Report on activities and findings under DOE grant "Collaborative research: An Interactive Multi-Model for Consensus on Climate Change"
#DE-SC0005238 (lead report[1])

Reporting period: 5/12/2012-10/30/2015[2]

Funding period: 10/1/2010-9/30/2013

No-cost extension: 10/1/2013-9/30/2014

PI/Co-PI information:

Greg Duane, Univ. of Colorado, lead PI

Anastasios Tsonis, Univ. of Wisconsin, PI

Ljupco Kocarev, UCSD, PI

Joseph Tribbia, NCAR, PI

Ben Kirtman, U. Miami, unfunded co-PI

Frank Selten, KNMI, unfunded collaborator

Noel Keenlyside, University of Bergen, unfunded collaborator (added after proposal)

Project synopsis

This project has sought to develop a new scheme for forming consensus among alternative climate models, that give widely divergent projections as to the *details* of climate change, that is more intelligent than simply averaging the model outputs, or averaging with *ex post facto* weighting factors. The method under development effectively allows models to assimilate data from one another in run time with

[1] This report primarily describes U. Colorado's research, some done in conjunction with KNMI and with U. Bergen (unfunded collaborating institutions), and the status of the multi-model development at NCAR. The information provided about travel, presentations, etc. is for U. Colorado only.

[2] Activities and findings from the previous funding period that were included in the previous report are briefly summarized here.

weights that are chosen in an adaptive training phase using 20th century data, so that the models synchronize with one another as well as with reality.

Results and Primary Activities

I. Summary

The project takes a hierarchical approach. The supermodeling scheme was first studied exhaustively with simple systems of ordinary differential equations. Results were described in detail in the previous report. The principal findings were that 1) for highly non-linear systems, such as Lorenz-63, including systems which describe phenomena on very different (atmosphere/ocean) times scales, supermodeling is far superior to any form of output-averaging; 2) negative coefficients can be used to advantage in situations where all models err in the same way, but to different degrees; 3) an interesting variant of supermodeling, "weighted supermodeling", is the limiting case where inter-model nudging coefficients in the originally conceived "connected supermodel" become infinite, but with fixed ratios, corresponding to a direct combination of the tendencies that appear in corresponding equations for the alternative models; 4) noise is useful for avoiding local optima in training the inter-model coefficients in the supermodel.

The supermodeling scheme was then investigated with simple quasigeostrophic (QG) models. As described in the previous report, it was found that QG models on a sphere can be coupled most efficaciously by working in a basis which captures the most variance, rather than the most instability, a somewhat unexpected result that still deserves scrutiny in a broader context. Further studies (since the last report) with QG channel models addressed the central question of when supermodeling is superior to output averaging in situations where nonlinearites are less extreme than with the ODEs initially studied. It was found that for realistic variations in a parameter in the QG model, output averaging is sufficient to capture all but the most subtle quantitative and qualitative behavior. Supermodeling helps when qualitative differences between the models result from unrealistically large parameter differences, or when very detailed spatial structure of the modes of variability are of interest. Therefore, the scheme may still be useful in the case of full climate models with qualitatively different parametrization schemes.

A supermodel was constructed from the intermediate-complexity SPEEDO model, a primitive equation model with ocean and land. Versions defined by different parameter choices, in a realistic range, were connected and the coefficients trained. Some improvement was found as compared to output averaging.

The learning algorithm used thus far gives sub-optimal, but still useful results when the CO2 level and other parameters are varied. Spatial structure remains to be studied.

The first use of supermodeling with full climate models has been with variants of the ECHAM model that use different convection schemes. As yet the models are only connected at the ocean-atmosphere interface, where weighted combinations of fluxes from the two atmospheres are passed to a common ocean, and the weights adapted during a training period. The supermodel was surprisingly successful at avoiding unrealistic features such as the double-ITCZ (Intertropical Convergence Zone), a problem that arises in both of the two models run separately.

The supermodels constructed thus far have not identified dynamical regime shifts in future climate. Thus the planned connection with the work of Tsonis on the relationship between regime shifts and synchronization/de-synchronization among the major climate modes (see U. Wisconsin report) has not yet been made. However the network analysis of the climate system, in observations and models, that was done in conjunction with that study, shows that models differ strongly from one another and from observations in regard to the dynamical structure described by correlation networks [Steinhaeuser and Tsonis 2013], providing a further justification for supermodeling.

Toward a general software framework for supermodeling, three versions of CAM (the Community Atmosphere Model) at NCAR were configured for inter-model nudging using the DART (Data Assimilation Research Testbed) capability to stop and re-start models in synchrony. It was clearly established that the inter-model nudging adds almost no computational burden to the runs, but there appears to be a problem with the re-initialization software that is still being debugged.

Publications: Several papers were published on the basic idea of the interactive multi-model (supermodel) including demonstrations with low-order ODEs. The last of these, a semi-philosophical review paper on the relevance of synchronization generally, encountered considerable resistance but was finally published in *Entropy* [Duane 2015]. A paper on the ECHAM/COSMOS supermodel, containing the most promising results so far [Shen et al. 2015] is presently under review.

Details of the QG channel model study, the initial SPEEDO supermodel results, the ECHAM/COSMOS supermodel results, and the work with the CAM supermodel are presented in separate sections following.

II. Supermodeling vs. Output-Averaging with QG Channel Models

If nonlinearities are strong enough so as to cause bifurcations in the climate systems as GHGs increase, it can be argued that output averaging will be insufficient to capture the effects and that supermodeling would be beneficial. However, there is little evidence for bifurcations of this type in model studies. But even without bifurcations, simple nonlinearity can still make the supermodel superior to an average of model outputs. This is perhaps most easily seen in the case where diagnostic properties depend non-monotonically on system parameters. Suppose we have two models of the form:

$$d\mathbf{x}/dt = F(\mathbf{x}, p_1)$$

$$d\mathbf{x}/dt = F(\mathbf{x}, p_2) \qquad (1)$$

where F is linear in the parameter p, and consider some diagnostic P(p), e.g. mean temperature. Further suppose that $P(p_1) = P(p_2)$, but that for some intermediate value p_i, $p_1 < p_i < p_2$, $P(p_i) > P(p_1) = P(p_2)$. Then any weighted average of model outputs will only give the first value $P(p_1)$. A weighted supermodel, on the other hand, could readily reproduce the correct dynamics, that is $F(\mathbf{x}, p_i) = w_1 F(\mathbf{x}, p_1) + w_2 F(\mathbf{x}, p_2)$ for appropriately chosen weights w_1 and w_2, since F is linear in p. It is hypothesized that a connected supermodel could also give the correct result.

Consider specifically a quasigeostrophic model of a re-entrant channel on a β-plane given by:

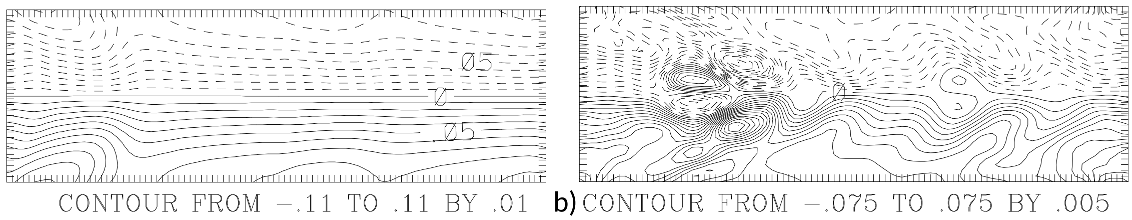
$$Dq_i/Dt \equiv \partial q_i / \partial t + J(\psi_i, q_i) = F_i + D_i \qquad (2)$$

where the layer i=1,2, ψ is streamfunction, and the Jacobian $J(\psi_i, q_i)$ gives the advective contribution to the Lagrangian derivative D/Dt (Vautard et al , 1988; Vautard and Legras , 1988). The forcing F is a relaxation term designed to induce a jet-like flow near the beginning of the channel:

$$F_i = \mu_0 (q_i - q_i) \qquad (3)$$

for q_i corresponding to a streamfunction ψ that defines a jet . The dissipation terms D, boundary conditions, and other parameter values are given in (Duane and Tribbia , 2004).

The QG channel model vacillates between two dynamical regimes corresponding to “blocked” and “zonal” flow, as illustrated in Fig. 1. The response of the blocking activity to the forcing parameter μ_0 in (2) provides a simple example of non-monotonic behavior. For zero forcing, blocking frequency is zero due to damping by the dissipative terms. For large forcing, the flow is consistently jet-like, and again there is no blocking. Typical flow fields for these two cases are shown in Fig. 2a,b. (The zero-forcing flow in Fig. 2a is turbulent, but of low amplitude, and includes no blocks.)

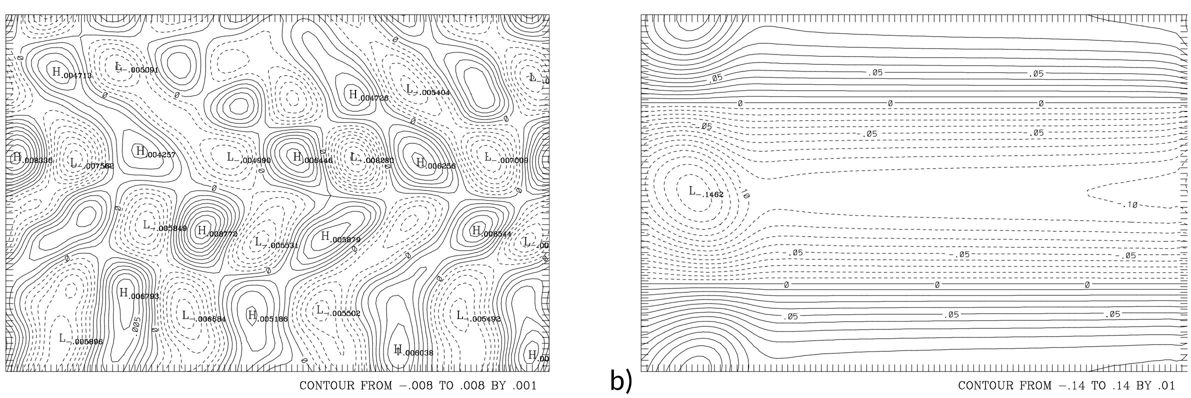


a) b)

Figure 1: Streamfunction (in dimensional units of $1.48{\times}10^9 m^2 s^{-1}$) describing a typical zonal flow state (a), and a typical blocked flow state (b) in the two-layer quasigeostrophic channel model. Parameter values are as in (Duane and Tribbia , 2004). An average streamfunction for the two vertical layers i = 1,2 is shown.

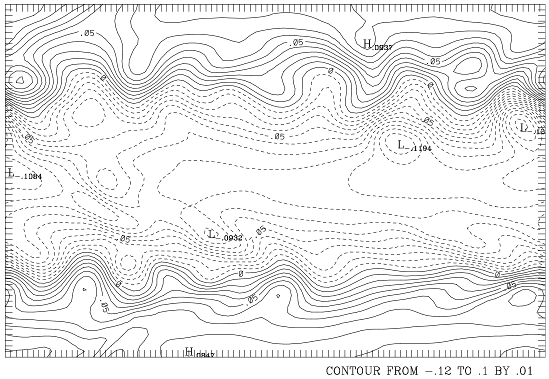
A weighted supermodel formed from the two individual models illustrated in Fig. 2 can reproduce the true dynamics exactly for any value of the forcing coefficient μ_0 between $\mu_0 = 0$ and $\mu_0 = 3$. For the typical value $\mu_0 = 0.3$ used previously, the behavior is as illustrated in Fig. 3. The supermodel flow spends much time in the blocked regime, unlike the flows in the individual models or any weighted average thereof. (If the actual flow fields of the individual models are averaged, instead of the blocking frequencies, the same conclusion is reached.)

The learning task for the weights is equivalent to that for determining the single parameter μ_0 directly. A previous algorithm for parameter learning in models that synchronize with identical parameters (Duane et al , 2007), for instance, is effective in the present context. While the argument applies exactly to a weighted supermodel, it seems likely that a connected supermodel could also be formed from the two individual models illustrated in Fig. 2 that would approximate the “true” behavior for arbitrary forcing coefficient.



a) b)

Figure 2: Typical flows in the QG channel model with very small forcing coefficient ($\mu_0 = 0$) (a), and very large forcing coefficient ($\mu_0 = 3.0$) (b). (The spatial domain in each panel includes two channels with flows in opposite directions).

a) 

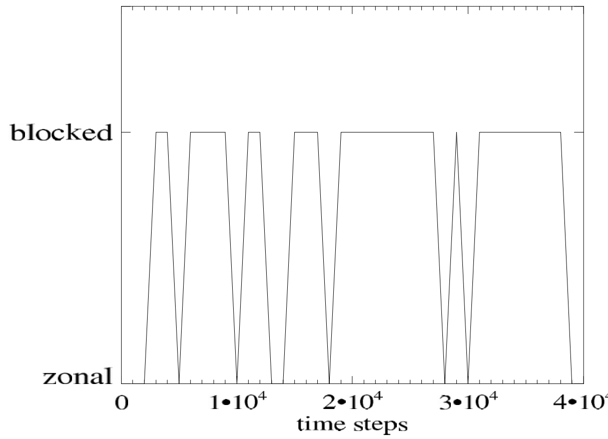
b) 

Figure 3: Typical flow in the QG channel model with a “realistic” forcing coefficient (μ_0 = 0.3) (a), and the history of vacillation of the flow in the bottom half of the domain between zonal and blocked regimes, sampled at low temporal resolution over the course of a simulation (b), using the blocking diagnostic defined in (Duane and Tribbia , 2004). The typical flow is also the exact solution to an appropriately weighted supermodel.

While a supermodel is clearly superior to an output average in the example given above, and in extreme cases generally, more linear behavior is expected for smaller inter-model differences as might occur in a realistic suite of models, such as the IPCC set. To construct a realistic experiment with toy models, a correspondence is needed between parameter differences among the toy models on the one hand, and differences among models or parameters used in actual climate projection on the other. Considering the forcing coefficient μ_0 in the QG models as representative of forcing generally, the question is about the relative magnitude of differences in forcing among the different models. External forcing in the different models is about the same, but the effective forcing, when differences in internal dynamics are included, varies significantly. The differences are manifest as differences in climate sensitivity - mean temperature change for given increase in greenhouse gas levels. Sensitivities of the climate models in the IPCC suite were determined from IPCC data [IPCC, 2007]. They were seen to vary by about ±1/3 of the average value. We take the QG forcing coefficient as analogous to these sensitivities, and introduce differences between the coefficients in the different models of the same relative magnitude. So we use models with μ_0 = 0.2 and μ_0 = 0.4 to form a supermodel.

Typical flow fields for the three values of the forcing coefficient μ_0 = 0.2, 0.3, 0.4 are shown in Fig. 4, along with blocked/zonal vacillation behavior. Unlike the case discussed above, it appears that if the two individual models err in their forcing coefficients only to a degree that seems realistic, a weighted average of their blocking frequencies could reproduce the “true” behavior. At least in regard to blocking frequency, the advantage of supermodeling is lost in this less extreme case.

If one pays more attention to the detailed modes of variability, a subtle advantage remains. It is known that there is a very weak anticorrelation between blocking activity in the Atlantic and in the Pacific [Duane and Tribbia ‘04]. That effect could not possibly occur in an output-average of models with Atlantic and Pacific forcing separately. It is thought that supermodeling will give improved predictions of other global multi-variable patterns of variability, where the relationships are stronger, as well.

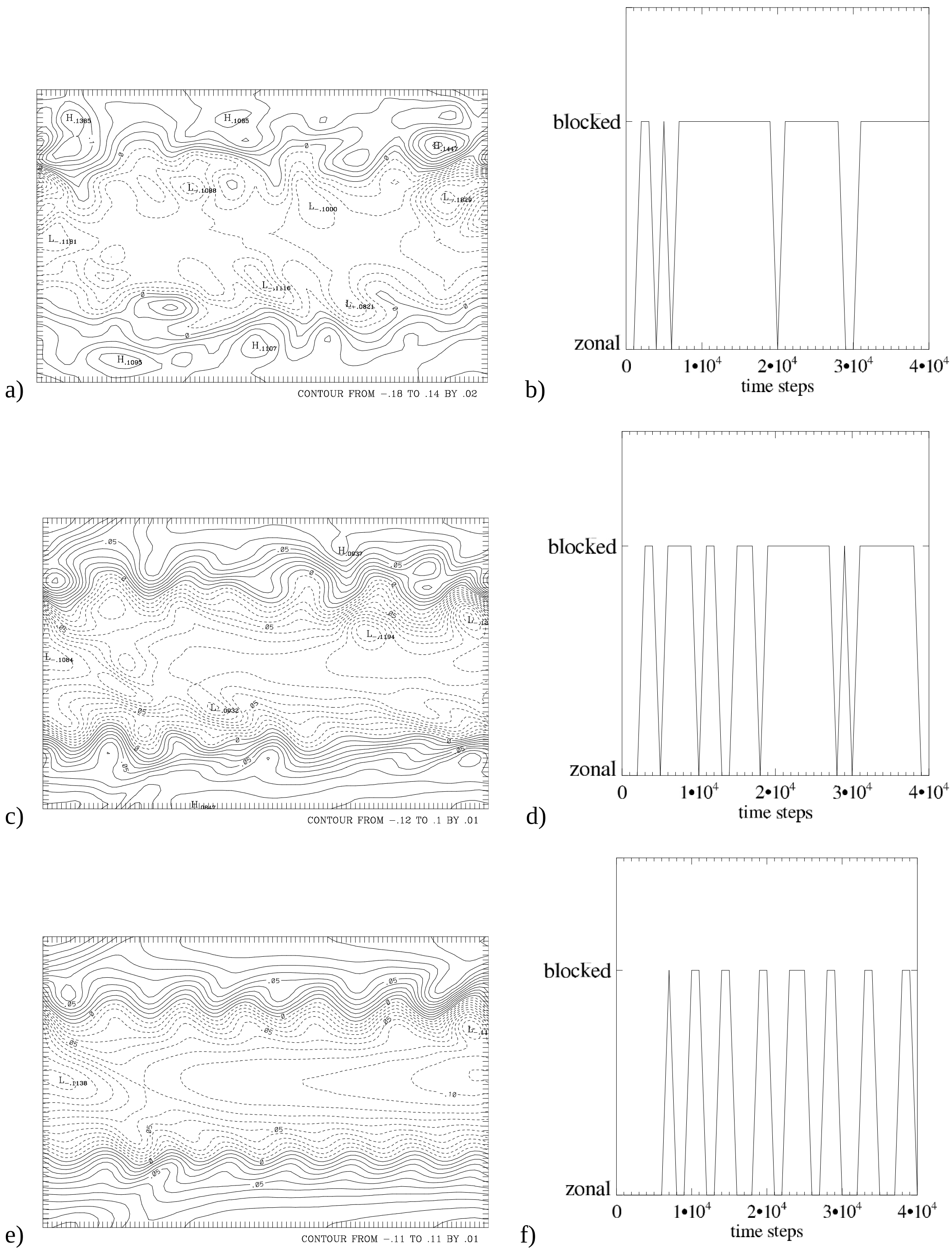


Figure 4: Typical flows in the QG channel model for realistically low forcing μ_0 = 0.2 (a), the "true" value μ_0 = 0.3 (c), and realistically high forcing μ_0 = 0.4 (e), with corresponding blocked/zonal vacillation histories (b),(d),(f). The "true" case is well approximated by an

average of the other two.

III. The SPEEDO Supermodel

To provide input for the state-of-the-art climate models with regard to connection strategies, we used the climate model of intermediate complexity SPEEDO [Severijns and Hazeleger , 2009]. The atmospheric component is the SPEEDY model that solves the primitive equations on a sphere using a spectral method. The spectral expansion is truncated at total wavenumber 30 which corresponds to a spatial resolution at the equator of about 500 km. It has 8 vertical levels and simple parameterizations for radiation, convection, clouds and precipitation The solar radiation follows the seasonal cycle but the diurnal cycle is not imposed. Instead daily mean solar radiation fluxes are prescribed. The total number of degrees of freedom is 38025: 31680 for the spectral coefficients of divergence, vorticity, temperature, specific humidity and log of surface pressure plus 6345 to describe the land temperature, land moisture and snow cover in the 2115 land points. The land component uses a simple bucket model to close the hydrological cycle over land and a heat budget equation that controls the land temperatures.

The ocean component is the CLIO model [Goosse and Fichefet, 1999]. The CLIO model is a primitive-equation, free-surface ocean general circulation model coupled to a thermodynamic-dynamic sea-ice model. The ocean component includes a relatively sophisticated parameterization of vertical mixing. A three-layer sea-ice model, which takes into account sensible and latent heat storage in the snow-ice system, simulates the changes of snow and ice thickness in response to surface and bottom heat fluxes. In the computation of ice-dynamics, sea ice is considered to behave asa viscous-plastic continuum. The horizontal resolution of CLIO is 3 degrees in latitude and longitude and there are 20 unevenly spaced vertical layers in the ocean. The CLIO model has a rotated grid over the North Atlantic ocean in order to circumvent the singularity at the pole. The total number of degrees of freedom is on the order of 200,000.

Three SPEEDY atmospheres, with different parameters chosen to reflect the typical range of behavior of different atmospheric models, were coupled to the same ocean and the same land, and also to one another, by adding inter-atmosphere coupling terms to the dynamical equations for each atmosphere. The modified equation for the temperature field for model i (i = 1 . . . 3), for instance, is:

$$\partial T_i/\partial t = (RT_i \;\; /c_p)(\dot{\sigma}_i/\sigma_i - \partial\dot{\sigma}_i/\partial\sigma_i - \nabla \cdot V_i) + \Sigma_{js} C^{ij}{}_s (T_j - T_i)\delta(x - x_s) \quad (4)$$

where all variables are evaluated at position x and $\{x_s\}$ is a set of discrete coupling points. In (4), R is the gas constant, c_p is the specific heat at constant pressure, σ is a vertical pressure coordinate scaled with surface pressure, σ its time-derivative, V is the horizontal wind velocity, and C^{ij} is a connection coefficient linking the temperature fields between models i and j at position x_s. Dynamical equations for the other independent variables, u (east-west velocity), v (north-south velocity), and q (humidity) are similarly modified to include coupling terms linking the different models.

In the present situation, regarding the PDE as a very high-order ODE, the general rule for adaptation of parameters, as applied to the connection coefficients C^{ij}, gives:

$$dC^{ij}/dt = a \int dx\ (T_j(x) - T_i(x))(T^t(x) - ⅓\ \Sigma_k T_k(x)) \quad (5)$$

where T^t is the true value of T, and *a* is an arbitrarily chosen learning rate. We assume spatially uniform connections C^{ij} that are independent of position s. Analogous rules are written to adapt the connections linking the other dynamical variables, with learning rates appropriate for their dynamics. The algorithm was tested by choosing one of the models to be a perfect replica of the "true" system; appropriate binary values for the connections did indeed result. All models are nudged to truth as the learning progresses; for the configuration studied, it was found that nudging to truth in the *u* field gave truth-model synchronization error rates that were useful in discriminating between good and bad models, so that the learning algorithm was effective.

Note that the last term in (4), connecting the models, tends to vanish as the models synchronize. This is desirable, so that each model satisfies its own physically motivated dynamical equation, without the influence of artificial coupling terms. Of, course, for each i, the parameters and hence the equations are different, so that the models cannot possibly synchronize completely. Typically, the differences in behavior are in small-scale processes that are not important for the large-scale behavior of interest.

The system was tested with the three arbitrarily chosen imperfect models of a "true" SPEEDO system, assuming ongoing nudging of the models to the "true" system, as in weather prediction. The "true" system also provided the land and ocean components for each of the imperfect models. Results are shown for the simple case of two identical models and a different third model in Fig. 5. It is seen that after 3 months, the truth-supermodel error, with adapted coefficients, is less than the error for each of the individual models, and less than the error for the supermodel with a choice of uniform connection coefficients that are not adapted.

Then the coefficients were frozen and atmospheric CO2 was doubled in the "true" system and in each of the models. Other parameters were also varied slightly. Results are shown in Fig. 6 . It is seen that the supermodel gives reduced error after three months as compared to the weighted averages of the separate models, but the coefficients learned from the single-CO2 runs are less than optimal. That is, a simple choice of uniform coefficients gives slightly better results than the learned coefficients.

It is thought that an improved learning algorithm may result from consideration of the detailed spatial structure of the various fields, with a view toward reproducing the qualitative features of the "true" dynamics. The spatial structure of the fields remains to be studied.

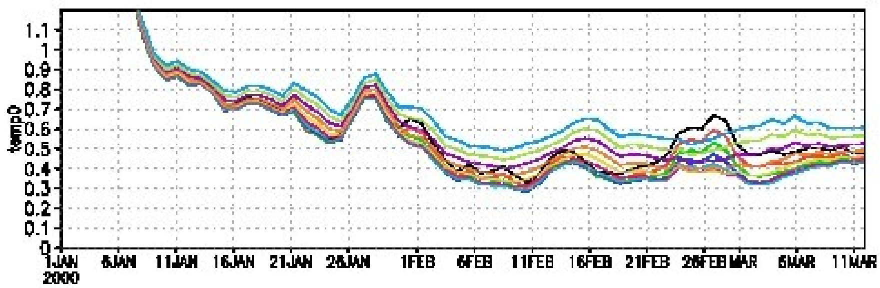
Additionally, the models could only be run for three months because of a "memory leak" in the software that is still being debugged. Longer adaptation periods might produce better results.

IV. A Weakly Connected Supermodel Formed From Full Climate Models Connected Only At the Ocean-Atmosphere Interface

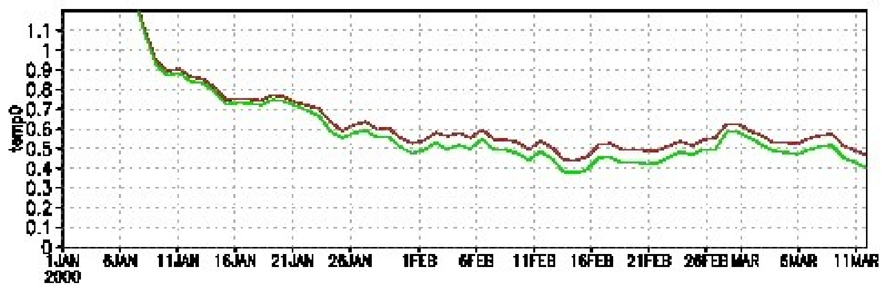
Investigations with full climate models have thus far reached a stage in which different atmosphere models are connected to a common ocean, as in the early work of Kirtman (2003) but not directly connected to each other. Yet even without the direct connections, the supermodel has been shown to be superior to any weighted combination of outputs of the individual models.

A climate model was built based on COSMOS (ECHAM5/MPIOM, developed at the Max-Planck-Institut fur Meteorologie, Germany [Jungclaus *et al.* 2006], and involved two atmospheric general circulation models (AGCMs). The two models differed in their cumulus parameterization schemes, Nordeng (1994) and Tiedtke (1989), to represent typical model diversity because cumulus convection schemes normally have a strong impact on the climate state [Kim *et al.* 2011; Klocke, Pincus, and Quaas 2011; Mauritsen *et al.* 2012]. The ocean model continuously interacts with the Nordeng atmosphere and Tiedtke atmosphere.

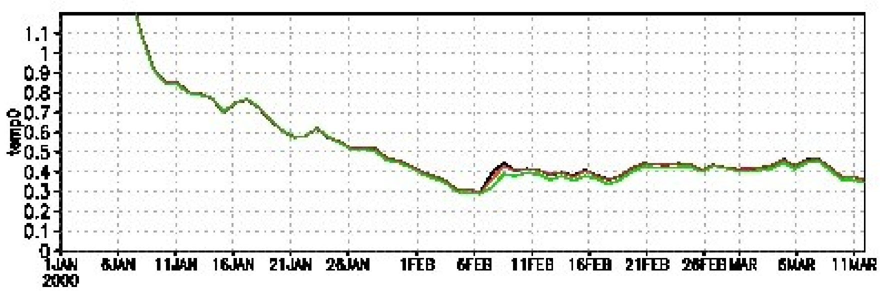
AGCMs are problematic in representing real air-sea fluxes to different degrees of accuracy. Some may be better in representing momentum flux (i.e. wind stress on the ocean) and some in energy (heat) flux [Kirtman *et al.* 2003]. Here we used different weights for the energy, momentum, and mass (i.e. precipitation) fluxes felt by the common ocean, with the sum of the weights over the two models, for each



a)



b)



c)

Figure 5: Truth-model synchronization error in surface temperature (in $^{\circ}$C) for a) three SPEEDO models with parameters perturbed away from their values in a "true" SPEEDO model to which the imperfect models are nudged via the u variable (with two of the models identically perturbed) and various weighted combinations of their outputs; b) a supermodel formed by connecting the three SPEEDO models through their dynamical equations according to Eq.(4) (for termperature) and analogous equations for u, v, and q, with constant and uniform connection coefficients C^{ij}; and (c) the same supermodel but with connections adapted according to (5) and analogous equations for the u, v, and q connections.

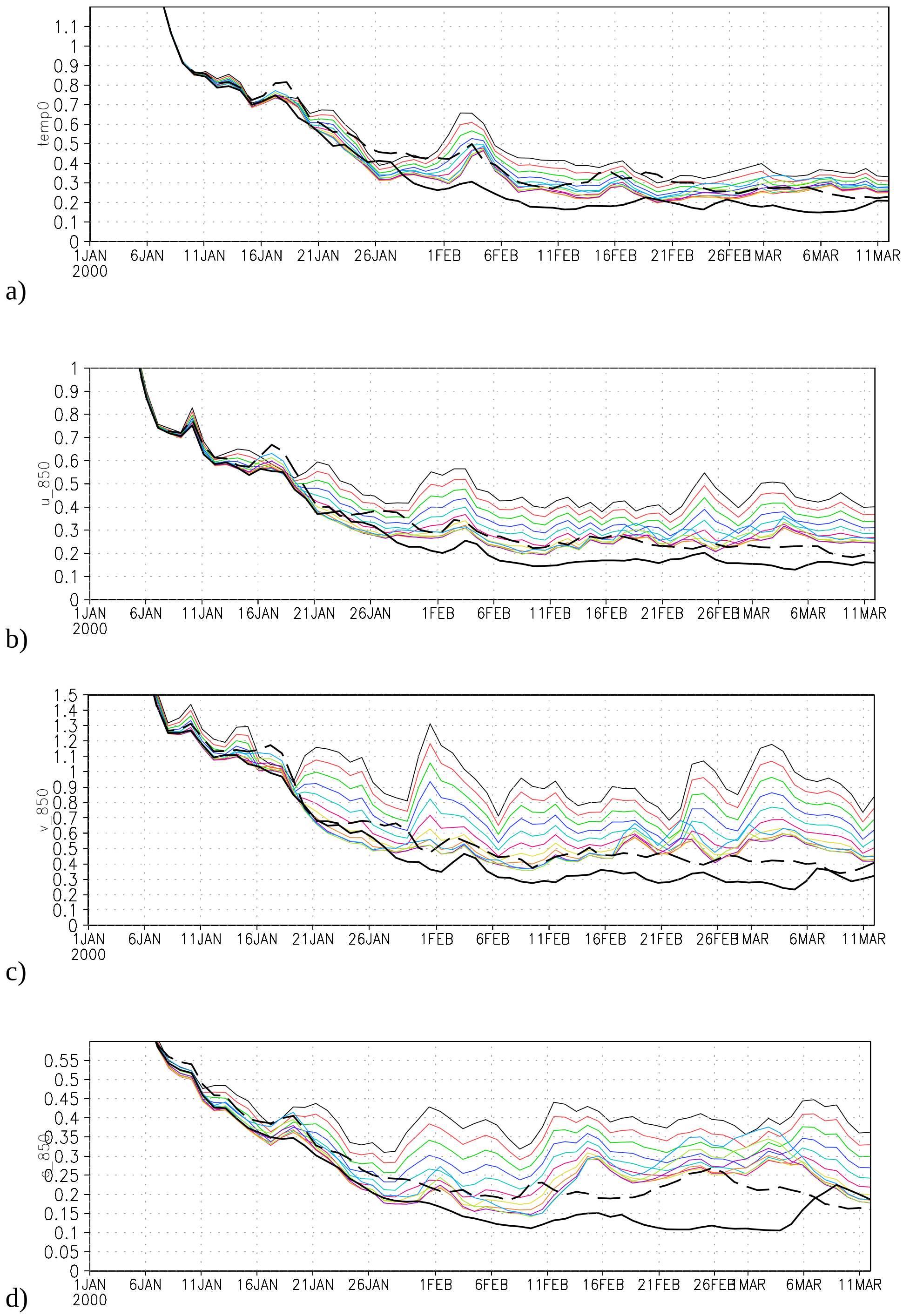


Figure 6: Truth-model synchronization error in surface temperature (in °C) a) for three SPEEDO models as in Fig. 5, but with doubled CO2 in both truth and models, for various weighted combinations of model outputs (colored lines), a supermodel with uniform connections (thick black line), and a supermodel using the connection strengths from the present-CO2 run (Fig. 5c) at final time (dashed line). Corresponding results are shown for error in zonal wind u at 850 mb (b), error in meridional wind v at 850 mb (c), and error in humidity q (d).

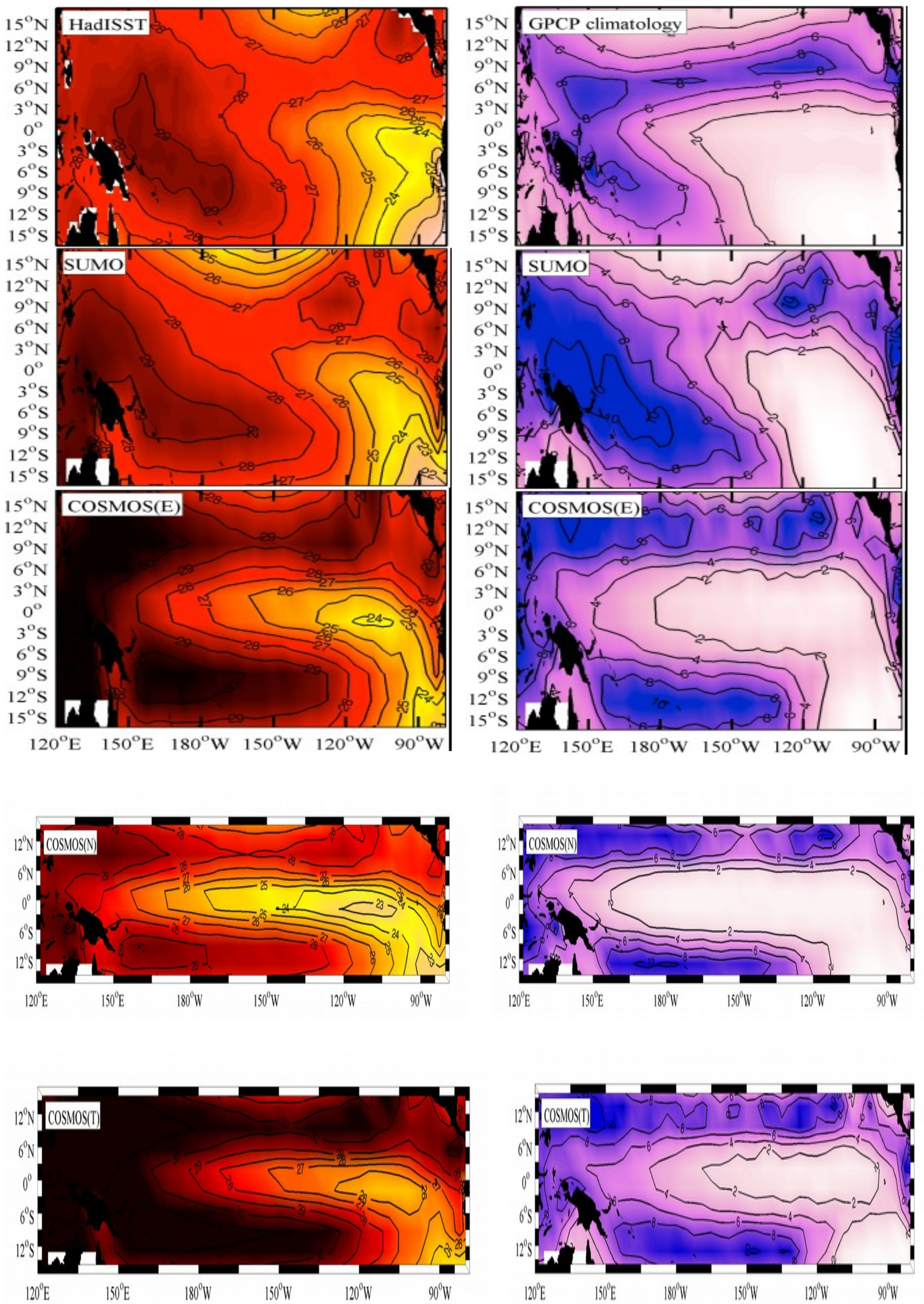


Figure 7: The climatology sea-surface temperature (SST) (left panel, scale in °C) and precipitation (right panel, scale in mm/day) in the Tropical Pacific from observation and from various models. The SST is from HadISST (1948-1979, the period used as a training set) and precipitation is from GPCP (1979-2012, due to the available data). Because the SST state over the equator is improved in the supermodel (SUMO), there is one ITCZ in SUMO.

type of flux, equal to unity. Each atmosphere feels only its own fluxes.

A machine learning technique, the Nelder-Mead method [Nelder and Meade 1965] was applied to optimize the weights for each of the fluxes. The Nelder-Mead method is also known as the simplex method, which is used to find a local minimum in multi-dimensional domain without having to compute gradients of a cost function. We used a performance index [Reichler and Kim 2008] computed over the Pacific region ($160^{\circ}E - 90^{\circ}W$, $10^{\circ}S - 10^{\circ}N$), as a metric because there is partial synchronization over the tropical Pacific in this configuration; hence it is reasonable to expect that improvement can only be achieved over this area. The assessment was started from equal weights and followed the weights suggested by the simplex method. Each case was spun up for ten years and run for another 30 years to get a reasonable climatology. Over 300 cases were tested along the path to optimality and the performance index (error) was reduced to a value at which the averaged SST bias over the metric area is only 0.48 $^{\circ}C$ and the correlation between zonal wind stress anomaly of two AGCMs is increased. Note that the variability of AGCMs tends to cancel over non-synchronized areas, thus reducing the ocean variability as well.

The behavior predicted by the supermodel was dramatically improved as shown in Figure 7, in which both the SST and precipitation have better agreement with observations. The cold tongue is stopped around the International Date Line, which suggests that a west-Pacific warm pool was formed in the supermodel, unlike the situation in COSMOS(N), COSMOS(T), or their averaged output, COSMOS(E), in all of which the cold tongue crossed the International Date Line to the western Pacific and the variability of SST is much larger (not shown). The supermodel differs from *both* component models in these regards because one of the weights is negative. The reduction of the SST bias in the supermodel implies that the whole dynamic is more realistic, suggesting that a much more realistic low level wind system exists in the supermodel, leading to a better latitudinal position of the Inter-tropical Convergence Zone (ITCZ). But it is still too wet in the South Pacific convergence zone.

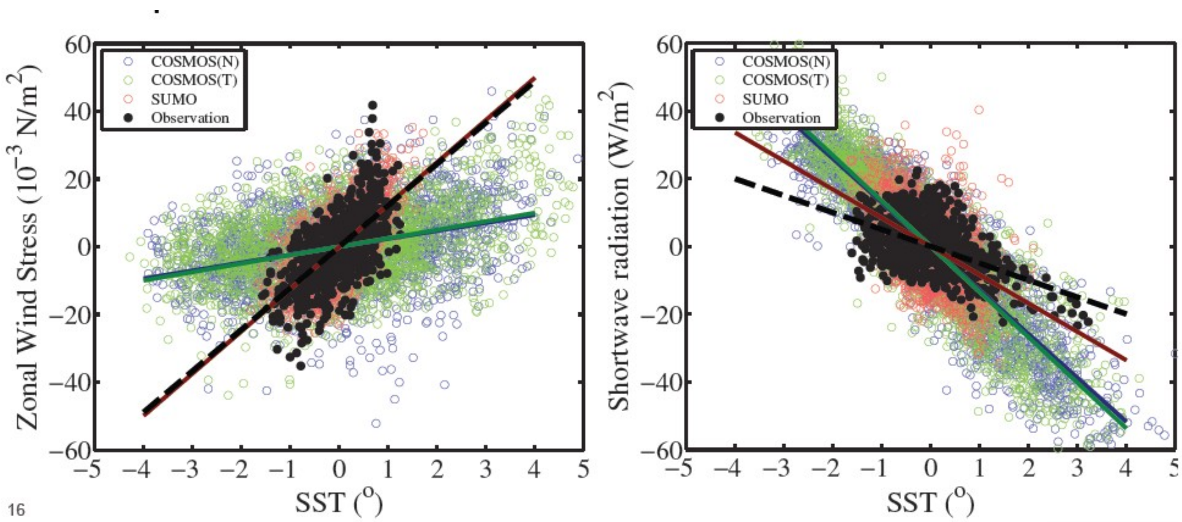


Figure 8: (a) The Bjerknes feedback (left panel), describing the relationship between the east Pacific SST anomaly (over 5°S-5°N, 150°W-90°W, Niño 3 region) and the remote wind stress over the west Pacific (5°S-5°N, 160°E-150°W, Niño 4 region); (b) the thermodynamic damping (right panel) over the Niño 3 area. Coefficients of regression and correlation are included in the legend.

The key to improved supermodel performance in this case appears to be in better representation of the air-sea feedbacks. In Figure 8, we show the Bjerknes feedback and the thermodynamic feedback for the supermodel (SUMO), the individual models, and observations. The Bjerknes feedback in the supermodel is almost perfect and the thermodynamic feedback is much improved.

It can be shown that the supermodel is superior to any weighted combination of the two model outputs. In Fig. 9, we present a Taylor diagram that shows the correlation between model and observations, as well as the normalized standard deviation of the model field, for the various models. It is seen that the supermodel has almost the same standard deviation of SST as in the observed data, unlike any of the models, and the correlation coefficient is higher.

An objection to supermodeling in the meteorological community is that ensembles of model runs (where the models are the same or different) are usually used to estimate spread as an indication of error. One loses this information with supermodeling if the models synchronize nearly completely. However, the ensemble of models in the usual practice can be replaced by an ensemble of weights. One can examine the learning history, or simply look at the performance metric for a random sample of weights, to infer a plateau in weight space along which the performance is close to optimal. Then weights on that plateau can be used to define an ensemble of supermodels. Results of this procedure, shown in Fig. 10 give a plausible ensemble of SST fields. The models effectively “agree to disagree”.

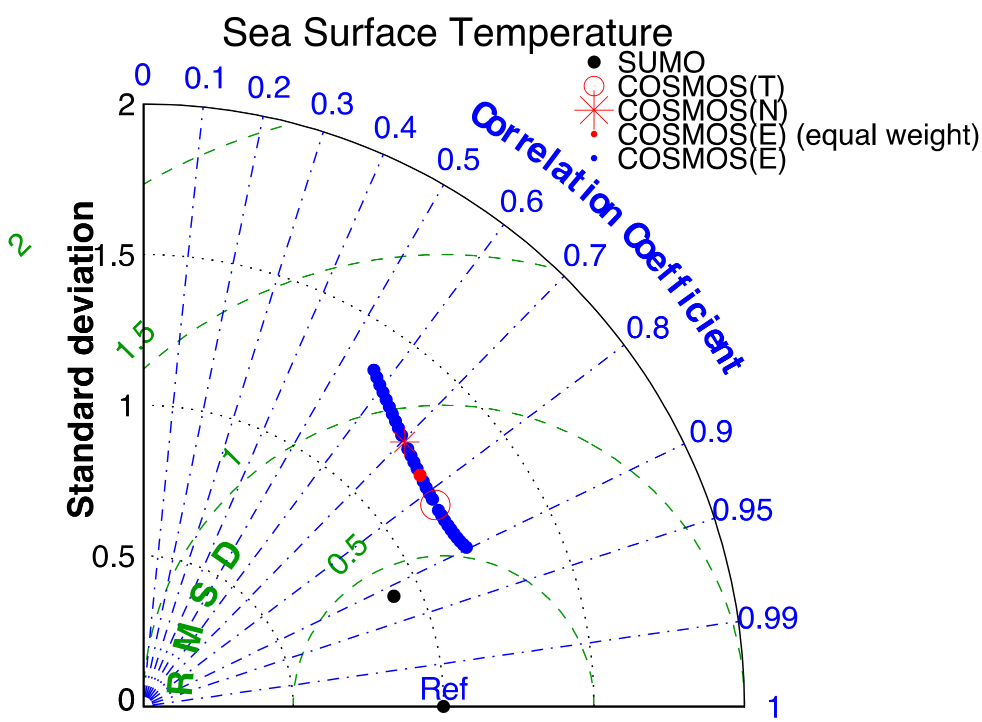


Figure 9: Taylor diagram showing the correlation between observed and modeled SST

over the Tropical Pacific, as well as the normalized standard deviation, for COSMOS(N), COSMOS(T), their equal-weighted combination COSMOS(E), all other weighted combinations (thick line), and the supermodel (SUMO). SUMO is clearly closer to observations (Ref) than any weighted average.

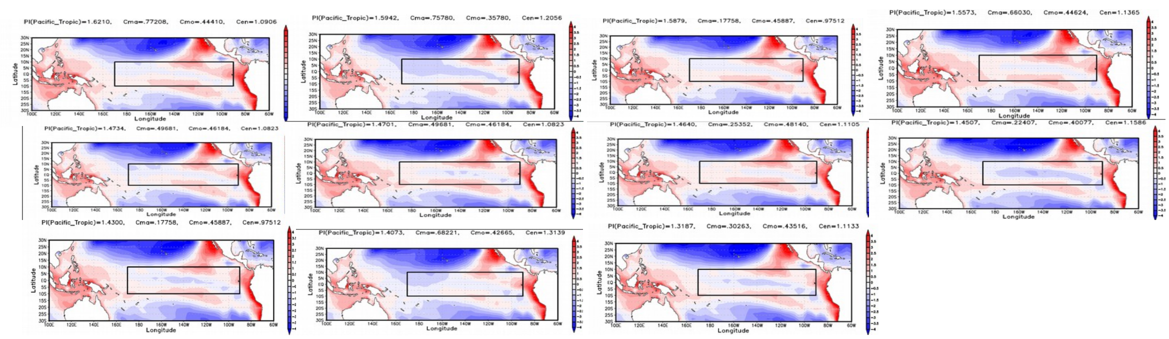


Figure 10: SST fields for an ensemble of supermodels defined by examining the learning history to select combinations of weights that give near optimal performance, each of which defines a different supermodel, giving a plausible spread in results.

The atmospheric models do not synchronize, unlike the SPEEDO models described in the previous section, except in a narrow region where indeed the supermodel results are superior. That is because the two atmospheres are connected only indirectly, through interactions with a common ocean. It is thought that direct connections between the atmospheres, suitably adapted, will further improve the supermodel performance.

V. A General Software Framework for Connecting Alternative Climate Models

In the previous stage of the project, it was hoped to extend NCAR's Data Assimilation Research Testbed (DART) to allow models to assimilate date from each other, as well as from reality, thus forming a supermodel. Subsequently, to provide a more accessible framework for interpretation and debugging of interactive ensembles formed from full climate models, it was decided to move the function of inter-model nudging to external interpreted scripts. These scripts were used in conjunction with existing DART/CESM software that time-evolves an ensemble consisting of multiple models or model versions (not just multiple instances of the same model as in ensemble Kalman filtering.) This strategy was applied to three versions of CAM:

- MODEL-1: CAM3-Aquaplanet
- MODEL-2: CAM4-Aquaplanet
- MODEL-3: CAM5-Aquaplanet 26 levels

The software configuration is schematized in Figure 11.

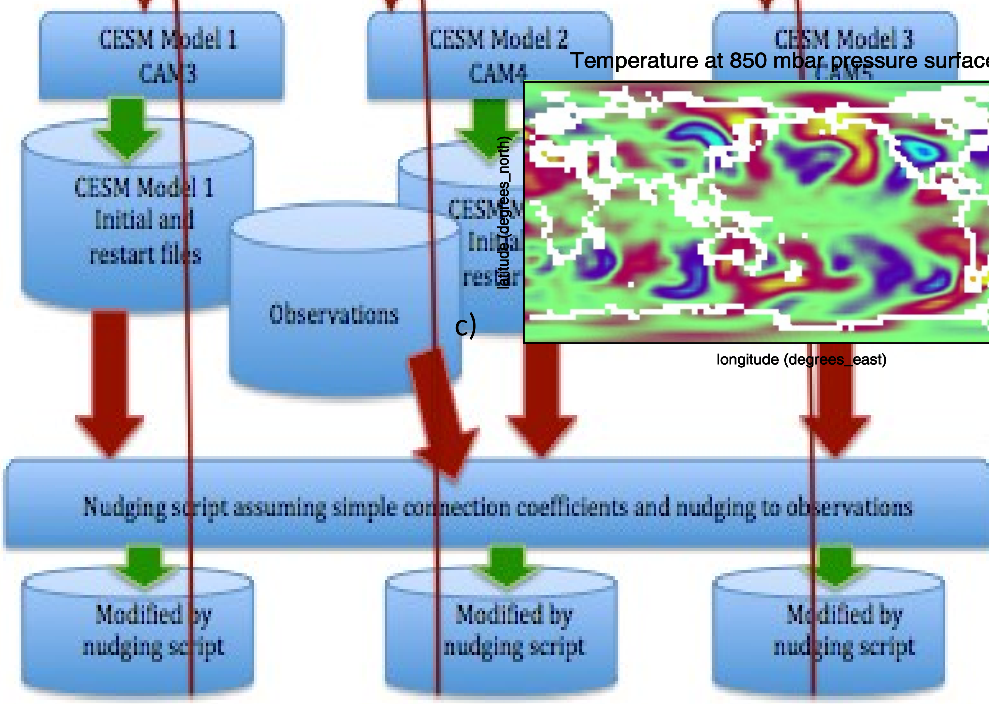


Figure 11: Three CESM models with different versions of the atmospheric component CAM are nudged to each other, and to observations in a training phase (during which the nudging coefficients are adapted).

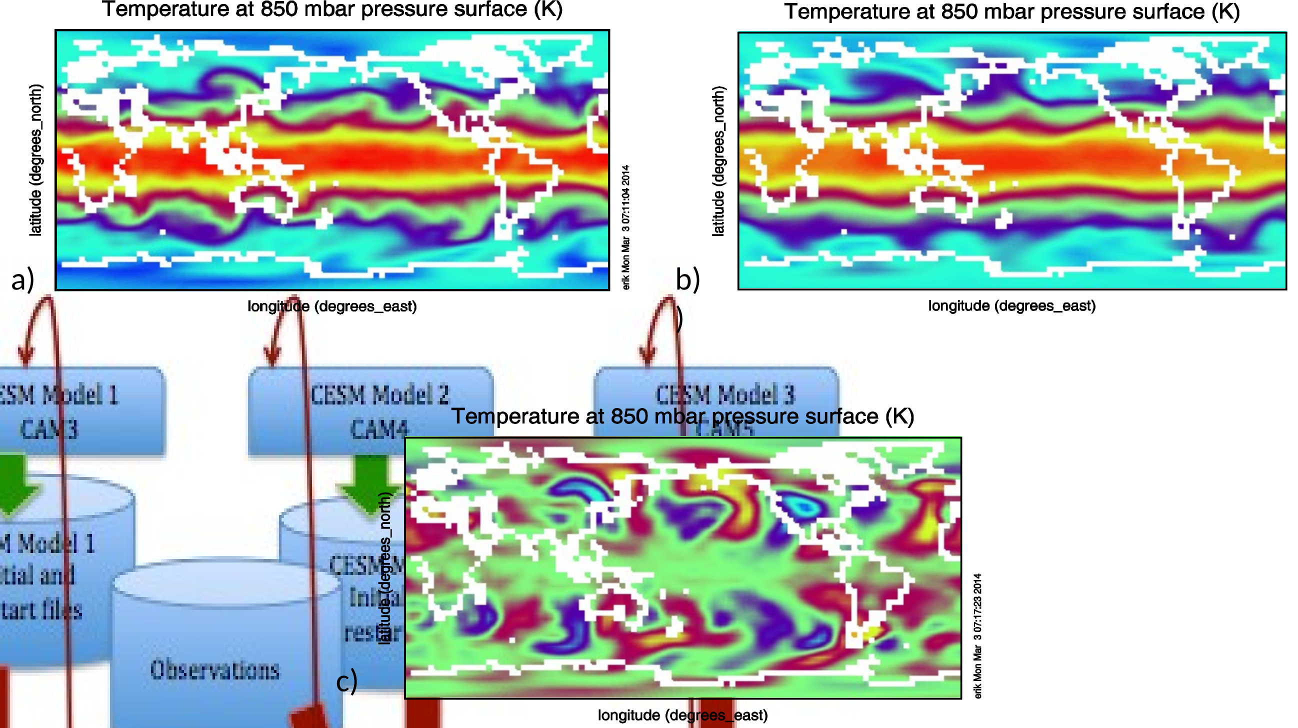


Figure 12: Temperature at 850 mb for a) CAM3 as nudged by other models in a supermodel, b) free-running CAM3, and c) the difference between the two cases, after 3 days.

At this stage of development inter-model connections are held fixed and uniform. At a later stage ancillary software would be added to adapt the coefficients in run time, during a training phase, based on a comparison with observed data. First results with the multi-CAM supermodel are shown in Fig. 12. Notably, the speed of the python scripts themselves is not an issue, compared to the speed of DART. If speed becomes an issue, we plan to use MPI modules for python to run the scripts distributed parallel. The scripts are built on top of CESM scripts, and so any component model that is part of CESM can be run this way. Further, any model that could be run from the scripts, with output that can be read by the scripts could be used as well (we are using a general python module for reading the datasets that can read multiple file formats). Therefore the current software, when fully tested and debugged, can be extended to be transparent to users, so that models that are plugged into DART/CESM are automatically available for supermodeling as well.

Toward the goal of incorporating models very different from CESM, we have been collaborating with Ben Kirtman at U. Miami (unfunded) to develop wrappers that would allow non-CESM models to be plugged into the supermodel framework. A coupling wrapper has been constructed that has been fully tested with a data model, and is currently being tested with ECHAM and the US Navy NoGAPS models.

At time of this writing, there remains a problem with the model re-initialization software that appears to use default values inappropriately. We are still trying to determine whether the problem lies with the DART/CESM software or with the newly written scripts.

Publications

1) van den Berge, L.A., Selten, F.M., Wiegerinck, W., and Duane, G.S., 2011: A multi-model ensemble method that combines imperfect models through learning, *Earth Syst. Dynam*., 2, 161-177.

2) Mirchev, M., Duane, G.S., Tang, W.S., and Kocarev, L., 2012: Improved modeling by coupling imperfect models, *Commun. Nonlin. Sci. and Num. Simulation*, 17, 2741-2751.

3) Duane, G.S. 2012: Data assimilation as artificial perception and supermodeling as artificial consciousness, chapter in *Consensus and Synchronization in Complex Networks*, pp. 209-226, ed. L. Kocarev, Springer.

4) Basnarkov, L., Duane, G.S., and Kocarev, L., 2014: Generalized synchronization in spatially extended systems, *Chaos, Soltions, and Fractals*, 59, 35-41.

5) Duane, G.S., 2015: Synchronicity from synchronized chaos, *Entropy,* 17, 1701-1733.

Presentations and Meetings Attended

1) attended kickoff meeting of the European project: Supermodeling by Combining Imperfect Models, Skopje, Macedonia, Nov. 2010.

2) G.Duane, F. Selten, N. Keenlyside, W. Wiegerinck, J. Kurths, and L. Kocarev: ``Supermodeling by adaptive synchronization of climate models" (poster), Annual Meeting of the European Geophysical Union, Vienna, Austria, April 2011.

4) G.Duane, F. Selten, N. Keenlyside, W. Wiegerinck, J. Kurths, and L. Kocarev: "Supermodeling by combining imperfect models" (poster), Future and Emerging Technologies Conference, Budapest, Hungary, May 2011.

5) G.S. Duane, L. Kocarev, and F. Selten: “Supermodeling” climate by combining alternative models, Annual Conference of the National Society of Black Physicists, Austin, TX, Sept. 2011.

6) G. S. Duane, L. Kocarev, and F. Selten: Consensus among climate models via synchronized chaos, DOE climate modeling PI meeting, Sept., 2011.

7) G. S. Duane: "Supermodeling: Consensus by Synchronization of Alternative Models", (video available online at videolectures.net) Josef Stefan Institute, Ljubljana, Oct. 2011

8) lead convener of session: "Supermodeling Climate by Combining Alternative Models", AGU (American Geophysical Union) Fall Meeting, San Francisco, Dec. 2011, with presentations[3]

a) G. S. Duane and L. Kocarev: “Supermodeling by synchronization of alternative climate models”

b) B. Kirtman, presented by J. Tribbia: “A new approach for coupled GCM sensitivity studies”

c) J. Tribbia, G. Duane, I. Trpevski, D. Trpevski, and A. Karspeck: “Toward a practical implementation of an interactive multimodel with full GCMs” (poster)

9) G. S. Duane: "Consensus by synchronization of alternative climate models", Chaos 2012, Athens, May 2012.

10) G.S. Duane, F. Selten, and P. Hiemstra: “Supermodeling climate by adaptive synchronization of alternative models”, Asia-Oceania Geosciences Society (AOGS) Joint Assembly, Singapore, August 2012

11) lead convener of session: “Ensemble Methods for Combining Alternative Models of Climate Change”, EGU (European Geosciences Union) General Assembly, Vienna, April 2013, with presentations:

a) G.S. Duane: “Interactive vs. non-interactive ensembles for weather prediction and climate projection”

b) L. Basnarkov, G.S. Duane, and L. Kocarev: “Supermodel- interactive ensemble of atmospheric models” (poster)

c) D. Trpevski, A. Karspeck, and G.S. Duane: “A computational framework for supermodeling as inter-model data assimilation” (poster)

12) organizer of minisymposium “Supermodeling Climate by Synchronization of Alternative Models” at the SIAM (Society for Industrial and Applied Mathematics) Conference on Dynamical Systems, Snowbird, UT, May 2013, with presentations:

a) G. S. Duane: “Truth-model synchronization and model-model synchronization: A path to intelligent compact representation of a high-dimensional reality”

[3] For conference sessions convened, presentations are listed that were given or co-authored by the PI’s and co-PI’s on this project only.

b) A. A. Tsonis and K. Steinhaeuser: “A climate model inter-comparison at the dynamics level”

c) J. Tribbia: **“**Synchronous coupling of large climate models for improved climate change projection”

13) G.S. Duane, M.-L. Shen, and N. Keenlyside: “Supermodeling climate by synchronization of **alternative** models”, seminar at McGill Univ., Montreal, June 2013.

14) G.S. Duane, M.-L. Shen, and N. Keenlyside: “Supermodeling climate by synchronization of alternative models”, seminar at Ecole Normale Superieur, Paris, June 2013.

15) G.S. Duane, M.-L. Shen, and N. Keenlyside: “Supermodeling climate by synchronization of alternative models”, seminar at COLA/George Mason U., July, 2013.

16) G.S. Duane: “Supermodeling for climate and weather prediction”, seminar at NCAR, Boulder, August 2013

17) G.S. Duane: **“**Improved climate modeling by synchronization of alternative models”, SUMO Summer School, Ohrid, Macedonia, September 2013

18) lead convener of session: “Unification of Alternative Models in Climate and Geophysics via Multi-Model Ensembles, Stochastic Parameterization, and Networks”, AGU Fall Meeting, San Francisco, December 2013, with presentations:

a) G.S. Duane: “Interactive vs. non-interactive multi-model ensembles”

b) M.-L. Shen, N. Keenlyside, F. Selten, W. Wiegerinck, G.S. Duane: “Reducing model systematic error through super-modeling; The tropical Pacific”

c) A. A. Tsonis and K. Steinhaeuser: “A climate model inter-comparison at the dynamics level”

d) E. Kluzek, G. Duane, J. Tribbia, M. Vertenstein: “Software engineering designs for supermodeling different versions of CESM models using DART” (poster)

19) lead convener of session: “Ensemble Methods for Combining Alternative Models of Climate Change”, EGU (European Geosciences Union) General Assembly, Vienna, April 2014, with presentations:

a) F. Selten: **“**Synchronization and super modeling experiments with a complex climate model”

b) M.-L. Shen, N. Keenlyside, F. Selten, W. Wiegerinck, G.S. Duane: “Reducing Model Systematic Error over Tropical Pacific through SUMO Approach”

c) N. Keenlyside, M.-L. Shen, F. Selten, W. Wiegerinck, G.S. Duane: Climate Change Projection with Reduced Model Systematic Error over Tropic Pacific”

d) G.S. Duane, W. Wiegerinck, and M.-L. Shen, “Interactive Ensembles Without Loss of Spread Information” (poster)

e) W. Wiegerinck and G.S. Duane, “Comparing error reduction in interactive and non-interactive ensemble approaches” (poster)

20) G.S. Duane: “Fusion of alternative models by synchronization: results with realistic models” (poster), DOE Climate Modeling PI Meeting, Potomac, May 2014

21) G.S. Duane, N. Keenlyside, M.-L. Shen: “Interactive ensembles with spread information”, AOGS 11th Annual Meeting, Sapporo, Japan, Jul-Aug 2014

22) G.S. Duane, M.-L. Shen, and N. Keenlyside: “Supermodeling climate by synchronization of alternative models”, seminar at UCLA, Los Angeles, August 2014

23) G.S. Duane, M.-L. Shen, N. Keenlyside, and F. Selten: “Supermodeling climate by synchronization of alternative models”, seminar at Weizmann Inst., Rehovot, Israel, June 2015

24) G.S. Duane, M.-L. Shen, N. Keenlyside, and F. Selten: “Supermodeling an objective process by synchronization of alternative models”, ITISE conference (Intl. Work-Conf. on Time Series Analysis), Granada, Spain, July 2015

Travel

Travel to all of the above meetings and presentations (some with external funding) and additionally:

1) collaboration with Jurgen Kurths, Potsdam Institute, Germany, January 2011

2) meeting with Ben Kirtman, U. Miami co-PI, regarding plan for software development and the use of ``wrappers” ; and regarding division of labor with NCAR and with European collaborators. Dec. 2011.

3) travel for Frank Selten, for collaboration in Boulder, August, 2011 (jointly sponsored by U. Colorado and NCAR).

4) travel for Frank Selten to AGU Fall Meeting, Dec. 2011 (jointly sponsored by U. Colorado and NCAR).

5) travel for Daniel Trpevski (student of Kocarev) to Boulder to work with Alicia Karspeck on software development for the revised DART, April 2012.

6) travel for PI to Utrecht for collaboration with Frank Selten, July 2012. (local expenses only)

7) travel for PI to Milwaukee for collaboration with co-PI Tsonis, August 2012.

8) travel for PI to Utrecht for collaboration with Frank Selten on quasigeostrophic models, August 2012.

9) travel for PI to Potsdam for meeting with Juergen Kurths and Frank Selten, March 2013.

10) travel for PI to Miami for collaboration with co-PI Kirtman regarding software development, December 2013.

References

Duane, G.S. and Tribbia, J.J., 2004: Weak Atlantic-Pacific teleconnections as synchronized chaos, *J. Atmos. Sci.* 61, 2149-2168.

Duane, G.S., Yu, D., and Kocarev, L., 2007: Identical synchronization, with translation in- variance implies parameter estimation, *Phys. Lett. A* 371, 416–420.

Duane, G.S., 2015: Synchronicity from synchronized chaos, *Entropy,* 17, 1701-1733.

Goosse H. and Fichefet, T., 1999: Importance of ice-ocean interactions for the global ocean circulation: a model study, *J. Geophys. Res.* 104, 23337-23355.

Intergovernmental Panel on Climate Change, 2007: *IPCC Fourth Assessment Report: Climate Change 2007 (AR4), Working Group I Report: The Physical Science Basis*.

Jungclaus, J. H. *et al.*, 2006: Ocean Circulation and Tropical Variability in the Coupled Model ECHAM5/MPI-OM, *J. Climate* 19, 3952–3972.

Kim, D. *et al.*, 2011; El Nio-Southern Oscillation Sensitivity To Cumulus Entrainment In A Coupled General Circulation Model. *J. Geophys. Res.* 116, D22112.

Kirtman, B. P., Min, D., Schopf, P. S., and Schneider, E. K., 2003: A New Approach for Coupled GCM Sensitivity Studies, *COLA Technical Report No.* 154.

Klocke, D., Pincus, R., and Quaas, J., 2011: On Constraining Estimates of Climate Sensitivity with Present-Day Observations through Model Weighting, *J. Climate* 24, 6092–6099.

Mauritsen, T. *et al.,* 2012: Tuning the Climate of A Global Model, *J. Adv. Model. Earth Syst.* 4, M00A01.

Nelder, J. A. and Mead, R., 1965: A Simplex Method for Function Minimization, *The Computer Journal* 7, 308–313.

Nordeng, T.-E., 1994: Extended Versions of the Convective Parametrization Scheme at ECMWF and Their Impact On the Mean and Transient Activity of the Model in the Tropics, *Technical Memorandum* No. 206, European Centre for Medium Range Weather Forecasts.

Reichler, T. and Kim, J., 2008: How Well Do Coupled Models Simulate Today's Climate?, *Bull. Amer. Meteor. Soc.* 89, 303–311.

Shen, M.-L., Keenlyside, N., Selten, F., Wiegerinck, W., and Duane, G.S., under review 2015: Dynamically Combining Climate Models to "Supermodel" the Tropical Pacific, submitted to *Geophys. Res. Lett.*

Severijns, C. and Hazeleger, W. , 2009: The efficient global primitive equation climate model Speedo., *Geoscientific Model Development Discussions* 2, 1115–1155.

Steinhaeuser, K. and Tsonis, A.A. 2013: A climate model intercomparison at the dynamics level, *Clim. Dyn.* Doi:10.1007/s00382-013-1761-5.

Tiedtke, M., 1989: A Comprehensive Mass Flux Scheme for Cumulus Parameterization in Large-Scale Models. Monthly Weather Rev. 117, 1779–1800.

Vautard, R. and Legras, B., 1988: On the source of mid-latitude low-frequency variability. Part II: Nonlinear equilibration of weather regimes, J. *Atmos. Sci.* 45, 2845-2867.

Vautard, R., Legras, B. and Deque, M., 1988: On the source of mid-latitude low-frequency variability. Part I: A statistical approach to persistence, *J. Atmos. Sci.* 45, 2811-2843.