The paper applied two approaches, CPT and synchronization based, to “combine” multiple imperfect models to form a so called supermodel to noisy and sparse observations. The general concept is quite interesting and potentially could be very valuable to multi-model forecasting. The authors inherited previous works on CPT and synch rule, and relaxed noise-free assumption, proposed adaptations and demonstrated in SPEEDO experiments. Overall I think this paper is publishable with some revisions. Please find my concerns listed below.

1) It seems that the “imperfect” scenario described in the paper (for example section 2) are limited in perfect model class with imperfect model parameters. Is the proposed methodology applicable to a more general scenario where the model class is imperfect, i.e. there is no parameter value would lead to a perfect model.

2) I am strongly against the idea (Line 28,44) that the benefit of using MME is due to “errors tend to cancel each other”, this misinterpretation is a result of only considering the ensemble mean. Note that the purpose of using the ensemble forecasts (Leutbecher and Palmer (2008)) is to account for uncertainty.

3) The authors proposed the possibility of negative weights when form the supermodel. Personally I am a little concern about this approach, especially how would one interpret the negative weights? Can we first bias correct the models (for example Line 158, models 1 and before combine them?

4) It is a little unclear what exactly is a supermodel. is the sentence at (Line 35) a formal definition of the "supermodel"?

5) Line 98-99, it seems that achieving a synchronized state is a good thing, please define what is a synchronized supermodel state and clarify why it is desirable.

6) Line 87, “observations were available”, this statement needs to be clarified as the observations were not mention early or in the equations 1a-c.

7) Section 3.3, please give more details about how the CPT trajectories are generated.

8) Paragraph at Line 240, It seems that the authors suggest that when the models are linear, the CPT approach works well. Why it doesn’t if the model is highly nonlinear?

9) Table 4 and Figure 3 shows the weights for supermodel trained by CPT and synch rule, it would be interesting to see the forecast performance of CPT compared with synch rule.