Supporting Information for "Quantifying interdecadal changes in large-scale patterns of surface air temperature variability"

Dario A Zappalà¹, Marcelo Barreiro², and Cristina Masoller¹

¹Departament de Física, Universitat Politècnica de Catalunya, Terrassa, Barcelona, Spain ²Departamento de Ciencias de la Atmosfera, Universidad de la Republica, Montevideo, Uruguay

Abstract. Here we provide supporting information for the results presented in the main text. We apply Hilbert analysis to synthetic data generated with an autoregressive process and compare the results with those obtained from real SAT data. We show how the significance test modifies the maps of relative change for different values of the significance threshold. We also compare the results obtained from two reanalysis datasets (ERA-Interim and NCEP-DOE) and show that Hilbert amplitude

5 and frequency uncover qualitatively similar spatial structures, but there are also some relevant differences between the two datasets.

Hilbert Analysis of Synthetic Data 1

n order to gain insight into the results obtained from SAT time series, we generated synthetic time series aimed at mimicking real SAT data, but with a parameter, α , that allows to control the level of noise.

As a minimal model of SAT time series we consider the sum of a sinusoidal term and a stochastic term that is an autoregres-10 sive (AR) process of order one. We have chosen an AR model because it is commonly used in the literature to model climate data (Hasselmann, 1976).

$$S(t) = \sqrt{1 - \alpha} C \sin(\omega_0 t) + \sqrt{\alpha} \xi_{\text{AR}(1)}(t).$$
⁽¹⁾

By choosing $\omega_0 = 2\pi/365$ oscillations/day, the sinusoidal term can mimic an annual oscillation with daily resolution. Here C is a normalisation factor such that $C\sin(\omega_0 t)$ has unit variance and $\xi_{AR(1)}(t)$ is an AR(1) process with zero mean and 15 unit variance (the parameter that expresses the persistence of the noise is $\beta = 0.5$). The control parameter $\alpha \in [0,1]$ allows to vary the level of noise, while keeping constant the first and second moments of the distribution of S(t) values (zero mean and unit variance). Synthetic time series are generated according to Eq. (1), with the same length as ERA daily reanalysis: T = 13696 days. From the synthetic time series we calculate the instantaneous Hilbert amplitude and frequency, following the same procedure as for the real SAT time series. 20

Figures 1(a) and 1(b) display the results obtained from synthetic series, the average amplitude and frequency respectively, as a function of α , with error bars that represent standard deviations. These results were computed from 10 realizations of the AR(1) process for each value of α . For comparison, the values obtained from SAT time series are also displayed (red dots).



Figure 1. Comparison between real SAT series and synthetic series. (a) Average amplitude and (b) average frequency as a function of level of noise, α , in Eq. (1). The error bars are computed from 10 realizations of the AR(1) process. The dots indicate the values computed from real SAT data.

We can see that there is a very good agreement between synthetic and SAT results, which suggests that, as a minimal model, we can consider SAT time series as the sum of a regular oscillation and an irregular noisy term represented by an AR process. As we have shown in *Zappala et al.* (2016), the regular term tends to prevail in the extratropics, while the noisy term prevails in the tropics and in some specific extratropical areas.

5 In the synthetic data we note that, as the noise level increases, the average Hilbert frequency increases while the average Hilbert amplitude decreases, a trend that is also observed in real SAT time series: the larger the average amplitude, the lower the average frequency.

This trend can be understood by considering the limiting values of α : if $\alpha = 0$, Eq. (1) is just a sine normalised to have unit variance, which gives an amplitude ≈ 1.4 ; if $\alpha = 1$, Eq. (1) is fully random, with a Gaussian distribution of unit variance that gives an amplitude ≈ 1.1 .

2 Significance Test

10

15

As we explained in the main text, we performed a significance test on the maps of relative change of the calculated quantities. We calculated 100 surrogate values of the same relative change, and from this ensemble we calculated the average μ and the standard deviation σ . Then, we considered the actual (no surrogate) relative change as statistically significant if its distance from μ is at least 2σ .

To see in more details how this technique works, in Figure 2 we show examples of the maps of change of amplitude and frequency, with different choices of the threshold value. As expected, we see that the higher the threshold is, the more sites get

erased from the map. Nonetheless, the main structures are still present even at 4σ , so we can conclude that they are robust with respect to the significance filtering.



AMPLITUDE

Figure 2. Maps of relative change of amplitude and frequency, with different values for the significance filter. (a, b) No filter. (c, d) Only values with a distance of at least 2σ from the average of the 100 surrogate values. (e, f) As (c, d), but with a threshold of 4σ .

3 **Comparison Between ERA-Interim and NCEP-DOE Reanalysis**

To test the robustness of our findings, in Fig. 3 we compare the maps of relative change of amplitude and frequency, obtained from two daily reanalysis datasets: ERA-Interim and NCEP-DOE. NCEP-DOE Reanalysis 2 covers a longer period and has 5

 $94 \times 192 = 18048$ geographical sites. In order to perform a precise comparison between the results of the two datasets, in the NCEP-DOE Reanalysis we consider the same time period as the ERA-Interim dataset.

In the first row (a,b) we present the maps of change of average amplitude, while in the second row (c,d) we have the maps of change of amplitude variance. A qualitative good agreement of spatial structures is seen, however, some differences can be

5 noticed, such as in the Indian Ocean, where NCEP-DOE reanalysis (panel b) gives an increase of average amplitude, while ERA (panel a) gives a decrease.

The third and fourth rows present the maps of change of average frequency, and of frequency variance. Here we can again see a qualitative agreement, but there are also some relevant differences. To investigate the underlying reasons, we inspected the time series in selected regions in which the differences between maps (e) and (f) are more pronounced (for example, one

- 10 map shows a small change while the other one shows a large change). We found that there were indeed significant differences between the two SAT time series, in the same geographical region. As an example, Fig. 4 displays the two SAT series in the region that is marked in South America. We see that the time series from ERA-Interim dataset maintains the same general trend throughout the entire length. On the other hand, the NCEP-DOE time series has a sudden change around year 2000, when the seasonal cycle becomes significantly smaller and the rapid fluctuations get a more dominant role on the series. Therefore,
- 15 Hilbert frequency is sensitive to this change and detects this difference between the two datasets, which should be due to different models used to perform the reanalysis.

Therefore, Hilbert analysis is a useful data analysis tool for performing model inter-comparisons, because it captures temporal variations of amplitude and frequency, which may not be detected by other analysis tools. It is an open question which reanalysis more closely represents the real SAT values.

References

Hasselmann, K. Stochastic climate models part I: Theory. Tellus (1976) 28, 473.

Zappala, D. A. et al. Global atmospheric dynamics investigated by using Hilbert frequency analysis. Entropy (2016) 18 (11): 408.









0

-30

-60

-90

0

90

180

(b) NCEP-DOE









Figure 3. Comparison of results obtained from ERA-Interim (left) and NCEP-DOE (right). Relative change of (**a**, **b**) average amplitude; (**c**, **d**) amplitude variance; (**e**, **f**) average frequency; (**g**, **h**) frequency variance.



Figure 4. Comparison between SAT series of the same point (50 S, 287.5 E) in the two datasets: (a) in ERA-Interim and (b) in NCEP-DOE.