

EcoPro-LSTM v0 : A Memory-based Machine Learning Approach to Predicting Ecosystem Dynamics across Time Scales in Mediterranean Environments

Mitra Cattry, Wenli Zhao, Juan Nathaniel, Jinghao Qiu, Yao Zhang, and Pierre Gentine

Please find our response to reviewer #2 comment. Our response is printed in blue while the comments are printed in black. The line numbers refer to the new version without track changes if not specified.

Dear Authors, Dear Editor,

find below my review. Please excuse the delay.

Summary:

This work introduces an LSTM network for simultaneously modeling four ecosystem variables measured at FluxNet sites: RECO, GPP, ET & SWC. A dataset of 17 FLUXNET sites in mediterranean climate is assembled from the FLUXNET2015 dataset. The LSTM model is then trained in 5-fold cross validation using temporal splitting (i.e. ensuring that the test data comes from a different time period than the train data). The resulting model performs well at capturing seasonal dynamics. Finally, the models learned sensitivities to input variables are investigated by computing Shapley values and integrated gradients.

Strengths:

I appreciate the effort to improve inter-annual variability of data-driven carbon flux estimates with deep neural networks. In addition, I enjoyed reading about the limitations of the two different interpretability methods, and in particular the analysis of correlated drivers and how due to them the applicability of SHAP may be hampered. A further strength of this paper is the inclusion of estimates of epistemic uncertainty through MC-Dropout, which highlights a large variability beyond the mean prediction.

Thank you for your thoughtful and constructive review. We appreciate your recognition of our efforts to improve inter-annual variability in carbon flux estimates using deep neural networks, as well as the insights on the limitations of SHAP versus integrated gradients. We're also grateful for your positive feedback on our incorporation of MC-Dropout for uncertainty integration.

Major:

1. I am unsure what benefit the presented EcoPro-LSTM brings. Since its training is on temporal splits, it appears to me to only be suitable for the 17 sites that it has been trained on in this study. The model would not be suitable for generalizing to other locations (within mediterranean ecosystems), and thus in particular not be suitable for generating a global map. See also Meyer & Pebesma for a related discussion <https://www.nature.com/articles/s41467-022-29838-9>
2. It remains unclear how this model compares against the state of the art. The comparisons against FluxCom (v1 and X-Base) are unfair, as cross validation sets are not chosen identical. E.g. While in Figure 9, the EcoPro-LSTM has been trained on all sites (just during different years), the

FluxCom and X-Base models have not seen any data from a particular site during training, but rather needed to extrapolate from other sites.

3. Hence the main claim "improved interannual variability" presented in Section 5.1 is in vain, as this improvement is not derived from generic understanding, but rather from learning site-specific patterns.

Thank you for this important point. We agree that it is essential to distinguish temporal improvements within the training sites from spatial generalization to new locations. We absolutely agree with the reviewer that our current assessment method evaluates the temporal rather than spatial generalization. In case, it is not clear, we have included this in the article. To ensure scientific clarity and soundness, we have revised the manuscript as follows:

- Clarified Scope in Introduction

“In this study, we introduce EcoPro-LSTMv0—a memory-infused LSTM architecture that improves prediction of sub-daily to interannual variability within a network of 17 Mediterranean FLUXNET towers. Spatial transferability to unseen sites is reserved for future work.”

- Included input data & improved code readability:

We have now included the input data on zenodo link, so you can directly paste it in input folder, and run the model. We further changed input names and included nomenclature to facilitate going through our code. In case you face any issues, please, do not hesitate to contact us through editor. The updated version will be available by Oct 12, 2025.

- Added “Next Steps”

Because our cross-validation employs only temporal splits, the spatial generalizability of EcoPro-LSTMv0 remains untested. Follow-up studies will apply leave-one-site-out validation and extend the model to larger FLUXNET datasets (all semi-arid regions) to assess broader applicability. The lines 504-14 (see line 656 in the version with tracked changes)

- Included additional analysis in the response letter to support soundness of our modelling assumptions.

Thank you for your thorough and insightful feedback. We appreciate your constructive suggestions and would like to clarify and respectfully address your points concerning the comparisons with FLUXCOM products and our claim regarding improved interannual variability.

Regarding your concerns about the fairness of comparison with FLUXCOM and FLUXCOM-X (X-base), we recognize and clearly acknowledge that our model employs temporal cross-validation splits, training on different years for the same set of 17 sites. In contrast, FLUXCOM and FLUXCOM-X models involve spatial extrapolation—predicting ecosystem dynamics at sites not included in their training set. We understand the methodological difference and explicitly stated in our revised manuscript that our current model setup assesses temporal rather than spatial generalization. However, we respectfully argue that this does not invalidate our findings regarding improved interannual variability.

Our central claim regarding enhanced interannual variability capture is founded on the inherent strengths of the EcoPro-LSTMv0:

1. **Time Dependencies:** LSTM architectures inherently capture long-term and short-term dependencies more effectively than traditional approaches, significantly enhancing temporal resolution of predictions. Previous literature (Besnard et al., 2019; Huang et al., 2024) supports this strength of LSTMs, specifically highlighting their superior ability in modeling ecosystem respiration over multiple months.
2. **Weighted RMSE (WRMSE):** Our use of weighted RMSE specifically addresses biases inherent to ecological datasets, which frequently display uneven data distributions and extreme values. This methodological choice ensures that our model is robust against skewed data distributions and thus contributes significantly to accurately capturing year-to-year variations.
3. **Multi-Timescale Integration:** Incorporating both hourly and daily data scales enables our model to leverage high-resolution dynamics (e.g., rapid responses to rainfall events) and low-resolution trends (seasonal and interannual fluctuations), which directly contributes to improved modeling of interannual variability.

Our detailed, per-site analysis (Tables 1, 2, B1, B2, and Figures 3, 9) supports these methodological choices by demonstrating consistent improvements in capturing seasonal and interannual fluctuations, surpassing the traditional FLUXCOM benchmarks within the limitations now clearly communicated in our manuscript.

Additionally, we transparently acknowledge the limitation concerning spatial extrapolation, explicitly stating that spatial generalization remains outside the scope of this particular study and is reserved for future research, as discussed under the newly added “Next Steps” section. Here, we clearly outline our intention to pursue leave-one-site-out validation and to incorporate mechanistic traits such as rooting depth, soil texture, and phenology, thereby addressing both spatial generalization and the representation of ecological mechanisms.

In summary, while our current evaluation method differs from FLUXCOM products' spatial extrapolation approach, our claims about improved temporal resolution and enhanced capture of interannual variability are rigorously supported by our methodological choices, previous literature, and explicit data-driven results provided in the manuscript. We hope this clarification addresses your concerns and respectfully reinforces our original contribution's validity and scientific soundness.

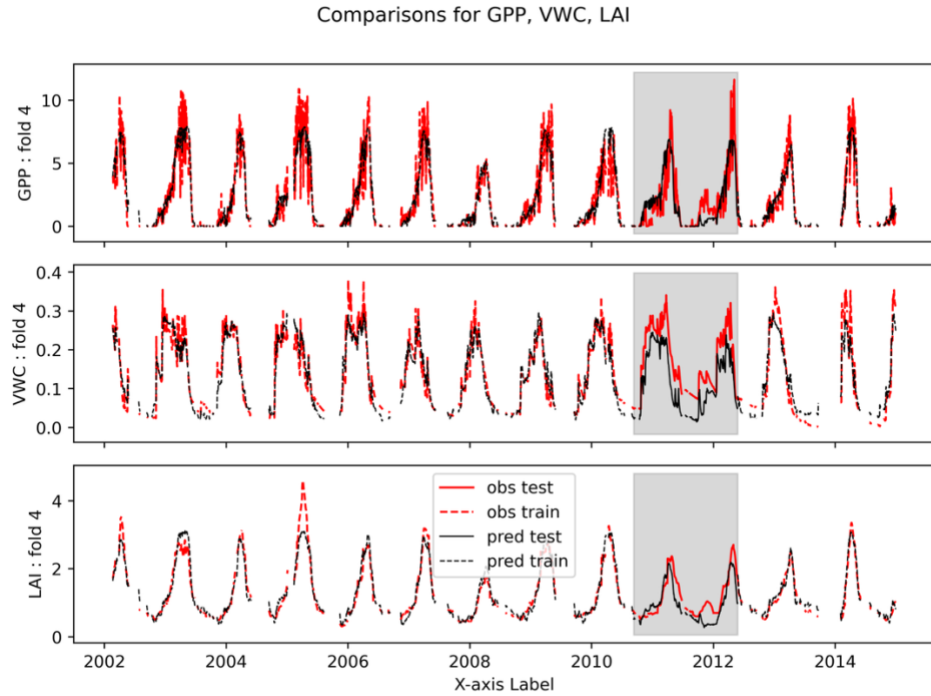
4. **No remote sensing predictors have been used.** However, many works have identified that for instance adding remotely sensed vegetation indices and LST can greatly enhance predictive performance. In fact, Kraft et al. <https://egusphere.copernicus.org/preprints/2024/egusphere-2024-2896/> argue that sequential deep learning models barely add any value beyond using remote sensing predictors.

The work by Kraft et al, 2024 and Martinuzzi et al 2023 are actually fantastic citations to emphasize that the temporal dynamics of simple LSTM model cannot capture extreme events. So now we have added these citations in lines 45-50 of our manuscript.

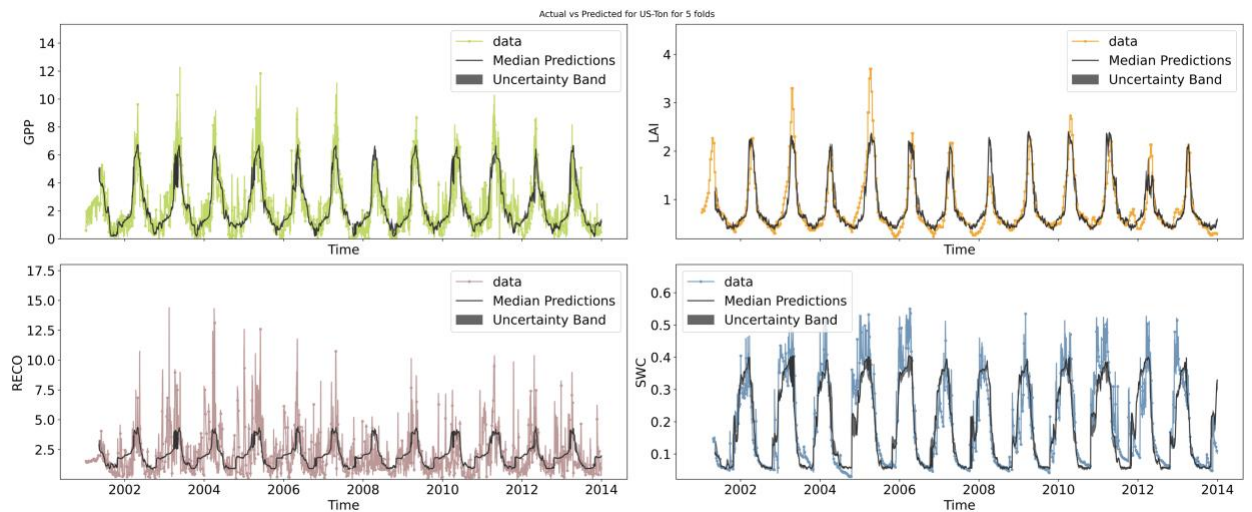
“And while Besnard et al. (2019); Huang et al. (2024) have compared tree-based against LSTM and demonstrated the superior performance of LSTM in predicting ecosystem respiration with a memory of up to six months, there is still a need for further model development for capturing rare and extreme ecosystem responses (Martinuzzi et al., 2023; Kraft et al., 2024).”

To show LAI did not help us improving capturing the extreme events against the multi-timescale architecture, below you can see GPP predictions when LAI is among the model outputs:

Fold 4: US-Var (gray dashed area indicates the test set)



Fold 1-5 for US-Ton



We also added the following text at lines 90-2 to reflect our discussion here (line 105 in the version with tracked changes):

“To test whether explicitly supplying leaf area index (LAI) would enhance the model’s memory of past conditions, we added LAI as both an input and an auxiliary output. However, this yielded no measurable improvement over the LSTM’s inherent memory capabilities.”

References:

1. Martinuzzi, F., Mahecha, M.D., Camps-Valls, G., Montero, D., Williams, T. and Mora, K., 2023. Learning extreme vegetation response to climate forcing: A comparison of recurrent neural network architectures. *EGUsphere*, 2023, pp.1-32.
<https://doi.org/10.5194/egusphere-2023-2368>
2. Kraft, B., Nelson, J.A., Walther, S., Gans, F., Weber, U., Duveiller, G., Reichstein, M., Zhang, W., Rußwurm, M., Tuia, D. and Körner, M., 2024. On the added value of sequential deep learning for upscaling evapotranspiration. *EGUsphere*, 2024, pp.1-30.
<https://egusphere.copernicus.org/preprints/2024/egusphere-2024-2896/>
5. Many modeling choices appear arbitrary, and thus would have to be ablated (show that they reduce the validation loss). For instance these include weighting the RMSE loss by target flux magnitude, modeling both temporal resolutions in a two-stage approach and the particular length of the temporal context windows (why not a full year e.g.?).

We have now expanded and explained our modelling choices in lines 130-135 (see line 150 in the version with tracked changes) and presented figures below in the response letter (see page 10 of this response letter). The choice of two timescale is the main components improving the prediction and it is well documented in literature just as you proposed before that a simple LSTM with RMSE or MAE fails to capture the interannual variability. The improvement made by the choice of WRMSE and two timescales.

6. Similarly, it would be necessary to properly ablate that multi-task-learning is in fact better than training 4 separate LSTM models, one per task.

Figure 10 directly compares the single-task learning (STL) approach with the multi-task learning (MTL) framework. While we do not observe a substantial change in predictive performance for GPP itself, MTL consistently improves the attribution of GPP variability to precipitation. This indicates that MTL captures cross-task relationships that single-task models overlook, leading to more physically interpretable and mechanistically meaningful feature attributions. The added interpretability — particularly the stronger and more consistent precipitation attribution across sites — is a key advantage of MTL, even in cases where raw predictive skill remains similar.

7. Since you train on only 17 sites, and include the site-class as a predictor, I would expect the model to perfectly be able to differentiate the different sites. Thus a meaningful baseline becomes a per-site-model, i.e. one that is trained only for one site. I'd be curious to see if your EcoPro-LSTM is able to outperform such a per-site-model.

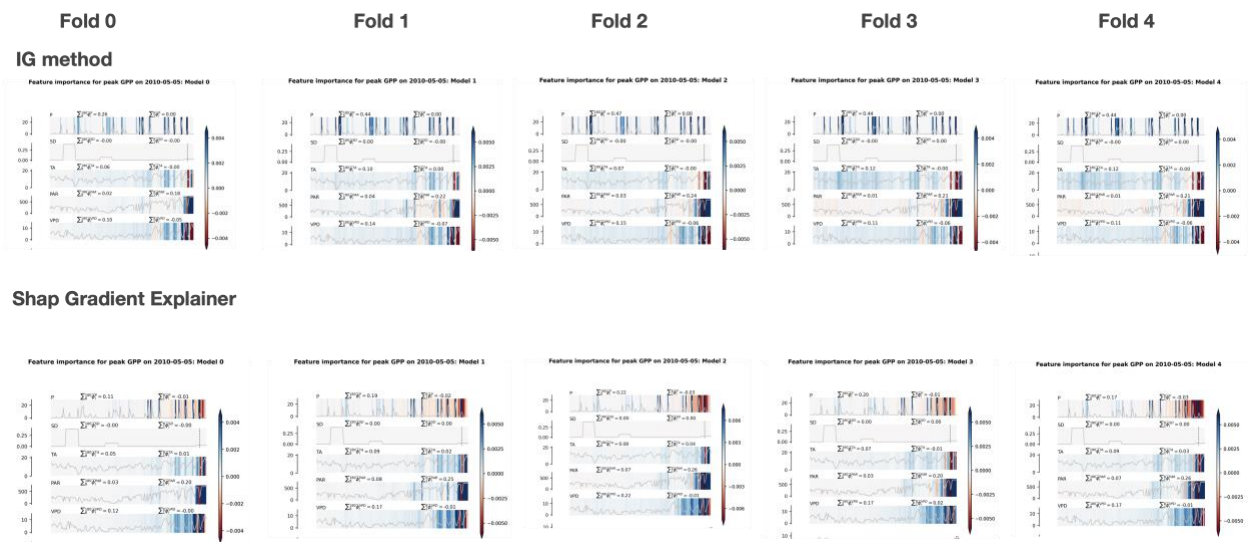
LSTM does not have an inherent sense of space, but rather of time. On the contrary to your claim, having only 17 points in space is never a good training set for a model to develop a sense of space even if we add another layer such as CNN or MLP to incorporate that feature. We have acknowledged this limitation in the article, at lines 505 and stated it is on our upcoming work under “Next Steps”. Thank you so much for bringing this to our attention.

8. How did you assess what is "the closest match to SHAP" in Fig. 5? It would be good to have a quantitative means to base this decision on.

Thank you for pointing out the lack of quantification. We have now moved this figure to the Supplementary Information. In our analysis, we referred primarily to visual comparisons and qualitative patterns rather than numerical estimates, as applying a single quantitative metric across many sites and dates would itself have been somewhat arbitrary and potentially misleading. Our aim was to illustrate the general consistency between methods rather than provide a one-to-one quantitative match. We now clarify this point in the text in line 233 (line 290 in the version with tracked changes).

9. L. 305f - To me it seems problematic to study integrated gradients for training set periods, and also to average over different models (especially if they use correlated drivers). This could result in multiple artifacts, such as spurious correlations solely due to over-fitting or canceling of contradicting explanations and thereby a misattribution of a given variable's contribution. In other words: please only plot explanations for individual models, and for test set data only.

Below you can see feature importance for one date 2010-05-05 per model (for each model, there is a different fold), and you can compare the method against Shap Gradient Explainer. As you can see: 1. IG method provides more stable estimates, 2: the pattern is very similar across folds. You are right as in Model 0, the test set, the feature importance attribution is slightly different, but overall, you can see the relative relation holds the same and the reason for this difference doesn't necessarily reflect one fold being test set, and could reflect computational error.



10. Please cite FluxNet data appropriately

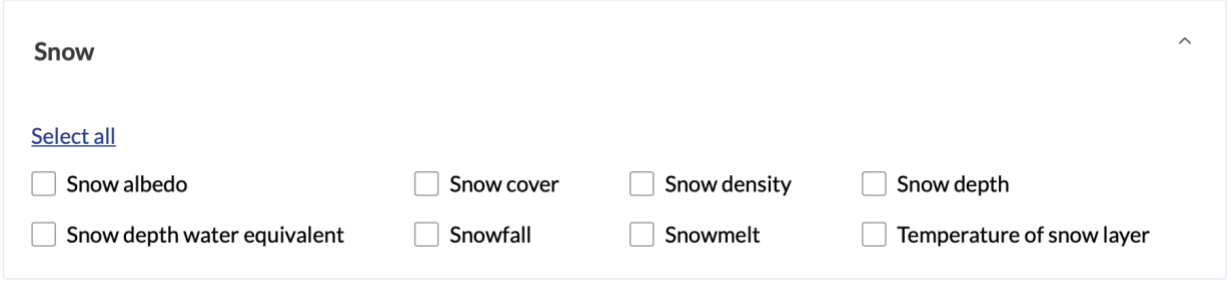
Corrected, thank you !

11. I could not run the code, because the data was not shared along with the code (maybe for licensing reasons?) - and there was no script or clear description provided on how to download the data. Please add this information, such that the work becomes reproducible.

We are sorry you could not run the code, to facilitate running the code for you, we have temporarily and while this code is under review prepared a zendo link (<https://zenodo.org/records/16323121>) to download the code. You need to download it and place it in the code folder and click run (to ensure folder placement is correct, or for further assist you can check: <https://ecoclimate-at-g-lab-columbia-university.github.io/MT-LSTM/>). If you face any issues, please, prior to posting a response on this platform contact us through the Editor, we will be more than happy to help. For easy follow up on parameters we have also included a nomenclature of what each single variable means in the site and inside the code “at appendix”. Thanks for your suggestion in improving the clarity and user-friendliness of our code.

Minor:

12. L. 64 "Snow depth data was retrieved hourly from the publicly available Copernicus platform" → please accurately say which product from which platform you are using.



The screenshot shows a web interface for selecting data products. At the top, the word "Snow" is displayed with an upward arrow icon. Below it, there is a link "Select all". A grid of checkboxes is shown, each followed by a data product name:

☐ Snow albedo	☐ Snow cover	☐ Snow density	☐ Snow depth
☐ Snow depth water equivalent	☐ Snowfall	☐ Snowmelt	☐ Temperature of snow layer

Above is a screen shot of data products from <https://cds.climate.copernicus.eu/datasets/reanalysis-era5-land?tab=download> as you can see “snow depth” is a data product. We rewrote this line as below to avoid any confusion

“Hourly snow-depth data were retrieved from the ERA5-Land reanalysis product via the Copernicus Climate Data Store (CDS; <https://cds.climate.copernicus.eu/>).”

References:

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H. and Martens, B., 2021. ERA5-Land: A state-of-the-art global reanalysis dataset for land applications. *Earth system science data*, 13(9), pp.4349-4383. <https://doi.org/10.5194/essd-13-4349-2021>

13 Please cite Gal et al. when you mention the use of MC-Dropout for uncertainty quantification L.108-110 <https://arxiv.org/abs/1506.02142>

Thank you for the prompt. While our implementation of dropout was developed independently, we have added the citation to Gal et al. (2016) as requested. Also added the citation to dropout itself as a regularization method (Srivastava et al., 2014)

14. Your code is very hard to read, partly because variable names like "combined_data_x_dict_DD" are not descriptive of their content or purpose. I recommend giving the code a refactor to improve its legibility.

Thank you for pointing out the readability issues with our variable names. We apologise for the confusion this has caused. Although our original intent was to make each name “self-documenting,” we recognise that chaining multiple conventions (e.g. _DD for daily, data_x for model inputs, combined_ for all sites, _dict for data type) can become opaque.

In our revised code we will:

- Simplify and standardise names so that each conveys its purpose at a glance (for example, combined_data_x_dict_DD will be refactored to inputs_daily_all_sites)
- Include a concise nomenclature table in the README and inline docstrings, outlining any remaining prefixes/suffixes
- Ensure consistency across modules, using clear, descriptive terms for inputs, outputs and data structures

We hope these changes will markedly improve legibility. If you encounter any further issues or would like additional examples of our refactored names, please let us know through the editor.

15. Figure 1 does not appear to be very useful in the current state. While I can see that the sites used in this study are all based in very few locations on Earth (California, Mid-West, Italy, South-Australia), it is hard to make out how many there are, exactly where they are located, which PFT they belong to and how much productivity they have on average. Please revise.

We thank the reviewer for pointing out the need for revising this image, so now, we split the data into 3 subpanels.

16. Figure 2 need a major rework, the fonts are not matching, alignment of boxes is not given, no legend provided and overall, it remains fuzzy how data flows.

We thank the reviewer for pointing out the need for greater visual consistency and clarity. We have redrawn Figure 2, aligned all boxes and arrows on a grid, and added numbered steps with more information explaining inputs, and outputs. LSTM cells has a wide known definition that is easily found in google search to avoid redundancy, we didn't define an LSTM cell as it is out of the scope of this paper. We also add similar information in the caption. We believe this new version makes the data flow is

transparent. Using a single font size is not required by journal, and is not possible at times, like in this figure, but we consider your comment and used only 7 font size.

17. L. 104ff. the description of the K-Fold cross-validation scheme used is incomplete. Please describe more precisely how the split is performed, and then ideally add a figure showing the Folds for all sites and visualizing which time periods belong to which fold.

Thank you for your comment. We had included the description of the K-fold cross-validation procedure in the revised manuscript line 250 (lines 310 in the version with tracked changes), and now expanded for further clarity. To maintain the temporal integrity of the time series, we applied a fold-wise split without shuffling. The dataset was first divided into 5 non-overlapping folds (each comprising 20% of the time series). One-fold was held out for testing, and the remaining 80% was further split into training and validation sets using a 90:10 ratio. Overlapping timestamps between training, validation, and test sets were explicitly removed to prevent data leakage.

While specific date ranges vary between sites and across Monte Carlo runs due to the embedded random seeds, we provide an illustrative example of how folds were distributed across time attached to this response letter (scroll down all the way to the end), and moreover, we have added an extra output “dates_range.txt” for the code, to print such dates as an additional output. We chose not to include exact date ranges for all sites in the main manuscript, as they are not fixed across runs and do not affect the validity of the general K-fold approach applied. Importantly, as shown in the results above, model performance remains consistent across folds. To reduce noise and the volume of information presented in the article, we chose to report fold-averaged results. Thanks for bringing this point to our attention which improves transparency of our code.

18. L. 152 "we further use interpretability" makes no sense, perhaps you mean "we further use methods to gain interpretability of the modeled functional relationship", or something alike?

19. Similar, L. 154 "can be combined", rather I would say it is more accurate to state, "can be applied to"

Corrected, thank you!

20. Can you elaborate what you mean by "suitable baseline" in L. 157

Expanded lines 230 (see line 290 in the version with tracked changes), thank you!

21. L. 159, "the GradientExplainer" is mentioned, without being properly introduced before.

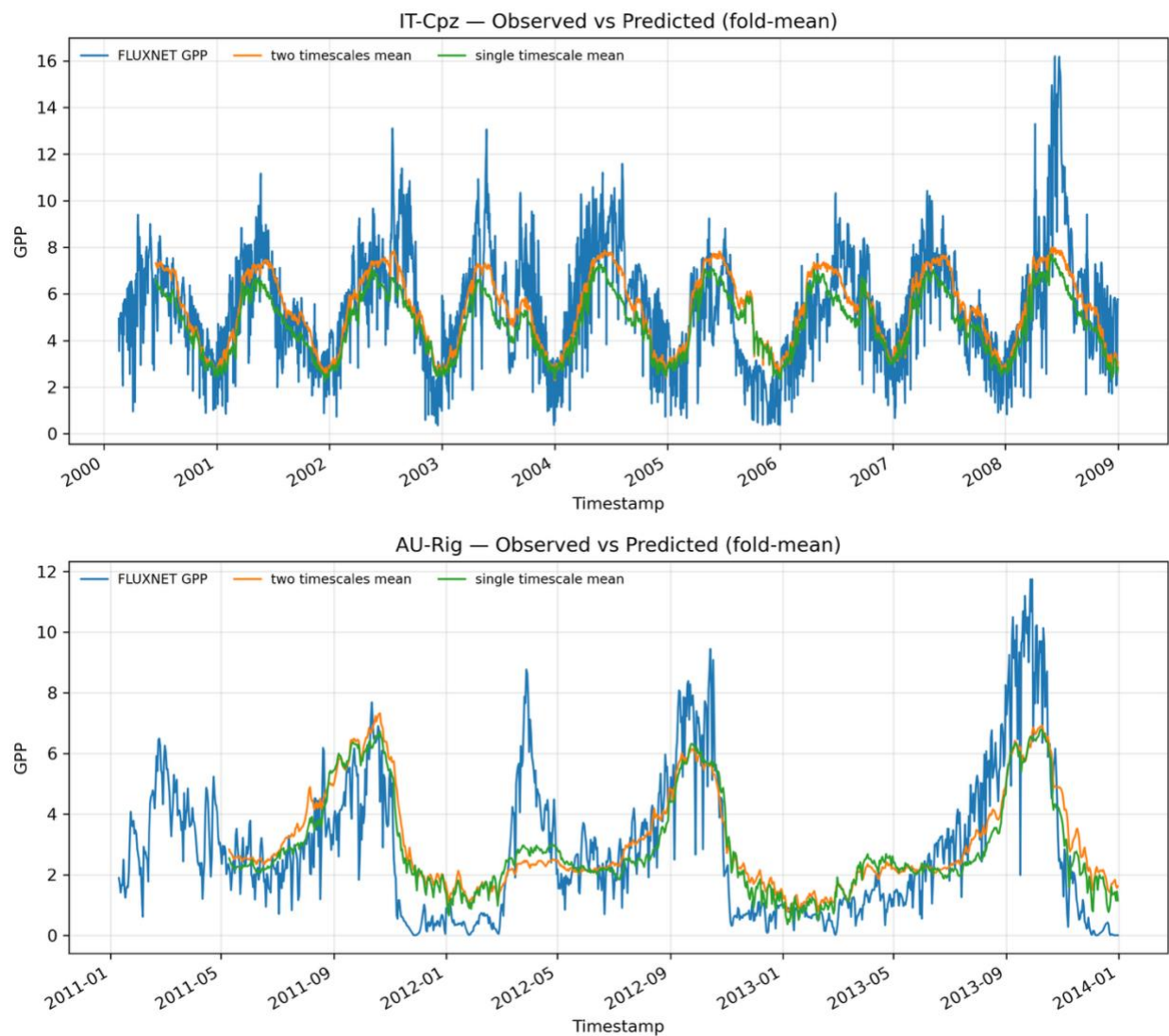
Explained as mentioned below, thank you!

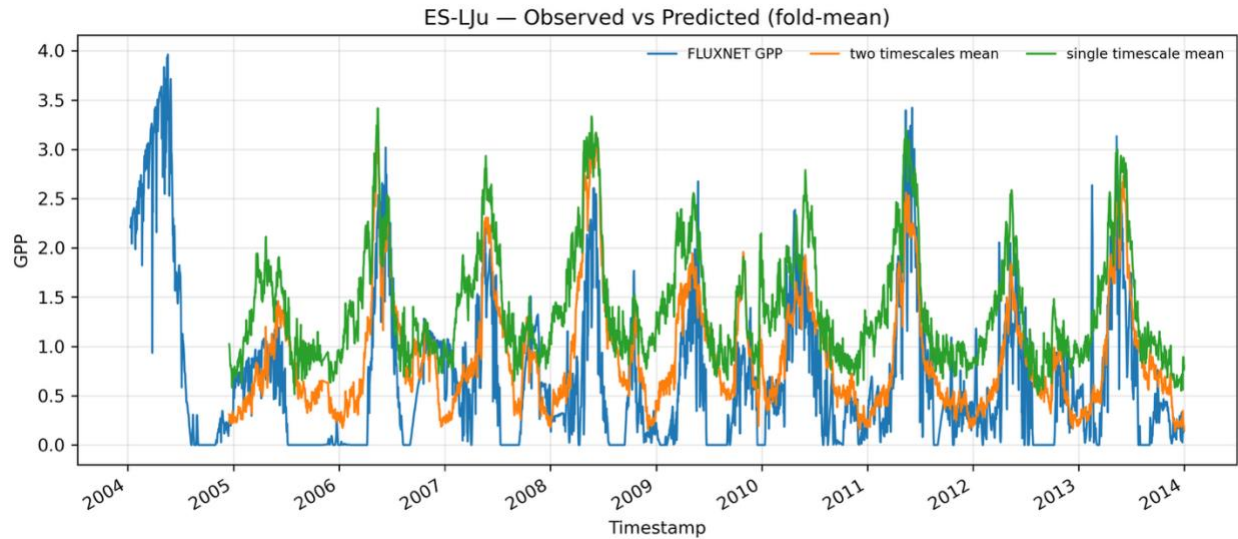
22. Section 3.2 you mention how important the choice of baseline is, but not, how to actually pick a baseline. Please add this information, to make the method section complete.

It was explained in the results section, now moved to method section, see lines 225-35, thanks.

23. L. 189f. - do you actually show somewhere how the inclusion of hourly data improves performance at daily resolution?

This behavior is well documented in the literature: standard LSTM models often struggle to reproduce peak dynamics, as discussed previously and cited at line 49 of the manuscript. Below, we show the performance of the same model trained with and without hourly data using a weighted RMSE objective function. The comparison illustrates that, in some sites, the model fails to reproduce short-term dynamics and consistently underestimates peaks, whereas in others the performance is more satisfactory. In sites such as IT-Cpz and AU-Rig, the improvements from incorporating hourly inputs are relatively minor, while at some others such as ES-LJu the model in absence of two timescale of data tends to overestimate productivity. This variation likely reflects the fact that weighted RMSE becomes less interpretable in the absence of sub-daily drivers, as it penalizes errors associated with variability that the model cannot resolve without hourly information.





24. L. 196f. - your evaluation methodology should be part of the method section

As per your suggestion, we have now included a note under method to explain better the evaluation, and a section in discussion as per Reviewer #1 request to summarize all evaluation methods. Thank you.

25. Fig. 3 (but also all other figures) please fix font sizes for better readability and consistency across figures. Probably Copernicus provides guidelines.

As per journal guideline as briefed in <https://www.geoscientific-model-development.net/submission.html>, under “Tables & Figures” the family font needs to be consistent, but there is no requirement for the font size. Still, to respect your request, we polished the figures for consistent font size as much as possible. Thank you.

26. Section 4.2 is extremely short and reads more like methodology. I recommend moving to the methods part, but then adding a longer discussion of Fig. 5 in the results part.

As per your and reviewer 1 suggestion, we have moved this section to SI as appendix A3, and Fig 5 is presented under this section as Fig A2. we hope this change improve the clarity of the manuscript.

27. The flow would be improved, if your evaluation in 4.1 and benchmarking in 5.1 would in fact be unified in a single section.

We followed the suggestion from Reviewer #1 in this regard and added to the start of section 5.1 a sub section which encapsulates all tables included in the manuscript (Fig 3, 9, Table 1 and 4, Table B1, B2). Also, model evaluation consists of 4 different variables, while benchmarking concerns only GPP product.

28. How did you tune the hyperparameters of your model?

We employed a manual, validation-driven hyperparameter tuning strategy rather than an exhaustive grid or random search. Added lines 220-5 for clarification and see Section B1 and Table B1 in the Appendix.

.

Following these remarks, I suggest a major revision of this work, to alleviate the major flaws related to the scientific content, but also to improve the overall presentation of the study.

However, I am also happy to be proven wrong should any of the points I raised solely be due to a misunderstanding from my side.

Thank you for your thorough reading of our manuscript and for the constructive feedback. With respect to the concern about “major flaws related to the scientific content,” we have carefully revisited each point and remain confident in the soundness of our methodology, the validity of our data analyses and the robustness of our conclusions. We acknowledge that our previous submission may give the wrong impression that we intend to claim our model is generic in space, which we did not intend to do in this work and have now revisited and acknowledged this in multiple locations throughout the manuscript. We therefore respectfully disagree that there are fundamental scientific errors in the work. That said, we wholeheartedly agree that the overall presentation needed to be strengthened and took all the raised points to address these issues.

Kindly,

the reviewer

Thank you so much,
Mitra Cattray on behalf of all co-authors

Fold dates: These values are exemplary and do not necessary reflect the exact dates used to produce the results

Fold 0:

Site 'AU-Rig':

Train: 2012-05-24 00:00:00 → 2014-05-17 00:00:00

Test : 2011-05-01 00:00:00 → 2012-01-24 00:00:00

Val : 2014-09-15 00:00:00 → 2014-12-31 00:00:00

Site 'ES-LJu':

Train: 2007-02-06 00:00:00 → 2012-12-11 00:00:00

Test : 2004-12-17 00:00:00 → 2006-10-08 00:00:00

Val : 2013-04-11 00:00:00 → 2013-12-31 00:00:00

Site 'IT-BCi':

Train: 2006-10-17 00:00:00 → 2013-10-25 00:00:00

Test : 2004-04-30 00:00:00 → 2006-06-18 00:00:00

Val : 2014-02-23 00:00:00 → 2014-12-31 00:00:00

Site 'IT-CA1':

Train: 2012-06-11 00:00:00 → 2014-05-19 00:00:00

Test : 2011-05-24 00:00:00 → 2012-02-11 00:00:00

Val : 2014-09-17 00:00:00 → 2014-12-31 00:00:00

Site 'IT-CA2':

Train: 2012-07-25 00:00:00 → 2014-05-24 00:00:00

Test : 2011-07-18 00:00:00 → 2012-03-26 00:00:00

Val : 2014-09-22 00:00:00 → 2014-12-31 00:00:00

Site 'IT-CA3':

Train: 2012-10-16 00:00:00 → 2014-06-01 00:00:00

Test : 2011-10-30 00:00:00 → 2012-06-17 00:00:00
 Val : 2014-09-30 00:00:00 → 2014-12-31 00:00:00
 Site 'IT-Cpz':
 Train: 2002-04-16 00:00:00 → 2007-12-19 00:00:00
 Test : 2000-03-14 00:00:00 → 2001-12-16 00:00:00
 Val : 2008-04-18 00:00:00 → 2008-12-31 00:00:00
 Site 'IT-Noe':
 Train: 2006-10-17 00:00:00 → 2013-10-25 00:00:00
 Test : 2004-04-30 00:00:00 → 2006-06-18 00:00:00
 Val : 2014-02-23 00:00:00 → 2014-12-31 00:00:00
 Site 'IT-PT1':
 Train: 2003-03-13 00:00:00 → 2004-06-16 00:00:00
 Test : 2002-05-02 00:00:00 → 2002-11-12 00:00:00
 Val : 2004-10-15 00:00:00 → 2004-12-31 00:00:00
 Site 'IT-Ro1':
 Train: 2002-05-24 00:00:00 → 2007-12-23 00:00:00
 Test : 2000-04-30 00:00:00 → 2002-01-23 00:00:00
 Val : 2008-04-22 00:00:00 → 2008-12-31 00:00:00
 Site 'IT-Ro2':
 Train: 2004-08-05 00:00:00 → 2008-12-31 00:00:00 and then 2010-01-01 00:00:00 → 2011-11-24 00:00:00
 Test : 2002-05-01 00:00:00 → 2004-04-06 00:00:00
 Val : 2012-03-24 00:00:00 → 2012-12-31 00:00:00
 Site 'IT-SRo':
 Train: 2002-11-24 00:00:00 → 2011-08-17 00:00:00
 Test : 1999-12-17 00:00:00 → 2002-07-26 00:00:00
 Val : 2011-12-16 00:00:00 → 2012-12-31 00:00:00
 Site 'US-AR1':
 Train: 2010-05-25 00:00:00 → 2012-05-17 00:00:00
 Test : 2009-05-01 00:00:00 → 2010-01-24 00:00:00
 Val : 2012-09-15 00:00:00 → 2012-12-31 00:00:00
 Site 'US-ARM':
 Train: 2005-08-05 00:00:00 → 2011-11-24 00:00:00
 Test : 2003-05-01 00:00:00 → 2005-04-06 00:00:00
 Val : 2012-03-24 00:00:00 → 2012-12-31 00:00:00
 Site 'US-Blo':
 Train: 1999-11-07 00:00:00 → 2006-08-27 00:00:00
 Test : 1997-06-12 00:00:00 → 1999-07-09 00:00:00
 Val : 2006-12-26 00:00:00 → 2007-10-24 00:00:00
 Site 'US-Ton':
 Train: 2004-05-24 00:00:00 → 2013-07-29 00:00:00
 Test : 2001-05-01 00:00:00 → 2004-01-24 00:00:00
 Val : 2013-11-27 00:00:00 → 2014-12-31 00:00:00
 Site 'US-Var':
 Train: 2003-12-12 00:00:00 → 2013-07-13 00:00:00
 Test : 2000-10-08 00:00:00 → 2003-08-13 00:00:00
 Val : 2013-11-11 00:00:00 → 2014-12-31 00:00:00
 Fold 1:
 Site 'AU-Rig':

Train: 2011-05-01 00:00:00 → 2011-09-26 00:00:00 and then 2013-02-16 00:00:00 → 2014-05-17 00:00:00
 Test : 2012-01-25 00:00:00 → 2012-10-18 00:00:00
 Val : 2014-09-15 00:00:00 → 2014-12-31 00:00:00
 Site 'ES-LJu':
 Train: 2004-12-17 00:00:00 → 2006-06-10 00:00:00 and then 2008-11-28 00:00:00 → 2012-12-11 00:00:00
 Test : 2006-10-09 00:00:00 → 2008-07-30 00:00:00
 Val : 2013-04-11 00:00:00 → 2013-12-31 00:00:00
 Site 'IT-BCi':
 Train: 2004-04-30 00:00:00 → 2006-02-18 00:00:00 and then 2008-12-05 00:00:00 → 2013-10-25 00:00:00
 Test : 2006-06-19 00:00:00 → 2008-08-06 00:00:00
 Val : 2014-02-23 00:00:00 → 2014-12-31 00:00:00
 Site 'IT-CA1':
 Train: 2011-05-24 00:00:00 → 2011-10-14 00:00:00 and then 2013-03-02 00:00:00 → 2014-05-19 00:00:00
 Test : 2012-02-12 00:00:00 → 2012-11-01 00:00:00
 Val : 2014-09-17 00:00:00 → 2014-12-31 00:00:00
 Site 'IT-CA2':
 Train: 2011-07-18 00:00:00 → 2011-11-27 00:00:00 and then 2013-04-04 00:00:00 → 2014-05-24 00:00:00
 Test : 2012-03-27 00:00:00 → 2012-12-04 00:00:00
 Val : 2014-09-22 00:00:00 → 2014-12-31 00:00:00
 Site 'IT-CA3':
 Train: 2011-10-30 00:00:00 → 2012-02-18 00:00:00 and then 2013-06-05 00:00:00 → 2014-06-01 00:00:00
 Test : 2012-06-18 00:00:00 → 2013-02-04 00:00:00
 Val : 2014-09-30 00:00:00 → 2014-12-31 00:00:00
 Site 'IT-Cpz':
 Train: 2000-03-14 00:00:00 → 2001-08-18 00:00:00 and then 2004-01-19 00:00:00 → 2007-12-19 00:00:00
 Test : 2001-12-17 00:00:00 → 2003-09-20 00:00:00
 Val : 2008-04-18 00:00:00 → 2008-12-31 00:00:00
 Site 'IT-Noe':
 Train: 2004-04-30 00:00:00 → 2006-02-18 00:00:00 and then 2008-12-05 00:00:00 → 2013-10-25 00:00:00
 Test : 2006-06-19 00:00:00 → 2008-08-06 00:00:00
 Val : 2014-02-23 00:00:00 → 2014-12-31 00:00:00
 Site 'IT-PT1':
 Train: 2002-05-02 00:00:00 → 2002-07-15 00:00:00 and then 2003-09-24 00:00:00 → 2004-06-16 00:00:00
 Test : 2002-11-13 00:00:00 → 2003-05-26 00:00:00
 Val : 2004-10-15 00:00:00 → 2004-12-31 00:00:00
 Site 'IT-Ro1':
 Train: 2000-04-30 00:00:00 → 2001-09-25 00:00:00 and then 2004-02-17 00:00:00 → 2007-12-23 00:00:00
 Test : 2002-01-24 00:00:00 → 2003-10-19 00:00:00
 Val : 2008-04-22 00:00:00 → 2008-12-31 00:00:00
 Site 'IT-Ro2':

Train: 2002-05-01 00:00:00 → 2003-12-08 00:00:00 and then 2006-07-13 00:00:00 → 2008-12-31 00:00:00 and then 2010-01-01 00:00:00 → 2011-11-24 00:00:00

Test : 2004-04-07 00:00:00 → 2006-03-14 00:00:00

Val : 2012-03-24 00:00:00 → 2012-12-31 00:00:00

Site 'IT-SRo':

Train: 1999-12-17 00:00:00 → 2002-03-28 00:00:00 and then 2005-07-04 00:00:00 → 2011-08-17 00:00:00

Test : 2002-07-27 00:00:00 → 2005-03-05 00:00:00

Val : 2011-12-16 00:00:00 → 2012-12-31 00:00:00

Site 'US-AR1':

Train: 2009-05-01 00:00:00 → 2009-09-26 00:00:00 and then 2011-02-17 00:00:00 → 2012-05-17 00:00:00

Test : 2010-01-25 00:00:00 → 2010-10-19 00:00:00

Val : 2012-09-15 00:00:00 → 2012-12-31 00:00:00

Site 'US-ARM':

Train: 2003-05-01 00:00:00 → 2004-12-07 00:00:00 and then 2007-07-13 00:00:00 → 2011-11-24 00:00:00

Test : 2005-04-07 00:00:00 → 2007-03-14 00:00:00

Val : 2012-03-24 00:00:00 → 2012-12-31 00:00:00

Site 'US-Blo':

Train: 1997-06-12 00:00:00 → 1999-03-11 00:00:00 and then 2001-12-04 00:00:00 → 2006-08-27 00:00:00

Test : 1999-07-10 00:00:00 → 2001-08-05 00:00:00

Val : 2006-12-26 00:00:00 → 2007-10-24 00:00:00

Site 'US-Ton':

Train: 2001-05-01 00:00:00 → 2003-09-26 00:00:00 and then 2007-02-17 00:00:00 → 2013-07-29 00:00:00

Test : 2004-01-25 00:00:00 → 2006-10-19 00:00:00

Val : 2013-11-27 00:00:00 → 2014-12-31 00:00:00

Site 'US-Var':

Train: 2000-10-08 00:00:00 → 2003-04-15 00:00:00 and then 2006-10-17 00:00:00 → 2013-07-13 00:00:00

Test : 2003-08-14 00:00:00 → 2006-06-18 00:00:00

Val : 2013-11-11 00:00:00 → 2014-12-31 00:00:00

Fold 2:

Site 'AU-Rig':

Train: 2011-05-01 00:00:00 → 2012-06-20 00:00:00 and then 2013-11-11 00:00:00 → 2014-05-17 00:00:00

Test : 2012-10-19 00:00:00 → 2013-07-13 00:00:00

Val : 2014-09-15 00:00:00 → 2014-12-31 00:00:00

Site 'ES-LJu':

Train: 2004-12-17 00:00:00 → 2008-04-01 00:00:00 and then 2010-09-19 00:00:00 → 2012-12-11 00:00:00

Test : 2008-07-31 00:00:00 → 2010-05-21 00:00:00

Val : 2013-04-11 00:00:00 → 2013-12-31 00:00:00

Site 'IT-BCi':

Train: 2004-04-30 00:00:00 → 2008-04-08 00:00:00 and then 2011-01-24 00:00:00 → 2013-10-25 00:00:00

Test : 2008-08-07 00:00:00 → 2010-09-25 00:00:00

Val : 2014-02-23 00:00:00 → 2014-12-31 00:00:00

Site 'IT-CA1':
 Train: 2011-05-24 00:00:00 → 2012-07-04 00:00:00 and then 2013-11-21 00:00:00 → 2014-05-19 00:00:00
 Test : 2012-11-02 00:00:00 → 2013-07-23 00:00:00
 Val : 2014-09-17 00:00:00 → 2014-12-31 00:00:00

Site 'IT-CA2':
 Train: 2011-07-18 00:00:00 → 2012-08-06 00:00:00 and then 2013-12-13 00:00:00 → 2014-05-24 00:00:00
 Test : 2012-12-05 00:00:00 → 2013-08-14 00:00:00
 Val : 2014-09-22 00:00:00 → 2014-12-31 00:00:00

Site 'IT-CA3':
 Train: 2011-10-30 00:00:00 → 2012-10-07 00:00:00 and then 2014-01-23 00:00:00 → 2014-06-01 00:00:00
 Test : 2013-02-05 00:00:00 → 2013-09-24 00:00:00
 Val : 2014-09-30 00:00:00 → 2014-12-31 00:00:00

Site 'IT-Cpz':
 Train: 2000-03-14 00:00:00 → 2003-05-23 00:00:00 and then 2005-10-23 00:00:00 → 2007-12-19 00:00:00
 Test : 2003-09-21 00:00:00 → 2005-06-24 00:00:00
 Val : 2008-04-18 00:00:00 → 2008-12-31 00:00:00

Site 'IT-Noe':
 Train: 2004-04-30 00:00:00 → 2008-04-08 00:00:00 and then 2011-01-24 00:00:00 → 2013-10-25 00:00:00
 Test : 2008-08-07 00:00:00 → 2010-09-25 00:00:00
 Val : 2014-02-23 00:00:00 → 2014-12-31 00:00:00

Site 'IT-PT1':
 Train: 2002-05-02 00:00:00 → 2003-01-26 00:00:00 and then 2004-04-06 00:00:00 → 2004-06-16 00:00:00
 Test : 2003-05-27 00:00:00 → 2003-12-07 00:00:00
 Val : 2004-10-15 00:00:00 → 2004-12-31 00:00:00

Site 'IT-Ro1':
 Train: 2000-04-30 00:00:00 → 2003-06-21 00:00:00 and then 2005-11-12 00:00:00 → 2007-12-23 00:00:00
 Test : 2003-10-20 00:00:00 → 2005-07-14 00:00:00
 Val : 2008-04-22 00:00:00 → 2008-12-31 00:00:00

Site 'IT-Ro2':
 Train: 2002-05-01 00:00:00 → 2005-11-14 00:00:00 and then 2008-06-19 00:00:00 → 2008-12-31 00:00:00 and then 2010-01-01 00:00:00 → 2011-11-24 00:00:00
 Test : 2006-03-15 00:00:00 → 2008-02-19 00:00:00
 Val : 2012-03-24 00:00:00 → 2012-12-31 00:00:00

Site 'IT-SRo':
 Train: 1999-12-17 00:00:00 → 2004-11-05 00:00:00 and then 2008-02-12 00:00:00 → 2011-08-17 00:00:00
 Test : 2005-03-06 00:00:00 → 2007-10-14 00:00:00
 Val : 2011-12-16 00:00:00 → 2012-12-31 00:00:00

Site 'US-AR1':
 Train: 2009-05-01 00:00:00 → 2010-06-21 00:00:00 and then 2011-11-12 00:00:00 → 2012-05-17 00:00:00
 Test : 2010-10-20 00:00:00 → 2011-07-14 00:00:00
 Val : 2012-09-15 00:00:00 → 2012-12-31 00:00:00

Site 'US-ARM':

Train: 2003-05-01 00:00:00 → 2006-11-14 00:00:00 and then 2009-06-19 00:00:00 → 2011-11-24 00:00:00

Test : 2007-03-15 00:00:00 → 2009-02-18 00:00:00

Val : 2012-03-24 00:00:00 → 2012-12-31 00:00:00

Site 'US-Blo':

Train: 1997-06-12 00:00:00 → 2001-04-07 00:00:00 and then 2003-12-31 00:00:00 → 2006-08-27 00:00:00

Test : 2001-08-06 00:00:00 → 2003-09-01 00:00:00

Val : 2006-12-26 00:00:00 → 2007-10-24 00:00:00

Site 'US-Ton':

Train: 2001-05-01 00:00:00 → 2006-06-21 00:00:00 and then 2009-11-12 00:00:00 → 2013-07-29 00:00:00

Test : 2006-10-20 00:00:00 → 2009-07-14 00:00:00

Val : 2013-11-27 00:00:00 → 2014-12-31 00:00:00

Site 'US-Var':

Train: 2000-10-08 00:00:00 → 2006-02-18 00:00:00 and then 2009-08-22 00:00:00 → 2013-07-13 00:00:00

Test : 2006-06-19 00:00:00 → 2009-04-23 00:00:00

Val : 2013-11-11 00:00:00 → 2014-12-31 00:00:00

Fold 3:

Site 'AU-Rig':

Train: 2011-05-01 00:00:00 → 2013-03-15 00:00:00

Test : 2013-07-14 00:00:00 → 2014-04-07 00:00:00

Val : 2014-09-15 00:00:00 → 2014-12-31 00:00:00

Site 'ES-LJu':

Train: 2004-12-17 00:00:00 → 2010-01-21 00:00:00 and then 2012-07-10 00:00:00 → 2012-12-11 00:00:00

Test : 2010-05-22 00:00:00 → 2012-03-11 00:00:00

Val : 2013-04-11 00:00:00 → 2013-12-31 00:00:00

Site 'IT-BCi':

Train: 2004-04-30 00:00:00 → 2010-05-28 00:00:00 and then 2013-03-13 00:00:00 → 2013-10-25 00:00:00

Test : 2010-09-26 00:00:00 → 2012-11-12 00:00:00

Val : 2014-02-23 00:00:00 → 2014-12-31 00:00:00

Site 'IT-CA1':

Train: 2011-05-24 00:00:00 → 2013-03-25 00:00:00

Test : 2013-07-24 00:00:00 → 2014-04-12 00:00:00

Val : 2014-09-17 00:00:00 → 2014-12-31 00:00:00

Site 'IT-CA2':

Train: 2011-07-18 00:00:00 → 2013-04-16 00:00:00

Test : 2013-08-15 00:00:00 → 2014-04-23 00:00:00

Val : 2014-09-21 00:00:00 → 2014-12-31 00:00:00

Site 'IT-CA3':

Train: 2011-10-30 00:00:00 → 2013-05-27 00:00:00

Test : 2013-09-25 00:00:00 → 2014-05-14 00:00:00

Val : 2014-09-30 00:00:00 → 2014-12-31 00:00:00

Site 'IT-Cpz':

Train: 2000-03-14 00:00:00 → 2005-02-24 00:00:00 and then 2007-07-28 00:00:00 → 2007-12-19 00:00:00

Test : 2005-06-25 00:00:00 → 2007-03-29 00:00:00

Val : 2008-04-18 00:00:00 → 2008-12-31 00:00:00
 Site 'IT-Noe':
 Train: 2004-04-30 00:00:00 → 2010-05-28 00:00:00 and then 2013-03-13 00:00:00 → 2013-10-25 00:00:00
 Test : 2010-09-26 00:00:00 → 2012-11-12 00:00:00
 Val : 2014-02-23 00:00:00 → 2014-12-31 00:00:00
 Site 'IT-PT1':
 Train: 2002-05-02 00:00:00 → 2003-08-09 00:00:00
 Test : 2003-12-08 00:00:00 → 2004-06-16 00:00:00
 Val : 2004-10-15 00:00:00 → 2004-12-31 00:00:00
 Site 'IT-Ro1':
 Train: 2000-04-30 00:00:00 → 2005-03-16 00:00:00 and then 2007-08-07 00:00:00 → 2007-12-23 00:00:00
 Test : 2005-07-15 00:00:00 → 2007-04-08 00:00:00
 Val : 2008-04-22 00:00:00 → 2008-12-31 00:00:00
 Site 'IT-Ro2':
 Train: 2002-05-01 00:00:00 → 2007-10-22 00:00:00 and then 2011-05-26 00:00:00 → 2011-11-24 00:00:00
 Test : 2008-02-20 00:00:00 → 2008-12-31 00:00:00 and then 2010-01-01 00:00:00 → 2011-01-25 00:00:00
 Val : 2012-03-24 00:00:00 → 2012-12-31 00:00:00
 Site 'IT-SRo':
 Train: 1999-12-17 00:00:00 → 2007-06-16 00:00:00 and then 2010-09-22 00:00:00 → 2011-08-17 00:00:00
 Test : 2007-10-15 00:00:00 → 2010-05-24 00:00:00
 Val : 2011-12-16 00:00:00 → 2012-12-31 00:00:00
 Site 'US-AR1':
 Train: 2009-05-01 00:00:00 → 2011-03-16 00:00:00
 Test : 2011-07-15 00:00:00 → 2012-04-07 00:00:00
 Val : 2012-09-15 00:00:00 → 2012-12-31 00:00:00
 Site 'US-ARM':
 Train: 2003-05-01 00:00:00 → 2008-10-21 00:00:00 and then 2011-05-26 00:00:00 → 2011-11-24 00:00:00
 Test : 2009-02-19 00:00:00 → 2011-01-25 00:00:00
 Val : 2012-03-24 00:00:00 → 2012-12-31 00:00:00
 Site 'US-Blo':
 Train: 1997-06-12 00:00:00 → 2003-05-04 00:00:00 and then 2006-01-26 00:00:00 → 2006-08-27 00:00:00
 Test : 2003-09-02 00:00:00 → 2005-09-27 00:00:00
 Val : 2006-12-26 00:00:00 → 2007-10-24 00:00:00
 Site 'US-Ton':
 Train: 2001-05-01 00:00:00 → 2009-03-16 00:00:00 and then 2012-08-06 00:00:00 → 2013-07-29 00:00:00
 Test : 2009-07-15 00:00:00 → 2012-04-07 00:00:00
 Val : 2013-11-27 00:00:00 → 2014-12-31 00:00:00
 Site 'US-Var':
 Train: 2000-10-08 00:00:00 → 2008-12-24 00:00:00 and then 2012-06-26 00:00:00 → 2013-07-13 00:00:00
 Test : 2009-04-24 00:00:00 → 2012-02-26 00:00:00
 Val : 2013-11-11 00:00:00 → 2014-12-31 00:00:00

