Response to Referee #3

I observe that the authors paid much attention to my previous comments and did a great job to integrate my remarks into the new version of the manuscript. The majority of my concerns have been addressed. There are still a few minor improvements required. Overall, the manuscript improved considerably. I suggest it for publication after minor revisions.

Response: We would like to thank the reviewer for the positive comments and the time invested to review our manuscript again. The revised manuscript will follow the reviewer's recommendations.

Specific comments:

1. Please add the description of the workflow (Fig. 2) to clarify the main steps in the 2.2 section.

Response: the description of the workflow is clarified in the 2.2. section: 'Typically, the flow of the NEE prediction modeling framework (Fig. 2) based on flux observations and machine learning is as follows: first, half-hourly scale NEE flux observations are aggregated into various time scale NEE data, and gap-filling techniques (Moffat et al., 2007) are often used in this step to obtain complete NEE series when data are missing. Various predictors including meteorological variables, remote sensing-based biophysical variables, etc. are extracted to match site-scale NEE series to generate a training dataset containing the target variable NEE and various covariates. Subsequently, various algorithms are used for the NEE prediction model construction and validated in different ways (e.g., leave-one-site-out validation (Zeng et al., 2020)). Finally, in some studies, prediction models were applied on gridded covariate data to map the regional or global-scale NEE spatial and temporal variations (Zeng et al., 2020; Papale and Valentini, 2003; Jung et al., 2020). The information of R-squared (at the validation phase) and the associated model features reported in the article are considered as one data record for the formal meta-analysis (i.e., each R-squared record corresponding to a prediction model). From the included papers, R-squared records and various features (Table 2) involved in the NEE modeling framework (Fig. 2) were extracted (including the used algorithms, modeling/validation methods, remote sensing data, meteorological data, biophysical data, and ancillary data). In some studies, multiple algorithms were applied to the same dataset, or models with different features were developed (Virkkala et al., 2021; Zhang et al., 2021; Cleverly et al., 2020; Tramontana et al., 2016). In these cases, multiple data records will be documented.' (Line 151)

2. The R2-based comparison with keeping other variables constant may cause potential uncertainty. Need to add statements and instructions in the Discussion section.

Response: elaborated in the Discussion section as 'We should pay more attention to the effect of features on model accuracy individually in future studies, and it may be valuable to keep other features as constants while changing the level of only one feature and assessing the difference. It may help us to understand the real sensitivity of model accuracy to different features in specific conditions.' (Line 480)

3. Please add an explanation for the reason why MLR, RF, SVM, and ANN are separately compared instead of all models and why PLSR with high R2 is removed in the Methodology

section. This response is similar to the response for L198.

Response: elaborated as 'Subsequently, the model accuracies corresponding to different levels of various features are compared in a cross-study fashion. In the evaluation of algorithms and time scales, we also implement comparisons within individual studies. For example, in the evaluation of the effects of the algorithms, we compare the accuracy of models using the same training data and keeping other features as constants in individual studies. In this intra-study comparison step, only algorithms with relatively large sample sizes in the cross-study comparisons were selected.' (Line 188)

References

Cleverly, J., Vote, C., Isaac, P., Ewenz, C., Harahap, M., Beringer, J., Campbell, D. I., Daly, E., Eamus, D., He, L., Hunt, J., Grace, P., Hutley, L. B., Laubach, J., McCaskill, M., Rowlings, D., Rutledge Jonker, S., Schipper, L. A., Schroder, I., Teodosio, B., Yu, Q., Ward, P. R., Walker, J. P., Webb, J. A., and Grover, S. P. P.: Carbon, water and energy fluxes in agricultural systems of Australia and New Zealand, 287, https://doi.org/10.1016/j.agrformet.2020.107934, 2020.

Jung, M., Schwalm, C., Migliavacca, M., Walther, S., Camps-Valls, G., Koirala, S., Anthoni, P., Besnard, S., Bodesheim, P., Carvalhais, N., Chevallier, F., Gans, F., S Goll, D., Haverd, V., Köhler, P., Ichii, K., K Jain, A., Liu, J., Lombardozzi, D., E M S Nabel, J., A Nelson, J., O'Sullivan, M., Pallandt, M., Papale, D., Peters, W., Pongratz, J., Rödenbeck, C., Sitch, S., Tramontana, G., Walker, A., Weber, U., and Reichstein, M.: Scaling carbon fluxes from eddy covariance sites to globe: Synthesis and evaluation of the FLUXCOM approach, 17, 1343–1365, https://doi.org/10.5194/bg-17-1343-2020, 2020.

Moffat, A. M., Papale, D., Reichstein, M., Hollinger, D. Y., Richardson, A. D., Barr, A. G., Beckstein, C., Braswell, B. H., Churkina, G., Desai, A. R., Falge, E., Gove, J. H., Heimann, M., Hui, D., Jarvis, A. J., Kattge, J., Noormets, A., and Stauch, V. J.: Comprehensive comparison of gap-filling techniques for eddy covariance net carbon fluxes, 147, 209–232, https://doi.org/10.1016/j.agrformet.2007.08.011, 2007.

Papale, D. and Valentini, R.: A new assessment of European forests carbon exchanges by eddy fluxes and artificial neural network spatialization, 9, 525–535, https://doi.org/10.1046/j.1365-2486.2003.00609.x, 2003.

Tramontana, G., Jung, M., Schwalm, C. R., Ichii, K., Camps-Valls, G., Ráduly, B., Reichstein, M., Arain, M. A., Cescatti, A., Kiely, G., Merbold, L., Serrano-Ortiz, P., Sickert, S., Wolf, S., and Papale, D.: Predicting carbon dioxide and energy fluxes across global FLUXNET sites with regression algorithms, Biogeosciences, 13, 4291–4313, https://doi.org/10.5194/bg-13-4291-2016, 2016.

Virkkala, A.-M., Aalto, J., Rogers, B. M., Tagesson, T., Treat, C. C., Natali, S. M., Watts, J. D., Potter, S., Lehtonen, A., Mauritz, M., Schuur, E. A. G., Kochendorfer, J., Zona, D., Oechel, W., Kobayashi, H., Humphreys, E., Goeckede, M., Iwata, H., Lafleur, P. M., Euskirchen, E. S., Bokhorst, S., Marushchak, M., Martikainen, P. J., Elberling, B., Voigt, C., Biasi, C., Sonnentag, O., Parmentier, F.-J. W., Ueyama, M., Celis, G., St.Louis, V. L., Emmerton, C. A., Peichl, M., Chi, J., Järveoja, J., Nilsson, M. B., Oberbauer, S. F., Torn, M. S., Park, S.-J., Dolman, H., Mammarella, I., Chae, N., Poyatos, R., López-Blanco, E., Christensen, T. R., Kwon, M. J., Sachs, T., Holl, D., and Luoto, M.: Statistical upscaling of ecosystem CO2 fluxes across the terrestrial tundra and boreal domain: Regional patterns and uncertainties, Global Change Biology, 27, 4040–4059, https://doi.org/10.1111/gcb.15659, 2021.

Zeng, J., Matsunaga, T., Tan, Z.-H., Saigusa, N., Shirai, T., Tang, Y., Peng, S., and Fukuda, Y.: Global terrestrial carbon fluxes of 1999–2019 estimated by upscaling eddy covariance data with a random forest, 7, https://doi.org/10.1038/s41597-020-00653-5, 2020.

Zhang, C., Brodylo, D., Sirianni, M. J., Li, T., Comas, X., Douglas, T. A., and Starr, G.: Mapping CO_2 fluxes of cypress swamp and marshes in the Greater Everglades using eddy covariance measurements and Landsat data, Remote Sensing of Environment, 262, https://doi.org/10.1016/j.rse.2021.112523, 2021.