Climate Dynamics (2022) 58:2825–2842
https://doi.org/10.1007/s00382-021-06034-x

Do CMIP models capture long-term observed annual precipitation trends?

S. M. Vicente-Serrano[1] · R. García-Herrera[2,3] · D. Peña-Angulo[1] · M. Tomas-Burguera[4] · F. Domínguez-Castro[5,6] · I. Noguera[1] · N. Calvo[2] · C. Murphy[7] · R. Nieto[8,9] · L. Gimeno[8,9] · J. M. Gutierrez[10] · C. Azorin-Molina[11] · A. El Kenawy[12,13]

Received: 8 June 2021 / Accepted: 29 October 2021 / Published online: 6 November 2021

Abstract
This study provides a long-term (1891–2014) global assessment of precipitation trends using data from two station-based gridded datasets and climate model outputs evolved through the fifth and sixth phases of the Coupled Model Intercomparison Project (CMIP5 and CMIP6, respectively). Our analysis employs a variety of modeling groups that incorporate low- and high-top level members, with the aim of assessing the possible effects of including a well-resolved stratosphere on the model's ability to reproduce long-term observed annual precipitation trends. Results demonstrate that only a few regions show statistically significant differences in precipitation trends between observations and models. Nevertheless, this pattern is mostly caused by the strong interannual variability of precipitation in most of the world regions. Thus, statistically significant model-observation differences on trends (1891–2014) are found at the zonal mean scale. The different model groups clearly fail to reproduce the spatial patterns of annual precipitation trends and the regions where stronger increases or decreases are recorded. This study also stresses that there are no significant differences between low- and high-top models in capturing observed precipitation trends, indicating that having a well-resolved stratosphere has a low impact on the accuracy of precipitation projections.

Keywords Observed precipitation · Historical simulations · CMIP · Trends · Stratosphere

✉ S. M. Vicente-Serrano
svicen@ipe.csic.es

[1] Instituto Pirenaico de Ecología, Consejo Superior de Investigaciones Científicas (IPE–CSIC), Zaragoza, Spain

[2] Departamento de Ciencias de la Tierra y Astrofísica, Facultad de Ciencias Físicas, Universidad Complutense de Madrid, Madrid, Spain

[3] Instituto de Geociencias (IGEO), Consejo Superior de Investigaciones Científicas–Universidad Complutense de Madrid, Madrid, Spain

[4] Estación Experimental de Aula Dei, Consejo Superior de Investigaciones Científicas (EEAD–CSIC), Zaragoza, Spain

[5] Aragonese Agency for Research and Development Researcher (ARAID), Zaragoza, Spain

[6] Department of Geography, University of Zaragoza, Zaragoza, Spain

[7] Irish Climate Analysis and Research UnitS (ICARUS), Department of Geography, Maynooth University, Maynooth, Ireland

[8] Environmental Physics Laboratory (EPhysLab), CIM-UVigo, Universidade de Vigo, Ourense, Spain

[9] Unidad Asociada CSIC-Universidad de Vigo: Grupo de Física de la Atmosfera y del Océano, Pontevedra, Spain

[10] Instituto de Física de Cantabria, Consejo Superior de Investigaciones Científicas (IFCA-CSIC), Santander, Spain

[11] Centro de Investigaciones Sobre Desertificación, Consejo Superior de Investigaciones Científicas (CIDE–CSIC), Moncada, Valencia, Spain

[12] Department of Geography, Mansoura University, Mansoura 35516, Egypt

[13] Department of Geography, Sultan Qaboos University, Al Khoud, Muscat, Oman

1 Introduction

Precipitation is one of the key climatic variables contributing to the water and energy balance across spatial and temporal scales, ranging from local catchments to global-scale hydrology. Due to their role in accelerating extreme weather events like drought and floods, precipitation anomalies can have significant hydrological, environmental and socioeconomic implications. As such, it is highly relevant to identify possible long term trends in precipitation, as this knowledge can contribute to a better understanding of many ecosystem processes and functions (e.g. land aridity, vegetation greening and browning, crop yield, water resources availability, biodiversity, etc.). In the same context, knowledge about possible future changes in precipitation is essential to inform societal and ecological adaptation. For this purpose, outputs from climate models, with different configurations and scenarios of greenhouse gas (GHG) emissions, have been widely used to assess possible future precipitation changes (e.g. Orlowsky and Seneviratne 2013; Huang et al. 2014; Swain and Hayhoe 2015; Raut et al. 2017; Martin 2018). These models simulate the possible complex feedbacks in the global climate system, including interactions between ocean, land, and atmosphere processes (Räisänen 2007; Dirmeyer et al. 2013). Simulations from these models span a wide variety of climate variables, including precipitation.

Different climate model simulation experiments underpin the reports of the Intergovernmental Panel on Climate Change (IPCC). The third phase of the Coupled Model Intercomparison Project (CMIP3) was developed for the Fourth Assessment Report (AR4) of the IPCC (Meehl et al. 2007). A few years later, simulations of the fifth phase (CMIP5) were delivered in 2011 (Taylor et al. 2012) to coincide with the IPCC's Fifth Assessment Report, followed by those of the sixth phase (CMIP6) (Eyring et al. 2016). Based on these model experiments, several studies have assessed future scenarios of precipitation at global (e.g. Orlowsky and Seneviratne 2013) and regional (e.g. Mariotti et al. 2008; Su et al. 2013) scales. For example, using CMIP5 models under high emission scenarios, different studies have suggested an increase in precipitation at high latitudes of the Northern Hemisphere over the coming century, while general dryness dominates over subtropical regions (e.g. the Mediterranean, southern North America, central America, and South Africa) (e.g. Dufresne et al. 2013; Knutti and Sedláček 2013; Orlowsky and Seneviratne 2013; Cook et al. 2014; Asadi Zarch et al. 2015). The uncertainty associated with these future simulations is commonly assessed using the spread among different emission scenarios, projections, and models (Schaller et al. 2011; Alexander and Arblaster 2017), with a high spread suggesting an increase in the uncertainty of projections (Orlowsky and Seneviratne 2013; Zhao et al. 2015b).

A necessary step to ensure the reliability and performance of climate models is to validate their skill and accuracy in reproducing near-present climate. This procedure is typically implemented by comparing model outputs for a historical period with available ground-based climate observations. In this context, several studies have confirmed the greater skill of climate models in reproducing characteristics (e.g. climatologies, anomalies, and magnitude and spatial patterns of trends) of air temperature over precipitation. (e.g. van Oldenborgh et al. 2013; Kumar et al. 2013; Zhao et al. 2015b; Lee et al. 2019). In fact, climate models show limited skill in capturing long-term observed precipitation variability (e.g. van Oldenborgh et al. 2013; Chen and Frauenfeld 2014).

Several comprehensive studies have attempted to compare long-term precipitation trends from observations with model simulations, mainly CMIP5 models (e.g. Kumar et al. 2013; Nasrollahi et al. 2015; Zhao et al. 2015a; Knutson and Zeng 2018). To date, a global assessment of long-term precipitation trends using CMIP6 simulations is still lacking, albeit with the availability of some preliminary regional evaluations (e.g. Rivera and Arnould 2020; Xin et al. 2020; Peña-Angulo et al. 2020). In addition, very few studies have assessed the possible role of model vertical structure in determining their capacity to reproduce precipitation trends. This includes, for example, the possible effects of the vertical resolution and top height of the models. While some models (so called low-top models) are characterized by their coarse vertical resolution in the stratosphere and associated limitations in explicitly modeling climate processes in this layer (possibly affecting stratosphere-troposphere coupling), other models (so called high-top models) include a well resolved stratosphere (Charlton-Perez et al. 2013; Hurwitz et al. 2014). Numerous studies have evaluated the effectiveness of high-top models to capture the observed signal in precipitation and the different atmospheric mechanisms driving multidecadal variability (e.g. Cagnazzo and Manzini 2009; Lee and Black 2014; Wei et al. 2018a; Haase et al. 2018). Others indicate that models without a well-resolved stratosphere show limitations in identifying atmospheric circulation mechanisms (Deser et al. 2012; Osprey et al. 2013). According to Scaife et al. (2012), future changes in stratospheric circulation are likely to double the increase in winter precipitation over Western and Central Europe, compared to projections from low-top models. Similarly, Marsh et al. (2013) compared simulations from low- and high-top versions of the NCAR Community Earth System Model (CESM), suggesting systematic differences between these two types of models in reproducing some variables like wind and precipitation, especially in the extratropics. These earlier findings emphasize that stratosphere–troposphere

coupling could be a major source of uncertainty in climate model projections of precipitation.

In the pursuit of this background, the main aims of this study are to: (i) explore the extent of agreement in both the magnitude and spatial distribution of precipitation trends from observations and climate simulations, (ii) compare precipitation signals in CMIP5 and CMIP6 using multi-model groups, and (iii) assess improvements in reproducing observed precipitation trends in climate models with and without a well-resolved stratosphere (i.e. low-top *vs.* high-top models).

2 Data

This study employs observed annual precipitation data from the Global Precipitation Climatology Centre (GPCC) dataset (https://www.dwd.de/EN/ourservices/gpcc/gpcc.html) at a 2.5° grid resolution. The GPCC dataset is maintained by the Deutscher Wetterdienst, German Weather Service under the auspices of the World Meteorological Organization (WMO). Due to the uneven distribution of gauged data from many national meteorological networks in the early decades of the twentieth century, our analysis for these earlier decades (i.e. from 1891) is restricted to regions with higher density of observations (e.g. North America, Europe, Middle East, North and South Africa, Australia, India, the La Plata basin in South America, and some sparse regions in Central Asia, Central America and East Asia) (See colored regions in the first column of Fig. 1) and, accordingly, more reliable outputs from interpolation algorithms. Our analysis is limited to the year 2014, which corresponds to the end of the historical CMIP6 model simulations.

We use the available historical monthly simulations with observed radiative forcing from a set of CMIP5 and CMIP6 models spanning the period 1891–2014. Because each model has a different number of ensemble members (ranging from one to ten, with the most common situation being one single member per model), a single member per model is used to maintain model homogeneity. A list of CMIP models employed is available in Table S1. Given that historical simulations of the CMIP5 are only available up to 2005, simulations from the RCP8.5 scenario are used for the period 2006–2014. This is justified as the evolution of CO_2 observations from 2006 to 2014 resembles those used in the RCP8.5 experiment more than other scenarios (Schwalm et al. 2020). The spatial resolutions of the CMIP models vary considerably, with the CMIP6 models having improved spatial resolution over the CMIP5 models (Table S1). High-resolution climate models were shown to

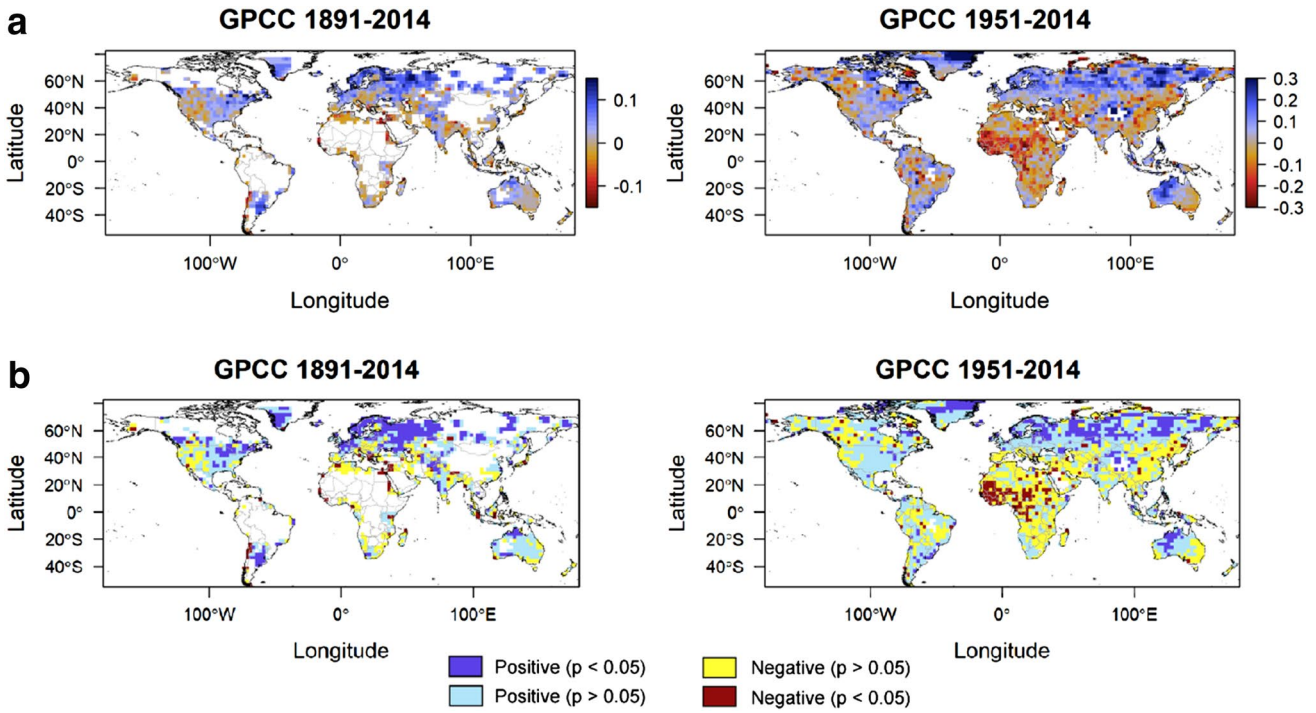


Fig.1 **a** Spatial distribution of the magnitude of change (in z-units/decade) of annual precipitation using GPCC observations and **b** Spatial distribution of their statistical significance (using a p-level of 0.05 for statistical significance, the dark blue and dark red colours show areas with statistically significant positive and negative trends, respectively, while the light blue and yellow colours denote regions with statistically non-significant increases and decreases, respectively). Results are presented for the periods 1891–2014 and 1951–2014. Due to the scarcity of GPCC data, white areas were not used for comparison

be capable of reproducing precipitation with the same skill as regional climate models (Demory et al. 2020). Nonetheless, all models were resampled to a common grid interval of 2.5° using a bilinear interpolation approach to match with the selected resolution of the GPCC data. The aim was to have a comparable spatial resolution to CMIP5 models and a common spatial resolution for all models, allowing for direct comparisons between all models.

As the gridded precipitation data, climate model data are aggregated into annual series. To assess the ability of the well-resolved stratosphere models to represent observed long-term precipitation trends, models are divided into two broad categories: high-top and low-top models. To define a model as a high-top or low-top model, the 1 hPa lid height was used following Charlton-Perez et al. (2013). Specifically, we assess whether there are significant differences in reproducing long-term observed precipitation trends between the two types of models (i.e. low-top vs. high-top models). We also derive trend results for the complete group of models (i.e. including both types of models).

3 Methods

In literature, it is well-established that the prediction and simulation of precipitation is a challenging task, with large uncertainties from different sources. On one hand, total precipitation can vary largely among models (Su et al. 2013). On the other hand, overestimation or underestimation biases affect large areas of the world (Brands et al. 2013; Mehran et al. 2014; Aloysius et al. 2016). In addition, precipitation is characterized by its high spatial and temporal variability. As we are more interested in trends than in the total volumes, and to account for these issues, we decided to standardize the precipitation series to have an average equal to zero and a standard deviation equal to one. This procedure makes precipitation series comparable over space and time, allowing consideration of the strong spatial and temporal variations of precipitation worldwide. Importantly, based on these standardized precipitation series, a comparison between observations and model outputs was feasible, both spatially and temporally. Following Knutson and Zeng (2018), precipitation series were standardized by means of the widely used procedure of calculating the Standardized Precipitation Index (SPI) (McKee et al. 1993) through fitting the data to a gamma distribution. This procedure was applied to the series for both observations and simulations of the different individuals of each group of models.

To analyze changes in annual precipitation series from both observations and model outputs, we employed a modified version of the non-parametric Mann–Kendall statistic (Hamed and Rao, 1998; Yue and Wang, 2004). This statistic accounts for the possible effect of autocorrelation on trend detection by returning the corrected p values after accounting for temporal pseudoreplication. To facilitate direct comparison between different regions worldwide, we presented the trend results using four categories of trends: positive and significant ($p < 0.05$), positive and non-significant ($p > 0.05$), negative and non-significant ($p > 0.05$) and negative and significant ($p < 0.05$).

It is quite difficult to establish a robust comparison between trends in series of observations and a group of model simulations. Some studies have made comparison between observations and model simulations by means of multimodel averages (e.g. Orlowsky and Seneviratne 2013; Sillmann et al. 2013; Kumar et al. 2013; Dai and Zhao 2017). However, in comparison to observed series, this approach dramatically reduces internal variability and the standard deviation of the resulting average series from a group of models, making it difficult to establish a comparison with observed series characterized by strong natural variability. For this reason, we calculated the trend for each independent model and computed an average of these magnitudes for all models. Although this approach reduces the range of precipitation changes derived from the models, as compared to those of observations, it allows for a comparison of the spatial patterns of the magnitudes of change between observations and climate models. The percentage of models showing the same trend category, as compared to a multi-model average, was also computed. Herein, four trend categories were considered, including statistically positive ($p < 0.05$), statistically negative ($p < 0.05$), statistically non-significant positive ($p > 0.05$), and statistically non-significant negative ($p > 0.05$) trends. To assess the magnitude of change in precipitation (z-unit/decade), we used the least squares regression model, in which time was considered as an independent variable, while precipitation represented the dependent variable. The slope of the regression indicated the amount of change, with a higher slope suggesting greater changes and vice versa. As the identification of hydroclimatic trends is strongly impacted by the selection of the base period for analysis (Hannaford et al. 2013; Vicente-Serrano et al. 2020; Peña-Angulo et al. 2020), trends were assessed for two different periods: 1891–2014, and 1951–2014.

In addition, we considered the series of observations as a member of each of the group of models. This method is used to determine the position of the trends in the observations in relation to the trends in the members (the independent model runs). This enables a statistical analysis of the possible differences between trends in observations and model outputs. In this context, it is acceptable to infer that changes in observed precipitation are significantly (p 0.05 two-tails) different from those of the group of models when the magnitude of change in the observations is above the 95th or below the 5th percentiles of changes suggested by the different members of a large-enough group of models. A similar

procedure has been adopted in earlier studies (e.g. Nasrollahi et al. 2015; Knutson and Zeng 2018).

Also, we compared the magnitude of change in the observations with the different independent model simulations by looking at the position (percentile) of the magnitude of trend in the observations, as compared to the entire range of the models. This approach allows determining if the models overestimate or underestimate precipitation patterns. The significance of the differences between the regression slopes of the observations and the independent models and also between the observations and the average group of models was assessed by means of a statistical test for the equality of regression at a confidence interval of 95% ($p < 0.05$)(Paternoster et al. 1998). Following this approach, we computed the percentage of independent models that showed significant differences in the magnitude of precipitation changes, compared to observed precipitation. Finally, to remove the noise at the spatial level, we compared trends for different latitude bands. To do so, we compared the observed trend with the trend of each group of the CMIP5 and CMIP6 models. We utilized a two-sample t-test, at the 95% significance level, to compare the precipitation trends between observations and models in the various latitudinal bands and explore whether there are significant differences between them.

4 Results

Figure 1 illustrates the spatial distribution of annual precipitation trends calculated from GPCC precipitation and their significance. Annual precipitation (GPCC) showed negative and significant trends in small areas of the Near East, West Africa, North Africa, and some regions of Chile from 1891 to 2014. On the other hand, Northern Europe, parts of northeastern North America and the La Plata basin in South America exhibited the most significant increase. Large areas of Central, Northern and Eastern Europe exhibited statistically significant increase in annual precipitation between 1891 and 2014. We also noted the dominance of statistically significant increasing trends in northwestern North America, North Australia, and Greenland. The dominant increase in North Eurasia was identified clearly from 1951 to 2014, while the main negative and significant trends were recorded for this period in Central and West Africa.

For the long term (1891–2014), the average change from the individual CMIP6 (Fig. 2) and CMIP5 (Figure S1) models depicts less spatial consistency with observations. In particular, simulated precipitation shows increases in North Eurasia, Northeast North America, and the La Plata basin, which are qualitatively consistent with observations. Rather, the majority of the world regions show more heterogeneous patterns with less agreement between models and observations. Accordingly, a non-significant statistical relationship was noted between patterns of change in observations and those of the various individual models (Fig. 3 and Figure S2). For 1891–2014, the Mediterranean region, east China, and the Philippines showed the largest decrease in the average change in the different CMIP6 models. This decrease is more pronounced in low-top models than in high-top models. This decrease was also evident for various CMIP5 groups of models, with precipitation decreasing in Central America, southern Africa, Bangladesh, and Indonesia. Given the internal climate variability and the inconsistent behavior of observations and models, the magnitude of trends between observations and models is expected to differ. A spatial comparison clearly demonstrates that the magnitude of change in annual precipitation from models tends to reinforce precipitation decrease in some regions of the world, particularly in subtropical areas in CMIP5's and in low-top models (Figs. 2 and S1). Between 1951 and 2014, there is noticeable spatial divergence in precipitation trends between model averages and observations, especially in South America, West North America, Africa, the Mediterranean, Australia, amongst other regions. Such divergence is expected to be larger in the trend analysis for shorter periods since the relative contribution of internal variability to the overall trend is typically larger than it is for longer trends (e.g., 1891–2014). The spatial relationship between model averages and observations is statistically non-significant for both CMIP5 and CMIP6 experiments, as well as low- and high-top models.

Thus, the average differences between the individual models and the observations seems to be substantial in several world regions, with a clear underestimation of long-term (1891–2014) observed precipitation trends, particularly in North and East Europe and, central and West Africa, most of Russia, South and North America. When analyzing trends for different latitudinal bands, the spatial differences between annual precipitation trends in observations and model simulations are clearly identified (Figs. 4 and S3–S5). For the period 1891 to 2014, the different groups of CMIP6 models show that models underestimate precipitation increases at high latitudes in the Northern Hemisphere (above 55° N). We found statistically significant differences between the average trends in observations and the individual model simulations for both low- and high-top models. A similar picture can be seen for the northern latitudes between 40° and 55°. The CMIP6 models perform better at mid-latitudes and in tropical and subtropical areas, with no statistically significant differences between trends in observations and models, regardless of using low- or high-top models (not shown). The CMIP5 models show a similar pattern, though they show statistically significant differences between observations and models in more latitudinal bands than the CMIP6 models (Figure S3). This pattern varies between latitudinal bands for the period 1951–2014 but there are no relevant

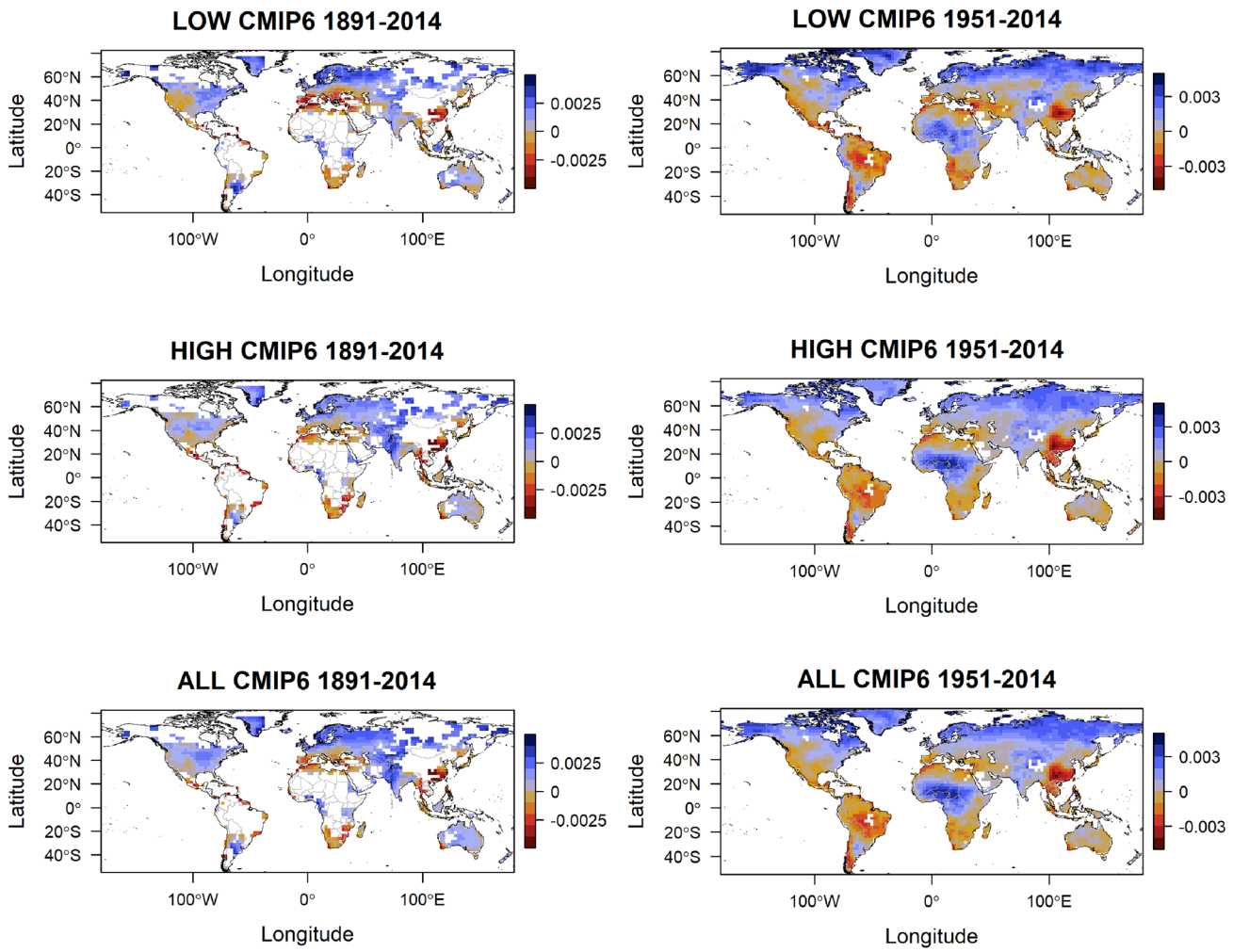


Fig. 2 Spatial distribution of the average magnitude of change (in z-units/decade) in annual precipitation using different CMIP6 models, considering low-top, high-top, and all models. Changes were analyzed for the periods 1891–2014 and 1951–2014. Due to the scarcity of GPCC data, white areas were not used for comparison

differences with the behavior of the CMIP6 models. No matter the latitudinal band, there is a domain of statistically significant differences in average trends between observations and models.

All these results contrast with the few significant differences in annual precipitation trends between models and observations (Figs. 5 and S6). This is simply because there is a dominance of non-significant differences between the trends in the individual models of each group of models and the series of observations (Figure S7 and Figure S8). Thus, for the period 1891–2014, with the exception of a few areas in Eastern Europe, Egypt and La Plata basin, the rest of the regions show a low percentage of models showing statistically significant differences with the trend of observations. With the exception of Central and West Africa, there are few areas where a large percentage of models (> 80%) show statistically significant differences in trends between 1951 and 2014 when compared to observations. The differences in the spatial patterns of trends and the substantial differences in the average magnitude of the changes between the individual models and the observations, suggest that the models' ability to reproduce trend in observations is low. Nevertheless, the small number of statistically significant differences in annual precipitation trends between models and observations seems to suggest the opposite. The problem is that precipitation is characterized by strong interannual variability, making it difficult to obtain statistically significant differences between two series characterized by large interannual variability. Nevertheless, although differences are mostly non-significant, in most of the world's regions, the annual

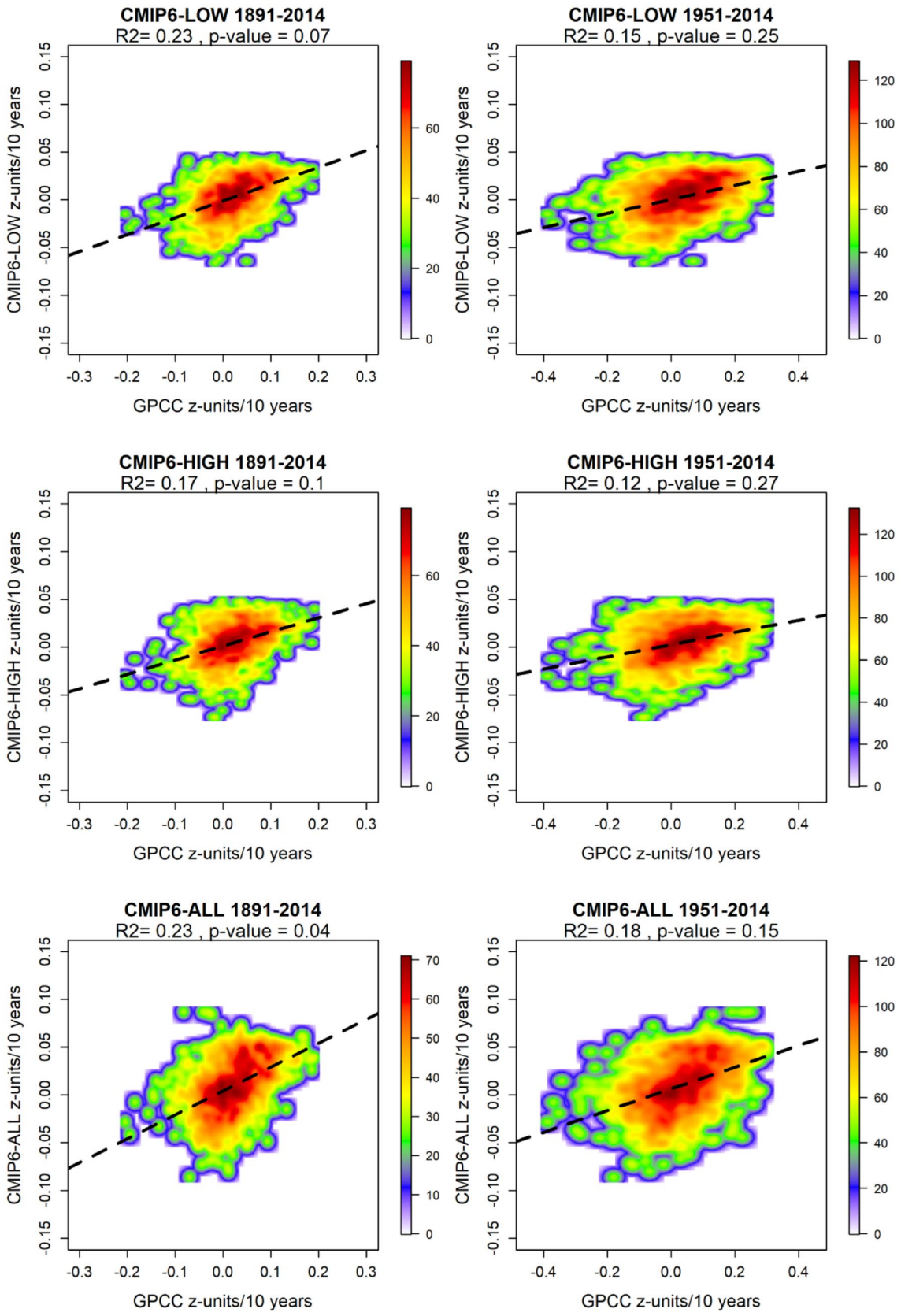
precipitation trend, as recorded in the majority of models, differs from the observed trends. Figure 6 and Figure S9 show the spatial distribution of the percentile of trends in annual precipitation observations in comparison to the trends in the independent models of each CMIP6 and CMIP5 group of models, respectively. The first remarkable feature is the small differences between the spatial patterns in the groups of low- and high-top models and between CMIP5 and CMIP6 models, which suggest similar skill among the different groups of models. For the period 1891–2014, in the majority of the world regions with available data, the models tend to underestimate or overestimate the magnitude of change of annual precipitation (Tables 1 and S1). This would affect both regions with increasing or decreasing annual precipitation in observations, and it means that, despite the fact that there are few statistical differences between trends in individual models and trends in observations, the models are primarily overestimating decreasing trends in annual precipitation in southwestern Europe, southern Africa, and western Australia, and underestimating increasing annual precipitation trends in western North America, the la Plata basin, western Europe, Scandinavia, and western Russia. For the period 1951–2014, most of the models overestimate precipitation increase in the Sahel, central and western Africa, the middle East and eastern Asia, and they underestimate precipitation increase in large areas of north Eurasia, France, West Australia and large areas of North and South America.

The misrepresentation of the annual precipitation trends in different groups of models is also identified, by means of the most common trend significance category (Fig. 7 and Figure S10). In comparison to the spatial patterns illustrated in Fig. 1B, where large areas of northern and western Europe, northern Eurasia, the La Plata basin, western North America, central Asia, Australia, etc., showed positive and statistically significant trends, the different CMIP5 and CMIP6 groups of models show a dominant pattern characterized by non-significant trends in these areas. Thus the dominant pattern of the models would clearly underestimate regions with significant annual precipitation trends. In comparison with the CMIP5 models, the CMIP6 models tend to reduce the areas with dominance of significant trends. Nonetheless, in some areas where observations do not show statistically significant changes, CMIP5 models show dominant negative significant trends (e.g. Bangla Desh, Indonesia, the Iberian Peninsula). The same dominant behavior is recorded for the most recent period (1951–2014). Thus, it is paradoxical that the regions where the models tend to record dominance of statistically and significant positive trends agree with areas in which observations show a dominant negative trend such as in the Sahel. Moreover, the large areas with positive and significant trends in North Eurasia, West China and Northwest Australia are characterized by the dominance of non-significant trends in both CMIP5 and CMIP6 model experiments. It is also a paradox that only the low-top CMIP5 models record the dominance of positive and significant trends in northern Eurasia, while the rest of the areas where low-top CMIP5 models show dominance of significant trends (negative in North South America and positive in West Africa and Central Asia) do not correspond to significant trends in observations. In general, the sign and significance of annual precipitation trends in observations are reproduced by few models in each group (Figs. 8 and S11). Few areas in the group of CMIP6 models have more than 60% of the models with the same sign and significance of observations. CMIP5 models tend to show the same pattern, although the group that includes both low- and high-top models has higher percentages in eastern North America and Western Europe for the period 1951–2014. Thus, there is dominance of the world regions, where only few models agree on the trend's sign and significance, compared to only a few areas where most models agree on the trend in observations (Figs. 9 and S12).

5 Discussion

This study provides a global evaluation of long-term annual precipitation trends from observations, as compared to model-based simulations from the CMIP5 and CMIP6 experiments. Specifically, this assessment considered data spanning the periods 1891–2014 and 1951–2014. Our analysis accounts for two different types of models depending on their representation of the stratosphere (i.e. low-top and high-top), with the aim of determining whether a good representation of this atmospheric layer may improve the ability of CMIP models to reproduce observed precipitation changes. An assessment of this possible influence is crucial for evaluating the capacity of climate models to determine long-term precipitation trends under future precipitation projections.

It should be noted that making a direct comparison between observed and modeled precipitation trends has some limitations. In this study, we included observations as an additional member in the different group of models: an approach that was adopted in earlier studies (e.g. Knutson and Zeng 2018; Peña-Angulo et al. 2020). Nevertheless, due to the large range of variability introduced in each group of models, this method can introduce some uncertainties. While there are noticeable differences in the spatial patterns of trends in observations and the dominant patterns of the different group of models (as represented by the average of the trends), no statistically significant differences exist


CMIP6-LOW 1891-2014
R2= 0.23 , p-value = 0.07
CMIP6-LOW 1951-2014
R2= 0.15 , p-value = 0.25
CMIP6-HIGH 1891-2014
R2= 0.17 , p-value = 0.1
CMIP6-HIGH 1951-2014
R2= 0.12 , p-value = 0.27
CMIP6-ALL 1891-2014
R2= 0.23 , p-value = 0.04
CMIP6-ALL 1951-2014
R2= 0.18 , p-value = 0.15
CMIP6-LOW z-units/10 years
CMIP6-HIGH z-units/10 years
CMIP6-ALL z-units/10 years
GPCC z-units/10 years

◂**Fig. 3** Scatterplots showing the spatial relationship between the magnitude of annual precipitation change in GPCC observations and the average magnitude of change derived from each CMIP6 individual models. Colours represent the density of points with red colours showing the highest densities (legends show the frequency of grid points). The significance of the correlation was obtained using a bootstrap method. A total of 1000 random samples of 30 data points each were extracted, from which correlations and p values were obtained. The final significance was assessed by means of the average of the p values

between trends in observations and model members. Exceptionally, in a few cases, the "observation member" might be placed outside the range of trends suggested by the majority of models (Nasrollahi et al. 2015).

Nevertheless, this result cannot support that models provide a good assessment of trends in observations. Although there are very few cases in which there are statistically significant differences in the magnitude of trends between models and observations, our findings suggest that, in large regions of the world, the majority of models tended to either overestimate or underestimate the trend magnitude found in observations. This is evident, given that the annual precipitation trend in observations corresponds mostly to very low or very high percentiles in comparison to the trend magnitudes suggested by the different members of the group. In this context, recalling that simulated changes in precipitation rarely differ significantly from observations, due to the low signal-to-noise ratio presented in local precipitation changes, the simple assessment of statistical significant differences between observations and models is a poor approach (Bhend and Whetton 2013). This effect was demonstrated in regions with high interannual precipitation variability, such as the Western Mediterranean, where earlier studies showed that both CMIP3 and CMIP5 models simulate more drying than observations, albeit with the absence of statistically significant differences in the magnitude of these trends between observations and individual ensemble members (Peña-Angulo et al. 2020). For this reason, although there is a prevalence of statistically non-significant differences in the trend magnitude of precipitation between observations and model simulations, it cannot be inferred that models are capable of reproducing observed trends. Indeed, our study found that model simulations generally show poor skill in replicating trends in observations and that annual observed precipitation trends in many parts of the world is generally overestimated or underestimated by the climate models. In addition, the patterns of observed and modeled precipitation trends were less consistent over space. Overall, our findings support previous studies that demonstrated low agreement between the magnitude of observed and modeled precipitation changes (e.g. Asadieh and Krakauer 2015; Bishop et al. 2018), as well as inconsistent spatial patterns of these changes (e.g. van Oldenborgh et al. 2013; Kumar et al. 2013; Sheffield et al. 2013; Zebaze et al. 2019).

Moreover, we find that the new CMIP6 models do not substantially improve the sign and statistical significance of the magnitude and spatial patterns of trends, compared to CMIP5 models. In some regions, like the West Sahel, the models simulate poorly the observed long-term precipitation trends. The dominant decrease in observed precipitation from 1951 to 2014 contrasts with the models' overall positive trend (both high- and low-top models in CMIP5 and CMIP6), which is characterized by a strong increasing trend. We also detected a long-term precipitation trend in areas of the high latitudes of the Northern Hemisphere that is consistent with anthropogenic influences, as indicated by some previous studies (e.g. Bhend and von Storch 2008; Wan et al. 2015). Nevertheless, we must emphasize that, even in these areas, there are inconsistencies with model simulations, mostly for the long term, with a general underestimation of the annual precipitation trend by the different groups of models. Also, there are substantial differences between the long-term evolution of observations and model simulations in the mid-latitude subtropical regions (i.e. southern North America, Central America, northern South America, East Asia and the Mediterranean). Models in these regions showed stronger drying, which contradicts long-term observations. Some exceptions (i.e. drying suggested by models coincide with drying from observations) were found in a few regions, mainly close to the Near East (Israel and Syria), where changes have been attributed to anthropogenic forcing (Knutson and Zeng 2018). The expansion of Hadley cells in the Southern and Northern Hemispheres could explain the significant drying of these regions in models (Staten et al. 2018). Nevertheless, the CMIP5 models show limitations in reproducing robustly subtropical expansion (Davis and Birner 2016), and they appear to amplify the possible effects of long-term anthropogenic forcing relative to natural variability that characterizes Hadley cells movement (Bronnimann et al. 2015; Grise et al. 2018).

In general, the existing differences between trends in observed precipitation and model simulations for different long- and short-term periods can probably be linked to a misrepresentation of the atmospheric processes driving decadal variability in model simulations (Sheffield et al. 2013; Chen and Frauenfeld 2014). For example, Gu et al. (2015) and Wei et al. (2018b) have shown that most models are inaccurate in simulating the impacts of the Pacific Decadal Oscillation (PDO) on the western Pacific subtropical high and the East Asian summer monsoon. Models also show

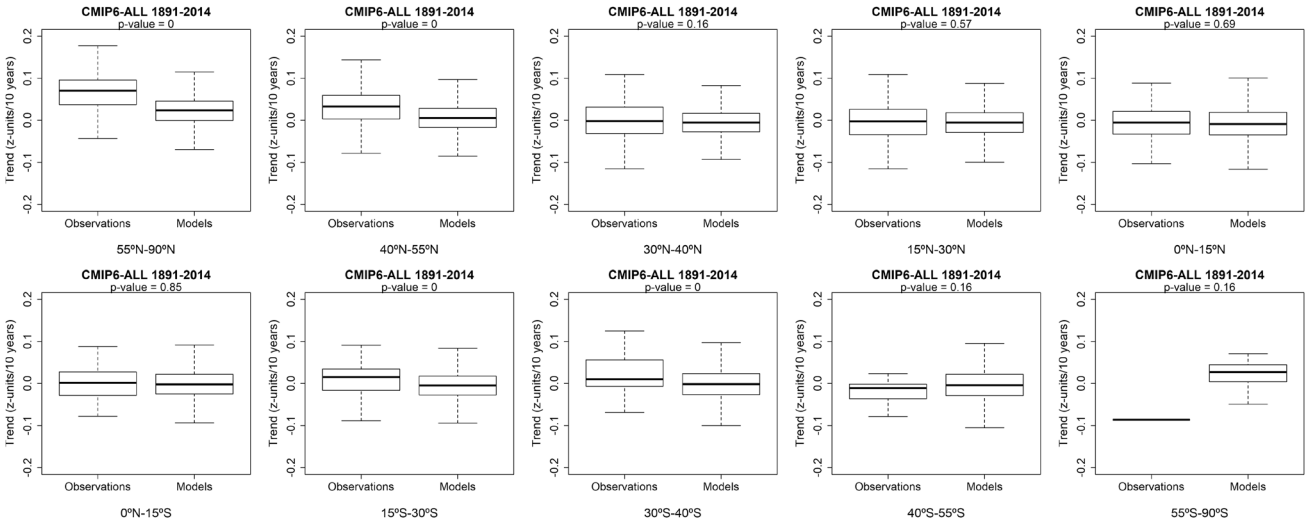


Fig. 4 Distribution of the magnitude of the annual precipitation trends in observations and the individual CMIP6 models from 1891 to 2014 distributed at different latitudinal bands. The significance of statistical differences between trends in observations and models is assessed by means of the T-test ($p < 0.05$)

limitations in replicating changes in the Indo-Pacific sea surface temperature (SST) and the Atlantic SST meridional gradients, which are key systems impacting precipitation variability in the Sahel (Biasutti et al. 2008). This could explain the discrepancy between observed and modelled precipitation trends from 1951 to 2014. The model representation of El Niño Southern Oscillation (ENSO) also shows clear differences among simulations. As such, precipitation biases in the CMIP5 model simulations over the western Pacific, the Indian Ocean, and the equatorial Pacific could be linked to the ENSO-related SST biases in the models (Dai and Arkin 2017). Moreover, the position of the Northern Annular Mode (NAM) and the North Atlantic Oscillation (NAO) is not well represented in precipitation simulation models, as reported by previous studies (Bladé et al. 2012; Kelley et al. 2012). This would reinforce, for example, the disparity between observed and modeled precipitation trends in Southern Europe. In the same context, the CMIP5 models do not adequately represent the Southern Annular Mode (SAM), limiting their ability to reproduce precipitation trends in semiarid mid-latitude regions of the Southern Hemisphere (Purich et al. 2013). Sheffield et al. (2013) showed a wide range of model performance in reproducing the displacement of the tropical synoptic-scale disturbances in North America, making tropical cyclones unpredictable. In addition, they showed a large spread in the model's ability to replicate ENSO impacts on the North American climate. In the Tibetan plateau, Duan et al. (2013) and Salunke et al. (2019) showed certain weaknesses of CMIP5 models for the reproduction of summer monsoon.

According to Deser et al. (2012), the internal atmospheric variability associated with the annular modes is a major source of uncertainty in model simulations in the middle and high latitudes. It could be reasonable to consider that high-top models could better represent trends of precipitation at these latitudes. This is simply because this mechanism has been a primary source to explain the interannual variability of precipitation and its long term trends over large regions (Vicente-Serrano and López-Moreno 2008; Bladé et al. 2012). Nevertheless, in this study, there were few differences in the skill of both CMIP5 and CMIP6 high and low-top models in reproducing long-term observed precipitation trends. Several studies have indicated that accurate representation of the stratosphere is required to properly replicate the various atmospheric processes in the troposphere (Cagnazzo and Manzini 2009; Wei et al. 2018a; Haase et al. 2018). Low-top models do not explicitly reproduce stratospheric dynamical variability, which may have consequences for the identification of tropospheric impacts of the SAM and NAM, which have a major influence on precipitation variability at high latitudes and in subtropical regions. Using the Hadley Center Global Environmental Model version 2 (HCGEM2), Osprey et al. (2013) suggested that high-top models could better replicate the Quasi-Biennial Oscillation (QBO), indicating that a better representation of stratospheric variability in tropical and extratropical stratospheric modes is required. Nevertheless, based on high-top model simulations using the Met Office Unified Model from

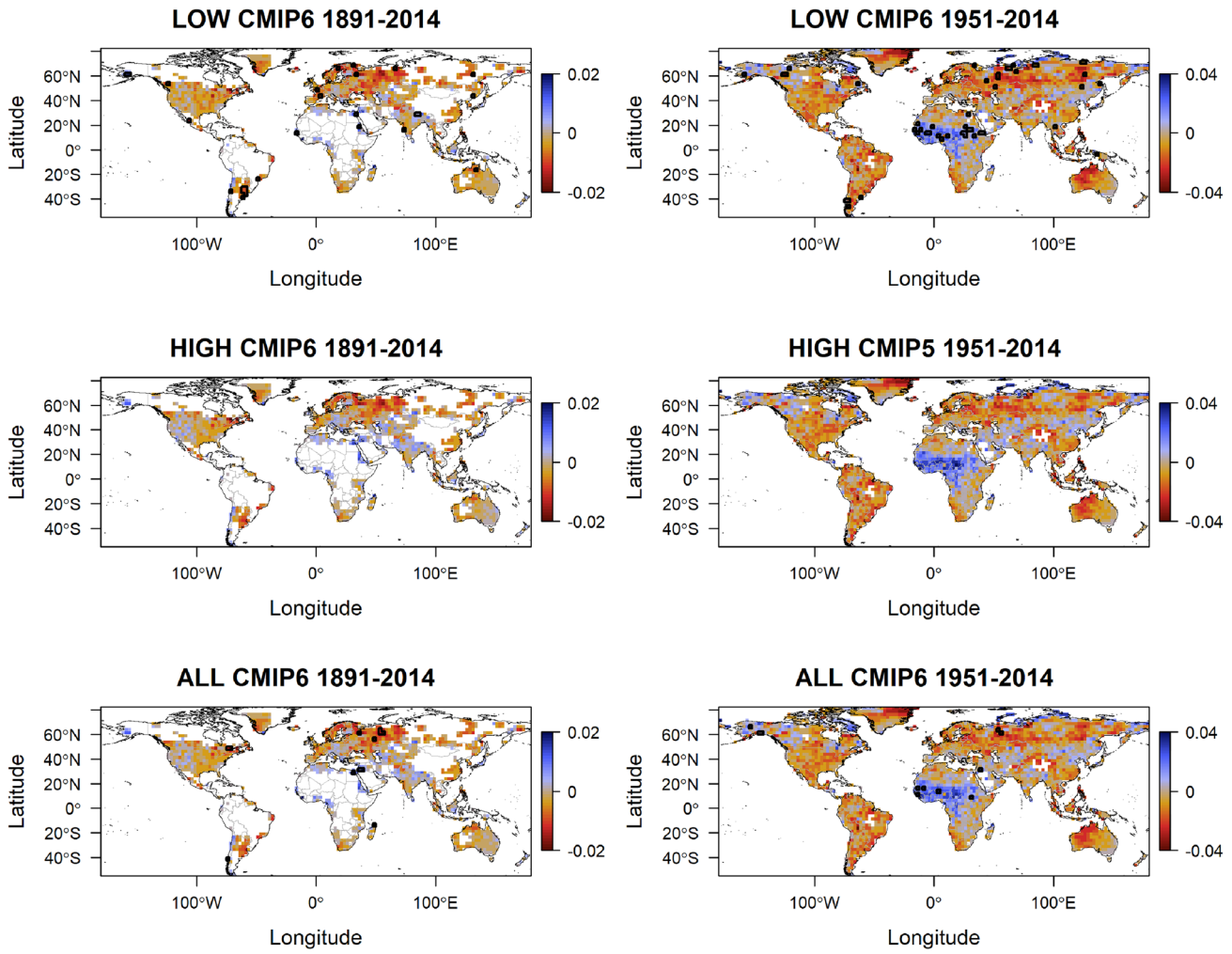


Fig. 5 Spatial distribution of the average differences between the magnitude of change in annual precipitation of the individual models of each CMIP6 group of models and the magnitude of change in the GPCC observations. Areas with statistically significant differences between observations and model groups are delineated by black lines (90% of the models)

CMIP5, Hardiman et al. (2012) showed that high-top model experiments enhanced the impact of atmospheric teleconnections on surface climate, particularly the response to ENSO. However, surprisingly, the addition of a well-resolved stratosphere had no effect on the annular mode index variability. This finding concurs with the study of Lee and Black (2014), which employed CMIP5 models. According to Scaife et al. (2012) changes in the stratospheric circulation induce a further shift in the tropospheric circulation. This shift in circulation could alter the Atlantic storm track and modify precipitation projections over Western and Central Europe. Our findings indicate minor differences between high- and low-top model groups, demonstrating that there is little improvement in the ability of the models to reproduce precipitation trends when considering a well-resolved stratosphere.

6 Conclusions

This study provides a comprehensive comparison of long-term precipitation trends using observations and a set of models from two CMIP experiments (CMIP5 and CMIP6). Making such a comparison is challenging in the sense that precipitation is characterized by high interannual and multidecadal variability. This study demonstrates that –in large regions worldwide- there are no statistically significant differences between observed and modeled

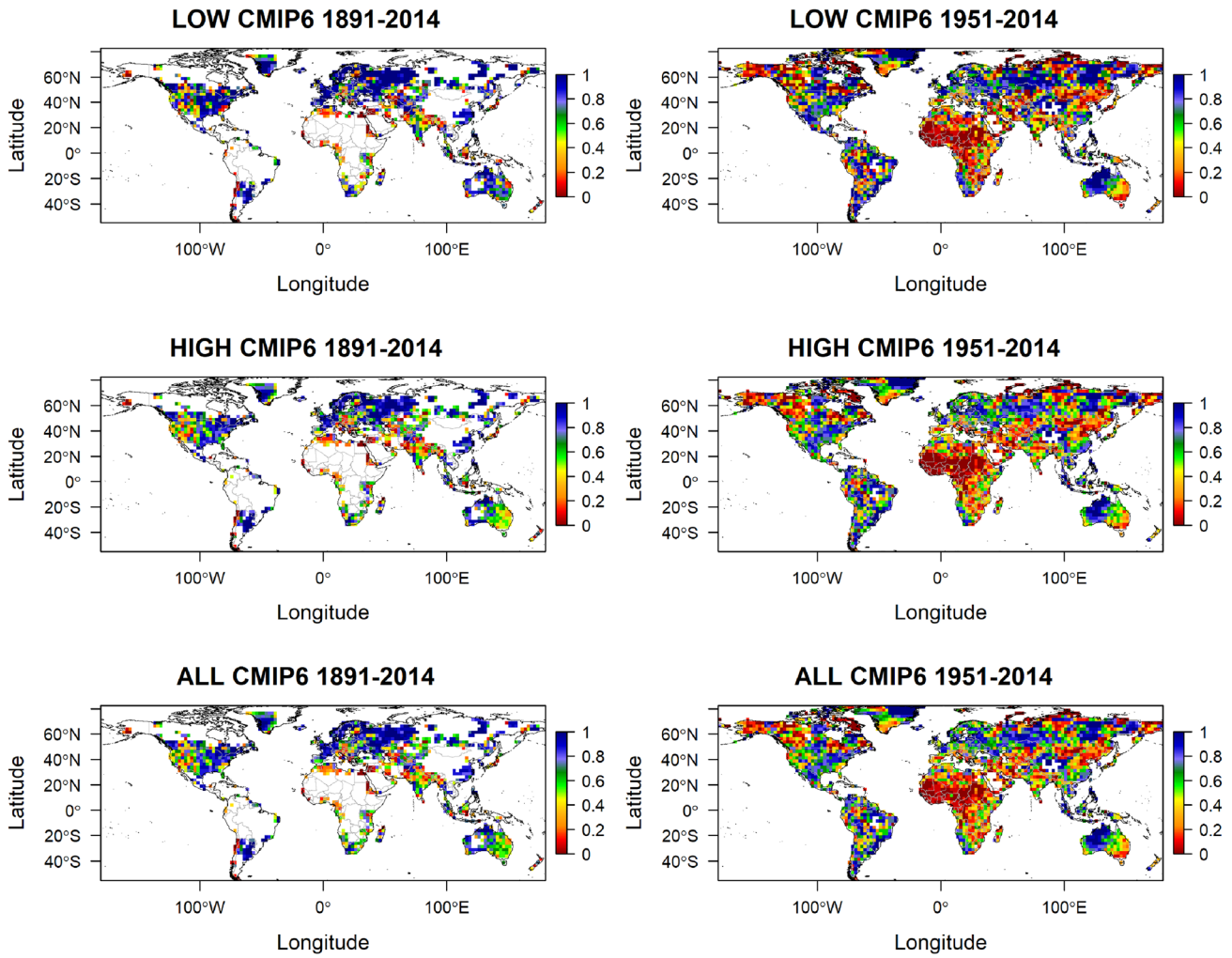


Fig. 6 Percentile of the magnitude of change in annual observations in comparison to the CMIP6 individual model simulations of each group. Percentiles were computed including observations as a member in all groups

precipitation trends. However, this does not imply that different model experiments can reproduce the observed trends. Rather, this finding is associated with the large interannual variability of precipitation, suggesting that the total variance explained by the dominant trend is low. This is evident for long-term (1891–2014) and short term (1951–2014) analysis.

Moreover, there is a statistically significant under-simulation of zonal mean observed increasing precipitation trends (1891–2014) by CMIP6 models over land regions with adequate data coverage (at least poleward of about 40º N in the northern hemisphere, and from 15º S to 40º S in the southern hemisphere). This new key finding generally supports previous findings of under-estimated increasing trends in similar regions using CMIP5 models.

In general, we found that the new experiments in CMIP6 do not significantly improve the representation of the magnitude and spatial patterns of global precipitation trends compared to CMIP5 models. Rather, we found that

Table 1 Percentage of the regions showing that the trend in observations is above the 0.8 percentile or below the 0.2 percentile of the trend of all models of the CMIP6 groups

	1891–2014		1951–2014	
	< 0.2	> 0.8	< 0.2	> 0.8
Low	13.3	43.03	27.9	26.5
High	16.3	42	32.2	26.8
All	15.2	42.7	30.7	27.0

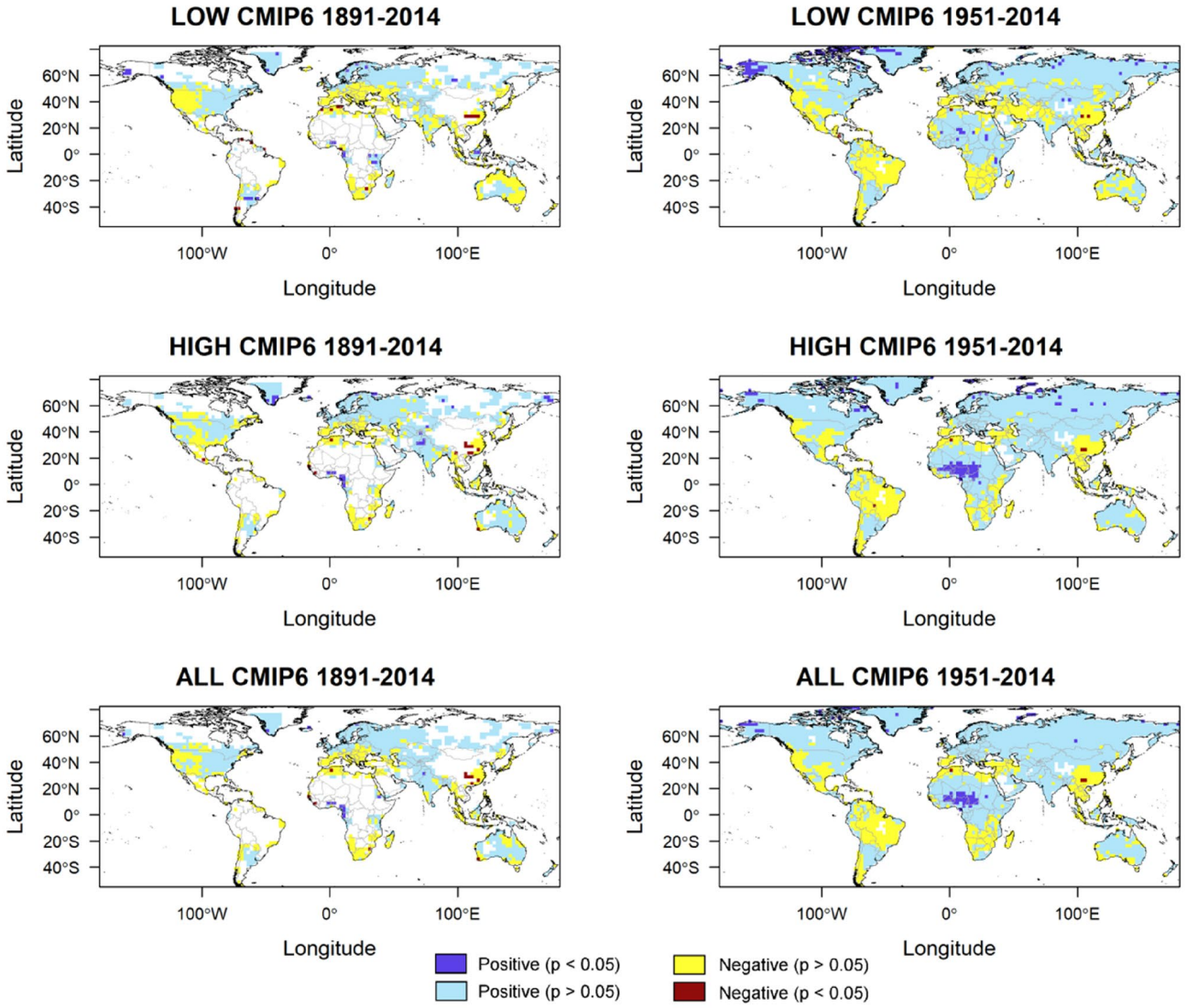


Fig. 7 Spatial distribution of the most frequent category of trend and statistical significance from the individual models in each one of the CMIP6 model groups

there is a spatial discrepancy between observed and modeled precipitation, which is stronger in CMIP5 models. Similarly, no significant differences were found between the high- and low-top models.

When comparing individual model trends to observations, there is a large spread in model trends. Models, in particular, overestimate drying in mid-latitude subtropical regions while underestimate observed precipitation increases in the Northern Hemisphere, particularly in North and Central Europe.

Based on comparisons of observations and model simulations, several studies have suggested human influence on precipitation trends (Zhang et al. 2007; Wan et al. 2015; Sarojini et al. 2016; Knutson and Zeng 2018). We argue that large uncertainties remain, mainly due to the strong interannual variability of precipitation, the lack of significant

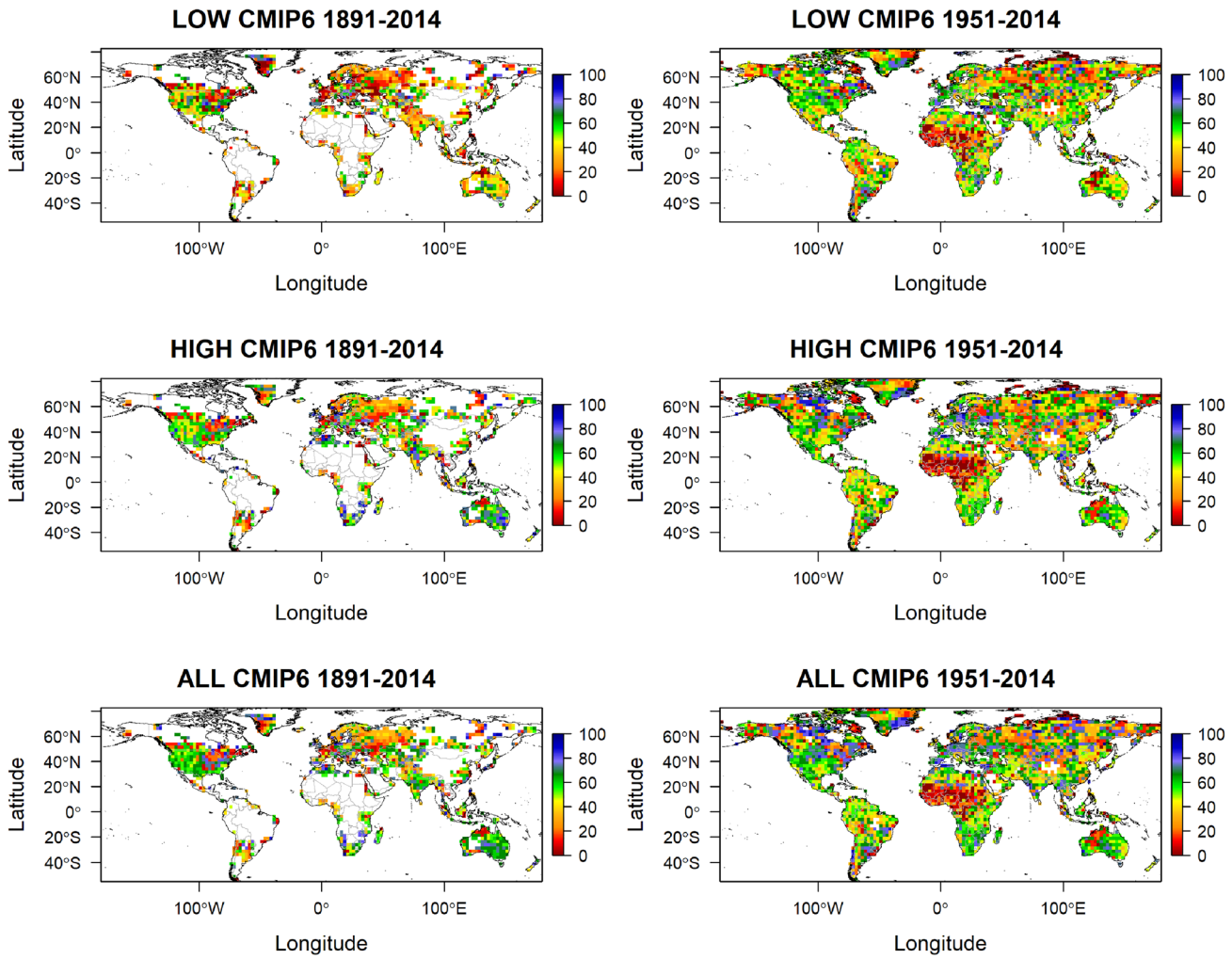


Fig. 8 Percentage of models in each CMIP6 group showing the same sign and significance of the observed annual precipitation trends

observed long-term trends (Nasrollahi et al. 2015; Knutson and Zeng 2018; Spinoni et al. 2019), the general overestimation or underestimation of precipitation changes at the regional scale, as well as the general spatial mismatch between observed and modelled precipitation trends.

This study emphasizes that future projections of precipitation trends from CMIP5 and CMIP6 should be viewed with caution because they do not adequately reproduce trends in observations. For example, several papers report that the general drying in large subtropical regions shown by the different groups of models is projected to magnify in future projections (Martin 2018; Spinoni et al. 2020; Ukkola et al. 2020). Nevertheless, even in these regions, these projections cannot be seen as fully robust, given the poor ability of models in reproducing long-term observed trends and correspondingly overestimating drying processes suggested by observations in these regions.

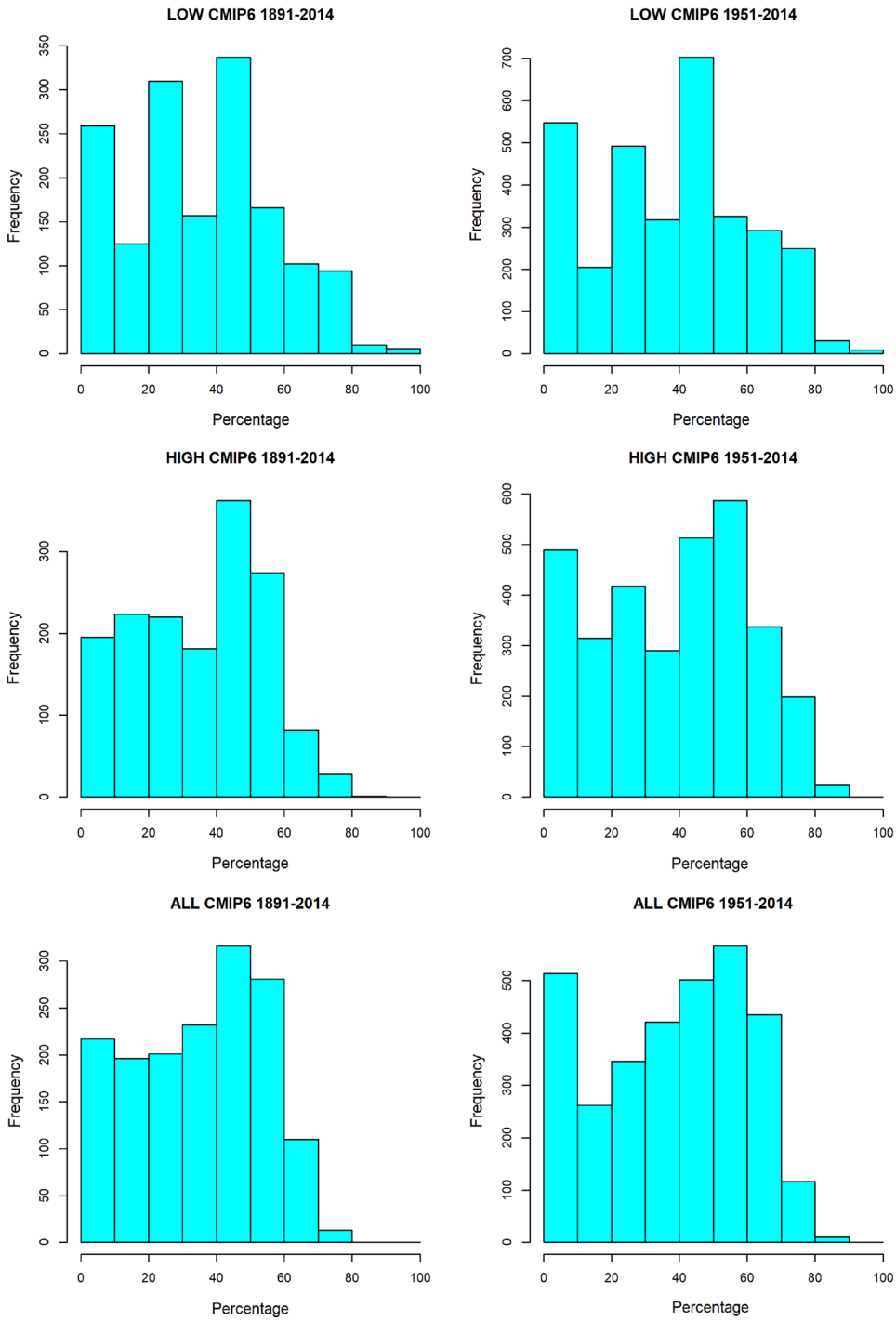


Fig. 9 Histograms depicting the absolute frequency of the world areas showing a percentage of agreement between categories of trend and significance of the individual models of the different CMIP6 groups in comparison to observations

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00382-021-06034-x.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. This work was supported by the research projects CGL2017-82216-R, PID2019-108589RA-I00 and PCI2019-103631 financed by the Spanish Commission of Science and Technology and FEDER and CROSSDRO project financed by the AXIS (Assessment of Cross(X)—sectoral climate Impacts and pathways for Sustainable transformation), JPI-Climate co-funded call of the European Commission. This study is also supported by "Unidad Asociada CSIC-Universidad de Vigo: Grupo de Fisica de la Atmosfera y del Océano". CM acknowledges funding from the Irish Environmental Protection Agency (2019-CCRP-MS.60). LG and RN received partial support from the Xunta de Galicia under the Project ED431C 2017/64–GRC Programa de Consolidación e Estructuración de Unidades de Investigación Competitivas (Grupos de Referencia Competitiva) and Consellería de Educación e Ordenación Universitaria.

Previous funders received cofounding from the ERDF, in the agenda of the Operational Program Galicia 2014–2020.

Data availability All the data used in this study is available at https://www.dwd.de/EN/ourservices/gpcc/gpcc.html; https://www.wcrp-climate.org/wgcm-cmip/wgcm-cmip5; and https://www.wcrp-climate.org/wgcm-cmip/wgcm-cmip6.

Declarations

Conflict of interest None.

References

Alexander LV, Arblaster JM (2017) Historical and projected trends in temperature and precipitation extremes in Australia in observations and CMIP5. Weather Clim Extrem 15:34–56

Aloysius NR, Sheffield J, Saiers JE et al (2016) Evaluation of historical and future simulations of precipitation and temperature in central Africa from CMIP5 climate models. J Geophys Res Atmos 121:130–152. https://doi.org/10.1002/2015JD023656

Asadi Zarch MA, Sivakumar B, Sharma A (2015) Droughts in a warming climate: a global assessment of Standardized precipitation index (SPI) and Reconnaissance drought index (RDI). J Hydrol 526:183–195. https://doi.org/10.1016/j.jhydrol.2014.09.071

Asadieh B, Krakauer NY (2015) Global trends in extreme precipitation: climate models versus observations. Hydrol Earth Syst Sci 19:877–891. https://doi.org/10.5194/hess-19-877-2015

Bhend J, Whetton P (2013) Consistency of simulated and observed regional changes in temperature, sea level pressure and precipitation. Clim Chang 118:799–810. https://doi.org/10.1007/s10584-012-0691-2

Bhend J, von Storch H (2008) Consistency of observed winter precipitation trends in northern Europe with regional climate change projections. Clim Dyn 31:17–28

Biasutti M, Held IM, Sobel AH, Giannini A (2008) SST forcings and sahel rainfall variability in simulations of the twentieth and twenty-first centuries. J Clim 21:3471–3486. https://doi.org/10.1175/2007JCLI1896.1

Bishop DA, Williams AP, Seager R et al (2018) Investigating the causes of increased twentieth-century fall precipitation over the southeastern United States. J Clim 32:575–590. https://doi.org/10.1175/JCLI-D-18-0244.1

Bladé I, Liebmann B, Fortuny D, van Oldenborgh GJ (2012) Observed and simulated impacts of the summer NAO in Europe: implications for projected drying in the Mediterranean region. Clim Dyn 39:709–727. https://doi.org/10.1007/s00382-011-1195-x

Brands S, Herrera García S, Fernández J, Gutiérrez J (2013) How well do CMIP5 Earth System Models simulate present climate conditions in Europe and Africa? A performance comparison for the downscaling community. Clim Dyn. https://doi.org/10.1007/s00382-013-1742-8

Bronnimann S, Fischer AM, Rozanov E et al (2015) Southward shift of the northern tropical belt from 1945 to 1980. Nat Geosci 8:969–974. https://doi.org/10.1038/ngeo2568

Cagnazzo C, Manzini E (2009) Impact of the stratosphere on the winter tropospheric teleconnections between ENSO and the North Atlantic and European Region. J Clim 22:1223–1238. https://doi.org/10.1175/2008JCLI2549.1

Charlton-Perez AJ, Baldwin MP, Birner T et al (2013) On the lack of stratospheric dynamical variability in low-top versions of the CMIP5 models. J Geophys Res Atmos 118:2494–2505. https://doi.org/10.1002/jgrd.50125

Chen L, Frauenfeld OW (2014) A comprehensive evaluation of precipitation simulations over China based on CMIP5 multimodel ensemble projections. J Geophys Res Atmos 119:5767–5786. https://doi.org/10.1002/2013JD021190

Cook BI, Smerdon JE, Seager R, Coats S (2014) Global warming and 21st century drying. Clim Dyn 43:2607–2627. https://doi.org/10.1007/s00382-014-2075-y

Dai N, Arkin PA (2017) Twentieth century ENSO-related precipitation mean states in twentieth century reanalysis, reconstructed precipitation and CMIP5 models. Clim Dyn 48:3061–3083. https://doi.org/10.1007/s00382-016-3251-z

Dai A, Zhao T (2017) Uncertainties in historical changes and future projections of drought. Part I: estimates of historical drought changes. Clim Change 144:519–533. https://doi.org/10.1007/s10584-016-1705-2

Davis N, Birner T (2016) Climate model biases in the width of the tropical belt. J Clim 29:1935–1954. https://doi.org/10.1175/JCLI-D-15-0336.1

Demory M-E, Berthou S, Fernández J et al (2020) European daily precipitation according to EURO-CORDEX regional climate models (RCMs) and high-resolution global climate models (GCMs) from the High-Resolution Model Intercomparison Project (HighResMIP). Geosci Model Dev 13:5485–5506. https://doi.org/10.5194/gmd-13-5485-2020

Deser C, Phillips A, Bourdette V, Teng H (2012) Uncertainty in climate change projections: the role of internal variability. Clim Dyn 38:527–546

Dirmeyer PA, Jin Y, Singh B, Yan X (2013) Trends in land-atmosphere interactions from CMIP5 simulations. J Hydrometeorol 14:829–849. https://doi.org/10.1175/JHM-D-12-0107.1

Duan A, Hu J, Xiao Z (2013) The Tibetan Plateau summer monsoon in the CMIP5 simulations. J Clim 26:7747–7766. https://doi.org/10.1175/JCLI-D-12-00685.1

Dufresne J-L, Foujols M-A, Denvil S et al (2013) Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5. Clim Dyn 40:2123–2165. https://doi.org/10.1007/s00382-012-1636-1

Eyring V, Bony S, Meehl GA et al (2016) Overview of the coupled model intercomparison project phase 6 (CMIP6) experimental design and organization. Geosci Model Dev 9:1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Grise KM, Davis SM, Simpson IR et al (2018) Recent tropical expansion: natural variability or forced response? J Clim 32:1551–1571. https://doi.org/10.1175/JCLI-D-18-0444.1

Gu G, Adler R, Huffman G (2015) Long-term changes/trends in surface temperature and precipitation during the satellite era (1979–2012). Clim Dyn. https://doi.org/10.1007/s00382-015-2634-x

Haase S, Matthes K, Latif M, Omrani N-E (2018) The importance of a properly represented stratosphere for northern hemisphere surface

variability in the atmosphere and the ocean. J Clim 31:8481–8497. https://doi.org/10.1175/JCLI-D-17-0520.1
Hannaford J, Buys G, Stahl K, Tallaksen LM (2013) The influence of decadal-scale variability on trends in long European streamflow records. Hydrol Earth Syst Sci 17:2717–2733. https://doi.org/10.5194/hess-17-2717-2013
Hardiman SC, Butchart N, Hinton TJ et al (2012) The effect of a well-resolved stratosphere on surface climate: differences between CMIP5 simulations with high and low top versions of the met office climate model. J Clim 25:7083–7099. https://doi.org/10.1175/JCLI-D-11-00579.1
Huang A, Zhou Y, Zhang Y et al (2014) Changes of the annual precipitation over Central Asia in the twenty-first century projected by multimodels of CMIP5. J Clim 27:6627–6646. https://doi.org/10.1175/JCLI-D-14-00070.1
Hurwitz MM, Calvo N, Garfinkel CI et al (2014) Extra-tropical atmospheric response to ENSO in the CMIP5 models. Clim Dyn 43:3367–3376. https://doi.org/10.1007/s00382-014-2110-z
Kelley C, Ting M, Seager R, Kushnir Y (2012) Mediterranean precipitation climatology, seasonal cycle, and trend as simulated by CMIP5. Geophys Res Lett. https://doi.org/10.1029/2012GL053416
Knutson TR, Zeng F (2018) Model assessment of observed precipitation trends over land regions: detectable human influences and possible low bias in model trends. J Clim 31:4617–4637. https://doi.org/10.1175/JCLI-D-17-0672.1
Knutti R, Sedláček J (2013) Robustness and uncertainties in the new CMIP5 climate model projections. Nat Clim Chang 3:369–373. https://doi.org/10.1038/nclimate1716
Kumar S, Merwade V, Kinter JL III, Niyogi D (2013) Evaluation of temperature and precipitation trends and long-term persistence in CMIP5 twentieth-century climate simulations. J Clim 26:4168–4185. https://doi.org/10.1175/JCLI-D-12-00259.1
Lee Y-Y, Black RX (2014) The structure and dynamics of the stratospheric northern annular mode in CMIP5 simulations. J Clim 28:86–107. https://doi.org/10.1175/JCLI-D-13-00570.1
Lee J, Waliser D, Lee H et al (2019) Evaluation of CMIP5 ability to reproduce twentieth century regional trends in surface air temperature and precipitation over CONUS. Clim Dyn 53:5459–5480. https://doi.org/10.1007/s00382-019-04875-1
Mariotti A, Zeng N, Yoon J-H, et al (2008) Mediterranean water cycle changes: transition to drier 21st century conditions in observations and CMIP3 simulations. Environ Res Lett 3
Marsh DR, Mills MJ, Kinnison DE et al (2013) Climate change from 1850 to 2005 simulated in CESM1(WACCM). J Clim 26:7372–7391. https://doi.org/10.1175/JCLI-D-12-00558.1
Martin ER (2018) Future projections of global pluvial and drought event characteristics. Geophys Res Lett. https://doi.org/10.1029/2018GL079807
McKee TB, Doesken NJ, Kleist J (1993) The relationship of drought frequency and duration to time scales. Eighth Conf Appl Climatol 179–184
Meehl GA, Covey C, Delworth T et al (2007) The WCRP CMIP3 multimodel dataset: a new era in climatic change research. Bull Am Meteorol Soc 88:1383–1394. https://doi.org/10.1175/BAMS-88-9-1383
Mehran A, AghaKouchak A, Phillips TJ (2014) Evaluation of CMIP5 continental precipitation simulations relative to satellite-based gauge-adjusted observations. J Geophys Res Atmos 119:1695–1707. https://doi.org/10.1002/2013JD021152
Nasrollahi N, Aghakouchak A, Cheng L et al (2015) How well do CMIP5 climate simulations replicate historical trends and patterns of meteorological droughts? Water Resour Res 51:2847–2864. https://doi.org/10.1002/2014WR016318
Orlowsky B, Seneviratne SI (2013) Elusive drought: uncertainty in observed trends and short-and long-term CMIP5 projections. Hydrol Earth Syst Sci 17:1765–1781. https://doi.org/10.5194/hess-17-1765-2013
Osprey SM, Gray LJ, Hardiman SC et al (2013) Stratospheric variability in twentieth-century CMIP5 simulations of the met office climate model: high top versus low top. J Clim 26:1595–1606. https://doi.org/10.1175/JCLI-D-12-00147.1
Paternoster R, BRAME R, Mazerolle P, Piquero A, (1998) Using the correct statistical test for equality of regression coefficients. Criminology 36:859–866. https://doi.org/10.1111/j.1745-9125.1998.tb01268.x
Peña-Angulo D, Vicente-Serrano SMM, Domínguez-Castro F et al (2020) Long-term precipitation in Southwestern Europe reveals no clear trend attributable to anthropogenic forcing. Environ Res Lett 15:094070. https://doi.org/10.1088/1748-9326/ab9c4f
Purich A, Cowan T, Min S-K, Cai W (2013) Autumn precipitation trends over southern hemisphere midlatitudes as simulated by CMIP5 models. J Clim 26:8341–8356. https://doi.org/10.1175/JCLI-D-13-00007.1
Räisänen J (2007) How reliable are climate models? Review Article Tellus A. https://doi.org/10.1111/j.1600-0870.2006.00211.x
Raut BA, Reeder MJ, Jakob C (2017) Trends in CMIP5 rainfall patterns over southwestern Australia. J Clim 30:1779–1788. https://doi.org/10.1175/JCLI-D-16-0584.1
Rivera JA, Arnould G (2020) Evaluation of the ability of CMIP6 models to simulate precipitation over Southwestern South America: climatic features and long-term trends (1901–2014). Atmos Res 241:104953. https://doi.org/10.1016/j.atmosres.2020.104953
Salunke P, Jain S, Mishra SK (2019) Performance of the CMIP5 models in the simulation of the Himalaya-Tibetan Plateau monsoon. Theor Appl Climatol 137:909–928. https://doi.org/10.1007/s00704-018-2644-9
Sarojini BB, Stott PA, Black E (2016) Detection and attribution of human influence on regional precipitation. Nat Clim Chang 6:669–675. https://doi.org/10.1038/nclimate2976
Scaife AA, Spangehl T, Fereday DR et al (2012) Climate change projections and stratosphere–troposphere interaction. Clim Dyn 38:2089–2097. https://doi.org/10.1007/s00382-011-1080-7
Schaller N, Mahlstein I, Cermak J, Knutti R (2011) Analyzing precipitation projections: a comparison of different approaches to climate model evaluation. J Geophys Res Atmos. https://doi.org/10.1029/2010JD014963
Schwalm CR, Glendon S, Duffy PB (2020) RCP8.5 tracks cumulative CO_2 emissions. Proc Natl Acad Sci 117:19656–19657. https://doi.org/10.1073/pnas.2007117117
Sheffield J, Camargo SJ, Fu R et al (2013) North American Climate in CMIP5 Experiments. Part II: evaluation of historical simulations of intraseasonal to decadal variability. J Clim 26:9247–9290. https://doi.org/10.1175/JCLI-D-12-00593.1
Sillmann J, Kharin VV, Zhang X et al (2013) Climate extremes indices in the CMIP5 multimodel ensemble: Part 1. Model evaluation in the present climate. J Geophys Res Atmos 118:1716–1733. https://doi.org/10.1002/jgrd.50203
Spinoni J, Barbosa P, De Jager A et al (2019) A new global database of meteorological drought events from 1951 to 2016. J Hydrol Reg Stud 22:100593. https://doi.org/10.1016/J.EJRH.2019.100593
Spinoni J et al (2020) Future global meteorological drought hot spots: a study based on CORDEX data. J Clim 33:3635–3661
Staten PWPW, Lu J, Grise KMKM et al (2018) Re-examining tropical expansion. Nat Clim Chang 8:768–775. https://doi.org/10.1038/s41558-018-0246-2
Su F, Duan X, Chen D et al (2013) Evaluation of the global climate models in the CMIP5 over the Tibetan Plateau. J Clim 26:3187–3208. https://doi.org/10.1175/JCLI-D-12-00321.1
Swain S, Hayhoe K (2015) CMIP5 projected changes in spring and summer drought and wet conditions over North America. Clim Dyn 44:2737–2750. https://doi.org/10.1007/s00382-014-2255-9

Taylor KE, Stouffer RJ, Meehl GA (2012) An overview of CMIP5 and the experiment design. Bull Am Meteorol Soc 93:485–498. https://doi.org/10.1175/BAMS-D-11-00094.1

Ukkola AM, De Kauwe MG, Roderick ML, Abramowitz G, Pitman AJ (2020) Robust future changes in meteorological drought in CMIP6 projections despite uncertainty in precipitation. Geophys Res Lett 47:e2020GL087820

van Oldenborgh GJ, Doblas Reyes FJ, Drijfhout SS et al (2013) Reliability of regional climate model trends. Environ Res Lett 8:14055. https://doi.org/10.1088/1748-9326/8/1/014055

Vicente-Serrano SM, López-Moreno JI (2008) Nonstationary influence of the North Atlantic Oscillation on European precipitation. J Geophys Res Atmos. https://doi.org/10.1029/2008JD010382

Vicente-Serrano SM, Domínguez-Castro F, Murphy C et al (2020) Long-term variability and trends in meteorological droughts in Western Europe (1851–2018). Int J Climatol 41:E690–E717

Wan H, Zhang X, Zwiers F, Min S-K (2015) Attributing northern high-latitude precipitation change over the period 1966–2005 to human influence. Clim Dyn 45:1713–1726. https://doi.org/10.1007/s00382-014-2423-y

Wei K, Cai Z, Chen W, Xu L (2018a) The effect of a well-resolved stratosphere on East Asian winter climate. Clim Dyn 51:4015–4028. https://doi.org/10.1007/s00382-016-3419-6

Wei Y, Yu H, Huang J et al (2018b) Comparison of the Pacific Decadal Oscillation in climate model simulations and observations. Int J Climatol 38:e99–e118. https://doi.org/10.1002/joc.5355

Xin X, Wu T, Zhang J et al (2020) Comparison of CMIP6 and CMIP5 simulations of precipitation in China and the East Asian summer monsoon. Int J Climatol. https://doi.org/10.1002/joc.6590

Zebaze S, Jain S, Salunke P et al (2019) Assessment of CMIP5 multi-model mean for the historical climate of Africa. Atmos Sci Lett 20:e926. https://doi.org/10.1002/asl.926

Zhang X, Zwiers F, Hegerl G et al (2007) Detection of human influence on twentieth-century precipitation trends. Nature 448:461–465. https://doi.org/10.1038/nature06025

Zhao L, Xu J, Powell AM, Jiang Z (2015a) Uncertainties of the global-to-regional temperature and precipitation simulations in CMIP5 models for past and future 100 years. Theor Appl Climatol 122:259–270. https://doi.org/10.1007/s00704-014-1293-x

Zhao T, Chen L, Ma Z (2015b) Simulation of historical and projected climate change in arid and semiarid areas by CMIP5 models Chin. Sci Bull 59(4):412–429