Tellus (2007)

TELLUS

Granger causality and Atlantic hurricanes

By JAMES B. ELSNER*, Department of Geography, The Florida State University, Tallahassee, FL, 32306, USA

(Manuscript received 2 August 2006; in final form 1 March 2007)

ABSTRACT

Atlantic tropical cyclones have been getting stronger recently with a trend that is related to an increase in the late summer/early fall sea-surface temperature over the North Atlantic. Some studies attribute the increasing ocean warmth and hurricane intensity to a natural climate fluctuation, known as the Atlantic Multidecadal Oscillation; others suggest that climate change related to anthropogenic greenhouse gases emissions is the cause. Noting that the only difference between these two hypotheses is the causal connection between global mean near-surface air temperature (GT) and Atlantic sea-surface temperature (SST), the author previously showed how to use statistical tests to examine this hypothesis. Here the author expands on this research. In particular, a more comprehensive explanation of the techniques and additional tests and checks against misspecification are provided. The earlier results are confirmed in showing that preceding GT anomalies have a significant statistical relationship to current SST anomalies but not conversely so that if causality exists between Atlantic SST and global temperature, the causal direction likely goes from GT to SST. The result is robust against a small amount of noise added to the data. Identical tests applied to surrogate time series fail to identify causality as expected. The work underscores the importance of using data models to understand relationships between hurricanes and climate.

1. Introduction

A major concern about the consequences of climate change is the potential increase in tropical cyclone activity. Indeed, in 2005 a couple of research papers showed the power of Atlantic tropical cyclones rising dramatically and correlated with an increase in the late summer/early fall sea-surface temperature over the North Atlantic (Emanuel, 2005; Webster et al., 2005). A debate ensued with some studies attributing the increase in hurricane intensity to a natural climate fluctuation, known as the Atlantic Multidecadal Oscillation (AMO) or mode (Goldenberg et al., 2001; Pielke et al., 2005), and others suggesting climate change related in part to anthropogenic increases in radiative forcing from greenhouse gases (Trenberth, 2005; Hoyos et al., 2006).

In 2006, I noted that the main difference between the two competing theories is the causal connection between global mean near-surface air temperature (hereafter GT) and Atlantic sea-surface temperature (hereafter SST) (Elsner, 2006). That is, given the large contemporaneous correlation between the GT and SST records, the climate change theory implies that SST is warming due to climate change while the AMO theory implies that GT is warming, at least in part, from warming SST. Thus, I suggested applying statistical tests of causality to see if some light could be shed on this debate. Specifically, I applied tests of Granger causality which determines whether time-series values of one variable can predict future values of another variable. In short, I showed that lagged values of GT are useful in predicting SST while lagged values of SST are not useful in predicting GT. Thus, GT causes SST in the Granger sense, supporting the theory that climate change influences hurricane intensity assuming changes in GT are due to climate change rather than longer term climate variability.

The space limitation of the Elsner (2006) paper precluded all but the essential details of the analysis. The purpose of the present paper is to include more explanation of the methodology and to provide support for the results using additional checks and tests. In particular, here I give a broader discussion of Granger causality and provide some illustrations. Test for normality and stationarity are described and a test using surrogate data is provided. A sensitivity test on the results using additive noise is also performed. This research is important for moving the hurricane-climate change debate away from simple trend analyses and towards unravelling physical mechanisms that can better account for the various observations.

In Section 2, I describe the data and sources used in Elsner (2006) and reexamined in the present study. In Section 3, I perform some preliminary analysis of the time series by checking for normality and stationarity. I create stationary time series' by using first differencing. The raw and transformed (first differenced) data are plotted as time series and the distributions are plotted as histograms. In Section 4, I describe how Granger causality tests can be used to examine the two competing

*Correspondence
e-mail: jelsner@fsu.edu
DOI: 10.1111/j.1600-0870.2007.00244.x

hypotheses concerning hurricanes. I provide a general description of the method and include a brief summary of climate studies that have applied it. In Section 5, I give more details including how to choose the proper model order using a vector autogressive model and the meaning of the F test. In Section 6, I present the results of the Granger tests expanding on those presented in Elsner (2006). In Section 7, I show results from additional checks that lend support to the original conclusions. I summarize the research in Section 8.

2. Data

As in Elsner (2006) my interest is in examining time series of GT and SST relevant to hurricane activity over the entire North Atlantic basin, which includes the Caribbean Sea and Gulf of Mexico. Therefore, I obtain monthly GT anomalies (1961–1990 base period) from the Intergovernmental Panel on Climate Change (IPCC) using the *Climatic Research Unit* (CRU) (Folland et al., 2001). The anomalies in $^\circ$C are accurate to ± 0.05 $^\circ$C for the period since 1951. I obtain monthly values of SST from the U.S. NOAA-CIRES *Climate Diagnostic Center*. The data are a blend of the Hadley model SST values and interpolated observed SST values from the U.S. National Oceanic and Atmospheric Administration (NOAA). Values are given as anomalies in $^\circ$C from the climatology based period of 1951–2000 (Enfield et al., 2001). The anomalies are spatially averaged over the entire North Atlantic Ocean.

Also as in Elsner (2006), I derive a total power dissipation index (PDI) using the HURricane DATa base (HURDAT or best-track) maintained by the National Hurricane Center. HURDAT is the official record of tropical storms and hurricanes for the Atlantic Ocean, Gulf of Mexico and Caribbean Sea. HURDAT consists of the 6-hourly position and intensity (maximum wind speed at an altitude of 10 m) estimates of tropical cyclones back to 1851 (Neumann et al., 1999). For storms in the period 1931–1956, the 6-h positions and intensities were interpolated from twice daily (00 and 12 UTC) observations. I compute the PDI by cubing the maximum wind speed for each 6 h observation. I consider only observations where the tropical cyclone is at hurricane intensity (33 m s^{-1}) or above and sum the cubed wind speeds over the entire hurricane season for the years 1871–2004. The vast majority of North Atlantic hurricanes occur from July to October. Annual values of this total PDI depend on the duration, frequency, and intensity of the strongest hurricanes.

Questions about the quality of the hurricane data are raised in Pielke et al. (2005), Landsea (2005) and Klotzbach (2006), but the concerns are largely about the potential under count of storms and possible underestimate of intensities prior to aircraft and satellite information. Here I am concerned with the interannual variability of activity, so a possible low bias (in counts and intensity) during the earlier part of the record is not a substantive issue. In fact, as will be seen, tests for Granger causality are applied on time series that have been differenced by subtracting successive time values thereby reducing the influence of a low frequency or trend variation due to nature or technology. Moreover, the principal focus of this work is the potential causal relationship between Atlantic SST and global temperature. Data representing these two variables do not contain the kind of observational bias inherent in the hurricane data.

Additionally, I am interested in investigating whether El Niño-Southern Oscillation (ENSO) can be used to causally explain both GT and Atlantic SST. Thus, I obtain monthly Niño 3.4 SST index values from the *Climate Diagnostic Center*. The values in $^\circ$C represent an area average from 5°S to 5°N latitude and from 170°W to 120°W longitude (Rayner et al., 2003).

3. Preliminary analysis

Time-series plots and histograms of the GT, SST and PDI data are shown in Fig. 1. The GT and SST anomalies are time averaged over the main Atlantic hurricane season months of August–October for $n = 135$ consecutive seasons from 1871 to 2005. The time-series plots show similar and coincident low and high frequency variation throughout the period. There is a marked overall increase in temperatures but it is not uniform with relatively larger increases noted since about 1920 and again since about 1980. The PDI, which only goes through 2004, shows larger interannal variations with a lower amplitude multidecadal variation that coincides with the lower frequency variation in both SST and GT.

The linear correlation between the two temperature series is 0.82. Warmer global temperatures are associated with warmer Atlantic SST. However, the value of the correlation and its statistical significance are influenced by trends and autocorrelation in the time series. More importantly, causality cannot be assessed from correlation analysis alone. Instead, tests of weak causality can be made by determining whether one time series is useful in predicting another. More specifically, causality in the Granger sense (Granger, 1969; Kaufmann and Stern, 1997) can be tested statistically by comparing two sets of models involving lagged values of the predictor variable(s). However, because the tests involve the use of regression models, it is important to first examine the data for stationarity and normality.

3.1. Normality

Classical regression models are applicable when the predictand is normally distributed. The histogram of SST (see Fig. 1) indicates values that are close to normally distributed. However, there is a positive skewness in the GT values and more so in the PDI values. That is, there tends to be more below normal temperature anomalies with fewer but more intense above normal anomalies. For the PDI record this characteristic is even more pronounced and is a consequence of using the third power of the wind speed. Non-normality of the predictand in a regression model may result in heteroscedasticity (non-constant variance) and non-normality

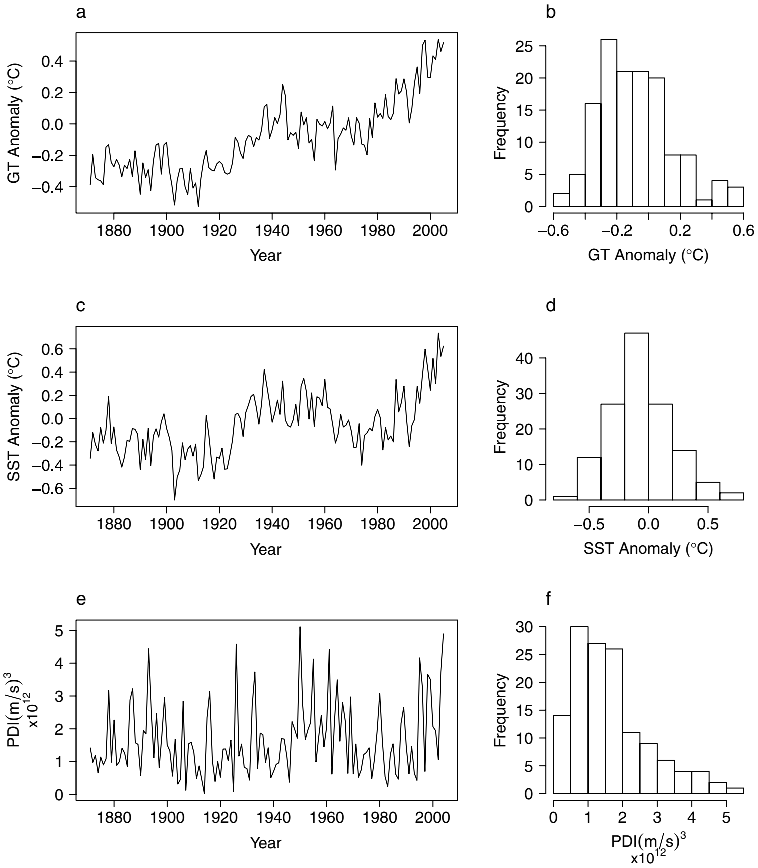

Fig. 1. Time series and histograms of GT (a, b), SST (c, d) and PDI (e, f). Data sources are given in the text. The GT and SST values are averaged over the months of August–October. The PDI values are totalled over the entire corresponding Atlantic basin hurricane season.

of the residuals (difference between the observed and predicted values), which can bias the forecasts and confidence intervals yielded by the statistical model.

To test for normality, I apply the well-known goodness-of-fit Shapiro–Wilks test. The test statistic W, is computed as

$$W = \frac{\sum_{i=1}^{n} \left[a_i x_{(i)}\right]^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}, \tag{1}$$

where the $x_{(i)}$'s are the ordered sample values [$x_{(1)}$ is the smallest] and the a_i's are constants generated from the means, variances and covariances of the order statistics of a sample of size n from a normal distribution. The value of W approximates the correlation coefficient of the values in a normal quantile–quantile (QQ) plot. Small values of W are evidence of departure from normality. Figure 2 shows the normal QQ plots for the raw time series where the quantiles of the data are plotted against corresponding quantiles of a normal distribution. The straight line indicates a perfect fit of the data to a normal distribution. For the SST data the fit is quite good, but for the GT and PDI, the fit is less so. Table 1 shows the results of the test which confirm rejecting the hypothesis of normality for the GT and PDI values, but not for the SST values. Next, I consider stationarity.

3.2. *Stationarity*

Tests of causal significance require stationary time series. Amultivariate time series is covariance stationary and ergodic if all its components are stationary and ergodic. Thus it is sufficient to consider stationarity as limited to time invariant in the mean. The trend and low frequency fluctuations noted in both the GT and SST records suggest that this is not likely the case. The KPSS test (Kwiatkowski et al., 1992) is commonly used to test for stationarity in time-series data. Details of the test are beyond the scope of this paper, but are given in Zivot and Wang (2002). The test statistic for a test of mean stationarity (stationarity about a constant level) is compared to right-tailed quantiles of asymptotic distributions constructed from standard Brownian motion. Large values of the statistic lead to a rejection of stationarity. Table 2 shows the results of the test applied to the GT and SST time series and indicate that I should reject the hypothesis of stationary for both series. Thus in order to proceed with examining causal significance trend removal is needed.

A common trend removal technique is to apply a first-difference operator to the series. Given a time series x_t, a forward first-difference operator can be defined as $\Delta x_t \equiv x_{t+1} - x_t$. I apply this operator to all three time series and plot them along with the corresponding distributions in Fig. 3. Unlike the raw

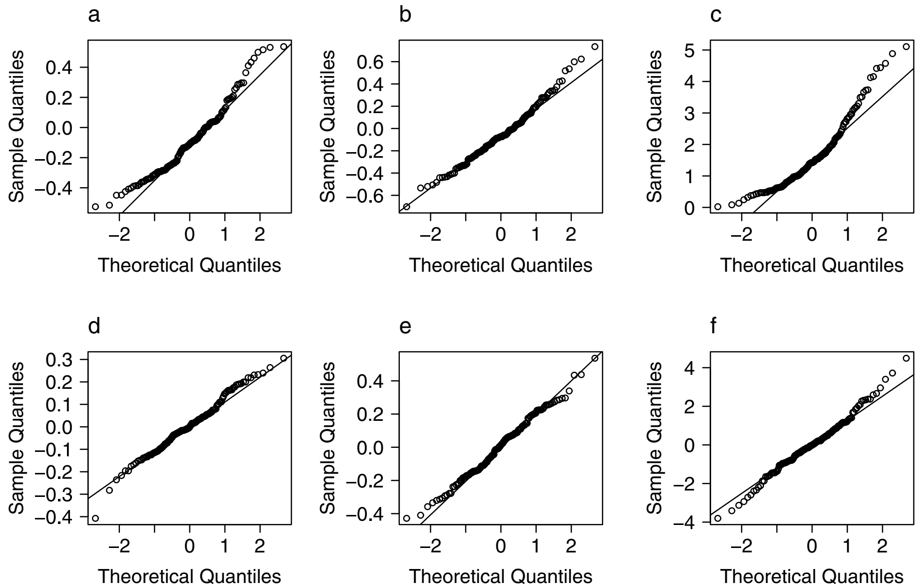

Fig. 2. Normal quantile plots of the raw and first differenced records of GT (a, d), SST (b, e) and PDI (c, f). The units of the sample quantile are °C for the GT and SST records and $m^3 s^{-3}$ for the PDI record.

Table 1. Goodness-of-fit test for normality. The Shapiro–Wilks test statistic and corresponding *p*-value for the GT, SST and PDI records and their first differences (ΔGT, ΔSST and ΔPDI). The value of the test statistic *W* is computed from eq. (1) and approximates the correlation between ranked quantiles of the data and ranked quantiles of values from a normal distribution. The null hypothesis is that the data come from a normal distribution, so a small *p*-value indicates a rejection of this hypothesis

Time series	*W*	*p*-value	Normal
GT	0.96	0.0005	No
SST	0.99	0.1765	Yes
PDI	0.91	<0.0001	No
ΔGT	0.99	0.4065	Yes
ΔSST	0.99	0.6644	Yes
ΔPDI	0.99	0.5333	Yes

Table 2. KPSS test for level stationarity. The test statistic and corresponding *p*-value for the GT, SST and PDI time series and their first differences (ΔGT, ΔSST and ΔPDI). The null hypothesis is that the data are stationary about some constant level (mean), so a small *p*-value indicates a rejection of this hypothesis

Time series	Test statistic	*p*-value	Stationary
GT	3.44	<0.01	No
SST	1.85	<0.01	No
PDI	0.335	>0.1	Yes
ΔGT	0.049	>0.1	Yes
ΔSST	0.041	>0.1	Yes
ΔPDI	0.045	>0.1	Yes

time series, all of the differenced series appear to fluctuate about a level mean. Moreover, the histograms indicate normally distributed values. Indeed, normality and stationarity are confirmed for all three differenced series (see Tables 1 and 2) using the tests described above.

4. Granger causality

Having transformed the time series to conform to the assumptions needed to test for causality, I focus on the problem of determining which of the two hypotheses concerning Atlantic hurricanes is more likely given the data and assumptions. Figure 4 illustrates the causality of the two candidate theories. The climate change hypothesis asserts that changes in radiative forcing resulting from increased greenhouse gas build up in the atmosphere increases GT and causes Atlantic SST to rise at least during the hurricane season months of August–October. On the other hand, the AMO hypothesis asserts that natural changes in the deep water circulation of the Atlantic Ocean drive hurricane season SST resulting in changes to hurricane activity. Under both hypotheses local SST plays a direct role in helping to power hurricanes by providing moist enthalpy and instability. Thus the point of departure for the two competing hypotheses is the causal connection between GT and Atlantic SST. The climate change hypothesis suggests the causality goes from GT to Atlantic SST whereas the AMO hypothesis implies it is the other way around. The implication that Atlantic SST is forcing GT is a consequence of their large contemporaneous correlation noted above.

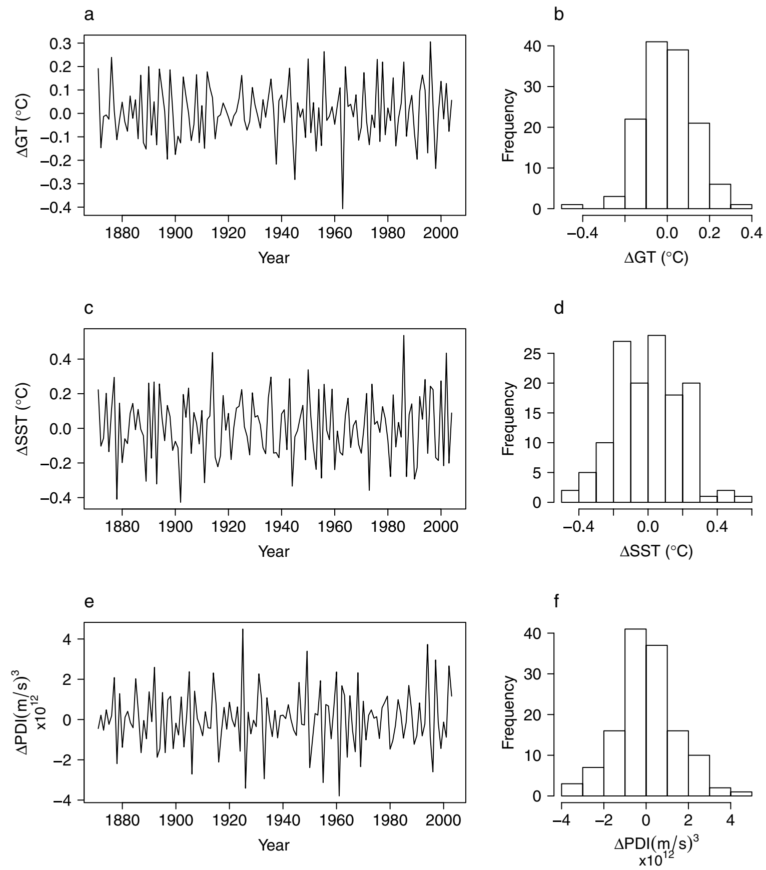

Fig. 3. Time series and histograms of the first-differenced GT (a, b), SST (c, d) and PDI (e, f) records.

Fig. 4. Two networks illustrating the competing hypothesis concerning Atlantic hurricane activity and global temperature. In the first network (a) global temperature (GT) predicts Atlantic sea-surface temperature which in turn predicts hurricane activity. In the second network (b) Atlantic sea-surface temperature (SST) predicts both hurricane activity and GT.

One way to compare hypotheses is to consider whether there is asymmetry in the predictive skill of statistical forecasts when one time series is used to predict the other and vice versa. The idea is that time's arrow is unidirectional in that only the past can cause the present. If lagged values of time series U are useful in predicting future values of time series V, but not the other way around, then U must come before V and is therefore a candidate for causing V, whereas V is eliminated as a candidate for causing U. Figure 5 illustrates the point graphically. The phase shift in the two time series with series U leading series V indicates that lagged values of U will be useful in predicting future values of V, but lagged values of V will be useless in predicting future values of U. It should be noted that this predictive causality (called Granger causality after Granger, 1969), though helpful to understand feedbacks and interactions in complex systems, does not necessarily imply true causality.

Somewhat more formally, a variable U is said to Granger-cause a variable V if it can be shown that time-series values of U provide statistically significant information on future values of V. The test works by first regressing U on lagged values of V to determine the maximum lag for V (reduced model). Then V is regressed on lagged values of V and lagged values of U for lags out to the maximum lag (full model). The full model is compared to the reduced model using an F-test to see if lagged values of U statistically improve upon the reduced model. If there is significant improvement with the full model by adding the U variable, then we say that U Granger causes V. This predictive definition of causality leaves open the possibility that causality is found between U and V when in fact they are uncoupled. This can be the case if they are both driven by a third variable, say W. As noted in Mosedale et al. (2006), a predictive definition of causality is quite relevant to climate science but it has not seen widespread use.

Richards (1993) is an early application of Granger causality tests to climate data showing that CO_2 is a significant forcing

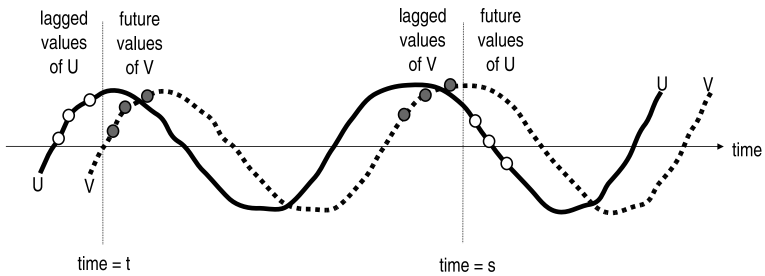

Fig. 5. A schematic showing two time series U and V with series U leading series V. Note that lagged values of U are useful in predicting future values of V (e.g. time $= t$), but not conversely (time $= s$).

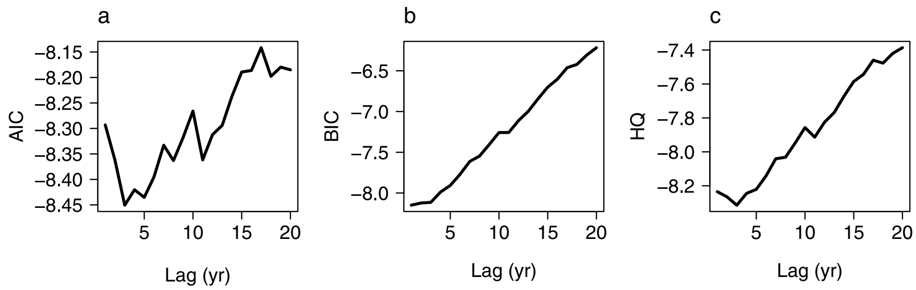

Fig. 6. Values of AIC (a), BIC (b) and HQ (c) as a function of model order (maximum lag) for the vector autoregressive model involving GT and SST. The function minimum identifies the order (L) to use in a test for causality.

for global temperature with contributions from solar irradiance and volcanic aerosol loading much smaller. Kaufmann and Stern (1997) use the tests on hemispheric temperature records and show that the Southern Hemisphere appears to lead the Northern Hemisphere as would be expected from anthropogenic climate change. Triacca (2001) provides a contrary opinion on the use of Granger causality tests in these contexts. The problem is drawing conclusions when more than two variables are involved in the tests. More recently Wang et al. (2004) use Granger causality tests to examine the relationship between the North Atlantic oscillation and Atlantic SST on the seasonal time scale. They find that the Gulf Stream extension into the Atlantic has a causal effect on the wintertime NAO. Mosedale et al. (2006) follow up this study using daily data and find that the tripole pattern of Atlantic SST Granger causes the NAO.

5. Methodology

The natural extension to a univariate autogression model is the vector autoregression (VAR) model, which is a flexible approach to the analysis of multivariate time series. The VAR model is especially useful for describing the dynamic behavior of economic time series and for structural inference (Zivot and Wang, 2002) as used here. Here I consider two time series, so I let $\mathbf{Y}_t = (y_{1t}, y_{2t})'$ denote the (2×1) vector of times-series variables. Then the L-lag vector autoregressive [VAR(L)] model is given by

$$\mathbf{Y}_t = \mathbf{c} + \Pi_1 \mathbf{Y}_{t-1} + \Pi_2 \mathbf{Y}_{t-2} + \cdots + \Pi_L \mathbf{Y}_{t-L} + \epsilon_t \tag{2}$$

for $t = 1, \ldots, T$, where Π_i are (2×2) coefficient matrices and ϵ_t is a (2×1) unobservable zero-mean white noise vector process that is serially uncorrelated with time invariant covariance matrix $\mathbf{\Sigma}$, and $\mathbf{c}$ is an offset to allow for non-zero means. My main focus is the bivariate time series consisting of the first-differenced values of GT (y_{1t}) and Atlantic SST (y_{2t}). Since both components of the bivariate VAR model have the same explanatory variables, each equation can be estimated separately by ordinary least squares without losing efficiency relative to generalized least squares (Zivot and Wang, 2002).

The first step is to determine the maximum lag (order) of the VAR model. This is done using a selection criteria with the value of the maximum lag (L) chosen that minimizes some criteria. Following Zivot and Wang (2002), let $\Sigma(L) = T^{-1} \sum_{t=1}^{T} \hat{\epsilon}_t \hat{\epsilon}_t'$ be the residual covariance matrix, then the three most common information criteria are the Akaike (AIC), Schwartz-Bayesian (BIC) and Hannan-Quinn (HQ) given by

$$\mathrm{AIC}(L) = \ln|\Sigma(L)| + \frac{2}{T} L n^2 \tag{3a}$$

$$\mathrm{BIC}(L) = \ln|\Sigma(L)| + \frac{\ln T}{T} L n^2 \tag{3b}$$

$$\mathrm{HQ}(L) = \ln|\Sigma(L)| + \frac{2 \ln \ln T}{T} L n^2. \tag{3c}$$

Values for the three criteria with maximum lags (L) 1–20 are shown in Fig. 6. All three criteria indicate a low-order model with the AIC and HQ suggesting order 3 and the BIC suggesting order 1. According to Zivot and Wang (2002), the AIC criterion asymptotically overestimates the order whereas the BIC and HQ criteria provide a more consistent estimate under fairly general conditions. Based on these results I consider Granger causality tests using maximum lags $L = 1$–5 years.

In a bivariate VAR(L) model for $\mathbf{Y}_t = (y_{1t}, y_{2t})'$, y_2 fails to Granger cause y_1 if all of the L VAR coefficient matrices $\Pi_1, \ldots, \Pi_L$ are lower triangular (Zivot and Wang, 2002). That is, all of the coefficients on lagged values of y_2 are zero in the equation for y_1 (reduced component model). Similarly, y_1 fails to Granger cause y_2 if all of the coefficients on the lagged values of y_1 are zero in the equation for y_2. To test whether the reduced model is statistically different from the full model I use an F test for nested models. Given T observations, where the full model has k coefficients and the reduced model has $k - L$ coefficients, the F-test statistic is given as

$$F = \frac{(\text{RSS}_{\text{reduced}} - \text{RSS}_{\text{full}})/L}{\text{RSS}_{\text{full}}/(T - k)}, \tag{4}$$

where $\text{RSS}_{\text{reduced}}$ (RSS_{full}) is the residual sum of squares of the reduced (full) component model. The number of predictors in the full model is k, the number of coefficients set to zero in the reduced model is L. The F statistic under the null hypothesis that the reduced model is correct comes from an F distribution with L and $n - k$ degrees of freedom.

6. Results

I perform two separate tests with the results taken together providing clues about the direction of causality. The first involves predicting SST from GT using time lagged values of SST and GT as predictors. In this case, SST is the response variable. The second involves predicting GT from SST (again using time lagged values of SST and GT). In this case GT is the response variable. Thus, for each test I entertain a set of two nested regression models. The full model contains lagged values of both the response and the explanatory variables and the reduced model contains lagged values of the response variable only. Table 3 shows the results of the Granger causality tests for a maximum lag $L = 1$ yr using the first differenced records of GT and SST.

The table shows that with SST as the response variable (Test 1) the reduced model with an extra degree of freedom ($L = 1$) is rejected at a significance level of less than 0.05. Thus I conclude that lagged values of GT improve the prediction of future values of SST. The Pr($>F$) arises from evaluating the likelihood of observing a value of F equal to or exceeding 7.072 from an F distribution with 1 and 131 degrees of freedom. In sharp contrast, with GT as the response variable (Test 2), I fail to reject the null hypothesis that the reduced model is adequate against the alternative of the full model. So in the second case I conclude that lagged values of SST do not improve the prediction of future values of GT.

These results taken together lead me to conclude that GT causes SST in the Granger sense. This implies that, if actual causality exists between Atlantic SST and global temperature, the causal direction likely goes from GT to SST and not the other way around. I repeat the tests using values of L from 2 to 5 (Table 4) and find that the results do not change appreciably. Models using maximum lagged values of SST (out to $L = 5$) do not significantly improve predictions of GT whereas models using lagged values of GT do improve predictions of SST and significantly for maximum lags $L = 1$, 2 and 5. For comparison, I repeat the analysis using the raw (undifferenced) temperature time series and find similar results with lagged values of Atlantic SST providing no improvement (p-value = 0.757, $L = 3$) in predicting GT over a model using only lagged values of GT, and lagged values of GT providing a significant improvement (p-value = 0.005, $L = 3$) in predicting SST over a model using only lagged values of SST.

Table 3. Granger causality tests for GT and SST. The tests are conducted using a maximum lag (model order) of $L = 1$ for the vector autoregressive model. The full models contain lagged values of both SST and GT whereas the reduced models contain only lagged values of the response variable (SST in Test 1 and GT in Test 2). *df.* stands for degrees of freedom and F is given in eq. (4). It is used to decide whether the full model is better at prediction than the reduced model. The null hypothesis is that the reduced model is correct. A value of Pr($>F$) less than 0.05 means I reject the reduced model in favor of the full model.

Model	Residual *df.*	*df.*	F	Pr($>F$)
Test 1: SST as the response				
Full	130			
Reduced	131	1	7.072	0.0088
Test 2: GT as the response				
Full	130			
Reduced	131	1	0.910	0.3419

Table 4. Results from a series of Granger causality tests using GT and SST. Lag refers to the maximum number of lags used in the models (model order). $F_{\text{GT}_{\text{response}}}$ ($F_{\text{SST}_{\text{response}}}$) is the value of the F statistic when GT (SST) is the response variable. Pr($> F_{\text{GT}_{\text{response}}}$) [Pr($> F_{\text{SST}_{\text{response}}}$)] is the probability of observing a value of F equal to or exceeding this value from an F distribution with degrees of freedom equal to lag and n minus lag.

Lag (L)	$F_{\text{GT}_{\text{response}}}$	Pr($> F_{\text{GT}_{\text{response}}}$)	$F_{\text{SST}_{\text{response}}}$	Pr($> F_{\text{SST}_{\text{response}}}$)
1	0.9100	0.3419	7.0716	0.0088
2	0.1367	0.8724	3.5413	0.0319
3	0.5246	0.6662	2.0763	0.1068
4	0.4163	0.7966	1.7083	0.1525
5	0.5509	0.7373	2.9478	0.0152

7. Additional checks

I perform additional checks to substantiate the above results. The first I call a sanity check where I run the tests on a relationship that I know is causal. The role of SST in modulating

Table 5. Same as Table 4 except SST and PDI. The numbers in parantheses are for the same tests using the shorter period of data (1950–2004)

Lag (L)	$F_{\mathrm{PDI}_{\mathrm{response}}}$	Pr($> F_{\mathrm{PDI}_{\mathrm{response}}}$)	$F_{\mathrm{SST}_{\mathrm{response}}}$	Pr($> F_{\mathrm{SST}_{\mathrm{response}}}$)
1	4.5232 (1.8698)	0.0353 (0.1776)	1.2876 (3.0973)	0.2586 (0.0845)
2	3.9497 (1.0669)	0.0217 (0.3523)	0.7377 (0.8796)	0.4803 (0.4217)
3	3.3975 (1.9387)	0.0200 (0.1372)	0.7039 (0.9304)	0.5515 (0.4341)
4	4.2139 (3.6060)	0.0032 (0.0131)	1.1316 (0.6970)	0.3449 (0.5984)
5	4.1608 (4.2011)	0.0016 (0.0039)	0.8719 (0.5630)	0.5023 (0.7277)

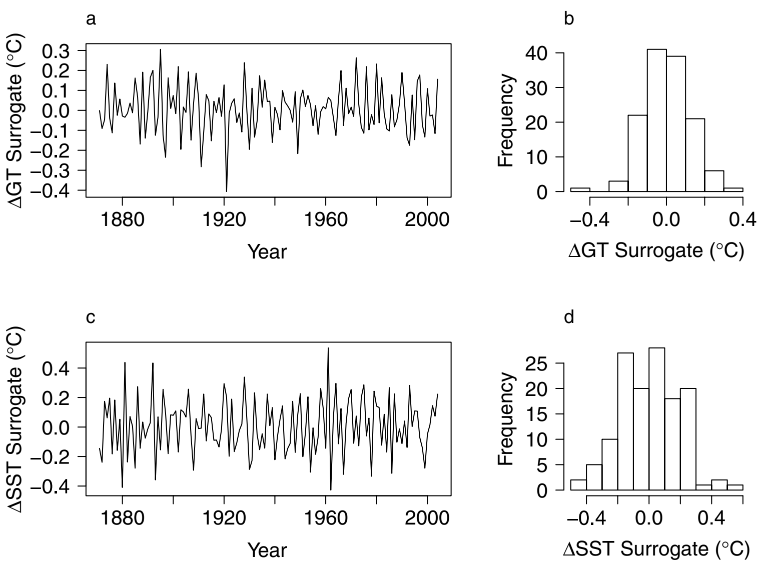

Fig. 7. Surrogate time series and histograms for the ΔGT (a, b) and ΔSST (c, d) series shown in Fig. 3.

hurricane activity is well established physically and statistically (Shapiro and Goldenberg, 1998; Elsner et al., 1999; DeMaria et al., 2001). The warm ocean provides the heat and moisture to sustain hurricane-force winds against friction. Thus, I expect that Atlantic SST causes greater hurricane activity rather than the other way around and I expect such causality to be detectable in the Granger sense at least on the annual time scale and using long enough time series'.

I examine this using the same test procedure as outlined above where I use PDI as a measure of Atlantic hurricane activity to account for intensity and duration of the hurricanes. Although there is no significant trend in PDI since 1871, there has been a notable increase over the past 30 yr or so (Emanuel, 2005). The same selection criteria for choosing the vector autoregressive model order is used and I find that the AIC suggests a maximum lag of 5 yr, with the BIC and HQ both suggesting a maximum lag of 2 yr. Results of the causality tests with SST and PDI using $L = 1$–5 are given in Table 5. They verify that indeed Atlantic SST Granger causes hurricane power as defined by the total PDI over the entire season. Results are somewhat ambiguous using the shorter period of record (1950–2004) for maximum lags 1–3, but for maximum lags 4 and 5 they are consistent with the results from the larger data set showing the causality arrow points from SST to hurricanes (see Table 5).

Another check involves repeating the Granger causality tests on surrogate time series. The surrogate time series retains the autocorrelation structure from the original time series but it removes the specific temporal ordering. On average Granger tests on such surrogate series should not detect causality in either direction. To generate the surrogate series I use the method of Theiler et al. (1992) whereby the phases from the series' Fourier spectrum are randomized. I do this separately for both the GT and SST differenced series. The surrogate time series (Fig. 7) has interannual variability that is indistinguishable from the original series.

As expected since the autocorrelation is preserved in each series, the vector autoregressive model order of the bivariate surrogate series is the same as that of the original bivariate series. However, the Granger tests show different results (Table 6). Here the tests fail to detect causality in either direction for $L = 1$. Results support the contention that the causality detected in the observed series here and in Elsner (2006) is due to the unique temporal ordering of the GT and SST anomalies and not a consequence of autocorrelation.

I also examine the likelihood that another variable is responsible for the causal link detected between GT and Atlantic SST. A leading candidate for this 'hidden' variable is the ENSO as it is known to influence both the Atlantic SST and GT. As with the

Table 6. Same as Table 3 except with surrogate time series. Similar results are found using models with $L = 2$–5 and other surrogate series

Model	Residual *df.*	*df.*	F	Pr(>F)
Test 1: SST surrogate as the response				
Full	130			
Reduced	131	1	0.0708	0.7905
Test 2: GT surrogate as the response				
Full	130			
Reduced	131	1	0.0033	0.9595

Atlantic SST and GT data, I apply a difference filter to the Niño 3.4 temperatures. I first test the possibility of Granger causality between ENSO and GT. The ENSO and GT differenced time series are modeled using a vector autoregressive model as before. The AIC suggests a maximum lag of 5 yr and the BIC and HQ suggest a maximum lag of 4 yr. The Granger test shows a significant model using lagged values of ENSO to predict GT (p-value $= 0.008$, $L = 4$) but also a significant model using lagged values of GT to predict ENSO (p-value $= 0.004$, $L = 4$), thus I conclude that the test is ambiguous in revealing causality with regard to ENSO and GT. I then apply a Granger test on the relationship between ENSO and Atlantic SST and find similar results.

Finally, I consider the influence of random data errors (additive noise) on the ability of the method to detect Granger causality. Here I return to the original test using GT and Atlantic SST first-differenced values. The range (maximum minus minimum) of first differences for the GT values is 0.71 °C and the range of first differences for the Atlantic SST values is 0.96 °C. I multiply these ranges by 5% to obtain an interval of (+0.036, −0.036) and (+0.048, −0.048) for the GT and SST errors, respectively. I then choose 134 random numbers from a uniform distribution in the first interval and add the numbers to the GT differenced values and I choose 134 different numbers from a uniform distribution in the second interval and add the numbers to the SST differenced values. I repeat the Granger tests on these additively perturbed records. In 19 of the 20 (95%) tests I find a significant model for the SST using lagged values of GT. I then increase the noise level to 10% and find only 45% of the tests show a significant model. I thus conclude the results are insensitive to random data errors for errors amounting to less than 10% of the range in values.

8. Conclusions

The power of Atlantic tropical cyclones has recently trended upward and the increases are correlated with increases in late summer/early fall SST over the North Atlantic. A debate concerns the nature of these increases with some studies attributing them to a natural climate fluctuation, known as the Atlantic Multidecadal Oscillation (AMO), and others suggesting climate change related to anthropogenic increases in radiative forcing from greenhouse-gases. Statistical models using GT and Atlantic SST records show that GT is useful in predicting Atlantic SST, but not the other way around suggesting that the causality is more likely to go from GT to SST (Elsner, 2006). Here I examined the data and methodology used in Elsner (2006) in significantly greater detail and demonstrated how statistical models can lead to insights into climate relationships not available from empirical studies.

Results from the Granger causality tests are consistent with the hypothesis that as climate change causes seas to warm, the ocean stores more energy that is converted to hurricane wind. As expected the tests fail to show causality when the temperature values are reordered in time (although preserving the autocorrelation). Results are robust to additive random error amounting to 5% of the range in temperature values (differenced), but degrade at a noise level of 10%. While the possibility is real that a third variable is causing both GT and Altantic SST, evidence for ENSO as that third variable is not strong.

While others argue that the warming of the oceans shows a clear signature of external forcing (Barnett et al., 2005; Mann and Emanuel, 2006; Trenberth and Shea, 2006), this work is the first to directly relate a climate change variable to hurricane activity. The results say nothing about the influence ocean circulation have on SST only that climate change is also likely playing a causative role. The results also say nothing about the magnitude of the effect that global warming has on SST. In fact, as pointed out in Elsner (2006), the causality found for the North Atlantic may not extend to other tropical cyclone regions where ocean circulations during the tropical cyclone season may play a greater role in warming and cooling the oceans and atmosphere. Moreover, the method provides only a necessary, but not sufficient, condition for a cause-and-effect relationship between global temperatures and Atlantic hurricane activity.

Research along these lines is important to the field of hurricane science for moving the debate away from simplistic trend analyses and towards unravelling the physical mechanisms that can account for the various observations. Statistical models provide a useful tool for understanding and predicting climate variability. Importantly, they provide a way to examine feedbacks in the coupled climate system (Mosedale et al., 2006). Unfortunately, statistical models are frequently overlooked in hurricane climate analysis where emphasis is on descriptive approaches like correlation, trend detection, and empirical orthogonal functions.

9. Acknowledgments

I thank Thomas Jagger for his help with the data. All statistical computations were done using R. The work is supported by the U.S. National Science Foundation and the Risk Prediction Initiative of the Bermuda Institute of Ocean Sciences.

References

Barnett, T. P., Pierce, D. W., AchutaRao, K. M., Gleckler, P. J., Santer, and co-authors. 2005. Penetration of human-induced warming into the world's oceans. *Science* **309**, 284–287.

DeMaria, M., Knaff, J. A. and Connell, B. H. 2001. A tropical cyclone genesis parameter for the tropical Atlantic. *Wea. Forecasting* **16**, 219–233.

Elsner, J. B. 2006. Evidence in support of the climate change-Atlantic hurricane hypothesis. *Geophys. Res. Lett.* **33**, L16705, doi:10.1029/2006GL026869.

Elsner, J. B., Kara, A. B. and Owens, M. A. 1999. Fluctuations in North Atlantic hurricane frequency. *J. Climate* **12**, 427–437.

Emanuel, K. A. 2005. Increasing destructiveness of tropical cyclones over the past 30 years. *Nature* **436**, 686–688.

Enfield, D. B., Mestas-Nuñez, A. M. and Trimble, P. J. 2001. The Atlantic multidecadal oscillation and it's relation to rainfall and river flows in the continental US. *Geophys. Res. Lett.* **28**, 2077–2080.

Folland, C. K. and co-authors, 2001. Global temperature change and its uncertainties since 1861. *Geophys. Res. Lett.* **28**, 2621–2624.

Goldenberg, S. B., Landsea, C. W., Mestas-Nuñez, A. M. and Gray, W. M. 2001. The recent increase in Atlantic hurricane activity: Causes and implications. *Science* **293**, 474–479.

Granger, C. W. J. 1969. Investigating causal relations by econometric methods and cross-spectral methods. *Econometrica* **34**, 424–438.

Hoyos, C. D., Agudelo, P. A., Webster, P. J. and Curry, J. A. 2006. Deconvolution of the factors contributing to the increase in global hurricane intensity. *Science* **312**, 94–97.

Kaufmann, R. K. and Stern, D. I. 1997. Evidence for human influence on climate from hemispheric temperature relations. *Nature* **388**, 39–44.

Klotzbach, P. J. 2006. Trends in global tropical cyclone activity over the past twenty years (1986–2005). *Geophys. Res. Lett.* **33**, L10805, doi:10.1029/2006GL025881.

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P. and Shin, Y. 1992. Testing the null hypothesis of stationarity against the alternative of a unit root. *J. Econometr.* **54**, 159–178.

Landsea, C. W. 2005. Hurricanes and global warming. *Nature* **438**, 11–13, doi:10.1038/nature04477.

Mann, M. E. and Emanuel, K. A. 2006. Atlantic hurricane trends linked to climate change. *EOS* **87**, 233, 238, 241.

Mosedale, T. J., Stephenson, D. B., Collins, M. and Mills, T. C. 2006. Granger causality of coupled climate processes: ocean feedback on the North Atlantic Oscillation. *J. Climate* **19**, 1182–1194.

Neumann, C. J., Jarvinen, B. R., McAdie, C. J. and Hammer, G. R. 1999. *Tropical cyclones of the North Atlantic Ocean, 1871–1998*. National Oceanic and Atmospheric Administration, 206 pp.

Pielke, R. A. Jr., Landsea, C., Mayfield, M., Laver, J. and Pasch, R. 2005. Hurricanes and global warming. *Bull. Am. Meteorol. Soc.* **86**, 1571–1575.

Rayner, N. A. and co-authors, 2003. Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.* **108**(D14), 4407, doi:10.1029/2002JD002670.

Richards, G. R. 1993. Change in global temperature: A statistical analysis. *J. Climate* **6**, 546–559.

Shapiro, L. J. and Goldenberg, S. B. 1998. Atlantic sea surface temperatures and tropical cyclone formation. *J. Climate* **11**, 578–590.

Theiler, J., Galdrikian, B., Longtin, A., Eubank, S. and Farmer, J. D. 1992. Using surrogate data to detect nonlinearity in time series. In: *Nonlinear Modelling and Forecasting* (eds M. Casdagli and S. Eubank), Santa Fe Institute, Addison Wesley, Boston, 163–188.

Trenberth, K. 2005. Uncertainty in hurricanes and global warming. *Science* **308**, 1753–1754.

Trenberth, K. and Shea, D. J. 2006. Atlantic hurricanes and natural variability in 2005. *Geophys. Res. Lett.* **33**, L12704, doi:10.1029/2006GL026894.

Triacca, U. 2001. On the use of Granger causality to investigate the human influence on climate. *Theor. Appl. Climatol.* **69**, 137–138.

Wang, W., Anderson, B. T., Kaufmann, R. K. and Myneni, R. B. 2004. The relation between the North Atlantic oscillation and SSTs in the North Atlantic basin. *J. Climate* **17**, 4752–4759.

Webster, P. J., Holland, G. J., Curry, J. A. and Chang, H. R. 2005. Changes in tropical cyclone number, duration, and intensity in a warming environment. *Science* **309**, 1844–1846.

Zivot, E. and Wang, J. 2002. *Modeling Financial Time Series with S-PLUS*. Insightful, Seattle, 653 pp.