Response to referee comments: Anonymous Referee #2

I only have one minor remaining question. For forecasting, there won't be any available measurements to determine the dominant model among the forecasting results from different models. Under this condition, your model should only rely on the previously training data to do the forecast. From your response, you mentioned that a dynamic strategy for choosing the dominant model for a specific location and time will be used for forecasting. I believe you used the strategy for validation, but I am not clear about the strategy. Has this strategy been described in your manuscript?

Response: Thank you for this important question. We interpret the term "forecasting" here as referring to the validation phase and the global application for ET estimation, which corresponds to scenarios where no ET measurements are available to determine the dominant model.

Yes, in our validation and application procedure, ET measurements were strictly excluded from the model selection process. Instead, we employed a separately trained classifier (a machine learning classification model) for dynamic model selection. During the training phase, the dominant model at each spatiotemporal point was identified based on its performance against ET measurements. These dominant model labels were then used as the training target for the classifier, with meteorological covariates as predictors. For consistency, our classifier uses the same input covariates as the other ML-based ET estimation models. It's important to note that its training target is not the ET value itself, but the category of the dominant model at each spatiotemporal point (i.e., three classes: 'ML model-dominated', 'hybrid model-dominated', and 'process-based algorithm-dominated'). In this way, the classifier learned the relationship between input covariates and the dominant model category. Finally, during validation and application, the classifier can dynamically select the dominant model using only input covariates, without the need for ET observations. The selected model is then applied to generate the ET estimate.

While Section 2.4 introduces this methodology, we realize that the original description was not sufficiently clear regarding this mechanism. We have revised the text around Line 170 to explicitly explain the training process as follows:

In the training phase, we utilized the classification results (i.e., the identified dominant model) as the training target for the ML classifier, using the same input covariates as the other ET models in this study. Through this process, the classifier learned the relationship between input covariates and the dominant model labels. Notably, after the training phase was completed, the classifier could operate independently of ET data. Therefore, during validation and application, we employed the classifier to identify the corresponding dominant model label at each spatiotemporal point based on the input covariates without requiring observed ET data. Subsequently,

the identified dominant model at each spatiotemporal point was utilized to generate the ET estimates.