# Autocorrelation - A Simple Diagnostic for Tropical Precipitation in Global Kilometer-Scale Climate Models

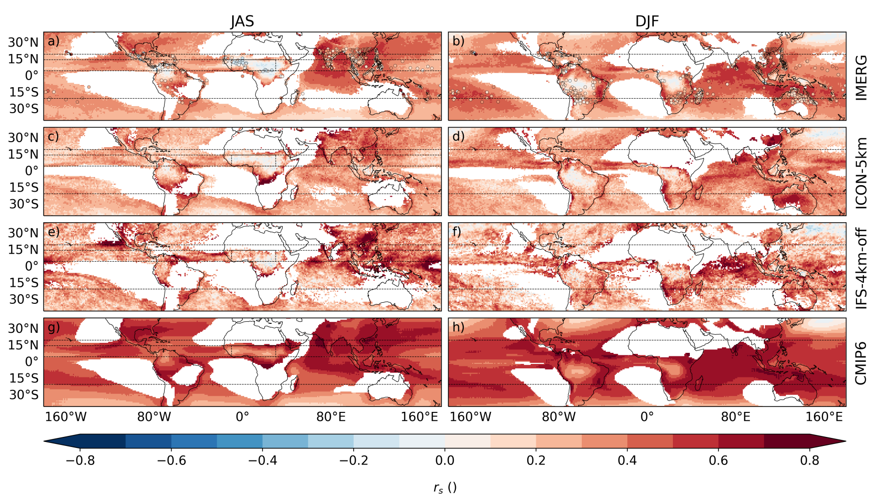
Dorian Spät[1], Michela Biasutti[2], David Schuhbauer[3], and Aiko Voigt[3]

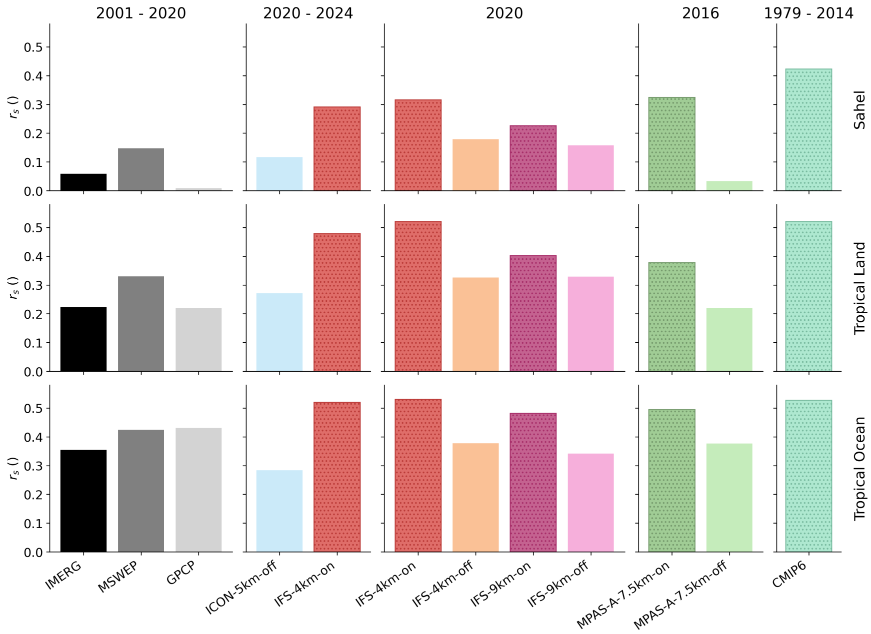
[1]University of Vienna
[2]Lamont-Doherty Earth Observatory
[3]Department of Meteorology and Geophysics, University of Vienna

March 05, 2024

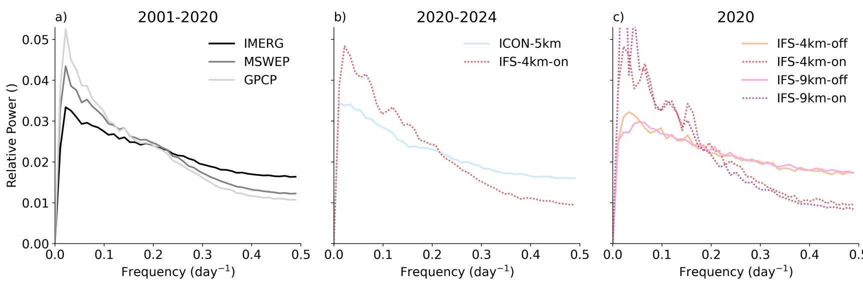
## Abstract

We propose the lag 1 autocorrelation of daily precipitation as a simple diagnostic of tropical precipitation in climate models. This metric generally has a relatively uniform distribution of positive values over the tropics. However, selected land regions are characterized by exceptionally low autocorrelation values. Low values correspond to the dominance of high-frequency variance in precipitation. Consistent with previous work, we show that CMIP6 climate models overestimate the autocorrelation. Global kilometer-scale models capture the observed autocorrelation pattern when deep convection is explicitly simulated. When a deep convection parameterization is used, the autocorrelation increases across the tropics, suggesting that land surface-atmosphere interactions are not responsible for the changes in precipitation variability. Furthermore, an accurate simulation of convectively coupled equatorial waves does not necessarily lead to a correct representation of the autocorrelation, and vice versa. This suggests other driving processes for the autocorrelation pattern.

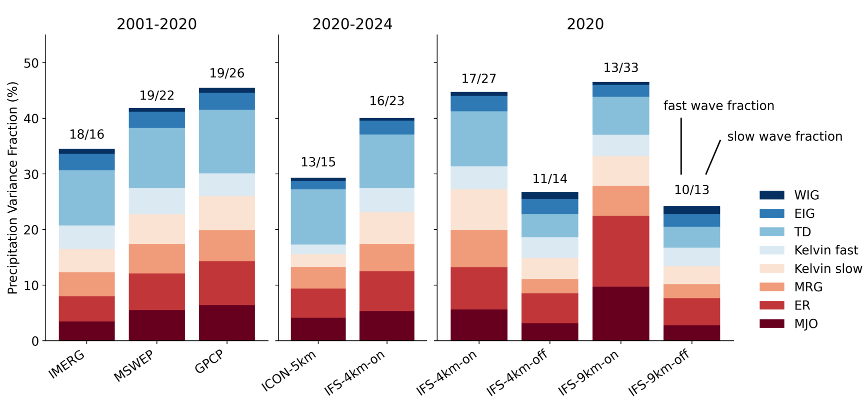




2001 - 2020
2020 - 2024
2020
2016
1979 - 2014
$r_s$ ()
Sahel
Tropical Land
Tropical Ocean
IMERG
MSWEP
GPCP
ICON-5km-off
IFS-4km-on
IFS-4km-on
IFS-4km-off
IFS-9km-on
IFS-9km-off
MPAS-A-7.5km-on
MPAS-A-7.5km-off
CMIP6

a)
2001-2020
b)
2020-2024
c)
2020
Relative Power ()
Frequency (day$^{-1}$)
IMERG
MSWEP
GPCP
ICON-5km
IFS-4km-on
IFS-4km-off
IFS-4km-on
IFS-9km-off
IFS-9km-on

2001-2020
2020-2024
2020
Precipitation Variance Fraction (%)
18/16
19/22
19/26
13/15
16/23
17/27
11/14
13/33
10/13
fast wave fraction
slow wave fraction
WIG
EIG
TD
Kelvin fast
Kelvin slow
MRG
ER
MJO
IMERG
MSWEP
GPCP
ICON-5km
IFS-4km-on
IFS-4km-on
IFS-4km-off
IFS-9km-on
IFS-9km-off

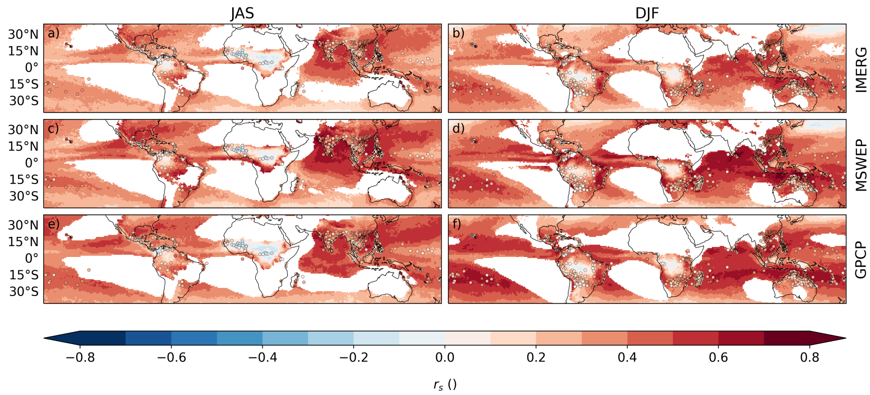
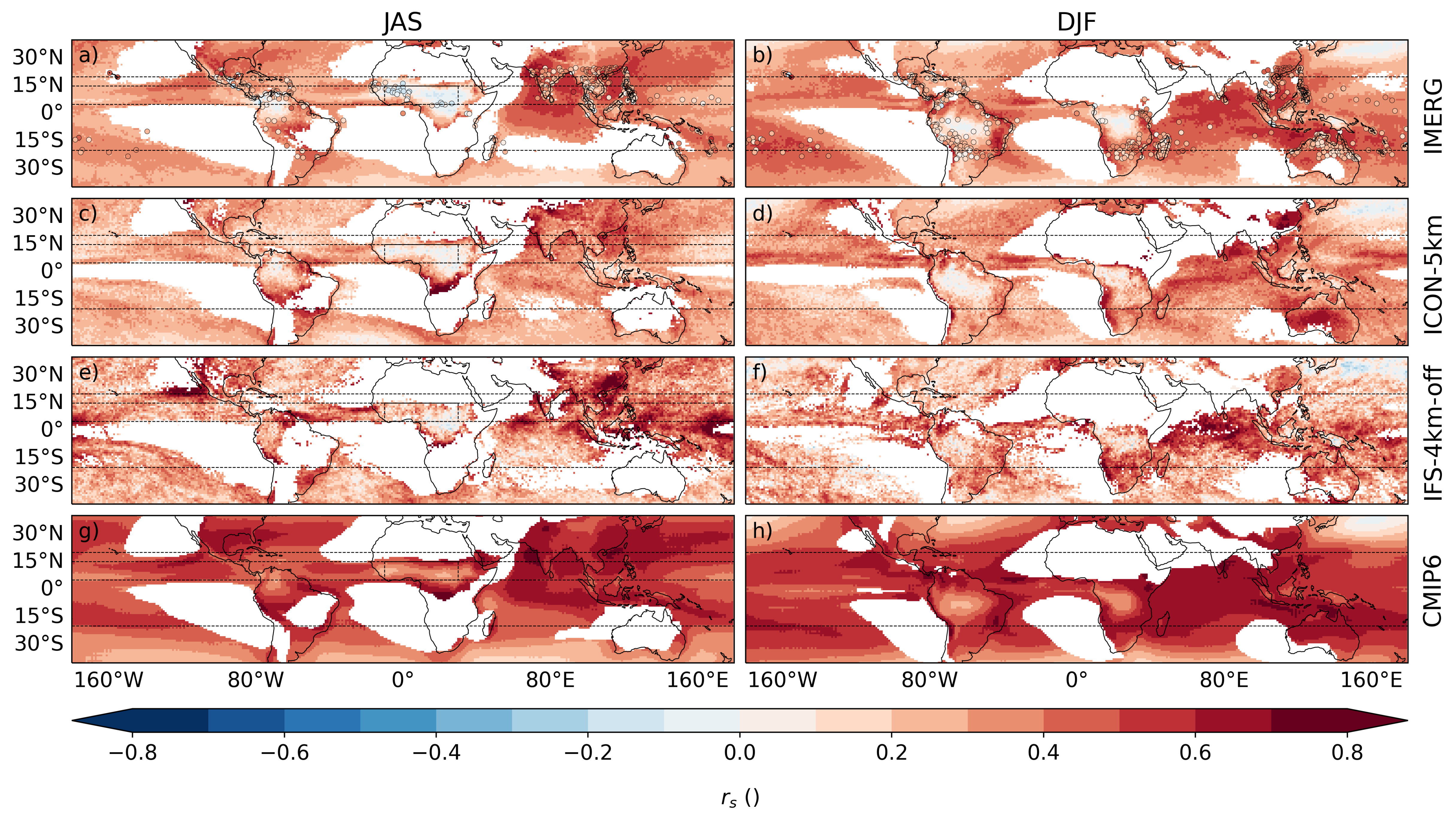
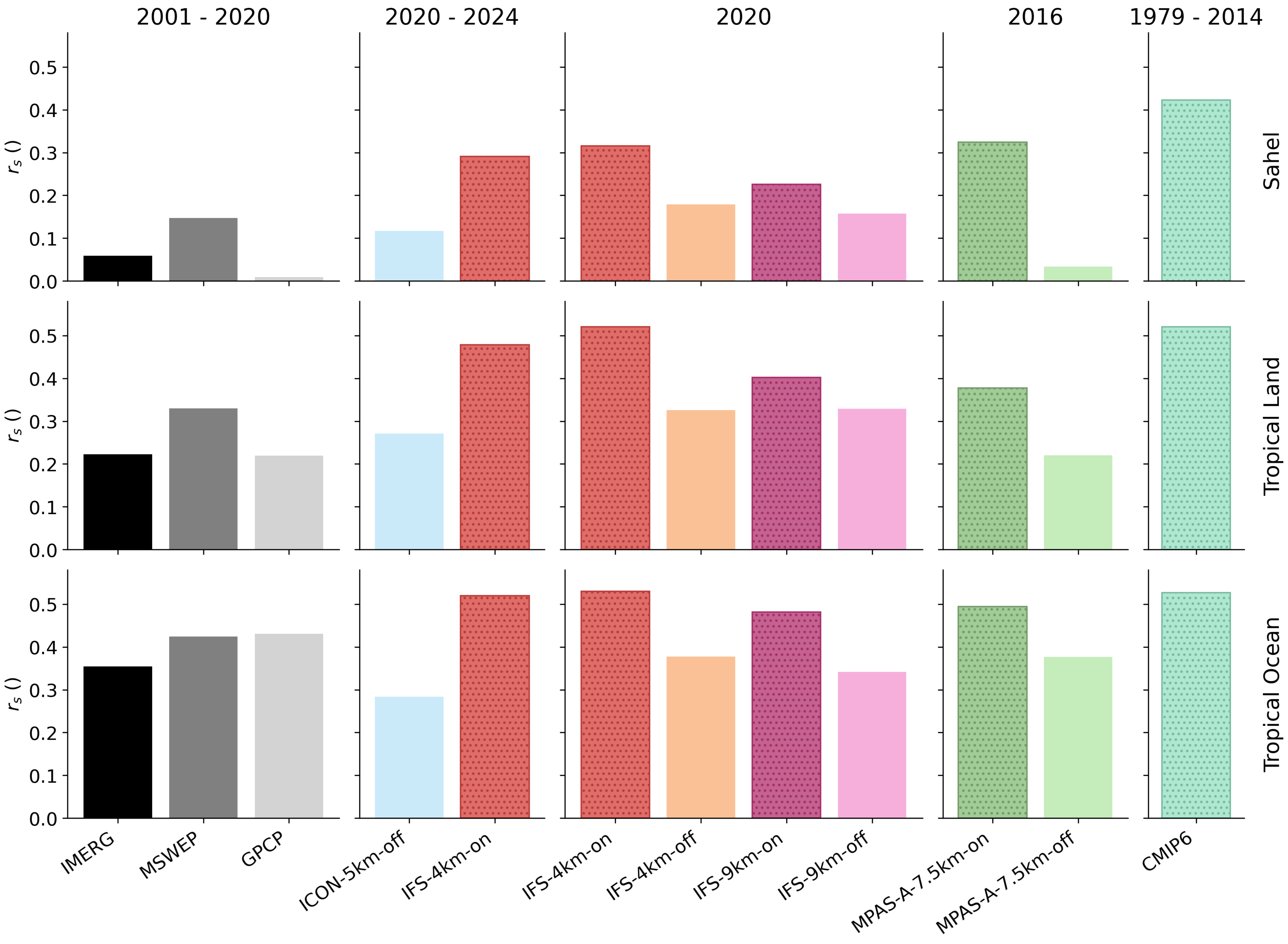

JAS
DJF
30°N
15°N
0°
15°S
30°S
a)
b)
c)
d)
e)
f)
IMERG
MSWEP
GPCP
−0.8
−0.6
−0.4
−0.2
0.0
0.2
0.4
0.6
0.8
$r_s$ ()

**Figure.**



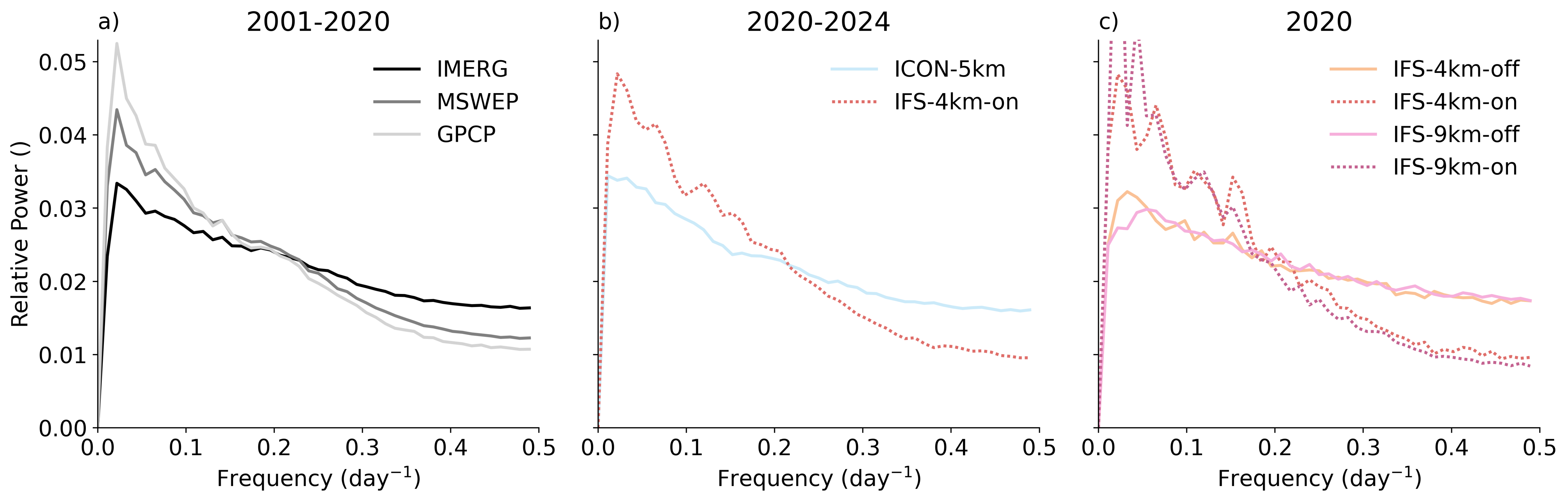
JAS
DJF
a)
b)
c)
d)
e)
f)
g)
h)
IMERG
ICON-5km
IFS-4km-off
CMIP6
30°N
15°N
0°
15°S
30°S
160°W
80°W
0°
80°E
160°E
−0.8
−0.6
−0.4
−0.2
0.0
0.2
0.4
0.6
0.8
$r_s$ ()

**Figure.**



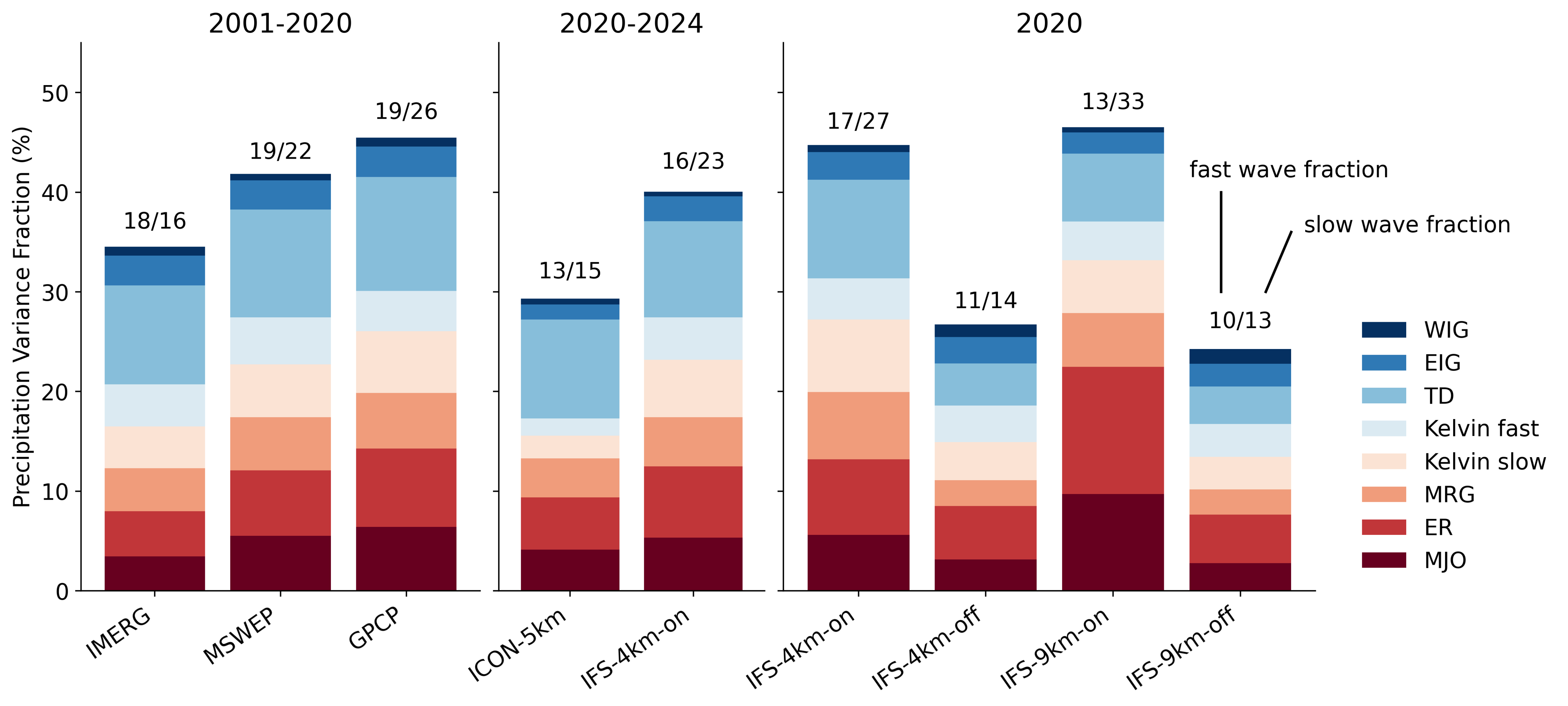
2001 - 2020
2020 - 2024
2020
2016
1979 - 2014
$r_s$ ()
Sahel
Tropical Land
Tropical Ocean
0.0
0.1
0.2
0.3
0.4
0.5
IMERG
MSWEP
GPCP
ICON-5km-off
IFS-4km-on
IFS-4km-on
IFS-4km-off
IFS-9km-on
IFS-9km-off
MPAS-A-7.5km-on
MPAS-A-7.5km-off
CMIP6

**Figure.**



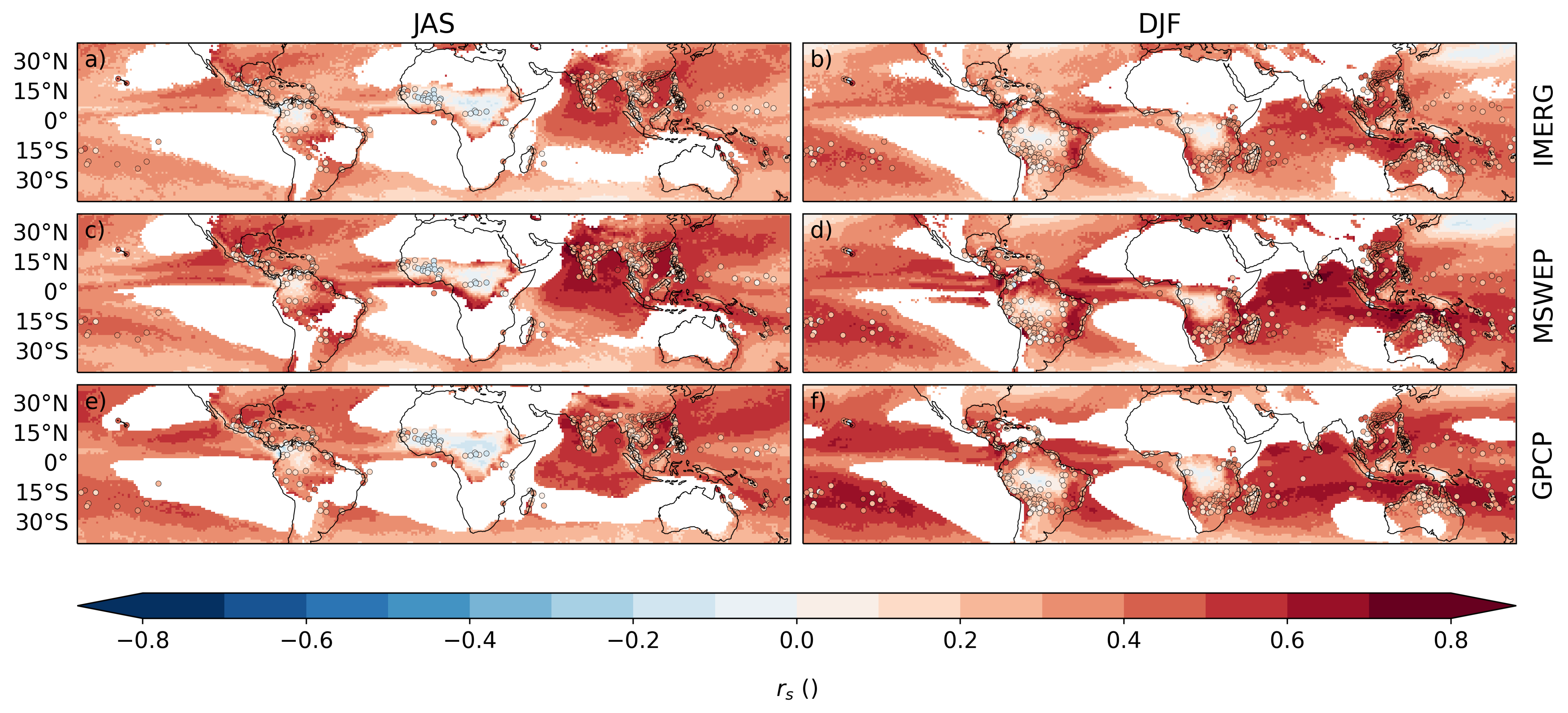
a)
2001-2020
b)
2020-2024
c)
2020
Relative Power ()
Frequency (day$^{-1}$)
IMERG
MSWEP
GPCP
ICON-5km
IFS-4km-on
IFS-4km-off
IFS-9km-off
IFS-9km-on
0.00
0.01
0.02
0.03
0.04
0.05
0.0
0.1
0.2
0.3
0.4
0.5

**Figure.**



2001-2020
2020-2024
2020
Precipitation Variance Fraction (%)
0
10
20
30
40
50
18/16
19/22
19/26
13/15
16/23
17/27
11/14
13/33
10/13
fast wave fraction
slow wave fraction
WIG
EIG
TD
Kelvin fast
Kelvin slow
MRG
ER
MJO
IMERG
MSWEP
GPCP
ICON-5km
IFS-4km-on
IFS-4km-on
IFS-4km-off
IFS-9km-on
IFS-9km-off

**Figure.**



JAS
DJF
a)
b)
c)
d)
e)
f)
30°N
15°N
0°
15°S
30°S
IMERG
MSWEP
GPCP
−0.8
−0.6
−0.4
−0.2
0.0
0.2
0.4
0.6
0.8
$r_s$ ()

# Autocorrelation – A Simple Diagnostic for Tropical Precipitation in Global Kilometer-Scale Climate Models

**Dorian Spät [1], Michela Biasutti [2], David Schuhbauer [1], Aiko Voigt [1]**

[1]University of Vienna, Department of Meteorology and Geophysics, Vienna, Austria
[2]Columbia University, Lamont-Doherty Earth Observatory, 10964 Palisades NY, United States

**Key Points:**

- A robust lag 1 autocorrelation pattern of daily precipitation exists in the tropics across different observations.
- Global kilometer-scale climate models capture the observed autocorrelation pattern, CMIP models do not.
- Convectively coupled equatorial waves are likely not driving the autocorrelation pattern.

Corresponding author: Dorian Spät, `dorian.spaet@univie.ac.at`

## Abstract

We propose the lag 1 autocorrelation of daily precipitation as a simple diagnostic of tropical precipitation in climate models. This metric generally has a relatively uniform distribution of positive values over the tropics. However, selected land regions are characterized by exceptionally low autocorrelation values. Low values correspond to the dominance of high-frequency variance in precipitation. Consistent with previous work, we show that CMIP6 climate models overestimate the autocorrelation. Global kilometer-scale models capture the observed autocorrelation pattern when deep convection is explicitly simulated. When a deep convection parameterization is used, the autocorrelation increases across the tropics, suggesting that land surface-atmosphere interactions are not responsible for the changes in precipitation variability. Furthermore, an accurate simulation of convectively coupled equatorial waves does not necessarily lead to a correct representation of the autocorrelation, and vice versa. This suggests other driving processes for the autocorrelation pattern.



## Plain Language Summary

Rainfall in the tropics is influenced by many atmospheric processes that depend on geographic location. We use the lag 1 autocorrelation a metric for the day-to-day persistence of rainfall. We find that rainfall is very persistent in most parts of the tropics with a few exceptions over land, e.g. the Sahel, where high-frequency rainfall events dominate. Our results show that models with a horizontal resolution of a few kilometers reproduce the autocorrelation pattern, in contrast to coarser climate models. We also analyze atmospheric waves and find that the degree to which a model captures these waves has no clear impact on its ability to capture rainfall persistence. Processes on smaller scales, like mesoscale convective systems and convective organization, could be key to understand the origin of the pattern.


## 1 Introduction

Precipitation in the tropics is, for the most part, the result of deep convection and is modulated by an intricate interplay of various different processes. The convective systems associated with most tropical rainfall are not of purely random nature, but are often organized by larger scale dynamics, ranging from mesoscale convective systems (MCSs) to synoptic scale convectively coupled equatorial waves (CCEWs) and the planetary scale Madden-Julian Oscillation (MJO) (Feng et al., 2021; Cheng et al., 2023; Wheeler & Kiladis, 1999; Cho et al., 2004; Kiladis et al., 2009). Thus, the day to day variability or rather persistence of tropical precipitation varies by location. Some of these geographic distinctions are captured by the lag 1 autocorrelation of daily precipitation, as presented by Roehrig et al. (2013) using precipitation observations from GPCP, and as shown below. Simply put, lag 1 autocorrelation is a measure of persistence. For example, high persistence of precipitation over several days will result in a high lag 1 autocorrelation. High variance of precipitation over the same period, on the other hand, will result in a low lag 1 autocorrelation.

Roehrig et al. (2013) further showed that the global climate models (GCMs) that are part of CMIP5 (Coupled Model Intercomparison Project Phase 5) all overestimate the lag 1 autocorrelation in the western Sahel region compared to GPCP. Consistently, Moon et al. (2019) found an overestimation of precipitation persistence in CMIP5 models compared to observations. This suggests that low resolution GCMs may be misrepresenting deep convection, its organization, or its coupling to the larger scale (e.g. CCEWs), which is consistent with previous work showing that CMIP models generally misinterpret tropical precipitation (Palmer & Stevens, 2019; Fiedler et al., 2020).

Unlike low resolution GCMs, kilometer-scale models are run with horizontal grid spacings of a few kilometers and with deep convection simulated explicitly. As a result, many precipitation characteristics, such as diurnal cycle, location and spatial propagation have been shown to be more accurately represented in kilometer-scale models (Stevens et al., 2020). Kilometer-scale models can also simulate many aspects of MCSs and related processes, which is not possible for most low resolution GCMs (see Feng et al. (2023) and sources within). The ability to correctly simulate MSCs is crucial, since MCSs produce more than 50 % of tropical precipitation (Feng et al., 2021). Considering larger scale dynamics, Judt and Rios-Berrios (2021) illustrated that the Model for Prediction Across Scales-Atmosphere (MPAS-A) produces much more realistic precipitation variances associated with what we will call fast CCEWs (lower frequency limit of 0.2 day$^{-1}$) when the resolution is refined to a few kilometers and deep convection is explicitly simulated. Moreover, Tomassini (2018) demonstrated that convection and African easterly waves (a type of fast CCEWs) are dynamically linked through mesoscale circulations, which poses a challenge to deep convection parameterizations. Kilometer-scale models, on the other hand, explicitly represent this mesoscale coupling mechanism and therefore could remedy shortcomings of traditional GCMs with respect to the lag 1 autocorrelation pattern in the tropics.

This study aims to introduce the lag 1 autocorrelation of daily precipitation as a simple diagnostic to assess tropical precipitation variability in kilometer-scale climate models. To this end, we first address the question of whether the lag 1 autocorrelation pattern is robust across different satellite-based observations and gauge station time series. We then investigate whether global kilometer-scale models are capable of producing a lag 1 autocorrelation pattern comparable to the observations. Finally, we analyze convectively coupled equatorial waves as a possible cause of the lag 1 autocorrelation pattern and discuss other possible drivers.

The paper is structured as follows. In Section 2 data and methods are detailed. We present the lag 1 autocorrelation of daily precipitation in Section 3.1 for multiple observations, global kilometer-scale simulations, and CMIP6 models. In Section 3.2 we examine the contribution of CCEWs to tropical precipitation variance and its connection to the lag 1 autocorrelation. Conclusions and further discussion are given in Section 4.

## 2 Data and Methods

### 2.1 Data

We use simulation data from two global kilometer-scale climate models, ICON-Sapphire (Hohenegger et al., 2023) and IFS coupled to the FESOM ocean model (Rackow et al., n.d.), produced as part of the nextGEMS project. We analyze the latest cycle 3 simulations (covering 2020-2024) but also include cycle 2 simulations (covering 2020) of the IFS model and coarser simulations with IFS coupled to the NEMO ocean model. The simulations span different periods, but results are robust already for one simulated season. In Table A1 we summarize key aspects of the nextGEMS simulations we use. While ICON simulations were always run without deep convection parameterization, IFS was run both with and without deep convection parameterization at both resolutions. In the IFS-4km-on simulation, a reduced cloud base mass flux was used in the deep convection parameterization (Rackow et al., n.d.).

Complementary, we also include global kilometer-scale simulations with MPAS-A (Skamarock et al., 2012) presented by Judt and Rios-Berrios (2021) (for the period August 1st to September 10th 2016). We analyze simulations at 3.75 km and at 7.5 km resolution, with and without deep convection parameterization. Furthermore, we utilize a number of amip style simulations of the CMIP6 ensemble (Eyring et al., 2016) (for the period 1979-2014, see Appendix B for the list of models) as low resolution model reference.

As observation reference we use three global precipitation datasets that are satellite-based and gauge-corrected: IMERG version 06 final run (Huffman, G.J. et al., 2019), MSWEP V2.2 (Beck et al., 2019) and GPCP version 1.3 (Adler et al., 2017). For the analysis we select an overlapping period from 2001 to 2020. In addition, we also include gauge measurement series from 1994 to 1999 (NOAA National Centers of Environmental Information., 1999).

To allow for an accurate comparison, all 2D fields are interpolated to a common 1° x 1° grid, before performing further analysis (except for the MPAS-A data, which are on a 2.5° x 2.5° grid).

### 2.2 Autocorrelation Diagnostic

Inspired by Roehrig et al. (2013), we calculate the lag $\tau$ autocorrelation $r_s$ for daily precipitation time series $X(t)$ for every grid cell. Although our results are robust to the choice of the correlation coefficient, we differ from Roehrig et al. (2013), in that we do not use the Pearson coefficient, but the Spearman coefficient. The latter measures how well a given relationship between two variables (in our case between a time series and a copy of that same time series shifted by lag=$\tau$) can be described by a monotone function, rather than a linear function. We believe that this approach is more suitable for daily precipitation. $r_s$ is defined as follows:

$$r_s = \rho_{R(X),R(X)}(\tau) \quad = \quad \frac{cov(R(X_{t+\tau}), R(X_t))}{var(R(X), R(X))}. \tag{1}$$

Here $cov(R(X_{t+\tau}), R(X_t))$ is the autocovariance of the rank $R$ of $X$, which is normalized by the variance $var(R(X), R(X))$, which means $r_s \in [-1, 1]$. Overbars denote time means. Like Roehrig et al. (2013), we use a lag of $\tau = 1$ because at this timescale the autocorrelation potentially captures the influence of large-scale processes such as CCEWs on precipitation, but no processes related to the diurnal cycle. We investigate individual seasons, so the ranks $R$ range from 1 to 92 for the JAS season, for example. From here on, we will refer to the lag 1 autocorrelation simply as autocorrelation.

### 2.3 Wave Filtering

We apply the wave filtering method developed by Wheeler and Kiladis (1999) to analyze six types of CCEWs and the MJO, here ordered from slow (low frequency) to fast (high frequency): MJO, equatorial Rossby (ER), mixed Rossby gravity (MRG), Kelvin, tropical depression (TD), eastward inertio-gravity waves (EIG, with $n = 0$) and westward inertio-gravity waves (WIG, with $n = 1$). For the filtering we use the full time-series, between 20°S and 20°N. We furthermore remove the first three harmonics of the seasonal cycle, detrend the signal and taper the ends to zero, as described by Wheeler and Kiladis (1999). To retain the full wave signal in regions where convection only occurs in one hemisphere, we do not apply a symmetric/anitsymmetric decomposition (Kiladis et al., 2009). The filter parameters for the different wave types are the same as in Schlueter et al. (2019) for all waves except WIG waves, for which we choose the parameters in Wheeler and Kiladis (1999). We moreover distinguish between slow and fast Kelvin waves, with a cutoff frequency of 0.2 day$^{-1}$, to make the lower frequency limit of the fast Kelvin waves match the one of the TD and the EIG waves. The two Kelvin wave types share the cut-off frequency, which results in a small overlap of the filters. The Kelvin wave splitting is still relevant, however, since the autocorrelation corresponds to the ratio of high to low frequency variance on the order of a few days, and the Kelvin filter band ranges from

0.05 to 0.4 day$^{-1}$. We will refer to MJO, ER, MRG and slow Kelvin as slow waves and to fast Kelvin, TD, EIG and WIG as fast waves. The filtering is done using a Python script, which we partly based on Miyachi (2023) and Medeiros (2023). We validated the script by comparing our results for the IMERG data against the Tropical Rainfall Measuring Mission based results in Schlueter et al. (2019). We estimate the variance of precipitation explained by CCEWs by calculating the squared Pearson correlation coefficient between the precipitation anomalies (without the first three harmonics of the seasonal cycle, detrended and tapered) and the wave filtered precipitation anomalies respectively (Schlueter et al., 2019).

# 3 Results

## 3.1 Lag 1 Autocorrelation of Daily Precipitation in the Tropics

The autocorrelation pattern for daily precipitation in the tropics described by Roehrig et al. (2013) for GPCP also holds for IMERG and MSWEP, with some differences in the absolute values (see Figure C1). For the subsequent analysis we focus on IMERG, since it compares best to the gauge data in regions of low autocorrelation (not shown) and unlike MSWEP is not influenced by reanalysis. The IMERG autocorrelation is therefore also depicted in (a) and (b) of Figure 1, together with gauge station data (indicated by the circles). There are some outliers when comparing the gauge stations with IMERG, e.g. in the Pacific, where individual stations show very low autocorrelation values. We do not believe that the differences are due to the different periods covered by the satellite and ground-based observations, since there is very little interannual variability in the autocorrelation. Instead, we believe that the differences are due to the local perspective of the stations, while the satellite-based grid cells cover an area of 1°x1°.

Overall, the autocorrelation is positive throughout the tropics and especially over the ocean, which is highlighted by the mean autocorrelation values over tropical land and ocean, depicted in Figure 2 for July-September (JAS). During JAS, there are two distinctive exceptional regions over land, the Sahel together with equatorial Africa, as well as the most northern part of South America at the border to Central America, with considerably lower (slightly negative) values of autocorrelation. During DJF, regions with similarly low autocorrelation are located in Africa, south of the equator and in South America. To some extend also parts of the maritime continent show considerably lower autocorrelation values than the surrounding ocean during DJF. Interestingly, these patterns can not simply be described as a land-sea contrast, as clarified by the pattern of autocorrelation over the Asian monsoon and the Gulf of Mexico regions. From here on, we will focus on JAS, since during this season the most pronounced low autocorrelation region over Africa exists, where we will more specifically focus on the Sahel domain.

The multi-model mean autocorrelation of a broad selection of CMIP6 simulations (see Figure 1) captures the observed pattern in broad strokes, but clearly overestimates the persistence of rainfall everywhere and especially over Africa. This is summarized by the region means for the Sahel domain, tropical land and tropical ocean in Figure 2. The drastic overestimation of the autocorrelation in the Sahel domain, is similar to the findings of Roehrig et al. (2013) for CMIP5 models. The CMIP6 models also overestimate the autocorrelation over tropical land and ocean, whereas the difference to the observations is smallest over the ocean. These low resolution GCM models all use deep convection parameterization schemes; we next contrast their behavior with kilometer-scale models without convective parameterizations.

Panels (c) and (d) of Figure 1 depict the autocorrelation maps for ICON-5km. The model captures key features that we also find in the observations. Especially the regions of low autocorrelation are captured by the ICON-5km simulation, with the African region during JAS being the most pronounced. The mean autocorrelation values are in the

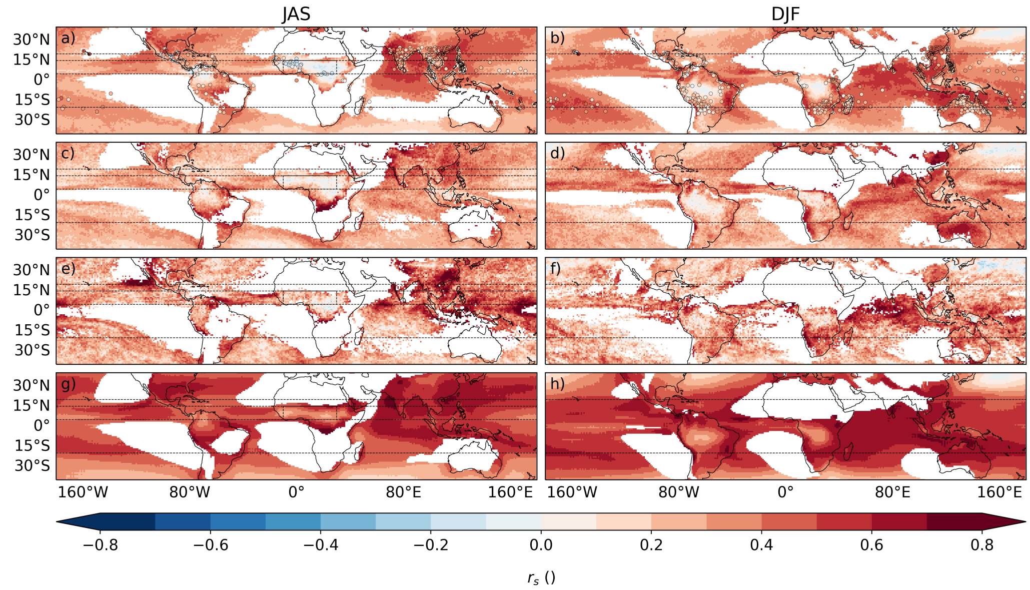



**Figure 1.** Lag 1 autocorrelation of daily precipitation on a 1° x 1° grid, datapoints with less than 1 mm/day mean precipitation are masked. IMERG for JAS in panel (a) and for DJF in panel (b). ICON-5km for JAS in panel (c) and for DJF in panel (d). IFS-4km-off for JAS in panel (e) and for DJF in panel (f). The circles in (a) and (b) indicate gauge station data. Stations with less than 1 mm/day mean precipitation are not shown. Black dashed lines indicate the domains for the further analysis: the Sahel domain (10°E-30°W, 5°N-15°N) the tropical domain (20°S-20°N) and the CCEW domain (5°N-15°N).

observational range in the Sahel region and generally also over tropical land (see Figure 2). However, the autocorrelation over the tropical ocean is lower than in the three observations.

The autocorrelation maps for IFS-4km-off are displayed in panels (e) and (f) of Figure 1. Again, the pattern matches the observations, but it is a little noisier than the IMERG and the ICON-5km data, since for this model setup we only have one simulated year. However, this shows that data from one single season is already enough for the autocorrelation pattern to appear. The mean autocorrelation values for IFS-4km-off in Figure 2 are close to the upper end of the observational range, or slightly above for the Sahel.

A number of additional simulations with the IFS model, produced within the nextGEMS project are available. IFS-4km-on, from the latest cycle 3 simulations is run with the deep convection scheme turned on, but a reduced vertical mass flux. Additionally, there are the 9 km simulations IFS-9km-off and IFS-9km-on. These simulations allow us to explore the influence of the deep convection paraterization scheme on the autocorrelation in the IFS model. Autocorrelation increases over the Sahel, tropical land and tropical ocean for both the 4.4 and the 9 km resolution IFS simulations (see Figure 2) when the deep convection parameterization is turned on. In fact, the increase in autocorrelation is relatively uniform across the tropics in both cases, with virtually no differences between land and ocean. This suggests that surface-atmosphere interactions are not driving factors for changes in the autocorrelation, when the deep convection parameterization is turned on.

MPAS-A also captures the autocorrelation pattern at both 3.75 and 7.5 km resolution when the deep convection is simulated explicitly (not shown). The mean auto-

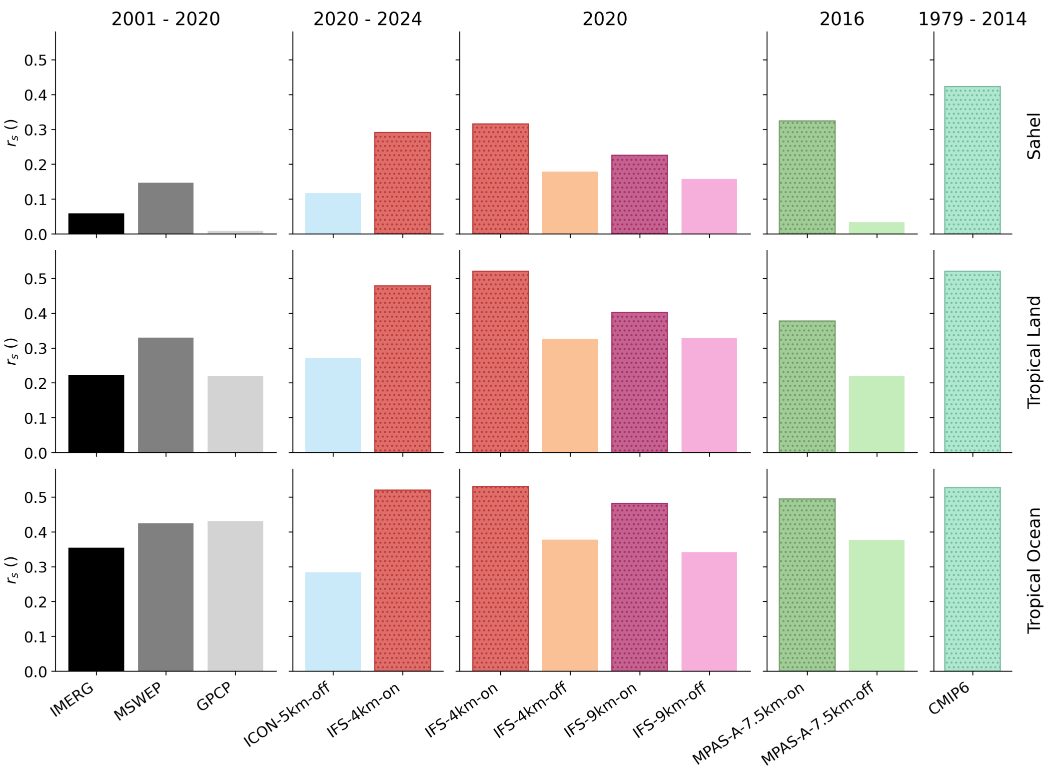



**Figure 2.** Mean lag 1 autocorrelation of daily precipitation calculated for July-September (JAS). In the first row calculated for the Sahel domain (10°E-30°W, 5°N-15°N), the second row for tropical land and in the third row for tropical ocean (20°S-20°N). Depicted in the first column from the left are the values for the observations, in the second column for the nextGEMS cycle 3 simulations, in the third column for different nextGEMS IFS simulations (cycle 2 and 3), in the fourth column for two MPAS-A simulations (simulation period only from August 1st to September 10th 2016) and in the fifth column the CMIP6 mean is depicted. Dotted bars indicate simulations with active deep convection parameterization. Gridpoints with less than 1 mm/day mean precipitation are not included.

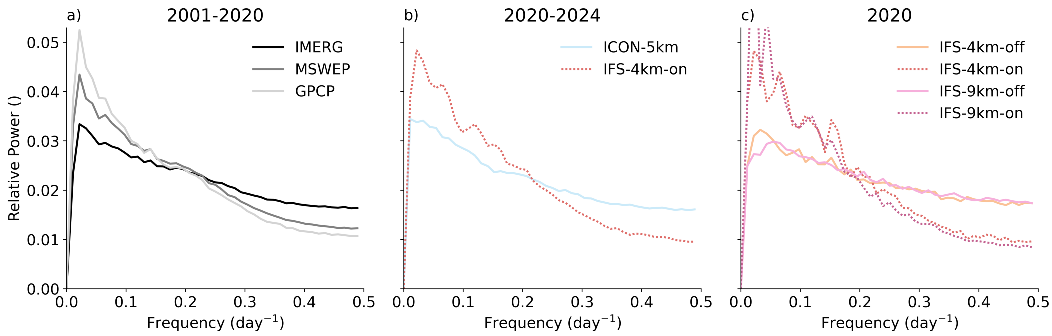



**Figure 3.** Relative power spectra calculated from daily precipitation, with the observations in (a), the nextGEMS cycle 3 simulations in (b) and the different nextGEMS IFS simulations (cycle 2 and 3) in (c). Gridpoints with less than 1 mm/day mean precipitation are not included.

correlation values for the MPAS-A-7.5km simulations in Figure 2 are consistent with the influence of the deep convection parameterization on the autocorrelation in IFS. As for the IFS model, the autocorrelation increases when the deep convection parameterization is turned on. The increase is much stronger over the Sahel than over the rest of the tropics in MPAS-A-7.5km (similarly in MPAS-A-3.75km, not shown). However, it should be kept in mind that these simulations only cover a range of 40 days.

The increase in autocorrelation when the deep convection parameterization is turned on is caused by an increase in high frequency variance (meaning on the time-scale of a few days) relative to low frequency variance in the IFS model. This is demonstrated in Figure 3, which depicts the tropical mean relative power spectra of daily precipitation for the nextGEMS simulations and the observations for JAS. When the IFS model is run with deep convection treated explicitly, the relative power spectra agree very well with ICON-5km and the observations. Whereas, when the parameterization is turned on, power shifts to lower frequencies and the spectra drop out of the observational range.

### 3.2 Relative CCEW Contribution to Precipitation Variance

A considerable fraction of the precipitation variance that underlies the power spectra presented in Figure 3 are related to CCEWs. To investigate the role of CCEWs in the origin of the autocorrelation pattern we therefore calculate the precipitation variance fraction on a daily time-scale for the different wave types in the CCEW band defined in Schlueter et al. (2019) for JAS (see Figure 1). Following Schlueter et al. (2019) this is achieved by calculating the squared Pearson correlation coefficients between precipitation anomalies and wave filtered precipitation anomalies of the latitudinal mean of the CCEW band. We subsequently take the longitudinal mean, to generate the mean values presented in Figure 4. Since this correlation method does not take interactions between waves into account, the sum of the precipitation variance fractions attributed to CCEWs can be interpreted as the maximum variance fraction explainable by CCEWs (Schlueter et al., 2019). It should be noted that only a small part of the WIG frequency spectrum can be filtered from daily data, which is why they only contribute relatively small variance fractions at this time-scale.

Differences in the summed CCEW variance fraction in the observations in Figure 4 are as large as 10 percentage points between IMERG and GPCP, which delivers the highest summed fraction of more than 45 %. These differences arise mainly due to differences in the slow CCEW fractions (MJO, ER and MRG).

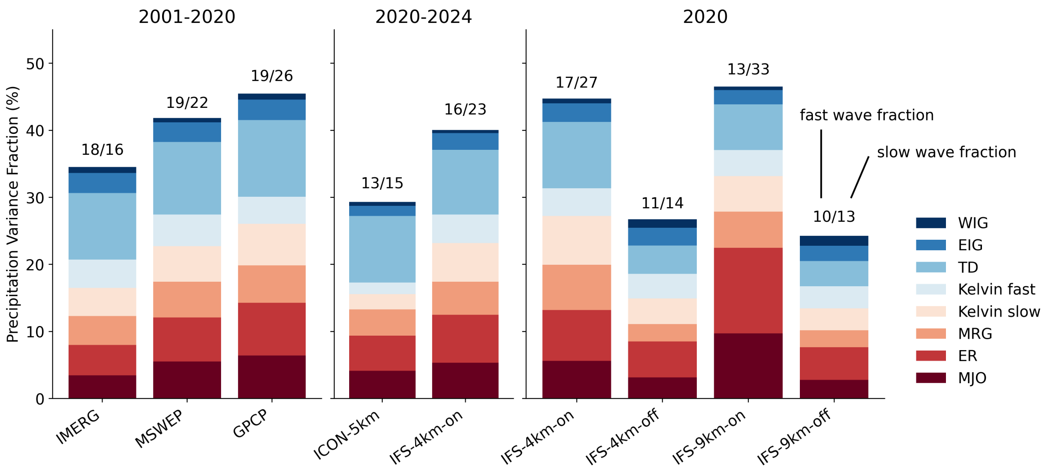



**Figure 4.** Precipitation variance fractions attributed to convectively coupled equatorial waves (CCEWs) calculated as the squared Pearson correlation between precipitation anomalies and wave filtered precipitation anomalies for July-September (JAS) in the CCEW band (5°N-15°N). From left to right, the first three bars depict the mean values for the observations, the next two bars the values for the nextGEMS cycle 3 simulations and the last four bars the values for different nextGEMS IFS simulations (cycle 2 and 3). The fractions over the bars denote the summed variance fractions of the fast waves over the summed variance fractions of the slow waves.

The ICON-5km simulations produce a summed CCEW variance fraction smaller than the observational range. The variance fractions for the slow CCEWs are comparable to IMERG, but the model produces lower values for both fast and slow Kelvin and EIG waves.

For the IFS simulations with the deep convection parameterization turned on, the summed CCEW variance fractions are close to the upper end of the observational range. The simulations with deep convection simulated explicitly, on the other hand, deliver summed CCEW variance fractions that are considerably smaller and closer to the one of ICON-5km. The changes between deep convection parameterization on and off simulations are not uniform across wave types. For example, there are differences in the variance fractions of the TD waves. The comparably small fractions for this wave type in IFS-4km-off and IFS-9km-off are also a difference to the ICON-5km simulation, which in turn delivers smaller fractions for Kelvin and EIG waves. The largest changes, however, are in the variance fractions of the slow waves, which are particularly pronounced for the IFS-9km simulations. These changes are related to the autocorrelation in the sense that the larger variance fraction in the slow waves, when the deep convection parameterization is turned on, increases the power in the lower frequencies of the power spectrum (see Figure 3).

The overall importance of CCEWs in the origin of the tropical autocorrelation pattern is difficult to accurately quantify. However, we find that the ICON-5km and the IFS-4km-off models, while accurately reproducing the autocorrelation pattern, produce summed CCEW variance fractions lower than the observational range. The IFS-4km-on model on the other hand, produces CCEW variance fractions close to the observations, while overestimating autocorrelation across the tropics.

## 4 Conclusions

We study the lag 1 autocorrelation of daily precipitation as an easy-to-compute diagnostic of tropical precipitation variability in kilometer-scale climate models. The diagnostic is helpful to investigate the persistence of precipitation across the tropics.

Our results shows that the autocorrelation pattern first presented by Roehrig et al. (2013) is robust across satellite-based observations and gauge stations. The pattern can be summarized as one of largest autocorrelation over ocean; smaller, but still positive autocorrelation over tropical land; and much smaller or even negative autocorrelation in the core of the rainy regions of Africa and South America. The kilometer-scale climate models of the nextGEMS project and MPAS-A produce a similar pattern, in contrast to a wide range of CMIP6 models. In the IFS model the autocorrelation is very comparable to observations when the deep convection is simulated explicitly, but overestimated across the tropics when the deep convection parameterization is turned on. This effect occurs over several horizontal resolutions and even if the parameterized vertical mass flux is strongly reduced. Hence, we find a distinct dependence on the representation of deep convection, rather than simply horizontal resolution, independent of the surface being land or ocean. This result is confirmed by our analysis of global kilometer-scale MPAS-A simulations, which show the same dependence on the representation of deep convection.

We attribute precipitation variance to different kinds of convectively coupled equatorial waves (CCEW). When deep convection is treated explicitly, the power spectrum of precipitation becomes whiter, reducing fractional variance from CCEW as a set, and specifically by depressing the role of slow waves. We show that a climate model with deep convection parameterization turned on, namely IFS-4km-on, produces CCEW related precipitation fractions that compare relatively well to the observations and at the same time overestimate the autocorrelation. On the other hand different kilometer-scale models, ICON-5km and IFS-4km-off, reproduce the autocorrelation pattern, but underestimate the CCEW related precipitation fraction. In the IFS model, turning off the deep convection parameterization leads to a decrease of variance fractions related to slow CCEWs and an increase in the variance fraction related to tropical depression type waves. The latter is in contrast to what Judt and Rios-Berrios (2021) reported for MPAS-A. Taking these findings into account, we assume that CCEWs, while probably playing a role, are not the main driving factor for the observed autocorrelation pattern.

The origin of the autocorrelation pattern could lie on scales finer than those of CCEWs. A comparison of the maps of the autocorrelation and observation-based statistics of mesoscale convective systems (MCSs) from Feng et al. (2021) indicates that the regions of low autocorrelation are also regions with high MCS precipitation fractions and MCS numbers. However, there are also regions over the ocean (e.g. the Indian ocean) where this association does not hold. Moreover, Mathon et al. (2002) showed that over the central Sahel, where autocorrelation is low, organized MCSs (OMCSs), account for 90 % of the seasonal precipitation, while only representing 12 % of the total MCS count. The low frequency of OMCSs may therefore indicate that MCSs are not the main drivers of the autocorrelation pattern either. A deeper analysis of MCSs and convective organization on even smaller scales in the nextGEMS simulations is necessary to further constrain the mechanisms behind the autocorrelation pattern.

Regardless the exact composition of phenomena that lead to the observed pattern of autocorrelation of daily precipitation, the autocorrelation is an easily computed metric that encapsulates these underlying atmospheric processes and provides a simple diagnostic for tropical precipitation in GCMs. Our finding that kilometer-scale models are able to reproduce the autocorrelation pattern furthermore demonstrates that these models offer a unique possibility to study tropical precipitation and the related atmospheric dynamics.

## Acknowledgments

nextGEMS is funded through the European Union's Horizon 2020 research and innovation program under the grant agreement number 101003470.

## Data Availability Statement

Selected data of the nextGEMS simulations can be found at `https://doi.org/10.26050/WDCC/nextGEMS_cyc2` (Wieners et al., 2023) for cycle 2 and at `https://doi.org/10.26050/WDCC/nextGEMS_cyc3` (Koldunov et al., 2023) for cycle 3. The CMIP6 model data can be obtained from the Earth System Grid Foundation. MPAS-A data is from `https://doi.org/10.5065/v05b-9e73`. IMERG data were accessed via NASA's Goddard Earth Sciences Data and Information Services Center (`https://disc.gsfc.nasa.gov/datasets/GPM_3IMERGDF_06/summary?keywords=imerg%20final%20v06`). MSWEP data were accessed via `https://www.gloh2o.org/mswep/`. GPCP data can be obtained from `https://www.ncei.noaa.gov/access/metadata/landing-page/bin/iso?id=gov.noaa.ncdc:C00999`. The gauge data were accessed via `http://iridl.ldeo.columbia.edu/SOURCES/.NOAA/.NCDC/.DAILY/.GLOBALSOD/?sem=iridl%3ADCAtmosphere`. The post-processed data will be made available and get a doi as soon as the paper is under review. The analysis scripts are available under `https://ucloud.univie.ac.at/index.php/s/ctGcFy3nw6yF4Zb` and will get a doi as soon as the paper is under review.

## Competing Interests

The authors declare that they have no conflict of interest.

## References

Adler, R., Wang, Jian-Jian, Sapiano, Mathew, Huffman, George, Bolvin, David, Nelkin, Eric, & NOAA CDR Program. (2017, May). *Global Precipitation Climatology Project (GPCP) Climate Data Record (CDR), Version 1.3 (Daily).* (NOAA National Centers for Environmental Information.) doi: 10.7289/V5RX998Z

Beck, H. E., Wood, E. F., Pan, M., Fisher, C. K., Miralles, D. G., Van Dijk, A. I. J. M., ... Adler, R. F. (2019, March). MSWEP V2 Global 3-Hourly 0.1° Precipitation: Methodology and Quantitative Assessment. *Bulletin of the American Meteorological Society*, *100*(3), 473–500. Retrieved 2024-01-15, from `https://journals.ametsoc.org/view/journals/bams/100/3/bams-d-17-0138.1.xml` doi: 10.1175/BAMS-D-17-0138.1

Cheng, Y., Dias, J., Kiladis, G., Feng, Z., & Leung, L. R. (2023, May). Mesoscale Convective Systems Modulated by Convectively Coupled Equatorial Waves. *Geophysical Research Letters*, *50*(10), e2023GL103335. Retrieved 2024-01-08, from `https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2023GL103335` doi: 10.1029/2023GL103335

Cho, H.-K., Bowman, K. P., & North, G. R. (2004, November). Equatorial Waves Including the Madden–Julian Oscillation in TRMM Rainfall and OLR Data. *Journal of Climate*, *17*(22), 4387–4406. Retrieved 2023-10-03, from `http://journals.ametsoc.org/doi/10.1175/3215.1` doi: 10.1175/3215.1

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016, May). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, *9*(5), 1937–1958. Retrieved 2024-01-30, from `https://gmd.copernicus.org/articles/9/1937/2016/` doi: 10.5194/gmd-9-1937-2016

Feng, Z., Leung, L. R., Hardin, J., Terai, C. R., Song, F., & Caldwell, P. (2023, February). Mesoscale Convective Systems in DYAMOND Global Convection-

Permitting Simulations. *Geophysical Research Letters*, *50*(4), e2022GL102603. Retrieved 2024-01-08, from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022GL102603 doi: 10.1029/2022GL102603

Feng, Z., Leung, L. R., Liu, N., Wang, J., Houze, R. A., Li, J., ... Guo, J. (2021, April). A Global High-Resolution Mesoscale Convective System Database Using Satellite-Derived Cloud Tops, Surface Precipitation, and Tracking. *Journal of Geophysical Research: Atmospheres*, *126*(8), e2020JD034202. Retrieved 2024-01-08, from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2020JD034202 doi: 10.1029/2020JD034202

Fiedler, S., Crueger, T., D'Agostino, R., Peters, K., Becker, T., Leutwyler, D., ... Stevens, B. (2020, September). Simulated Tropical Precipitation Assessed across Three Major Phases of the Coupled Model Intercomparison Project (CMIP). *Monthly Weather Review*, *148*(9), 3653–3680. Retrieved 2024-01-31, from http://journals.ametsoc.org/doi/10.1175/MWR-D-19-0404.1 doi: 10.1175/MWR-D-19-0404.1

Hohenegger, C., Korn, P., Linardakis, L., Redler, R., Schnur, R., Adamidis, P., ... Stevens, B. (2023, January). ICON-Sapphire: simulating the components of the Earth system and their interactions at kilometer and subkilometer scales. *Geoscientific Model Development*, *16*(2), 779–811. Retrieved 2023-06-20, from https://gmd.copernicus.org/articles/16/779/2023/ doi: 10.5194/gmd-16-779-2023

Huffman, G.J., E.F. Stocker, D.T. Bolvin, E.J. Nelkin, & Jackson Tan. (2019). *GPM IMERG Final Precipitation L3 1 day 0.1 degree x 0.1 degree V06, Edited by Andrey Savtchenko, Greenbelt, MD, Goddard Earth Sciences Data and Information Services Center (GES DISC).* doi: 10.5067/GPM/IMERGDF/DAY/06

Judt, F., & Rios-Berrios, R. (2021, July). Resolved Convection Improves the Representation of Equatorial Waves and Tropical Rainfall Variability in a Global Nonhydrostatic Model. *Geophysical Research Letters*, *48*(14). Retrieved 2023-08-22, from https://onlinelibrary.wiley.com/doi/10.1029/2021GL093265 doi: 10.1029/2021GL093265

Kiladis, G. N., Wheeler, M. C., Haertel, P. T., Straub, K. H., & Roundy, P. E. (2009, April). Convectively coupled equatorial waves. *Reviews of Geophysics*, *47*(2), RG2003. Retrieved 2023-06-09, from http://doi.wiley.com/10.1029/2008RG000266 doi: 10.1029/2008RG000266

Koldunov, N., Kölling, T., Pedruzo-Bagazgoitia, X., Rackow, T., Redler, R., Sidorenko, D., ... Ziemen, F. A. (2023). *nextGEMS: output of the model development cycle 3 simulations for ICON and IFS.* World Data Center for Climate (WDCC) at DKRZ. Retrieved from https://doi.org/10.26050/WDCC/nextGEMS_cyc3 doi: 10.26050/WDCC/nextGEMS_cyc3

Mathon, V., Laurent, H., & Lebel, T. (2002, November). Mesoscale Convective System Rainfall in the Sahel. *Journal of Applied Meteorology*, *41*(11), 1081–1092. Retrieved 2023-07-06, from http://journals.ametsoc.org/doi/10.1175/1520-0450(2002)041<1081:MCSRIT>2.0.CO;2 doi: 10.1175/1520-0450(2002)041⟨1081:MCSRIT⟩2.0.CO;2

Medeiros, B. (2023, October). *wavenumber_frequency.* Retrieved from https://github.com/brianpm/wavenumber_frequency

Miyachi, T. (2023, January). *mcclimate.* Retrieved from https://github.com/tmiyachi/mcclimate/tree/master

Moon, H., Gudmundsson, L., Guillod, B. P., Venugopal, V., & Seneviratne, S. I. (2019, October). Intercomparison of daily precipitation persistence in multiple global observations and climate models. *Environmental Research Letters*, *14*(10), 105009. Retrieved 2024-02-14, from https://iopscience.iop.org/article/10.1088/1748-9326/ab4169 doi: 10.1088/1748-9326/ab4169

NOAA National Centers of Environmental Information. (1999). *Global Surface Sum-*

*mary of the Day - GSOD. 1.0.*

Palmer, T., & Stevens, B. (2019, December). The scientific challenge of understanding and estimating climate change. *Proceedings of the National Academy of Sciences*, *116*(49), 24390–24395. Retrieved 2024-01-31, from `https://pnas.org/doi/full/10.1073/pnas.1906691116` doi: 10.1073/pnas.1906691116

Rackow et al. (n.d.). Multi-year simulations at kilometre scale with the Integrated Forecasting System coupled to FESOM2.5/NEMOv3.4.
(in preparation)

Roehrig, R., Bouniol, D., Guichard, F., Hourdin, F., & Redelsperger, J.-L. (2013, September). The Present and Future of the West African Monsoon: A Process-Oriented Assessment of CMIP5 Simulations along the AMMA Transect. *Journal of Climate*, *26*(17), 6471–6505. Retrieved 2023-05-11, from `http://journals.ametsoc.org/doi/10.1175/JCLI-D-12-00505.1` doi: 10.1175/JCLI-D-12-00505.1

Schlueter, A., Fink, A. H., Knippertz, P., & Vogel, P. (2019, March). A Systematic Comparison of Tropical Waves over Northern Africa. Part I: Influence on Rainfall. *Journal of Climate*, *32*(5), 1501–1523. Retrieved 2023-05-11, from `http://journals.ametsoc.org/doi/10.1175/JCLI-D-18-0173.1` doi: 10.1175/JCLI-D-18-0173.1

Skamarock, W. C., Klemp, J. B., Duda, M. G., Fowler, L. D., Park, S.-H., & Ringler, T. D. (2012, September). A Multiscale Nonhydrostatic Atmospheric Model Using Centroidal Voronoi Tesselations and C-Grid Staggering. *Monthly Weather Review*, *140*(9), 3090–3105. Retrieved 2024-01-29, from `http://journals.ametsoc.org/doi/10.1175/MWR-D-11-00215.1` doi: 10.1175/MWR-D-11-00215.1

Stevens, B., Acquistapace, C., Hansen, A., Heinze, R., Klinger, C., Klocke, D., . . . Zängl, G. (2020). The Added Value of Large-eddy and Storm-resolving Models for Simulating Clouds and Precipitation. *Journal of the Meteorological Society of Japan. Ser. II*, *98*(2), 395–435. Retrieved 2023-05-11, from `https://www.jstage.jst.go.jp/article/jmsj/98/2/98_2020-021/_article` doi: 10.2151/jmsj.2020-021

Tomassini, L. (2018, December). Mesoscale Circulations and Organized Convection in African Easterly Waves. *Journal of the Atmospheric Sciences*, *75*(12), 4357–4381. Retrieved 2023-05-11, from `https://journals.ametsoc.org/doi/10.1175/JAS-D-18-0183.1` doi: 10.1175/JAS-D-18-0183.1

Wheeler, M., & Kiladis, G. N. (1999, February). Convectively Coupled Equatorial Waves: Analysis of Clouds and Temperature in the Wavenumber–Frequency Domain. *Journal of the Atmospheric Sciences*, *56*(3), 374–399. Retrieved 2023-10-03, from `http://journals.ametsoc.org/doi/10.1175/1520-0469(1999)056<0374:CCEWAO>2.0.CO;2` doi: 10.1175/1520-0469(1999)056⟨0374:CCEWAO⟩2.0.CO;2

Wieners, K.-H., Ziemen, F. A., Koldunov, N., Pedruzo-Bagazgoitia, X., Rackow, T., Redler, R., . . . Kölling, T. (2023). *nextGEMS: output of the model development cycle 2 simulations for ICON and IFS.* World Data Center for Climate (WDCC) at DKRZ. Retrieved from `https://doi.org/10.26050/WDCC/nextGEMS_cyc2` doi: 10.26050/WDCC/nextGEMS_cyc2

## Appendix A Overview of nextGEMS Simulations

The key settings and the simulation acronyms of the nextGEMS simulations we analyze are presented in Table A1.

**Table A1.** Settings of the nextGEMS simulations used in our analysis.

| Simulation acronym | Atmosphere model | Cycle | Horizontal resolution | Simulation years | Deep convection parameterization |
|---|---|---|---|---|---|
| ICON-5km | ICON-Sapphire | 3 | 5 km | 2020-2024 | off |
| IFS-4km-on | IFS | 3 | 4.4 km | 2020-2024 | on (reduced) |
| IFS-4km-off | IFS | 2 | 4.4 km | 2020 | off |
| IFS-9km-on | IFS | 2 | 9 km | 2020 | on |
| IFS-9km-off | IFS | 2 | 9 km | 2020 | off |

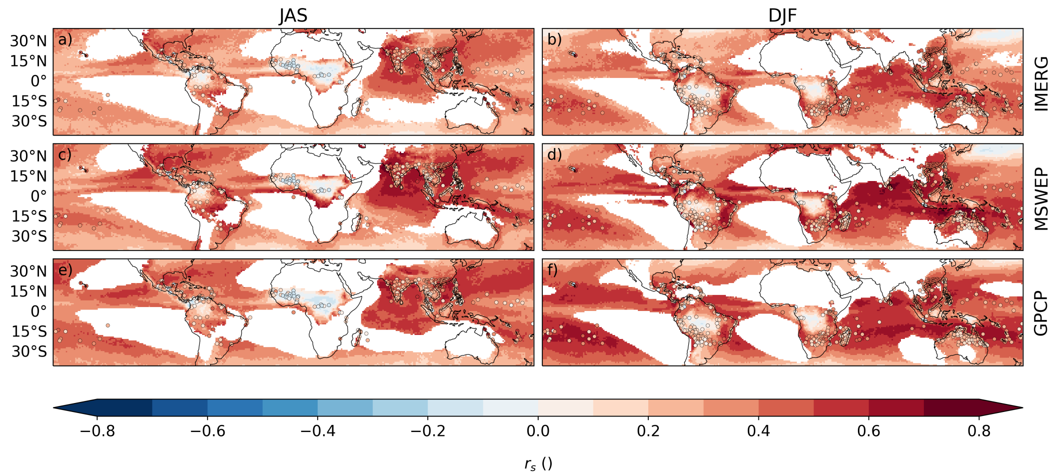



**Figure C1.** Lag 1 autocorrelation of daily precipitation on a 1° x 1° grid, datapoints with less than 1 mm/day mean precipitation are masked. IMERG for JAS in panel (a) and for DJF in panel (b). MSWEP for JAS in panel (c) and for DJF in panel (d). GPCP for JAS in panel (e) and for DJF in panel (f). Circles indicate gauge station data. Stations with less than 1 mm/day mean precipitation are not shown.

## Appendix B List of CMIP6 Simulations

We use amip r1i1p1f1 CMIP6 simulations (Eyring et al., 2016) for our analysis, included are the following models:

ACCESS-CM2, ACCESS-ESM1-5, BCC-ESM1, CAMS-CSM1-0C, ESM2-FV2, CESM2-WACCM, CESM2, CMCC-CM2-HR4, CMCC-CM2-SR5, CanESM5, E3SM-1-0, E3SM-2-0, EC-Earth3-AerChem, EC-Earth3-CC, EC-Earth3-Veg-LR, EC-Earth3-Veg, EC-Earth3, FGOALS-f3-L, FGOALS-g3, GFDL-CM4, GFDL-ESM4, ICON-ESM-LR, IITM-ESM, INM-CM4-8, INM-CM5-0, IPSL-CM6A-LR, KACE-1-0-G, MIROC6, MPI-ESM-1-2-HAM, MPI-ESM1-2-HR, MPI-ESM1-2-LR, MRI-ESM2-0, NESM3, NorCPM1, NorESM2-LM and TaiESM1.

## Appendix C Autocorrelation Maps for Observations

Figure C1 depicts the lag 1 autocorrelation maps for the analyzed observations.

# Autocorrelation – A Simple Diagnostic for Tropical Precipitation in Global Kilometer-Scale Climate Models

**Dorian Spät [1], Michela Biasutti [2], David Schuhbauer [1], Aiko Voigt [1]**

[1]University of Vienna, Department of Meteorology and Geophysics, Vienna, Austria
[2]Columbia University, Lamont-Doherty Earth Observatory, 10964 Palisades NY, United States

**Key Points:**

- A robust lag 1 autocorrelation pattern of daily precipitation exists in the tropics across different observations.
- Global kilometer-scale climate models capture the observed autocorrelation pattern, CMIP models do not.
- Convectively coupled equatorial waves are likely not driving the autocorrelation pattern.

Corresponding author: Dorian Spät, `dorian.spaet@univie.ac.at`

## Abstract

We propose the lag 1 autocorrelation of daily precipitation as a simple diagnostic of tropical precipitation in climate models. This metric generally has a relatively uniform distribution of positive values over the tropics. However, selected land regions are characterized by exceptionally low autocorrelation values. Low values correspond to the dominance of high-frequency variance in precipitation. Consistent with previous work, we show that CMIP6 climate models overestimate the autocorrelation. Global kilometer-scale models capture the observed autocorrelation pattern when deep convection is explicitly simulated. When a deep convection parameterization is used, the autocorrelation increases across the tropics, suggesting that land surface-atmosphere interactions are not responsible for the changes in precipitation variability. Furthermore, an accurate simulation of convectively coupled equatorial waves does not necessarily lead to a correct representation of the autocorrelation, and vice versa. This suggests other driving processes for the autocorrelation pattern.



## Plain Language Summary

Rainfall in the tropics is influenced by many atmospheric processes that depend on geographic location. We use the lag 1 autocorrelation a metric for the day-to-day persistence of rainfall. We find that rainfall is very persistent in most parts of the tropics with a few exceptions over land, e.g. the Sahel, where high-frequency rainfall events dominate. Our results show that models with a horizontal resolution of a few kilometers reproduce the autocorrelation pattern, in contrast to coarser climate models. We also analyze atmospheric waves and find that the degree to which a model captures these waves has no clear impact on its ability to capture rainfall persistence. Processes on smaller scales, like mesoscale convective systems and convective organization, could be key to understand the origin of the pattern.


## 1 Introduction

Precipitation in the tropics is, for the most part, the result of deep convection and is modulated by an intricate interplay of various different processes. The convective systems associated with most tropical rainfall are not of purely random nature, but are often organized by larger scale dynamics, ranging from mesoscale convective systems (MCSs) to synoptic scale convectively coupled equatorial waves (CCEWs) and the planetary scale Madden-Julian Oscillation (MJO) (Feng et al., 2021; Cheng et al., 2023; Wheeler & Kiladis, 1999; Cho et al., 2004; Kiladis et al., 2009). Thus, the day to day variability or rather persistence of tropical precipitation varies by location. Some of these geographic distinctions are captured by the lag 1 autocorrelation of daily precipitation, as presented by Roehrig et al. (2013) using precipitation observations from GPCP, and as shown below. Simply put, lag 1 autocorrelation is a measure of persistence. For example, high persistence of precipitation over several days will result in a high lag 1 autocorrelation. High variance of precipitation over the same period, on the other hand, will result in a low lag 1 autocorrelation.

Roehrig et al. (2013) further showed that the global climate models (GCMs) that are part of CMIP5 (Coupled Model Intercomparison Project Phase 5) all overestimate the lag 1 autocorrelation in the western Sahel region compared to GPCP. Consistently, Moon et al. (2019) found an overestimation of precipitation persistence in CMIP5 models compared to observations. This suggests that low resolution GCMs may be misrepresenting deep convection, its organization, or its coupling to the larger scale (e.g. CCEWs), which is consistent with previous work showing that CMIP models generally misinterpret tropical precipitation (Palmer & Stevens, 2019; Fiedler et al., 2020).

Unlike low resolution GCMs, kilometer-scale models are run with horizontal grid spacings of a few kilometers and with deep convection simulated explicitly. As a result, many precipitation characteristics, such as diurnal cycle, location and spatial propagation have been shown to be more accurately represented in kilometer-scale models (Stevens et al., 2020). Kilometer-scale models can also simulate many aspects of MCSs and related processes, which is not possible for most low resolution GCMs (see Feng et al. (2023) and sources within). The ability to correctly simulate MSCs is crucial, since MCSs produce more than 50 % of tropical precipitation (Feng et al., 2021). Considering larger scale dynamics, Judt and Rios-Berrios (2021) illustrated that the Model for Prediction Across Scales-Atmosphere (MPAS-A) produces much more realistic precipitation variances associated with what we will call fast CCEWs (lower frequency limit of 0.2 day$^{-1}$) when the resolution is refined to a few kilometers and deep convection is explicitly simulated. Moreover, Tomassini (2018) demonstrated that convection and African easterly waves (a type of fast CCEWs) are dynamically linked through mesoscale circulations, which poses a challenge to deep convection parameterizations. Kilometer-scale models, on the other hand, explicitly represent this mesoscale coupling mechanism and therefore could remedy shortcomings of traditional GCMs with respect to the lag 1 autocorrelation pattern in the tropics.

This study aims to introduce the lag 1 autocorrelation of daily precipitation as a simple diagnostic to assess tropical precipitation variability in kilometer-scale climate models. To this end, we first address the question of whether the lag 1 autocorrelation pattern is robust across different satellite-based observations and gauge station time series. We then investigate whether global kilometer-scale models are capable of producing a lag 1 autocorrelation pattern comparable to the observations. Finally, we analyze convectively coupled equatorial waves as a possible cause of the lag 1 autocorrelation pattern and discuss other possible drivers.

The paper is structured as follows. In Section 2 data and methods are detailed. We present the lag 1 autocorrelation of daily precipitation in Section 3.1 for multiple observations, global kilometer-scale simulations, and CMIP6 models. In Section 3.2 we examine the contribution of CCEWs to tropical precipitation variance and its connection to the lag 1 autocorrelation. Conclusions and further discussion are given in Section 4.

## 2 Data and Methods

### 2.1 Data

We use simulation data from two global kilometer-scale climate models, ICON-Sapphire (Hohenegger et al., 2023) and IFS coupled to the FESOM ocean model (Rackow et al., n.d.), produced as part of the nextGEMS project. We analyze the latest cycle 3 simulations (covering 2020-2024) but also include cycle 2 simulations (covering 2020) of the IFS model and coarser simulations with IFS coupled to the NEMO ocean model. The simulations span different periods, but results are robust already for one simulated season. In Table A1 we summarize key aspects of the nextGEMS simulations we use. While ICON simulations were always run without deep convection parameterization, IFS was run both with and without deep convection parameterization at both resolutions. In the IFS-4km-on simulation, a reduced cloud base mass flux was used in the deep convection parameterization (Rackow et al., n.d.).

Complementary, we also include global kilometer-scale simulations with MPAS-A (Skamarock et al., 2012) presented by Judt and Rios-Berrios (2021) (for the period August 1st to September 10th 2016). We analyze simulations at 3.75 km and at 7.5 km resolution, with and without deep convection parameterization. Furthermore, we utilize a number of amip style simulations of the CMIP6 ensemble (Eyring et al., 2016) (for the period 1979-2014, see Appendix B for the list of models) as low resolution model reference.

As observation reference we use three global precipitation datasets that are satellite-based and gauge-corrected: IMERG version 06 final run (Huffman, G.J. et al., 2019), MSWEP V2.2 (Beck et al., 2019) and GPCP version 1.3 (Adler et al., 2017). For the analysis we select an overlapping period from 2001 to 2020. In addition, we also include gauge measurement series from 1994 to 1999 (NOAA National Centers of Environmental Information., 1999).

To allow for an accurate comparison, all 2D fields are interpolated to a common 1° x 1° grid, before performing further analysis (except for the MPAS-A data, which are on a 2.5° x 2.5° grid).

### 2.2 Autocorrelation Diagnostic

Inspired by Roehrig et al. (2013), we calculate the lag $\tau$ autocorrelation $r_s$ for daily precipitation time series $X(t)$ for every grid cell. Although our results are robust to the choice of the correlation coefficient, we differ from Roehrig et al. (2013), in that we do not use the Pearson coefficient, but the Spearman coefficient. The latter measures how well a given relationship between two variables (in our case between a time series and a copy of that same time series shifted by lag=$\tau$) can be described by a monotone function, rather than a linear function. We believe that this approach is more suitable for daily precipitation. $r_s$ is defined as follows:

$$r_s = \rho_{R(X),R(X)}(\tau) \quad = \quad \frac{cov(R(X_{t+\tau}), R(X_t))}{var(R(X), R(X))}. \tag{1}$$

Here $cov(R(X_{t+\tau}), R(X_t))$ is the autocovariance of the rank $R$ of $X$, which is normalized by the variance $var(R(X), R(X))$, which means $r_s \in [-1, 1]$. Overbars denote time means. Like Roehrig et al. (2013), we use a lag of $\tau = 1$ because at this timescale the autocorrelation potentially captures the influence of large-scale processes such as CCEWs on precipitation, but no processes related to the diurnal cycle. We investigate individual seasons, so the ranks $R$ range from 1 to 92 for the JAS season, for example. From here on, we will refer to the lag 1 autocorrelation simply as autocorrelation.

### 2.3 Wave Filtering

We apply the wave filtering method developed by Wheeler and Kiladis (1999) to analyze six types of CCEWs and the MJO, here ordered from slow (low frequency) to fast (high frequency): MJO, equatorial Rossby (ER), mixed Rossby gravity (MRG), Kelvin, tropical depression (TD), eastward inertio-gravity waves (EIG, with $n = 0$) and westward inertio-gravity waves (WIG, with $n = 1$). For the filtering we use the full time-series, between 20°S and 20°N. We furthermore remove the first three harmonics of the seasonal cycle, detrend the signal and taper the ends to zero, as described by Wheeler and Kiladis (1999). To retain the full wave signal in regions where convection only occurs in one hemisphere, we do not apply a symmetric/anitsymmetric decomposition (Kiladis et al., 2009). The filter parameters for the different wave types are the same as in Schlueter et al. (2019) for all waves except WIG waves, for which we choose the parameters in Wheeler and Kiladis (1999). We moreover distinguish between slow and fast Kelvin waves, with a cutoff frequency of 0.2 day$^{-1}$, to make the lower frequency limit of the fast Kelvin waves match the one of the TD and the EIG waves. The two Kelvin wave types share the cut-off frequency, which results in a small overlap of the filters. The Kelvin wave splitting is still relevant, however, since the autocorrelation corresponds to the ratio of high to low frequency variance on the order of a few days, and the Kelvin filter band ranges from

0.05 to 0.4 day$^{-1}$. We will refer to MJO, ER, MRG and slow Kelvin as slow waves and to fast Kelvin, TD, EIG and WIG as fast waves. The filtering is done using a Python script, which we partly based on Miyachi (2023) and Medeiros (2023). We validated the script by comparing our results for the IMERG data against the Tropical Rainfall Measuring Mission based results in Schlueter et al. (2019). We estimate the variance of precipitation explained by CCEWs by calculating the squared Pearson correlation coefficient between the precipitation anomalies (without the first three harmonics of the seasonal cycle, detrended and tapered) and the wave filtered precipitation anomalies respectively (Schlueter et al., 2019).

## 3 Results

### 3.1 Lag 1 Autocorrelation of Daily Precipitation in the Tropics

The autocorrelation pattern for daily precipitation in the tropics described by Roehrig et al. (2013) for GPCP also holds for IMERG and MSWEP, with some differences in the absolute values (see Figure C1). For the subsequent analysis we focus on IMERG, since it compares best to the gauge data in regions of low autocorrelation (not shown) and unlike MSWEP is not influenced by reanalysis. The IMERG autocorrelation is therefore also depicted in (a) and (b) of Figure 1, together with gauge station data (indicated by the circles). There are some outliers when comparing the gauge stations with IMERG, e.g. in the Pacific, where individual stations show very low autocorrelation values. We do not believe that the differences are due to the different periods covered by the satellite and ground-based observations, since there is very little interannual variability in the autocorrelation. Instead, we believe that the differences are due to the local perspective of the stations, while the satellite-based grid cells cover an area of 1°x1°.

Overall, the autocorrelation is positive throughout the tropics and especially over the ocean, which is highlighted by the mean autocorrelation values over tropical land and ocean, depicted in Figure 2 for July-September (JAS). During JAS, there are two distinctive exceptional regions over land, the Sahel together with equatorial Africa, as well as the most northern part of South America at the border to Central America, with considerably lower (slightly negative) values of autocorrelation. During DJF, regions with similarly low autocorrelation are located in Africa, south of the equator and in South America. To some extend also parts of the maritime continent show considerably lower autocorrelation values than the surrounding ocean during DJF. Interestingly, these patterns can not simply be described as a land-sea contrast, as clarified by the pattern of autocorrelation over the Asian monsoon and the Gulf of Mexico regions. From here on, we will focus on JAS, since during this season the most pronounced low autocorrelation region over Africa exists, where we will more specifically focus on the Sahel domain.

The multi-model mean autocorrelation of a broad selection of CMIP6 simulations (see Figure 1) captures the observed pattern in broad strokes, but clearly overestimates the persistence of rainfall everywhere and especially over Africa. This is summarized by the region means for the Sahel domain, tropical land and tropical ocean in Figure 2. The drastic overestimation of the autocorrelation in the Sahel domain, is similar to the findings of Roehrig et al. (2013) for CMIP5 models. The CMIP6 models also overestimate the autocorrelation over tropical land and ocean, whereas the difference to the observations is smallest over the ocean. These low resolution GCM models all use deep convection parameterization schemes; we next contrast their behavior with kilometer-scale models without convective parameterizations.

Panels (c) and (d) of Figure 1 depict the autocorrelation maps for ICON-5km. The model captures key features that we also find in the observations. Especially the regions of low autocorrelation are captured by the ICON-5km simulation, with the African region during JAS being the most pronounced. The mean autocorrelation values are in the

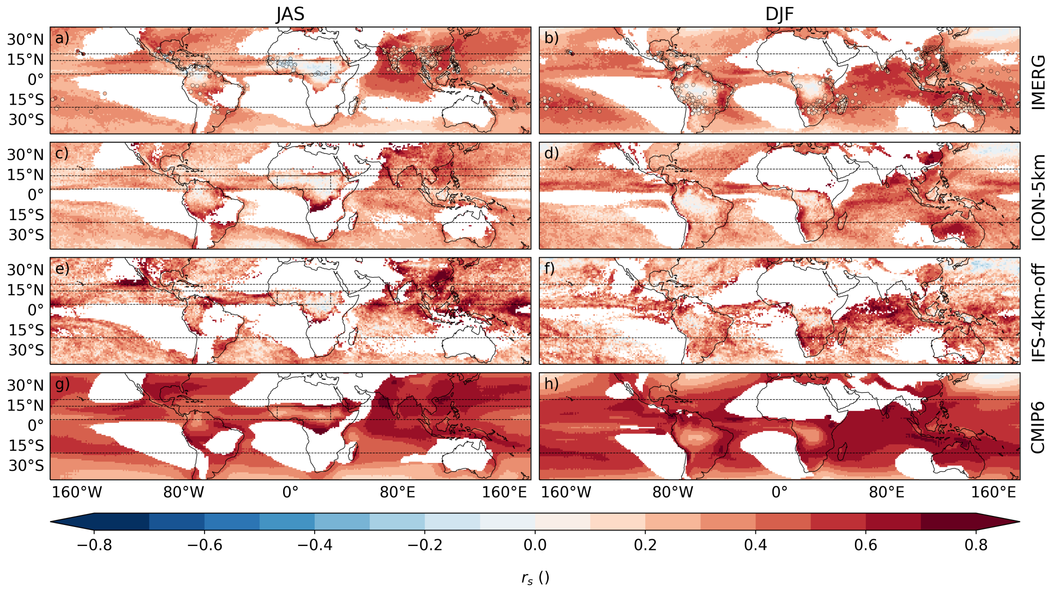



**Figure 1.** Lag 1 autocorrelation of daily precipitation on a 1° x 1° grid, datapoints with less than 1 mm/day mean precipitation are masked. IMERG for JAS in panel (a) and for DJF in panel (b). ICON-5km for JAS in panel (c) and for DJF in panel (d). IFS-4km-off for JAS in panel (e) and for DJF in panel (f). The circles in (a) and (b) indicate gauge station data. Stations with less than 1 mm/day mean precipitation are not shown. Black dashed lines indicate the domains for the further analysis: the Sahel domain (10°E-30°W, 5°N-15°N) the tropical domain (20°S-20°N) and the CCEW domain (5°N-15°N).

observational range in the Sahel region and generally also over tropical land (see Figure 2). However, the autocorrelation over the tropical ocean is lower than in the three observations.

The autocorrelation maps for IFS-4km-off are displayed in panels (e) and (f) of Figure 1. Again, the pattern matches the observations, but it is a little noisier than the IMERG and the ICON-5km data, since for this model setup we only have one simulated year. However, this shows that data from one single season is already enough for the autocorrelation pattern to appear. The mean autocorrelation values for IFS-4km-off in Figure 2 are close to the upper end of the observational range, or slightly above for the Sahel.

A number of additional simulations with the IFS model, produced within the nextGEMS project are available. IFS-4km-on, from the latest cycle 3 simulations is run with the deep convection scheme turned on, but a reduced vertical mass flux. Additionally, there are the 9 km simulations IFS-9km-off and IFS-9km-on. These simulations allow us to explore the influence of the deep convection paraterization scheme on the autocorrelation in the IFS model. Autocorrelation increases over the Sahel, tropical land and tropical ocean for both the 4.4 and the 9 km resolution IFS simulations (see Figure 2) when the deep convection parameterization is turned on. In fact, the increase in autocorrelation is relatively uniform across the tropics in both cases, with virtually no differences between land and ocean. This suggests that surface-atmosphere interactions are not driving factors for changes in the autocorrelation, when the deep convection parameterization is turned on.

MPAS-A also captures the autocorrelation pattern at both 3.75 and 7.5 km resolution when the deep convection is simulated explicitly (not shown). The mean auto-

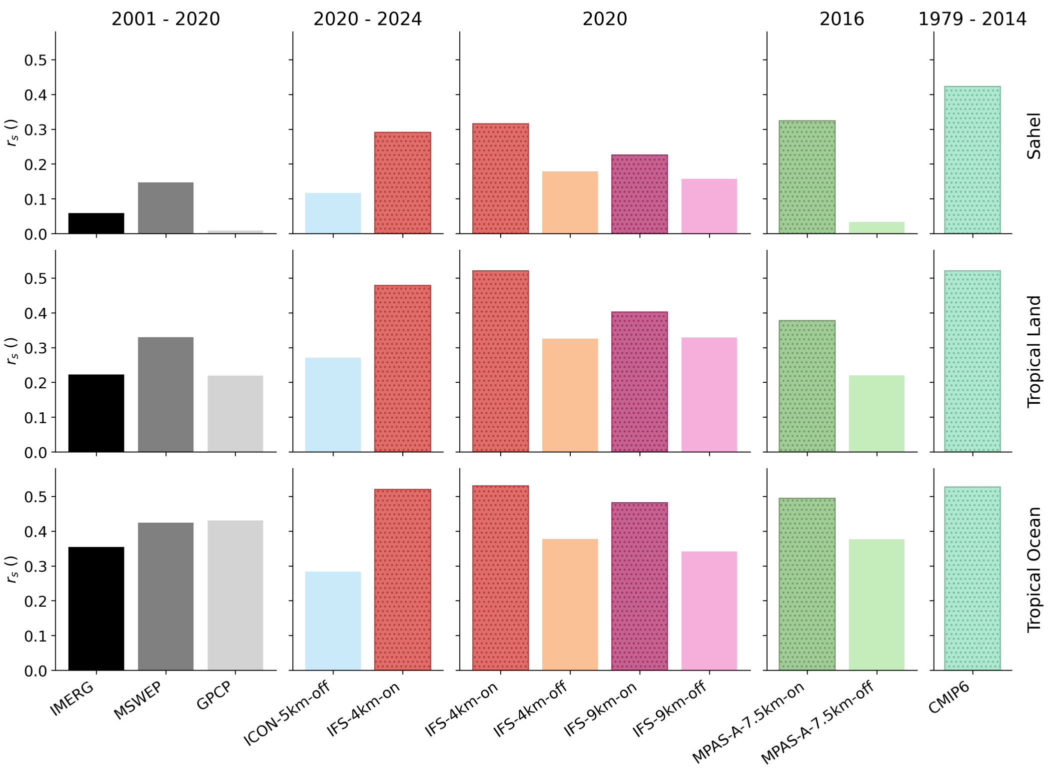



**Figure 2.** Mean lag 1 autocorrelation of daily precipitation calculated for July-September (JAS). In the first row calculated for the Sahel domain (10°E-30°W, 5°N-15°N), the second row for tropical land and in the third row for tropical ocean (20°S-20°N). Depicted in the first column from the left are the values for the observations, in the second column for the nextGEMS cycle 3 simulations, in the third column for different nextGEMS IFS simulations (cycle 2 and 3), in the fourth column for two MPAS-A simulations (simulation period only from August 1st to September 10th 2016) and in the fifth column the CMIP6 mean is depicted. Dotted bars indicate simulations with active deep convection parameterization. Gridpoints with less than 1 mm/day mean precipitation are not included.

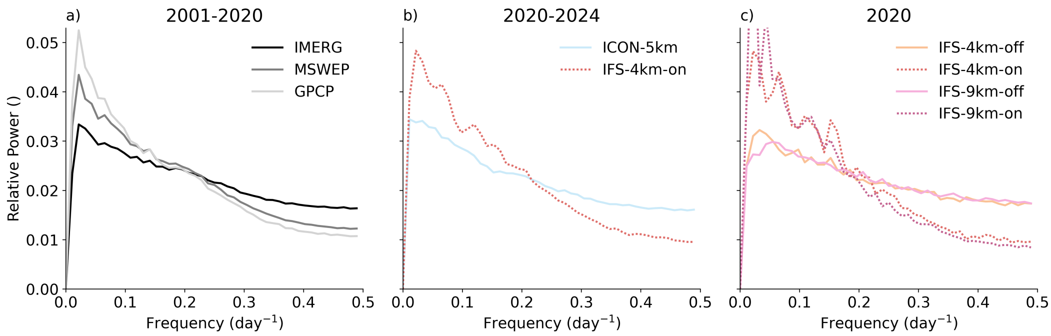



**Figure 3.** Relative power spectra calculated from daily precipitation, with the observations in (a), the nextGEMS cycle 3 simulations in (b) and the different nextGEMS IFS simulations (cycle 2 and 3) in (c). Gridpoints with less than 1 mm/day mean precipitation are not included.

correlation values for the MPAS-A-7.5km simulations in Figure 2 are consistent with the influence of the deep convection parameterization on the autocorrelation in IFS. As for the IFS model, the autocorrelation increases when the deep convection parameterization is turned on. The increase is much stronger over the Sahel than over the rest of the tropics in MPAS-A-7.5km (similarly in MPAS-A-3.75km, not shown). However, it should be kept in mind that these simulations only cover a range of 40 days.

The increase in autocorrelation when the deep convection parameterization is turned on is caused by an increase in high frequency variance (meaning on the time-scale of a few days) relative to low frequency variance in the IFS model. This is demonstrated in Figure 3, which depicts the tropical mean relative power spectra of daily precipitation for the nextGEMS simulations and the observations for JAS. When the IFS model is run with deep convection treated explicitly, the relative power spectra agree very well with ICON-5km and the observations. Whereas, when the parameterization is turned on, power shifts to lower frequencies and the spectra drop out of the observational range.

### 3.2 Relative CCEW Contribution to Precipitation Variance

A considerable fraction of the precipitation variance that underlies the power spectra presented in Figure 3 are related to CCEWs. To investigate the role of CCEWs in the origin of the autocorrelation pattern we therefore calculate the precipitation variance fraction on a daily time-scale for the different wave types in the CCEW band defined in Schlueter et al. (2019) for JAS (see Figure 1). Following Schlueter et al. (2019) this is achieved by calculating the squared Pearson correlation coefficients between precipitation anomalies and wave filtered precipitation anomalies of the latitudinal mean of the CCEW band. We subsequently take the longitudinal mean, to generate the mean values presented in Figure 4. Since this correlation method does not take interactions between waves into account, the sum of the precipitation variance fractions attributed to CCEWs can be interpreted as the maximum variance fraction explainable by CCEWs (Schlueter et al., 2019). It should be noted that only a small part of the WIG frequency spectrum can be filtered from daily data, which is why they only contribute relatively small variance fractions at this time-scale.

Differences in the summed CCEW variance fraction in the observations in Figure 4 are as large as 10 percentage points between IMERG and GPCP, which delivers the highest summed fraction of more than 45 %. These differences arise mainly due to differences in the slow CCEW fractions (MJO, ER and MRG).

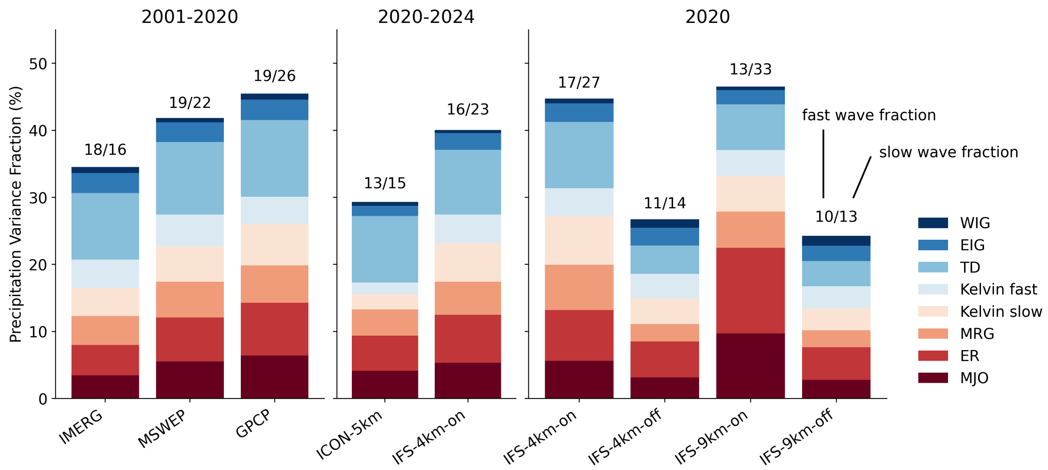



**Figure 4.** Precipitation variance fractions attributed to convectively coupled equatorial waves (CCEWs) calculated as the squared Pearson correlation between precipitation anomalies and wave filtered precipitation anomalies for July-September (JAS) in the CCEW band (5°N-15°N). From left to right, the first three bars depict the mean values for the observations, the next two bars the values for the nextGEMS cycle 3 simulations and the last four bars the values for different nextGEMS IFS simulations (cycle 2 and 3). The fractions over the bars denote the summed variance fractions of the fast waves over the summed variance fractions of the slow waves.

The ICON-5km simulations produce a summed CCEW variance fraction smaller than the observational range. The variance fractions for the slow CCEWs are comparable to IMERG, but the model produces lower values for both fast and slow Kelvin and EIG waves.

For the IFS simulations with the deep convection parameterization turned on, the summed CCEW variance fractions are close to the upper end of the observational range. The simulations with deep convection simulated explicitly, on the other hand, deliver summed CCEW variance fractions that are considerably smaller and closer to the one of ICON-5km. The changes between deep convection parameterization on and off simulations are not uniform across wave types. For example, there are differences in the variance fractions of the TD waves. The comparably small fractions for this wave type in IFS-4km-off and IFS-9km-off are also a difference to the ICON-5km simulation, which in turn delivers smaller fractions for Kelvin and EIG waves. The largest changes, however, are in the variance fractions of the slow waves, which are particularly pronounced for the IFS-9km simulations. These changes are related to the autocorrelation in the sense that the larger variance fraction in the slow waves, when the deep convection parameterization is turned on, increases the power in the lower frequencies of the power spectrum (see Figure 3).

The overall importance of CCEWs in the origin of the tropical autocorrelation pattern is difficult to accurately quantify. However, we find that the ICON-5km and the IFS-4km-off models, while accurately reproducing the autocorrelation pattern, produce summed CCEW variance fractions lower than the observational range. The IFS-4km-on model on the other hand, produces CCEW variance fractions close to the observations, while overestimating autocorrelation across the tropics.

## 4 Conclusions

We study the lag 1 autocorrelation of daily precipitation as an easy-to-compute diagnostic of tropical precipitation variability in kilometer-scale climate models. The diagnostic is helpful to investigate the persistence of precipitation across the tropics.

Our results shows that the autocorrelation pattern first presented by Roehrig et al. (2013) is robust across satellite-based observations and gauge stations. The pattern can be summarized as one of largest autocorrelation over ocean; smaller, but still positive autocorrelation over tropical land; and much smaller or even negative autocorrelation in the core of the rainy regions of Africa and South America. The kilometer-scale climate models of the nextGEMS project and MPAS-A produce a similar pattern, in contrast to a wide range of CMIP6 models. In the IFS model the autocorrelation is very comparable to observations when the deep convection is simulated explicitly, but overestimated across the tropics when the deep convection parameterization is turned on. This effect occurs over several horizontal resolutions and even if the parameterized vertical mass flux is strongly reduced. Hence, we find a distinct dependence on the representation of deep convection, rather than simply horizontal resolution, independent of the surface being land or ocean. This result is confirmed by our analysis of global kilometer-scale MPAS-A simulations, which show the same dependence on the representation of deep convection.

We attribute precipitation variance to different kinds of convectively coupled equatorial waves (CCEW). When deep convection is treated explicitly, the power spectrum of precipitation becomes whiter, reducing fractional variance from CCEW as a set, and specifically by depressing the role of slow waves. We show that a climate model with deep convection parameterization turned on, namely IFS-4km-on, produces CCEW related precipitation fractions that compare relatively well to the observations and at the same time overestimate the autocorrelation. On the other hand different kilometer-scale models, ICON-5km and IFS-4km-off, reproduce the autocorrelation pattern, but underestimate the CCEW related precipitation fraction. In the IFS model, turning off the deep convection parameterization leads to a decrease of variance fractions related to slow CCEWs and an increase in the variance fraction related to tropical depression type waves. The latter is in contrast to what Judt and Rios-Berrios (2021) reported for MPAS-A. Taking these findings into account, we assume that CCEWs, while probably playing a role, are not the main driving factor for the observed autocorrelation pattern.

The origin of the autocorrelation pattern could lie on scales finer than those of CCEWs. A comparison of the maps of the autocorrelation and observation-based statistics of mesocale convective systems (MCSs) from Feng et al. (2021) indicates that the regions of low autocorrelation are also regions with high MCS precipitation fractions and MCS numbers. However, there are also regions over the ocean (e.g. the Indian ocean) where this association does not hold. Moreover, Mathon et al. (2002) showed that over the central Sahel, where autocorrelation is low, organized MCSs (OMCSs), account for 90 % of the seasonal precipitation, while only representing 12 % of the total MCS count. The low frequency of OMCSs may therefore indicate that MCSs are not the main drivers of the autocorrelation pattern either. A deeper analysis of MCSs and convective organization on even smaller scales in the nextGEMS simulations is necessary to further constrain the mechanisms behind the autocorrelation pattern.

Regardless the exact composition of phenomena that lead to the observed pattern of autocorrelation of daily precipitation, the autocorrelation is an easily computed metric that encapsulates these underlying atmospheric processes and provides a simple diagnostic for tropical precipitation in GCMs. Our finding that kilometer-scale models are able to reproduce the autocorrelation pattern furthermore demonstrates that these models offer a unique possibility to study tropical precipitation and the related atmospheric dynamics.

## Acknowledgments

nextGEMS is funded through the European Union's Horizon 2020 research and innovation program under the grant agreement number 101003470.

## Data Availability Statement

Selected data of the nextGEMS simulations can be found at `https://doi.org/10.26050/WDCC/nextGEMS_cyc2` (Wieners et al., 2023) for cycle 2 and at `https://doi.org/10.26050/WDCC/nextGEMS_cyc3` (Koldunov et al., 2023) for cycle 3. The CMIP6 model data can be obtained from the Earth System Grid Foundation. MPAS-A data is from `https://doi.org/10.5065/v05b-9e73`. IMERG data were accessed via NASA's Goddard Earth Sciences Data and Information Services Center (`https://disc.gsfc.nasa.gov/datasets/GPM_3IMERGDF_06/summary?keywords=imerg%20final%20v06`). MSWEP data were accessed via `https://www.gloh2o.org/mswep/`. GPCP data can be obtained from `https://www.ncei.noaa.gov/access/metadata/landing-page/bin/iso?id=gov.noaa.ncdc:C00999`. The gauge data were accessed via `http://iridl.ldeo.columbia.edu/SOURCES/.NOAA/.NCDC/.DAILY/.GLOBALSOD/?sem=iridl%3ADCAtmosphere`. The post-processed data will be made available and get a doi as soon as the paper is under review. The analysis scripts are available under `https://ucloud.univie.ac.at/index.php/s/ctGcFy3nw6yF4Zb` and will get a doi as soon as the paper is under review.

## Competing Interests

The authors declare that they have no conflict of interest.

## References

Adler, R., Wang, Jian-Jian, Sapiano, Mathew, Huffman, George, Bolvin, David, Nelkin, Eric, & NOAA CDR Program. (2017, May). *Global Precipitation Climatology Project (GPCP) Climate Data Record (CDR), Version 1.3 (Daily).* (NOAA National Centers for Environmental Information.) doi: 10.7289/V5RX998Z

Beck, H. E., Wood, E. F., Pan, M., Fisher, C. K., Miralles, D. G., Van Dijk, A. I. J. M., ... Adler, R. F. (2019, March). MSWEP V2 Global 3-Hourly 0.1° Precipitation: Methodology and Quantitative Assessment. *Bulletin of the American Meteorological Society*, *100*(3), 473–500. Retrieved 2024-01-15, from `https://journals.ametsoc.org/view/journals/bams/100/3/bams-d-17-0138.1.xml` doi: 10.1175/BAMS-D-17-0138.1

Cheng, Y., Dias, J., Kiladis, G., Feng, Z., & Leung, L. R. (2023, May). Mesoscale Convective Systems Modulated by Convectively Coupled Equatorial Waves. *Geophysical Research Letters*, *50*(10), e2023GL103335. Retrieved 2024-01-08, from `https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2023GL103335` doi: 10.1029/2023GL103335

Cho, H.-K., Bowman, K. P., & North, G. R. (2004, November). Equatorial Waves Including the Madden–Julian Oscillation in TRMM Rainfall and OLR Data. *Journal of Climate*, *17*(22), 4387–4406. Retrieved 2023-10-03, from `http://journals.ametsoc.org/doi/10.1175/3215.1` doi: 10.1175/3215.1

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016, May). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, *9*(5), 1937–1958. Retrieved 2024-01-30, from `https://gmd.copernicus.org/articles/9/1937/2016/` doi: 10.5194/gmd-9-1937-2016

Feng, Z., Leung, L. R., Hardin, J., Terai, C. R., Song, F., & Caldwell, P. (2023, February). Mesoscale Convective Systems in DYAMOND Global Convection-

Permitting Simulations. *Geophysical Research Letters*, *50*(4), e2022GL102603. Retrieved 2024-01-08, from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022GL102603 doi: 10.1029/2022GL102603

Feng, Z., Leung, L. R., Liu, N., Wang, J., Houze, R. A., Li, J., … Guo, J. (2021, April). A Global High-Resolution Mesoscale Convective System Database Using Satellite-Derived Cloud Tops, Surface Precipitation, and Tracking. *Journal of Geophysical Research: Atmospheres*, *126*(8), e2020JD034202. Retrieved 2024-01-08, from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2020JD034202 doi: 10.1029/2020JD034202

Fiedler, S., Crueger, T., D'Agostino, R., Peters, K., Becker, T., Leutwyler, D., … Stevens, B. (2020, September). Simulated Tropical Precipitation Assessed across Three Major Phases of the Coupled Model Intercomparison Project (CMIP). *Monthly Weather Review*, *148*(9), 3653–3680. Retrieved 2024-01-31, from http://journals.ametsoc.org/doi/10.1175/MWR-D-19-0404.1 doi: 10.1175/MWR-D-19-0404.1

Hohenegger, C., Korn, P., Linardakis, L., Redler, R., Schnur, R., Adamidis, P., … Stevens, B. (2023, January). ICON-Sapphire: simulating the components of the Earth system and their interactions at kilometer and subkilometer scales. *Geoscientific Model Development*, *16*(2), 779–811. Retrieved 2023-06-20, from https://gmd.copernicus.org/articles/16/779/2023/ doi: 10.5194/gmd-16-779-2023

Huffman, G.J., E.F. Stocker, D.T. Bolvin, E.J. Nelkin, & Jackson Tan. (2019). *GPM IMERG Final Precipitation L3 1 day 0.1 degree x 0.1 degree V06, Edited by Andrey Savtchenko, Greenbelt, MD, Goddard Earth Sciences Data and Information Services Center (GES DISC).* doi: 10.5067/GPM/IMERGDF/DAY/06

Judt, F., & Rios-Berrios, R. (2021, July). Resolved Convection Improves the Representation of Equatorial Waves and Tropical Rainfall Variability in a Global Nonhydrostatic Model. *Geophysical Research Letters*, *48*(14). Retrieved 2023-08-22, from https://onlinelibrary.wiley.com/doi/10.1029/2021GL093265 doi: 10.1029/2021GL093265

Kiladis, G. N., Wheeler, M. C., Haertel, P. T., Straub, K. H., & Roundy, P. E. (2009, April). Convectively coupled equatorial waves. *Reviews of Geophysics*, *47*(2), RG2003. Retrieved 2023-06-09, from http://doi.wiley.com/10.1029/2008RG000266 doi: 10.1029/2008RG000266

Koldunov, N., Kölling, T., Pedruzo-Bagazgoitia, X., Rackow, T., Redler, R., Sidorenko, D., … Ziemen, F. A. (2023). *nextGEMS: output of the model development cycle 3 simulations for ICON and IFS.* World Data Center for Climate (WDCC) at DKRZ. Retrieved from https://doi.org/10.26050/WDCC/nextGEMS_cyc3 doi: 10.26050/WDCC/nextGEMS_cyc3

Mathon, V., Laurent, H., & Lebel, T. (2002, November). Mesoscale Convective System Rainfall in the Sahel. *Journal of Applied Meteorology*, *41*(11), 1081–1092. Retrieved 2023-07-06, from http://journals.ametsoc.org/doi/10.1175/1520-0450(2002)041<1081:MCSRIT>2.0.CO;2 doi: 10.1175/1520-0450(2002)041⟨1081:MCSRIT⟩2.0.CO;2

Medeiros, B. (2023, October). *wavenumber_frequency.* Retrieved from https://github.com/brianpm/wavenumber_frequency

Miyachi, T. (2023, January). *mcclimate.* Retrieved from https://github.com/tmiyachi/mcclimate/tree/master

Moon, H., Gudmundsson, L., Guillod, B. P., Venugopal, V., & Seneviratne, S. I. (2019, October). Intercomparison of daily precipitation persistence in multiple global observations and climate models. *Environmental Research Letters*, *14*(10), 105009. Retrieved 2024-02-14, from https://iopscience.iop.org/article/10.1088/1748-9326/ab4169 doi: 10.1088/1748-9326/ab4169

NOAA National Centers of Environmental Information. (1999). *Global Surface Sum-*

*mary of the Day - GSOD. 1.0.*

Palmer, T., & Stevens, B. (2019, December). The scientific challenge of understanding and estimating climate change. *Proceedings of the National Academy of Sciences*, *116*(49), 24390–24395. Retrieved 2024-01-31, from https://pnas.org/doi/full/10.1073/pnas.1906691116 doi: 10.1073/pnas.1906691116

Rackow et al. (n.d.). Multi-year simulations at kilometre scale with the Integrated Forecasting System coupled to FESOM2.5/NEMOv3.4. (in preparation)

Roehrig, R., Bouniol, D., Guichard, F., Hourdin, F., & Redelsperger, J.-L. (2013, September). The Present and Future of the West African Monsoon: A Process-Oriented Assessment of CMIP5 Simulations along the AMMA Transect. *Journal of Climate*, *26*(17), 6471–6505. Retrieved 2023-05-11, from http://journals.ametsoc.org/doi/10.1175/JCLI-D-12-00505.1 doi: 10.1175/JCLI-D-12-00505.1

Schlueter, A., Fink, A. H., Knippertz, P., & Vogel, P. (2019, March). A Systematic Comparison of Tropical Waves over Northern Africa. Part I: Influence on Rainfall. *Journal of Climate*, *32*(5), 1501–1523. Retrieved 2023-05-11, from http://journals.ametsoc.org/doi/10.1175/JCLI-D-18-0173.1 doi: 10.1175/JCLI-D-18-0173.1

Skamarock, W. C., Klemp, J. B., Duda, M. G., Fowler, L. D., Park, S.-H., & Ringler, T. D. (2012, September). A Multiscale Nonhydrostatic Atmospheric Model Using Centroidal Voronoi Tesselations and C-Grid Staggering. *Monthly Weather Review*, *140*(9), 3090–3105. Retrieved 2024-01-29, from http://journals.ametsoc.org/doi/10.1175/MWR-D-11-00215.1 doi: 10.1175/MWR-D-11-00215.1

Stevens, B., Acquistapace, C., Hansen, A., Heinze, R., Klinger, C., Klocke, D., … Zängl, G. (2020). The Added Value of Large-eddy and Storm-resolving Models for Simulating Clouds and Precipitation. *Journal of the Meteorological Society of Japan. Ser. II*, *98*(2), 395–435. Retrieved 2023-05-11, from https://www.jstage.jst.go.jp/article/jmsj/98/2/98_2020-021/_article doi: 10.2151/jmsj.2020-021

Tomassini, L. (2018, December). Mesoscale Circulations and Organized Convection in African Easterly Waves. *Journal of the Atmospheric Sciences*, *75*(12), 4357–4381. Retrieved 2023-05-11, from https://journals.ametsoc.org/doi/10.1175/JAS-D-18-0183.1 doi: 10.1175/JAS-D-18-0183.1

Wheeler, M., & Kiladis, G. N. (1999, February). Convectively Coupled Equatorial Waves: Analysis of Clouds and Temperature in the Wavenumber–Frequency Domain. *Journal of the Atmospheric Sciences*, *56*(3), 374–399. Retrieved 2023-10-03, from http://journals.ametsoc.org/doi/10.1175/1520-0469(1999)056<0374:CCEWAO>2.0.CO;2 doi: 10.1175/1520-0469(1999)056⟨0374:CCEWAO⟩2.0.CO;2

Wieners, K.-H., Ziemen, F. A., Koldunov, N., Pedruzo-Bagazgoitia, X., Rackow, T., Redler, R., … Kölling, T. (2023). *nextGEMS: output of the model development cycle 2 simulations for ICON and IFS.* World Data Center for Climate (WDCC) at DKRZ. Retrieved from https://doi.org/10.26050/WDCC/nextGEMS_cyc2 doi: 10.26050/WDCC/nextGEMS_cyc2

## Appendix A Overview of nextGEMS Simulations

The key settings and the simulation acronyms of the nextGEMS simulations we analyze are presented in Table A1.

**Table A1.** Settings of the nextGEMS simulations used in our analysis.

| Simulation acronym | Atmosphere model | Cycle | Horizontal resolution | Simulation years | Deep convection parameterization |
|---|---|---|---|---|---|
| ICON-5km | ICON-Sapphire | 3 | 5 km | 2020-2024 | off |
| IFS-4km-on | IFS | 3 | 4.4 km | 2020-2024 | on (reduced) |
| IFS-4km-off | IFS | 2 | 4.4 km | 2020 | off |
| IFS-9km-on | IFS | 2 | 9 km | 2020 | on |
| IFS-9km-off | IFS | 2 | 9 km | 2020 | off |

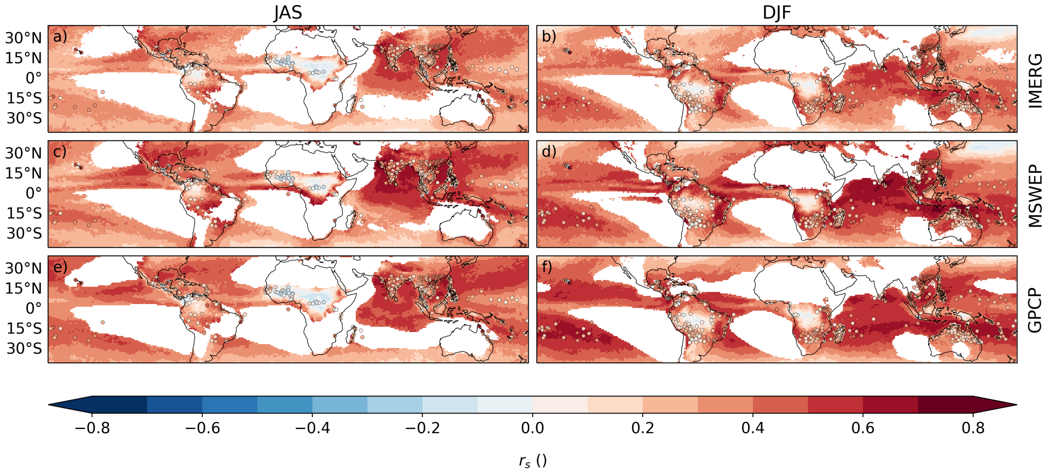



**Figure C1.** Lag 1 autocorrelation of daily precipitation on a 1° x 1° grid, datapoints with less than 1 mm/day mean precipitation are masked. IMERG for JAS in panel (a) and for DJF in panel (b). MSWEP for JAS in panel (c) and for DJF in panel (d). GPCP for JAS in panel (e) and for DJF in panel (f). Circles indicate gauge station data. Stations with less than 1 mm/day mean precipitation are not shown.

## Appendix B List of CMIP6 Simulations

We use amip r1i1p1f1 CMIP6 simulations (Eyring et al., 2016) for our analysis, included are the following models:

ACCESS-CM2, ACCESS-ESM1-5, BCC-ESM1, CAMS-CSM1-0C, ESM2-FV2, CESM2-WACCM, CESM2, CMCC-CM2-HR4, CMCC-CM2-SR5, CanESM5, E3SM-1-0, E3SM-2-0, EC-Earth3-AerChem, EC-Earth3-CC, EC-Earth3-Veg-LR, EC-Earth3-Veg, EC-Earth3, FGOALS-f3-L, FGOALS-g3, GFDL-CM4, GFDL-ESM4, ICON-ESM-LR, IITM-ESM, INM-CM4-8, INM-CM5-0, IPSL-CM6A-LR, KACE-1-0-G, MIROC6, MPI-ESM-1-2-HAM, MPI-ESM1-2-HR, MPI-ESM1-2-LR, MRI-ESM2-0, NESM3, NorCPM1, NorESM2-LM and TaiESM1.

## Appendix C Autocorrelation Maps for Observations

Figure C1 depicts the lag 1 autocorrelation maps for the analyzed observations.