https://doi.org/10.5194/cp-2020-103
Preprint. Discussion started: 27 August 2020

Technical Note: Characterising and comparing different palaeoclimates with dynamical systems theory

Gabriele Messori[1, 2] and Davide Faranda[3, 4, 5]

[1]Department of Earth Sciences, Uppsala University, and Centre of Natural Hazards and Disaster Science (CNDS), Uppsala, Sweden.
[2]Department of Meteorology and Bolin Centre for Climate Research, Stockholm University, Stockholm, Sweden
[3]Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, Gif-sur-Yvette, France.
[4]London Mathematical Laboratory, London, U. K.
[5]LMD/IPSL, Ecole Normale Superieure, PSL research University, Paris, France

Correspondence: Gabriele Messori (gabriele.messori@geo.uu.se)

Abstract. Numerical climate simulations produce vast amounts of high-resolution data. This poses new challenges to the palaeoclimate community – and indeed to the broader climate community – in how to efficiently process and interpret model output. The palaeoclimate community also faces the additional challenge of having to characterise and compare a much broader range of climates than encountered in other subfields of climate science. Here we propose an analysis framework, grounded in dynamical systems theory, which may contribute to overcome these challenges. The framework enables to characterise the dynamics of a given climate through a small number of metrics. These may be applied to individual climate variables or to diagnose the coupling between different variables. To illustrate its applicability, we analyse three numerical simulations of mid-Holocene climates over North Africa, under different boundary conditions. We find that the three simulations produce climate systems with different dynamical properties, which are reflected in the dynamical systems metrics. We conclude that the dynamical systems framework holds significant potential for analysing palaeoclimate simulations. At the same time, an appraisal of the framework's limitations suggests that it should be viewed as a complement to more conventional analyses, rather than as a wholesale substitute.

1 Motivation

Numerical climate models have enjoyed widespread use in palaeoclimate studies, from early investigations based on simple thermodynamic or general circulation models (e.g. Gates, 1976; Donn and Shaw, 1977; Barron et al., 1980) to the state-of-the-art models being used in the fourth phase of the Paleoclimate Modelling Intercomparison Project (PMIP4, Kageyama et al., 2018). Compared to data from palaeo-archives, which is typically geographically sparse and with a low temporal resolution even for the more recent palaeoclimates (e.g. Bartlein et al., 2011), numerical climate simulations produce a vast amount of

https://doi.org/10.5194/cp-2020-103
Preprint. Discussion started: 27 August 2020

horizontally gridded, vertically resolved and temporally high-resolution data. This poses new challenges to the palaeoclimate community in how to efficiently process and interpret model output – indeed an issue which is faced by the broader climate community (Schnase et al., 2016).

A related, yet distinct, challenge faced by the palaeoclimate community are the large uncertainties often found in palaeo-simulations. These reflect the uncertainties in palaeo-archives and in our knowledge of the boundary conditions and forcings affecting past climates (e.g. Kageyama et al., 2018). Thus, different simulations of the climate in the same period and region may yield very different results. This emerges in both reconstructions of climates from millions of years ago, such as the mid-Pliocene warm period over 3 Myr BP (e.g. Haywood et al., 2013) and in climates much closer to us, such as the mid-Holocene around 6000 yr BP (e.g. Pausata et al., 2016). Characterising and understanding these discrepancies, requires analysis tools which may efficiently distil the differences between the simulated palaeoclimates.

Here, we propose an analysis framework which addresses the challenges of efficiently processing and interpreting large amounts of model output to compare different simulated palaeoclimates. The framework is grounded in dynamical systems theory, and enables to characterise the dynamics of a given climate through a small number of metrics. These may be applied to individual climate variables or to diagnose the coupling between different variables. In other words, the dynamical information embedded in 3-dimensional (latitude, longitude and time) or 4-dimensional (latitude, longitude, pressure level and time) data, commonly produced by climate models, can be projected onto one-dimensional metrics. These may then be interpreted and compared with relative ease.

The rest of this technical note is structured as follows: in Section 2 we briefly describe the theory underlying the dynamical systems framework, and provide both a qualitative and a technical description of the metrics. We further provide a link to a repository from which MatLab code to implement the metrics may be obtained. In Section 3, we illustrate the application of the metrics to palaeoclimate data and their interpretation by using a set of recent numerical simulations for the mid-Holocene climate in North Africa. This is not meant to be a comprehensive analysis, but rather provide a flavour of the information provided by our framework. We conclude in Section 4 by reflecting on the framework's strengths and limitations and by outlining potential applications in future palaeoclimate studies.

2 A qualitative overview and theoretical underpinnings of the dynamical systems framework

2.1 A qualitative overview of the dynamical systems framework

The dynamical systems framework we propose rests on three indicators. All are instantaneous in time, meaning that given a long time series of model data, they provide a value for each timestep. For example, if we were to analyse daily latitude-longitude sea-level pressure (SLP) over 30 years, we would then have 30×365 values for each indicator.

The first indicator, termed *local dimension* (d), provides a proxy for the number of active degrees of freedom of the system about a state of interest (Lucarini et al., 2016; Faranda et al., 2017a). In other words, the value of d for a given day in our SLP dataset tells us how the SLP in the chosen geographical region can evolve to or from the pattern it displays on that day. The number of different possible evolutions is proportional to the number of degrees of freedom, and therefore days with a low

https://doi.org/10.5194/cp-2020-103
Preprint. Discussion started: 27 August 2020

(high) local dimension correspond to SLP patterns that derive from and may evolve into a small (large) number of other SLP patterns in the preceding and following days.

The second indicator, termed *persistence* (θ^{-1}), measures the mean residence time of the system around a given state, and is bounded to the range $1 \leq \theta^{-1} < +\infty$. In other words, if a given day in our SLP dataset has a high (low) persistence, the SLP pattern on that day has evolved slowly (rapidly) from and will evolve slowly (rapidly) to a different SLP pattern. The higher the persistence, the more likely it is that the SLP patterns on the days immediately preceding and following the chosen day will resemble the SLP pattern of that day. This metric is related to, yet distinct from, the notion of persistence issuing from weather regimes and similar partitionings of the atmospheric variability (Hochman et al., 2019, 2020).

Both indicators may be used to characterise the dynamics underlying complex systems, including the Earth's climate (e.g. Faranda et al., 2017a; Buschow and Friederichs, 2018; Brunetti et al., 2019). On a more practical level, they can also be linked to the notion of intrinsic predictability of the system's different states. A state with a low d and high θ^{-1} will afford a better predictability than one with a high d and low θ^{-1}. For more detailed discussions on this topic, and a comparison to the conventional idea of predictability as evaluated through numerical weather forecasts, see Messori et al. (2017); Scher and Messori (2018) and Faranda et al. (2019a).

Both d and θ^{-1} may in principle be computed for more than one climate variable jointly (Faranda et al., 2020; De Luca et al., 2020) but here we will focus on their univariate implementation. Unlike the first two, the third metric we propose here, termed *co-recurrence ratio* (α), is exclusively defined for two or more variables, and is bounded to the range $0 \leq \alpha \leq 1$. Given two climate variables, α diagnoses the extent to which their recurrences co-occur, and hence provides a measure of the coupling between the two variables (Faranda et al., 2020). As example, imagine that we now have both our SLP dataset and a corresponding precipitation dataset, which also provides daily values on a latitude-longitude grid. If α on a given day is large, then every time we have an SLP pattern on another day which closely resembles the SLP pattern of the chosen day (i.e. a *recurrence*), the precipitation pattern on that other day will also resemble the precipitation pattern of the chosen day. In other words, recurrences of similar SLP patterns lead to recurrences of similar precipitation patterns, which would suggest that the two variables are highly coupled. If α on a given day is small, then every time we have an SLP pattern on another day which closely resembles the SLP pattern of the chosen day, the precipitation pattern on that other day will not resemble the precipitation pattern of the chosen day. In other words, recurrences of similar SLP patterns do not lead to recurrences of similar precipitation patterns, which would suggest that the two variables are weakly coupled. We note that α may not be interpreted in terms of causation. However, since the joint recurrence of two fields implies the existence of a common underlying dynamics, the information it provides is nonetheless grounded in the physics of the system being analysed. Finally, α provides a very different information from many other conventional statistical dependence measures, since it gives a value for every timestep in the dataset.

https://doi.org/10.5194/cp-2020-103
Preprint. Discussion started: 27 August 2020

2.2 Theoretical underpinnings of the dynamical systems framework

The three dynamical systems metrics described above, issue from the combination of extreme value theory with Poincaré recurrences (Freitas et al., 2010; Lucarini et al., 2012, 2016). Given a dynamical system with a trajectory $x(t)$, and a state of interest ζ_x, we define logarithmic returns as:

$$g(x(t),\zeta_x) = -\log[\mathrm{dist}(x(t),\zeta_x)] \tag{1}$$

In this note, $dist$ is the Euclidean distance between two vectors. More generally, $dist$ can be a distance function which tends to zero as the two vectors increasingly resemble each other. For the implications of using $dist$ other than the Euclidean distance, the reader is referred to Lucarini et al. (2016) and Faranda et al. (2019b). The $-\log$ implies that $g(x(t),\zeta_x)$ attains large values when $x(t)$ and ζ_x are close to one another. We thus have a timeseries g of logarithmic returns, which is large if x at a specific time resembles the state of interest ζ_x.

We next define a high threshold $s(q,\zeta_x)$ as the q*th* quantile of $g(x(t),\zeta_x)$, and define exceedances $u(\zeta_x) = g(x(t),\zeta_x) - s(q,\zeta_x) \; \forall \; g(x(t),\zeta_x) > s(q,\zeta_x)$. We then leverage the Freitas-Freitas-Todd theorem (Freitas et al., 2010; Lucarini et al., 2012), which states that the cumulative probability distribution $F(u,\zeta)$ converges to the exponential member of the Generalised Pareto Distribution:

$$F(u,\zeta_x) \simeq \exp\left[-\vartheta(\zeta_x)\frac{u(\zeta_x)}{\sigma(\zeta_x)}\right] \tag{2}$$

Here, u and σ are parameters of the distribution which depend on the chosen ζ_x, while ϑ is the extremal index (Moloney et al., 2019). We estimate the latter using the Süveges Maximum Likelihood Estimator (Süveges, 2007). We then obtain the persistence as: $\theta^{-1}(\zeta_x) = \Delta t/\vartheta(\zeta_x)$, where Δt is the timestep of the data being analysed, and the local dimension as $d(\zeta_x) = 1/\sigma(\zeta_x)$.

Finally, we define the co-recurrence ratio by considering two trajectories $x(t)$ and $y(t)$, and a corresponding joint state of interest $\zeta = (\zeta_x, \zeta_y)$. We then have that:

$$\alpha(\zeta) = \frac{\nu[g(x(t)) > s_x(q) \cup g(y(t)) > s_y(q)]}{\nu[g(x(t)) > s_x(q)]} \tag{3}$$

Here, $\nu[-]$ is the number of events satisfying condition $[-]$, and all other variables are defined as before.

The analytical derivation of the above framework assumes an underlying stationary Axiom A system, in the limit of infinitely long time-series. Amongst other things, this implies that it assumes a dynamical system with homogeneous properties along its trajectory. When computing the indicators for climate data, one has to take into account both the finite length of the datasets, and non-stationarities such as those issuing from internal low-frequency variability or varying external forcing. A formal justification of the applicability of the dynamical systems metrics to finite data issues from the results of Caby et al. (2020).

https://doi.org/10.5194/cp-2020-103
Preprint. Discussion started: 27 August 2020

There, the authors show that finite-time deviations of d and θ from the asymptotic, unknown values contain information about the underlying system, since they are linked to the presence of unstable or periodic points of the dynamics. Similarly, both analytical and empirical evidence from Pons et al. (2020) shows that, although affected by the curse of dimensionality, estimates of d from finite timeseries may be used in a relative sense to characterise the dynamics of a system – i.e. by comparing values of d to one another. The conclusions drawn from these more theoretical results match those issuing from empirical tests on climate timeseries of finite length conducted by Buschow and Friederichs (2018). In practice, the two metrics may thus be applied to a variety of datasets issuing from chaotic dynamical systems, including (non-stationary) climate datasets (e.g. Faranda et al., 2019c, 2020; Brunetti et al., 2019).

MatLab code to compute d, θ^{-1} and α is provided at the end of this paper under "code availability".

3 Dynamical systems in action: an example from the mid-Holocene Green Sahara

3.1 The mid-Holocene Green Sahara: background and data

Today, the Sahara is the largest hot desert on Earth. Most of the precipitation in North-Western Africa is associated with the West African Monsoon (WAM), which reaches to around 16-17 °N (e.g. Sultan and Janicot, 2003) and effectively sets the boundary between the semiarid Sahel and the Sahara. However, the region has repeatedly experienced momentous hydroclimatic shifts in the past. In particular, there have been several periods when the Sahara was wetter and greener than today, often termed *African Humid Periods* (AHPs, see Claussen et al. (2017) and Pausata et al. (2020) for recent reviews on the topic).

The most recent AHP peaked during the mid-Holocene (MH), approximately 9000 yr – 6000 yr BP. It is thought to have coincided with an intensification and northward shift of the WAM, allowing the presence of vegetation, lakes and wetlands in areas that today are desert (e.g. Holmes, 2008, and references therein). Palaeo-archives suggest that during the MH AHP, summer precipitation reached the northern parts of the present-day desert (e.g. Sha et al., 2019) and that tropical vegetation may have extended as far as 24 °N (Hély et al., 2014).

Numerical climate simulations of the MH have struggled to reproduce the full extent of the monsoonal intensification suggested by the palaeo-archives, and commonly suffer from a dry bias (Harrison et al., 2014). Early investigations on the topic highlighted the large sensitivity of the simulations to land-surface characteristics (e.g. Kutzbach et al., 1996; Kutzbach and Liu, 1997; Claussen and Gayler, 1997). More recent modelling efforts have confirmed this, and have further highlighted the potential role of an incorrect representation of atmospheric aerosols in favouring the dry bias (Pausata et al., 2016; Gaetani et al., 2017; Messori et al., 2019). Such hypothesis has triggered a lively discussion in the literature (cf. Thompson et al., 2019; Hopcroft and Valdes, 2019).

Here, we analyse the simulations used in Messori et al. (2019), performed with the EC-Earth Earth System Model v3.1 (Hazeleger et al., 2010). The atmospheric model has a T159 horizontal spectral resolution and 62 vertical levels. The ocean model has a nominal horizontal resolution of 1° and 46 vertical levels. In all simulations, the vegetation and aerosol concentrations are prescribed.

https://doi.org/10.5194/cp-2020-103
Preprint. Discussion started: 27 August 2020

To illustrate the dynamical systems approach described in Sect. 2, we consider three different simulations. The first is a MH control simulation (MH$_{CNTL}$), which follows the PMIP3 protocol in imposing pre-industrial vegetation and atmospheric dust concentrations (Braconnot et al., 2011). The second is a Green Sahara simulation (MH$_{GS+PD}$), which imposes shrubland over the region 11—33 °N and 15 °W – 35 °E. The third is a Green Sahara simulation that, in addition to the vegetation, also imposes a strongly reduced atmospheric dust loading (MH$_{GS+RD}$). Indeed, a greening of the Sahara would intuitively correspond to decreased dust emissions and hence to a lower atmospheric loading, as also supported by palaeo-archives (Demenocal et al., 2000; McGee et al., 2013) and modelling studies (Egerer et al., 2016).

We analyse 30 years of daily data of sea-level pressure (SLP), 500 hPa geopotential height (Z500) and precipitation frequency (prp) for each simulation. Precipitation frequency is constructed by assigning a value of 1 to grid points and time steps with non-zero precipitation and a value of 0 otherwise. This is preferable to using raw precipitation data for estimating the dynamical systems metrics (and d in particular), as discussed further in Langousis et al. (2009) and Faranda et al. (2017a). We define the pre-monsoon season as March, April and May (MAM) and the monsoon season as June, July, August and September (JJAS).

3.2 A dynamical systems view of the Mid-Holocene Green Sahara

The main interest in analysing the above simulations lies in understanding whether and why they reproduce different hydroclimates over the Sahelian-Saharan region. Our aim in this Section is not to provide a comprehensive answer to these two aspects, but rather to illustrate how the dynamical systems framework proposed here can be used to characterise the individual simulations and provide a concise overview of the differences between them. We argue that such an approach can provide a valuable complement to conventional analyses, and we relate our results to those obtained in earlier studies (e.g. Pausata et al., 2016; Gaetani et al., 2017; Messori et al., 2019).

A simple composite of JJAS average precipitation immediately highlights large differences in the precipitation regimes, with the MH$_{GS+PD}$ simulation showing a large northward shift and intensification of the monsoonal precipitation compared to MH$_{CNTL}$ and the MH$_{GS+RD}$ simulation showing an additional, albeit smaller, precipitation increase (Fig. 1). However, this time-mean picture hides a number of complex dynamical changes in the WAM, which we investigate using our dynamical systems framework. We focus on the Northern WAM region (12.5 – 30 °N, 10 °W – 20 °E, black box in Fig. 1a). This domain is chosen to reflect the region of seasonal monsoon rainfall which we expect to be most affected by the changes in land surface and atmospheric dust loading. Results for a more geographically extended domain are shown in Appendix A.

We begin by studying the seasonality of d and θ^{-1} for precipitation data. In the MH$_{CNTL}$ (Fig. 2a, blue curve), the local dimension displays two peaks, roughly matching the onset and withdrawal phases of the monsoon, somewhat lower values during the height of the summertime monsoon and the lowest values during the dry season. Previous studies have noted how transition seasons can display an increase in the local dimension of atmospheric fields, because the atmosphere explores configurations belonging to more than one season (Faranda et al., 2017b). In more technical terms, this would reflect a saddle-like point of the atmospheric dynamics. We therefore interpret the two local maxima as reflecting the northward shift and retreat of the monsoonal rainfall. The local dimension in the MH$_{GS+PD}$ and MH$_{GS+RD}$ simulations (red and orange curves, respectively) presents a similar seasonal cycle, yet with the first local maximum shifted to earlier in the year and the second

https://doi.org/10.5194/cp-2020-103
Preprint. Discussion started: 27 August 2020

local maximum shifted to later in the year (Fig. 2b). This points to a lengthening of the monsoon season, with an earlier rainfall onset and a later withdrawal. The timing of the first local maximum in d indeed coincides with a rapid increase in the zonally averaged precipitation at the southern edge of the domain in the MH$_{GS+PD}$ and MH$_{GS+RD}$ simulations (Fig. 2c). Such a lenghtening of the monsoonal period under a greening of the Sahara was previously noted in Pausata et al. (2016) by adopting a monsoon duration algorithm. The seasonal cycle of θ in MH$_{CNTL}$ (Fig. 2b, blue curve) displays a very different pattern. Low values (high persistence) occur during the monsoon season while higher values (lower persistence) occur during the dry season, albeit with a very large spread. This may reflect sporadic rainfall events at the edges of the domain outside of the monsoon season, with more persistent precipitation patterns during the monsoon season. The MH$_{GS+PD}$ and MH$_{GS+RD}$ simulations (red and orange curves, respectively) display a similar seasonality, albeit with a longer high-persistence monsoonal period, and a more marked difference in values between the monsoonal and dry phases. This chiefly results from lower values during the monsoonal period, likely reflecting a more geographically extensive and persistent precipitation regime. One may further hypothesise that this underlies a decrease in importance of transient, mesoscale convective systems for driving the monsoonal precipitation, in favour of a regional re-organisation of precipitation into larger-scale persistent features. This would also explain the decrease in d during the Monsoon season in the MH$_{GS+PD}$ and MH$_{GS+RD}$ simulations relative to the MH$_{CNTL}$ case. In agreement with this hypothesis, Gaetani et al. (2017) indeed found that the greening of the Sahara and dust reduction suppress African Easterly Waves and their role in triggering precipitation. The above qualitative considerations are mostly insensitive to the exact choice of geographical domain (cf. Figs. 2 and A1).

The seasonal variations in d and θ can also be related to variations in the dynamical indicators on shorter timescales. The fact that a rapid increase in d and a corresponding decrease in θ coincide with the northward progression of monsoonal rainfall indeed suggests that concurring high d values and low θ values on daily timescales may correspond to specific spatial precipitation patterns. To verify this, we compute composite rainfall anomalies during JJAS on days with concurrent d anomalies above the 70*th* percentile and θ anomalies below the 30*th* percentile of the respective JJAS distributions (Fig. 3). These relatively broad ranges are needed to ensure a good sample of dates, since here we are imposing a condition on each of the two metrics simultaneously. The anomalies are defined as deviations from a daily seasonal cycle. For example, the climatological value of a given variable in a given simulation for the 22nd July, is the mean of that variable across all 22nd July days in the simulation. Applying a smoothing to the climatology leads to very minor quantitative changes in our results (not shown). In MH$_{CNTL}$ (Fig. 3a), the anomalies are limited to the southern part of the domain, as the bulk of the Sahara receives little or no precipitation even at the peak of the monsoon (see Fig. 1a). The spatial pattern of the anomalies is wave-like, pointing to the fact that the dynamical systems metrics may reflect modulations in African Easterly Wave activity (see e.g. Fig. 8 in Gaetani et al. (2017) and the discussion above). The MH$_{GS+PD}$ and MH$_{GS+RD}$ simulations instead display clear anomaly dipoles, oriented in a predominantly meridional direction but with some zonal asymmetry. These correspond to a nortwhard shift of the monsoonal precipitation relative to the climatology (Fig. 3b and c, respectively). The dipoles span the whole domain, and display the largest anomaly values in the MH$_{GS+RD}$ simulation. This is indeed the simulation showing the largest total rainfall, as well as the strongest northward shift of the monsoonal precipitation range (cf. Figs. 1b, c). Very similar results are obtained

https://doi.org/10.5194/cp-2020-103
Preprint. Discussion started: 27 August 2020

if the same calculation is repeated over a larger domain (Fig. A2) or over the whole year (Fig. A3), albeit in the latter case with weaker precipitation anomalies, as may be expected from the inclusion of the dry season.

We next try to understand the physical processes underlying the differences in precipitation in the three simulations, by computing the co-recurrence ratio α between prp and SLP (Fig. 4a). In MH$_{CNTL}$ (blue line), as the monsoonal precipitation progresses northwards the coupling between the two variables increases, peaking in the middle of the monsoon season and waning thereafter. The dry season is characterised by overall low coupling values. In the MH$_{GS+PD}$ and MH$_{GS+RD}$ simulations (red and orange curves, respectively), α displays two local minima in the pre-monsoon season and in fall. During the northward progression of precipitation and the peak monsoonal phase, the values are mostly higher than for the MH$_{CNTL}$ simulation. Both simulations also show higher α values than MH$_{CNTL}$ during the dry season, although these values are generally lower than in the monsoonal period. Similar results are found when extending the geographical domain (cf. Figs. 4 and A5), albeit with slightly higher coupling values for the extended domain during the dry season. These are likely associated to the presence of more abundant wintertime precipitation at the latter domain's southern boundary (Fig. A4). The stronger coupling in the MH$_{GS+PD}$ and MH$_{GS+RD}$ simulations compared to the MH$_{CNTL}$ during the pre-monsoon and monsoon seasons, points to the role of circulation anomalies – reflected in the SLP field – in favouring the northwards extension of the monsoonal precipitation. This was indeed noted in Pausata et al. (2016) by analysing changes in lower-level atmospheric thickness related to the Saharan Heat Low (see also Lavaysse et al., 2009). The higher α values during wintertime in the MH$_{GS+PD}$ and MH$_{GS+RD}$ simulations, may once again be related to the presence of limited amounts of winter precipitation in the domain while precipitation is almost entirely absent in the MH$_{CNTL}$ simulation (Fig. A4). A similar picture is found for the co-recurrence ratio between prp and Z500 (Fig. 4b), highlighting the robust nature of the increased coupling between precipitation and large-scale atmospheric circulation features in the MH$_{GS+PD}$ and MH$_{GS+RD}$ simulations.

As for d and θ above, one may relate the seasonal variations in α to the daily anomalies associated with large or small values of the metric. We specifically consider precipitation, SLP and Z500 anomalies (computed as in Fig. 3) on JJAS days when α exceeds the 95*th* percentile of its anomaly distribution. These "strong coupling" days may be conceptualised as days on which recurrent spatial large-scale circulation anomalies favour recurrent spatial precipitation anomalies. In MH_{CNTL}, this takes the form of stronger precipitation across the southern portion of the domain, favoured by negative SLP and Z500 anomalies to the North of the strongest precipitation anomalies (Figs. 5a, 6a). These are likely the footprint of a strengthened heat low (see e.g. Figs. 2 and 8 in Lavaysse et al. (2009)), which favours a northward progression of the monsoonal precipitation. As noted above, the signal being limited to the southern part of the domain is due to the MH_{CNTL} simulation displaying little or no precipitation in the more northerly parts of the domain. The MH_{GS+PD} simulation shows a predominantly zonal dipole, with positive precipitation anomalies in the eastern part of the domain and negative anomalies further west (Figs. 5b, 6b). On strong coupling days, the large-scale circulation therefore favours an eastward extension of precipitation into a region that, even under a vegetated Sahara, receives little precipitation (see Fig. 1b). Similar spatial patterns are seen for both the SLP and Z500 composites, and in both cases the footprint of an enhanced heat low is visible. The MH_{GS+RD} simulation resembles the MH_{GS+PD} simulation for the Z500 case, albeit with weaker geopotential height anomalies (Fig. 6c). An inverted precipitation dipole, with a drier eastern part of the domain and a wetter North-Western part, is instead seen for the SLP composite (Fig.

5c). Comparable results are found when extending the geographical domain, with some differences that we partly ascribe to the effect of α capturing some tropical precipitation patterns at the southern edge of the domain (cf. Figs. 5 and 6 with Figs. A6 and A7). A hypothesis to explain the contrasting patterns found in the MH_{GS+RD} simulation is that thermal lows play a comparatively less important role there than in the MH_{GS+PD} case, in favour of enhanced deep convection triggered by large upward heat fluxes over the Sahara (Gaetani et al., 2017). This is in agreement with the increased amount of locally recycled moisture over the Sahara driven by dust reduction under a vegetated Sahara, noted by (Messori et al., 2019).

The above results illustrate some of the strengths and limitations of the analysis framework we propose in this note, which we discuss further in Section 4 below. If applied in the context of a full-length research paper, some of the hypotheses expounded here could be verified through additional analyses. These could include, for example, the use of lower-level atmospheric thickness or other tailored indicators of heat low activity, of atmospheric radiative and heat fluxes, and of moist static energy as an indicator of convection.

4 An appraisal of the dynamical systems framework in a palaeoclimate context

Palaeoclimate simulations of the same period and region may yield very different results, the understanding of which requires analysis tools that may efficiently distil the discrepancies and point to possible underlying drivers. In this technical note, we have outlined an analysis framework which can efficiently compare the salient dynamical features of different simulated palaeoclimates. The framework is grounded in dynamical systems theory, and rests on computing three metrics: the local dimension d, the persistence θ^{-1} and the co-recurrence ratio α. The first two metrics inform on the evolution of a system about a given state of interest – for example how the atmosphere evolves to or from a given large-scale configuration. The third metric describes the coupling between different variables.

From a theoretical standpoint, the dynamical systems framework presents a number of advantages over other statistical approaches for the analysis of large amounts of data. First, the three metrics we use are rooted in the underlying dynamics of the system being analysed. In other words, their values are projections of mathematical properties of the underlying equations of the system, even when these are unknown. Second, the choice of the free parameters needed to estimate the metrics is basically limited to selecting a threshold to define recurrences. Finally, the metrics provide one value for every timestep in the analysed data, and may be conveniently used to investigate seasonality, oscillatory behaviours, high-frequency variability and more. This is especially valuable for the co-recurrence ratio, as a number of other measures of coupling or correlation between two variables only provide a single value for the whole time period being considered.

Because of these characteristics, the dynamical systems metrics can be particularly helpful when processing large datasets (see e.g. Rodrigues et al. (2018); Faranda et al. (2019a)). To illustrate their practical applicability in palaeoclimate studies, we have analysed three numerical simulations of the mid-Holocene climate over North Africa: a control simulation with pre-industrial vegetation and atmospheric dust loading, a Green Sahara simulation with shrubland imposed over a broad swath of what is today the Sahara desert, and a second Green Sahara simulation which additionally features heavily reduced atmospheric dust loading. Our aim is to show that the different hydroclimates in these simulations correspond to different dynamical

https://doi.org/10.5194/cp-2020-103
Preprint. Discussion started: 27 August 2020

properties of the modelled climate systems, which are reflected in the three dynamical systems metrics. Both d and θ^{-1} underscore changes in the duration and geographical extent of the monsoon, and further hint to the physical drivers of these changes, such as modulations in atmospheric wave activity. The co-recurrence ratio α enriches the picture by enabling to contextualise precipitation changes relative to large-scale atmospheric circulation anomalies.

As a caveat, obtaining good estimates of d, θ^{-1} and α requires relatively long time series, limiting their applicability to palaeo-archives. There are no fixed rules in this sense, but current best practice is to have several good recurrences of the patterns of interest in the data (e.g. Faranda et al., 2011). Non-stationary data, such as may be found in transient palaeoclimate simulations, also requires some care in verifying that recurrences can be identified (see also Sect. 2). A further difficulty that may be encountered in applying the dynamical systems framework pertains its interpretation. While the three metrics lend themselves to making relatively intuitive heuristic inferences, they may sometimes provide counterintuitive results, such as Figs. 5c and 6c here, and there is no universally valid approach to overcome these interpretative difficulties. Furthermore, expounding formal arguments to support the results obtained requires a detailed knowledge of the underlying theoretical bases, which may initially be daunting.

In this technical note, we aimed to give a flavour of the dynamical systems framework's possible application to palaeoclimate simulations, as opposed to presenting a systematic analysis. We specifically wished to highlight its potential for comparing different palaeoclimates, while also providing an appraisal of its limitations. To do so, we focussed on three existing simulations and on a small number of atmospheric variables. However, the approach may be applied to a very broad range of climate variables, not limited to the atmosphere. We envisage that the most effective application of the framework would be for the analysis of very large datasets, such as those issuing from the PMIP initiative or from downscaling efforts on very long transient simulations (e.g. Lorenz et al., 2016). At the same time, we stress that we do not view the framework as a wholesale substitute for conventional analyses of palaeoclimate dynamics. Rather, it is intended as a complement that may help to strengthen mechanistic interpretations and rapidly identify features deserving further investigation.

Code availability. The code to compute the three dynamical systems indicators used in this study is made freely available through the cloud storage of the *Centre National de la Recherche Scientifique* (CNRS), under a CC BY-NC 3.0 license:
https://mycore.core-cloud.net/index.php/s/pLJw5XSYhe2ZmnZ

Author contributions. G. Messori conceived the study and performed the analysis. D. Faranda provided the publicly available code. Both authors contributed to drafting the manuscript.

Competing interests. The authors have no competing interests to declare.

https://doi.org/10.5194/cp-2020-103
Preprint. Discussion started: 27 August 2020

Climate of the Past
Discussions

Acknowledgements. The authors thank F. Pausata, Q. Zhang and M. Gaetani for making the palaeoclimate simulations available. G. Messori was partly supported by the Swedish Research Council Vetenskapsrådet (grant no. 2016-03724) and the Swedish Research Council FORMAS (grant no. 2018-00968). DF is supported by a CNRS-INSU LEFE/MANU grant (DINCLIC project).

Appendix A: Additional Figures

In this appendix, we provide figures illustrating the sensitivity of our results to the choice of geographical domain and season. The figures are discussed in the main text.

https://doi.org/10.5194/cp-2020-103
Preprint. Discussion started: 27 August 2020

References

Barron, E. J., Sloan II, J., and Harrison, C.: Potential significance of land—sea distribution and surface albedo variations as a climatic forcing factor; 180 my to the present, Palaeogeography, Palaeoclimatology, Palaeoecology, 30, 17–40, 1980.

Bartlein, P. J., Harrison, S., Brewer, S., Connor, S., Davis, B., Gajewski, K., Guiot, J., Harrison-Prentice, T., Henderson, A., Peyron, O., et al.: Pollen-based continental climate reconstructions at 6 and 21 ka: a global synthesis, Climate Dynamics, 37, 775–802, 2011.

Braconnot, P., Harrison, S., Otto-Bliesner, B., Abe-Ouchi, A., Jungclaus, J., and Peterschmitt, J.: The paleoclimate modeling Intercomparison project contribution to CMIP5, CLIVAR Exchanges, 56, 15—-19, 2011.

Brunetti, M., Kasparian, J., and Vérard, C.: Co-existing climate attractors in a coupled aquaplanet, Climate Dynamics, 53, 6293–6308, 2019.

Buschow, S. and Friederichs, P.: Local dimension and recurrent circulation patterns in long-term climate simulations, Chaos: An Interdisciplinary Journal of Nonlinear Science, 28, 083 124, 2018.

Caby, T., Faranda, D., Vaienti, S., and Yiou, P.: Extreme value distributions of observation recurrences, arXiv preprint arXiv:2002.10873, 2020.

Claussen, M. and Gayler, V.: The greening of the Sahara during the mid-Holocene: results of an interactive atmosphere-biome model, Global Ecology and Biogeography Letters, pp. 369–377, 1997.

Claussen, M., Dallmeyer, A., and Bader, J.: Theory and modeling of the African humid period and the green Sahara, in: Oxford Research Encyclopedia of Climate Science, Oxford University Press, 2017.

De Luca, P., Messori, G., Pons, F. M., and Faranda, D.: Dynamical systems theory sheds new light on compound climate extremes in Europe and Eastern North America, Quarterly Journal of the Royal Meteorological Society, 2020.

Demenocal, P., Ortiz, J., Guilderson, T., Adkins, J., Sarnthein, M., Baker, L., and Yarusinsky, M.: Abrupt onset and termination of the African Humid Period:: rapid climate responses to gradual insolation forcing, Quaternary science reviews, 19, 347–361, 2000.

Donn, W. L. and Shaw, D. M.: Model of climate evolution based on continental drift and polar wandering, Geological Society of America Bulletin, 88, 390–396, 1977.

Egerer, S., Claussen, M., Reick, C. H., and Stanelle, T.: The link between marine sediment records and changes in Holocene Saharan landscape: Simulating the dust cycle, Climate of the Past, 12, 1009–1027, 2016.

Faranda, D., Lucarini, V., Turchetti, G., and Vaienti, S.: Numerical convergence of the block-maxima approach to the generalized extreme value distribution, Journal of statistical physics, 145, 1156–1180, 2011.

Faranda, D., Messori, G., Alvarez-Castro, M. C., and Yiou, P.: Dynamical properties and extremes of Northern Hemisphere climate fields over the past 60 years, Nonlinear Processes in Geophysics, 24, 713, 2017a.

Faranda, D., Messori, G., and Yiou, P.: Dynamical proxies of North Atlantic predictability and extremes, Scientific reports, 7, 41 278, 2017b.

Faranda, D., Alvarez-Castro, M. C., Messori, G., Rodrigues, D., and Yiou, P.: The hammam effect or how a warm ocean enhances large scale atmospheric predictability, Nature communications, 10, 1–7, 2019a.

Faranda, D., Messori, G., and Vannitsem, S.: Attractor dimension of time-averaged climate observables: insights from a low-order ocean-atmosphere model, Tellus A: Dynamic Meteorology and Oceanography, 71, 1–11, 2019b.

Faranda, D., Sato, Y., Messori, G., Moloney, N. R., and Yiou, P.: Minimal dynamical systems model of the northern hemisphere jet stream via embedding of climate data, Earth System Dynamics, 10, 555–567, 2019c.

Faranda, D., Messori, G., and Yiou, P.: Diagnosing concurrent drivers of weather extremes: application to warm and cold days in North America, Climate Dynamics, 54, 2187–2201, 2020.

https://doi.org/10.5194/cp-2020-103
Preprint. Discussion started: 27 August 2020

Freitas, A. C. M., Freitas, J. M., and Todd, M.: Hitting time statistics and extreme value theory, Probability Theory and Related Fields, 147, 675–710, 2010.

Gaetani, M., Messori, G., Zhang, Q., Flamant, C., and Pausata, F. S.: Understanding the mechanisms behind the northward extension of the West African Monsoon during the Mid-Holocene, Journal of Climate, 30, 7621–7642, 2017.

Gates, W. L.: The numerical simulation of ice-age climate with a global general circulation model, Journal of the Atmospheric Sciences, 33, 1844–1873, 1976.

Harrison, S., Bartlein, P., Brewer, S., Prentice, I., Boyd, M., Hessler, I., Holmgren, K., Izumi, K., and Willis, K.: Climate model benchmarking with glacial and mid-Holocene climates, Climate Dynamics, 43, 671–688, 2014.

Haywood, A., Hill, D., Dolan, A., Otto-Bliesner, B., Bragg, F., Chan, W.-L., Chandler, M., Contoux, C., Dowsett, H., Jost, A., et al.: Large-scale features of Pliocene climate: results from the Pliocene Model Intercomparison Project, Climate of the Past, 9, 191–209, 2013.

Hazeleger, W., Severijns, C., Semmler, T., Ştefănescu, S., Yang, S., Wang, X., Wyser, K., Dutra, E., Baldasano, J. M., Bintanja, R., et al.: EC-Earth: a seamless earth-system prediction approach in action, Bulletin of the American Meteorological Society, 91, 1357–1364, 2010.

Hély, C., Lézine, A.-M., et al.: Holocene changes in African vegetation: tradeoff between climate and water availability, Climate of the Past, 10, 681, 2014.

Hochman, A., Alpert, P., Harpaz, T., Saaroni, H., and Messori, G.: A new dynamical systems perspective on atmospheric predictability: Eastern Mediterranean weather regimes as a case study, Science advances, 5, eaau0936, 2019.

Hochman, A., Alpert, P., Kunin, P., Rostkier-Edelstein, D., Harpaz, T., Saaroni, H., and Messori, G.: The dynamics of cyclones in the twentyfirst century: the Eastern Mediterranean as an example, Climate Dynamics, 54, 561–574, 2020.

Holmes, J. A.: How the Sahara became dry, Science, 320, 752–753, 2008.

Hopcroft, P. O. and Valdes, P. J.: On the Role of Dust-Climate Feedbacks During the Mid-Holocene, Geophysical Research Letters, 46, 1612–1621, 2019.

Kageyama, M., Braconnot, P., Harrison, S. P., Haywood, A. M., Jungclaus, J., Otto-Bliesner, B. L., Peterschmitt, J.-Y., Abe-Ouchi, A., Albani, S., Bartlein, P. J., et al.: PMIP4-CMIP6: the contribution of the Paleoclimate Modelling Intercomparison Project to CMIP6, Geoscientific Model Development Discussions, 11, 1033–1057, 2018.

Kutzbach, J., Bonan, G., Foley, J., and Harrison, S.: Vegetation and soil feedbacks on the response of the African monsoon to orbital forcing in the early to middle Holocene, Nature, 384, 623–626, 1996.

Kutzbach, J. E. and Liu, Z.: Response of the African monsoon to orbital forcing and ocean feedbacks in the middle Holocene, Science, 278, 440–443, 1997.

Langousis, A., Veneziano, D., Furcolo, P., and Lepore, C.: Multifractal rainfall extremes: Theoretical analysis and practical estimation, Chaos, Solitons & Fractals, 39, 1182–1194, 2009.

Lavaysse, C., Flamant, C., Janicot, S., Parker, D., Lafore, J.-P., Sultan, B., and Pelon, J.: Seasonal evolution of the West African heat low: a climatological perspective, Climate Dynamics, 33, 313–330, 2009.

Lorenz, D. J., Nieto-Lugilde, D., Blois, J. L., Fitzpatrick, M. C., and Williams, J. W.: Downscaled and debiased climate simulations for North America from 21,000 years ago to 2100AD, Scientific data, 3, 1–19, 2016.

Lucarini, V., Faranda, D., and Wouters, J.: Universal behaviour of extreme value statistics for selected observables of dynamical systems, Journal of statistical physics, 147, 63–73, 2012.

Lucarini, V., Faranda, D., de Freitas, J. M. M., Holland, M., Kuna, T., Nicol, M., Todd, M., Vaienti, S., et al.: Extremes and recurrence in dynamical systems, John Wiley & Sons, 2016.

https://doi.org/10.5194/cp-2020-103
Preprint. Discussion started: 27 August 2020

McGee, D., deMenocal, P. B., Winckler, G., Stuut, J.-B. W., and Bradtmiller, L.: The magnitude, timing and abruptness of changes in North African dust deposition over the last 20,000 yr, Earth and Planetary Science Letters, 371, 163–176, 2013.

Messori, G., Caballero, R., and Faranda, D.: A dynamical systems approach to studying midlatitude weather extremes, Geophysical Research Letters, 44, 3346–3354, 2017.

Messori, G., Gaetani, M., Zhang, Q., Zhang, Q., and Pausata, F. S.: The water cycle of the mid-Holocene West African monsoon: The role of vegetation and dust emission changes, International Journal of Climatology, 39, 1927–1939, 2019.

Moloney, N. R., Faranda, D., and Sato, Y.: An overview of the extremal index, Chaos: An Interdisciplinary Journal of Nonlinear Science, 29, 022 101, 2019.

Pausata, F. S., Messori, G., and Zhang, Q.: Impacts of dust reduction on the northward expansion of the African monsoon during the Green Sahara period, Earth and Planetary Science Letters, 434, 298–307, 2016.

Pausata, F. S., Gaetani, M., Messori, G., Berg, A., de Souza, D. M., Sage, R. F., and deMenocal, P. B.: The Greening of the Sahara: Past Changes and Future Implications, One Earth, 2, 235–250, 2020.

Pons, F. M. E., Messori, G., Alvarez-Castro, M. C., and Faranda, D.: Sampling hyperspheres via extreme value theory: implications for measuring attractor dimensions, Journal of Statistical Physics, pp. 1–20, 2020.

Rodrigues, D., Alvarez-Castro, M. C., Messori, G., Yiou, P., Robin, Y., and Faranda, D.: Dynamical properties of the North Atlantic atmospheric circulation in the past 150 years in CMIP5 models and the 20CRv2c reanalysis, Journal of Climate, 31, 6097–6111, 2018.

Scher, S. and Messori, G.: Predicting weather forecast uncertainty with machine learning, Quarterly Journal of the Royal Meteorological Society, 144, 2830–2841, 2018.

Schnase, J. L., Lee, T. J., Mattmann, C. A., Lynnes, C. S., Cinquini, L., Ramirez, P. M., Hart, A. F., Williams, D. N., Waliser, D., Rinsland, P., et al.: Big data challenges in climate science: Improving the next-generation cyberinfrastructure, IEEE geoscience and remote sensing magazine, 4, 10–22, 2016.

Sha, L., Ait Brahim, Y., Wassenburg, J. A., Yin, J., Peros, M., Cruz, F. W., Cai, Y., Li, H., Du, W., Zhang, H., et al.: How far north did the African Monsoon fringe expand during the African Humid Period?–Insights from Southwest Moroccan speleothems, Geophysical Research Letters, 2019.

Sultan, B. and Janicot, S.: The West African monsoon dynamics. Part II: The "preonset" and "onset" of the summer monsoon, Journal of climate, 16, 3407–3427, 2003.

Süveges, M.: Likelihood estimation of the extremal index, Extremes, 10, 41–55, https://doi.org/10.1007/s10687-007-0034-2, https://doi.org/10.1007/s10687-007-0034-2, 2007.

Thompson, A. J., Skinner, C. B., Poulsen, C. J., and Zhu, J.: Modulation of mid-Holocene African rainfall by dust aerosol direct and indirect effects, Geophysical Research Letters, 46, 3917–3926, 2019.

https://doi.org/10.5194/cp-2020-103
Preprint. Discussion started: 27 August 2020

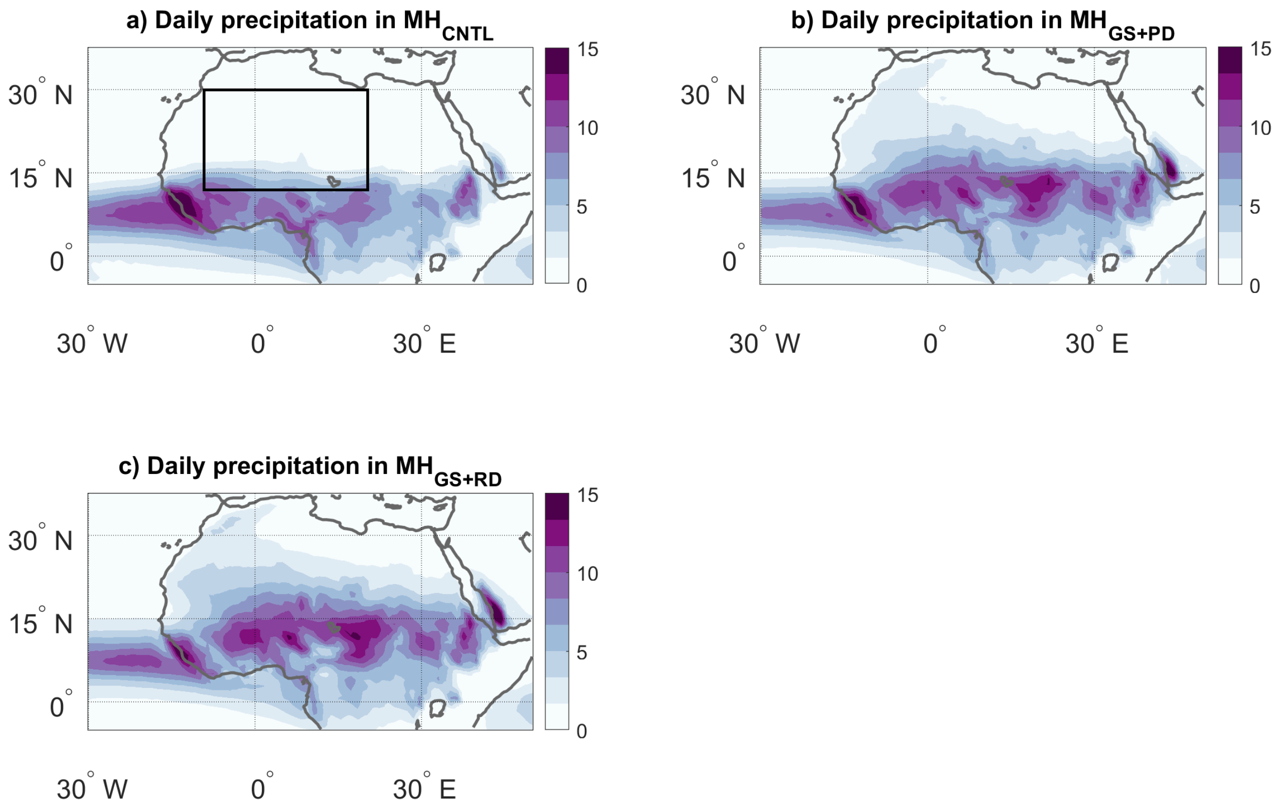

Figure 1. JJAS precipitation (mm day^{-1}) for the: (a) MH$_{CNTL}$, (b) MH$_{GS+PD}$ and (c) MH$_{GS+RD}$ simulations. The black box in (a) marks the domain used to perform the dynamical systems analysis (12.5 – 30 °N, 10 °W – 20 °E).

https://doi.org/10.5194/cp-2020-103
Preprint. Discussion started: 27 August 2020

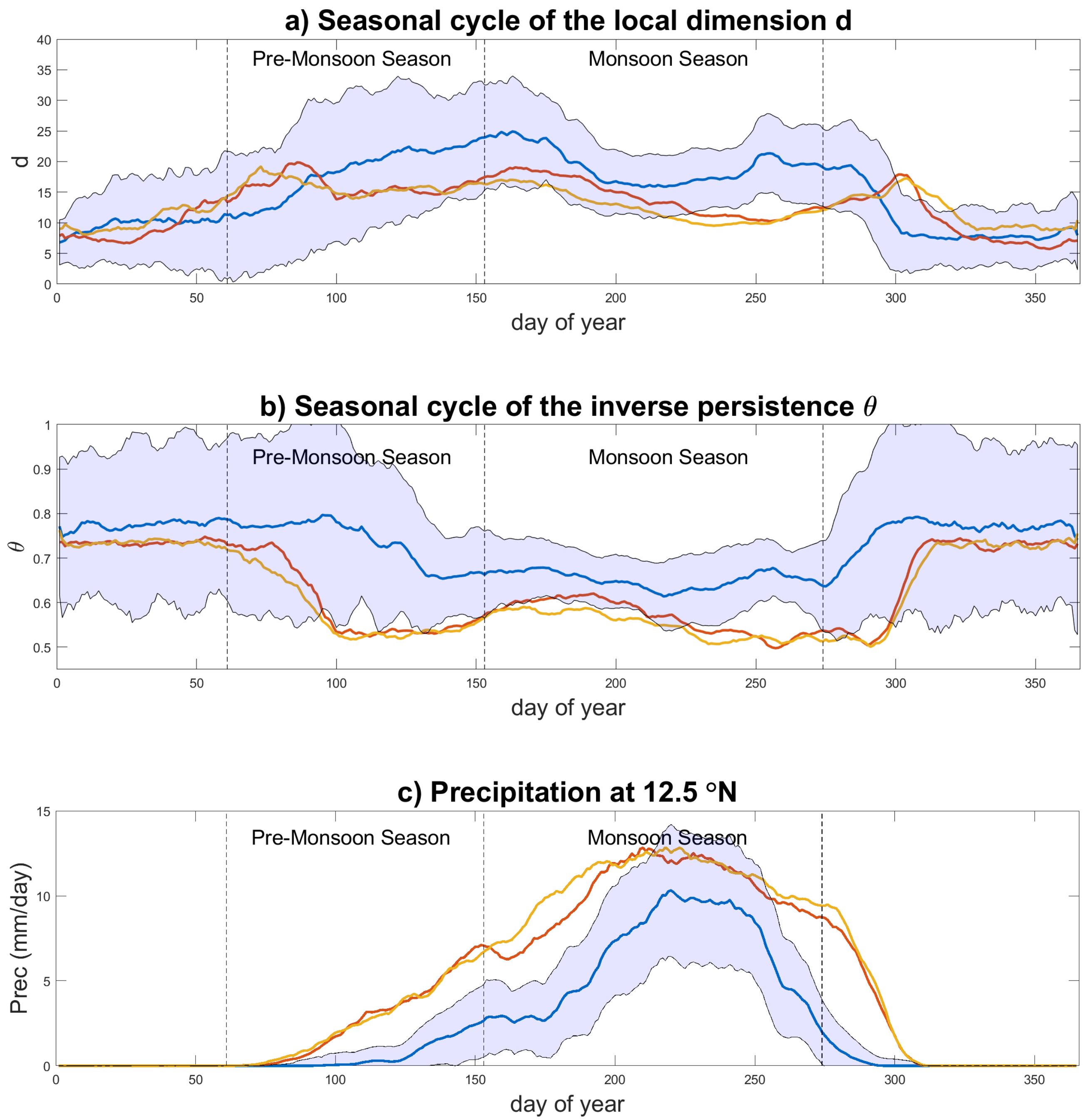

Figure 2. Seasonal cycle of median (a) d, (b) θ and (c) zonally averaged daily precipitation at 12.5 °N for the MH_{CNTL} (blue), MH_{GS+PD} (red) and MH_{GS+RD} (orange) simulations. The blue shading marks ± 1 std from the MH_{CNTL}. The vertical dashed lines mark the pre-monsoon (MAM) and monsoon (JJAS) seasons. The data is smoothed with a 10-day moving average.

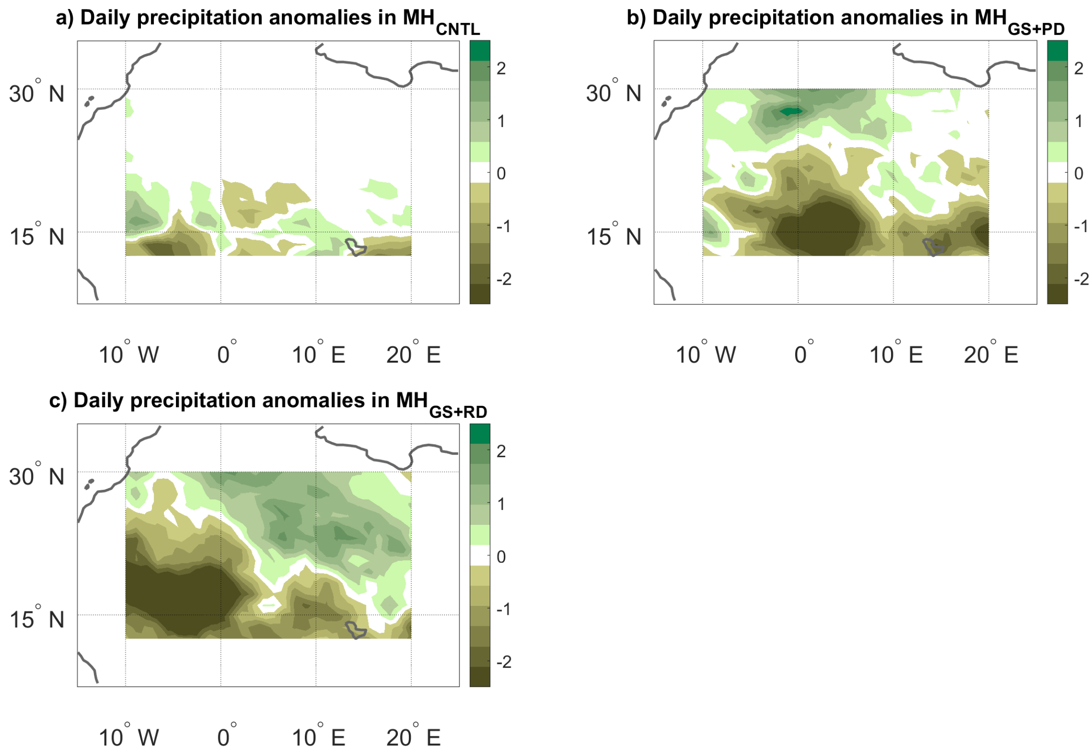

Figure 3. JJAS precipitation anomalies (mm day^{-1}) on days with high d and low θ (see text) for the: (a) MH$_{CNTL}$, (b) MH$_{GS+PD}$ and (c) MH$_{GS+RD}$ simulations. The anomalies are only shown over the domain used to perform the dynamical systems analysis (see black box in Fig. 1a).

https://doi.org/10.5194/cp-2020-103
Preprint. Discussion started: 27 August 2020

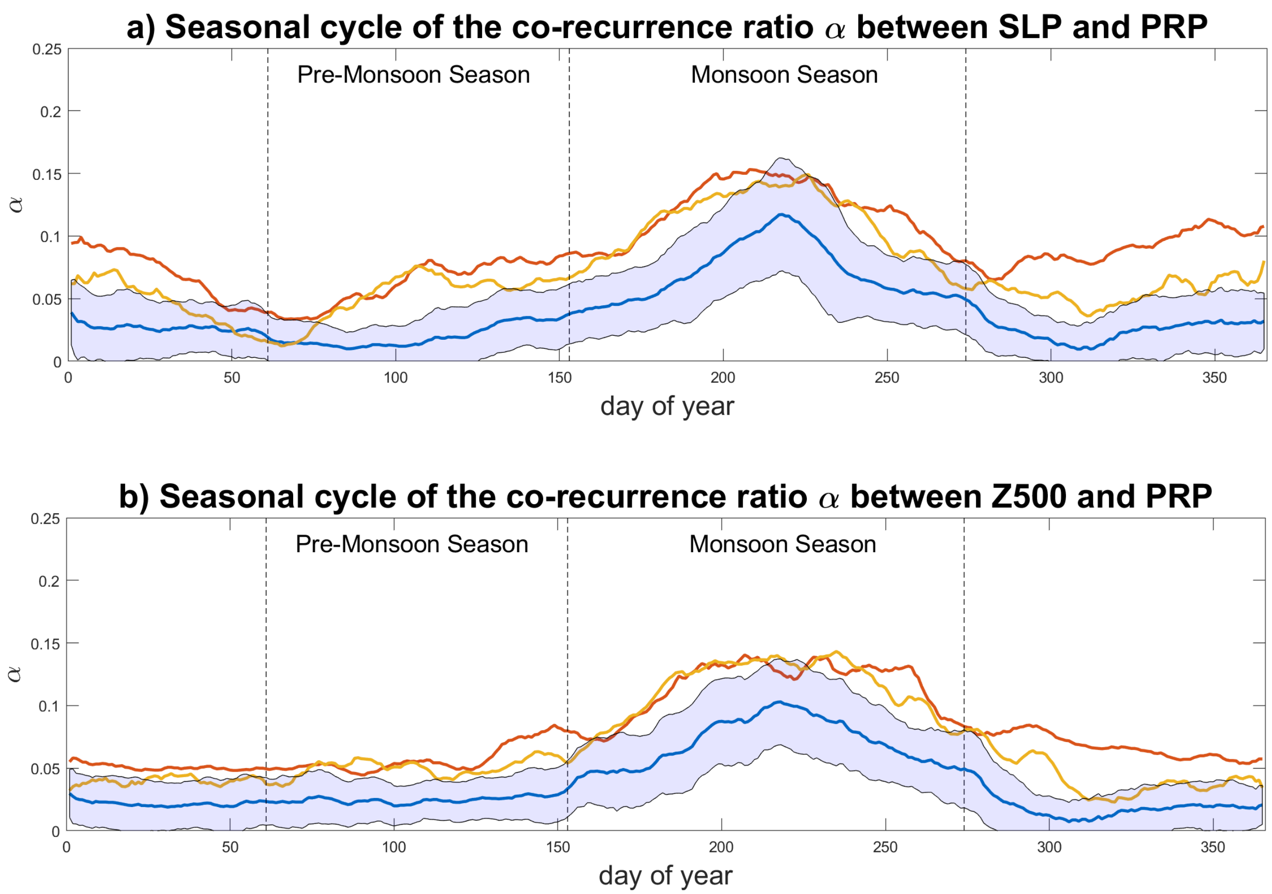

Figure 4. Seasonal cycle of median (a) $\alpha_{SLP,PRP}$ and (b) $\alpha_{Z500,PRP}$ for the MH$_{CNTL}$ (blue), MH$_{GS+PD}$ (red) and MH$_{GS+RD}$ (orange) simulations. The blue shading marks $\pm$ 1 std from the MH$_{CNTL}$. The vertical dashed lines mark the pre-monsoon (MAM) and monsoon (JJAS) seasons. The data is smoothed with a 10-day moving average.

https://doi.org/10.5194/cp-2020-103
Preprint. Discussion started: 27 August 2020

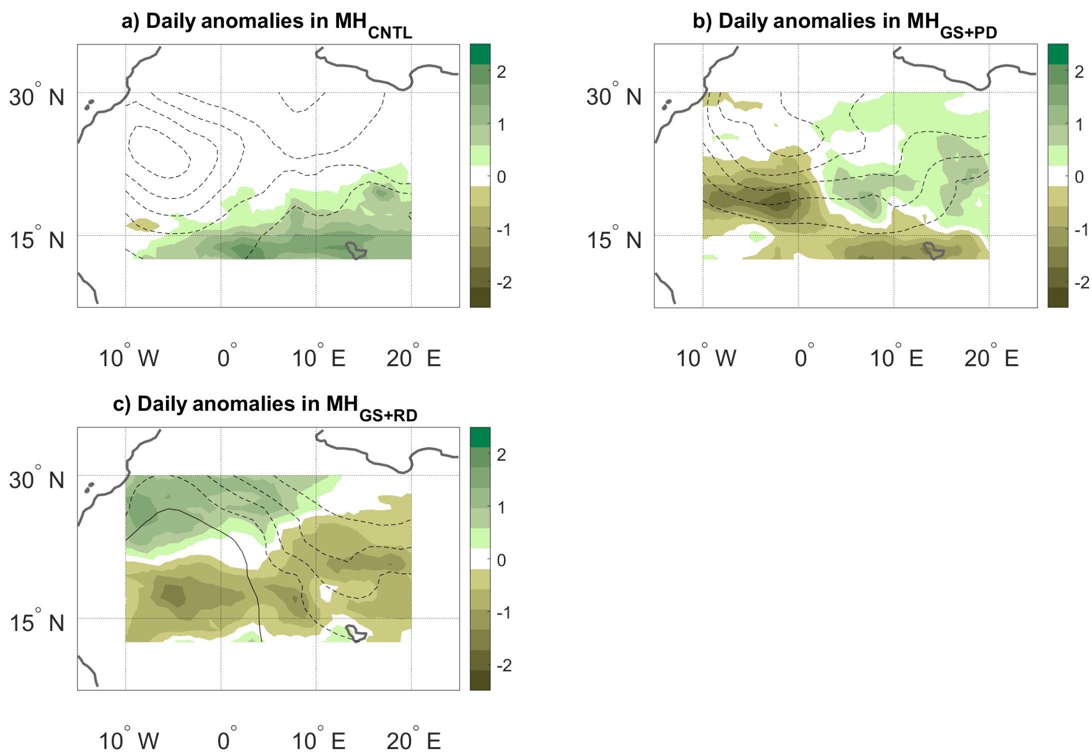

Figure 5. JJAS precipitation (colours, mm day^{-1}) and SLP (contours, hPa) anomalies on days with high α (see text) for the: (a) MH$_{CNTL}$, (b) MH$_{GS+PD}$ and (c) MH$_{GS+RD}$ simulations. The contour lines have an interval of 0.25 hPa. Continuous contours show positive anomalies, dashed contours show zero and negative anomalies. The anomalies are only shown over the domain used to perform the dynamical systems analysis (see black box in Fig. 1a).

https://doi.org/10.5194/cp-2020-103
Preprint. Discussion started: 27 August 2020

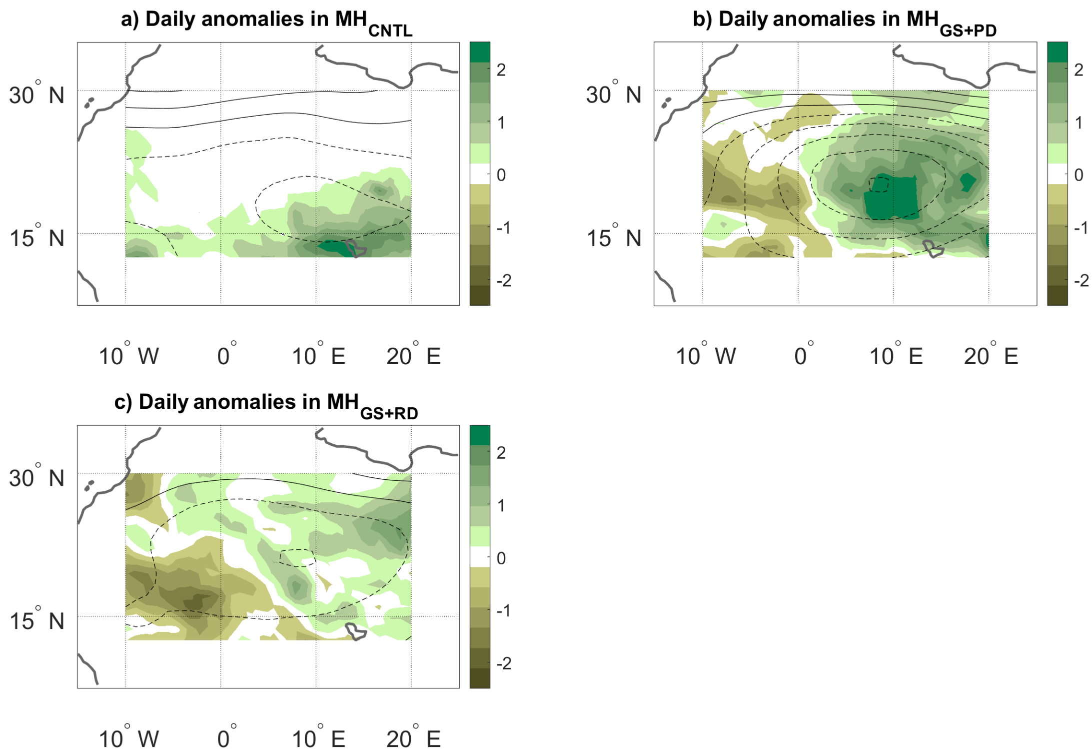

Figure 6. JJAS precipitation (colours, mm day^{-1}) and Z500 (contours, m) anomalies on days with high α (see text) for the: (a) MH$_{CNTL}$, (b) MH$_{GS+PD}$ and (c) MH$_{GS+RD}$ simulations. The contour lines have an interval of 25 m. Continuous contours show positive anomalies, dashed contours show zero and negative anomalies. The anomalies are only shown over the domain used to perform the dynamical systems analysis (see black box in Fig. 1a).

https://doi.org/10.5194/cp-2020-103
Preprint. Discussion started: 27 August 2020

Climate of the Past
Discussions

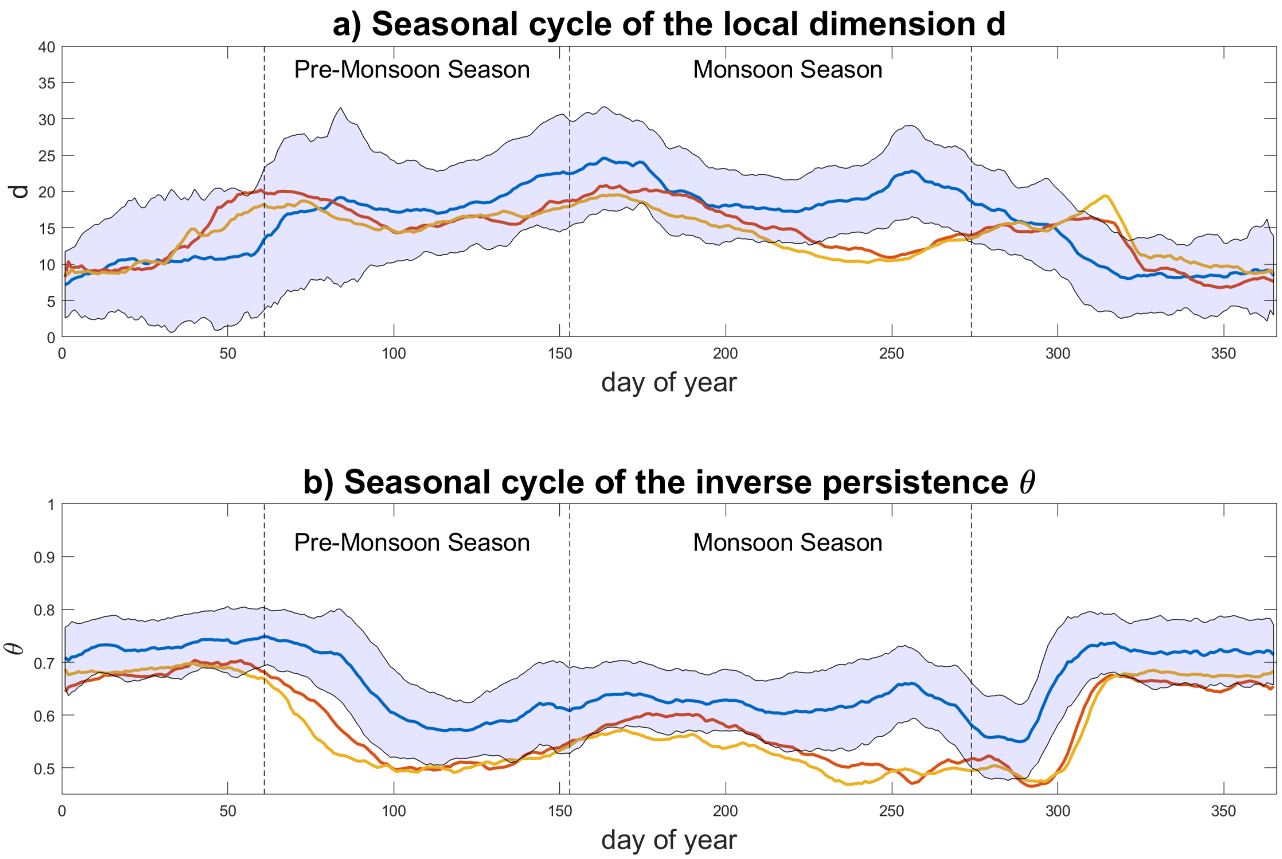

Figure A1. Same as Fig. 2a, b, but using the domain (10 – 30 °N, 15 °W – 20 °E).

https://doi.org/10.5194/cp-2020-103
Preprint. Discussion started: 27 August 2020

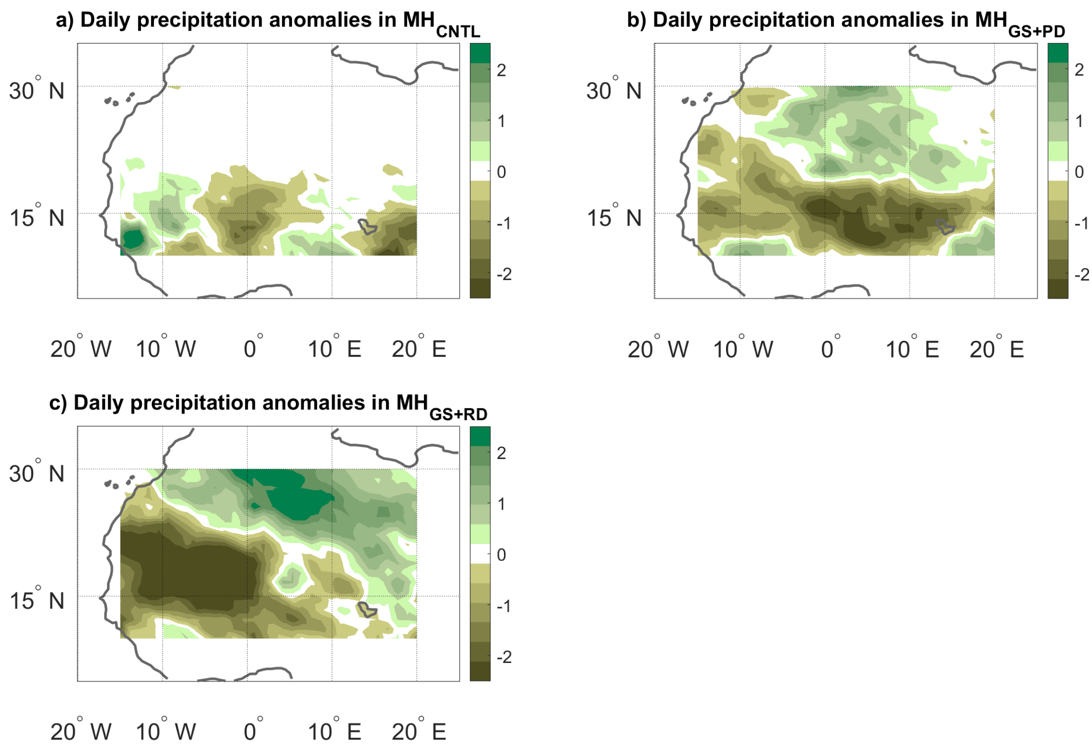

Figure A2. Same as Fig. 3, but using the domain (10 – 30 °N, 15 °W – 20 °E).

https://doi.org/10.5194/cp-2020-103
Preprint. Discussion started: 27 August 2020

Climate of the Past
Discussions
Open Access

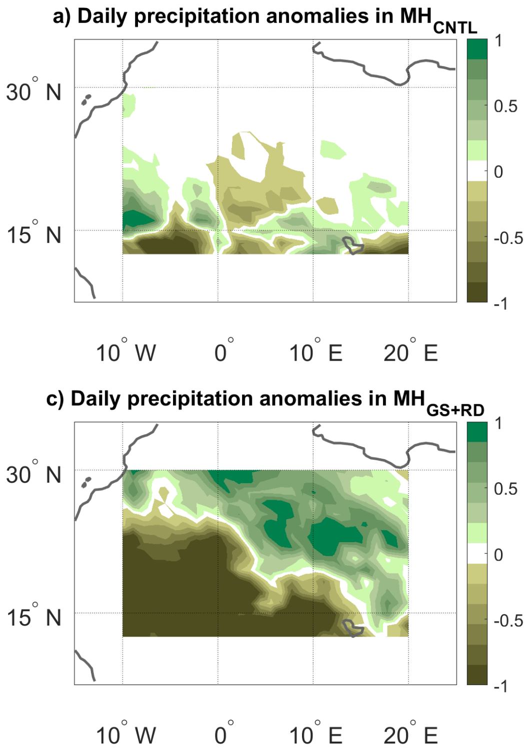

Figure A3. Same as Fig. 3, but for the whole year.

https://doi.org/10.5194/cp-2020-103
Preprint. Discussion started: 27 August 2020

Climate of the Past
Discussions
Open Access
EGU

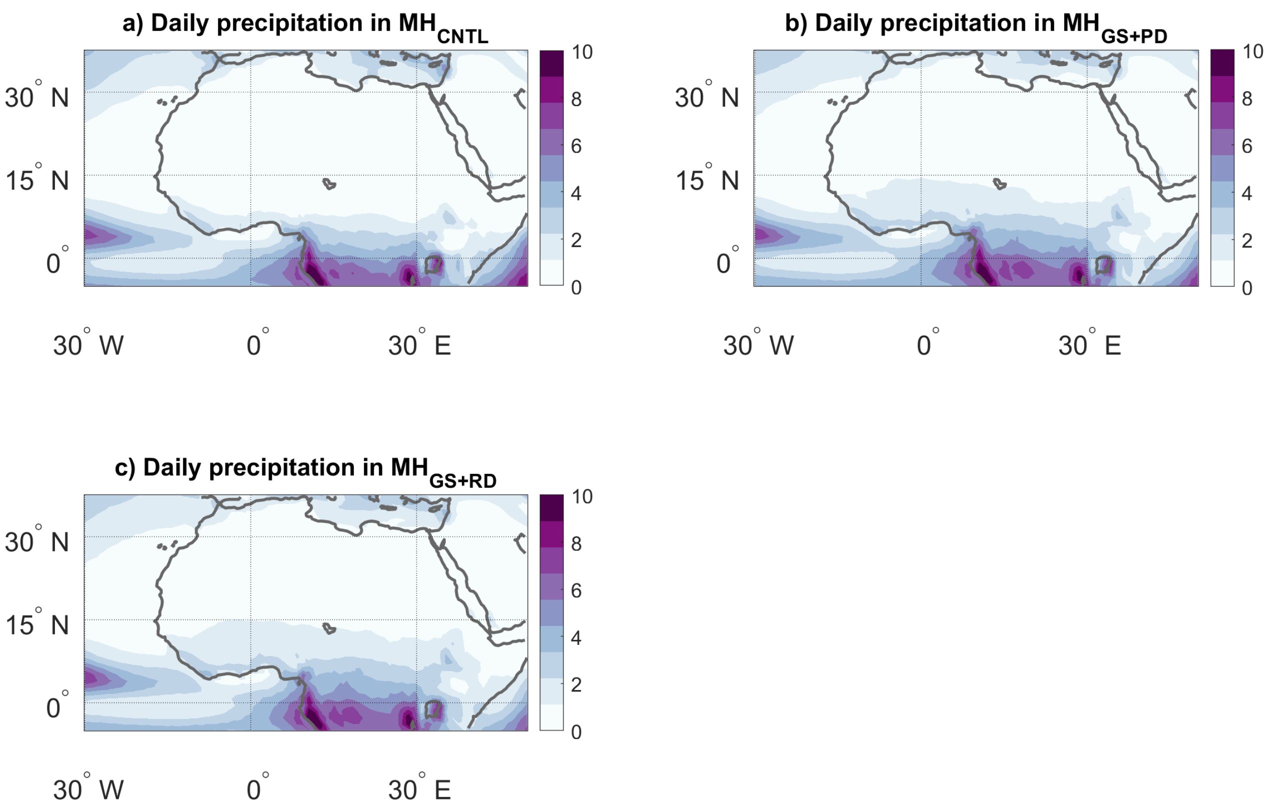

Figure A4. Same as Fig. 1, but for the October-February period.

https://doi.org/10.5194/cp-2020-103
Preprint. Discussion started: 27 August 2020

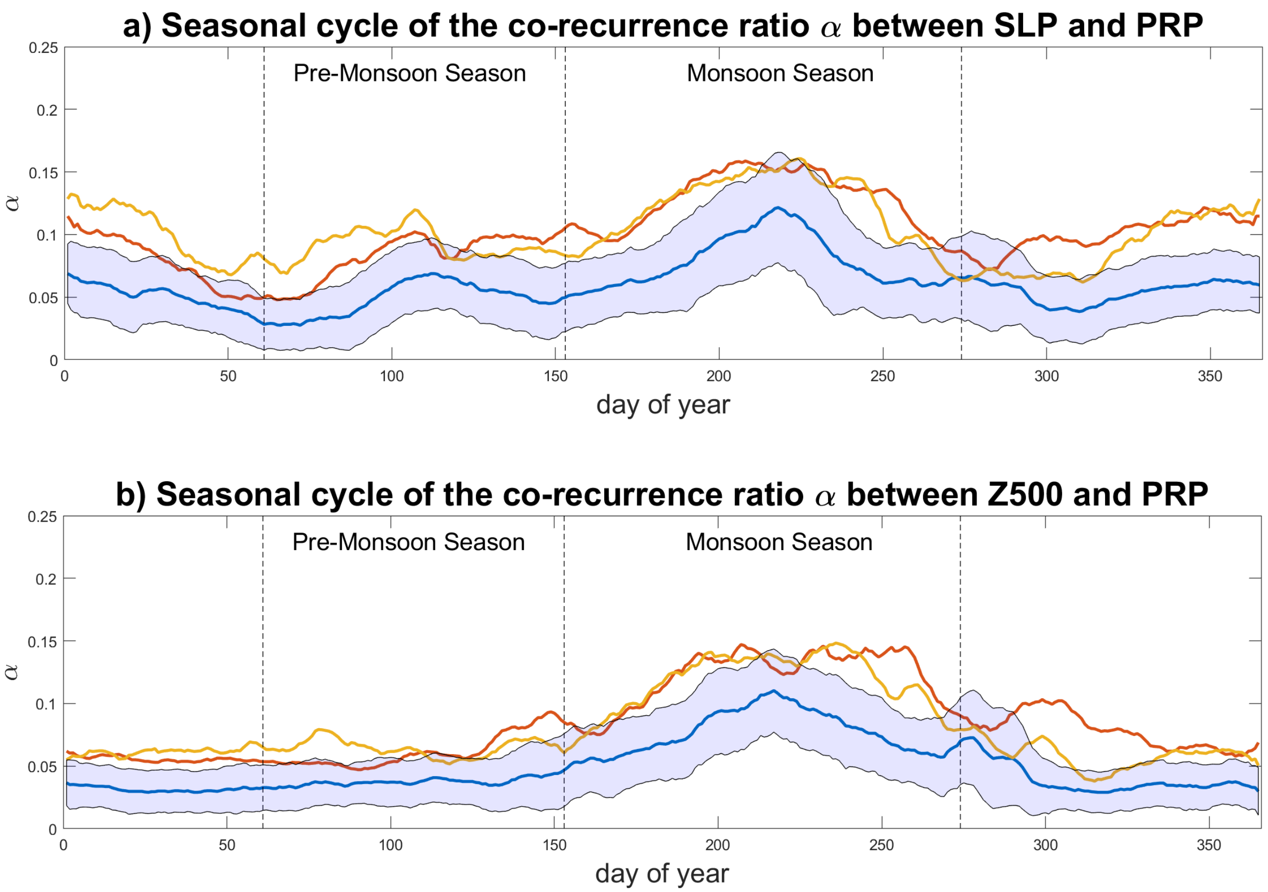

Figure A5. Same as Fig. 4, but using the domain (10 – 30 °N, 15 °W – 20 °E).

https://doi.org/10.5194/cp-2020-103
Preprint. Discussion started: 27 August 2020

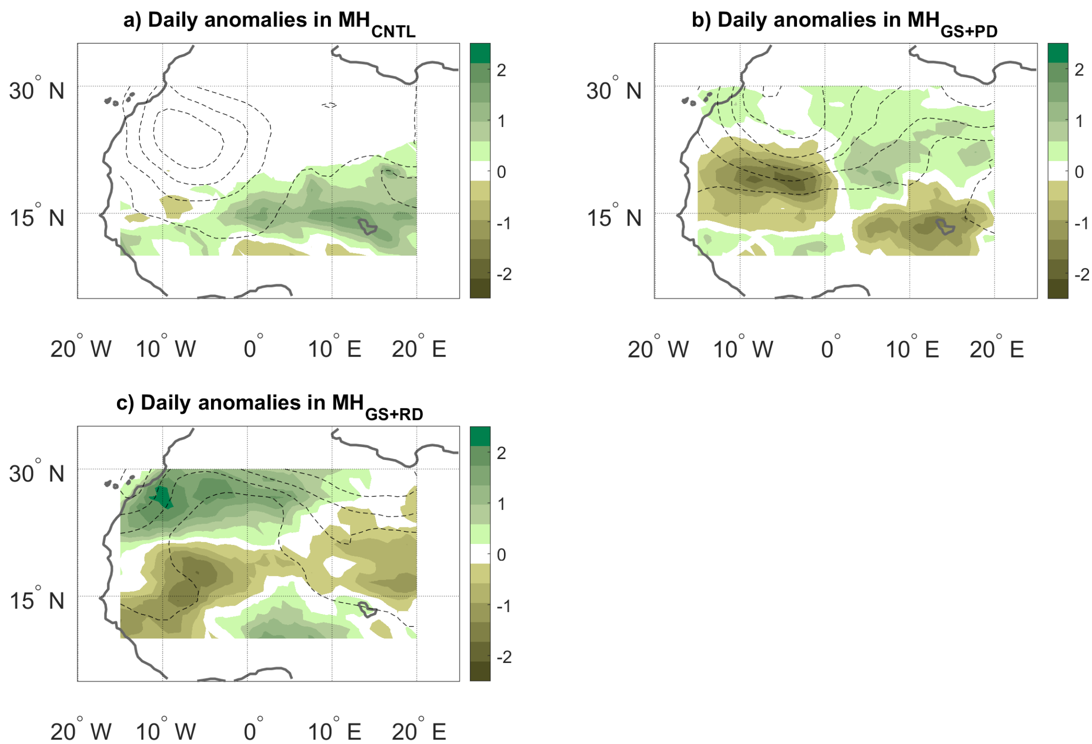

Figure A6. Same as Fig. 5, but using the domain (10 – 30 °N, 15 °W – 20 °E).

https://doi.org/10.5194/cp-2020-103
Preprint. Discussion started: 27 August 2020

Climate of the Past
Discussions

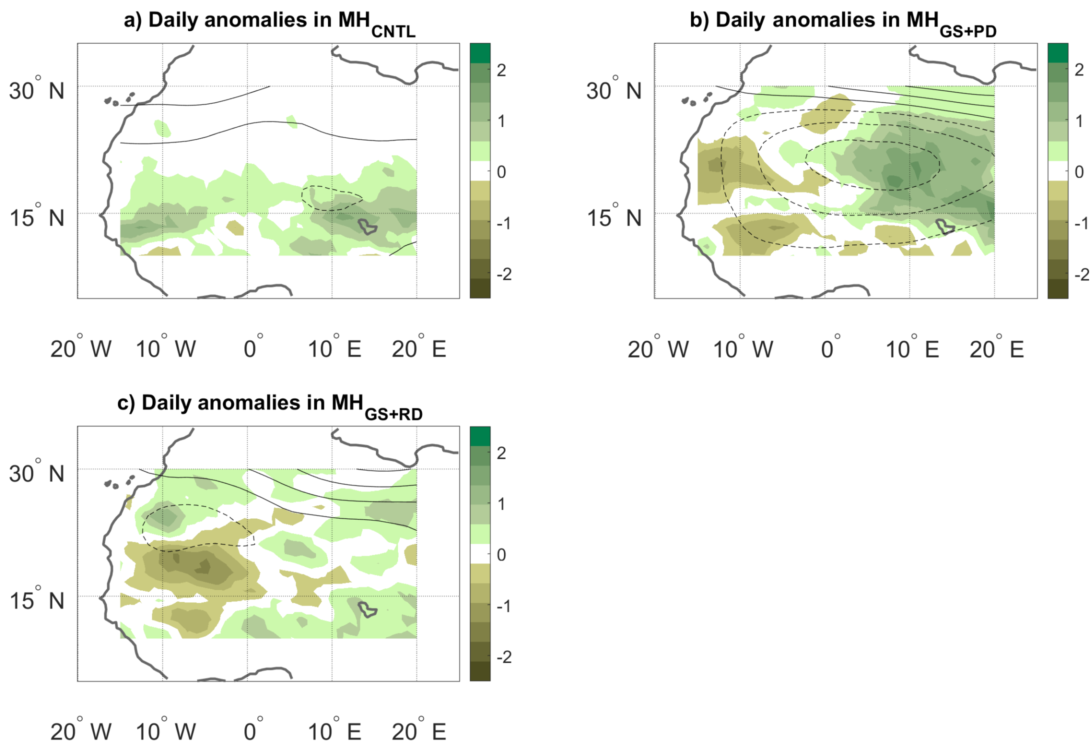

Figure A7. Same as Fig. 6, but using the domain (10 – 30 °N, 15 °W – 20 °E).