Response to Topic editor

Q1. Lack of Explanation of FLAML Model Structures: From the current description, it is only mentioned that FLAML uses an automated framework (AutoML) for building the models, but there is no description on the type of models that are tested, and the type of models that were selected by the procedure. The article only mentions the variables that were selected, but not the type of model found by the AutoML. I believe this is an important piece of information missing from the entire article. Right now, the article only focuses on evaluating model performance through a set of metrics, but there are no new insights regarding the value of FLAML in selecting certain types (structures) of models. I recommend the authors to better describe the type of models that are tested by FLAML in section 2.3.2. The focus should be on the structural characteristics of the model, not on the predictive variables, which are well explained in section 2.3.3. and do not need further elaboration.

We sincerely appreciate your valuable comments. In response to the two main concerns raised, we have revised the manuscript accordingly and provide the detailed replies below.

Firstly, the lack of structural description of the models tested by FLAML. Accordingly, we have revised Section 2.3.2 to provide more detailed explanations of the candidate model types used by FLAML, focusing on their structural characteristics and differences in learning strategy, randomness, and ensemble behavior. The below changes have been made to Section 2.3.2 (Lines 265 – 287):

"For our regression tasks, AutoML was configured with the "auto" option for the estimator list, focusing on optimizing the R^2 metric and using a time budget of 120 seconds per run. Under this "auto" setting, FLAML explores a variety of built-in regression estimators, including:

1. LightGBM (Ke et al., 2017): a histogram-based gradient boosting method designed for speed and scalability;
2. XGBoost (Chen and Guestrin, 2016): a regularized gradient boosting framework known for its robustness and accuracy;
3. CatBoost (Prokhorenkova et al., 2018): efficiently handles categorical features and reduces overfitting via ordered boosting;
4. Random Forest (Breiman, 2001): an ensemble method utilizing bootstrap aggregation of decision trees;
5. Extra Trees (Geurts et al., 2006): enhances randomness in split point selection for tree construction;
6. Histogram-based Gradient Boosting (Brownlee, 2020), accelerate training through feature binning;
7. K-Nearest Neighbors (Cover and Hart, 1967): a non-parametric distance-based algorithm relying on local data density;
8. Transformer models (Vaswani et al., 2023), deep learning architectures leveraging self-attention mechanisms, adapted here for structured data regression.

Collectively, these estimators span a broad algorithmic spectrum, including ensemble learning, distance-based methods, and neural networks, enabling FLAML to automatically identify the optimal model architecture for the dataset and objective."

Secondly, to address the issue that the manuscript previously did not report which models AutoML ultimately selected, we have added new content in Section 3.1 (Lines 351 – 360):

"To evaluate the model performance at the site level, the accuracy of the 18 FLAML-LUE models was assessed using test datasets from individual flux tower sites. The algorithms selected by each FLAML-LUE model are listed in **Table S1**. Notably, the Extra-Trees algorithm was most frequently chosen as the best-performing model. Extra Trees is an ensemble method that constructs multiple unpruned decision trees and introduces high randomness in both feature and threshold selection, which enhances generalization and reduces overfitting, particularly in noisy or high-dimensional datasets. The consistent selection of Extra Trees suggests that FLAML tends to favor models with higher stochasticity and ensemble structures under the given data and computational constraints."

We believe these revisions adequately address the reviewers' concerns by clarifying the range of candidate models tested and by highlighting FLAML's model selection process in our study.

Table S1

Optimal algorithms of the FLAML-LUE Model for GPP Simulation Across Different Temporal Scales and Predictor Combinations.

FLAML	Daily	8-day	16-day	Monthly
FLAML00	extra_tree	extra_tree	extra_tree	extra_tree
FLAML01	extra_tree	extra_tree	extra_tree	extra_tree
FLAML02	extra_tree	extra_tree	extra_tree	extra_tree
FLAML03	extra_tree	extra_tree	xgb_limitdepth	extra_tree
FLAML04	rf	rf	extra_tree	extra_tree
FLAML05	extra_tree	extra_tree	extra_tree	extra_tree
FLAML10	extra_tree	extra_tree	extra_tree	extra_tree
FLAML11	extra_tree	rf	extra_tree	extra_tree
FLAML12	extra_tree	extra_tree	extra_tree	extra_tree
FLAML13	extra_tree	extra_tree	extra_tree	extra_tree
FLAML14	lgbm	extra_tree	extra_tree	extra_tree
FLAML15	extra_tree	extra_tree	extra_tree	extra_tree
FLAML20	extra_tree	extra_tree	extra_tree	extra_tree
FLAML21	lgbm	extra_tree	extra_tree	extra_tree
FLAML22	extra_tree	extra_tree	extra_tree	extra_tree
FLAML23	extra_tree	extra_tree	extra_tree	extra_tree
FLAML24	rf	rf	rf	extra_tree
FLAML25	extra_tree	extra_tree	extra_tree	extra_tree

Q2. Insufficient Discussion on the Value of AutoML (FLAML): The discussion section should include something about the structure of the selected models and why the AutoML procedure has advantages over other ML methods.

Thank you for your valuable suggestion. In the revised Discussion section, we have further expanded on the structure of the models selected by FLAML and emphasized the advantages of using AutoML. Specifically, we highlight FLAML's ability in conducting automatic search across a diverse set of model families and configurations, adaptively selecting optimal models based on their performance metrics, and minimizing the need for manual trial-and-error. These features make FLAML a practical and efficient tool for environmental modeling tasks. Specifically, we have added the following text to the revised manuscript (Lines 788 – 810):

"In this study, FLAML (Wang et al., 2021) selected the Extra Trees algorithm as the best-performing model for GPP simulation in China. Extra Trees is an ensemble learning method that builds multiple unpruned decision trees and incorporates randomization in features selection and split thresholds determination. Compared to traditional decision tree ensembles such as Random Forests, Extra Trees typically achieves minimal variance while maintaining low bias, which makes it particularly well-suited for complex, high-dimensional datasets (Geurts et al., 2006).

The adoption of FLAML provides several significant advantages. First, it automates the model selection and hyperparameter tuning process, eliminating the need for extensive manual trial-and-error and reducing reliance on domain expertise (Nakano and Liu, 2025; Wang et al., 2022). Instead of manually evaluating various algorithms and their configurations, FLAML efficiently explores a broad search space and identifies the most appropriate model for the dataset.

Moreover, FLAML employ a cost-aware hyperparameter optimization strategy, enabling it to find high-performing models with relatively low computational cost (Zhang et al., 2023; Wang et al., 2021). This feature is particularly advantageous in scenarios with limited computational resources or the need for rapid prototyping.

Compared to conventional machine learning workflows, FLAML significantly reduces human bias in model selection, improves reproducibility, and lowers the barrier to applying advanced modeling techniques (He et al., 2021). Overall, the use of FLAML in this study not only improved model performance but also streamlined the modeling process, supporting its broader applicability in ecological and climate-related research."

Q3. Inadequate Clarification on LUE Model Use: Reviewer 2 in Q2 made an important comment on the use of LUE in the text but not using an explicit LUE model. The answer you provided is a good answer, but I didn't see an explicit text in the manuscript on this topic. In your response you mentioned that lines 122 and 272 include this information, but the text in the revised version is too vague. Please include a similar text as what you provided in your answer to reviewer 2 in section 2.3.3.

Thank you for pointing this out. In response to your comment, we have revised Section 2.3.3 of the manuscript to provide a clearer and more explicit explanation of how LUE theory is incorporated into our modeling approach. Specifically, we have added the following text (Lines 289－299):

“Eighteen FLAML-LUE model variations were constructed for all sites by combining different permutations of six input factor groups, as described in Eq. (3) and detailed in Table 3. Technically, the term "FLAML-LUE" does not refer to a direct implementation of a mechanistic LUE model. Instead, it reflects a hybrid modeling strategy, through which we incorporate key explanatory variables that originate from LUE theory—such as fPAR, light-use efficiency modifiers, and environmental stress indicators (e.g., VPD, temperature, and water stress indices)—into an automated machine learning framework (FLAML). These variables capture the main drivers of vegetation productivity in traditional LUE models. Their integration enables FLAML to build models that are both ecologically grounded and predictive, effectively balancing model interpretability and accuracy.”

We sincerely appreciate the editor’s constructive suggestions, which have substantially improved the clarity and rigor of our manuscript. We hope that the revised version now meets the expectations for publication.