



HAL
open science

Diagnosing concurrent drivers of weather extremes: application to hot and cold days in North America

Davide Faranda, Gabriele Messori, Pascal Yiou

► **To cite this version:**

Davide Faranda, Gabriele Messori, Pascal Yiou. Diagnosing concurrent drivers of weather extremes: application to hot and cold days in North America. *Climate Dynamics*, 2020, 54, pp.2187-2201. 10.1007/s00382-019-05106-3 . hal-02095205v2

HAL Id: hal-02095205

<https://hal.science/hal-02095205v2>

Submitted on 7 Nov 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Diagnosing concurrent drivers of weather extremes: application to warm and cold days in North America

Davide Faranda · Gabriele Messori ·
Pascal Yiou

Received: date / Accepted: date

Abstract A fundamental challenge in climate science is decomposing the concurrent drivers of weather extremes in observations. Achieving this can provide insights into the drivers of individual extreme events as well as into possible future changes in extreme event frequencies under greenhouse forcing. In the present work, we exploit recent results from dynamical systems theory to study the co-variation and recurrence statistics of different atmospheric fields. Specifically, we present a methodology to quantify the recurrences of bivariate fields and the coupling between distinct univariate fields in terms of their joint recurrences. The coupling is defined by a parameter which varies according to the chosen fields, season, and domain and can be understood in terms of the underlying physics of the atmosphere. For suitably chosen fields, this approach enables to decompose the different drivers of weather extremes. Here, we compute the above metrics for near-surface temperature and sea level pressure, and use them to study warm or cold days over North America. We first identify states where temperature extremes are strongly or weakly coupled to the large-scale atmospheric circulation, and then elucidate the interplay between coupling and the occurrence of temperature extremes.

Keywords Weather extremes · Analogues · Climate Dynamics · Dynamical Systems Theory

D. Faranda

Laboratoire des Sciences du Climat et de l'Environnement LSCE-IPSL, CEA Saclay l'Orme des Merisiers, UMR 8212 CEA-CNRS-UVSQ, Université Paris-Saclay, 91191 Gif-sur-Yvette, France & London Mathematical Laboratory, 8 Margravine Gardens, London, W68RH, UK. E-mail: davide.faranda@lsce.ipsl.fr

G. Messori

Department of Earth Sciences, Uppsala University, Uppsala, Sweden & Department of Meteorology and Bolin Centre for Climate Research, Stockholm University, Stockholm, Sweden

P. Yiou

Laboratoire des Sciences du Climat et de l'Environnement LSCE-IPSL, CEA Saclay l'Orme des Merisiers, UMR 8212 CEA-CNRS-UVSQ, Université Paris-Saclay, 91191 Gif-sur-Yvette, France

1 Introduction

Weather extremes are complex events, driven by a variety of physical processes. Disentangling these processes is a fundamental challenge in climate science and is essential in order to reconstruct the chain leading to individual or co-occurring extremes (Davies, 2015; Messori et al, 2016; Harnik et al, 2016; Kornhuber et al, 2019). Early statistical studies of weather extremes adopted a univariate framework, projecting extreme events onto peak-over-threshold (Pickands III, 1975) or block-maxima (Gnedenko, 1943) of single observables (e.g. storm surges, wind speeds or temperatures (Galambos et al, 1994)). Multivariate approaches were introduced to take into account combinations of variables that achieve simultaneously large or small values (e.g. wind-speed and storm surge). Recently, the awareness has grown that individual variables may not be extreme themselves, but that their joint occurrence may yield an extremal behavior (see e.g. Bevacqua et al (2017)).

In order to understand the root drivers of large-scale weather extremes, one must therefore follow all the relevant variables in space and time over extended regions. For example, heatwaves or cold days reflect the interaction between persistent circulation regimes and temperature patterns. In some cases additional confounding factors, such as soil-moisture levels, further complicate the picture (Zscheischler and Seneviratne, 2017; Hirschi et al, 2011). Studying these interplays a priori necessitates a large number of numerical experiments with multi-parameter models.

In this paper, we aim to combine a multivariate view of extremes with analytical and computational efficiency, by constraining the analyses with the generic behavior of chaotic systems. We propose a new methodology, based on the adaptation and extension of recent mathematical results from dynamical systems theory, and apply it to the study of temperature extremes (warm or cold days) over North America. Our work builds upon the numerous studies that have tried to connect the dynamical properties of the atmosphere to the occurrence of temperature extremes in the region (e.g. Grotjahn (2016) and references therein, Messori et al (2016); Harnik et al (2016)), as well as the broader literature looking at mid-latitude extremes (e.g. Von Storch and Zwiers (2001); Coumou et al (2014); Palmer (2013); Gálfi et al (2017); Messori et al (2019)). These point to the degree of coupling between atmospheric circulation variables and temperature as a key step to both understand the physical drivers of the extremes and constrain the variability of climate models in order to better simulate such events in future climates. Here, we specifically diagnose the dynamical features and coupling of the sea-level pressure (SLP) and 2-m temperature (T2M) fields in reference to the above-mentioned North American temperature extremes. The purpose is to provide a proof-of-concept for the applicability of our novel analysis approach – which is entirely general and may in theory be extended to any number of variables – to the study of multi-variate atmospheric configurations.

We begin by providing a theoretical definition of three dynamical systems metrics: d , θ^{-1} and α . The local dimension d can be intuitively interpreted

73 as a proxy for the number of active degrees of freedom of the system around
 74 a given state. The persistence θ^{-1} measures the residence time around such
 75 state. Finally, the coupling parameter (also referred to as co-recurrence ratio)
 76 α informs on the dependence structure between instantaneous configurations
 77 of different variables. The underlying concept is akin to that of joint recurrence
 78 plots (Marwan et al, 2007). We next test these metrics on simple stochastic
 79 processes and discrete dynamical systems. Following this, we apply our frame-
 80 work to study the drivers underlying the occurrence of temperature extremes
 81 in North America.

82 **2 Dynamical systems indicators**

83 2.1 A dynamical systems framework

84 The attractor of a dynamical system is a geometrical object defined in the
 85 space hosting all the possible states of the system (phase space). For atmo-
 86 spheric flows, it is impossible to obtain all *variables* of the system, as this
 87 would require the knowledge of the properties of each fluid parcel. What is in-
 88 stead available is a set, or sequences, of *observables* – namely transformations
 89 of the variables of the system. By making some assumptions on the dimen-
 90 sionality of the system, it is possible to retrieve its entire dynamics from the
 91 observables (Huke, 2006). This allows the computation of the dimension of
 92 the system (Grassberger, 1983; Grassberger and Procaccia, 1984). The major
 93 caveat of such approach lies in the underlying mathematical assumptions, that
 94 render the application to geophysical datasets unreliable (Eckmann and Ruelle,
 95 1992). Recently, an alternative approach to determine a system’s dynamical
 96 properties has been proposed, which does not require a priori assumptions on
 97 the dimensionality of the system. This approach is based on an analysis of the
 98 recurrences of trajectories of the dynamical system, and has been successfully
 99 tested on various complex datasets (Faranda et al, 2017b,a; Messori et al, 2017;
 100 Faranda et al, 2019a). One of the outcomes is a characterization of each point
 101 ζ along the sequence of observables of the system by two dynamical indicators:
 102 the local dimension (d) and the persistence (θ^{-1}) (Faranda et al, 2017b). There
 103 have been theoretical (Barros et al, 2019) and experimental (Faranda et al,
 104 2017a,c) arguments that show that recurrences of observables are useful to
 105 investigate properties of the underlying phase space of atmospheric motions,
 106 even though all the variables defining the phase space of the system cannot be
 107 taken into account for obvious computational reasons.

108 Before detailing the computation of d and θ^{-1} , we give a physical interpre-
 109 tation to these seemingly abstract dynamical indicators. The local dimension
 110 d measures the geometry of the trajectories in a small region of the phase
 111 space. It is therefore linked to the number of degrees of freedom that a system
 112 can locally explore. In other terms, a region of the phase space with a low
 113 (high) dimension, allows a small (large) number of different configurations of
 114 the system. In Faranda et al (2017b) and Messori et al (2017), we showed

115 that lower dimensional states of the atmospheric circulation over the North
 116 Atlantic correspond to a predominantly zonal flow. Higher local dimensions
 117 are instead found for blocked flows. The persistence θ^{-1} informs on the per-
 118 sistence of the system in the same small region of the phase space. A higher
 119 (lower) persistence implies that trajectories will evolve more slowly (rapidly)
 120 in phase space. Hochman et al (2019) found that the dynamical systems per-
 121 sistence may be directly linked to the more conventional notion of persistence
 122 of weather regimes.

123 We consider a dynamical system with an observed trajectory $x(t)$, and a
 124 point ζ in phase space along this trajectory. Taking as example a daily SLP
 125 dataset, $x(t)$ would be a succession of daily SLP latitude-longitude maps, and
 126 ζ would be an SLP latitude-longitude map for a specific day (Fig. 1a). We are
 127 interested in the behavior of the system near ζ . To characterise this, we define
 128 a logarithmic return as:

$$g(x(t), \zeta) = -\log[\text{dist}(x(t), \zeta)]. \quad (1)$$

129 Where dist is a distance function between two vectors, which tends to zero
 130 when x and y are close to each other. The values of $g(x(t), \zeta)$ define a time series
 131 of logarithmic returns which takes large values when the distance between $x(t)$
 132 and ζ is small. Given $s(q, \zeta)$ a high q -th quantile of $g(x(t), \zeta)$, we introduce
 133 the exceedances $u(t, \zeta) = g(x(t), \zeta) - s(q, \zeta)$ with the condition $g(x(t), \zeta) >$
 134 $s(q, \zeta)$. The latter condition defines a recurrence, namely a case when the
 135 trajectory $x(t)$ approaches ζ (Fig. 1b). Going back to the SLP example, these
 136 would be days when the SLP map resembles closely that of the reference
 137 state ζ . The cumulative probability distribution of $F(u, \zeta)$ then converges to
 138 the exponential member of the Generalized Pareto Distribution (Freitas et al,
 139 2010; Lucarini et al, 2012):

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

140 The parameters ϑ and σ depend on the point ζ chosen on the attractor.
 141 This means that, for our series of daily SLP latitude-longitude maps, each
 142 day will have a different value of ϑ and σ . As discussed below, ϑ and σ are
 143 fundamental quantities to describe the dynamical properties of the system.
 144 When applying this approach to a given dataset, all the maps in the data
 145 series serve in turn as reference point ζ .

146 2.2 Local dimension

147 The local dimension $d(\zeta)$ is obtained via the simple relation $d(\zeta) = 1/\sigma(\zeta)$. A
 148 short MatLab code to compute d is provided in Appendix B of Faranda et al
 149 (2019b). This result holds when $x(t)$ contains the system's full set of phase-
 150 space variables. In this case, d is independent of the chosen dist for all ζ . If
 151 $x(t)$ is an *observable* of the system, i.e. some smooth function of a variable

152 in the full phase space, $d(\zeta)$ can still be computed in this way (Rousseau and
 153 Saussol, 2010; Rousseau, 2014), but the quantitative results depend on the
 154 distance definition.

155 For practical reasons, here we use the Euclidean distance, which is also used
 156 to compute circulation analogues (e.g. Yiou et al (2013)). We also emphasize
 157 that the value of d is bounded by the number of coordinates of the observable
 158 and the dynamics of the system.

159 Given two observables (or two sets of observables) $x(t)$ and $y(t)$ of a
 160 "larger" system, we can define $d_x(\zeta)$ and $d_y(\zeta)$ (from now on we will drop
 161 the dependence on ζ). They are the dimensions of the Poincaré sections defined
 162 by x and y around ζ , with respect to the chosen dist. This would be
 163 equivalent to, for example, computing daily values of d for SLP and T2M
 164 fields separately, over a given geographical domain. We can further consider
 165 the Poincaré section jointly spanned by x and y . A state on this section is
 166 defined by the pair $\zeta = \{\zeta_x, \zeta_y\}$. The joint logarithmic returns can then be
 167 defined as:

$$g(x(t), y(t)) = -\log \left[\text{dist} \left(\frac{x(t)}{\|x\|}, \frac{\zeta_x}{\|\zeta_x\|} \right)^2 + \text{dist} \left(\frac{y(t)}{\|y\|}, \frac{\zeta_y}{\|\zeta_y\|} \right)^2 \right]^{\frac{1}{2}} \quad (3)$$

168 Here, $\|\cdot\|$ is the average root mean square norm of the coordinates of a
 169 vector. For example, $\|x\| = E_t \left(\left[\sum_i^K x_i(t)^2 \right]^{\frac{1}{2}} \right)$, where K is the number of
 170 components of x and E_t is an average over time t . Based on Eq. (3), one
 171 can then compute $d_{x,y}$, namely the *co-dimension* between x and y . This is
 172 equivalent to considering the SLP and T2M fields jointly, and requiring that
 173 a recurrence – a day similar to the day of interest ζ – is such in both fields
 174 simultaneously.

175 The co-dimension can be used as a first quantity to characterise the mutual
 176 dependence of two observables. For two observables x and y of the system, the
 177 following properties hold:

$$\min(d_x, d_y) \leq d_{x,y} \leq d_x + d_y \quad (4)$$

178 where the min domain consists of the dimensions computed $\forall t$. The fact
 179 that the co-dimension is bounded by the sum of the individual dimensions
 180 suggests that it is a measure of the coupling of the two observables, namely
 181 the dimension required to describe their joint local behaviour. If x and y
 182 are uncoupled, then $d_{x,y} = d_x + d_y$. If x and y are linked with a deterministic
 183 function, then $d_{x,y} = \min(d_x, d_y)$. We exemplify these behaviours for a discrete
 184 dynamical system in Sec. 3.2.

185 2.3 Local persistence

186 The persistence of the state ζ is measured via the extremal index $0 < \vartheta(\zeta) < 1$,
 187 an adimensional parameter, from which we extract $\theta(\zeta) = \vartheta(\zeta)/\Delta t$. $\theta(\zeta)$ is

188 therefore the inverse of the average residence time of trajectories around ζ
 189 in units of frequency (in this study $\Delta t = 1$ day). If ζ is a fixed point of the
 190 attractor then $\theta(\zeta) = 0$ (trajectories stay at ζ). For a point that immediately
 191 leaves the neighborhood of ζ , then $\theta = 1$. Intermediate values of θ are obtained
 192 for regions of the attractor that are close to fixed points, and that trajectories
 193 leave "slowly". To estimate θ , we adopt the Suveges estimator (Suveges, 2007).
 194 Returning to our previous example, we consider all days on which the SLP
 195 map resembles the reference SLP state ζ , and we estimate the average number
 196 of consecutive days such a resemblance lasts for. Since we are investigating
 197 meteorological time scales, the choice of daily fields is appropriate to compute
 198 θ . Different applications might necessitate the use of a smaller (e.g. in the
 199 study of convective systems) or a larger (e.g. in the study of ocean dynamics)
 200 Δt .

201 As for d above, the procedure can be extended to more than one vari-
 202 able, such that one may define the inverse co-persistence $\theta_{x,y}$. The physical
 203 interpretation for atmospheric fields is analogous to that of the inverse per-
 204 sistence above, except that we now consider two fields jointly. Taking again
 205 SLP and T2M fields as examples, we thus consider series of consecutive days
 206 on which both resemble the respective reference states ζ_{SLP} and ζ_{T2M} . $\theta_{x,y}$
 207 is a weighted average of θ_x and θ_y , where the weights depend on the density of
 208 observed neighbours of ζ . For a more detailed discussion we refer the reader
 209 to Abadi et al (2018).

210 2.4 Local co-recurrence ratio

211 Given two observables x and y , we define the co-recurrence ratio $0 \leq \alpha(\zeta) \leq 1$
 212 of a state $\zeta = \{\zeta_x, \zeta_y\}$ as:

$$213 \alpha(\zeta) = \frac{\Pr [g(x(t)) > s_x(q) | g(y(t)) > s_y(q)]}{\Pr [g(x(t)) > s_x(q)]} \quad (5)$$

214 In terms of the SLP and T2M latitude-longitude fields, we compute the
 215 probability of having a day when the SLP map resembles that of the reference
 216 state ζ_{SLP} , given that the T2M map resembles that of the reference state ζ_{T2M} ,
 217 versus the probability of only the SLP map resembling the relevant reference
 218 state. The thresholds $s_x(q)$ and $s_y(q)$ determine how close the monovariate
 219 recurrences have to be. Whenever x and y do not have the same units, a
 220 normalization $x/||x||$ and $y/||y||$ must be performed before computing α . When
 221 $\alpha(\zeta) = 0$, there are no co-recurrences of $\zeta = \{\zeta_x, \zeta_y\}$ when we observe a
 222 recurrence of ζ_x . When $\alpha(\zeta) = 1$, all the co-recurrences of $\zeta = \{\zeta_x, \zeta_y\}$ also
 223 correspond to recurrences of ζ_x .

224 From Bayes' theorem on conditional probabilities, we note that α does not
 225 depend on the order of x and y , which can hence be exchanged. This implies
 226 that α cannot be interpreted in terms of causation. We further note that α
 227 is not necessarily correlated with the co-dimension $d_{x,y}$, as discussed further
 in Sect 3.2. Finally, we underscore that α cannot be directly compared to

228 statistical dependence measures because of its local nature in phase space.
 229 Here, we have defined α in the case of two variables x and y , but the approach
 230 can be extended to more complex multivariate cases.

231 3 Application to stochastic processes and dynamical systems

232 3.1 Gaussian Bivariate copula

233 To illustrate the interpretation of α , we draw data from random Gaussian
 234 bivariate copula distributions, namely:

$$\phi(z, \mu, \Sigma) = \frac{1}{\sqrt{4\pi^2|\Sigma|}} \exp \left[-\frac{1}{2}(z - \mu)' \Sigma^{-1} (z - \mu) \right] \quad (6)$$

235 where $z = (x, y)$ is an independent and identically distributed (i.i.d.) vari-
 236 able vector in $[0, 1] \times [0, 1]$, μ is the mean vector, Σ is a 2×2 symmetric matrix
 237 and $|\Sigma|$ is its norm. The diagonal elements of Σ contain the variances for each
 238 variable (x, y) , while the off-diagonal elements of Σ contain the covariances
 239 between variables (x, y) . In the example we set $\mu = \{0, 0\}$ and the diagonal
 240 elements of Σ equal to one. We term the non-diagonal elements ρ .

241 By varying ρ , we obtain a range of different behaviours. There are two
 242 limiting cases: for $\rho = 0$, $\{x, y\}$ the data cover a large portion of the unit
 243 square; for $\rho = 1$, $x = y$ and the data are aligned along the first diagonal. We
 244 illustrate graphically four different cases: $\rho = 0.5$, $\rho = 0.75$, $\rho = 0.9992$, and
 245 $\rho = 1$. We generate 10^5 values of z for each case. For clarity we report 2×10^3
 246 points in Figure 2. These are colored by the values of α , the co-recurrence
 247 ratio, computed by fixing the quantile $q = 0.98$.

248 α can be interpreted as follows: when the variables are locally independent
 249 (Figure 2a, b), $\alpha = 0$ almost everywhere. When ρ is close to 1 (Figure 2c),
 250 near the edges $(0, 0)$ and $(1, 1)$ the dependence is very strong and $\alpha=1$. Near
 251 the centre $(x = y = 0.5)$ the dependence is lower and $\alpha < 1$. This behavior
 252 is determined by the very nature of the distribution considered here, which is
 253 constrained to be close to the diagonal near the edge points. In the perfect
 254 coupling case $\rho = 1$ (Figure 2d), $x = y$ and $\alpha = 1$ along the full length of the
 255 diagonal.

256 3.2 Baker's map

257 To further illustrate the properties of the dynamical indicators, we analyse
 258 a modified version of baker's map, defined on the unit square $[0, 1] \times [0, 1]$
 259 iteratively for $n \geq 1$, by:

$$x_{n+1} = f(x_n) = \begin{cases} ax_n, & \text{for } y_n < c \\ \frac{1}{2} + bx_n, & \text{for } y_n > c \end{cases} \quad (7)$$

260 and

$$y_{n+1} = f(y_n) = \begin{cases} y_n/a, & \text{for } y_n < c \\ (y_n - c)/(1 - c), & \text{for } y_n > c \end{cases} \quad (8)$$

261 with $c = 1/3$, $a = 1/5$ and $b = 1/4$. For these parameter values, the map is
 262 expanding on x and contracting on y . Figure 3a shows the first 10^5 iterations
 263 of the map. Along the x direction, the map consists of a set with dimension
 264 $d_x < 1$ while along y there is a collection of lines yielding a dimension $d_y = 1$.

265 We can add coupling between x and y as:

$$x_{n+1} = (1 - \epsilon)f(x_n) + \epsilon f(y_n)$$

266 with $0 \leq \epsilon \leq 1$. When ϵ is zero, x and y are uncoupled. As the noise inten-
 267 sity ϵ increases, the dynamics of x, y becomes more and more synchronized.
 268 Finally, when $\epsilon = 1$, the dynamics is synchronized and concentrated on the
 269 diagonals (Figure 3b).

270 To illustrate the properties of the co-dimension and the co-recurrence coef-
 271 ficients in this set-up, we change the values of the noise intensity ϵ in the range
 272 $10^{-10} < \epsilon < 1$ and perform 10^6 iterations of the maps starting from random
 273 initial conditions on the unit square. The first 10^4 iterations are discarded.
 274 We then compute the dynamical indicators at 500 points ζ . Figure 4 shows
 275 the results of this computation for the dimensions d_x , d_y , the co-dimension
 276 $d_{x,y}$ and the co-recurrence coefficient α . In the limit $\epsilon \rightarrow 0$, $d_x + d_y = d_{x,y}$
 277 and $\alpha \simeq 0$. When increasing the coupling, $d_{x,y} < d_x + d_y$. In the limit for
 278 large noise intensity $\epsilon = 1$, $x \propto y$, $d_{x,y} = 1$ and the co-recurrence coefficient
 279 α approaches 1. The relatively sharp transition happening for noise intensity
 280 $\epsilon \simeq 10^{-3}$ can be explained by drawing from the behavior of stochastically
 281 perturbed systems (Faranda et al, 2013). In the latter, when noise spanning
 282 several orders of magnitude is added, it has an appreciable effect only when
 283 stochastic fluctuations (of order ϵ) have the same order as those induced by
 284 the deterministic dynamics. If the noise fluctuations are larger than those of
 285 the deterministic system, the dimension d reflects the stochastic system.

286 This simple example highlights the difference in the information provided
 287 by the co-dimension and co-recurrence ratio. The co-dimension generally de-
 288 creases for increasing coupling, while the co-recurrence ratio increases. We
 289 further note that the co-recurrence ratio provides a normalised coupling mea-
 290 sure, while the value of the co-dimension depends on the relative dimension
 291 on the x and y manifold. The behavior for θ is not shown as, for this system,
 292 we have trivially that $\theta_x = \theta_y = \theta_{x,y} \forall \epsilon$.

293 4 Application to North American temperature extremes

294 4.1 Data and Methods

295 We base our study on NCEP/NCAR reanalysis data (Kalnay et al, 1996) over
 296 the period 1948-2018, with a horizontal resolution of 2.5° . We consider a do-
 297 main spanning North America ($170^\circ\text{W} \leq \text{Long.} \leq 40^\circ\text{W}$, $22.5^\circ\text{N} \leq \text{Lat.} \leq 70^\circ\text{N}$).

298 We adopt daily SLP as the meteorological variable to describe the large scale
 299 atmospheric circulation. A wealth of atmospheric features, ranging from tele-
 300 connection patterns to storm track activity to atmospheric blocking can be
 301 diagnosed from the SLP field (e.g. Murray and Simmonds (1991); Yiou et al
 302 (2013); Comas-Bru and McDermott (2014)). We diagnose temperature ex-
 303 tremes using T2M. Throughout the analysis, we consider winter (December-
 304 January-February: DJF) and summer (June-July-August: JJA) seasons sepa-
 305 rately.

306 We define warm or cold extremes following the procedure of Messori et al
 307 (2018). This amounts to using area-averaged temperature anomalies relative to
 308 a running daily climatology and applying some smoothing. Here, we consider
 309 the domain $100^{\circ}\text{W} \leq \text{Long.} \leq 70^{\circ}\text{W}$, and $30^{\circ}\text{N} \leq \text{Lat.} \leq 45^{\circ}\text{N}$. This corresponds
 310 to a densely populated part of western North America (see Figure 5). We then
 311 define two anomaly distributions for each season: one for all days with negative
 312 area-averaged temperature anomalies and one for all days with positive area-
 313 averaged temperature anomalies. The percentiles discussed in the text below
 314 are relative to these distributions. We specifically define cold days as days
 315 below the 10th percentile of the negative-anomalies distributions and warm
 316 days as days above the 90th percentile of the positive-anomalies distributions.
 317 We refer to these throughout the text as *warm days* or *warm extremes* and *cold*
 318 *days* or *cold extremes*, respectively. To avoid double-counting a single warm
 319 or cold event, following Messori et al (2016) we impose a minimum separation
 320 between successive warm or cold days – here set at one week. All geographical
 321 plots show single-gridpoint T2M anomalies defined relative to a daily seasonal
 322 cycle, obtained by averaging all days available in the dataset. For example, the
 323 climatological value at a given gridbox for the 3rd December is the average
 324 value at that gridbox for all 68 3rd Decembers in the data.

325 4.2 Dynamical properties and seasonality

326 We analyse the relations between local dimension d , persistence θ^{-1} and co-
 327 recurrence ratio α for SLP and T2M. We consider the cases where d and θ are
 328 computed on each variable individually and on both variables at the same time.
 329 α is naturally always computed on both variables. The results are reported
 330 in Figure 6 in the form of dimension-persistence diagrams. Each point in the
 331 diagram represents a pair (d, θ) corresponding to the patterns observed on a
 332 given day for the variable(s) of interest. The colours show the values of α .

333 α is negatively correlated with d_{SLP} (Figure 6a), meaning that configu-
 334 rations with lower d typically favour a higher coupling. The fact that d and
 335 θ are themselves correlated implies that the latter configurations are, on av-
 336 erage, also highly persistent. The picture for T2M is less clear (Figure 6b):
 337 strongly coupled states often correspond to high d and low-persistence con-
 338 figurations. Finally, the joint analysis highlights a similar pattern as for SLP,
 339 with α showing a clear dependence on the co-dimension $d_{SLP, T2M}$.

340 Faranda et al (2017a) and Rodrigues et al (2018) found that the correla-
 341 tions evident in dimension-persistence diagrams, as those shown in Figure 6,
 342 reflected the strong seasonal dependence of the metrics. We therefore investi-
 343 gate the seasonal cycle in the above quantities (Figure 7). All three indicators
 344 are subject to a marked seasonality, with α presenting a clear peak in sum-
 345 mertime, when mean values are almost twice those of the rest of the year. The
 346 local dimension d instead displays maxima in the shoulder seasons for SLP,
 347 and in summer for T2M. The broad summertime maximum in d_{T2M} closely
 348 matches the timing of the α peak; d_{SLP} displays a minimum at roughly the
 349 same time. This explains the very different relations between d and α shown
 350 in Figures 6a and b for SLP and T2M, respectively. The variability in local
 351 dimensions is not, however, as marked as that seen in α . The co-dimension
 352 $d_{SLP,T2M}$ does not follow the seasonal cycle of either d_{SLP} or d_{T2M} . It peaks
 353 in late spring/early summer and early autumn, while displaying mid-summer
 354 and winter minima. θ_{SLP} is mostly in phase with d_{SLP} , as is θ_{T2M} with d_{T2M} ,
 355 albeit with a larger seasonal variability. The inverse co-persistence $\theta_{SLP,T2M}$
 356 again does not follow either of the single-variable cycles, but rather displays
 357 an oscillatory behavior throughout the year, somewhat reminiscent of the vari-
 358 ability in $d_{SLP,T2M}$. The largest values tend to occur during the autumn and
 359 winter months, while the lowest values are mostly found in spring.

360 The seasons during which most of the highest and lowest values of the three
 361 indicators occur reflect these seasonal cycles (Figure 8). Even though the mean
 362 of α has a very clear summertime peak, the variability peaks instead in winter
 363 (Figure 7a), implying that a small number of α maxima are seen during this
 364 season. This matches the timing of minima in both d_{SLP} and θ_{SLP} , consistent
 365 with what shown in Figure 6a. The α minima occur instead predominantly
 366 in winter and spring, when d_{SLP} and θ_{SLP} display relatively large values.
 367 Analogous conclusions can be drawn concerning the seasonal cycles of the
 368 co-dimension and co-persistence (Figure 8).

369 4.3 Temperature extremes

370 The above results show that d , θ and α display both a large intra-seasonal
 371 variability and a strong seasonal cycle. The most striking features are the dis-
 372 tinct wintertime and summertime behaviours in α . We ascribe the seasonality
 373 of the three metrics to the fact that they reflect the very different atmospheric
 374 dynamics found in the different seasons (see Faranda et al (2017b,a)). Even
 375 though the mean values of the metrics and an analysis of their seasonal variabil-
 376 ity can provide interesting physical insights, such as assessing which seasons
 377 present more or less coupling between SLP and T2M, the real added value in
 378 having instantaneous metrics is the possibility of using them to study climate
 379 extremes. Here, we specifically investigate warm and cold extremes over North
 380 America, during both the summer (JJA) and winter (DJF) seasons.

381 *Summer extremes.* We begin by considering summertime warm and cold days
 382 (see Section 4.1). The co-dimension and co-persistence, as a function of the
 383 temperature percentile over the selection domain, show little change for ex-
 384 treme warm and extreme cold events relative to their respective seasonal cycles
 385 (continuous lines in Figure 9a,b). Similarly, α (Figure 9c) remains close to its
 386 seasonal cycle, regardless of the temperature percentile. This points to the fact
 387 that the joint d and θ metrics do not reflect the evolution of the temperature
 388 field and that, consistently with this, the co-recurrence ratio is also largely
 389 insensitive to temperature extremes. The d and θ metrics computed for the
 390 SLP and T2M fields individually (dashed and dashed-dotted lines in Figure
 391 9a,b) display a similar behaviour.

392 To highlight the link between the dynamical systems metrics and the un-
 393 derlying physical properties of the atmospheric circulation associated with
 394 different events, we define an average daily baroclinic vector as:

$$B = \overline{\nabla(SLP) \times \nabla(T2M)} \quad (9)$$

395 where \times denotes the vector product and the overbar a spatial mean over the
 396 domain highlighted in black in Figure 5. B is in units of $[\text{hPa} \times \text{K}] / [^\circ \text{lon} \times ^\circ \text{lat}]$.
 397 Values close to zero indicate a predominantly barotropic atmosphere i.e. the
 398 isolines of pressure are aligned with those of temperature. Positive and neg-
 399 ative values correspond instead to warm and cold air advection, respectively.
 400 Two caveats are that we do not account for cancellation between opposing
 401 advectons in the domain, and may give excessive weight to zonal advection
 402 at northerly latitudes. Domain-averaged values for B are shown in Figure 9d.
 403 In summer, all the values of B are generally close to zero. This decoupling
 404 of the temperature anomalies from specific large-scale advection patterns is
 405 also reflected in the T2M and SLP composites for α extremes. Regardless
 406 of whether coupling maxima or minima are chosen, the anomalies are weak
 407 throughout the domain, and the sign agreement between the different extreme
 408 α occurrences is low (Figure 10).

409 *Winter extremes.* The situation in winter is radically different: $d_{SLP,T2M}$,
 410 $\theta_{SLP,T2M}$ and α for the cold extremes all show significant deviations from
 411 the climatology. The warm extremes show weak deviations in $d_{SLP,T2M}$ and
 412 $\theta_{SLP,T2M}$, and a moderate deviation in α . The signs of the d and θ deviations
 413 for the two sets of temperature extremes are opposite: cold extremes show
 414 negative anomalies, while the warm extremes show weak positive anomalies
 415 (continuous lines in Figure 11a-c). The d and θ computed for the two vari-
 416 ables individually show a similar behaviour to the bivariate indicators for the
 417 cold extremes (blue dashed and dashed-dotted lines in Figure 11a, b). For the
 418 warm extremes, d_{SLP} and θ_{SLP} are broadly consistent with $d_{SLP,T2M}$ and
 419 $\theta_{SLP,T2M}$ (red dashed lines in Figure 11a, b), while d_{T2M} and θ_{T2M} show
 420 anomalies of a different sign to those of the bivariate indicators (red dash-
 421 dotted lines in Figure 11a, b). However, they still show a clear deviation from
 422 the climatology. The co-recurrence ratio also shows deviations of the opposite

423 sign for the two sets of extremes: negative for warm extremes and positive for
424 cold extremes (Figure 11c). The analysis of the baroclinic vector B (Figure
425 11d) provides clear insights on such asymmetry: cold anomalies systematically
426 correspond to a strong cold advection ($B < 0$), reflecting the high value of
427 the co-recurrence ratio α , while the warm extremes display a significant warm
428 advection ($B > 0$) only for the highest percentiles.

429 By conditioning purely on low and high values of α , one recovers strong
430 temperature anomalies over the domain of interest (Figure 12b, d). This is
431 particularly evident for the high α , which correspond to cold anomalies across
432 a broad swath of the North American continent. Both low and high α are
433 associated with anomalous SLP patterns favoring strong meridional advection
434 (Figure 12a, c). The SLP and T2M fields therefore appear to display recurring
435 joint large-scale configurations which favour cold extremes and, to a lesser
436 extent, warm extremes. This may seem in contradiction with the information
437 provided by α , as one may expect no coherent large scale SLP patterns to
438 be associated with α minima (and hence warm extremes). We expand on this
439 aspect in the discussion section below.

440 5 Discussion and conclusions

441 We have proposed two metrics to diagnose the properties of instantaneous
442 configurations of atmospheric variables, namely local dimension d and inverse
443 persistence θ . The first is a proxy for the number of active degrees of freedom
444 of a given atmospheric configuration, while the second measures the average
445 time over which a configuration is maintained. These indicators have been pre-
446 viously applied to individual atmospheric variables (e.g. Faranda et al (2016,
447 2017b,a, 2019c); Messori et al (2017); Rodrigues et al (2018); Scher and Messori
448 (2018, 2019); Hochman et al (2019)). Here, we present their use in a multi-
449 variate context for both simple dynamical systems and climate reanalysis data.
450 We further introduce a novel metric which measures the co-occurrence of anal-
451 ogue states of two (or more) variables: the co-recurrence ratio α . A small
452 α indicates rare co-recurrences, while a large α indicates that recurrences of
453 a given state in one of the two variables systematically match recurrences
454 of a corresponding state in the other. This metric thus elucidates the mutual
455 dependence between the dynamical evolutions of the two variables. Unlike sta-
456 tistical dependence measures, α is grounded in the topology of the variables'
457 phase-space, and therefore in the local geometry of the underlying attrac-
458 tor. Locality in phase space, which translates to simultaneity in time, makes
459 the above framework well-suited to the study of atmospheric variability and
460 extremes. From the numerical analysis performed in this work on simple dy-
461 namical systems and climate data, we infer that the co-recurrence ratio is an
462 indicator of physical coupling and underscores interactions between variables.
463 This understanding comes from the observation that α relates non-trivially
464 to the co-dimension (cf. the baker's map and atmospheric data). The results
465 obtained for atmospheric data suggest that, when α is high, the phase space

466 around ζ_x and ζ_y shrinks yielding lower co-dimensions. In other words, when
 467 α is high the variables are dynamically dependent. We expect this behaviour
 468 to be representative of high-dimensional systems.

469 As an illustration on atmospheric data, we have investigated the occurrence
 470 of cold or warm days in North America, linking temperature anomalies over
 471 the eastern part of the continent to the larger-scale T2M and SLP fields. We
 472 identify strongly and weakly coupled states in these two variables, and then
 473 elucidate the interplay between these and the occurrence of the temperature
 474 extremes.

475 Summertime temperature extremes appear to be unrelated to the joint dy-
 476 namical properties of the T2M and SLP large-scale fields. This is also reflected
 477 in the d and θ metrics computed on the individual variables and in the weak
 478 co-recurrence between SLP and T2M. The recurrences of T2M fields associ-
 479 ated with temperature extremes are not matched by recurrences of the same
 480 SLP fields, or at least not more so than what is typical for any day in the
 481 summer season. Indeed, the sign agreement of the large-scale SLP anomalies
 482 associated with the warm and cold temperature extremes are relatively weak
 483 and display very limited sign agreement (see Fig. 13 for the upper/lower 10
 484 percentiles). Similarly, a proxy for temperature advection shows near-zero val-
 485 ues (B in Fig. 9d). This points to local factors, not captured by the large-scale
 486 SLP field analysed here, as contributors to these events (e.g. local soil moisture
 487 anomalies, small-scale precipitation events, local sensible heat fluxes etc.).

488 In winter, the picture is different (see Fig. 14 for the SLP anomalies corre-
 489 sponding to the upper/lower 10 percentiles of T2M anomalies). Both the joint
 490 and univariate d and θ discriminate warm and cold extremes, and these also
 491 display anomalous α values. Cold extremes are characterised by an anoma-
 492 lously high persistence, low local dimension and strong coupling. This points
 493 to the need for persistent circulation patterns and large-scale cold advection
 494 (Fig. 14c) for a cold day to occur, as well as to the fact that whenever cold
 495 days occur, similar large-scale T2M and SLP patterns are found (Cellitti et al,
 496 2006; Grotjahn, 2016; Walsh et al, 2001; Messori et al, 2016). Here we identify
 497 these as: a large anticyclone over North America, advecting cold, dry air from
 498 the Arctic region (as reflected by the baroclinic vector B in Fig. 11d), and two
 499 low-pressure cores on either flank (Fig. 14c). Such pattern leads to widespread
 500 low temperatures as far south as Texas and Northern Florida (Fig. 14d). Warm
 501 extremes are instead characterised by weak negative anomalies in persistence,
 502 weak positive anomalies in local dimension and a below-climatology coupling.
 503 The large-scale pattern is roughly inverse to that seen for the cold days, with
 504 a cyclonic anomaly over the continent favouring advection of warm, moist air
 505 from the low latitudes (Figs. 11d and 14a) and leading to high temperatures
 506 on the eastern part of the North American continent (Fig. 14b). Wintertime
 507 warm days are thus associated with a clear large-scale SLP pattern, yet the
 508 SLP and T2M fields do not necessarily display frequent joint recurrences, this
 509 might mean that warm winter days are not dynamically different from the
 510 other “average” days of the season, contrary to cold episodes.

511 There is an evident ambiguity in an apparently coherent large scale anomalous SLP pattern and the claim of a weak coupling to the T2M field. A possible explanation is that this issues from the dynamical systems analysis being performed on the absolute fields rather than on the anomalies shown in the figures. Since the co-recurrence ratio is sensitive to the seasonal cycle (Figure 7a), this may introduce a discrepancy between absolute and anomalous recurrence patterns. This points to the need to perform a systematic analysis on the anomaly fields, although this is far from trivial in a dynamical systems context. There is no currently accepted framework for removing the average attractor components and then characterising the residual fluctuations. Formally, this would correspond to analysing recurrences within hyper-ellipsoids whose the eccentricity depends on the position in phase space, rather than hyper-spheres.

524 Our analysis has shown insights that can be provided by our local (in phase-space, instantaneous in time) and multi-variate dynamical systems analysis as well as some ongoing challenges. As such, it should be viewed as a proof of concept for the utility of this approach. More generally, there is no requirement for the different variables to be geographically co-located or temporally coincident, as was the case here. For example, one could select SLP over a region upstream of the target temperature region, or could use SLP fields lagged by a few days relative to the T2M data. This would enable to study forcing instead of coupling between variables. The approach may also be used to study changes in the coupling and the interaction of variables under anthropogenic forcing, in a similar fashion as Faranda et al (2019a).

535 In this study we have considered bivariate applications and climate data on latitude-longitude grids. However the dynamical systems indicators can be generalised to higher dimensional variables or to a larger number of variables. We note, however, that in higher dimensional space we lose the x, y symmetry in the definition of the co-recurrence ratio, which then needs to be defined with respect to only one of the variables of the system. Our approach is thus flexible, and is applicable to observables of any complex system.

542 **Acknowledgements** DF, PY and GM acknowledge the support of the ERC grant No. 338965-A2C2. GM was partly supported by the Swedish Research Council Vetenskapsrådet No. 2016-03724. DF and PY were further supported by a CNRS-INSU LEFE/MANU grant (DINCLIC project).

546 References

- 547 Abadi M, Freitas ACM, Freitas JM (2018) Dynamical counterexamples regarding the extremal index and the mean of the limiting cluster size distribution. arXiv preprint arXiv:180802970
- 550 Barros V, Liao L, Rousseau J (2019) On the shortest distance between orbits and the longest common substring problem. *Advances in Mathematics* 344:311–339

- 553 Bevacqua E, Maraun D, Hobæk Haff I, Widmann M, Vrac M (2017) Multivari-
554 ate statistical modelling of compound events via pair-copula constructions:
555 analysis of floods in ravenna (italy). *Hydrology and Earth System Sciences*
556 21(6):2701–2723
- 557 Cellitti MP, Walsh JE, Rauber RM, Portis DH (2006) Extreme cold air out-
558 breaks over the united states, the polar vortex, and the large-scale circula-
559 tion. *Journal of Geophysical Research: Atmospheres* 111(D2)
- 560 Comas-Bru L, McDermott F (2014) Impacts of the ea and sca patterns on
561 the european twentieth century nao–winter climate relationship. *Quarterly*
562 *Journal of the Royal Meteorological Society* 140(679):354–363
- 563 Coumou D, Petoukhov V, Rahmstorf S, Petri S, Schellnhuber HJ (2014) Quasi-
564 resonant circulation regimes and hemispheric synchronization of extreme
565 weather in boreal summer. *Proceedings of the National Academy of Sciences*
566 111(34):12,331–12,336
- 567 Davies HC (2015) Weather chains during the 2013/2014 winter and their sig-
568 nificance for seasonal prediction. *Nature Geoscience* 8(11):833
- 569 Eckmann JP, Ruelle D (1992) Fundamental limitations for estimating dimen-
570 sions and lyapunov exponents in dynamical systems. *Physica D: Nonlinear*
571 *Phenomena* 56(2-3):185–187
- 572 Faranda D, Freitas JM, Lucarini V, Turchetti G, Vaienti S (2013) Extreme
573 value statistics for dynamical systems with noise. *Nonlinearity* 26(9):2597
- 574 Faranda D, Masato G, Moloney N, Sato Y, Daviaud F, Dubrulle B, Yiou
575 P (2016) The switching between zonal and blocked mid-latitude atmo-
576 spheric circulation: a dynamical system perspective. *Climate Dynamics*
577 47(5-6):1587–1599
- 578 Faranda D, Messori G, Alvarez-Castro MC, Yiou P (2017a) Dynamical prop-
579 erties and extremes of northern hemisphere climate fields over the past 60
580 years. *Nonlinear Processes in Geophysics* 24(4):713
- 581 Faranda D, Messori G, Yiou P (2017b) Dynamical proxies of north atlantic
582 predictability and extremes. *Scientific reports* 7:41,278
- 583 Faranda D, Sato Y, Saint-Michel B, Wiertel C, Padilla V, Dubrulle B, Daviaud
584 F (2017c) Stochastic chaos in a turbulent swirling flow. *Physical review*
585 *letters* 119(1):014,502
- 586 Faranda D, Alvarez-Castro MC, Messori G, Rodrigues D, Yiou P (2019a)
587 The hammam effect or how a warm ocean enhances large scale atmospheric
588 predictability. *Nature communications* 10(1):1316
- 589 Faranda D, Messori G, Vannitsem S (2019b) Attractor dimension of time-
590 averaged climate observables: insights from a low-order ocean-atmosphere
591 model. *Tellus A: Dynamic Meteorology and Oceanography* 71(1):1–11
- 592 Faranda D, Sato Y, Messori G, Moloney NR, Yiou P (2019c) Minimal dynam-
593 ical systems model of the northern hemisphere jet stream via embedding of
594 climate data. *Earth System Dynamics* 10(3):555–567
- 595 Freitas ACM, Freitas JM, Todd M (2010) Hitting time statistics and extreme
596 value theory. *Probab Theor Rel* 147(3-4):675–710
- 597 Galambos J, Lechner J, Simiu E, Hagwood C (1994) Extreme value theory for
598 applications:. NIST technical publications

- 599 Gálfi VM, Bódai T, Lucarini V (2017) Convergence of extreme value statistics
600 in a two-layer quasi-geostrophic atmospheric model. *Complexity* 2017
- 601 Gnedenko B (1943) Sur la distribution limite du terme maximum d'une serie
602 aleatoire. *Annals of mathematics* pp 423–453
- 603 Grassberger P (1983) Generalized dimensions of strange attractors. *Physics*
604 *Letters A* 97(6):227–230
- 605 Grassberger P, Procaccia I (1984) Dimensions and entropies of strange attrac-
606 tors from a fluctuating dynamics approach. *Physica D: Nonlinear Phenom-
607 ena* 13(1-2):34–54
- 608 Grotjahn R (2016) Western north american extreme heat, associated large
609 scale synoptic-dynamics, and performance by a climate model. *Dynamics
610 and Predictability of Large-Scale, High-Impact Weather and Climate Events*
611 2:198
- 612 Harnik N, Messori G, Caballero R, Feldstein SB (2016) The circumglobal
613 north american wave pattern and its relation to cold events in eastern north
614 america. *Geophysical Research Letters* 43(20)
- 615 Hirschi M, Seneviratne SI, Alexandrov V, Boberg F, Boroneant C, Christensen
616 OB, Formayer H, Orłowsky B, Stepanek P (2011) Observational evidence
617 for soil-moisture impact on hot extremes in southeastern europe. *Nature
618 Geoscience* 4(1):17
- 619 Hochman A, Alpert P, Harpaz T, Saaroni H, Messori G (2019) A new dy-
620 namical systems perspective on atmospheric predictability: Eastern mediter-
621 ranean weather regimes as a case study. *Science advances* 5(6):eaau0936
- 622 Huke J (2006) Embedding nonlinear dynamical systems: A guide to takens'
623 theorem
- 624 Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M,
625 Saha S, White G, Woollen J, et al (1996) The ncep/ncar 40-year reanalysis
626 project. *B Am Meteorol Soc* 77(3):437–471
- 627 Kornhuber K, Osprey S, Coumou D, Petri S, Petoukhov V, Rahmstorf S,
628 Gray L (2019) Extreme weather events in early summer 2018 connected
629 by a recurrent hemispheric wave-7 pattern. *Environmental Research Letters*
630 14(5):054,002
- 631 Lucarini V, Faranda D, Wouters J (2012) Universal behaviour of extreme value
632 statistics for selected observables of dynamical systems. *Journal of statistical
633 physics* 147(1):63–73
- 634 Marwan N, Romano MC, Thiel M, Kurths J (2007) Recurrence plots for the
635 analysis of complex systems. *Physics reports* 438(5-6):237–329
- 636 Messori G, Caballero R, Gaetani M (2016) On cold spells in north america and
637 storminess in western europe. *Geophysical Research Letters* 43(12):6620–
638 6628
- 639 Messori G, Caballero R, Faranda D (2017) A dynamical systems approach
640 to studying midlatitude weather extremes. *Geophysical Research Letters*
641 44(7):3346–3354
- 642 Messori G, Woods C, Caballero R (2018) On the drivers of wintertime tem-
643 perature extremes in the high arctic. *Journal of Climate* 31(4):1597–1618

- 644 Messori G, Davini P, Alvarez-Castro MC, Pausata FS, Yiou P, Caballero R
645 (2019) On the low-frequency variability of wintertime euro-atlantic plane-
646 tary wave-breaking. *Climate Dynamics* 52(3-4):2431–2450
- 647 Murray RJ, Simmonds I (1991) A numerical scheme for tracking cyclone centres
648 from digital data. *Australian Meteorological Magazine* 39(3):155–166
- 649 Palmer T (2013) Climate extremes and the role of dynamics. *Proceedings of*
650 *the National Academy of Sciences* 110(14):5281–5282
- 651 Pickands III J (1975) Statistical inference using extreme order statistics. *the*
652 *Annals of Statistics* pp 119–131
- 653 Rodrigues D, Alvarez-Castro MC, Messori G, Yiou P, Robin Y, Faranda D
654 (2018) Dynamical properties of the north atlantic atmospheric circulation
655 in the past 150 years in cmip5 models and the 20crv2c reanalysis. *Journal*
656 *of Climate* (2018)
- 657 Rousseau J (2014) Hitting time statistics for observations of dynamical systems.
658 *Nonlinearity* 27(9):2377
- 659 Rousseau J, Saussol B (2010) Poincaré recurrence for observations. *Transactions*
660 *of the American Mathematical Society* 362(11):5845–5859
- 661 Scher S, Messori G (2018) Predicting weather forecast uncertainty with
662 machine learning. *Quarterly Journal of the Royal Meteorological Society*
663 144(717):2830–2841
- 664 Scher S, Messori G (2019) Weather and climate forecasting with neural networks:
665 using general circulation models (gcms) with different complexity as
666 a study ground. *Geoscientific Model Development* 12(7):2797–2809
- 667 Süveges M (2007) Likelihood estimation of the extremal index. *Extremes* 10(1-
668 2):41–55
- 669 Von Storch H, Zwiers FW (2001) *Statistical analysis in climate research*. Cambridge
670 university press
- 671 Walsh JE, Phillips AS, Portis DH, Chapman WL (2001) Extreme cold
672 outbreaks in the united states and europe, 1948–99. *Journal of climate*
673 14(12):2642–2658
- 674 Yiou P, Salameh T, Drobinski P, Menut L, Vautard R, Vrac M (2013) Ensemble
675 reconstruction of the atmospheric column from surface pressure using
676 analogues. *Climate dynamics* 41(5-6):1333–1344
- 677 Zscheischler J, Seneviratne SI (2017) Dependence of drivers affects risks associated
678 with compound events. *Science Advances* 3(6):e1700263

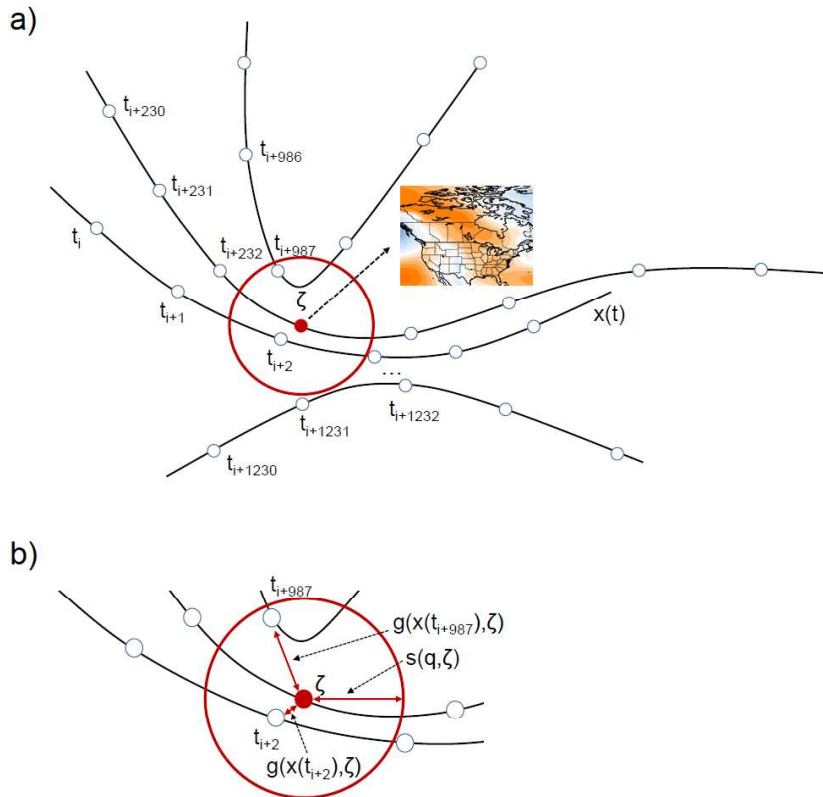


Fig. 1 (a) Idealised schematic of a long trajectory $x(t)$ representing the evolution of a sea-level pressure field. Note that all segments shown in the panel are part of a single, long trajectory. The white circles along the trajectory represent discrete, instantaneous measurements of the continuous evolution of the field, such as may be provided by a reanalysis dataset. The red point ζ is the state of interest, for example a specific sea-level pressure map at a specific time. (b) Visual illustration of the hyper-sphere determined by the high threshold $s(q, \zeta)$, which defines recurrences, and the distances between measurements defined by $g(x(t), \zeta)$. Note that g is defined so as to take large values for small separations. Thus $g(x(t), \zeta) > s(q, \zeta)$ holds for all points within the hyper-sphere. In the idealised schematic, only two measurements satisfy this condition.

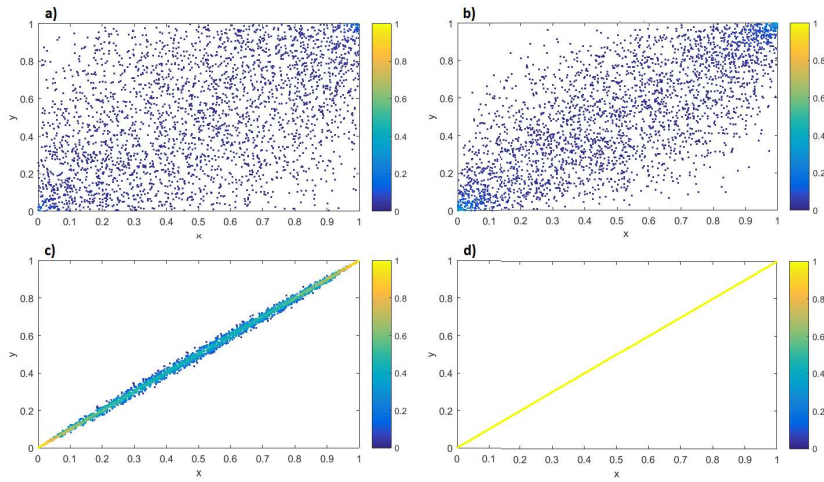


Fig. 2 Four different realizations of the Gaussian bivariate copula (Eq. 6) with $\rho = 0.5$ (a), $\rho = 0.75$ (b), $\rho = 0.9992$ (c), $\rho = 1$ (d). The co-recurrence ratio α is displayed in colours.

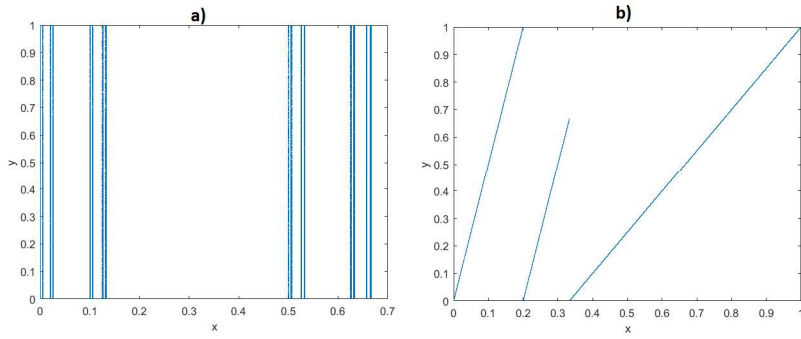


Fig. 3 Two different realizations of 10^5 iterations of the baker's map (Eqs. 7-8) with $\epsilon = 0$ (a), and $\epsilon = 1$ (b).

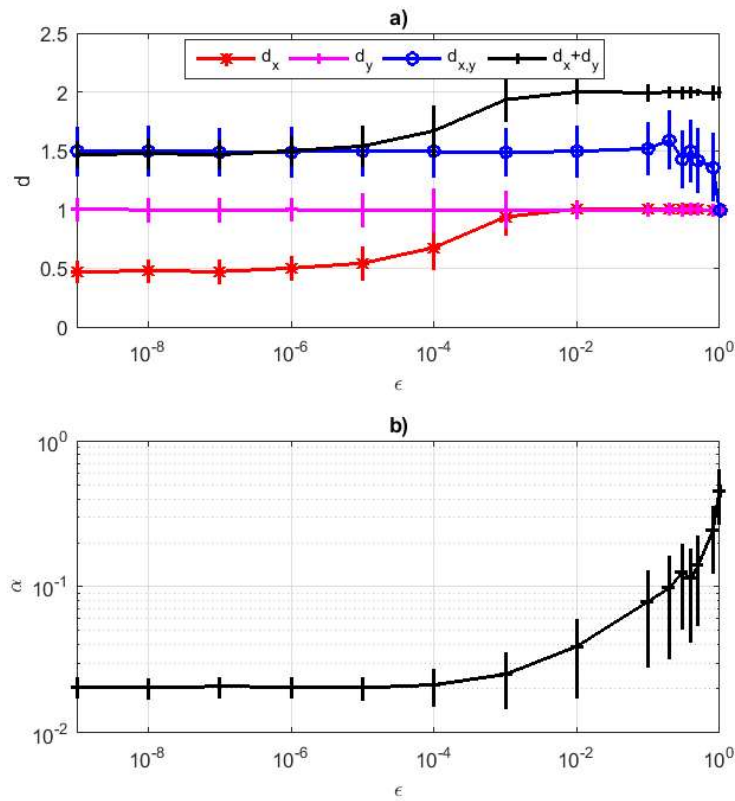


Fig. 4 Dynamical indicators for the baker's map (Eqs. 7-8). a) Dimensions d_x (red line, asterisks) and d_y (pink line, no markers), co-dimensions $d_{x,y}$ (blue line, asterisks) and d_x+d_y (black line, no markers) for different values of ϵ . b) Co-recurrence ratio α for different values of ϵ . Error bars show one standard deviation of the d, α distributions.

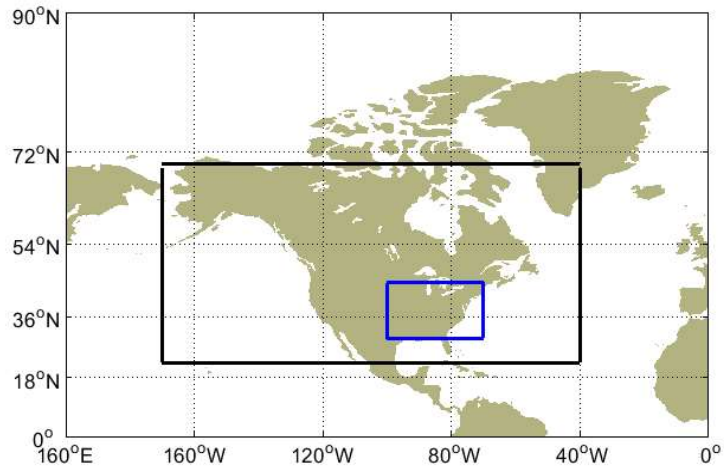


Fig. 5 The region studied in this article. The black rectangle highlights the domain over which the dynamical indicators are computed; the blue rectangle highlights the domain over which the temperature anomalies are computed. Note that only land points are taken into account in the area-averaged temperature anomalies over the domain.

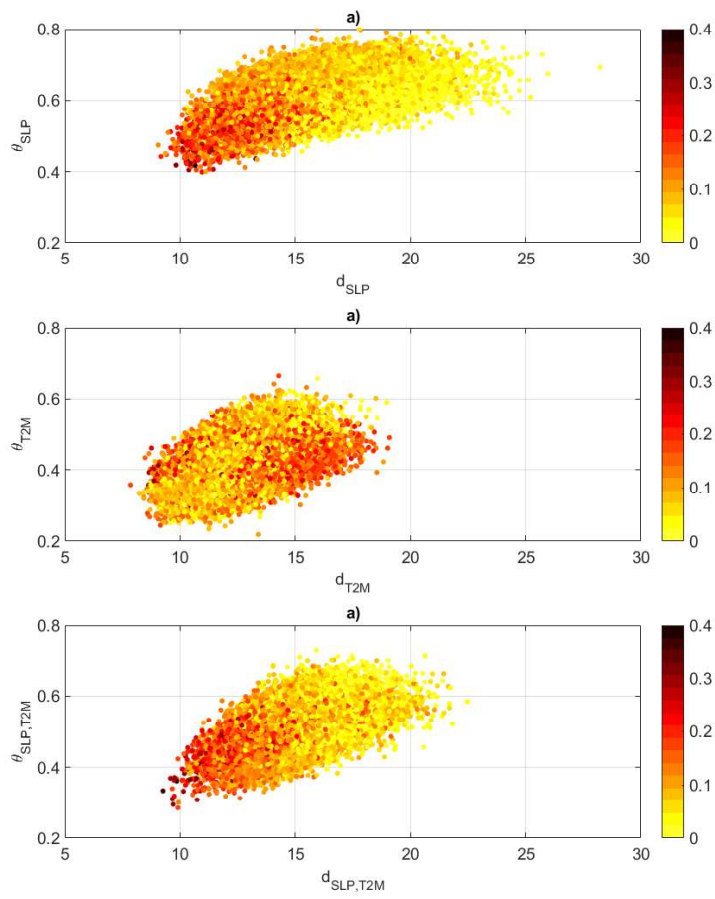


Fig. 6 Local dimension d and inverse persistence θ scatter plots coloured with the values of the co-recurrence ratio α . a) d_{SLP} and θ_{SLP} . b) d_{T2M} and θ_{T2M} . c) $d_{SLP,T2M}$ and $\theta_{SLP,T2M}$.

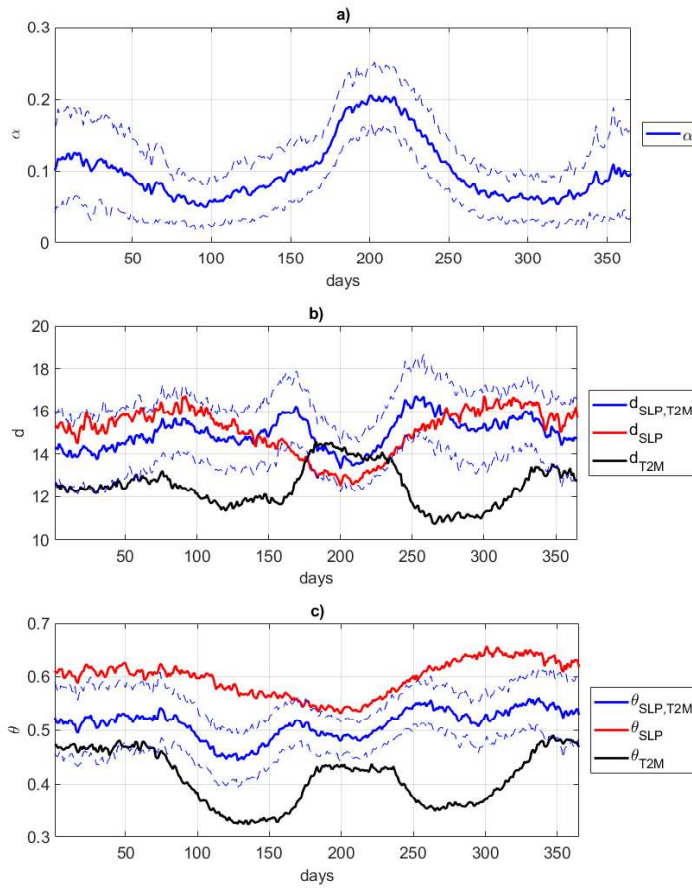


Fig. 7 Seasonal cycles of co-recurrence ratio α (a), local dimension d (b) and inverse persistence θ (c). The continuous red, black and blue lines in panels (b, c) correspond to d , θ computed on SLP, T2M and both variables jointly, respectively. The dashed lines mark one standard deviation of the quantities represented by the blue lines.

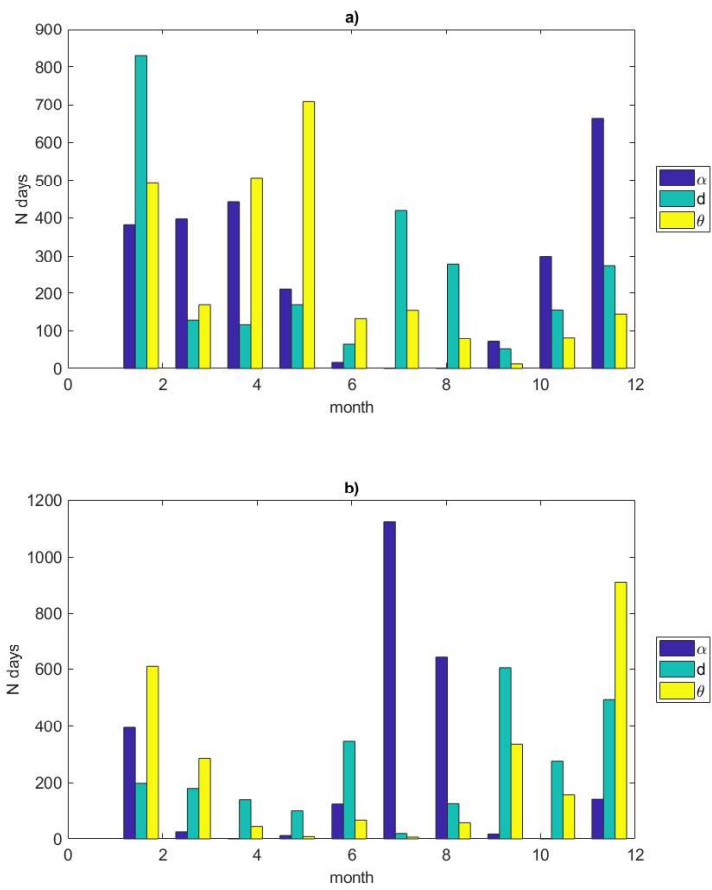


Fig. 8 Number of minima (a) and maxima (b) of co-recurrence ratio α , local co-dimension $d_{SLP,T2M}$ and inverse co-persistence $\theta_{SLP,T2M}$ per month. These are defined as the 10% highest and lowest values of the relevant distributions, respectively.

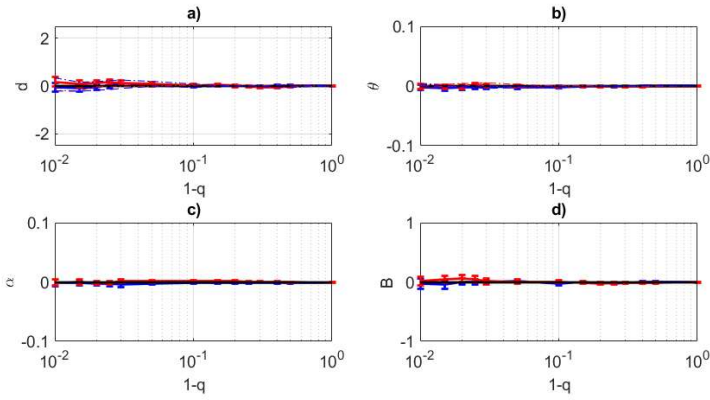


Fig. 9 Local dimensions d_{SLP} (dashed), d_{T2M} (dashed-dotted) and co-dimension $d_{SLP,T2M}$ (continuous) (a); inverse persistences θ_{SLP} (dashed), θ_{T2M} (dashed-dotted) and inverse co-persistence $\theta_{SLP,T2M}$ (continuous) (b); co-recurrence ratio α (c); and average baroclinic vector (d) for summer temperature extremes (cold in blue, warm in red, all data in black) over Eastern North America as a function of the extreme co-percentile q . The values are expressed as average deviations from the respective seasonal cycles. The error bars indicate one standard deviation of the mean and are only shown for the continuous lines.

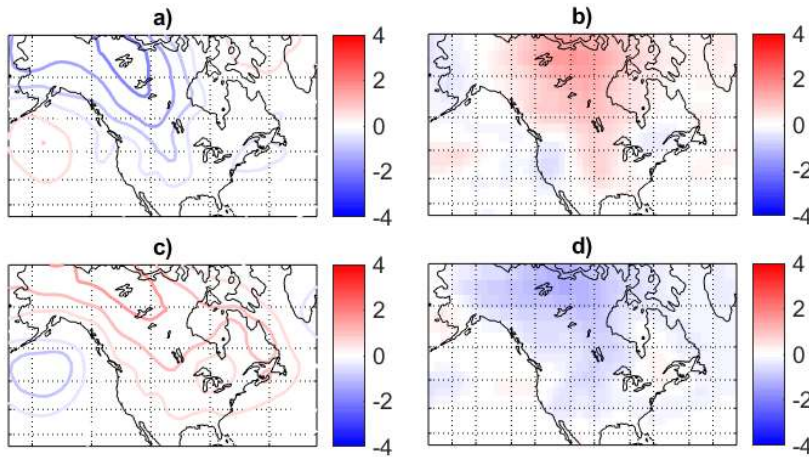


Fig. 10 Composite sea-level pressure (SLP) anomalies in hPa (a,c) and temperature anomalies (T2M) in K (b,d) corresponding to the 10% highest (a,b) and lowest (c,d) values of deviations from the seasonal cycle of the co-recurrence ratio α during JJA. There are no regions where at least $2/3$ of the composited anomalies have the same sign.

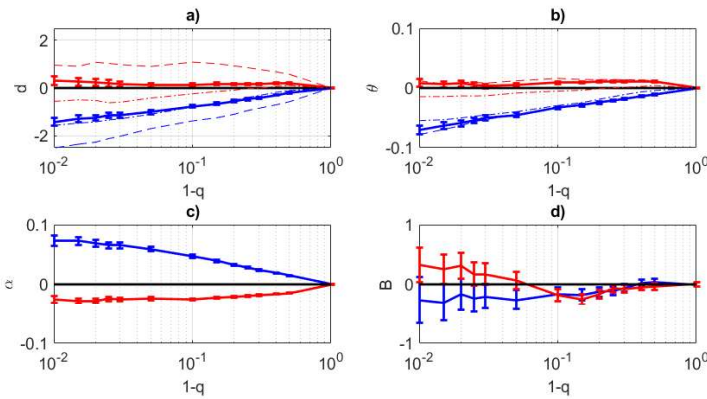


Fig. 11 Local dimensions d_{SLP} (dashed), d_{T2M} (dashed-dotted) and co-dimension $d_{SLP,T2M}$ (continuous) (a); inverse persistences θ_{SLP} (dashed), θ_{T2M} (dashed-dotted) and inverse co-persistence $\theta_{SLP,T2M}$ (continuous) (b); co-recurrence ratio α (c); and average baroclinic vector (d), for winter temperature extremes (cold in blue, warm in red, all data in black) over Eastern North America as a function of the extreme co-percentile q . The values are expressed as deviations from the respective seasonal cycles. The error bars indicate one standard deviation of the mean and are only shown for the continuous lines.

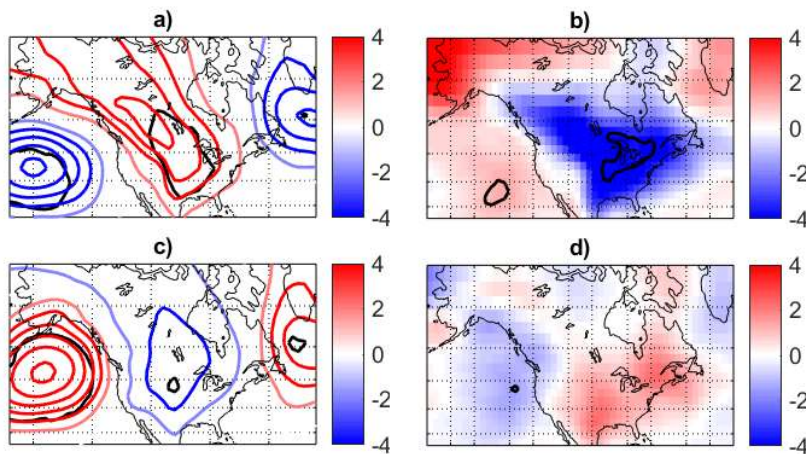


Fig. 12 Composite sea-level pressure (SLP) anomalies in hPa (a,c) and temperature anomalies (T2M) in K (b,d) corresponding to the 10% highest (a,b) and lowest (c,d) values of deviations from the seasonal cycle of the co-recurrence ratio α , during DJF. The black lines bound regions where at least 2/3 of the composited anomalies have the same sign.

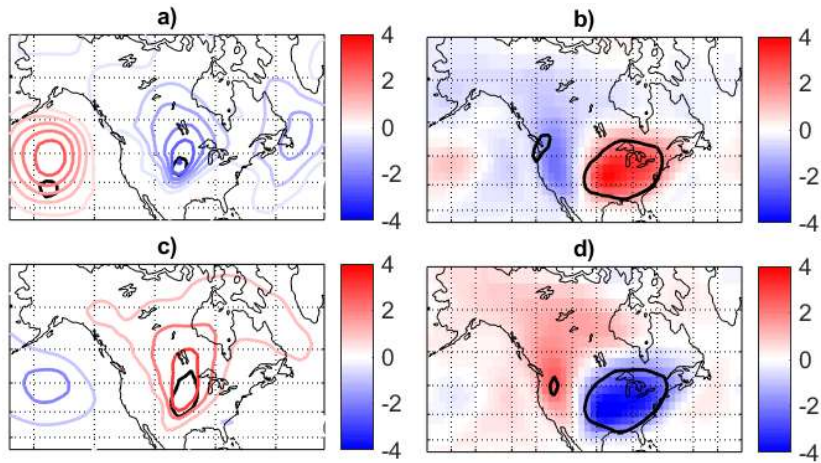


Fig. 13 Composite sea-level pressure (SLP) anomalies in hPa (a,c) and temperature anomalies (T2M) in K (b,d) corresponding to the 10% coldest (a,b) and warmest (c,d) days during JJA. The values are expressed as deviations from the respective seasonal cycles. The black lines bound regions where at least $2/3$ of the composited anomalies have the same sign.

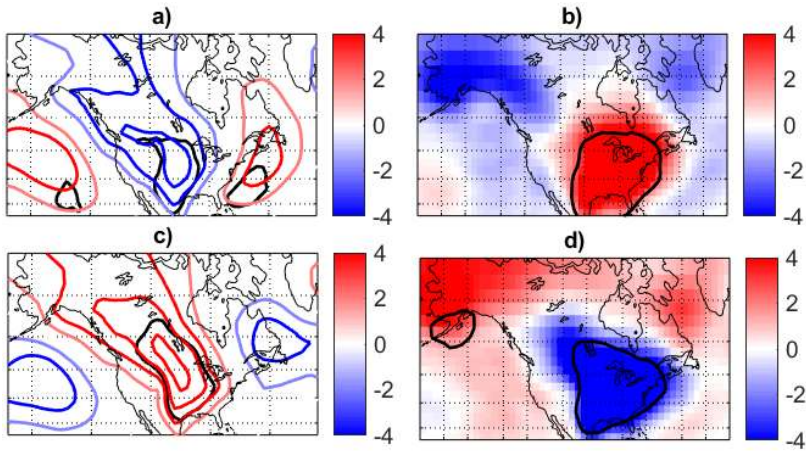


Fig. 14 Composite sea-level pressure (SLP) anomalies in hPa (a,c) and temperature anomalies (T2M) in K (b,d) corresponding to the 10% coldest (a,b) and warmest (c,d) days during DJF. The values are expressed as deviations from the respective seasonal cycles. The black lines bound regions where at least $2/3$ of composited anomalies have the same sign.