The paper devoted to analysis of quality of terrestrial carbon cycle simulation using Earth System Models from CMIP6. Improvements of CMIP6 models comparing to CMIP5 and empirical datasets are shown. Data compared using a set of statistical parameters and colorful maps. Methods and the aim of the paper are clear.

Specific comments

Soil carbon storage, net primary productivity and carbon turnover time were selected as variables responsible for terrestrial soil carbon estimations. According to suggestions NPP related with soil carbon through plant ang root litter (line 30-35), but empirical datasets have negligible correlation between these values (line 458). Pleas, give more attention for the support of your idea on relations of soil carbon and NPP.

Carbon turnover time determined as a ratio of carbon amount and heterotrophic respiration. According to presented results soil carbon estimations were improved in CMIP6 comparing CMIP5, but soul carbon turnover time estimations is not good enough. Likely the issue is related with heterotrophic respiration. Could you check the hypothesis and present an analysis of quality of HR simulations?

Changes in soil carbon storage occurs through changes in fluxes. The accuracy of simulation of carbon fluxes will result in total estimations of soil carbon. You have shown only one flux (NPP) not directly related with soil system and give a complex parameter related with heterotrophic respiration. Is it possible to demonstrate the quality of simulations of carbon fluxes relates with soil system (i.e. heterotrophic respiration, ecosystem respiration, dissolved carbon runoff, decay rate, litterfall, etc)
The paper contains a lot of statistical information about comparison of results from CMIP6/5 ESMs. Total estimations and spatial variability of parameters are shown. But the meaning of obtained estimations and relations with land ecosystem is missed. In the present form the paper is more suitable for Geoscientific Model Development journal where ESM and their characteristics are discussed. Understanding of reasons of ESM errors requires identification of an ecosystem types where highest discrepancies observed. Clear, that highest soil carbon is typical for peatlands. Proper simulation of peatland water, thermal and nutrient regime will giver more impact to the global carbon estimations than for other ecosystems. I suggest to emphasize the role of ecosystems in soil carbon formation and discuss the errors and improvements of ESMs not only at global scale but at ecosystem scale too.