Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia Computer Science 7 (2011) 261–263

Procedia Computer Science

The European Future Technologies Conference and Exhibition 2011

# Supermodeling by Combining Imperfect Models

F.M. Selten[a], G.S. Duane[b,c], W. Wiegerinck[d], N. Keenlyside[e], J. Kurths[f], L. Kocarev[c,g]

[a] *Royal Netherlands Meteorological Institute, De Bilt, The Netherlands*
[b] *Department of Atmospheric and Oceanic Sciences, University of Colorado, Boulder, USA*
[c] *Macedonian Academy of Sciences and Arts, Skopje, Macedonia*
[d] *Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands*
[e] *Leibniz Institute of Marine Sciences at Kiel University, Kiel, Germany*
[f] *Potsdam Institute for Climate Impact Research, Potsdam, Germany*
[g] *BioCircuits Institute, University of California San Diego, La Jolla, CA, USA*

**Abstract**

SUMO (Supermodeling by combining imperfect models) is a three-year project funded under FET Open program with a starting date October, 1, 2010. We review some basic facts and findings of the SUMO project.

Selection and peer review under responsibility of FET11
*Keywords:* nonlinear dynamics; machine learning; climate science

Climate scientists are faced with the challenge to provide society with credible scenarios of future climate change. At a dozen or so institutes around the world, comprehensive climate models are being developed and improved. Over time, the capability of these models to simulate the observed climate of the last 100 years has improved due to increased spatial resolution and improved descriptions of unresolved processes, but large systematic errors remain and the models are still far from perfect. Nevertheless, these models are time-integrated into the future to simulate the response of the climate system to so-called scenarios of future anthropogenic emissions of greenhouse gasses. The outcomes of these models are archived at a central repository and form the basis of extensive studies on future climate change by climate scientists (http://www-pcmdi.llnl.gov/). Studies based on these outcomes are used by policy-makers to assess the impact of climate change on society.

A crucial question that arises is how the outcomes of these different models should be combined to get the most realistic estimate of the unknown true response of the climate system to the expected future emissions of greenhouse gases, etc. Large differences do indeed exist in model-based predictions of future climate, in regard to the overall extent and regional characteristics of warming. Common practice is that some form of a weighted average is applied to all model outcomes with greater weight given to those models that in some sense better reproduce the historically observed climate evolution.

We question that such a posteriori combination of the individual imperfect model outcomes uses the full potential of the model ensemble. Instead, we propose in SUMO a radically different computational strategy that combines ideas from the machine learning, dynamical systems and climate science community to improve our ability to simulate the observed, historical evolution of climate and obtain more realistic estimates of future climate change using the existing ensemble of state-of-the-art climate models. This approach is new and timely since the models are there, the need for

*E-mail address:* lkocarev@ucsd.edu (L. Kocarev).

1877-0509 
Selection and peer review under responsibility of FET11 doi:10.1016/j.procs.2011.09.025

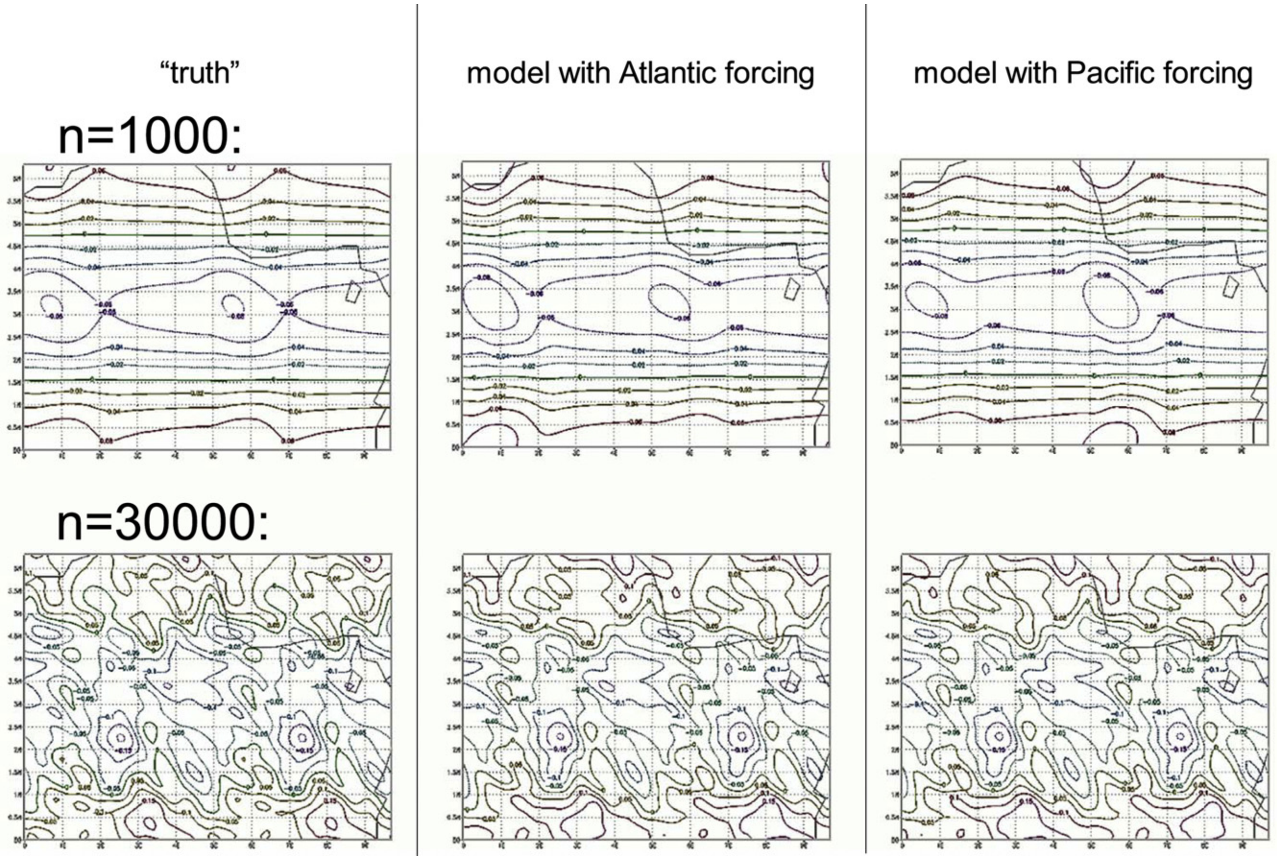


Figure 1. Two quasigeostrophic channel models [2], one with a better representation of the Atlantic jet stream and the other with a better representation of the Pacific jet stream, were fused to form a supermodel by adapting a connection coefficient so as to reproduce "truth" in a training set. After $n$ = 30000 time steps, the models were synchronized with each other and with truth. Contours represent streamlines of the flow, as in [2].

better-constrained predictions is urgent and most importantly the computational resources will hopefully soon be there through for instance advanced grid technologies to allow the simultaneous integration of an ensemble of state-of-the-art climate models that exchange information on a time-step basis.

Our targeted breakthrough is to provide a more realistic simulation of the observed historical evolution of climate and make actual predictions of our changing climate system with a climate "supermodel", consisting of an ensemble of interconnected state-of-the-art climate models, that has "learned" to reproduce the observed, historical evolution of the climate system. The supermodel will be superior to any of the individual models in the ensemble since it has learned to combine the strengths of the individual models. Information is exchanged between the models on a time-step basis in such a way that the observed evolution of the climate system is best reproduced in a suitably chosen measure. Our long-term vision is that the methods to be developed in our project will be applied to combine expert models of objective processes generally. Complex systems of partial differential equations (PDEs) or lattice maps can represent social, economic, biological, ecological, or physical processes.

It has been observed by climate researchers that the imperfections of their models are often complementary; for instance one atmospheric model might produce more realistic heat fluxes for the ocean, while another produces more realistic momentum fluxes. An improved simulation was obtained in which two atmospheric models were coupled to one oceanographic model, one provided the more realistic heat fluxes at every time step, the other the more realistic momentum fluxes. Importantly, it was not known a priori that one model produced a more realistic heat flux at the ocean surface and that the other model produced a more realistic momentum flux. Rather, those relative advantages became apparent only in the context of the behavior of the entire model with exchanged variables. Running models in parallel and allowing a dynamic exchange of information appears a feasible avenue to increase the simulation performance by combining the strength of the individual models.

The learning of supermodels is being also studied in the context of small chaotic dynamical systems described by ODEs, like the well-known Lorenz'63 system. Imperfect models are created by perturbing the standard parameter values. These models are combined into one super-model, through the introduction of connections between the model equations. The connection coefficients are learned from data from the unperturbed model, that is regarded as the ground truth. The main result of this study [1] is that after learning the supermodel is a very good approximation to the truth, much better than each imperfect model separately, and better than a weighted combination of their outputs. Furthermore, the supermodel turned out to be robust against parameter disturbances, e.g. after doubling a certain parameter in the ground-truth model and the corresponding parameters in the supermodel, the supermodel remains still close to the new assumed ground truth, even without additional learning. These results suggest that the supermodeling approach is a promising strategy to improve climate simulations. In an initial test of our adaptive model fusion strategy, two simplified PDE models of the midlatitude atmospheric circulation, with complementary representations of the jet stream, were fused to form a single supermodel (Fig. 1).

## References

[1] L.A. van den Berge, F.M. Selten, W. Wiegerinck, G.S. Duane, A multi-model ensemble method that combines imperfect models through learning, Earth Syst. Dynam. Discuss. 1 (2010) 247–296, doi:10.5194/esdd-1-247-2010.

[2] G.S. Duane, J.J. Tribbia, Weak Atlantic-Pacific teleconnections as synchronized chaos, J. Atmos. Sci. 61 (2004) 2149–2168.