Interactive comment on "Optimal model complexity for terrestrial carbon cycle prediction" by Caroline A. Famiglietti et al.

Enqing Hou (Referee)

enqing.hou@nau.edu

Famiglietti and colleagues explored relationship between model complexity and forecast skill either with or without assimilated data using a data assimilation system. The authors found that without assimilated data, a complex model has a poorer forecast skill than a simple model; with assimilated data, the opposite is true. The findings make sense and highlight the importance of using data to inform model before forecasting. The manuscript is very interesting and well written. I have only a few minor concerns about the manuscript below. L230-232: Will there be any difference in key results and conclusion obtained between using the histogram interaction and using the more familiar metrics? L233-239: n value (number of bins) used is? L405: "assimilate diverse data types" operates blindly. Some datasets are more useful to constrain

a specific variable than other datasets. We can do better than just "diverse". Table 1. Explain the meanings of the IDs (e.g., C groups, S groups). Why the sub-models ordered in the current way in the Table? Fig. 5a: "(a) All runs", do you mean all runs without assimilating data? You may make it clearer. Fig. 6: arrangement of the panels are not in a good logic to me. Probably as (a) None ... (f) NEE, LAI, biomass. Description of skill metric and complexity metric as well as the model structure are clear, but deposit the code to produce this manuscript could be more helpful to others to use the approaches here.

Interactive comment on Biogeosciences Discuss., https://doi.org/10.5194/bg-2020-478, 2020.