Received: 18 July 2020 | Revised: 30 November 2020 | Accepted: 4 January 2021

DOI: 10.1002/joc.7003

RESEARCH ARTICLE

International Journal of Climatology RMetS

Evaluation of gridded rain-gauge-based precipitation datasets: Impact of station density, spatial resolution, altitude gradient and climate

Andrés Merino[1] | **Eduardo García-Ortega**[1] | **Andrés Navarro**[1] | **Sergio Fernández-González**[2] | **Francisco J. Tapiador**[3] | **José Luis Sánchez**[1]

[1]Universidad de León (ULE), Atmospheric Physics Group (GFA), Environmental Institute (IMA), León, Spain

[2]State Meteorological Agency (AEMET), Santander, Spain

[3]Faculty of Environmental Sciences and Biochemistry, University of Castilla-La Mancha (UCLM), Institute of Environmental Sciences (ICAM), Earth and Space Sciences (ESS) Group, Toledo, Spain

Correspondence
Andrés Merino, Universidad de León (ULE), Atmospheric Physics Group (GFA), Environmental Institute, 24071 León, Spain.
Email: andres.merino@unileon.es

Funding information
Consejería de Educación, Junta de Castilla y León, Grant/Award Number: LE240P18; Ministerio de Ciencia e Innovación, Grant/Award Numbers: CGL2016-78702-C2-1-R, PID2019-108470RB-C22, CGL201

Abstract

Gridded precipitation datasets have been developed for data assimilation and evaluation tasks of weather and climate models and for climate analyses. Gridded data uncertainty evaluation is crucial to understand the limitations and feasibility. The development of high-resolution daily gridded precipitation datasets is desirable, but several factors need to be considered, namely rain gauge station availability, their spatial distribution, and orographic and climate characteristics of a study area. Quality assessment of gridded datasets can present difficulties when the influence of these factors is not thoroughly analysed. The main objective of this study was a detailed validation of precipitation grids based on four factors, that is, station density used for grid construction, grid spatial resolution, station altitude, and climate type. To this end, 18 grids were built using six spatial resolutions (0.01°, 0.025°, 0.05°, 0.1°, 0.2° and 0.4°) and three station densities (25, 50 and 75% of all available stations). Results indicate larger differences among the grids as a function of analysed factors. Station density was found to be the main factor, whereas grid spatial resolution had minor importance. However, the latter factor becomes more relevant in areas with strong altitude gradients and when a high station density is available. In addition, weak and moderate precipitation is overestimated on daily grids, whereas heavy precipitation cells are less frequent, reducing data variability. On the contrary, monthly and annual aggregates present less deviation from the observed distribution than daily comparisons. These findings question the applicability of the daily grid datasets for validation studies and climate analysis on a grid cell level.

KEYWORDS

daily observations, gridded precipitation, spatial resolution, station density, uncertainties

1 | INTRODUCTION

Calibration and evaluation tasks of precipitation estimation products, and of weather and climate models, require observational precipitation datasets. The grid cell data of these products should be compared with observed precipitation and, as a result, the need to build observational gridded datasets emerges. Thus, gridded precipitation products based on observed rain gauge data are widely used in climate studies (Royé and Martin-Vide, 2017; Cardell *et al*., 2020), model verification (Azorin-Molina *et al*., 2014), water resource management (Tramblay *et al*., 2019), and many other applications.

A set of precipitation gridded datasets have recently emerged. Globally, one highlight is the Global Precipitation Climatology Centre Full Data Reanalysis (Schneider *et al*., 2011; Becker *et al*., 2013). At continental scale, examples include the E-OBS dataset (Klein Tank *et al*., 2002) developed in the framework of the ENSEMBLES project for Europe, APHRODITE precipitation dataset for Asia (Yatagai *et al*., 2012), and North America regular gridded dataset (Chen *et al*., 2008; Daly *et al*., 2008).

High-quality gridded datasets require a sufficient number of stations within a grid cell to account for subgrid variability. Nevertheless, the density of available rain stations is not uniform and so most databases have large number of pixels without observational data. In addition, precipitation is not a continuous variable and has strong spatial gradients, especially in convective precipitation. All these facts increase errors on the grid, and its use for verification is risky.

Previous studies have analysed the effect of station density and interpolation methods on grid data quality. Hofstra *et al*. (2010) found substantial over-smoothing when fewer stations are used for interpolation. Rauthe *et al*. (2013) evaluated the relationship between mean absolute error and the percentage of stations included in the grid. Herrera *et al*. (2019) examined the effect of station density, interpolation method, and spatial resolution on observation-based gridded datasets, finding station density the most influential factor. Different interpolation methods of rain gauge data have been developed and implemented for gridded precipitation dataset development. Deterministic methods such as inverse distance weighting and nearest neighbour plus geostatistical methods such as kriging have been widely used (Herrera *et al*., 2012; Feki *et al*., 2017). However, several grid evaluations (Hewitson and Crane, 2005; Hofstra *et al*., 2008) have revealed that station density and spatial resolution leads to more sensitivity in the results than the choice of interpolation method. In summary, according to the literature, a larger mean absolute error and stronger smoothing is expected as the number of stations decreases.

In this study, the R package "reddprec" (https://cran.r-project.org/web/packages/reddPrec; Serrano-Notivoli *et al*., 2017a) was selected for grid construction. This method is based on the creation of reference values using generalized linear models (Serrano-Notivoli *et al*., 2017c). Also, the package contains the functions required to reconstruct original daily precipitation series and create grids. This method has been applied to build high-resolution daily gridded precipitation datasets in the same study area (Serrano-Notivoli *et al*., 2017b) and others (Decuyper *et al*., 2020). Serrano-Notivoli *et al*. (2017b) presented a new high-resolution daily gridded precipitation dataset for Spain, and some extreme precipitation indices were used as an example of climate applicability. Decuyper *et al*. (2020) reconstructed precipitation series to do a spatiotemporal assessment of beech growth. On the contrary, in our work, a validation of different grids made with the "reddprec" package was done to study their reliability and performance based on various factors.

The main objective of the present study was a detailed validation of daily precipitation grids based on four factors, that is, station density used for grid construction, grid spatial resolution, station altitude, and climate type. To this end, 18 grids were built using six different grids with horizontal resolutions 0.01°, 0.025°, 0.05°, 0.1°, 0.2° and 0.4°, and three station density levels (25, 50 and 75% of all available stations).

The work is organized as follows. A description of the study area and details of datasets are provided in Section 2. Section 3 describes the methods used. Results are in Section 4, with grid verification setting at daily scale (4.1) and multiple temporal aggregation (4.2) based on Köppen climates (4.3) and re-gridding methods (4.4). Finally, a discussion and conclusions are presented in Section 5.

2 | STUDY AREA AND DATASET

The study area corresponds to the Ebro basin (~85,000 km^2) on the northeastern Iberian Peninsula. Several reasons justify this selection, namely, a high density of rain gauges, strong altitude gradient due to complex orography (sea level to 3,400 m), and diverse types of climate. Based on the Köppen climate classification, there are cold continental climates in the Pyrenees at the north end of the basin (Dfb and Dfc), dry and hot Mediterranean climates in the lowlands (BSk, Csa and Csb), and humid temperate climates across the northwestern plains (Cfb and Cfa).

A set of 367 rain gauge tipping-bucket type stations was used. The stations are distributed throughout the

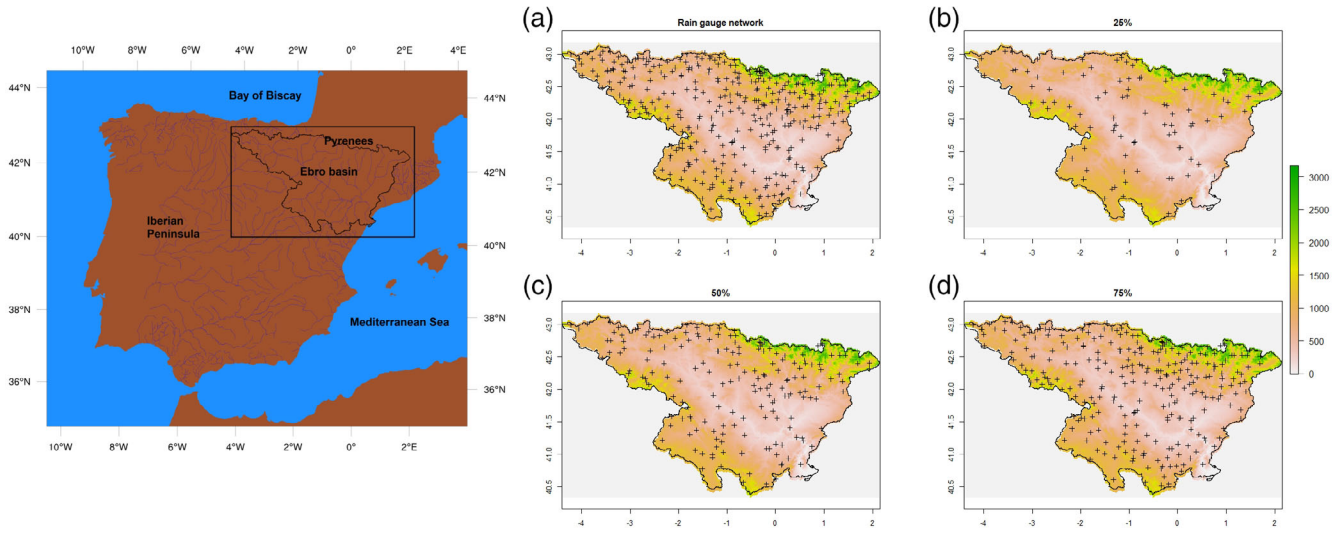

FIGURE 1 Topographic map of Ebro River basin. (a) Rain gauge network; (b–d) random selection of 25, 50, and 75% of rain gauges respectively for gridded interpolation

basin, with a greater density in the highlands (Figure 1a). Solid precipitation is frequent in winter, so highlands stations measure both solid and liquid precipitation. Daily precipitation data were retrieved from the 367 stations between 2008 and 2018. However, the number of available stations is not fixed and changed over time, as seen in Figure 2a.

3 | METHODS

The grids were built from 367 original stations by creating reference values using generalized linear models based on the 10 nearest observations, using altitude, longitude and latitude as covariates. These calculations were performed using the R package reddprec. That package contains functions to apply quality control (QC), identifying and removing suspect data from the original dataset, daily precipitation reconstruction for filling gaps, and creating grids (Serrano-Notivoli *et al*., 2017a).

The original dataset was filtered using the QC. The process uses five criteria to flag and remove suspect data: (a) Suspect data; (b) suspect zero; (c) suspect outlier; (d) suspect dry day; (e) suspect wet day (details in Serrano-Notivoli *et al*., 2017a). The QC process flagged and removed an annual average of 1.0% of the data (Figure 2b). There were no major differences in the number of removed data by year; only the first year showed slightly larger values. Suspect data and suspected zero were the main reasons to remove data in all years; only in the first years were there suspect outliers and wet values in substantial amounts, decreasing over the years. The percentage with suspect data found was smaller than that obtained using other databases with the same QC scheme (Serrano-Notivoli *et al*., 2017b).

Then, original missing data and data filtered by QC were filled using reference values. Total missing data after QC are shown in Figure 2c. Typically, the precipitation databases used for grid construction have an irregular number of available stations over time. In this case, the number of available stations increased in the first years, so the number of missing data declined markedly. Changes in the number of precipitation stations could be challenging, but the robustness of the package is based on the individual calculation of Reference Values (RVs) for each day and location.

Grid evaluation was conducted based on four factors, that is, station density used for grid construction, grid spatial resolution, station altitude, and climate type.

Spatial resolution was evaluated by defining six different grids with horizontal resolutions: 0.01°, 0.025°, 0.05°, 0.1°, 0.2° and 0.4°. The evaluation of station density was done by random selection of the stations included in the grid. Thus, for each defined spatial resolution three interpolations were done randomly using 25, 50, and 75% of all available stations (Figure 1b–d). The resolution was selected taking into account its use for the verification of climate models, with horizontal resolutions greater than 0.1°, or an NWP model with horizontal resolution less than 0.05°. The stations removed were not used for grid development. In this case 275, 183 and 92 were taken as reference for validation (validation set).

Grid validation in terms of station density, grid spatial resolution, and station altitude was portrayed by scatter plots and some statistical goodness-of-fit measures at daily, monthly and yearly scales. Grid precipitation was

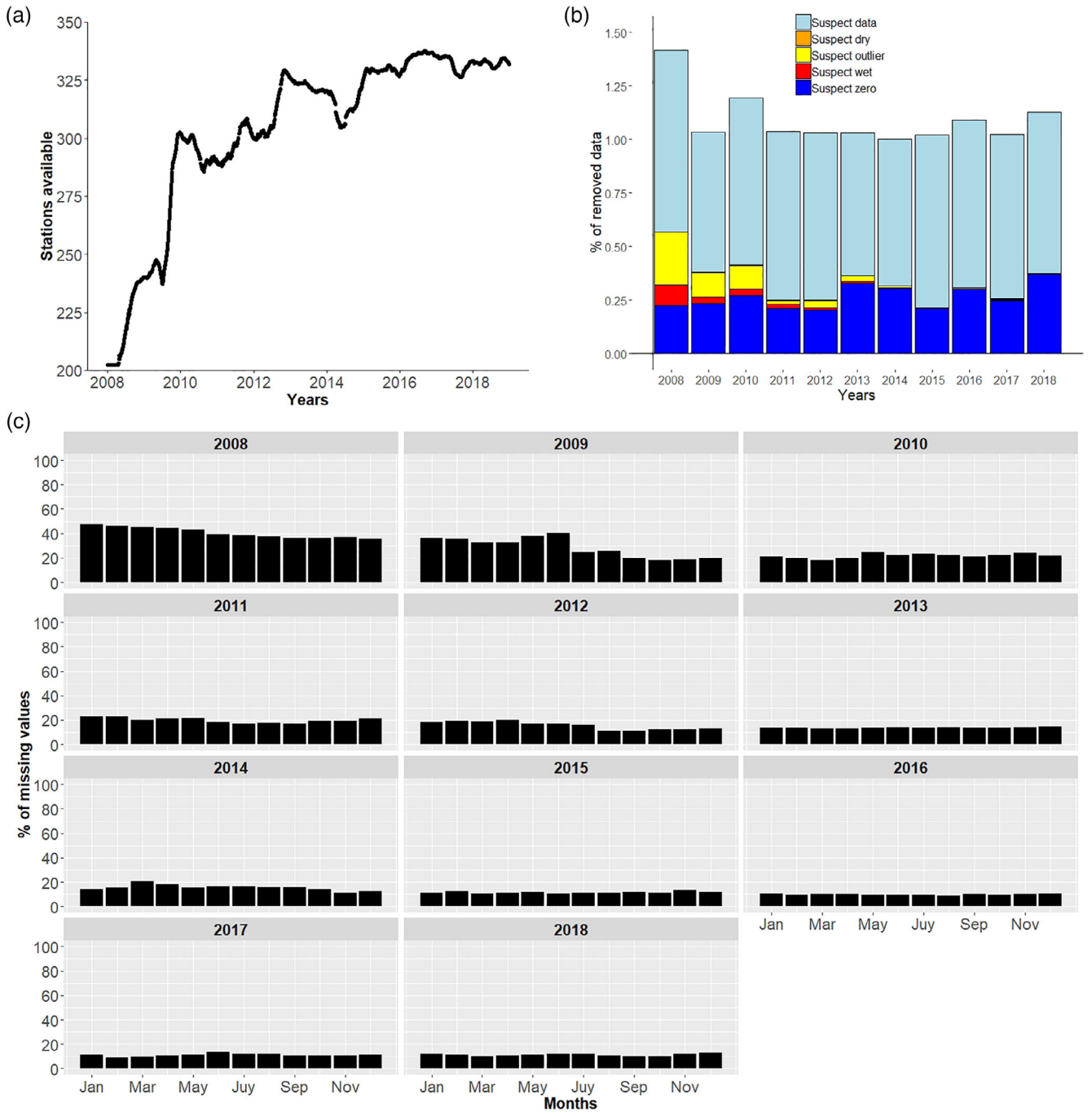

FIGURE 2 (a) Time evolution of number of stations available. (b) Percentage of removed data using five criteria of QC. (c) Percentage of missing data in the original files

taken at grid pixels where validation stations were located and the density of scatter points grouped in hexagons were represented using a colour scale.

To assess the performance of grids, the modified Kling-Gupta efficiency (KGE; Gupta *et al.*, 2009, Kling *et al.*, 2012) was selected. This index compares observed precipitation with grid precipitation, decomposing the total performance into three components: linear correlation (r), bias ratio (β), and variability ratio (γ). r is the Pearson product–moment correlation coefficient; β measures the total precipitation compared to ground-based observations, indicating the average tendency of the grid precipitation to underestimate ($\beta < 1$) or overestimate ($\beta > 1$); γ measures the relative dispersion between the gridded and the ground-based measurements. The optimal value for the KGE and all its components is one. It has been widely used to evaluate the performance of precipitation products (e.g., Lievens *et al.*, 2015;

TABLE 1 Accuracy of the wet/dry day estimates: Percent observed dry ($P = 0$) and wet ($P > 0$) days, and percent predicted dry (predicted = 0) and wet (predicted > 0) days on observed dry and wet days

		75%		50%		25%	
		$P = 0$	$P > 0$	$P = 0$	$P > 0$	$P = 0$	$P > 0$
	Observed	66.20	33.80	66.92	33.08	66.82	33.18
0.04	Predicted = 0	93.19	15.68	92.24	16.93	91.91	17.08
	Predicted > 0	6.81	84.32	7.76	83.07	8.09	82.92
0.2	Predicted = 0	94.62	12.61	92.72	15.61	92.40	16.09
	Predicted > 0	5.38	87.39	7.28	84.39	7.60	83.91
0.1	Predicted = 0	95.34	11.05	92.65	14.91	92.57	15.79
	Predicted > 0	4.66	88.95	7.35	85.09	7.43	84.21
0.05	Predicted = 0	95.75	10.00	92.94	14.56	92.40	15.50
	Predicted > 0	4.25	90.00	7.06	85.44	7.60	84.50
0.25	Predicted = 0	96.13	9.17	92.90	14.22	92.64	14.73
	Predicted > 0	3.87	90.83	7.10	85.78	7.36	85.27
0.01	Predicted = 0	96.29	8.43	92.77	14.24	92.80	14.30
	Predicted > 0	3.71	91.57	7.23	85.76	7.20	85.70

TABLE 2 Extended Köppen climates (Tapiador, 2019) for 75% of stations used in high station-density grids and mean station distances

Köppen climate	Stations	Mean distance (m)
BSk	51	10,812
Cfa	52	10,241
Cfb	134	9,282
Csa	17	13,770
Csb	14	10,746
Dfb	6	4,400
Dfc	1	8,119

Baez-Villanueva *et al*., 2018; Wang *et al*., 2018). Also, grid precipitation was evaluated by the mean absolute error (MAE) using the validation set for each case, and the probability of detection and ability to predict dry/wet days is shown in Table 1.

Total grid variance was evaluated, portraying the precipitation distribution by means of density plots, which represent a smoothed version of the histogram. For better understanding of the results, a logarithmic scale was used in these plots. Because there was a high frequency of days with little rainfall, the log transformation was done because precipitation values less than zero were impossible. Subsequently, the temporal and spatial variance was evaluated by computing the *SD* for the validation set on the basis of time and station, respectively.

In addition, considering the climate diversity of the study area, we evaluated the results based on the extended Köppen classification (Tapiador, 2019). In this case, the grids built with high station density (75%) were selected, and the stations used in the grid were classified into seven types of climates (Table 2). Subsequently, scatter plots of gridded daily precipitation versus station were developed for each climate type. The result facilitated an analysis of grid performance based on climatic characteristics, with the findings potentially applicable to other study areas.

Finally, we compared the grids built using the reddprec package and re-gridding from the higher spatial resolution grid. Sometimes grid products for climatic purposes are unavailable at the desired spatial resolution. This is usually resolved by re-gridding the finer grid dataset on the coarser resolution grid using an aggregation re-gridding method. In this case, the first-order conservative remapping method was used because it is most commonly used for precipitation (Jones, 1999; Philip, 1999). This method is based on the ratio of source cell area overlapped by the corresponding destination cell area. The objective was to determine if the quality of the grid changed after application of the re-gridding technique.

4 | RESULTS

4.1 | Daily scale

Scatter plots between observed and grid precipitation at daily scale, together the statistical goodness-of-fit measures (KGE, r, β and γ) are shown in Figure 3. All analysed grids show average underestimation compared

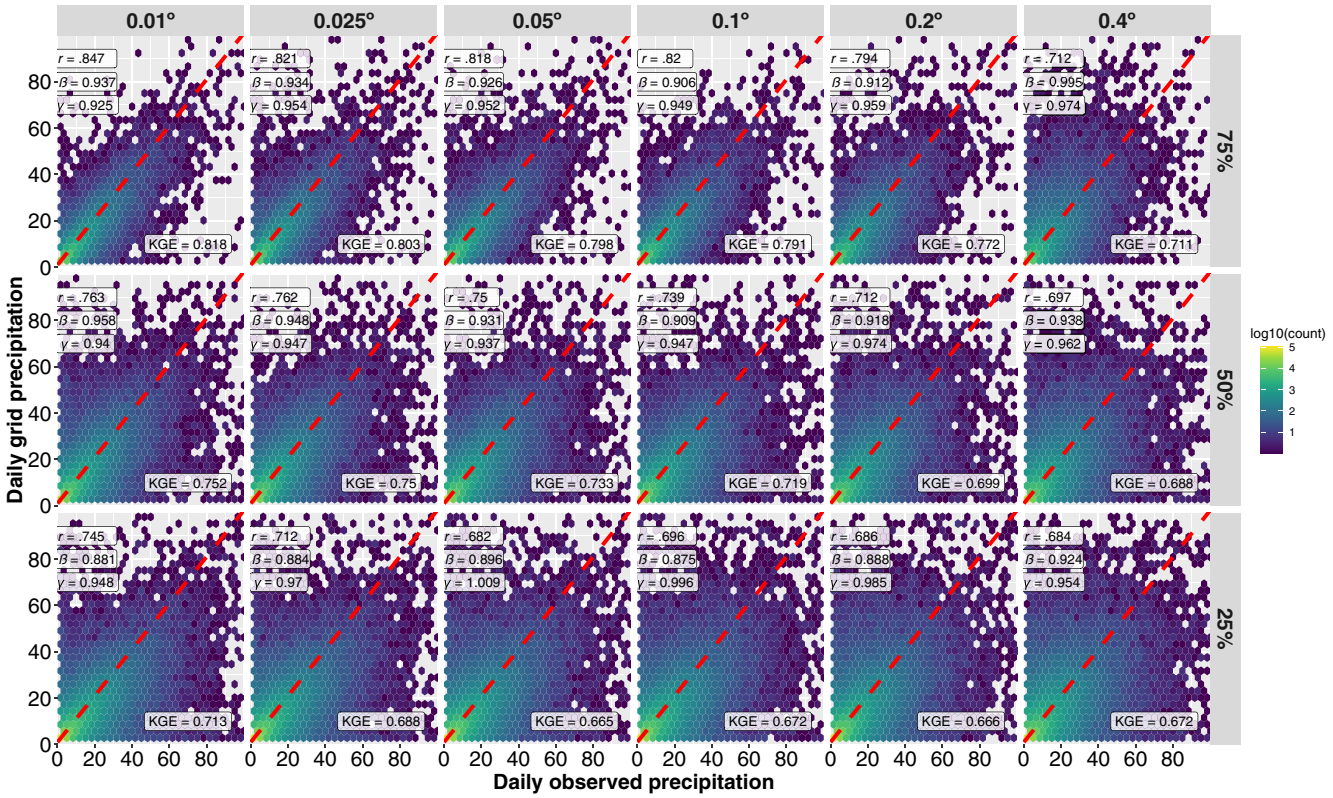

FIGURE 3 Scatter plots of observed daily precipitation versus grid precipitation. Rows show different station densities (75, 50, and 25%) and columns grid spatial resolutions (0.01°, 0.025°, 0.05°, 0.1°, 0.2°, and 0.4°)

with the observed precipitation. This is particularly clear because for all cases, the bias ratio (β) is less than 1. Maxima of daily precipitation are often due to convective events that are frequent in the Ebro Valley in summer. Convective precipitation has strong spatiotemporal variability, so the smoothing of large values on the grid could have contributed to the observed underestimation. This is supported by the fact that γ has similar values and are slightly less than one in all grids, resulting in underestimation of the observed variability. Analysing separately the effects of grid spatial resolution and station density, it seems clear that station density had greater influence on grid data quality. When 75% of available stations were used (first row), the KGE index showed better results (0.82–0.71), worsening as the number of stations diminished (0.75–0.68 for 50% of the stations and 0.7–0.67 for 25%). Conversely, the grid spatial resolution did not have such influence, although improvements were observed when the resolution was increased. This was more prominent when a greater station density was used (KGE = 0.71 for 0.4° and 0.82 for 0.01°). This suggests that the use of high spatial resolution is only useful when there is a high station density.

The mean absolute error for each grid was computed seasonally and for the entire period (Figure 4). General results affirm the conclusions previously described. Errors at daily scale became strongly dependent on station density. The behaviour of spatial resolution was remarkable; for the complete precipitation series, improvement was only noticeable when station density was high (MAE of 1.1–0.7 mm/day using 75% of stations and 1.25–1.05 mm/day using 25%). Analysing the results at seasonal scale, an increase in resolution of the grid gave improvement regardless of station density for winter precipitation. In contrast, spatial resolution had no noticeable effect in summer. Furthermore, the effect of station density was linear for all seasons, except in summer when substantial improvement was only produced when 75% of stations were used.

Overall results show larger errors in spring (MAE of 0.8–1.4 mm/day). In that season, rainfall is abundant across the entire basin, in contrast with winter and summer, when droughts affect large areas in the central study area. Small errors in summer (MAE = 0.7–1.1 mm/day) are notable, when larger values might be expected because of the convective precipitation. However, this can be explained by the prevailing dry conditions during that season.

These results reveal the outstanding role of dominant rainfall patterns for selecting the most suitable grid. When stratiform precipitation is dominant, both station

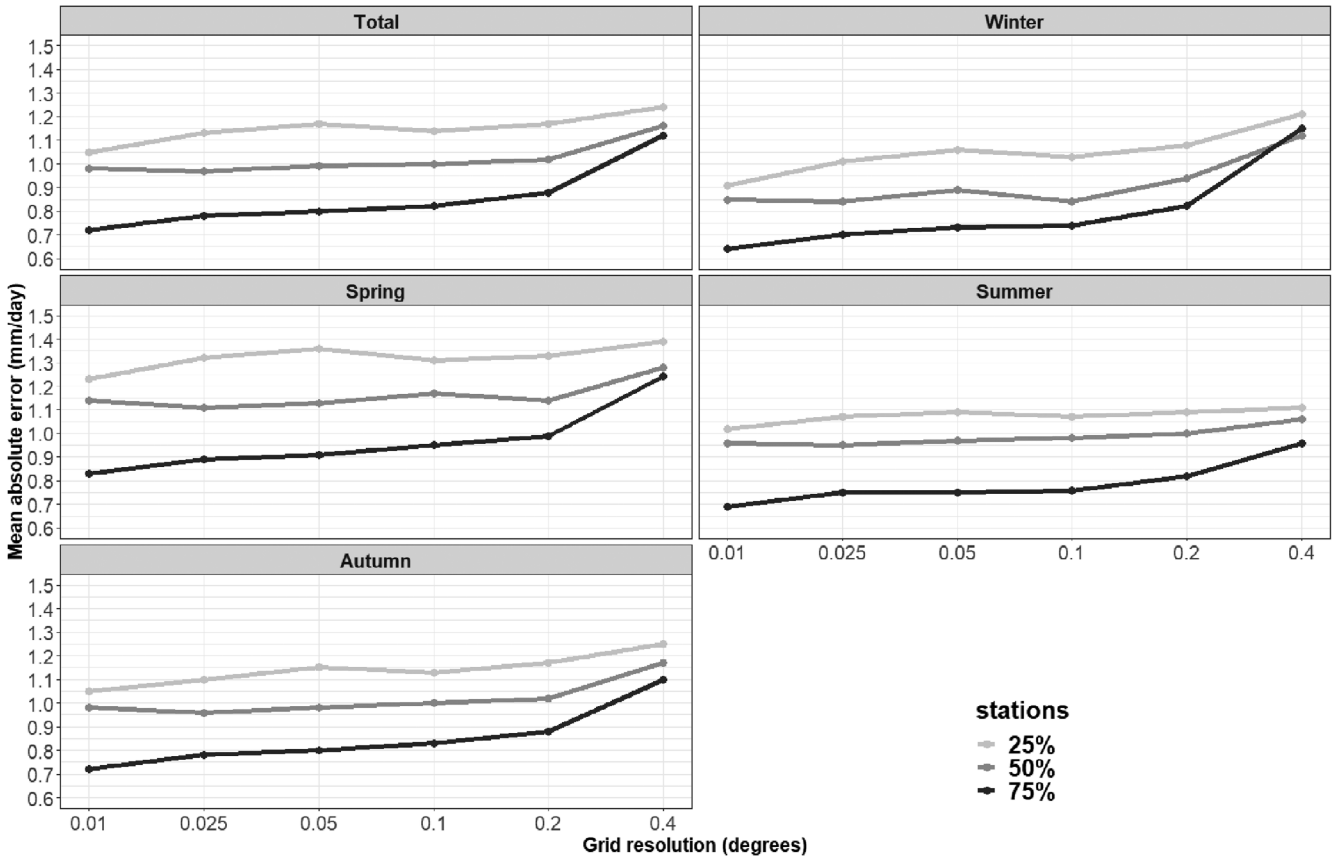

FIGURE 4 Mean absolute error (mm) of daily precipitation for each grid (station density as lines and spatial resolutions on *x*-axis), for season and entire period

density and spatial resolution are relevant factors in grid construction. On the contrary, for convective precipitation, station density is the most important factor, and a greater density is necessary.

One of the main features of the study area is its complex orography. It is well known that in elevated areas, spatial precipitation gradients are more pronounced, so greater uncertainties in the grid are anticipated there. To check this, station altitude was represented as a function of the grid mean absolute error (Figure 5). The results are only shown for grids built with 75% of the stations; for the remaining grids, the results are similar. Although errors tend to increase with station altitude, the strong dispersion of the results does not permit categorical confirmation of this relationship. The relationship is more evident upon increasing the grid spatial resolution, because of the smoothing effect of low resolutions. On the other hand, the initial station distribution, with greater density in mountainous areas, can reduce the errors in elevated areas. This means that when the observation stations are irregularly distributed, prioritizing areas of complex orography over those on the plains, it promotes a more uniform distribution of grid uncertainties.

The use of precipitation grids for climate applications has been widespread. In this regard, it is not only important to obtain an acceptable mean precipitation value but also data variance. When the grid is unable to reproduce the observed variance, a given statistical analysis can yield uncertain results, such as a trend analysis. To evaluate the grid variance, we used density plots with log scale to construct the daily precipitation distribution (Figure 6a). The observed deviations of rainfall <1 mm between observation (black line) and gridded (colour lines) are due to rain gauge resolution (0.2 mm), producing jumps in the density representation. The results are clear, irrespective of station density and spatial resolution. The grids have a notably higher frequency of daily precipitation around 2–10 mm, decreasing the frequency of heavier rainfall relative to the observed distribution. This means that weak and moderate precipitation are overrepresented on the grid, while heavy precipitation is less frequent, decreasing data variability. Surprisingly, the lower-resolution grids (green) underestimate to a lesser extent daily precipitation >50 mm. Furthermore, the fact that the number of stations changes over time does not affect evolution of the variance (not shown), confirming the non-dependence of station density on grid variance.

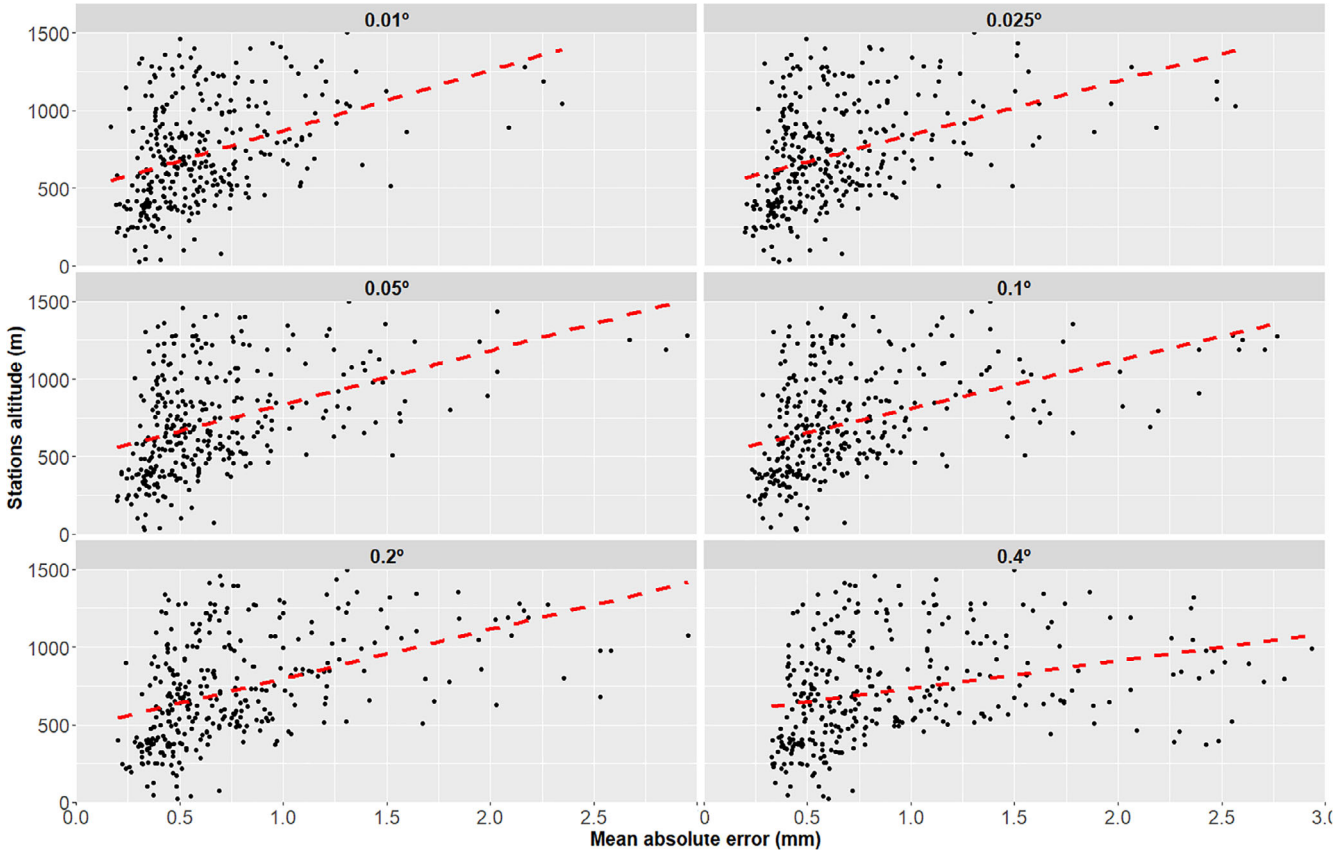

FIGURE 5 Scatter plots of mean absolute error of daily precipitation versus station altitude by grid spatial resolution, using 75% of available stations. Linear regression model (line dashed)

Subsequently, the temporal (Figure 6b) and spatial (Figure 6c) variance was evaluated by computing density plots of *SD* for the validation set relative to time and station, respectively. Temporal smoothing of the grids is more evident than spatial smoothing. The temporal *SD* distribution of observed precipitation (black line) is clearly shifted toward values larger than those of the grid distributions, regardless of station density and spatial resolution. This result is a consequence of the changes in precipitation distribution mentioned above. Spatial smoothing of the grids is also observed but to a lesser degree. This is because the spatial *SD* distribution of the grids had more frequent small values than the observed. In this case, station density appeared to have greater weight in the preservation of spatial variability.

Finally, the probability of detection and ability to predict dry/wet days is evaluated in Table 1. The probability of detection of wet days is ~80% and that of dry days is ~90%. At all grids, more dry days were detected than those observed. The detection of wet days was improved for high spatial resolution and station density. Consistent with the previous outcomes, spatial resolution only appeared to matter when a high density of stations was used.

4.2 | Multiple temporal aggregation

Scatter plots between observed and grid precipitation at monthly scale, together the statistical goodness-of-fit measures, are shown in Figure 7. Monthly aggregates show similar parameters in comparison to daily precipitation (KGE = 0.65–0.82). In addition, conclusions regarding the daily analysis are similar to those from the monthly aggregate analysis, namely: Station density is the dominant factor with regard to grid spatial resolution; higher spatial resolutions improve the results only when there is a high density of stations; irrespective of station density and spatial resolution, grid precipitation tends to underestimate the observed precipitation ($\beta < 1$) and the observed variance ($\gamma > 1$). The yearly aggregate analysis (Figure 8) confirms the slight improvement in the KGE parameter for greater temporal period aggregation (KGE = 0.6–0.85). The reproduction of annual accumulated precipitation is particularly acceptable for grids with high station density and spatial resolutions greater than 0.1° ($r \sim .87$). These results support grid suitability for use in climate analysis. For grids built with lower station density, there is a predominant underestimation ($\beta \sim 0.88$), while the bias is close to

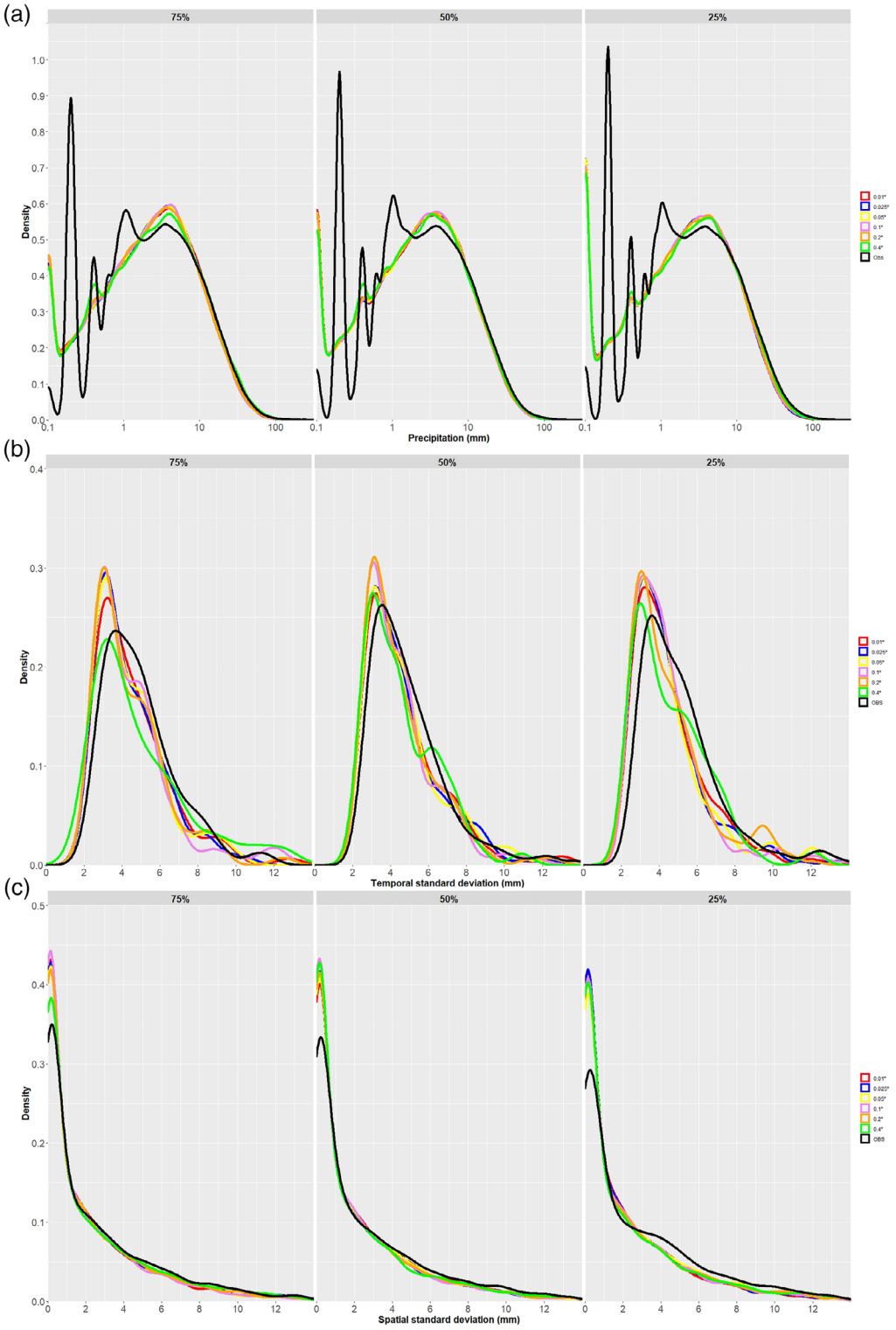

FIGURE 6 Density plots for different grids (columns: Stations density; colours within plot: Spatial resolution) and observed precipitation (Obs): (a) daily precipitation; (b) *SD* over time; (c) *SD* over stations

1 when high station density is used. The variance results remain unrelated to the spatial resolution and station density. In this case, the spatial resolution factor becomes more important relative to the monthly and daily analyses. For any station density, the results improve as spatial resolution increases, reaching 0.1–0.05°, from which improvements are imperceptible. It is not surprising that these values match the average distances of stations used in each density category studied.

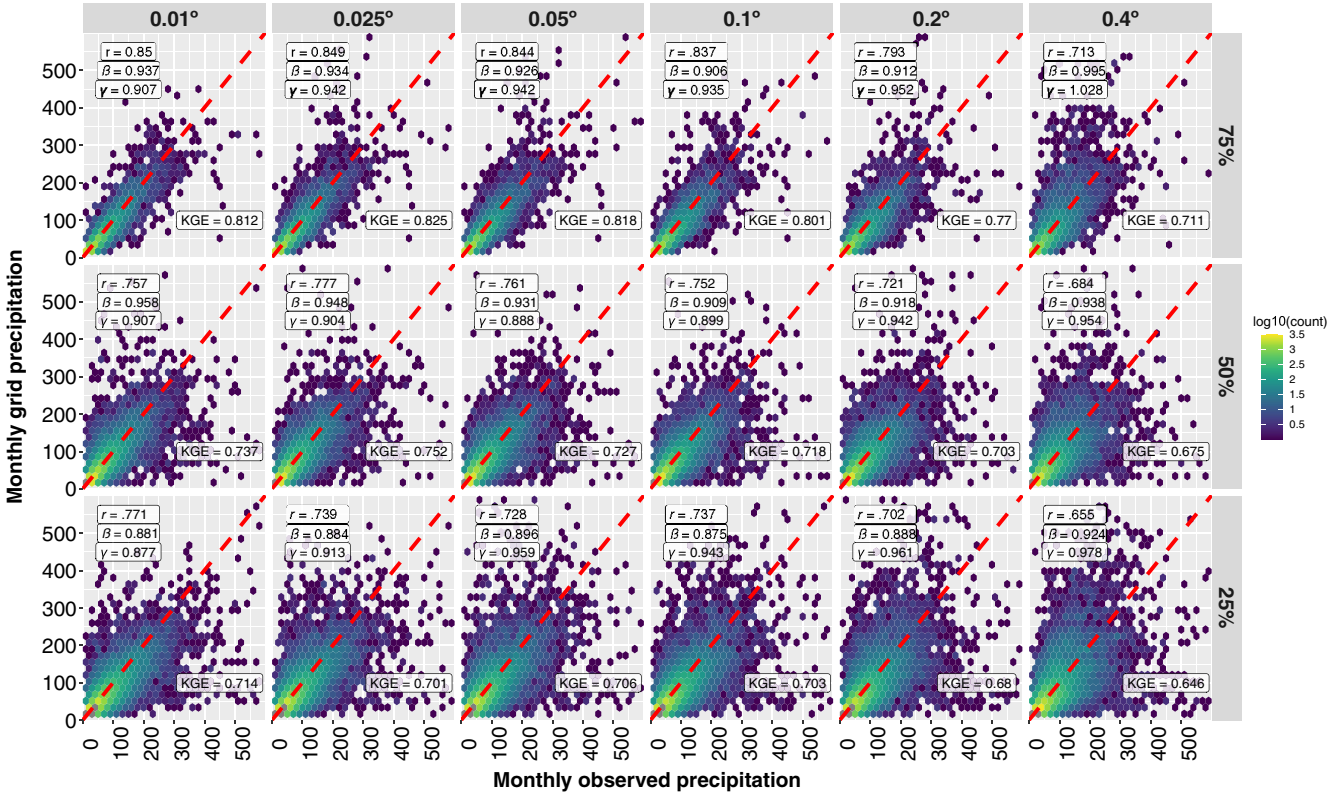

FIGURE 7 As in Figure 3 but for monthly precipitation aggregates

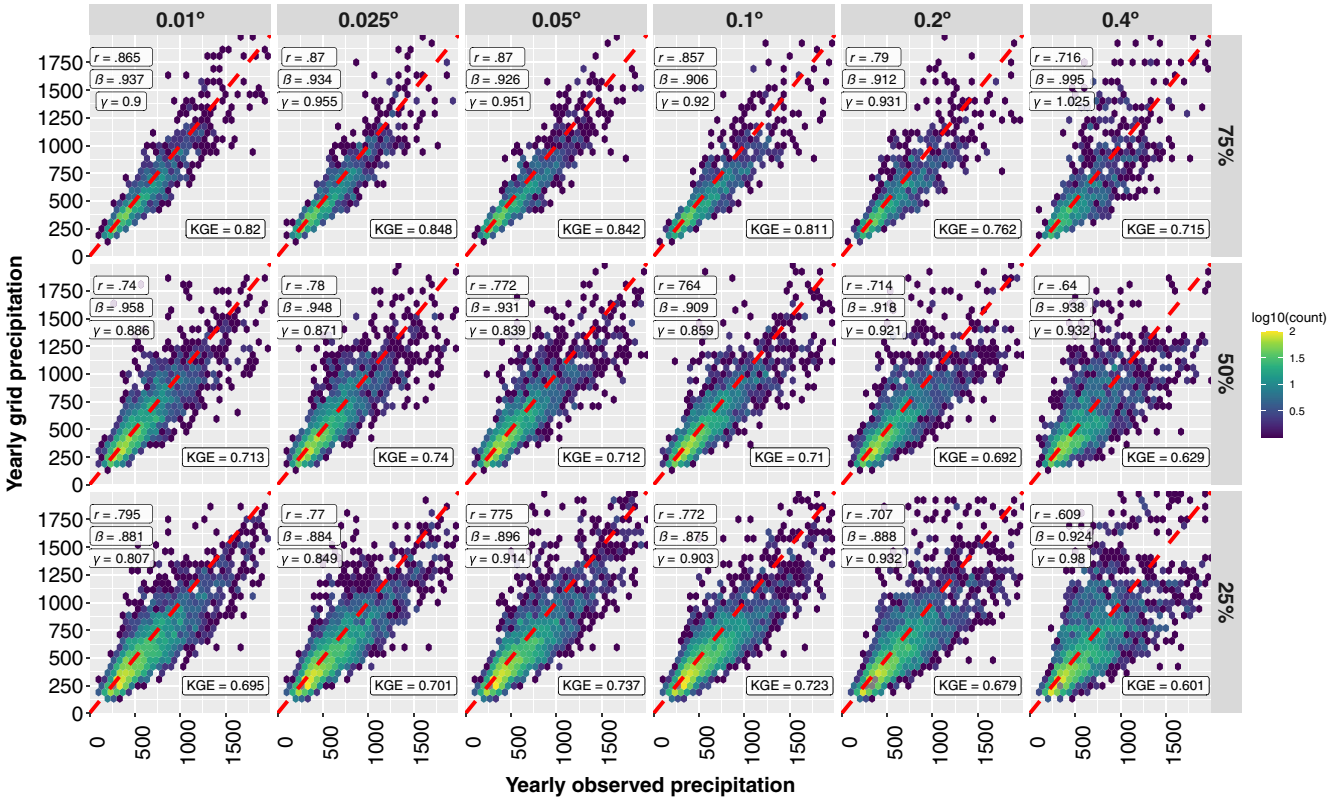

FIGURE 8 As in Figure 3 but for annual precipitation aggregates

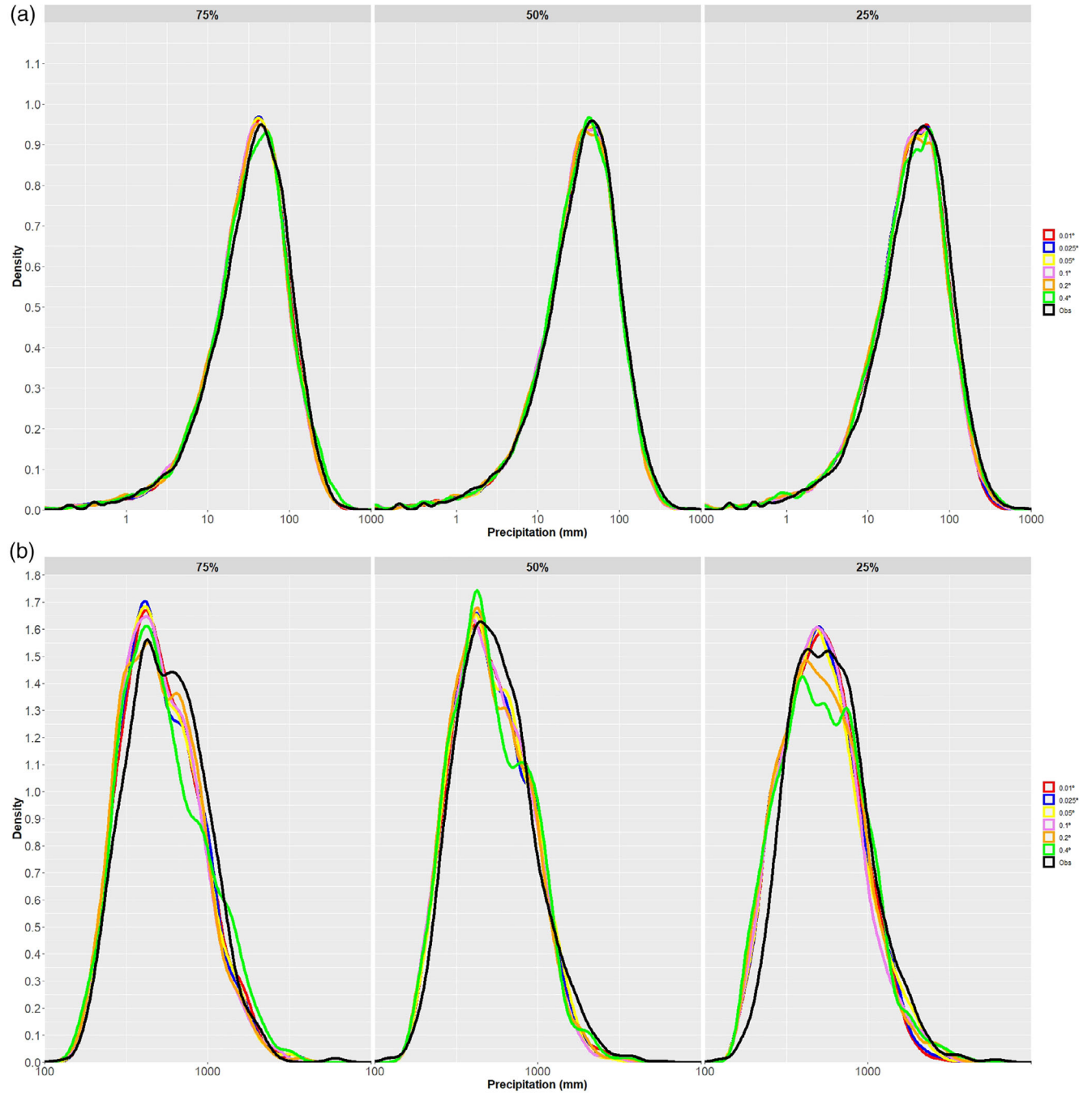

FIGURE 9 Density plots of monthly (a) and yearly (b) precipitation aggregates for different grids (columns: Stations density; colours within plot: Spatial resolution) and observed precipitation (Obs)

Similar to the daily analysis, the assessment of variance at monthly (Figure 9a) and yearly (Figure 9b) scales was performed using density plots with log scale. The results of observed distributions for monthly and annual precipitation were preserved accurately by the grids, especially for monthly precipitation. A slight shift is still observed in the observed distribution (black line) to the right of the plot, as well as overrepresentation of intermediate values of accumulated annual precipitation. This leads to extreme value smoothing, but in the monthly distribution, there is reasonable similarity in the value distribution, suggesting the potential use of grids for climate analysis.

Finally, mean absolute errors were compared for monthly, seasonal and yearly aggregates on each grid (Figure 10). A reduction of errors for larger temporal aggregates is evident, although this was not linear. Grid uncertainties decreased substantially for the monthly

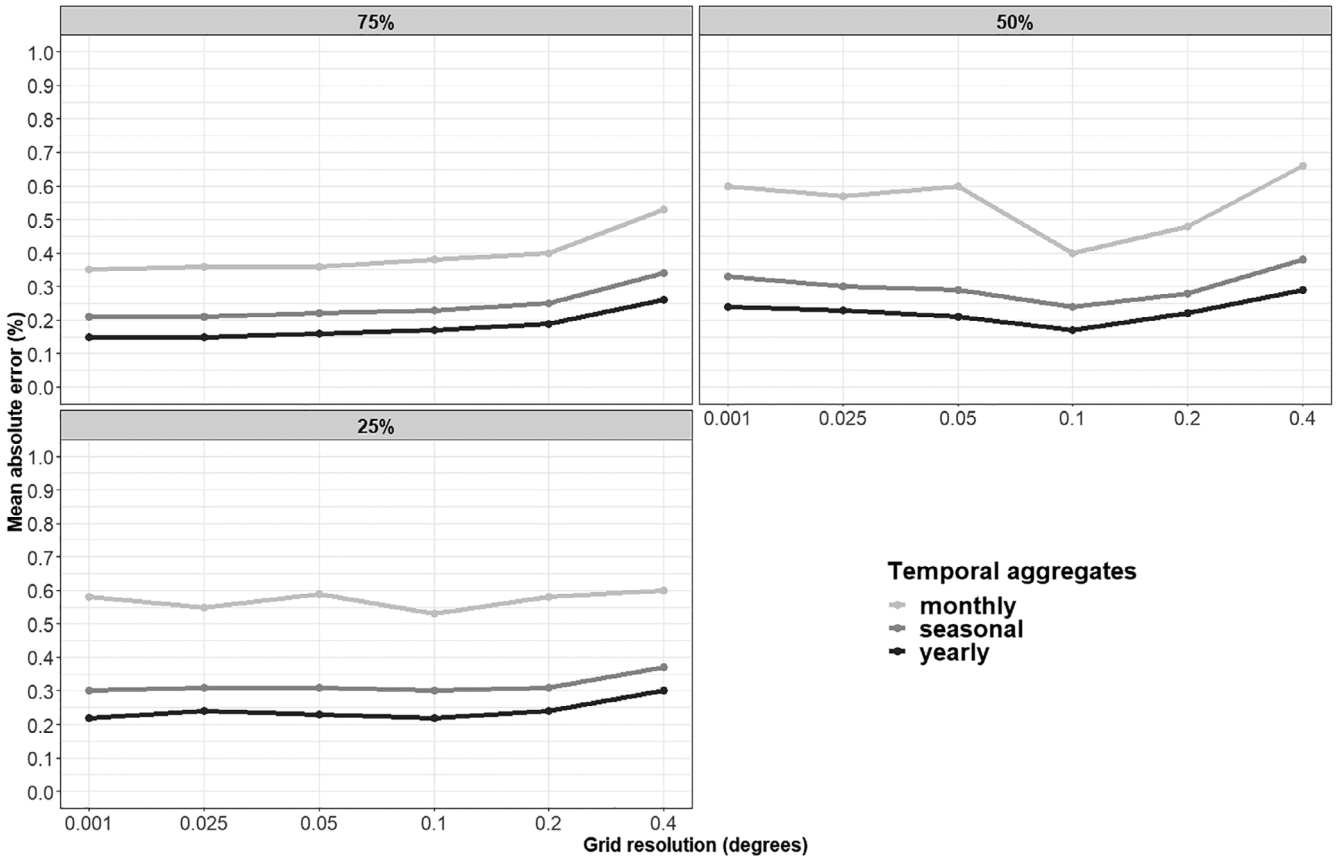

FIGURE 10 Mean absolute error of precipitation (%) for several temporal aggregates (lines) and each grid (station density in each plot and spatial resolution on x-axis)

aggregates (30–60%) relative to seasonal (20–35%), but there were only small additional improvements shown by the annual aggregates (15–25%). Regarding the performance of each grid, the role of grid spatial resolution was surprising. Not only were there no improvements seen for high-resolution grids built at low station densities, but in some cases, such as using 50% of the stations, the results of high resolutions worsened. It is clear once again that the use of high resolutions only makes sense when high station densities are available, as observed when using 75% of the stations in the network. Furthermore, improvements are evident as station density increases, with this being the main factor affecting grid data quality.

4.3 | Köppen climates

There are seven different climates in the study area according to the extended Köppen climate classification (Table 2; Tapiador, 2019): Cold continental climates in the Pyrenees at the north end of the basin (Dfb, Dfc); dry and hot Mediterranean climates in the lowlands (BSk, Csa and Csb), and humid temperate climates in the northwestern plains (Cfb and Cfa). The evaluation of daily precipitation is shown in Figure 11 via scatter plots. The first highlight is the difference in grid performance between the Cfa (KGE = 0.83–0.88) and Cfb (KGE = 0.63–0.80) climates, despite a similar mean station distance used in grid construction (Table 2). The Cfa climate is in a transition zone between a temperate oceanic climate (northwestern edge of the basin) and dry and hot Mediterranean climate (basin centre). Neither climate has a dry season but precipitation is heavier with Cfb, with changes of precipitation pattern during the year. In addition, Cfb areas show a pronounced precipitation gradient as a consequence of the strong altitude gradient. These characteristics appear to have an important weight in the behaviour of the grid, even more important than station density, a factor that heretofore has been shown as dominant.

In dry and hot Mediterranean climates of the lowland basin (BSk, Csa and Csb), with similar mean distances between stations (Table 2), the results are similar (KGE ~ 0.85 for higher grid spatial resolution). These areas do not have major altitude changes and spatial precipitation

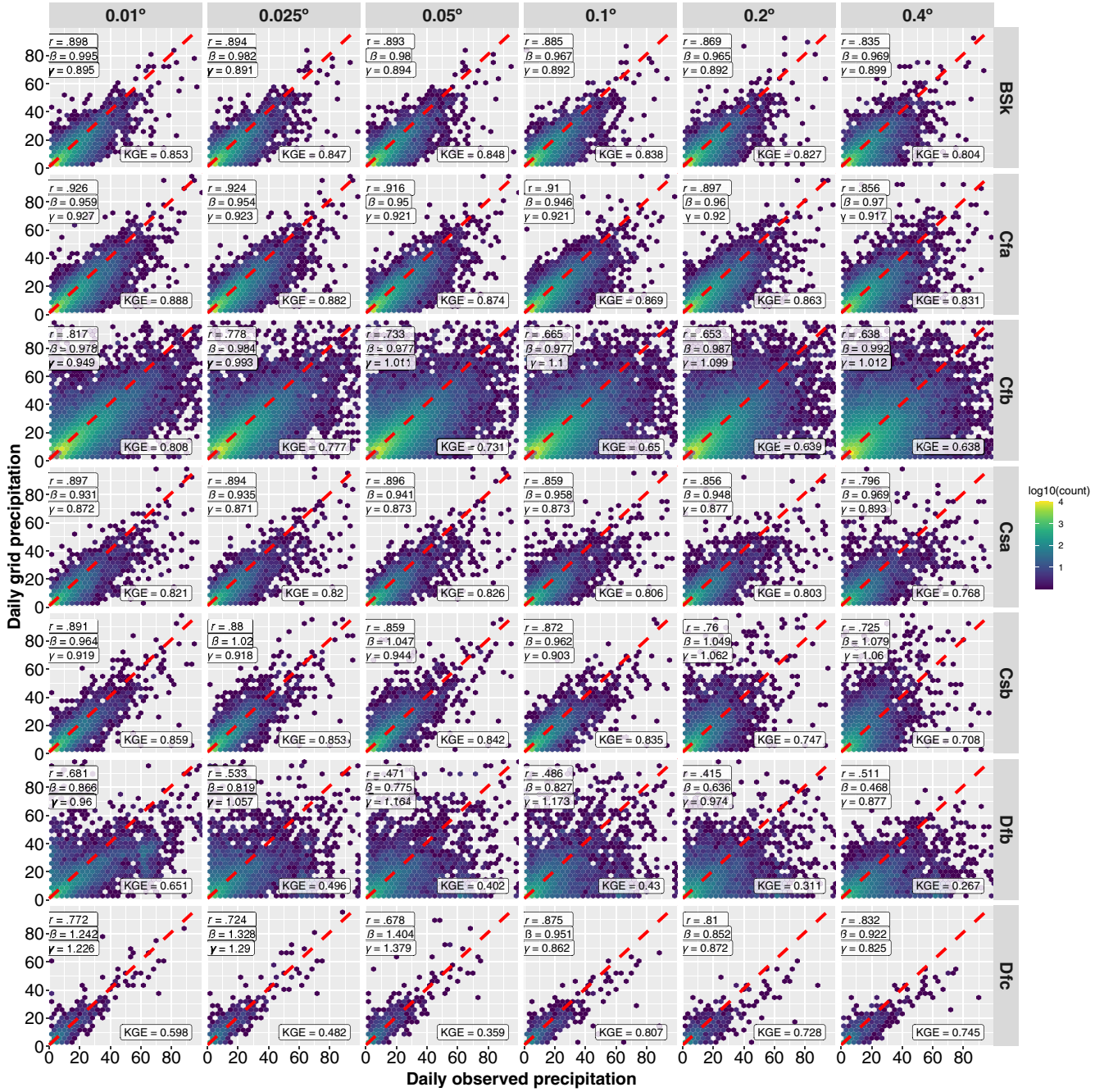

FIGURE 11 Scatter plots of observed daily precipitation versus grid precipitation by climate, using grids built with high station density. Rows show Köppen climates (Tapiador, 2019) and columns grid spatial resolution

gradients are only strong for convective precipitation. Consequently, the grids have similar results and an acceptable quality.

The poorest results were for cold climates of the high Pyrenean mountains (Dfb), despite having the minimum distance between stations (KGE = 0.26–0.65). Dfc had better performance, but there was only one station in this area, so the results cannot be considered representative. In Dfb areas, there was marked underestimation of precipitation by the grid, and only for high spatial resolution grids was precipitation reasonably represented.

The above results emphasize the importance of climate type in grid behaviour. These climate classifications are closely linked to orography and rainfall regime. Despite this, station density appeared as the main influence on grid quality. Spatial resolution is relevant in areas with strong altitude gradients (Cfb, Dfb, and Dfc). Precipitation regime appears not to play a decisive role. Climates where convective precipitation is dominant (BSk, Csa) do not present poorer behaviour than those where stratiform precipitation is predominant (Cfb, Dfb). The reasons for this can be attributed to the location of

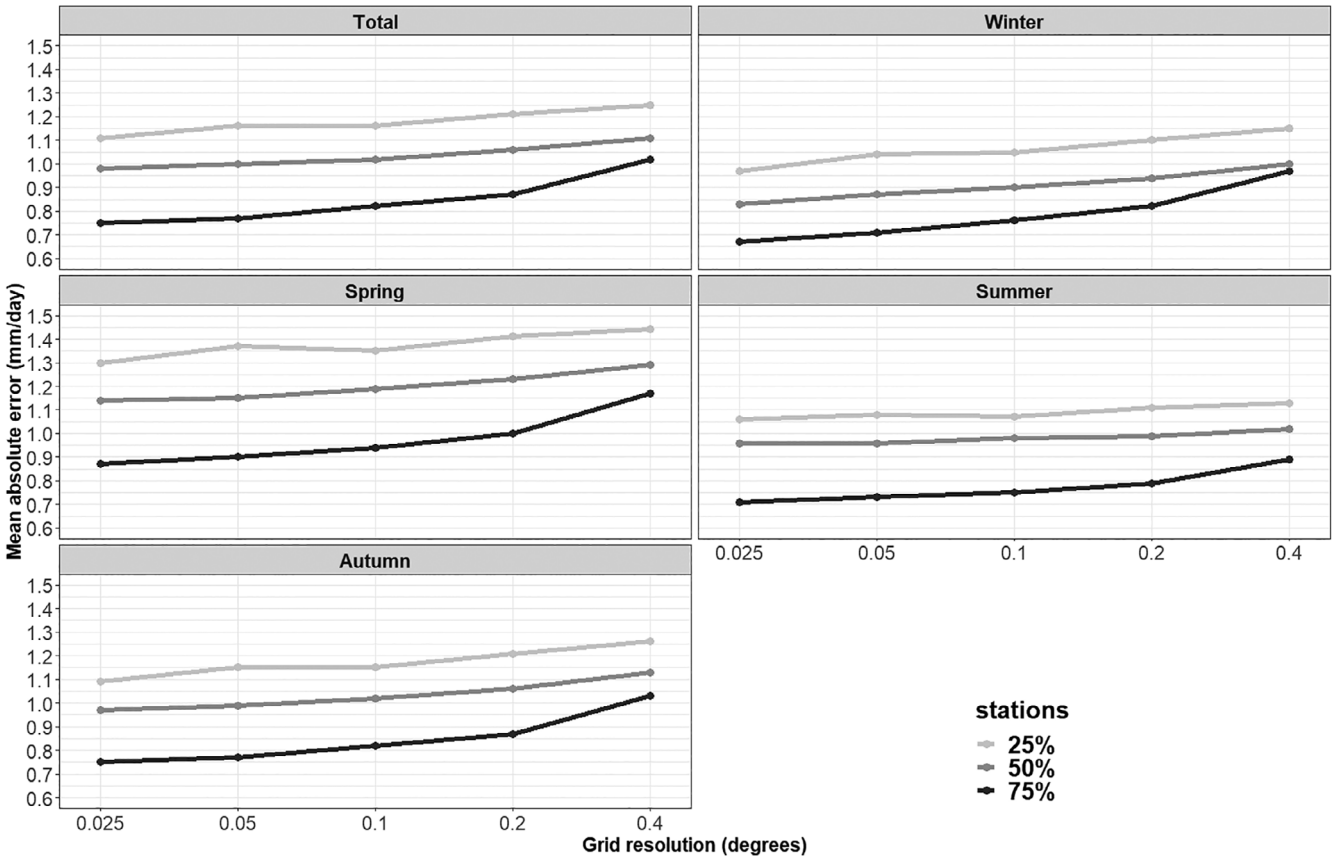

FIGURE 12 As in Figure 4 but for conservative interpolation from 0.01° spatial resolution

the latter areas. Their complex relief favours strong spatial precipitation gradients. As a result, grid behaviour is similar to that of convective precipitation.

4.4 | Re-gridding from the original grids

Frequently, users of grid precipitation find that the available grids do not have the proper resolution for their studies. The ideal process would be to rebuild the grid with the desired resolution, but sometimes the original observational data are unavailable. Thus, the need for re-gridding techniques emerges. We evaluated the application of conservative interpolation, most commonly used for precipitation, to original grids of 0.01° resolution by re-gridding at lower resolutions (Figure 12). Comparing these results with those from the grid validation at daily scale (Figure 4), one can see similar mean absolute errors. Initial grid quality determines the re-gridding errors. Therefore, performance of the station density factor analysed in Figure 4 is preserved, as well as the seasonal behaviour of the errors. The main differences emerged upon analysing the re-gridding effect on spatial resolution. Here, the errors increased linearly as the grid resolution decreased, without assessing any particular seasonal behaviour. Considering these results, it is clear that re-gridding from high to low resolution does not generate a substantial decrease in quality relative to grids constructed with a specific resolution. However, it should be emphasized that performing interpolations from low to high resolution is not useful, since errors are retained from the original grid at low resolution.

5 | DISCUSSION AND CONCLUSIONS

The construction and validation of gridded precipitation has seen increasing interest in recent years, owing to its applications to weather, hydrology and climate studies (Gervais *et al.*, 2014; Avila *et al.*, 2015; Bianchi *et al.*, 2016). However, the use of grid products requires prior in-depth analysis to assess their uncertainties. The quality of such datasets could be unreliable for the established goal.

The present study provides an in-depth evaluation of precipitation grid performance based on four factors, that is, station density used for grid construction, grid spatial resolution, station altitude and climate type. Interpolation methods were not evaluated. Several authors

(Hewitson and Crane, 2005; Hofstra *et al*., 2008) have claimed that these factors are conducive to greater sensitivity in the results than the choice of interpolation method. All grid calculations were performed using the R package reddprec (Serrano-Notivoli *et al*., 2017a), because it had already been used in our study area (Serrano-Notivoli *et al*., 2017b).

The results described above support and extend those of Herrera *et al*. (2019), who also highlighted station density as the fundamental factor and suggested different grid performance due to climatic variability and the complex orography. Both studies show that grid spatial resolution has minor importance. However, we have established herein how this factor becomes more relevant in areas with strong altitude gradients and when there is high station density.

The mean absolute error at daily scale was found to be ~0.7 mm/day on the best grid (75% of stations and 0.01° spatial resolution) and ~1.2 mm/day on the poorest grid (25% of stations and 0.4° spatial resolution). These results contrast with those of Rauthe *et al*. (2013), who built a gridded dataset with a high spatial resolution of 1 km^2 and daily temporal resolution for central Europe. Those authors had errors ~2 mm/day, but with strong spatial and temporal variability. However, although we also had temporal variability, with larger errors in spring, this was not as pronounced. The contrasting rainfall patterns between the two study areas may explain these differences. The sensitivity analysis of station density by Rauthe *et al*. (2013) showed a linear increase in mean absolute error as the percentage of stations included in the grid decreased. Those results agree with those presented herein and again highlight the importance of this factor. Regarding the probability of detection of dry and wet days, the results are similar to those of Serrano-Notivoli *et al*. (2017b) for the Iberian Peninsula, despite the fact that the frequency of wet days in the database used in that study was lower (20 vs. 33%). Similar results were also reported by Serrano-Notivoli *et al*. (2017c).

However, the general results (MAE, bias, correlation and variability) indicate that gridded daily precipitation may not be sufficiently accurate for use in calibration and verification. This statement concurs with the findings of other studies (Hofstra *et al*., 2010; Kysely and Plavcova, 2010; Maraun *et al*., 2012) that question the applicability of the grid datasets for validation studies on a grid-cell level, particularly in mountainous regions. Another technical problem with daily precipitation grids is the smoothed surfaces relative to reality, that is, they underestimate the field's variance (Beguería *et al*., 2016). Regardless of factors analysed, we observed a greater frequency of daily precipitation 2–10 mm and a smaller frequency of heavier precipitation. The smoothing is observed both spatially and temporally, although the latter is more noticeable. Thus, in daily precipitation, but also in monthly and annual aggregates, there was a decrease in the variability of the grids, with variability ratios less than one. The weak dependence of station density on grid variability makes the use of databases with no fixed station number over time, less problematic. This scenario is frequent when a long time series of precipitation is used. Similar results have been found by analysing several gridded precipitation datasets (Rauthe *et al*., 2013; Sungmin *et al*., 2016). Furthermore, using the same method in grid construction, Serrano et al. Serrano-Notivoli *et al*., 2017b; Serrano-Notivoli *et al*., 2017c observed an overestimation of small values and underestimation of large values, yielding a global displacement in the distribution of daily precipitation. This result is fully consistent with that described in the present work, and it has also been shown that the spatial resolution and station density are largely irrelevant in capturing the precipitation distribution. These results, together with γ parameter, show a slight reduction in grid variability and could bring into question the applicability of the grid datasets at daily scale for certain climate applications, such as trend analysis.

Precipitation temporal aggregates yielded better MAE results and less deviation from the observed distribution than daily comparisons, as expected, because precipitation data are by nature zero-inflated (Bruno *et al*., 2014; Monteiro *et al*., 2016). The monthly aggregates still showed large mean absolute error percentages, although this metric may be affected by the large number of months with scarce rainfall in the study area. Deviations considerably decreased with seasonal and yearly aggregates. Therefore, the applicability of these grids in climate model evaluation at these temporal scales is clear.

Usually, the grid evaluation results were non-uniform throughout the study area. The reasons are well known, namely, non-uniform station density, complex orography, and different precipitation patterns. In this respect, the study area, because of its strong climatic variability and altitude gradients resulting from complex orography, make it suitable for evaluating the influence of these factors. The assessment of several precipitation grids revealed poorer results in areas with complex orography (Hofstra *et al*., 2010; Maraun *et al*., 2012; Hiebl and Frei, 2018). The results shown here do not indicate a general relationship with station elevation. Only areas corresponding to the Dfb (high Pyrenees mountains) climate showed poor results. Nevertheless, Cfb climate areas, also with strong altitude gradients, generated results comparable to low-lying areas. The explanation is based on the fact that convective precipitation is dominant in the latter areas. Therefore, orography and

precipitation patterns have a simultaneous influence on the spatial variability of grid data quality.

Finally, conservative re-gridding from high to low resolution was done, but this did not substantially reduce the quality as compared to grids constructed with a specific resolution. Thus, this solution may be interesting when the grid is unavailable at the desired resolution.

The findings of this study provide valuable information for the establishment of necessary factors when building a grid. Knowledge of grid quality based on station density, spatial resolution, and orographic and climatic characteristics facilitates the evaluation and proper use of the grid.

ACKNOWLEDGEMENTS

Funding came from projects LE240P18 (Consejería de Educación, Junta de Castilla y León) and CGL2016-78702-C2-1-R, PID2019-108470RB-C22, CGL2016-80609-R and PID2019-108470RB-C21 (Ministerio de Economía y Competitividad). The authors acknowledge the Automatic Hydrological Information System (SAIH-CHEBRO) for the rain gauge data.

ORCID

Andrés Merino https://orcid.org/0000-0001-8806-6263

REFERENCES

Avila, F.B., Dong, S., Menang, K.P., Rajczak, J., Renom, M., Donat, M.G. and Alexander, L.V. (2015) Systematic investigation of gridding-related scaling effects on annual statistics of daily temperature and precipitation maxima: a case study for south-East Australia. *Weather and Climate Extremes*, 9, 6–16. https://doi.org/10.1016/j.wace.2015.06.003.

Azorin-Molina, C., Tijm, S., Ebert, E.E., Vicente-Serrano, S.M. and Estrela, M.J. (2014) Sea breeze thunderstorms in the eastern Iberian Peninsula. Neighborhood verification of HIRLAM and HARMONIE precipitation forecasts. *Atmospheric Research*, 139, 101–115. https://doi.org/10.1016/j.atmosres.2014.01.010.

Baez-Villanueva, O.M., Zambrano-Bigiarini, M., Ribbe, L., Nauditt, A., Giraldo-Osorio, J.D. and Thinh, N.X. (2018) Temporal and spatial evaluation of satellite rainfall estimates over different regions in Latin-America. *Atmospheric Research*, 213, 34–50. https://doi.org/10.1016/j.atmosres.2018.05.011.

Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., Schamm, K., Schneider, U. and Ziese, M. (2013) A description of the global land-surface precipitation data products of the Global Precipitation Climatology Centre with sample applications including centennial (trend) analysis from 1901–present. *Earth System Science Data*, 5, 71–99. https://doi.org/10.5194/essd-5-71-2013.

Beguería, S., Vicente-Serrano, S.M., Tomás-Burguera, M. and Maneta, M. (2016) Bias in the variance of gridded data sets leads to misleading conclusions about changes in climate variability. *International Journal of Climatology*, 36, 3413–3422. https://doi.org/10.1002/joc.4561.

Bianchi, E., Villalba, R., Viale, M., Couvreux, F. and Marticorena, R. (2016) New precipitation and temperature grids for northern Patagonia: advances in relation to global climate grids. *Journal of Meteorological Research*, 30(1), 38–52. https://doi.org/10.1007/s13351-015-5058-y.

Bruno, F., Cocchi, D., Greco, F. and Scardovi, E. (2014) Spatial reconstruction of rainfall fields from rain gauge and radar data. *Stochastic Environmental Research and Risk Assessment*, 28, 1235–1245. https://doi.org/10.1007/s00477-013-0812-0.

Cardell, M.F., Amengual, A., Romero, R. and Ramis, C. (2020) Future extremes of temperature and precipitation in Europe derived from a combination of dynamical and statistical approaches. *International Journal of Climatology*, article in press, 40, 4800–4827. https://doi.org/10.1002/joc.6490.

Chen, M., Xie, P. and Shi, W. (2008) CPC unified gauge-based analysis of global daily precipitation. In: Cairns, Australia: Western Pacific Geophysics Meeting.

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

Decuyper, M., Chávez, R.O., Čufar, K., Estay, S.A., Clevers, J.G.P.W., Prislan, P., Gričar, J., Črepinšek, Z., Merela, M., de Luis, M., Notivoli, R.S., del Castillo, E.M., Rozendaal, D.M.A., Bongers, F., Herold, M., Sass-Klaassen, U. (2020). Spatio-temporal assessment of beech growth in relation to climate extremes in Slovenia–an integrated approach using remote sensing and tree-ring data. *Agricultural and Forest Meteorology*, 287, 107925. https://doi.org/10.1016/j.agrformet.2020.107925

Feki, H., Slimani, M. and Cudennec, C. (2017) Geostatistically based optimization of a rainfall monitoring network extension. Case of the climatically-heterogeneous Tunisia. *Hydrology Research*, 48(2), 514–541. https://doi.org/10.2166/nh.2016.256.

Gervais, M., Tremblay, L.B., Gyakum, J.R. and Atallah, E. (2014) Representing extremes in a daily gridded precipitation analysis over the United States: impacts of station density, resolution, and gridding methods. *Journal of Climate*, 27(14), 5201–5218. https://doi.org/10.1175/JCLI-D-13-00319.1.

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

Herrera, S., Gutiérrez, J.M., Ancell, R., Pons, M.R., Frías, M.D. and Fernández, J. (2012) Development and analysis of a 50-year high-resolution daily gridded precipitation dataset over Spain (Spain02). *International Journal of Climatology*, 32, 74–85. https://doi.org/10.1002/joc.2256.

Herrera, S., Kotlarski, S., Soares, P.M.M., Cardoso, R.M. Jaczewski, A., Gutiérrez, J.M. and Maraun, D. (2019) Uncertainty in gridded precipitation products: influence of station density, interpolation method and grid resolution. *International Journal of Climatology*, 39, 3717–3729. https://doi.org/10.1002/joc.5878

Hewitson, B.C. and Crane, R.G. (2005) Gridded area-averaged daily precipitation via conditional interpolation. *Journal of Climate*, 18, 41–57. https://doi.org/10.1175/JCLI3246.1.

Hiebl, J. and Frei, C. (2018) Daily precipitation grids for Austria since 1961—development and evaluation of a spatial dataset for hydroclimatic monitoring and modelling. *Theoretical and Applied Climatology*, 132(1–2), 327–345. https://doi.org/10.1007/s00704-017-2093-x.

Hofstra, N., Haylock, M., New, M., JONES, P. and Frei, C. (2008) Comparison of six methods for the interpolation of daily, European climate data. *Journal of Geophysical Research*, 113 (21), D21110. https://doi.org/10.1029/2008JD010100.

Hofstra, N., New, M. and McSweeney, C. (2010) The influence of interpolation and station network density on the distribution and extreme trends of climate variables in gridded data. *Climate Dynamics*, 35, 841–858. https://doi.org/10.1007/s00382-009-0698-1.

Jones, P.W. (1999) First- and second-order conservative remapping schemes for grids in spherical coordinates. *Monthly Weather Reviews*, 127, 2204–2210. https://doi.org/10.1175/1520-0493 (1999.

Klein Tank, A.M.G., Wijngaard, J.B., Können, G.P., Böhm, R., Demarée, G., Gocheva, A., Mileta, M., Pashiardis, S., Hejkrlik, L., Kern-Hansen, C., Heino, R., Bessemoulin, P., Müller-Westermeier, G., Tzanakou, M., Szalai, S., Pálsdóttir, T., Fitzgerald, D., Rubin, S., Capaldo, M., Maugeri, M., Leitass, A., Bukantis, A., Aberfeld, R., van Engelen, A.F.V., Forland, E., Mietus, M., Coelho, F., Mares, C., Razuvaev, V., Nieplova, E., Cegnar, T., Antonio, L.J., Dahlström, B., Moberg, A., Kirchhofer, W., Ceylan, A., Pachaliuk, O., Alexander, L.V. and Petrovic, P. (2002) Daily dataset of 20th-century surface air temperature and precipitation series for the European climate assessment. *International Journal of Climatology*, 22, 1441–1453. https://doi.org/10.1002/joc.773.

Kling, H., Fuchs, M. and Paulin, M. (2012) Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. *J. Hydro.*, 424, 264–277. https://doi.org/10.1016/j.jhydrol.2012.01.011.

Kysely, J. and Plavcova, E. (2010) A critical remark on the applicability of e-obs european gridded temperature dataset for validating control climate simulations. *Journal of Geophysical Research*, 115, D23118. https://doi.org/10.1029/2010JD014123.

Lievens, H., Tomer, S.K., Al Bitar, A., De Lannoy, G.J., Drusch, M., Dumedah, G., Franssen, H.-J.H., Kerr, Y.H., Martens, B., Pan, M., et al. (2015) SMOS soil moisture assimilation for improved hydrologic simulation in the Murray Darling Basin, Australia. *Remote Sensing Environment*, 168, 146–162. https://doi.org/10.1016/j.rse.2015.06.025.

Maraun, D., Osborn, T.J. and Rust, H.W. (2012) The influence of synoptic airflow on UKdaily precipitation extremes. Part II: regional climate model and E-OBS data validation. *Climate Dynamics*, 39(1–2), 287–301. https://doi.org/10.1007/s00382-011-1176-0.

Monteiro, J.A.F., Strauch, M., Srinivasan, R., Abbaspour, K. and Gücker, B. (2016) Accuracy of grid precipitation data for Brazil: application in river discharge modelling of the Tocantins catchment. *Hydrological Processes*, 30(9), 1419–1430. https://doi.org/10.1002/hyp.10708.

Philip, W.J. (1999) First- and second-order conservative remapping schemes for grids in spherical coordinates. *Monthly Weather Reviews*, 127, 2204–2210. https://doi.org/10.1175/1520-0493 (1999)127<2204:FASOCR>2.0.CO;2.

Rauthe, M., Steiner, H., Riediger, U., Mazurkiewicz, A. and Gratzki, A. (2013) A central European precipitation climatology - part I: generation and validation of a high-resolution gridded daily dataset (HYRAS). *Meteorologische Zeitschrift*, 22(3), 235–256. https://doi.org/10.1127/0941-2948/2013/0436.

Royé, D. and Martin-Vide, J. (2017) Concentration of daily precipitation in the contiguous United States. *Atmospheric Research*, 196, 237–247. https://doi.org/10.1016/j.atmosres.2017.06.011.

Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B. and Ziese, M. (2011) GPCC full data reanalysis version 6.0 at 0.5°: monthly land-surface precipitation from rain-gauges built on GTS-based and historic data. *Global Precipitation Climatology Centre*. https://doi.org/10.5676/DWD_GPCC/FD_M_V6_050.

Serrano-Notivoli, R., Beguería, S., Saz, M. Á., Longares, L. A. and de Luis, M. (2017b) SPREAD: a high-resolution daily gridded precipitation dataset for Spain – an extreme events frequency and intensity overview. *Earth System Science Data*, 9, 721–738. https://doi.org/10.5194/essd-9-721-2017.

Serrano-Notivoli, R., De Luis, M. and Beguería, S. (2017a) An R package for daily precipitation climate series reconstruction. *Environ. Model. Softw.*, 89, 190–195. https://doi.org/10.1016/j.envsoft.2016.11.005.

Serrano-Notivoli, R., De Luis, M., Saz, M.A. and Beguería, S. (2017c) Spatially based reconstruction of daily precipitation instrumental data series. *Climate Research*, 73(3), 167–186. https://doi.org/10.3354/cr01476.

Sungmin, O., Foelsche, U., Kirchengast, G. and Fuchsberger, J. (2016) Validation and correction of rainfall data from theWegenerNet high density network in Southeast Austria. *Journal of Hydrology*, 556, 110–1122. https://doi.org/10.1016/j.jhydrol.2016.11.049.

Tapiador, F.J. (2019) *The Geography of Spain*. Switzerland AG: Springer Nature.

Tramblay, Y., Feki, H., Quintana-Seguí, P. and Guijarro, J.A. (2019) The SAFRAN daily gridded precipitation product in Tunisia (1979–2015). *International Journal of Climatology*, 39(15), 5830–5838. https://doi.org/10.1002/joc.6181.

Wang, C., Tang, G., Han, Z., Guo, X. and Hong, Y. (2018) Global intercomparison and regional evaluation of GPM imerg version-03, version-04 and its latest version-05 precipitation products: similarity, difference and improvements. *Journal of Hydrology*, 564, 342–356. https://doi.org/10.1016/j.jhydrol.2018.06.064.

Yatagai, A., Kamiguchi, K., Arakawa, O., Hamada, A., Yasutomi, N. and Kitoh, A. (2012) APHRODITE constructing a long-term daily gridded precipitation dataset for Asia based on a dense network of rain gauges. *Bulletin of the American Meteorological Society*, 93 (9), 1401–1415. https://doi.org/10.1175/BAMS-D-11-00122.1.

How to cite this article: Merino A, García-Ortega E, Navarro A, Fernández-González S, Tapiador FJ, Sánchez JL. Evaluation of gridded rain-gauge-based precipitation datasets: Impact of station density, spatial resolution, altitude gradient and climate. *Int J Climatol*. 2021;1–17. https://doi.org/10.1002/joc.7003