

Dear Editor,

first of all, we would like to express our sincere thanks for your overall positive evaluation of our manuscript and the highly valuable comments of the two reviewers who have evaluated our work. With this letter, we are submitting a revised version of our manuscript entitled “Multiscale fractal dimension analysis of a reduced order model of coupled ocean-atmosphere dynamics”. We carefully considered and addressed all Referees’ comments and suggestions to improve our manuscript.

We are confident that the revised version allows us to present our results in a more detailed and appropriate way, improving the clarity and the readability of our manuscript. In the following, we provide a point-by-point reply (in *italics*) to all comments (in normal font) of both Referees.

Sincerely,

Tommaso Alberti, Reik Donner, and Stéphane Vannitsem

### **Referee #1**

**C1.** The work by Alberti et al. is a very intensive and information paper. It shows how to model Atmosphere and ocean dynamics within the scope of ESD. The authors extended the concept of multiscale generalized fractal dimensions employing Multivariate Empirical Mode Decomposition to analyze multiscale and multivariate behavior of the ocean-atmosphere coupled dynamics. Although the concept is not new to the scientific community, it is interesting to know how such a process is applicable for elucidating atmospheric behavior. The one important thing is that they tried to give more credits to the relevant works as much as possible.

The paper is well written with proper usage of English and scientific jargon. However, for the general audience, some of the terminologies need to be explained simpler. For example, the readers may not necessarily need to know about the Hausdorff dimension.

**A1.** *We really appreciate the positive evaluation of our manuscript both in terms of scientific/methodological and formal points of view. To address some of the Referees’ comments we decided to add Supplementary Material in which we introduce more precisely some relevant concepts and terms around the notion of generalized fractal dimensions. This is done at lines 50-60 where the supplementary material is also referenced.*

**C2.** Although they are making some valid assumptions in the methodology, some statements are a bit confusing. For instance, the authors mentioned that mathematical properties of completeness, convergence, linearity, and stationarity are usually not met when real-world geophysical data are analyzed. But it is not clear the reason behind this and what makes the use of adaptive methods. How is the complexity of data suitable for such methods?. Likewise, while the shifting process needs careful implementation for multivariate techniques, Mandic (2010) proposed an alternative way to cubic spline interpolation in each direction with a quasi-Monte Carlo-based approach. But the reviewer does not fully agree with it as such interpolation may lose the data's intrinsic properties since this approach produces smoother dynamics that do not exist in the data.

**A2.** *We really appreciate this comment that allows us to make more clear some statements, trying to reduce some possible confusion.*

*First of all, we thank the Reviewer for highlighting an imprecise sentence that we wrote in our manuscript. Indeed, the mentioned properties that are usually not met in real-world geophysical data should only include linearity and stationarity, and of course not convergence and completeness issues. For this reason, we have corrected this sentence accordingly.*

*Furthermore, adaptive methods can be helpful for overcoming some limitations of fixed-basis methods that are generally characterized by linearity and stationarity assumptions (as for Fourier analysis, for example). In addition, fixed-basis methods implicitly assume that a given natural phenomenon or a superposition of physical processes can be represented in terms of a priori defined mathematical functions like sine and/or cosine or some other kinds of wave functions. This cannot be assured, thus adaptive methods (as the MEMD) could be more suitable for reducing some mathematical assumptions and a priori constraints.*

*Finally, concerning the quasi-Monte Carlo-based approach it is used only to provide a more uniform set of direction vectors over which to compute the local mean of envelopes, and not to interpolate maxima and minima and/or to manipulate the data introducing a smoother dynamics. Indeed the core of the MEMD algorithm proposed by Rehman and Mandic (2010) consists on the following procedures:*

- 1. given a  $n$ -dimensional space we need to find the direction vectors by considering that these vectors reduce to points in a  $(n-1)$ -dimensional space;*
- 2. the simplest choice is to employ uniform angular sampling on an  $n$ -dimensional hypersphere but this will lead to a non-uniform filling of the  $n$ -dimensional space (there would be a higher density of points near the poles of the  $n$ -dimensional hypersphere);*
- 3. a quasi-Monte Carlo method is then used to provide a more uniform distribution of direction vectors;*
- 4. once the direction vectors are chosen, the signal is projected onto these vectors, the extrema of the resulting projected signals are evaluated and interpolated component-wise to yield multidimensional envelopes that are then averaged to obtain the multivariate mean.*

*The quasi-Monte Carlo method is needed only for selecting a uniform sampling of direction vectors, thus to avoid implicitly preferred directions that could be more dominant with respect to the others, which could introduce a source of errors in evaluating signal projections.*

*We have modified both Section 3 “Methods” and Section 3.1 “Multivariate Empirical Mode Decomposition (MEMD)” to add all these details and corrections.*

**C3.** The authors tried to interpret most of the results efficiently. However, some of the interpretation is very unclear and hard to follow. For example, the authors did not mention what is the physical meaning behind the correlation dimension. It is just a kind of statistics of the data. Without understanding the physical meaning, it is not clear why it is a function of time. Another issue is that some of the figures are not interpreted well. e.g., the description of Figures 5 and 6 are not marched. They are not clear, as seen in the figures. For the general audience, they are confusing. Even though the multiscale correlation dimension for each MIMF decreases with an increasing timescale, as seen in panel (a), the other two panels are not well elaborated.

**A3.** *We thank the Referee for this comment.  $D_0$ ,  $D_1$ , and  $D_2$  are strictly related to different properties of physical systems: (i) one purely geometric measure ( $D_0$ ) providing us information on the coverage of the phase-space by the studied system’s dynamics, (ii) one information measure ( $D_1$ )*

*giving us a measure of the information gained on the phase-space with a given accuracy , and (iii) one measure of correlations, i.e., mutual dependence, between phase-space points ( $D_2$ ). Since the collective behavior of a system is given by physical processes operating at different scales, it is straightforward to look how they contribute to the topology of the phase-space, not only singularly (as in panel (a) of Figs. 5-6) but especially when considering all processes occurring below a selected scale (as in panel (b) of Figs. 5-6) and by looking separately at the atmosphere and ocean (as in panel (c) of Figs. 5-6).*

*We modified both the “Introduction” and Section 4.2 “Multiscale generalized fractal dimensions” to take care of this comment, introducing more details on the fractal dimensions in terms of their physical meaning (at least for  $D_0$ ,  $D_1$ , and  $D_2$ ), together with a more detailed description of Figures 5-6.*

**C4.** The authors cleverly described the experiments. To reproduce the work, one needs to understand all the mathematical formulas. In the scientific method, some time calculation and mathematical expression do not match as most of the calculation procedure follow fundamental statistical programming. It needs a concise explanation of calculating all these quantities like system attractors, phase space, and correlation dimensions. The description of these quantities introduced in the manuscript is very dubious and complex to replicate. The reviewer is thankful for providing data sets. But it becomes worthy if it includes an explanation of how to reanalyze these data sets.

**A4.** *We thank the Referee for this suggestion. Our full system consists of 36 variables, thus we are working on a 36-D space. For visual purposes, we reduced our 36-D space to a 3-D subspace by looking at the behavior of the three selected dynamical variables (i.e.,  $T_{o,2}$ ,  $\Psi_{o,2}$ , and  $\psi_{a,1}$ ) in a 3D plot. This allows us to investigate the 3D projection of the full system phase-space attractor, i.e., the set of values toward which our system tends to evolve.*

*Concerning the calculation of the generalized fractal dimensions, we used the approach proposed by Hentschel and Procaccia (1983) consisting of partitioning the phase-space into hypercubes and then measuring the probability of finding a given hypercube filled by points and/or its generalization to a statistical order  $q$  (as also described at lines 40-52 of the submitted version).*

*This comment has been addressed by adding more information and details as Supplementary Material.*

**C5.** Finally, the reviewer appreciates the work of the authors. Still, it needs a bit more simplification and incorporating the issues mentioned above.

**A5.** *We really thank the Referee for his/her nice words on our manuscript and we have done our best to be more precise and more clear in the revised version of the manuscript.*

## Referee #2

### General comments:

The authors propose combining two apparently contradictory analysis techniques to the outputs of a low (36) dimensional dynamical ocean - atmosphere model. The first, makes a nontrivial decomposition of the 36 dimensional signal into series with well-defined time scales, the second analyses the phase spaces assuming the existence of scale invariant properties. The justification and interpretation of this is opaque.

While the authors question the utility of conventional analysis techniques, at least the latter can be interpreted in straightforward manners. The interpretation of their results is nontrivial.

*We thank the Referee for raising some points that can be helpful for improving the presentation and clarity of our findings. Most of all, we did not mean to question the utility of conventional analysis techniques, but rather acknowledge their intrinsic limitations and attempt to explore the potentials of a combination of two “non-conventional” techniques to provide additional information.*

*We also want to stress that the two methods should not be seen as “apparently contradictory” as emphasized by the reviewer. The modes extracted in the first analysis step have no well-defined time scales but are instead characterized by scales that are time-dependent. This is one of the main novelties of the Empirical Mode Decomposition and its multivariate extension we used here (i.e., the MEMD) as compared to fixed-scale decomposition methods like wavelets. The extracted modes can be seen as representative of fluctuations at a typical scale that is the average of the instantaneous scales derived from a given mode via the Hilbert Transform. Moreover, the second analysis step, i.e., the generalized fractal dimensions, requires to have scale invariant properties in the phase-space of a given system, thus working (essentially) on measuring the geometric properties of the system trajectory and information on how to reconstruct it by measuring the information dimension  $D_1$  and  $q$ -tuple correlations  $D_{q>1}$ . This means that there are no a priori constraints on understanding a system using  $D_q$ . Thus, the two methods are not contradictory but rather complementary.*

### Detailed comments:

**C1.** The notation is not easy to follow. Please explain the curly bracket notation used throughout:

On the left, a bold symbol “s” is used which is standard for indicating a vector. Why do the authors (apparently needlessly) add curly brackets and then an explicit restriction as a subscript?

Further, there is the bizarre looking symbol that is also not adequately explained.

**A1.** *We thank the Referee for this suggestion. Indeed, we agree that the notation using curly brackets has been partly misleading, since the left-hand side of the equation was originally intended to represent a sequence of multivariate observation vectors, while the right-hand side was supposed*

*to clarify the structure of each of those vectors composed of  $k$  scalar properties, the latter of which however was lacking clarity in our notation. This aspect has been clarified in our revised manuscript by removing the additional sequence notation. Moreover, the “bizarre looking” symbol has been changed to  $D_q$ , following the notation used in Alberti et al. (Chaos, 2020). We modified Section 3.1 “Multivariate Empirical Mode Decomposition MEMD)” accordingly as well as we took care of modifying other parts through the manuscript.*

**C2.** When discussing the mathematical properties of the usual decompositions (“completeness, convergence, linearity, and stationarity”) it is stated that “these conditions are not usually met when real-world geophysical data are analyzed”. This is confusing since the mathematical properties of Fourier or other decompositions are valid irrespective of any application. I think the authors meant to question the appropriateness of such decompositions for their specific application? However, this is a mathematical question that cannot be answered without reference to a specific assumed mathematical framework. In the paper the authors do not analyze empirical data at all but rather model outputs. Contrary to real empirical series, their series are therefore taken from a well-defined mathematical framework given by dynamical systems theory. Please explain why standard decompositions are not adequate for studying such model outputs and why there is a need for them to be replaced by decompositions with quite nontrivial interpretations and properties.

*A2. We thank the Referee for this important suggestion. As also highlighted in our reply to Referee #1 we actually had to clarify the sentence on properties met by real-world data (note that our manuscript does not exclusively utilize low-order model output, but also reanalysis data, which in our opinion would qualify as “empirical”) for which linearity and stationarity assumptions are often not met. Indeed, we fully agree that mathematical properties of the decomposition methods themselves are surely valid irrespective of any application. As also suggested by the Referee, we referred to the use of adaptive methods that can be justified to overcome some limitations of fixed-basis methods such as linearity and stationarity assumptions. Moreover, adaptive methods (as the MEMD) could be more suitable for reducing some mathematical assumptions and a priori constraints.*

*Although we use the MEMD on a well-defined framework derived from dynamical systems theory, the reduced a priori constraints and the limited number of intrinsic components that can be visually inspected could be an advantage with respect to standard decompositions. Another advantage concerns the combination with generalized fractal dimensions: if we, for example, use Fourier decomposition we would have a large number of (harmonic) oscillating components at different fixed frequencies that should be summed up for exploiting our proposed procedure. Furthermore, if we, for example, use wavelets we would deal with some a priori assumptions on the decomposition basis onto which we are projecting our data that could produce misleading results in our procedure of evaluating fractal measures on a priori fixed scales. Thus, we do not question the appropriateness of conventional analysis techniques, but rather acknowledge that they (as well as any other) have intrinsic limitations in what we can learn from them. We modified accordingly Section 3 “Methods”.*

**C3.** Also in the Methods section, it is stated that the authors “put forward a novel approach based on combining two different data analysis methods for Multivariate Empirical Mode Decomposition and generalized fractal dimensions”. What is confusing is that while the MEMD analyzes time series in real space, in their application, the generalized fractal dimensions analysis is carried out in a quite different space - the phase space of each series. The result is that for each time series with characteristic time scale  $t$ , that the corresponding phase spaces are assumed to be scaling. In other words, while there are essentially no scaling properties in real space, it is assumed that there will be nontrivial scaling properties in the corresponding phase space. The approach is presumably justified if the characterizing these scaling properties via generalized fractal dimensions will help understand the system. At this point one wonders whether the conventional Fourier spectrum of each  $t$  scale series might have been easier to interpret, to understand. All this needs explanation, clarification.

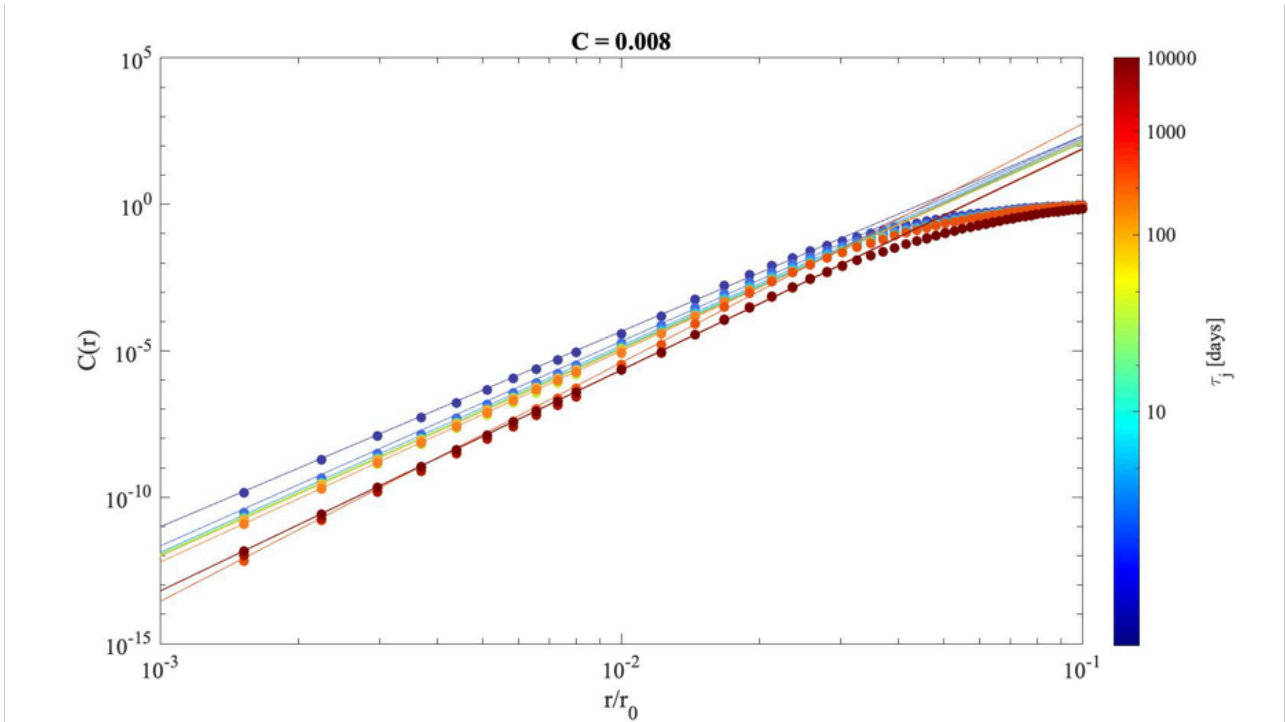
*A3. We really appreciate this comment since it allows us to better underline our main aim. We have been interested in investigating how phase-space properties (geometry, correlations) change when dynamical components at different mean scales with different dynamics are considered. In other words, we have been interested in looking at the role of scale-dependent phenomena in defining the whole properties of a system. Global measures proposed in the past only allow us to investigate the statistical, topological, or geometric scaling properties of the whole system; conversely, our proposed approach allows us to investigate how the different scales contribute to the global properties of a system. Moreover, our framework also provides consistency with established measures for characterizing time series from a global (not scale-resolved) perspective, since the scale-dependent measures we evaluate converge to the associated global measures as all scales are considered, i.e., when the full system dynamics, composed by all accessible scales, is reached. Within this framework, our approach could be promising for investigating scale-dependent properties, as measured by fractal dimensions, of the system. We are indeed interested in nonlinear variability characteristics at different time scales, thus employing for example Fourier decomposition would leave us with perfectly linear and stationary harmonic functions as components, which do not carry any information on nonlinear dynamics, unless when studying their mutual phase relationships, leaving out the high-order statistical properties and only focusing on the autocorrelation function (i.e., the second-order moment). Otherwise, by looking at the behavior of fractal dimensions we can explore how the different scales contribute to change the phase-space properties that cannot be highlighted by using the conventional Fourier spectrum.*

*We have added more explanations and clarifications on those aspects at the end of Section 3.2 “Multivariate and multiscale generalized fractal dimensions” (lines 220-234).*

**C4.** In particular, when the generalized fractal dimensions are estimated, the authors need to show that there are indeed some phase space scaling properties. Using mathematical definitions such as eqs. 1-3 - where the small scale limits are taken - has only a formal validity when the definitions are applied to numerical model outputs, especially when the latter has been subjected to cubic spline interpolation which makes the small scales artificially smooth. In practice, one needs to display scaling behaviour over at least an order of magnitude or so in scale in order for any fractal dimension estimates to be convincing. The authors must therefore display some of their scaling plots - not just logarithmic slopes that have already been interpreted in terms of dimensions.

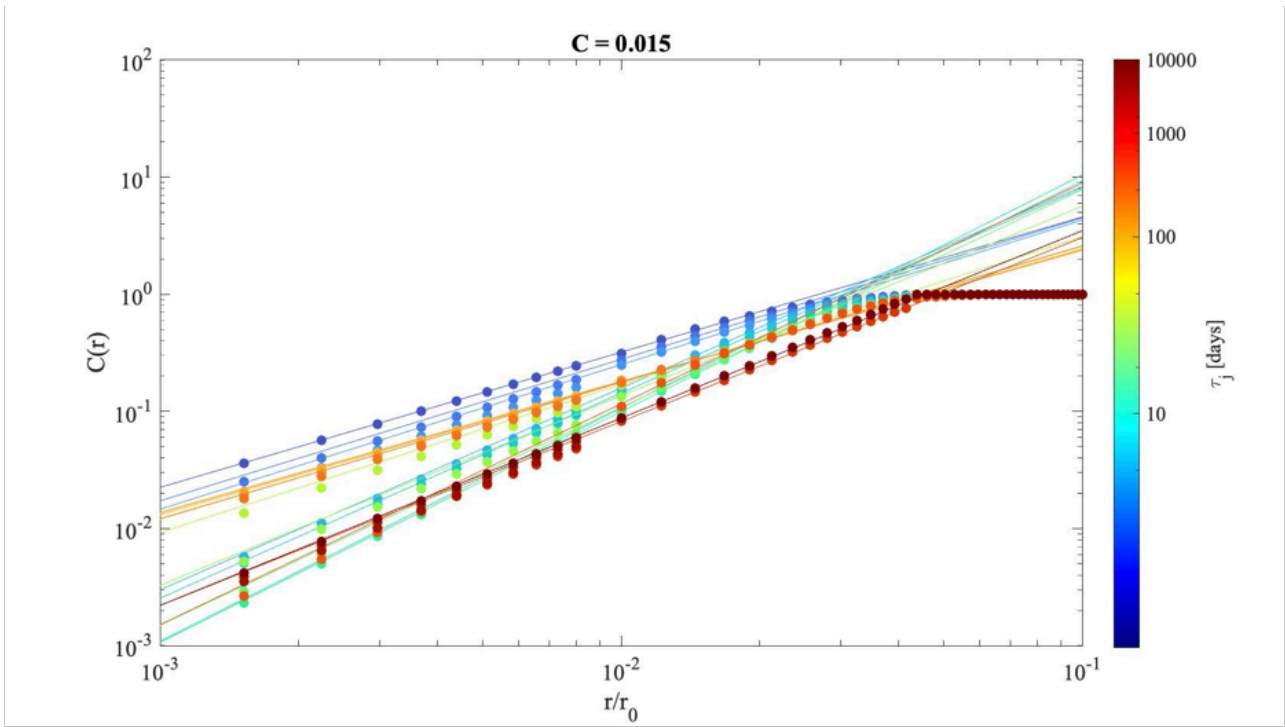
*A4. We agree with this comment. To be clearer and more convincing, we display in Figs. 1 and 2 of this response letter the scaling behavior for the correlation integral for the two cases  $C=0.008$  and  $C=0.015$  at different timescales. We choose to show here only the correlation integral since it can be faster evaluated than other moments (cfr. Grassberger and Procaccia, 1983). We show here that there exists at least an order of magnitude in scale over which a scaling behavior is observed. A similar behavior is also observed when considering the reanalysis data as shown in Fig. 3 of this response letter for the different regions. Taking also into consideration a corresponding comment by Referee #1, we added Supplementary Material with more details on the computation of fractal dimensions and scaling plots to the revised version of our manuscript.*

*We would further like to remark that the cubic spline interpolation does not produce artificially smoothed small scales since it does not act on the data themselves but only on local extreme values of the data to extract intrinsic oscillating components from the data. Thus, the shape of the raw data is not changed and generally the (M)EMD extracts scale-dependent components that are smoother as the largest scales are approached. We also fixed this point in Section 3.1 “Multivariate Empirical Mode Decomposition (MEMD)”.*

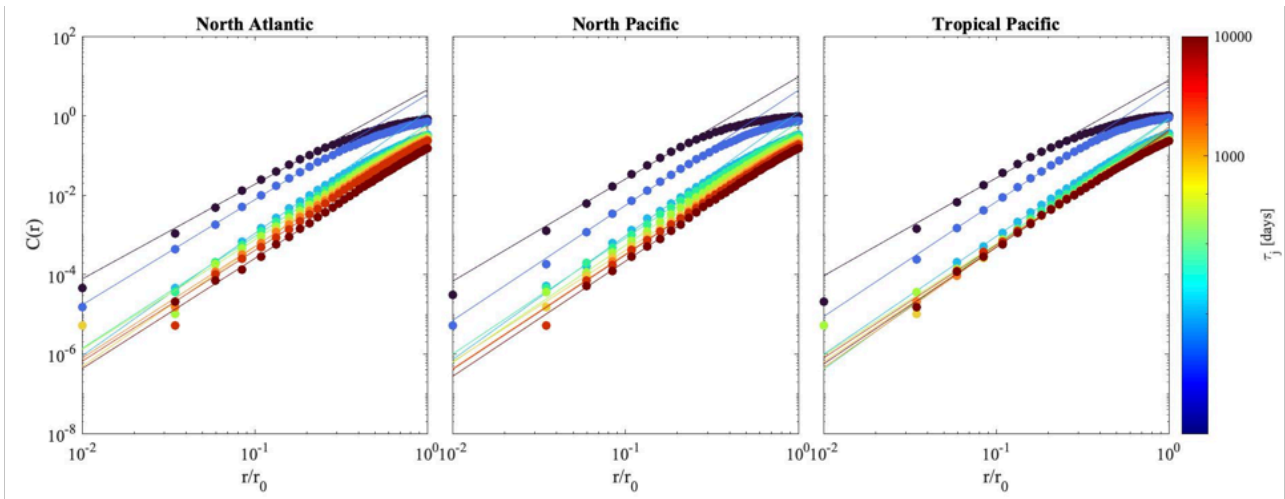


**Fig. 1** The log-log scaling plots of the correlation integral  $C(r)$  as a function of  $r$  (normalized with the respect to the largest possible separation between points in the phase-space represented by  $r_0$ ) at different scales represented by colors for the case  $C=0.008$ .  
The lines refer to the power law fit in the limit  $r \rightarrow 0$ .





**Fig. 2** The log-log scaling plots of the correlation integral  $C(r)$  as a function  $r$  (normalized with the respect to the largest possible separation between points in the phase-space represented by  $r_0$ ) at different scales represented by colors for the case  $C=0.015$ .  
The lines refer to the power law fit in the limit  $r \rightarrow 0$ .



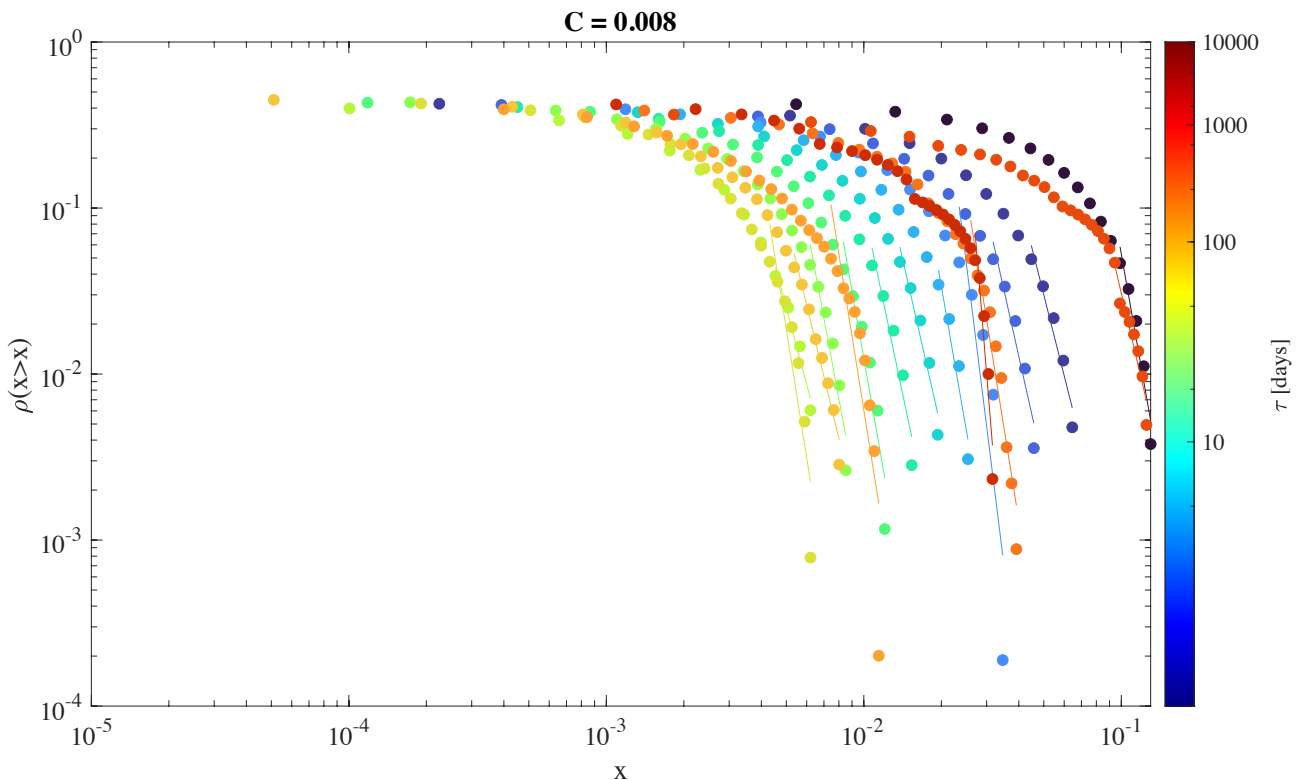
**Fig. 3** The log-log scaling plots of the correlation integral  $C(r)$  as a function  $r$  (normalized with the respect to the largest possible separation between points in the phase-space represented by  $r_0$ ) at different scales represented by colors for the reanalysis data.  
The lines refer to the power law fit in the limit  $r \rightarrow 0$ .



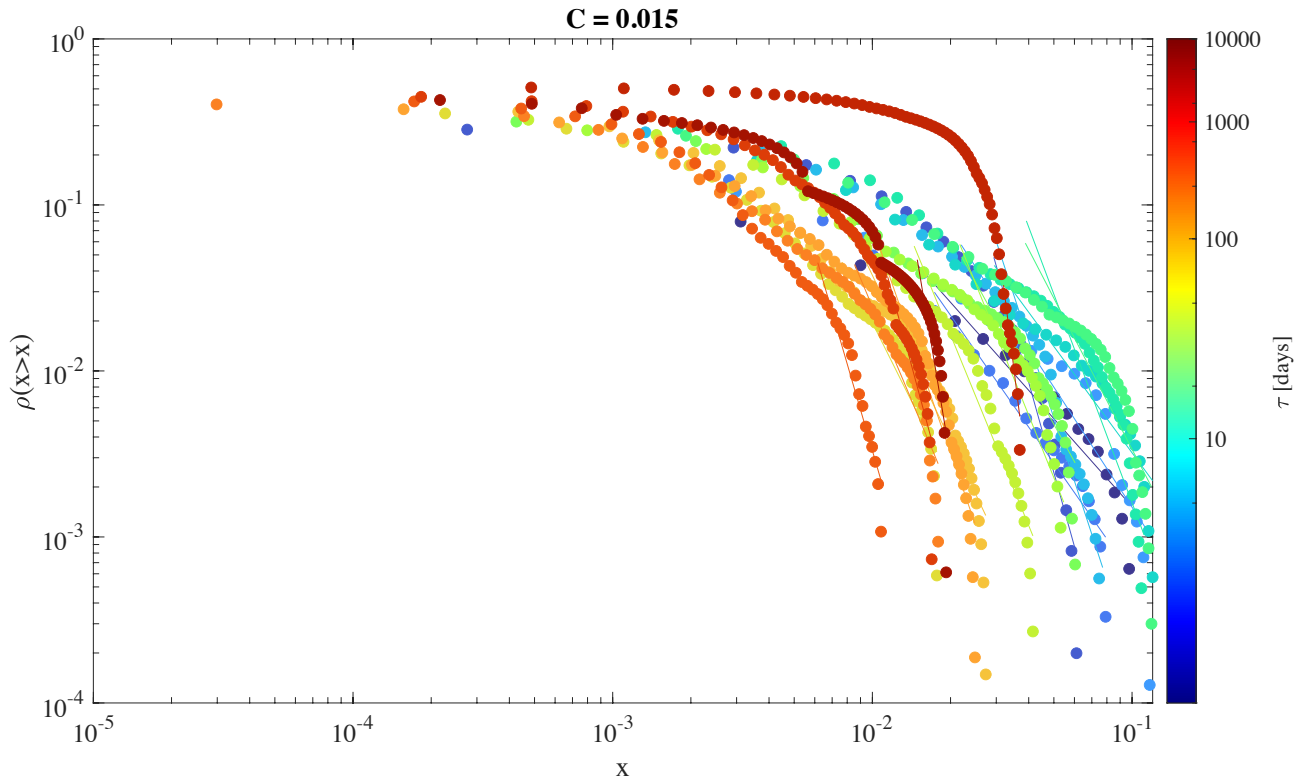
**C5.** In this regard, I could also add that figs. 9 and 11 are almost certainly largely spurious. This is because typically for moments of order  $q \gtrsim 3-4$ , the moments are completely dominated by a single hypercube (a “second order multifractal phase transition”) so that for larger  $q$ , the values will depend sensitively on the exact details of the input series. Similarly for  $q < 0$  most if not all the values will likely be spurious essentially due to the statistics of the very sparsely populated regions of phase space (the very low probability regions, see e.g. the discussion in ch. 5 of [Lovejoy and Schertzer, 2013]). In other words over most of the range of moments given in the figure ( $-20 < q < 20$ ), the dimensions are likely to be spurious.

**A5.** We thank the Referee for raising this important point on the statistical significance of higher-order moments. We are aware that this is a crucial point, especially when working with scale invariant features measured via structure functions, detrended fluctuation analysis, and spectral methods (as for wavelets). To deal with this problem and to support the statistical significance of our results we have followed the approach described in Ch. 5 of Lovejoy and Schertzer (2013) to evaluate the maximum moments as those derived from the tail of the cumulative distribution function of the data. Since we deal with the investigation of scale-dependent fractal dimensions, we evaluate the cumulative statistics at different scales. As shown in Figs. 4-6 in this response letter, we observe that extreme fluctuations follow a power law decay leading to the divergence of the 6-th order moment and the 4-th order moment for  $C=0.008$  and  $C=0.015$ , respectively. Thus we have now fixed our range of moments  $-6 < q < 6$  and  $-4 < q < 4$  for  $C=0.008$  and  $C=0.015$ , respectively, and modified Figs. 9-11 accordingly, without changing the previously described results qualitatively. Similar results have also been obtained for the reanalysis data (see Figs. 7-8 in this response letter), thus we fix here our range of moments to  $-3 < q < 3$ .

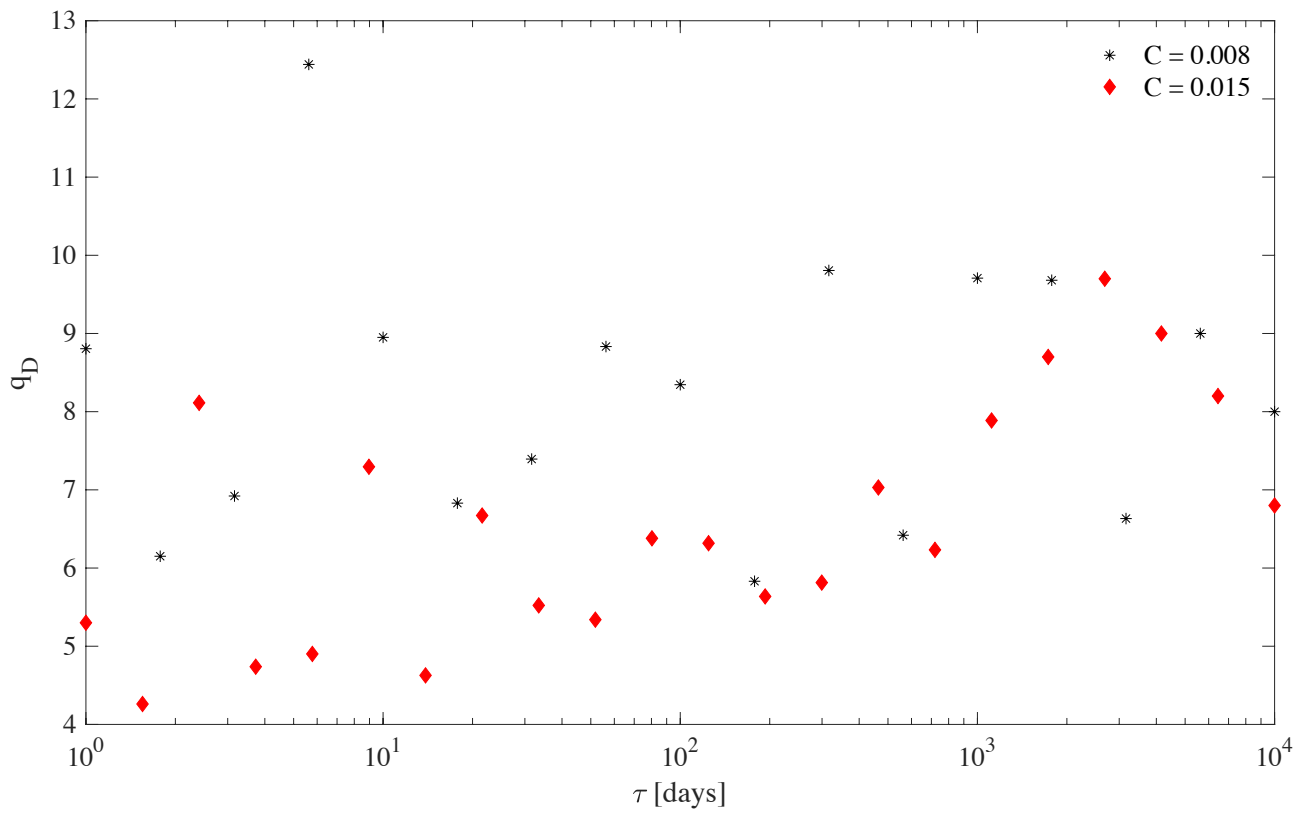
We also inserted Figs. 4-8 of this response letter as parts of our new Supplementary Material (see Figs. S6-S10).



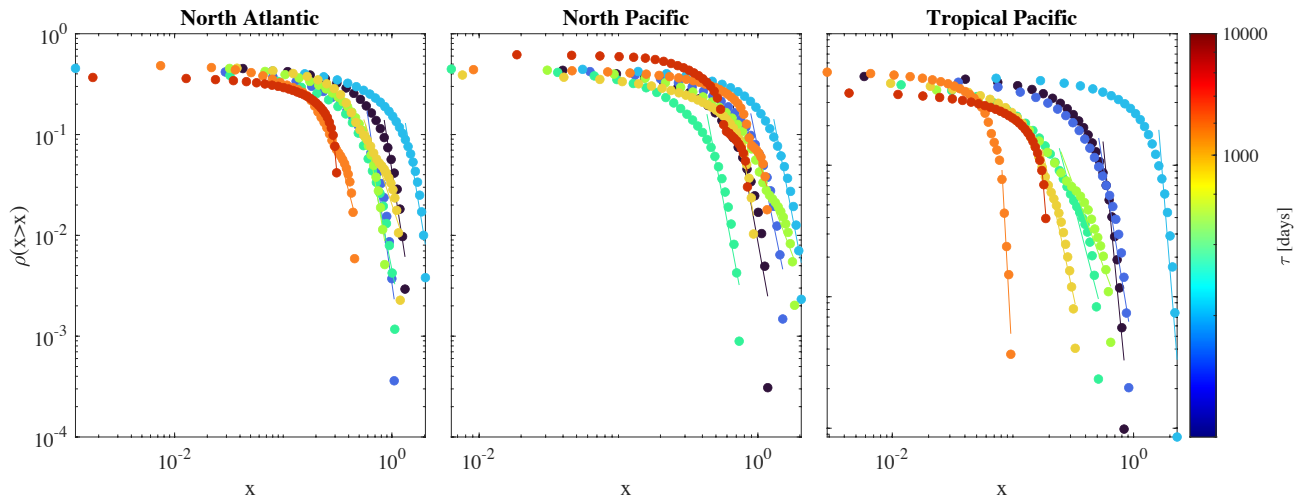
**Fig. 4** Complementary cumulative distribution functions at different scales as reported by different colors for the case  $C = 0.008$ . The lines refer to the power law fit of the tail.



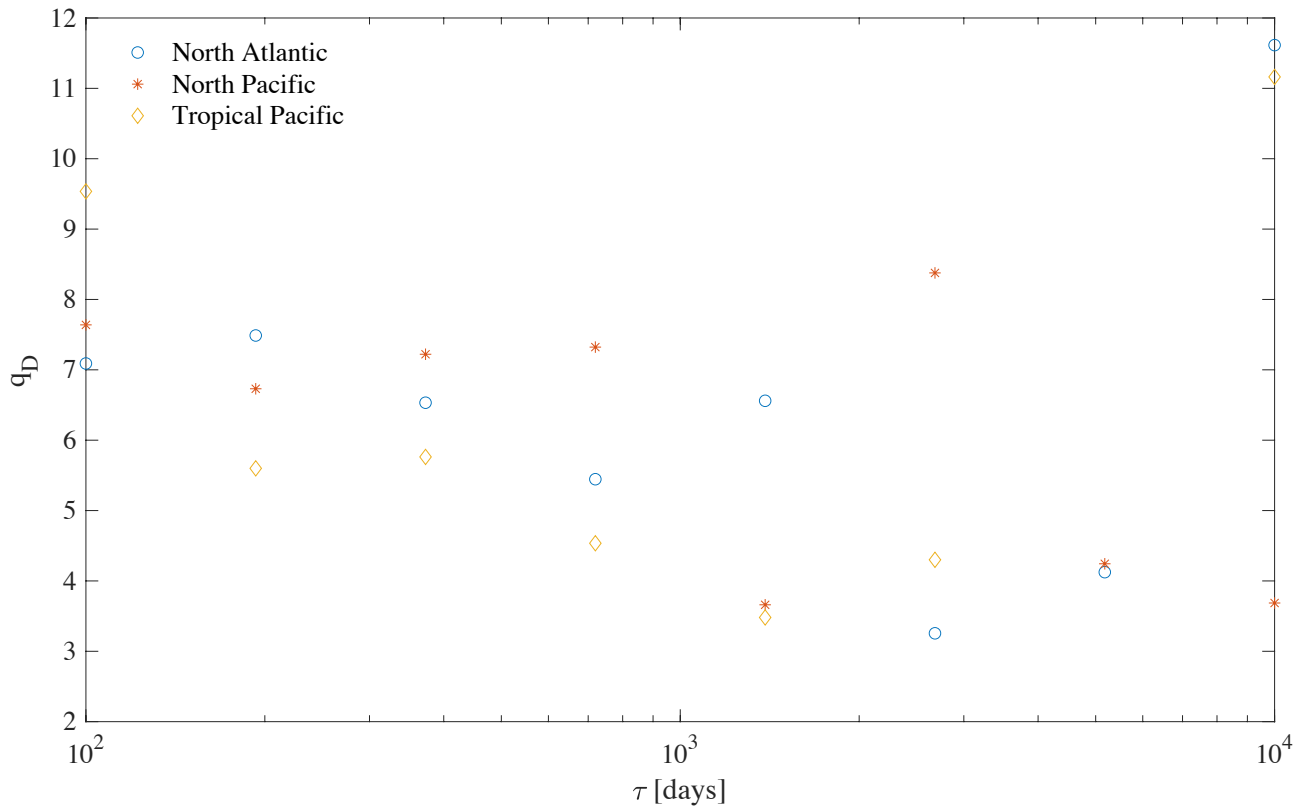
**Fig. 5** Same as in Fig. 4, but for  $C = 0.015$ .



**Fig. 6** The power law scaling exponent  $q_D$  as a function of the different scales for the case  $C = 0.008$  (black asterisks) and  $C = 0.015$  (red diamonds). The minimum  $q_D$  has been chosen to set the range of statistically significant moments.



**Fig. 7** Same as in Fig. 4, but for the reanalysis data.



**Fig. 8** Same as in Fig. 6, but for the reanalysis data.

**C6.** Finally, the interpretation of the key figures 5-8 is not at all obvious. Calling these characterizations “topological, geometric” is unhelpful and/or misleading since they are actually statistical exponents without any straightforward relationship to the phenomenon under study. The authors could note that whereas a white noise signal would give a correlation dimension equal to the dimension of the phase space itself (it is space filling), that a Brownian motion in a space  $d \geq 2$  has a constant dimension = 2.

*A6. We thank the Referee for this comment. We have attempted to further improve the clarity of our manuscript during the revision, especially in the paragraphs introducing some key concepts and/or describing key features. We are referring to topological and geometric since some measures are able to give us information on phase-space properties. For example,  $D_0$  is a measure of the filling of the phase-space, thus providing a measure on the coverage of the phase-space by the studied system's dynamics,  $D_1$  provides a measure of the information gained on the phase-space with a given accuracy, and the  $D_{q>1}$  provide measures of  $q$ -tuple correlations, i.e., mutual dependence, between phase-space points. This explains why we used the terms topological and geometric in our manuscript. We modified accordingly our “Introduction” as well as we give a deeper description of our key figures 5-8 (see Section 4.2 “Multiscale generalized fractal dimensions”, lines 307-352).*

Reference:

Lovejoy, S., and Schertzer, D., The Weather and Climate: Emergent Laws and Multifractal Cascades, 496 pp., Cambridge University Press, 2013.

*Thanks a lot for this reference that we have also cited in the revised version of our manuscript.*