New Cointegration Methods for Detection and Attribution of Climate Trends

Alemtsehai A. Turasie, David B. Stephenson, Tim E. Jupp and James Davidson

University of Exeter, UK

EGU2012

Detection

IPCC defined detection of climate change as a process of demonstrating that climate or a system affected by climate has changed in some defined statistical sense without providing a reason for that change

Attribution

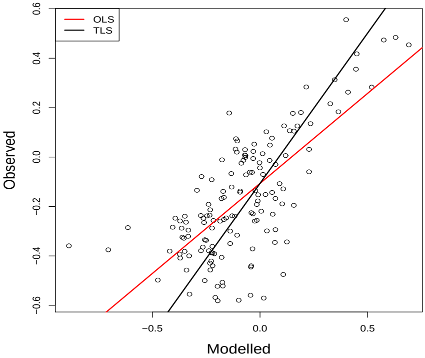
Attribution is a process of evaluating the relative contributions of multiple causal factors to a change or event with an assignment of statistical confidence.

Reliable detection and attribution of changes in climate is fundamental in enabling decision makers to manage climate-related risk (Hegerl et al., 2010).

Existing regression approach

The ordinary least squares (OLS) and/or total least squares (TLS) regression methods have been employed in detection-attribution studies (Allen and Stott, 2003; Stott et al., 2003).

The OLS and TLS fits



$$\left.\begin{aligned} \text{OLS} : \mathbf{y} &= \mathbf{X}\beta + \epsilon_t \\ \text{TLS} : \mathbf{y} &= (\mathbf{X} - \mathbf{w})\beta + \epsilon_t \end{aligned}\right\} \begin{aligned} &\mathbf{y} - \textit{observed}, \mathbf{X} - \textit{modelled}, \\ &\epsilon - \text{climate noise \& } \mathbf{w} - \text{model uncertainty} \end{aligned}$$

Detection: reject $H_0 : \beta = \mathbf{0}$
Attribution: reject $H_0 : \beta_f = 0$ for a particular forcing f

Why New Method?

- Most climatic time series are non-stationary.
- Non-stationary time series can lead to spurious regressions (Granger and Newbold, 1974).
- Better to see if one time series x can be used to detrend another time series y
- Vector autoregressive (VAR) models are useful in capturing the linear interdependencies among multiple time series.

Cointegrating time series model: VAR(2) model

VAR(2)

For a bivariate vector $\mathbf{z}_t = \begin{pmatrix} y_t \\ x_t \end{pmatrix}$, we use a VAR(2) model:

$$\mathbf{z}_t = \mathbf{\Pi}_1 \mathbf{z}_{t-1} + \mathbf{\Pi}_2 \mathbf{z}_{t-2} + \epsilon_t, \qquad t = 1, 2, ..., T$$

VECM

The vector error correction model, VECM

$$\Delta \mathbf{z}_t = \mathbf{\Pi} \mathbf{z}_{t-1} + \mathbf{\Gamma} \Delta \mathbf{z}_{t-1} + \epsilon_t$$

is estimated using Johansen's maximum likelihood method.

Given the rank of $\mathbf{\Pi}$ matrix is 1, the long-run coefficient β can be estimated from

$$\mathbf{\Pi} = \alpha \beta'$$

Synthetic and Real Data

Simple "toy" models are used to generate synthetic data.

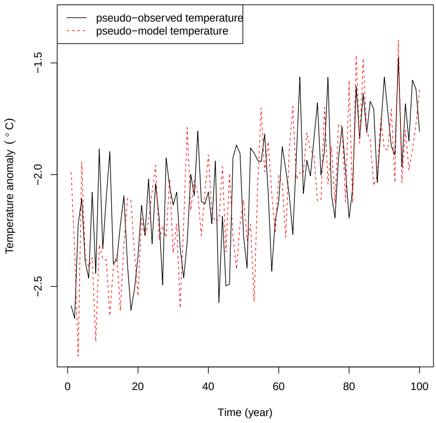
$$\left.\begin{aligned} c_t &= e + c_{t-1} + \varepsilon_t \\ y_t &= \alpha_0 + \alpha_1 c_t + \xi_t \\ x_t &= \gamma_0 + \gamma_1 c_t + \nu_t \end{aligned}\right\} \quad \begin{aligned} &c_t - CO_2 \text{ in year } t, e - \textit{emissions}, \\ &y_t - \text{observed temperature}, x_t - \text{modelled temp} \end{aligned}$$

Parameters estimated using the 20^{th} century annual observation from HadCRUT3 and GISS model simulations.

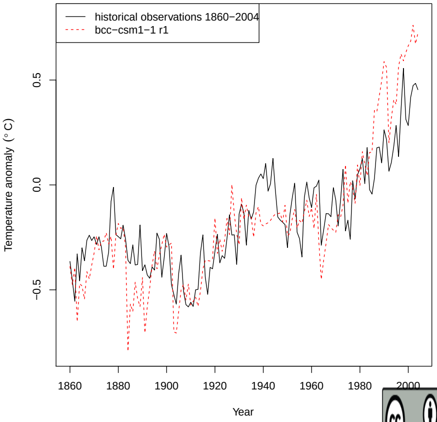
Historical near surface temperature (Brohan et al., 2006) and 16 of the models included in CMIP5 simulations are also used.

Time series of synthetic and real data

Example of synthetic data



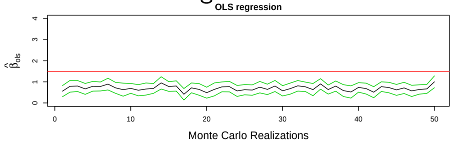
Historical global mean temperature and example of CMIP5 simulation



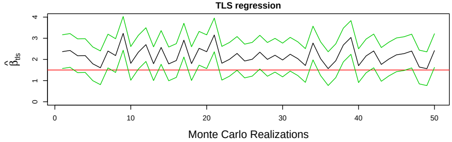
⇒ Non-stationary

Estimates for 50 MC realizations

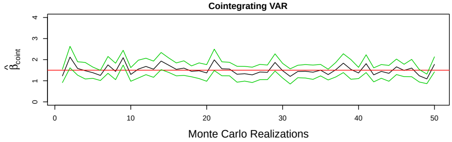
Estimates for 50 Monte Carlo realizations of length 100



Negatively biased?

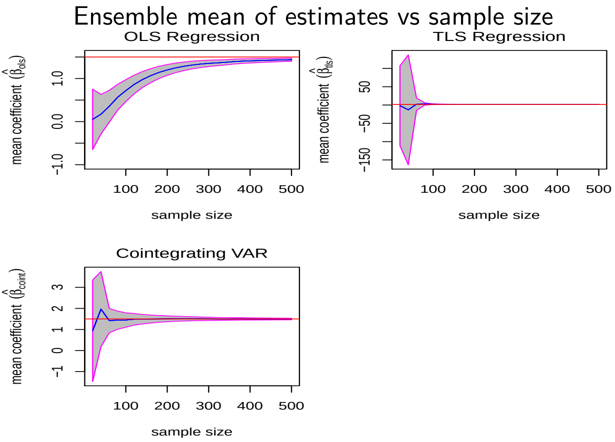


Less efficient?



Less biased and more efficient?

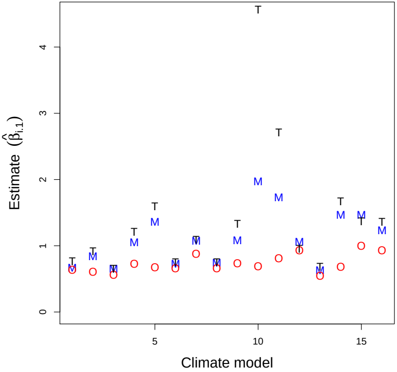
Sampling properties of the estimators



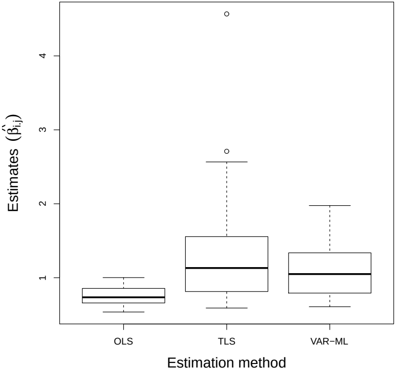
⇒ OLS-biased, TLS-very uncertain, VAR-ML-better than the two

Distribution of estimates: using real observations and CMIP5 simulations

Estimates-one run from each model:O-OLS, M-VAR-ML, and T-TLS



Distribution of estimates: for all realizations



⇒ The cointegrating VAR estimates are not biased to either extremes.

⇒ The pattern/distribution of estimates do back results from synthetic data.

Summary

- The static regression methods could end-up with spurious results for non-stationary climatic variables- high risk of misleading policy makers.
- The dynamic VAR based MLE are less biased and more efficient than those of the OLS and TLS estimates of static regression.
- We can do better with cointegrating VAR method in detection and attribution studies.

End

Thank you

Questions?

Contact: aatt201@exeter.ac.uk
alemtsehaia@yahoo.co.uk