

Answer Letter: Minimal dynamical systems model of the northern hemisphere jet stream via embedding of climate data by Faranda et al.

esd-2018-80

Dear Editor,

We are pleased to resubmit a new version of our paper “Minimal dynamical systems model of the northern hemisphere jet stream via embedding of climate data” for consideration in ESD. We have undertaken the extensive changes recommended by the reviewers and yourself and rewritten/extended the paper to make it more readable to the climate scientists community. The specific changes are given in the answers below, however we would like to underline the major changes in this manuscript version:

- The introduction has been revised: a better overview of the existing literature on the jet dynamics is provided as well as a motivation for the use of low dimensional models. The added value of stochastic modelling is also underlined
- The methods section has been rewritten to be more readable
- The model section has been rewritten and each single term is motivated on physical basis and linked to the existing literature, as recommended by the reviewers and yourself.
- The validation of the model has been rewritten: we have looked in the phase space for the best parameters and presented indicators that are more readable for the climate community.
- Following the request to provide a rigorous introduction on coupled map lattices and the embedding procedure, we have added two appendices.

At the end of the referees answers we also provide a marked-up version of the manuscript, useful for review purpose. We hope that this new version of the article will be suitable for publication in ESD.

Best Regards,
Davide Faranda,
On behalf of the authors

Referee 1

General Comments:

The average functional form $f(x)$ and the three stochastic subgrid terms are purely heuristic so unlike the case of other reduction techniques and subgrid modeling the connection with the physics of the problem is unclear. How would the results change with different, perhaps more physical, subgrid terms? Without a physical basis for the driving terms it seems unlikely that the simple model will be seen as any more than a curve fitting exercise.

The presentation of the article is substandard and not in a form that would appeal to the audience of ESD. The paper lacks motivation, the mathematics is poorly presented with terms undefined and too

many typos and has the feel of a first draft. Perhaps unfortunately, the mathematical nomenclature for what are really very simple concepts (new words for old), would most likely put off an audience of largely data analysts. For this audience the authors should make the article more pedagogical and stand alone.

We thank the reviewer for this comment and we take into serious consideration the criticism that our paper should be better motivated. We disagree with the statement that the embedding methodology we apply amounts to curve-fitting (indeed, the only part of our analysis where this definition may be argued for is for obtaining the return map). This model is motivated by geometrical and temporal evolutions of the jet (in a reanalysis). Virtually any models, including conventional numerical simulations of large-scale atmospheric flows, require some arbitrarily chosen parameters, and our case is no different. We also underline that our coupled map lattice model rests on clear physical hypotheses, such as: 1) the eastward propagation of information within the jet stream, 2) the presence of anticyclones and cyclones (baroclinic activity) i.e. sinuosity of the jet, 3) the presence of geographical constraints, 4) small-scale turbulent disturbances. This again sets it apart from curve-fitting exercises. We indeed view our approach as complementary to other idealised approaches which have attempted to formalise atmospheric waves and sinuosity, such as that of Petoukhov et al. In contrast to the latter, our model does not rely on regular wave decomposition hypotheses that are difficult to verify in practice. We now address this aspect more directly in the introduction to clarify the motivations underlying our analysis.

However, we do agree with the Reviewer that our illustration of the physical principles underlying some of our choices were not as clear as they should have been. We now detail how our choices have solid physical underpinnings, issued from both laboratory tank experiments, numerical simulations in the literature and scale arguments applied to fundamental concepts in atmospheric dynamics. Concerning the physical processes specific to sub-grid scales, we highlight that a key advantage of the return plot methodology is that it enables us to ignore detailed microphysics, focusing instead on the largest scale effective dynamics which is observed in the real data. Indeed, such small-scale processes are only indirectly present in the data we use to build our model, through assimilation of observations, but would be largely parametrized in the numerical model underlying the reanalysis dataset. Finally, we would like to stress that our purpose is to develop a minimal model to study the phenomenology of the effective dynamics of the jet flow, emerging from the complex underlying physics. Such a model does not require physical sub-grid terms a priori, but only if they were found to be essential to capture the large-scale phenomenology – which we show is not the case. This approach is fundamentally different from Direct Numerical Simulation (DNS) based studies, which typically start from the detailed microphysics at the cost of not incorporating real large-scale data. The phenomenological properties of our model, such as its bifurcation structure, are largely independent of the selected parameters, except for a few leading terms such as κ , β , and ϵ . We thus believe that our approach is a valid complement to the classical DNS-like approaches.

Concerning the second part of the Reviewer's comment, we address this in more detail in the responses to the specific comments below. One point we would like to highlight is that we have taken very seriously the Reviewer's encouragement to refocus our article to appeal to ESD readership, especially when explaining the derivation and implementation of the model. To this effect, amongst

other changes we have added two appendices providing background on some key concepts leveraged in the paper: Appendix A: Coupled map lattice, and Appendix B: Average return map and noise.

Specific Comments:

P2, line8: Perhaps references to Charney and De Vore (1979) and Wiin-Nielsen (1979) would be appropriate.

We have added these references, as suggested.

Section 3: The mathematics is surprisingly poorly presented given that one of the authors is from a Department of Mathematics and Statistics. For example, you need to define n as the time step, i as the longitude and define $N=360$ when it first appears. You need to check your equations for typos as in equation (3). Also, the equations keep changing until you eventually settle on the system that you eventually address.

We will replace n with t , and now describe all variables and the full model at the beginning of Section 3. We will further include a brief background review on coupled map lattices (CMLs), to highlight why they are appropriate in our context, in addition to providing a more detailed background in Appendix A.

P4, lines 2&3: Northern hemisphere blocking occurs in preferred regions so why does the return map not reflect that?

The local topography is represented as small shifts in the return maps. These small differences in the maps induce a sudden large change of the dynamics, corresponding to shifting the jet towards the north or the south and therefore triggering blocking in selected regions. We omit detailed landscape factors such as high mountains in this particular model for studying global phenomenology. However, these may be included via the boundary condition $r^{\{i\}}$. We now show that the inclusion of $r^{\{i\}}$ is able to highlight preferred regions where the jet shifts towards northern or southern directions (Figure 1 of this answer). We further argue that it is a strength of the model that it reproduces jet-like phenomenology, independently of the choice of the location, and that geographic effects can be introduced via boundary conditions. This discussion will be added to the new version of the paper.

P4, lines 4-31: Why is necessary to have separate stochastic processes for the effects of (1) convection and gravity waves, (2) effects of topography and (3) effects of baroclinic Rossby waves, rather than combine the three?

Given the CML model, the external perturbations to each local dynamics are categorized as: (0) initial conditions (1) local noise, (2) spatial boundary conditions, and (3) global noise. In this study, we made models for factors (1), (2), and (3). Term (0), the initial conditions, is chosen from a stationary state in the model. These are based mainly on (1) effects of convection and gravity waves, (2) effects of topography, and (3) effects of baroclinic Rossby waves. In brief, we started with modeling

phenomenological external perturbations, and then verified the underlying physics which affects factors (1), (2), and (3). To answer the question of the Reviewer, we cannot combine the three terms into one because they act on different spatial scales. More specifically: without (1) the system will be stacked in only one of the three states with no transitions; without (2) the jet position will not have a geographical dependence (see again Figure 1) and would thus not match the patterns observed in the ERA-Interim data; without (3) there will not be persistent blocking. We will add to the paper the new Figure 2, which shows temporal and spatial cluster size distribution for different models and the data, once they have binarized as follows: “1” means a northern shift of the jet with respect to its central position, “0” means a southern shift. The different model runs show the effect of the suppression of noise terms. The figure clearly shows that by suppressing one of the noise ingredients, the spatiotemporal cluster distributions of the data cannot be reproduced. The motivation of our work is precisely to show that these three ingredients are essential to reproduce the features of the jet dynamics.

Also why are these parameterizations purely stochastic when more systematic subgrid parameterizations indicate that they should be represented by a combination of deterministic and stochastic terms (e.g., Kisios and Frederiksen 2018 and references therein). In general, the authors should relate their subgrid parameterizations at least in broad terms to physically based parameterizations.

We will add a discussion about the stochastic vs deterministic parametrization based on Kisios and Frederiksen 2018 and references therein. Our goal here is to have only the large scales described by a deterministic term, as we build a global model of the jet dynamics. Indeed, a mixture of deterministic and stochastic terms improve the dynamical description of the jet dynamics, but the addition of other terms will not make our model a minimal model of the jet dynamics.

P4, lines 14-19:

The impression that the authors convey here is that the topography is a stochastic term in their model in which case it should be multiplicative noise rather than additive noise. However, according to the above reference, deterministic topography interacting with eddies produces an additive noise contribution as well as contributions from barotropic and baroclinic Rossby waves.

In our model, the topography is given as a boundary condition and therefore is a deterministic term. We will rephrase the model description to say that it is included in the perturbations term, where the perturbations are split into deterministic (topography) and stochastic (turbulence and baroclinic waves) contributions.

P4, line 20:

baroclinic → baroclinic and barotropic

Corrected.

P4, line 21:

$10^{-3} \rightarrow 10^3$

Corrected.

P5, line 5:

What exactly is the form of the non-autonomous force? What is the explicit time dependence? You should define your terms for an audience of largely data analysts.

The model of the local dynamics at location i can be rewritten as:

$$x_{t+1}^{(i)} = f(x_t^{(i)}) + p_t^{(i)}.$$

The non-autonomous term $p_t^{(i)}$ includes all driving forces other than $f(x_t^{(i)})$ at position i , and its explicit time dependence cannot be given simply. Non-autonomous dynamical systems theory can be applied to dynamical systems with such “unknown” driving forces. Here, we approximated it as a random variable $p_t^{(i)}$ in $[-\kappa, \kappa]$, assuming a bounded external force, and analyzed the bifurcation structure of the approximated model. We clearly explain the above mathematical approach in the revised Section 3 in the manuscript.

Section 4:

Again the mathematics is poorly presented. I would expect precision and elegance from mathematicians. You will need to explain your terminology for the major audience of ESD. The authors need to carefully check their manuscript for a number of typos.

We will add the mathematical details on the dynamical indicators in two appendices mentioned in the reply to the Reviewer’s general feedback.

References: Kitsios, V., and J. Frederiksen, 2018: Subgrid parameterizations of the eddy-eddy, eddy-meanfield, eddy-topographic, mean field-mean field and mean field topographic interactions in atmospheric models. J. Atmos. Sci. doi:10.1175/JAS-D-18-0255.

Referee 2

The manuscript needs to be substantially improved before I can recommend it for publication. In particular, presentation of the CML model lacks clarity for general readership, as well as interpretation and significance of some results are overstated.

Both Reviewers have highlighted lack of clarity for ESD readership as a key shortcoming of our manuscript, and we have taken this comment very seriously. In addition to the changes detailed in the

replies to the individual comments below, we have added two appendices providing background on some key concepts leveraged in the paper: Appendix A: Coupled map lattice, and Appendix B: Average return map and noise.

Comments:

1. Please provide some background on CML and why it has been chosen for this study.

We will describe the full model at the beginning of Section 3, and further include a brief background review of CMLs, motivating their use here, as an appendix in the new version of the manuscript.

2. Please provide more mathematical details on return map in Section 3 and how it can be used to estimate $f(x)$.

We have added a clearer mathematical background on return maps in Section 3, and also highlight its advantages in the present context.

3. Why the particular form of Eq (4) is chosen and how these coefficients are estimated?

The particular form of equation 4 is chosen as the one best fitting the data and presenting a stable state around 0. We have tried other functional forms for the maps given in Eq. 4. An account of this will be given in Section 3 in the revised version of the study.

4. What about uncertainties in the model coefficients? Fig. 3 shows that red line (Eq.4) seem to be missing excursions that are very few to begin with.

The return map is obtained by adapting the model to the data. The phenomenological properties, such as bifurcation structure, are largely independent of the selected parameters. The excursions are modelled via the stochastic escape that will add fluctuations on top of the red line. We will explain this in the new version of the manuscript.

5. It is rather hard to follow the discussion of the stochastic noise terms and it leaves impression that they are tuned without much mathematical guidance.

In the new version we will explain the rationale behind: (1) local noise, (2) spatial boundary conditions, and (3) global noise, more clearly in the manuscript. We have further made an effort to highlight that these issue from both mathematical and physical considerations, which ground our model in both dynamical systems theory and atmospheric dynamics.

6. The Fig.7 comparison of summary statistics (ACF and PDF) for the optimal value of $\epsilon = 0.4$ does not show much qualitative agreement between the modeled and observed dynamics (also in P15 in conclusions). The space-time patterns also look visibly rather different. It makes look weaker the rest of results on bifurcation analysis and dynamical indicators.

Indeed, we realized how important it is to provide a quantitative characterization of the spatio-temporal properties of the model versus data. We proceed as follows: we binarize the data so that “1” is any shift towards the northern jet state and “0” is a shift towards the southern jet state. We then compute the time and spatial cluster size distributions for different models, including or not the noise terms and show their importance in matching the distribution of the jet position observed in the data. We will replace Figure 7 in the paper with the new Figure 2 here in which we show these analyses. The best model is now obtained for the parameters $\eta=1.2$, $\epsilon=0.33$, $r_{\text{oceans}}=0$, $r_{\text{mountains}}=0.02$,

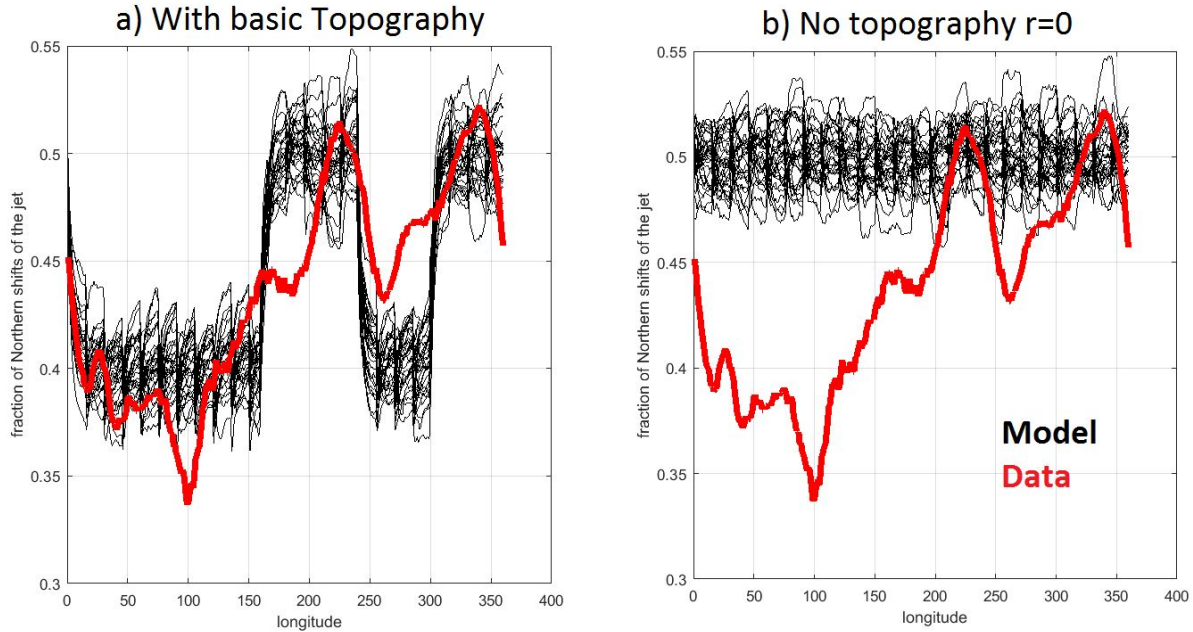


Figure 1: Role of the term r_i in the shift of the jet position towards northern or southern latitudes. The figures show the fraction of shifts towards the north: a value >0.5 indicates that the jet’s preferred position is to the north, a value <0.5 that the jet’s preferred position is to the south. a) Model with $r_i=0$ over oceans and $r_i=-0.02$ over the mountains (same domains as given in the previous version of the paper). b) Model with $r_i=0$ for all the latitudes. Red: shift frequency from data. Black: shift frequency from the model: each line corresponds to a realization of the system.

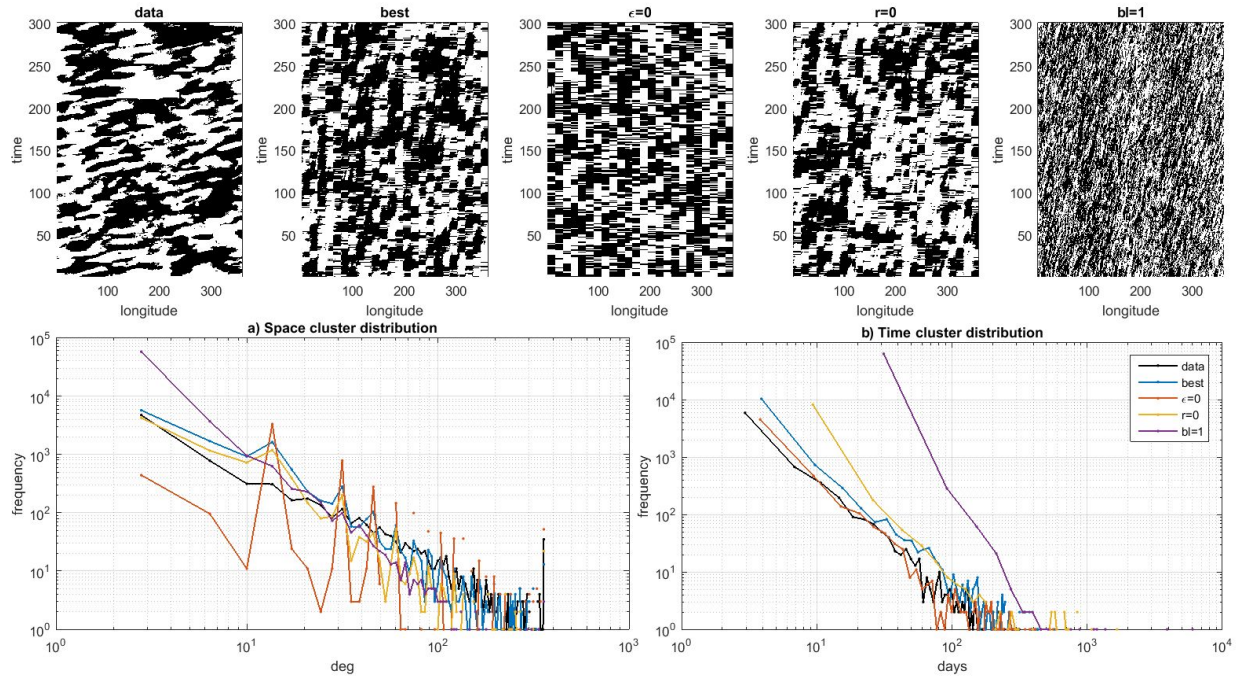


Figure 2: Upper plots: Temporal and spatial cluster size distribution for the ERA Interim data (top left), and few different model runs. The clusters are obtained once the data are binarized: 1 corresponds to a northern shift of the jet with respect to its central position; 0 corresponds to a southern shift. The different model runs show the effect of the suppression of noise terms. Lower plots: Space and time cluster distributions for ERA interim data (black) and different model runs (colors).

Minimal dynamical systems model of the northern hemisphere jet stream via embedding of climate data

Davide Faranda^{1,2}, Yuzuru Sato^{3,2}, Gabriele Messori^{4,5}, Nicholas R. Moloney^{6,2}, and Pascal Yiou¹

¹Laboratoire des Sciences du Climat et de l'Environnement, UMR 8212 CEA-CNRS-UVSQ, Université Paris-Saclay, IPSL, 91191 Gif-sur-Yvette, France

²London Mathematical Laboratory, 8 Margravine Gardens London, W6 8RH, UK

³RIES / Department of Mathematics, Hokkaido University, Kita 20 Nichi 10, Kita-ku, Sapporo 001-0020, Japan

⁴Department of Earth Sciences, Uppsala University, 752 33, Uppsala, Sweden

⁵Department of Meteorology and Bolin Centre for Climate Research, Stockholm University, 106 91, Stockholm, Sweden

⁶Department of Mathematics and Statistics, University of Reading, Reading RG6 6AX, UK

Correspondence: Davide Faranda (davide.faranda@lscce.ipsl.fr)

Abstract. We derive a minimal dynamical systems model for the northern hemisphere mid-latitude jet dynamics by embedding atmospheric data, and investigate its properties (bifurcation structure, stability, local dimensions) for different atmospheric flow regimes. ~~We derive our model according to the following steps: i) The derivation is a three-step process: first, we~~ obtain a 1-D description of the mid-latitude jet-stream by computing the position of the jet at each longitude using the ERA-Interim reanalysis. ~~ii) Next, we~~ use the embedding procedure to derive a map of the local jet position dynamics. ~~iii) Finally, we~~ introduce the coupling and stochastic effects deriving from both atmospheric turbulence and topographic disturbances to the jet. We then analyze the dynamical properties of the model in different regimes: ~~i) one that gives the closest representation of the properties extracted from real data, ii); one featuring a stronger jet (strong coupling), iii); one featuring a weaker jet (low coupling), iv) weak coupling); and one with modified topography. We argue that such a simple model provides a useful Our~~ model, notwithstanding its simplicity, provides an instructive description of the dynamical properties of the atmospheric jet.

Copyright statement. TEXT

1 Introduction

Jet streams are narrow, fast-flowing westerly air currents near the tropopause. They are a major feature of the large-scale atmospheric circulation and modulate the frequency, severity and persistence of weather events across the extra-tropics (e.g. Röthlisberger et al. (2016)). Their location and intensity also affects commercial aviation and shipping (Reiter and Nania, 1964; Hadlock and Kreitzberg, 1988; Williams and Joshi, 2013). Two types of atmospheric jets can be identified: thermally-driven subtropical jets, and eddy-driven jets associated with baroclinic instability at the polar front. In the Northern Hemisphere (NH), the two are not always clearly separated (Lee and Kim, 2003), and when considering monthly or longer ~~time averages~~

time-averages a single, spiral-shaped jet structure emerges (e.g. Archer and Caldeira, 2008). In this paper we consider a single NH jet (NHJ), rather than attempting to separate the subtropical and eddy-driven jets (e.g. Belmecheri et al., 2017).

Even though the climatological NHJ is a westerly flow, it can present large meanders on ~~a synoptic timescale~~ synoptic timescales (Koch et al., 2006; Röthlisberger et al., 2016). These can cause the local flow to become predominantly meridional, or can even determine a splitting or breaking of the jet (Haines and Malanotte-Rizzoli, 1991). The occurrence of these large meanders in the jet is often associated with ~~extreme climate~~ events such as temperature and precipitation extremes (e.g. Dole et al., 2011; Screen and Simmonds, 2014). Although jet dynamics are well understood in a climatological sense, our ~~understanding of~~ insights into dynamical features such as jet ~~breaking and meandering is still incomplete~~ splitting or meandering are still limited.

The dynamics of meanders and ~~discontinuities in the jet~~ split jets has often been framed in terms of transitions between zonal and blocked flows since the seminal work by Charney and DeVore (1979). Legras and Ghil (1985) and Ghil (1987) used an intermediate complexity barotropic model with dissipation forcing and topography, and observed two distinct equilibria associated with the zonal and blocked flows. Similar mechanisms have been proposed by Mo and Ghil (1988) using experimental facilities (Weeks et al., 2000). However, there is no consensus about the nature of flow multistability, and a wide range of theoretical explanations and models have been proposed (e.g. Tung and Lindzen, 1979; Simmons et al., 1983; Frederiksen, 1982; Faranda et al., 2016b). Moreover, jet ~~breaking has~~ dynamics have been described as a manifestation of multiple equilibria in asymmetrically forced flows (Hansen, 1986) or as a result of ~~soliton~~ soliton-modon structures (McWilliams et al., 1981).

~~Here, the focus is not on the persistence and transitions between zonal and blocked regimes. Rather, we propose an alternative framework to diagnose the instantaneous meridional location of~~ In order to progress in our fundamental understanding of the jet dynamics, we employ a low-dimensional dynamical systems model derived from reanalysis data. The best-known example of a low dimensional model for atmospheric phenomena is Lorenz' simple three-dimensional system representing some features of Rayleigh-Bénard convection (Lorenz, 1963). Thereafter, simple dynamical systems models have been devised to study El Niño (Penland and Matrosova, 1994), ocean-atmosphere interactions (Dijkstra and Ghil, 2005), climate tipping points (Stommel, 1961; Benzi et al., 1982), large scale atmospheric motions (Lorenz, 1984, 1996) and many other problems. The goal of these investigations was not to provide the most realistic representation of the relevant systems, but rather to capture key emerging behaviour (such as chaos, intermittency, multistability). The main drawback of those investigations was the lack of connection between models and real data due to the ~~jet from data and model this location using a simple stochastic coupled map lattice. We follow the approach outlined by Faranda et al. (2017c) to build~~ scarcity of observations as well as theoretical limitations. Until very recently, there was a strong case against the use of embedding techniques to derive low dimensional models from experimental data (Letellier et al., 2006). This opposition was motivated by a long sequel of papers that appeared between 1984 and 1991. The initial claim (see e.g. Nicolis and Nicolis (1984); Fraedrich (1986)) that low dimensional models for complex phenomena could be derived using a very small numbers of variables was disrupted by rigorous numerical computations by Grassberger (1986) and Lorenz (1991).

Progress in data quality and availability and the advent of stochastic dynamical systems have renewed the attention for data embedding. Recently, Faranda et al. (2017c) have shown that embedding techniques can yield effective low-dimensional dynamics provided that the chosen observables reflect the symmetries of the system and that small-scale (subgrid) dynamics

is lumped in stochastic perturbations. Here, we use these results to develop a minimal model ~~for the jet position by embedding atmospheric data obtained from the ERA-Interim dataset. This model~~ of the effective dynamics of the mid-latitude jet. This is useful to explore a range of possible behaviors beyond those displayed in the available data, that could have appeared in past climates and could appear again in future climates. In analogy to the model derived in Faranda et al. (2017c) for the von Karman turbulent flow, the jet model is based on a coupled map lattice (CML, see also Appendix A). Each element of the lattice reflects the dynamics of the jet at each longitude. ~~This CML is stochastically perturbed to account for baroclinic eddies and turbulence. The goal of the paper is to~~ Such a model does not require physical sub-grid terms a priori, but only if they are found to be essential to capture the large-scale phenomenology — which we show is not the case. We then evaluate how this model represents the key dynamical features of the jet, namely its stability, the statistics of splitting/breaking and the response to geographical features, and relate the results back to the original ERA Interim data.

The paper is organized as follows: we first give ~~First, we provide~~ the details of the ERA-Interim data and of the jet detection algorithm (Section 2). We then present the stochastic coupled lattice map model and compute its bifurcation structure (Section 3). ~~We then~~ Next, we introduce some instantaneous dynamical indicators (Section 4) and use them to relate the conceptual model to state-of-the-art global more complex climate models and reanalysis data (Section 5). Finally, we highlight the open questions our results can answer and the new questions they pose (Section 6).

2 ERA-Interim data ~~Data~~ and jet position algorithm ~~Methods~~

2.1 ERA-Interim data and jet position algorithm

The analysis is based on the ~~ERA-Interim reanalysis of the~~ European Centre for Medium Range Weather Forecasts' ~~ERA-Interim reanalysis~~ Dee et al. (2011). We consider daily data with a 1° horizontal resolution over the period 1979–2016.

The jet position is diagnosed through a modified version of the approach ~~of by~~ Woollings et al. (2010). We take daily mean wind-speed averaged over (200–400 hPa) and apply a 10-day low-pass Lanczos filter (Duchon, 1979). We then identify the latitudinal position of the jet at every longitude as the location of the strongest wind, over the band 15° – 75° N. This approach is intended to capture a "raw" measure of the jet variability. We then consider the longitude and time dependence of the latitude of the jet. ~~This allows us to monitor the waviness of the jet to monitor its waviness.~~

We define an index of large jet meanders, or breaks, (Breaking Index, *BRI*) ~~by, as~~ the daily number of meridional variations in jet position of more than 10° of latitude across adjacent longitude gridpoints, except at longitude 0. The analysis presented in this paper has been repeated for meridional variations between 5° and 15° , with no significant qualitative differences.

~~A Figure 1 shows a~~ snapshot of the jet position, ~~obtained for on~~ Feb 04th 1979, ~~is shown in Figure 1,~~ together with the time series of the daily jet position recorded in 1979 at longitude 120° W. An animation of the jet location for the year 1980 is provided ~~in the as~~ supplementary video. Both the time series and the snapshot show jumps in the jet position, reflected in the *BRI*. ~~A qualitative analysis of the jet position data suggests that the jet fluctuates around a central latitude (Central Jet, CJ) and seldom shifts to more northerly (NJ) or southerly (SJ) latitudes.~~

In order to embed the data and derive the effective maps of the dynamics, we remove the seasonal cycle from the data by subtracting, longitude by longitude, the average meridional position for each calendar day and dividing by the standard deviation. For the deseasonalized data, the dimensionless threshold for the computation of *BRI* corresponding to about 10° latitude is $1.2|x| > 1$.

5

3 Derivation of lattice jet model

2.1 Local dynamical systems metrics

Our analysis leverages two recently developed dynamical systems metrics, namely: the local dimension of the attractor d and the stability of phase-space trajectories θ^{-1} . Instantaneity in time corresponds to locality in phase-space, such that a value of d and θ can be computed for a given variable (in our case the jet position data) at every timestep. d is a proxy for the system's active number of degrees of freedom. It provides information on how the system can reach a given state, and how it can evolve from such state. θ^{-1} describes the persistence of a state in time, thus providing complementary information to d .

10

~~We focus on the local dynamics divided into 360 cells, assuming that the dynamics in each cell is almost autonomous, but perturbed by the external environment, i. e., the adjacent cells. The spatial continuity of the jet stream is modeled via the coupling of the local dynamics. The local dynamics are represented as a discrete map and the jet dynamics as a coupled map lattice (CML) Kaneko (1983). A diffusively coupled map is given by:-~~

15

2.1.1 Local Dimension

The local dimension is estimated by making use of extreme value statistics applied to recurrences. The Freitas et al. (2010) theorem and its modification by Lucarini et al. (2012), states that the probability of entering a ball with a small radius centred on a state ζ on a chaotic attractor obeys a generalized Pareto distribution (Pickands III, 1975). In order to compute this probability empirically, we first calculate the series of distances $\text{dist}(x(t), \zeta)$ between the point on the attractor ζ and all other points $(x(t))$ on the trajectory. This series is transformed via the distance function:

20

$$g(x(t)) = -\log(\text{dist}(x(t), \zeta)), \tag{1}$$

such that nearby recurrences to ζ correspond to large values of $g(x(t))$ (Collet and Eckmann, 2009). Thus, the probability of entering a small ball at ζ is transformed into the probability of exceeding a high threshold $s(q)$, where q is a percentile of

25

the series $g(x(t))$ itself. In the limit of an infinitely long trajectory, it can be shown that the choice $g(x(t))$ in Eq. (1) locks this probability into the exponential member of the generalized Pareto distribution:

$$x_{n+1}^{(i)} = \text{Pr}(1 - \epsilon) f z \geq s(x_n^{(i)} q) + \frac{\epsilon}{2} \simeq \exp \left[\frac{f - \vartheta(x_n^{(i-1)} \zeta) + f(x_n^{(i+1)})}{\beta(\zeta)} \left(\frac{z - \mu(\zeta)}{\beta(\zeta)} \right) \right], \quad (2)$$

where $i = 1, 2, \dots, N$, $\epsilon \in [0, 1]$ and $x_n^{(i)} \in \mathbb{R}$. For the jet dynamics, we adopt the open flow model with uni-directional coupling (Kaneko, 1985). We also include additive noise $\xi_n^{(i)}$, which perturbs the local dynamics $\{x_n^{(i)}\}$ at each cell according to:

$$x_{n+1}^{(i)} = (1 - \epsilon) f(x_n^{(i)}) + \epsilon f(x_n^{(i-1)}) + \xi_n^{(i)}.$$

Here, we assume that the local dynamics are mostly affected by (upstream) coupling via the left-hand cell, with weak noisy perturbations. Periodic boundary conditions are applied with $N = 360$

where where $z = g(x(t))$, and μ and β (obtained via fitting) depend on the point ζ . These are the location and the scale parameters of the distribution. Remarkably, $\beta(\zeta) = 1/d(\zeta)$, where $d(\zeta)$ is the local dimension around the point ζ . This result has been recently applied to a range of atmospheric and oceanic fields (e.g. Faranda et al., 2017b, a; Messori et al., 2017; Faranda et al., 2019). In this paper, we use the quantile 0.975 of the series $g(x(t))$ to determine q . Our results are robust with respect to reasonable changes in this quantile.

To extract the local dynamics, we construct an average return map. We first coarse-grain the state space into 500 partitions with median $\bar{x}_n^{(i,k)}$, where $k \in \{1, \dots, 500\}$.

2.1.2 Local Persistence

Extreme value statistics also provide a way of estimating the persistence of a given state ζ , by inspecting the temporal evolution of the dynamics around ζ . A measure of persistence around ζ can be obtained from the mean residence time of the trajectory within the neighborhood of ζ . To measure this quantity, we employ the so-called extremal index ϑ (Freitas et al., 2012; Faranda et al., 2016a): an adimensional parameter $0 < \vartheta(\zeta) < 1$ which can be interpreted as the inverse of the mean residence time. We can then compute $\theta^{-1}(\zeta) = dt/\vartheta(\zeta)$, where dt is the timestep of our data. Heuristically, if the i th visit to the neighbourhood of ζ lasts τ_i consecutive time steps, and N such visits are made in total, then $\theta^{-1} \approx (1/N) \sum_i \tau_i$. In practice, instead of this naive estimator, we compute the extremal index using the likelihood estimator of Süveges (2007). $\theta = 0$ corresponds to a stable fixed point of the dynamics so that the trajectory resides an infinite amount of time in the neighbourhood of ζ . $\theta = 1$ corresponds to residing in the neighbourhood of ζ for only one time step per visit. The estimate of θ is thus sensitive to the dt used. If dt is too large, the time dependence structure is unresolved and θ will be close to 1. Conversely, if dt is too small, θ is close to zero. In Faranda et al. (2017b) it has been observed for sea-level pressure fields over the North Atlantic that θ varies between 0.3 and construct a return map \tilde{f} via the first return plot of $(\bar{x}_n^{(i,k)}, \bar{x}_{n+1}^{(i,k)})$ as follows:-

$$\langle \bar{x}_{n+1}^{(i,k)} \rangle = \tilde{f}(\bar{x}_n^{(i,k)}), \quad k \in \{1, \dots, 500\}.$$

From the observed data, the average return map $\tilde{f}(x)$ is approximated as

$$\tilde{f}(x) = \begin{cases} -\frac{A(A+x)}{A-c}, & x < -c, \\ \sinh(\beta x), & -c \leq x \leq c, \\ \frac{A(A-x)}{A-c}, & c < x, \end{cases}$$

where $\beta = 0.75$, $A = 3$, and $c = \sinh^{-1}(A)/\beta \approx 2.4246$, which is estimated from the extracted average return map. In 0.5, when $dt = 1$ day. In this work, we use the same dt .

5 3 Derivation of the lattice jet model

3.1 Model Framework

While not directly issued from the Navier-Stokes equations, our framework builds on concrete physical hypotheses, namely that: (i) the physics of the jet is the region where $|x| > c$, we have linear reflection effects. The functional forms of the return map are approximately the same at each longitude every longitude and it is only slightly modified by the presence of geographical constraints, (ii) the jet can experience sudden breaks and shifts from its central position (central jet, CJ) to northern (northern jet, NJ) or southern latitudes (southern jet, SJ), (iii) the jet must propagate to the west, (iv) smaller scale phenomena, such as turbulence and baroclinic waves, will be introduced in the model only if necessary to reproduce the effective dynamics in the data. This latter point is fundamentally different from the philosophy of direct numerical simulations.

The noise ξ is a fundamental ingredient for the breaking of We construct our model starting from the local time series of the jet and the transition between zonal and blocked states, as has also been shown through tank experiments and numerical simulations Jacoby et al. (2011). Indeed, in the absence of noise, the purely deterministic map results in a jet that settles on a central jet position (CJ) $x = 0$ non-dimensionalised jet position x measured at each longitude i and time t . Here we use the simplest possible embedding procedure (see Appendix B), which consists in plotting the return map $x_t^{(i)}$ vs. $x_{t+1}^{(i)}$ (an example is shown in Figure 2) and searching for a function $f^{(i)}$ such that $x_{t+1}^{(i)} = f^{(i)}(x_t^{(i)})$. The first thing to verify is that the same function form $f^{(i)}$ can be used at all the longitudes i . This is equivalent to asking that there is only one dynamics driving the jet independently of the location. With the choice:

$$f^{(i)}(x) = \begin{cases} -\frac{A(A+x)}{A-c} + r^{(i)}, & x < -c, \\ \sinh(\beta x) + r^{(i)}, & -c \leq x \leq c, \\ \frac{A(A-x)}{A-c} + r^{(i)}, & c < x. \end{cases} \quad (3)$$

where we have dropped the dependencies of x for clarity, the parameters can be fixed at all longitudes as: $\beta = 0.75$, $A = 3$, and $c = \sinh^{-1}(A)/\beta \approx 2.4246$. Even though the functional form of $f^{(i)}$ is independent of longitude, a dependence on i remains in the form of the parameter $r^{(i)}$, which represents the effects of topography in terms of spatial inhomogeneities of the local dynamics. As first order approximation, we consider only the difference between land and ocean, and assign one of two discrete

values to each $r^{(i)}$. The choice of the function $f^{(i)}$ is not unique; however the one we propose here is a suitable option that satisfies hypotheses (i) and (ii) above. In order to reproduce the eastward propagation of the jet (iii), we introduce the coupled map lattice (CML, see Kaneko (1983) and Appendix A):

$$x_{t+1}^{(i)} = (1 - \epsilon)f^{(i)}(x_t^{(i)}) + \epsilon f^{(i-1)}(x_t^{(i-1)}), \quad (i = 1, 2, \dots, N; t = 1, 2, \dots). \quad (4)$$

5 With this geometry, the dynamics is divided into $N = 360$ cells. Periodic boundary conditions are applied at $N = 360$. The dynamics in each cell i is time-independent, but perturbed by the cell $i - 1$ (i.e. its neighbour to the west) with intensity ϵ , which we estimate and scale based on the observed data. This further implies that our reference length-scale in the model corresponds to that of 1° longitude in the mid latitudes, namely order 100km.

10 3.2 Subgrid feedbacks to jet dynamics

If we perform a numerical simulation of Eq. (4), the dynamics is fixed to one of the three states, depending on the value of ϵ . This means that the role of small scales perturbations in triggering the transitions between CJ, SJ and NJ is fundamental. We therefore have to include an additive noise term $\xi_t^{(i)}$ in Eq. (4):

$$x_{t+1}^{(i)} = (1 - \epsilon)f^{(i)}(x_t^{(i)}) + \epsilon f^{(i-1)}(x_t^{(i-1)}) + \xi_t^{(i)}, \quad (i = 1, 2, \dots, N; t = 1, 2, \dots). \quad (5)$$

15 The noise is a fundamental ingredient for the breaking of the jet and the transition between zonal and blocked states, as shown in tank experiments and numerical simulations (Jacoby et al., 2011). ~~But with the addition of noise, jumps are observed between a northern jet position (NJ) $x \simeq +2$, a central jet position (CJ) $x = 0$, and a southern jet (SJ) position $x \simeq -2$. Physically, noise arises from key sub-grid processes that affect the jet dynamics, such as convection or the interaction between the jet stream and gravity waves. These have been recognized as fundamental ingredients of jet breaking in both numerical~~
20 ~~simulations and experimental studies Williams et al. (2003, 2005). (Williams et al., 2003, 2005).~~ Translated to our model ~~, with reference spatial scale of order 100km~~ these phenomena, ranging from a few meters to a few kilometers, ~~implies imply~~ a perturbation in the range $10^{-4} < \nu < 10^{-3}$, ~~where ν is a random variable draw from a uniform distribution in $[-\delta, \delta]$. However, the model only exhibits jet breaking when $\delta \gtrsim 0.4$, which is too large to represent realistic corrections to the jet dynamics. We are therefore led to introduce further noise perturbations.~~ Several subgrid parametrization of turbulence exists: the seminal works
25 of Kraichnan (1961) and Thomson (1987) showed that if large scales are represented by a deterministic term, a single random variable can drive the turbulence term. This means that Langevin model representations are appropriate to describe turbulent eddies (McComb, 1990; Frederiksen and Davies, 1997). Following these ideas, we model $\nu_t^{(i)} \in [-\delta, \delta]$ as an uniform random variable.

30 ~~Another ingredient of~~ However, considering small scale turbulent disturbances to the jet dynamics is ~~the presence of topographic obstacles to~~ not sufficient to reproduce the blocking and breaking of the jet. Even if the introduction of ν

as a stochastic term can account for the direct Kolmogorov turbulent cascade (Kolmogorov, 1941), the mid-latitude zonal flow. Mountain ranges and land-sea boundaries cause meridional deviations in the mean jet location (Tibaldi et al., 1980). This inhomogeneity can be modeled via a parameter r that mimics “spatial noise.” Since the topography is at most a few kilometers in height, this translates to a perturbation of the order of 10^{-3} in the model. Reasonable geographical constraints are therefore: $r^{(i)} = 0.02$ ($i \in \text{land}$) and $r^{(i)} = -0.04$ ($i \in \text{ocean}$), where land spans the ranges $0 \leq i < 161$ and $239 \leq i < 301$, and ocean spans the ranges $161 \leq i < 239$ and $301 \leq i < 360$. Other parameter values are presented and tested in the next section. Jet dynamics is also driven by the effect of an inverse cascade transferring energy to large scales via baroclinic waves (Held and Larichev, 1996).

A further important ingredient is the baroclinic activity. Baroclinic activity is associated with extra-tropical cyclones and anticyclones. These atmospheric features, on scales of order 10^{-3} km, 10^3 km. These can affect the jet position by several degrees of latitude. In our model, Again, there is not a unique way to model baroclinic waves in our framework. We follow the rationale of multiscale parametrizations as they can be theoretically justified (e.g. Wouters and Lucarini (2013); Kitsios and Frederiksen (2018)) and are numerically efficient (Faranda et al., 2014). The simple introduction of another source of noise $\eta^{(i)}$, acting at intermediate scales (i.e. between the scale of the effect of baroclinic waves is introduced through the block noise η^{bl} , where η takes jet and the scale of turbulence), is enough to obtain a reliable jet breaking dynamics (see Sect. 4.1). The simplest choice for $\eta_t^{(i)} \in [-\mu, \mu]$ is a block noise taking the same value over bl blocks with an amplitude of order 1 cells (the one-dimensionalized size of cyclones/anticyclones). To determine a realistic length for bl , we reason as follows: if $N = 360$, each grid cell is about 100km wide. Assuming a typical scale of about 3000 km, the extra-tropical cyclones are therefore ≈ 30 blocks wide. However, the perturbations are associated with the cyclone radius rather than diameter: upstream of the cyclone, the jet will mostly be deviated southwards, while downstream of the cyclone, the jet will mostly be deviated northwards. We therefore take the block perturbation to be of size $15 < bl < 20$ blocks. In conclusion, the noise term in our model is comprised of, see also Appendix B, and obeying the uniform distribution. Another choice for modelling baroclinic disturbances to the jet could be to introduce a second deterministic equation, weakly coupled with the jet position. However, this choice requires additional hypotheses and parameters and does not emerge naturally from the embedding procedure used to derive the dynamics of x .

The minimal sub-grid parametrization can thus be written in the form:

$$\xi_{nt}^{(i)} = \nu^{(i)} \underline{n}_t + r^{(i)} + \eta_{nt}^{(i)}, \quad (6)$$

where ν , r and η represent

where $\nu_t^{(i)}$ and $\eta_t^{(i)}$ model the effects of small turbulent disturbances, spatial inhomogeneities, and the baroclinic eddies, respectively (see the schematic picture in Figure 3). Additional aspects of the noise terms are discussed further in Appendix B.

3.3 Local dynamics

Owing to the uni-directional coupling in ~~the model~~ our model and to the large N , the local dynamics $\{x_n^{(i)}\}$ at i can be approximated by ~~the following a~~ non-autonomous ~~(or random) dynamical system;~~ dynamical system:

$$x_{n+1}^{(i)} \approx f^{(i)}(x_n^{(i)}) + r^{(i)} + p_n^{(i)}, \quad (7)$$

- 5 where $r^{(i)} \in \{-0.04, 0.02\}$, $p_n^{(i)} = \epsilon \Delta_n^{(i)} + \nu_n^{(i)} + \eta_n^{(i)}$, and $\Delta_n^{(i)} = \tilde{f}(x_n^{(i-1)}) - \tilde{f}(x_n^{(i)})$. The new variable $p_n^{(i)}$ represents the external perturbation from the adjacent environment. Assuming that the time averages $\langle \Delta_n^{(i)} \rangle$, $\langle \nu_n^{(i)} \rangle$, and $\langle \eta_n^{(i)} \rangle$ are all 0 by symmetry, then $\langle p_n^{(i)} \rangle \approx 0$, so that we recover the average return map at cell i given in Eq. 4.

where a non-autonomous external force $p_t^{(i)}$ is given by:

In the absence of interactions, $|p_n^{(i)}| \rightarrow 0$, there are three invariant sets:

$$10 \quad p_t^{(i)} = \epsilon \left[f^{(i-1)}(x_t^{(i-1)}) - f^{(i)}(x_t^{(i)}) \right] + \nu_t^{(i)} + \eta_t^{(i)}. \quad (8)$$

When $|p_t^{(i)}| \rightarrow 0$, the local dynamics has a stable fixed point at $x \approx 0$, and two unstable chaotic sets at near $x \approx \pm 2$. With a non-autonomous external force $p_n^{(i)}$ When $p_t^{(i)} > 0$, the resulting perturbed dynamics may exhibit escape behaviour from the fixed point to the chaotic regions with positive Lyapunov exponents. The external perturbation $p_n^{(i)}$ can be approximated by Approximating the external perturbation $p_t^{(i)}$ as a random variable drawn from a obeying the uniform distribution in $[-\kappa, \kappa]$:

- 15 The bifurcation diagrams as a function of κ over land ($r^{(i)} = -0.04$) $r^{(i)} = -0.02$, and ocean ($r^{(i)} = 0.02$) are shown in Figure 4. They both indicate a bifurcation to chaotic and partially chaotic behaviour (Sato et al., submitted). The different values of $r^{(i)}$ over land and ocean give rise to an asymmetry in the minimal invariant sets, which delimit the namely the sets delimiting the accessible region of the dynamics with respect to all possible external perturbations $p_t^{(i)}$. In Figure 4, these minimal invariant sets are dynamically reachable regions are depicted in grey, while a realization of the dynamics with $\{p_n^{(i)}\}$ is given is depicted by the black dots. With $r^{(i)} = -0.04$ ($r^{(i)} = 0.02$) and $0.136 < \kappa < 0.217$ ($0.156 < \kappa < 0.196$) $r^{(i)} = -0.02$ over land and $0.1574 < \kappa < 0.1985$, there is a small chance to reach SJ (NJ) positions and no chance to reach NJ (SJ) positions. This is reflected in the skewed x distribution for larger values of κ . In the interest distribution of $x_t^{(i)}$. For the sake of conciseness, we do not report here the exact the detailed bifurcation analysis of the local dynamics here. A brief analysis for the global dynamics is presented in Sect. 5.

25

4 Dynamical indicators Validation of the model against ERA Interim data

We assess the validity

In this section we compare the ERA Interim deseasonalized jet position data with numerical simulations of our model by computing two instantaneous dynamical systems metrics for both ERA-Interim data and the coupled map lattice. Specifically,

we consider: the local dimension d of θ^{-1} of phase-space trajectories. We briefly outline the physical meaning of these quantities and the way they are computed below. In order to have the same statistical sample as for the attractor and the stability analysis, we simulate 37 years of daily snapshots of the jet position. The best fit of our model to the data is obtained for the parameters: $\eta = 1.2$, $bl = 15$, $\epsilon = 0.33$, $\delta = 10^{-4}$. We further compare the results of model runs containing all noise terms with runs where individual terms are suppressed in turn: the coupling ($\epsilon = 0$), the geography ($r = 0$), and baroclinic waves ($bl = 1$).

4.1 Local Dimension Spatiotemporal Dynamics

The local dimension is estimated by making use of extreme value statistics applied to recurrences. The Freitas et al. (2010) theorem and its modification by Lucarini et al. (2012) states that the probability of entering a ball of small radius centered on ζ obeys a generalized Pareto distribution (Pickands III, 1975) for chaotic attractors. In order to compute this probability empirically, we first calculate the series of distances $\text{dist}(x(t), \zeta)$ between the point on the attractor ζ and all other points $(x(t))$ on the trajectory. This series is transformed via the distance function:-

$$g(x(t)) = -\log(\text{dist}(x(t), \zeta)),$$

such that nearby recurrences to ζ correspond to large values of $g(x(t))$ Collet and Eckmann (2009). Thus, the probability of entering a small ball at ζ is transformed into the probability of exceeding a high threshold q . In the limit of an infinitely long trajectory, it can be shown that the choice $g(x(t))$ in Eq. locks this probability into the exponential member of the generalized Pareto distribution:-

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

where μ and β (obtained via fitting) depend on the point ζ . Remarkably, $\beta(\zeta) = 1/d(\zeta)$, where $d(\zeta)$ is the local dimension around the point ζ . This result has been recently applied to sea-level pressure fields in Faranda et al. (2017b). In this paper, we use the quantile 0.975 of the series $g(x(t))$ to determine q . We have checked that our results are robust with respect to reasonable changes in this quantile. We first consider the latitudinal distribution of the yearly median jet positions at each longitude and their interannual mean. The ERA Interim dataset (Figure 5-a) presents a negative interannual mean jet position at almost all longitudes, with noticeable zonal asymmetries. The best model run (Figure 5-b) captures both the interannual variability and, thanks to the term r , the longitudinal variations in average location. A run without coupling ($\epsilon = 0$) is shown in Figure 5-c). In this simulation the dynamics is local, except for the presence of block noise, resulting in a discontinuous jet profile. Unlike the ERA Interim data, the run with no geography (Figure 5-d) has median values which are roughly symmetric around zero. Finally, the run with suppressed baroclinic activity (Figure 5-e) has a smaller interannual variability than the ERA Interim data and sharp changes in the median values of x following the geographic constraints.

4.2 Local Persistence

Extreme value statistics also provides a way of estimating local stability, by inspecting the temporal evolution of the dynamics around ζ . In particular, it is of interest to know the mean residence time of the trajectory within the neighborhood of ζ . To measure this quantity, we employ the so-called extremal index θ (Freitas et al., 2012; Faranda et al., 2016a), which can be thought of as the inverse of this mean residence time. Heuristically, if the i th visit to the neighbourhood of ζ lasts τ_i (i. e. τ_i consecutive time steps), and N such visits are made in total, then $\theta^{-1} \approx (1/N) \sum_i \tau_i$, where $\theta \in [0, 1]$. $\theta = 0$ corresponds to a stable fixed point of We next consider the NHJ's shifts from CJ towards NJ or SJ positions. We binarize the dynamics by the detecting all the events such that $|x| > 1$. Note that this corresponds to breaks in the jet position with the same threshold defined by the breaking index BRI . We then assign '0' to all the observations with $|x| < 1$ (CJ) and '1' to all the others (NJ or SJ). This procedure, known as *coding*, is widely used in dynamical systems analysis to identify different dynamical phases in complex systems (Kaneko, 1990). The so-obtained binary spatio-temporal dynamics are shown in Figure 6a-e) for all the previously described runs. In the ERA Interim data (Figure 6a), the dynamics so that the trajectory resides an infinite amount of time in the neighbourhood of ζ . $\theta = 1$ corresponds to residing in the neighbourhood of ζ for only one time step per visit. In practice, the estimate of θ is sensitive to the value of the time step dt used. If dt is too large, the time dependence structure is unresolved and θ will be close to 1. Conversely, if dt is too small, θ is close to zero. In Faranda et al. (2017b) it has been observed for sea-level pressure fields over the North Atlantic that θ varies between 0.3 and 0.5, when $dt = 1$ day. In this work we use the same dt . The extremal index is estimated using the likelihood estimator of Süveges (2007). switch from CJ to NJ and SJ phases occurs in clusters displaying characteristic longitudinal extent and temporal persistence. There is also some indication of westerly propagation of the clusters. The best model fit captures the qualitative aspects of this behavior, although the longitudinal coherence is weaker (see section 4.2.3 below). In the remaining model simulations, the suppression of different noise terms alters either the cluster size or the westerly propagation of the clusters (6b-e). A quantitative analysis of the cluster size distributions is presented in Figure-6f) for space clusters and Figure-6g) for time clusters. There is a clear power law behavior, reminiscent of a multiscale structure (Schertzer et al., 1997). This is coherent with the claim that the underlying jet dynamics is turbulent, with energy at all scales. Despite its simplicity, our model is capable to reproduce this power law behavior. The theoretical reasons are non-trivial and can be related to the possibility of building turbulent cascades starting from simple Langevin equations Wouters and Lucarini (2013); Faranda et al. (2014). We underline here the necessity of adding ϵ and having $bl > 1$. Indeed when $\epsilon = 0$ the spatial cluster distributions consist of discrete peaks with the energy concentrated at precise scales. These are a resonance of the block noise size. When instead we impose $bl = 1$, we have still a power-law behavior for time clusters but the slope for the temporal clustering largely deviates from that observed for ERA Interim data.

30 4.2 Dynamical properties of ERA-Interim jet position data

We illustrate the use of the above metrics on the ERA-Interim jet position data. Figure ??-a) plots the local dimension-

4.2 Dynamical indicators

We further assess our model by means of the d versus inverse persistence and θ metrics described in Sect. 2.1, computed on both ERA-Interim data and the coupled map lattice. We also compare here the statistics of the spatial breaks of the jet, detected via the indicator BRI .

Figure ?? show the box-plots of d (Figure 7-a), and θ (Figure 7-b) for each day of the data set. The colour scale represents and the breaking index BRI . Panels b) and c) show the cross-correlation between BRI and d and θ , respectively. BRI is highly correlated with d . The more breaks in the zonal flow, the higher the local dimension. This result is consistent with the findings of Faranda et al. (2017b) for the North Atlantic region. BRI is also correlated with θ : the more breaks, the lower the persistence of the flow. The shape and (Figure 7-c). The range of values of the d and θ diagram are very close to for the ERA Interim data resemble closely those found for sea-level pressure fields over the Northern Hemisphere (Faranda et al., 2017a). This supports the claim that the position of the jet is indicative of large-scale features of the NH atmospheric circulation. Similar claims about the relevance of low dimensional projections in describing the atmospheric circulation at mid-latitudes are presented by Madonna et al. (2017). In the following, we will use these dynamical properties as guidelines when choosing the best parameters of our model.

5 Results: Comparison of model and reanalysis dynamical properties

We next analyze the bifurcation structure, the spatio-temporal dynamics and The model runs can produce average dimensions comparable to those observed in the ERA Interim data, except for the $bl = 1$ case. There, the fragmented dynamics leads to a much higher dimension. This is coherent with the spatio-temporal diagrams shown in Figure 6. The models' inverse persistence θ are slightly larger than those observed in reanalysis, but still in the order of 2 days. Here we can clearly see the effect of the noise suppression ($\epsilon = 0$ and $bl = 1$) in modifying the dynamical properties (local dimension, persistence and by leading to lower persistence. Finally, we remark that the number of breaks is correlated to the local dimension. This result is consistent with Faranda et al. (2017b), who found that high d match blocking-like atmospheric configurations in the North Atlantic region. For the limiting $bl = 1$ case, BRI of the coupled lattice model defined in Sec. 3, is also correlated with θ : the more breaks, the lower the persistence of the flow.

5 Bifurcation diagram and jet regimes

The bifurcation diagram in Figure 8 is constructed by plotting the empirical density $\rho(x)$ of the jet position $\rho(x)$ at fixed longitude (240° W) at all longitudes as a function of ϵ for a) $r = 0$ (spatial homogeneity) and b) $r^{(i)} = 0.02 (i \in \text{land})$ and $r^{(i)} = -0.04 (i \in \text{ocean})$. The parameter values used are $\eta = 1$, $bl = 15$, $\delta = 10^{-4}$. Three different regimes emerge: i) for small ϵ . The vertical gray line corresponds to the value of ϵ ; CJ fluctuates around a fixed latitude with rare excursions towards NJ and SJ; ii) for large ϵ , the system is pushed from NJ to SJ, and never stabilizes around the CJ; iii) for intermediate ϵ , the jet is centered around its central position CJ on average, but with relatively frequent excursions towards NJ and SJ. We claim, as discussed below, that this is the state that qualitatively resembles the BRI from the reanalysis. The addition of the geographical

inhomogeneity (Figure 8 b)) does not change this qualitative picture, but modifies the region of bifurcations and that best fits the ERA Interim data. The diagram would look symmetric with respect to $x = 0$ if $r = 0$ everywhere, but the addition of geography via $r^{(i)}$ alters the relative proportions of time spent in SJ versus NJ. Moreover, an asymmetric land/ocean distribution also implies a southward shift of the average CJ position with increasing coupling. This is reminiscent of the behavior-behaviour in the stochastic bifurcation obtained from the approximated local dynamics (Figure 4).

In order to verify the claim that for intermediate ϵ the dynamics of the model matches that of the ERA-Interim data, we analyze other properties of the flow for three different values of ϵ . The results are summarized in Figure ???. The first column of panels (a,e,i,m) display ERA-Interim detrended data, while all other panels display model data with $\epsilon = 0.4$ (b,f,j,n), $\epsilon = 0.01$ (c,g,k,o), $\epsilon = 0.8$ (d,h,l,p). Panels (a-d) display spatial-temporal dynamics for the last 300 days of data or simulations; (e-h) plot auto-correlation functions for the jet position time series at location 240° W; (i-l) are snapshots of the jet, where the CJ is displayed in black, NJ in red and SJ in green, and land (model) in magenta; (m-p) plot histograms. By analysing the bifurcation structure of the conceptual model as a function of the coupling coefficient — which mimics the coherence of the jet position (for all time). From a visual analysis of the spatio-temporal diagrams, we obtain the best qualitative match between model and ERA-Interim data for $\epsilon = 0.4$. Over large regions (see the snapshots in panels (i-l)) the transition to SJ and NJ are zonally extensive and progress eastwards. However, we note that the propagation speed in the model is significantly slower than in the — we identify three behaviours: (i) a strong and uniform jet where large meridional excursions in the jet location are relatively rare events ($\epsilon < 0.35$), which is close to the jet dynamics as inferred from the ERA-Interim data. This is reflected in the longer tail of the auto-correlation function (compare panels e, f). The distribution of jet positions for $\epsilon = 0.4$ is also qualitatively similar to that of the data (cf. panels m-n), although the latter has a larger positive skewness. It is evident that the output from $\epsilon = 0.01$ (c,g,k,o) and $\epsilon = 0.8$ (d,h,l,p) does not match the qualitative features of the data. When ϵ is small we do not observe SJ and NJ but only small fluctuation around CJ, mostly due to the spatial noise η mimicking the effects of baroclinic waves. When ϵ is large, the system shows very persistent jet deviations.

A more quantitative analysis can be performed by looking at d and θ and their dependence on the BRI for the different simulations discussed above (Figure ??). These can then be compared to the corresponding diagram for the ERA-Interim data; (ii) a state with sharp meridional excursions in which the jet is very unstable and on average shifted far to the South ($0.6 < \epsilon < 0.9$); and (Figure ??), which we show in Figure ??a for ease of comparison. In support of the previous analysis, the dependence of d and θ on BRI for $\epsilon = 0.01$ (Figure ??b) is very close to that obtained for the data, although the dimensions of the model are slightly higher and the persistence lower (i.e. higher θ), and both show a suppressed variance. The simulations with $\epsilon = 0.01$ and $\epsilon = 1$ (Figure ??c, d) again show a larger difference from the reanalysis. For $\epsilon = 0.01$ the persistence is lower (higher θ) and d is constrained to a very narrow range of values, because this state features just one trivial noisy fixed point. In the case $\epsilon = 1$, the persistence is high (θ small) and the range of d is large. Moreover, the BRI index is large because the system experiences very frequent breaks between the SJ and NJ. The few events with low BRI correspond to maxima of d , whereas for the data these events correspond to relatively high values of BRI (iii) an intermediate state of transition between the two. These jet regimes are broadly consistent with those obtained in idealised atmospheric simulations (Lachmy and Harnik, 2016; Son and Lee, 2005), although here we do not delve into the physical mechanisms underlying the

different behaviours. It is also noteworthy that our model qualitatively reproduces a southern jet configuration, even though we provide it with a single NHJ and do not distinguish between eddy-driven and subtropical jets.

When averaging the instantaneous indicators over all the configurations, regardless of whether they are simulated or observed, we obtain the dimension of the attractor D , and the average persistence Θ . For the ERA-Interim data, $D \simeq 13$ and $\theta \simeq 0.43$.

5 We can therefore scan through the model parameters to find the best fit to the data. The values of D and Θ computed for different model runs are shown in Figure ???. We scan the parameter space by varying the coupling, block length, and noise amplitude. The closest values to those of the reanalysis are obtained with $bl = 20$, $\epsilon = 0.4$, $\eta = 0.6$, $r^{(i)} = 0.01$ ($i \in \text{land}$) and $r^{(i)} = -0.02$ ($i \in \text{ocean}$). This means that, compared to our first guess, the noise amplitude for a realistic representation of the data should be lower. In any case, the spatio-temporal diagram, the snapshot and the distribution for this optimized parameter set looks qualitatively similar to that for $\epsilon = 0.4$ and $\eta = 1$.

6 Conclusions

We have derived a minimal model of the jet stream position dynamics, based on a stochastic coupled map lattice, by embedding data extracted from the ERA-Interim reanalysis data set. In comparison to the results obtained for the von Karman flow in Faranda et al. (2017c), the procedure applied here is new — we make use of coupled map lattice derived from a local embedding of the data — and could be adapted to systems with several degrees of freedom. Instead of embedding the data of a global observable in a high-dimensional space, we have constructed the return map for the local position of the jet and then added, via coupling and noise, the physical ingredients recognized-identified in previous studies as drivers of the jet dynamics. The conceptual model is then validated and tuned using local-and-global dynamical indicators of the data-jets dimension and persistence in the reanalysis data.

20 By analyzing the bifurcation structure of the conceptual model as a function of the coupling coefficient — which mimics the coherence — Future analyses could apply this approach to the southern hemisphere, where the role of topography is less important than in its northern counterpart. This would allow us to better constrain the influence of topography on the dimension-persistence diagrams. Another possibility would be to use the low-dimensional model to build a surrogate data set of the jet — we identify three behaviors: (i) a strong and uniform jet where large meridional excursions in the jet location are rare extreme events, (ii) a state which is close to the jet dynamics as inferred from the ERA-Interim data, and (iii) a state with sharp meridional excursions, in which the jet is very unstable. Surprisingly, the bifurcation diagram and jet regimes match the possible behaviours obtained from climate change experiments in global climate models, as the solar input (Lachmy and Harnik, 2016) or the obliquity of the Earth (Armstrong et al., 2014; Linsenmeier et al., 2015) are varied positions and then apply this to atmospheric analogues, so as to construct realistic atmospheric dynamics. Finally, it would be interesting to study whether further projections of the atmospheric dynamics to a lower dimensional space are possible, beyond the model developed here, and to test possible relations between different atmospheric blocking indices and the breaking index BRI defined here.

This study answers The analysis we have conducted can however already answer some of the questions left open in Faranda et al. (2017a) and Madonna et al. (2017) about-concerning the possibility of reducing the complex mid-latitude

circulation dynamics to low-dimensional representations given by blocking indices or conceptual models. The fact that the dimension-persistence diagram of ~~the conceptual-our minimal~~ model qualitatively matches many features of ~~that-those~~ obtained for the ERA-Interim jet position and ~~the entire~~ sea-level pressure fields ~~of the North Atlantic shows that all the dynamics~~ ~~substantially project~~ ~~shows that a substantial part of the dynamics projects~~ along a single line (the jet position). This explains why ~~we observe~~ ~~previous investigations observed~~ relatively low dimensions when considering the full sea-level pressure fields (Faranda et al., 2017b, a). It also suggests that breaks in the jet are responsible for higher dimensions.

~~This study poses a number of further research questions. The same approach could be applied to the southern hemisphere, where the role of topography is less important.~~

Appendix A: Coupled map lattice

10 A coupled map lattice (CML, Kaneko (1983)) is given by:

$$\underline{x_{t+1}^{(i)} = (1 - \epsilon)f(x_t^{(i)}) + \frac{\epsilon}{2} [f(x_t^{(i-1)}) + f(x_t^{(i+1)})]}, \quad (i = 1, 2, \dots, N; t = 1, 2, \dots). \quad (9)$$

where $\epsilon \in [0, 1]$, $x_t^{(i)} \in \mathbb{R}$, and $f(x) : \mathbb{R} \rightarrow \mathbb{R}$. For our jet dynamics, we adopt the open flow model, which is class of CML with uni-directional coupling (Kaneko, 1985):

$$\underline{x_{t+1}^{(i)} = (1 - \epsilon)f(x_t^{(i)}) + \epsilon f(x_t^{(i-1)})}, \quad (i = 1, 2, \dots, N; t = 1, 2, \dots). \quad (10)$$

15 The CML is a phenomenological model to study complex spatio-temporal dynamics in systems with large numbers of degrees of freedom. The idea is to discretize the dynamics in space and time, while capturing the global phenomenology of physical systems. CMLs have been successfully applied to processes such as turbulence in thermal convection (Yanagita and Kaneko, 1993) and turbulent puff in pipe flow (Avila and Hof, 2013). It is convenient for us to model the jet dynamics leveraging the CML approach because we can extract a local one-dimensional map from the observed time series.

20 Appendix B: Average return map and noise

To extract the local jet dynamics, we construct an average return map. We first coarse-grain the state space into M partitions $L_k^{(i)}$ ($k = 1, 2, \dots, M$) and let $\bar{x}^{(i,k)}$ be the midpoint of $L_k^{(i)}$. Then, we construct a set $Y^{(i,k)} = \{x_t^{(i)} | x_{t-1}^{(i)} \in L_k^{(i)}\}$ ($t = 2, 3, \dots$) and a return map $f^{(i)}$ via the return plot of $(\bar{x}^{(i,k)}, \langle Y^{(i,k)} \rangle)$, where $\langle \cdot \rangle$ is the average over the elements of $Y^{(i,k)}$ at each longitude i , and at each partition k :

$$25 \quad \underline{\langle Y^{(i,k)} \rangle = f^{(i)}(\bar{x}^{(i,k)})}, \quad i = 1, 2, \dots, N, \quad k = 1, 2, \dots, M, \quad (11)$$

where $|x| \leq c$. In the region $|x| > c$, we assume linear reflection effects. As a result, we have the return map $f^{(i)}$ in Eq. (3). Figure 2 illustrates the construction for $i = 1$ and $M = 500$. This would allow us to verify some of the hypotheses made here and confirm the influence of topography on the dimension-persistence diagram. Another possibility would be to use the low-dimensional model to build a surrogate data set of the jet positions and then apply this to atmospheric analogues, so as to construct realistic atmospheric dynamics. Finally, it would be interesting to study whether further projections of the atmospheric dynamics to a lower dimensional space are possible, beyond the model developed here.

An important ingredient of the jet dynamics is the presence of topographic obstacles to the mid-latitude zonal flow. Mountain ranges and land-sea boundaries cause meridional deviations in the mean jet location (Tibaldi et al., 1980). This inhomogeneity can be modeled via a parameter $r^{(i)}$ that mimics this “spatial noise”. Since the topography is at most a few kilometers in height, this translates to a perturbation of the order of 10^{-3} in the model. Reasonable geographical constraints are therefore: $r^{(i)} = -0.02$ ($i \in \text{land}$) and $r^{(i)} = 0.0$ ($i \in \text{ocean}$), where land spans the ranges $0 \leq i < 161$ and $239 \leq i < 301$, and ocean spans the ranges $161 \leq i < 239$ and $301 \leq i < 360$. The negative sign for the jet shifts over land is justified by the negative median values of the ERA Interim jet position anomalies over land (Compare Figure 5-a) and b) with Figure 5-c) where no topography is present).

As discussed in Sect. 3.2, noise is a fundamental ingredient in the jet dynamics. The “turbulent noise” term ν relates to physical phenomena in the range from a few meters to a few kilometers, implying a perturbation in the range $10^{-4} < \nu < 10^{-3}$, where ν is a random variable obeying the uniform distribution. The second noise term, η , relates to baroclinic activity and we model it as a block noise taking the same value over bl blocks (the one-dimensionalized size of cyclones/anticyclones in our model) with an amplitude of order 1. The latter value is determined empirically as it is an indicative magnitude of the large shifts mid-latitude baroclinic systems can induce in the jet. To determine a realistic length for bl , we reason as follows: given that our model has a reference scale of about 100km, and assuming a typical scale for extra-tropical cyclones of about 3000 km, we then have that $bl \approx 30$ blocks. However, the perturbations are associated with the cyclone radius rather than diameter: upstream of the cyclone, the jet will mostly be deviated southwards, while downstream of the cyclone, the jet will mostly be deviated northwards. We therefore take the block perturbation to be of size $bl = 15$ blocks.

Owing to the uni-directional coupling in our lattice jet model and to the large N , the local dynamics can be approximated by a non-autonomous dynamical system $x_{t+1}^{(i)} \simeq f^{(i)}(x_t^{(i)}) + p_t^{(i)}$, where the external force $p_t^{(i)} = \epsilon \left[f^{(i-1)}(x_t^{(i-1)}) - f^{(i)}(x_t^{(i)}) \right] + \nu_t^{(i)} + \eta_t^{(i)}$. Assuming that the time averages $\langle f^{(i-1)}(x_t^{(i-1)}) - f^{(i)}(x_t^{(i)}) \rangle$, $\langle \nu_t^{(i)} \rangle$, and $\langle \eta_t^{(i)} \rangle$ are all 0 by symmetry, we have $\langle p_t^{(i)} \rangle \approx 0$. Thus we recover the average return map given in Eq. (11).

30 7 Acknowledgments

All the authors were supported by the ERC Grant No. 338965-A2C2. DF and YS were supported by the CNRS PICS Grant No. 74774. YS is supported by the Grant in Aid for Scientific Research (C) No. 18K03441, JSPS, Japan. YS and DF are supported by London Mathematical Laboratory External Fellowship, United Kingdom and by a CNRS PICS grant. GM is supported by

~~VR~~ grant No. 2016-03724, ~~Sweden,~~ [from the Swedish Research Council Vetenskapsrådet](#), and the Department of Meteorology of Stockholm University. YS, [GM](#) and NRM thank the LSCE for their hospitality.

Author contributions. DF and YS performed the analysis and derived the conceptual model. GM computed the jet position data. All the authors participated in the writing and the discussions.

5 *Competing interests.* The authors declare no competing interest.

References

- Archer, C. L. and Caldeira, K.: Historical trends in the jet streams, *Geophysical Research Letters*, 35, 2008.
- Armstrong, J., Barnes, R., Domagal-Goldman, S., Breiner, J., Quinn, T., and Meadows, V.: Effects of extreme obliquity variations on the habitability of exoplanets, *Astrobiology*, 14, 277–291, 2014.
- 5 Avila, M. and Hof, B.: Nature of laminar-turbulence intermittency in shear flows, *Physical Review E*, 87, 063 012, 2013.
- Belmecheri, S., Babst, F., Hudson, A. R., Betancourt, J., and Trouet, V.: Northern Hemisphere jet stream position indices as diagnostic tools for climate and ecosystem dynamics, *Earth Interactions*, 21, 1–23, 2017.
- Benzi, R., Parisi, G., Sutera, A., and Vulpiani, A.: Stochastic resonance in climatic change, *Tellus*, 34, 10–16, 1982.
- Charney, J. G. and DeVore, J. G.: Multiple flow equilibria in the atmosphere and blocking, *Journal of the atmospheric sciences*, 36, 1205–
10 1216, 1979.
- Collet, P. and Eckmann, J.-P.: *Iterated maps on the interval as dynamical systems*, Springer Science & Business Media, 2009.
- Dee, D. P., Uppala, S., Simmons, A., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M., Balsamo, G., Bauer, d. P., et al.: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system, *Quarterly Journal of the royal meteorological society*, 137, 553–597, 2011.
- 15 Dijkstra, H. A. and Ghil, M.: Low-frequency variability of the large-scale ocean circulation: A dynamical systems approach, *Reviews of Geophysics*, 43, 2005.
- Dole, R., Hoerling, M., Perlwitz, J., Eischeid, J., Pegion, P., Zhang, T., Quan, X.-W., Xu, T., and Murray, D.: Was there a basis for anticipating the 2010 Russian heat wave?, *Mon Weather Rev*, 38, 2011.
- Duchon, C. E.: Lanczos filtering in one and two dimensions, *Journal of applied meteorology*, 18, 1016–1022, 1979.
- 20 Faranda, D., Pons, F. M. E., Dubrulle, B., Daviaud, F., Saint-Michel, B., Herbert, É., and Cortet, P.-P.: Modelling and analysis of turbulent datasets using Auto Regressive Moving Average processes, *Physics of Fluids*, 26, 105 101, 2014.
- Faranda, D., Alvarez-Castro, M. C., and Yiou, P.: Return times of hot and cold days via recurrences and extreme value theory, *Climate Dynamics*, 47, 3803–3815, 2016a.
- Faranda, D., Masato, G., Moloney, N., Sato, Y., Daviaud, F., Dubrulle, B., and Yiou, P.: The switching between zonal and blocked mid-latitude
25 atmospheric circulation: a dynamical system perspective, *Climate Dynamics*, 47, 1587–1599, 2016b.
- Faranda, D., Messori, G., Alvarez-Castro, M. C., and Yiou, P.: Dynamical properties and extremes of Northern Hemisphere climate fields over the past 60 years, *Nonlinear Processes in Geophysics*, 24, 713, 2017a.
- Faranda, D., Messori, G., and Yiou, P.: Dynamical proxies of North Atlantic predictability and extremes, *Scientific reports*, 7, 41 278, 2017b.
- Faranda, D., Sato, Y., Saint-Michel, B., Wiertel, C., Padilla, V., Dubrulle, B., and Daviaud, F.: Stochastic chaos in a turbulent swirling flow,
30 *Physical review letters*, 119, 014 502, 2017c.
- Faranda, D., Alvarez-Castro, M. C., Messori, G., Rodrigues, D., and Yiou, P.: The hammam effect or how a warm ocean enhances large scale atmospheric predictability, *Nature communications*, 10, 1316, 2019.
- Fraedrich, K.: Estimating the dimensions of weather and climate attractors, *Journal of the atmospheric sciences*, 43, 419–432, 1986.
- Frederiksen, J.: A unified three-dimensional instability theory of the onset of blocking and cyclogenesis, *J Atmos Sci*, 39, 969–982, 1982.
- 35 Frederiksen, J. S. and Davies, A. G.: Eddy viscosity and stochastic backscatter parameterizations on the sphere for atmospheric circulation models, *Journal of the atmospheric sciences*, 54, 2475–2492, 1997.
- Freitas, A. C. M., Freitas, J. M., and Todd, M.: Hitting time statistics and extreme value theory, *Probab Theor Rel*, 147, 675–710, 2010.

- Freitas, A. C. M., Freitas, J. M., and Todd, M.: The extremal index, hitting time statistics and periodicity, *Adv Math*, 231, 2626–2665, 2012.
- Ghil, M.: Dynamics, statistics and predictability of planetary flow regimes, in: *Irreversible Phenomena and Dynamical Systems Analysis in Geosciences*, pp. 241–283, Springer, 1987.
- Grassberger, P.: Do climatic attractors exist?, *Nature*, 323, 609, 1986.
- 5 Hadlock, R. and Kreitzberg, C. W.: The Experiment on Rapidly Intensifying Cyclones over the Atlantic (ERICA) field study: Objectives and plans, *Bulletin of the American Meteorological Society*, 69, 1309–1320, 1988.
- Haines, K. and Malanotte-Rizzoli, P.: Isolated anomalies in westerly jet streams: A unified approach, *Journal of the atmospheric sciences*, 48, 510–526, 1991.
- Hansen, A. R.: Observational characteristics of atmospheric planetary waves with bimodal amplitude distributions, *Adv Geophys*, 29, 101–133, 1986.
- 10 Held, I. M. and Larichev, V. D.: A scaling theory for horizontally homogeneous, baroclinically unstable flow on a beta plane, *Journal of the Atmospheric Sciences*, 53, 946–952, 1996.
- Jacoby, T., Read, P., Williams, P. D., and Young, R.: Generation of inertia–gravity waves in the rotating thermal annulus by a localised boundary layer instability, *Geophysical & Astrophysical Fluid Dynamics*, 105, 161–181, 2011.
- 15 Kaneko, K.: Transition from Torus to Chaos Accompanied by Frequency Lockings with Symmetry Breaking: In Connection with the Coupled-Logistic Map, *Progress of Theoretical Physics*, 69, 1427–1442, 1983.
- Kaneko, K.: Spatial period-doubling in open flow, *Physics Letters A*, 111, 321–325, 1985.
- Kaneko, K.: Clustering, coding, switching, hierarchical ordering, and control in a network of chaotic elements, *Physica D: Nonlinear Phenomena*, 41, 137–172, 1990.
- 20 Kitsios, V. and Frederiksen, J.: Stochastic Subgrid Turbulence Parameterisation of Eddy-Eddy, Eddy-Meanfield, Eddy-Topographic and Meanfield-Meanfield Interactions in the Atmosphere, 2018.
- Koch, P., Wernli, H., and Davies, H. C.: An event-based jet-stream climatology and typology, *International Journal of Climatology*, 26, 283–301, 2006.
- Kolmogorov, A.: AN Kolmogorov, *Dokl. Akad. Nauk SSSR* 30, 301 (1941)., in: *Dokl. Akad. Nauk SSSR*, vol. 30, p. 301, 1941.
- 25 Kraichnan, R. H.: Dynamics of nonlinear stochastic systems, *Journal of Mathematical Physics*, 2, 124–148, 1961.
- Lachmy, O. and Harnik, N.: Wave and Jet Maintenance in Different Flow Regimes, *Journal of the Atmospheric Sciences*, 73, 2465–2484, 2016.
- Lee, S. and Kim, H.-k.: The dynamical relationship between subtropical and eddy-driven jets, *Journal of the Atmospheric Sciences*, 60, 1490–1503, 2003.
- 30 Legras, B. and Ghil, M.: Persistent anomalies, blocking and variations in atmospheric predictability, *J Atmos Sci*, 42, 433–471, 1985.
- Letellier, C., Aguirre, L., and Maquet, J.: How the choice of the observable may influence the analysis of nonlinear dynamical systems, *Communications in Nonlinear Science and Numerical Simulation*, 11, 555–576, 2006.
- Linsenmeier, M., Pascale, S., and Lucarini, V.: Climate of Earth-like planets with high obliquity and eccentric orbits: implications for habitability conditions, *Planetary and Space Science*, 105, 43–59, 2015.
- 35 Lorenz, E. N.: Deterministic nonperiodic flow, *Journal of the atmospheric sciences*, 20, 130–141, 1963.
- Lorenz, E. N.: Irregularity: A fundamental property of the atmosphere, *Tellus A*, 36, 98–110, 1984.
- Lorenz, E. N.: Dimension of weather and climate attractors, *Nature*, 353, 241, 1991.
- Lorenz, E. N.: Predictability: A problem partly solved, in: *Proc. Seminar on predictability*, vol. 1, 1996.

- Lucarini, V., Faranda, D., Turchetti, G., and Vaienti, S.: Extreme value theory for singular measures, *Chaos*, 22, 023 135, 2012.
- Madonna, E., Li, C., Grams, C. M., and Woollings, T.: The link between eddy-driven jet variability and weather regimes in the North Atlantic-European sector, *Quarterly Journal of the Royal Meteorological Society*, 143, 2960–2972, 2017.
- McComb, W. D.: *The physics of fluid turbulence*, Chemical physics, 1990.
- 5 McWilliams, J. C., Flierl, G. R., Larichev, V. D., and Reznik, G. M.: Numerical studies of barotropic modons, *Dynam Atmos Oceans*, 5, 219–238, 1981.
- Messori, G., Caballero, R., and Faranda, D.: A dynamical systems approach to studying midlatitude weather extremes, *Geophysical Research Letters*, 44, 3346–3354, 2017.
- Mo, K. and Ghil, M.: Cluster analysis of multiple planetary flow regimes, *J Geophys Res-Atmos*, 93, 10 927–10 952, 1988.
- 10 Nicolis, C. and Nicolis, G.: Is there a climatic attractor?, *Nature*, 311, 529, 1984.
- Penland, C. and Matrosova, L.: A balance condition for stochastic numerical models with application to the El Nino-Southern Oscillation, *Journal of climate*, 7, 1352–1372, 1994.
- Pickands III, J.: Statistical inference using extreme order statistics, *the Annals of Statistics*, pp. 119–131, 1975.
- Reiter, E. R. and Nania, A.: Jet-stream structure and clear-air turbulence (CAT), *Journal of Applied Meteorology*, 3, 247–260, 1964.
- 15 Röthlisberger, M., Pfahl, S., and Martius, O.: Regional-scale jet waviness modulates the occurrence of midlatitude weather extremes, *Geophysical Research Letters*, 43, 2016.
- Sato, Y., Doan, T., Rasmussen, M., and Lamb, J. S.: Dynamical characterization of stochastic bifurcations in a random logistic map, [arXiv:1811.03994](https://arxiv.org/abs/1811.03994).
- Schertzer, D., Lovejoy, S., Schmitt, F., Chigirinskaya, Y., and Marsan, D.: Multifractal cascade dynamics and turbulent intermittency, *Fractals*, 20 5, 427–471, 1997.
- Screen, J. A. and Simmonds, I.: Amplified mid-latitude planetary waves favour particular regional weather extremes, *Nature Climate Change*, 4, 704, 2014.
- Simmons, A., Wallace, J., and Branstator, G.: Barotropic wave propagation and instability, and atmospheric teleconnection patterns, *J Atmos Sci*, 40, 1363–1392, 1983.
- 25 Son, S.-W. and Lee, S.: The response of westerly jets to thermal driving in a primitive equation model, *Journal of the atmospheric sciences*, 62, 3741–3757, 2005.
- Stommel, H.: Thermohaline convection with two stable regimes of flow, *Tellus*, 13, 224–230, 1961.
- Süveges, M.: Likelihood estimation of the extremal index, *Extremes*, 10, 41–55, 2007.
- Thomson, D.: Criteria for the selection of stochastic models of particle trajectories in turbulent flows, *Journal of Fluid Mechanics*, 180, 30 529–556, 1987.
- Tibaldi, S., Buzzi, A., and Malguzzi, P.: Orographically induced cyclogenesis: Analysis of numerical experiments, *Monthly Weather Review*, 108, 1302–1314, 1980.
- Tung, K. and Lindzen, R.: A theory of stationary long waves. I-A simple theory of blocking. II-Resonant Rossby waves in the presence of realistic vertical shears, 1979.
- 35 Weeks, E. R., Crocker, J. C., Levitt, A. C., Schofield, A., and Weitz, D. A.: Three-dimensional direct imaging of structural relaxation near the colloidal glass transition, *Science*, 287, 627–631, 2000.
- Williams, P. D. and Joshi, M. M.: Intensification of winter transatlantic aviation turbulence in response to climate change, *Nature Climate Change*, 3, 644, 2013.

Williams, P. D., Read, P., and Haine, T.: Spontaneous generation and impact of inertia-gravity waves in a stratified, two-layer shear flow, *Geophysical research letters*, 30, 2003.

Williams, P. D., Haine, T. W., and Read, P. L.: On the generation mechanisms of short-scale unbalanced modes in rotating two-layer flows with vertical shear, *Journal of Fluid Mechanics*, 528, 1–22, 2005.

5 Woollings, T., Hannachi, A., and Hoskins, B.: Variability of the North Atlantic eddy-driven jet stream, *Quarterly Journal of the Royal Meteorological Society*, 136, 856–868, 2010.

Wouters, J. and Lucarini, V.: Multi-level dynamical systems: Connecting the Ruelle response theory and the Mori-Zwanzig approach, *Journal of Statistical Physics*, 151, 850–860, 2013.

Yanagita, T. and Kaneko, K.: Coupled map lattice model for convection, *Physics Letters A*, 175, 415–420, 1993.

10 ~~Schematic representation of noise contributions to the CML model (Eq.6): ν represents local turbulent disturbances, r geographical features, η baroclinic eddies, and i spatial positions.~~

~~Bifurcation diagrams as a function of κ for (a) land ($r^{(i)} = -0.04$) and (b) ocean ($r^{(i)} = 0.02$). The grey regions are minimal invariant sets, which delimit the accessible region of the dynamics with respect to all possible external perturbations. A realization of the dynamics with $\{p_n^{(i)}\}$ is given by the black dots. For $r^{(i)} = -0.04$ ($r^{(i)} = 0.02$) and $0.136\dots < \kappa < 0.217\dots$~~

15 ~~($0.156\dots < \kappa < 0.196\dots$), there is a small chance to reach SJ (NJ) positions and no chance to reach NJ (SJ) positions.~~

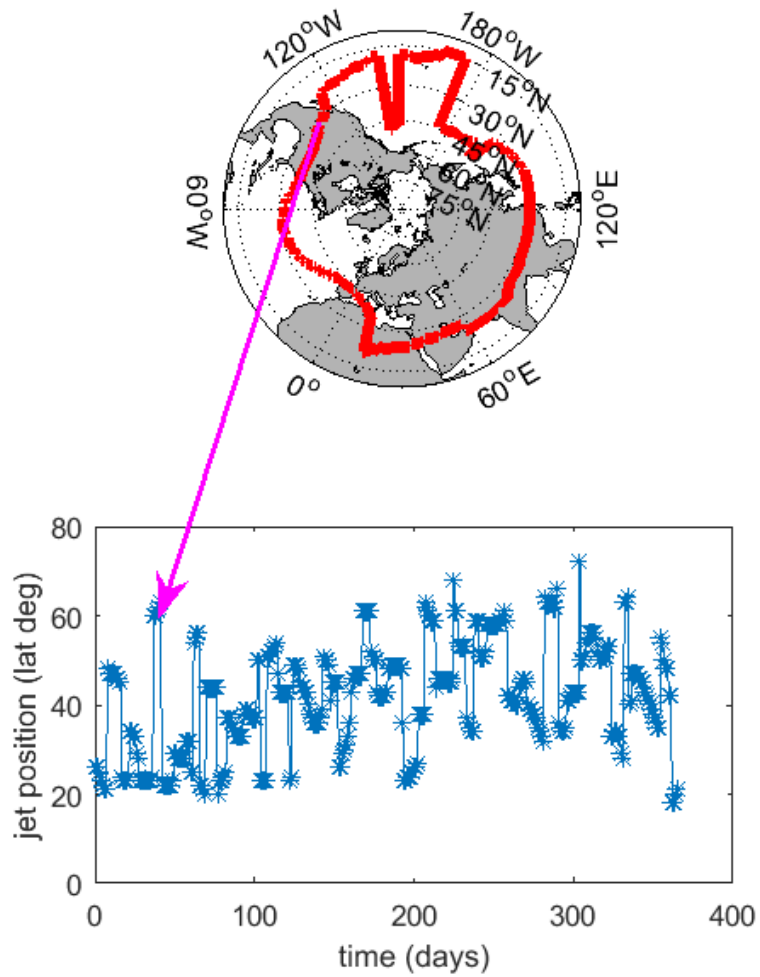


Figure 1. Snapshot of the jet position extracted from the ERA-Interim dataset on Feb 4th 1979 and time series of the jet position for the year 1979, recorded at longitude 120° W.

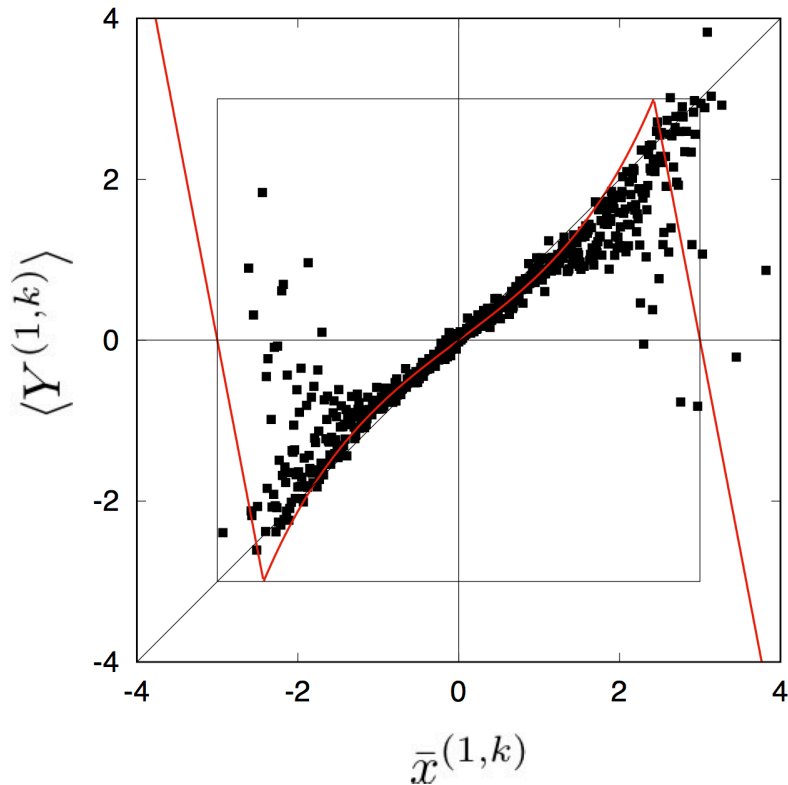


Figure 2. The average return map extracted from the data at longitude $i=0$ $i=1$. This is constructed by coarse-graining the state space at $i=1$ into M partitions $L_k^{(1)}$ ($k=1, 2, \dots, M$). We then define $\bar{x}^{(1,k)}$ as the midpoint of the partition $L_k^{(1)}$, and $Y^{(1,k)} = \{x_t^{(1)} | x_{t-1}^{(1)} \in L_k^{(1)}\}$ ($t=2, 3, \dots$). The black dots are represent $(\bar{x}^{(1,k)}, \langle Y^{(1,k)} \rangle)$ for $k=1, 2, \dots, 500$, where $\langle \cdot \rangle$ is the average over the elements of $Y^{(1,k)}$ computed based on the observed data $\langle x_{n+1} \rangle$ averaged over each median \bar{x}_n . The red line represents the approximated averaged-average return map $\langle Y^{(1,k)} \rangle = f^{(1)}(\bar{x}^{(1,k)})$, when $|x| \leq c$. In the region $|x| > c$, we assume linear reflection effects. As a result, we have the return map $f^{(1)}$ in Eq. 4(3).

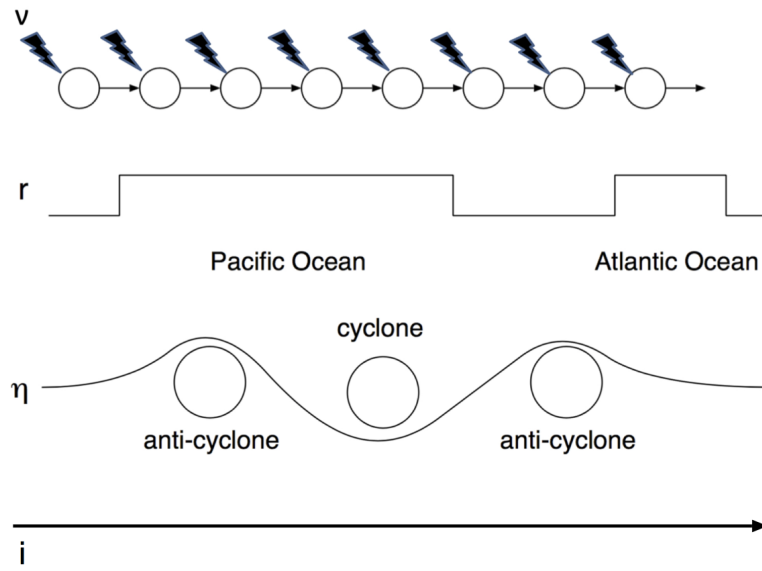


Figure 3. a) Dimension-persistence diagram for the ERA-Interim jet data from 1979-2016. Each point represents the local dimension d and inverse persistence θ for a given day in the data set. The colour scale indicates the number of breaks measured by noise contributions to the BRICML model (Eq. b3, 6). b) Cross-correlation between BRI and v represents local dimension d . c) Cross-correlation between BRI turbulent disturbances, r geographical features, η baroclinic eddies, and inverse persistence θ ; longitudinal positions.

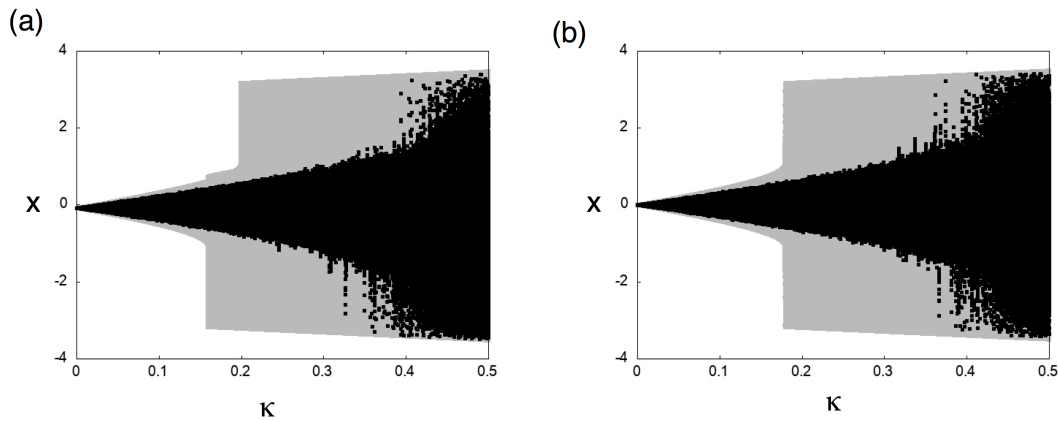


Figure 4. Bifurcation diagram in terms of the density $\rho(x)$, at longitude 120° W, diagrams as a function of ϵ - κ for (a) $r = 0$ land (no spatial inhomogeneity $r^{(i)} = -0.02$), and (b) $r^{(i)} = 0.02$ ($i \in \text{land}$) and $r^{(i)} = -0.04$ ($i \in \text{ocean}$) ($r^{(i)} = 0.0$). The parameter values are $\eta = 1$. Grey regions delimit the accessible region of the dynamics with respect to all possible external forcings. A realization of the dynamics is depicted by the black dots. For $r^{(i)} = -0.02$ and $0.1574 < \kappa < 0.1985$, $bl = 15$ there is a small chance to reach SJ positions and $\delta = 10^{-4}$ no chance to reach NJ positions.

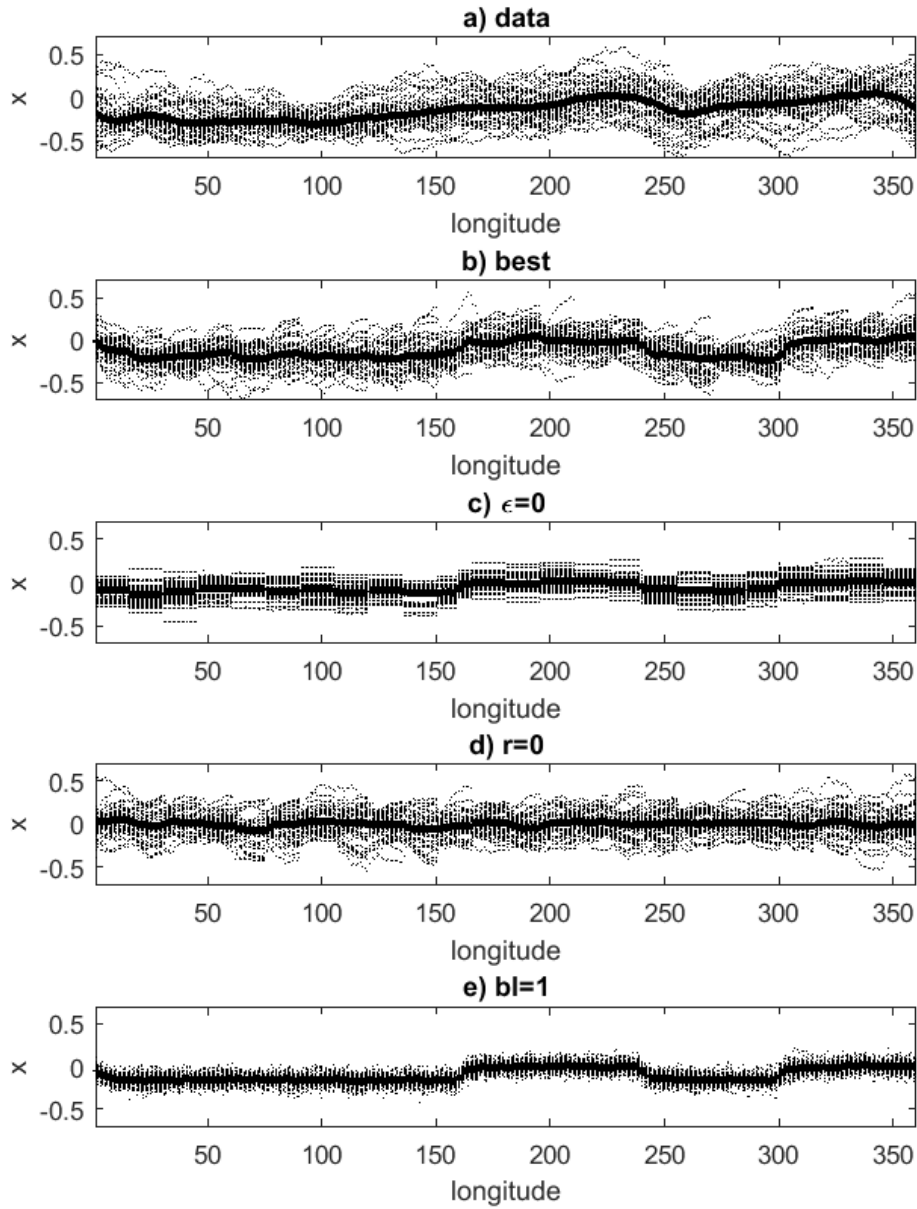


Figure 5. Comparison between ERA-Interim data and model for different parameter sets: (a-d) spatial-temporal diagrams for the last 300 days of data; (e-h) auto-correlation functions for the jet position time-series at [Single year median](#) location 120°W ; (i-l) [dotted points](#) snapshots of the jet, with CJ in black, NJ in red, SJ in green and [land position multi-year average \(model solid lines\)](#) in magenta; (m-p) histograms of the [meridional](#) jet position at all latitudes and times. Plots for ERA Interim (a, e, blue) and model (b-c, i, magenta) refer to the ERA-Interim detrended data; other plots refer to the: b) Best-fit model, obtained with $\eta = 1.2$, $bl = 15$, $\epsilon = 0.33$, $\delta = 10^{-4}$, and $\epsilon = 0.4$ (c) as in b; f, j, n) but with $\epsilon = 0$, $\epsilon = 0.01$ (e, g, k, o, d) as in b) but with $r = 0$, $\epsilon = 0.8$ (d, h, l, p, e) as in b) but with $bl = 1$. The simulations consists of 37 years of daily jet positions.

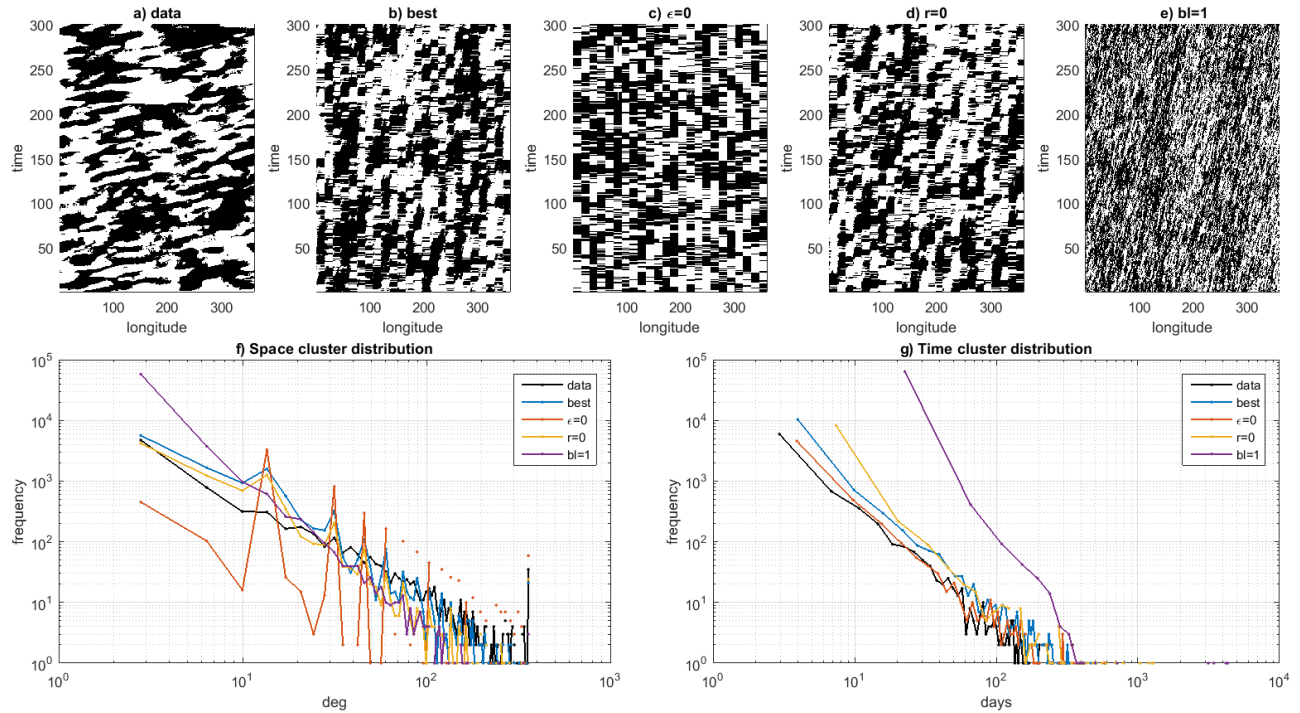


Figure 6. Dimension-persistence diagrams for a-e) Space-time daily representation of the binarized jet dynamics: 1 corresponds to a NJ or SJ shift ($|x(t)| > 1$) ERA-Interim and 0 corresponds to a CJ position. The results are for the ERA Interim data (same as Figure ??-aa) and model runs (b-db-e). The latter are the three model simulations shown same as in Figure ?? with b); $\epsilon = 0.4$; 5. Space (ef), $\epsilon = 0.01$, and time (dg), $\epsilon = 0.8$. Each point represents cluster distributions for the value of the local dimension d -binarized ERA interim data (black) and inverse persistence θ for a given day in the dataset different model runs (colors). The colour scale indicates the number of breaks measured by the BRI.

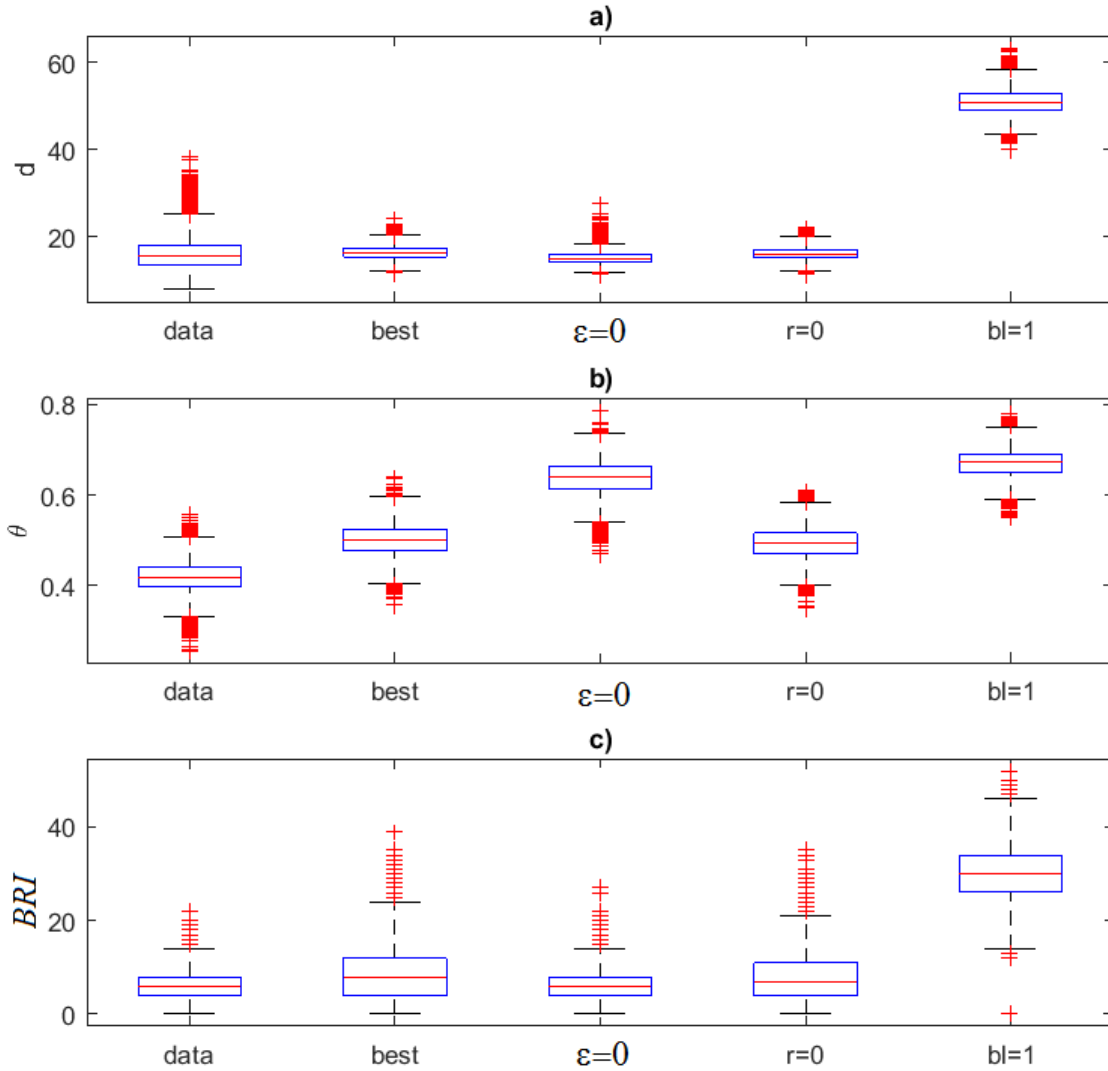


Figure 7. Attractor Boxplots of the local dimension $D-d$ (a,e) and average inverse persistence $\Theta-\theta$ (b,d,f) and breaking index BRI (c) for different parameters of the model. a,b) $\delta = 10^{-3}$, $r^{(i)} = 0.02 (i \in \text{land})$ ERA Interim data and $r^{(i)} = -0.04 (i \in \text{ocean})$, $bl = 15$ four numerical simulations as in Figure 5. e) In each box, d) $\delta = 10^{-4}$ the central mark is the median, $r^{(i)} = 0.01 (i \in \text{land})$ the edges of the box are the 25th and $r^{(i)} = -0.02 (i \in \text{ocean})$ 75th percentiles, $bl = 15$. e) the whiskers extend to the most extreme data points not considered outliers, f) $\delta = 10^{-4}$, $r^{(i)} = 0.01 (i \in \text{land})$ and $r^{(i)} = -0.02 (i \in \text{ocean})$, $bl = 20$ outliers are plotted individually.

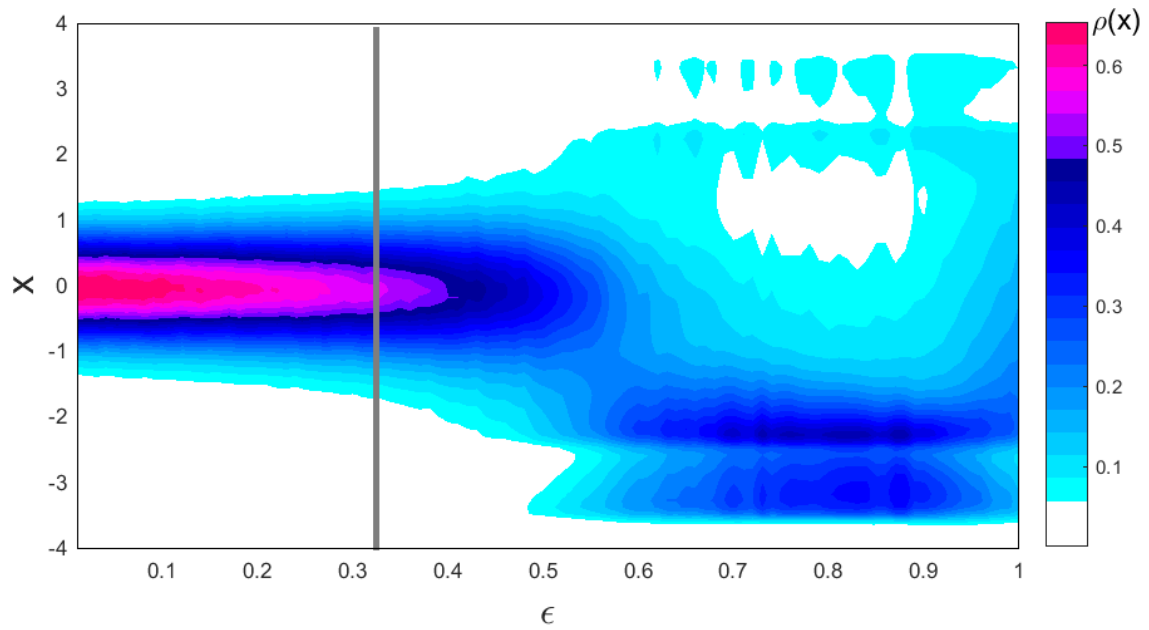


Figure 8. Bifurcation diagram of the global dynamics obtained for $\eta = 1.2$, $bl = 15$, $0 < \epsilon < 1$, and $\delta = 10^{-4}$. The diagram represents the density of states $\rho(x)$ obtained by varying ϵ . The vertical grey line indicates the value used as best fit to the ERA Interim data.