Small-number statistics, Common Sense, and Profit: Challenges and Non-challenges for Hurricane Forecasting

Alex Jarman and Leonard A. Smith

Centre for the Analysis of Time Series, London School of Economics
Email: *a.s.jarman@lse.ac.uk*

Abstract

When making only one forecast per year, or per decade, it can take some time to establish statistical confidence in the skill of a given forecast scheme. Must a risk tolerant decision maker wait decades until skill is "proven" if that decision maker believes the system to have value? What of a risk neutral decision maker? A methodology is illustrated to demonstrate there are imperfect forecast systems which almost certainly have nontrivial value long before one might establish that their skill was statistically significant.

1 Identifying skill with small datasets

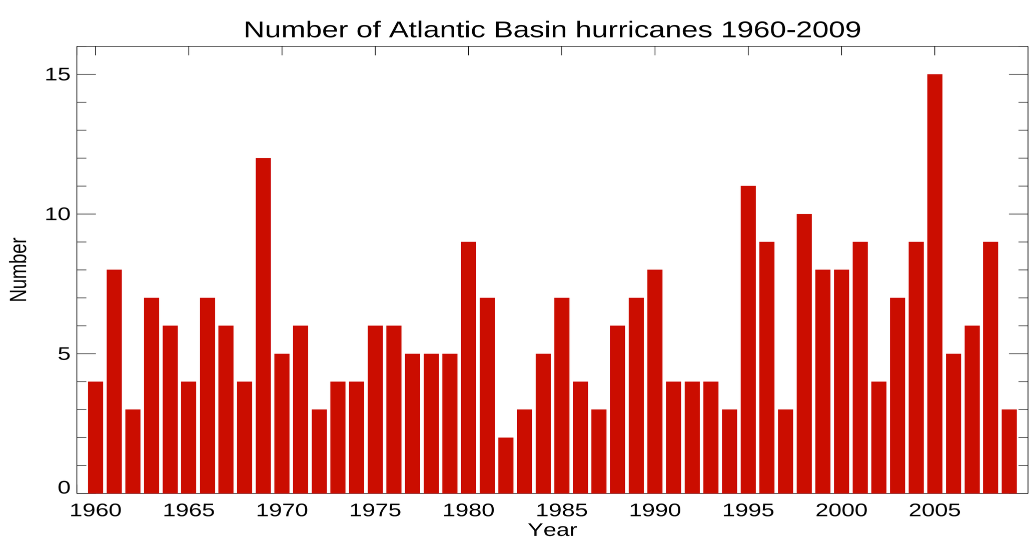

Figure 1: HURDAT data: Number of Atlantic Basin hurricanes from 1960 to 2008 [3])

The forecasting of US hurricanes has become a high profile endeavour over recent years largely due to its potential applications for the insurance industry, and the search for climate change signals in hurricane activity data. Establishing out-of-sample skill in an annual hurricane forecasting system poses a challenge on decadal timescales due to the slow rate new information is gathered with which to verify forecasts. The range of uncertainty in a sample of forecast model evaluations increases with decreasing time duration, and thus genuine skill cannot be reliably ascribed to a forecast model which is verified with a small dataset.

It has been argued that to robustly assess the predictive skill of a hurricane forecasting system, it would need to sustain an accurate enough performance over at least a period of several decades [4]. Otherwise, there is no way of knowing whether any skilful predictions made by the model are attributable to the quality of the model or to just chance alone. This raises two interesting questions: 1) would it truly take several decades to establish skill in practice? And 2) even if so, should the lack of established skill deter a decision-maker from using a forecast they believe to be valuable? The second question is investigated in this poster, and we argue that the answer to this question is "no".

2 Does a decision-maker need to wait?

If demonstrating genuine skill with limited datasets is not possible, then should a decision-maker wait for proof of skill in a model before using it? Might they be forgoing the opportunity to benefit from forecast information whilst seeking statistical reassurance? We examine the cost of waiting. If the decision-maker believes in the skill of their model, they might rationally choose to begin implementing it and will begin to receive value before those who choose to delay. The chance to profit before proving can be conceptualised in the context of what is called the "Swindled Statistician Scam" (see Box 1).

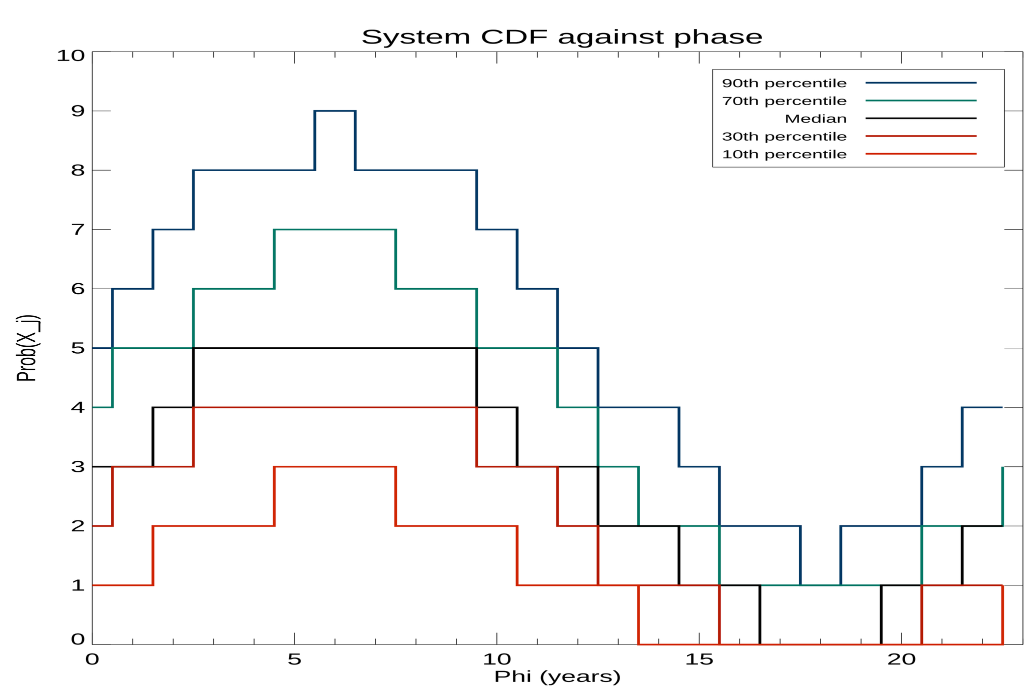

Figure 2: Contours of the cumulative distribution function showing how the system probability distribution changes as a function of the phase of the 24 year cycle. Note that occurrences of higher hurricane numbers are more likely in years with phase 5, 6 or 7 than those with phase 17, 18 or 19.

3 Profit or Proof: how should a decision-maker decide?

Consider a toy hurricane system in which the mean number of storms follows a 24 year cycle, while the number of storms in any given year is determined at random. To illustrate that structural model error does not preclude value to a decision-maker, consider an imperfect model of that system with the same cycle period, but where the probability distribution function used is incorrect in shape, not merely in parameter. This model will then be used in games of Hurricane Roulette, where the imperfect (but time dependent) model probabilities are used to place bets against odds set by a house using the correct (but not time dependent) climatological probability distribution. The results can be reported in either bits of information or as an expected annual return (see [1]).

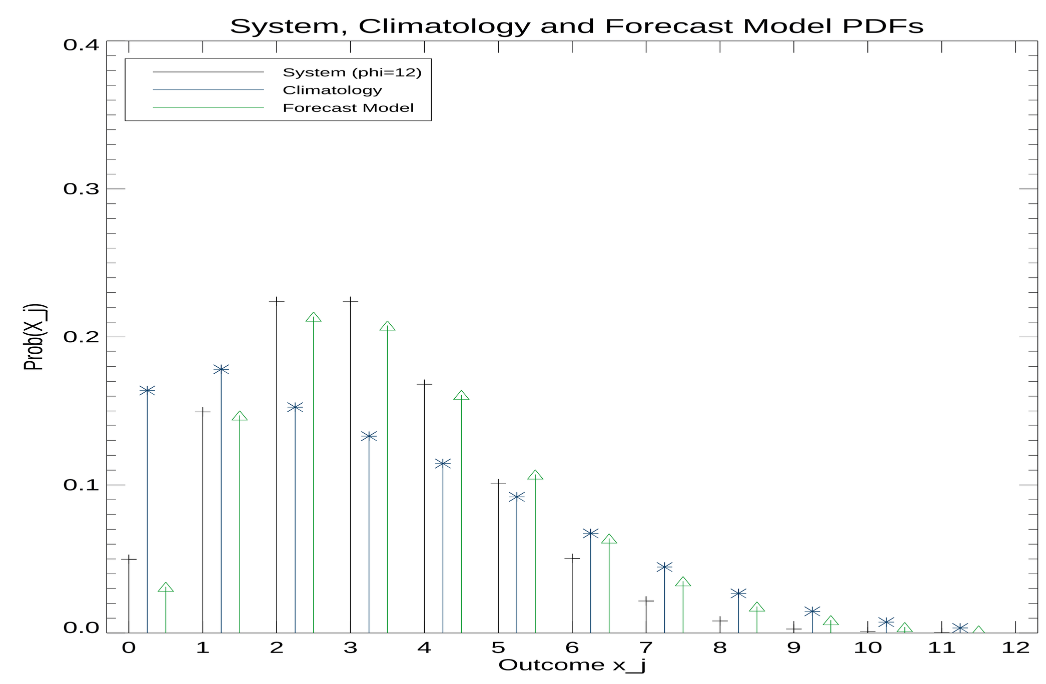

Figure 3: This figure shows the probability distributions for the system (black), and an imperfect model (green) for phase year 12 of the 24 year cycle. The climatological distribution (computed over all values of phase) is also shown in blue. Note that one need not know the true system PDF in order to outperform climatology. In this case the imperfect model PDF is expected to have a better ignorance score than the climatology; 0.18 bits on average (corresponding to a 1.13% per year interest rate).

Box 1. The Swindled Statistician Scam: A wily underwriter approaches a non-Floridian statistician with a business deal: the statistician will produce a probability forecast of the number of destructive events in the coming year, the underwriter will use her market contacts to bet on the forecast. As soon as the statistician can prove the forecast really does have skill, the underwriter will pay royalties. Will this leave the statistician swindled out of a small fortune?

The system is defined as a Poisson process, $X \sim Pois(\lambda(t))$, where X is the number of hurricanes in a given year and has a sinusoidal time-dependent mean determined by the equation $\lambda(t) = 2.5sin(2\pi t/T + \phi) + 3.0$. The probability distribution of the system over each of the 24 phases is illustrated in Figs 2 and 3 which describe the cumulative distribution function for all phases, and the probability density function corresponding to phase year 12 respectively.

The forecast model is defined by a squared Gaussian distribution with a mean equal to the climatological hurricane number average i.e. if $V \sim \mathcal{N}(\mu_{clim}, \sigma^2)$ then the random variable $Y = \lfloor V^2 + 0.5 \rfloor$ represents the distribution of annual forecast hurricane numbers. In addition, the model parameter σ has been fitted to each phase of the 24-year system cycle by minimising the expected ignorance of the forecast.

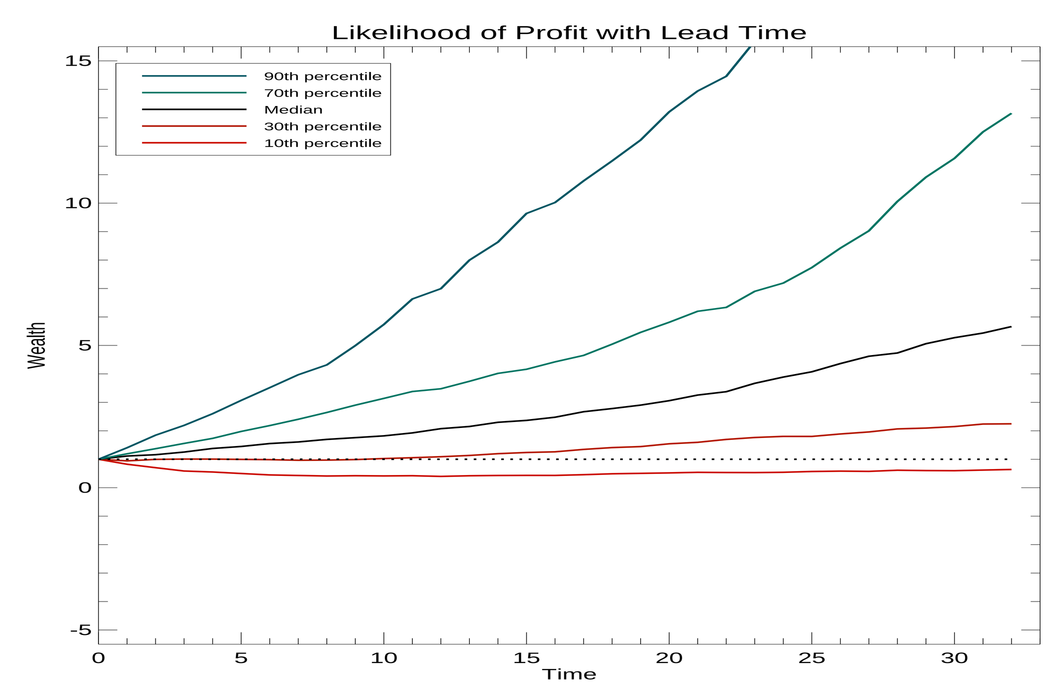

Figure 4: Frequency distribution of wealth of 2048 decision-makers who bet using the imperfect hurricane forecast model, as a function of the duration of time over which they bet. The dotted line shows their initial wealth (set equal to one unit), so the fraction of those above this line reflects the likelihood of being in profit by using the model. Note that the 30th percentile (brown) crosses and stays clearly in profit after about year 12: this is long before one could establish that the model had statistically significant skill.

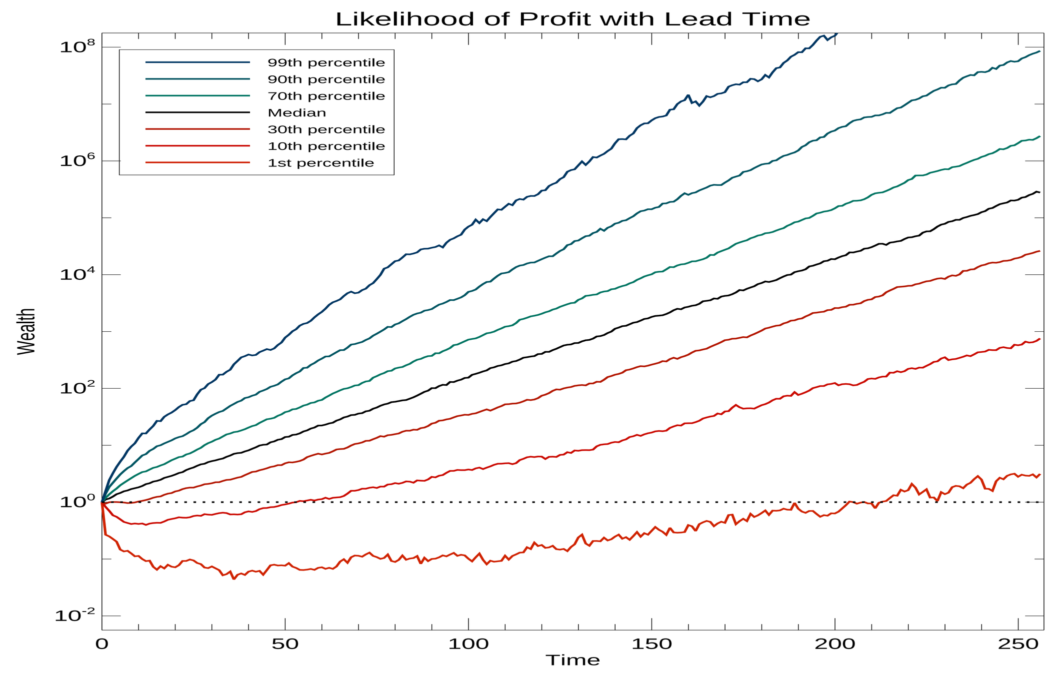

Figure 5: Frequency distribution of wealth as in Fig. 4, but shown over a longer time period and on logarithmic scales. Note that on the time scales on which one could establish statistically significant skill by time series analysis given a model with a period of 24 years, most of the decision-makers have made a nontrivial profit.

Hurricane Roulette proceeds as follows: at the start of each annual hurricane season a punter, or in this case the decision-maker, is offered odds defined by the climatology PDF (equally-weighted sum of the 24 system phase PDFs). She then places her bet by distributing all of her current wealth (based on the Kelly betting strategy [2]) according to the forecast probabilities assigned to each possible hurricane number outcome (X_j). The actual outcome determines the pay-off on each annual bet.

The results of a sample of 2048 realisations (or worlds) of this game of Hurricane Roulette demonstrate that the decision-maker would be very likely to have made a non-trivial profit even before two system cycles have completed - which is much shorter than the time it would take to reach a sufficient level of statistical significance (NB: the phase, ϕ, is selected at random for each realisation to avoid bias). This is evident in Fig 4 which shows the frequency distribution of wealth of 2048 betting decision-makers.

If the game of Hurricane Roulette is extended out to 256 years, the likelihood of profit significantly increases and we find that 99% of decision-makers will have made a non-trivial profit before 250 years have passed. This is shown in Figs 5 and 6. Even though this is an idealised situation, it does demonstrate why a decision-maker might consider and potentially benefit from the use of a forecasting system they believe to be skilful. The results above illustrate a case in which the fleeced statistician would be kicking himself for a long time before the first royalty payment came in.

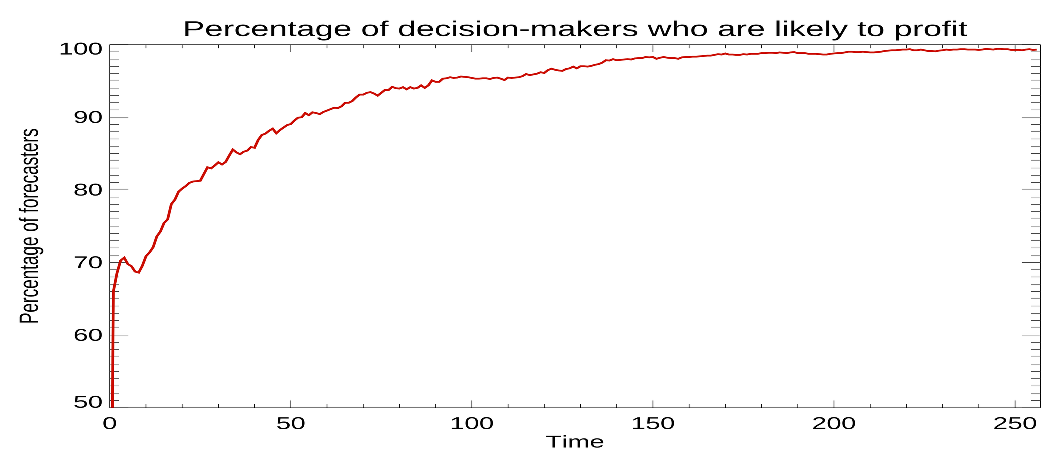

Figure 6: Percentage of decision-makers expected to make a profit with time when betting against climatology using the imperfect model in a game of hurricane roulette; the sample size is 2048 worlds.

Box 2. Discussion Points

- Establishing Forecast Skill on Systems with long time scales poses a challenge
- A decision-maker may take accept risk in order to gain profit, rather than first wait to fist establish statistical confidence in the forecast. In such cases, any difficulty in establishing forecast skill is a non-challenge
- Ways of benefiting from an imperfect model can be demonstrated through the use of games like Hurricane Roulette

References

[1] Renate Hagedorn and Leonard A. Smith. Communicating the value of probabilistic forecasts with weather roulette. *Meteorological Applications*, 16(2):143–155, October 2009.

[2] J. L. Kelly. A new interpretation of information rate. *Bell Systems Tech.*, 35:917–926, 1956.

[3] National Oceanic and Atmospheric Administration. Atlantic hurricane database re-analysis project. source: http://www.aoml.noaa.gov/hrd/datasub/re\verbanal.html.

[4] R. A. Pielke. United states hurricane landfalls and damages: Can one- to five-year predictions beat climatology? *Environmental Hazards*, 8:187–200(14), September 2009.

Acknowledgements: The support of EQUIP, of the Economic and Social Research Council (ESRC), and of Munich Re is gratefully acknowledged.