



## **The scaling of wild events in stochastic models: The Fisher limit, the Mandelbrot limit, and FARIMA as a model of the intermediate cases**

Nicholas Watkins (1,2,3)

(1) British Antarctic Survey, Cambridge, CB3 0ET, UK, (2) Centre for the Analysis of Time Series, LSE, Houghton Street, London, UK, (3) University of Warwick, Coventry, CV4 7AL, UK

Stochastic modelling is of increasing importance, both specifically in climate science and more broadly across the whole of nonlinear geophysics. Traditionally, the noise components of such models would be spectrally white (delta-correlated) and Gaussian in amplitude, and their variance (first named by Fisher in 1918) would well characterise the likely size of fluctuations. Integration, for example in autoregressive models like AR(1), would redden a noise spectrum, while multiplication in turbulent cascades could greatly increase the range of fluctuation amplitudes, but such processes would still inherit aspects of their finite variance building blocks.

In the 60s and 70s, however, Mandelbrot and others [see e.g. Watkins, GRL Frontiers, 2013] began to present evidence in nature for much stronger departures from Gaussianity (via very heavy tailed, infinite variance, distributions) and from white noise (through long range dependence (LRD) in time). He also observed intermittency, defined here as correlations between absolute magnitudes in some time series, in, for example, finance and turbulence. He proposed various models, including self-similar ones for heavy tails and LRD, and multifractal cascades for intermittency.

In this presentation we compare contrasting types of model by looking at the "wild" events that they produce. The notion of a "wild" event can be made more precise in many ways, including by its duration in time, peak amplitude, and spatial extent. Our chosen measure will be the "burst", defined as the area of a time series above a fixed threshold. We will compare burst scaling in a self-similar, LRD, heavy tailed model (LFSM, e.g. Watkins et al, PRE, 2009] with our newer results for multifractal random walks [with M. Rypdal and O. Løvstetten], and for the heavy tailed extended version of the FARIMA (1,d,0) process, which combines long range dependence with the high frequency structure familiar from AR(1). We will also discuss the physical meaning of FARIMA and its potential as a modelling tool.



## Bayesian Analysis of Non-Gaussian Long-Range Dependent Processes

Timothy Graves (1,4), Nicholas Watkins (2), Christian Franzke (2), and Robert Gramacy (3)

(1) University of Cambridge, Statistics Laboratory, Cambridge, United Kingdom. , (2) British Antarctic Survey, Cambridge, United Kingdom (nww62@yahoo.co.uk) , (3) Booth School of Business, The University of Chicago, Chicago, IL, United States. , (4) URS Corporation, 6-8 Greencoat Place, London, United Kingdom

Recent studies [e.g. the Antarctic study of Franzke, J. Climate, 2010] have strongly suggested that surface temperatures exhibit long-range dependence (LRD). The presence of LRD would hamper the identification of deterministic trends and the quantification of their significance. It is well established that LRD processes exhibit stochastic trends over rather long periods of time. Thus, accurate methods for discriminating between physical processes that possess long memory and those that do not are an important adjunct to climate modeling. As we briefly review, the LRD idea originated at the same time as H-selfsimilarity, so it is often not realised that a model does not have to be H-self similar to show LRD [e.g. Watkins, GRL Frontiers, 2013].

We have used Markov Chain Monte Carlo algorithms to perform a Bayesian analysis of Auto-Regressive Fractionally-Integrated Moving-Average ARFIMA(p,d,q) processes, which are capable of modeling LRD. Our principal aim is to obtain inference about the long memory parameter,  $d$ , with secondary interest in the scale and location parameters. We have developed a reversible-jump method enabling us to integrate over different model forms for the short memory component. We initially assume Gaussianity, and have tested the method on both synthetic and physical time series.

Many physical processes, for example the Faraday Antarctic time series, are significantly non-Gaussian. We have therefore extended this work by weakening the Gaussianity assumption, assuming an alpha-stable distribution for the innovations, and performing joint inference on  $d$  and  $\alpha$ . Such a modified FARIMA(p,d,q) process is a flexible, initial model for non-Gaussian processes with long memory. We will present a study of the dependence of the posterior variance of the memory parameter  $d$  on the length of the time series considered. This will be compared with equivalent error diagnostics for other measures of  $d$ .



## **Compound Extremes and Bunched Black (or Grouped Grey) Swans.**

Nicholas Watkins (1,2,3)

(1) British Antarctic Survey, Cambridge, United Kingdom (nww@bas.ac.uk), (2) Centre for the Analysis of Time Series, LSE, London, UK, (3) Centre for Fusion, Space and Astrophysics, University of Warwick, Coventry, UK

Observed “wild” natural fluctuations may differ substantially in their character. Some events may be genuinely unforeseen (and unforeseeable), as with Taleb’s “black swans”. These may occur singly, or may have their impact further magnified by being “bunched” in time. Some of the others may, however, be the rare extreme events from a light-tailed underlying distribution. Studying their occurrence may then be tractable with the methods of extreme value theory [e.g. Coles, 2001], suitably adapted to allow correlation if that is observed to be present.

Yet others may belong to a third broad class, described in today’s presentation [ reviewed in Watkins, GRL Frontiers, 2013, doi: 10.1002/grl.50103]. Such “bursty” time series may show comparatively frequent high amplitude events, and/or long range correlations between successive values. The frequent large values due to the first of these effects, modelled in economics by Mandelbrot in 1963 using heavy- tailed probability distributions, can give rise to an “IPCC type I” burst composed of successive wild events. Conversely, long range dependence, even in a light-tailed Gaussian model like Mandelbrot and van Ness’ fractional Brownian motion, can integrate “mild” events into an extreme “IPCC type III” burst.

I will show how a standard statistical time series model, linear fractional stable motion (LFSM), which descends from the two special cases advocated by Mandelbrot, allows these two effects to be varied independently, and will present results from a preliminary study of such bursts in LFSM. The consequences for burst scaling when low frequency effects due to dissipation (FARIMA models), and multiplicative cascades (such as multifractals) are included will also be discussed, and the physical assumptions and constraints associated with making a given choice of model.



## **A spatio-temporal analysis of US station temperature trends over the last century**

Vincenzo Capparelli (1), Christian Franzke (2), Antonio Vecchio (1), Mervyn P Freeman (2), Nicholas W. Watkins (2,3,4), Vincenzo Carbone (1,5)

(1) Dipartimento di Fisica, Università della Calabria, Cosenza, Italy (vincenzo.capparelli@fis.unical.it), (2) British Antarctic Survey, Cambridge, UK, (3) Centre for the Analysis of Time Series, London School of Economics and Political Science, Houghton Street, London, UK, (4) Centre for Fusion, Space and Astrophysics, University of Warwick, Coventry, UK, (5) IPCF/CNR Università della Calabria, Cosenza, Italy

This study presents a nonlinear spatio-temporal analysis of 1167 station temperature records from the United States Historical Climatology Network covering the period from 1898 through 2008. We use the Empirical Mode Decomposition (EMD) method to extract the generally nonlinear trends of each station. The statistical significance of each trend is assessed against three null models of the background climate variability, represented by stochastic processes of increasing temporal correlation length. We find strong evidence that more than 50 percent of all stations experienced a significant trend over the last century with respect to all three null models. A spatio-temporal analysis reveals a significant cooling trend in the South-East and significant warming trends in the rest of the contiguous US. It also shows that the warming trend appears to have migrated equatorward and possibly also in altitude. This shows the complex spatio-temporal evolution of climate change at local scales



## Mapping the changing pattern of local climate as an observed distribution

Sandra Chapman (1,4), David Stainforth (1,2), Nicholas Watkins (1,2,3)

(1) University of Warwick, Physics Dept., Coventry CV4 7AL, United Kingdom (s.c.chapman@warwick.ac.uk, +44 2476 692016), (2) CATS, London School of Economics, London, United Kingdom, (3) British Antarctic Survey, Cambridge, United Kingdom, (4) University of Tromsø, Dept. of Mathematics and Statistics, Norway

It is at local scales that the impacts of climate change will be felt directly and at which adaptation planning decisions must be made. This requires quantifying the geographical patterns in trends at specific quantiles in distributions of variables such as daily temperature or precipitation. Here we focus on these local changes and on the way observational data can be analysed to inform us about the pattern of local climate change. We present a method[1] for analysing local climatic timeseries data to assess which quantiles of the local climatic distribution show the greatest and most robust trends. We demonstrate this approach using E-OBS gridded data[2] timeseries of local daily temperature from specific locations across Europe over the last 60 years. Our method extracts the changing cumulative distribution function over time and uses a simple mathematical deconstruction of how the difference between two observations from two different time periods can be assigned to the combination of natural statistical variability and/or the consequences of secular climate change. This deconstruction facilitates an assessment of the sensitivity of different quantiles of the distributions to changing climate. Geographical location and temperature are treated as independent variables, we thus obtain as outputs the pattern of variation in sensitivity with temperature (or occurrence likelihood), and with geographical location. We find as an output many regionally consistent patterns of response of potential value in adaptation planning. We discuss methods to quantify and map the robustness of these observed sensitivities and their statistical likelihood. This also quantifies the level of detail needed from climate models if they are to be used as tools to assess climate change impact.

[1] S C Chapman, D A Stainforth, N W Watkins, 2013, On Estimating Local Long Term Climate Trends, Phil. Trans. R. Soc. A, in press

[2] Haylock, M.R., N. Hofstra, A.M.G. Klein Tank, E.J. Klok, P.D.

Jones and M. New. 2008: A European daily high-resolution gridded dataset of surface temperature and precipitation. J. Geophys. Res (Atmospheres), 113, D20119, doi:10.1029/2008JD10201



## **An observationally centred method to quantify local climate change as a distribution**

David Stainforth (1,2), Sandra Chapman (1,4), Nicholas Watkins (1,2,3)

(1) University of Warwick, Physics Dept., Coventry CV4 7AL, United Kingdom (s.c.chapman@warwick.ac.uk, +44 2476 692016), (2) CATS, London School of Economics, London, United Kingdom, (3) British Antarctic Survey, Cambridge, United Kingdom, (4) University of Tromsø, Dept. of Mathematics and Statistics, Norway

For planning and adaptation, guidance on trends in local climate is needed at the specific thresholds relevant to particular impact or policy endeavours. This requires quantifying trends at specific quantiles in distributions of variables such as daily temperature or precipitation. These non-normal distributions vary both geographically and in time. The trends in the relevant quantiles may not simply follow the trend in the distribution mean. We present a method[1] for analysing local climatic timeseries data to assess which quantiles of the local climatic distribution show the greatest and most robust trends. We demonstrate this approach using E-OBS gridded data[2] timeseries of local daily temperature from specific locations across Europe over the last 60 years. Our method extracts the changing cumulative distribution function over time and uses a simple mathematical deconstruction of how the difference between two observations from two different time periods can be assigned to the combination of natural statistical variability and/or the consequences of secular climate change. This deconstruction facilitates an assessment of the sensitivity of different quantiles of the distributions to changing climate. Geographical location and temperature are treated as independent variables, we thus obtain as outputs how the trend or sensitivity varies with temperature (or occurrence likelihood), and with geographical location. These sensitivities are found to be geographically varying across Europe; as one would expect given the different influences on local climate between, say, Western Scotland and central Italy. We find as an output many regionally consistent patterns of response of potential value in adaptation planning. We discuss methods to quantify the robustness of these observed sensitivities and their statistical likelihood. This also quantifies the level of detail needed from climate models if they are to be used as tools to assess climate change impact.

[1] S C Chapman, D A Stainforth, N W Watkins, 2013, On Estimating Local Long Term Climate Trends, Phil. Trans. R. Soc. A, in press

[2] Haylock, M.R., N. Hofstra, A.M.G. Klein Tank, E.J. Klok, P.D. Jones and M. New. 2008: A European daily high-resolution gridded dataset of surface temperature and precipitation. J. Geophys. Res (Atmospheres), 113, D20119, doi:10.1029/2008JD10201