Supporting Information for ”Time-scales of the European surface air temperature variability: The role of the 7–8 year cycle”

Nikola Jajcay,* Jaroslav Hlinka,† Sergey Kravtsov,‡ Anastasios A. Tsonis,‡ and Milan Paluš§

I. INTRODUCTION

Here we provide supporting information regarding the methods of data analysis used in the study, presented in the letter ”Time-scales of the European surface air temperature variability: The role of the 7–8 year cycle” (thereafter referred to as “the main text”), including statistical tests with other types of null hypotheses, analysis of surface air temperature (SAT) data in specific seasons and finally, we present results for different stations – Hamburg-Fuhlsbüttel and Potsdam, in addition to the results for Prague - Klementinum, presented in the main text.

II. SAT ANOMALIES

While the variability of the annual cycle is studied using the raw SAT data, the variability on all other scales (referred to as the “overall variability”) is studied using the SAT anomalies (SATA). The anomaly is a digression from a long-term seasonal mean. For the daily SAT data we compute the mean SAT for each calendar day over the actually analyzed segment of the record. For instance, the results in Figs. 4a, b of the main paper were computed from SATA obtained in the following way: Take SAT values of all days with the calendar date January 1 from the analyzed period, i.e., from January 1, 1962 till January 1, 2009. Compute the mean of these values and subtract it from the SAT values of all days with the calendar date January 1. Repeat the process for all other calendar days.

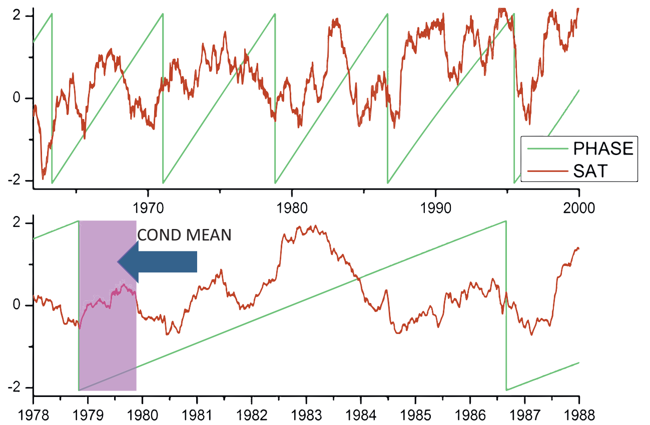

FIG. 1. (top) Example of smoothed (1-year moving average) SATA time series with the phase of the 8-year cycle derived from CCTW and (bottom) a zoomed view to one 8-year cycle: smoothed SATA (red), the phase of the 8-year cycle (green) and the first bin used for the computation of the conditional mean (violet). The phase is rescaled to the range of the smoothed SATA.

III. CONDITIONAL MEANS

In order to estimate the effect of the 8-year oscillatory mode, we employed the conditional mean technique, where the means (or other statistical measures such as variance or standard deviation) are taken conditionally on the phase of the 8-year cycle. The instantaneous phase $\phi_{8y}(t)$ of the 8-year cycle, estimated from the studied SATA time series using continuous complex wavelet transform (CCWT) (Fig. 1, green sawtooth pattern), is confined into the interval $(-\pi, \pi)$. The phase interval $(-\pi, \pi)$ (representing the full cycle) is equidistantly divided into 8 bins. Since the CCWT assigns a value of the phase $\phi_f(t)$ to each day (time sample t), each day of the recording period falls into one of the 8 phase bins.

The data falling into the first bin (which represents approximately the first year of each cycle) are stacked together over all cycles and the mean (or other statistical measure) is computed from that set. The same is done for the remaining 7 bins, resulting in the comparison of the statistics of temperature in various stages of the cycle.

---

[*] Department of Nonlinear Dynamics and Complex Systems, Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague, Czech Republic; Department of Atmospheric Physics, Faculty of Mathematics and Physics, Charles University in Prague, Czech Republic

[†] Department of Nonlinear Dynamics and Complex Systems, Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague, Czech Republic

[‡] Department of Mathematical Sciences, Atmospheric Science Group, University of Wisconsin-Milwaukee, Milwaukee, Wisconsin, USA

[§] Corresponding author; Department of Nonlinear Dynamics and Complex Systems, Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague, Czech Republic; mp@cs.cas.cz; http://www.cs.cas.cz/mp/

IV. CLIMATOLOGICAL AMPLITUDE OF THE ANNUAL CYCLE

A “climatological amplitude” which is not derived from CCWT, (depicted by dark blue dots in Fig. 1a of the main text) is defined not as the difference between mean seasonal (DJF vs. JJA) temperatures, but rather as the difference between the mean temperature of the warmest 25% (representing 91, not necessarily continuously connected days with the highest temperature in a particular year) and the coldest 25% of a particular year. This climatological amplitude is in good agreement with the amplitude $A_{1y}(t)$ acquired from the wavelet transform.

V. TEMPORAL EVOLUTION AND STATISTICAL TESTING

The temporal evolution of differences between the maximum and minimum conditional means was characterized using the sliding window of 16384 daily SATA samples. (The amplitude of the annual cycle was estimated from the raw – not anomalized – SAT data.) The analysis was performed as follows: take 16384 daily data points, use CCTW to extract the instantaneous phase from the SATA data, cut off the data and the phase in order to suppress the edge effects of CCTW (cone-of-influence) by 4 years at the beginning and 4 years at the end. Thus the effective window length used for estimating the conditional means is 36.86 years, although 44.86 years (16384 daily samples) are used. Then slide the window by one year and repeat until possible due to the time series length. The result for each window is plotted in the middle point of the window, thus the time range on the abscissa in the figures is shorter by 22 years from each side, comparing with the data range.

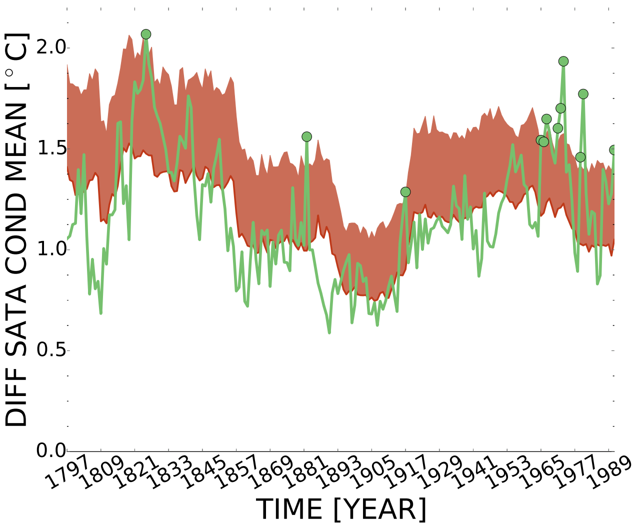

FIG. 2. Temporal evolution of the effect of the 8-year cycle on the daily SAT anomalies in Prague-Klementinum. Differences between the maximum and minimum SATA conditional means (green curve), tested using 1000 FT surrogates (means in red curve, 95th percentile of the surrogate distribution in light red). Windows with statistically significant differences are marked by the green dots, plotted in the middle of the window of the effective length 36.86 yr.

Randomization procedures known as “surrogate data” [Paluš, 2007] were used for the statistical testing. Always deseasonalized SAT data were used as the input to a randomization procedure. The deseasonalization for the daily SAT was performed by the computations of means and standard deviations (SD) for each calendar

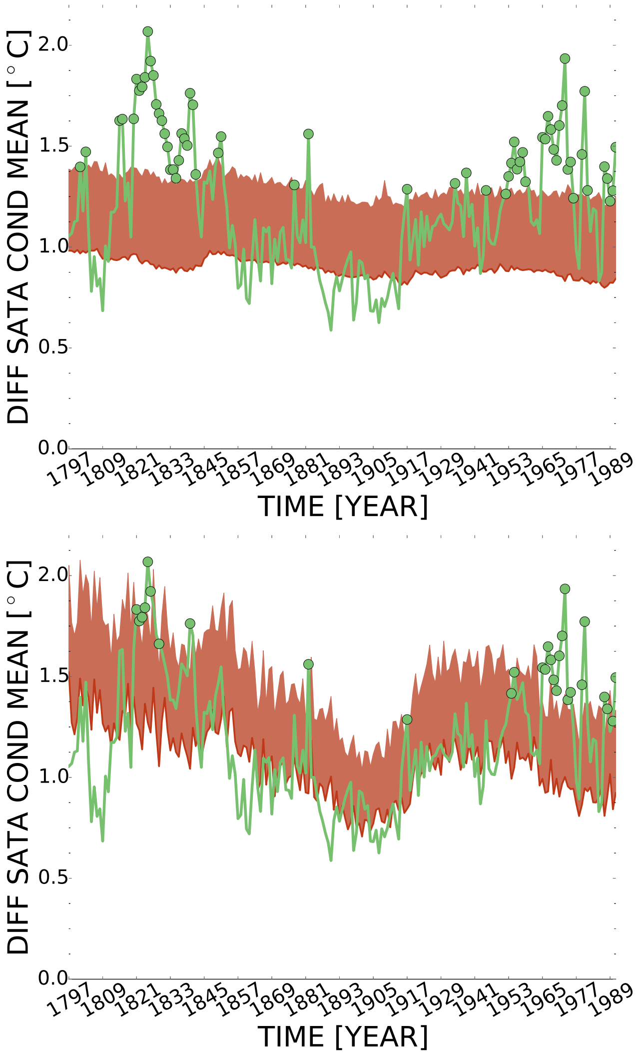

FIG. 3. Temporal evolution of the effect of the 8-year cycle on the daily SAT anomalies in Prague-Klementinum. Differences between the maximum and minimum SATA conditional means (green curve), tested using (top) 1000 AR surrogates and (bottom) MF surrogates (means in red curve, 95th percentile of the surrogate distribution in light red). Windows with statistically significant differences are marked by the green dots, plotted in the middle of the window of the effective length 36.86 yr.

day over the analyzed record. Then, the mean for January 1 was subtracted from all samples with the calendar date January 1, and the resulting anomalies were divided by the SD for this date. The same procedure was done for all calendar days. After the randomization, the seasonal means and variances were returned to the surrogate data, when a statistics for the SAT data were computed. For statistics based on the SAT anomalies, only the seasonal variances were returned to the surrogate data realizations (multiplying each sample by the SD obtained for its calendar day).

The statistical testing was done using three different types of surrogate data representing three different null hypotheses. The weakest null hypothesis supposes that no cycles are present in the data. It is represented by autoregressive (AR) surrogates generated as realizations of an AR process of order 1 (Allen and Smith [1996]). Fitting the AR model of order 1 to the deseasonalized data yields the AR coefficients and the respective residuals. Each surrogate data realization is generated using the estimated AR1 model with innovations (noise part) obtained as shuffled (randomly permuted) residuals.

The next type of the surrogate data used were the Fourier Transform (FT) surrogates (Theiler et al. [1992]) representing the null hypothesis of a linear stochastic process which has the same spectrum as the sample spectrum of tested experimental data, however, no interactions between different temporal scales (frequencies) can exist. In this case, the deseasonalized data are transformed to the Fourier domain (using the Fast Fourier Transform, FFT thereafter), the Fourier phases (phases of the complex Fourier coefficients) are randomized by adding a random number at each frequency, and transforming back to the temporal domain (using the inverse FFT). Note that the complex Fourier coefficients have not been changed in their magnitudes (the power spectra are preserved), but their Fourier phases were randomized. The randomization destroyed all non-linear properties of the original data, including possible relations between oscillatory modes with different frequencies.

The most sophisticated null hypothesis is represented by the multifractal (MF) surrogate data [Paluš, 2008] in which possible information transfer from larger to smaller scales, explained by random cascades on wavelet dyadic trees, is preserved. The surrogate data are generated as follows: deseasonalized data are decomposed using discrete wavelet transform. On each scale (level of the discrete wavelet) the wavelet coefficients are shuffled in the way preserving the multifractal spectrum of the original data, and finally, using the inverse discrete wavelet transform, the surrogate data are obtained. This randomization procedure is based on a model of a turbulent cascade in which a dynamical mode on the time scale S (having a main frequency f) influences (transfers energy or information onto) a mode on the scale $S/2$ (having the main frequency $2f$), but no other cross-scale relations are present.

In all three cases, the seasonal variance and mean are returned to the surrogate data after they are generated. In the following figures, as well as in Fig. 2 of the main text, we illustrate the surrogate range in each time window by its mean and the 95th percentile of the surrogate data distribution. The latter is obtained as follows: In each time window (16384 daily samples) one thousand surrogate data realizations are generated and processed in the same way as the original data. Then the surrogate values are sorted into the ascending order and the value of the 950th element is considered as the estimate of the 95th percentile of the surrogate distribution. If the data value is greater than the 95th percentile, we consider it statistically significant with $p < 0.05$. Note that the tests are not corrected for the multiplicity of tests, so we can encounter 5% of false positive results.

Figure 2 presents the sliding window analysis with the FT surrogates testing, while Fig. 3 provides the comparison of multifractal and autoregressive surrogate data testing. The differences between the maximum and minimum SATA conditional means (considered as a measure of the effect of the 8-year cycle on the SATA variability) are ranging from approximately 0.6 °C to a maximum about 2.2 °C. The effect of the 8-year cycle on the SATA variability reaches the values 1.5–2.2 °C in the first decades of the 19th century. Then it is weaker for almost a century and increases again over 1.5 °C (up to 1.7 °C) from the 1950's, when its values are mostly statistically significant. The FT surrogate data provide the most conservative null hypothesis and are used in all subsequent statistical tests.

In order to understand what the significant results in the above statistical tests mean, let us return to Fig. 4 of the main text. The difference approximately 1.5 °C between the maximum and minimum conditional SATA means (Fig. 4a of the main text) is a measure of the average (the average over the period 1962–2009) effect of the cycle with the period close to 8 years on the overall temperature variability, reflected in annual SATA means. This value (1.5 °C) is illustrated by the black vertical bar in Fig. 4c of the main text. The grey bars in the same figure illustrate the histogram of the same quantity estimated from 1000 realizations of the FT surrogate data. The figure shows that the value 1.5 °C lies on the tail of the surrogate distribution, while its statistical significance is quantitatively estimated using the sorted surrogate values as it was described above. The value 1.5 °C is greater than the 98th percentile of the surrogate distribution, i.e. it is statistically significant with $p < 0.02$. It means that the probability of occurrence of the difference 1.5 °C between the maximum and minimum conditional means as a result of random variability is less than 0.02. The cycle with the period of approximately 8 years, however, is a part of the overall SATA variability. Its amplitude, estimated using the regressed CCWT component, is less than 0.5 °C and is well reproduced in the FT surrogate data (Fig. 4d of the main text). It is not reproduced exactly since the FT surrogates reproduce exactly the amplitudes of harmonic oscillations with well defined, constant frequency. The 7–8 year cycle has a fluctuating frequency which is spread over a number of frequency bins in the Fourier spectrum. The inverse FFT could exactly reproduce the amplitude of the 8-year cycle only if both the magnitudes and phases of the Fourier coefficients were preserved. The randomization of the Fourier phases leads to fluctuations of the 8-year cycle amplitude. These fluctuations are small (Fig. 4d of the main text) and cannot be the source of the significant result in the test in Fig. 4c of the main text. In summary, these results support the hypothesis that the 1.5 °C difference in the SATA annual means is not a result of random variability, neither can be explained by an 8-year component, linearly added to a background variability. Since the amplitude of the 8-year cycle itself is smaller than 0.5 °C, the difference 1.5 °C of the annual means during the 8-year cycle is mainly the result of the cross-scale interactions of the 8-year cycle with the variability on shorter time scales.

VI. SEASONALITY

In order to account for seasonality, the data was divided to four seasons: spring - MAM, summer - JJA, autumn - SON and winter - DJF; and each seasonal data set underwent the evolution of conditional means difference analysis. The result for the summer and winter seasons in the Prague-Klementinum data is presented in Fig. 4. As seen in the figure, the effect of the 8-year cycle on the overall SATA variability is markedly stronger in the winter season where the differences between the cold and warm bins are ranging over 4 °C in the 1950's, and over 5 °C in the 1820's and are almost everywhere statistically significant (with the exception of the period 1910–1940) with respect to the surrogate testing using 1000 FT surrogates. On the other hand, the effect on the summer season is quite weak, the differences range up to about 1.5 °C in the majority of the record, with almost no statistically significant results, i.e., during the summer season the effect of the 8-year cycle is not discernible from random temperature variability. The other seasons, spring and autumn (not shown), resemble the whole year analysis (Fig. 2) with differences up to 2 °C and being significant mostly in the last decades.

VII. SPATIAL PATTERN OF THE INFLUENCE OF THE 7–8 YEAR CYCLE ON SAT VARIABILITY IN EUROPE

The spatial pattern of the influence of the cycle with the period close to 8 years on the variability of the amplitude of the annual cycle (AAC) in the SAT is mapped in Fig. 3 of the main text and the effect on the overall SATA variability is illustrated in Fig. 5 of the main text. There is a difference in the localization of maxima, where the maximum in the case of the AAC conditional means differences is located in Norway/Sweden and the maximum for the overall variability in Finland/Russia. Both maxima, however, ale located to the north of the 60th parallel. Otherwise the patterns in Figs. 3 and 5 are similar and in central and eastern Europe they resemble mountain topography, as we state in the main text. The pattern is similar due to the individual scales of the figures. The individual scales are necessary since the sizes of the effect on the AAC and on the overall variability are different, however, in both cases the effect has similar geographical distribution.

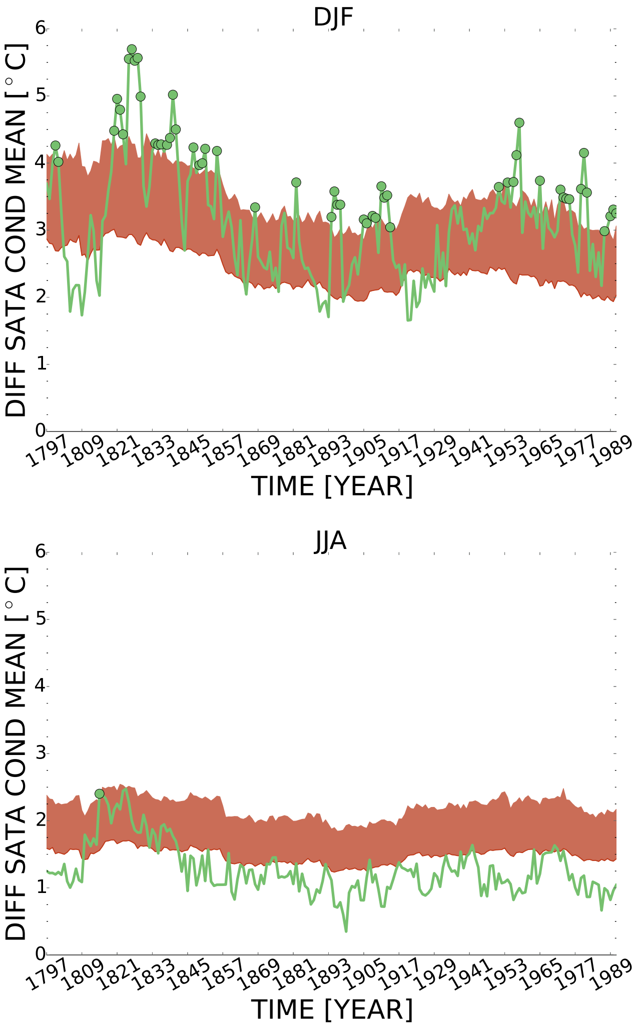

FIG. 4. Temporal evolution of the effect of the 8-year cycle on the daily temperature anomalies in Prague-Klementinum in different seasons. The differences between the maximum and minimum SATA conditional means (green curve) for (top) winter - DJF season and (bottom) summer - JJA season, tested using 1000 FT surrogates (means in red curve, 95th percentile of the surrogate distribution in light red). Windows with statistically significant differences are marked by the green dots, plotted in the middle of the window of the effective length 36.86 yr.

The areas where our measures quantifying the effect of

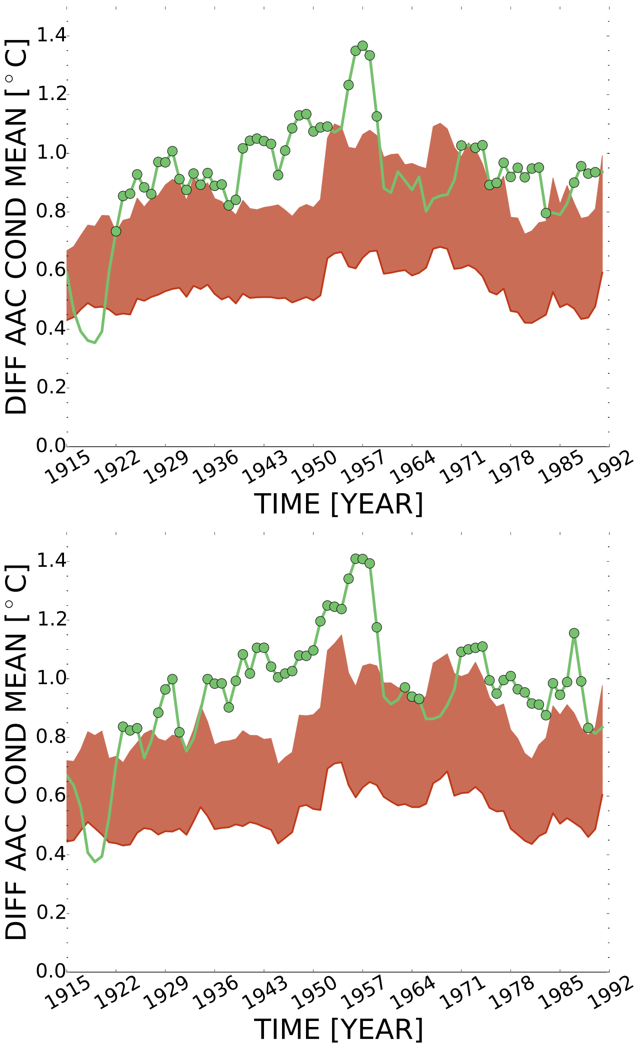

FIG. 5. Temporal evolution of the effect of the 8-year cycle on the amplitude of the annual cycle in (top) Hamburg - Fuhlsbüttel and (bottom) Potsdam daily SAT. Differences between the maximum and minimum AAC conditional means (green curve), tested using 1000 FT surrogates (means in red curve, 95th percentile of the surrogate distribution in light red). Windows with statistically significant differences are marked by the green dots, plotted in the middle of the window of the effective length 36.86 yr.

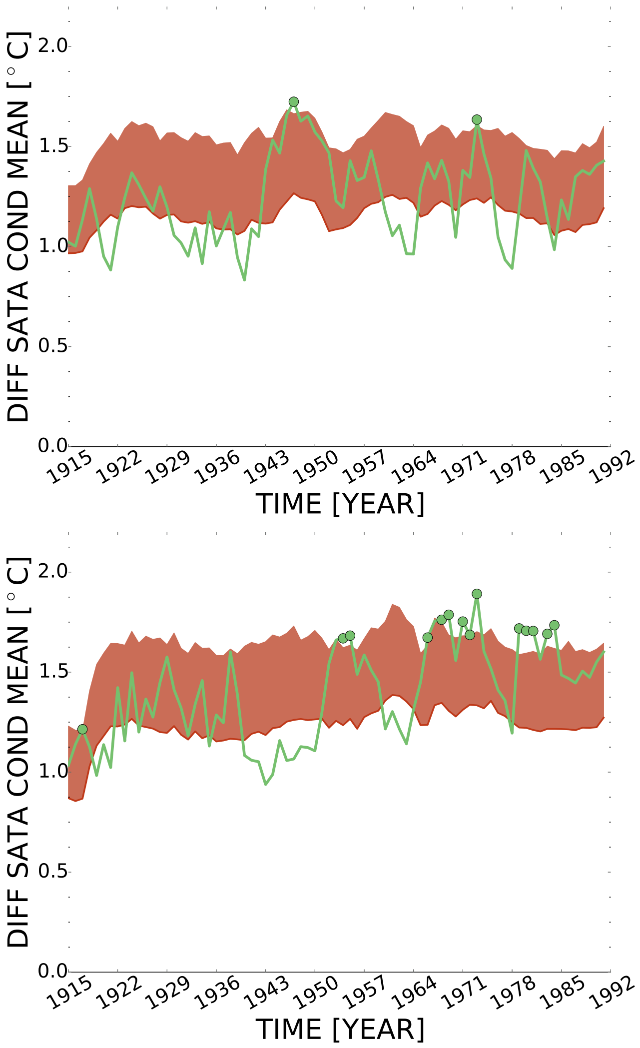

FIG. 6. Temporal evolution of the effect of the 8-year cycle on the daily SAT anomalies in (top) Hamburg - Fuhlsbüttel and (bottom) Potsdam. Differences between the maximum and minimum SATA conditional means (green curve), tested using 1000 FT surrogates (means in red curve, 95th percentile of the surrogate distribution in light red). Windows with statistically significant differences are marked by the green dots, plotted in the middle of the window of the effective length 36.86 yr.

the 7–8 year cycle are statistically significant are marked by the hatch pattern in Figs. 3 and 5 of the main text.

We can see that these areas are not the same. This is the result of the well-known problem that there is no straightforward relation between the statistical significance and the size of the effect. Having the figures in individual scales we can see that the effect is similar in the relative sense. The statistical significance is influenced not only by the effect itself, but also by the background variability. The latter is apparently larger in the overall variability than in the variability of AAC and therefore we observe the smaller area of the statistically significant results in the overall variability than in the AAC variability.

VIII. SAT DATA FROM OTHER STATIONS

For comparison we present the results from the SAT data from two German stations - Hamburg - Fuhlsbüttel and Potsdam which, however, are shorter than the record from Prague-Klementinum, analyzed in the main text. Firstly, the German station data underwent the analysis of the response of the amplitude of the annual cycle (AAC) to the driving of the 8-year cycle's phase. In agreement with the ECA&D reanalysis dataset, it can be

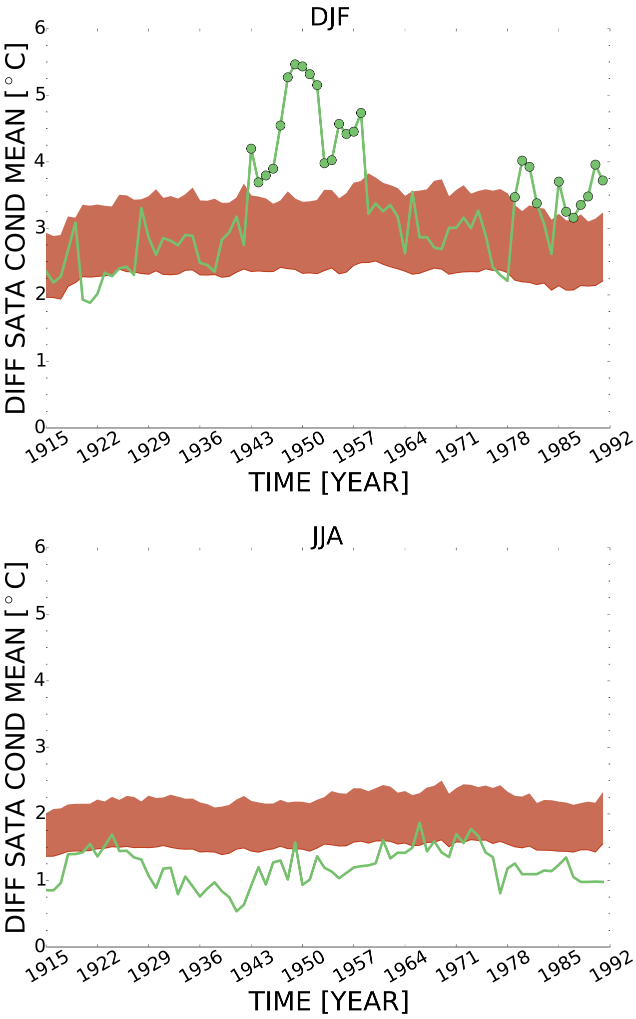

FIG. 7. Temporal evolution of the effect of the 8-year cycle on the daily SAT anomalies in Hamburg - Fuhlsbüttel in different seasons. Differences between the maximum and minimum SATA conditional means (green curve) for (top) winter - DJF season and (bottom) summer - JJA season, tested using 1000 FT surrogates (means in red curve, 95th percentile of the surrogate distribution in light red). Windows with statistically significant differences are marked by the green dots, plotted in the middle of the window of the effective length 36.86 yr.

seen in Fig. 5 that the evolution of the AAC differences driven by 8-year cycle resemble each other and also resemble the AAC difference evolution in Prague (Fig. 2 in the main article), with higher values reaching 1.4 °C around 1950's followed by rather an obvious descend. Note that the AAC differences are statistically significant (FT surrogate tests) in almost all time windows.

The overall variability of the surface air temperature anomalies from the two German stations was also tested using 1000 FT surrogates (Fig. 6). The magnitude of the differences between the maximum and minimum SATA conditional means in Potsdam peaks over 1.7 °C and is mostly statistically significant from the 1950's. This is not observed in Hamburg, however the seasonal SATA responses to the 8-year cycle in the two German stations (Fig. 7 and 8) resemble that of Prague (Fig. 4) in the 20th century. The winter - DJF - season SATA differences range up to 5 °C and are statistically significant in the 1950's (in Hamburg also in the 1940's), and again from the 1980's, while the summer - JJA - season differences are lower (around 1.5 °C) and not significant.

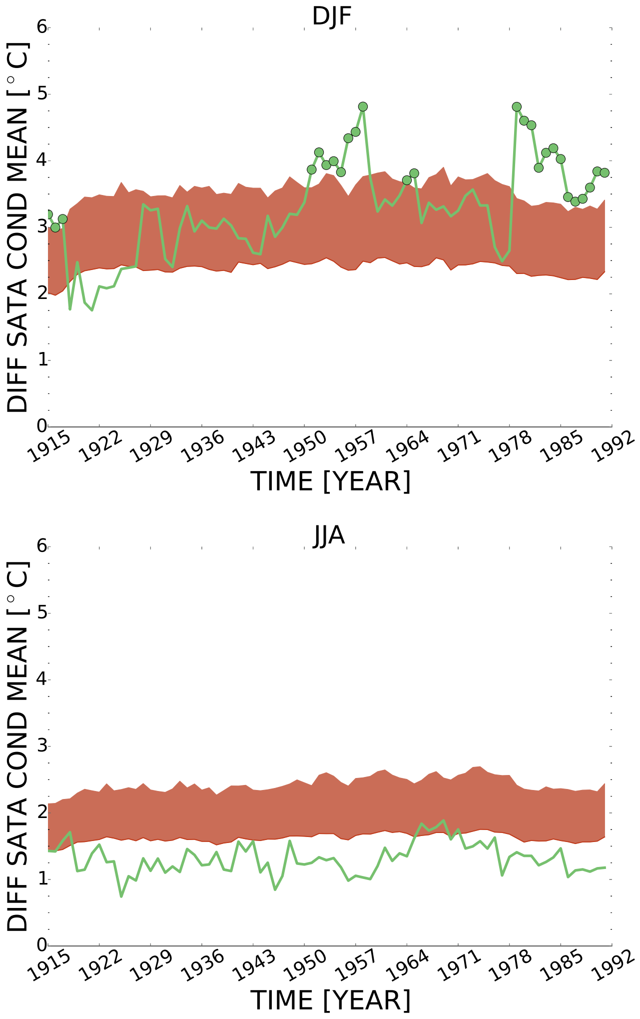

FIG. 8. Temporal evolution of the effect of the 8-year cycle on the daily SAT anomalies in Potsdam in different seasons. Differences between the maximum and minimum SATA conditional means (green curve) for (top) winter - DJF season and (bottom) summer - JJA season, tested using 1000 FT surrogates (means in red curve, 95th percentile of the surrogate distribution in light red). Windows with statistically significant differences are marked by the green dots, plotted in the middle of the window of the effective length 36.86 yr.

[1] Allen, M. R., and L. A. Smith (1996), Monte Carlo SSA: Detecting irregular oscillations in the presence of colored noise, *J. Climate*, *9*(12), 3373–3404, doi:10.1175/1520-0442(1996)009¡3373:MCSDIO¿2.0.CO;2.

[2] Paluš, M. (2008), Bootstrapping multifractals: Surrogate data from random cascades on wavelet dyadic trees, *Phys. Rev. Lett.*, *101*, 134,101, doi: 10.1103/PhysRevLett.101.134101.

[3] Paluš, M. (2007), From nonlinearity to causality: statistical testing and inference of physical mechanisms underlying complex dynamics, *Contemp. Phys.*, *48*(6), 307–348, doi:10.1080/00107510801959206.

[4] Theiler, J., S. Eubank, A. Longtin, B. Galdrikian, and J. Doyne Farmer (1992), Testing for nonlinearity in time series: The method of surrogate data, *Physica D*, *58*(1), 77–94, doi:10.1016/0167-2789(92)90102-S.