



**HAL**  
open science

# Dynamical properties of the North Atlantic atmospheric circulation in the past 150 years in CMIP5 models and the 20CRv2c Reanalysis

David Rodrigues, M Carmen Alvarez-Castro, Gabriele Messori, Pascal Yiou, Yoann Robin, Davide Faranda

## ► To cite this version:

David Rodrigues, M Carmen Alvarez-Castro, Gabriele Messori, Pascal Yiou, Yoann Robin, et al.. Dynamical properties of the North Atlantic atmospheric circulation in the past 150 years in CMIP5 models and the 20CRv2c Reanalysis. *Journal of Climate*, 2018, 10.1175/JCLI-D-17-0176.1 . hal-01504478v2

**HAL Id: hal-01504478**

**<https://hal.science/hal-01504478v2>**

Submitted on 30 Oct 2017

**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.



Distributed under a Creative Commons Attribution 4.0 International License

1 **Dynamical properties of the North Atlantic atmospheric circulation in the**  
2 **past 150 years in CMIP5 models and the 20CRv2c Reanalysis**

3 David Rodrigues<sup>1</sup>, M. Carmen Alvarez-Castro<sup>1</sup>, Gabriele Messori<sup>2</sup>,

4 Pascal Yiou<sup>1</sup>, Yoann Robin<sup>1</sup>, Davide Faranda<sup>1,3\*</sup>

5 <sup>1</sup>*Laboratoire de Sciences du Climat et de l'Environnement, UMR 8212 CEA-CNRS-UVSQ,IPSL,*

6 *Universite Paris-Saclay, 91191 Gif-sur-Yvette, France.*

7 <sup>2</sup>*Department of Meteorology and Bolin Centre for Climate Research, Stockholm University,*

8 *Stockholm, Sweden.*

9 <sup>3</sup>*London Mathematical Laboratory, 14 Buckingham Street, London WC2N 6DF, UK*

10 *\*Corresponding author address: Laboratoire de Sciences du Climat et de l'Environnement, UMR*

11 *8212 CEA-CNRS-UVSQ,IPSL, Universite Paris-Saclay, 91191 Gif-sur-Yvette, France.*

12 *E-mail: davide.faranda@lsce.ipsl.fr*

## ABSTRACT

13 It is of fundamental importance to evaluate the ability of climate models to  
14 capture the large-scale atmospheric circulation patterns and, in the context of  
15 a rapidly increasing greenhouse forcing, the robustness of the changes sim-  
16 ulated in these patterns over time. Here we approach this problem from an  
17 innovative point of view based on dynamical systems theory. We characterize  
18 the atmospheric circulation over the North Atlantic in the CMIP5 historical  
19 simulations (1851 to 2000) in terms of two instantaneous metrics: local di-  
20 mension of the attractor and stability of phase-space trajectories. We then use  
21 these metrics to compare the models to the 20CRv2c reanalysis over the same  
22 historical period. The comparison suggests that: i) most models capture to  
23 some degree the median attractor properties and models with finer grids gener-  
24 ally perform better; ii) in most models the extremes in the dynamical systems  
25 metrics match large-scale patterns similar to those found in the reanalysis; iii)  
26 changes in the attractor properties observed for the ensemble-mean 20CRv2c  
27 reanalysis might be artifacts due to inhomogeneities in the standard deviation  
28 of ensemble over time; iv) the long-term trends in local dimension observed  
29 among the 56 members of the 20-CR ensemble have the same sign as those  
30 observed in the CMIP5 multimodel mean.

## 31 1. Introduction

32 One of the main sources of uncertainty in determining the impact of climate change on ex-  
33 treme events is the forced response of atmospheric dynamics (Shepherd 2014; Field 2012). While  
34 changes in observables such as surface temperature are easily diagnosed, shifts in the mid-latitude  
35 atmospheric patterns have proved very difficult to quantify. Some advances have been made by  
36 focussing on specific features such as atmospheric blocking (Kay et al. 2015; Cassou and Cattiaux  
37 2016; Faranda et al. 2016b), which in turn influence the occurrence of European cold spells and  
38 heat waves, but the broader appreciation of circulation changes is still unsatisfactory. Here we  
39 shed some light on this knowledge gap by using a dynamical systems framework. We illustrate  
40 the power of such an approach by considering the well-known Lorenz (1963) system, a conceptual  
41 model of atmospheric convection consisting of three differential equations<sup>1</sup>

$$\dot{x} = \sigma(y - x), \quad \dot{y} = rx - y - xz, \quad \dot{z} = xy - bz, \quad (1)$$

42 where  $x, y, z$  represent respectively the convection strength, the temperature difference between  
43 the surface and the top of the troposphere and the asymmetry of the convection cells. The  
44 parameters  $\sigma, r$  are the Prandtl and the Rayleigh numbers, while  $b$  is a ratio of critical parameters.  
45 A trajectory of the Lorenz (1963) attractor is shown in blue in Figure 1. The figure consists of  
46 2000 points obtained by iterating the Lorenz equations with  $dt \simeq 0.035$ ,  $\sigma = 28$ ,  $r = 10$ ,  $b = 8/3$   
47 with a Runge Kutta scheme of order 4. These are the classical parameters used by Lorenz (1963).

48  
49 To study the effects of an external forcing, we now increase  $\sigma$  by 2% with respect to the classical  
50 value, such that  $\sigma = 28.5$  (magenta trajectory in Fig. 1). This new trajectory on average favors  
51 higher values of  $z$ , but the changes relative to the original trajectory depend on the point being

---

<sup>1</sup>a list of symbol used in the manuscript is provided in the supplementary material Table S1.

52 considered: some points are not displaced, while some others are mapped elsewhere. Assuming  
53 no knowledge of the system other than the trajectories' paths, is there a way to determine that  
54 they belong to Lorenz attractors with different forcings? To answer this question we would  
55 need: i) to measure the dynamical properties of an ensemble of trajectories representing the two  
56 configurations; and ii) to estimate the distance between these trajectories and determine if the shift  
57 has changed the properties of the points in a detectable way. As a further complicating factor, we  
58 note that changes in natural systems are often gradual, and a sudden jump from one forcing to a  
59 different one is unlikely. A more realistic example discussing a Lorenz attractor with a smoothly  
60 varying forcing is presented in Section 2.

61  
62 The atmospheric equivalent of a point on the Lorenz attractor is the ensemble of instantaneous  
63 fields describing the atmosphere at a time  $t$ . Here we study the atmospheric circulation over the  
64 North Atlantic and focus on a single field: the sea-level pressure (SLP). The SLP field reflects the  
65 major modes of variability affecting the North Atlantic (Hurrell 1995; Moore et al. 2013) and can  
66 further be used to diagnose a wealth of other atmospheric features, ranging from teleconnection  
67 patterns to storm track activity to atmospheric blocking e.g. (Rogers 1997; Comas-Bru and  
68 McDermott 2014). The trajectories of our dynamical systems are the successions of daily SLP  
69 fields from 26 CMIP5 models and the 20CRv2c reanalysis over the period 1851 to 2000 (Compo  
70 et al. 2011). The choice of the North Atlantic domain is motivated by the better observational  
71 coverage over the region in the first part of the analysis period compared to other regions of the  
72 globe (Krueger et al. 2013a). In order to measure changes in the systems, one must be able to  
73 specify at each point (each day) the local (in phase-space, daily in time) dynamical properties  
74 and track their evolution. Recent contributions to dynamical systems analysis have proven that  
75 local properties of the trajectories are characterized by two quantities: the local dimension and

76 the stability of the field considered (Lucarini et al. 2016; Faranda et al. 2017). They correspond  
77 respectively to the rarity and the typical persistence of the configuration. Faranda et al. (2017) and  
78 Messori et al. (2017) have also shown that these two metrics can be connected to the predictability  
79 of a given atmospheric state and that their extremes match climate extremes.

80

81 In this work we will first assess whether the models and reanalysis present similar attractor prop-  
82 erties over the full time period considered. To do this, we compute daily values for the dimension  
83 and stability of the SLP fields and study their average and extreme properties. Next, we quantify  
84 their changes over the analysis period. We then compare the changes seen in the models to those  
85 observed in the reanalysis.

## 86 2. Data & Methods

87 We use daily SLP model output from the historical simulations of 26 CMIP5 models (Table 1)  
88 available from the CMIP5 archive (Taylor et al. 2012). We then compare these simulations to  
89 the 20th Century Reanalysis (20CRv2c) dataset (Compo et al. 2011), studying both the ensemble  
90 mean and the 56 individual members. The analysis focuses on the region  $22.5^{\circ}N - 70^{\circ}N$  and  
91  $80^{\circ}W - 50^{\circ}E$  and the period 1850-2000. SLP anomalies are defined as deviations from a daily  
92 climatology. For example, the SLP anomaly at a given geographical point on the 5th December  
93 2000 is computed relative to the mean SLP value at that same location for all 5th Decembers over  
94 the analysis period.

95

96 In order to compute the dynamical systems metrics, we combine the statistical tools of extreme  
97 value theory with the results obtained by Freitas et al. (2010) for Poincaré recurrences. The pa-  
98 rameters mentioned in the introduction (local dimension  $d$  and stability  $1/\theta$ ) are computed for the

99 points  $\zeta$  on the attractor obtained as a sequence of states of the system. The dynamical indicators  
 100 are linked to the probability  $p$  that a trajectory  $x(t)$  explores a ball centered on  $\zeta$  with diameter  $2\varepsilon$ ,  
 101 i.e. the recurrence rate of the configuration  $\zeta$ . We briefly outline the physical meaning of these  
 102 quantities and the way they are computed below.

103 (i) *Local Dimension:* The Freitas et al. (2010) theorem and its modification in Lucarini et al.  
 104 (2012) states that the probability  $p$  of entering a ball centred on  $\zeta$  with a radius  $\varepsilon$  for chaotic  
 105 attractors obeys a generalized Pareto distribution (Pickands III 1975). In order to compute such  
 106 probability, we first calculate the series of distances  $\delta(x(t), \zeta)$  between the point on the attractor  
 107  $\zeta$  and all other points ( $x(t)$ ) on the trajectory. We then put a logarithmic weight on the time series  
 108 of the distance:

$$g(x(t)) = -\log(\delta(x(t), \zeta)).$$

109 The reason for taking the logarithm is explained by Collet and Eckmann (2009): in the dynamical  
 110 system set-up the negative logarithm increases the discrimination of small values of  $\delta(x, y)$   
 111 which correspond to large values of  $g(x(t))$ . The probability of entering a ball of radius  $\varepsilon$  centered  
 112 on  $\zeta$  is now translated in the probability  $p$  of exceeding a threshold  $q$ . In the limit of an infinitely  
 113 long trajectory, such probability is the exponential member of the generalized Pareto distribution:

$$p = \Pr(g(x(t)) > q, \zeta) \simeq \exp(-[x - \mu(\zeta)]/\beta(\zeta))$$

114 whose parameters  $\mu$  and  $\beta$  depend on the point  $\zeta$  chosen on the attractor. Remarkably,  
 115  $\beta(\zeta) = 1/d(\zeta)$ , where  $d(\zeta)$  is the local dimension around the point  $\zeta$ . This result has recently  
 116 been applied to SLP fields in Faranda et al. (2017). In this paper we use the quantile 0.975 of  
 117 the series  $g(x(t))$  to determine  $q$ . We have further checked the stability of the results against

118 reasonable changes in the quantile.

119

120 (ii) *Local Stability*: The probability  $p$  contains information about the geometry of the ball around  
121  $\zeta$  but provides no insight on the temporal evolution of the dynamics around  $\zeta$ . In particular, it  
122 is interesting to know the mean residence time of the trajectory within the neighborhood of  $\zeta$ .  
123 To measure this quantity, we employ the extremal index  $\theta$  (Freitas et al. 2012; Faranda et al.  
124 2016a), namely the inverse of this mean residence time. Heuristically, if the trajectory enters  $N$   
125 times the neighborhood of  $\zeta$  and at each time  $i$  the length of the cluster (that is, the number of  
126 successive timesteps when the trajectory is within the neighborhood) is  $\tau_i$ , a simple estimate is:  
127  $\theta^{-1} = (1/N) \sum_i \tau_i$ , such that  $\theta$  varies between 0 and 1. The value  $\theta = 0$  corresponds to a stable  
128 fixed point of the dynamics where the observation  $\zeta$  is repeated infinite times (as for a pendulum  
129 left in its equilibrium position). A value of  $\theta = 1$  indicates a point immediately leaving the  
130 neighborhood of  $\zeta$ . Since  $\theta$  is the inverse of a persistence time, it depends on the  $dt$  used. If  $dt$   
131 is too large, the time dependence structure is hidden and  $\theta$  will fail to be close to 1. If  $dt$  is too  
132 small,  $\theta$  is close to zero. In Faranda et al. (2017) it has been observed that  $\theta$  for SLP fields over  
133 the North Atlantic varies between 0.3 and 0.5, when  $dt = 1$  day. In this work we use the same  $dt$ .  
134 The extremal index is estimated with the likelihood estimator by Süveges (2007).

135

136 Figure 2 illustrates the meaning of the indicators for atmospheric flows: the local dimension  $d$   
137 is the number of degrees of freedom needed to describe the dynamics of the system linearized  
138 around the state  $\zeta$  and is therefore proportional to the number of degrees of freedom of  $\zeta$ . In  
139 the pictogram,  $d = 2$  since we only consider two possible origins and evolutions of the state  $\zeta$ .  
140 The inverse of the persistence time  $\theta$  is linked to the probability that the trajectory follows a path



141 where each field resembles those of the previous and subsequent days. In the figure we present  
 142 two possibilities: i) along the green trajectory, the pattern changes every day so that  $\theta(t) = 1$ ;  
 143 ii) along the red trajectory, the same pattern is observed for three consecutive days so that  $\theta = 1/3$ .

144

145 In order to evaluate how close the  $d$  and  $\theta$  of the different models are to those of the reanalysis,  
 146 we adopt a number of distance metrics. The simplest metrics that can be defined are:  $R(d) =$   
 147  $\Delta(d)/\max(\Delta(d))$  and  $R(\theta) = \Delta(\theta)/\max(\Delta(\theta))$ , where  $\Delta$  represents the difference between the  
 148 median values of  $d$  or  $\theta$  of each model and the 20CRv2c values. We further compute a global  
 149 score:

$$R_{tot} = (R(d) + R(\theta))/2. \quad (2)$$

150 To check the validity of the global score, we compare  $R_{tot}$  with the Wasserstein distance  $\mathscr{W}$ ,  
 151 computed as described in Robin et al. (2017).

152

153 The Wasserstein distance  $\mathscr{W}$  measures how two multivariate probability density functions  $\mu$  and  
 154  $\eta$  differ from each other. Probability distributions are like distributions of mass (normalized to  
 155 1) across space. The Wasserstein distance is proportional to the minimal work that is needed to  
 156 transport one distribution of mass into another. It is therefore rooted in optimal transport theory  
 157 (Villani 2008; Santambrogio 2015). In practice, the two measures  $\mu$  and  $\eta$  are discretized into  
 158 multivariate histograms, which are sums of Dirac masses:

$$\mu = \sum_i \mu_i \delta_{x_i}, \quad \eta = \sum_j \eta_j \delta_{x_j}.$$

159 The distribution  $\mu$  is transformed into  $\eta$  by the transport plan  $\gamma$  such that  $\gamma_{ij}$  transports the mass  
 160  $\mu_i$  at  $x_i$  to  $\eta_j$  at  $x_j$ . Therefore the  $\gamma_{ij} \geq 0$  verify:

$$\mu_i = \sum_j \gamma_{ij}, \quad \eta_j = \sum_i \gamma_{ij}, \quad \sum_{ij} \gamma_{ij} = 1.$$

161 If  $\Gamma$  is the set of all possible transport plans of  $\mu$  to  $\eta$ , then the Wassertein distance is:

$$\mathcal{W}(\mu, \eta) = \sqrt{\inf_{\gamma \in \Gamma} \sum_{ij} \gamma_{ij} \|x_i - x_j\|^2},$$

162 where  $\|\cdot\|$  is the usual Euclidean distance.

163

164 Before beginning the analysis of the model and reanalysis data, it is necessary to outline if the  
 165 dynamical properties are sensitive to the changes of the attractor due to continuous modifications  
 166 of the underlying forcing and if these changes are statistically detectable. There are few theoretical  
 167 results on non-stationary statistics of dynamical systems, as well as on non-stationary extreme  
 168 value theory. Luckily, the recurrence approach we use here to estimate  $d$  and  $\theta$  allows to bypass  
 169 most of the technical difficulties linked to non-stationarity, because the dynamical properties  
 170 are measured with respect to each individual state  $\zeta$  of the attractor. If the change affects the  
 171 neighborhood of a state, it will change its dynamical properties. If most of the states are affected  
 172 by the changes in the dynamics, then the average dimension of the attractor and the average  
 173 persistence will change accordingly.

174

175 In order to test this idea, we consider the Lorenz (1963) systems discussed in the Intro-  
 176 duction and perform a set of 30 realizations (trajectories), where  $\sigma$  varies continuously over  
 177  $28 < \sigma < 28.54$  according to:  $\sigma(t) = \sigma(t - dt) + \delta\sigma$ , with  $\delta\sigma = 10^{-5}$  Each realization con-

178 sists of about 70000 iterations with time-step  $dt = 0.02$ . The number of time-steps and the  $dt$  are  
179 chosen to mimic the persistence properties of the SLP field over the North Atlantic, which dis-  
180 plays a median  $\theta$  of around 0.5 (see Section 3). If our methodology can indeed detect the gradual  
181 change from  $\sigma = 28$  to  $\sigma = 28.54$ , then the  $d$  and  $\theta$  distributions for the first and second half of  
182 the simulations should be significantly different. In order to provide a visually immediate picture,  
183 we show distributions of the medians of  $(d, \theta)$  for each half of each simulation (Figure 3). It is  
184 straightforward to verify that the data forms two clouds of well-separated median centroids. This  
185 analysis therefore confirms that the indicators are sensitive to small changes in the attractor prop-  
186 erties and that we can attempt to use them to detect long term changes in the dynamical properties  
187 of reanalysis and CMIP5 data.

### 188 **3. Aggregate analysis of model and reanalysis attractors**

189 We begin the analysis of the daily SLP fields from 1851 to 2000 by presenting the scatterplot  
190 of  $d$  versus  $\theta$  for the ensemble mean of the 20CRv2c reanalysis (Figure 4). Hereafter we will  
191 call this run 20CR-EM. With respect to the computations done for the Lorenz (1963) system,  
192  $\zeta$  is now a daily SLP map and distances are computed using an euclidean metrics at each grid  
193 point as in Faranda et al. (2017). The average of  $d$  is proportional to the number of degrees of  
194 freedom needed to represent the systems' dynamics (this quantity is called attractor dimension  
195 in dynamical systems theory) while the average of  $\theta$  is the inverse of the mean persistence  
196 time of a given SLP configuration. Maxima (minima) of  $d$  therefore correspond to the most  
197 complex (simple) SLP configurations. Maxima (minima) of  $\theta$  correspond to the most unstable  
198 (stable) configurations (Messori et al. 2017; Faranda et al. 2017). Panels A-D in Figure 4 show  
199 the composite SLP anomalies for dynamical extremes - namely days beyond the 0.98 and 0.02  
200 quantiles of the  $d$  and  $\theta$  distributions. These closely – albeit not exactly – resemble the canonical

201 North Atlantic weather regimes (Vautard 1990). In particular, maxima of  $\theta$  (A) reproduce an  
202 Atlantic Ridge pattern, while minima of  $\theta$  (B) correspond to a negative North Atlantic Oscillation  
203 (NAO) phase. Similarly, maxima of  $d$  (C) correspond to a blocking pattern and minima of  $d$  (D) to  
204 a positive NAO. This is in agreement with previous results from Faranda et al. (2017), who further  
205 found that dynamical extremes occur mostly in the winter season. The patterns are stable if the  
206 definition of dynamical extremes is stable up to the 20th and the 80th percentiles of the relevant  
207 distributions, although the magnitude of the composite anomalies reduces when including days  
208 corresponding to lower percentiles (not shown). We note that the values of  $\theta$  shown in the Figure  
209 should not be compared directly to the persistence of the traditional weather regimes defined using  
210 clustering algorithms, as the requirement that the flow does not leave the neighborhood of the state  
211  $\zeta$  is a more restrictive condition than continued permanence within a given cluster. For example,  
212 intense or frequent mobile synoptic systems leading to substantial day-to-day fluctuations in  
213 sea-level pressure could cause the dynamics to leave the ball centered on the state  $\zeta$  in phase space  
214 while remaining within a weather regime cluster. Indeed, if one considers the typical partition of  
215 the North Atlantic atmospheric variability into 4 weather regimes, the probability of being in one  
216 of them is of order 0.25, whereas the probability of being close to  $\zeta$  is set by the threshold  $q$  in  
217 our case 0.0225 (see Section 2).

218  
219 The 20CR-EM data analyzed above is constructed by averaging instantaneous fields from a  
220 56-member ensemble of simulations. The ensemble is less constrained at the beginning of the  
221 period, when surface observations were scarce, than at the end. This implies that the 20CR-EM  
222 fields are smoother at the beginning of the period than at the end, because the latter are obtained by  
223 averaging over an ensemble with smaller differences between individual members. This may affect  
224 the effective number of degrees of freedom as measured by  $d$ . For this reason, we also measure  $d$

225 and  $\theta$  for the 56 individual members and then average them to obtain a single daily value. We will  
226 refer to these quantities as the 20CRv2c means of ensemble (20CR-ME). Schematically:

- 227 • 20CR-EM indicates the daily dynamical properties computed for the 20CRv2c ensemble  
228 mean.
  
- 229 • 20CR-ME indicates the average of daily dynamical properties computed for each single en-  
230 semble member.

231 We can now compare the distributions of  $d, \theta$  for 20CR-EM to  $d, \theta$  for 20CR-ME. The  
232 scatterplot of  $d$  versus  $\theta$  for the 20CR-ME is shown in Figure 5. A number of differences relative  
233 to the 20CR-EM appear (cf. Figure 4). First, the median value of  $d$  is lower for 20CR-EM than  
234 for 20CR-ME (Table 2), which indicates that the averaging of SLP fields in the Ensemble Mean  
235 has suppressed some degrees of freedom. Similarly, the ensemble mean has higher persistence  
236 (lower  $\theta$ ) because smoother fields tend to have slower variations. Although the numerical values  
237 of  $d$  and  $\theta$  differ, the cross-correlation coefficient between the time series for the two data-sets  
238 are 0.93 and 0.97, respectively. This suggests that features such as the seasonality and interannual  
239 variability of the  $d$  and  $\theta$  time series are preserved with the EM averaging. We next look at the  
240 SLP anomaly fields corresponding to the dynamical extremes of the 20CR-ME (Panels A-D in  
241 Fig. 5). We find similar patterns to those observed for 20CR-EM (Panels A-D in Fig. 5)). Indeed,  
242 70% of  $d$  maxima from 20CR-EM match those from 20CR-ME, while this percentage is 61% for  
243 the minima. For  $\theta$  maxima and minima we find values of 81% and 55%, respectively. We further  
244 investigate the differences between 20CR-EM and 20CR-ME by looking at the changes of the  
245 dynamical properties over time in Section 4.

246

247 We next compare the  $(d, \theta)$  bivariate histograms obtained for the 20CR-EM (Figure 6a) with  
248 those computed for the CMIP5 models (Figure 6c,e). Two different behaviors emerge: some of  
249 the models (e.g. CMCC-CMS, Figure 6c) yield a unimodal distribution resembling that obtained  
250 for the 20CR-EM; other models (e.g. the IPSL-CM5A, Figure 6e) show bimodal distributions.  
251 We find these different behaviors to be related to different seasonal cycles: in Figure 6-b,d,f), we  
252 plot the  $(d, \theta)$  scatterplots for the same models by coloring each point according to the month  
253 of the year it occurs in. In the 20CR-EM and the CMCC-CMS model, the different seasons are  
254 spread across the cloud, although maxima of  $\theta$  mostly occur in winter and the summer season is  
255 biased towards low  $d$  and  $\theta$ . The IPSL-CM5A displays a much stronger seasonal discrimination,  
256 with two distinct  $(d, \theta)$  clouds for the winter and for the summer seasons corresponding to the  
257 different modes of the bivariate histograms. This implies that both the bulk statistics and the  
258 extremes are modified by the seasonal cycle. For a more detailed discussion of the seasonality of  
259 the dynamical extremes, we refer the reader to Faranda et al. (2017).

260  
261 Given the variety of the possible behaviors, we will analyze separately the mean and the  
262 extreme behavior of the CMIP5 dynamical properties. We report the aggregate analysis in Table  
263 2 and in Figure 7. In the latter figure, the colored numbers correspond to the median values  
264 for each model (numbered as in Table 1), the red dot to the 20CR-EM median and the black  
265 dots to the medians of each of the 56 members of the 20CRv2c ensemble. For 20CR-EM and  
266 20CR-ME we also draw the ellipses whose semiaxes correspond to the standard deviation of the  
267 mean. We note that the median values of the 20CR-ME are so close to each other as to be almost  
268 indistinguishable, meaning that the individual members provide a coherent representation of the  
269 SLP field. As noted above, the 20CR-EM  $d$  and  $\theta$  median values are both smaller than those of  
270 20CR-ME. All of the models fall within the ellipse of 20CR-EM and most of the models within

271 that of 20CR-ME. In Table 2 we provide the median values and standard deviations of  $d$  and  $\theta$   
272 for all the models and also present the distance metrics  $R(d)$ ,  $R(\theta)$  and  $R_{tot}$  between models and  
273 reanalysis as introduced in Section 2. In order to verify their robustness, we also compare  $R_{tot}$  to  
274 the Wasserstein distance  $\mathcal{W}_2$ , also defined in Section 2 (Fig. 8). The two indicators provide very  
275 similar information (Pearson coefficient:  $r_{pear} = 0.90$  and Spearman coefficient  $r_{spear} = 0.85$   
276 (Von Storch 1999)); from now on we will therefore use the simpler  $R_{tot}$  when discussing our  
277 results.  $R_{tot}$  further indicates the direction of the changes (larger or smaller  $d, \theta$ ) while  $\mathcal{W}_2$  only  
278 provides this information if the transport plan is computed (Villani 2008). The latter would be  
279 particularly complex to compute for the dataset analyzed here.

280

281 All the models are within one standard deviation of the 20CR-EM (Figure 7), while 5 are not  
282 withing the 20CR-ME ellipse. Both Figure 7 and Table 2 further indicate that in many cases  
283 models with a higher horizontal resolution have median values closer to those of the reanalysis.  
284 We note, however, that the median values in  $d$  and  $\theta$  are statistically different from those found in  
285 20CR-EM for all models except model 24 and 26 for  $d$  and model 23 for  $\theta$ . For 20CR-ME, the  
286 medians are different for all models except 13. The statistical significance is determined using the  
287 Wilcoxon ranksum test (Von Storch 1999). The null hypothesis is that the variables are samples  
288 from continuous distributions with equal medians. A rejection of the null hypothesis therefore  
289 indicates a significant difference between the medians.

290

291 We also find that different models have different spreads in  $d$  and  $\theta$  (Table 2), suggesting  
292 investigating the extremes of these quantities and their relation with the weather regimes found  
293 in 20CR-ME and 20CR-EM. Figures S1-S26 in the supplementary material display the  $(d, \theta)$   
294 histograms and the composite SLP anomalies for the  $d$  and  $\theta$  extremes computed as in Figures

295 4.5. A quantitative analysis is reported in Table S2 using the Root Mean Square Error (RMSE)  
296 between 20CR-EM, 20CR-ME and the CMIP5 SLP composite anomalies. In general, we find the  
297 B and D (NAO-like) patterns to have a higher RMSE whereas A and C are better represented. This  
298 is counterintuitive, since one might naively expect models to better reproduce low-dimensional  
299 rather than high-dimensional patterns. At the same time, the NAO is the dominant mode of  
300 variability in the North Atlantic region, which could explain why a higher RMSE is found in that  
301 region of the phase space. Another interesting observation is that models with low  $R_{tot}$  scores  
302 do not always show the best match in the spatial patterns corresponding to the dynamical extremes.

303

#### 304 **4. Changes in the attractor properties**

305 We now investigate whether the SLP's dynamical indicators have changed over time by com-  
306 puting 30-year moving time averages (denoted by  $\langle \cdot \rangle$ ) for 20CR-EM, 20CR-ME (and separately  
307 for the 56 members) and the 26 CMIP5 models. These results are presented in Figure 9. The  
308 abscissa show the final year of each averaging window. In the left panels, we consider the full  
309 150 years and subtract from each model/reanalysis the respective mean values of the indicators in  
310 1880 (that is, averaged over 1851-1880); in the right panels we focus on the last 25 years of data  
311 and subtract the values of 1979 (that is, averaged over 1940-1979).

312

313 20CR-EM (red) shows an increase in  $d$  over time which is opposite in sign and an order of  
314 magnitude larger than the changes observed in 20CR-ME (blue), the single members (light blue),  
315 the CMIP5 multimodel mean (black) and the single CMIP5 models (gray; Figure 9a). When  
316 focusing on the last 25 years of the datasets, the curves scale similarly and most of them show a  
317 moderate decrease of the dimension (Figure 9b). We hypothesize that the rapid increase of the



318 dimension for 20CR-EM during the pre-satellite era is caused by changes in the variance of the  
319 ensemble members which, as discussed in Section 3 above, become more closely constrained as  
320 we approach the present day. To test this hypothesis, we compute the daily values of the standard  
321 deviation of the 56 SLP fields and plot the moving average of such quantity against that of the  
322 dimension of 20CR-EM (Figure 10). The standard deviation of the ensemble is in turn linked to  
323 the number of observations incorporated in the analysis (e.g. Krueger et al. (2013a)). The linear  
324 relation points to the increase in  $\langle d \rangle$  20CR-EM not being a physical trend but rather an artifact due  
325 to the scarcity of observations during the first part of the reanalysis period. However, from 1979,  
326 the variance stabilizes and both the 20CR-EM and 20CR-ME show a decrease in local dimension.  
327 This decrease is also observed in most of CMIP5 models as reflected by the multi-model mean  
328 (Fig. 9b).

329

330 The inverse persistence  $\theta$  of 20CR-EM shows a similar behavior (Figure 9c).  $\theta$  increases  
331 up to 1950 and then levels off. In terms of persistence time this increase is small compared to  
332 the resolution of the datasets used for the analysis (1 day). 20CR-ME shows a similar, albeit  
333 weaker, trend, while the CMIP5 models mostly oscillate around zero. No systematic trends can  
334 be identified after 1970 (Figure 9d).

335

336 We conclude by noting that the changes found in CMIP5 models are mostly of the same sign  
337 as those found for 20CR-ME, but systematically smaller in magnitude. This is associated with  
338 a much larger spread between the different models than between the different members of the  
339 reanalysis, with a number of models showing trends of opposite sign to those of the multi-model  
340 mean.

341

## 342 **5. Discussion and Conclusions**

343 We have computed the instantaneous dynamical properties of the SLP fields over the North  
344 Atlantic Region for the 20CRv2c and the CMIP5 historical runs, over the period 1851-2000.  
345 The goal of our analysis was to assess whether different models with different physics and  
346 resolutions quantitatively represent the *same* dynamical system and therefore possess attractors  
347 with similar characteristics. The metrics we use are the local dimension  $d$  and the inverse  
348 of the persistence time  $\theta$ . As described in Faranda et al. (2017), these two quantities give a  
349 complete characterization of the attractor of the system. To take into account the possibility that  
350 inhomogeneities in the assimilated data create artifacts in the 20CRv2c ensemble-mean SLP  
351 fields, we have computed two sets of dynamical properties. The first is composed of the dynamical  
352 properties of the ensemble-mean field (20CR-EM). The second is computed as the average of the  
353 instantaneous dynamical properties of each of the 56 ensemble members (20CR-ME).

354  
355 When the whole analysis period is considered, we find that the models successfully capture  
356 many of the dynamical systems features identified in the reanalysis, such as the range and  
357 variability of the  $d$  and  $\theta$  metrics. The SLP fields corresponding to extremes in  $d$  and  $\theta$  are also  
358 similar across the models and reanalysis. At the same time, some models exaggerate the effects of  
359 the seasonal cycle on the dynamical indicators, and the statistical agreement in the median values  
360 of the metrics is generally poor. Models with higher horizontal resolutions mostly perform better.

361  
362 To detect the changes in the attractor properties with time, we have analysed the 30-year moving  
363 averages of  $d$  and  $\theta$  for the models and reanalysis. We find a number of interesting behaviors: up  
364 to 1950, the 20CR-EM shows a rapid increase in  $d$  and  $\theta$  whereas 20CR-ME shows a decrease in

365  $d$  and a slower increase in  $\theta$ . After 1950, there is no trend in  $\theta$  and a weakly decreasing trend in  $d$   
366 for both datasets. The CMIP5 multi-model mean shows a weakly decreasing trend in  $d$  throughout  
367 the analysis period, in agreement with 20CR-ME, and no trend in  $\theta$ . These results suggest that  
368 one should be very careful in using ensemble means for studying the atmospheric dynamics of  
369 the late 19th and beginning of the 20th century, even over the North Atlantic (Krueger et al.  
370 2013b; Ferguson and Villarini 2012, 2014). The 20CRv2c ensemble members are increasingly  
371 constrained by a growing number of SLP observations as one approaches the present day. This  
372 causes a decrease of the ensemble spread with time, since the system is more closely pinned to a  
373 specific manifold (the observations) without the possibility of exploring the full phase space. We  
374 hypothesize this is the root cause of the upward trend in the 20CR-EM local dimension.

375

376 The next natural question is whether we can trust the results obtained for single ensemble  
377 members. We believe that the answer is affirmative, as: i) the dataset has a sufficiently high  
378 horizontal resolution to obtain a good estimate of the local dimension distribution (Faranda et al.  
379 2017) and ii) we focus here on the North Atlantic sector, which can be expected to perform better  
380 than elsewhere since most of the observations used to constrain 20CRv2c in the first part of the  
381 dataset are located in Europe or eastern North America (Cram et al. 2015). We therefore argue  
382 that the results obtained for the 20CR-ME and the multimodel ensemble are valuable, and that  
383 the decrease in dimension with time is a real and interesting feature of the atmospheric dynamics  
384 which merits a more detailed analysis in further studies. In fact, a decreasing dimension implies a  
385 more predictable atmosphere.

386

387 As a final caveat we note that our analysis does not attempt to separate the forced variability  
388 from natural low-frequency oscillations and that, especially during the first part of the analysis

389 period, it is unclear whether the greenhouse forcing can be clearly discerned above the background  
390 "climate noise" (Paeth et al. 1999; Lyu et al. 2015). We must therefore take into account the  
391 possibility that the data's internal variability dominates over the long-term forcing trends for the  
392 time period considered.

393

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

## 397 **References**

398 Cassou, C., and J. Cattiaux, 2016: Disruption of the european climate seasonal clock in a warming  
399 world. *Nature Climate Change*, **6 (6)**, 589–594.

400 Collet, P., and J.-P. Eckmann, 2009: *Iterated maps on the interval as dynamical systems*. Springer  
401 Science & Business Media.

402 Comas-Bru, L., and F. McDermott, 2014: Impacts of the ea and sca patterns on the european  
403 twentieth century nao–winter climate relationship. *Quarterly Journal of the Royal Meteorolog-  
404 ical Society*, **140 (679)**, 354–363.

405 Compo, G. P., and Coauthors, 2011: The twentieth century reanalysis project. *Quarterly Journal  
406 of the Royal Meteorological Society*, **137 (654)**, 1–28.

407 Cram, T. A., and Coauthors, 2015: The international surface pressure databank version 2. *Geo-  
408 science Data Journal*, **2 (1)**, 31–46, doi:10.1002/gdj3.25, URL [http://dx.doi.org/10.1002/gdj3.  
409 25](http://dx.doi.org/10.1002/gdj3.25).

- 410 Faranda, D., M. C. Alvarez-Castro, and P. Yiou, 2016a: Return times of hot and cold days via re-  
411 currences and extreme value theory. *Climate Dynamics*, 1–13, doi:10.1007/s00382-016-3042-6.
- 412 Faranda, D., G. Masato, N. Moloney, Y. Sato, F. Daviaud, B. Dubrulle, and P. Yiou, 2016b: The  
413 switching between zonal and blocked mid-latitude atmospheric circulation: a dynamical system  
414 perspective. *Climate Dynamics*, **47 (5-6)**, 1587–1599.
- 415 Faranda, D., G. Messori, and P. Yiou, 2017: Dynamical proxies of north atlantic predictability and  
416 extremes. *Scientific Reports*, **7**, 41 278.
- 417 Ferguson, C. R., and G. Villarini, 2012: Detecting inhomogeneities in the twentieth century reanal-  
418 ysis over the central united states. *Journal of Geophysical Research: Atmospheres*, **117 (D5)**,  
419 n/a–n/a, doi:10.1029/2011JD016988, URL <http://dx.doi.org/10.1029/2011JD016988>, d05123.
- 420 Ferguson, C. R., and G. Villarini, 2014: An evaluation of the statistical homogeneity of  
421 the twentieth century reanalysis. *Climate Dynamics*, **42 (11)**, 2841–2866, doi:10.1007/  
422 s00382-013-1996-1, URL <http://dx.doi.org/10.1007/s00382-013-1996-1>.
- 423 Field, C. B., 2012: *Managing the risks of extreme events and disasters to advance climate change*  
424 *adaptation: special report of the intergovernmental panel on climate change*. Cambridge Uni-  
425 versity Press.
- 426 Freitas, A. C. M., J. M. Freitas, and M. Todd, 2010: Hitting time statistics and extreme value  
427 theory. *Probability Theory and Related Fields*, **147 (3-4)**, 675–710.
- 428 Freitas, A. C. M., J. M. Freitas, and M. Todd, 2012: The extremal index, hitting time statistics and  
429 periodicity. *Advances in Mathematics*, **231 (5)**, 2626–2665.
- 430 Hurrell, J. W., 1995: Decadal trends in the north atlantic oscillation: Regional tem-  
431 peratures and precipitation. *Science*, **269 (5224)**, 676–679, doi:10.1126/science.269.5224.

432 676, URL <http://science.sciencemag.org/content/269/5224/676>, [http://science.sciencemag.org/](http://science.sciencemag.org/content/269/5224/676.full.pdf)  
433 [content/269/5224/676.full.pdf](http://science.sciencemag.org/content/269/5224/676.full.pdf).

434 Kay, J., and Coauthors, 2015: The community earth system model (cesm) large ensemble project:  
435 A community resource for studying climate change in the presence of internal climate variabil-  
436 ity. *Bulletin of the American Meteorological Society*, **96** (8), 1333–1349.

437 Krueger, O., F. Schenk, F. Feser, and R. Weisse, 2013a: Inconsistencies between long-term trends  
438 in storminess derived from the 20cr reanalysis and observations. *Journal of Climate*, **26** (3),  
439 868–874.

440 Krueger, O., F. Schenk, F. Feser, and R. Weisse, 2013b: Inconsistencies between long-term trends  
441 in storminess derived from the 20cr reanalysis and observations. *Journal of Climate*, **26** (3),  
442 868–874, doi:10.1175/JCLI-D-12-00309.1.

443 Lorenz, E. N., 1963: Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences*, **20** (2),  
444 130–141.

445 Lucarini, V., D. Faranda, G. Turchetti, and S. Vaienti, 2012: Extreme value theory for singular  
446 measures. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, **22** (2), 023 135.

447 Lucarini, V., and Coauthors, 2016: *Extremes and recurrence in dynamical systems*. John Wiley &  
448 Sons.

449 Lyu, K., X. Zhang, J. A. Church, and J. Hu, 2015: Quantifying internally generated and ex-  
450 ternally forced climate signals at regional scales in cmip5 models. *Geophysical Research*  
451 *Letters*, **42** (21), 9394–9403, doi:10.1002/2015GL065508, URL [http://dx.doi.org/10.1002/](http://dx.doi.org/10.1002/2015GL065508)  
452 [2015GL065508](http://dx.doi.org/10.1002/2015GL065508), 2015GL065508.

- 453 Messori, G., R. Caballero, and D. Faranda, 2017: A dynamical systems approach to studying  
454 mid-latitude weather extremes. *Submitted: Geophysical Research Letters*.
- 455 Moore, G., I. A. Renfrew, and R. S. Pickart, 2013: Multidecadal mobility of the north atlantic  
456 oscillation. *Journal of Climate*, **26** (8), 2453–2466.
- 457 Paeth, H., A. Hense, R. Glowienka-Hense, S. Voss, and U. Cubasch, 1999: The north at-  
458 lantic oscillation as an indicator for greenhouse-gas induced regional climate change. *Cli-  
459 mate Dynamics*, **15** (12), 953–960, doi:10.1007/s003820050324, URL [http://dx.doi.org/10.  
460 1007/s003820050324](http://dx.doi.org/10.1007/s003820050324).
- 461 Pickands III, J., 1975: Statistical inference using extreme order statistics. *the Annals of Statistics*,  
462 119–131.
- 463 Robin, Y., P. Yiou, and P. Naveau, 2017: Detecting changes in forced climate attractors with  
464 wasserstein distance. *Nonlinear Processes in Geophysics Discussions*, **2017**, 1–19, doi:10.5194/  
465 npg-2017-5, URL <http://www.nonlin-processes-geophys-discuss.net/npg-2017-5/>.
- 466 Rogers, J. C., 1997: North atlantic storm track variability and its association to the north atlantic  
467 oscillation and climate variability of northern europe. *Journal of Climate*, **10** (7), 1635–1647,  
468 doi:10.1175/1520-0442(1997)010<1635:NASTVA>2.0.CO;2.
- 469 Santambrogio, F., 2015: *Optimal Transport for Applied Mathematicians*, Vol. 87. Birkhäuser  
470 Basel, XXVII, 353 pp.
- 471 Shepherd, T. G., 2014: Atmospheric circulation as a source of uncertainty in climate change  
472 projections. *Nature Geoscience*, **7** (10), 703–708.
- 473 Süveges, M., 2007: Likelihood estimation of the extremal index. *Extremes*, **10** (1), 41–55.

- 474 Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the experiment  
475 design. *Bulletin of the American Meteorological Society*, **93** (4), 485–498.
- 476 Vautard, R., 1990: Multiple weather regimes over the north atlantic: Analysis of precursors and  
477 successors. *Monthly Weather Review*, **118** (10), 2056–2081.
- 478 Villani, C., 2008: *Optimal transport: old and new*, Vol. 338. Springer Science & Business Media.
- 479 Von Storch, H., 1999: Misuses of statistical analysis in climate research. *Analysis of Climate*  
480 *Variability*, Springer, 11–26.



481 **LIST OF TABLES**

482 **Table 1.** List of CMIP5 Models analysed and 20CRv2c reanalysis from 1851 to 2001. Models are  
483 approximately ordered by increasing horizontal resolution. . . . . 25

484 **Table 2.** List of median values and standard deviations of dimension  $d$  and inverse per-  
485 sistence  $\theta$  for 20CR-EM, 20CR-ME and the 26 CMIP5 models. The table  
486 further presents the relative distances  $R(d) = \delta(d)/\max(\delta(d))$  and  $R(\theta) =$   
487  $\delta(\theta)/\max(\delta(\theta))$  normalized with respect to the farthest model and the global  
488 score  $R_{tot} = (R(d) + R(\theta))/2$ . . . . . 26

TABLE 1. List of CMIP5 Models analysed and 20CRv2c reanalysis from 1851 to 2001. Models are approximately ordered by increasing horizontal resolution.

No. <sup>1</sup>	Model	Institution/ID	Country	Resolution <sup>2</sup>
0	20CRv2c ME	NOAA-CIRES	USA	2 x 2
1	20CRv2c EM	NOAA-CIRES	USA	2 x 2
2	CMCC-CESM	Centro Euro-Mediterraneo sui Cambiamenti Climatici	Italy	3.75x3.75
3	CanESM2	Canadian Centre for Climate Modelling and Analysis, CCCMa	Canada	2.81x 2.79
4	MIROC-ESM-CHEM	MIROC <sup>4</sup>	Japan	2.81x 2.79
5	MIROC-ESM	MIROC <sup>4</sup>	Japan	2.81x 2.79
6	BCC-CSM1-1	Beijing Climate Center	China	2.81x 2.79
7	IPSL-CM5B-LR	Institute Pierre Simon Laplace, IPSL	France	3.75x1.89
8	NorESM1-M	Norwegian Climate Center	Norway	2.5x1.89
9	FGOALS-2	Institute of Atmospheric Physics, Chinese Academy of Sciences	China	2.81x2.81
10	MPI-ESM-P	Max Planck Institute for Meteorology, MPI	Germany	1.87x1.87
11	MPI-ESM-LR	Max Planck Institute for Meteorology, MPI	Germany	1.87x1.87
12	CSIRO-MK3-6-0	CSIRO-BOM <sup>5</sup>	Australia	1.87x1.87
13	CMCC-CMS	Centro Euro-Mediterraneo sui Cambiamenti Climatici	Italy	1.87x1.87
14	MPI-ESM-MR	Max Planck Institute for Meteorology, MPI	Germany	1.87x1.87
15	IPSL-CM5A-MR	Institute Pierre Simon Laplace, IPSL	France	2.5x1.26
16	INM-CM4	Institute for Numerical Mathematics, INM	Russia	2x1.5
17	ACCESS 1-0	CSIRO-BOM <sup>5</sup>	Australia	1.87x1.25
18	MIROC5	MIROC <sup>4</sup>	Japan	1.40x1.40
19	CNRM-CM5	CNRM-CERFACS <sup>3</sup>	France	1.40x1.40
20	MRI-ESM1	Meteorological Research Institute, MRI	Japan	1.125x1.125
21	BCC-CSM1-M	Beijing Climate Center	China	1.125x1.125
22	MRI-CGCM3	Meteorological Research Institute, MRI	Japan	1.125x1.125
23	EC-EARTH	Danish Meteorological Institute, DMI	Denmark	1.125x1.125
24	CESM1-FASTCHEM	Community Earth System Model Contributors, NCAR	USA	1.25x0.94
25	CESM1-CAM5	Community Earth System Model Contributors, NCAR	USA	1.25x0.94
26	CESM1-BGC	Community Earth System Model Contributors, NCAR	USA	1.25x0.94
27	CCSM4	National Center for Atmospheric Research, NCAR	USA	1.25x0.94

<sup>1</sup> Order by horizontal resolution (Decreasing)

<sup>2</sup> Longitude x Latitude (°)

<sup>3</sup> Centre National de Recherches Meteorologiques - Centre Europeen de Recherche et de Formation Avancee en Calcul Scientifique

<sup>4</sup> Atmosphere and Ocean Research Institute (University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology

<sup>5</sup> Commonwealth Scientific and Industrial Research Organisation(CSIRO), Bureau of Meteorology(BOM)

489 TABLE 2. List of median values and standard deviations of dimension  $d$  and inverse persistence  $\theta$  for  
 490 20CR-EM, 20CR-ME and the 26 CMIP5 models. The table further presents the relative distances  $R(d) =$   
 491  $\delta(d)/\max(\delta(d))$  and  $R(\theta) = \delta(\theta)/\max(\delta(\theta))$  normalized with respect to the farthest model and the global  
 492 score  $R_{tot} = (R(d) + R(\theta))/2$ .

Model	median( $d$ )	median( $\theta$ )	std( $d$ )	std( $\theta$ )	$R(d)$ EM	$R(\theta)$ EM	$R_{tot}$ EM	$R(d)$ ME	$R(\theta)$ ME	$R_{tot}$ ME
00-20CR-ME	12.26	0.53	1.85	0.05						
01-20CR-EM	11.56	0.50	2.06	0.06						
02-CMCC-CESM	12.22	0.51	1.90	0.06	0.54	0.13	0.33	0.04	0.23	0.14
03-CanESM2	11.99	0.51	2.06	0.07	0.35	0.02	0.19	0.34	0.30	0.32
04-MIROC-ESM-CHEM	12.54	0.47	1.87	0.06	0.80	0.75	0.77	0.35	0.83	0.59
05-MIROC-ESM	12.48	0.47	1.95	0.05	0.75	0.76	0.76	0.27	0.84	0.56
06-BCC-CSM1	12.12	0.51	1.95	0.05	0.46	0.12	0.29	0.16	0.24	0.20
07-IPSL-CM5B	12.73	0.46	1.79	0.06	0.95	0.93	0.94	0.58	0.95	0.77
08-NorESM1-M	12.12	0.48	2.27	0.06	0.46	0.44	0.45	0.18	0.62	0.40
09-FGOALS-S2	11.63	0.45	1.89	0.06	0.06	1.00	0.53	0.78	1.00	0.89
10-MPI-ESM-P	12.17	0.51	1.82	0.07	0.50	0.06	0.28	0.11	0.28	0.19
11-MPI-ESM-LR	12.13	0.51	1.91	0.05	0.47	0.14	0.30	0.16	0.23	0.19
12-CSIRO-MK3-6-0	12.66	0.50	1.91	0.05	0.90	0.02	0.46	0.50	0.34	0.42
13-CMCC-CMS	11.95	0.52	2.14	0.06	0.32	0.22	0.27	0.38	0.17	0.28
14-MPI-ESM-MR	12.09	0.51	1.86	0.06	0.43	0.09	0.26	0.21	0.26	0.24
15-IPSL-CM5A	11.86	0.48	1.90	0.06	0.25	0.51	0.38	0.49	0.67	0.58
16-INM-CM4	12.79	0.47	2.14	0.07	1.00	0.70	0.85	0.65	0.80	0.72
17-ACCESS '1-0	11.74	0.49	1.92	0.06	0.15	0.31	0.23	0.64	0.53	0.58
18-MIROC5	12.58	0.49	2.00	0.07	0.83	0.33	0.58	0.39	0.54	0.47
19-CNRM-CM5	12.36	0.47	2.08	0.05	0.65	0.59	0.62	0.12	0.72	0.42
20-MRI-ESM1	11.72	0.51	2.09	0.05	0.13	0.12	0.13	0.67	0.23	0.45
21-BCC-CSM1-M	11.45	0.55	2.00	0.07	0.09	0.91	0.50	1.00	0.30	0.65
22-MRI-CGCM3	11.74	0.51	1.66	0.07	0.15	0.07	0.11	0.64	0.27	0.45
23-EC-EARTH	11.87	0.50	1.99	0.07	0.26	0.01	0.13	0.47	0.33	0.40
24-CESM1-FASTCHEM	11.56	0.51	1.90	0.06	0.00	0.18	0.09	0.86	0.19	0.53
25-CESM1-CAM5	11.88	0.51	1.83	0.07	0.26	0.13	0.20	0.47	0.23	0.35
26-CESM1-BGC	11.53	0.51	2.01	0.06	0.02	0.11	0.06	0.90	0.24	0.57
27-CCSM4	11.57	0.51	1.84	0.07	0.01	0.12	0.07	0.85	0.24	0.54

493 **LIST OF FIGURES**

494 **Fig. 1.** Two realizations of the Lorenz attractor. Blue: classic attractor with:  $dt \simeq 0.035$ ,  $\sigma = 28$ ,  
495  $r = 10$ ,  $b = 8/3$ ; Violet:  $\sigma = 28.5$  . . . . . 29

496 **Fig. 2.** Schematic representation of the dynamical indicators on fictive SLP maps. H indicates high  
497 and L low pressure systems; the arrows mimic horizontal wind fields. The local dimension  
498  $d$  is proportional to the number of possible configurations preceding and following the day  
499 analyzed (here  $d = 2$ ).  $\theta$  is the inverse of the persistence time of a given configuration. If  
500 the pattern persist for three days (red path), then  $\theta = 1/3$ . If the patterns change every day  
501 (green path), then  $\theta = 1$ . . . . . 30

502 **Fig. 3.** Medians of  $d$  and  $\theta$  (dots) and standard deviation (semiaxes of the ellipses) for 30 realiza-  
503 tions of the Lorenz attractor with  $\sigma$  varying continuously over  $28 < \sigma < 28.54$ . The blue  
504 markers correspond to the first half of each simulation  $\sigma_1$ ; the red markers to the second half  
505  $\sigma_2$ . The asterisks mark the medians of the ensemble of realizations. . . . . 31

506 **Fig. 4.** Scatterplot of the daily values of instantaneous dimension  $d$  and inverse persistence  $\theta$  for the  
507 20CRv2c ensemble mean SLP (20CR-EM). The straight black lines mark the 0.02 and 0.98  
508 quantiles of  $d$  and  $\theta$ . The composite anomalies in SLP obtained averaging the days beyond  
509 the quantiles can be associated with known weather regimes: A) Atlantic Ridge (maxima  
510 of  $\theta$ ), B) NAO- (minima of  $\theta$ ), C) Blocking (maxima of  $d$ ), D) NAO+ (minima of  $d$ ). The  
511 black contours in panels A-D indicate regions where at least 2/3 of the composite members  
512 display sea-level pressure anomalies of the same sign. . . . . 32

513 **Fig. 5.** Scatterplot of the daily values of instantaneous dimension  $d$  and inverse persistence  $\theta$  for  
514 the SLP fields of the 56 individual 20CRv2c members (20CR-ME). The straight black lines  
515 mark the 0.02 and 0.98 quantiles of  $d$  and  $\theta$ . The composite anomalies in SLP obtained  
516 averaging the days beyond the quantiles can be associated with known weather regimes: A)  
517 Atlantic Ridge (maxima of  $\theta$ ), B) NAO- (minima of  $\theta$ ), C) Blocking (maxima of  $d$ ), D)  
518 NAO+ (minima of  $d$ ). The black contours in panels A-D indicate regions where at least 2/3  
519 of the composite members display sea-level pressure anomalies of the same sign. . . . . 33

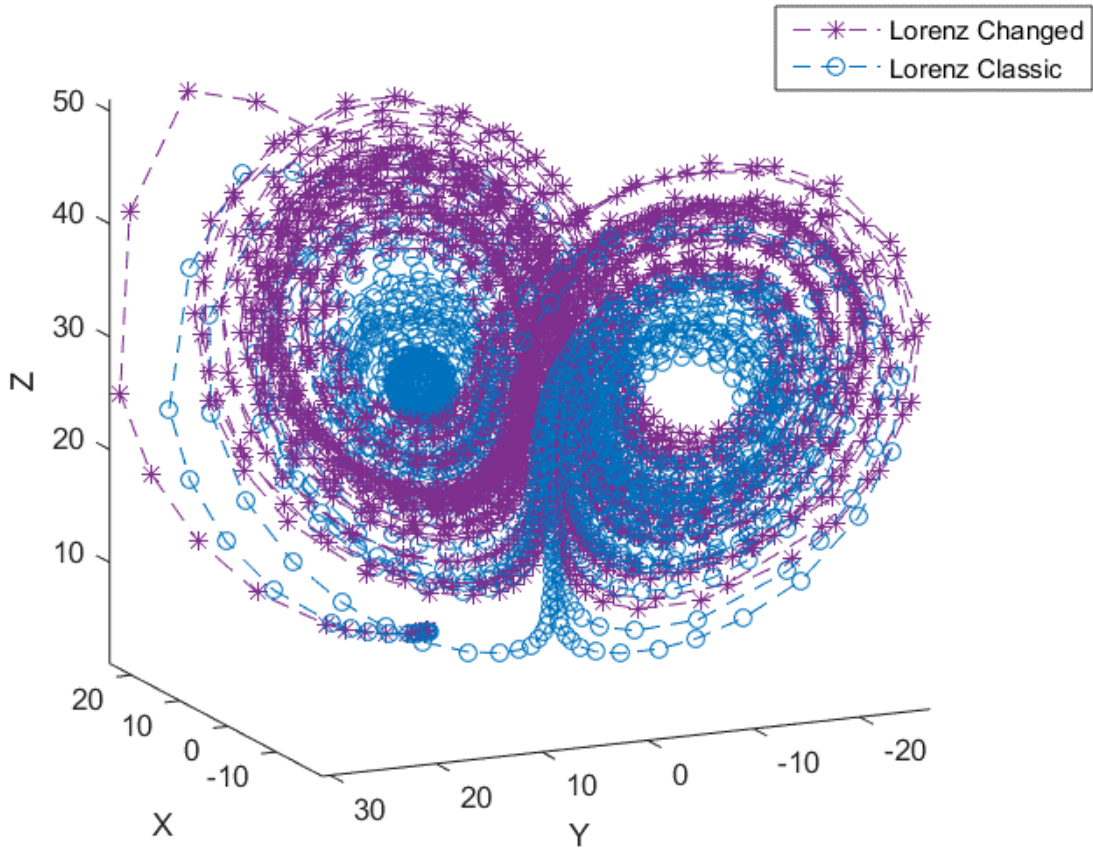
520 **Fig. 6.**  $(d, \theta)$  bivariate histograms (a,c,e) and scatter-plots (b,d,f) for the 20CR-EM reanalysis (a,b),  
521 CMCC-CMS (c,d) and IPSL-CM5A models (e,f). The color scales in (a,c,e) indicate the  
522 frequency of observations in number of days. The color scales in (b,d,f) indicate the month  
523 of the observation and show the dependence of the  $(d, \theta)$  diagrams on the seasonal cycle. . . . . 34

524 **Fig. 7.** Comparison between the 56 20CR-ME median values of  $(d, \theta)$  (black points whose average  
525 is denoted by 0), the 20CR-EM (in red and numbered by 1) and all the CMIP5 models (pro-  
526 gressive numbers 2-27, see table 1 for details). The semiaxes of the two ellipses represent  
527 one standard deviation of  $d$  and  $\theta$  for 20CR-EM (red) and 20CR-ME (black). . . . . 35

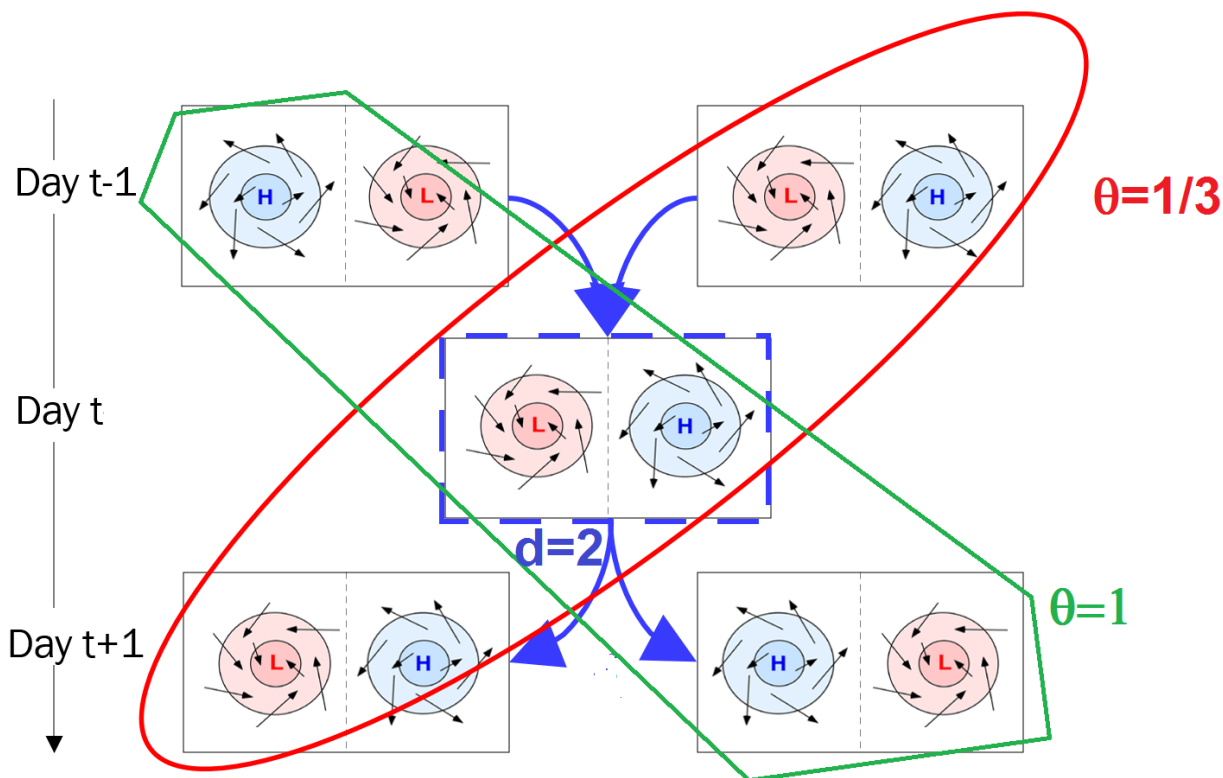
528 **Fig. 8.** Comparison between  $R_{tot}$  values (black) and Wasserstein distances  $\mathscr{W}$  (red) between the  
529  $(d, \theta)$  of 20CR-EM (a) and 20CR-ME (b) and of the CMIP5 models. The distances are  
530 normalized by the maximum value. . . . . 36

531 **Fig. 9.** 30-year moving averages  $\langle \cdot \rangle$  of instantaneous dimension  $d$  and inverse persistence  $\theta$  minus  
532 the respective mean values over 1851-1880 (a,c) and 1950-1979 (b,d). Note that labels on  
533 the abscissa mark the last year in each 30-year averaging window. (a,b): local dimension  $d$ ;  
534 (c,d): inverse persistence  $\theta$ . Red: 20CR-EM; Blue: 20CR-ME; Light blue: single 20CRv2c  
535 members; Black: CMIP5 multimodel mean; Grey: single CMIP5 models; Green: zero line. . . . . 37

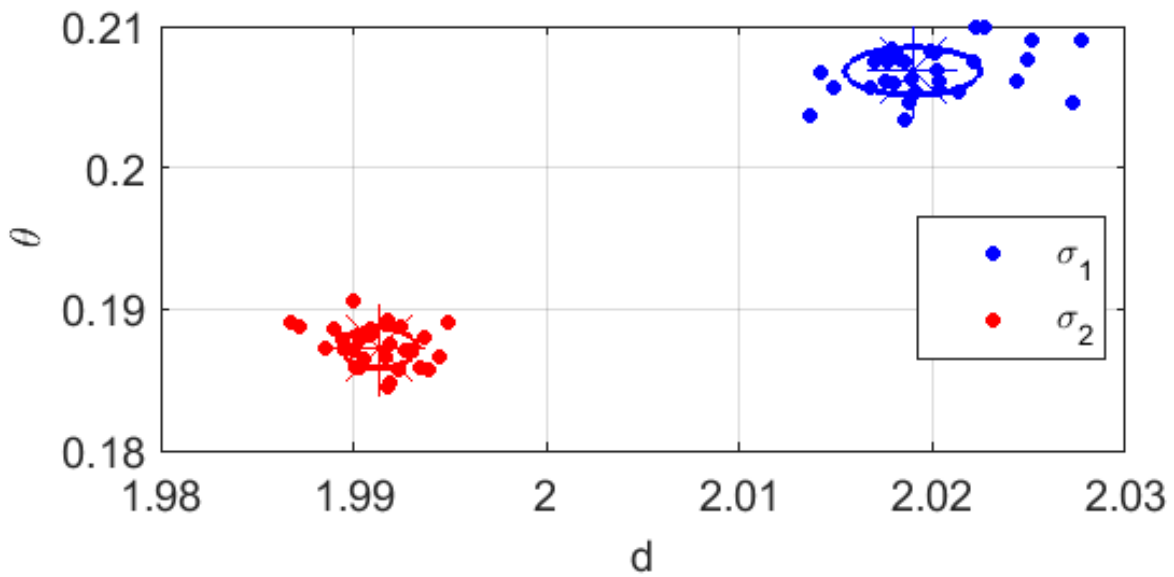
536 **Fig. 10.** 30-year moving average of the standard deviation of SLP fields across the 20CRv2c ensem-  
537 ble  $\langle std(SLP) \rangle$  versus  $\langle d \rangle$  for 20CR-EM. The color scale shows time. . . . . 38



538 FIG. 1. Two realizations of the Lorenz attractor. Blue: classic attractor with:  $dt \simeq 0.035$ ,  $\sigma = 28$ ,  $r = 10$ ,  
 539  $b = 8/3$ ; Violet:  $\sigma = 28.5$

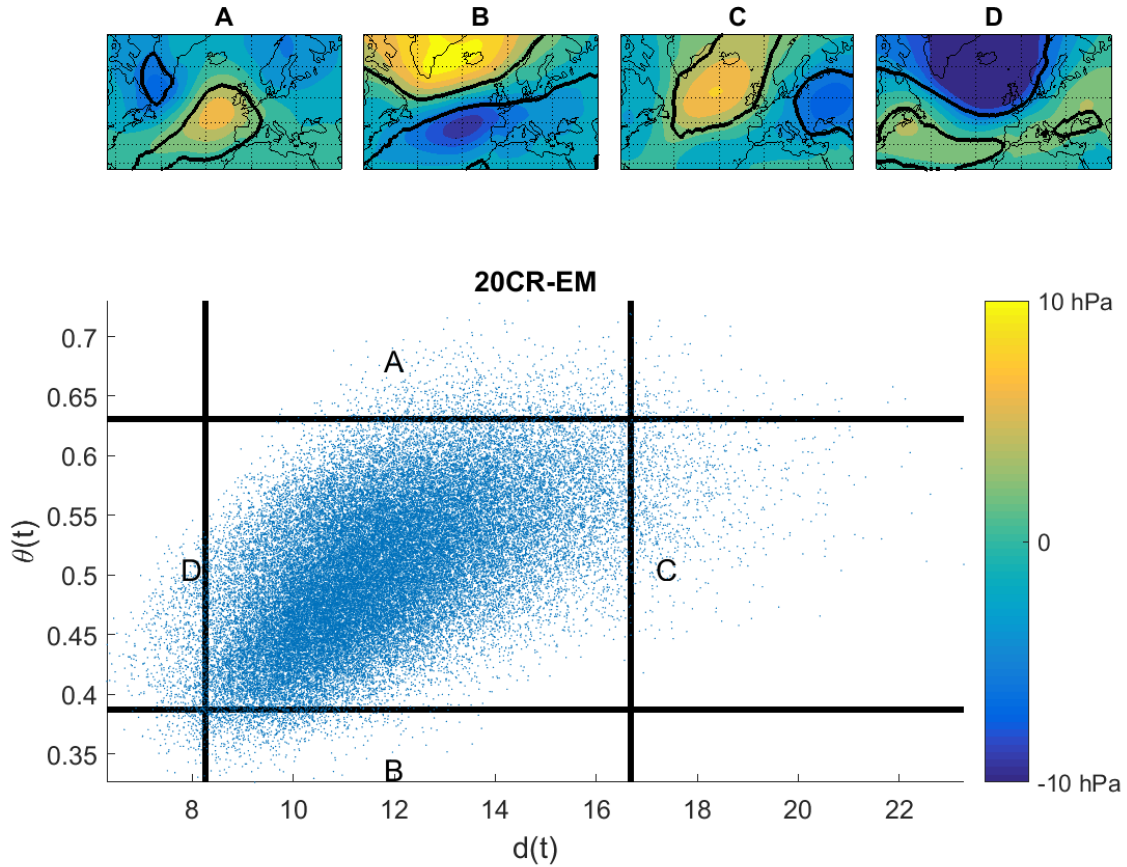


540 FIG. 2. Schematic representation of the dynamical indicators on fictive SLP maps. H indicates high and L  
 541 low pressure systems; the arrows mimic horizontal wind fields. The local dimension  $d$  is proportional to the  
 542 number of possible configurations preceding and following the day analyzed (here  $d = 2$ ).  $\theta$  is the inverse of  
 543 the persistence time of a given configuration. If the pattern persist for three days (red path), then  $\theta = 1/3$ . If the  
 544 patterns change every day (green path), then  $\theta = 1$ .

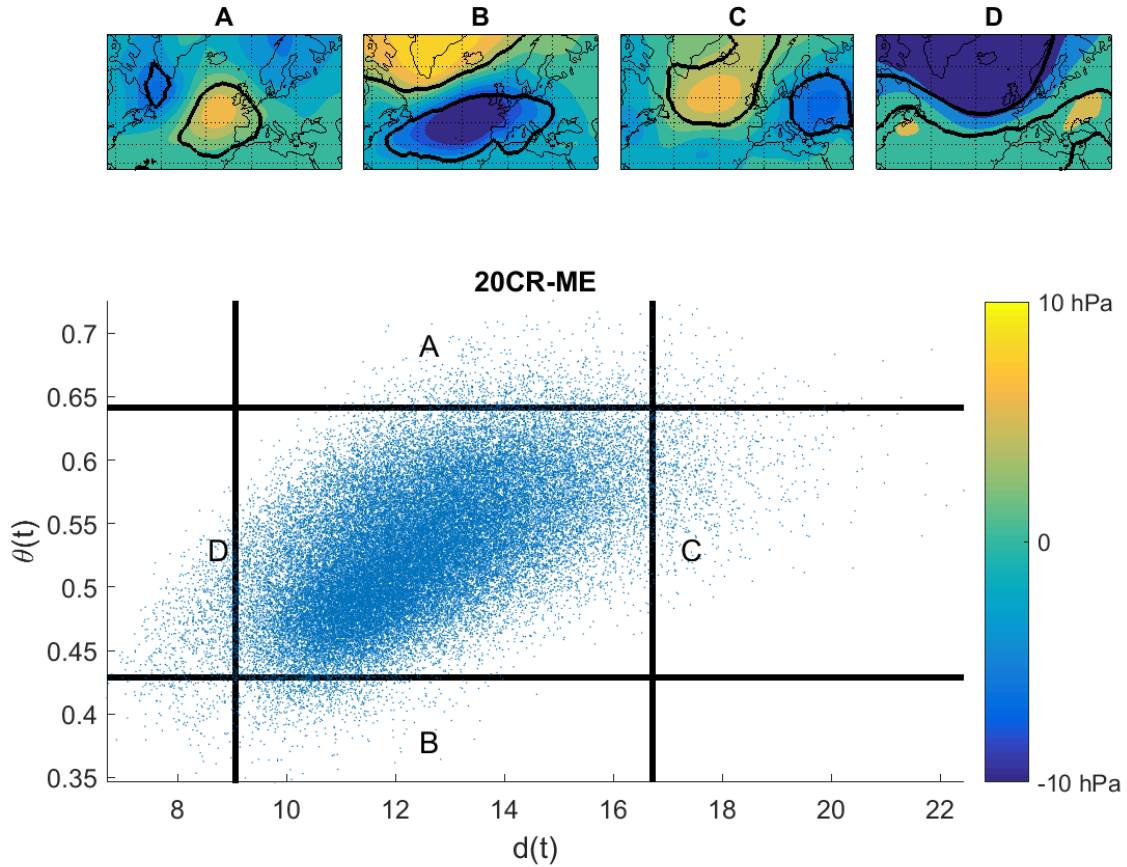


545 FIG. 3. Medians of  $d$  and  $\theta$  (dots) and standard deviation (semiaxes of the ellipses) for 30 realizations of the  
 546 Lorenz attractor with  $\sigma$  varying continuously over  $28 < \sigma < 28.54$ . The blue markers correspond to the first half  
 547 of each simulation  $\sigma_1$ ; the red markers to the second half  $\sigma_2$ . The asterisks mark the medians of the ensemble of  
 548 realizations.

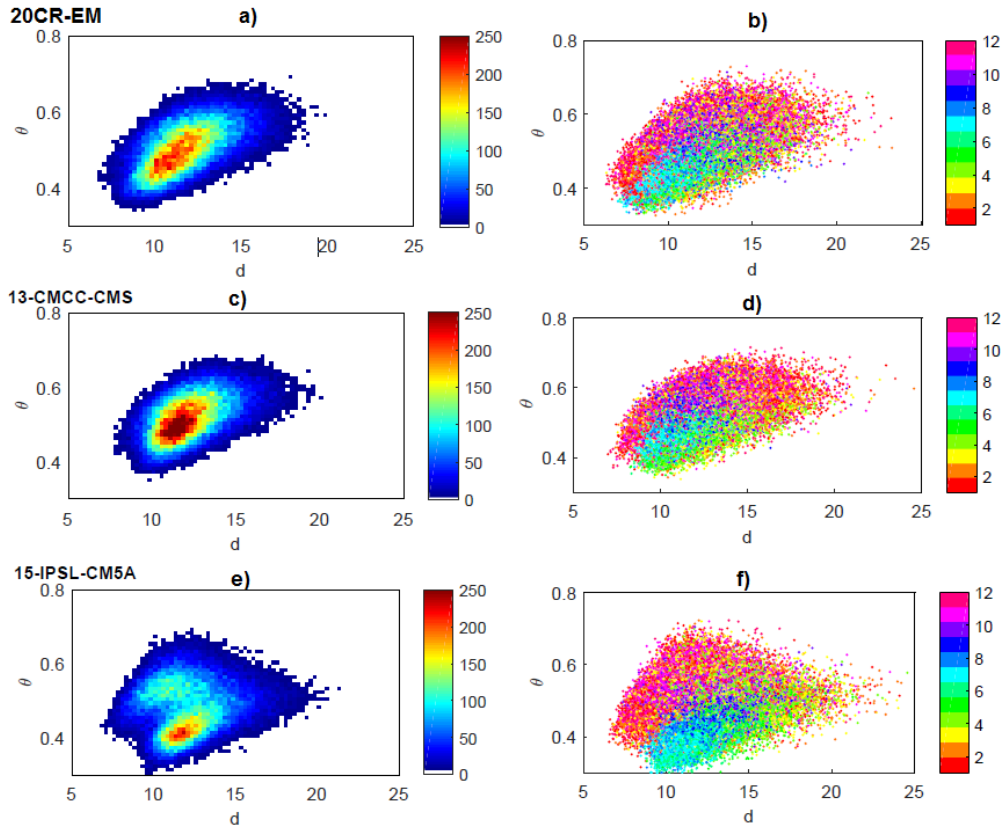




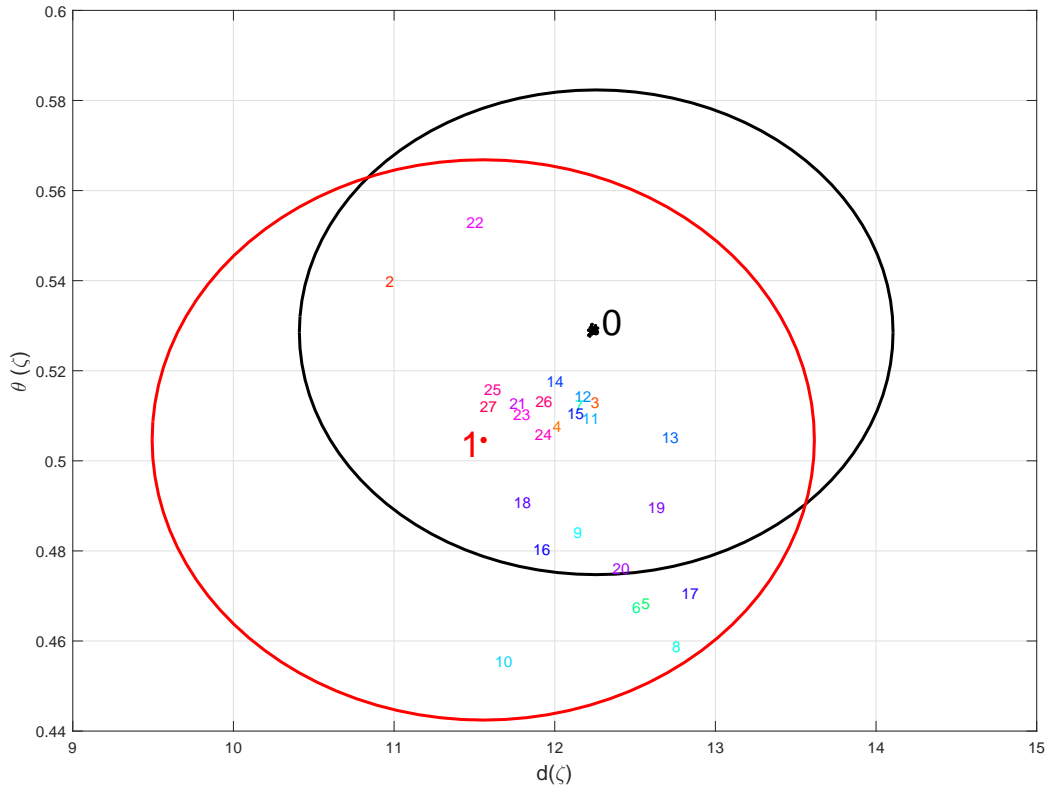
549 FIG. 4. Scatterplot of the daily values of instantaneous dimension  $d$  and inverse persistence  $\theta$  for the 20CRv2c  
 550 ensemble mean SLP (20CR-EM). The straight black lines mark the 0.02 and 0.98 quantiles of  $d$  and  $\theta$ . The  
 551 composite anomalies in SLP obtained averaging the days beyond the quantiles can be associated with known  
 552 weather regimes: A) Atlantic Ridge (maxima of  $\theta$ ), B) NAO- (minima of  $\theta$ ), C) Blocking (maxima of  $d$ ), D)  
 553 NAO+ (minima of  $d$ ). The black contours in panels A-D indicate regions where at least 2/3 of the composite  
 554 members display sea-level pressure anomalies of the same sign.



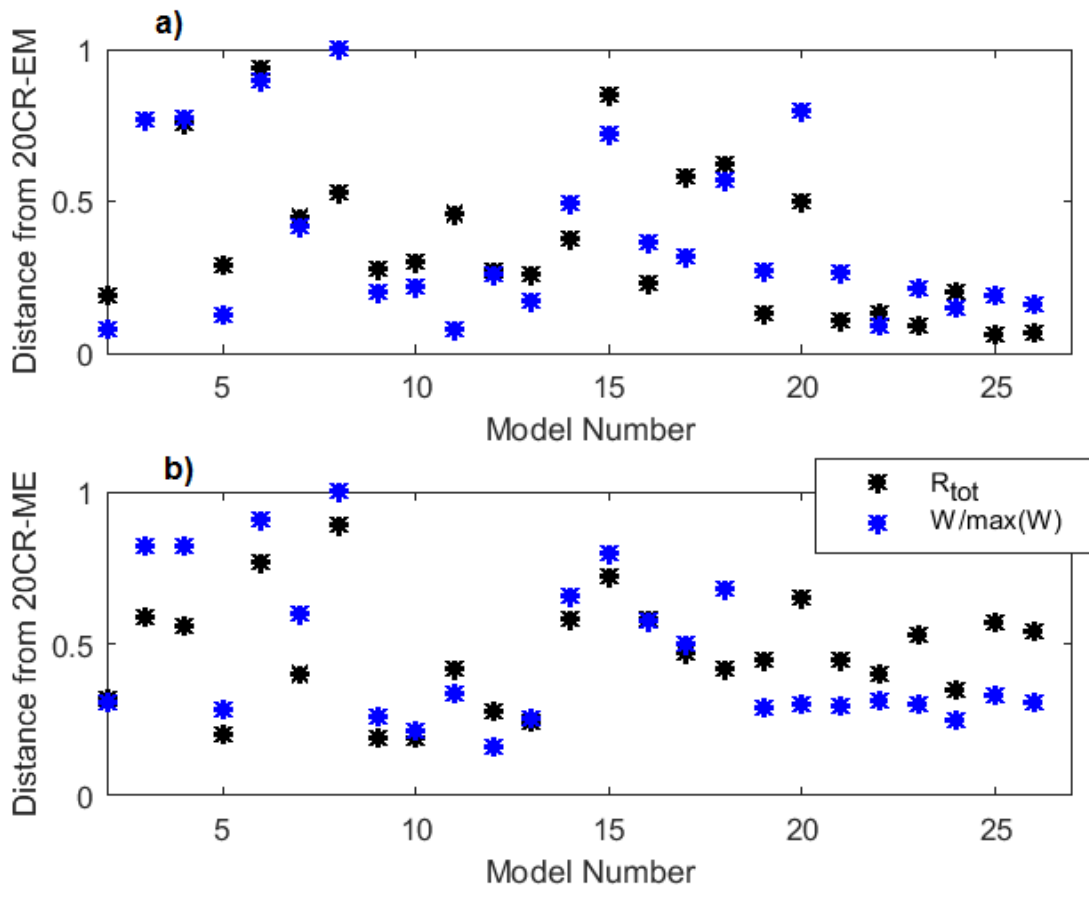
555 FIG. 5. Scatterplot of the daily values of instantaneous dimension  $d$  and inverse persistence  $\theta$  for the SLP  
 556 fields of the 56 individual 20CRv2c members (20CR-ME). The straight black lines mark the 0.02 and 0.98  
 557 quantiles of  $d$  and  $\theta$ . The composite anomalies in SLP obtained averaging the days beyond the quantiles can be  
 558 associated with known weather regimes: A) Atlantic Ridge (maxima of  $\theta$ ), B) NAO- (minima of  $\theta$ ), C) Blocking  
 559 (maxima of  $d$ ), D) NAO+ (minima of  $d$ ). The black contours in panels A-D indicate regions where at least 2/3  
 560 of the composite members display sea-level pressure anomalies of the same sign.



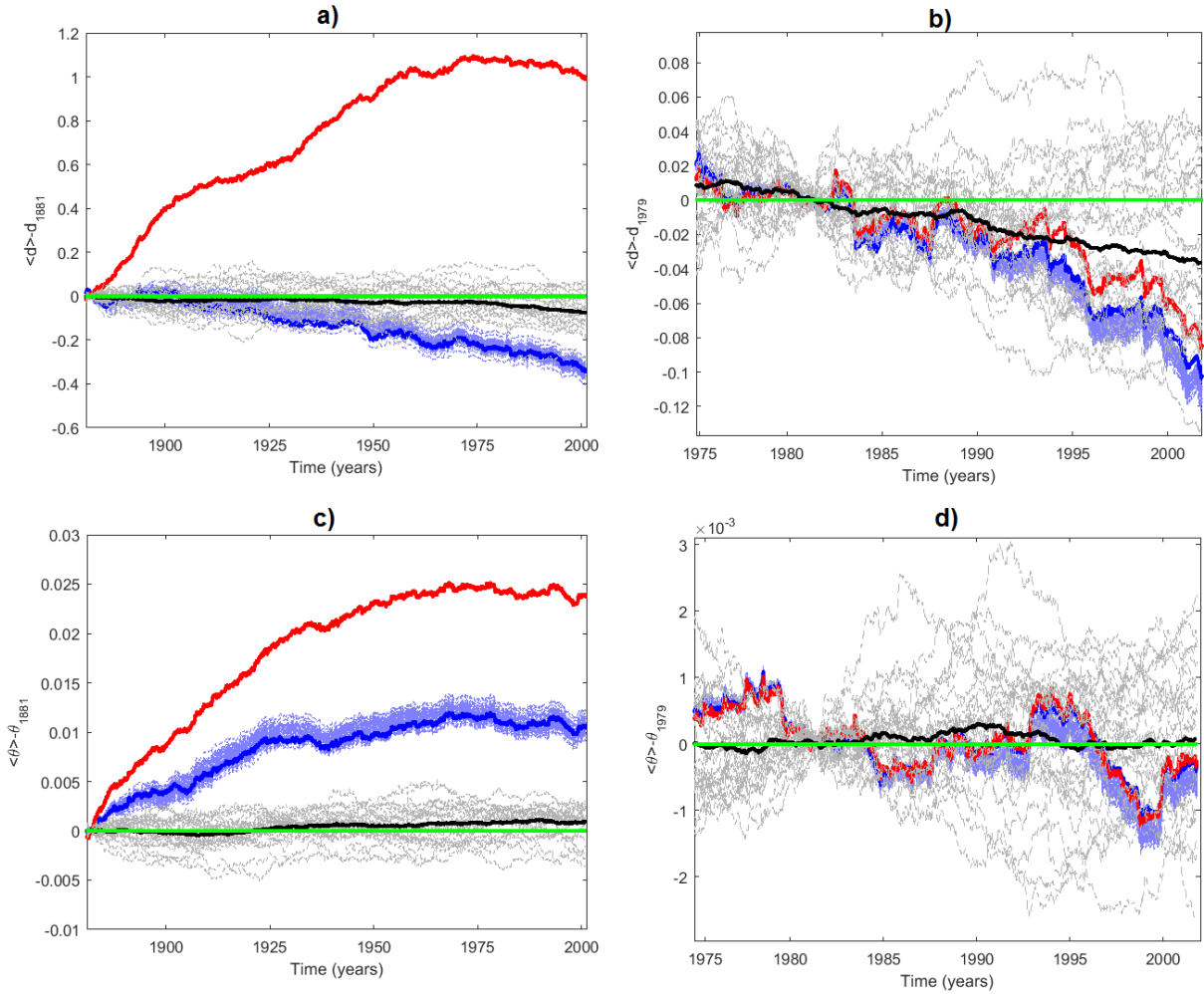
561 FIG. 6.  $(d, \theta)$  bivariate histograms (a,c,e) and scatter-plots (b,d,f) for the 20CR-EM reanalysis (a,b), CMCC-  
 562 CMS (c,d) and IPSL-CM5A models (e,f). The color scales in (a,c,e) indicate the frequency of observations in  
 563 number of days. The color scales in (b,d,f) indicate the month of the observation and show the dependence of  
 564 the  $(d, \theta)$  diagrams on the seasonal cycle.



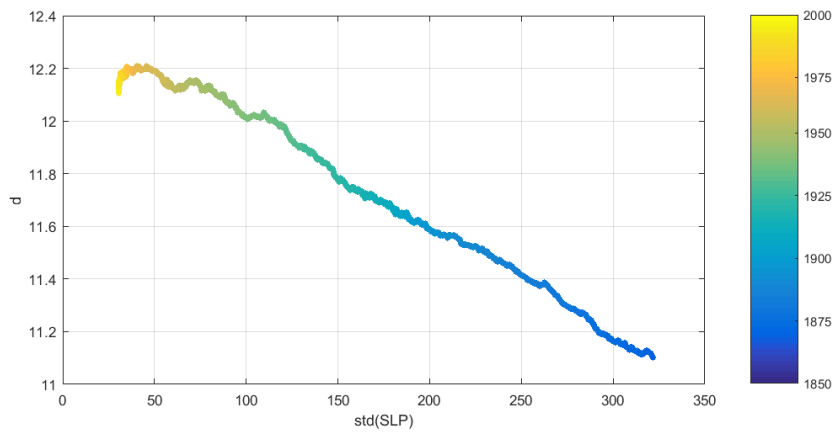
565 FIG. 7. Comparison between the 56 20CR-ME median values of  $(d, \theta)$  (black points whose average is denoted  
 566 by 0), the 20CR-EM (in red and numbered by 1) and all the CMIP5 models (progressive numbers 2-27, see table  
 567 1 for details). The semiaxes of the two ellipses represent one standard deviation of  $d$  and  $\theta$  for 20CR-EM (red)  
 568 and 20CR-ME (black).



569 FIG. 8. Comparison between  $R_{tot}$  values (black) and Wasserstein distances  $\mathcal{W}$  (red) of  
 570 20CR-EM (a) and 20CR-ME (b) and of the CMIP5 models. The distances are normalized by the maximum  
 571 value.



572 FIG. 9. 30-year moving averages  $\langle \cdot \rangle$  of instantaneous dimension  $d$  and inverse persistence  $\theta$  minus the re-  
573 spective mean values over 1851-1880 (a,c) and 1950-1979 (b,d). Note that labels on the abscissa mark the last  
574 year in each 30-year averaging window. (a,b): local dimension  $d$ ; (c,d): inverse persistence  $\theta$ . Red: 20CR-EM;  
575 Blue: 20CR-ME; Light blue: single 20CRv2c members; Black: CMIP5 multimodel mean; Grey: single CMIP5  
576 models; Green: zero line.



577 FIG. 10. 30-year moving average of the standard deviation of SLP fields across the 20CRv2c ensemble  
 578  $\langle std(SLP) \rangle$  versus  $\langle d \rangle$  for 20CR-EM. The color scale shows time.