Does sunspot numbers cause global temperatures? A reconsideration using non-parametric causality tests

Hossein Hassani, Xu Huang, Rangan Gupta, Mansi Ghodsi

PII: S0378-4371(16)30135-2
DOI: http://dx.doi.org/10.1016/j.physa.2016.04.013
Reference: PHYSA 17093

To appear in: *Physica A*

Received date: 9 January 2016
Revised date: 9 March 2016

Does Sunspot Numbers Cause Global Temperatures? A Reconsideration Using Non-parametric Causality Tests[*]

Hossein Hassani[1], Xu Huang[2], Rangan Gupta[3], Mansi Ghodsi[1]

1- Institute for International Energy Studies (IIES), Tehran 1967743 711, Iran

2- The Statistical research Centre, Bournemouth University, UK

3- Department of Economics, University of Pretoria, Pretoria, 0002, South Africa

ABSTRACT

In a recent paper, Gupta et al., (2015), analyzed whether sunspot numbers cause global temperatures based on monthly data covering the period 1880:1-2013:9. The authors find that standard time domain Granger causality test fails to reject the null hypothesis that sunspot numbers does not cause global temperatures for both full and sub-samples, namely 1880:1-1936:2, 1936:3-1986:11 and 1986:12-2013:9 (identified based on tests of structural breaks). However, frequency domain causality test detects predictability for the full-sample at short (2 to 2.6 months) cycle lengths, but not the sub-samples. But since, full-sample causality cannot be relied upon due to structural breaks, Gupta et al., (2015) concludes that the evidence of causality running from sunspot numbers to global temperatures is weak and inconclusive. Given the importance of the issue of global warming, our current paper aims to revisit this issue of whether sunspot numbers cause global temperatures, using the same data set and sub-samples used by Gupta et al., (2015), based on an nonparametric Singular Spectrum Analysis (SSA)-based causality test. Based on this test, we however, show that sunspot numbers have predictive ability for global temperatures for the three sub-samples, over and above the full-sample. Thus, generally speaking, our non-parametric SSA-based causality test outperformed both time domain and frequency domain causality tests and highlighted that sunspot numbers have always been important in predicting global temperatures.

Keywords: Causality; Singular Spectrum Analysis; frequency domain; global temperatures predictability; sunspot numbers

JEL classification: C32

[*] We would like to thank two anonymous referees for many comments. However, any remaining errors are solely ours.

1. Introduction

Global warming, i.e., rising temperature of the earth's surface, is undoubtedly the biggest topic of research amongst researchers working on environment. While, analyzing the impact of global warming cannot be ignored, but what factors drive it is perhaps more important, as it not only allows us to predict global warming, but also takes measures to control it. It is quite well-accepted that global warming is due to greenhouse gases, additionally, there is a large literature[1] that relates the same with solar activity. However, the evidence from this literature is, at best, mixed. While there are studies (see for example, Lean and Rind, 1998, 2009; Scafetta and West, 2003, 2005; Scafetta et al., 2004; Scafetta, 2009, 2011; Folland et al., 2013; Zhou and Tung, 2013) that find significant relationships between solar radiation and global temperatures, one hand. On the other hand, there are some authors who claim that the two variables are unrelated (see for example, Pittock, 1978, 1983, 2009; Love et al., 2011; Usoskin, et al., 2004). Thus, there is no clear-cut consensus about the possibility of a relationship between solar irradiance and global temperatures (Gil-Alana et al., 2014).

Against this backdrop, using sunspot numbers as a proxy for solar activity, Gupta et al., (2015), recently analyzed whether sunspot numbers cause global temperatures based on monthly data covering the period 1880:1-2013:9. However, at this stage, it is important to point out, as indicated by Scafetta (2014), sunspot numbers can only be considered as a "partial proxy" for solar activity. This is because time intervals between major solar flares, cosmic ray records, ACRIM composite of total solar irradiance satellite measurement, multi-scale thermal models of several total solar irradiances, and solar and astronomical oscillations are also possible, and perhaps, better proxies for solar activity than sunspot numbers. In addition, one must be cautious in suggesting that sunspot numbers are linearly and positively related to solar activity due to the intrinsic complexity of solar

[1] The reader is referred to Gray et al. (2010) and Gupta et al., (2015) for further details.

dynamics and of its multiple coupled phenomena, as discussed in detail in Scafetta (2014). Gupta et al., (2015) find that standard time domain Granger causality test fails to reject the null hypothesis that sunspot numbers does not cause global temperatures for both full and sub-samples, namely 1880:1-1936:2, 1936:3-1986:11 and 1986:12-2013:9 (identified based on tests of structural breaks). However, frequency domain causality test detects predictability for the full-sample at short (2 to 2.6 months) cycle lengths. Interestingly however, the study could not detect any causality for the sub-samples. Gupta et al., (2015) thus, highlights the importance of analysing causality using the frequency domain tests, which, unlike the time domain Granger causality test, allows one to decompose causality by different time horizons, and hence, possibly detect predictability at certain cycle lengths even when the time domain causality test might fail to pick up any causality. However, given that there exists structural breaks in the sample, Gupta et al., (2015), suggests that the relationship could be spurious based on a full-sample analysis, since a full-sample analysis assumes stability of the parameters of a VAR, which is clearly not the case in the presence of breaks, and which is also vindicated by the fact that there is no evidence of causality over the sub-samples.

Given the importance of the issue of global warming, and more importantly the lack of evidence in favor of sunspot numbers leading to global temperatures in linear models, our current paper aims to revisit this issue of whether sunspot numbers cause global temperatures, using the same data set and sub-samples used by Gupta et al., (2015), based on Singular Spectrum Analysis (SSA) technique, which is a new nonparametric technique known for both time series analysis and forecasting (as discussed further in Hassani, 2007; Hassani and Thomakos, 2010; Hassani et al., 2009, 2010, 2013a, 2013b; Hassani and Mahmoudvand, 2013). The reason behind using a nonparametric technique is to capture possible nonlinearities that could exist in the data generating processes of the global temperatures and sunspots individually (Scafetta, 2014), as well as, in the relationship between global temperatures and sunspot activity, for instance due to the structural breaks detected by Gupta et al., (2015). The SSA being a nonparametric method captures the possible nonlinearities using a data-driven approach, without specifying any known functional nonlinear model to the relationship, which in turn, could be incorrectly specified in the first place, just like the linear model, on which time domain and frequency domain Granger causality tests are based on. Further, as pointed out by Aguirre et al., (2008), the difficulties encountered in modeling sunspot numbers and global temperature data are due to the apparent nonstationarity property of the series and the complex dynamic fluctuations in the cycle amplitude of the sunspot number series. In other words, these complexities could be driving the mixed results discussed above in terms of the relationship between these two variables. In light of this, the importance of the nonparametric SSA-based causality cannot be underestimated, which besides being a nonlinear data-driven approach, also does not require pretesting to ensure that the variables under consideration is stationary (Hassani, 2007; Hassani and Thomakos, 2010; Hassani et al., 2009, 2010, 2013a, 2013b; Hassani and Mahmoudvand, 2013).

The paper is structured as follows: Given that time and frequency domain causality tests were already discussed in Gupta et al., (2015), the details of the frequency domain causality test have been relegated to the appendix for the sake of completeness, with Section 2 introducing the SSA-based causality test (following the works of Hassani and Mahmoudvand, 2013). Section 3 presents the data and empirical results. Finally, Section 4 concludes.

2. Methodology: The SSA-based causality test (MSSA)

Multivariate singular spectrum analysis (MSSA) is an extension of the standard Singular Spectrum Analysis (SSA) to the case of multivariate time series (Hassani et al., 2013), in which SSA is a relatively new nonparametric technique known for both time series analysis and forecasting, detail description can be found in (Hassani, 2007). After Broomhead and King (1986) theoretically proposed the MSSA technique in the context of nonlinear dynamics for the first time, it has been widely applied on a range of different fields and a multitude of fairly precise results proved it as powerful and applicable technique, numerous applications and examples can be found in (Hassani, 2007; Hassani et al., 2009, 2010, 2013a, 2013b; Ghodsi et al., 2010; Hassani and Thomakos, 2010; Hassani and Mahmoudvand, 2013; Sanei and Hassani, 2015). From the perspective of MSSA, two main concerns that make the problem more complex are: i) similarity and orthogonality among series play an important rule for selecting the window length L and the number of eigenvalues r, and ii) MSSA deals with a block trajectory Hankel matrix with special features rather than one simple Hankel matrix (Hassani and Mahmoudvand, 2013). Briefly descriptions of MSSA and causality criteria are listed in following subsections.

2.1 Algorithm Description of MSSA

In this subsection of brief description of MSSA algorithm, we mainly follow the paper by Hassani and Mahmoudvand (Hassani and Mahmoudvand, 2013). Consider M time series with different series length N_i as $Y_{N_i}^{(i)} = \left(y_1^{(i)}, \ldots, y_{N_i}^{(i)}\right)(i = 1, \ldots, M)$. By the embedding process that transfer a one-dimensional series $Y_{N_i}^{(i)} = (y_1^{(i)}, \ldots, y_{N_i}^{(i)})$ in to a multidimensional matrix $[X_1^{(i)}, \ldots, X_{K_i}^{(i)}]$ with vectors$X_j^{(i)} = (y_j^{(i)}, \ldots, y_{i+L_i-1}^{(i)})^T \in R^{L_i}$, where $L_i (2 \leq L_i \leq N_i)$ is the window length for each series with length N_i and $K_i = N_i - L_i + 1$. We can then get the trajectory matrix $X^{(i)} = [X_1^{(i)}, \ldots, X_{K_i}^{(i)}] = (x_{mn})_{m,n=1}^{L_i,K_i}$ after this step. The above procedure for each series separately provides M different $L_i \times K_i$ trajectory matrices $X^{(i)}(i = 1, \ldots, M)$. To construct a block Hankel matrix, we need to have $L_1 = L_2 = \cdots = L_M = L$. Therefore, we have different values of K_i and series length N_i, but similar L_i. The result of this step is $X_H = [X^{(1)}: X^{(2)}: \ldots : X^{(M)}]$. Hence, the structure of the matrix

$X_H X_H^T$ is as follows: $X_H X_H^T = X^{(1)}X^{(1)^T} + \cdots + X^{(M)}X^{(M)^T}$ and the sum of $X^{(i)}X^{(i)^T}$ provides the new block Hankel matrix, which can be subsequently converted to a time series.

2.2 Causality criteria based on forecasting accuracy

Granger (1969) proposed and formalized the causality concept to address the question that whether one variable can help in predicting another. The criterion we use is based on out-of-sample forecasting, which is very common in the framework of Granger causality. Here, we compare the forecast values obtained by the univariate procedure, SSA and MSSA. If the forecasting errors using MSSA are significantly smaller than those of univariate SSA, we can conclude that there is a causal relationship between these series. Brief introduction is listed below which we mainly follow Hassani et al. (2010)[2].

Let us consider the procedure for constructing vectors of forecasting error for out-of-sample tests in a two variable case X_T and Y_T by both univariate and multivariate SSA techniques respectively. In the first step we divide the series $X_T = (x_1, \ldots, x_T)$ into two separate subseries X_R and X_F : $X_T = (X_R, X_F)$, where $X_R = (x_1, \ldots, x_R)$ and $X_F = (x_{R+1}, \ldots, x_T)$. Same procedure is conducted for Y_T. The subseries X_R and Y_R are used in the reconstruction step to provide the noise-free series $\widetilde{X}_R$ and $\widetilde{Y}_R$. The noise-free series are then used for forecasting the subseries X_F and $Y_F$with the help of the recursive formula using SSA and MSSA respectively. For variable X_T, two different forecasting values of $\widehat{X}_F = (\hat{x}_{R+1}, \ldots, \hat{x}_T)$ by SSA and MSSA are then used for computing the forecasting errors accordingly, which will be the same process for variable Y_T. Therefore, in a multivariate system like this, the vectors of forecasts obtained can be used in computing the forecasting accuracy and therefore examining the association between the two variables.

[2] The readers are referred to Hassani et al. (2010) for more details.

The length of out-of-sample does not have specific limitation, generally considering the simulation scenario, the length of time series for reconstruction will take 2/3 of the whole series and the rest 1/3 is considered as out-of-sample for constructing forecasting error. The separate point to define the out-of-sample size for different series can be chosen respectively, whilst it is important that when it goes to comparing the performances of different techniques based on constructed forecasting error of one specific series , the sizes of reconstruction and out-of-sample for all techniques should be identical. In addition, the choices of window length L and the referring options of numbers of eigenvalues r should also be carefully evaluated in practice of SSA-based causality test respectively. In order to conduct the most accurate causality detection results, all the possibilities of L and its referring choices of r should be applied for both univariate SSA and MSSA processes, then the optimal ones with best performance of forecasting will be chosen to construct the finally causality detection procedure.

Therefore, here we define the criterion $F_{X|Y} = \Delta X_F|Y / \Delta X_F$ corresponding to the forecast of the series X_T in the presence of the series Y_T. If $F_{X|Y}$ is small, then having information obtained from the series Y can help us to have better forecasts of the series X. If $F_{X|Y} < 1$, we conclude that the information provided by the series Y can be regarded as useful or supportive for forecasting the series X. Alternatively, if the values of $F_{X|Y} \geq 1$, then either there is no detectable association between X and Y or the performance of the univariate SSA is better than of the MSSA (this may happen, for example, when the series Y has structural breaks misdirecting the forecasts of X).

3. Data and empirical results

The data are at monthly frequency for global land-ocean temperatures (GT) and sunspot numbers (SS), and cover the period from January 1880 to September 2013, with the start and end-points being maintained the same as that of Gupta et al., (2015) for the sake of comparison. Empirical results, for the time-domain causality and the SSA tests listed in this section are conducted by R programming based on source code, while the frequency domain causality tests are performed in GAUSS. In terms of the data, the global temperatures were obtained from the National Aeronautics and Space Administration's (NASA), Goddard Institute for Studies (GISS) (http://data.giss.nasa.gov/gistemp), while the sunspot numbers were obtained from the Solar Influences Data Analysis Centre (SIDC: http://www.sidc.be/sunspot-data). The data for temperatures are anomalies relative to the base period 1951-1980. Figures 1(a) and 1(b) plot the two variables. As can be seen, the plot of the global temperature seems to be non-stationary, though it could well be trend-stationary, while that of the sunspot looks stationary with a cyclical pattern completed at about 10/11 years.

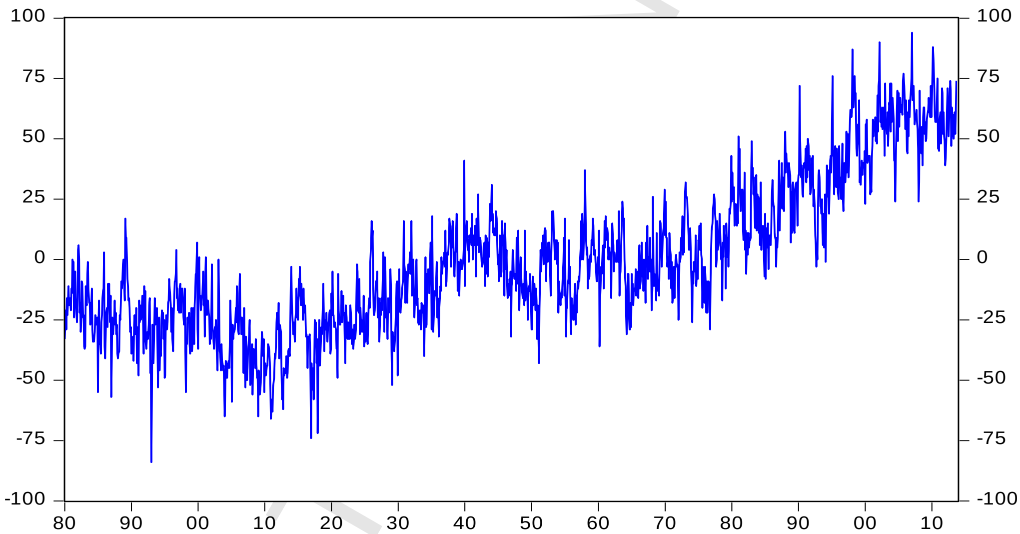


Figure 1(a): Plot of Global Temperatures (1880:1-2013:9)

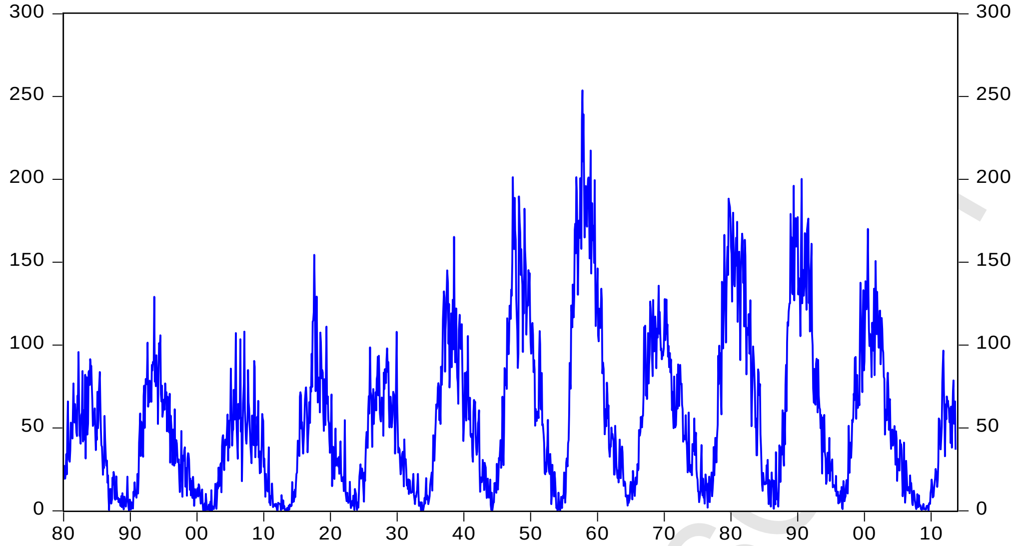


Figure 1(b): Plot of Sunspot Numbers (1880:1-2013:9)

As in Gupta et al., (2015), we start off with unit root tests to verify whether the two series are stationary I(0) or not. As can be seen, based on the Kwiatkowski-Phillips-Schmidt-Shin (1992, KPSS), augmented Dickey-Fuller (1981, ADF), Dickey-Fuller test with Generalised least Squares detrended residuals (Elliot et al., 1996, DF-GLS) Phillips and Perron (1988, PP), and Ng and Perron (2001, NP) unit root tests, the null of a unit root is overwhelmingly rejected (except for KPSS test the null of being stationary, it cannot be overwhelmingly rejected), for the total sample of SS. However, for total sample of GT, while all the tests support that the variable is trend-stationary, the ADF and DF-GLS test tends to suggest non-stationarity of the series when the unit root test-equation has only a constant (or neither a constant and trend in case of the ADF test). The PP and the NP tests, though, indicate stationarity even under the assumption of constant only (and neither a constant and trend in case of the PP test).

As Gupta et al., (2015) points out, among the unit root tests conducted, the NP test is believed to have overwhelmingly stronger power relative to the other tests, and hence, one would tend to rely on the results from this test. Also, given the nature of GT, it is evident that the unit root equation should in fact include a trend, while that for SS, it should only be with a constant. In light of this as in Gupta et al., (2015), we can conclude that GT is stationary as well, and hence, we do not need to transform the data further for either GT or SS. In addition, we do not need to account for the possibility cointegration, and hence error-correction, between the two variables.

Note that Gupta et al., (2015) applied Bai and Perron's (2003) sequential and repartition tests of multiple structural breaks on the GT equation of the VAR (4) model comprising of GT and SS. The GT equation on which the tests were performed involved a constant and four lags each of GT and SS. Now since structural breaks were detected in the full-sample at 1936:3 and 1986:12, we also conducted the unit root tests over the sub-samples, which are reported in Table 1. In general, for subsample A and subsample B we have overwhelming evidence of stationary (especially based on the results of NP test, which mentioned above that have stronger power compared to the other tests). For sample C, while GT is found to be stationary in general, the evidence of stationarity, surprisingly, is quite weak for SS, barring the PP and NP tests, at 10 % level of significance. But given the cyclical pattern of SS, it is very difficult to believe that the variable is non-stationary. In fact, we can conclude that the variable is weakly stationary for sub-sample C. In summary, for the full sample and all sub-samples, both variables are stationary.

Table 1: Unit Root Test Results

Sample Size	Series	Methods	None		Intercept		Intercept and Trend	
			Level	Decision	Level	Decision	Level	Decision
Total Sample (1605 Obs) 1880:1-2013:9	GT	KPSS	-----------	-----------	4.136*** (31)	**I(1)**	0.638***(30)	**I(1)**
		ADF	-1.620 (17)	**I(1)**	-1.598 (17)	**I(1)**	-3.707** (24)	I(0)
		PP	-6.231*** (12)	I(0)	-6.222*** (12)	I(0)	-18.761*** (23)	I(0)
		DF-GLS	-----------	-----------	-1.539 (6)	**I(1)**	-6.868***(3)	I(0)
		NP	-----------	-----------	-33.684***(12)	I(0)	-537.250*** (23)	I(0)
	SS	KPSS	-----------	-----------	0.494**(31)	**I(1)**	0.119 (31)	I(0)
		ADF	-2.499**(3)	I(0)	-4.055***(3)	I(0)	-4.109***(3)	I(0)
		PP	-3.457*** (13)	I(0)	-6.488*** (12)	I(0)	-6.800*** (13)	I(0)

		DF-GLS	-----------	-----------	-3.303***(3)	I(0)	-4.029***(3)	I(0)
		NP	-----------	-----------	-52.985***(12)	I(0)	-81.5259***(13)	I(0)
Subsample A (674 Obs) 1880:1-1936:2	**GT**	KPSS	-----------	-----------	0.455*(19)	**I(1)**	0.430***(19)	I(0)
		ADF	-2.710***(3)	I(0)	-7.207***(2)	I(0)	-7.228***(2)	I(0)
		PP	-4.313***(2)	I(0)	-13.397***(14)	I(0)	-13.424***(14)	I(0)
		DF-GLS	-----------	-----------	-6.325***(2)	I(0)	-7.076***(2)	I(0)
		NP	-----------	-----------	-234.149***(14)	I(0)	-275.304***(14)	I(0)
	SS	KPSS	-----------	-----------	0.053(21)	I(0)	0.051(21)	I(0)
		ADF	-1.819*(3)	**I(1)**	-3.451***(3)	I(0)	-3.447**(3)	I(0)
		PP	-3.226***(18)	I(0)	-6.075***(8)	I(0)	-6.075***(8)	I(0)
		DF-GLS	-----------	-----------	-3.043***(3)	I(0)	-3.322**(3)	I(0)
		NP	-----------	-----------	-52.499***(8)	I(0)	-57.985***(8)	I(0)
Subsample B (609 Obs) 1936:3-1986:11	**GT**	KPSS	-----------	-----------	0.794***(17)	**I(1)**	0.321***(16)	**I(1)**
		ADF	-7.121***(1)	I(0)	-7.211***(1)	I(0)	-7.515***(1)	I(0)
		PP	-12.979***(13)	I(0)	-13.102***(13)	I(0)	-13.678***(13)	I(0)
		DF-GLS	-----------	-----------	-3.287***(2)	I(0)	-6.454***(1)	I(0)
		NP	-----------	-----------	-92.270***(13)	I(0)	-229.775***(13)	I(0)
	SS	KPSS	-----------	-----------	0.061(18)	I(0)	0.052(18)	I(0)
		ADF	-1.690*(2)	**I(1)**	-2.720*(2)	**I(1)**	-2.741(2)	**I(1)**
		PP	-1.932*(11)	**I(1)**	-3.600***(2)	I(0)	-3.614**(2)	I(0)
		DF-GLS	-----------	-----------	-2.718***(2)	I(0)	-2.754*(2)	**I(1)**
		NP	-----------	-----------	-24.056***(2)	I(0)	-24.089***(2)	I(0)
Subsample C (322 Obs) 1986:12-2013:9	**GT**	KPSS	-----------	-----------	1.651***(14)	**I(1)**	0.126*(12)	I(0)
		ADF	-0.682 (3)	**I(1)**	-4.604***(1)	I(0)	-6.618***(1)	I(0)
		PP	-1.203 (26)	**I(1)**	-6.835***(8)	I(0)	-9.997***(8)	I(0)
		DF-GLS	-----------	-----------	-1.178*(3)	**I(1)**	-5.614***(1)	I(0)
		NP	-----------	-----------	-16.711***(8)	I(0)	-106.142***(8)	I(0)
	SS	KPSS	-----------	-----------	0.534**(15)	**I(1)**	0.093 (14)	I(0)
		ADF	-0.936 (3)	**I(1)**	-1.812 (3)	**I(1)**	-2.415 (3)	**I(1)**
		PP	-1.497(12)	**I(1)**	-2.761*(2)	**I(0)**	-3.394*(2)	**I(0)**
		DF-GLS	-----------	-----------	-1.138(3)	**I(1)**	-1.356(3)	**I(1)**
		NP	-----------	-----------	-6.718*(2)	**I(0)**	-8.802(2)	**I(1)**

Notes: *, ** and *** indicates significance at the 10%, 5% and 1% level, respectively. The critical values are as follows:

- None: -2.566, -1.941 and -1.616 for ADF and PP at 1%, 5% and 10% level of significance, respectively.
- Intercept: -3.434, -2.863 and -2.567 (-2.566, 1.941, 1.617) [-13.8, -8.1 and -5.7] {0.739, 0.463, 0.347} for ADF and PP (DF-GLS) [NP] {KPSS} at 1%, 5% and 10% level of significance, respectively.
- Intercept and Trend: -3.963, -3.412 and -3.128 (3.48, 2.89, 2.57) [-23.80, -17.3 and -14.2] {0.216, 0.146, 0.119} for ADF and PP (DF-GLS) [NP] {KPSS} at 1%, 5% and 10% level of significance, respectively.

Numbers in parentheses for ADF, PP and DF-GLS tests indicates lag-lengths selected based on the Schwarz Information Criterion (SIC). For the NP test and the KPSS test, based on the Bartlett kernel spectral estimation method, the corresponding numbers are the Newey-West bandwidth.

Though our primary interest is to analyze causality between global temperatures and sunspot numbers using the SSA approach, for the sake of completeness, we also present here the results in time and frequency domains, as used in Gupta et al., (2015).

As shown in Table 2 the null hypothesis that SS does not Granger cause GT cannot be rejected for both full and the sub-samples – a result also pointed out by Gupta et al., (2015). This result continues to hold when we also detrend GT.[3,4]

Table 2. Time-Domain Granger Causality Test Results

Sample and Number of Observation		**Total Sample (1605 Obs)**		**Subsample A (674 Obs)**		**Subsample B (609 Obs)**		**Subsample C (322 Obs)**	
Referring Periods		1880:1-2013:9		1880:1-1936:2		1936:3-1986:11		1986:12-2013:9	
Causality Direction		SS-->GT		SS-->GT		SS-->GT		SS-->GT	
Tested Series	**Original**	F	p-value	F	p-value	F	p-value	F	p-value
		1.0107	0.3642	0.947	0.3884	1.1374	0.3213	1.5871	0.2062
	De-trended	F	p-value	F	p-value	F	p-value	F	p-value
		1.3569	0.2287	1.2343	0.2949	1.6907	0.1505	0.9201	0.4525

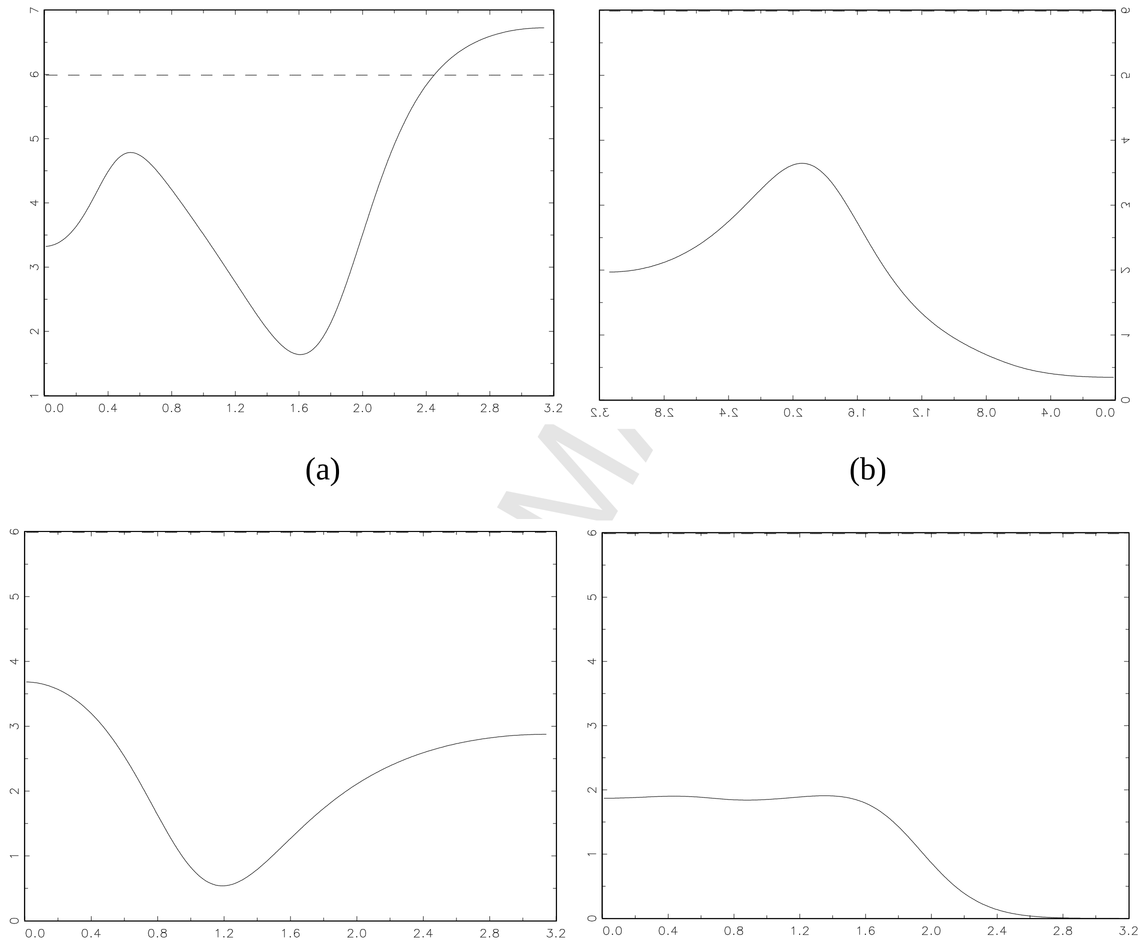
Next, we repeat and present the frequency domain causality results of Gupta et al., (2015) for the full and the sub-samples in Figures 2, with the same lag-structure as used in the time domain Granger causality tests. The figures depict the test statistics (solid line) along with their 5 percent critical values (broken line) for all frequencies in the interval $(0, \pi)$, to assess the predictive content of SS for GT. For the full-sample (1880:1-2013:9), the null hypothesis of non-predictability is rejected for ω greater than 2.45 corresponding to a cycle length between 2 and 2.6 months.[5] For the sub-samples 1 (1880:1-1936:2), 2 (1936:3-1986:11) and 3 (1986:12-2013:9), however, the null of no

[3] Given the weak evidence of stationarity for SS for sub-sample C, we repeated the Granger causality test with first differences of SS and GT without and with detrending. The null of non-causality still continued to hold with *p*-values of 0.7279 and 0.6597, respectively. Further details on these results are available upon request from the authors.

[4] Base don the suggestions of an anonymous referee, we also conducted the nonparametric rank Granger causality tests which is robust to non-normal errors of Holmes and Hutton (1990). However, as with the standard Granger causality tests, the null of no-causality could not be rejected at the conventional 5 percent level of significance. Complete details of these results are available upon request from the authors.

[5] Recall that, the frequency (ω) on the horizontal axis can be translated into a cycle or periodicity of T months by $T = (2\pi / \omega)$, where T is the period.

predictability cannot be rejected for any frequency. So, as in the time domain Granger causality tests for the sub-samples, the frequency domain tests too fail to reject the null that SS has no predictability for GT in the sub-samples. Since in the presence of structural breaks, the full-sample causality results cannot be relied upon, our frequency domain causality tests, as in Gupta et al., (2015), tend to suggest that there is no causality running from SS to GT



(a) (b)

(c) (d)

Figure 2: Frequency Domain Causality of -- (a) Total Sample (1880:1-2013:9), (b) Subsample A (1880:1-1936:2), (c) Subsample B (1936:3-1986:11), (d) Subsample C (1986:12-2013:9).

As with the time domain tests, we also present below, in Figures 3, the results from the frequency domain test of GT after detrending. As can be seen, now there is no evidence of causality either for the full-sample or sub-samples any frequency. This result could imply that results based on trending GT series for the full-sample could have been spurious in the frequency domain as reported in Gupta et al., (2015).

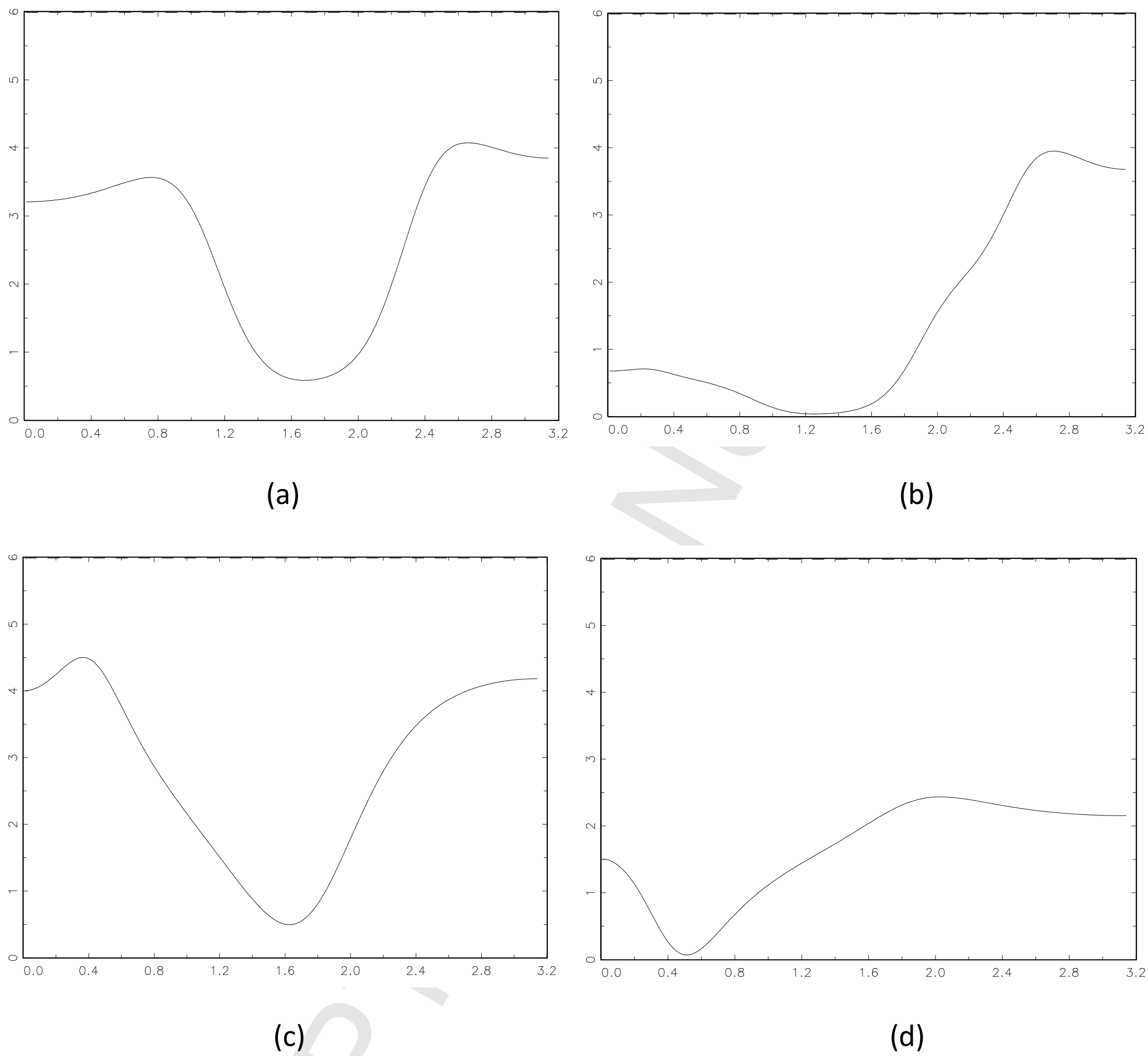


Figure 3: Frequency Domain Causality with Detrended GT of – (a) Total Sample (1880:1-2013:9), (b) Subsample A (1880:1-1936:2), (c) Subsample B (1936:3-1986:11), (d) Subsample C (1986:12-2013:9).

Against this background of lack of evidence of causality in the time and frequency domains, we now next turn our attention to the causality using the SSA-based approach. As mentioned in 2.2, in order to conduct the SSA-based Causality Test for the sunspot and global temperature data, the out-of-sample size for each subsample series is 1/3 of the whole series. In addition, before the last step which determines causality by causality criterion $F_{GT|SS}$ in 2.2, all the forecasting results of both SSA and MSSA steps are the optimal choice chosen respectively after considering all the possibilities of window length L and its corresponding choices of number of eigenvalues r. The following table summarizes the causality test results based on SSA technique. As what is mentioned in 2.2, if the causality criterion $F_{GT|SS} \geq 1$, then either there is no detectable association between GT and SS or the performance of the univariate SSA is better than of the MSSA, this may happen, for example, when one of the series has structural breaks misdirecting the forecasts; If $F_{GT|SS} < 1$, then we conclude that the information provided by the series Y can be regarded as useful or supportive for forecasting the series X. According to the following table, when the whole sample is considered, the test statistics is very close to 1 and could not provide strong information to determine the causality between GT and SS. This is possibly affected by the structural breaks we detected in GT, which misleads the forecasts. Comparing with the empirical evidence of Gupta et al., (2015), whereby the authors detected causality only in for the full-sample, our SSA-based causality tests, provides strong evidence of causality for all the-subsamples as well, to go on with the weak evidence of causality for the full-sample. In addition, considering the detrended *GT* series in comparison with using the *GT* with trend for our tests, the causality for all subsamples and the weak evidence for total samples still hold. Recall, when we repeated the frequency domain analysis for Gupta et al., (2015) using detrended *GT*, we could not detect causality even for the full-sample – a result also obtained for the time-domain version of the test.[6] In more details, subsample A show the strongest effect comparing to other subsamples regardless of the original and de-trended series; followed by subsample C with slightly weaker causal effect from SS to GT; moreover, the weakest causal effect holds for subsample B according to tests of both original and de-trended series.

[6] For a discussion of causality based on cross-spectrum analysis between GT and SS, refer to the Appendix of this paper.

Table 3. SSA-based Causality Test Results

Sample and Number of Observation		Total Sample (1605 Obs)	Subsample A (674 Obs)	Subsample B (609 Obs)	Subsample C (322 Obs)
Referring Periods		1880:1-2013:9	1880:1-1936:2	1936:3-1986:11	1986:12-2013:9
Test Statistics		$F_{GT\|SS}$	$F_{GT\|SS}$	$F_{GT\|SS}$	$F_{GT\|SS}$
Series	Original	0.998	0.284	0.399	0.308
	De-trended	0.967	0.400	0.800	0.465

Note that $F_{GT|SS}$ is the criterion of SSA-based causality test based on forecasting accuracy (see **2.2**).

4. Concluding remarks

Global warming is undoubtedly the biggest topic of research amongst researchers working on environment. What drives global temperatures is understandably an interesting area of research. While greenhouse gases emissions are believed to be a major cause, there is also a large literature that tends to suggest that solar activity also drives global temperatures. However, the evidence in terms of the latter line of reasoning is mixed. Given this, in a recent paper, Gupta et al., (2015) analyzed whether sunspot numbers cause global temperatures based on monthly data covering the period 1880:1-2013:9, using not only time-domain, but also frequency domain causality tests. The authors find that standard time domain Granger causality test fails to reject the null hypothesis that sunspot numbers does not cause global temperatures for both full and sub-samples, namely 1880:1-1936:2, 1936:3-1986:11 and 1986:12-2013:9 (identified based on tests of structural breaks). However, frequency domain causality test detects predictability for the full-sample at short (2 to 2.6 months) cycle lengths. As with the time domain results, no causality however, could be detected for the sub-samples. But since, full-sample causality cannot be relied upon due to structural breaks, as Granger causality tests assumes constancy of parameters during the sub-sample, which is of course not the case with structural breaks, Gupta et al., (2015) concludes that the evidence in favour of sunspot numbers causing global temperatures is weak, if not non-existent.

Given the importance of the issue of global warming, our current paper aims to revisit the question of whether sunspot numbers cause global temperatures, using the same data set and sub-samples used by Gupta et al., (2015), but now, based on an advanced new nonparametric technique -- the Singular Spectrum Analysis (SSA)-based causality test. Our nonparametric technique is able to capture possible nonlinearities that could exist in the data generating processes of the global temperatures and sunspots, but also, in the relationship between global temperatures and sunspot activity, for instance due to the structural breaks. The SSA being a nonparametric method captures the possible nonlinearities using a data-driven approach, without specifying any known functional nonlinear model to the relationship, which in turn, could be incorrectly specified in the first place, as is possibly the linear model. Using the SSA-based causality tests, we show that sunspot numbers have predictive ability for global temperatures for the all three sub-samples, over and above the full-sample, even if the latter result can be ignored due to structural instability. Thus, the non-parametric SSA-based causality test outperforms both time domain and frequency domain causality tests, and, more importantly, highlights that sunspot numbers have always been important in predicting global temperatures. In other words, researchers working on global warming can predict movements of the global temperatures based on movements in sunspot activity, but for this, they need to rely on a nonlinear data-driven, i.e., nonparametric approach.

Given the importance global warming, two areas of future research would be: (1) Since there is evidence of causality in the full-sample, it is clear that there must be causality at certain specific points in time, even if it is not for the sub-samples identified based on structural breaks. In light of this, one needs to undertake a time-varying or rolling sub-samples based test of causality. Also, in this regard, it is important to analyze the direction or the sign of the effect of this causal relationship if it exists at specific points in time, to design environmental policies better, and (2) Here we analyze in-sample predictability, in the future it would be interesting to compare linear and nonlinear models in forecasting out-of-sample global temperatures based on sunspot numbers. This would provide information, ahead of time as to where global temperatures are headed given an existing set of information on sunspot numbers.

Finally, as a cautionary note, it is important to highlight, something that we have touched above as well, that the Earth's climate is regulated by anthropogenic emissions like CO_2, volcanoes and other greenhouse gases, which need to be factored in as well to properly identify the contribution of solar activity (Scafetta, 2014). Ignoring these issues could also lead to spurious, in other words, more significant influence from sunspot numbers on global temperatures. However, in our case, the objective was replicating the work of Gupta et al., (2015), and over the same sample period data on CO_2 emissions were only available at annual frequency. In this regard, an interesting piece of recent work can be found in Hassani et al., (2015). In addition, while we are only analyzing causality and not correlation between sunspot numbers and global temperatures, we must be careful in saying that sunspot numbers used as a partial proxy for solar activity are positively (and linearly), since this might not be the case, and hence. In other words, our evidence of causality between sunspot numbers and global temperatures should not be associated with positive correlation between these two variables. The sign of this relationship is beyond the scope of this paper.

References

Aguirre, L.A., Letellier, C. and Maquet, J. (2008). Forecasting the time series of sunspot numbers. Solar Physics, 249, 103-120.

Bai, J. and Perron, P. (2003). Computation and analysis of multiple structural change models. Journal of Applied Econometrics, 18, 1-22.

Breitung, J. and Candelon, B. (2006). Testing for short- and long-run causality: A frequency-domain approach. Journal of Econometrics, 132, 363-378.

Broomhead, D.S. and King, G.P. (1986). Extracting qualitative dynamics from experimental data. Physica D: Nonlinear Phenomena, 20(2-3), 217-236.

Dickey, D. and Fuller, W. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. Econometrica, 49, 1057-1072.

Elliot, G., Rothenberg, T.J. and Stock, J.H. (1996). Efficient tests for an autoregressive unit root. Econometrica, 64, 813-836.

Folland, C.K., Colman, A.W., Smith, D.W., Boucher, O., Parker, D.E. and Vernier, J-P. (2013). High predictive skill of global surface temperature a year head. Geophysical Research Letters, 40(4), 761-767.

Geweke, J. (1982). Measurement of linear dependence and feedback between multiple time series. Journal of the American Statistical Association, 77, 304-324.

Ghodsi, M., Hassani, H. and Sanei, S. (2010). Extracting fetal heart signal from noisy maternal ECG by singular spectrum analysis. Journal of Statistics and its Interface, Special Issue on the Application of SSA, 3(3), 399-411.

Gil-Alana, L.A., Yaya, O.S., and Shittu, O.I. (2014). Global temperatures and sunspot numbers. Are they related? Physica A: Statistical Mechanics and its Applications, 396(15), 42-50.

Granger C.W.J. (1969). Investigating causal relations by econometric models and cross-spectral methods. Econometrica, 37, 424-439.

Gray, L.J., Beer, J., Geller, M., Haigh, J.D., Lockwood, M., Matthes, K., Cubasch, U., Fleitmann, D., Harrison, G., Hood, L., Luterbacher, J., Meehl, G.A., Shindell, D., van Geel, B. and White, W. (2010). Solar influences on climate. Review of Geophysics, 48(4). DOI: 10.1029/2009RG000282.

Gupta, R., Gil-Alana, L. And Yaya, O. (2015). Does sunspot numbers cause global temperatures? Evidence from a frequency domain causality test. *Applied Economics*, 47 (8), 798-808.

Hassani, H. (2007). Singular spectrum analysis: methodology and comparison. Journal of Data Science, 5, 239-257.

Hassani, H., Heravi, S. and Zhigljavsky, A. (2009). Forecasting European industrial production with singular spectrum analysis. International journal of forecasting, 25(1), 103-118.

Hassani H. and Thomakos, D. (2010). A review on singular spectrum analysis for economic and financial time series. Statistics and its Interface, 3, 377-397.

Hassani H., Zhigljavsky, A., Patterson, K., et al. (2010). A comprehensive causality test based on the singular spectrum analysis. Causality in Science, 379-406.

Hassani H., Soofi A. S. and Zhigljavsky, A. (2013a). Predicting inflation dynamics with singular spectrum analysis. Journal of the Royal Statistical Society: Series A (Statistics in Society), 176(3), 743-760.

Hassani, H., Heravi, S. and Zhigljavsky, A. (2013b), Forecasting UK Industrial Production with Multivariate Singular Spectrum Analysis. Journal of Forecasting, 32, 395–408.

Hassani, H. and Mahmoudvand, R. (2013). Multivariate singular spectrum analysis: A general view and new vector forecasting approach. International Journal of Energy and Statistics, 1, 55-83.

Hassani, H. Silva, E. S., Gupta, R., and Das, S. (2015). Predicting Global Temperature Anomaly: A Definitive Investigation Using an Ensemble of Twelve Competing Forecasting Models. Department of Economics, University of Pretoria, Working Paper No. 201561.

Holmes, J.M., and Hutton, P.A. (1990). On the Casual Relationship between Government Expenditures and National Income. The Review of Economics and Statistics, 72(1), 87-95.

Hosoya, Y. (1991). The decomposition and measurement of the interdependence between second-order stationary processes. Probability Theory and Related Fields, 88, 429-444.

Kwiatkowski, D.P., Phillips, C.B., Schmidt, P. and Shin, Y. (1992). Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root. Journal of Econometrics, 54, 159-178.

Lean, J.L. and Rind, D.H. (1998). Climate forcing by changing solar radiation. Journal of Climate 11(12), 3069-3094.

Lean, J.L. and Rind, D.H. (2009). How will earth's surface temperature change in future decades? Geophysical Research Letters, 36(15). DOI: 10.1029/2009GL038932.

Love, J.J., Mursula, K., Tsai, V.C. and Perkins, D.M. (2011). Are secular correlations between sunspots, geomagnetic activity, and global temperature significant? Geophysic Research Letters, 38, 1-6.

Ng, S. and Perron, P. (2001). Lag length selection and the construction of unit root tests with good size and power. Econometrica, 69, 1519-1554.

Phillips, P.C.B. and Perron, P. (1988). Testing for a unit root in time series regression. Biometrika 75, 335-346.

Pittock, B. (1978). A critical look at long term sum-weather relationships. Review of Geophysics and Space Physics, 16(3), 400-420.

Pittock, B. (1983). Solar variability, weather and climate: an update. Quarterly Journal of the Royal Meteorological Society, 109, 23-55.

Pittock, B. (2009). Can solar variations explain variations in Earth's climate? Climate Change, 96, 483-487.

Sanei, S. and Hassani, H. (2015). Singular spectrum analysis of biomedical signals [M]. CRC Press.

Scafetta, N. (2009). Empirical analysis of the solar contribution to global mean air surface temperature change. Journal of Atmospheric and Solar-Terrestrial Physics, 71, 1916-1923.

Scafetta, N. (2011). A shared frequency set between the historical mid-latitude aurora records and the global surface temperature. Journal of Atmospheric and Solar-Terrestrial Physics, 102, 368-371.

Scafetta N., Grigolini, P. Imholt, T., Roberts, J.A. and West, B.J. (2004). Solar turbulence in earth's global and regional temperature anomalies. Physical Review E, 69, 026303.

Scafetta, N., and West, B.J. (2003). Solar flare intermittency and the earth temperature anomalies. Physical Review Letters, 90, 248701.

Scafetta, N. and West, B.J. *(*2005*). E*stimated solar contribution to the global surface warming using the ACRIM TSI satellite composite. Geophysical Research Letters, 32, L18713.

Scafetta, N. (2014), Global temperatures and sunspot numbers. Are they related? Yes, but non linearly. A reply to Gil-Alana et al. (2014), Physica A: Statistical Mechanics and its Applications, 413, 329-342.

Usoskin, I.G., Schussler, M., Solanki, S.K. and Mursula, K. (2004). Solar activity over the last 1150 years. Does it correlate with climate? Proceedings of 13th Cool Stars Workshop, Hamburg.

Zhou, J. and Tung, K-K. (2013). Deducing Multidecadal Anthropogenic Global Warming Trends Using Multiple Regression Analysis. Journal of Atmospheric Science, 70, 3-8.

Appendix

The Frequency Domain Causality Test

Breitung and Candelon' (2006) presented that in a two-dimensional vector of time series $Z_t = (X_t, Y_t)$ observed at time $t = 1, \ldots, T$, where Z_t is a finite-order VAR process, is of the form:

$$\Theta(B) Z_t = \varepsilon_t, \quad t = 1, 2, \ldots, \qquad \text{(A1)}$$

where $\Theta(B) = 1 - \theta_1 B - \ldots - \theta_p B^p$ is a 2×2 lag polynomial with $B^k Z_t = Z_{t-k}$. The error vector ε_t is a white noise process, with $E(\varepsilon_t) = 0$ and $E(\varepsilon_t \varepsilon_t') = \Sigma$, where Σ is a positive definite variance matrix. The VAR process may include a constant, a trend or dummy variables. The matrix Σ is then decomposed as $G'G = \Sigma^{-1}$ where G is the lower triangular matrix of the Cholesky decomposition. With the assumption that the system is stationary, the moving average (MA) representation of the process is,

$$Z_t = \Phi(B)\varepsilon_t = \begin{pmatrix} \phi_{11}(B) & \phi_{12}(B) \\ \phi_{21}(B) & \phi_{22}(B) \end{pmatrix} \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} = \begin{pmatrix} \psi_{11}(B) & \psi_{12}(B) \\ \psi_{21}(B) & \psi_{22}(B) \end{pmatrix} \begin{pmatrix} \xi_{1t} \\ \xi_{2t} \end{pmatrix} = \Psi(B)\xi_t \qquad \text{(A2)}$$

where $\Psi(B) = \Phi(B) G^{-1}$. Then, the spectral density of X_t can be expressed as:

$$f_x(\omega) = \frac{1}{2\pi}\left[\left|\psi_{11}(e^{-i\omega})\right|^2 + \left|\psi_{12}(e^{-i\omega})\right|^2 \right]. \qquad \text{(A3)}$$

Using the following measure of causality, as in Geweke (1982) and Hosoya (1991):

$$M_{Y \to X}(\omega) = \log\left[1 + \frac{2\pi f_x(\omega)}{\left|\psi_{11}(e^{-i\omega})\right|^2} \right]. \qquad \text{(A4)}$$

Replacing (A3) into (A4) gives,

$$M_{Y \to X}(\omega) = \log\left[1 + \frac{\left|\psi_{12}(e^{-i\omega})\right|^2}{\left|\psi_{11}(e^{-i\omega})\right|^2}\right]. \qquad (A5)$$

Note that, Equation (A5) is zero, if $\left|\psi_{12}(e^{-i\omega})\right|^2 = 0$, which implies that *Y* does not Granger-cause *X* at frequency ω.

The null hypothesis that *Y* does not Granger-cause *X* at frequency ω is then given as:

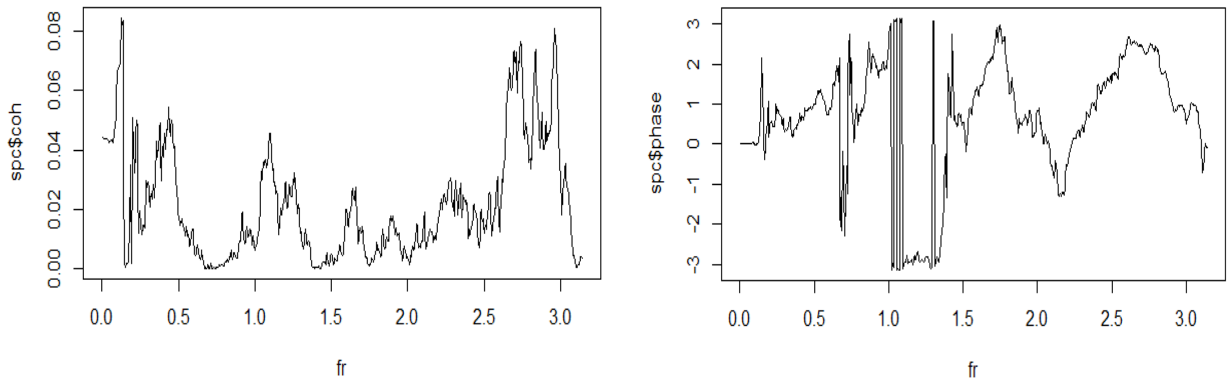
$$H_0: M_{Y \to X}(\omega) = 0. \qquad (A6)$$

The statistic $M_{Y \to X}(\omega)$ is then obtained by replacing $\left|\psi_{11}(e^{-i\omega})\right|$ and $\left|\psi_{12}(e^{-i\omega})\right|$ in (A5) by the estimated values obtained from the fitted VAR.

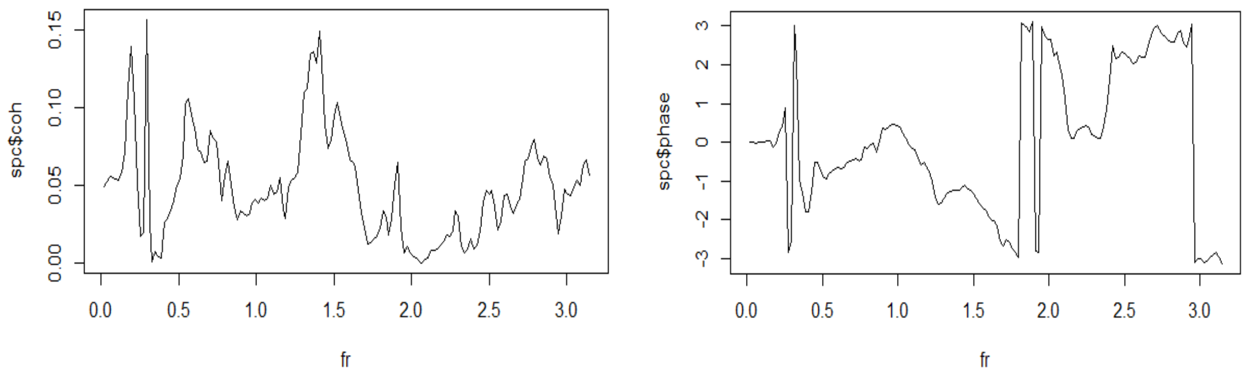
Cross Spectrum Analysis

For the total sample and all subsamples, we performed the cross spectrum analysis on SS and GT, as well as on SS and detrended GT series for comparison. Briefly, the cross spectrum analysis is the Fourier transformation of cross-covariance of two series, which gives us the degree of relationship between two series at different frequency. For each case, i.e., SS and GT and SS and detrended GT, while conducting the cross spectrum analysis, two types of figures are provided: the squared coherency by frequency and the phase spectrum by frequency. If the squared coherency is large at some specific frequencies, it implies that we can probably consider linear relationship between two tested series at these frequencies. Therefore, we then refer to the figure of the phase spectrum by frequency at these frequencies with relatively large squared coherency. If the phase spectrum is approximately linear with a positive slope, it will suggest the first variable lead changes in the duration of the second variable. When we change the order of variables in the beginning, the final results will be identical for the squared coherency, but an opposite slope should emerge for the phase spectrum accordingly. Here we only provide the results where sunspots numbers is the first variable. As can be seen from the results below, we can generally conclude that, unlike the SSA-based approach, there is not much clear-cut evidence of SS causing GT based on the cross-spectrum analysis.

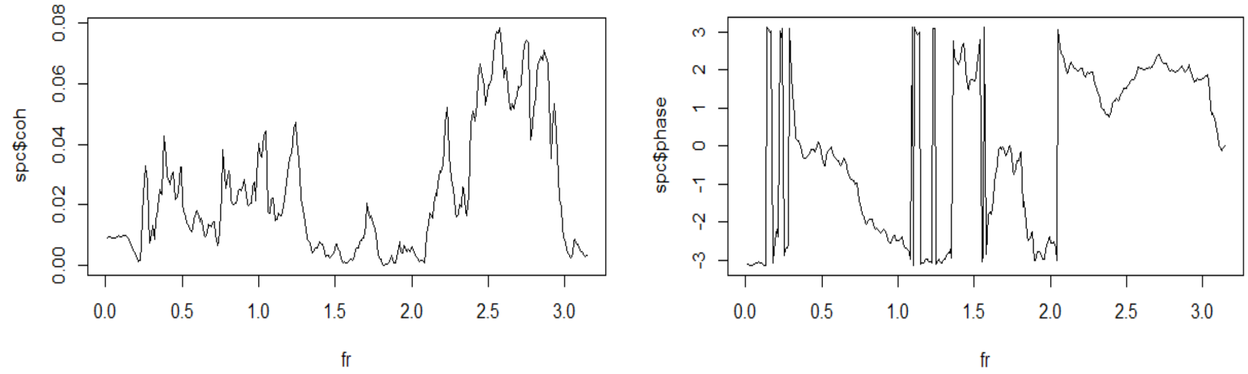
Total Sample (original series)



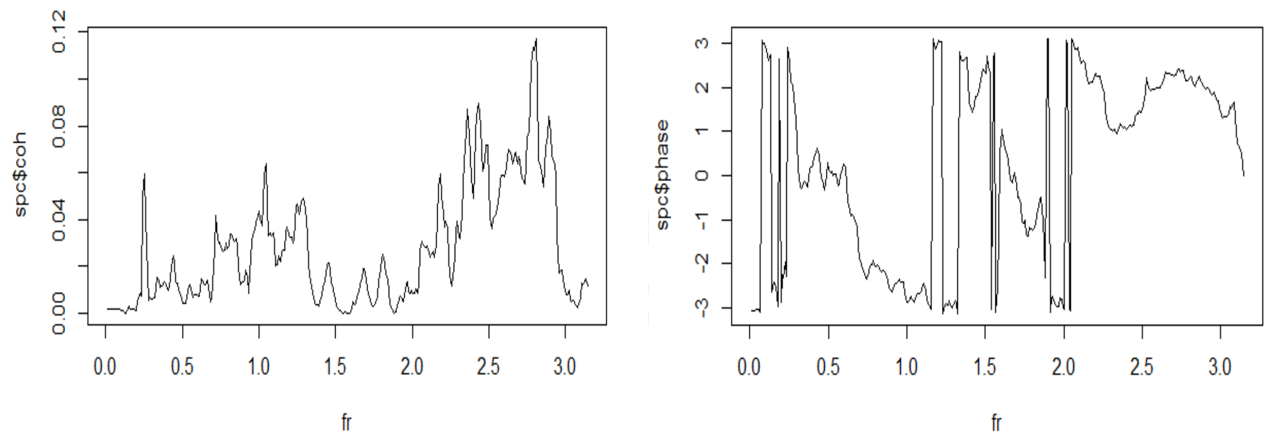
Total Sample (detrended series)



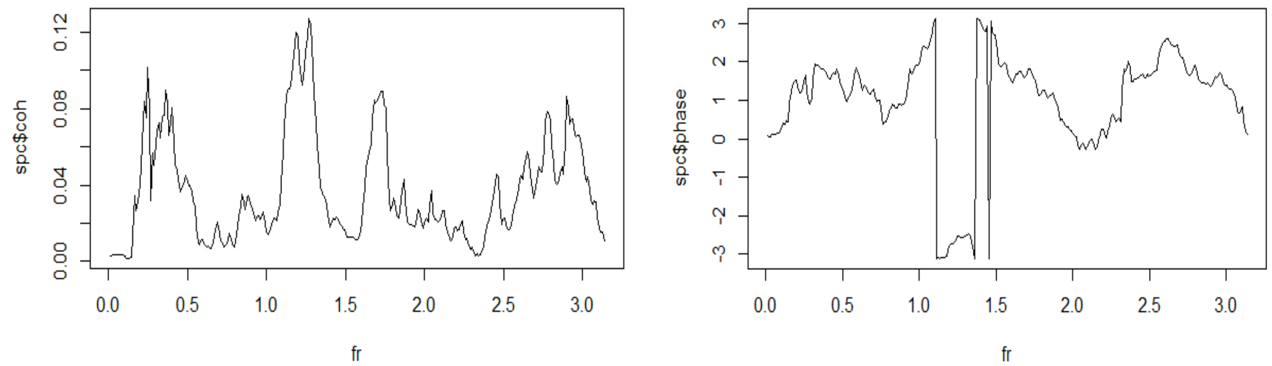
Subsample A (original series)



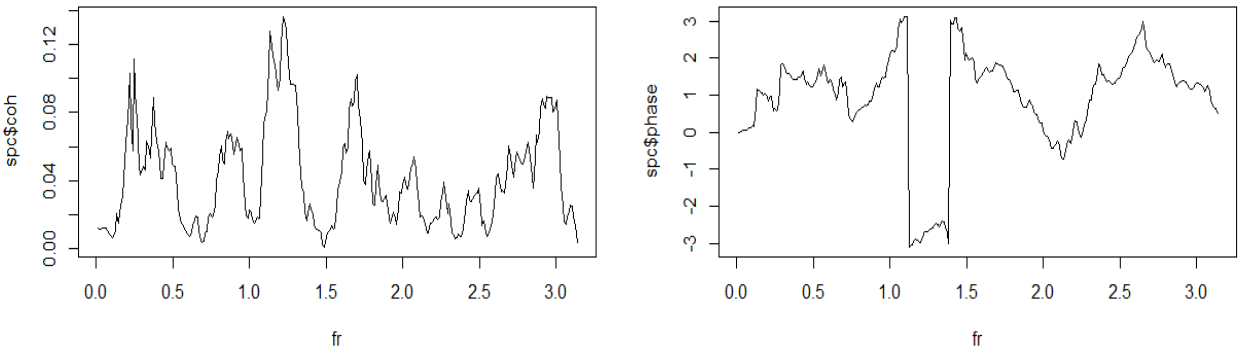
Subsample A (detrended series)



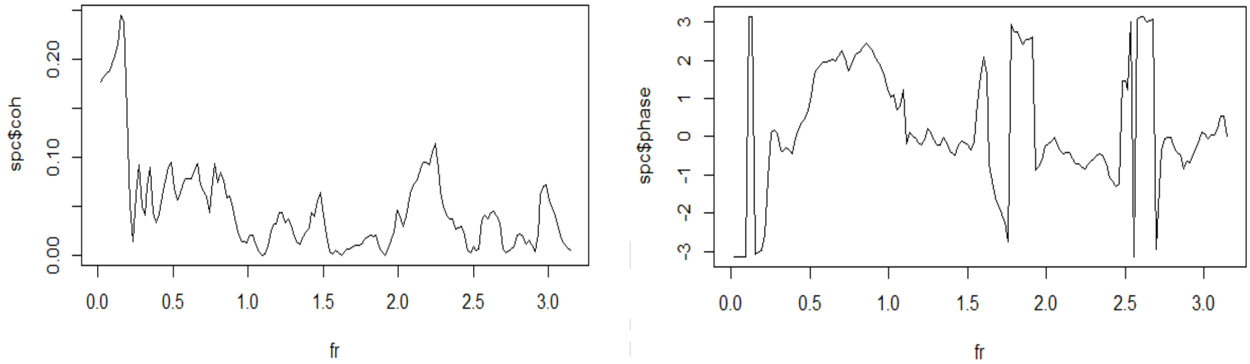
Subsample B (original series)



Subsample B (detrended series)



Subsample C (original series)



Subsample C (detrended series)