Climate Generators: Benchmarks for THORPEX Forecasts Beyond Week One

Anna Andrianova[1] and Leonard A. Smith[1,2]

[1] Centre for the Analysis of Time Series, Dept. of Statistics, London School of Economics, U.K.
[2] Oxford Centre for Industrial and Applied Mathematics, Oxford University, U.K.

www.lsecats.org

INTRODUCTION

A 'climate generator' is presented, which is based upon the Random Analog Prediction [1] approach. ERAP generates distributions of arbitrary size allowing enhanced sampling statistics 'consistent' with historical data. The aim is to generate background (unconditioned) distributions. ERAP provides a useful 'zero-skill' benchmark for more complex simulation-model-based forecasts, and has application in other areas, such as pricing weather derivatives. Application to a controlled experiment and European temperature are discussed.

ERAP: the algorithm

The ERAP approach searches for patterns in the past data close to the most recent data, and uses these to generating trajectories. Close patterns are found by searching for the nearest neighbours of the learning set after it has been arranged into a delay space and filtered using Singular Spectrum Analysis.

Sets of nearest neighbours identified both for short time scales and for long time scales. This is done by using two delay spaces of different dimensions (say 10 and 90). The set of nearest neighbours for the long time scale are used to restrict the choice of the short term nearest neighbours.

An ensemble is formed by picking one neighbour at random from the set of short term nearest neighbours, where each neighbour has a probability related to its distance from the present state. The trajectory is extended using the 'future' (or first differences) associated with the chosen neighbour. The length of the extension varies randomly to avoid artificial cycles.

Controlled Experiment: ERAP ensemble for the synthetic weather data

In order to conduct a controlled experiment of the ensemble produced by ERAP a non-linear process has been chosen to generate synthetic weather data that roughly resembles the daily temperature data. The ERAP approach was then applied to a learning set consisting of 100 years of synthetic weather data. Which one is the real data?

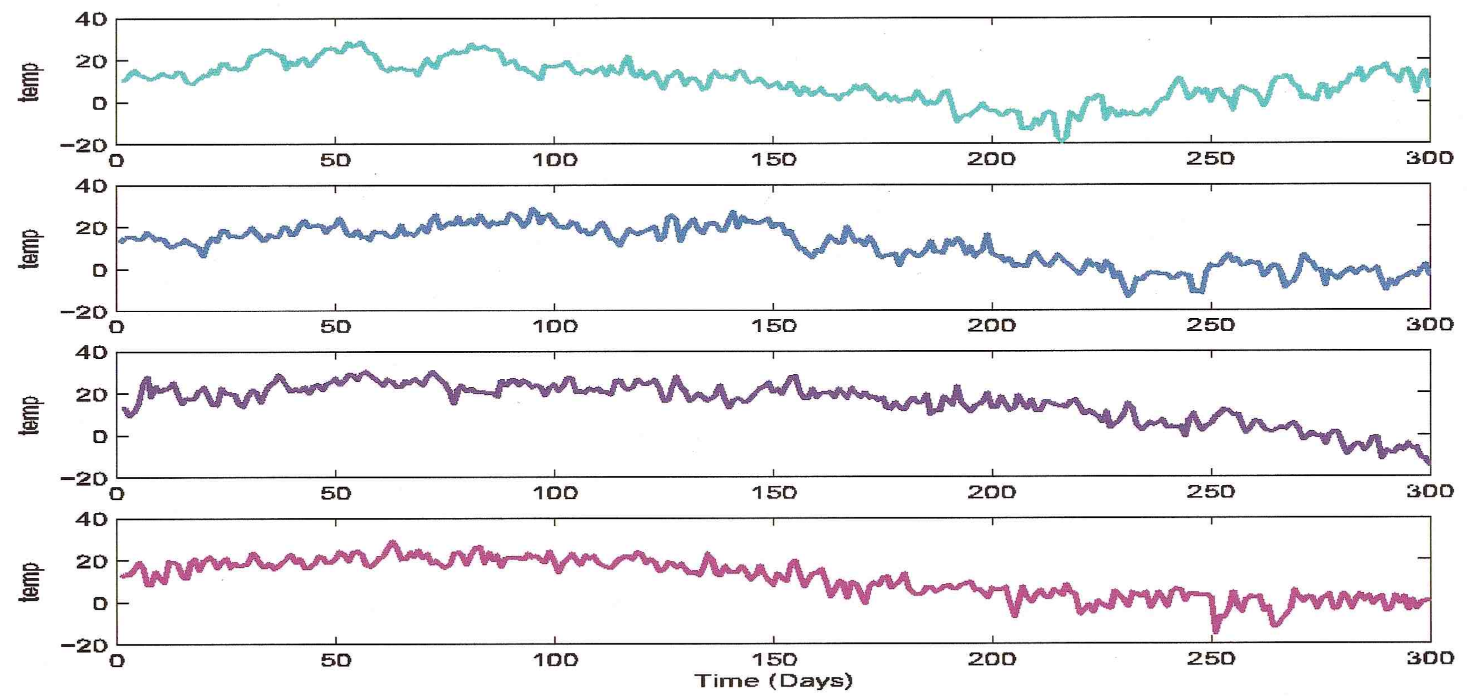

FIG. 1: Three trajectories from the ERAP based on the 100 years of the synthetic weather data and the raw data

Which is which?

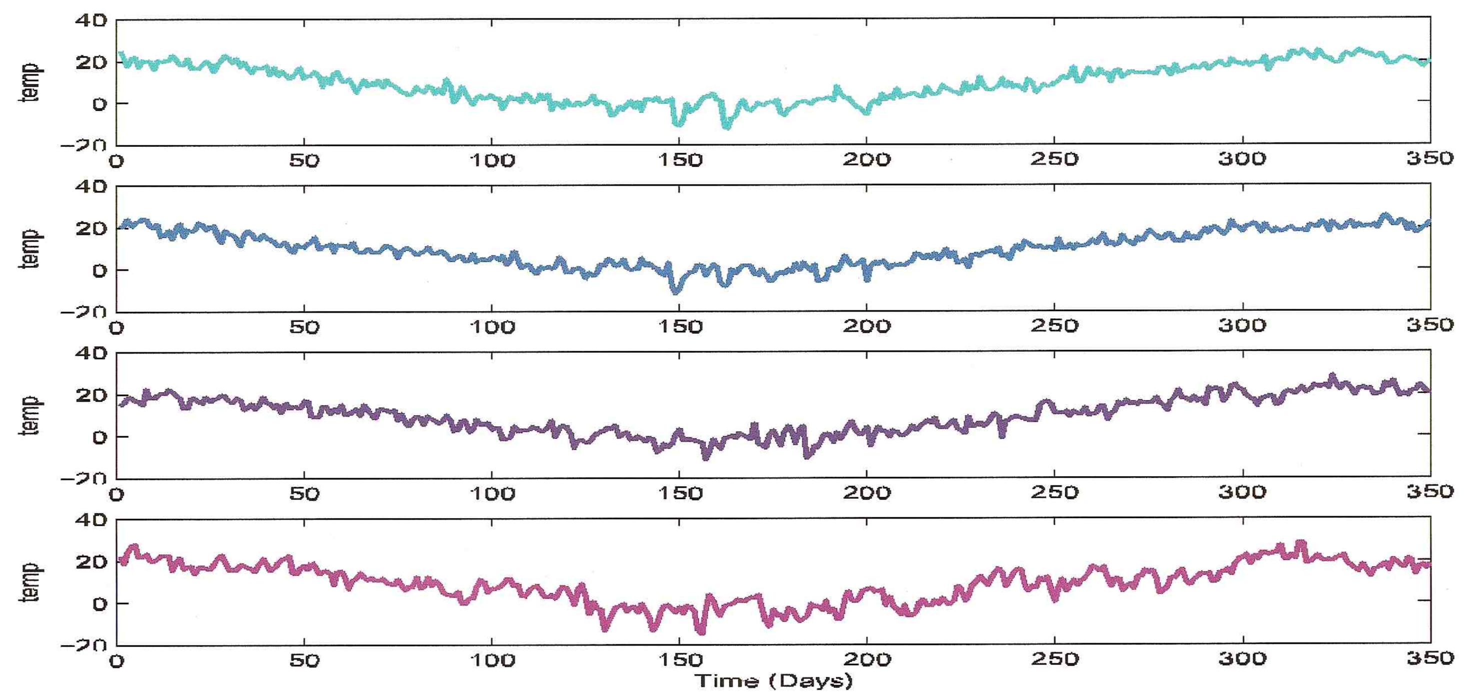

FIG. 2: One trajectory from the ERAP based on the 100 years of the synthetic weather data and one trajectory of the raw data

ERAP ensemble for the Berlin daily temperature data

Next, the ERAP approach was also applied to the Berlin daily maximum temperature data. The size of the learning set used was 50 years.

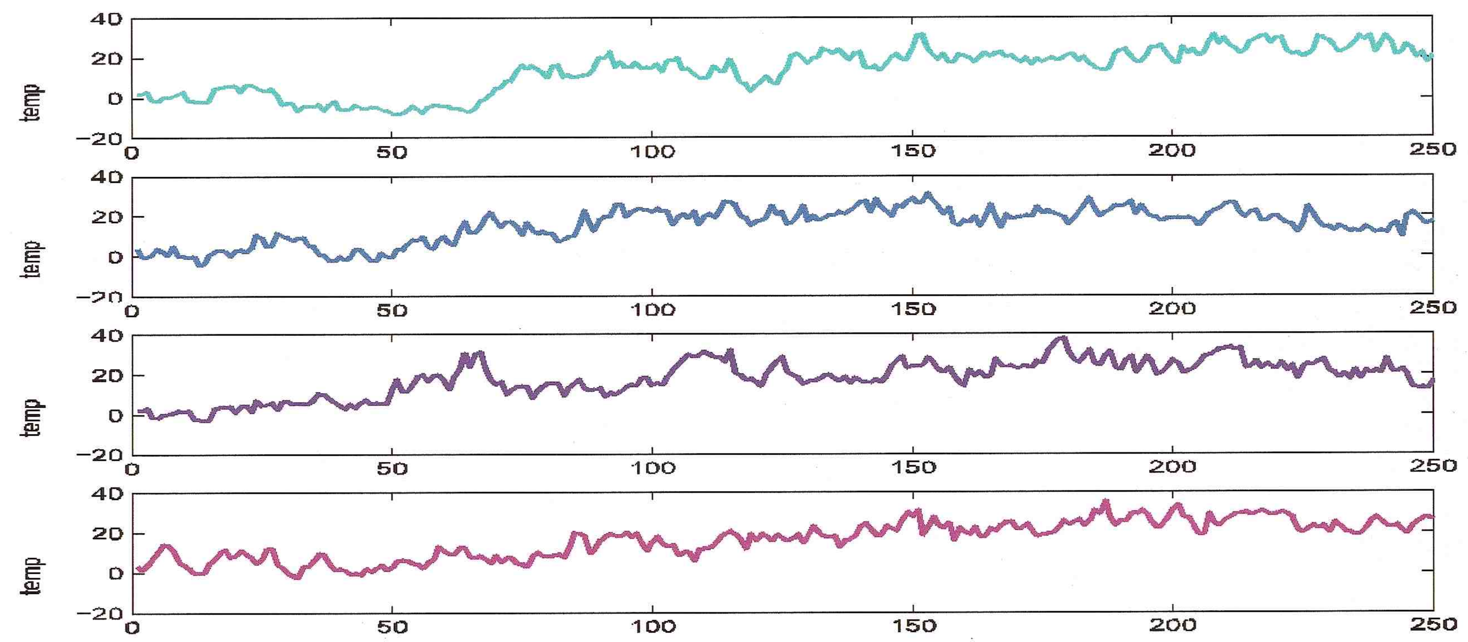

FIG. 3:Three one year trajectories from the ERAP based on the 50 years of the Berlin and the raw data

Future work will illustrate how the ERAP approach compares with traditional methods (like matching some autocorrelation function). The synthetic weather process allows us to test different approaches where 'the answer' is known. In practice, of cause, one is always extrapolating.

References

[1] Random Analog Prediction, Paparella et al Phys.Lett.A, 235(3 233-240), 1997

[2] Fermi Notes, can be found on: www.lsecats.org. The Maintenance of Uncertainty, Leonard A. Smith, 2003.

a.andrianova@lse.ac.uk