

Unit roots, Granger causality and global warming: bridging the gap between econometric theory and climatology

Domenico Vitale*

Massimo Bilancia*

*Department of Statistics “C. Cecchi”, Bari University
domenicovitale1981@libero.it, mabil@dss.uniba.it

Abstract: In the last few years, improved mathematical modelling and abundance of data, strengthened the conclusion that human emissions are very likely to cause serious climate change. The adoption of the Granger causality test applied to temperature and CO₂ emissions series, allowed us to analyze the nature of such associations.

Keywords: Granger causality, Environmental Statistics.

1. Introduction

The forth IPCC report, published in 2007, affirmed that by the end of the century the planet's average temperature is expected to rise anywhere between about 1.4 and 6.0°C. Many scientists agree about the fact that CO₂ emissions play a positive and predominant role in global climate warming but many controversies exist. Given the aforementioned scenario, advanced time series methods rooted in econometrics have been applied to climatic time series having this objective in mind. The main tool of our analysis is the notion of Granger causality between two time series (Granger and Lee, 1989).

2. Methods

Before the application of any causality test, the series have to be made stationary. There are two approaches to describing data non-stationarities: the *trend stationary* model (TS) in which we have the constant mean of the stationary process replaced by a linear function of the date and the *unit root* or *difference stationary* process (DS). To distinguish between the two specifications the augmented Dickey-Fuller (ADF) test was adopted (Dickey and Fuller, 1981). The complete setting used in this paper for the ADF procedure considers three different regression equations that can be used to test for the presence of a unit root in Y_t : the model with trend plus drift (1), the one with drift only (2) and the one where neither a drift nor a trend (3) is included:

$$(1 - \rho L)Y_t = \mu + \delta t + \{(\lambda_1 L + \lambda_2 L^2 + \dots + \lambda_{p-1} L^{p-1})(1 - L)\}Y_t + \varepsilon_t \quad (1)$$

$$(1 - \rho L)Y_t = \mu + \{(\lambda_1 L + \lambda_2 L^2 + \dots + \lambda_{p-1} L^{p-1})(1 - L)\}Y_t + \varepsilon_t \quad (2)$$

$$(1 - \rho L)Y_t = \{(\lambda_1 L + \lambda_2 L^2 + \dots + \lambda_{p-1} L^{p-1})(1 - L)\}Y_t + \varepsilon_t \quad (3)$$

Dolado et al. (1990) note that the additional estimated parameters in the model, as in (1), reduce the power of tests, where reduced power means that the researcher may conclude very often that the process contains a unit root when this is not the case. For these reasons, they proposed a systematic testing strategy between the three alternative equations that was adopted in this paper.

Assume now that Y_t and Z_t are second-order stationary time series, and let the information set $\mathfrak{S}_t$ have the form $(Y_t, Z_t, Y_{t-1}, Z_{t-1}, \dots, Y_1, Z_1)$: we say that Z_t is *Granger causal* for Y_t with respect to $\mathfrak{S}_t$ if the variance of the optimal linear predictor of Y_{t+s} based on $\mathfrak{S}_t$ has smaller variance for any forecasting horizon than the optimal linear predictor of Y_{t+s} based on $(Y_t, Y_{t-1}, \dots, Y_1)$. In other words Z_t is Granger causal for Y_t if Z_t helps to linearly predict Y_t at some stage in the future. If Granger causality in one direction is found, the shorter terms “causes” can be used as a shorthand for “Granger causes”. Assuming a particular lag-length m and given the information set $\mathfrak{S}_t$, we estimate by OLS the following (unrestricted) linear model:

$$Y_t = c_1 + \alpha_1 Y_{t-1} + \dots + \alpha_m Y_{t-m} + \beta_1 Z_{t-1} + \dots + \beta_m Z_{t-m} + u_t \quad (4)$$

where $\{u_t\}$ is a disturbance term. Under the null hypothesis $H_0 : \beta_1 = \beta_2 = \dots = \beta_m = 0$ that Z_t does not Granger (linearly) cause Y_t we have the following restricted linear model estimated by OLS:

$$Y_t = c_0 + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \dots + \alpha_m Y_{t-m} + \varepsilon_t \quad (5)$$

For fixed regressors and Gaussian disturbances standard F -test of the null hypothesis has an exact F distribution, whereas an asymptotically equivalent test valid for lagged dependent regressors is given in Hamilton (1994, p. 305).

3. Main results

The analyses were based on 158-years historical data from 1850 to 2007. Global CO_2 emissions series were available from the Carbon Dioxide Information Analysis Center (<http://cdiac.ornl.gov/>). The temperature anomalies series over the same period were available from the UK Met Office Hadley Centre (<http://hadobs.metoffice.com/>).

As shown in Figure 1, the stationary condition, need to the causality test application, is not satisfied for both variables. The CO_2 emissions have an increasing trend imputable to an untangled mixture of long-period natural cyclic waves and increasing anthropogenic gas emissions: half of these emissions have occurred since the mid 1970s. Concerning the climate series, an increasing trend is apparent from 1980s until today, and the last few years resulted to be the warmest during the whole period taken into consideration. The scatter plot between the series (Fig. 2) denotes an high degree of correlation between the two phenomena, albeit this is likely to be a case of spurious regression.

Following the to Dolado's procedure, both series were tested for the presence of unit roots by means of a sequential application of ADF test. The lag-length p was set in agreement to the Schwert formula (Schwert, 1989).

Series *Temp* and $\log CO_2$ resulted to be $I(1)$, but this conclusion is likely to be biased from the fact that $\log CO_2$ records presented a non-stationarity in variance. In order to remove such temporal heteroschedasticity (from which inconsistent results may follow), we divided the observation period into three sub-windows: up to 1899, from 1900 up to 1950 and after 1950 (Fig. 3). Such new series were tested for the presence of unit roots: for each of the three periods considered *Temp* resulted to be $I(1)$, while $\log CO_2$ was judged to be $I(2)$, except for the mid period in which a first-order integration resulted to be appropriate. Approximate variance homogeneity was assessed on the ground of the Brown-Forsythe-Levene type test by splitting each sub-window into $k = 4$ equally spaced intervals (Brown and Forsythe, 1974). Differencing the series ensures the stationary conditions and eliminates stochastic trends. After these transformations the degree of dependence between the new variables appears lower negligible (Fig. 4).

Subsequently, Granger causality between integrated time series was assessed: results of tests are shown in Table 1. The choice of lag-length m in (5) was based on several data information criteria: Akaike (AIC), Hannan and Quinn (HQ), Schwarz (SC) and *final prediction error* (FPE). *P-values* suggested that $\log CO_2$ Granger causes *Temp* series in the first and last sub-window, when m is set respectively equal to 7 and 4. In the mid interval a feedback system was observed, a form of association which is very difficult to interpret in the present context.

4. Conclusions

In this paper we applied several tools from econometric theory to the analysis of bivariate climatic time series: we found some weak evidences for the presence of Granger causality from CO_2 emissions to global surface temperature (see also Triacca, 2005, based on different data sources). It must be pointed out that our conclusions are not definitive, and suffers of weaknesses inherent in our approach: in particular, unit roots and Granger causality tests are known to be extremely sensitive to lag-length choice (Hamilton, 1994). Structural analysis based on cointegration theory is a far more robust approach that will be pursued in future papers, in order to confirm the presence of an unidirectional association between CO_2 emissions and global warming.

Table 1: Granger causality test: $Z \not\rightarrow Y$ stands for “Z fails to Granger cause Y”

Period	Null Hypothesis	m	P-value	m	P-value	m	P-value
from 1850 up to 1899	$\log CO_2 \not\rightarrow Temp$	1	0.510	7	0.096	15	0.740
	$Temp \not\rightarrow \log CO_2$	1	0.706	7	0.690	15	0.724
from 1900 up to 1950	$\log CO_2 \not\rightarrow Temp$	1	0.188	15	0.023		
	$Temp \not\rightarrow \log CO_2$	1	0.780	15	0.031		
from 1951 up to 2007	$\log CO_2 \not\rightarrow Temp$	2	0.248	4	0.031	15	0.138
	$Temp \not\rightarrow \log CO_2$	2	0.365	4	0.668	15	0.844

Figure 1: Global Temperature Anomalies and CO_2

Figure 2: Scatter plot

emissions series

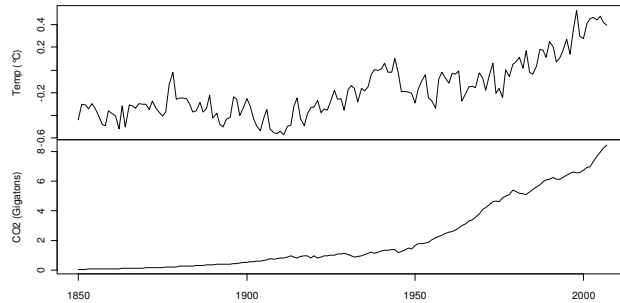


Figure 3: Global Temperature Anomalies and logCO₂ emissions after differencing

between original series

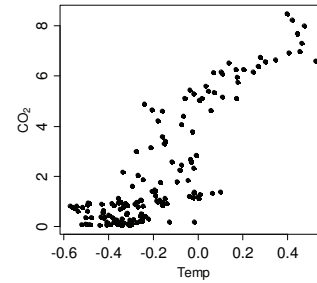
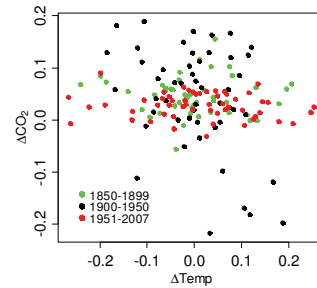
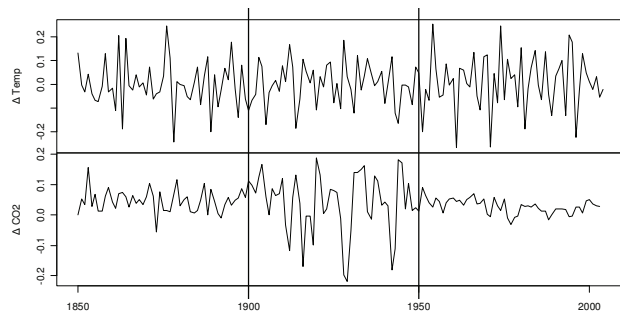


Figure 4: Scatter plot between differenced series



References

- Brown M.B., Forsythe A.B. (1974) Robust tests for equality of variances, *Journal of the American Statistical Association*, 69, 364-367.
- Dickey D.A., Fuller W.A. (1981) Likelihood ratio statistics for autoregressive time series with a unit root, *Econometrica*, 49, 1057-1072.
- Dolado J.J., Jenkinson T., Sosvilla-Rivero S. (1990) Cointegration and unit roots, *Journal of Economic Surveys*, 4(3), 249-273.
- Granger C.W.J., Lee T.H. (1989) Investigation of production, sales and inventory relationships using multicointegration and non-symmetric error correction models, *Journal of Applied Econometrics*, 4, 145-159.
- Hamilton J.D. (1994) *Time Series Analysis*, Princeton University Press, New Jersey.
- Schwert G.W. (1989) Tests for Unit Roots: A Monte Carlo Investigation, *Journal of Business & Economic Statistics*, 7(5), 147-159.
- Triacca U. (2005) Is Granger causality analysis appropriate to investigate the relationship between atmospheric concentration of carbon dioxide and global surface air temperature?, *Theoretical and Applied Climatology*, 81, 133-135.