Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Computational Physics 206 (2005) 334–362
JOURNAL OF COMPUTATIONAL PHYSICS
www.elsevier.com/locate/jcp

Quantifying predictability through information theory: small sample estimation in a non-Gaussian framework

Kyle Haven, Andrew Majda, Rafail Abramov *

Department of Mathematics and Center for Atmosphere/Ocean Science, Courant Institute of Mathematical Sciences, 251 Mercer Street, New York University, New York, NY 10012, USA

Received 10 September 2004; received in revised form 10 December 2004; accepted 15 December 2004
Available online 25 January 2005

Abstract

Many situations in complex systems require quantitative estimates of the lack of information in one probability distribution relative to another. In short term climate and weather prediction, examples of these issues might involve the lack of information in the historical climate record compared with an ensemble prediction, or the lack of information in a particular Gaussian ensemble prediction strategy involving the first and second moments compared with the non-Gaussian ensemble itself. The relative entropy is a natural way to quantify the predictive utility in this information, and recently a systematic computationally feasible hierarchical framework has been developed. In practical systems with many degrees of freedom, computational overhead limits ensemble predictions to relatively small sample sizes. Here the notion of predictive utility, in a relative entropy framework, is extended to small random samples by the definition of a sample utility, a measure of the unlikeliness that a random sample was produced by a given prediction strategy. The sample utility is the minimum predictability, with a statistical level of confidence, which is implied by the data. Two practical algorithms for measuring such a sample utility are developed here. The first technique is based on the statistical method of null-hypothesis testing, while the second is based upon a central limit theorem for the relative entropy of moment-based probability densities. These techniques are tested on known probability densities with parameterized bimodality and skewness, and then applied to the Lorenz '96 model, a recently developed "toy" climate model with chaotic dynamics mimicking the atmosphere. The results show a detection of non-Gaussian tendencies of prediction densities at small ensemble sizes with between 50 and 100 members, with a 95% confidence level.

* Corresponding author.
E-mail address: abramov@cims.nyu.edu (R. Abramov).

0021-9991/$ - see front matter
doi:10.1016/j.jcp.2004.12.008

1. Introduction

Complex systems with many spatial degrees of freedom arise in environmental science in diverse contexts such as atmosphere/ocean general circulation models (GCMs) for climate or weather prediction, pollution models, and models for the spread of hazardous biological, chemical, or nuclear plumes, as well as biological molecular dynamics, complex microfluids, etc. These nonlinear models are intrinsically chaotic over many time scales with sensitive dependence on initial conditions. Given both the uncertainty in a deterministic initial condition as well as the intrinsic chaos in solutions of such systems, it is natural instead to consider an ensemble of initial data representing uncertainty in measurements and characterized by a probability density. Monitoring the propagation of such a forecast ensemble in time gives one the potential to quantify the uncertainty and measure the confidence interval and average predictive power of a single deterministic solution, whose initial condition is randomly drawn from the initial spread. Apparently, small ensemble spread at a certain time is a strong evidence of dynamics with good predictive utility, and large spread denotes otherwise. However, there are at least two major problems with ensemble simulations which are often encountered in large complex systems. The first problem arises from the ensemble prediction strategy itself: even though a qualitative estimate of ''small'' and ''large'' ensemble spreads might be enough to give some basic insight into the nature of predictability, how one can quantify the predictive utility of a forecast ensemble in a rigorous manner? The conventional way is to measure the mean and variance of an ensemble, which is equivalent to approximating the internal structure of an ensemble by a Gaussian probability density. Central issues of practical importance in an ensemble prediction such as bimodality or skewness in a forecast ensemble require a general non-Gaussian description of predictive utility. The second problem becomes important for complex systems: certainly, a larger ensemble size means better quality of a prediction. However, usually ensembles of very limited size are affordable in complex systems for making real-time forecasts, largely due to enormous consumption of computational power. Thus, the natural question arises – whether or not one can trust the information provided by a forecast ensemble with relatively small size? In other words, how one can quantify the credibility of a forecast ensemble depending on its sample size? The current work systematically addresses these two problems within the framework of information theory and rigorous predictability estimates via relative entropy.

The applicability of information theory for weather or climate prediction has been studied previously by Carnevale and Holloway [1], Schneider and Griffies [2], Roulston and Smith [3], Leung and North [4]. Recently, Kleeman [5] has suggested the relative entropy as an estimate of predictive utility in an ensemble forecast relative to the climatological record, as well as a signal-dispersion decomposition. The Gaussian framework of relative entropy and signal-dispersion decomposition has been tested by Kleeman et al. [6] for a simple 100-mode truncated Burgers–Hopf model with chaotic behavior and well-understood spectrum and autocorrelation time scaling (for complete model description and climatology see Majda and Timofeyev [7,8], and Abramov et al. [9]). Majda et al. [10] developed a more sophisticated framework of predictability through relative entropy for non-Gaussian probability density functions, which includes a hierarchy of rigorous lower bounds on relative entropy through the statistical moments beyond the mean and covariance through maximum entropy optimization (Mead and Papanicolaou [11]). Abramov and Majda [12] converted the non-Gaussian predictability framework into a practical tool through the hierarchy of lower bounds and a rapid numerical optimization algorithm. Recently, Cai et al. [13] exhaustively tested several facets of the non-Gaussian information theoretic predictability framework in a simple chaotic mapping model with an explicit attractor ranging from Gaussian to fractal as parameters are varied. Kleeman and Majda [14] have quantified the loss of information in coarse-grained ensemble estimators and applied these ideas to geophysical turbulence. Different applications of relative entropy as a predictability tool were developed in Abramov and Majda [12]; besides a straightforward measure of lack of information in the climate relative to the prediction