



Attractor Learning In Interactive Ensembles

Wim Wiegerinck (1,2) and Lasko Basnarkov (3,4)

(1) Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands, (2) SNN, Nijmegen, The Netherlands, (3) Faculty of Computer Science and Engineering, SS. Cyril and Methodius University, Skopje, Macedonia, (4) Macedonian Academy of Sciences and Arts, Skopje, Macedonia

Recently methods for model fusion by dynamically combining model components in an interactive ensemble have been proposed. Although different in detail, these interactive ensembles can be generally considered as a supermodel, which have the different original models as fixed basis functions, and which is parameterized by the fusion parameters. In most of the proposals, the fusion parameters are optimized based on a short time scale prediction error. In general this will improve weather prediction skill, but not necessarily climate projection skill. Expressed in terms of nonlinear dynamical systems, reducing error on the level of vector fields does not necessarily lead to a better attractor.

We demonstrate this in a low dimensional dynamical system toy example. The example consists of three models. One model is the assumed ground truth. The other two are “imperfect models” of the ground truth. The ground truth is represented by a chaotically forced Lorenz 63 model. The chaotic forcing plays the role of unresolved scales and is assumed not directly observable. The two imperfect models, named model 1 and model 2, are both represented by a Lorenz 63 system with perturbed parameters and a constant forcing. The perturbations and forcings in model 1 and model 2 are such that the vector field of imperfect model 1 is closest to the true vector field. However the long term statistics of imperfect model 2 is closest to the true long term statistics.

The two models, model 1 and 2, are fused into a single supermodel. The fusion parameters are optimized on the basis of a finite data set of observables generated by the ground truth dynamics, the so-called training set. After optimization, the resulting supermodel skills are evaluated on the basis of a test set, which is a second, larger data set of observables generated by the ground truth dynamics. If, in the example, vector field error is used as optimization criterion, optimization indeed leads to an improved short term prediction skill on the test set. However it turns out to strongly degrade the prediction skill of the long term statistics, e.g. the mean and the variance of the supermodel attractor are very different from the test set mean and variance.

A notion of attractor (training/test) error is introduced by considering metrics between probability densities, one of which is estimated on the basis of the given (training/test) data set. The other density is estimated from the data generated by a long term (super) model simulation. With this notion we define attractor learning as the optimization of the attractor training error. Attractor learning is demonstrated in the example. Compared to vector field learning, attractor learning leads to a significantly reduced attractor test error and improved long term statistics of the supermodel, while the resulting vector field test error is hardly increased.