The DOI for this manuscript is doi: 10.1175/JCLI-D-14-00337.1

The final published version of this manuscript will replace the preliminary version at the above DOI once it is available.

If you would like to cite this EOR in a separate work, please use the following full citation:

Volosciuk, C., D. Maraun, V. Semenov, and W. Park, 2014: Extreme Precipitation in an Atmosphere General Circulation Model: Impact of Horizontal and Vertical Model Resolution. J. Climate. doi:10.1175/JCLI-D-14-00337.1, in press.

Extreme Precipitation in an Atmosphere General Circulation Model: Impact of Horizontal and Vertical Model Resolution

CLAUDIA VOLOSCIUK * AND DOUGLAS MARAUN

GEOMAR Helmholtz Centre for Ocean Research Kiel, Kiel, Germany

VLADIMIR A. SEMENOV

GEOMAR Helmholtz Centre for Ocean Research Kiel, Kiel, Germany

A.M. Obukhov Institute of Atmospheric Physics, Russian Academy of Sciences, Moscow, Russia,

P.P. Shirshov Institute of Oceanology, Russian Academy of Sciences, Moscow, Russia,

and Institute of Geography, Russian Academy of Sciences, Moscow, Russia

WONSUN PARK

GEOMAR Helmholtz Centre for Ocean Research Kiel, Kiel, Germany

* *Corresponding author address:* Claudia Volosciuk, GEOMAR Helmholtz Centre for Ocean Research Kiel, Duesternbrooker Weg 20, 24105 Kiel, Germany.

E-mail: cvolosciuk@geomar.de

Abstract

To investigate the influence of atmospheric model resolution on the representation of daily precipitation extremes, ensemble simulations with the atmospheric general circulation model ECHAM5 at different horizontal (T213 to T31) and vertical (L31 to L19) resolutions and forced with observed sea surface temperatures and sea ice concentrations have been carried out for 01/1982 - 09/2010. All results have been compared with the highest resolution, which has been validated against observations.

Resolution affects both the representation of physical processes and the averaging of precipitation across grid boxes. The latter, in particular, smoothes out localized extreme events. These effects have been disentangled by averaging precipitation simulated at the highest resolution to the corresponding coarser grid. Extremes are represented by seasonal maxima, modeled by the generalized extreme value distribution.

Effects of averaging and representation of physical processes vary with region and season. In the tropical summer hemisphere, extreme precipitation is reduced by up to 30% due to the averaging effect, and a further 65% owing to a coarser representation of physical processes. Towards mid- to high latitudes, the latter effect reduces to 20%; in the winter hemisphere it vanishes towards the poles. A strong drop is found between T106 and T63 in the convection dominated tropics. At the lowest resolution, northern hemisphere winter precipitation extremes, mainly caused by large scale weather systems, are in general represented reasonably well. Coarser vertical resolution causes an equatorward shift of maximum extreme precipitation in the tropics. The impact of vertical resolution on mean precipitation is less pronounced; for horizontal resolution it is negligible.

1. Introduction

Much of our knowledge about future changes in extreme weather events and the mechanisms causing these changes is based on global climate model simulations that employ general circulation models (GCMs). There is confidence that climate models provide credible quantitative estimates of future climate change, particularly at larger scales, because of their physical basis and the ability of models to reproduce observed climate and past climate changes (Flato et al. 2013). The representation of mean precipitation patterns has steadily improved between each phase of the Coupled Model Intercomparison Project (CMIP) used for the Intergovernmental Panel on Climate Change (IPCC) assessment reports (Flato et al. 2013). However, confidence in projections of extremes is generally weaker than for projections of long-term averages (Seneviratne et al. 2012). Extreme precipitation intensities (e.g., Sun et al. 2006), frequencies (e.g., Allan and Soden 2008) and return levels (Wehner et al. 2010) are generally underestimated by GCMs.

The simulation of precipitation is much more complex than that of temperature; anisotropic multifractal behavior over a wide range of scales has been attributed to precipitation (e.g., Lovejoy and Schertzer 1995) and the simulation of precipitation depends heavily on processes that are parameterized in current GCMs (Flato et al. 2013). To accurately represent extreme precipitation, models must correctly simulate atmospheric humidity as well as a number of relevant processes, such as evapotranspiration, condensation and transport processes (Randall et al. 2007). There are uncertainties in the simulation of the water cycle in most CMIP3 GCMs due to a time varying imbalanced atmospheric moisture budget. These biases in turn imply biases in the energy balance (Liepert and Previdi 2012; Lucarini and Ragone 2011).

Along with the increase of computational capacity since the first assessment report (FAR) of the IPCC, typical model resolution for short term climate simulations has increased from T21 (∼500 km) in the FAR to T106 (∼110 km) in the fourth assessment report (AR4) (Le Treut et al. 2007). Vertical resolution has also increased, from ten atmospheric layers in the FAR to about 30 layers in the AR4 (Le Treut et al. 2007). Nevertheless, resolving all important spatial and temporal scales remains beyond current capabilities for transient global climate change simulations (Le Treut et al. 2007). Biases thus remain, particularly on smaller scales and in the tropics, where the regional distribution of precipitation is strongly determined by convection, on a wide range of spatial and temporal scales, and on interactions between convective processes and the large scale circulation (Flato et al. 2013). For high resolution projections of precipitation extremes, different approaches have been employed: high-resolution GCMs, dynamical downscaling using regional climate models (RCMs) (Rummukainen 2010) and statistical downscaling (Maraun et al. 2010).

Several studies have investigated the resolution dependence of spatial precipitation patterns in atmospheric general circulation models (AGCMs). For example, patterns of seasonal mean precipitation in the NCAR AGCM CCM3 (Duffy et al. 2003; Iorio et al. 2004), as well as patterns of extreme precipitation (20-yr return levels) in the NCAR AGCM fvCAM2 (Wehner et al. 2010), are better represented over the USA with enhanced model resolution. Wehner et al. (2010) suggest $0.5^{\circ}\times0.625^{\circ}$, their highest resolution, to be a breakthrough resolution for the representation of extreme precipitation. However, precipitation intensity is still limited at this resolution, particularly for tropical cyclones (Wehner et al. 2010). Kopparla et al. (2013) have found biases in high percentiles ($>$95th) of daily precipitation in the NCAR AGCM CAM4 to decrease with finer resolution over the USA and Europe, whereas their highest resolution (0.25°) overestimates these high percentiles over Australia. Li et al. (2011) have shown in aqua-planet simulations with the CAM3 model that total precipitation increases at higher resolutions, especially in the tropics. The larger scales of the zonal average precipitation converge with increasing resolution for T85 and higher in the aqua-planet version of CAM3 (Williamson 2008). Seasonal differences in resolution dependence of extreme precipitation are indicated by Prein et al. (2013), who have found different mechanisms to be responsible for higher resolution requirement in June, July and August (JJA) (more small scale convective events) than in December, January and February (DJF) in an RCM over the Colorado Headwaters.

Changing horizontal model resolution has two effects on the representation of precipitation, in particular on its extremes. First, GCM simulated precipitation represents grid box area averages (e.g., Osborn and Hulme 1997; Chen and Knutson 2008); the coarser the model resolution, the more strongly localized events are smoothed out. To account for this "averaging effect", Chen and Knutson (2008) advise to compare extreme rainfall for different model resolutions after all data have been averaged to the lowest considered model resolution. Second, coarser model resolution involves reduced precision in the simulation of various features, especially feedbacks from smaller to larger scales. These feedbacks, including the impact of changes in resolved scales as well as in subgrid scales represented by parameterizations, deteriorates with coarser resolution. Hence, we refer to this effect as the "scale interaction effect". For instance, transient vertical velocities, and accordingly vertical moisture transport, are simulated more accurately with enhanced horizontal resolution (Pope and Stratton 2002; Li et al. 2011). A better representation of orography, due to higher horizontal resolution, improves local precipitation patterns (e.g., Smith et al. 2013; Pope and Stratton 2002; Duffy et al. 2003; Iorio et al. 2004) and has remote effects on the storm tracks as well as on the mean circulation (Pope and Stratton 2002; Jung et al. 2006). In general, changes in resolution mostly affect resolved scales, but there are also impacts on the parameterized physics (Roeckner et al. 2004). The more realistic representation of resolved dynamical properties provides, in turn, improved input to the parameterization schemes. Also, the interaction between parameterization schemes (e.g., between the convection and cloud microphysics schemes) is more detailed at higher resolution. Finally, truncation causes artificial separation of resolved and unresolved (i.e., parameterized) processes (Arakawa 2004). When changing the horizontal model resolution, one faces the combined effects of averaging and scale interaction. We call these overall effects "resolution effects"[1].

Changing vertical resolution affects several physical processes, particularly those related

[1]Note that resolution effects include changing grid size as well as changing the resolution dependent tunable parameters, see section 2b.

to the hydrological cycle. Higher vertical resolution leads to a marked redistribution of humidity and clouds (Roeckner et al. 2006). Most notable is the drying of the upper troposphere, which is related to a lowering of the tropopause and hygropause (Roeckner et al. 2006). In the tropics, the response of humidity and clouds to increased vertical resolution is related to changes in cloud top detrainment of water vapor and cloud water/ice (Roeckner et al. 2006). These improvements are largely due to the smaller numerical diffusion at higher vertical resolution, allowing for a larger, and also more realistic, vertical moisture gradient to be maintained throughout the troposphere (Hagemann et al. 2006). These changes in humidity and clouds in turn influence precipitation. On the global scale, both precipitation and evaporation are smaller at higher vertical resolution over land, in better agreement with observations (Hagemann et al. 2006). Finally, the sensitivity of the hydrological cycle to vertical resolution might be closely related to the tropospheric moisture changes caused by a more accurate vertical moisture transport at higher vertical resolution (Hagemann et al. 2006).

Which minimum resolution of GCMs is sufficient to represent patterns and characteristics of extreme precipitation at the global scale remains an open question. To our knowledge, there is no study investigating the resolution dependence of (1) extreme precipitation on (2) the global scale, with (3) realistic topography and (4) separately for different seasons. We are also not aware of any study investigating the impact of vertical resolution on extreme precipitation. While it is widely acknowledged that the averaging effect plays an important role when evaluating extreme precipitation on gridded datasets, and therefore should be removed before any comparisons of extreme precipitation from different sources are carried out, its separation from the overall resolution effect and quantification across different scales remains an open question.

Here, we study the dependency of extreme precipitation on horizontal and vertical model resolution. In particular, we address the following questions:

i. What is the importance of the averaging effect to the overall resolution effect when simulating extreme precipitation?

ii. To what extent does representation of extreme precipitation at different resolutions depend on season?

iii. At which resolution, compared with the highest considered resolution, is the strongest deterioration in the representation of extreme precipitation evident?

iv. Are there regions where the dependence of extreme precipitation on resolution is weak or where the scale interaction effect can be neglected?

v. What is the influence of vertical resolution on the representation of extreme precipitation?

In section 2 of the paper, we describe the setup of the atmospheric model, the design of the resolution experiment and the statistical model used to analyze extremes. In section 3, modeled extreme precipitation return levels at different horizontal and vertical resolutions are compared for different seasons. Finally, section 4 contains the conclusions.

2. Data and Methods

We consider daily precipitation simulated by the AGCM ECHAM5. A key part of our study is to disentangle the averaging and scale interaction effects. To this end, we consider simulations at different resolutions and compare them with the highest resolution simulation, averaged to the corresponding lower spatial scales as recommended by Chen and Knutson (2008).

a. The Atmospheric General Circulation Model

We use the AGCM ECHAM5 (Roeckner et al. 2003), developed at the Max Planck Institute for Meteorology, Germany. ECHAM5 is a global spectral model and calculates precipitation fluxes on the Gaussian transform grid (Roeckner et al. 2003). The sensitivity of ECHAM5 to horizontal and vertical resolution has been studied for mean climate characteristics (Roeckner et al. 2006) and the hydrological cycle (Hagemann et al. 2006). Notable deficiencies in the hydrological cycle are a dry bias over Australia and a lack of a rainforest climate in central Africa, where precipitation is too low during the dry season (Hagemann et al. 2006). The ECHAM5 model overestimates precipitation over the oceans, especially in high-resolution simulations. This bias is a general problem in current GCMs that could possibly be related to insufficient atmospheric absorption of solar radiation by aerosols, water vapor, or clouds (Hagemann et al. 2006). The bias of basic climate variables decreases monotonically with increasing horizontal resolution from T42 to T159 (Roeckner et al. 2006).

As the L31 vertical resolution versions are superior to their L19 counterparts, except for T42 horizontal resolution, Roeckner et al. (2006) recommend the vertical resolution L19 for the horizontal resolutions T31 and T42, and the vertical resolution L31 for higher horizontal resolutions. Enhanced vertical resolution is more beneficial than increased horizontal resolution for the simulation of mean precipitation in ECHAM5 (Hagemann et al. 2006).

b. Experiments

We carried out simulations covering the period 01/1982 - 09/2010 (29 years), driven with the same transient present day boundary forcing for all resolutions. Sea surface temperatures (SSTs) and sea ice concentrations (SICs) were interpolated to the corresponding horizontal resolutions from optimal interpolation 1/4 degree daily SST analysis (OISST), version 2, (Reynolds et al. 2007) and high resolution (12.7 km) observed SIC from Grumbine (1996) of the National Oceanic and Atmospheric Administration (NOAA). Greenhouse gas forcing was kept constant at present day concentrations (348 ppm). An overview of the different horizontal and vertical resolutions of these simulations is given in Table 1. Three ensemble realizations of the resolutions T106L31, T63L31, T42L19 and T31L19 were run to assess internal variability. The top four and bottom two vertical levels of L31 and L19 are similar.

The greatest difference (doubling) in vertical resolution occurs between approximately 70 and 500 hPa (Roeckner et al. 2003). In all resolutions we used the default ECHAM5 parameterization and the parameter settings recommended by Roeckner et al. (2004, 2006) for the respective resolution. Note that our aim is not to isolate the sensitivity of the dynamical and physical response to pure grid spacing from the sensitivity of modeled precipitation to tunable parameters. Such intention would require experiments with fixed parameterizations and tuning parameter values such as proposed in Leung et al. (2013) and applied by, e.g., Rauscher et al. (2013). Our objective is rather to quantify the effect of changing the model resolution, and to separate this effect into the contribution of spatial averaging and the residual scale interaction effect. Our definitions of both scale interaction and resolution effect thus are not limited to changing the grid spacing, but additionally include the adaptation of tunable parameters to recommended values as feedback from parameterizations also interact with different scales. Nevertheless, additional experiments showed that the sensitivity of extreme precipitation to parameter choice is negligible in the range of considered resolutions (not shown).

c. *Statistical Model*

We modeled daily precipitation extremes with the block maxima approach, following the Fisher-Tippet theorem: Given a sequence of n independent identically distributed random variables X_i, $i = 1, \ldots, n$, the properly rescaled maximum of this sequence M_n converges for large n - in case a limiting distribution exists - to the Generalized Extreme Value (GEV) family of distributions (Fisher and Tippett 1928; Gnedenko 1943; Coles 2001):

$$G(z) = \exp\left\{ -\left[1 + \xi \left(\frac{z-\mu}{\sigma} \right) \right]^{-1/\xi} \right\} \quad (1)$$

with the location parameter μ, the scale parameter σ and the shape parameter ξ. The tail of the distribution is determined by ξ as follows: $\xi \to 0$: infinite smooth tail; $\xi > 0$: infinite heavy tail; $\xi < 0$: bounded tail (Coles 2001). The independence assumption of the Fisher-

Tippet theorem can be relaxed to a wide class of stationary, but not necessarily independent, processes (Coles 2001; Rust 2009; Faranda et al. 2011, 2013).

Extreme quantiles are obtained by inverting Eq. 1:

$$z_p = \begin{cases} \mu - \frac{\sigma}{\xi}\left[1 - \{-\log(1-p)\}^{-\xi}\right] & \text{for } \xi \neq 0 \\ \mu - \sigma \log\{-\log(1-p)\} & \text{for } \xi = 0 \end{cases}, \tag{2}$$

where $G(z_p) = 1 - p$. The return level z_p associated with the return period $1/p$ is expected to be exceeded on average once in $1/p$ blocks, i.e., z_p is exceeded in any particular block with probability p (Coles 2001).

Parameters of the GEV distribution (Eq. 1) were estimated with Probability Weighted Moments (PWM) (Hosking et al. 1985) using the "fExtremes" package (Wuertz 2009) in R (R Development Core Team 2011). PWM performs well for small sample sizes and is computational efficient (Hosking et al. 1985). The analysis was carried out seasonwise. A block length of one season (i.e., three months) turned out to be a good compromise between an appropriate fit for most regions and a sufficiently long maxima time series of 29 years to keep sampling uncertainties reasonably low. To avoid a misfit of the GEV distribution in very dry regions, we excluded time series from our analysis that contained more than one zero in the seasonal maxima time series. As a representation of extreme events, we considered the 20 season return level of daily precipitation (RL20S). For example, the RL20S for DJF is exceeded in any DJF season with the probability 1/20, i.e., on average every 20th DJF season. The RL20S is already reasonably extreme, but still low enough to avoid biases caused by the estimation procedure (Hosking et al. 1985) or undesirably high estimation uncertainty. Sampling uncertainties of RL20S were assessed by a bootstrap method (see appendix for details).

d. Separation of averaging and scale interaction effects

The results of the simulations at different model resolutions are compared with our highest resolution (T213L31). Chen and Knutson (2008) advise that, when comparing extreme precipitation from different sources, precipitation should be averaged to the same spatial scale beforehand, as climate models provide grid box averages of precipitation (e.g., Roeckner et al. 2003, for ECHAM5), which includes the averaging effect if precipitation is compared on different grids. We averaged daily precipitation at the highest resolution (T213) to coarser grids for comparison with the coarser resolutions on similar spatial scales (see Table 2). Statistics were calculated after daily precipitation had been averaged to the appropriate spatial scale. In the following, we refer to the simulations carried out at different model resolutions (Table 1) as coarser resolution simulations (CRS). The averaged T213 resolutions $T213_{2\times2}, \ldots, T213_{7\times7}$ (Table 2) are referred to as averaged high resolution simulations (AHS). The averaging effect was approximately disentangled from the scale interaction effect by comparing RL20S in CRS with those in AHS on similar spatial scales.

3. Results and Discussion

The highest resolution, T213L31, has been validated against observational datasets: globally for seasonal mean precipitation and over the USA, Europe, Russia, the Middle East and southeast Asia for extreme precipitation. The global pattern of seasonal mean precipitation, as well as many features of the regional spatial distribution of RL20S, are well represented (see appendix for details).

a. Resolution and averaging effect

Fig. 1 illustrates the global pattern of RL20S as a function of resolution for DJF and JJA. The first and third rows (panels a - c and f - h) show CRS and, hence, the full resolution effect, including both averaging and scale interaction. The second and fourth rows (panels d - e and i - k) show AHS and, thus, represent solely the averaging effect in relation to the respective first panel (a and f). The differences between the first (third) and second (fourth) rows illustrate the scale interaction effect. The middle panels differ in horizontal resolution, while the right panels differ in horizontal and vertical resolution. The general global pattern of the RL20S is captured by all resolutions: differences are rather small and mainly related to reduced magnitudes[2]. The differences between RL20S in CRS and in AHS are in general smaller for T63L31 than for T31L19, see, e.g., the south Pacific in DJF and Siberia in JJA. These differences indicate a better performance of T63L31 in both DJF and JJA.

Fig. 2 demonstrates the different effects for four example regions: the tropical Amazon region, which is governed by deep convection; the southeastern USA, a subtropical climate with mild winters; eastern Asia, a continental climate with cold snowy winters; northern Europe, where winter precipitation is mainly caused by large scale weather systems AHS (black) represents the averaging effect of RL20S, i.e., this scaling dependence is caused by increased grid size. CRS (blue) shows the overall resolution effects of the RL20S. The difference between the RL20S in AHS and in CRS is a first order estimate of the scale interaction effect. The pure averaging effect in general causes a decrease of RL20S in AHS with increasing spatial length scale. The same holds for CRS. Three different horizontal scaling dependencies of RL20S are found. CRS is either below (e.g., Amazon region), approximately equal (95 % confidence intervals overlap; e.g., southeastern USA) or above (e.g., eastern Asia) AHS. This finding indicates that the dominant mechanism strongly influences the scaling behavior and thereby also determines the minimal required horizontal resolution. Different vertical resolutions (blue and red) are compared in section c.

[2]Note that regional differences are masked by the logarithmic scale.

b. *Influence of horizontal resolution*

To quantify the differences between the RL20S in the CRS and in the AHS, grid box wise ratios of the RL20S at each resolution to the corresponding averaged high resolution[3] were computed (see Fig. 3). The colors in Fig. 3 correspond to the different scaling types in Fig. 2 as follows: (a) red: RL20S in CRS below RL20S in AHS, (b) yellow: both curves approximately equal, and (c) blue: RL20S in CRS above RL20S in AHS. RL20S strongly decreases between T106 and T63 over an almost entire zonal band. This behavior is particularly pronounced in regions where deep convection is the main mechanism causing extreme precipitation, i.e., close to the intertropical convergence zone (ITCZ). This big difference between T106 and T63 suggests that T106 is an efficient horizontal resolution for simulating extreme precipitation at these latitudes. However, for all resolutions, parts of the northern hemisphere's landmass remain in the range of ± 20% from T213 in DJF, indicating that extreme precipitation is still represented comparably well at T31L19 resolution.

Fig. 4 shows the impact of all resolution effects in CRS compared to the high resolution at its original resolution - not to those in AHS - on the representation of extreme precipitation. T106 resolution is again good enough for simulating extreme precipitation. The deterioration of return level representation from T106 to T63 is even more pronounced and extends to a wider area as when compared with AHS (see Fig. 3). Yet still, wide areas in the northern hemisphere in DJF are not sensitive to changes in resolution. In these regions, both scale interaction and averaging effects are negligible.

To illustrate the benefit of choosing a higher resolution, compared with the nearest coarser resolution, the overall difference of extreme precipitation return level representation without "removing" the averaging effect between consecutive resolutions is provided in Fig. 5. Again, T106 is an efficient resolution for simulating extreme precipitation.

[3]For resolutions which do not have an exactly corresponding averaged T213 resolution (T159, T63, T31), the corresponding value was linearly interpolated between the two surrounding averaged T213 resolutions (e.g., $T213_{3\times3}$ and $T213_{4\times4}$ for T63).

Fig. 6 provides zonal means of the RL20S for all considered resolutions. Panels a and b show zonal means of the RL20S covering the overall resolution effect. In panels c and d the zonal means (a, b) are normalized by the zonal mean of the RL20S of the corresponding averaged high resolution[4], i.e., the averaging effect is approximately removed and only the residual scale interaction effect is shown. As expected, meridional variation decreases at coarser resolution. The highest relative reduction occurs in the belt of extreme tropical summer precipitation related to the ITCZ: here the RL20S decreases by about 75% from T213 to T31 (a and b). This reduction is dominated by the scale interaction effect. After removing the averaging effect, the decrease still amounts to 65% (c and d). The averaging effect alone thus causes a decrease of approximately $1 - \frac{0.25}{0.35} = 29\%$. In the mid-to higher latitudes of the summer hemisphere, the scale interaction effect reduces to a decrease of about 20%; in the winter hemisphere it vanishes towards the poles.

The most noticeable differences are again found between the RL20S in T106 and in T63. For instance, the RL20S peaks just off the equator, towards the winter hemisphere, vanish at T63 and lower resolutions (a and b). The corresponding dips in panels c and d indicate that this reduction is caused by the scale interaction effect. However, consistent with the ratios in Fig. 3 - 5, the zonal means of the RL20S in the mid- and high latitudes in winter are not sensitive to changes in resolution.

[4]For resolutions which do not have an exactly corresponding averaged T213 resolution (T159, T63, T31), the corresponding averaged T213 zonal mean was approximated as follows: Initially, both surrounding averaged T213 zonal means (e.g., $T213_{3\times3}$ and $T213_{4\times4}$ for T63) were interpolated to the latitudinal scale of the coarser horizontal resolution (e.g., T63) to have an equal number of values. Subsequently, a weighted mean between the averaged T213 zonal means was taken. The weights were chosen according to the position of the coarser horizontal resolution's latitudinal length scale in relation to each surrounding averaged T213 resolution's latitudinal length scale.

c. *Vertical resolution*

Fig. 2 shows that vertical resolution also has a regionally varying impact on the representation of extreme precipitation. Over northern Europe in DJF, differences between the area averages of the RL20S at different vertical resolutions are negligible, whereas in the other regional examples the area average of the RL20S at coarser vertical resolution is less than the area average of the RL20S at higher vertical resolution. This difference is more pronounced at T63 than at T42.

To further investigate the structure of changes in the RL20S with vertical resolution, zonal means of the RL20S (Fig. 6) of high vertical resolution (solid lines) are compared with the RL20S of the low vertical resolution (dashed lines). Coarser vertical resolution causes a decrease in the RL20S. Additionally, the peak of extreme tropical summer precipitation associated with the ITCZ is shifted equatorwards at coarser vertical resolution. This effect is stronger in boreal summer (JJA) than in austral summer (DJF). The spatial structure of changes in extreme precipitation return levels with vertical resolution is shown in Fig. 7. The impact of vertical resolution is higher at T63 than at T42, consistent with the regional examples (Fig. 2). High vertical resolution is particularly important in a zonal band around the ITCZ. For extreme precipitation associated with the Asian monsoon, high vertical resolution is crucial. However, over parts of the northern hemisphere in DJF, coarser vertical resolution is sufficient for the representation of the RL20S.

d. *Comparison with mean precipitation*

Fig. 8 shows zonal means of mean precipitation totals (a, b), mean precipitation intensities (c, d) and the mean number of wet days (e, f) for DJF and JJA to study differences to the scale dependence of extreme precipitation. The impact of horizontal resolution on mean precipitation totals and mean precipitation intensity is negligible. Peaks of the high resolutions T213, T159 and T106 are similar, however coarser resolutions show slightly decreased peaks. Even though these differences are small compared to those of extremes, there is consistency regarding the large differences between T106 and T63 which were observed for extremes. As zonal means of coarser vertical resolution show a slightly different structure, higher vertical resolution is beneficial for the representation of mean precipitation totals and intensities as well. However, these differences are less pronounced than for extremes.

The mean number of wet days increases with coarser resolution due to small scale events being averaged over a larger area ("drizzle effect"). The differences in the mean number of wet days between resolutions are most pronounced in the mid- and high latitudes of the northern hemisphere in DJF, as well as in JJA. Most landmasses are located in this area, leading to different representations of orography at different resolutions, which influences, e.g., precipitation induced by orographic lifting. In JJA, over the mid- and high latitudes of the northern hemisphere, vertical resolution appears to be an important factor, in addition to horizontal resolution. In DJF, vertical resolution does not appear to play an important role in the mean number of wet days. These results suggest that spatial resolution also has an impact on the representation of dry spells in the model we use.

Discussion

The strong dependence of extreme precipitation on model resolution is consistent with Wehner et al. (2010), Chen and Knutson (2008) and Kopparla et al. (2013). Wehner et al. (2010) found $0.5° \times 0.675°$ (similar to T213) of the fvCAM2 to be a breakthrough resolution for the representation of 20-yr return level patterns over the USA, particularly for precipitation intensities of tropical cyclones in the southeastern USA, by validating the model with observational patterns of 20-yr return levels on similar spatial scales. We found that return levels at T106 ($1.13° \times 1.13°$) were comparable to those of the highest resolution T213 ($0.56° \times 0.56°$) in most regions. Thus, in general, at least T106 appears to be required for the representation of extreme precipitation. Consistent with our results, their coarsest resolution $2° \times 2.5°$ (between T63 and T42) is too coarse to represent the main features of extreme precipitation return levels, compared with observations (Wehner et al. 2010).

The efficiency of ECHAM5 in simulating extreme precipitation at different resolutions varies with season and region. These differences are likely due to a varying convective contribution to total precipitation and a changing height of the convective cell. Areas where deep convection is an important process generally require higher horizontal resolution than regions where extreme precipitation is mainly due to large scale weather systems. For the representation of extreme precipitation resulting from large scale weather systems, the scale interaction effect is negligible and higher horizontal resolution only reduces the averaging effect. These differences, which are related to different underlying mechanisms, were identified by studying seasonal instead of annual return levels.

Roeckner et al. (2006) found an adequate representation of climate in ECHAM5 with a vertical resolution of L19 for T42 and T31. In contrast to these findings, Hagemann et al. (2006) found a higher vertical resolution of L31 to improve the representation of mean precipitation in ECHAM5. Here we show that this effect is even more pronounced for extreme precipitation. Our results demonstrate that, in general, higher vertical resolution is necessary to study extreme precipitation: L31 outperforms L19 at all horizontal resolutions, except for parts of the mid- and high latitudes in winter. Mean precipitation, as well as evaporation, at coarser vertical resolution is higher over land and lower over the ocean in ECHAM5 (Hagemann et al. 2006), whereas dependence of extreme precipitation on vertical resolution varies with latitude and season over ocean as well as land.

We show that for mean precipitation, the impact of horizontal resolution is negligible, which is consistent with Hagemann et al. (2006) and Kopparla et al. (2013). A comparison of mean precipitation totals and intensities with extreme precipitation yields completely different structures of resolution dependence and, hence, extreme precipitation cannot be estimated directly from mean precipitation intensities or from a distribution that was estimated or corrected according to the mean.

4. Conclusions

We analyzed the impact of horizontal and vertical resolution on the representation of extreme precipitation return levels in the AGCM ECHAM5. ECHAM5 was driven with the same transient present day boundary forcings for all resolutions.

Decreasing horizontal resolution has several impacts on extreme precipitation. First, increasing grid size has the effect that precipitation is averaged over a larger area (averaging effect). Second, in lower horizontal resolutions the coarser representation of, e.g., physical processes and orography yields inferior representation of extreme precipitation (scale interaction effect). Note that we do not intend to identify the pure grid spacing effect, but rather define the resolution effect as the overall effect of changing grid spacing and tunable parameters. If one were interested in a separation of the pure grid spacing, one would have to carry out experiments as proposed by Leung et al. (2013) and applied by, e.g., Rauscher et al. (2013). The highest resolution (T213) averaged to coarser grid sizes ($T213_{1\times1}$ - $T213_{7\times7}$: averaged high resolution simulation - AHS) was compared with coarser resolutions (T159 - T31: coarser resolution simulations - CRS). Differences between AHS and CRS provide an approximate first order discrimination between these two effects. Thereby, the relative importance of both effects was determined. 20 season return levels of daily precipitation (RL20S) in different resolutions were compared, derived from a generalized extreme value (GEV) distribution.

Horizontal, as well as vertical, model resolution were found to affect the representation of extreme precipitation. The averaging effect contributes considerably to decreasing return levels with resolution. In the belt of tropical summer extreme precipitation associated with the ITCZ, averaging from T213 to T31 reduces the RL20S by almost 30%. Hence, in accordance with Chen and Knutson (2008), we strongly recommend to compare extreme precipitation from different sources (e.g., different models, observations) only after averaging to the same spatial scale. The scale interaction effect is strongest in the summer hemisphere. In the band of extreme precipitation associated with the ITCZ, the reduction amounts to around 65% when changing the model resolution from T213 to T31. Towards mid-to higher latitudes, the scale interaction effect reduces to a decrease of about 20%. In the winter hemisphere it vanishes towards the poles.

The minimum required horizontal resolution for extreme precipitation was found to depend on season and region and, thus, mainly on the underlying process(es). In general, extreme precipitation caused by small scale convective events requires higher horizontal resolution than extreme precipitation caused by synoptic scale weather systems. Particularly in the tropics, but also in the extratropics during summer, at least T106 is required to represent comparable return levels to the highest resolution T213. Only marginal changes to RL20S, caused by the averaging effect, were found in the mid- and high latitudes in winter, such as over parts of the northern hemisphere's landmass in DJF; here RL20S in T31L19 are comparable to those in the highest resolution (T213) on similar spatial scales. Over wide areas of the mid- and high latitudes during winter (e.g., Canada and Asia in DJF), extreme precipitation was even found to be insensitive to changes in resolution when comparing T31 with the highest resolution (T213) at its original resolution.

Higher vertical resolution is crucial for the representation of precipitation (consistent with Hagemann et al. 2006). This applies particularly to the extremes, as coarser vertical resolution causes an equatorward shift of maximum extreme precipitation, as well as a decrease in return levels. Therefore, we recommend the use of higher vertical resolution for extreme precipitation, even for relatively coarse horizontal resolutions such as T42 or T63. Yet, the impact of vertical resolution is more pronounced in T63 than in T42. An exception is during winter in the mid- and high latitudes where RL20S in coarser vertical resolution are comparable to those in high vertical resolution.

Extreme precipitation shows a completely different scale dependence to mean precipitation. The impact of horizontal resolution on mean precipitation is negligible, whereas higher vertical resolution is still meaningful but less pronounced than for the extremes. This implies that extreme precipitation cannot be estimated directly from mean precipitation intensities or from a distribution that was estimated or corrected according to the mean.

Here we present a model study where we take the highest model resolution as reference for comparison with the coarser model resolutions. This reference simulation, in general, compares well with gridded observations, but also shows deficiencies in simulating Asian monsoon as well as orographic extreme precipitation, which both tend to be overestimated. By construction, we disregard effects not correctly simulated by the highest considered resolution of the chosen model. In all considered resolutions, convection is parameterized. Thus, related dynamical feedbacks are not resolved. Other relevant processes for extreme precipitation that might need even higher resolution than all considered resolutions, such as tropical cyclones (Wehner et al. 2010), are beyond the scope of our study. Furthermore, climate models may not fully capture important features of atmospheric dynamics related to extremes, in particular persistent weather regimes (Petoukhov et al. 2013; Palmer 2013). Finally, as we have employed an atmosphere only model with prescribed ocean boundary conditions, ocean feedbacks are likewise not represented. Any recommendations for minimum resolutions refer solely to the representation of RL20S in an AGCM and do not imply that the above listed phenomena are well represented at these resolutions.

Although we have only studied the scaling behavior of extreme precipitation in one AGCM, i.e., ECHAM5, we believe that our results are also valid for other AGCMs as physical explanations for the scale dependence of extreme precipitation could be identified.

Acknowledgments.

The authors acknowledge help with the simulations by W. Tseng, N. Keenlyside and G. Zhou. We thank M. Latif, A. Schindler and E. Meredith for helpful discussions as well as V. Lucarini and three anonymous reviewers for comments on the manuscript. Simulations were run at the North-German Supercomputing Alliance (HLRN). This study was funded by the EUREX project of the Helmholtz Association (HRJRG-308) and supported by Russian Foundation for Basic Research (14-05-00518) and Russian Ministry of Education and Science (grant no. 14.B25.31.0026). The GPCP combined precipitation data were developed and computed by the NASA/Goddard Space Flight Center's Laboratory for Atmospheres as a contribution to the GEWEX Global Precipitation Climatology Project. GPCP and CPC US Unified Precipitation data are provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at http://www.esrl.noaa.gov/psd/. We acknowledge the E-OBS dataset from the EU-FP6 project ENSEMBLES (http://ensembles-eu.metoffice.com) and the data providers in the ECA&D project (http://eca.knmi.nl) as well as the APHRODITE dataset and the data providers in the APHRODITE's Water Resources project (http://www.chikyu.ac.jp/precip/).

APPENDIX

Uncertainties in the Return Levels

a. Internal model variability

One source of uncertainty in the estimation of return levels is internal variability of the climate system. To assess this unforced internal variability of the climate model, long time series are required. As our model runs are only 29 years long, due to limited availability of the high resolution boundary conditions, we performed three ensemble members with slightly different initial conditions for the resolutions T106L31, T63L31, T42L19 and T31L19 which are each 29 years long. The difference between RL20S in these three ensemble members yields uncertainties in the return level estimation due to the climate model's internal variability. Fig. 9 shows zonal means and the respective zonal standard deviations of RL20S in these three ensemble members for different resolutions. Rather small differences between the zonal means of the three ensemble members in all resolutions in DJF as well as in JJA indicate that the forced climate is reliably represented.

b. GEV sampling uncertainty

In this study, GEV parameters were estimated from 29 data points of three month long blocks. This rather small sample size may cause uncertainties in the return levels. To assess these uncertainties, we applied a parametric bootstrap method (Efron and Tibshirani 1993) to the highest (T213L31) and coarsest resolution (T31L19) as follows. 1000 random time series (size: 29 data points, as in the actual sample), distributed according to the fitted GEV distribution, were generated for each grid box. Subsequently, GEV parameters for each time series were estimated. The 95% confidence interval of the empirical distribution of RL20S in these 1000 realizations quantifies the GEV parameter uncertainties of RL20S. Fig. 10 shows the zonal mean of RL20S in this study (solid lines) and the zonal mean of the grid box wise 95% confidence intervals derived from the bootstrap method (dashed lines), i.e., the latitude dependent mean parameter uncertainty of a grid box is shown. The confidence intervals are quite symmetric and indicate an acceptable spread, which gives us confidence in our return level estimates. Note that this is the parameter uncertainty of the mean grid box at a given latitude. Under the assumption that the empirical distribution is symmetric and the samples are independent, the parameter uncertainty of the zonal mean is related to the zonal mean of the parameter uncertainty by a scaling factor of $\frac{1}{\sqrt{n}}$ (according to Gaussian error propagation). Thus, sampling uncertainties for the zonal mean (see Fig. 10) are negligible.

c. Validation of the highest resolution of ECHAM5 with observational datasets

To assess the performance of the highest resolution (T213L31) of ECHAM5 which is used as reference for the coarser resolutions in our study, we validated model precipitation with gridded observational datasets. As no global daily precipitation dataset with sufficient density of rain gauges is available to reliably estimate extreme precipitation return levels, the latter were only validated for regions where daily precipitation gridded datasets with a high density of rain gauges are available. On a global level we validated seasonal mean precipitation using the global precipitation climatology project (GPCP) dataset (Adler et al. 2003). The GPCP gridded dataset is a globally complete monthly analysis of surface precipitation at 2.5°×2.5° resolution (Adler et al. 2003). It incorporates precipitation estimates from low-orbit satellite microwave data, geosynchronous-orbit satellite infrared data and surface rain gauge observations (Adler et al. 2003). Precipitation of the ECHAM5 model output was averaged by area conservative remapping to the GPCP grid. 20 season return levels (RL20S) were validated over the USA, Europe, Russia, the Middle East and southeastern Asia. For the USA, the NOAA CPC (Climate Prediction Center) “US Unified Precipitation” dataset (Higgins et al. 2000) was used. This is based on approximately 35 000 rain gauges over the whole continental USA, sparsest in the western USA, and gridded to 0.25°×0.25° (Higgins et al. 2000). RL20S over Europe is validated with the European daily high-resolution (0.25°×0.25°) gridded data set (E-OBS, version 9) of precipitation (Haylock et al. 2008). This has been developed in the framework of the ENSEMBLES project. The density of rain gauges is irregular and, in some regions, sparse (Haylock et al. 2008). To estimate RL20S over Asia, the "Asian precipitation - highly-resolved observational data integration towards evaluation of the water resources" (APHRODITE) dataset (Yatagai et al. 2012) was employed. The APHRODITE dataset comprises Global Telecommunication System-based data (the global summary of the day), data precompiled by other projects or organizations, and APHRODITE's own collection (Yatagai et al. 2012). The number of included rain gauges varies considerably over the domain (Yatagai et al. 2012). From all observational datasets the same time period as in the model runs was used for the validation, with the exception of the APHRODITE datasets which cover a slightly shorter time period up to 2007. Precipitation in the gridded datasets was averaged by area conservative remapping to the T213 grid.

Fig. 11 shows seasonal mean precipitation in ECHAM5 (T213L31) and in the GPCP dataset. In both seasons, the global pattern is well captured by ECHAM5. However, regional biases can be seen, such as an overestimation of monsoon precipitation over southeastern Asia in JJA. Large uncertainties in the simulation of the Asian summer monsoon have been shown by Hasson et al. (2013) for CMIP3-GCMs. Precipitation over parts of the oceans in both seasons is also too high. Over the western Asian continent and Australia in DJF, precipitation is underestimated by ECHAM5. These biases are consistent with the validation of the hydrological cycle in ECHAM5 by Hagemann et al. (2006).

In Fig. 12 and 13, RL20S of daily precipitation as simulated by ECHAM5 at T213L31 resolution and different high resolution observational gridded datasets are provided over the USA, Europe, Russia, the Middle East and southeastern Asia for DJF and JJA, respectively. In Tab. 3, the root mean squared errors of the spatial mean of RL20S over these analyzed regions of the ECHAM5 model at T213L31 resolution are displayed. The pattern of RL20S in the USA (panels a - b) is generally well captured by ECHAM5 at T213L31 resolution. The major deficiencies are a wet bias in the east in DJF and too dry regions in JJA in Florida and north of the Gulf of Mexico. The latter is in accordance with Wehner et al. (2010), who suggested that this high resolution is still too coarse to capture precipitation intensities that are related to tropical cyclones which might not be resolved. Over Europe (panels c - d), the pattern of RL20S is well captured by the ECHAM5 model compared to the E-OBS dataset. RL20S in mountainous regions (e.g., the Alps) are overestimated. In JJA, some regions are slightly too wet, such as eastern Europe. Yet, rain gauge density in the E-OBS dataset is sparsest in this region (Haylock et al. 2008), and hence, extreme precipitation might be underrepresented in the E-OBS dataset, especially in summer when many heavy rainfall events are caused by small scale convective events. The patterns of RL20S over Russia (panels e - g) in ECHAM5 and in the APHRODITE dataset are similar, but the model is slightly too wet, especially in eastern Russia, in JJA. Again, the sparse density of rain gauges in eastern Russia (Yatagai et al. 2012) might contribute to this difference. In the Middle East (panels g - h), the RL20S pattern around the Black Sea is reasonably captured. However, a wet bias in DJF as well as in JJA can be identified, which is particularly pronounced in the southwest of the Arabian peninsula in JJA and in the Iranian plateau in DJF. Although the rain gauge density in the APHRODITE dataset over the Arabian peninsula is quite sparse as well (Yatagai et al. 2012), this wet region in the southwest of the Arabian peninsula with high RL20S appears to be mainly due to a bias in the model, as in the observations no evidence for this wet region is visible. Panels i - k show patterns of RL20S over southeastern Asia in ECHAM5 and the APHRODITE dataset. Many features of the RL20S pattern are captured by the model. However, this region exhibits the largest deficiencies of the analyzed regions which is in accordance with the wet bias in the summer monsoon that is also visible in seasonal mean precipitation totals (see Fig. 11). The Himalayas are too wet in DJF as well as in JJA, of which no considerable part can be attributed to the rain gauge density as this region is well covered with rain gauges (Yatagai et al. 2012). A wet bias over India can be identified in the monsoon season - with India being well covered with rain gauges as well. Heavy precipitation associated with the summer monsoon is not well captured, which is a general problem in current GCMs (Hasson et al. 2013). This is underlined by the high RMSE for southeastern Asia in JJA (42.5 mm d^{-1}; see also Tab. 3), the RMSEs in all other regions are considerably lower.

Summarized, the ECHAM5 model at T213L31 resolution well represents the large scale pattern of seasonal mean precipitation, as well as many features of the regional spatial distribution of RL20S. In most regions, the range of RL20S is well captured, but over parts of southeastern Asia (e.g., the monsoon region) and in mountainous regions (e.g., Himalayas, Sierra Nevada, Alps, Iranian plateau), RL20S is overestimated by a factor of two. This validation of RL20S is limited by the availability of high quality observational datasets with suitable rain gauge density. Generally, it is difficult to produce reliable gridded precipitation datasets for the analysis of extremes due to spatial and temporal inhomogeneity of precipitation - especially of precipitation extremes (Teegavarapu 2012).

References

Adler, R., and Coauthors, 2003: The version-2 global precipitation climatology project (GPCP) monthly precipitation analysis (1979-present). *Journal of Hydrometeorology*, **4**, 1147–1167.

Allan, R. P., and B. J. Soden, 2008: Atmospheric warming and the amplification of precipitation extremes. *Science (New York, N.Y.)*, **321 (5895)**, 1481–1484.

Arakawa, A., 2004: The cumulus parameterization problem: Past, present, and future. *Journal of Climate*, 2493–2525.

Chen, C.-T., and T. Knutson, 2008: On the verification and comparison of extreme rainfall indices from climate models. *Journal of Climate*, **21 (7)**, 1605–1621.

Coles, S., 2001: *An introduction to statistical modeling of extreme values*. Springer-Verlag, London, 208 pp.

Duffy, P. B., B. Govindasamy, J. P. Iorio, J. Milovich, K. R. Sperber, K. E. Taylor, M. F. Wehner, and S. L. Thompson, 2003: High-resolution simulations of global climate, part 1: present climate. *Climate Dynamics*, **21 (5-6)**, 371–390.

Efron, B., and R. Tibshirani, 1993: *An introduction to the bootstrap*. Chapman and Hall, 436 pp.

Faranda, D., J. M. Freitas, V. Lucarini, G. Turchetti, and S. Vaienti, 2013: Extreme value statistics for dynamical systems with noise. *Nonlinearity*, **26 (9)**, 2597–2622.

Faranda, D., V. Lucarini, G. Turchetti, and S. Vaienti, 2011: Numerical convergence of the block-maxima approach to the generalized extreme value distribution. *Journal of Statistical Physics*, **145 (5)**, 1156–1180.

Fisher, R., and L. Tippett, 1928: On the estimation of the frequency distributions of the largest or smallest member of a sample. *Proceedings of the Cambridge Philosophical Society*, **24**, 180–190.

Flato, G., and Coauthors, 2013: Evaluation of Climate Models. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. Midgley, Eds., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, chap. 9, 741–866.

Gnedenko, B., 1943: Sur la distribution limite du terme maximum d'une série aléatoire. *Ann. Math.*, **44 (3)**, 423–453.

Grumbine, R., 1996: Automated passive microwave sea ice concentration analysis at NCEP. *NOAA Tech. Note*, **120**.

Hagemann, S., K. Arpe, and E. Roeckner, 2006: Evaluation of the hydrological cycle in the ECHAM5 model. *Journal of Climate*, **19 (16)**, 3810–3827.

Hasson, S., V. Lucarini, and S. Pascale, 2013: Hydrological cycle over south and southeast Asian river basins as simulated by PCMDI/CMIP3 experiments. *Earth System Dynamics Discussions*, **4 (1)**, 109–177.

Haylock, M. R., N. Hofstra, a. M. G. Klein Tank, E. J. Klok, P. D. Jones, and M. New, 2008: A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. *Journal of Geophysical Research*, **113 (D20119)**.

Higgins, R. W., W. Shi, E. Yarosh, and R. Joyce, 2000: Improved US precipitation quality control system and analysis. *NCEP/Climate Prediction Center Atlas No. 7, National Centers for Environmental Prediction, Climate Prediction Center, Camp Springs, Maryland. Available at http://www.cpc.ncep.noaa.gov/research.*

Hosking, J., J. Wallis, and E. Wood, 1985: Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. *Technometrics*, **27 (3)**, 251–261.

Iorio, J., P. Duffy, B. Govindasamy, S. Thompson, M. Khairoutdinov, and D. Randall, 2004: Effects of model resolution and subgrid-scale physics on the simulation of precipitation in the continental United States. *Climate Dynamics*, **23 (3-4)**, 243–258.

Jung, T., S. K. Gulev, I. Rudeva, and V. Soloviov, 2006: Sensitivity of extratropical cyclone characteristics to horizontal resolution in the ECMWF model. *Quarterly Journal of the Royal Meteorological Society*, **132 (619)**, 1839–1857.

Kopparla, P., E. M. Fischer, C. Hannay, and R. Knutti, 2013: Improved simulation of extreme precipitation in a high-resolution atmosphere model. *Geophysical Research Letters*, **40 (21)**, 5803–5808.

Le Treut, H., U. Cubasch, and M. Allen, 2007: Historical Overview of Climate Change Science. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, S. Solomon, D. Qin, M. Manning, M. Marquis, K. B. Averyt, M. Tignor, H. L. Miller, and Z. Chen, Eds., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, chap. 1, 93–128.

Leung, L., T. Ringler, W. Collins, M. Taylor, and M. Ashfaq, 2013: A hierarchical evaluation of regional climate simulations. *Eos, Transactions, American Geophysical Union*, **94 (34)**, 297–298.

Li, F., W. D. Collins, M. F. Wehner, D. L. Williamson, J. G. Olson, and C. Algieri, 2011: Impact of horizontal resolution on simulation of precipitation extremes in an aqua-planet version of Community Atmospheric Model (CAM3). *Tellus A*, **63 (5)**, 884–892.

Liepert, B. G., and M. Previdi, 2012: Inter-model variability and biases of the global water cycle in CMIP3 coupled climate models. *Environmental Research Letters*, **7 (1)**, 014 006.

Lovejoy, S., and D. Schertzer, 1995: Multifractals and rain. *New uncertainty concepts in hydrology and water resources*, Z. Kundzewicz, Ed., Cambridge University Press, Cambridge, UK, chap. III.2, 61–103.

Lucarini, V., and F. Ragone, 2011: Energetics of climate models: Net energy balance and meridional enthalpy transport. *Reviews of Geophysics*, **49**.

Maraun, D., and Coauthors, 2010: Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user. *Reviews of Geophysics*, **48 (3)**.

Osborn, T., and M. Hulme, 1997: Development of a relationship between station and grid-box rainday frequencies for climate model evaluation. *Journal of Climate*, **10 (8)**, 1885–1908.

Palmer, T. N., 2013: Climate extremes and the role of dynamics. *Proceedings of the National Academy of Sciences of the United States of America*, **110 (14)**, 5281–5282.

Petoukhov, V., S. Rahmstorf, S. Petri, and H. Schellnhuber, 2013: Quasiresonant amplification of planetary waves and recent Northern Hemisphere weather extremes. *Proceedings of the National Academy of Sciences of the United States of America*, **110 (14)**, 5336–5341.

Pope, V., and R. Stratton, 2002: The processes governing horizontal resolution sensitivity in a climate model. *Climate Dynamics*, **19 (3-4)**, 211–236.

Prein, A. F., G. J. Holland, R. M. Rasmussen, J. Done, K. Ikeda, M. P. Clark, and C. H. Liu, 2013: Importance of regional climate model grid spacing for the simulation of heavy precipitation in the Colorado Headwaters. *Journal of Climate*, **26 (13)**, 4848–4857.

R Development Core Team, 2011: *R: A Language and Environment for Statistical Computing*. Vienna, Austria, R Foundation for Statistical Computing.

Randall, D., and Coauthors, 2007: Climate models and their evaluation. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. Averyt, M. Tignor, and H. Miller, Eds., Cambridge University Press Cambridge, UK and New York, Cambridge, United Kingdom and New York, NY, USA, chap. 8, 589–662.

Rauscher, S., T. Ringler, W. Skamarock, and A. Mirin, 2013: Exploring a Global Multiresolution Modeling Approach Using Aquaplanet Simulations*. *Journal of Climate*, **26 (8)**, 2432–2452.

Reynolds, R. W., T. M. Smith, C. Liu, D. B. Chelton, K. S. Casey, and M. G. Schlax, 2007: Daily high-resolution-blended analyses for sea surface temperature. *Journal of Climate*, **20 (22)**, 5473–5496.

Roeckner, E., and Coauthors, 2003: The atmospheric general circulation model ECHAM5. Part I: Model description. *Rep. 349*, Vol. No. 349, Max-Planck-Institute for Meteorology, Hamburg, Germany.

Roeckner, E., and Coauthors, 2004: The atmospheric general circulation model ECHAM5 - Part II: Sensitivity of simulated climate to horizontal and vertical resolution. *Rep. 354*, Vol. No. 354, Max-Planck-Institute for Meteorology, Hamburg, Germany.

Roeckner, E., and Coauthors, 2006: Sensitivity of simulated climate to horizontal and vertical resolution in the ECHAM5 atmosphere model. *Journal of Climate*, **19**, 3771–3791.

Rummukainen, M., 2010: State-of-the-art with regional climate models. *Wiley Interdisciplinary Reviews: Climate Change*, **1**, 82–96.

Rust, H. W., 2009: The effect of long-range dependence on modelling extremes with the generalised extreme value distribution. *The European Physical Journal Special Topics*, **174 (1)**, 91–97.

Seneviratne, S. I., and Coauthors, 2012: Changes in climate extremes and their impacts on the natural physical environment. *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change (IPCC).*, C. Field, V. Barros, T. Stocker, D. Qin, D. Dokken, K. Ebi, M. Mastrandrea, K. Mach, G.-K. Plattner, S. Allen, M. Tignor, and P. Midgley, Eds., Cambridge University Press, Cambridge, UK, and New York, NY, USA, chap. 3, 109–230.

Smith, I., A. Moise, J. Katzfey, K. Nguyen, and R. Colman, 2013: Regional-scale rainfall projections: Simulations for the New Guinea region using the CCAM model. *Journal of Geophysical Research: Atmospheres*, **118**, 1271–1280.

Sun, Y., S. Solomon, A. Dai, and R. Portmann, 2006: How often does it rain? *Journal of Climate*, **19 (6)**, 916–934.

Teegavarapu, R. S., 2012: *Floods in a changing climate*. Cambridge University Press, New York, 269 pp.

Wehner, M. F., R. L. Smith, G. Bala, and P. Duffy, 2010: The effect of horizontal resolution on simulation of very extreme US precipitation events in a global atmosphere model. *Climate Dynamics*, **34 (2-3)**, 241–247.

Williamson, D. L., 2008: Convergence of aqua-planet simulations with increasing resolution in the Community Atmospheric Model, version 3. *Tellus*, **60A (5)**, 848–862.

Wuertz, D., 2009: Package ‘ fExtremes ’ - Rmetrics - Extreme Financial Market Data.

Yatagai, A., K. Kamiguchi, O. Arakawa, A. Hamada, N. Yasutomi, and A. Kitoh, 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.

List of Tables

Table 1. List of horizontal and vertical resolutions of the ECHAM5 simulations used in this study. Horizontal resolution is given as spectral resolution and Gaussian transform grid resolution. Vertical resolution is given as the number of vertical levels.

Horizontal resolution		Vertical resolution
Spectral	Gaussian	
T213	0.56°×0.56°	L31
T159	0.75°×0.75°	L31
T106	1.13°×1.13°	L31
T63	1.88°×1.88°	L31/L19
T42	2.81°×2.81°	L31/L19
T31	3.75°×3.75°	L19

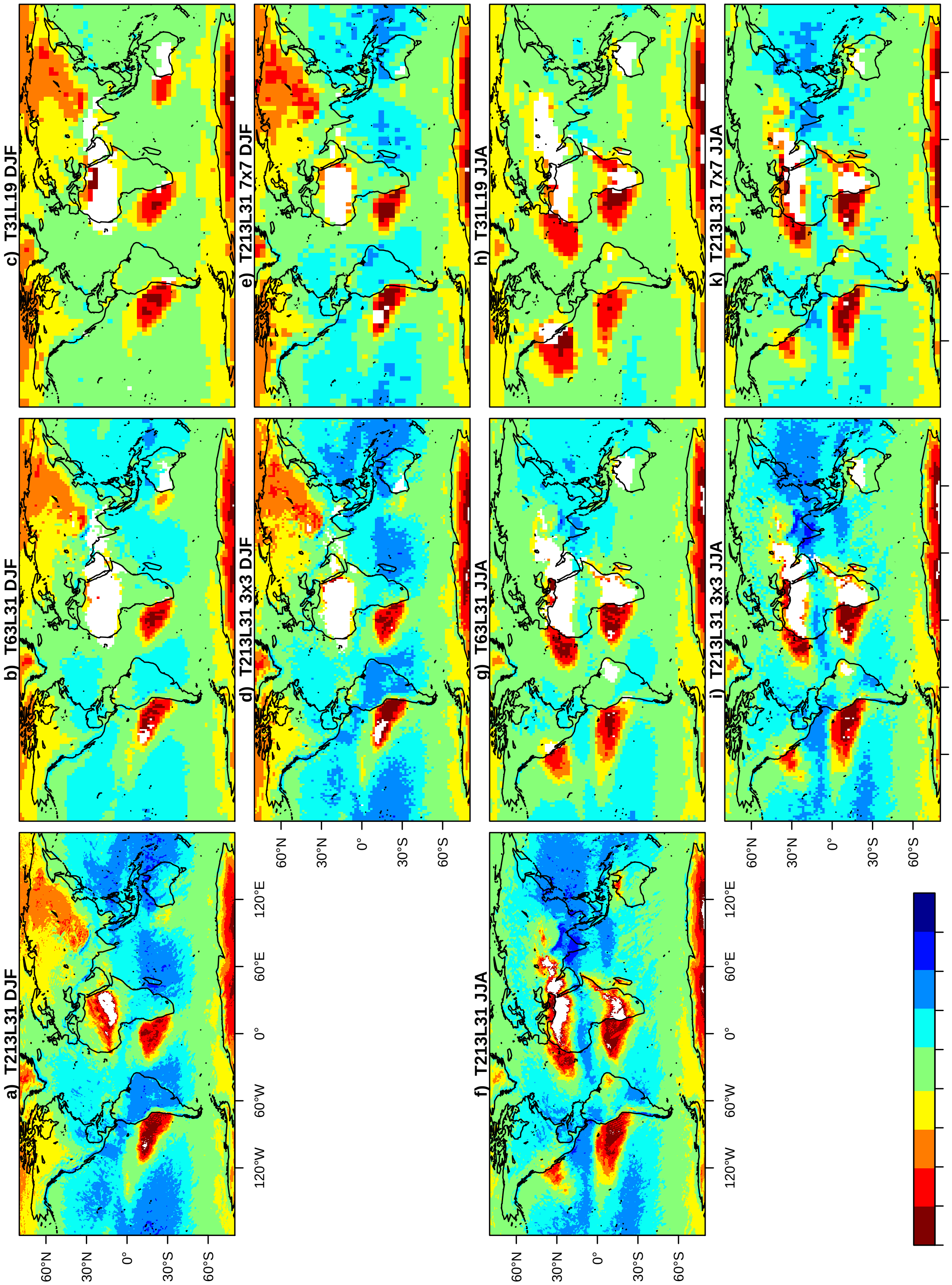
Table 2. Spatial averaging of the highest used ECHAM5 resolution T213L31: Number of averaged grid boxes and resulting Gaussian grid box size.

Spatial averaging	Gaussian grid box size
2×2	1.125°×1.125°
3×3	1.69°×1.69°
4×4	2.25°×2.25°
5×5	2.81°×2.81°
6×6	3.38°×3.38°
7×7	3.94°×3.94°

TABLE 3. Root mean squared error of simulated 20 season return levels (RL20S) [mm d^{-1}] in the highest used ECHAM5 resolution T213L31 validated by CPC, E-OBS (version 9) and APHRODITE gridded precipitation datasets.

	USA	Europe	Russia	Middle East	Monsoon Asia
DJF	14.02	6.59	7.19	9.74	15.48
JJA	4.87	8.08	26.60	6.47	42.54

List of Figures



a) T213L31 DJF
b) T63L31 DJF
c) T31L19 DJF
d) T213L31 3x3 DJF
e) T213L31 7x7 DJF
f) T213L31 JJA
g) T63L31 JJA
h) T31L19 JJA
i) T213L31 3x3 JJA
k) T213L31 7x7 JJA
60°N
30°N
0°
30°S
60°S
120°W
60°W
0°
60°E
120°E
1
2
5
10
20
50
100
200
500

FIG. 1. 20 season return level (RL20S) [mm d^{-1}] maps for (a - e) DJF and (f - k) JJA; logarithmic color scale, a - c and f - h: changing model resolution, d - e and i - k: averaged high resolution. White: seasonal maxima time series contain more than one zero value.

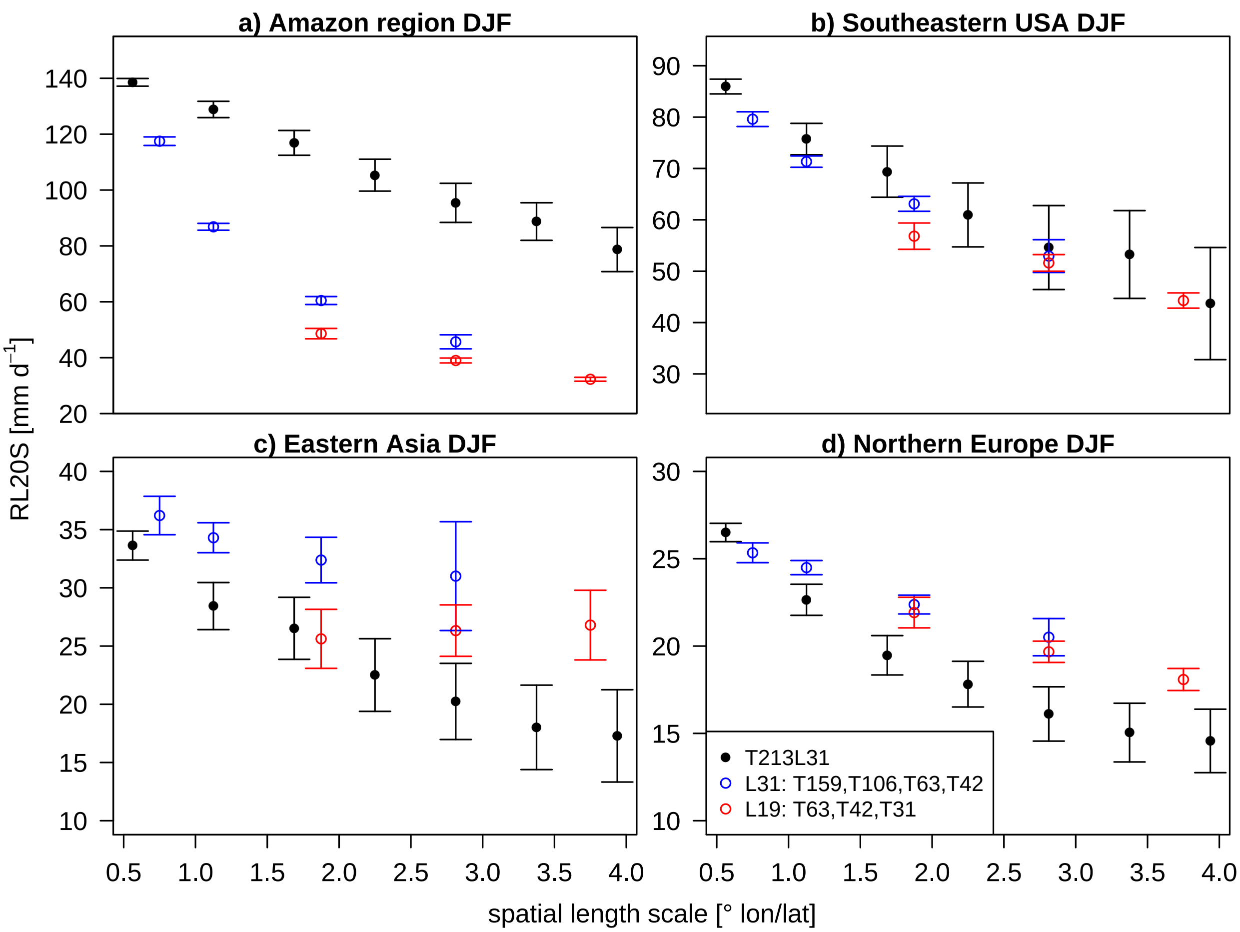


FIG. 2. Scaling behavior for example regions; area averages (with 95% confidence interval, as 1.96×area standard deviation) of 20 season return levels (RL20S). Black: averaged high resolution, blue: coarser horizontal resolutions in high vertical resolution, red: coarser horizontal resolutions in low vertical resolution.

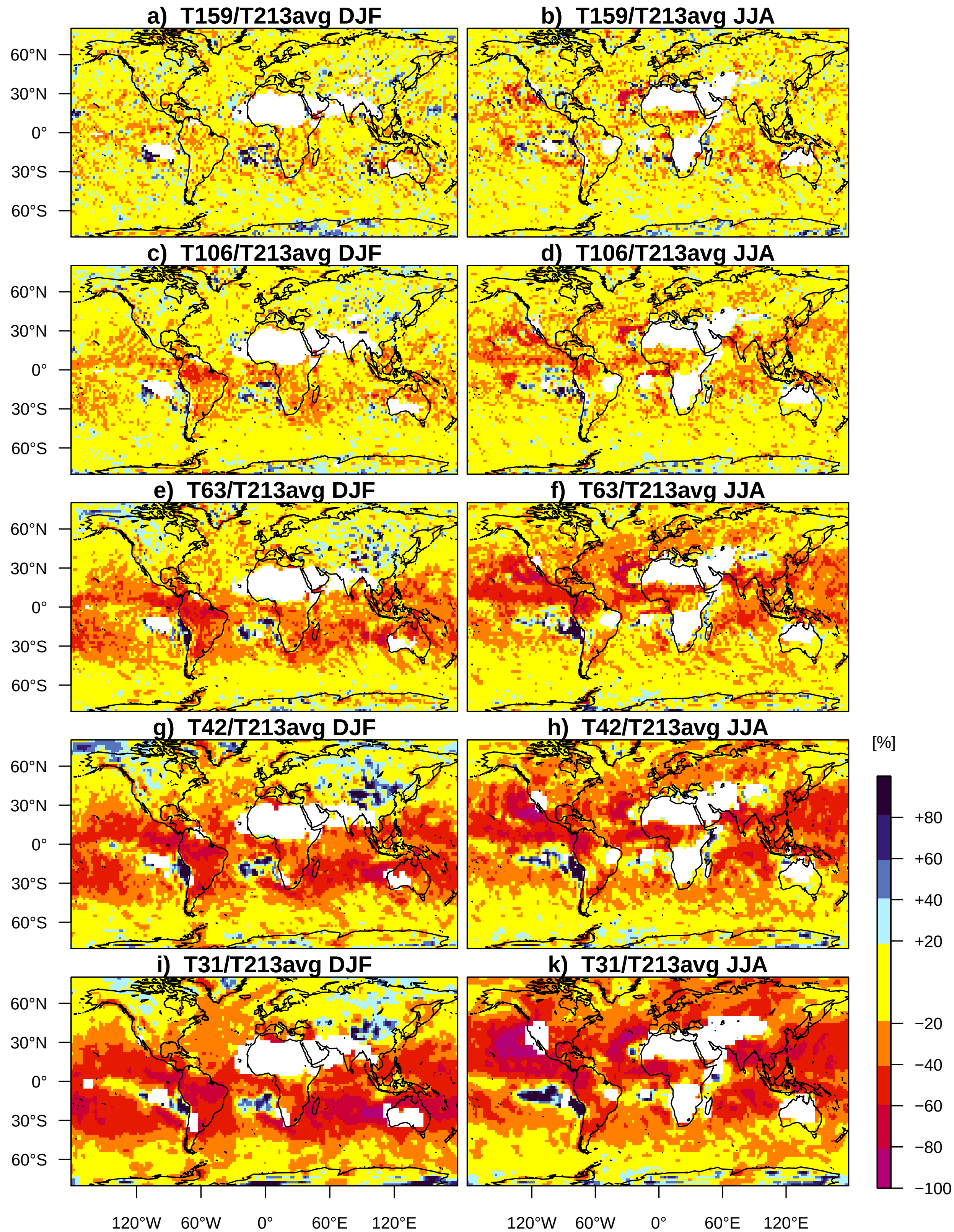

a) T159/T213avg DJF
b) T159/T213avg JJA
c) T106/T213avg DJF
d) T106/T213avg JJA
e) T63/T213avg DJF
f) T63/T213avg JJA
g) T42/T213avg DJF
h) T42/T213avg JJA
i) T31/T213avg DJF
k) T31/T213avg JJA
60°N
30°N
0°
30°S
60°S
120°W
60°W
0°
60°E
120°E
[%]
+80
+60
+40
+20
−20
−40
−60
−80
−100

Fig. 3. Ratios between 20 season return levels (RL20S) at coarser horizontal resolutions (for T63 and T42 the L31 simulations are shown) and RL20S at the respective averaged high resolution for DJF (left hand column) and JJA (right hand column). White: seasonal maxima time series contain more than one zero value. Before computing the ratios, RL20S in all resolutions were interpolated bilinearly to a T63 grid.

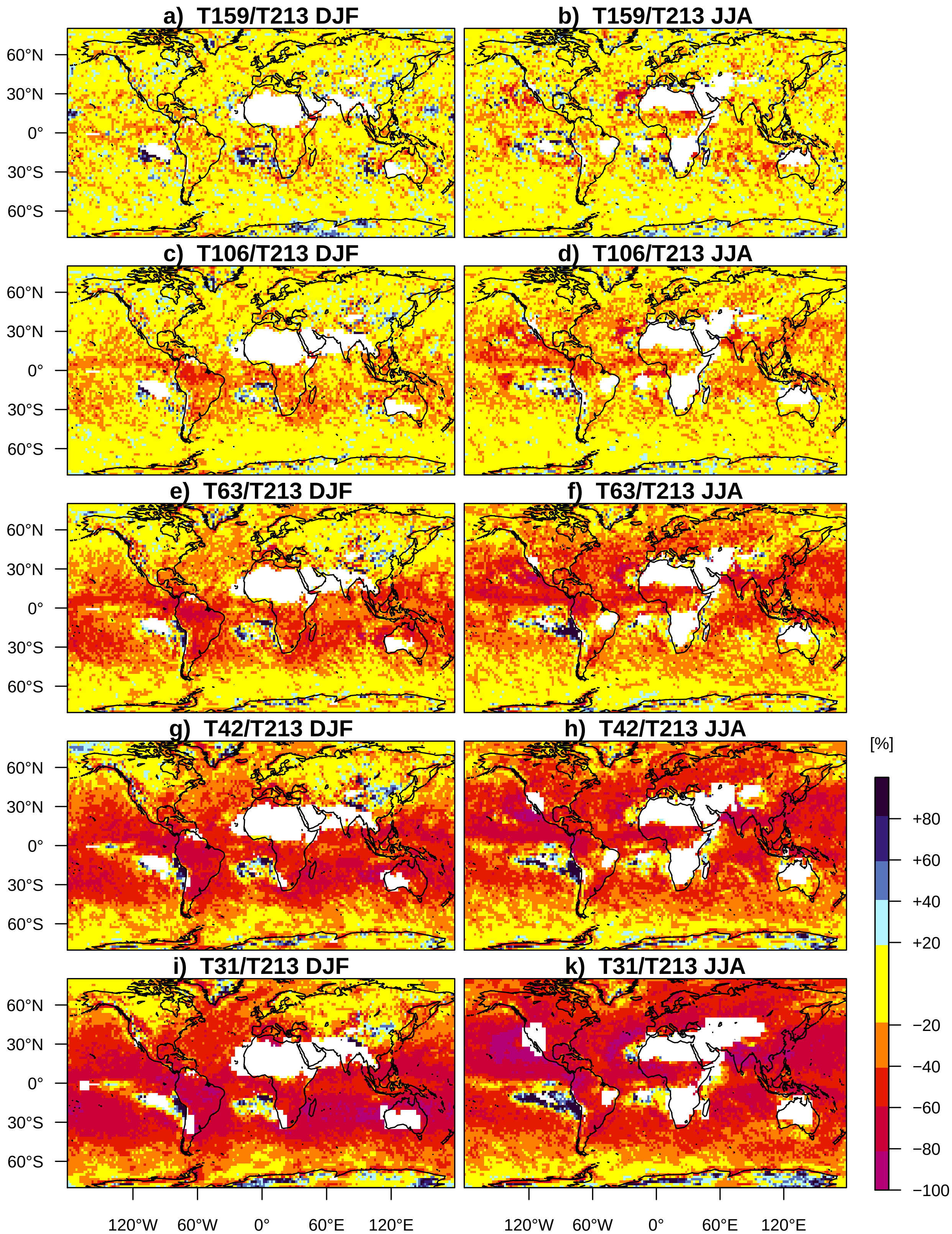

a) T159/T213 DJF
b) T159/T213 JJA
c) T106/T213 DJF
d) T106/T213 JJA
e) T63/T213 DJF
f) T63/T213 JJA
g) T42/T213 DJF
h) T42/T213 JJA
i) T31/T213 DJF
k) T31/T213 JJA
60°N
30°N
0°
30°S
60°S
120°W
60°W
0°
60°E
120°E
[%]
+80
+60
+40
+20
−20
−40
−60
−80
−100

FIG. 4. Ratios between 20 season return levels (RL20S) at coarser horizontal resolutions (for T63 and T42 the L31 simulations are shown) and RL20S at the highest resolution at its original resolution for DJF (left hand column) and JJA (right hand column). White: seasonal maxima time series contain more than one zero value. Before computing the ratios, RL20S at all resolutions were interpolated bilinearly to a T63 grid.

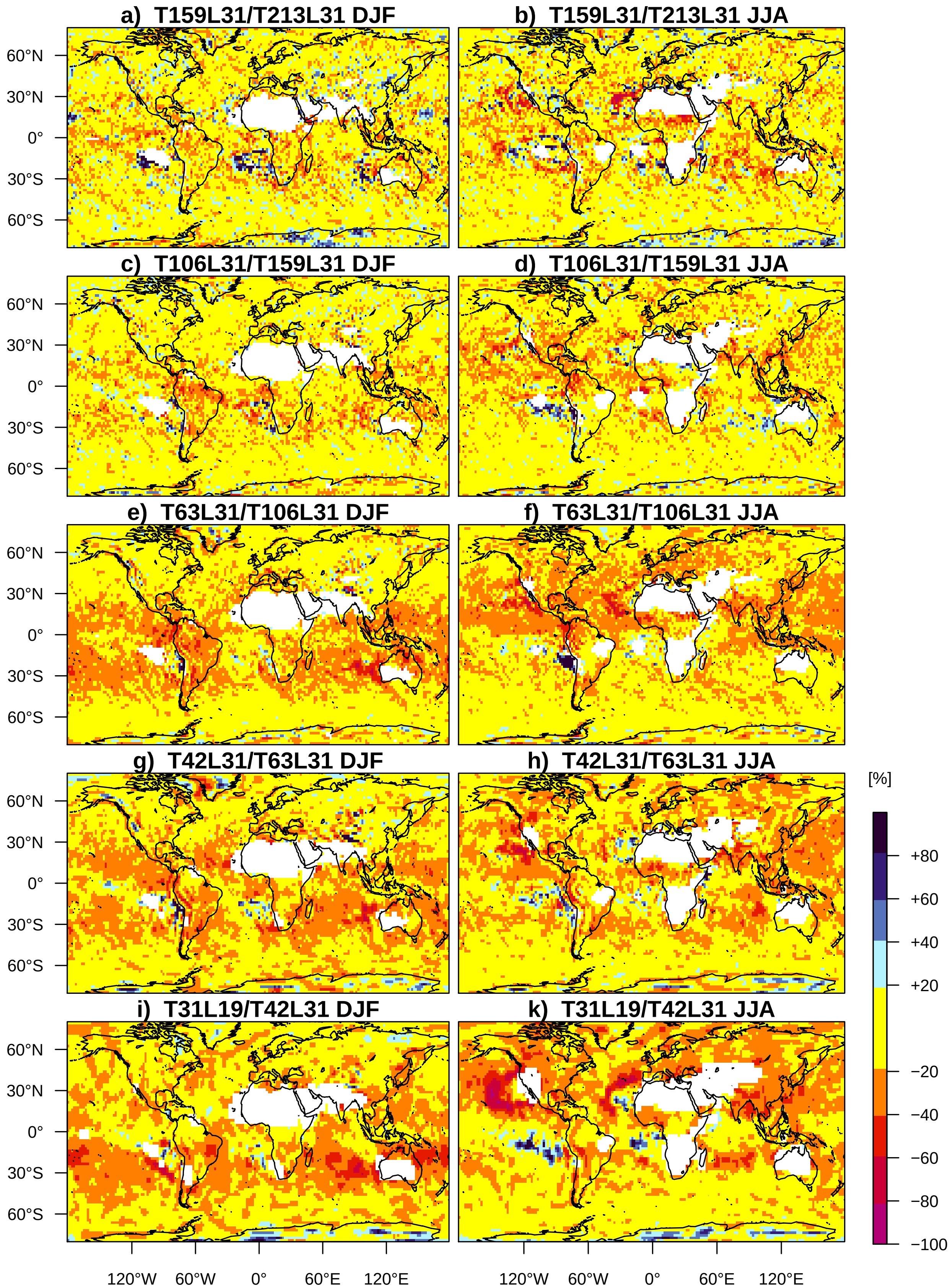

a) T159L31/T213L31 DJF
b) T159L31/T213L31 JJA
c) T106L31/T159L31 DJF
d) T106L31/T159L31 JJA
e) T63L31/T106L31 DJF
f) T63L31/T106L31 JJA
g) T42L31/T63L31 DJF
h) T42L31/T63L31 JJA
i) T31L19/T42L31 DJF
k) T31L19/T42L31 JJA
60°N
30°N
0°
30°S
60°S
120°W
60°W
0°
60°E
120°E
[%]
+80
+60
+40
+20
−20
−40
−60
−80
−100

Fig. 5. Ratios between 20 season return levels (RL20S) in consecutive horizontal resolutions (for T63 and T42, the L31 simulations are shown) at their original resolutions for DJF (left hand column) and JJA (right hand column). White: seasonal maxima time series contain more than one zero value. Before computing the ratios, RL20S in all resolutions were interpolated bilinearly to a T63 grid.

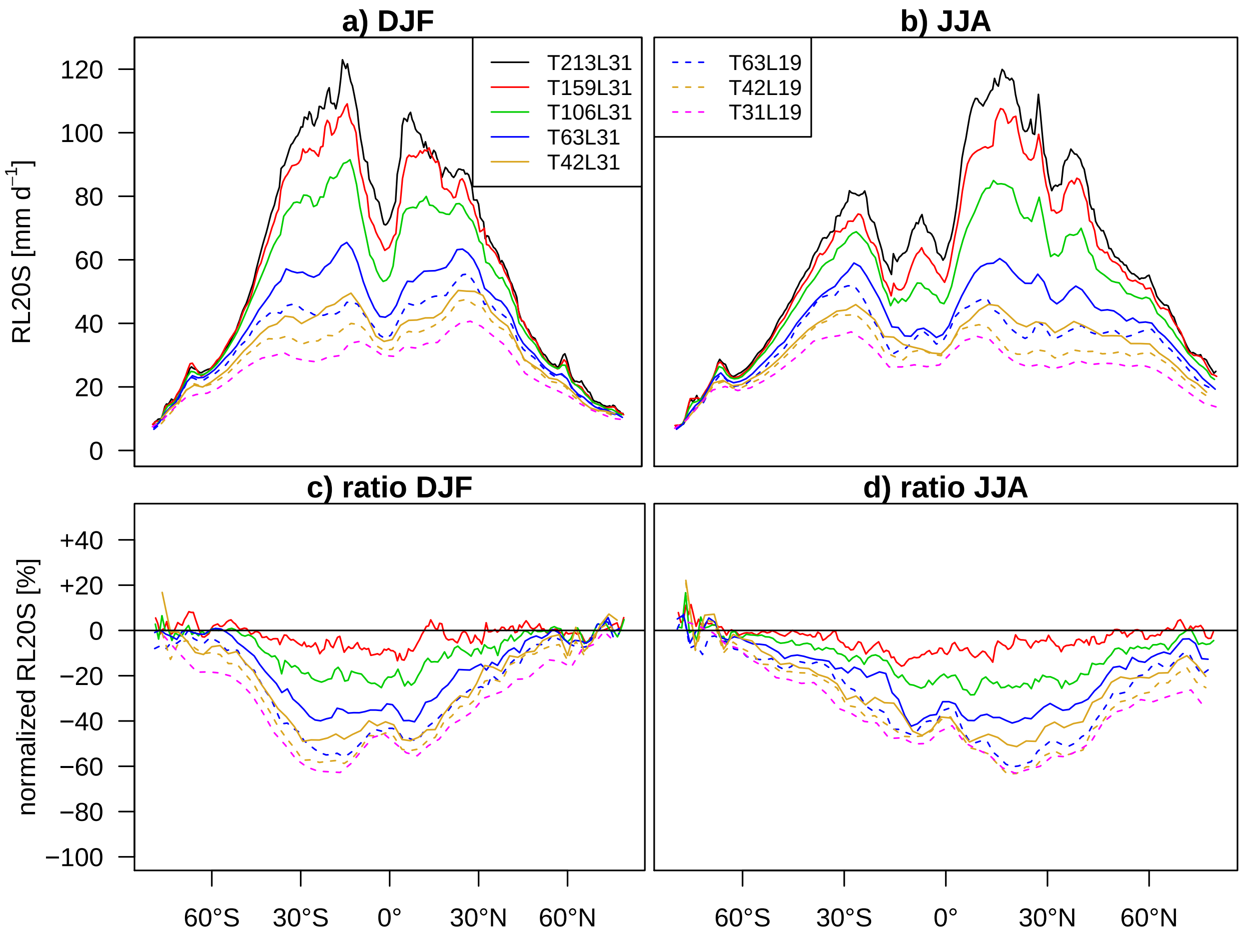


FIG. 6. Zonal means of 20 season return levels (RL20S); (a, b): different horizontal (solid lines) and vertical (dashed lines) resolutions, (c ,d): additionally normalized with the zonal mean of RL20S in the respective averaged high resolution. Grid boxes whose seasonal maxima time series contain more than one zero value in at least one resolution are excluded in all resolutions.

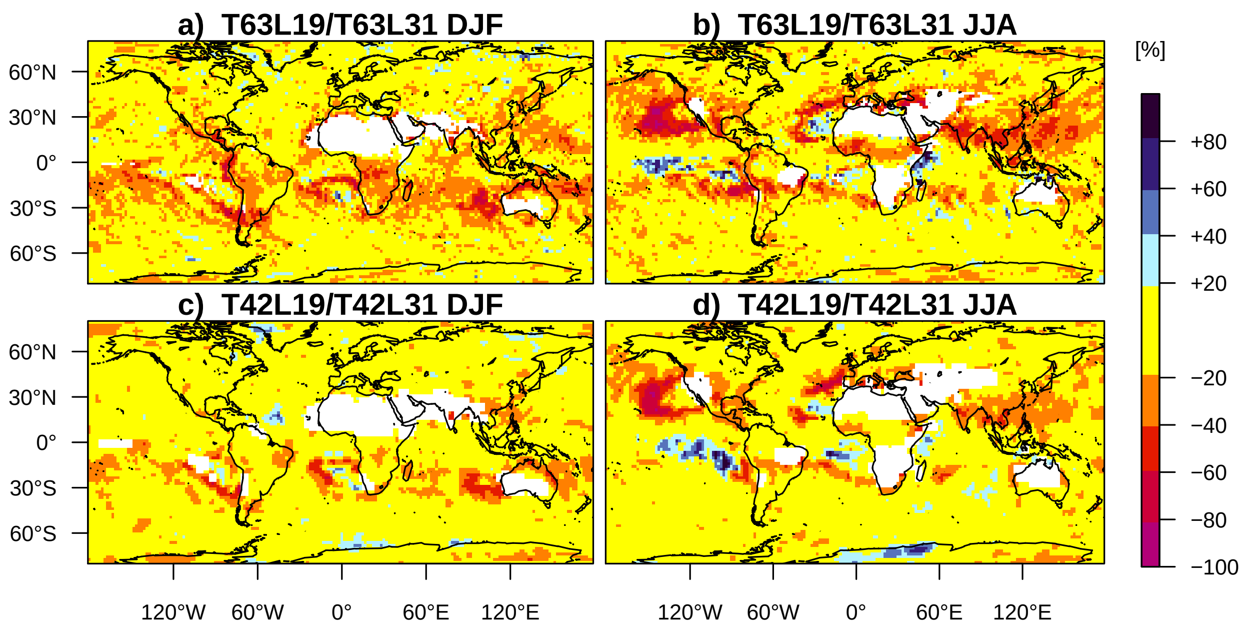


FIG. 7. Ratios of 20 season return levels (RL20S) between different vertical resolutions at the same horizontal resolution, for DJF (left hand column) and JJA (right hand column). White: seasonal maxima time series containing more than one zero value. Before computing the ratios, RL20S in all resolutions were interpolated bilinearly to a T63 grid.

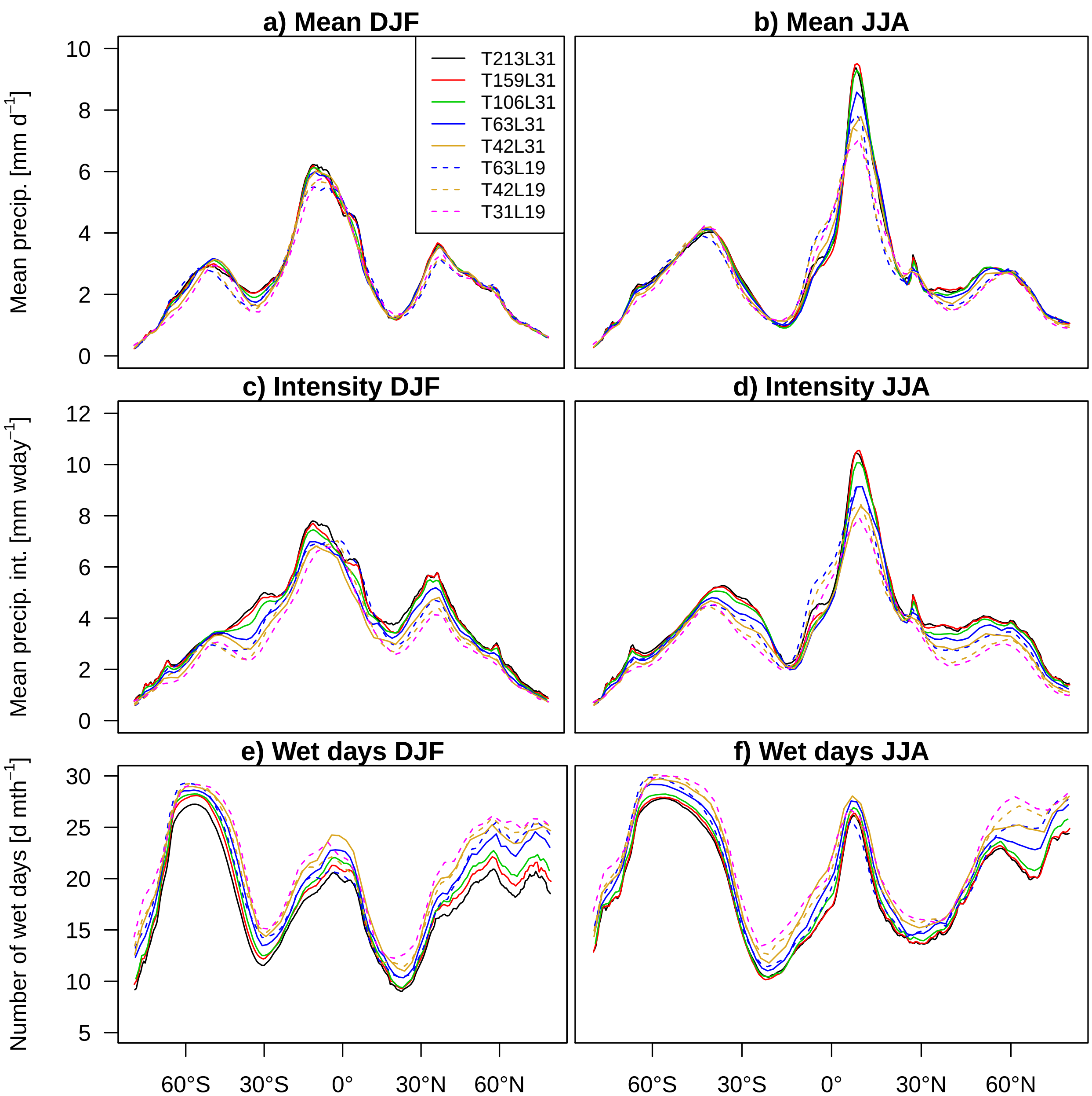


FIG. 8. Zonal means of (a, b) daily mean precipitation totals, (c, d) mean precipitation intensity (mean precipitation on wet days) and (e, f) the mean number of wet days per month (days with $\geq$ 0.1 mm precipitation) in different horizontal (solid lines) and vertical (dashed lines) resolutions for DJF (left hand column) and JJA (right hand column).

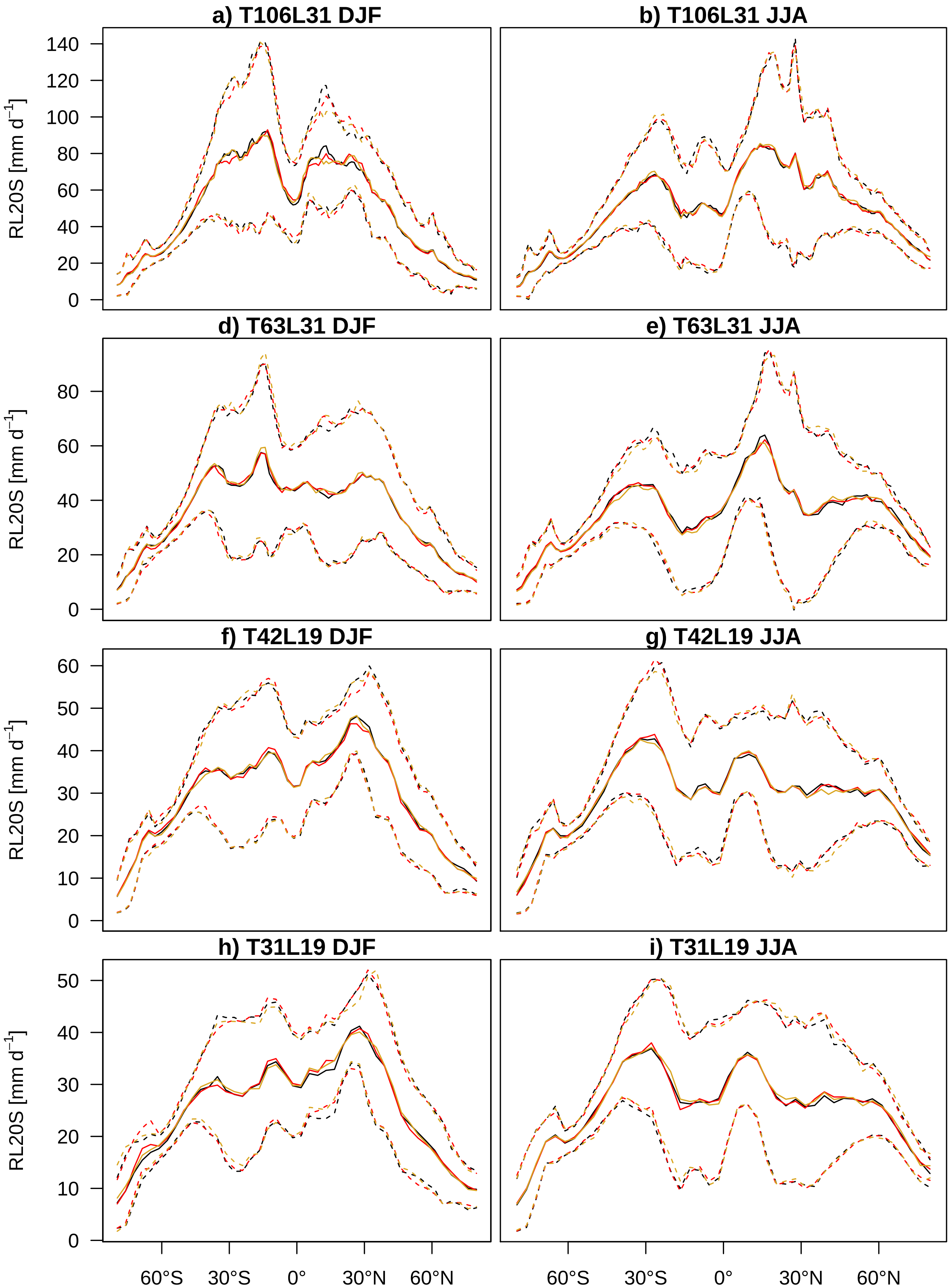

a) T106L31 DJF
b) T106L31 JJA
d) T63L31 DJF
e) T63L31 JJA
f) T42L19 DJF
g) T42L19 JJA
h) T31L19 DJF
i) T31L19 JJA
RL20S [mm d⁻¹]
60°S
30°S
0°
30°N
60°N

Fig. 9. Zonal means (solid lines) and zonal standard deviations (dashed lines) of 20 season return levels (RL20S) for three ensemble members with slightly different initial conditions in the resolutions T106L31, T63L31, T42L19 and T31L19 for DJF (left hand column) and JJA (right hand column).

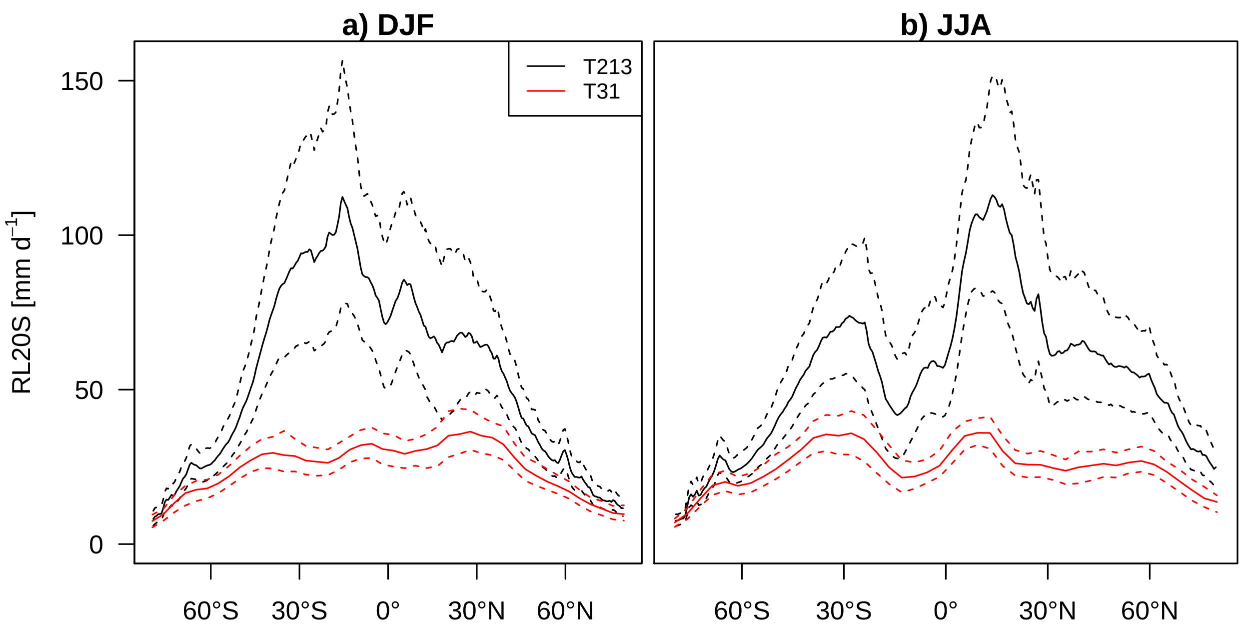


FIG. 10. Zonal means of 20 season return levels (RL20S) of this study (solid lines) and zonal means of 95% confidence intervals (dashed lines) for RL20S in DJF and JJA; Confidence intervals are computed with a parametric bootstrap method.

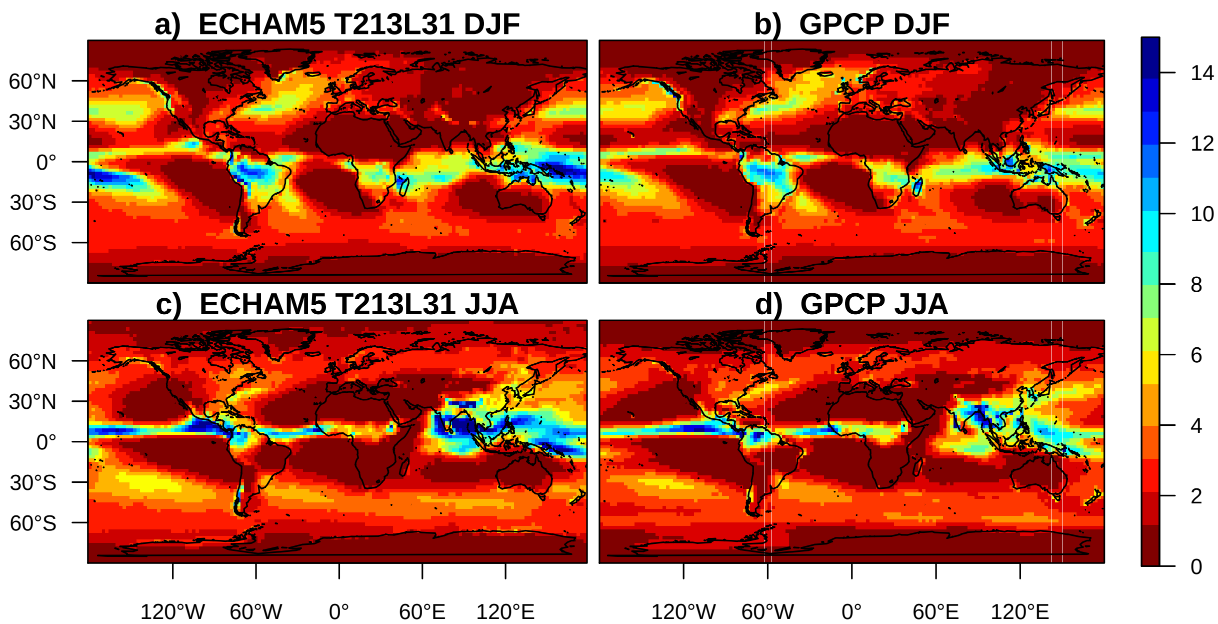


FIG. 11. Simulated (T213L31, left hand panels) and observed (GPCP, right hand panels) monthly mean precipitation totals [mm d^{-1}] in DJF (a, b) and JJA (c, d).

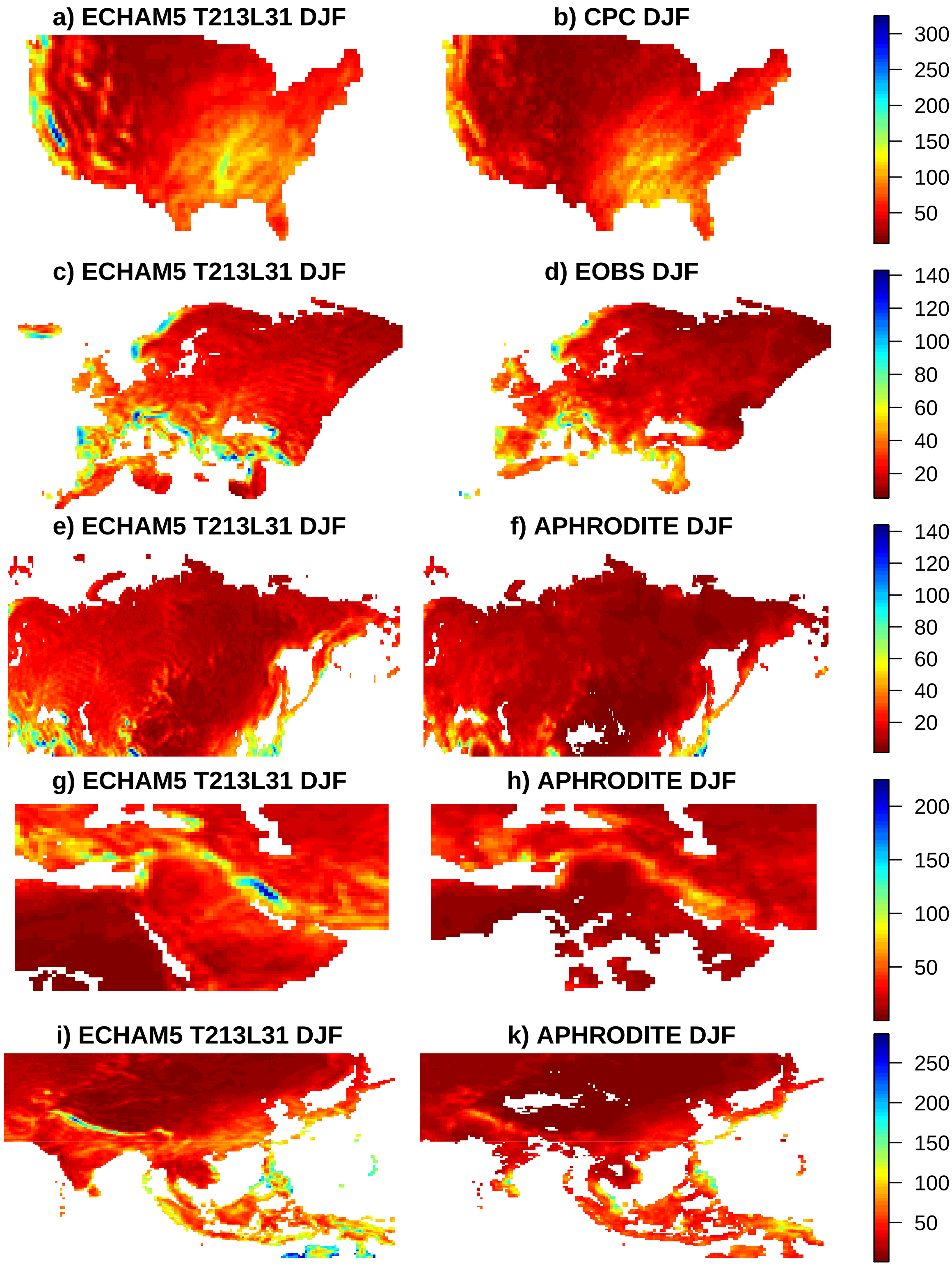

a) ECHAM5 T213L31 DJF
b) CPC DJF
300
250
200
150
100
50
c) ECHAM5 T213L31 DJF
d) EOBS DJF
140
120
100
80
60
40
20
e) ECHAM5 T213L31 DJF
f) APHRODITE DJF
140
120
100
80
60
40
20
g) ECHAM5 T213L31 DJF
h) APHRODITE DJF
200
150
100
50
i) ECHAM5 T213L31 DJF
k) APHRODITE DJF
250
200
150
100
50

FIG. 12. Simulated (T213L31, left hand panels) and observed (right hand panels) 20 season return levels (RL20S) [mm d^{-1}] in DJF; Observational datasets are b) CPC, d) E-OBS version 9, f) APHRODITE Russia, h) APHRODITE Middle East, k) APHRODITE Monsoon Asia. White: missing values in observational dataset or seasonal maxima time series contain more than one zero value.

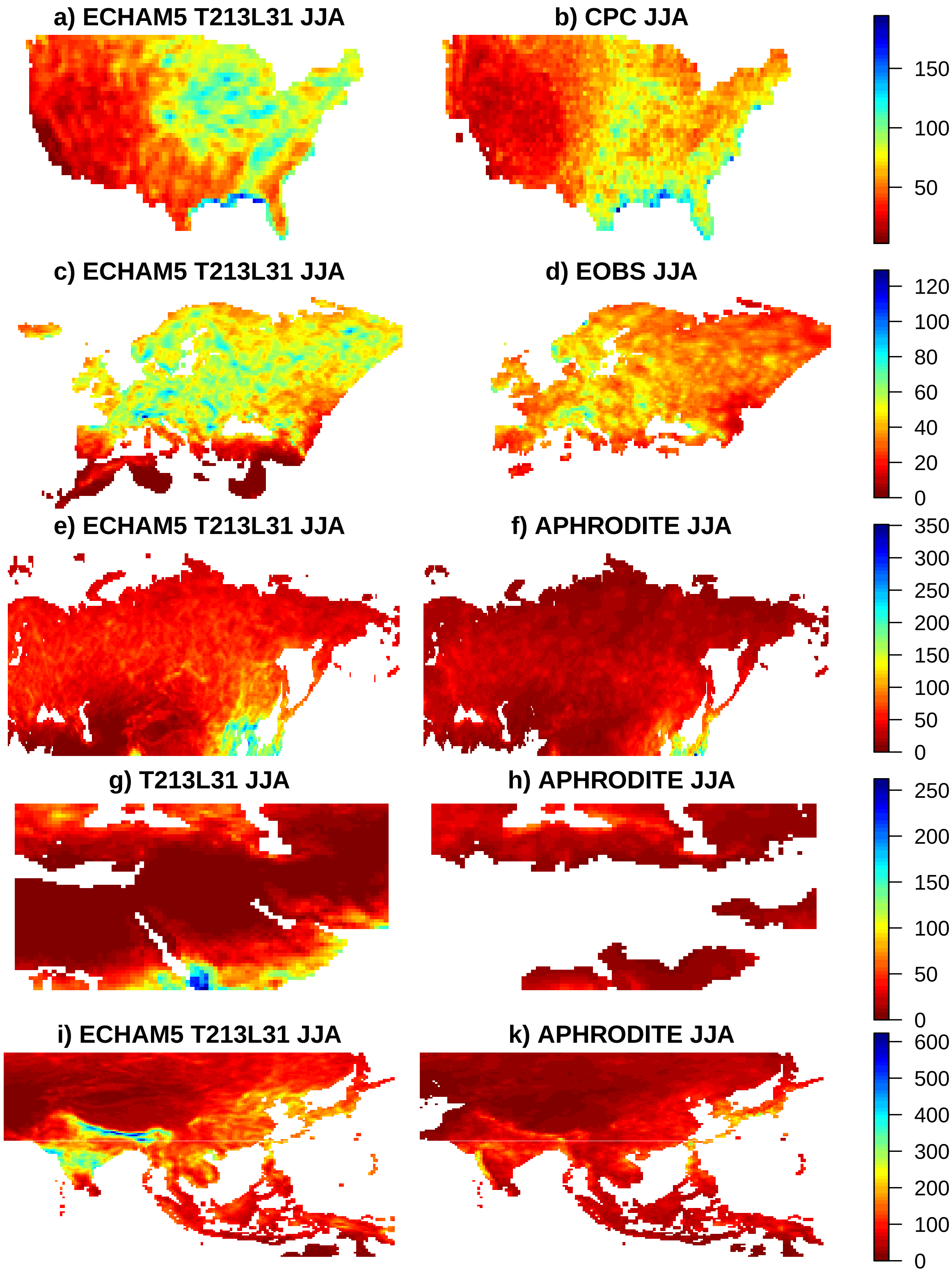

a) ECHAM5 T213L31 JJA
b) CPC JJA
150
100
50
c) ECHAM5 T213L31 JJA
d) EOBS JJA
120
100
80
60
40
20
0
e) ECHAM5 T213L31 JJA
f) APHRODITE JJA
350
300
250
200
150
100
50
0
g) T213L31 JJA
h) APHRODITE JJA
250
200
150
100
50
0
i) ECHAM5 T213L31 JJA
k) APHRODITE JJA
600
500
400
300
200
100
0

Fig. 13. Simulated (T213L31, left hand panels) and observed (right hand panels) 20 season return levels (RL20S) [mm d^{-1}] in JJA; Observational datasets are b) CPC, d) E-OBS version 9, f) APHRODITE Russia, h) APHRODITE Middle East, k) APHRODITE Monsoon Asia. White: missing value in observational dataset or seasonal maxima time series contain more than one zero value.