Supplement of Geosci. Model Dev., 18, 4317–4333, 2025
https://doi.org/10.5194/gmd-18-4317-2025-supplement

Geoscientific Model Development
Open Access

Supplement of

Emulating grid-based forest carbon dynamics using machine learning: an LPJ-GUESS v4.1.1 application

Carolina Natel et al.

Correspondence to: Carolina Natel (carolina.moura@kit.edu)

S1. LPJ-GUESS Model Modifications, Experimental Setup and Forcing Data

We used LPJ-GUESS version 4.1.1 (Nord et al., 2021) for our simulations. To ensure reproducibility, we provide the model code with necessary modifications used to generate data for emulation development through the Zenodo repository at: https://zenodo.org/records/15065248.

- **Modification 1:** Implemented an eXtended NetCDF input (cfxinput.h/cpp).
- **Modification 2:** Added output for annual climate data.
- **Modification 3:** Introduced a parameter to apply a fixed nitrogen deposition across simulations (fixed_ndep and fixed_ndep_year).
- **Modification 4:** Enabled the output of spin-up period results (if_spinup_outputs).
- **Modification 5:** Added spinup_clear2_year parameter to control stand-replacing disturbances.

Experimental Setup

Each LPJ-GUESS simulation began with a 500-year spin-up to stabilize carbon pools, using the 1901 atmospheric CO_2 concentration and repeating, detrended 1901–1930 climate data. Following the spin-up, a stand-replacing disturbance (via spinup_clear2_year) simulated a clear-cut, removing all vegetation and exposing soil. Vegetation residues were left on-site, contributing to litter and soil carbon pools. Post-disturbance, natural vegetation regrew under historical (1850–2014) and future (2015–2100) conditions. Land-use changes and fire disturbance were not modelled. In LPJ-GUESS, besides fire disturbances, we account for other external disturbances (e.g. windstorms, plant diseases etc) using a generic patch-destroying regime with a stochastic probability interval of the expected return time. Disturbance return time varies substantially across the global forest area (Pugh et al., 2019), and the interval we have chosen is a simplification that has been adopted in a number of previous studies using LPJ-GUESS and other vegetation models (Zaehle et al. 2008). The LPJ-GUESS parameter settings are presented below:

- Fire model: Disabled.
- Nitrogen deposition: Held constant at 2015 levels, following Lamarque et al. (2013).
- Disturbance interval: Default LPJ-GUESS setting of 100 years.
- Replicate number of patches: 50.
- Vegetation type: Potential natural vegetation only, to simplify ecosystem carbon responses and isolate climate-driven impacts.

Forcing Data

The simulations used both historical (1850–2014) and future (2015–2100) climate data. Future runs began from the end state of the historical period.

- Climate Data: Bias-corrected CMIP6 data from the ISIMIP 3b project (Lange, 2019) was used, including five Earth System Models (ESMs) to cover climate sensitivity variations: IPSL-CM6A-LR, MPI-ESM1-2-HR, MRI-ESM2-0, GFDL-ESM4, and UKESM1-0-LL. Simulations included four Representative Concentration Pathways (RCPs): RCP2.6, RCP4.5, RCP7.0, and RCP8.5.
- Nitrogen Deposition: Set at 2015 levels based on Lamarque et al. (2013) data.
- Atmospheric CO_2 Concentrations: Aligned with observed CO_2 mixing ratios for each RCP scenario.

Table S1. Random forests hyperparameters, showing the values tested during the hyperparameter grid search and the best values for each task (Carbon stocks: C stocks and Carbon fluxes: C fluxes).

Hyperparameter	Description	Values	Best value (C stocks | C fluxes)
n_estimators	Number of trees in the random forest	350, 500, 600, 700, 1000	1000 | 1000
max_samples	The number of samples to draw from data to train each decision tree	0.2, 0.4, 0.6, 0.8, 1.0	0.2 | 0.2
max_features	Number of features to consider when looking for the best split	0.2, 0.4, 0.6, 0.8, 1.0	0.8 | 0.8
max_depth	Maximum depth of the decision tree	200, 1000, 2000	200 | 200
min_samples_split	Minimum number of samples required to split an internal node	10, 20, 250, 400	250 | 250

Table S2. Neural network hyperparameters, showing the values tested during the hyperparameter grid search and the best values for each task (Carbon stocks: C stocks and Carbon fluxes: C fluxes).

Hyperparameter	Name	Description	Values	Best value (C stocks | C fluxes
learning_rate	Learning rate	Controls the step size at each iteration while moving toward a minimum of the loss function.	0.001, 0.01, 0.1	0.001 | 0.001
layers	Number of layers	Defines the depth of the neural network. Each layer encapsulates a state (weights) and some computation.	1, 2, 3	2 | 2
neurons	Number of neurons	The basic computational units in a neural network layer. More neurons can capture more complex patterns.	32, 64, 128	64 | 128
activation_function	Activation function	Introduces non-linearity into the network, allowing it to learn complex patterns.	'relu', 'tanh'	‘tanh’ | ‘relu’
dropout_rate	Dropout rate	A regularization technique to prevent overfitting. Determines the proportion of neurons randomly set to zero during training.	0, 0.2, 0.5	0.2 | 0.2
batch_size	Batch size	Determines the number of samples processed before the model is updated. Affects training speed and stability.	32, 64, 128	32 | 128

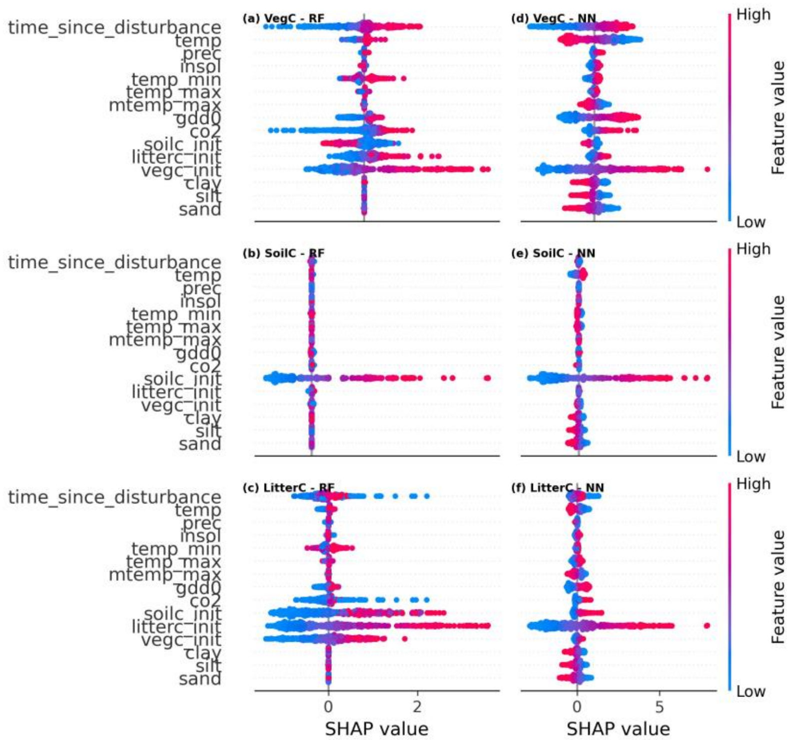

Figure S1. SHAP values per feature for carbon stock predictions (vegetation carbon (VegC), soil carbon (SoilC), and litter carbon (LitterC)) using the (a - c) Random Forest (RF) and (d - f) Neural Network (NN) emulators. The Y-axis lists the features used in the model. The X-axis displays SHAP values, which quantify the impact of each feature on the model's prediction. Positive SHAP values indicate that a feature increases the prediction, while negative SHAP values suggest a decrease. The color gradient represents the feature values (red for high, blue for low). Each point on the plot corresponds to a single data point from the dataset, and its position along the X-axis shows the contribution of that feature to the prediction for that instance. For example, in 6a, low values (blue) of time elapsed since last disturbance (time_since_disturbance) decrease the predicted VegC, while high values (red) increase the predicted VegC by up to 2.feature values (red for high, blue for low).

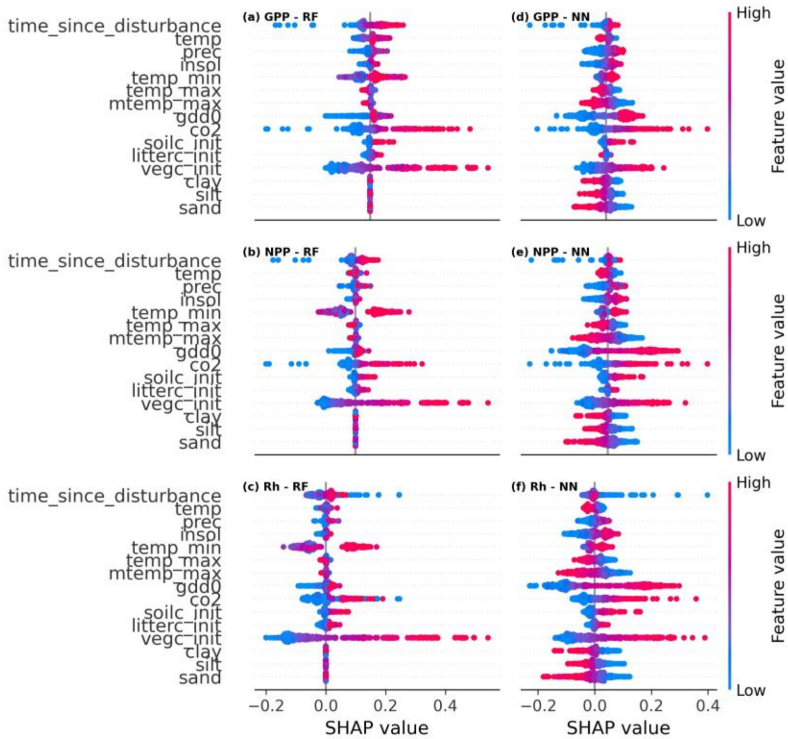

Figure S2: SHAP values per feature for carbon flux predictions (gross primary productivity (GPP), net primary productivity (NPP), and heterotrophic respiration (Rh)) using the (a - c) Random Forest (RF) and (d - f) Neural Network (NN) emulators. For a detailed explanation of the SHAP plot, refer to the caption of Fig. S1.