BAMS Article

Supermodeling

Improving Predictions with an Ensemble of Interacting Models

Francine Schevenhoven, Noel Keenlyside, François Counillon, Alberto Carrassi, William E. Chapman, Marion Devilliers, Alok Gupta, Shunya Koseki, Frank Selten, Mao-Lin Shen, Shuo Wang, Jeffrey B. Weiss, Wim Wiegerinck, and Gregory S. Duane

KEYWORDS:
Climate prediction;
Superensembles;
Forecasting techniques;
Data assimilation;
Model errors;
Machine learning

ABSTRACT: The modeling of weather and climate has been a success story. The skill of forecasts continues to improve and model biases continue to decrease. Combining the output of multiple models has further improved forecast skill and reduced biases. But are we exploiting the full capacity of state-of-the-art models in making forecasts and projections? Supermodeling is a recent step forward in the multimodel ensemble approach. Instead of combining model output after the simulations are completed, in a supermodel individual models exchange state information as they run, influencing each other's behavior. By learning the optimal parameters that determine how models influence each other based on past observations, model errors are reduced at an early stage before they propagate into larger scales and affect other regions and variables. The models synchronize on a common solution that through learning remains closer to the observed evolution. Effectively a new dynamical system has been created, a supermodel, that optimally combines the strengths of the constituent models. The supermodel approach has the potential to rapidly improve current state-of-the-art weather forecasts and climate predictions. In this paper we introduce supermodeling, demonstrate its potential in examples of various complexity, and discuss learning strategies. We conclude with a discussion of remaining challenges for a successful application of supermodeling in the context of state-of-the-art models. The supermodeling approach is not limited to the modeling of weather and climate, but can be applied to improve the prediction capabilities of any complex system, for which a set of different models exists.

https://doi.org/10.1175/BAMS-D-22-0070.1
Corresponding author: Francine Schevenhoven, francine.schevenhoven@uib.no
In final form 6 July 2023

AFFILIATIONS: Schevenhoven—Department of Atmospheric and Oceanic Sciences, University of Colorado Boulder, Boulder, Colorado, and Geophysical Institute, University of Bergen, and Bjerknes Centre for Climate Research, Bergen, Norway; **Koseki, Shen, and Wang**—Geophysical Institute, University of Bergen, and Bjerknes Centre for Climate Research, Bergen, Norway; **Keenlyside and Counillon**—Geophysical Institute, University of Bergen, and Bjerknes Centre for Climate Research, and Nansen Environmental and Remote Sensing Center, Bergen, Norway; **Carrassi**—Department of Physics and Astronomy "Augusto Righi," University of Bologna, Bologna, Italy; **Chapman**—National Center for Atmospheric Research, Boulder, Colorado; **Devilliers**—Danish Meteorological Institute, Copenhagen, Denmark; **Gupta**—NORCE Norwegian Research Centre, and Bjerknes Centre for Climate Research, Bergen, Norway; **Selten**—Royal Netherlands Meteorological Institute, De Bilt, Netherlands; **Weiss and Duane**—Department of Atmospheric and Oceanic Sciences, University of Colorado Boulder, Boulder, Colorado; **Wiegerinck**—SNN Adaptive Intelligence, and Department of Biophysics, Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, and HAN University of Applied Sciences, Arnhem, Netherlands

The modeling of our climate is a challenging task given the complexity of the system. A myriad of processes, not all perfectly understood, interact across scales over many orders of magnitude. Computational constraints limit the accuracy of solving the equations numerically on a grid with finite resolution. The effect of unresolved processes is approximated by parameterized descriptions, often based on empirical relationships. Model development is thus a continuously ongoing process that has resulted in a steady, yet relatively slow, improvement of weather and climate models (Bauer et al. 2015). The climate models participating in the Coupled Model Intercomparison Project (CMIP) improve from generation to generation (Reichler and Kim 2008; Bock et al. 2020). Still, systematic biases persist. These biases include the double intertropical convergence zone (ITCZ) in the Pacific (Tian and Dong 2020), a poor simulation of tropical and subtropical low-level clouds (Stouffer et al. 2017) ("too few, too bright" low-level cloud problem; Nam et al. 2012), underestimation of atmospheric blocking over Europe (Davini and D'Andrea 2016), and warm tropical biases in sea surface temperature (SST) (Richter 2015).

An ensemble of simulations of different models—a multimodel ensemble (MME)—is a useful approach to account for model uncertainty (Hagedorn et al. 2005; Lee et al. 2021). Averaging over the MME tends to reduce systematic errors. In addition, predictions can be improved by combining MME forecasts through weighting different models to form a consensus forecast (Krishnamurti et al. 2016), a posteriori. The weights (as many as 10 million) vary in space and time and are trained on retrospective predictions. These statistically determined weights are then used in prediction mode under the assumption that the processes are autonomous (i.e., they are not a function of time dependent forcing such as the anthropogenic forcing). However, such postprocessing methods might not be optimal for nonlinear processes or for predicting extrema of types or magnitudes that did not occur during training. Correcting ex post facto averages for historical biases does therefore not necessarily improve the skill of predictions and climate projections.

In this paper, we will show that predictions and projections can be further improved by reducing initial error growth through combining the models of an MME dynamically into a so-called supermodel. The models in a supermodel interact with each other during their runs. The interaction is frequent enough such that the models can compensate for each

other's errors before they extend to larger scales and affect other variables in other regions. Essentially, a new dynamical system is formed. We will formalize the supermodel definition, identify the potential and promise of supermodeling, and identify the challenges to developing this vision and provide potential solutions.

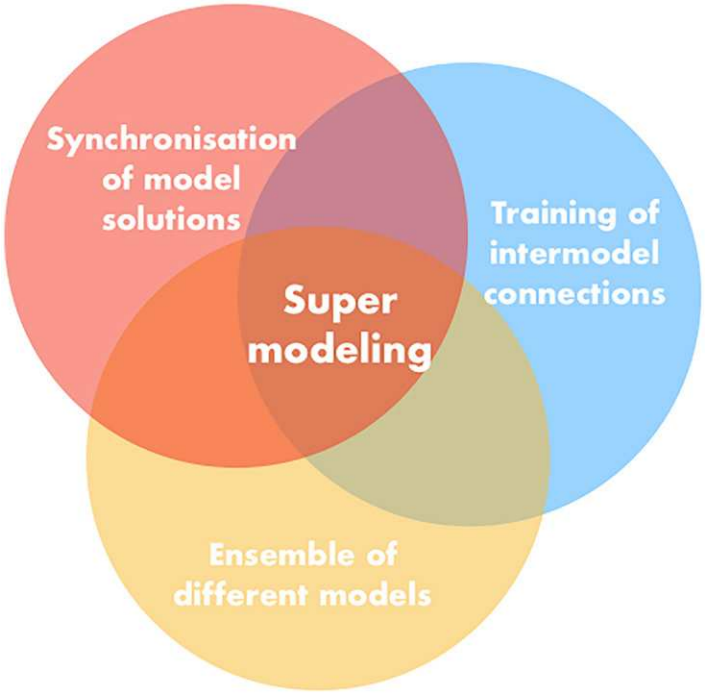

Fig. 1. Key elements of supermodeling.

What is supermodeling?

A supermodel is an ensemble of different models interacting so as to induce synchronization with each other as they run, with the interaction terms trained on the basis of data to achieve a superior model. The frequent interaction among models is central to supermodeling; the models use each other's states to continue their simulation. The interaction can for instance take place by nudging the state of each model in the ensemble to the state of every other model. Instead of nudging terms, we use the more general phrase "connection terms" to describe the interaction between models. Depending on the form and strength of such connections, the models synchronize (i.e., are in approximately the same state at the same time). Synchronization within a supermodel is important to maintain internal variability: combining states of models that are in different phases results in smoothing and reduction of variability. Synchronization is therefore a key element of supermodeling. The optimal strength of the connection terms between the models is found by training the connection terms on the basis of historical observations. Training is therefore another key element of supermodeling. Whether the ensemble of models can synchronize on a solution that is close to the observed evolution depends on whether the models can compensate for each other's shortcomings. Below we will explain these different elements in detail. To summarize, supermodeling is based on an ensemble of different models that have compensating errors, the synchronization of the trajectories of these models, and the training of the free parameters introduced in the intermodel connections to achieve a model superior to any of the individual models. (Fig. 1).

An example of a supermodel based on Lorenz 63. A simple supermodel using an ensemble of three parameter-perturbed Lorenz 63 systems (Lorenz 1963) was advanced by van den Berge et al. (2011), following a formulation originally introduced by Duane (et al. 2009, EGU conference presentation).[1]

[1] Only the abstract (Duane et al. 2009) is published. For additional details, see, e.g., the review in Duane (2015), section 5.

In the Lorenz 63 model, Eq. (1), a chaotic attractor emerges for the standard values of the parameters $\sigma = 10$, $\rho = 28$, and $\beta = 8/3$ (Lorenz 1963). The state space is described by coordinates x, y, and z. In van den Berge et al. (2011), the model with standard parameter values is regarded as the "perfect model," alias "truth," and three imperfect models are generated by perturbing these standard parameter values (see Table 1). We will elaborate in the section

"An ensemble of different models" on the choice of parameters to form an ensemble of different models:

$$\dot{x} = \sigma (y - x), \tag{1a}$$

$$\dot{y} = x (\rho - z) - y, \tag{1b}$$

$$\dot{z} = xy - \beta z. \tag{1c}$$

Table 1. Standard and perturbed parameter values for the Lorenz 63 system.

Model	σ	ρ	β
Perfect model	10	28	8/3
Model 1	13.25	19	3.5
Model 2	7	18	3.7
Model 3	6.5	38	1.7

The perturbed parameters lead to very distinct behavior (Fig. 2). In the first two models, the trajectories converge onto stable fixed points, while the third has a chaotic attractor with larger variations and a shift to larger z values as compared to the perfect model.

We form a supermodel from the three imperfect models by adding linear connection terms to the equations:

$$\dot{x}_i = \sigma_i (y_i - x_i) + \sum_{j \neq i} C_{ij}^x (x_j - x_i), \tag{2a}$$

$$\dot{y}_i = x_i (\rho_i - z_i) - y_i + \sum_{j \neq i} C_{ij}^y (y_j - y_i), \tag{2b}$$

$$\dot{z}_i = x_i y_i - \beta_i z_i + \sum_{j \neq i} C_{ij}^z (z_j - z_i), \tag{2c}$$

where i indexes the three imperfect models, and $\mathbf{C}_{ij} = \left(C_{ij}^x, C_{ij}^y, C_{ij}^z\right)$ denote the connection coefficients. These determine the strength of the nudging of model i toward model j. In the examples in this paper the connection coefficients are time independent, but they could be time dependent, for example, to better represent seasonal cycles.

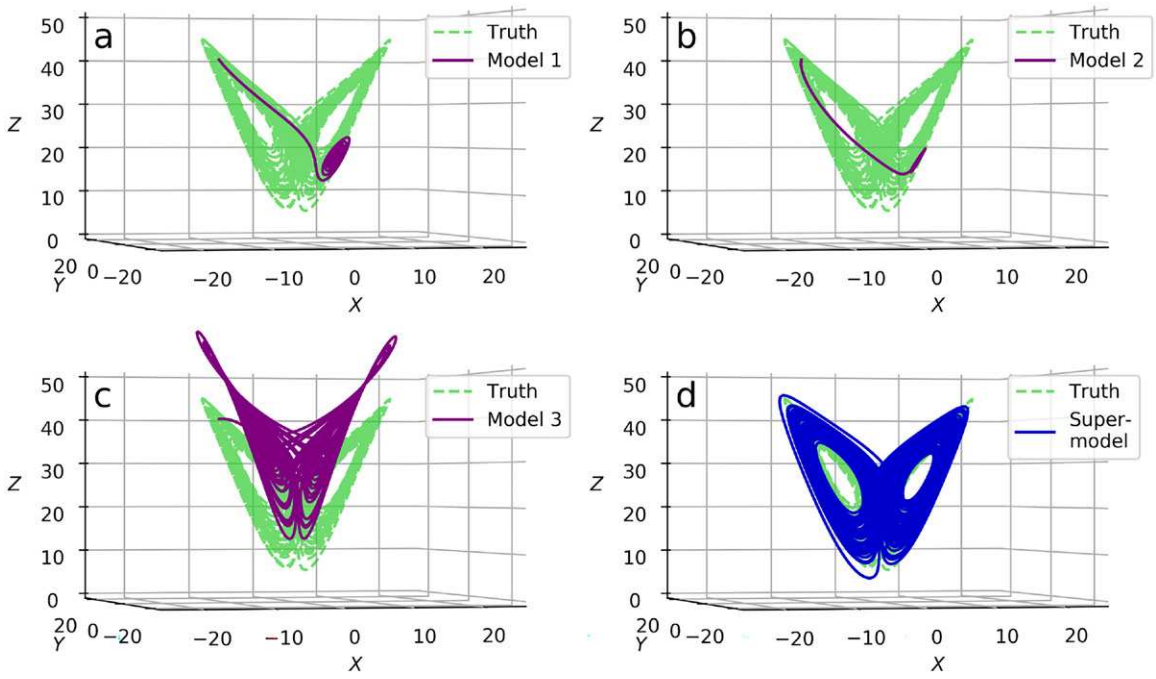

Fig. 2. Trajectories for (a)–(c) the three unconnected imperfect models (purple) and (d) the supermodel (blue), and the standard Lorenz 63 system (green). The trajectory for the imperfect models includes the transient evolution from the initial condition toward the attractor. Replotted from van den Berge et al. (2011).

The connection coefficients $\mathbf{C}_{ij}$ in this example are trained based on the minimization of a cost function calculating the error between the supermodel trajectory and the perfect model trajectory. However, more efficient training methods have been developed, as introduced in the section "Training of intermodel connections." After the connection coefficients $\mathbf{C}_{ij}$ have been established by training, Eq. (2) defines a new model, the supermodel. The connection coefficients are sufficient to cause the three imperfect models in the supermodel to synchronize on a chaotic attractor. Since the models are not perfectly synchronized, the supermodel solution is defined as the ensemble average of the imperfect models (Fig. 3). The supermodel is robust against the choice of the initial condition. Once we run the supermodel from out-of-sample initial states, we see that the supermodel attractor is very close to that of the perfect model, and hence very different from the original imperfect models (Fig. 2d).

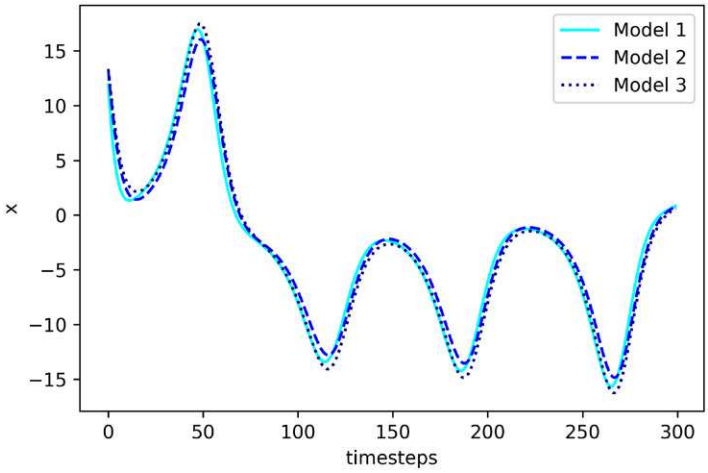

Fig. 3. Time series for the *x* coordinates of the three connected imperfect models forming the supermodel of van den Berge et al. (2011). The synchronization between the models is not perfect, but the models are in the same phase and close to each other.

The Lorenz 63 experiment was the first proof of concept of supermodeling. It suggested that it is possible to synchronize different models on a common solution by introducing intermodel connections. The model combination can be trained using observed trajectories resulting in a supermodel with more realistic dynamical behavior and improved predictive skill.

Supermodeling in comparison to the MME and machine learning

Supermodeling versus MME. In a standard MME the models do not interact. Their individual model outputs are combined after the simulation period is completed. In a supermodel, on the other hand, the models exchange states with each other during their simulations. If the interaction between the models in a supermodel is strong and frequent enough, the models synchronize their evolution. The models form a "consensus solution" that reflects the average behavior and internal variability of the climate system better than the ensemble mean of the MME. Therefore, an advantage of the supermodel approach is that it not only provides improved estimates of climate statistics as a standard MME can do, but due to the synchronization, it also generates a unique, dynamically and physically consistent continuous simulation, including the occurrences of extreme events such as for instance multiyear droughts. In analogy with "classical" probabilistic prediction as in an MME, the uncertainty associated with the supermodel simulation can be represented by an ensemble of differently trained supermodels (i.e., an MME of supermodels). The supermodel can also be used to create a single model ensemble by perturbing initial conditions of the supermodel. Figure 4 schematically shows the difference between the construction of an MME (Fig. 4a) and that of a supermodel (Fig. 4b).

Supermodeling as a form of machine learning. The study of machine learning (ML) has resulted in various methods for generating models solely through algorithmic analysis of data. The algorithms are based on learned weights and biases, serving as connections between nodes in a large network, enabling them to generalize and replicate any functional relationship. Modern ML methods have been used to bias correct models both offline (e.g., Chapman et al. 2022) and online (Watt-Meyer et al. 2021). Unless explicitly informed by

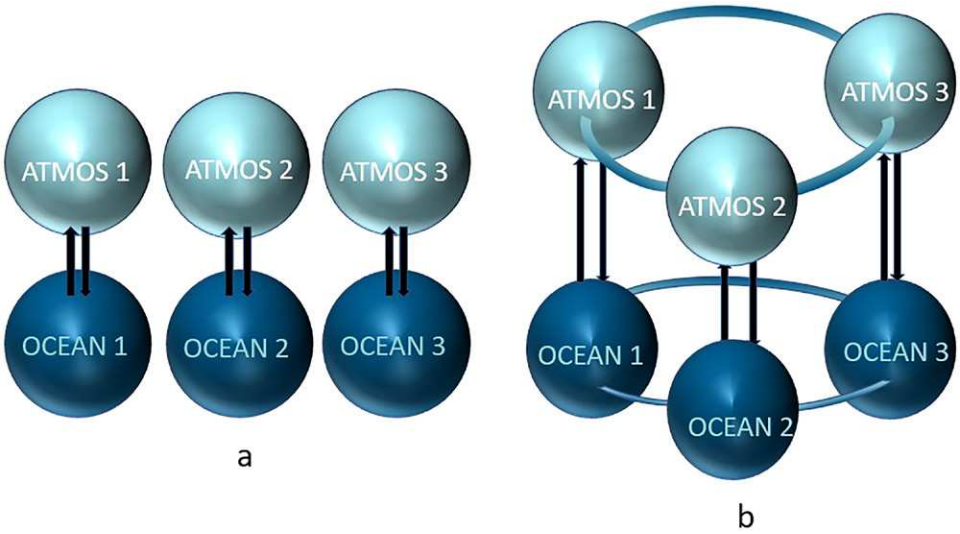

Fig. 4. Difference between (a) a multimodel ensemble (MME) of independently run coupled atmosphere–ocean models, and (b) an atmosphere–ocean-connected supermodel with interactions during the run (indicated by the circular bands) between the atmosphere components of the models and between the ocean components of the models.

physics (Cheng et al. 2023; Jakhar et al. 2023), ML models learn their knowledge about the climate, including basic physics, empirically from data alone. This learning task requires large amounts of data, even if training is restricted to connections leading to output units only, as in the "reservoir computing" approach (Schrauwen et al. 2007).

A supermodel can be compared to an ML model in which state-of-the-art climate models are used as modular learning blocks (an analogous neural network example is a residual block; He et al. 2016) within an ML network. However, in contrast to standard ML, the weights are physics constrained, and expert-user built. They are thus physics aware: the individual models, the ingredients of the training, bring their physical consistency to the supermodel. In a supermodel, only the weights to steer the interaction between the individual models need to be trained from data. This gives supermodels an advantage in the learning task compared to standard ML approaches.

Synchronization of model solutions

Synchronization of model solutions to evolve simultaneously in a similar manner is one of the key elements of supermodeling. The concept of synchronization also applies to data assimilation (Yang et al. 2006; Duane et al. 2006). In data assimilation, a model synchronizes with reality by assimilating a relatively sparse set of observations (Carrassi et al. 2018). In a supermodel, the models assimilate data from other models and synchronize on a common solution. Here we discuss some practical issues regarding the intermodel connections in relation to the requirement of a synchronized solution.

In supermodeling, where nonidentical models (i.e., models not defined by exactly the same equations) are connected, synchronization is always imperfect. In practice, however, synchronization errors are small (e.g., Fig. 3). The supermodel solution is defined by averaging the synchronized model solutions. Since synchronization errors are small, there is hardly any loss of variance in doing so. Perhaps surprisingly, models can synchronize when connected strongly and frequently enough, despite the well-known sensitive dependence of weather on initial conditions ("butterfly effect"). It should be recognized that much of the inspiration for supermodeling came from outside of meteorology, from nonlinear dynamics, where the synchronization of general chaotic systems has been an active area of research (Pecora and Carroll 2015), including a generalization to chaotic systems that are very different (Rulkov et al. 1995).

A crucial feature of synchronization is that not every prognostic variable needs to be connected (Pecora et al. 1997); those that are connected do not need to communicate at every time step, nor at every location. This reduces the data transfer needs between models to a practical level. In synchronization experiments with a supermodel consisting of different parametric versions of a global primitive equation model it was found that it was sufficient to connect the temperature and momentum equations. A degree of synchronization was obtained such that the variance of the supermodel was not reduced or smoothed compared to the perfect model (Selten et al. 2017). Nudging the remaining prognostic variables, such as specific humidity and surface pressure, disturbed the synchronized state, probably due to the introduction of imbalances and fast spurious adjustments (Schevenhoven et al. 2019).

The minimum connection frequency necessary to synchronize models sufficiently has been determined by experimentation (Selten et al. 2017; Schevenhoven and Carrassi 2022). Connecting atmosphere models every 6 h results in very small synchronization errors (Selten et al. 2017). Ocean models can be connected with a much lower frequency due to the longer time scales of the ocean dynamics.

In the next example, we will show that even in early stages of supermodel development, with a limited degree of synchronization between the models, the supermodel is already able to reduce long-term standing climatological biases. Kirtman et al. (2003) were the first to combine dynamically different models during integration. They connected two different atmosphere models to a single ocean model by using the heat flux from one atmosphere model and the momentum flux from the other, in a combination of the least erroneous fluxes. This resulted in a new dynamical system. In the areas of strong air–sea interaction, such as the tropical oceans, the atmosphere models exhibited synchronized behavior. Combining the two atmosphere models in this way led to a reduced climatological error in SSTs and a better representation of interannual variability in the tropics. This result inspired the development of supermodels for climate.

Shen et al. (2016) improved on the Kirtman et al. (2003) approach by training the model connections (Fig. 5a). In this case, the ocean general circulation model (OGCM) receives a weighted mean of the surface fluxes of two atmospheric general circulation models (AGCMs). One atmospheric model used the Tiedtke (1989) and the other the Nordeng (1994) convection scheme. Both model versions suffered from the well-known double-ITCZ bias with too much precipitation in the Southern Hemisphere tropical Pacific and a narrow cold tongue that extends too far west along the equator accompanied by an ITCZ on its northern and southern flank (Fig. 6). Despite this fact, the trained supermodel, in which the atmospheres again exhibit synchronized behavior in the tropical Pacific (Fig. 7), simulated a much more realistic climatological cold tongue, the double-ITCZ error was alleviated (Fig. 6), and in general there was an improvement of equatorial Pacific dynamics (Shen et al. 2017).

This solution was achieved by compensating for errors in surface winds before ocean–atmosphere interactions lead to

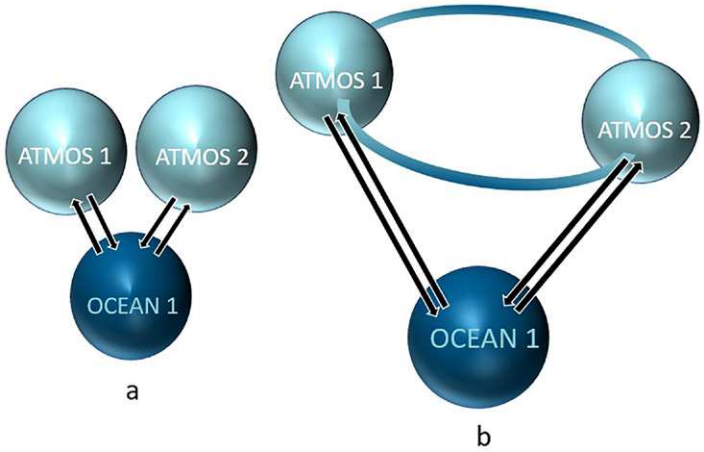

Fig. 5. (a) The supermodel setup of Shen et al. (2016), with two atmospheric models with different convection schemes combined by sharing their fluxes with one ocean model. (b) The supermodel setup of Selten et al. (2017) and Schevenhoven et al. (2019), in which multiple atmospheric models are interconnected. Here also, the atmospheric fluxes are shared with one ocean model.

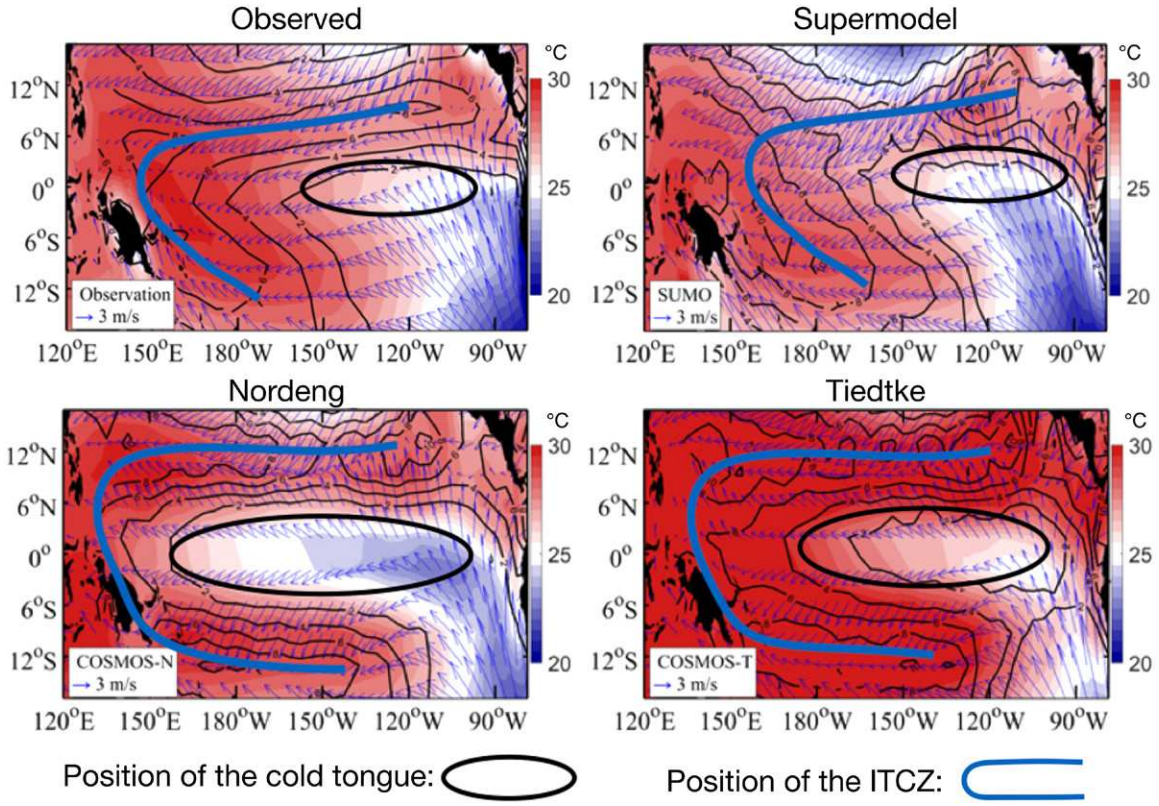

Fig. 6. SST (shading; °C), precipitation (contours; mm day^{-1}), and 10-m wind climatology for (top left) the observations, (bottom left) the model with the Nordeng convection scheme, (bottom right) the model with the Tiedtke convection scheme, and (top right) the supermodel constructed from the Nordeng and Tiedtke models. Schematically indicated are the Pacific cold tongue and ITCZ. The Nordeng and Tiedtke models show the double-ITCZ model biases. Figure adapted from Shen et al. (2017).

the development of the large-scale biases (Shen et al. 2017). This level of error reduction is not possible by any weighted mean of the unconnected Nordeng and Tiedtke model solutions (Shen et al. 2016). The implication is that nonlinear effects of the interactions between the models during the run of the supermodel impacted the final result. The combination of fluxes impacts the ocean and in return the ocean impacts the atmosphere in a way that could not have been achieved by a linear combination of models a posteriori.

The effects of these interactions were particularly evident in the tropical Pacific, where the atmosphere and ocean are subject to stronger interplay. To have effective interactions between the models outside the tropics, the models should exchange more information with each other, beyond ocean–atmosphere fluxes. The models could use this information to improve the synchronization of their states. A natural next step is to exchange variables between the atmospheric models themselves, as in Fig. 5b.

Training of intermodel connections

Efficient training methods. An essential component of the supermodel approach is the training of connections, the second key element of the supermodel concept (Fig. 1). Training ensures that the supermodel performs at least as well as any of the individual models and maximizes the expected improvement. A supermodel is a new dynamical system, and therefore can display unexpected behavior if not properly trained.

A standard training approach such as minimization of a cost function requires many model simulations, which is computationally very expensive. It also requires a huge amount of observational data, which are not always available. Therefore, two new methods were developed to efficiently train supermodels. The first method (Schevenhoven and Selten 2017)

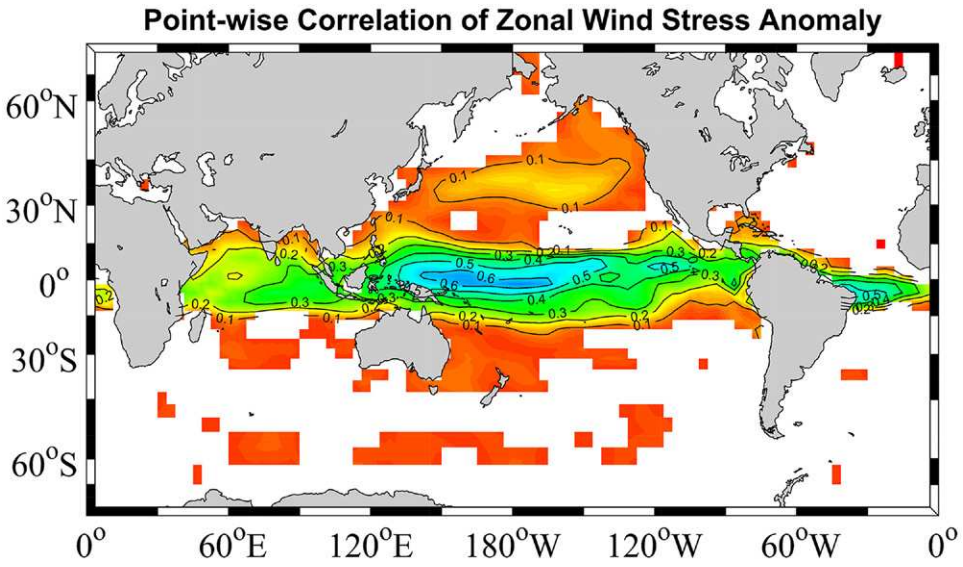

Fig. 7. Pointwise correlation of zonal wind stress between two connected AGCMs. Areas with insignificant correlation are blank. Approximate synchronization was found over the tropical Pacific. There is no significant correlation between the unconnected AGCMs (not shown) (Shen et al. 2016, supporting information).

is based on an idea called cross pollination in time (CPT) (Smith 2001), where models exchange their states during the training. The second method (Duane 2015; Selten et al. 2017) is a synchronization-based learning rule (synch rule), originally developed for parameter estimation (Duane et al. 2007). Both methods worked very well in the context of supermodels built from climate models of intermediate complexity, such as the SPEEDO model (Selten et al. 2017; Schevenhoven et al. 2019; Schevenhoven and Carrassi 2022). Both methods are online learning methods, meaning that the connection terms between the models are adjusted while the models are running. As in other fitting approaches, in supermodeling we train and evaluate on distinct sets of observations.

The two new methods are very efficient, because they are designed to keep the models in the vicinity of the observations by synchronizing them with the observations during training. This limits the search area for optimized connection terms compared to standard minimization methods. Both methods only required a 1-yr training simulation in order to train the SPEEDO climate supermodel. This short training period was sufficient to improve both the short-term forecasts and long-term climate predictions as compared to the individual models in a perfect model setting. This makes the methods promising for training state-of-the-art models, since the cost of training is comparable to a few short model runs. Often short-term training can reduce climatological errors as these tend to develop within the first 10 days of a forecast (Rodwell and Palmer 2007). However, training for short-term improvement does not always imply improved long-term behavior (Wiegerinck and Selten 2017).

Improvement in dynamics by training. A well-trained supermodel displays improved dynamics compared to the individual models. Selten et al. (2017) presented an example of a supermodel constructed from two versions of the global coupled atmosphere–ocean–land model SPEEDO (Severijns and Hazeleger 2010), to demonstrate supermodeling in the presence of parametric error in the imperfect models. Two SPEEDY atmosphere models are connected to each other and to the ocean model CLIO and the LBM land model, as in the setup of Fig. 5b. The climatology of the supermodel is better than that of either individual model, for example, in terms of temperature, precipitation, wind, and cloud cover (not shown). The supermodel also simulates the response to climate change (characterized by CO_2 doubling) more realistically, i.e., closer to the model we define as "perfect," given by certain set parameter

values (Fig. 8). In this perfect model, there is a clear shift in the east–west component of the wind at 850 hPa (Fig. 8a). The Southern Hemispheric jet strengthens, while the wind on the Maritime Continent decreases. The imperfect models are able to show the shift in the Southern Hemisphere, but do not see the same change on the Maritime Continent as the perfect model (Figs. 8c,d). The supermodel, however, is able to capture the changes in both regions (Fig. 8b). Although the supermodel has been trained on the basis of past observations from the perfect model, the dynamics of the supermodel have improved as compared to the individual models. A multimodel mean of the imperfect models with positive global weights would lead to a larger global mean error than the supermodel, since the errors of the individual imperfect models have a similar spatial structure and have the same sign.

An ensemble of different models

To create a supermodel, the individual models need to be able to compensate for each other's dynamical imperfections: the third element of the supermodel concept (Fig. 1). Formally the models should define a convex hull around the true dynamics. In case the models differ simply in parameter values, the convex hull can be described as an envelope of imperfect parameter values, containing the true parameter values. This is the case, for example, in the Lorenz 63 experiment of van den Berge et al. (2011) (Table 1). Plotting the parameter values shows that they form an envelope around the perfect parameter values (Fig. 9).

We further illustrate the convex hull principle using SPEEDO, with four imperfect versions that overestimate global temperature compared to the "perfect" model with standard parameter values (positive biases). Nevertheless, since the imperfect model parameters formed a convex hull around the true parameters, we can obtain a supermodel with an

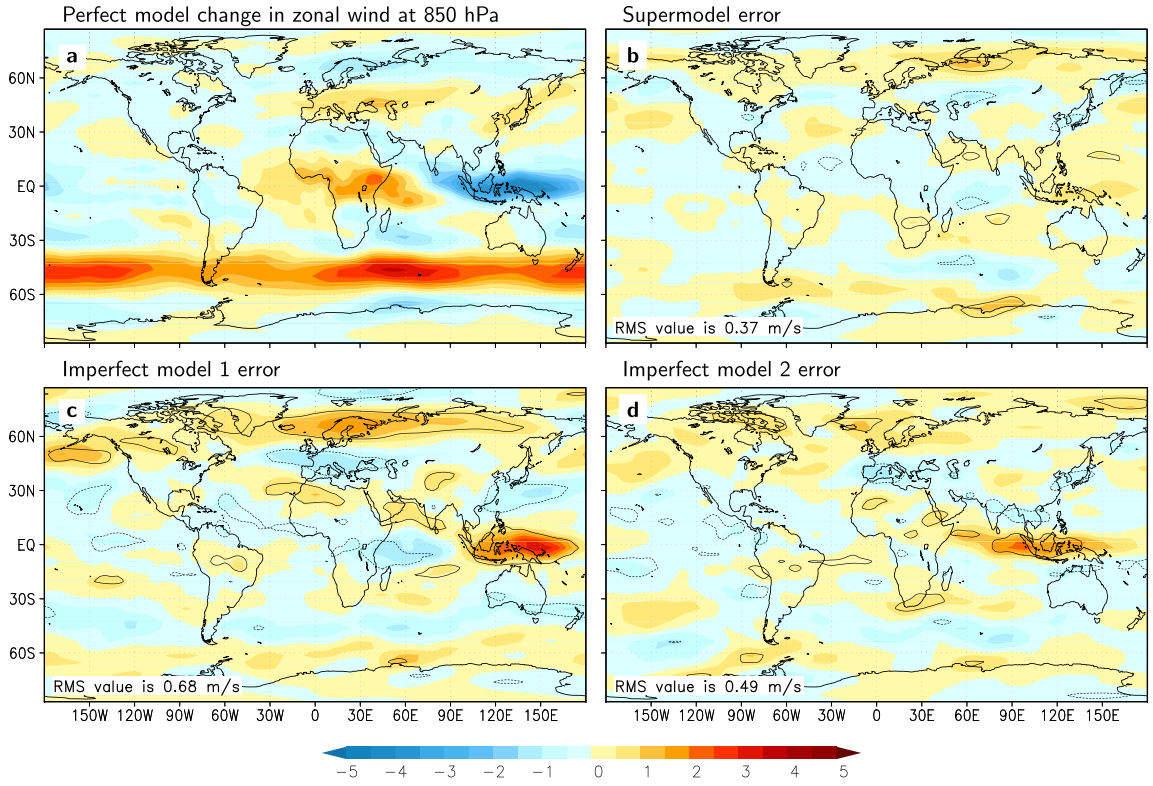

Fig. 8. (a) Change in the east–west component of the wind at 850 hPa due to a CO_2 doubling in the perfect model and the error in the simulated wind change for (b) the supermodel and for (c),(d) the two imperfect models. The change is calculated by subtracting the average wind before CO_2 doubling during model years 2016–40 from the average wind during 2056–70. The contours indicate regions where the difference is statistically significant at the 95% confidence level. The root of the global mean squared error is given in the lower-left corner of each panel (Selten et al. 2017).

average temperature much closer to the perfect model. This is because the imperfect model tendencies are able to compensate for each other's errors (Fig. 10) (Schevenhoven et al. 2019). Had those same imperfect models been included in an MME, the solution would not have been closer to the perfect model than the closest individual model.

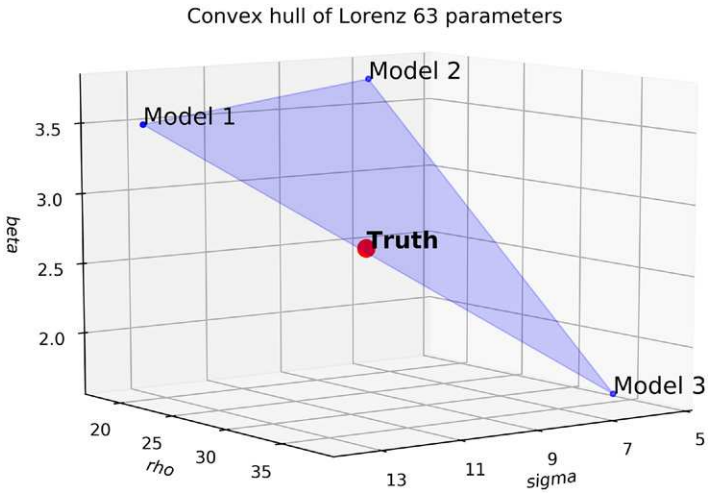

Fig. 9. Convex hull for the parameter values of the imperfect models 1–3, as used in the Lorenz 63 supermodel of van den Berge et al. (2011). The imperfect parameter values surround the perfect parameter values, defining the "perfect model."

If models differ not only in parameter values but also in structure (i.e., they have different equations, state vectors, and phase space), the parametric convex hull will not ensure having a better supermodel. In this case, the imperfect models themselves need to form a "convex hull," defined in a more general way. In other words, their dynamical behavior should envelop the true dynamics so that the imperfect model tendencies are able to compensate for each other. Interestingly, to form a convex hull, adding a "bad" model can create a superior supermodel, as long as it has dynamical behavior that complements the other models (Schevenhoven and Selten 2017).

Current developments and challenges

One of the key challenges in supermodeling is to combine models with different architectures. Recently, Counillon et al. (2023) have introduced a supermodeling framework for Earth system models (ESMs) based on data assimilation technology. It was used to connect the Norwegian (NorESM), the Community (CESM), and the Max Planck Institute (MPI-ESM) ESMs in their CMIP5 configurations via ocean components (Fig. 11). Differences in the grid type, resolution, and coordinate systems (Table 1 in Counillon et al. 2023) are handled using a methodology comparable to that of Du and Smith (2017): 1) constructing pseudo observations from the individual models (their weighted mean) on a common grid and 2) assimilating them back into the individual models to propagate the information to the unobserved state variables dynamically (Carrassi et al. 2018). In consideration of the practical limitations on the volume and frequency of the data exchanged, only the oceans of the ESMs are connected, on a monthly basis,

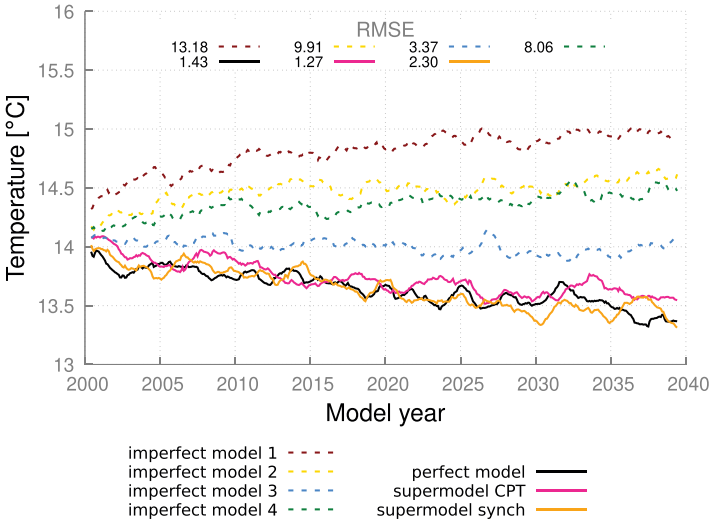

Fig. 10. Global mean time series for surface air temperature for the perfect model, the imperfect models, and the two supermodels trained by CPT and the synch rule. The normalized RMSE in the climatology of model years 2011–40 with respect to the climatology of the perfect model is given. The normalization is such that the expected value of the perfect model error is 1. Note that also the perfect model exhibits error when started from a perturbed initial condition. Figure adapted from Schevenhoven et al. (2019).

by connecting their SSTs. Note that the oceans are here distinct, unlike the earlier examples of Shen et al. (2016) and Selten et al. (2017). The pseudo observations are assimilated into the models using model specific ensemble optimal interpolation data assimilation methods (Counillon and Bertino 2009) that update the full ocean states in their native coordinate systems. The computational cost of the supermodel framework is increased by 28% compared to that of running the three models independently. Still, this increased cost can be brought down to less than 10% by parallelizing the synchronization step.

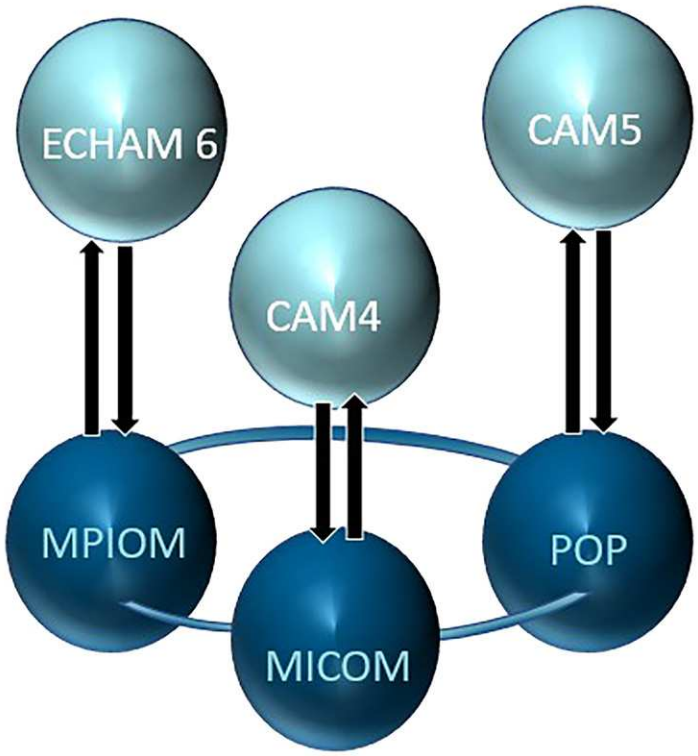

Fig. 11. Schematic of an ocean-connected supermodel based on state-of-the art Earth system models (Counillon et al. 2023). The oceans exchange state information during the simulations.

Here, we present results from the trained supermodel [unlike in Counillon et al. (2023), where no training was done] based on these three ESMs with ocean connections as described above. The supermodel is compared to the noninteractive (NI) ensemble. The weight of each model used to construct the pseudo observation should ideally adjust recursively as the minimization process that brings the supermodel to its optimal regime is generally nonlinear. However, in these experiments, weights are estimated on the basis of the monthly climatological SST bias of each of the models, calculated against NOAA Optimum Interpolation (OI) SST version 2 (OISSTV2; Reynolds et al. 2002) for the period 1982–2005. The triplet of weights (one for each model) is estimated at each pseudo-observation grid cell and for each calendar month; their sum is 1, and they vary smoothly in space and time (not shown).

The variability of the supermodel (Fig. 12c) is damped because the three ESMs weakly synchronize in the current configuration—only the oceans are connected every month (Counillon et al. 2023). The SST variability of the supermodel is comparable to that of the NI ensemble mean (Fig. 12b). In regions with strong ocean–atmosphere interactions, such as the equatorial Pacific, the supermodel achieves a fair degree of synchronization and an improved SST variability (Fig. 12c). In the equatorial Pacific, the band of high variability in the NI ensemble extends too far to the west and is too low (Fig. 12b). On the other hand, in the supermodel (Fig. 12c) it has an extent comparable to that of the observed band, although it is still too weak. In the North Atlantic Subpolar Gyre (SPG) the supermodel variability is substantially enhanced and more realistic than in the NI ensemble, as SST assimilation can effectively constrain ocean dynamics (Counillon et al. 2016) and synchronization can be achieved.

The supermodel is able to reduce long-standing climate model biases in SST and precipitation. This is shown in Fig. 13 for the period 2006–21, a different time span than the one used for training the weights (1980–2005). The SST bias of the supermodel is reduced over most regions compared to the NI ensemble mean (Fig. 13). All individual models show a pronounced double ITCZ (Tian and Dong 2020), also seen in the NI multimodel mean (Fig. 14b). In agreement with Shen et al. (2017) (Fig. 6), supermodeling suppresses this

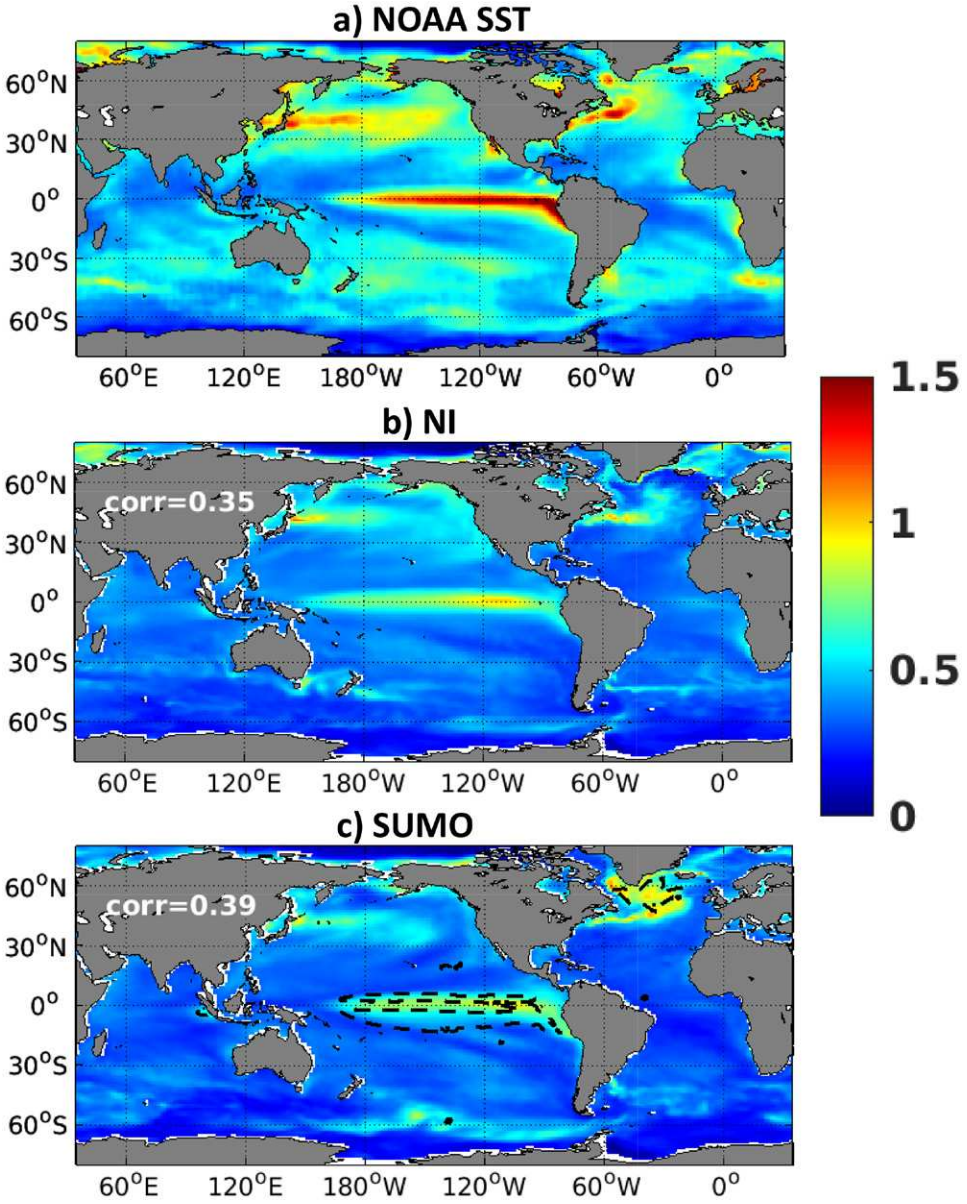

Fig. 12. (a) Deseasoned time standard deviation of SST in the NOAA OISST2 observations, (b) the multimodel average of NI SST and that of (c) the supermodel SST over the period 1982–2021. The values of the spatial correlation with observations are given in white. The black dashed lines highlight regions where synchronization is achieved, i.e., where the deseasoned time standard deviation of the multimodel mean is twice the time average of the monthly deseasoned intermodel spread—see Counillon et al. (2023) for details.

bias (Fig. 14c) and there is good agreement with the climatological observed pattern (cf. Figs. 14a and 14c). There is still some disagreement with the observations—with a too intense and wide northern precipitation band—but the aforementioned achievements are encouraging considering the rather premature stage of development of the system and pragmatic choices made in the implementation. Compared to Counillon et al. (2023), we have shown that training can reduce the rainfall biases.

Our current agenda for further development of the supermodel will allow us to substantially enhance its performance globally. We foresee that increasing the frequency of the information exchange, connecting atmospheres as well as other components of the ESMs, and improving the methodology of the training of the connections would improve the supermodel performance and enable a better representation of diverse phenomena (Du and Smith 2017; Schevenhoven and Carrassi 2022; Bach and Ghil 2022).

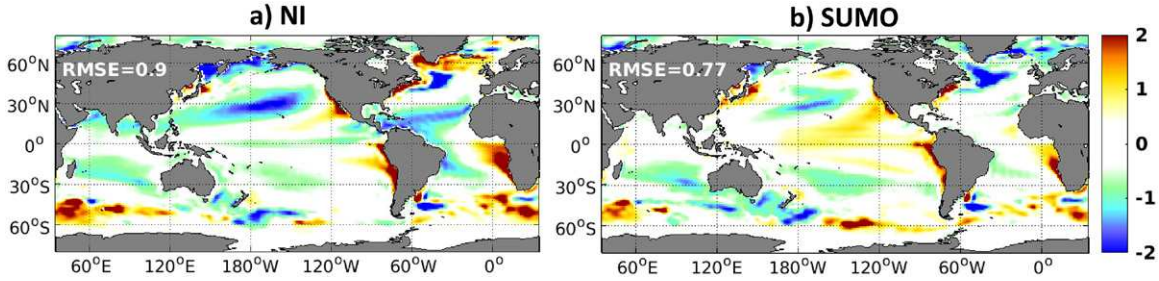

Fig. 13. Climatological SST bias of the multimodel mean in (a) the NI ensemble multimodel mean and (b) the supermodel against OISST2 observations over the period 2006–21. The quantity in white is the global spatial RMSE normalized by grid cell area.

Outlook and conclusions

In summary, the past decade has seen the development of supermodeling—a new framework for modeling Earth's climate in which models are combined interactively to improve forecasts and reduce systematic errors. Several examples have demonstrated the advantages of supermodeling over the standard multimodel ensemble in mitigating model biases. Supermodeling can be viewed as an advanced physics-informed machine learning approach. It takes advantage of existing models based on physical laws and applies machine learning techniques (such as CPT and the synch rule introduced in the section "Training of intermodel connections") to train the combination of models based on observational data.

Supermodeling is unique in the sense that an ensemble of different models is integrated in time simultaneously and state information is exchanged during the simulation. The resulting synchronized solution depends on the specifics of the interactions. By training these interactions, using historical observations, it is possible to achieve a solution that better matches the observed evolution. The trained supermodel has smaller errors compared to the individual models in the ensemble and as compared to any average of their outputs. The technique improves the dynamics of the resolved scales and better represents the interaction between the resolved and unresolved scales. Supermodeling is in principle not much more expensive than running a multimodel ensemble. There is an added cost in the training of the connections between the models, but this needs to be done only once, and efficient training schemes have been developed.

Other methods aimed at reducing model error have been developed such as parameter optimization techniques (e.g., Ruckstuhl and Janjić 2018) or stochastic parameterizations (Berner et al. 2017). The main difference with supermodeling is that these methods mostly focus on reducing the error of a single model. Supermodeling, on the other hand, rather than

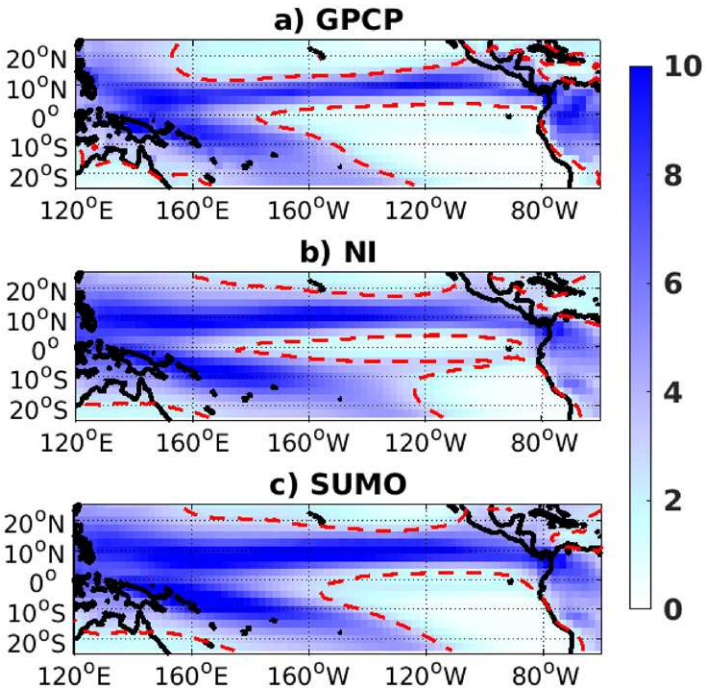

Fig. 14. Mean of precipitation in the tropical band over 2006–21 in the (a) GPCP observations, (b) NI ensemble multimodel mean, and (c) supermodel. The 2.5-mm isoline (dashed red line) delimits the ITCZ.

correcting the error in the individual models, tries to take maximum advantage of an ensemble of imperfect models: each of them will have something useful to contribute to the skill of the supermodel. Supermodeling can therefore be seen as an additional step to improve forecasts: if an individual model is improved, the supermodel can also be improved.

There are also other methods to correct biases in single models, such as flux correction and anomaly coupling (Sausen et al. 1988; Kirtman et al. 1997; Danforth et al. 2007; Kröger and Kucharski 2011; Toniazzo and Koseki 2018). These approaches apply static corrections (e.g., to the surface energy budget) to maintain the modeled climate close to observations, rather than correcting the source of these errors. These corrections can inhibit important climate feedbacks, leading to unstable model behavior (e.g., Gent 2018). Supermodeling focuses on reducing the development of errors, which occur within the first couple of weeks of the forecast (Jung et al. 2005; Knutti et al. 2010), before they develop into large-scale climate model biases. As such it not only maintains key climate feedbacks, but it aims at improving them.

There are multiple technical challenges with supermodeling. The models need to exchange information while running, and this currently limits supermodeling to a single computer. Further developments to connect models over the cloud would make the supermodel more widespread, as one could directly capitalize on existing models and data assimilation rather than porting them to a single high-performance computer. Another technical challenge is that the state representation differs between models: for instance, the vertical and horizontal discretization of the atmosphere in state-of-the-art climate models is typically different. These issues can in principle be dealt with by interpolation, data assimilation, or projection techniques, as in operational numerical weather prediction systems (e.g., Magnusson et al. 2022). Another venue for potential improvement is to use machine learning during the data assimilation step to emulate the missing processes in the pseudo observations (Brajard et al. 2020; Sonnewald et al. 2021).

The results shown in this paper give us the confidence that supermodeling is suitable to apply to state-of-the-art models and can truly mitigate long-standing biases and improve weather and climate predictions. Furthermore, supermodeling is not restricted to weather and climate modeling. It is also applicable to other areas of research that involve computational modeling of complex systems, such as hydrology (Santos 2018), ecological economy (Sendera 2019), and medicine (Paszyński et al. 2022).

Acknowledgments. We would like to acknowledge the efforts and contributions of the members of the EU-funded project SUMO led by Ljupco Kocarev. This work was supported by the EU Project 648982—STERCP, NSF Project 2015618—Coherent Precipitation Extremes in a Supermodel of Future Climate, ERC PoC Grant 101101037, EU Horizon Impetus4Change (101081555), and the RCN under Project 310098. AC acknowledges support of the project SASIP funded by Schmidt Futures—a philanthropic initiative that seeks to improve societal outcomes through the development of emerging science and technologies. This work has also received a grant for computer time (Projects nn9207k, nn9385k, and nn9039k) as well as storage space (Project ns9207k) from the Norwegian research infrastructure services.

Data availability statement. The full monthly output simulation of the three ESMs with ocean connections is accessible from https://archive.sigma2.no/pages/public/datasetDetail.jsf?id=10.11582/2023.00114.

References

Bach, E., and M. Ghil, 2022: A multi-model ensemble Kalman filter for data assimilation and forecasting. arXiv, 2202.02272v2, https://doi.org/10.48550/arXiv.2202.02272.

Bauer, P., A. Thorpe, and G. Brunet, 2015: The quiet revolution of numerical weather prediction. *Nature*, **525**, 47–55, https://doi.org/10.1038/nature14956.

Berner, J., and Coauthors, 2017: Stochastic parameterization: Toward a new view of weather and climate models. *Bull. Amer. Meteor. Soc.*, **98**, 565–588, https://doi.org/10.1175/BAMS-D-15-00268.1.

Bock, L., and Coauthors, 2020: Quantifying progress across different CMIP phases with the ESMValTool. *J. Geophys. Res. Atmos.*, **125**, e2019JD032321, https://doi.org/10.1029/2019JD032321.

Brajard, J., A. Carrassi, M. Bocquet, and L. Bertino, 2020: Combining data assimilation and machine learning to emulate a dynamical model from sparse and noisy observations: A case study with the Lorenz 96 model. *J. Comput. Sci.*, **44**, 101171, https://doi.org/10.1016/j.jocs.2020.101171.

Carrassi, A., M. Bocquet, L. Bertino, and G. Evensen, 2018: Data assimilation in the geosciences: An overview of methods, issues, and perspectives. *Wiley Interdiscip. Rev.: Climate Change*, **9**, e535, https://doi.org/10.1002/wcc.535.

Chapman, W. E., L. D. Monache, S. Alessandrini, A. C. Subramanian, F. M. Ralph, S.-P. Xie, S. Lerch, and N. Hayatbini, 2022: Probabilistic predictions from deterministic atmospheric river forecasts with deep learning. *Mon. Wea. Rev.*, **150**, 215–234, https://doi.org/10.1175/MWR-D-21-0106.1.

Cheng, S., and Coauthors, 2023: Machine learning with data assimilation and uncertainty quantification for dynamical systems: A review. arXiv, 2303.10462v1, https://doi.org/10.48550/arXiv.2303.10462.

Counillon, F., and L. Bertino, 2009: Ensemble optimal interpolation: Multivariate properties in the Gulf of Mexico. *Tellus*, **61A**, 296–308, https://doi.org/10.1111/j.1600-0870.2008.00383.x.

——, N. Keenlyside, I. Bethke, Y. Wang, S. Billeau, M. L. Shen, and M. Bentsen, 2016: Flow-dependent assimilation of sea surface temperature in isopycnal coordinates with the Norwegian Climate Prediction Model. *Tellus*, **68A**, 32437, https://doi.org/10.3402/tellusa.v68.32437.

——, ——, S. Wang, M. Devilliers, A. Gupta, S. Koseki, and M.-L. Shen, 2023: Framework for an ocean-connected supermodel of the Earth system. *J. Adv. Model. Earth Syst.*, **15**, e2022MS003310, https://doi.org/10.1029/2022MS003310.

Danforth, C. M., E. Kalnay, and T. Miyoshi, 2007: Estimating and correcting global weather model error. *Mon. Wea. Rev.*, **135**, 281–299, https://doi.org/10.1175/MWR3289.1.

Davini, P., and F. D'Andrea, 2016: Northern Hemisphere atmospheric blocking representation in global climate models: Twenty years of improvements? *J. Climate*, **29**, 8823–8840, https://doi.org/10.1175/JCLI-D-16-0242.1.

Du, H., and L. A. Smith, 2017: Multi-model cross-pollination in time. *Physica D*, **353–354**, 31–38, https://doi.org/10.1016/j.physd.2017.06.001.

Duane, G. S., 2015: Synchronicity from synchronized chaos. *Entropy*, **17**, 1701–1733, https://doi.org/10.3390/e17041701.

——, J. J. Tribbia, and J. B. Weiss, 2006: Synchronicity in predictive modelling: A new view of data assimilation. *Nonlinear Processes Geophys.*, **13**, 601–612, https://doi.org/10.5194/npg-13-601-2006.

——, D. Yu, and L. Kocarev, 2007: Identical synchronization, with translation invariance, implies parameter estimation. *Phys. Lett.*, **371A**, 416–420, https://doi.org/10.1016/j.physleta.2007.06.059.

——, J. Tribbia, and B. Kirtman, 2009: Consensus on long-range prediction by adaptive synchronization of models. *EGU General Assembly 2009*, Vienna, Austria, EGU, Abstract 13324, https://meetingorganizer.copernicus.org/EGU2009/EGU2009-13324-1.pdf.

Gent, P. R., 2018: A commentary on the Atlantic meridional overturning circulation stability in climate models. *Ocean Modell.*, **122**, 57–66, https://doi.org/10.1016/j.ocemod.2017.12.006.

Hagedorn, R., F. J. Doblas-Reyes, and T. Palmer, 2005: The rationale behind the success of multi-model ensembles in seasonal forecasting—I. Basic concept. *Tellus*, **57A**, 219–233, https://doi.org/10.3402/tellusa.v57i3.14657.

He, K., X. Zhang, S. Ren, and J. Sun, 2016: Deep residual learning for image recognition. *2016 IEEE Conf. on Computer Vision and Pattern Recognition*, Los Alamitos, CA, IEEE, 770–778, https://doi.org/10.1109/CVPR.2016.90.

Jakhar, K., Y. Guan, R. Mojgani, A. Chattopadhyay, P. Hassanzadeh, and L. Zanna, 2023: Learning closed-form equations for subgrid-scale closures from high-fidelity data: Promises and challenges. arXiv, 2306.05014v1, https://doi.org/10.48550/arXiv.2306.05014.

Jung, T., A. M. Tompkins, and M. J. Rodwell, 2005: Some aspects of systematic error in the ECMWF model. *Atmos. Sci. Lett.*, **6**, 133–139, https://doi.org/10.1002/asl.105.

Kirtman, B. P., J. Shukla, B. Huang, Z. Zhu, and E. K. Schneider, 1997: Multiseasonal predictions with a coupled tropical ocean–global atmosphere system. *Mon. Wea. Rev.*, **125**, 789–808, https://doi.org/10.1175/1520-0493(1997)125<0789:MPWACT>2.0.CO;2.

——, D. Min, P. S. Schopf, and E. K. Schneider, 2003: A new approach for coupled GCM sensitivity studies. COLA Tech. Rep. 154, 50 pp.

Knutti, R., R. Furrer, C. Tebaldi, J. Cermak, and G. A. Meehl, 2010: Challenges in combining projections from multiple climate models. *J. Climate*, **23**, 2739–2758, https://doi.org/10.1175/2009JCLI3361.1.

Krishnamurti, T. N., V. Kumar, A. Simon, A. Bhardwaj, T. Ghosh, and R. Ross, 2016: A review of multimodel superensemble forecasting for weather, seasonal climate, and hurricanes. *Rev. Geophys.*, **54**, 336–377, https://doi.org/10.1002/2015RG000513.

Kröger, J., and F. Kucharski, 2011: Sensitivity of ENSO characteristics to a new interactive flux correction scheme in a coupled GCM. *Climate Dyn.*, **36**, 119–137, https://doi.org/10.1007/s00382-010-0759-5.

Lee, J.-Y., and Coauthors, 2021: Future global climate: Scenario-based projections and near-term information supplementary material. *Climate Change 2021: The Physical Science Basis*, V. Masson-Delmotte et al., Eds., Cambridge University Press, 553–672.

Lorenz, E., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.*, **20**, 130–141, https://doi.org/10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2.

Magnusson, L., and Coauthors, 2022: Skill of medium-range forecast models using the same initial conditions. *Bull. Amer. Meteor. Soc.*, **103**, E2050–E2068, https://doi.org/10.1175/BAMS-D-21-0234.1.

Nam, C., S. Bony, J.-L. Dufresne, and H. Chepfer, 2012: The 'too few, too bright' tropical low-cloud problem in CMIP5 models. *Geophys. Res. Lett.*, **39**, L21801, https://doi.org/10.1029/2012GL053421.

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. ECMWF Tech. Memo. 206, 41 pp., www.ecmwf.int/en/elibrary/75843-extended-versions-convective-parametrization-scheme-ecmwf-and-their-impact-mean.

Paszyński, M., L. Siwik, W. Dzwinel, and K. Pingali, 2022: Supermodeling, a convergent data assimilation meta-procedure used in simulation of tumor progression. *Comput. Math. Appl.*, **113**, 214–224, https://doi.org/10.1016/j.camwa.2022.03.025.

Pecora, L. M., and T. L. Carroll, 2015: Synchronization of chaotic systems. *Chaos*, **25**, 097611, https://doi.org/10.1063/1.4917383.

——, ——, G. A. Johnson, D. J. Mar, and J. F. Heagy, 1997: Fundamentals of synchronization in chaotic systems, concepts, and applications. *Chaos*, **7**, 520–543, https://doi.org/10.1063/1.166278.

Reichler, T., and J. Kim, 2008: How well do coupled models simulate today's climate? *Bull. Amer. Meteor. Soc.*, **89**, 303–312, https://doi.org/10.1175/BAMS-89-3-303.

Reynolds, R.-W., N.-A. Rayner, T.-M. Smith, D.-C. Stokes, and W. Wang, 2002: An improved in situ and satellite SST analysis for climate. *J. Climate*, **15**, 1609–1625, https://doi.org/10.1175/1520-0442(2002)015<1609:AIISAS>2.0.CO;2.

Richter, I., 2015: Climate model biases in the eastern tropical oceans: Causes, impacts and ways forward. *Wiley Interdiscip. Rev.: Climate Change*, **6**, 345–358, https://doi.org/10.1002/wcc.338.

Rodwell, M. J., and T. N. Palmer, 2007: Using numerical weather prediction to assess climate models. *Quart. J. Roy. Meteor. Soc.*, **133**, 129–146, https://doi.org/10.1002/qj.23.

Ruckstuhl, Y. M., and T. Janjić, 2018: Parameter and state estimation with ensemble Kalman filter based algorithms for convective-scale applications. *Quart. J. Roy. Meteor. Soc.*, **144**, 826–841, https://doi.org/10.1002/qj.3257.

Rulkov, N. F., M. M. Sushchik, L. S. Tsimring, and H. D. I. Abarbanel, 1995: Generalized synchronization of chaos in directionally coupled chaotic systems. *Phys. Rev.*, **51E**, 980–994, https://doi.org/10.1103/PhysRevE.51.980.

Santos, L., 2018: Perspectives for hydrological supermodels: Evaluation of a method based on a dynamical combination of rainfall-runoff models. Ph.D. thesis, L'Institut des Sciences et Industries du Vivant et de l'Environnement, 256 pp.

Sausen, R., K. Barthel, and K. Hasselmann, 1988: Coupled ocean-atmosphere models with flux correction. *Climate Dyn.*, **2**, 145–163, https://doi.org/10.1007/BF01053472.

Schevenhoven, F. J., and F. M. Selten, 2017: An efficient training scheme for supermodels. *Earth Syst. Dyn.*, **8**, 429–438, https://doi.org/10.5194/esd-8-429-2017.

——, and A. Carrassi, 2022: Training a supermodel with noisy and sparse observations: A case study with CPT and the synch rule on SPEEDO—v.1. *Geosci. Model Dev.*, **15**, 3831–3844, https://doi.org/10.5194/gmd-15-3831-2022.

——, F. Selten, A. Carrassi, and N. Keenlyside, 2019: Improving weather and climate predictions by training of supermodels. *Earth Syst. Dyn.*, **10**, 789–807, https://doi.org/10.5194/esd-10-789-2019.

Schrauwen, B., D. Verstraeten, and J. Van Campenhout, 2007: An overview of reservoir computing: Theory, applications and implementations. *Proc. 15th European Symp. on Artificial Neural Networks*, Bruges, Belgium, ESANN, 471–482, www.esann.org/sites/default/files/proceedings/legacy/es2007-8.pdf.

Selten, F. M., F. J. Schevenhoven, and G. S. Duane, 2017: Simulating climate with a synchronization-based supermodel. *Chaos*, **27**, 126903, https://doi.org/10.1063/1.4990721.

Sendera, M., 2019: Data adaptation in HANDY economy-ideology model. arXiv, 1904.04309v1, https://doi.org/10.48550/arXiv.1904.04309.

Severijns, C. A., and W. Hazeleger, 2010: The efficient global primitive equation climate model SPEEDO v2.0. *Geosci. Model Dev.*, **3**, 105–122, https://doi.org/10.5194/gmd-3-105-2010.

Shen, M.-L., N. Keenlyside, F. Selten, W. Wiegerinck, and G. S. Duane, 2016: Dynamically combining climate models to 'supermodel' the tropical Pacific. *Geophys. Res. Lett.*, **43**, 359–366, https://doi.org/10.1002/2015GL066562.

——, ——, B. C. Bhatt, and G. S. Duane, 2017: Role of atmosphere-ocean interactions in supermodeling the tropical Pacific climate. *Chaos*, **27**, 126704, https://doi.org/10.1063/1.4990713.

Smith, L. A., 2001: Disentangling uncertainty and error: On the predictability of nonlinear systems. *Nonlinear Dynamics and Statistics*, A. I. Mees, Ed., Birkhäuser, 31–64, https://doi.org/10.1007/978-1-4612-0177-9_2.

Sonnewald, M., R. Lguensat, D. C. Jones, P. Dueben, J. Brajard, and V. Balaji, 2021: Bridging observations, theory and numerical simulation of the ocean using machine learning. *Environ. Res. Lett.*, **16**, 073008, https://doi.org/10.1088/1748-9326/ac0eb0.

Stouffer, R. J., V. Eyring, G. A. Meehl, S. Bony, C. Senior, B. Stevens, and K. E. Taylor, 2017: CMIP5 scientific gaps and recommendations for CMIP6. *Bull. Amer. Meteor. Soc.*, **98**, 95–105, https://doi.org/10.1175/BAMS-D-15-00013.1.

Tian, B., and X. Dong, 2020: The double-ITCZ bias in CMIP3, CMIP5, and CMIP6 models based on annual mean precipitation. *Geophys. Res. Lett.*, **47**, e2020GL087232, https://doi.org/10.1029/2020GL087232.

Tiedtke, M., 1989: A comprehensive mass flux scheme for cumulus parameterization in large-scale models. *Mon. Wea. Rev.*, **117**, 1779–1800, https://doi.org/10.1175/1520-0493(1989)117<1779:ACMFSF>2.0.CO;2.

Toniazzo, T., and S. Koseki, 2018: A methodology for anomaly coupling in climate simulation. *J. Adv. Model. Earth Syst.*, **10**, 2061–2079, https://doi.org/10.1029/2018MS001288.

van den Berge, L. A., F. M. Selten, W. Wiegerinck, and G. S. Duane, 2011: A multi-model ensemble method that combines imperfect models through learning. *Earth Syst. Dyn.*, **2**, 161–177, https://doi.org/10.5194/esd-2-161-2011.

Watt-Meyer, O., N. D. Brenowitz, S. K. Clark, B. Henn, A. Kwa, J. McGibbon, W. A. Perkins, and C. S. Bretherton, 2021: Correcting weather and climate models by machine learning nudged historical simulations. *Geophys. Res. Lett.*, **48**, e2021GL092555, https://doi.org/10.1029/2021GL092555.

Wiegerinck, W., and F. M. Selten, 2017: Attractor learning in synchronized chaotic systems in the presence of unresolved scales. *Chaos*, **27**, 126901, https://doi.org/10.1063/1.4990660.

Yang, S.-C., and Coauthors, 2006: Data assimilation as synchronization of truth and model: Experiments with the three-variable Lorenz system. *J. Atmos. Sci.*, **63**, 2340–2354, https://doi.org/10.1175/JAS3739.1.