DYNAMICAL PROPERTIES AND EXTREMES OF NORTHERN HEMISPHERE CLIMATE FIELDS OVER THE PAST 60 YEARS

Davide Faranda[1,2], Gabriele Messori[3,4], M. Carmen Alvarez-Castro[1], and Pascal Yiou[1]

[1]LSCE-IPSL, CEA Saclay l'Orme des Merisiers, CNRS UMR 8212 CEA-CNRS-UVSQ, Université Paris-Saclay, 91191 Gif-sur-Yvette, France
[2]London Mathematical Laboratory, 14 Buckingham Street, London, WC2N 6DF, UK
[3]Department of Meteorology, Stockholm University and Bolin Centre for Climate Science, Stockholm, Sweden
[4]Currently: visiting researcher at LSCE-IPSL, CEA Saclay l'Orme des Merisiers, CNRS UMR 8212 CEA-CNRS-UVSQ, Université Paris-Saclay, 91191 Gif-sur-Yvette, France

Correspondence to: D Faranda (davide.faranda@cea.fr)

Abstract. Atmospheric dynamics is described by a set of partial differential equation yielding an infinite dimensional phase space. However, the actual trajectories followed by the system appear to be constrained to a finite dimensional phase space, i.e. a strange attractor. The dynamical properties of this attractor are difficult to determine due to the complex nature of atmospheric motions. A first step to simplify the problem is to focus on observables which affect — or are linked to phenomena which affect — human welfare and activities, such as sea level pressure, 2 meter temperature and precipitation frequency. We make use of recent advances in dynamical systems theory to estimate two instantaneous dynamical properties of the above fields for the Northern Hemisphere: local dimension and persistence. We then use these metrics to characterise the seasonality of the different fields and their interplay. We further analyse the large-scale anomaly patterns corresponding to phase-space extremes — namely timesteps at which the fields display extremes in their instantaneous dynamical properties. The analysis is based on the NCEP/NCAR reanalysis data, over the period 1948–2013. The results show that: i) despite the high dimensionality of atmospheric dynamics, the Northern Hemisphere sea level pressure and temperature fields can on average be described by roughly twenty degrees of freedom; ii) the precipitation field has a higher dimensionality; iii) the seasonal forcing modulates the variability of the dynamical indicators and affects the occurrence of phase-space extremes. We further identify a number of robust correlations between the dynamical properties of the different variables.

1 Introduction

Atmospheric motions are governed by a web of complex interactions among the different components of the earth system (Charney, 1947). Solar radiation and the earth's rotation are the primary large-scale drivers of the chaotic atmospheric dynamics, while turbulent motions add a layer of complexity at small scales. This picture is further complicated by the presence of features such as ocean-land interactions, vegetation, anthropocentric forcing and the hydrological cycle. Understanding both the transient (i.e. meteorological) and mean (i.e. climatic) properties of this system is one of today's major scientific challenges.

Since Lorenz (1963)'s seminal work, dynamical systems techniques have been widely applied to the study of the atmosphere. For example, the use of tools such as the Lyapunov exponents or the Kolmogorov-Sinai entropy have led to important advances in our understanding of atmospheric predictability (Zeng et al., 1993). An important result has been to show that atmospheric motions are chaotic but not random: their trajectories stay close to a high dimensional object called an attractor (Lorenz, 1969; Carrassi et al., 2008; Ghil et al., 2008; Vannitsem, 2014). This object occupies only a fraction of the atmospheric phase space, meaning that its dimension D is smaller than the number of variables used to describe the system. D is an important quantity because it represents the number of degrees of freedom of the system, namely the minimum number of variables needed to represent the dynamics. The computation of D for atmospheric attractors has posed a challenged to the dynamical systems community for several decades. Whereas in the early '80s several estimates pointed to a low dimensional $D < 10$ attractor (Fraedrich, 1986), a later review of the numerical limitations of the available techniques suggested that they tended to underestimate D for complex systems (Lorenz, 1991). However, further estimates of D were hardly attempted, because $D > 10$ implies that low dimensional models should fail in describing the atmospheric dynamics.

D is a mean property, since it describes the dimension of the attractor for the whole atmospheric trajectory. However, it is often more useful to determine instantaneous dynamical systems metrics that describe transient states ζ of complex attractors. A quantity that contains such information is the local dimension $d(\zeta)$ (Lucarini et al., 2016). The value of d is proportional to the active number of degrees of freedom and provides information on how predictable the state ζ and its future evolution are (Faranda et al., 2017). By averaging d over all possible ζ, one recovers the attractor dimension D. Unfortunately, the computation of d has posed even greater challenges than that of D. The original method developed by Liebovitch and Toth (1989) used box counting techniques. First, a small portion of the phase space is partitioned in hypercubes of different sizes. One then looks at the amount of space filled up in each hypercube. The scaling of this quantity across different scales is proportional to d. The complexity of this technique prevented computations for high dimensional systems such as atmospheric flows. Very recently, some of the authors of this paper have contributed to develop an alternative way to obtain d, based on the universal behavior of Poincaré recurrences in chaotic systems (Freitas et al., 2010; Faranda et al., 2011, 2013). In a few words (see Methodology and Data for the details), the recurrences of a state ζ of a chaotic dynamical system of arbitrary dimension have a universal asymptotic distribution in the limit of infinite recurrences. The parameters of this distribution are linked to the instantaneous dimension $d(\zeta)$ and to another important dynamical quantity, namely the inverse of the average persistence time

of the trajectory around ζ (Freitas et al., 2012). Estimating these parameters via Poincaré recurrences is easier than with the box counting algorithms because the method avoids altogether computations in scale space. Since the asymptotic distribution is known, one can numerically check that enough recurrences are taken into account for the parameter estimates by performing standard statistical tests (see Faranda et al. (2017).

Having overcome the technical difficulties inherent to the calculation of instantaneous dynamical systems metrics, the remaining step is the choice of the states ζ of interest. Since it is unpractical — not to say impossible — to consider all atmospheric observables at once, we focus our analysis on variables which are representative of events which affect human welfare and society, namely: sea-level pressure (slp; cyclones, windstorms etc.), 2 meter temperature (t2m; heat waves, cold spells) and precipitation frequency (prp; droughts, floods). In dynamical systems terms, these fields represent projections of the full phase space dynamics onto specific subspaces, called Poincaré sections. In Faranda et al. (2016) and Messori et al. (2017) we have shown that d and θ can be used to characterise regional-scale atmospheric fields and further provide information on the predictability linked to a given atmospheric state. It is therefore important to investigate these indicators at different spatial scales to fully understand the insights they can provide. In this study we present a novel analysis based on d and θ computed for the whole Northern Hemisphere (NH). We further investigate for the first time the mutual correlations between the dynamical properties of different climate variables.

The paper is organised as follows: in section 2 we give an overview of the dynamical indicators, the methodology to compute them and the data used. In section 3 we present and discuss the dynamical properties of each of the three atmospheric fields separately, while in section 4 we analyse them jointly. Finally, we discuss our results and summarise our conclusions in section 5.

2 Methodology and Data

The attractor of a dynamical system is a geometrical object defined in the space hosting all the possible states of the system (the so-called phase space) (Milnor, 1985). Each point on the attractor ζ can be characterized by two dynamical quantities: i) the local dimension $d(\zeta)$, which provides the number of degrees of freedom active locally around ζ and ii) the inverse persistence of the state ζ: $\theta(\zeta)$, which is a measure of the mean residence time of the system around ζ.

2.1 Local dimensions

The term *attractor dimension* D usually refers to a global measure (Grassberger and Procaccia, 1983). D indicates the average number of degrees of freedoms of a dynamical system. Several methods to measure D were developed in the '80s (Grassberger and Procaccia, 1984; Halsey et al., 1986). These techniques have a certain number of adjustable parameters and require the system to be embedded in a subspace of the phase space. They provide good estimates of D only when the trajectories are sufficiently long to estimate the embedding parameters. Such computations are therefore problematic in systems with large

numbers of degrees of freedom and give biased results when applied to atmospheric flows (Grassberger, 1986; Lorenz, 1991).

The technique we exploit here results from the application of extreme value theory to Poincaré recurrences in dynamical systems (Freitas et al., 2010; Faranda et al., 2011). In this approach, the returns to points on chaotic attractors are fully characterised by extreme value laws. In practice, one needs long trajectories $x(t)$ that approximate sequences of states on the attractor. One then fixes a point ζ on the trajectory and computes the probability P that $x(t)$ returns in a ball of radius ϵ centered on the point ζ. The Freitas et al. (2010) theorem, modified in Lucarini et al. (2012b), states that logarithmic returns: $g(x(t)) = -\log(dist(x(t), \zeta))$ are distributed as:

$$P(g(x(t)) > s(q), \zeta) \simeq \exp\left[-\frac{x(t) - \mu(\zeta)}{\sigma(\zeta)}\right]$$

Here s is a high threshold associated to a quantile q of the series $g(x(t))$ itself, linked to the radius ϵ via $s = g^{-1}(\epsilon)$. In other words, requiring that the orbit falls within a ball of radius ϵ around the point ζ is equivalent to asking that the series $g(x(t))$ is over the threshold s. Here we adopt $q = 0.98$ to determine s. The resulting distribution is the exponential member of the Generalized Pareto Distribution family. The parameters μ and σ depend on the point ζ chosen on the attractor. $\sigma(\zeta)$ then provides the local dimensions $d(\zeta)$ via the simple relation $\sigma = 1/d(\zeta)$. This result is very powerful because it provides a new way to compute local dimensions on the attractor and to recover D as the average of d on all the ζs without the need for embedding, as is required in most dimension computation algorithms Grassberger and Procaccia (1984). We want to stress that this procedure is not just a statistical fitting. The reason why it provides good estimates of d that were impossible to obtain with previous techniques derives from the universality of the extreme value statistics for Poincaré recurrences: one knows a priori the statistics of such recurrences and can then check that they are achieved for the numerical trajectory examined.

2.2 Local persistence

The stability of the state ζ is measured by $\theta(\zeta)$: namely the inverse of the average residence time of trajectories around ζ. For discrete maps θ is uniquely defined (see Freitas et al. (2012) for details): if ζ is a fixed point of the dynamics, $\theta(\zeta) = 0$. For a point that leaves the neighborhood of ζ immediately, $\theta = 1$. For continuous flows, the definition of θ depends on the Poincaré map chosen and precisely on the Δt chosen to discretize the flow. Since θ is the inverse of the average residence time, it is measured in units of $1/\Delta t$. In general, the higher the persistence of the point ζ, the longer the previous and subsequent states of the system will resemble ζ. The residence time can be computed by introducing a further parameter in the previous law. This parameter, known as extremal index, is such that:

$$P(g(x(t)) > s(q)) \simeq \exp\left[-\theta\left(\frac{x - \mu(\zeta)}{\sigma(\zeta)}\right)\right]$$

To estimate θ, we adopt the Süveges estimator (Süveges, 2007). For a fixed quantile q, the estimator is defined as:

$$\theta = \frac{\sum_i^{N_c}(1-q)S_i + N + N_c - \sqrt{\left(\sum_i^{N_c}(1-q)S_i + N + N_c\right)^2 - 8Nc\sum_i^{N_c}(1-q)S_i}}{2\sum_i^{N_c}(1-q)S_i},$$

where N is the number of recurrences above the chosen quantile, N_c the number of observations which form a cluster of at least two consecutive recurrences and S_i the length of each cluster i. This length is the number of consecutive timesteps during which the trajectory remains within a radius ϵ of ζ. For further details on the derivation of this estimator, the reader is referred to Süveges (2007).

2.3 Data

We use daily fields from the NCEP/NCAR reanalysis (Kalnay et al., 1996), with a horizontal resolution of 2°. The analysis is carried out over the whole Northern Hemisphere for all days of the year over the period 1948–2013. The observables of interest are sea-level pressure, 2 meter temperature and precipitation frequency. A previous study (Faranda et al., 2017) has shown that the results obtained are largely independent of the dataset used and of its spatial resolution.

Anomalies are defined as deviations from the long-term daily mean. So, for example, the anomaly of t2m at a given location on the 5^{th} December 2000 is computed relative to the mean value of all 5^{th} Decembers in the dataset at that location.

The relevance of the composite anomaly maps for the different variables is evaluated using a sign test. This identifies geographical areas where at least 2/3 of the composite maps have the same-sign anomaly. Assuming a binomial process with the same number of draws as the composite maps and equal chances of positive or negative outcomes (binomial distribution with success rate 0.5), a 2/3 threshold is beyond the 99.99th percentile of the distribution. In Figures 5, 8 and 11, such regions are marked by the thick black lines.

3 Dynamical properties of individual observables

3.1 Sea-Level Pressure (d, θ)

The local dimension, d, of the slp field shows a marked variability throughout the analysis period, with values ranging from as low as 8.6 to as high as 33.2 (Figure 1a). The average dimension D, which in this case is roughly 19.4, therefore provides incomplete information concerning the field of interest, since the number of locally active degrees of freedom (identified by d) can vary by a factor of almost 4. The autocorrelation function (ACF) of d (Figure 1c) highlights a robust variability pattern which is not immediately evident from the raw time series. There is a clear semi-annual cycle, with peak autocorrelation values in excess of 0.21. Over a full year there are therefore two positive and two negative peaks in autocorrelation, with the second positive peak typically displaying a larger magnitude than the first. This is consistent with previous analyses which have identified a strong seasonal dependence in d (Faranda et al., 2017; Rodrigues et al., 2017). The presence of a semi-annual cycle leads us to interpret the ACF as being modulated by the four seasons, with the first positive peak corresponding to cross-season

correlation and the second, larger peak corresponding to correlation between the same seasons in successive years. This periodicity could be linked to semi-annual slp variability features in the mid-latitudes (Schwerdtfeger and Prohash, 1956).

The inverse persistence, θ, shows a marked variability with values ranging from 0.28 to 0.65 (i.e. 1.6 to 3.6 days in terms of $1/\theta$) (Figure 1b). We note that these values should not be compared directly to the persistence of the traditional weather regimes defined using clustering algorithms as: i) here we consider a full hemispheric domain, while weather regimes are typically computed for specific regions and ii) the requirement that the flow does not leave the neighborhood of the state ζ is a more restrictive condition than continued permanence within a given cluster. Indeed, if one considers the typical partition f the atmospheric patterns over the North Atlantic into 4 weather regimes, the probability of being in one of them is of order 0.25, whereas the probability of being close to zeta is set by the threshold s — in our case 0.02 (see section 2.1). Concerning the ACF (Figure 1d), θ shows a very different pattern to d. The year-to-year correlation between same seasons is still large and positive, but the semi-annual oscillation seen in d is almost entirely absent. Indeed, the winter values appear to be anticorrelated to those of the other three seasons.

An analysis of monthly mean values confirms the strong seasonal control on the dynamical characteristics of the field. In the summer months, both the magnitude (Figure 2a,b) and variability of the two metrics reaches a minimum (Figure 2c,d). With the autumn season, the local dimension and θ increase rapidly while the variability remains low. As winter progresses the variability increases, θ remains roughly constant while the local dimension shows a marked decrease, albeit remaining well above the summertime values. In spring d grows back to values similar to those seen in autumn, while its variability peaks and θ starts decreasing. This picture is consistent with the ACFs described above. The annual cycle of θ can be explained as follows: stability peaks in summer when the mid-latitude storm tracks and wave activity are comparatively weak, decreases in autumn and winter and starts increasing again during spring. For the annual cycle of d, the maxima occurring in the intermediate season can be explained as follows: assuming that there is a *winter* and a *summer* attractor, the transitional seasons are more unstable because the atmospheric flow can explore both the summer and winter configurations. In dynamical systems terminology, the spring/autumn atmospheric flow sits on a saddle-like point of the dynamics.

One can further look at the slp anomalies corresponding to extremes in d (Figure 3), here defined as events beyond the 0.02 and 0.98 percentiles of the full distribution (see dashed lines in Figure 4a). The maxima of d occur primarily during the spring and autumn months, while the minima are mainly found in summer and winter (Figure 3). This mirrors the monthly-mean values discussed above. The d maxima correspond to a complex anomaly pattern spanning the whole hemisphere, but a sign test shows that in both the spring and autumn seasons there is very little agreement between the individual events (Figure 5a). This suggests that there is no single, dominant hemispheric-scale slp configuration leading to large dimensional extremes. The wintertime d minima again display very limited sign agreement, with the only significant feature being an intensification and eastward extension of the climatological Aleutian Low-Pressure Centre (Figure 5b). Such result is very different from what has recently been observed in the North Atlantic region, where both d minima and d maxima systematically correspond to precise

large-scale features (Faranda et al., 2017). On the contrary, the summer minima display an extensive and significant region of negative anomalies over the pole. The former pattern is an enhancement of the relatively low climatological slp values seen over the Arctic basin during the summer months. This results in a strengthened climatological meridional gradient (and presumably a strengthened polar vortex and a reduced airmass exchange between the mid and high latitudes, although we recognise that slp is not the optimal field to diagnose this) and we therefore hypothesise that it matches a relatively predictable configuration.

We next analyze slp anomalies corresponding to extremes in θ. The θ maxima occur predominantly during the autumn and winter months, while the minima are mostly found in spring and summer (Figure 3,c,d). The high mean persistence found in the summer months therefore also corresponds to instantaneous maxima in this quantity. The maxima of θ in both autumn and winter correspond to a circum-hemispheric wave-like structure and show very little sign agreement (Figure 5c). Since θ maxima are by definition unstable states, the lack of sign agreement might simply be due to the zonal propagation of the wave-like anomalies in time, although we do not explore this idea further here. The θ minima correspond to a mostly zonally symmetric pattern with a significant positive slp anomaly over the pole and locally significant negative anomalies throughout the mid and low latitudes (Figure 5d). Over the North Atlantic, this results in a negative NAO-like dipole which is consistent with the anomaly pattern found for regional persistence maxima in the Euro-Atlantic domain (Faranda et al., 2017). The reversal in the sign of the polar anomalies relative to summertime d minima is difficult to interpret. However, we note that Faranda et al. (2017)'s regional analysis also showed opposite-sign anomalies in the high latitudes for d and θ minima. More generally, our results highlight that persistent, predictable states are primarily associated with zonally symmetric slp anomalies and therefore modulations of the zonal flow. Indeed, past studies have interpreted the hemispheric circulation as being dominated by a zonal-flow attractor with blocked or wavy states being associated with an unstable fixed point (Faranda et al., 2016). Similarly, an enhanced zonality of the large-scale flow has been linked to increased downstream predictability on regional scales (Messori and Caballero, 2015).

The seasonal control on the two dynamical systems metrics we discuss here can be further investigated through a d–θ scatterplot (Figure 4a). This highlights how each season forms a distinct diagonal band of relatively well-correlated d and θ values. Figure 4(b) indeed confirms that the two metrics have high cross-correlation values, with the lag-0 correlation approaching 0.7. The lagged cross-correlation shows a semi-yearly cycle, with the peak correlation values reflecting integer year shifts and peak anticorrelation values reflecting a shift of approximately one season. This latter feature can be easily understood in terms of the above analysis. Both d and θ peak in autumn and spring. If we imagine shifting the d curve forwards or backwards by one season, the autumn θ peak will now match a dimension trough, thus leading to a negative correlation. The smaller positive and negative peaks in the cross-correlation function correspond to shifts of one and three seasons, respectively, such that the two cycles are roughly in quadrature.

3.2 Temperature (d, θ)

The local dimension d of the t2m field shows a marked variability throughout the analysis period, with a range similar to that of $d(\text{slp})$: $8.9 < d(\text{t2m}) < 33.3$ (Figure 6a). The average dimension D is roughly 17.6, slightly lower than $D(\text{slp})$. The ACF of d (Figure 6c) again displays a semi-annual cycle, albeit with larger ACF values than those seen for the slp. We note that the ACF structure should not be linked directly to the large seasonal cycle in temperature, since here we are considering $d(\text{t2m})$, which is not necessarily linked to the absolute value of the field. The inverse persistence θ spans a range corresponding to periods between 1.9 and 6.3 days ($0.16 < \theta < 0.54$, Figure 6b), indicating a higher persistence than slp. θ's ACF (Figure 6d) is again different to that seen for d. The inter-year same-season correlation is still large and positive, but the semi-annual oscillation seen in d is entirely absent. Indeed, the winter values appear to be anticorrelated to those of the other three seasons, albeit with some weak modulation in the negative correlation values on seasonal scales. This difference is driven by the small offset between the seasonal cycles of d and θ, as discussed below.

An analysis of monthly mean values (Figure 2) confirms the strong seasonal control on the dynamical characteristics of the field, but also highlights a radically different picture from that seen for the slp. In the summer months, d and its variability peak while θ and its variability display a local maximum. With the autumn season, both the local dimension and θ reach a local minimum, only to increase again during wintertime. During spring, both metrics display a second minimum before returning to their high summertime values. The seasonal cycle in the variability of both indicators roughly matches that of the indicators themselves. The fact that the monthly-mean minima in d are broader and occurr with a one-month shift relative to those in θ accounts for why the semi-annual ACF cycle is only seen in the former variable. The general picture is therefore consistent with the ACFs described above. We hypothesise that the summertime and wintertime local maxima in d are associated to the inherent difficulty in forecasting the onset and duration of warm and cold spells (Sillmann et al., 2013; Matsueda, 2011). This is presumably linked to a high-dimensional atmospheric configuration, namely one with a large number of allowed preceding and future evolutions. The annual cycle of θ also displays summertime and wintertime local maxima and suggests that the winter (and to a lesser degree the summer) temperature fields are comparatively unstable, while the transitional seasons have a more sluggish dynamical evolution. This can be linked to the presence of wintertime cold spells and summertime heat waves which are usually non-stationary and locally short-lived, although notable exceptions can occur. One can therefore picture the t2m dynamics as following a single potential well configuration, with the extremes located in winter and summer. The dynamics of slp and t2m are very different, and the interaction between the two could be akin to a Langevin-like model (Lucarini et al., 2012a). The slp would be the variable pushed into the *winter* or *summer* potential wells, while the temperature acts as a forcing noise term with extremes in winter and summer.

One can further look at the t2m anomalies corresponding to extremes in d and θ, again defined as events beyond the 0.02 and 0.98 percentiles of the full distribution (see dashed lines in Figure 7a). The d maxima occur predominantly during summer, while the minima are mostly found in autumn (Figure 3e,f), consistently with the seasonal cycle described above. The spatial

anomalies corresponding to d maxima are weak and display very little sign agreement (Figure 8a), suggesting that there is no single large-scale pattern matching these extremes. This is consistent with the theory that the high dimensionality of the temperature field may be associated with warm and cold spells, which are highly non-stationary and therefore will not emerge in a composite plot. Interestingly, the only region showing strong sign agreement is over East Africa and also emerges in the composite anomalies for the dynamical extremes of prp (see below). The anomalies associated with minima in d are stronger, but show a similarly low sign agreement (Figure 8b). The θ maxima and minima occur predominantly during winter and spring, respectively (Figure 3g,h), and again show low sign agreement (Figure 8c, d). While, as discussed above, the dynamical extremes elucidate a number of features of the temperature's seasonality and variability, they therefore seem to afford relatively little insight concerning its geographical nature.

In the d–θ scatterplot for t2m, the winter months form a cluster corresponding to high θ, relatively high d values, while the spring and autumn seasons form a low θ, low d cluster (Figure 7a). Summertime forms a continuation of the spring/autumn band, extending it toward higher d and θ values. The relatively broad scatter of the cloud of points to a weak correspondence between d and θ. Indeed, the two metrics show lower cross-correlation values than those seen for slp, with a lag-0 correlation of just above 0.5 (Figure 7b). The lagged cross-correlation shows a semi-yearly cycle, with the peak correlation values reflecting full and half-year shifts, in agreement with the synchronous double peak in both metrics shown in Figure 8. Similarly, the two large negative peaks in the cross-correlation function correspond to shifts of one and three seasons, respectively, leading to situations in which the two yearly cycles are in anti-phase.

3.3 Precipitation Frequency (d, θ)

We construct a daily precipitation frequency variable as follows: we assign a value of 1 to each grid point and timestep for non-zero precipitation rates and a value of 0 otherwise. For this variable the statistical fit of the recurrences to the expected distribution is better than for the precipitation rate itself (not shown). Another motivation to use the precipitation frequency data can be found in the multifractal analysis performed by Langousis et al. (2009), where the authors warn against using the precipitation data directly to measure dimensionality. The local dimension d of the precipitation frequency (prp) shows a large variability throughout the analysis period, with markedly higher values than those of the previous variables: $48 < d(\text{prp}) < 132$ (Figure 9a). This is also reflected in the average dimension $D = 83.1$. These high values are consistent with the very scattered, noisy nature of the precipitation field. The autocorrelation function of d displays a semi-annual cycle with large autocorrelation values at full-year lags and near-zero values at 6-month lags (Figure 9c). This is very different from the positive 6-month autocorrelation values seen for $d(slp)$ and $d(t2m)$. The inverse persistence θ (Figure 9b) spans a range corresponding to periods between 1.0 and 2.1 days ($0.3 < \theta < 0.7$), indicating a lower persistence than the previous variables, compatible with the precipitation's noisy nature. θ's autocorrelation function displays features similar to those discussed in Sections 3.1 and 3.2 (Figure 9d), namely a yearly ACF peak with negative autocorrelations at intermediate lags.

An analysis of monthly mean values (Figure 2) reveals a marked semi-seasonal cycle in d, whose absolute values and variability both peak during spring and late summer/early autumn. These are seasons with enhanced convective precipitation at mid-latitudes. θ displays a similar variability behavior, while the magnitude has a minimum in summer and an extended period of higher values from autumn into early spring. During the summer months, the mature phase of NH monsoon systems provide comparatively persistent and predictable precipitation patterns. The high persistence might also be favored by the predominantly dry summers in the Mediterranean and other mid-latitude regions, with long dry spells at regional scale being the norm. This picture is consistent with the autocorrelation functions described above. In particular, the low absolute ACF values of d seen at lags of 2 to 10 months can be linked to the asymmetry in the positive and negative peaks. For example, for a lag of six months the first peak will roughly match the second peak, but the first trough will not match the second trough.

Notwithstanding the local and noisy nature of precipitation, a number of coherent features emerge from the geographical composites corresponding to extremes in the two dynamical systems metrics, again defined as events beyond the 0.02 and 0.98 percentiles of the full distribution (see dashed lines in Figure 10a). The bulk of the d and θ maxima occur during spring, while the minima are primarily associated with the summertime (Figure 3i-l). d maxima show predominantly negative precipitation anomalies (Figure 11a). Significant features include decreases over the East Asian Monsoon region, the Indian Ocean and continental and eastern Europe. Weak positive anomalies are mainly found over the central Pacific ocean, parts of the U.S.A. and the Canadian Arctic Archipelago. The anomalies associated with θ maxima closely track the d maxima throughout the northern hemisphere, albeit with a stronger preference for negative anomalies (Figure 11c). Since the positive extremes in both metrics occur predominantly during spring, we hypothesise that the anomalies over East Asia could be linked to the shift between the northerly flow associated with the winter monsoon and the southerly flow associated with the summer monsoon. It is reasonable to expect that this shift between very different large-scale flow configurations could contribute to the high local dimensions. Similarly, the anomalies over the Indian Ocean could be associated with the onset of the Indian Monsoon's large-scale flow. d minima (Figure 11b) show strong negative anomalies over the continental U.S.A. and southern Indochina and the strong positive anomalies over East Africa and the mid-latitude Pacific. θ minima again track closely the patterns seen for the d minima (Figure 11d). The East African positive anomalies, roughly corresponding to Northern Ethiopia, might indicate a modulation of the late-summer rainfall peak in the region (Gissila et al., 2004). Similarly, the widespread negative anomalies across eastern continental North America suggest a modulation of the wet season over the great plains and Northern Mexico.

The d–θ scatterplot (Figure 10a) shows two clouds of points and a clear separation between the late spring/summer and autumn/winter seasons. While both clusters span a wide range of d and θ, there is a clear vertical and horizontal offset between the two, with the former seasons corresponding to lower values than the latter. The two prp metrics show high cross-correlation features (Figure 10b), with the lag-0 correlation exceeding 0.6. The lagged cross-correlation shows a yearly cycle, a negative correlation for lags of 2-4 months and a local maximum for lags of around 7 months.

4 Cross-analysis of the dynamical properties

We next address the co-variability of the dynamical indicators of the different variables. In physical space, there is an obvious link between anomalies in the large-scale slp and 2 meter temperature fields. A similarly close link can be found between precipitation and temperature or slp anomalies. There are therefore strong grounds to expect some systematic relationships to emerge.

We begin by analysing the cross-correlation functions between the slp and t2m (Figure 12a,b). $d(\text{slp})$ and $d(\text{t2m})$ are anti-correlated at zero lag, as might be expected by their contrasting seasonal cycles described above. The lagged cross-correlations display a roughly regular semi-yearly cycle, which derives from the fact that both local dimensions display a double peak, albeit in different seasons. The cross-correlation between $\theta(\text{slp})$ and $\theta(\text{t2m})$ is more nuanced, owing to the fact that $\theta(\text{slp})$ displays high values throughout the autumn and winter while $\theta(\text{t2m})$ displays two well-separated peaks, one of which partially overlaps the months of high $\theta(\text{slp})$ values. The lag-0 correlations are positive, albeit low, and peak negative cross-correlations are achieved at lags of approximately 6-7 months. The lag-0 anticorrelation of the local dimensions points to the fact that it is rare to find co-occurring slp and t2m fields both displaying high predictability. This is compounded by the fact that positive correlations between the θ are generally weak, suggesting that persistent slp configurations do not necessarily match equally persistent t2m patterns. An example of this are wintertime cold spells in the mid-latitudes: while the large-scale circulation anomalies are often very persistent, the temperature can evolve rapidly with a build-up of cold t2m masses leading to a rapidly cooling region which then relaxes back to near-climatological values as soon as the anomalous circulation pattern weakens (Messori et al., 2016).

The cross-correlation between $d(\text{slp})$ and $d(\text{prp})$ is shown in Figure 12(c,d). At lag-0, the two variables have a moderate positive correlation, with peak positive values being reached for negative shifts of 1-2 months (i.e. slp leading prp). Indeed, $d(\text{prp})$ has a broad peak during the spring, then decreases rapidly through the summer season and peaks again in early autumn. These three features precede by roughly one month the corresponding ones in the $d(slp)$ signal. However, we note that peak cross correlation values are lower than those seen between $d(\text{slp})$ and $d(\text{t2m})$. The two persistence metrics, on the contrary, display peak correlation at lag-0, since they both display a spring maximum and a summer minimum. The large-scale circulation changes associated with the onset of the monsoonal precipitation over Asia and Africa and the high summertime persistence in the precipitation field therefore have a clear correspondence in $d(slp)$ and $\theta(slp)$, albeit with a small temporal lag in the case of the first.

The t2m–prp pair is analysed in Figure 12(e,f). In this case, the lag-0 cross-correlation between the local dimensions is very small, and peaks for prp leading temperature by roughly two months. A similar picture is seen for the persistence metrics, since the summer peak in $\theta(\text{t2m})$ is out of phase with the summer minimum in $\theta(\text{prp})$. From a dynamical systems perspective there therefore seems to be a significant lag between changes in the monthly-mean properties of the large-scale temperature

and precipitation signals.

The cross-correlations consider the time series of the different metrics as a whole, but provide little insight on the correlation between dynamical extremes. We conclude our analysis by looking at the d-θ scatterplots for the local dimensions and persistences of the three observables (Figure 13). The negative lag-0 correlation found for $d(\text{slp})$ and $d(\text{t2m})$ is evident (Figure 13a), while the other two d scatterplots (Figure 13c,e) show a more diffuse distribution, consistent with the low correlation values previously discussed. The strongest match between both positive and negative d extremes is found for the slp–prp pair. θ shows generally higher co-occurrences of extremes across all pairs, with the most frequent match being for the late springtime low θ extremes of the prp–t2m pair (Figure 13b, d, f). This indicates that: i) rapidly shifting slp patterns can lead to equally rapid shifts in the large-scale temperature and precipitation fields, and vice versa for persistent configurations; ii) that persistent 2 meter temperature and precipitation configurations show a systematic co-occurrence during the spring months.

5 Conclusions

In the present study we have applied recent advances in dynamical systems theory to estimate the local dimension and inverse persistence of instantaneous atmospheric fields over the Northern Hemisphere. Persistence is a very intuitive metric, which quantifies the average residence time of the system's trajectory in phase space within the neighbourhood of the point of interest. Local dimension is a proxy for the number of locally active degrees of freedom in the system, and can thus be directly linked to the number of possible configurations preceding and following the instantaneous field being analysed. We have specifically focused on three observables: sea-level pressure, 2 meter temperature and precipitation frequency. Despite the high dimensionality of atmospheric dynamics, we find that the Northern Hemisphere sea level pressure and low-level temperature fields can on average be described by roughly 15–20 degrees of freedom, while the noisier precipitation field has an average dimension of over 80. We further note that the dimension of the instantaneous fields can vary by almost a factor of 4 for a given observable. The links between the local dimension and persistence of a given variable can be complex. While the two generally show a positive lag-0 correlation, they can display very different seasonal cycles.

This study further analyses dynamical extremes, namely the instances where one – or both – dynamical systems metrics are at the positive or negative edge of their respective distributions. The dynamical extremes in d and θ of a given variable occur independently for 2 meter temperature but they coincide almost always for precipitation frequency and sea level pressure. Both d and θ are linked to atmospheric predictability, since a persistent, low-dimensional state is intrinsically easier to forecast than a rapidly shifting, high-dimensional situation. Fields where the co-occurrence of d and θ extremes is more frequent – such as is the case for slp – therefore provide more highly predictable (or unpredictable) configurations than those where the two occurrences are rarer.

We further identify a number of robust correlations between the dynamical properties of the different variables. For example, low-persistence cases in slp often indicate a similarly low persistence in t2m and prp. This is an intuitive relationship since rapidly shifting slp patterns can lead to equally rapid shifts in the large-scale temperature and precipitation fields. Similarly, persistent prp and t2m configurations often co-occur. Other links do not have a similarly straightforward physical interpretation. For example, the local dimensions of prp and t2m seem to be mostly uncorrelated, suggesting that predictable large-scale precipitation patterns do not directly affect the predictability of the t2m field, unlike what seen for persistence.

Our results do not always match those obtained in Faranda et al. (2017) for the North Atlantic region. Indeed, there is a strong dependence on the region chosen and the dynamics of the northern hemisphere include other degrees of freedom than the North Atlantic dynamics. This is coherent with the increase of the average dimension found. The dynamical indicators provide a different answer according to the region selected, meaning that predictability is overall different for the hemisphere than for regional mid-latitude dynamics. We conclude that the dynamical systems metrics we adopt here provide a wealth of information concerning the large-scale atmospheric processes and dynamics. We are convinced that this analysis framework will find applications in a wide number of climate studies.

Code availability. The code to perform the analysis is available upon request to the corresponding author.

Data availability. The dataset NCEP-NCAR is public available at https://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis.html

Author contributions. DF and GM performed the analysis. MCAC downloaded and organised the datasets. All the authors contributed to the writing.

Competing interests. No competing interests are present

Acknowledgements. P. Yiou, D. Faranda and G. Messori were supported by ERC grant No. 338965, M.C. Alvarez-Castro was supported by Swedish Research Council grant No. C0629701 and G. Messori was further supported by a grant from the Department of Meteorology of Stockholm University.

References

Carrassi, A., Ghil, M., Trevisan, A., and Uboldi, F.: Data assimilation as a nonlinear dynamical systems problem: Stability and convergence of the prediction-assimilation system, Chaos: An Interdisciplinary Journal of Nonlinear Science, 18, 023 112, 2008.

Charney, J. G.: The dynamics of long waves in a baroclinic westerly current, Journal of Meteorology, 4, 136–162, 1947.

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., Freitas, J. M., Lucarini, V., Turchetti, G., and Vaienti, S.: Extreme value statistics for dynamical systems with noise, Nonlinearity, 26, 2597, 2013.

Faranda, D., Masato, G., Moloney, N., Sato, Y., Daviaud, F., Dubrulle, B., and Yiou, P.: The switching between zonal and blocked mid-latitude atmospheric circulation: a dynamical system perspective, Climate Dynamics, 47, 1587–1599, 2016.

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

Fraedrich, K.: Estimating the dimensions of weather and climate attractors, Journal of the atmospheric sciences, 43, 419–432, 1986.

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.

Freitas, A. C. M., Freitas, J. M., and Todd, M.: The extremal index, hitting time statistics and periodicity, Advances in Mathematics, 231, 2626–2665, 2012.

Ghil, M., Chekroun, M. D., and Simonnet, E.: Climate dynamics and fluid mechanics: Natural variability and related uncertainties, Physica D: Nonlinear Phenomena, 237, 2111–2126, 2008.

Gissila, T., Black, E., Grimes, D. I. F., , and Slingo, J. M.: Seasonal forecasting of the Ethiopian summer rains, International Journal of Climatology, 24, 1345–1358, 2004.

Grassberger, P.: Do climatic attractors exist?, Nature, 323, 609–612, 1986.

Grassberger, P. and Procaccia, I.: Characterization of strange attractors, Physical review letters, 50, 346, 1983.

Grassberger, P. and Procaccia, I.: Dimensions and entropies of strange attractors from a fluctuating dynamics approach, Physica D: Nonlinear Phenomena, 13, 34–54, 1984.

Halsey, T. C., Jensen, M. H., Kadanoff, L. P., Procaccia, I., and Shraiman, B. I.: Fractal measures and their singularities: the characterization of strange sets, Physical Review A, 33, 1141, 1986.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., et al.: The NCEP/NCAR 40-year reanalysis project, Bulletin of the American meteorological Society, 77, 437–471, 1996.

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

Liebovitch, L. S. and Toth, T.: A fast algorithm to determine fractal dimensions by box counting, physics Letters A, 141, 386–390, 1989.

Lorenz, E. N.: Deterministic nonperiodic flow, Journal of the Atmospheric Sciences, 20, 130–141, 1963.

Lorenz, E. N.: Atmospheric predictability as revealed by naturally occurring analogues, Journal of the Atmospheric sciences, 26, 636–646, 1969.

Lorenz, E. N.: Dimension of weather and climate attractors, Nature, 353, 241, 1991.

Lucarini, V., Faranda, D., and Willeit, M.: Bistable systems with stochastic noise: virtues and limits of effective one-dimensional Langevin equations, Nonlinear Processes in Geophysics, 19, 9–22, 2012a.

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, 2012b.

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

Matsueda, M.: Predictability of Euro-Russian blocking in summer of 2010, Geophysical Research Letters, 38, n/a–n/a, doi:10.1029/2010GL046557, http://dx.doi.org/10.1029/2010GL046557, l06801, 2011.

Messori, G. and Caballero, R.: On double Rossby wave breaking in the North Atlantic, Journal of Geophysical Research: Atmospheres, 120, 2015.

Messori, G., Caballero, R., and Gaetani, M.: On cold spells in North America and storminess in western Europe, Geophysical Research Letters, 43, 6620–6628, 2016.

Messori, G., Caballero, R., and Faranda, D.: A Dynamical Systems Approach to studying Mid-Latitude Weather Extremes, Submitted: Geophysical Research Letters, 2017.

Milnor, J.: On the concept of attractor, Communications in Mathematical Physics, 99, 177–195, 1985.

Rodrigues, D., Alvarez-Castro, M. C., Messori, G., Yiou, P., Robin, Y., and Faranda, D.: Changes in the dynamical properties of the North Atlantic atmospheric circulation in the past 150 years, 2017.

Schwerdtfeger, W. and Prohash, F.: The semi-annual pressure oscillation, its cause and effects, Journal of Meteorology, 13, 217–218, 1956.

Sillmann, J., Kharin, V., Zhang, X., Zwiers, F., and Bronaugh, D.: Climate extremes indices in the CMIP5 multimodel ensemble: Part 1. Model evaluation in the present climate, Journal of Geophysical Research: Atmospheres, 118, 1716–1733, 2013.

Süveges, M.: Likelihood estimation of the extremal index, Extremes, 10, 41–55, 2007.

Vannitsem, S.: Dynamics and predictability of a low-order wind-driven ocean-atmosphere coupled model, Climate dynamics, 42, 1981, 2014.

Zeng, X., Pielke, R. A., and Eykholt, R.: Chaos theory and its applications to the atmosphere, Bulletin of the American Meteorological Society, 74, 631–644, 1993.

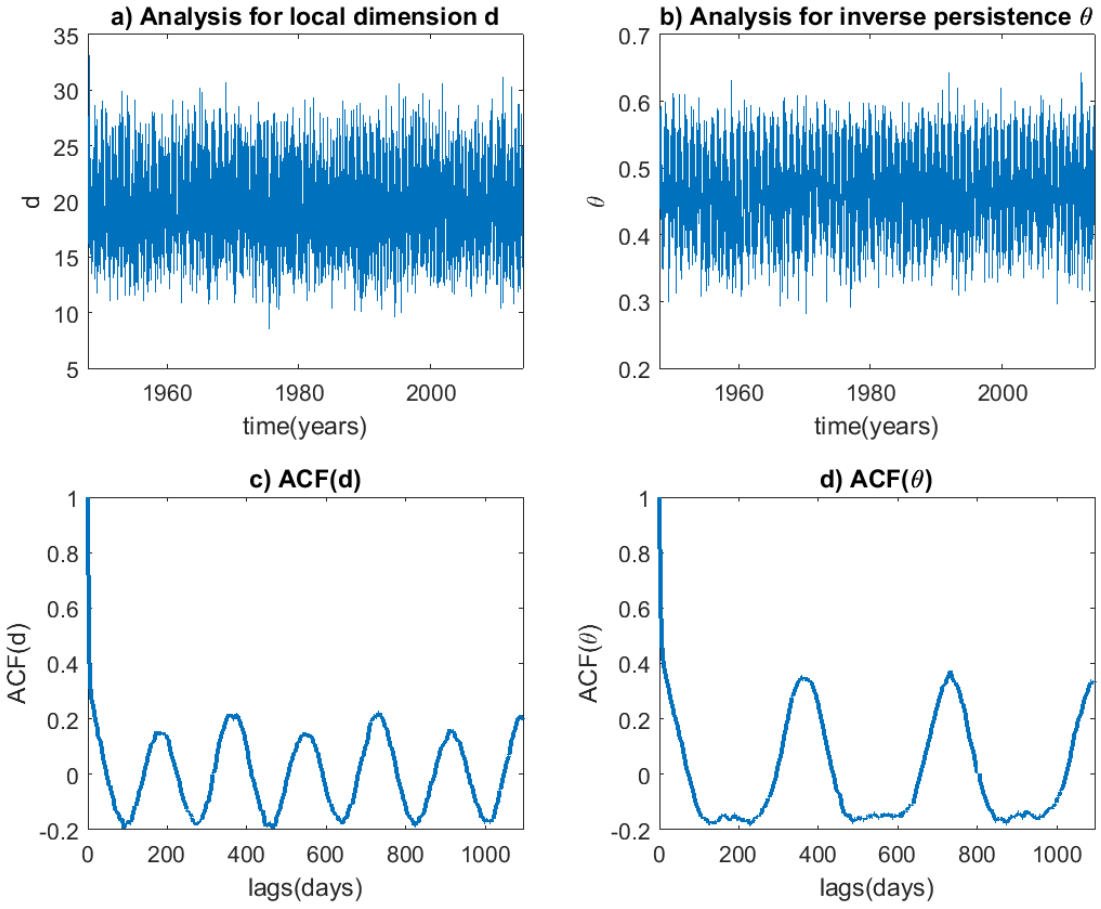

Figure 1. Statistics of local dimension d and persistence θ for daily sea-level pressure (slp) data from the NCEP-NCAR reanalysis. Time series of daily values of d (a) and θ (b). Autocorrelation function ACF(d) (c) and ACF(θ) (d) for three years in daily lags.

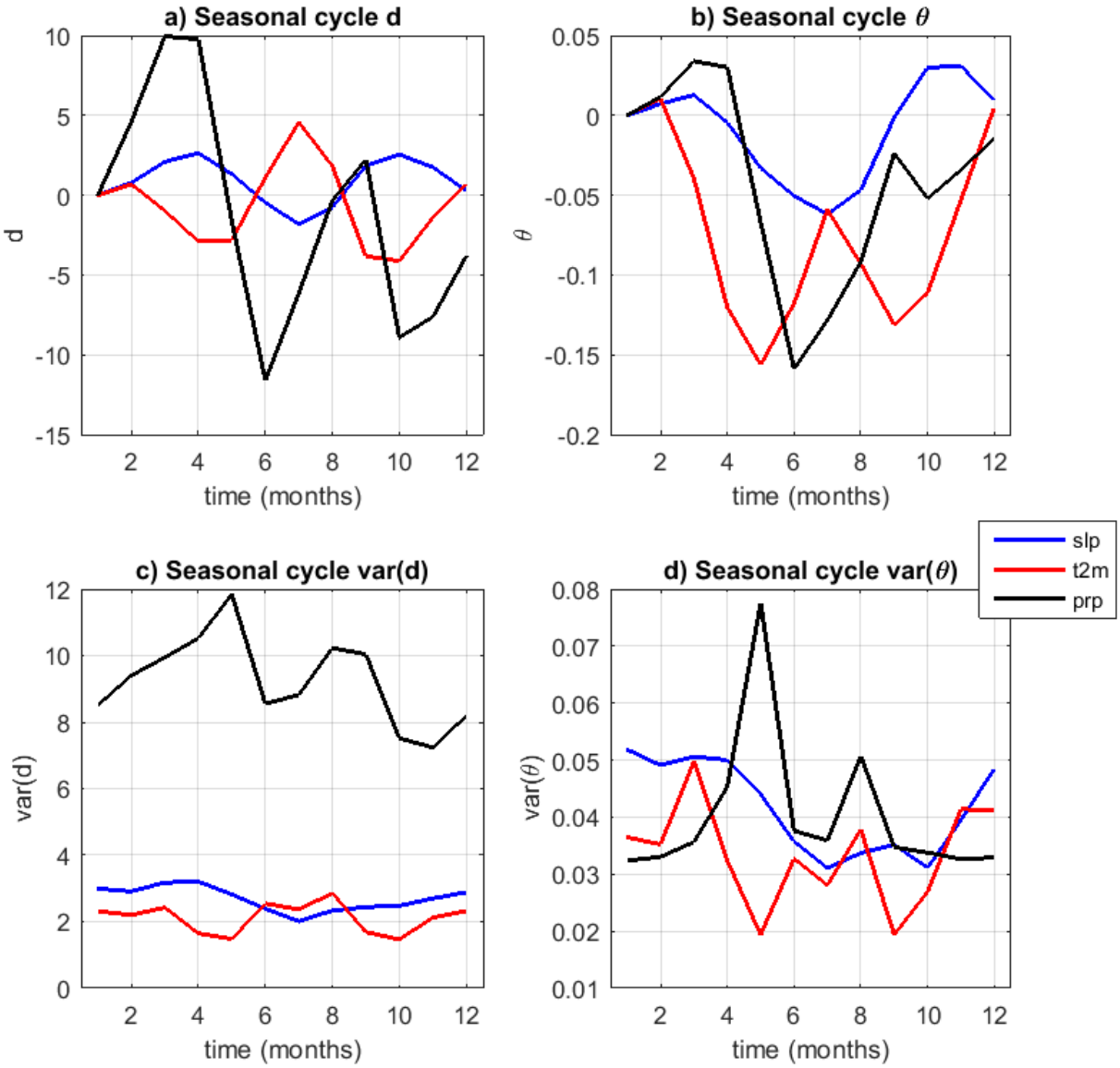

Figure 2. Monthly average values (a,b) and standard deviation (c,d) for the local dimension d (a,c) and inverse persistence θ (b,d) of sea-level pressure (slp, blue), 2 meter temperature (t2m, red) and precipitation frequency (prp, black) data.

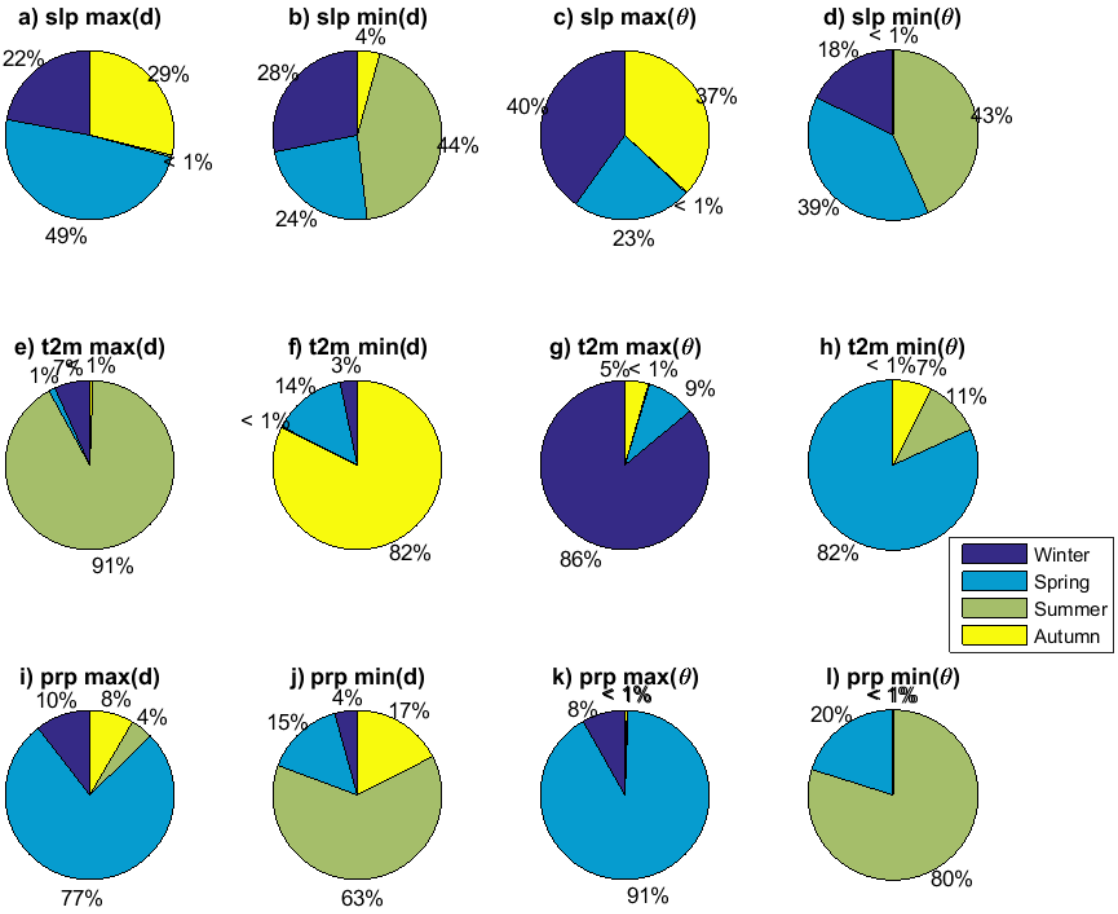

Figure 3. Occurrence of extremes (observation beyond the 0.02 (minima) and 0.98 (maxima) quantiles) of local dimension d and inverse persistence θ in different seasons. a-d) sea level pressure (slp), (e-h) 2 meter temperature (t2m), (i-l) precipitation frequency (prp).

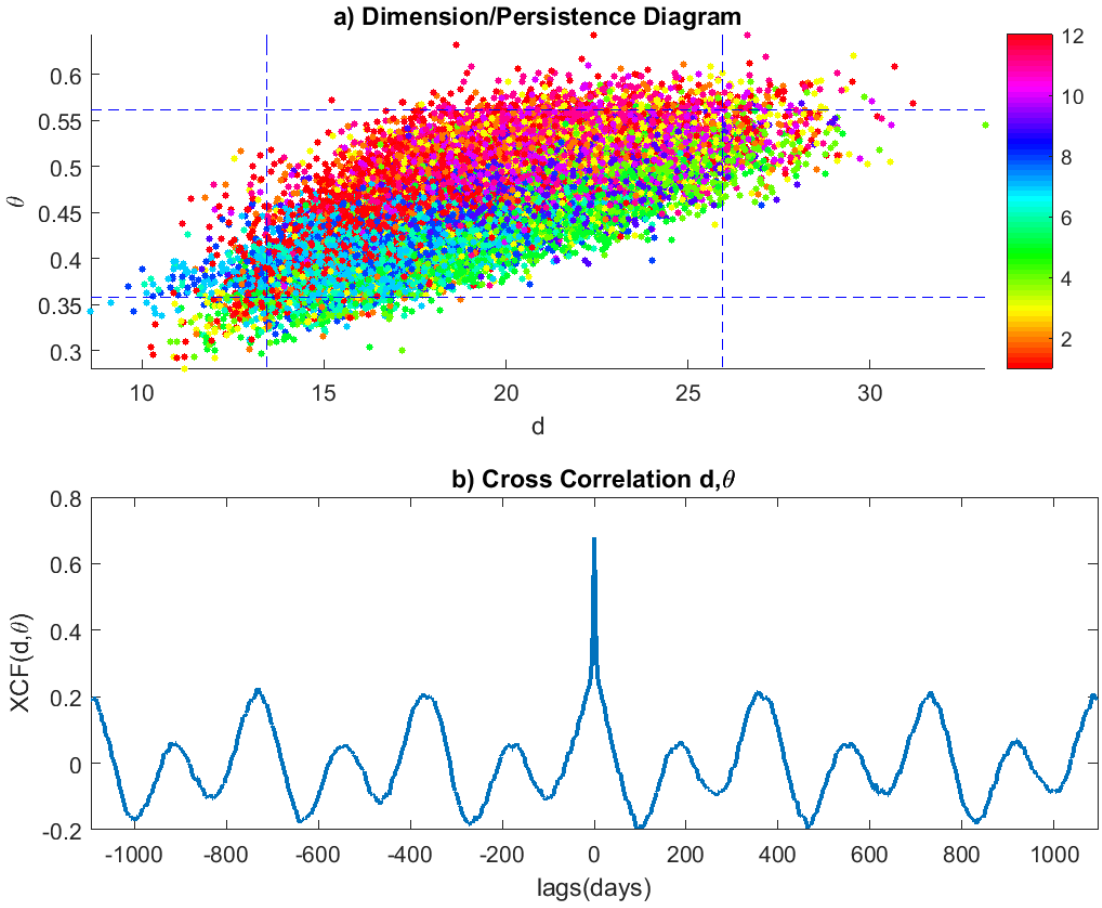

Figure 4. (a) Scatter-plot of local dimension d and inverse persistence θ for slp data. Each point represents the value corresponding to one day in the NCEP-NCAR reanalysis. The color indicates the month of the year the datapoint falls in. Blue dotted lines indicate the 0.02 and 0.98 percentiles of the d, θ distributions. (b) Corresponding cross correlation function between d and θ.

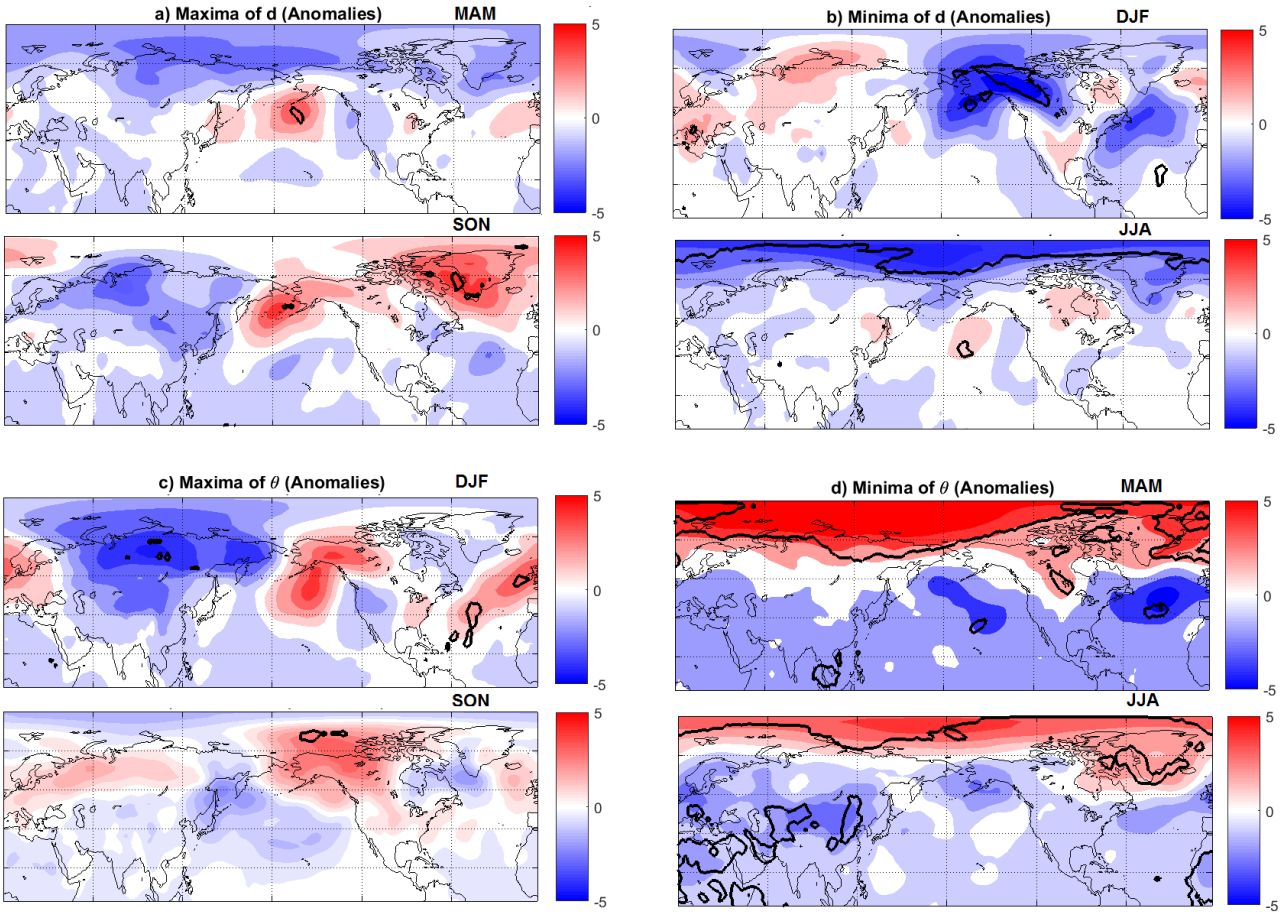

Figure 5. Composite anomalies with respect the seasonal cycle in sea-level pressure (slp) for the four phase-space regions delimited by the blue dotted lines in Fig. 4. Maxima of d (a), minima of d (b), maxima of θ (c), minima of θ (d). Geographical composites are shown for seasons which account for 25% or more of the extreme occurrences (Fig. 3). Units: hPa

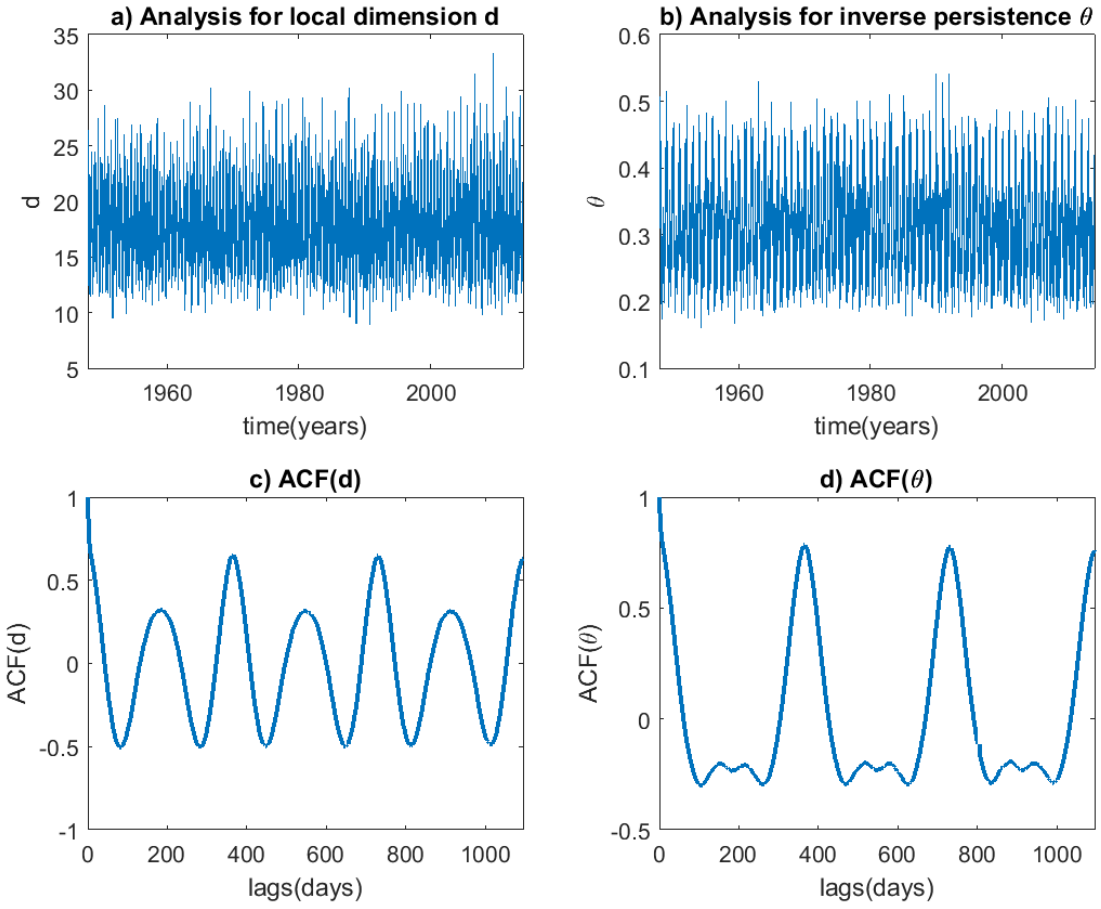

Figure 6. Statistics of local dimension d and persistence θ for daily 2 meter temperature (t2m) data from the NCEP/NCAR reanalysis. Time series of daily values of d (a) and θ (b). Autocorrelation function ACF(d) (c) and ACF(θ) (d) for three years in daily lags.

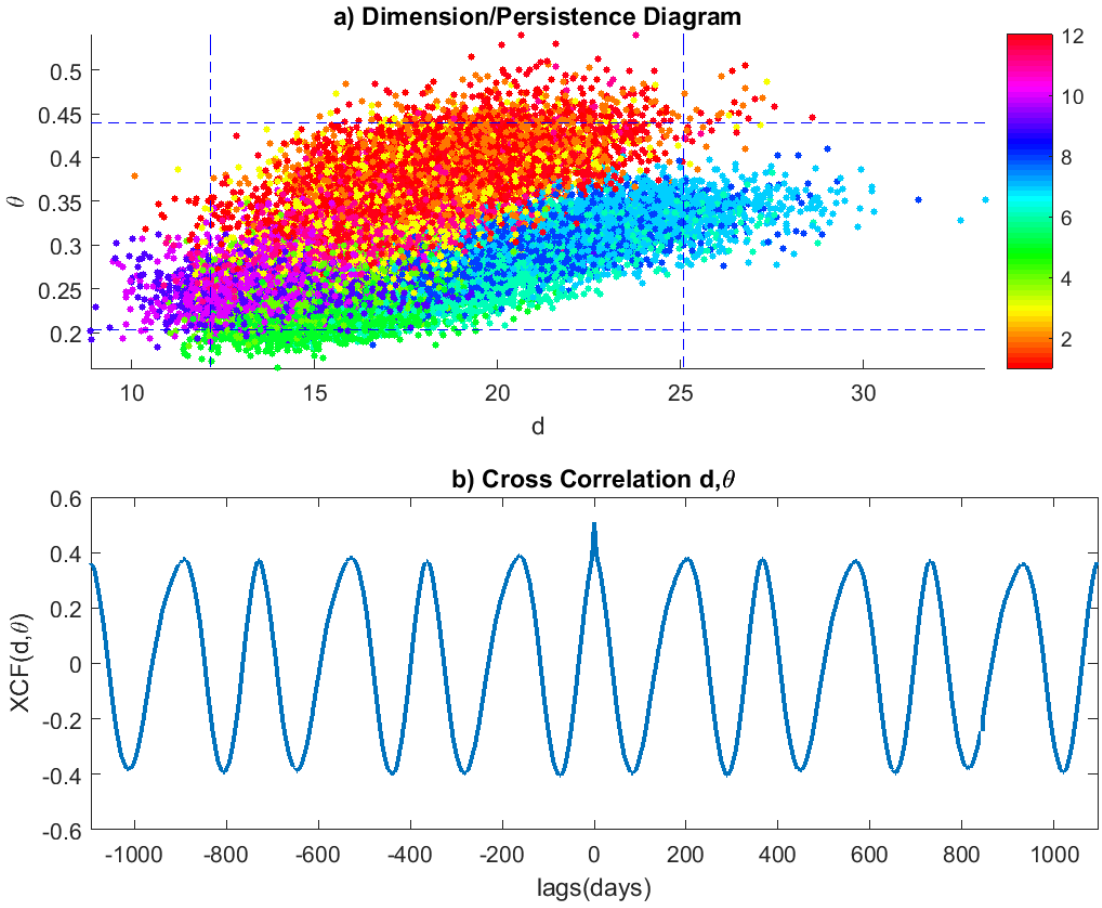

Figure 7. (a) Scatter-plot of local dimension d and inverse persistence θ for 2 meter temperature (t2m) data. Each point represents the value corresponding to one day in the NCEP/NCAR reanalysis. The color indicates the month of the year the datapoint falls in. Blue dotted lines indicate the 0.02 and 0.98 percentiles of the d, θ distributions. (b) Corresponding cross correlation function between d and θ.

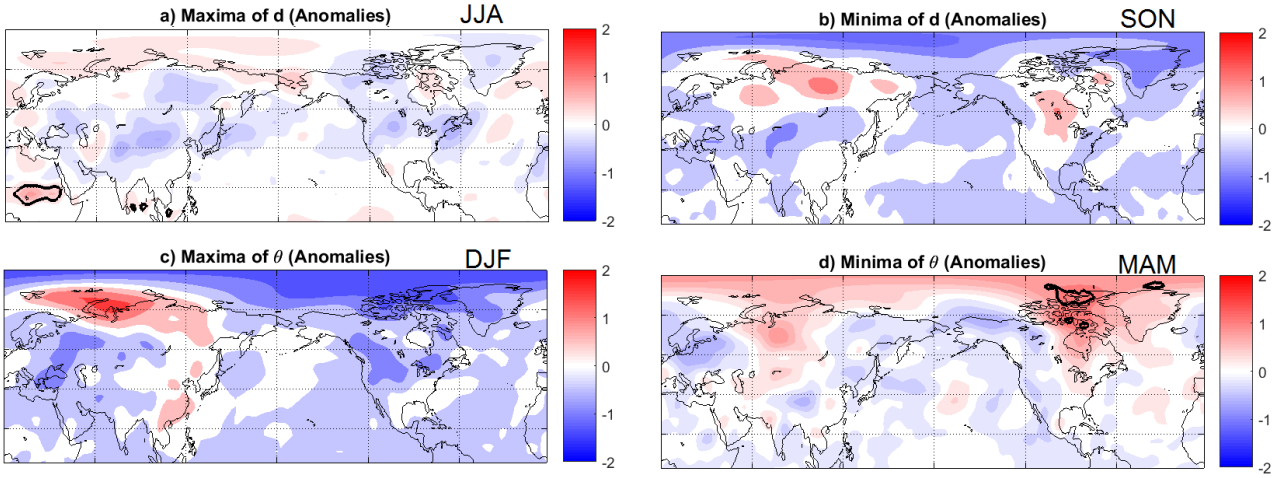

Figure 8. Composite anomalies with respect the seasonal cycle in 2 meter temperature (t2m) for the four phase-space regions delimited by the blue dotted lines in Fig. 7. Maxima of d (a), minima of d (b), maxima of θ (c), minima of θ (d). Geographical composites are shown for seasons which account for 25% or more of the extreme occurrences (ref Fig. 3). Units: $^\circ$ C

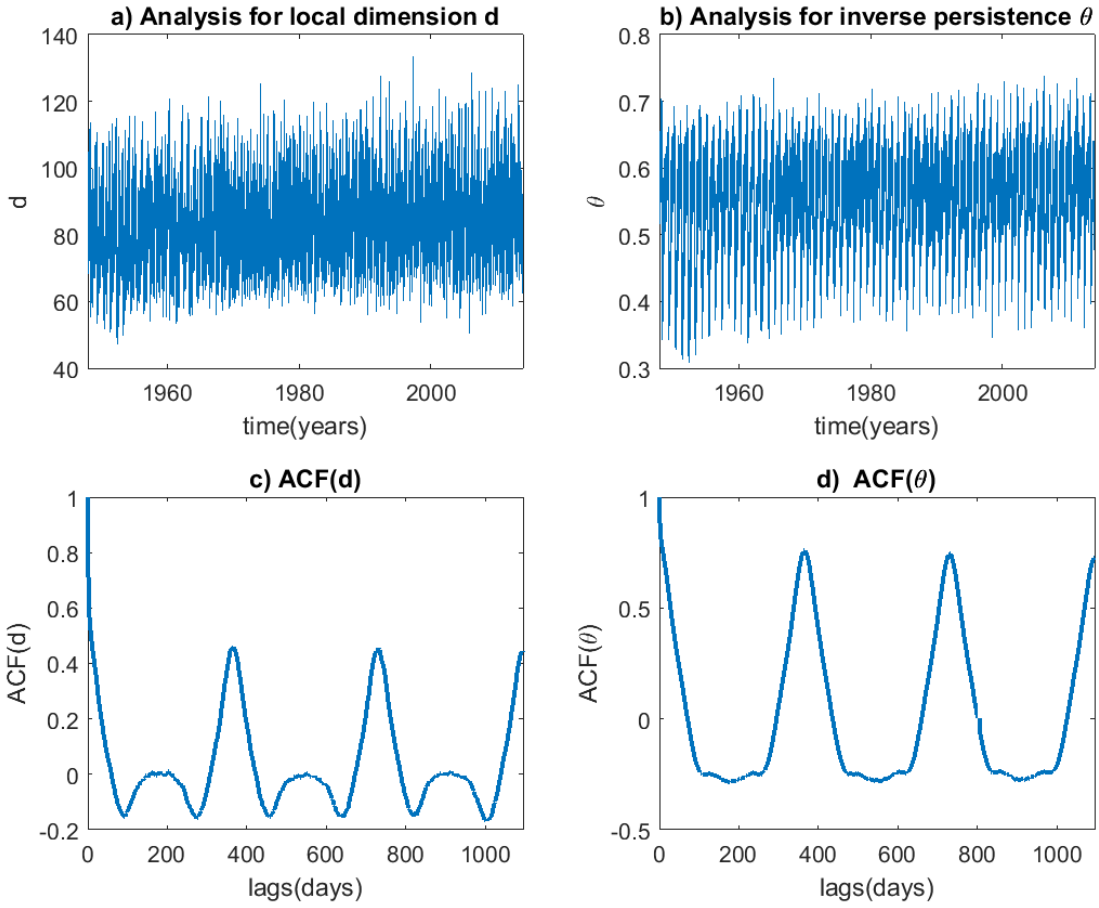

Figure 9. Statistics of local dimension d and persistence θ for precipitation frequency (prp) data from the NCEP/NCAR reanalysis. Time series of daily values of d (a) and θ (b). Autocorrelation function ACF(d) (c) and ACF(θ) (d) for three years in daily lags.

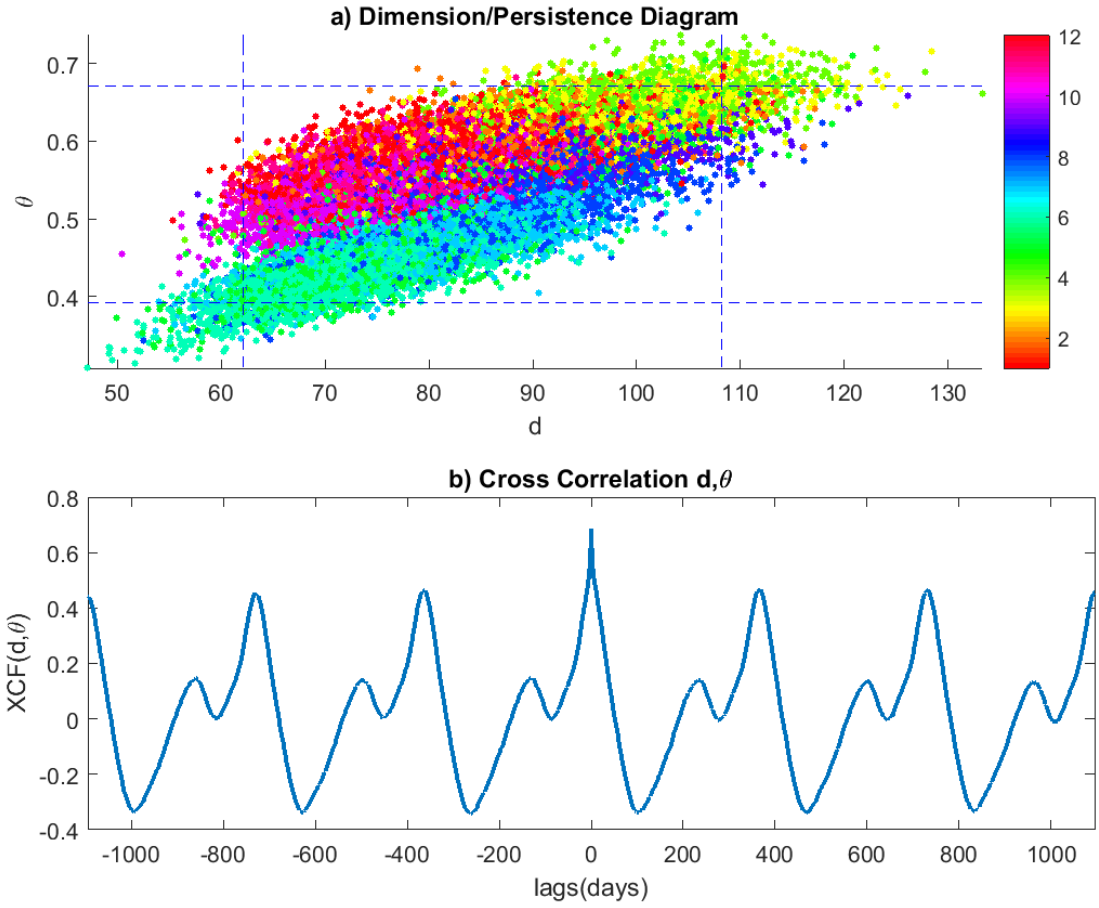

Figure 10. (a) Scatter-plot of local dimension d and inverse persistence θ for precipitation frequency (prp) data. Each point represents the value corresponding to one day in the NCEP/NCAR reanalysis. The color indicates the month of the year the datapoint falls in. Blue dotted lines indicate the 0.02 and 0.98 percentiles of the d, θ distributions. (b) Corresponding cross correlation function between d and θ.

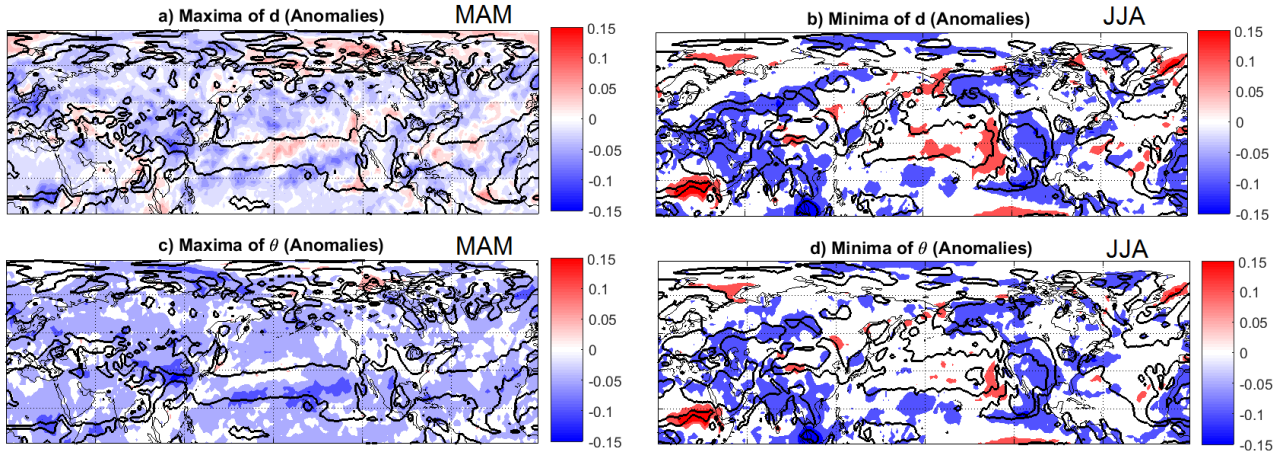

Figure 11. Composite anomalies with respect the seasonal cycle in precipitation frequency (prp) for the four phase-space regions delimited by the blue dotted lines in Fig. 10. Maxima of d (a), minima of d (b), maxima of θ (c), minima of θ (d). Geographical composites are shown for seasons which account for 25% or more of the extreme occurrences (Fig.3). Units: dimensionless quantity

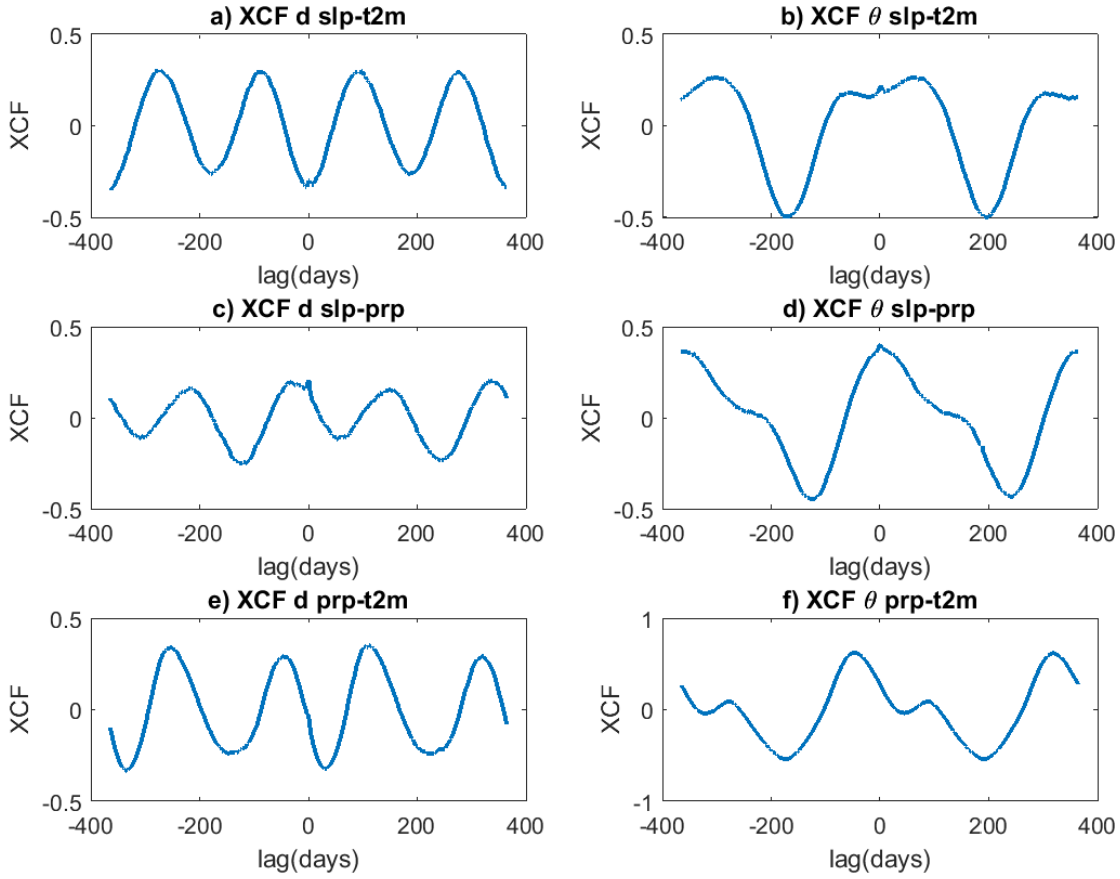

Figure 12. Cross correlation functions for the local dimensions d (a,c,e) and inverse persistences θ (b,d,f) of sea-level pressure (slp), 2 meter temperature (t2m) and precipitation frequency (prp) data. Positive lags indicate that the first variable precedes the second.

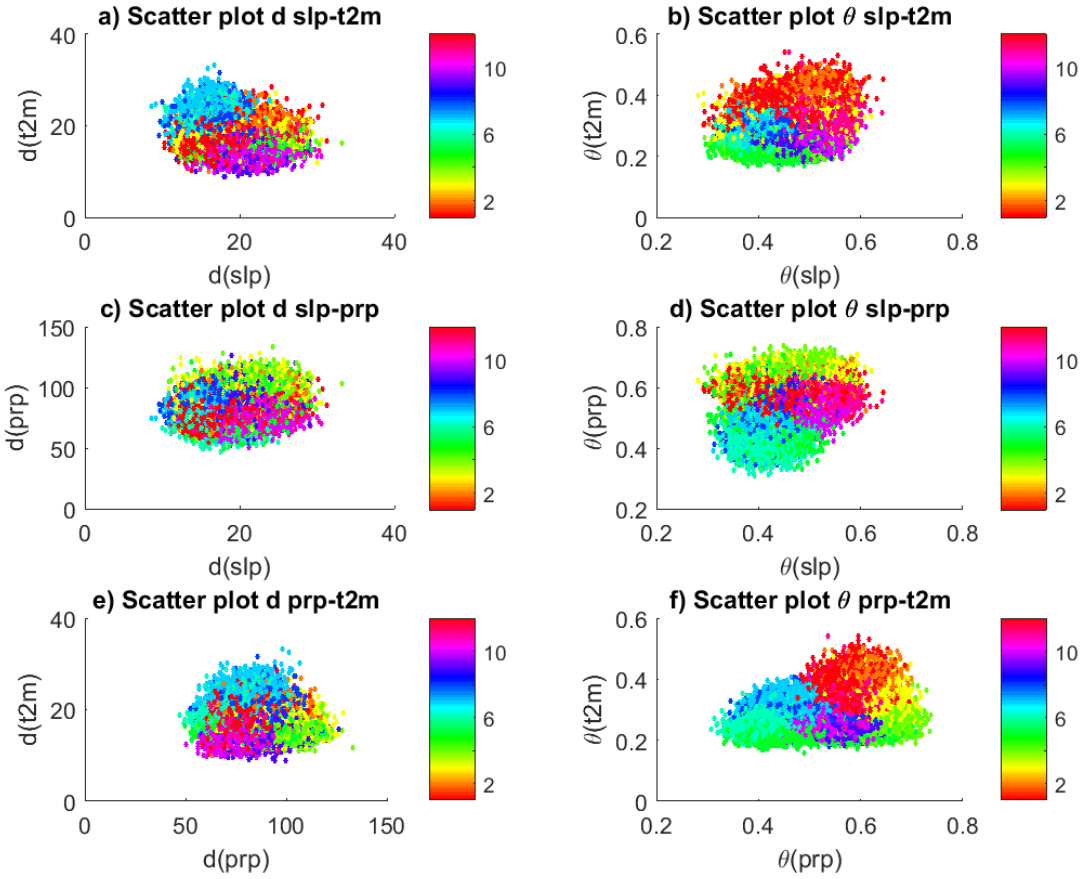

Figure 13. Scatter plots of the local dimensions d (a,c,e) and inverse persistences θ (b,d,f) of sea-level pressure (slp), 2 meter temperature (t2m) and precipitation frequency (prp) data