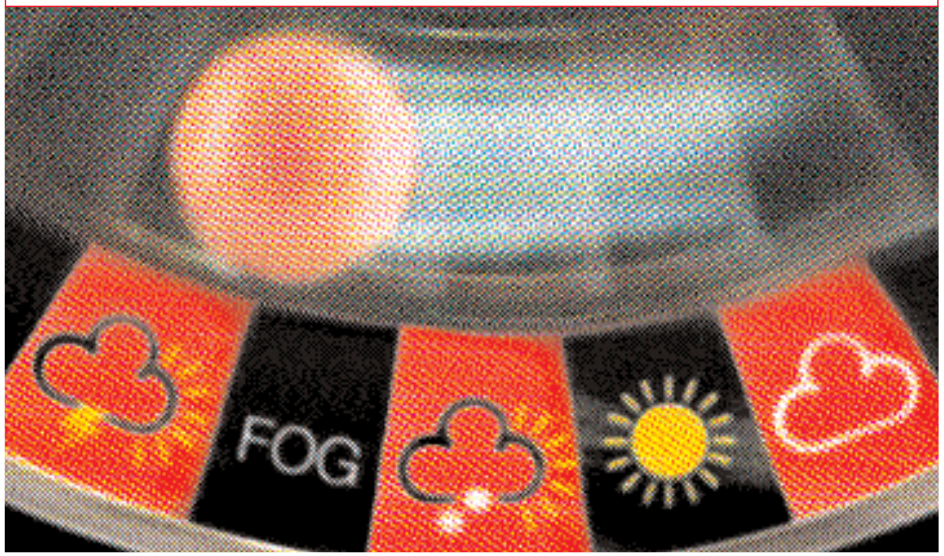

Weather roulette

The Centre for the Analysis of Time Series (CATS) was established within the Department of Statistics at LSE in 2000. The School has a long and distinguished history in time series analysis and continues to develop a world class centre of excellence in this area. **Leonard Smith** explains how weather and epilepsy fit into the picture.

Imagine if you knew exactly what the weather would be like each day this summer. How much easier it would be for people planning days out, or for hotel and theme park owners, transport chiefs and ice cream companies, if they knew that 4 July was going to fall in the middle of a heatwave. Or that on the last Monday in August, the UK's national summer holiday and the main day of London's Notting Hill Carnival, it would rain mid-afternoon.

But even if you never know exactly what will happen, knowing fair odds, or ideally the relative probability of different scenarios, allows much better decision-making: you can hedge your bets.

Can we be more successful at 'weather roulette' than we currently are? At CATS, we believe the answer is yes. This summer our Faraday DIME project has launched new technology, developed while working with data from the Met Office, the European Centre for Medium-range Weather Forecasts and the US National Center for Environmental Protection (NCEP). While we are still a long way from predicting what will happen on 25 August, we are now posting two-week probability forecasts for the weather in the UK.

Weather forecasts are made by feeding an estimate of the current state of the atmosphere into a computer programme and then simulating how the atmosphere will behave over the next couple of weeks. Sometimes a small error in the estimate of the current state will make a big difference to how the forecast turns out. Other times, because the atmosphere is in a more stable state, small errors are less important. And then, sometimes, the forecast goes wrong due to shortcomings of the model.

The DIME software is based on nonlinear mathematics. What myself and colleagues Drs Mark Roulston, Pat McSharry and Liam Clarke have done is take the ensemble weather simulations and reconfigured these simulations into forecasts which reflect the relative probability of it being hot, normal, or cold. Everyone knows that, sometimes, the forecast is really good, even many days in advance. And sometimes it is not. The DIME forecasts provide a way to see if the model thinks its forecast will be precise.

How does this fit in with epilepsy research, you may be thinking? The answer is invention and justification of new nonlinear statistical techniques. In linear systems, the effect is proportional to the strength of the action taken: in a nonlinear system this isn't always the case. Sometimes pushing harder can cause the opposite of pushing softly. So nonlinear analysis can often see things that are literally invisible to linear statistical analysis. The key of course is to be relatively certain that what is 'seen' is really there.

CATS has four main research themes. Two of these reflect two different aspects of predictability. The first attempts to trace out what will happen in the near term – for example, weather forecasts. The second is more concerned with general behaviour and possible events – for example, climate modelling. The third main research theme is nonlinear time series analysis. This is where most of the research you'd expect to find in a statistics department happens, although the approach, while broad, is closely linked to real data and importantly includes comparison with traditional time series techniques. Finally, we consider novel modelling methods, in particular comparing simple observation based models with complicated simulation models. We are interested not only in merging techniques and physical insight, and in evaluating the result in real applications, but also the philosophical underpinning of model building and data analysis – hence we maintain active links with the School's Centre for Philosophy of Natural and Social Science.

And economics? Well, new time series analysis techniques are always applied to financial data, and the provision of tests of goodness, or at least of internal consistency tests, is in great demand. In addition to targeting a better understanding of how to price weather-derivative or weather-related insurance, we promote an end to end approach to weather forecasting, tracing our current uncertainty about the state of the system all the way through the modelling process to enable a user to quantify more clearly the uncertainty in the future economic variables of interest to them, be it airport runway usage, wind energy or the efficiency of electricity generation. The DIME website also includes forecasts for heating degree days.

In practice therefore, our aim is to make the application of statistics much more relevant and immediately useful to a variety of professionals. This has in itself brought us into a debate over methods and scientific ethics. In a letter published in *Nature* magazine in March, myself and my CATS colleague Patrick McSharry, together with Oxford professor Lionel Tarresinko, argued that traditional statistics sometimes contain as much information as new fangled nonlinear creations, the behaviour of which is much less well understood.

Our concern was about research which suggested a new nonlinear way in which epileptic seizures could be more accurately predicted. Epilepsy is one of the most common serious neurological disorders, affecting up to five per cent of the population at some point. Reliable, robust detection of precursors to seizures would improve the quality of life of many epilepsy sufferers. We were concerned that some recent research might have been oversold, in particular that novel methods were presented without a strenuous attempt to compare them with more traditional methods. There is a critical social understanding of science issue here. If novel methods are widely advertised simply because they are novel, without being shown to add value to existing knowledge, how are people supposed to interpret flashy new scientific results?

It may appear odd that researchers who focus on the invention and application of nonlinear analysis techniques should argue against their use but the implications are much broader. The social question is 'How are doctors supposed to tell the difference?' At a time when decisions about whether to vaccinate children or to convict a criminal, based on statistical evidence, are increasingly linking back to science, it is ever more crucial that professionals and the public be given enough information to keep nonlinearity from joining the list of 'lies, damned lies and statistics'.

This emphasis on the social impact and philosophical basis is, I believe, one of the main things which distinguishes CATS research at LSE. We aim to stress the importance of establishing that new methods offer a real advance – the kind of benchmark used regularly in CATS to judge new methods. That means, of course, that we often end up shooting down our own ideas. But ultimately the use of any complicated statistical method – whether in medical diagnostics, for weather forecasting, or in a range of other economic applications – can be justified only by showing with real data that it outperforms well understood traditional statistics or provides complementary, new value added information. ■

Dr Leonard Smith
is director of CATS, a reader in statistics and senior research fellow at LSE, and a senior research fellow (mathematics) of Pembroke College, Oxford.

CATS is supported by UK research councils, in particular the Engineering and Physical Sciences Research Council, and Department of Trade and Industry Faraday Partners including London Electricity and Risk Management Solutions, and the US Department of Commerce.

CATS works jointly with the Smith Institute, which aims to enhance the academic/industry interface of mathematics and statistics in the UK. Staff have worked on a number of consultancies including research for Shell, the National Grid Transco, the Passport Agency, and the German Weather Service.

For more information, see www.lse.ac.uk/collections/cats

For weather forecasts and climate modelling, see www.dime.lse.ac.uk and www.climateprediction.net