Validation and ground truth in stochastic forecasts of dynamical systems

Valerie Livina[1]

[1]National Physical Laboratory, Teddington, United Kingdom

We apply stochastic modelling and forecasts for dynamical systems, in particular, those that experience critical transitions (nonstationary behaviour). In validating such forecasts, it is crucial to have ground truth of the system trajectory.

Verification procedures assess how complicated the model linking between the measured quantities and the quantity of interest. This assumes checking the model output for cases where the solution is known exactly: the mathematical specification of the calculations, the source code if available, the qualitative behaviour of results (e.g., linearity, monotonicity, scaling exponents, fractality, etc.).

The larger the number of observed variables, the more detailed the underlying model, but there is a certain practical limit on the number of measurements. Using a fully resolved physics-based model with thousands of degrees of freedom for such a purpose is unfeasible because of complexity and limited resources. It is necessary to note that less complex models may still perform with sufficient accuracy. Therefore, the purpose of model validation is not to identify which model is most sophisticated but to assess how accurate each model is and how close to a known ground truth, no matter the approach they are using.

In validation, we compare calculations with reality, which could be defined by a different – trusted - method that is being used to measure the same object, or by an artefact with well-characterised properties. A controlled artificial dataset with induced properties can be used for validation of a model or testing technique, and in this case, it is possible to obtain sufficient statistics of performance, when validating various detection techniques.

Propagation of uncertainty, whether associated with the measured quantities or other parameters within the model, informs the decision on whether the data (measured quantities and internal parameters) are good enough. A validation that does not take uncertainty into account is not reliable since the difference between two values is only meaningful when compared with the variation of those values.

We propose the approach based on hindcast and artificial datasets for validation of forecasts of dynamical systems. Applications include paleo and modern climate datasets, artificial data simulations and sensor data from environmentally affected installations, in which we apply techniques for uncertainty quantification and detection of statistically significant transitions.

References:

[1] Livina & Lenton, GRL 2007;
[2] Livina et al, CoP 2010;
[3] Livina et al, ClimDyn 2011;
[4] Livina et al, Physica A 2013;
[5] Livina et al, Chaos 2015;
[6] Livina et al, NPG 2018;
[7] Livina et al, JET 2020;
[8] Billuroglu & Livina, JFAP 2022;
[9] Livina, Nature Climate Change 2023;
[10] Livina et al, JIMDS 2025;
[11] Livina, Chaos, in press.