Earth Syst. Dynam. Discuss.,
doi:10.5194/esd-2017-6-RC2, 2017

Earth System Dynamics Discussions
Open Access
EGU

Interactive comment on "An efficient training scheme that improves the forecast skill of a supermodel" *by* Francine Schevenhoven and Frank Selten

G. Hu (Referee)

guannan.hu@studium.uni-hamburg.de

Received and published: 21 February 2017

This research article proposes a new method to determine the weights of imperfect models in a supermodel, and proves that such a supermodel outperforms every individual imperfect model. Further more the article also shows that supermodel using different weights performs better than using equal weights. The method used to determine weight is called Cross Pollination in Time (CPT). This is a good idea and satisfied results are obtained by applying this method in Lorenz 63 system and a quasi-geostrophic model. The article is clearly structured and the language is fluent and concise, but some parts of the article should be in more detail and clarified. Tables and figures are clear and relevant, although they are not so plentiful. I recommend that this

Printer-friendly version

Discussion paper

article needs minor revisions.

Major comment: 1. Section 2 Training the supermodel: my understanding is that every time CPT takes place, it is counted which model is closest to the truth. If we have done 100 times CPT, then we have 100 predictions, and we count for each model in these 100 predictions, how often they have closest predictions to the truth. If I understand correctly, there are several questions. First is CPT interval time should be shown (how often does CPT take place?). Second is whether iterative method is necessary. What if increasing the training time steps or training the model from many different initial conditions. I think these two ways also make the results converge. 2. Page 4, Table 1: if the parameter values for the imperfect models are all smaller or greater than standard values, how well will the supermodel perform? 3. Page 10, Table 5 and page 4, Table 1: an important issue is how to set the parameter values of imperfect models and how to decide the number of imperfect models. These problems are not touched in the article. As shown in conclusion (page 13, line 8 and 9) the number of imperfect models is related to the number of uncertain parameters. The uncertain parameters in Lorenz 63 and quasi-geostrophic model are both 3, but imperfect models used to construct supermodel are 2 and 4, respectively. I wonder if the number of imperfect models depends on the degrees of freedom. As shown in Table 5 each two imperfect models have a same parameter value, why? Is it possible to randomly select parameter values for imperfect models with given error. 4. Page 12, 4.3 Use of fewer imperfect models: I am interested in the difference between supermodel without the worst imperfect model and supermodel without the best imperfect model. In such a way, it can be clarified whether the accuracy of imperfect model affects the performance of supermodel. Clearly, the number of imperfect models influences the skill of the supermodel. While the relationship is unclear. The differences between supermodel and truth as a function of number of imperfect models can used to illustrate this relationship.

Minor comment: 1. Abstract, page 1, line 7 and 8: the supermodel obtained by the CPT training should not only compared to the individual imperfect models, it is not the topic

of this article. As indicated by the title: the method improves the supermodel, so the comparison should between the supermodels with different weights. 2. Introduction: I'm wondering if there are still other methods have been used to determine the weights except the example in the second paragraph of page 2. And I also want to know the advantages of CPT, since the multi-model ensemble method (MME) can also enlarge the state space. 3. Page 4, line 12: why choose deviation 30%? 4. Page 5, Figure 2 and page 7, Figure 3: I suggest figure of 3D-phase space. 5. Page 7, 3.1 Climate measures: I suggest using probability density function and autocorrelation function to compare the supermodel and truth, not only the mean, standard deviation and covariance. 6. Page 9, Figure 5: (a) of Figure 5 is not necessary. I suggest combine (a) and (b) of Figure 5. 7. Page 9, line 15 and page 10, line 1: in the experiments with Lorenz 63, the model 1 has two stable fixed points and very different behavior compared to standard Lorenz 63 system. But it still contributes to form supermodel. 8. Page 11, Table 7: I suggest add tables for other height (200hPa, 500hPa) and in table supermodel with same weights should be included. 9. Page 11, line 10: what is the minimal time steps for training? If the training time steps increases, does the weights change? 10. The supermodel with different weights performs better than supermodel with equal weights, but it requires computational time. Increasing the number of imperfect models also improves the skill of the supermodel with equal weights. So we need to compare in which way the computational demand is lower in certain application.

Typos and misspelling: 1. Page 1, end of line 8: "state-of-the art", "-" before "art" is missing 2. Page 1, line 16: add "," before because 3. Page 4, title of section 3: "Results Lorenz 63", "for" missing

---

Interactive comment on Earth Syst. Dynam. Discuss., doi:10.5194/esd-2017-6, 2017.