Response to reviewer comments

Comment 2.1 *The authors have put great effort in an attempt to bring together all the different blocs of persistence. Persistence is a loose concept as there is no definite definition for it, and the authors propose to bring together its different facets. I had to go more than once to get a clear picture of the different bits and pieces. For example, in the 6 sections there are 13 subsections (excluding subsub-subsections). Perhaps to aid readers, especially early career researchers, an extra figure showing a tree-like diagram linking the different concepts of persistence would be welcome.*

Answer: Thank you for your comment. This is an excellent suggestion which we will include in a revised version. It will certainly be helpful to readers (i) to get a clear overview of the paper and (ii) to easily navigate through the maze of methods which we discuss.

Comment 2.2 *Some related papers are missing from this review. Two particular references related to persistence and therefore predctability: predictive oscillation patterns (Kooperberg and O'Sullivan 1996), and a related paper: optimally interpolated patterns (Hannachi 2008), making use of the power spectra (ie autocorrelation.) In global stationarity, third-order statistics based, eg on bispectrum (Pires and Hannachi 2021) would complement the persistence description. In relation to extremes and persistence archetypal analysis Hannachi and Trendafilov 2017, Chapman et a. 2022) also identifies 'quasi-stationary' states or regimes.*

Answer: Thank you for pointing out to us these additional papers, which we will include in the revision.

Comment 2.3 *Pg. 1, abstract: delete acronym S2S. It is abbreviated in section 1.*

Answer: If the abstract needs to stand on its own, then the acronym should remain. So far as we know, the rule is that if an acronym is used in the abstract, it must be defined in the abstract, and then defined again the first time it is used in the main text.

Comment 2.4 *Pg 3, l75: you mean Fig. 1c.*

Answer: Correct, thanks for noticing.

Comment 2.5 *Pg7, l175: delete 'of'*

Answer: Thanks, will do.

Comment 2.6 *Pg 13, eq (15): may be use* $0 < |\alpha| < 1$ *to include the case of anti-persistence.*

Answer: Good suggestion, thanks.

Comment 2.7 *l358: add the following reference on short timescale of precipitation (Hannachi 2014).*

Answer: Thanks for pointing out this reference to us.

Comment 2.8 *l367, 369: may be 'persistence' is more convenient here than 'stationary'.*

Answer: Good point. We will change to "persistent".

Comment 2.9 *Pg 14, last paragraph 4.1.2: Fig.3 does not seem to have been mentioned.*

Answer: This is an oversight; we will add the reference to Fig. 3 at l. 389 when mentioning the Fraedrich and Larnder study.

Comment 2.10 *Pg 15, l407: consider adding an earlier reference of Baur (1951) on Grosswetterlagen.*

Answer: Good suggestion, we will add this reference to the sentence.

Comment 2.11 *Pg18, l461: add "by projecting simplified dynamics (eg quasi-geostrophy) onto the leading modes of variability of the GCM simulation" after "January conditions".*

Answer: Thanks for this detauil, we will add it to the revision.

Comment 2.12 *Pg 20, eq (13): x is not specified, and I think there is some confusion in the original reference (Faranda et al.) I think x represents the log of the distance between state at time $t_{\mathbf{x(t)}}$ and $t_{\mathbf{x_0}}$.*

Answer: You are right (and there is indeed a confusion in the original reference). The distribution is fitted to the distance values, and we will modify the corresponding equation as follows:

$$\mathbb{P}\left(d(\mathbf{x}(t), \mathbf{x}_0) \leq \epsilon\right) \simeq \exp\left\{-\theta(\mathbf{x}_0)\frac{d(\mathbf{x}(t), \mathbf{x}_0) - \mu(\mathbf{x}_0)}{\sigma(\mathbf{x}_0)}\right\} \qquad (1)$$

Comment 2.13 *Pg 23, eq(21): my understanding is that α is a probability ($P(x_0) = P(x_{t+1} = x_0|x_t = x_0)$), therefore eq(21) < 0, and also $\log(\alpha) < 0$, please clarify.*

Answer: You are right, we made a mistake in this equation... We have

$$\mathbb{P}\left(R_{\mathbf{x}_0} \geq n\right) = \beta \times \alpha^{n-1}$$

and therefore

$$\mathbb{P}\left(R_{\mathbf{x}_0} = n\right) = \mathbb{P}\left(R_{\mathbf{x}_0} \geq n\right) - \mathbb{P}\left(R_{\mathbf{x}_0} \geq n+1\right) = \beta\alpha^{n-1}(1-\alpha) > 0$$

The slope is indeed negative with respect to n since the probability of the residence time being equal to n decreases with n.

Comment 2.14 *Pg 31, l646: Ripley's K function was considered in Stephenson et al. (2004), and Hannachi (2010).*

Answer: Thanks for pointing out these two references, we will add the following sentence to the section: "*Stephenson et al. (2004) and Hannachi (2010) also used Ripley's K to characterise the clustering of climate modes in the phase space.*"

Comment 2.15 *Pg 31, l686: " ... as the variance of successive event counts over an interval of ..."*

Answer: We will correct it, thanks.

Comment 2.16 *Pg 37, l791: full stop before 'Note'*

Answer: Will correct, thanks.

Comment 2.17 *l793: 'too'*

Answer: Will correct, thanks.

Comment 2.18 *l801: use 'measures' instead of 'metrics'*

Answer: Will correct, thanks.

Comment 2.19 *Fig. 15: are a, b, c 1-day apart?*

Answer: yes, indeed. We will make sure to make it explicit in the figure caption.

Comment 2.20 *Section 5: I suggest renumbering/relabelling the subsections as:*

5.1 Diagnostic methods
- *5.1.1 Window counts*
- *5.1.2 Dispersion metrics*
- *5.1.3 Ripley's K function*
- *5.1.4 Distribution of inter-event times*

5.2 Stochastic modeling of recurrence
- *5.2.1 Events series*
- *5.2.2 Recurrence plots*

Answer: This is an excellent suggestion which brings more structure to this section. Thanks!