{N} \sum_{i=1}^N y_i - o_i $	≥ 0 (ideal)
Normalized Mean Absolute Error	NMAE	$\frac{\sum_{i=1}^N y_i - o_i }{\sum_{i=1}^N o_i}$	≥ 0 (ideal)
Brier Score	BS	$\frac{1}{N} \sum_{i=1}^N (Y_i - O_i)^2$	0 (ideal)—1
Pearson’s correlation coefficient	cor_P	$\frac{cov(y,o)}{\sigma_y \sigma_o} = \frac{\sum_{i=1}^N y_i o_i - N \bar{y} \bar{o}}{\sqrt{\sum_{i=1}^N y_i^2 - N \bar{y}^2} \sqrt{\sum_{i=1}^N o_i^2 - N \bar{o}^2}}$	0–1 (ideal)
Spearman’s rank correlation	cor_S	$\frac{cov(R(y),R(o))}{\sigma_{R(y)} \sigma_{R(o)}}$	0–1 (ideal)
Kendall’s rank correlation	cor_K	$\frac{N_C - N_D}{N(N-1)/2}$	0–1 (ideal)
Tail dependence measure	$\chi(u = 0.95)$	$P(y > u o > u)$	0–1 (ideal)
Accuracy	ACC	$\frac{TP + TN}{P + N}$	0–1 (ideal)
Heidke Skill Score	HSS	$\frac{2((TP \cdot TN) - (FN \cdot FP))}{P(FN + TN) + N(TP + FP)}$	≤ 1 (ideal)

Table 3. Indicators used to evaluate the precipitation statistics of the test and reference series.

Name	Short Name	Formula	Range Values
mean precipitation value over wet days	mwet	$\left[\left(\frac{\sum_{i=1}^N Y_i}{\sum_{i=1}^N Y_i} - \frac{\sum_{i=1}^N O_i}{\sum_{i=1}^N O_i} \right) / \left(\frac{\sum_{i=1}^N O_i}{\sum_{i=1}^N O_i} \right) \right] \cdot 100$	$\geq -100\%$ (0 ideal)
frequency of wet days	freq	$\left[\frac{\sum_{i=1}^N Y_i - \sum_{i=1}^N O_i}{\sum_{i=1}^N O_i} \right] \cdot 100$	$\geq -100\%$ (0 ideal)

- Root-Mean-Square Error (RMSE) is the quadratic mean of the differences between the observations and the values predicted by the model (in this case, the adjustment methods):

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (y_i - o_i)^2}{N}}$$

where N is the number of observations, y_i is the value predicted by the adjustment method considered, and o_i is the observed one.

- Mean Absolute Error (MAE) is a common indicator to measure the errors between values predicted by the model and the observations:

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - o_i|$$

- Normalized Mean Absolute Error (NMAE) is a validation metric to compare the MAE of (time) series with different scales. As the precipitation series of the stations listed in Tabel 1 have different temporal averages, both MAE and NMAE were calculated. NMAE is the ratio of MAE to mean daily precipitation:

$$NMAE = \frac{\sum_{i=1}^N |y_i - o_i|}{\sum_{i=1}^N o_i}$$

- Brier Score (BS) compares the predicted probability of an event to observations. As precipitation reconstruction does not provide probabilities, Y_i and O_i are both binary with 1 = rain and 0 = no rain [31]. Therefore, BS is the percentage of time steps wrongly assigned as wet or dry, calculated as

$$BS = \frac{1}{N} \sum_{i=1}^N (Y_i - O_i)^2$$

Only the mismatches between Y_i and O_i (wet the first and dry the second or viceversa) contribute with non-zero terms.

- Pearson’s correlation coefficient:

$$cor_P = \frac{cov(y, o)}{\sigma_y \sigma_o} = \frac{\sum_{i=1}^N y_i o_i - N \bar{y} \bar{o}}{\sqrt{\sum_{i=1}^N y_i^2 - N \bar{y}^2} \sqrt{\sum_{i=1}^N o_i^2 - N \bar{o}^2}}$$

where cov is the covariance, σ_y and σ_o are the standard deviations of y and o , and \bar{y} and \bar{o} are the mean values $\frac{1}{N} \sum_{i=1}^N y_i$ and $\frac{1}{N} \sum_{i=1}^N o_i$, respectively.

- Spearman’s rank correlation coefficient is defined similarly but the variables y_i and o_i are converted to ranks $R(y_i)$ and $R(o_i)$:

$$cor_S = \frac{cov(R(y), R(o))}{\sigma_{R(y)} \sigma_{R(o)}}$$

- Kendall’s rank correlation coefficient measures the correspondence between the ranking of y_i and o_i : the number of possible pairings of y_i and o_i is $N(N - 1)/2$; if the

pairs are ordered by the o_i values, then, for each y_i , we count the number of $y_j > y_i$ (N_C , total number of concordant pairs) and the number of $y_j < y_i$ (N_D , total number of discordant pairs); hence, the correlation coefficient is defined as

$$\text{cor_K} = \frac{N_C - N_D}{N(N - 1)/2}$$

- Tail dependence (χ) takes in input y and o and evaluates the dependence on the tail of the distribution of two series about a set quantile; therefore, it investigates how the adjustment method affects the temporal alignment of extreme days: in this work, 0.95 was chosen, following Oyler et al. [14] and Weller et al. [32]. It is defined as

$$\chi(u = 0.95) = P(y > u | o > u)$$

This indicator is directly available in the R extRemes package [33], while new scripts were created to calculate the others.

- Accuracy was derived by the confusion matrix [34], which takes the binary variables Y_i and O_i as input and is defined as

$$\text{ACC} = \frac{TP + TN}{P + N}$$

where P and N are the total positive (wet days) and negative (dry days) cases, and TP and TN are the true positive and true negative cases, respectively. A “true positive” is a day correctly identified by the adjustment method as a wet day, while a “true negative” is a day correctly identified as a dry day.

- Heidke Skill Score (HSS) quantifies the alignment of precipitation occurrence and is defined as

$$\text{HSS} = \frac{2((TP \cdot TN) - (FN \cdot FP))}{P(FN + TN) + N(TP + FP)}$$

where FP and FN are the false positive and false negative cases, i.e., days incorrectly identified as wet or dry, respectively. The confusion matrix and the indicator ACC were calculated using the R caret package [35], while HSS was calculated using the elements of the confusion matrix.

- Mean precipitation over wet days (mwet) is the difference between the mean of the predicted values y_i and the mean of the observed values o_i of precipitation. Mwet is expressed as a percentage, calculated with respect to the observed values o_i ; since all $0 < y_i < 1$ are set to zero, for the reason explained in Section 2.3, the calculation of the mean values over only the wet days (when the binary variables are 1) is simplified:

$$\text{mwet} = \left[\left(\frac{\sum_{i=1}^N y_i}{\sum_{i=1}^N Y_i} - \frac{\sum_{i=1}^N o_i}{\sum_{i=1}^N O_i} \right) / \left(\frac{\sum_{i=1}^N o_i}{\sum_{i=1}^N O_i} \right) \right] \cdot 100$$

- Frequency over wet days (freq) is the percentage difference between the number of predicted wet days and the observed wet days. The percentage is calculated with respect to the observed wet days:

$$\text{freq} = \left[\frac{\sum_{i=1}^N Y_i - \sum_{i=1}^N O_i}{\sum_{i=1}^N O_i} \right] \cdot 100$$

Finally, the impact of the adjustment methods on the extreme days was investigated by analyzing the trend over the years of the percentile distributions of the original and adjusted series, with particular attention to the upper percentiles.

2.6. Multivariate Approach

A crucial problem when dealing with time series is missing data. Over time, a wide variety of methods has been developed, with the percentage of gaps and data missingness mechanism being the main factors limiting their applicability [36]. The most performing techniques require the availability of data from neighboring locations [37], and their success depends on the extent of the correlation between the target and predictor stations [38]. From this perspective, it is interesting to investigate whether, in the case of a 9–9 series, it is preferable to apply an adjustment method to convert the 9–9 series in a 0–24 series or leave it and use the 0–24 series of another station close to the target one. Therefore, the performance indicators described in Section 2.5 were also evaluated for the 0–24 series of the stations listed in Table 1 and located near Padua. To make the interpretation and visualization of the results easier, an exploratory analysis technique was employed, Principal Component Analysis (PCA, [39]). PCA is a dimension reduction method, used to capture the relevant information and to visualize major trends and structure of data. PCA was applied to the dataset of the indicators calculated: (i) for the 9–9 series adjusted using the methods described in Section 2.4, and (ii) for the 0–24 original series of the stations near Padua. The dependence of the results on the month of the year was also investigated. To manage this further variability element, Parallel Factor Analysis (PARAFAC), which is a generalization of PCA to higher-order arrays [40], was applied. In PARAFAC, any source of variability constitutes a so-called “mode” and the variation in each mode can be described by a low number of factors. PARAFAC was mainly used to improve and simplify the visualization of the results. PCA and PARAFAC were both performed using the software PLS Toolbox 8.1 (Eigenvector Research, Inc., Wenatchee, WA, USA) for Matlab © R2017b.

3. Results

The homogenization tests applied to the datasets listed in Table 1 indicate that all the series are homogeneous.

NOAA reconstructed data are available in 3 h steps, in UTC format; therefore, the 00 UTC value of d_j actually covers the interval from 22 LT of d_{j-1} to 01 LT of d_j . As it is not possible to disaggregate the amount of this 3 h interval, and to allocate it between the two subsequent days, the 00 UTC amount was entirely assigned to the first day d_j . The comparison with the daily observations calculated as “1-1” sums showed that the differences are negligible, i.e., the 1 h shift of the 00 UTC 3-h value does not significantly alter the indicators.

3.1. Comparison between Methods at Daily Resolution

Figure 4a–e show the scatter plots of the 0–24 vs. the 9–9 series adjusted using the adjustment methods described in Section 2.4. Linear regressions were added with the resulting equations and R^2 values. The simple 1-day method significantly improves the linearity between the original and adjusted series, but the methods based on reanalysis perform better than the others. The same comparison at the monthly level (Figure 4f) indicates, as expected, that the choice of adjustment method is not as crucial as at the daily level, in particular in terms of linearity (Table 4). Nevertheless, the methods based on reanalysis give a lower RMSE than the other methods (Table 4).

Table 4. Significant parameters of the linear regression applied to original and adjusted Padua monthly series.

Adjustment Method	R^2	RMSE (mm)
9–9	0.979	7.9
1–day	0.991	5.2
unif	0.994	4.2
ERA5	0.998	2.3
NOAA	0.997	2.6

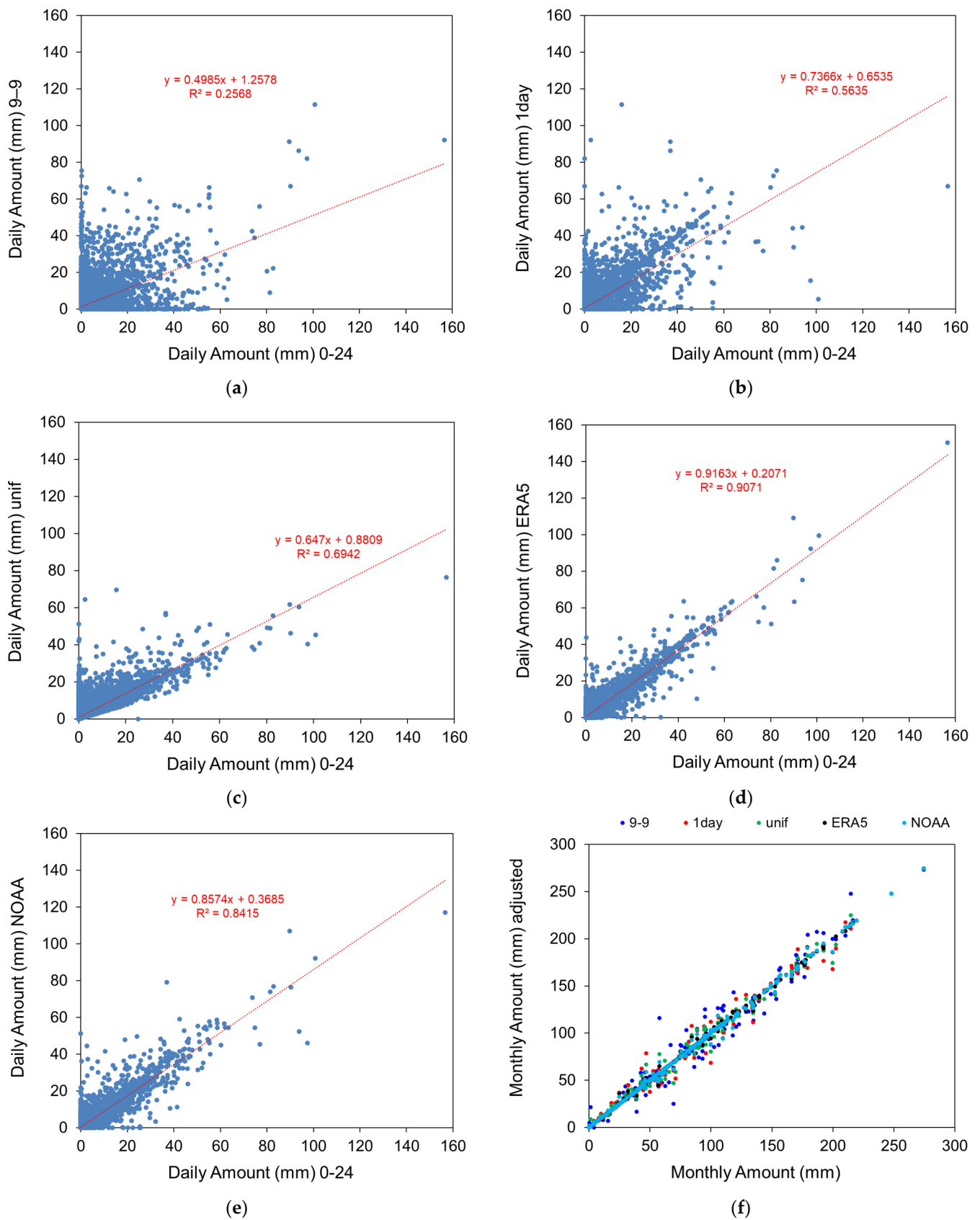


Figure 4. Scatter plots of the 0–24 series of Padua compared to (a) 9–9 series; 9–9 adjusted series (b) using 1-day method, (c) uniform method, (d) ERA5 method, and (e) NOAA method; (f) scatter plot of the monthly original Padua series compared to the adjusted series using the 4 methods.

The performance of the different adjustment methods can be discussed based on the values of the indicators reported in Figure 5, where the 0–24 datasets were used as reference. For each method, the average calculated for all the stations was also provided. Cor_P, cor_S, and cor_K indicate the Pearson, Spearman, and Kendall correlation coefficients, respectively. Two different color scales were used for the columns: (i) a three-color scale (green-yellow-red) for the indicators related to temporal alignment, i.e., from RMSE to HSS; and (ii) a double-ended (white-violet) color scale for the indicators related to precipitation statistics, i.e., mean value over wet days (mwet) and frequency (freq). The indicators of temporal alignment were evaluated considering their relative value—the ones related to precipitation statistics in their absolute value. In fact, a method performs better the larger or smaller the temporal alignment indicators are in relative value, depending on the indicator; the green color indicates the best performing method, the red color the worst performing one. As an example, a good method has a low RMSE and MAE and a high cor_P, cor_S, and cor_K. At the same time, a method performs better the smaller the indicators of precipitation statistics are in absolute value (i.e., white color), and worse the higher they are in absolute value (i.e., violet color).

9-9	RMSE	MAE	NMAE	BS	cor_P	cor_S	cor_K	$\chi(0.95)$	ACC	HSS	mwet	freq
Pd	7.33	2.69	1.08	0.19	0.51	0.52	0.47	0.39	0.81	0.55	2.5	-2.2
Lg	6.88	2.45	1.09	0.19	0.47	0.52	0.47	0.36	0.82	0.59	1.7	-2.3
Cm	7.67	2.79	1.07	0.19	0.50	0.53	0.48	0.36	0.81	0.60	2.6	-2.4
Mr	7.33	2.58	1.08	0.20	0.51	0.52	0.47	0.34	0.80	0.58	1.7	-1.6
Cd	7.24	2.40	1.09	0.20	0.47	0.52	0.48	0.38	0.80	0.59	2.0	-2.0
Tr	5.98	2.12	1.08	0.20	0.47	0.51	0.47	0.39	0.80	0.58	3.1	-2.7
average	7.07	2.50	1.08	0.19	0.49	0.52	0.47	0.37	0.81	0.58	2.3	-2.2
1day	RMSE	MAE	NMAE	BS	cor_P	cor_S	cor_K	$\chi(0.95)$	ACC	HSS	mwet	freq
Pd	5.21	1.52	0.61	0.11	0.75	0.76	0.71	0.66	0.89	0.74	2.3	-2.4
Lg	4.77	1.36	0.61	0.15	0.74	0.75	0.71	0.65	0.85	0.67	2.2	-1.9
Cm	5.36	1.59	0.61	0.16	0.76	0.76	0.72	0.65	0.84	0.65	2.6	-2.4
Mr	5.55	1.49	0.62	0.18	0.72	0.75	0.71	0.64	0.82	0.63	1.8	-1.6
Cd	5.19	1.41	0.64	0.17	0.73	0.75	0.70	0.62	0.83	0.64	2.4	-2.1
Tr	3.99	1.20	0.61	0.16	0.76	0.76	0.71	0.63	0.84	0.66	3.4	-3.1
average	5.01	1.43	0.62	0.16	0.74	0.76	0.71	0.64	0.84	0.66	2.5	-2.3
unif	RMSE	MAE	NMAE	BS	cor_P	cor_S	cor_K	$\chi(0.95)$	ACC	HSS	mwet	freq
Pd	4.14	1.54	0.62	0.13	0.83	0.79	0.74	0.67	0.87	0.73	-26.8	35.6
Lg	3.82	1.41	0.63	0.14	0.82	0.79	0.73	0.66	0.86	0.73	-26.8	35.7
Cm	4.22	1.57	0.60	0.14	0.84	0.80	0.74	0.68	0.86	0.73	-26.1	34.9
Mr	4.32	1.50	0.63	0.15	0.82	0.79	0.73	0.65	0.85	0.70	-26.9	36.2
Cd	4.10	1.42	0.64	0.15	0.81	0.78	0.73	0.62	0.85	0.70	-26.2	34.9
Tr	3.15	1.21	0.62	0.15	0.84	0.79	0.73	0.63	0.85	0.71	-26.3	34.9
average	3.96	1.44	0.62	0.14	0.83	0.79	0.73	0.65	0.86	0.72	-26.5	35.4
ERA5	RMSE	MAE	NMAE	BS	cor_P	cor_S	cor_K	$\chi(0.95)$	ACC	HSS	mwet	freq
Pd	2.27	0.67	0.27	0.08	0.95	0.90	0.86	0.87	0.92	0.82	-5.8	5.5
Lg	2.32	0.65	0.29	0.11	0.94	0.89	0.86	0.84	0.89	0.77	-7.2	7.4
Cm	2.70	0.75	0.29	0.12	0.94	0.90	0.86	0.84	0.88	0.75	-6.2	6.3
Mr	2.40	0.68	0.28	0.14	0.95	0.89	0.85	0.83	0.86	0.72	-7.1	7.4
Cd	2.39	0.66	0.30	0.13	0.94	0.89	0.86	0.83	0.87	0.74	-7.2	7.5
Tr	2.00	0.58	0.29	0.13	0.94	0.89	0.85	0.83	0.87	0.74	-5.7	5.7
average	2.35	0.66	0.29	0.12	0.94	0.89	0.86	0.84	0.88	0.76	-6.5	6.6
NOAA	RMSE	MAE	NMAE	BS	cor_P	cor_S	cor_K	$\chi(0.95)$	ACC	HSS	mwet	freq
Pd	3.09	0.97	0.37	0.10	0.92	0.85	0.81	0.79	0.90	0.79	-11.9	12.7
Lg	2.94	0.90	0.39	0.12	0.90	0.85	0.81	0.79	0.88	0.75	-11.9	13.0
Cm	3.26	1.02	0.39	0.13	0.91	0.86	0.81	0.77	0.87	0.73	-12.2	13.6
Mr	3.38	0.98	0.40	0.14	0.90	0.84	0.80	0.77	0.86	0.71	-12.7	14.3
Cd	3.01	0.89	0.40	0.14	0.91	0.85	0.80	0.76	0.86	0.72	-11.4	12.5
Tr	2.41	0.79	0.39	0.13	0.91	0.85	0.81	0.81	0.87	0.73	-11.1	12.0
average	3.01	0.93	0.39	0.13	0.91	0.85	0.81	0.78	0.87	0.74	-11.9	13.0

Figure 5. Values of the indicators calculated for each adjustment method applied to each station. RMSE and MAE are expressed in mm, and mwet and freq are expressed in percentage.

The results of the various methods applied to different stations are consistent between them, as the indicators for the same method show no significant differences between stations. The reanalysis-based methods, especially ERA5, produce the greatest increases in temporal alignment. In fact, the ERA5 method is characterized by the highest correlation coefficients (i.e., cor_P, cor_S, and cor_K), $\chi(0.95)$, accuracy, and HSS, and by the lowest errors (i.e., RMSE, MAE, NMAE) and BS. Also, the 1-day and unif methods produce an

improvement in temporal alignment. Therefore, in the absence of reanalysis data, they can be considered valid alternatives to adjust the 9–9 series. Concerning the precipitation statistic, the values averaged over all stations reported in Figure 5 are better visualized in Figure 6. The unif method produces large changes in frequency and mean value over wet days, increasing the former and decreasing the latter. The reanalysis-based methods introduce changes in the same directions but to a smaller extent than the unif method. Finally, the 1-day method produces inconsistent improvements in the statistics.

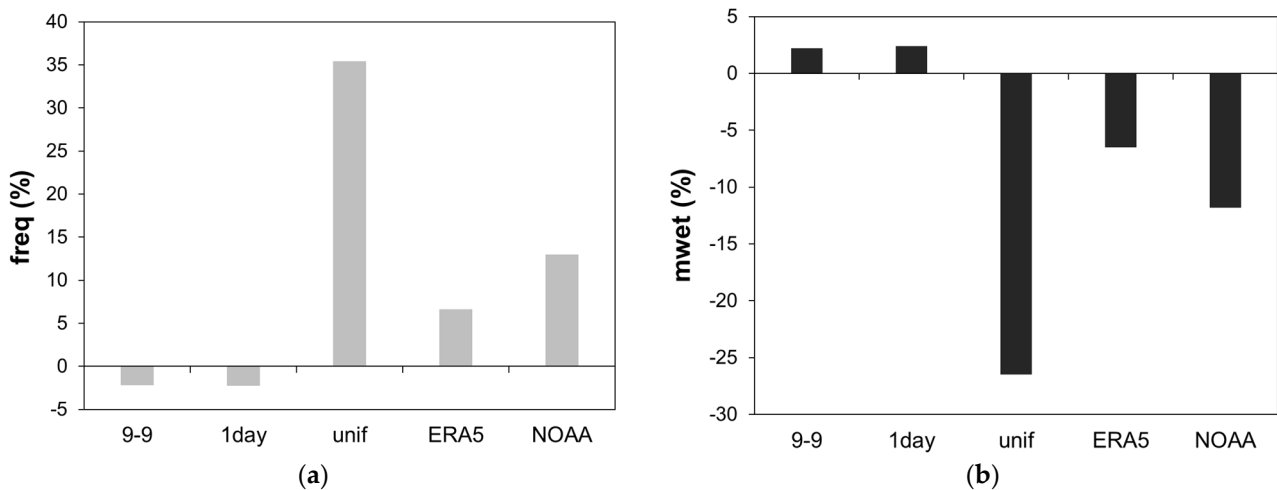


Figure 6. Values averaged over all stations of precipitation: (a) frequency; (b) mean value over wet days.

3.2. Comparison between Methods and Stations at Daily Resolution

The same analysis was applied to the 0–24 datasets of the stations listed in Table 1, again using the 0–24 dataset of Padua as reference. Results are shown in Figure 7.

	RMSE	MAE	NMAE	BS	cor_P	cor_S	cor_K	$\chi(0.95)$	ACC	HSS	mwet	freq
9-9	13.54	8.68	1.11	0.18	0.51	0.52	0.47	0.39	0.82	0.49	2.5	-2.2
1day	10.27	5.44	0.55	0.10	0.75	0.76	0.71	0.66	0.91	0.73	2.3	-2.4
unif	7.45	4.86	0.61	0.11	0.83	0.79	0.74	0.67	0.89	0.73	-26.8	35.6
ERA5	4.45	2.50	0.25	0.05	0.95	0.90	0.86	0.87	0.95	0.87	-5.8	5.5
NOAA	5.82	3.38	0.35	0.07	0.92	0.85	0.81	0.79	0.93	0.81	-11.9	12.7
Cm	3.82	1.05	0.43	0.05	0.87	0.88	0.84	0.74	0.95	0.85	1.9	2.9
Cd	5.00	1.49	0.61	0.08	0.76	0.81	0.76	0.61	0.92	0.78	-8.6	-5.0
Lg	3.29	0.93	0.49	0.05	0.90	0.90	0.86	0.78	0.95	0.87	-6.9	-3.7
Mr	4.39	1.20	0.49	0.06	0.83	0.86	0.82	0.71	0.94	0.83	-3.7	-2.1
Tr	4.76	1.44	0.59	0.08	0.77	0.80	0.75	0.62	0.92	0.77	-16.0	-7.4

Figure 7. Values of the indicators calculated for the adjusted Padua series compared to the 0–24 series of other stations. RMSE and MAE are expressed in mm.

To capture the most relevant information, PCA was applied to the two-dimensional matrix 10×12 of Figure 7, in which the adjustment methods and the stations were considered as “samples” and the performance indicators as variables. Mean centering and variance scaling were applied as data pretreatments. The number of principal components (PCs) to be retained was selected based on the percentage of total variance explained, not to be lower than 90%. The total variance accounted for by the first two PCs was around 92%; therefore, the discussion of the results focuses on PC1 and PC2. Figure 8a shows the loading plot of PC1 vs. PC2. PC1, which is responsible for the description of 77% of the variance, measures the temporal alignment because it has large (in absolute value) association with the indicators related to this aspect. In particular, PC1 shows positive loadings for cor_P, cor_S, cor_K, $\chi(0.95)$, ACC, and HSS, with negative loadings for RMSE, MAE, NMAE, and BS. Looking at the position of the indicators in the loading plot, it is

evident that the indicators of temporal alignment can be divided into two groups. The former is referred to as “correlation”, the latter to as “error” group, looking at the meaning of the indicators forming each group. In fact, the most performing method is characterized by a low value of the indicators that have negative loadings on PC1 (i.e., RMSE, MAE, NMAE, BS) and high values of the indicators that have positive loadings on PC1 (i.e., cor_P, cor_S, cor_K, $\chi(0.95)$, ACC, HSS). PC2 instead measures precipitation statistics, as both freq and mwet have high (in absolute value) loadings on PC2, with the former possessing positive ones and the latter possessing negative ones.

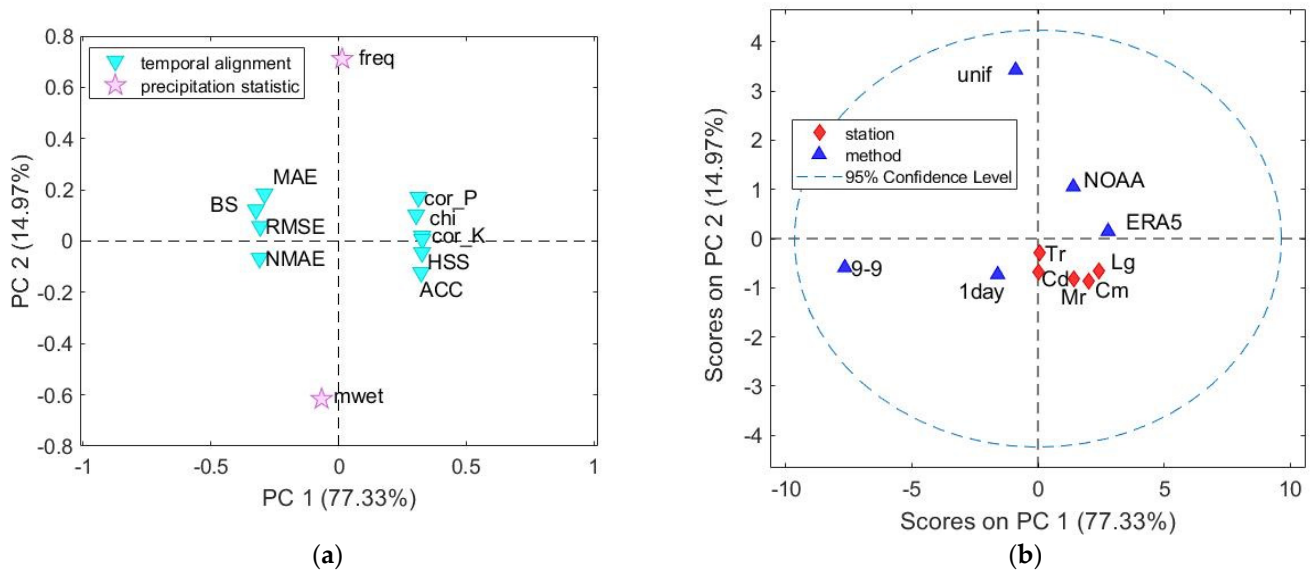


Figure 8. Results of PCA applied to the dataset 10 (stations/methods) × 12 (indicators): (a) loading plot of PC1 vs. PC2; (b) scores plot on PC1 vs. PC2.

The score plot of PC1 vs. PC2 in Figure 8b makes the comparison between methods and stations easier than that in Figure 6. Three out of five stations are characterized by positive scores on PC1. Hence, concerning temporal alignment, using one of these station’s datasets gives better results than the 9–9, 1-day, and unif adjustment methods. There is no significant difference between applying the most performing method, i.e., ERA5, to the 9–9 dataset and taking the data from Legnaro station, as the two points (ERA 5 and Lg) are both characterized by the highest values of PC1. The scores on PC1 of the stations Campodarsego and Mira are placed in the middle between ERA5 and NOAA methods. Regarding precipitation statistics, all the stations are characterized by negative scores on PC2. Therefore, they exhibit slightly lower values of freq and higher values of mwet with respect to the 0–24 series, giving similar results than 9–9 series and 1-day methods. Anyhow, using another station’s dataset improves the precipitation statistics with respect to unif and NOAA methods, characterized by higher scores on PC2. The results obtained with PCA agree with the conclusions drawn using the traditional statistical data analysis and visualization in Section 3.1.

3.3. Monthly Analysis

The performance indicators calculated at a daily resolution were then aggregated on a monthly basis to investigate the eventual dependence of the results on the month of the year. Only the adjustment methods applied to the Padua series were considered. Since there are now three elements of variability, i.e., the performance indicators, the adjustment methods, and the month of the year, Parallel Factor Analysis (PARAFAC) [40] was preferred to PCA. The input data were organized in a three-way array that reports the methods in the first mode, the indicators in the second mode, and the months in the third mode, i.e., array with 5 × 12 × 12 dimensions. The choice to build a three-way array was due to the need to

highlight clear information about differences among months. Preprocessing of the three-way arrays is much more complicated than in the two-way case, as centering and scaling across each mode are not independent [41,42]. The variable “indicator” is not homogeneous, i.e., the performance indicators are of very different typologies and their definitions include the comparison with the reference series in different ways (see Section 2.5). Hence, no data preprocessing was applied, to avoid the introduction of artifacts in the analysis. For the choice of the right number of PARAFAC factors, several different criteria were evaluated, such as core consistency [40], percentage of explained variance, and sum of squared errors. The one-factor model with an explained variance of 94% was chosen for the three-way array because of its high core consistency (100%) and its robustness considering the low values of the sum of the squared residuals.

The loading plots of the first (adjustment method), second (performance indicators) and third modes (months) of the first factor are reported in Figure 9. In the first mode plot (Figure 9a), unif, NOAA, and ERA5 methods have positive scores values, with 9–9 and 1 day exhibiting negative ones. The first factor mainly differentiates the unif method (characterized by the highest score value) from the others. In particular, it is characterized by the most remarkable difference in precipitation statistics, i.e., freq and mwet, with respect to the reference series 0–24 (Figure 9b). This behavior is particularly true for the two central summer months, i.e., July and August (Figure 9c), which exhibit the highest positive scores values in Mode 3.

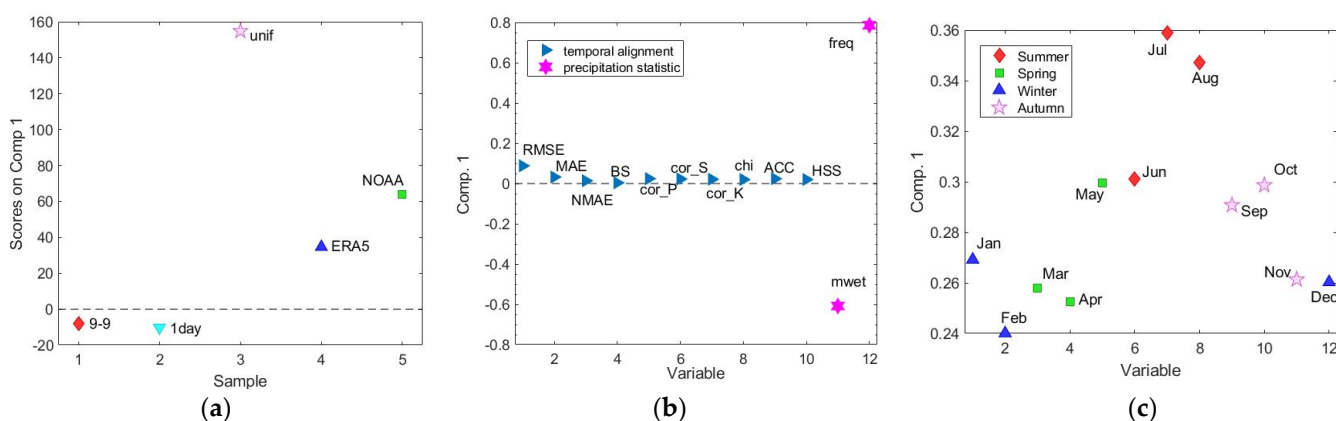


Figure 9. Three-way PARAFAC model of monthly values of performance indicators for Padua. Loadings on factor 1 of the three modes of data analysis: (a) Mode 1—adjustment methods; (b) Mode 2—performance indicators; (c) Mode 3—month of the year.

From an explorative point of view, Figure 9c shows the presence of three groups of months, according to the values of the loadings on the first factor. Starting from the lower values to the higher ones, the first group includes the months from late autumn to early spring (from November to April); the second one the months of late spring/early summer and early autumn, i.e., May, June, September, and October; and the third one the two central summer months, i.e., July and August.

The PARAFAC model allows some preliminary conclusions to be drawn up on the “month” variable: in fact, it seems that the adjustment methods, mainly unif and NOAA among the others, show poor performance concerning precipitation statistics in the warmer part of the year, from late spring to early autumn.

Mode 2 (Figure 9b) confirms the result of the PCA, i.e., that the two categories of indicators behave differently and are internally consistent. This would make it possible to reduce the number of indicators needed to assess temporal alignment on one side and precipitation statistics on the other. Nevertheless, with Mode 2 being dominated by the precipitation statistics, the monthly variability shown by Mode 3 is mainly referred to regarding this aspect.

To investigate more in depth the monthly dependence of temporal alignment, a new three-way array was created, $5 \times 10 \times 12$, with the only difference with respect to the previous one being that the second mode included only the indicators related to temporal alignment. Following the same criteria already explained, the one-factor model with an explained variance of 95% was chosen.

The new loading plots of the three modes of the first factor are reported in Figure 10a–c, respectively. Mode 2 is dominated by RMSE (Figure 10b), as it is the indicator that has the highest loading on factor 1, while the other indicators have similar lower values. The performance ranking of the different methods, represented by Mode 1, throughout the months, represented by Mode 3, is mainly related to this indicator. Figure 10c shows the presence of two groups of months, according to the values of the loadings on the first factor; the first group includes the months from December to April and corresponds to a lower RMSE, i.e., better performance, than the second group, which includes months from May to November. Therefore, the series adjusted using the methods represented by Mode 1 are less aligned to the 0–24 series in summer and autumn, as these seasons (Mode 3) are characterized by higher values of RMSE (Mode 2). This is particularly true for the 9–9 method, characterized by the highest loading on factor 1 (Figure 10a), followed in scale by the methods 1–day, unif, NOAA, and ERA5. The most performing adjustment method is ERA5, characterized by the lower loading on factor 1, i.e., the lower RMSE, in particular in winter and early spring.

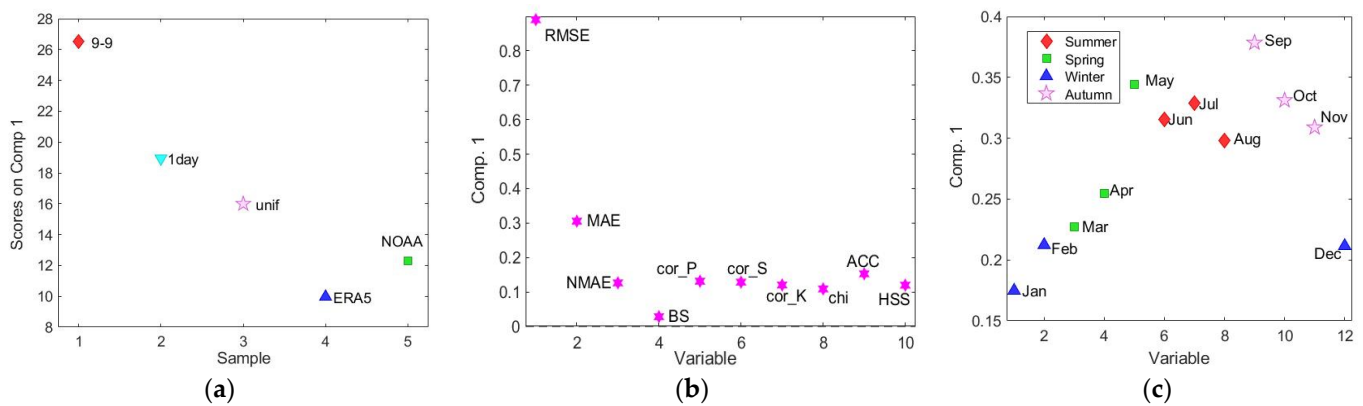


Figure 10. Three-way PARAFAC model of monthly values of performance indicators related to temporal alignment for Padua. Loadings on factor 1 of (a) Mode 1—adjustment methods; (b) Mode 2—performance indicators; (c) Mode 3—month of the year.

To obtain more robust results, as the PARAFAC model was dominated by RMSE, PCA was also run on the two-dimensional matrix 60×10 , in which the monthly adjustment methods were considered as “samples” and the performance indicators related to temporal alignment were considered as variables. There were 60 samples in total, as each method was composed of 12 rows, one for each month. Mean centering and variance scaling were applied as data pretreatments. The model with one principal component was selected as the variance explained by PC2 was brought by outliers, as revealed by the Hotelling’s T-squared test [43]. Results are summarized in Figure 11. PC1, which is responsible for the description of 84% of the variance, shows positive loadings for the “correlation” group of indicators, with negative loadings for the “error” group. The best performing methods, i.e., ERA5 and NOAA, are characterized by a low value of the indicators that have negative loadings on PC1, i.e., “error” indicators, and a high value of the indicators that have positive loadings on PC1, i.e., “correlation” indicators. The interpretation of the monthly dependence is less immediate than with PARAFAC analysis, but the results completely agree. In general, the adjustment methods show less temporal alignment with the original series in summer and autumn (Figure 11c), and this is particularly evident for the two methods characterized by the highest errors, i.e., mainly 9–9, followed by 1 day (Figure 11b).

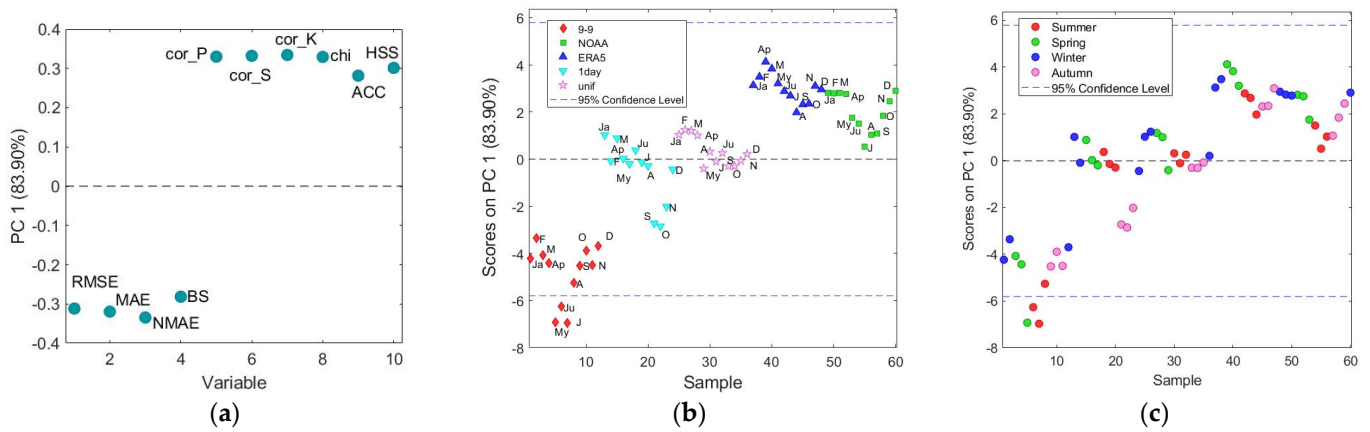


Figure 11. Results of PCA applied to the dataset 60 (12 months × 5 methods) × 10 (indicators): (a) loading plot of PC1; scores plot on PC1 differentiating (b) the methods and (c) the seasons.

3.4. Percentiles Distribution

Figure 12 visualizes the results of the analysis of the percentile distribution to assess the effect of the adjustment methods on daily extremes. In Figure 12a, the values of the percentiles from the 50-ile to the 100-ile calculated for the original and adjusted daily series of Padua are compared. All the adjustment methods underestimate the percentile values, except ERA5 that outperforms all the others. The unif method exhibits the greatest difference with the 0–24 series, as it halves the values of the percentiles above the 95-ile. The same analysis carried out separately for the other stations gave similar results. Then, the adjustment methods applied to Padua series were compared to the series of the neighboring stations: the difference from the 0–24 Padua series of the percentile values from the 90-ile to the 100-ile was calculated for the adjusted series and the series of the neighboring stations. Tribano was excluded as its dataset is 3 years shorter. The results are shown in Figure 12b as percentage and in Figure 12c as absolute values. The columns in Figure 12c are colored using a three-color scale, i.e., red for the highest difference, green for the lowest one, and yellow for what is in the middle. The only series that comes close to the ERA5 method is Legnaro, the station that is most correlated with Padua (Figure 7).

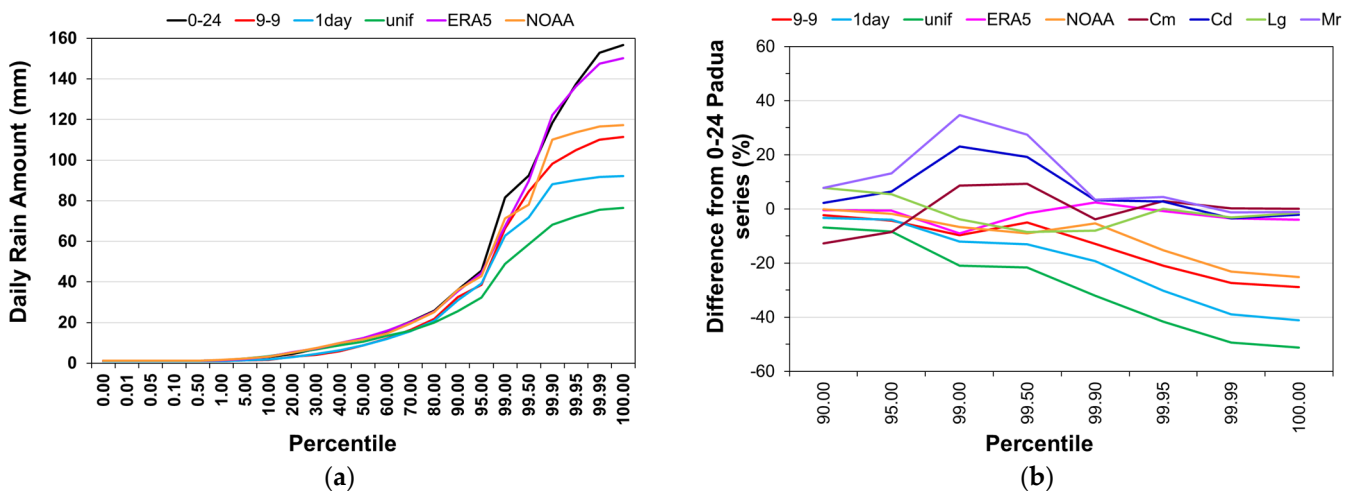


Figure 12. Cont.

method/station	percentile							
	100.00	99.99	99.95	99.90	99.50	99.00	95.00	90.00
9-9	28.86	27.26	20.87	12.88	4.94	9.66	4.37	2.30
1day	41.12	38.94	30.20	19.28	13.11	11.98	3.88	3.32
unif	51.21	49.30	41.65	32.08	21.58	20.92	8.38	6.74
ERA5	4.02	3.39	0.83	2.35	1.68	8.99	0.57	0.47
NOAA	25.16	23.17	15.22	5.29	9.04	6.65	1.70	0.04
Cm	12.77	8.50	8.56	9.29	3.78	2.87	0.26	0.00
Cd	2.30	6.46	23.13	19.20	3.19	2.79	3.39	2.18
Lg	7.79	5.47	3.79	8.43	7.94	0.01	3.18	1.40
Mr	7.79	13.15	34.61	27.54	3.51	4.41	1.26	1.16

(c)

Figure 12. Percentiles of the daily series: (a) values of the percentiles from the 50-ile to the 100-ile for the original and adjusted Padua series; (b) percentage difference between the percentile values from the 90-ile to the 100-ile of the adjusted Padua series and the neighboring series, and the 0–24 Padua series; (c) absolute values of the differences in (b).

4. Discussion

The adjustment methods applied in the present study to the 9–9 precipitation series of Padua exhibit different performances depending on the point of view, temporal alignment, or precipitation statistics, which confirm that they are two distinct aspects, according to both the traditional statistical analysis and the multivariate approach. The comparison between the results of the present and other studies is limited, as the adjustment methods have never been tested in all the possible scenarios; in particular, the climatological day starting at 9 LT has never been considered. Based on the results of this work and of previous studies, all methods clearly improve the temporal alignment. The reanalysis-based methods, especially ERA5, that have been tested here for the first time produce the greatest improvement. But, for the periods in which reanalysis data are not available, 1 day and uniform methods can be considered valid alternatives to adjust the 9–9 series from this point of view. Concerning the precipitation statistics, the adjustment methods are not without drawbacks, as already pointed out in previous studies. In particular, the uniform shift method, which reapporitions uniformly daily precipitation observations, is confirmed as the method that has the highest potential of artificially increasing the precipitation frequency and decreasing the mean value over wet days. The reanalysis-based methods also introduce these changes, but to a smaller extent than the uniform shift method. Therefore, for the period in which reanalysis data are not available, the 1-day shift method is confirmed as better performing than the uniform shift method also in the 9–9 scenario, if both the aspects of temporal alignment and precipitation statistics are considered.

Moving from the perspective of one single station, and considering several stations close to the target one—an operation that is mandatory in the case of missing data—the distance from the adjusted series to the 0–24 series of the same station has been compared to the 0–24 series of the other stations. It is difficult to generalize the results, as they depend on the method and station. For Padua, using another station dataset gives similar or better results than any adjustment method applied to the 9–9 series, except for the uniform shift method that significantly changes the precipitation statistics.

The multivariate approach allows for better visualizing whether the results obtained for a single station depend on the month or season of the year. All the adjustment methods introduce the most relevant changes in the precipitation statistics in summer, in particular the uniform shift method (Figure 9). Analyzing separately the temporal alignment results of each method, the adjusted series are less aligned to the 0–24 series in summer and autumn, and this is particularly true for the 9–9 method (Figure 10). This result can be interpreted considering the precipitation regime in Padua, where heavy rains are frequent especially in summer and autumn (Figure S1). At the same time, the reanalysis methods are less performing in summer than in other seasons (Figure 11), because the reanalysis has limitations in correctly simulating thunderstorms. Since summer thunderstorms mainly occur in the late afternoon/evening, the 9–9 method attributes them to the wrong day, i.e.,

the day after the target one, which is not the case with the 1-day method. In autumn, the 1-day shift method performs worse than in summer (Figure 11). This can be explained considering that autumn rainfall is quite homogeneous, with no time preference; therefore, the effect on the 9–9 method is not as dramatic as for months with convective rainfall, i.e., summer. Anyhow, the 9–9 method in autumn is still worse than the 1-day method because the former takes only 9 h of the target day, while the latter takes 15 out of 24, 25% of the total (Figure 3).

Concerning the impact of the adjustment methods on the daily precipitation percentile distribution and consequently on the identification and characterization of extreme days, the results showed that all the methods underestimate the percentile values, except ERA5 that simulates daily extremes better than taking the dataset of a neighboring station (Figure 12). When regular time series are considered, e.g., regular daily precipitation amounts, selected percentiles are directly related to the return period (RP) [44]. The precipitation amounts related to 10, 20, and 30-year RPs were evaluated for all the methods and stations considered in this study, taking advantage of a specific function (i.e., fevd) of the R extRemes package, version 2.1-3. [33]. In Figure 13, the RPs of the different adjustment methods applied to Padua datasets are compared between them and with Legnaro, the station mostly correlated with Padua; the results are expressed as a percentage difference with respect to the 0–24 Padua series. It is evident that the length of the period considered is not as important as the method. ERA5 is the method that better reproduces the RPs of the original series, followed by Legnaro; both datasets can be considered reliable candidates to fill the gap of the Padua series. The uniform adjustment is confirmed to markedly decrease the extremes and consequently increase the RPs. The same analysis carried out for the other stations gives similar results concerning the performance of the adjustment methods in terms of RPs.

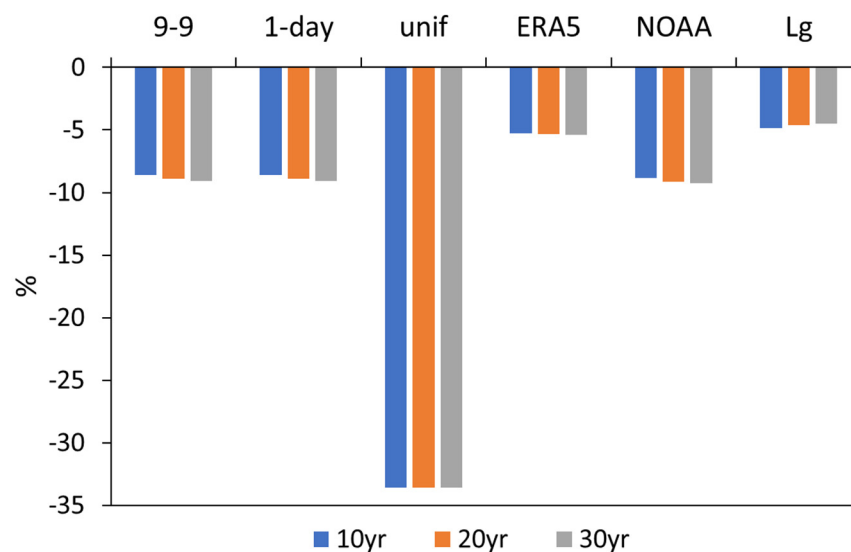


Figure 13. Percentage difference with respect to the 0–24 series of Padua of the precipitation amounts related to the 10, 20, and 30-year return periods for the different adjustment methods and the 0–24 series of Legnaro.

5. Conclusions

The evaluation of the time of observation adjustment methods is not a simple task, as their performance depends on the type and entity of the temporal misalignment between the datasets, and their application. In this study, five adjustment methods were applied to the 9–9 daily precipitation series recorded by ARPAV in Padua, and in five nearby stations. Two out of five methods are based on reanalysis and have never been applied before.

The selected indicators evaluate the methods in terms of temporal alignment and precipitation statistics. The results of both traditional statistical analysis and the multi-

variate approach confirm that they are two distinct aspects and indicate that none of the methods considered is the best in either aspect. Nevertheless, the reanalysis-based methods, especially ERA5, significantly improve the temporal alignment of the 9–9 series. At the same time, they increase the precipitation frequency and reduce the mean value over wet days, NOAA much more than ERA5. Overall, using the 0–24 dataset of another station close to Padua gives similar or better results than applying any adjustment method to the 9–9 series. This finding can be hardly generalized, as it depends on the method, station, and local climatology.

While the time of observation misalignment can cause problems with daily precipitation, it becomes less of an issue at coarser temporal resolutions, e.g., monthly or seasonally. In general, all the adjustment methods introduce the most relevant changes in precipitation statistics in summer. In addition, they show less temporal alignment with the original series in summer and autumn, which are the two seasons mainly affected by heavy rains in Padua. Finally, all the adjustment methods underestimate the percentile values, to a greater extent the higher the percentile, except ERA5 that outperforms all the others. Among the stations near Padua, the only series that come close to the Padua series adjusted with the ERA5 method is Legnaro, the station most correlated with Padua.

The methodology described in this work can be extended to broader contexts, as it is applicable to precipitation datasets of any country and/or period. Nevertheless, the identification of the best performing adjustment method, or the choice of the series of a nearby station instead of the adjusted series, depends on the specific dataset under study and local climate conditions. The new method based on ERA5 reanalysis showed good potential as an adjustment method, as it was successfully applied to the modern precipitation series of Padua and nearby stations. As a future perspective, this method can be extended to all datasets recorded by the Meteorological Observatory of the Water Magistrate, which constitutes, for Italy, a precious source of instrumental data for the 20th century. The alignment of these datasets characterized by a different definition of the climatological day with respect to the modern standard, will allow for extending the daily precipitation series of several Italian locations, increasing the availability of data for climate research. Last but not least, the results obtained in this work allow the completion of the reconstruction of the 300-year precipitation series of Padua, one of the longest series in the world.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/atmos15040412/s1>, Figure S1: Seasonal precipitation intensity in Padua in the period 1994–2022; Table S1: Yearly precipitation amount (mm) in the 1993–2022 period for the ARPAV stations in proximity of Padua; Table S2: Number of rainy days in the 1993–2022 period for the ARPAV stations in proximity of Padua.

Author Contributions: Conceptualization, D.C. and F.B.; methodology, F.B. and C.S.; validation, F.B. and C.S.; formal analysis, F.B. and C.S.; investigation, F.B. and C.S.; data curation, A.d.V., F.R. and F.Z.; writing—original draft preparation, F.B.; writing—review and editing, F.B., C.S., D.C., A.d.V., F.R. and F.Z.; supervision, D.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

Acknowledgments: Hersbach, H. et al., (2018) [45] was downloaded from the Copernicus Climate Change Service (C3S) (2023). The results contain modified Copernicus Climate Change Service information 2020. Neither the European Commission nor ECMWF is responsible for any use that may be made of the Copernicus information or data it contains. Support for the Twentieth Century Reanalysis Project version 3 dataset is provided by the U.S. Department of Energy, Office of Science Biological and Environmental Research; by the National Oceanic and Atmospheric Administration Climate Program Office; and by the NOAA Earth System Research Laboratory Physical Sciences Laboratory.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Coll, J.; Domonkos, P.; Guijarro, J.; Curley, M.; Rustemeier, E.; Aguilar, E.; Walsh, S.; Sweeney, J. Application of homogenization methods for Ireland's monthly precipitation records: Comparison of break detection results. *Int. J. Climatol.* **2020**, *40*, 6169–6188. [[CrossRef](#)] [[PubMed](#)]
- Cristiano, E.; ten Veldhuis, M.-C.; van de Giesen, N. Spatial and temporal variability of rainfall and their effects on hydrological response in urban areas—A review. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 3859–3878. [[CrossRef](#)]
- Hutchinson, M.F.; McKenney, D.W.; Lawrence, K.; Pedlar, J.H.; Hopkinson, R.F.; Milewska, E.; Papadopol, P. Development and Testing of Canada-Wide Interpolated Spatial Models of Daily Minimum–Maximum Temperature and Precipitation for 1961–2003. *J. Appl. Meteorol. Climatol.* **2009**, *48*, 725–741. [[CrossRef](#)]
- Werner, A.T.; Schnorbus, M.A.; Shrestha, R.R.; Cannon, A.J.; Zwiers, F.W.; Dayon, G.; Anslow, F. A long-term, temporally consistent, gridded daily meteorological dataset for northwestern North America. *Sci. Data* **2019**, *6*, 180299. [[CrossRef](#)] [[PubMed](#)]
- Poschod, B.; Ludwig, R.; Sillmann, J. Ten-year return levels of sub-daily extreme precipitation over Europe. *Earth Syst. Sci. Data* **2021**, *13*, 983–1003. [[CrossRef](#)]
- Camuffo, D.; Becherini, F.; della Valle, A.; Zanini, V. Three centuries of daily precipitation in Padua, Italy, 1713–2018: History, relocations, gaps, homogeneity and raw data. *Clim. Change* **2020**, *162*, 923–942. [[CrossRef](#)]
- Morbideilli, R.; Saltalippi, C.; Dari, J.; Flammini, A. A Review on Rainfall Data Resolution and Its Role in the Hydrological Practice. *Water* **2021**, *13*, 1012. [[CrossRef](#)]
- Daly, C.; Doggett, M.K.; Smith, J.I.; Olson, K.V.; Halbleib, M.D.; Zlatko Dimcovic, Z.; Keon, D.; Loiselle, R.A.; Steinberg, B.; Ryan, A.D.; et al. Challenges in Observation-Based Mapping of Daily Precipitation across the Conterminous United States. *J. Atmos. Ocean. Technol.* **2021**, *38*, 1979–1992. [[CrossRef](#)]
- Guidelines on the Calculation of Climate Normal (WMO-No. 1203). Available online: <https://library.wmo.int/records/item/55797-wmo-guidelines-on-the-calculation-of-climate-normals> (accessed on 1 March 2024).
- Guidelines on the Definition and Characterization of Extreme Weather and Climate Event (WMO-No. 1310). Available online: <https://library.wmo.int/records/item/58396-guidelines-on-the-definition-and-characterization-of-extreme-weather-and-climate-events> (accessed on 5 July 2023).
- della Valle, A.; Camuffo, D.; Becherini, F.; Zanini, V. Recovering, correcting and reconstructing precipitation data affected by gaps and irregular readings: The Padua series from 1812 to 1864. *Clim. Change* **2023**, *176*, 9. [[CrossRef](#)]
- Camuffo, D.; Becherini, F.; della Valle, A.; Zanini, V. A comparison between different methods to fill gaps in early precipitation series. *Environ. Earth Sci.* **2022**, *81*, 345. [[CrossRef](#)]
- Holder, C.; Boyles, R.; Syed, A.; Niyogi, D.; Raman, S. Comparison of collocated automated (NCECONet) and manual (COOP) climate observations in North Carolina. *J. Atmos. Ocean. Technol.* **2006**, *23*, 671–682. [[CrossRef](#)]
- Oyler, J.W.; Nicholas, R.E. Time of observation adjustments to daily station precipitation may introduce undesired statistical issues. *Int. J. Climatol.* **2018**, *38* (Suppl. S1), e364–e377. [[CrossRef](#)]
- Maurer, E.P.; Wood, A.W.; Adam, J.C.; Lettenmaier, D.P.; Nijssen, B. A Long-Term Hydrologically Based Dataset of Land Surface Fluxes and States for the Conterminous United States. *J. Clim.* **2002**, *15*, 3237–3251. [[CrossRef](#)]
- Kim, J.W.; Pachepsky, Y.A. Reconstructing missing daily precipitation data using regression trees and artificial neural networks for SWAT streamflow simulation. *J. Hydrol.* **2010**, *394*, 305–314. [[CrossRef](#)]
- Buishand, T.A. Some Methods for Testing the Homogeneity of Rainfall Records. *J. Hydrol.* **1982**, *58*, 11–27. [[CrossRef](#)]
- Pettitt, A.N. A Non-Parametric Approach to the Change-Point Detection. *Appl. Stat.* **1979**, *28*, 126–135. [[CrossRef](#)]
- von Neumann, J. Distribution of the Ratio of the Mean Square Successive Difference to the Variance. *Ann. Math. Stat.* **1941**, *13*, 367–395. [[CrossRef](#)]
- Peterson, T.C.; Easterling, D.R.; Karl, T.R.; Groisman, P.; Nicholls, N.; Plummer, N.; Torok, S.; Auer, I.; Boehm, R.; Gullett, D.; et al. Homogeneity Adjustments of in Situ Atmospheric Climate Data: A Review. *Int. J. Climatol.* **1998**, *18*, 1493–1517. [[CrossRef](#)]
- Yozgatligil, C.; Yazici, C. Comparison of homogeneity tests for temperature using a simulation study. *Int. J. Climatol.* **2016**, *36*, 62–81. [[CrossRef](#)]
- Guide to Climatological Practices (WMO-No. 100). Available online: <https://library.wmo.int/records/item/60113-guide-to-climatological-practices> (accessed on 1 March 2024).
- User's Guide of the Climatol R Package (Version 4). Available online: <https://www.climatol.eu/climatol4-en.pdf> (accessed on 25 July 2023).
- Kuya, E.K.; Gjeltén, H.M.; Tveito, O.E. Homogenization of Norwegian monthly precipitation series for the period 1961–2018. *Adv. Sci. Res.* **2022**, *19*, 73–80. [[CrossRef](#)]
- Alexandersson, H. A Homogeneity Test Applied to Precipitation Test. *J. Climatol.* **1986**, *6*, 661–675. [[CrossRef](#)]
- Hawkins, M. Testing a sequence of observations for a shift in location. *J. Am. Stat. Assoc.* **1977**, *72*, 180–186. [[CrossRef](#)]
- Camuffo, D.; della Valle, A.; Becherini, F. How the rain-gauge threshold affects the precipitation frequency and amount. *Clim. Change* **2022**, *170*, 7. [[CrossRef](#)]

28. Strumenti e Criteri di Osservazione e di Gestione Dei Dati. La Serie Pluviometrica 1984–2010 Dell’arpav. Available online: https://www.arpav.veneto.it/temi-ambientali/agrometeo/file-e-allegati/atlante-precipitazioni/20_strumenti-e-criteri-di-osservazione-e-di-gestione-dei-dati---la-serie-pluviometrica-1984-2010-dell2019arpav.pdf/@@display-file/file (accessed on 1 March 2024).
29. ERA5 Hourly Data on Single Levels from 1940 to Present. Available online: <https://cds.climate.copernicus.eu/cdsapp#!/dataset/10.24381/cds.adbb2d47> (accessed on 7 March 2023).
30. NOAA/CIRES/DOE 20th Century Reanalysis (V3). Available online: https://www.psl.noaa.gov/data/gridded/data.20thC_ReanV3.html (accessed on 7 March 2023).
31. Pfister, L.; Brönnimann, S.; Schwander, M.; Isotta, F.A.; Horton, P.; Rohr, C. Statistical reconstruction of daily precipitation and temperature fields in Switzerland back to 1864. *Clim. Past* **2020**, *16*, 663–678. [[CrossRef](#)]
32. Weller, G.B.; Cooley, D.S.; Sain, S.R. An investigation of the pineapple express phenomenon via bivariate extreme value theory. *Environmetrics* **2012**, *23*, 420–439. [[CrossRef](#)]
33. Gilleland, E.; Katz, R.W. extRemes 2.0: An Extreme Value Analysis Package in R. *J. Stat. Soft.* **2016**, *72*, 1–39. [[CrossRef](#)]
34. Ting, K.M. Confusion Matrix. In *Encyclopedia of Machine Learning*; Sammut, C., Webb, G.I., Eds.; Springer: Boston, MA, USA, 2011; p. 209.
35. confusionMatrix: Create a Confusion Matrix. Available online: <https://rdrr.io/cran/caret/man/confusionMatrix.html> (accessed on 7 March 2023).
36. Aguilera, H.; Guardiola-Albert, C.; Serrano-Hidalgo, C. Estimating extremely large amounts of missing precipitation data. *J. Hydroinform.* **2020**, *22*, 578–592. [[CrossRef](#)]
37. Bellido-Jiménez, J.A.; Gualda, J.E.; García-Marín, A.P. Assessing machine learning models for gap filling daily rainfall series in a semiarid region of Spain. *Atmosphere* **2021**, *1*, 1158. [[CrossRef](#)]
38. Longman, R.J.; Newman, A.J.; Giambelluca, T.W.; Lucas, M. Characterizing the uncertainty and assessing the value of gap-filled daily rainfall data in Hawaii. *J. Appl. Meteorol. Climatol.* **2020**, *59*, 1261–1276. [[CrossRef](#)]
39. Wold, S.; Esbensen, K.; Geladi, P. Principal Component Analysis. *Chemom. Intell. Lab. Syst.* **1987**, *2*, 37–52. [[CrossRef](#)]
40. Bro, R.; Andersson, C.A.; Kiers, H.A.L. N-way principal component analysis theory, algorithm and applications. *J. Chemom.* **1999**, *13*, 295–309. [[CrossRef](#)]
41. Bro, R. PARAFAC. Tutorial and applications. *Chemom. Intell. Lab. Syst.* **1997**, *38*, 149–171. [[CrossRef](#)]
42. Bro, R.; Smilde, A.K. Centering and scaling in component analysis. *J. Chemom.* **2003**, *17*, 16–33. [[CrossRef](#)]
43. Jolliffe, I.T. *Principal Component Analysis (Springer Series in Statistics)*, 2nd ed.; Springer: New York, NY, USA, 2002; pp. 1–488.
44. Camuffo, D.; Becherini, F.; della Valle, A. Relationship between selected percentiles and return periods of extreme events. *Acta Geophys.* **2020**, *68*, 4. [[CrossRef](#)]
45. Hersbach, H.P.; de Rosnay, B.; Bell, D.; Schepers, A.; Simmons, C.; Soci, S.; Abdalla, M.; Alonso Balmaseda, G.; Balsamo, P.; Bechtold, P.; et al. Operational Global Reanalysis: Progress, Future directions and Synergies with NWP, ECMWF ERA Report Series 27. 2018. Available online: <https://www.ecmwf.int/sites/default/files/elibrary/2018/18765-operational-global-reanalysis-progress-future-directions-and-synergies-nwp.pdf> (accessed on 17 February 2024).

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.