

ESF EXPLORATORY WORKSHOP

Econometric Time-Series Analysis Applied To Climate Research

Frascati, Italy 6-7 September 2007

Scientific report by:

Torben Schmith

Søren Johansen

Hans von Storch

Peter Thejll

Executive summary

This workshop was convened by Mr. Torben Schmith, Dr. Peter Thejll from Danish Meteorological Institute, prof. Søren Johansen from Department of Statistics and Operations Research, Institute of Mathematical Sciences, University of Copenhagen, Denmark and prof., dir. Hans von Storch from Institute of Coastal Research, GKSS Research Center, Germany and was held on 6. and 7. September 2007 in Hotel Antica Colonia in Frascati outside Rome, Italy. There was an attendance of 27 people, mainly within the expertise fields of climate science and econometrics.

The purpose of the workshop was to bring together scientists from basically two different fields, namely climate science and econometric statistics, in order to explore the possibilities of transferring methods of econometric time series analysis of non-stationary series to the field of climate series. The motivation was that time series from the two fields often have in common that they are nonstationary in time. This often leads to confusion in climate science and standard methods valid only for stationary time series are uncritically used for the analysis of non-stationary time series. Within econometrics, methods have been developed over the recent decades to handle relationships among nonstationary time series, including estimation and inference, and it is appealing to investigate the possibilities for using these methods on climate problems.

Therefore, the workshop was fundamentally 'asymmetric' in its nature, in that one part, the climate scientists, supplied relevant climate problems whereas the other part, the econometricians, delivered suitable methods.

The fact that the workshop was held in a hotel, where the attendees also stayed created a very intense atmosphere with discussions during meals and other breaks.

Through the support granted by the European Science Foundation it was possible to cover living costs (accommodation and all meals) for all attendees of the workshop, whereas travel was paid by the attendees themselves, except for one overseas exception.

The scientific program on the two workshop days was from 9.00 am to 5.30 pm, with alterations between longer review presentations (typically 45 min) of central themes and group discussions inspired by these. In these discussions there were a lively discussion among the attendees, and the session chairmen had a hard time keeping to the time schedule. The afternoon of the last workshop day was reserved for summing up the workshop, identifying research directions and discussing a future framework of collaboration.

1 Scientific content of the event

The changing climate of Earth due to emission of radiatively active gases, like carbon dioxide, methane and others, into the atmosphere has become a matter of international concern on the political agenda. Climate science is characterised by long time scales making the usual scientific approach with hypothesis formulation and controlled experiments and testing on independent data impossible. Instead, 'experiments' with complicated climate models based on the physical laws of Nature are extensively used for testing hypotheses, studying mechanisms etc. Such model experiments produce a vast amount of data, from which knowledge needs to be condensed, and in that process statistics is an essential tool. Complementary to experiments using climate models is the analysis of instrumental data and the reconstruction of past climate from indirect evidence, so-called climate proxies and also here statistical analysis is widely used.

In summary, in climate science statistical analysis is essential. Below we describe some topics presented at the conference, some statistical tools and several topics in climate science, where statistical analysis is essential.

1.1 Statistical analysis of non-stationary time series

Spurious regression

Regression analysis of time series conducted by standard 'ordinary least squares' (OLS) methods can be very misleading if applied to time series data. This is due to the fact that many time series exhibit trends and thus violate the assumption of data stationarity. Both economic as well as climate series share this problem.

Even purely random series may under certain circumstances correlate strongly without representing a meaningful economic or physical relationship. This leads to a phenomenon called 'spurious regression' which is explained using different concepts of non-stationarity (difference stationarity, trend stationarity). Spurious regression was first noted by Yule (1925) and later Granger and Newbold (1974) demonstrated spurious regression in the case of integrated processes, where the time-differenced series but not the series themselves are stationary.

Basic testing procedures for stationarity exists, such as the Dickey-Fuller test and the augmented Dickey-Fuller test, but a problem is the low power of these tests.

The demand for `non-spurious' regressions leads to the idea of an equilibrium relationship between the variables involved. It is shown, that economically or physically meaningful regressions must contain two elements, a long-run

equilibrium relationship in variable levels and a short-run adjustment mechanism. The concept of cointegration is briefly introduced.

The problems and methods were demonstrated with economic and financial time series as well as with climate data.

Statistical analysis of multivariate nonstationary time series

Both global temperature and sea level have been rising through the past century or so. In a recently published work in *Science*, this has led to fitting a statistical model and reporting too optimistic skills of the model. This is an example of a spurious correlation due to the non-stationary nature of the ingoing time series. A fruitful concept is to distinguish between sample correlation/regression (calculated by the usual formulas) and population correlation/regression. For stationary processes will sample values converge to population values, whereas for non-stationary processes this is not guaranteed.

There are different types of non-stationary processes. One class is trend stationary; by subtracting a linear trend the process becomes stationary. Another type is difference-stationary; the timewise differences becomes stationary. A given time series can exhibit both types of non-stationarity

Vector autoregressive models provide a general framework for a simultaneous description of different types of stationarity/non stationarity in vector time series. This gives the possibility to see new features of the data, to distinguish between short-run and long-run dynamics, and to conduct statistical inference on causal relationships.

Fractional integration and long range memory

The time series of global average surface temperature is obviously nonstationary, the question is whether it is stationary (I(0)) around a deterministic trend or it is difference-stationary (I(1)). Looking at the residuals, there are some indications of I(1)-ness but on the other hand, it is not really a random walk. This calls for considering a wider range of models, fractionally integrated models, of which stationary I(0) and I(1) are two particular cases. Fractional models contain a parameter, the fractional integration order d, and the value d=0.5 discriminates between stationary and non-stationary processes. I any case, fractionally processes possesses long-range memory, which means the autocorrelation function falls off slower than exponential, i.e. it is not summable. This as opposed to I(0)-processes, which have short memory, i.e. the correlation function is summable/falls off faster than exponential. In frequency domain, short memory processes are characterised by the spectral density being bounded for all frequencies, while long memory processes are unbounded at zero (or other) frequency.

As an example, previous analyses of global average surface temperature have been based on assuming either I(0)- or I(1)-ness (plus a deterministic trend). Assuming fractional integration yields a value of d near 0.5, which makes it difficult to conclude about stationarity/non-stationarity. The analysis bases on fractional integration gives a different value for the global warming rise over the past 150 years, than the above mentioned 'classical' analyses.

1.2 Climate analyses where these methods have been used

Granger causality of air-sea interactions

There is a decade-long discussion in the literature on the interaction between ocean atmosphere in the Atlantic. Whereas the atmosphere's influence of ocean temperatures are well-known, the opposite, i.e. whether anomalies in sea surface temperatures influences atmospheric circulation is still under debate. First, it is noted, that the autocorrelation of the NAO-index does not falls off exponentially, but have shoulders around monthly time scale. This is taken as an indication of ocean's influence of atmospheric circulation. Further evidence for a two-way interaction is obtained by fitting vector-autoregressive models to NAO and associated tripole pattern in the sea surface temperature field.

Applying the cointegrated VAR approach to climate data

Using long simulations with a coupled climate model forced with historical forcings solar radiation, greenhouse gases and volcanic eruptions provide a testbed for the ability of the statistical techniques to extract relevant information, like the overall climate sensitivity. This is because we in the model world, as opposed to the real world, are not hampered by incomplete and inhomogeneous observations. In addition, in the model world we also know the `future'.

Cointegration and VAR-models are applied to the time series of forcing and temperature. A number of inference procedures appropriate in integratedcointegrated vector autoregressive processes (VARs) are illustrated. Particular attention is paid to the properties of VARs, to the modeling of deterministic terms, to the determination of the number of cointegration vectors, and to testing hypotheses on long-run relation and short-run adjustment.

Emissions, Concentrations, and Temperature: A Time Series Approach

In 1990 the International Panel on Climate Change noted that "rigorous statistical tools do not exist to show whether relationships between statistically

non-stationary data of this kind are truly statistically significant.... ". This statement calls for the application for econometric techniques to evaluate questions like:

- Does human activity affect global surface temperature ?
- Does global surface temperature affect atmospheric carbon dioxide?

Both questions are answered in the affirmative based on the results of an cointegration / error correction model estimated from historical data starting in 1860. Results indicate that global surface temperature and radiative forcing cointegrate and that changes in radiative forcing move temperature towards the long run equilibrium. The long-run equilibrium effect represented by the cointegrating relationship indicates that doubling radiative forcing increases surface temperature by 2.1°C. This effect represents the transient climate response—as indicated by values estimated by using the same statistical technique to analyse climate scenarios assembled by the coupled model intercomparison project. The cointegration/error correction model can separate the temperature effect of human activity, which is largely responsible for the ongoing increase, from natural variability, which causes temperature to rise and fall over short periods.

Econometric techniques are used to evaluate the effects of global surface temperature on atmospheric concentrations of carbon dioxide. A different methodology is required because scientists cannot balance the global carbon cycle—the unknown carbon sink contains a stochastic trend. A cointegration/error correction model indicates that the unknown carbon sink is associated with anthropogenic carbon emissions and summer temperatures in the Northern Hemisphere. To avoid the effect of the unknown carbon sink on cointegration among emissions, concentrations, and temperature, we use instrumental variables to estimate a first difference equation. The results indicate that increases in global surface temperature on net, raise the atmospheric concentrations by about 1.5 ppm per degree Celsius.

On fractional integration detected in observed climate series

Climatic time series like water levels have long-term correlations. This has implications for the statistics of extreme events and return intervals between events above a given threshold. One must distinguish between (a) long-term correlated (monofractal) records where the autocorrelation function decreases by a power law and (b) in multifractal data sets where the linear autocorrelation function vanishes and only non-linear correlations are present. Both monofractal and multifractal data sets play an important role not only in climate but also in the financial markets. The long-term correlations affect the statistics of the return intervals, leading to a stretched exponential decay of the pdf and to long-term correlations between the return intervals, which then in turn leads to a clustering of extreme events. These characteristic features can be seen in long natural and reconstructed climate records as well as in financial records. Finally, in multifractal data sets that are characterized by an absence of linear correlations (examples are financial records) the non-linear correlations lead to pronounced long-term correlations of the return intervals and to a pdf that decays by a power law. Current issues in climate research where econometric methods have a potential

Climate reconstruction from proxies

Instrumental evidence of climate change (e.g. temperatures measured with a thermometer) exists only since the mid-19th century, with a few exceptions. A longer record of climate is necessary in order to provide evidence against which hypotheses of causes for climate change and models of the climate system functionality can be tested. This calls for indirect evidence in the form of information from natural and documentary archives which somehow contain information (proxies) of past variations in climate.

There are several steps from proxy series to climate reconstruction. The first step is dating, for tree rings e.g., this is a simple counting of rings. Next step is preprocessing, for tree rings this is measuring the width of each annual ring. Finally a calibration, i.e. a statistical relationship between proxy and instrumental series is established in the overlapping period of proxy and instrumental series. Finally, the calibration relationship is used to find the instrumental variable from proxies.

A range of statistical methods exist to establish the proxy reconstruction. Since 1998 there has been an intense debate about methods of global temperature reconstruction of millennium length with annual resolutions, through which much has been learned.

Statistical downscaling

Physically based climate models are used to give scenarios for future climate change due to anthropogenic change of atmospheric composition. 'Present day' models, however have a spatial resolution of a few 100 km. This means that smaller scale details are not always well simulated. This applies in particular in areas with complex topography, like the Alps.

There are two basic methods to bridge the gap between the global climate model scale and the local. One method, called *dynamical downscaling*, consists of embedding a regional climate with high resolution and 'driven' by the original global model. Alternatively, one can use *statistical downscaling*, which are ad-hoc statistical methods to relate simulated changes at large scale (predictors) to associated changes at local or regional scales (predictands). A priori any statistical model would be possible, but a reasonable choice strongly depends on: the regional variable, the timescale (monthly, daily,..), the skill of the climate model to simulate the predictor and the skill of the model to reproduce the regional variable over the period where observations are available.

The steps in the statistical downscaling procedure are first to estimate the statistical model relating *observed* predictors and predictands. Secondly, the statistical model is applied to simulated predictors from the 'present' and 'future' climate to find the climate change signal.

Different statistical methods are applied for establishing relationships between predictor(s) and predictand(s). Linear methods, like PC-regression or canonical correlation are often applied and usually yield very similar results. Another class are classification methods, like: clustering, fuzzy-logic-based and analogue methods. Some degree of subjectivity in defining the model is present and range of output is restricted to the range of past observations. Finally, neural networks are black-box type and usually too few observations are present to allow for a complex network design.

Detection of climate change and attribution of causes

The 20th century has seen climate changes: abundant evidence of rising temperatures, some evidence of changed hydrological cycles and some evidence of changed atmospheric circulation. This raises a number of questions: Do they arise from natural internal variability, or are there external causes? And if the latter, which causes, and how do we test our hypotheses? In systematic attribution studies one compares output fields (of e.g. temperature) from a global simulation with a climate model forced with observed changes in forcing with the corresponding observed fields. This involves fitting of a statistical model, often an 'error in variables' on the low-pass filtered (in space and time) data.

Sensitivity of Earth's climate to external forcings

The climate sensitivity is the equilibrium increase in global average temperature of the climate system for a doubling of carbon dioxide concentration relative to pre-industrial values, or other external forcings. Most studies of climate sensitivity are based on climate models, which are based on the physical laws of nature. The viewpoint was put forward, that this made them more reliable than the semi-empirical models used in economy, which rely predominantly on curve fitting.

In climate science there has been a shift towards the production of probability density functions. However, climate models are not independent – they share methods, parameterizations, code etc.

One big ensemble effort, *climateprediction.net*, consists of 2500 different realizations of 400 model versions, from which a histogram of climate sensitivities was presented. This stimulated a lively discussion of how this should be interpreted; does it reflect our uncertainty or is it a 'probability'?

2 Assessment of the results, contribution to the future direction of the field

The purpose of the workshop was to bring together scientists from two different fields with the aim of trying to use the methods from one field (econometric) in the other field (climate science). Judged from the scientific presentations given at the workshop and the lively discussions among its participants, both during the sessions and during meals etc., the workshop was a success.

One central issue in the debate was whether the econometric methods could bring any added value when applied to problems in climate research. To this, the view was brought forward, that correct time series modelling could supply correct error estimates. This could deliver valuable input to the ongoing discussion on climate reconstruction methods.

The question was put forward, whether econometric methods could help solving specific problems in the analysis of climate data or, alternatively, should be seen as a general way to raise the quality level of statistics in climate research. No definite answer was arrived at.

There was general enthusiasm to proceed with the interaction between econometrics and climate in some way. Different possibilities were discussed in the light of the presentation by Tiina Nöges from ESF. At the present stage, it was decided to have a similar workshop in one years time and find funding sources for that. By that time there would probably be a number of draft or published papers arising from the present Frascati-workshop. If the next workshop turns out to be a success too, it may be time to apply for a Research Networking Programme to be funded by the ESF

Finally, it was decided to maintain and extend the workshop-website to a 'focal point' for next years' workshop. The website will contain workshop-presentations, scientific papers and manuscript, relevant conferences and other material.

September 5, 2007 Arrival

September 6, 2007

09:00-09:15 Welcome

Session 1: Current issues in climate research (chair: Hans von Storch)

09:15-10:00	Anders Moberg : Reconstruction methods of past climate from proxies
10:00-10:45 10:45-11:15	Eduardo Zorita: Statistical downscaling methods General discussion

11:15-11:30 **Coffee break**

Session 2: Regressions, dynamic models and causality (chair: Jörg Breitung)

11:30-12:15	Richard Reichel: Spurious regression
12:15-13:00	Søren Johansen: Statistical analysis of multivariate
	nonstationary time series

13:00-14:15 Lunch break

Session 2 cont. (chair: Dag Tjøstheim)

14:15-15:15
 Katarina Juselius: Applying the cointegrated VAR approach to climate data
 15:15-16:00
 Robert Kaufmann: Emissions, Concentrations, and Temperature: A Time Series Approach

16:00-16:15 **Coffee break**

- 16:15-17:30 General discussion of correlations versus causality
- 19:00- **Dinner**

September 7, 2007

Session 3: Applications to climate data (chair: Hans Wackernagel)

09:00-09:45	Dave Stainforth : Sensitivity of Earth's climate to external forcings
09:45-10:30	Francis Zwiers: Detection of climate change and attribution of causes
10:30-11:00	David Stephenson : Granger causality in North Atlantic air- sea interactions
11:00-11:30	Discussion
11:30-11:45	Coffee break

Session 4: Fractional integration and climate data (Chair: Olavi Kärner)

11:45-12:15	Jörg Breitung: Fractional integration and long range memory
12:15-12:45	Armin Bunde: On fractional integration detected in observed climate series
12:45-13:15	Discussion
13:15-14:30	Lunch

Session 5: Plans for the future (Chair: Katarina Juselius)

14:30-15:00	Tiina Nöges (ESF representative): Possibilities for further development of the research area
15:00-17:30	 Series of 5-minutes presentations of current projects/ideas by workshop participants Discussion of future plans for collaboration or proposals. Summer schools, ERC and ESF possibilities discussed. A few solicited presentations of `what next' scenarios.
19:30 -	DINNER

September 8, 2007

Departure

4 Statistical information on participants

Gender: female: 4 male: 24 Age brackets: Senior: 24 Ph.D./post.doc.: 4 Country of Research Institution: Denmark: 6 Germany: 5 France: 1 Norway: 2 Switzerland: 2 Italy: 2 USA: 1 Canada: 1 Great Britain: 2 Sweden: 1

Estonia: 1 Romania: 1 Poland: 2 Spain: 1

ESF Representative not included.

5 The Final list of participants

Prof. Søren Johansen Department of Economics University of Copenhagen Studiestræde 6 1455 København K Denmark Phone: +45 35323071 Fax: +45 35320678 Email: <u>sjo@math.ku.dk</u>

Prof. Katarina Juselius Department of Economics University of Copenhagen Studiestræde 6 1455 København K Denmark Phone: 0045 35323068 Fax: 0045 35323064 Email: <u>katarina.juselius@econ.ku.dk</u>

Prof. Anders Rahbek Department of Mathematical Sciences University of Copenhagen Universitetsparken 5 2100 Copenhagen Denmark Phone: 0045 353 20682 Fax: 0045 35 32 07 04 Email: <u>rahbek@math.ku.dk</u>

Mr. Theis Lange Department of Mathematical Sciences University of Copenhagen Universitetsparken 5 2100 Copenhagen Denmark Phone: 0045 353 20682 Fax: 0045 35 32 07 04 Email: <u>lange@math.ku.dk</u> Dr. Peter Thejll Danish Meteorological Institute Lyngbyvej 100 2100 Copenhagen Denmark Phone: 0045 39 157 477 Fax: 0045 39 27 10 80 Email: <u>pth@dmi.dk</u>

Mr. Torben Schmith Centre for Ocean and Ice Danish Meteorological Institute Lyngbyvej 100 2100 Copenhagen Denmark Phone: 0045 39 15 74 44 Fax: 0045 39 27 10 80 Email: <u>ts@dmi.dk</u>

Dr. Olavi Kärner Tartu Observatory 61602, Tõravere Tartumaa Estonia Phone: (372) 7410-265 Fax: (372) 7410-205 Email: <u>olavi@aai.ee</u>

Dr. Hans Wackernagel Centre de Geostatistique de l'Ecole des Mines de Paris 35 Rue St Honore 77300 Fontainebleau France Phone: 0033 1 64 69 47 60 Fax: 0033 1 64 69 47 05 Email: <u>hans.wackernagel@ensmp.fr</u> Prof. Hans von Storch GKSS-Forschungszentrum Geesthacht GmbH Max-Planck-Straße 1 21502 Geesthacht Germany Phone: 0049 4152 / 87-1831 Fax: 0049 4152 / 87-2832 Email: hvonstorch@web.de

Dr. Eduardo Zorita GKSS-Forschungszentrum Geesthacht GmbH Max-Planck-Straße 1 21502 Geesthacht Germany Phone: 0049 4152 87-1856 Fax: 0049 4152 87-1888 Email: <u>eduardo.zorita@gkss.de</u>

Prof. Armin Bunde Fachgebiet Physik im Fachbereich 07 University of Giessen Heinrich-Buff-Ring 16 35392 Gießen Germany Phone: 0049 641 99 33000 Fax: 0049 641 99 33009 Email: <u>Armin.Bunde@physik.uni-</u> <u>giessen.de</u>

Prof. Paolo Paruolo Department of Economics University of Insubria Via Ravasi 2 21100 Varese Italy Phone: 0039 332 215500 Fax: 0039 332 215509 Email: <u>paolo.paruolo@uninsubria.it</u> Prof. Dag Tjøstheim Department of Mathematics University of Bergen Johannes Bruns gate 12 5008 Bergen Norway Phone: 0047 55 58 28 85 Fax: 0047 55 58 96 72 Email: <u>dagt@mi.uib.no</u>

Prof. Anders Rygh Swensen Department of Mathematics University of Oslo Postboks 1053, Blindern 0316 Oslo Norway Phone: 0047 228 55871 Fax: 0047 22 85 43 49 Email: <u>swensen@math.uio.no</u>

Dr. Krzysztof Fortuniak Katedra Meteorologii i Klimatologii Uniwersytet Łódzki Narutowicza 88 09-139 Łódz Poland Phone: 0042-6655954 Fax: 0042-6655951 Email: kfortun@uni.lodz.pl

Dr. Aristida Busuioc National Institute of Meteorology, Hydrology and Water Management Bucharest Sos. Bucureti-Ploiesti 97 71552 Bucharest Romania Phone: 0040 1 230 31 43 Fax: 0040 1 230 31 16 Email: busuioc@meteo.inmh.ro Dr. Anders Moberg Institutionen för naturgeografi och kvartärgeologi Stockholm universitet 106 91 Stockholm Sweden Phone: 0046 6747814 Fax: 0046 08-16 48 18 Email: <u>anders.moberg@natgeo.su.se</u>

Dr. Marcel Kuettel University of Berne Institute of Geography Hallerstrasse 12 3012 Bern Switzerland Phone: 0041 31 631 85 45 Fax: 0041 31 631 43 38 Email: kuettel@giub.unibe.ch

Dr. Dave Stainforth Dept of Physics AOPP University of Oxford Parks Road Oxford, OX1 3PU UK Phone: 0044 1865 272 093 Fax: 0044 -1865-272400 Email: das@atm.ox.ac.uk

Prof. David Stephenson University of Exeter School of Engineering,Computer Science and Mathematics, Harrison Building Room 334, North Park Road Exeter EX4 4QF UK Phone: +44 (0) 1392 269275 Fax: +44 (0) 1392 217965 Email: d.b.stephenson@exeter.ac.uk Dr. Manuel Blanco Meteologica SA C. Almansa 110 - esc. 2 - 1C 28040 Madrid Spain Phone: 0034 91 456 10 01 Fax: 0034 91 456 10 02 Email: <u>mbb@meteologica.es</u>

Prof. Richard Reichel Volkswirtschaftliches Institut Universität Erlangen-Nürnberg Postfach 3931 90020 Nürnberg Germany Phone: 0049 911 - 5302 323 Fax: 0049 911 - 5302 323 Email: <u>richard.reichel@wiso.uni-</u> <u>erlangen.de</u>

Dr. Joanna Wibig Dept. of Meteorology and Climatology University of Lodz Narutowicza 88 91-139 Lodz Poland Phone: 0048 42-6655 957 Fax: 0048 42-6655951 Email: <u>zameteo@uni.lodz.pl</u>

Prof. Robert Kaufmann Center for Energy & Environmental Studies 675 Commonwealth Avenue Boston University Boston, MA 02215 USA Phone: 617 353-3940 Fax: 617 353-5986 Email: <u>kaufmann@bu.edu</u> Mrs. Nadja Riedwyl University of Bern Institute of Geography Hallerstrasse 12 3012 Bern Switzerland Phone: 0041 31 631 88 43 Fax: 0041 31 631 43 38 Email: <u>riedwyl@giub.unibe.ch</u>

Prof. Jörg Breitung University of Bonn Department of Economics Adenauerallee 24-26 53113 Bonn Germany Phone: +49 (0)228 73-9220 Fax: +49 (0)228 73-9221 Email: <u>breitung@uni-bonn.de</u> Director Francis Zwiers Climate Research Division Environment Canada 4905 Dufferin St. Toronto, Ont. M3H 5T4 Phone: (416)739-4767 Fax: (416)739-5700 Email: francis.zwiers@ec.gc.ca

Valerio Lucarini Department of Mathematics and Computer Science University of Camerino CINFAI Unit of Camerino Via Madonna delle Carceri 9 62032 Camerino (MC), Italy. tel:(+39) 0737 40 2581 fax: (+39) 0737 63 2525 Email: lucarini@adgb.df.unibo.it