Reducing Uncertainties in Coupled Carbon-water Cycle Predictions in the Southeastern United States Using Perturbed Parameter Experiments and Machine Learning Approaches

By

Thomas Mutiso Kavoo

A thesis submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Master of Science

Auburn, Alabama
December 9, 2023

Keywords: Parameter perturbation experiments, sensitivity analysis, uncertainties, carbon-water cycle predictions, Emulator, Calibration

Approved by

Sanjiv Kumar, Chair, Assistant Professor College of Forestry, Wildlife and Environment
Daniel Kennedy, Research Associate/Project Scientist, University of California/NCAR
Latif Kalin, Professor, College of Forestry, Wildlife and Environment
Lana Narine, Assistant Professor, College of Forestry, Wildlife and Environment

Abstract

Carbon-water cycle interactions are considerably uncertain in the climate models. Parameter sensitivity studies are extensively utilized to reduce uncertainty in hydrological predictions. However, because of its high computation requirement, parameter sensitivity study has limited in-roads in climate modeling applications. A new National Center for Atmospheric Research – National Ecological Observatory Network (NCAR-NEON) simulation system provides a renewed opportunity to investigate parameter sensitivity in climate modeling applications. Two main reasons are: (1) the NCAR-NEON tool is ultra-computationally efficient, and (2) the availability of quality controlled and high-frequency NEON observations. Using the NCAR-NEON system, we perform parameter perturbation experiments at three flux tower sites utilizing a subset of 30 parameters affecting carbon-water cycle processes. We use the result to develop a land surface model Gaussian process regression emulator. About 244,800 years of CLM5 simulations are realized, including 200 years of spin-up runs and four years of production runs for each of the 400 parameter sets and at three sites (3 sites x 400 parameter sets x 204 years). The best-performing parameter set is selected as the optimized one that improved CLM5 performance in simulating carbon-water cycle processes from 0.37 to 0.78 on a normalized score from 0 to 1. Sensitivity analysis shows that photosynthetic parameters significantly influence carbon-water cycle interactions. The optimized parameters were examined at a regional scale using a high-resolution CLM configuration and forcing data. Upscaling of point-scale calibration at the regional scale shows mixed results, suggesting the need for separate parameter calibration for each plant functional type distribution. In part 2, we developed a Gaussian process regression emulator to mimic the behavior of CLM5 and provide predictions based on given inputs. The NCAR-NEON system, although computationally efficient, faces limitations due to the initialization process and resource consumption, limiting thorough parameter exploration for better alignment with observations. To overcome this barrier, we train an emulator that requires no initialization and is fast and efficient in mimicking land surface process behavior. The emulator improved CLM5 performance from 0.78 to 0.84.

Acknowledgments

I want to express my profound gratitude to God Almighty for guiding me through this significant milestone in my academic journey. I sincerely appreciate my academic supervisor, Dr. Sanjiv Kumar, for his excellent supervision, guidance, support, and motivation throughout my research and writing process. He was always available whenever I needed clarification or had a question, and his constant encouragement and inspiration have been invaluable throughout my academic pursuits. I would also like to extend my sincere gratitude to Dr. Daniel Kennedy, Professor Latif Kalin, and Dr. Lana Narine for their valuable contributions and as committee members to my research project. Special appreciation and mention to the National Center for Atmospheric Research Institution for their generous support, providing access to crucial resources, including the supercomputer and analysis resources, and technical guidance—my special appreciation to Dr. Will Wieder and Dr. Keith Oleson for their technical assistance.

I am deeply grateful to Dr. Daniel Kennedy for his essential role in setting up the experiments, offering guidance, and providing technical assistance whenever needed. My appreciation extends to Dr. Linnia Hawkins for her support and suggestions. I thank the Climate Water and Society lab (CWS) Lab for their support and encouragement throughout my research journey. I would also like to thank the Hatch Program of the National Institute of Food and Agriculture, United States Department of Agriculture, for funding this research. I express my gratitude to Md. Mozahidul Islam and Montasir Maruf for their support and Suggestions.

Finally, I want to thank my family and friends. A special thank you to my parents, Boniface Kavoo and Hanna Boniface, and my brothers, sisters, nephews, and nieces. Throughout my educational journey, my loved ones provided unwavering support.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AVHRR	Advanced very high-resolution radiometer.
CDF	Cumulative density function
CLM5	Community Land Model version 5
C02	Carbon dioxide
D0_	NEON Domain
ERA5	European Centre for Medium-Range Weather Forecasts Reanalysis version 5
ESM	Earth system models
ET	Evapotranspiration
FACE	Free air C02 enrichment experiment
GPP	Gross primary productivity
GP	Gaussian processes
GPR	Gaussian process regression
ILAMB	International land model benchmarking system
LAI	Leaf area index
LSM	land surface models
MBE	Mean Bias Error
NCAR	National Center for Atmospheric Research
NEON	National Ecological Observatory network
NLCD	National Landcover dataset
NLDAS	North American Land Data Assimilation System

ORNL	Oak Ridge National Laboratory
OSBS	Ordway-Swisher Biological Station
PFT	Plant functional type
PPE	Parameter perturbation experiment
PRISM	Parameter Elevation Regressions on Independent Slopes Model
RMSE	Root mean squared error.
TALL	Talladega National Forest NEON Tower site
VISCOUS	Variance-based sensitivity analysis using Copulas.

Chapter One: General Introduction

1.1 Background

The carbon-water cycle processes are integral to Land surface models (LSMs), enabling the simulation of complex interactions that drive climate and land use change (Arora et al., 2013). Carbon-water cycle processes within LSMs are intricately connected. Changes in water availability significantly influence plant growth and carbon sequestration through photosynthesis, while atmospheric CO2 variations affect plant water use via stomatal regulation (Gentine et al., 2019). These interactions create a feedback loop with the terrestrial biosphere, wherein climate changes disrupt carbon-water dynamics, driving shifts in vegetation patterns that subsequently affect carbon storage capacity and local climate conditions, resulting in a complex interplay of processes within terrestrial ecosystems (Heimann & Reichstein, 2008).

The interaction between atmospheric CO2 and terrestrial carbon storage shapes future climate dynamics, highlighting the complex interplay between biological systems and environmental conditions (Andrew et al., 2014). Particularly significant within this context is the potential for dynamic changes in vegetation, observed in tropical forests, to function as triggers for biome shifts (Lewis, 2006). These shifts underline the complex nature of regulating carbon stocks and carry implications for amplifying feedback mechanisms linked to climate change (Cox et al., 2000). The impact of CO2 on terrestrial carbon storage emerges as a perspective of negative feedback within the broader context of future climate dynamics and is essential for modern land surface modeling (Sabine et al., 2013).

Existing land surface models conceptually represent vegetation, likely to constrain the ecosystem's response to climate change. They use fixed vegetation distribution that does not represent functional diversity in vegetation, especially in tropical forests (Koven et al., 2020). These assumptions are likely to amplify the biases in the model's results. It is impractical to project long-term ecosystem changes and abrupt changes from vegetation changes using prescribed biogeography (Cox et al., 2000). The inability to represent vegetation dynamics and resolve to work on the assumption that a particular ecosystem contains specific vegetation attributes is likely to exaggerate changes to an imposed forcing as compared to methods that accommodate vegetation dynamics (Boit et al., 2016; Deb et al., 2018; Powell et al., 2018; van der Sande et al., 2017). Poor vegetation representation will give an erroneous response as it does not capture traits such as

vegetation mortality rates from resource competition and climate change (Le Quéré et al., 2015; Schuur et al., 2015).

LSMs are resource-intensive as they require significant data and energy to run—the complexities of the Earth's climate system challenge models' ability to replicate Earth system processes accurately. Regardless of how best the model captures these processes, achieving an accurate representation result from the model without any adjustments is almost impossible. Model adjustments help minimize uncertainties and improve model performance. Land surface model uncertainties could result from numerical representations of processes, parameterization schemes, and discrepancies between regional and synoptic scales (Wu et al., 2022). Some LSM parameters take prescribed values that are not explicitly determined and require approval from existing knowledge. To improve performance and minimize simulation uncertainties, model parameters are adjusted to calibrate the model against observation data, and parameter uncertainty is usually quantified through sensitivity analysis based on sampling within the acceptable parameter space (Prihodko et al., 2008)

A significant challenge facing climate modeling is the complexity and demand for intensive computer infrastructure, especially for global simulations. This challenge limits the ability to experiment on a global scale. However, point scale simulations offer more flexibility to experiment even with limited computer infrastructure as they are computationally efficient compared to global simulations.

1.2 Problem Statement

Land surface models play a significant role in understanding and projecting the Earth's climate. By simulating the interactions between various earth system components, these models provide insights into how humans, plants, animals, and microbes contribute to and are affected by the Earth's climate. However, despite their significance, LSMs grapple with significant uncertainties, from climate system variability measurement errors to earth system processes and model biases. The complex nature of these uncertainties is further emphasized by the obscured interactions between several factors affecting earth system model performance. For instance, the response of plant productivity, allocation, and turnover to atmospheric CO2 concentrations remains a significant source of uncertainty (Vicente-Serrano et al., 2022), as does the nuanced reaction of soil systems to land-use and land-cover changes (J. Li et al., 2018).

While efforts have been made to minimize uncertainties and improve the land surface model performance through data assimilation, it has been observed that the effectiveness of such approaches varies significantly across regions and vegetation types, further highlighting the complexity of earth system processes. Considering these challenges, this research addresses the complexities of reducing uncertainty in land surface models, specifically focusing on the Community Land Model version 5 (CLM5). By adopting a systematic methodology that integrates perturbed parameter experiments (PPE), sensitivity analysis, and machine learning, we aim to enhance the accuracy of land surface model predictions and bridge the gap between point-scale calibration and its application to larger-scale simulations. The core hypothesis guiding this study posits that point-scale calibrations improve land surface model performance at a regional scale. The study seeks to answer the following research questions to address this hypothesis.

1) To what extent can we quantify carbon-water cycle projection uncertainties in CLM5 using perturbed parameter experiments?
2) What key processes drive uncertainties in the carbon-water cycle projections in CLM5?
3) How do site-specific calibrations perform when transferred and applied to larger-scale (continental) CLM5 simulations?
4) Can machine-learning parameter optimization techniques yield optimized parameter values that bridge the gap between CLM5 simulated and observed carbon-water cycle processes?

To answer the first question, we performed exhaustive perturbed parameter experiments using CLM5 for three National Ecological Observatory Network sites within the southeastern United States. This was realized by setting up a single-site simulation for each site using different parameter configurations and evaluating the simulation outputs (gross primary productivity and latent heat flux) against NEON observations. The simulation's performance was evaluated using three performance metrics: root mean squared error (rmse), mean bias error (MBE), and R-squared (R^2). These metrics were standardized to a value between zero and one, with values close to one suggesting better performance and values close to zero suggesting poor performance. They were then combined, and the simulations for each site were ranked based on the overall performance according to the standardized metric. To answer the second question, we apply a variance-based global sensitivity to quantify parameter contribution to the overall uncertainty in CLM5 model outputs. This was realized by recycling the already performed model simulations without

necessitating further model runs as the model recycles the existing model input-output combinations to quantify parameter uncertainty. To answer the third question, the single-site CLM5 calibrations (optimized parameters from PPEs) were transferred and applied to a more extensive CLM5 simulation (continental scale). Since CLM5 point scale simulations only apply a single plant functional type (dominant vegetation type), it is essential to understand how this reflects when transferred to a larger scale with heterogeneous land cover. An emulator is designed to emulate land surface processes using CLM5 output to constrain water cycle simulation uncertainties.

1.3 Objectives

The overarching goal of this research is to constrain carbon-water cycle projection uncertainties in Earth system models. The study intends to enhance the accuracy, reliability, and applicability of CLM5 simulations across varying spatial scales by employing a comprehensive methodology that integrates perturbed parameter experiments, sensitivity analysis, and machine-learning-based parameter optimization techniques. By refining the key drivers contributing to overall uncertainties and their interactions, current research seeks to advance the field of Earth system modeling and contribute to more robust climate projections.

1.3.1 Specific Objectives

1) Point-scale simulations configurations: Initiate CLM5 point-scale simulations for three National Ecological Observatory Network (NEON) sites – TALL, ORNL, and OSBS. These simulations will serve as the baseline for subsequent analyses.
2) Parameter Selection: Review existing literature to identify a subset of parameters that influence carbon-water cycle processes within CLM5.
3) Perturbed Parameter Experiments Implementation: Perform parameter perturbation experiment individually at each of the selected NEON sites by sampling the identified parameter set using Latin Hypercube sampling techniques and conducting CLM5 simulations across the sampled parameter space to assess the impacts of varied parameter configurations.
4) Ensemble performance evaluation and selection: Assess the performance of each ensemble member from the perturbation experiments using targeted performance metrics to identify

the top 10% of simulations that exhibit superior performance concerning carbon-water cycle processes, specifically latent heat flux and gross primary productivity.

5) High-resolution simulation at a Larger Scale: Perform and execute a high-resolution CLM5 simulation encompassing a more considerable spatial extent using parameter values derived from the top-performing simulations. Evaluate the model's performance against reference datasets to assess the model's fidelity in capturing carbon-water cycle dynamics.
6) Emulator development: Construct a machine-learning-based emulator leveraging parameter values from the perturbation experiments as inputs and mean monthly evapotranspiration as the output. The goal is to establish an efficient means of constraining model parameters and minimizing water-cycle projection biases. Subsequently, validate the effectiveness of parameter optimization using an emulator by performing CLM5 simulation.

1.4 Thesis Organization

The current work is designed to address carbon-water cycle projection uncertainties in land surface models through parameter perturbation experiments and machine learning techniques. The work is grouped into two independent but related studies. Chapter two outlines the implementation of the parameter perturbation experiment, which is outlined by specific objectives 1 through five. The chapter explores a Parameter Sensitivity Study of the Community Land Model at three National Ecological Observatory Network Tower Sites within the Southeastern United States. Chapter three presents a machine learning algorithm (advancing Land Surface Modeling using Gaussian Process Emulators) to constrain carbon-water cycle projection uncertainties. Summary and Conclusions are presented in Chapter 4. Chapter five lists all the literature sources used in the current study. Chapter six summarizes the CLM5 parameters used in this work.

Chapter Two: Perturbed Parameter experiments using CLM5 at three NEON flux tower sites in the Southeastern US.

2.1 Introduction

Land surface models' representation of vegetation dynamics significantly contributes to uncertainties in the carbon-water cycle projection. For instance, the leaf area index (LAI), which regulates water, carbon dioxide (CO2), and energy flux exchanges between plants and the atmosphere, is usually estimated as it is particularly challenging to measure at the ecosystem scale (Ryu et al., 2012). Its estimation is highly uncertain, especially in regions with diverse vegetation and complex topography (Jin et al., 2017). As a result, significant discrepancies exist between model simulations and observed growth and dynamics of LAI, further exacerbating the uncertainties in carbon-water cycle projections. For example, the "LAI paradox" is explained by differences in simulated and observed LAI where many ecosystem models predict an increase in LAI with increasing atmospheric CO2 concentrations, while satellite measurements of LAI from the last few decades show little to no increase in many regions of the world, suggesting a mismatch between model predictions and real-world observations(Fang et al., 2019; Fisher et al., 2019; Winkler et al., 2019)

These uncertainties and biases in vegetation dynamics representation simulated by the Community Land Model version 5 (CLM5) (Lawrence et al., 2019) intensify the uncertainty in carbon-water cycle projections. Recent studies have demonstrated that vegetation dynamics representation in CLM5 is particularly uncertain and biased, especially in regions with high vegetation heterogeneity. For instance, Singh et al. (2020) found that CLM5 overestimates the observed LAI trend within the Southeastern region of the United States. Their study compared the CLM5 LAI trend against the AVHHR-LAI trend from 1982 to 2010. CLM5 shows a trend of 0.13 m^2/m^2 per decade (statistically significant at $p < 0.05$), while AVHHR-LAI shows a trend of 0.03 m^2/m^2 per decade (not statistically significant at $p = 0.11$) (Singh et al., 2020). The CLM5 LAI trend significantly differs from observations, potentially suggesting inadequate vegetation representation in the model. These findings highlight the need for a better understanding and representation of the underlying processes and parameters that control vegetation dynamics, including competition for light, nutrients, water, and the effects of disturbances.

Adequate representation of vegetation dynamics in land surface models will help understand how vegetation responds to elevated atmospheric CO2 concentrations, significantly impacting the future trajectory of global climate change. The Free Air CO2 Enrichment (FACE) experiment is a well-established experimental approach to examine vegetation response to elevated CO2 concentrations (Ainsworth & Long, 2005). In a study to assess the performance of CLM5 in simulating LAI, Wieder et al. (2019) found a considerable positive bias in CLM5 simulations of LAI growth in response to elevated CO2 concentrations against the FACE experiment, indicating uncertainty in capturing the physiological responses of vegetation to elevated CO2 concentrations. This uncertainty can overestimate terrestrial ecosystems' potential carbon sink capacity under future climate scenarios.

Reducing uncertainty in carbon-water cycle interactions is essential for using earth system models for water cycle projections. These models face significant challenges due to the many processes and parameters that control these interactions. The many processes and parameters represented in these models present a computational challenge, especially global-scale parameter calibration and sensitivity experiments, making it difficult to improve model performance. Fisher and Koven (2020) highlight the computational challenges these models face, particularly in the context of global-scale parameter calibration and sensitivity experiments and acknowledge that the multitude of processes and parameters represented in these models contribute to the complexity and difficulty in improving model performance.

Numerous calibration efforts have been undertaken to improve the performance of various ESMs by adjusting model structure configurations, including parameter tuning. These calibration studies focus on a subset of model parameters, given the model's complexity, and aim to assess the model's ability to reproduce observed land surface states and fluxes. A study by Yan et al. (2023) explored CLM5 parameterization and uncertainties with 15 hydrological parameters using five meteorological datasets across the United States. Their findings led to the development of valuable datasets, including benchmark performance evaluation, parameter sensitivities, and large-ensemble outputs, which can aid in further calibration of CLM5 and support various applications, such as drought and flood vulnerability assessment. The realized datasets provide significant insights into the hydroclimatological conditions that drive hydrological predictions. They also

highlight areas for further research on the interaction between hydrological prediction uncertainties and other Earth system processes of specific parameters and their impact on model performance.

The adequacy of vegetation representation in CLM5 through parameter tuning has been explored. Kennedy et al. (2019) explored the representation of vegetation dynamics through parameter calibration by examining plant hydraulic architecture, stomatal conductance, and root distribution to improve plant dynamics representation in land surface models. The results imply that calibration improves model simulations of water fluxes and plant productivity. Similarly, Gao et al. (2021) conducted a calibration study on the CLM's soil hydraulic properties, demonstrating the potential for reducing uncertainty in soil moisture simulations. These studies illustrate the value of parameter calibration in improving model performance and reducing uncertainties in specific aspects of carbon-water cycle interactions.

Considering these challenges, this research addresses the complexities of minimizing uncertainty in earth system models, specifically focusing on the Community Land Model version 5. The goal is to improve the accuracy of land surface models by adopting a systematic approach (Fig. 2.1 below) that combines perturbation parameter experiments and sensitivity analysis. The goal is to bridge the gap between point-scale calibration and its application to regional-scale simulations. The core hypothesis guiding this study posits that point-scale calibrations improve land surface model performance at a regional scale. The study seeks to answer the following research questions to address this hypothesis.

1. To what extent can we quantify carbon-water cycle projection uncertainties in CLM5 using perturbed parameter experiments?
2. What key processes drive uncertainties in the carbon-water cycle projections in CLM5?
3. How do site-specific calibrations perform when transferred and applied to larger-scale (continental) CLM5 simulations?

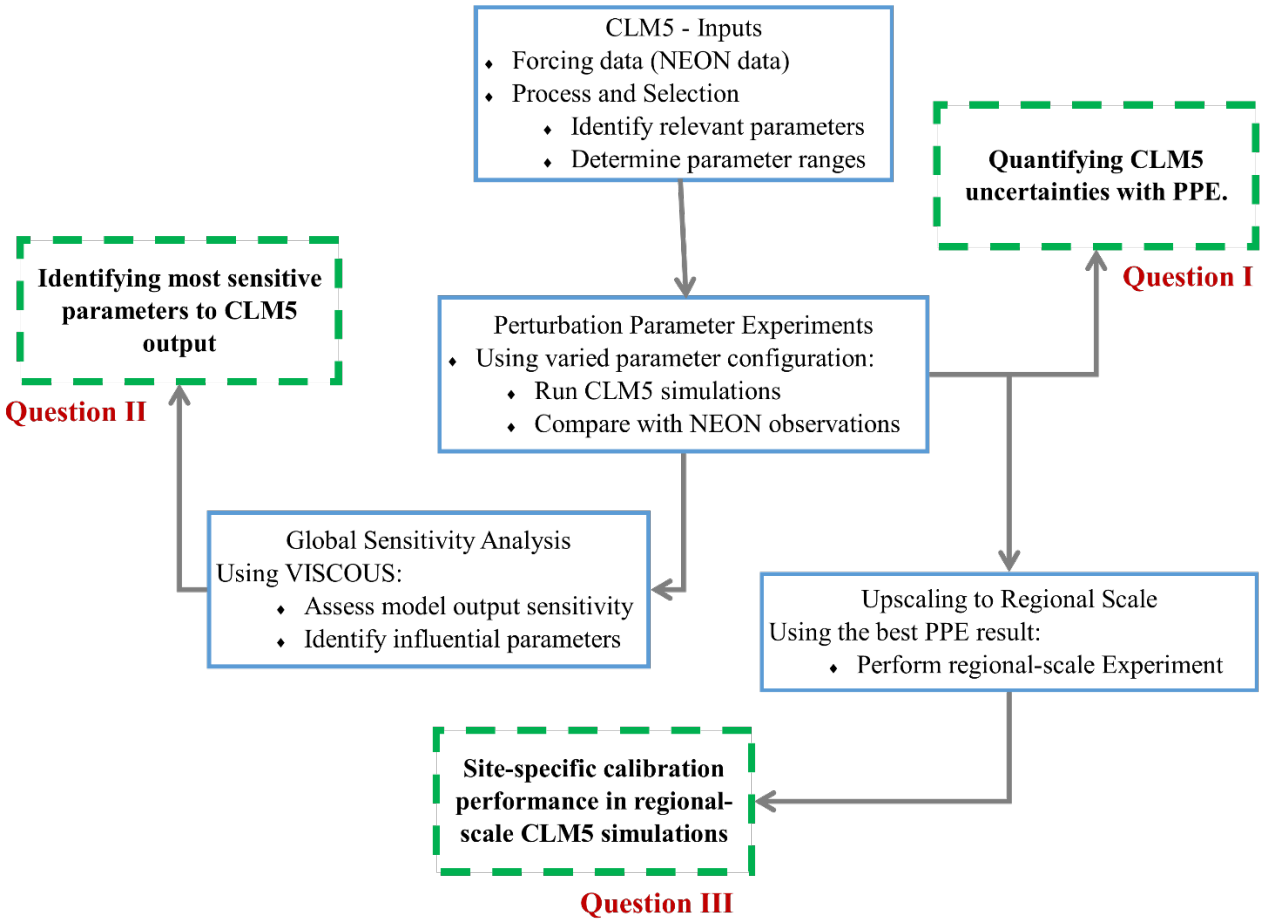

Figure 2.1: A data science workflow for CLM5 parameter uncertainty quantification

To answer the first question, we performed exhaustive perturbed parameter experiments using CLM5 for three National Ecological Observatory Network sites within the southeastern United States. This was realized by setting up a single-site simulation for each site using different parameter configurations and evaluating the simulation outputs (GPP and ET) against NEON observations. The simulation's performance was evaluated using a composite metric derived from three performance metrics: root means squared error (rmse [eq. 2.1]), mean bias error (MBE [eq. 2.2]), and r-squared (r^2 [eq. 2.3]). The composite metric [eq. 2.6] is between zero and one, with values close to one suggesting better performance and values close to zero suggesting inferior performance. The simulations are then ranked for each site based on their performance against NEON observations (gross primary productivity and evapotranspiration). To answer the second question, we apply a variance-based global sensitivity to quantify parameter contribution to the overall uncertainty in CLM5 model outputs. This was realized using the PPEs inputs (parameter

configurations) and outputs (CLM5 evapotranspiration and gross primary productivity) without necessitating further model runs as the applied methodology recycles the existing model input-output components to quantify parameter uncertainty. The single-site CLM5 calibrations (optimized parameters from PPEs) were transferred and applied to a regional scale to answer the third question. Since CLM5 point scale simulations only apply a single plant functional type, it is essential to understand how this affects model performance when transferred to a regional scale with heterogeneous land cover characteristics.

2.2 Materials and Methods

2.2.1 Study Area

The current study is conducted within the southeastern United States (Fig. 2.2), a region ideal for examining carbon-water cycle projections uncertainties in land surface models. The area features diverse soils, including coastal sands and nutrient-rich loam in river valleys, supporting various agricultural practices. Forested land dominates land use activities within the region, with deciduous forests (21%) and evergreen forests (13%), mixed forests (7%), and wetland vegetation (14%), and agricultural lands account for about 26% of the total land area (Dewitz & U.S. Geological Survey, 2021).

The region's climate is humid subtropical, with hot and humid summers, temperate winters, and even precipitation distribution. The topography ranges from coastal plains to hilly Piedmont, the Appalachian Mountains, and fertile river valleys. These diverse characteristics are primary for understanding various ecological processes. Our study is specifically designed for three National Ecological Observatory Networks (NEON) ecoclimatic domain regions: Ozarks Complex (D08), Appalachians (D07), and Southeast (D03).

NEON divides the continental U.S., Hawaii, and Puerto Rico into 20 ecoclimatic domains, each representing distinct ecological and climatic regions. The network collects data on various ecological components, including plants, animals, soil, nutrients, hydrology, and the atmosphere through sensors and field observations. NEON operates 81 strategically located field sites across these ecoclimatic domains and provides free open access to data supporting local to continental-scale research. The data collection strategy allows for local and regional analyses, addressing topics such as biogeochemistry and ecohydrology. NEON's standardized data collection methods

and open-access approach contribute to a comprehensive understanding of ecosystem processes, supporting long-term studies of climate impacts and ecological changes.

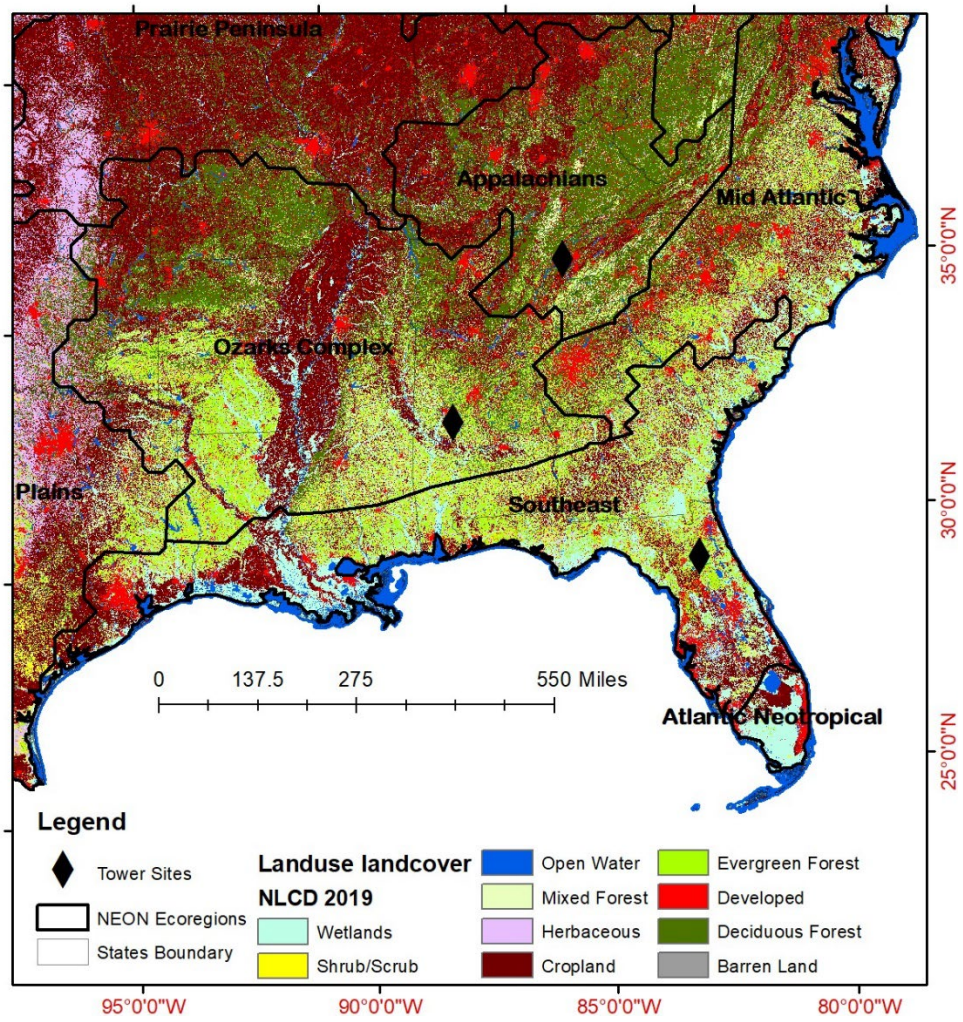

Figure 2.2: The location of the study area showing landcover characteristics according to the National Land Cover Dataset (NLCD) 2019.

Talladega (TALL) is a terrestrial NEON site that is part of the Ozarks Complex ecoregion (D08) located within the Talladega National Forest in west-central Alabama and covers an area of 52.3 km2. This site experiences a subtropical climate with hot summers and mild winters, receiving an average annual precipitation of 1380 mm. The geology consists of marine sediments from the Cretaceous age Tuscaloosa Group. Soils in the area are primarily sandy, clayey, and loamy. The hydrology of the site is typical of that of forested streams. Vegetation is dominated by conifers, mainly longleaf pine (Pinus palustris) and loblolly pine (Pinus taeda), with some hardwoods and

wetlands. NEON site characterization at TALL was completed in 2014, with all operations starting in 2016 (NEON (National Ecological Observatory Network), n.d.-c).

The Oak Ridge National Laboratory (ORNL) terrestrial site extends an area of 57.4 km^2. It is located in the Cumberland Plateau in Anderson and Roane Counties, TN, within the Appalachians and Cumberland Plateau ecoregion (D07). The region's climate is characterized by moderate temperature variations and abundant precipitation (mean annual precipitation 1340 mm), with hot and humid summers and mild winters. The region's geology consists of limestone, Cambrian-aged dolostone, and some parent materials. The region is dominated by the Fullerton cherty silt loam soils (Jennifer, 2017). Its vegetation includes hardwood forests dominated by oaks and hickories, deciduous species, shortleaf pine, and Virginia pine. Initial site establishment began in 2014 but was fully operational in the summer of 2015 for instrumentation sampling and terrestrial observations (NEON (National Ecological Observatory Network), n.d.-a).

The Ordway-Swisher Biological Station (OSBS) terrestrial site (38.5 km^2) is in Putnam County, Florida. The site is within the NEON Southeast ecoregion (D03), an ecologically diverse region bounded by the Atlantic Ocean and the Gulf of Mexico, featuring a mix of forest, grassland, and wetland communities. The site's climate is characterized by hot and humid summers and mild, drier winters, with annual precipitation averages around 1302 mm. Its geology consists of Miocene and Pliocene formations, including limestone, dolomite, sand, mud, clay, and quartz sands. Soils are primarily well-drained and sandy, with some organic soils in low-lying areas, especially near lakes and swamps. Vegetation is dominated by pine and turkey oak, with Longleaf Pines and Lobolly Pines being prevalent, along with grass and forb groundcover. Site establishment began in 2012, and data collection began in 2013 for most systems but was fully functional in August 2017 (NEON (National Ecological Observatory Network), n.d.-b).

2.2.2 The NCAR-NEON System

The NCAR-NEON system (D. L. Lombardozzi et al., 2023) integrates site-specific and time-resolved NEON's observational data into the CLM5 model. As described (section 2.2.1), NEON collects quality data of high-resolution and high frequency across multiple ecological sites all over the United States, which is critical to understanding various ecosystem processes. These datasets are significant for modeling and simulating the complexities of the carbon-water cycle processes. The current research employs the NEON-NCAR system to comprehensively investigate

carbon-water cycle processes at three NEON ecoclimatic domains. The NEON-NCAR system integrates NEON observational datasets into the land surface model (CLM5), providing an opportunity to evaluate model performance against observational data. This presents opportunities for model calibration and parameterization with the potential of guiding land surface modeling at a localized scale.

Land surface models (CLM5) can be configured for various scales, including global, regional, and local (single site). When configured for a single site, the model is adjusted to match the site-specific characteristics, including dominant vegetation, soil type, and climate. The NCAR-NEON site has already processed surface characteristics data for each NEON site, and by defining the specific site you want to analyze, the model appends these datasets automatically. To capture vegetation characteristics, CLM uses plant functional types. The three sites identified for the current study represent two distinct pfts, with TALL and OSBS sharing the needle leaf evergreen ecosystem while the ORNL represents the deciduous forest. From the site characteristics (section 2.1), these two ecosystems represent about 35% of the total land cover area under study (Dewitz & U.S. Geological Survey, 2021). We used the NCAR-NEON system to perform a Parameter Perturbation Experiment (PPE) by initializing the model and a four-year production run at each site.

The NCAR-NEON simulation system provides a new opportunity to experiment with parameter sensitivity to various model processes. The tool is computationally efficient, with point-scale simulations being 15,000 times or more computationally efficient than global simulations. The system leverages the quality control observations of water, energy, and carbon exchange fluxes, including evapotranspiration and gross primary productivity from NEON flux tower sites. By selecting a sub-set of 30 parameters (Table 2.1) and generating 400 parameter sets using Latin hypercube sampling methods, we could perform an exhaustive simulation of 244,800 years (3 sites × 400 parameter sets × 200 spinup +4 production years). This level of simulation would not have been feasible without the ultra-computational efficiency of the NCAR-NEON System.

Table 2.1: Selected CLM5 parameters and their description.

Parameter name	Description and units
nstem (pft)	Stem number: number of individuals per meter squared (like stocking number) [#/m^2]
leafcn (pft)	Leaf carbon to nitrogen ratio ($gCgN^{-1}$)

Parameter name	Description and units
stem_leaf (pft)	Stem: leaf ratio (gC/gC)
froot_leaf (pft)	Allocation parameter: new fine root C per new leaf C (gC/gC)
FUN_fracfixers (pft)	Fraction of N fixers
aq_sp_yield_min	Minimum specific yield
baseflow_scalar	The scalar multiplier for the base flow rate
fff	Decay factor for the fractional saturated area (m^{-1})
hksat_sf	The scalar adjustment factor for hksat
liq_canopy_storage_scalar	Maximum storage of liquid water on the leaf surface (kg/m^2 mm)
maximum_leaf_wetted_fraction	The maximum fraction of leaf that may be wet before drip occurring
sucsat_sf	The scalar adjustment factor for sucsat
watsat_sf	The scalar adjustment factor for watsat
zbedrock_sf	The scalar adjustment factor for zbedrock
wc2wjb0	the baseline ratio of rubisco limited rate vs. light limited photosynthetic rate (Wc: Wj)
jmaxb1	determines the response of electron transport rate to light availability
jmaxb0	Baseline proportion of nitrogen allocated for electron transport (J)
slatop (pft)	The specific leaf area measured at the canopy top (m^2/gC)
kmax (pft)	Plant segment maximum conductance (s^{-1})
krmax (pft)	Root segment maximum conductance (s^{-1})
rootprof_beta (pft)	Rooting beta parameter for carbon and nitrogen vertical discretization
psi50 (pft)	Water potential at 50% loss of conductance (mm H_2O)
cv	Turbulent transfer coefficient between canopy surface and canopy air($ms^{-1/2}$)
d_max	Dry surface layer (DSL) parameter (mm)
dleaf (pft)	The characteristic dimension of leaves in the direction of wind flow (m)
frac_sat_soil_dsl_init	Fraction of saturated soil for moisture value at which DSL initiates
act25	The specific activity of Rubisco at 25°C (µmol (CO_2)/ (gRubisco s)
medlynintercept (pft)	Medlyn intercept of conductance-photosynthesis relationship (µmol H_2O/ m^2s))
medlynslope (pft)	The slope of stomatal conductance-photosynthesis relationship (µmol H_2O µmol CO_2^{-1})
n_baseflow	Drainage power exponent

2.2.3 Model Description

The Community Land Model Version-5 (CLM5) is a versatile tool that can be run at both large scales (region and global) Lawrence 2019 and at a single-point configuration mimicking an eddy covariance tower's footprint (Bonan et al., 2021; Lombardozzi et al., 2012; Wieder et al., 2017). The model's adaptability and ability to simulate localized ecological phenomena make it valuable for advancing ecological understanding and addressing region-specific research questions. We performed CLM5 point-scale BGC simulations using the new NCAR-NEON system that enables seamless integration between the model and NEON datasets.

The NCAR-NEON System (Lombardozzi et al., 2023) integrates site-specific and time-resolved NEON's observational data into the CLM5 model. NEON collects quality data of high-resolution and high frequency across multiple ecological sites all over the United States, which is vital to understanding various ecosystem processes. These datasets are significant for modeling and simulating the complexities of the water and carbon cycle processes. Our research employed this tool to comprehensively investigate carbon-water cycle processes at three flux tower sites. By integrating NEON's rich observational datasets into the CLM simulations, we aimed to constrain uncertainties and gain a deeper understanding of these critical ecosystem dynamics. Using the NCAR-NEON tool, we conducted perturbation parameter experiments (PPE) by initializing the model, followed by a four-year production run at each site.

2.2.4 Experimental Design and Implementation

The study is designed to quantify and minimize uncertainty in carbon-water cycle projection in CLM5 with a specific focus on gross primary productivity and evapotranspiration. Considering the challenges associated with land surface models' global scale parameter calibration and sensitivity analysis studies, the point-scale experiments exploited the NCAR-NEON system (Lombardozzi et al., 2023) to constrain CLM5 uncertainty. Specifically, we focused on the Southeastern United States, including Talladega National Forest (TALL - longleaf pine ecosystem), Oakridge National Laboratory (ORNL - hardwood deciduous forest), and Ordway-Swisher Biological Station (OSBS - longleaf pine ecosystem) sites.

The study employs two different model configurations: a parameter perturbation experiment to quantify overall model uncertainty and use the knowledge gained to perform a regional scale simulation using the best-performing simulation parameters. Initializing CLM5 state variables was realized by spinning the first four years of data (1st January 2018 to 21st December 2021) for 200 years for the sampled parameter set as recommended by (Wieder et al. 2019).

The perturbation parameter experiment setup includes a set of 400 parameter files generated using the Latin Hypercube technique sampled from the subset of 30 parameters. The simulations are based on standard 2000 cases with fixed CO2 concentrations (408.83 ppm CO2) and validated against NEON observations (gross primary productivity and latent heat fluxes). We performed an exhaustive analysis by running 244,800 years (3 sites × 400 parameter sets × (200 spinup +4 production) years) simulations. This level of simulation would not have been feasible

without the ultra-computational efficiency of the NCAR-NEON Tool. The simulation's performance is evaluated against NEON observations to identify the best parameter configuration that captures NEON observations.

We used three evaluation metrics to evaluate the model's performance: root mean square error (RMSE) [eq. 2.1], mean bias error (MBE) [eq. 2.2], and coefficient of determination (R) [eq. 2.3]. These performance evaluation metrics compare CLM5 model output against NEON observations for carbon-water cycle processes, including evapotranspiration and gross primary productivity. By comparing the model results with NEON observations, we identify the top 10% best-performing models. We extracted the best model from this subset, which exhibited the most accurate representation of evapotranspiration and gross primary productivity.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(S_i - O_i)^2} \tag{2.1}$$

$$MBE = \frac{1}{n}\sum_{i=1}^{n}(S_i - O_i) \tag{2.2}$$

$$R^2 = 1 - \frac{Sum\ squared\ residual\ (SSR)}{Total\ sum\ of\ squares\ (SST)} \tag{2.3}$$

The RMSE and MBE metrics were standardized to 0 to 1 [eq. 2.4 & 2.5]. In this standardized scale, values close to 1 indicate the best model performance, while values close to zero suggest poor model performance. The standardized RMSE and MBE values and R-squared (R^2) were then combined and averaged to obtain a single performance evaluation metric that also falls within the 0 to 1 range.

Applying a composite metric that combines different performance measures such as RMSE, MBE, and R^2 offers a comprehensive evaluation of model performance compared to a single specific metric. Performance evaluation metrics provide various insights, especially when comparing simulations against observations. For instance, RMSE measures accuracy, MBE determines systematic error, while R^2 indicates the variance proportion explained by the model. Combining all these different views presents a more holistic view of how the simulated data compares with observational data.

A composite metric allows a tradeoff balance between bias, accuracy, and explanatory power. Using a single performance metric does not provide this flexibility and cannot capture such tradeoffs. A model that excels in one metric but fails in others might not be ideal for practical applications, and combining different performance measures helps address this issue and minimize overfitting risk to a specific measure.

A single performance measure gives a limited perspective and lacks the adequacy of capturing model behavior. Therefore, applying a composite metric helps address the limitations of using a single metric, leading to a more robust model selection and better alignment with observations. To evaluate the robustness of our composite performance evaluation metric, we compared our results with existing model evaluation metrics such as t-statistic (Jacovides & Kontoyiannis, 1995) and Kling-Gupta Efficiency model (Gupta et al., 2009). Our composite performance metric [eq. 2.6] offers a comprehensive evaluation of model performance. We give equal importance to all three metrics while computing the M_composite value. The emphasis on a specific metric depends on the objective of the study. For instance, when focusing on predictions – you will put more weight on RMSE and R^2, and when the main focus is interpretability, you will emphasize the MBE (Brovkin et al., 2013; Jung et al., 2017). A higher M_composite value indicates that the simulations performed exceptionally well, while a lower value suggests poor performance.

Standardized RMSE:

$$RMSE_{std} = 1 - \left(\frac{RMSE - RMSE_{min}}{RMSE_{range}}\right) \tag{2.4}$$

Standardized MBE:

$$MBE_{std} = 1 - \left(\frac{MBE - MBE_{min}}{MBE_{range}}\right) \tag{2.5}$$

Composite Performance Metric:

$$M_{composite} = \left(\frac{R^2 + RMSE_{std} + MBE_{std}}{3}\right) \tag{2.6}$$

RMSEmin refers to the minimum RMSE value, and $RMSE_{range}$ refers to the difference between maximum RMSE and minimum RMSE values. If RMSE=$RMSE_{min}$, we have a good model; when RMSE = RMSEmax, it represents a poor-performing model.

To examine the most influential parameters driving carbon-water cycle projections uncertainties in CLM5, we performed a sensitivity analysis using a variance-based sensitivity analysis technique using data generated from the parameter perturbation experiment. The applied method uses existing data (CLM5 inputs – parameter configurations and CLM5 outputs – GPP and ET) to measure the contribution of each parameter and interaction among different parameters to the overall uncertainty of model output.

The best input parameter configuration (optimized parameters) obtained from the parameter perturbation experiment is then used to configure the CLM5 simulation at a regional scale. The CLM5 is prepared and executed for the continental United States (25-53N, 235-293E) using a data-driven high-resolution (0.125-degree) NLDAS2 dataset. The model is initialized for 460 years by recycling the first thirty-eight years (1981 – 2018) and a production run from 1981 – 2021 using transient CO2. The model performance is evaluated against reanalysis datasets obtained from the ILAMB package.

2.2.5 The Parameter Perturbation Experiment

The Parameter Perturbation Experiment (Fig. 2.3) was set using a subset of 30 parameters (Table 1) affecting various carbon-water cycle processes, including Photosynthesis (7), Hydrology (10), Plant Hydraulics (4), C/N Allocation (3), Biomass Heat Storage (1), Fix/Uptake Nitrogen (1), and Surface Fluxes (4). The parameters are sampled using the Latin hypercube sampling technique to generate 400 parameter configurations.

The Latin hypercube sampling technique is a stratified Monte Carlo (McKay et al., 2000). The method has been proven to generate results with less variability from a smaller sample size than other sampling techniques, such as the conventional Monte Carlo (Williamson, 2015). Latin hypercube has a unique way of extracting samples that ensures better representation and has been widely used in both climate and hydrological modeling studies (Gray et al., 2016; Loeppky et al., 2009; Lucas et al., 2013; Murphy et al., 2006; Pang et al., 2021; Williamson, 2015). The method partitions the sampling region by splitting the parameter ranges. The Latin hypercube sampling technique rule of thumb states that a sample size ten times the number of the design parameters is

sufficient to model the population mean (Loeppky et al., 2009). It has an advantage in its efficiency in saving computer processing time and adequacy in sampling the parameter space over the conventional Monte Carlo technique.

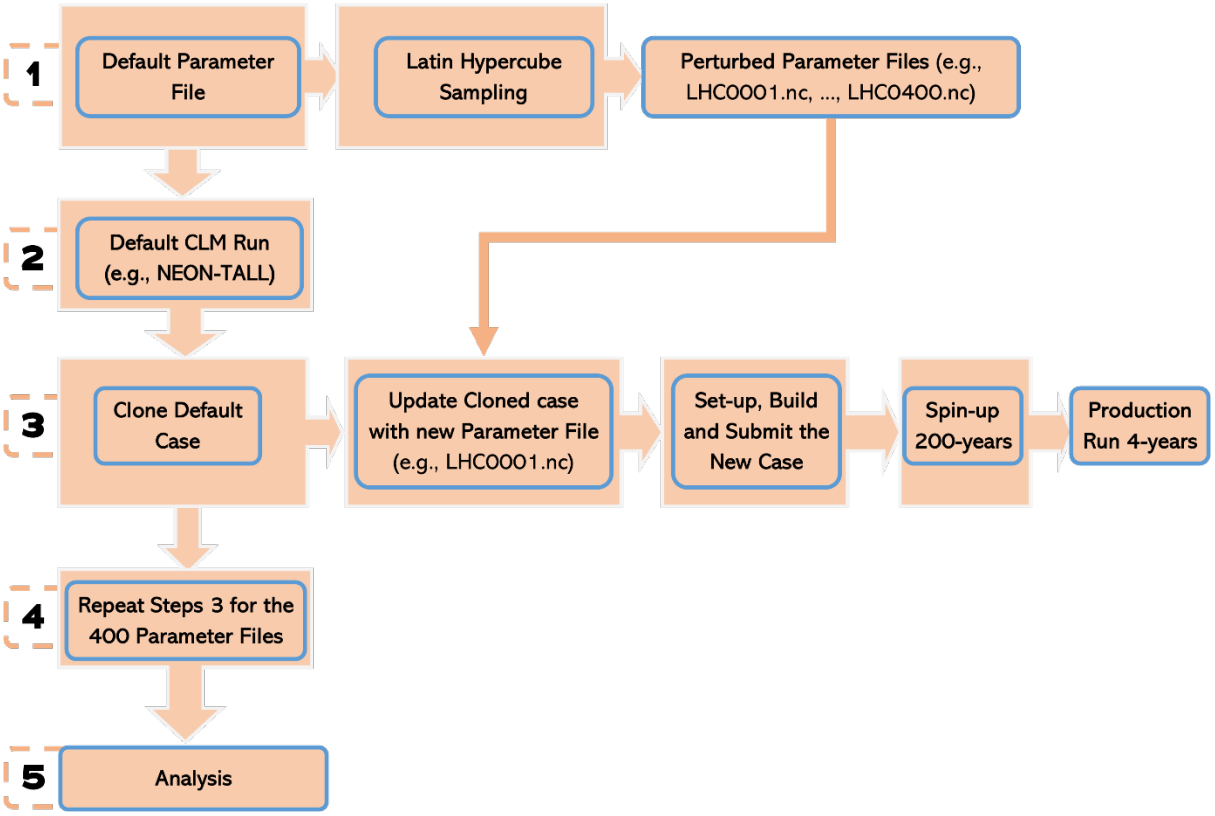

Figure 2.3: A flowchart summarizing the steps for implementing the CLM-NEON Parameter perturbation experiment.

The generated parameter samples are used to update the default CLM5 parameter file to create 400 parameters for the PPE approach. The implementation of the PPE approach includes CLM5 model initialization, which for our case, was run for 200 years by cycling over the initial four years of forcing data (2018 - 2021) to establish initial conditions Lawrence 2019 The simulations are based on standard year 2000 cases with fixed CO2 concentrations (408.83 ppm CO2) and are validated against observations with a focus on gross primary productivity, and latent heat fluxes. Each parameter file performs a CLM5 simulation for the Three NEON sites with a 200-year model initialization and 4-year production run (2018 – 2021). The simulation's performance is evaluated using the three defined performance evaluation metrics (RMSE, MBE,

and R^2) to identify the optimal conditions for gross primary productivity and evapotranspiration in CLM5.

2.2.6 Parameter Sensitivity Analysis

A variance-based global sensitivity analysis technique is applied to break down the uncertainty in carbon-water cycle process simulations in CLM5. The sensitivity analysis helps quantify the contributions of each of the 30 parameters and their interactions to the overall uncertainty. We refer to parameter sensitivity as how each parameter influences the observed variance in carbon-water cycle processes. We calculate two sensitivity measures: the main effect and the total effect. The main effect represents the percentage of total variance (uncertainty) in carbon-water cycle processes that can be reduced if a specific parameter is precisely known. The total effect encompasses all variance components associated with the parameter, including the main effect and the interaction effect (how the parameter interacts with all other parameters), expressed as a percentage of the total variance. We can assess how sensitive carbon-water cycle processes are to interactions between uncertain parameters by analyzing these two sensitivity measures. The primary and total effect measures become equivalent if no interactions exist between a specific parameter and other parameters.

Our analysis uses a variance-based global sensitivity analysis using the copulas (VISCOUS) approach (Razi et al., 2021). The choice of this sensitivity analysis technique lies in its ability to recycle existing data to quantify parameter uncertainties to carbon-water cycle projections in CLM5 without the need to run new simulations. It is a non-parametric technique that uses few assumptions and does not require a sample from a specific distribution (Razi et al., 2021). The approach only requires knowledge of conditional probability densities and the capacity to sample from the same in which the Gaussian copulas technique generates the conditional samples. Here, we summarize some critical steps implemented for the sensitivity analysis, but a detailed description of how the method works is outlined (Razi et al., 2021).

The VISCOUS estimates the sensitivity index by extracting the input-output matrix from the model outputs (y) and inputs (parameters: x) (Razi et al., 2021). The kernel density function transforms the extracted matrix into a marginal cumulative density function (CDF) $u_i = F_{X_i}(x)$ and $F_Y(y) = v$ (*for i = 1, 2, …, m*). The CDF is then transformed into a normal distribution and its inverse distribution is estimated: $\psi^{-1}(u_1), \psi^{-1}(u_2), \psi^{-1}(u_3), \ldots, \psi^{-1}v$. The Gaussian mixture

copula method estimates the probability density function of the copula method (Tewari et al., 2011)

$$c_g(v,\mu;\boldsymbol{\Theta}) = \frac{f_Z(\psi^{-1}(v),\psi^{-1}(u_1),\psi^{-1}(u_2),\dots\ \psi^{-1}(u_m);\boldsymbol{\Theta})}{\emptyset\big(\psi^{-1}(v)\big)\times\emptyset\big(\psi^{-1}(u_1)\big)\ \times\ \emptyset\big(\psi^{-1}(u_2)\big)\times\dots\times\emptyset\big(\psi^{-1}(u_m)\big)} \tag{2.7}$$

Where ψ^{-1} is the standard normal distribution inverse function, f_Z is the Gaussian mixture model probability density function.

$$f_Z(z;\Theta) = \sum_{k=1}^{\eta} \omega_k \times \emptyset\big(z;\mu^k,\Sigma^{(k)}\big) \tag{2.8}$$

Where $z = (\psi^{-1}(v),\psi^{-1}(u_1),\psi^{-1}(u_2),\dots\ \psi^{-1}(u_m)$, ω_k is the weighted Gaussian mixture model *k-th* component, the mean vector is μ^k, the *k-th* component covariance matrix is $\Sigma^{(k)}$, and all the mean vectors, covariance matrices and weights are represented by $\boldsymbol{\Theta}$ determined as

$$\boldsymbol{\Theta} = \big(\omega_k,\mu^k\Sigma^{(k)}\big)\ for\ k = 1,2,\dots,\eta \tag{2.9}$$

The next step of the VISCOUS model is building a Gaussian mixture model (GMM) f_z by fitting η Gaussian density components to the data obtained from the inverse of the standard normal distributions. After determining the global sensitivity analysis indices *(p)*, an equal number of copula models are built to estimate each index. Finally, equations 2.7 and 2.8 build a copula model c_g to compute the sensitivity indices using equations 2.10 and 2.11.

$$E(Y|X-x_c) \approx \frac{1}{N_{MC}}\sum_{j=1}^{N_{MC}} {F_Y}^{-1}\big(v^{(j)}\big)c_g\big(v^j,\boldsymbol{u}_c;\boldsymbol{\Theta}\big) \tag{2.10}$$

$$Var[E(Y|X=x_c)] \approx \frac{1}{2{N_{MC}}^3}\sum_{i=1}^{N_{MC}}\sum_{j=1}^{N_{MC}}\left(\sum_{k=1}^{N_{MC}} {F_Y}^{-1}\big(v^{(k)}\big)c_g\big(v^{(k)},{\boldsymbol{u}_c}^i;\boldsymbol{\Theta}\big) - \sum_{k=1}^{N_{MC}} {F_Y}^{-1}\big(v^{(k)}\big)c_g\big(v^{(k)},{\boldsymbol{u}_c}^j;\boldsymbol{\Theta}\big)\right)^2 \tag{2.11}$$

Where the given random input factor for estimating Sobol's sensitivity indices is $(X = x_c)$, the Monte Carlo integration samples are N_{MC}, and the inverse cumulative distribution density of Y estimated using kernel density is ${F_Y}^{-1}(v)$.

2.2.7 Regional Simulation

The best parameter set from the point scale simulations is evaluated at a regional scale. At each of the three sites, we evaluated all the 400 simulations and using the evaluation metric, we identified the best performing simulation with results showing close agreement to NEON observations. Once identified, the parameter set for the best simulations for each site was extracted and then used for regional-scale CLM5 configuration. These values were used to update the CLM5 default parameter file by replacing the default parameter value with the optimized value for all 30 parameters, including the pft parameters. The regional setup examines whether the calibration and parameterization at the point scale improve the performance of the regional model. The experiment is set up with a standard 2000 case and transient CO2 using a high-resolution CLM5 model with NLDAS-2 forcing datasets for the continental United States. The model is initialized for 460 years by recycling the initial thirty-eight years (1980 – 2018), followed by a production run from 1980 to 2021.

The regional model performance is evaluated using data from the International Land Model Benchmarking (ILAMB) package (Collier et al., 2018). ILAMB is an open-source benchmarking framework that generates graphical diagnostics and scores model performance based on benchmark datasets, allowing for comprehensive assessment across various land variables. We employed data obtained from the ILAMB framework: ECMWF Reanalysis v5 (ERA5) to validate ET and FLUXCOM data to validate for GPP to evaluate the significant contribution of point-scale CLM5 parameter optimization efforts on regional scale simulations.

2.3 Results

2.3.1 Parameter Perturbation Experiments

CLM5 consistently underestimates evapotranspiration and overestimates gross primary productivity. The default CLM5 output significantly differs from the NEON observations (Fig. 2.4). This is supported by the mean bias estimate, which shows that latent heat flux has a negative mean bias (-2.28 W/m²), Standard Deviation (16.6 W/m²) and a mean (56.4 W/m²). In contrast, gross primary productivity has a positive mean bias (2.15 μmol C/m²/s), a Standard deviation (1.14 μmol C/m²/s), and a mean of (3.37 μmol C/m²/s). The high standard deviation, especially for latent heat flux, suggests significant variability in the model's predictions, reflecting the spread of individual latent heat flux measurements around the mean. The smaller standard deviation for GPP

suggests less variability in the model's predictions, with individual measurements clustered more closely around the gross primary productivity mean. Higher standard deviation values indicate more spread-out model errors from the mean bias and suggest higher variability or inconsistency in the model's predictions. A significantly +ve/-ve bias and higher standard deviation are vital indicators for potential systematic issues within the model predictions.

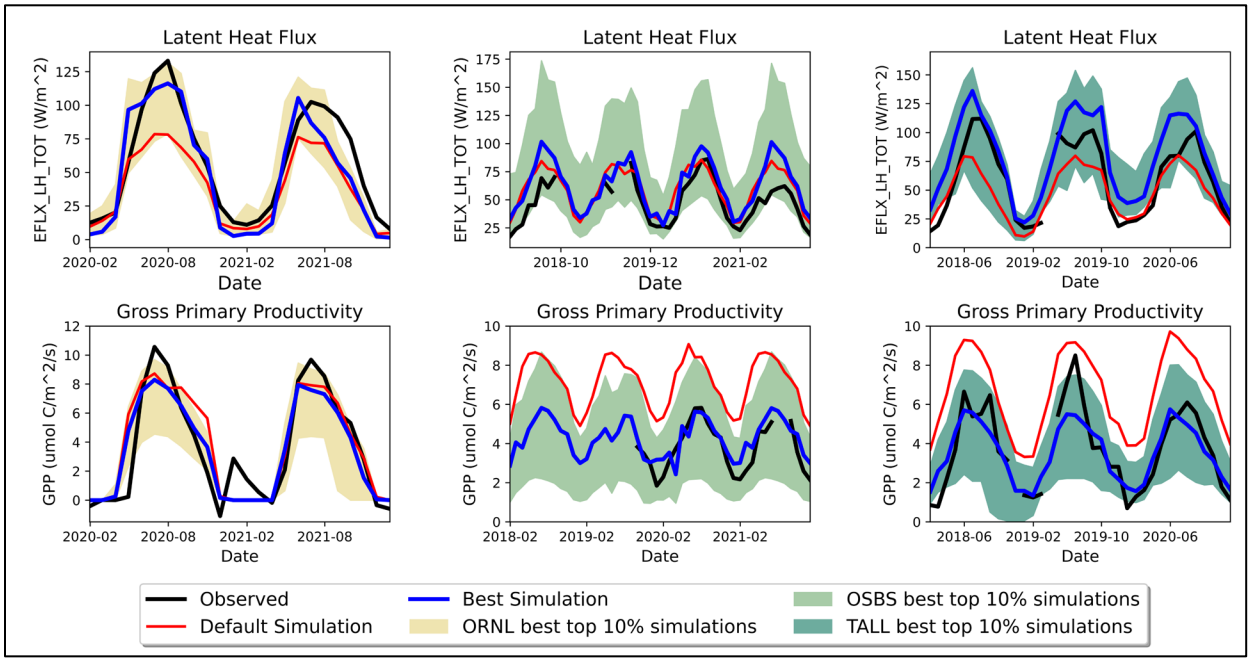

Figure 2.4: CLM5 Calibration of evapotranspiration and gross primary productivity using Parameter Perturbation experiment at three NEON Sites. The shaded region shows the top 10% best-performing simulations at each site.

Carbon-water cycle projection uncertainties can be constrained in CLM5. The PPE simulations were ranked against NEON observations according to their performance on ET and GPP. The PPEs significantly improve CLM5's ability to simulate carbon-water cycle predictions. The performance evaluation metric, which ranks the simulations from best to worst, shows that the CLM5 performance at the TALL site improved from 0.19 (default) to 0.81 (best simulation). Comparable results are also noted for both OSBS and ORNL sites: improving from 0.59 (default) to 0.73 (best simulation) and from 0.33 (default) to 0.8 (best simulation) at both OSBS and ORNL sites, respectively. The top 10% of the best simulations (shaded region Fig. 2.4) from the PPEs show improved performance compared to default simulations.

Comparing model results with observations allows for evaluating model performance across the parameter space explored. The parameter space for the top 10%, the best, and the default parameter values are all compared (Fig. 2.5). Some parameter values from the PPEs are significantly different from the default values and could be constrained. These parameters include hksat_sf, jmaxb0, n_baseflow, leafcn, psi50, rootprof_beta, and baseflow_scalar. This subset represents a constrained parameter space where the best and top 10% simulation outputs significantly deviate from the default parameter value. Constrained parameters improve carbon-water cycle projections in CLM5. The assessment of the top 10% best-performing simulations reveals interesting patterns of these highlighted parameters.

Photosynthesis capacity (LUNA): The *jmaxb0* parameter is crucial in determining photosynthesis capacity in vegetation in CLM5. The parameter represents the allocated nitrogen proportion for electron transport and signifies the maximum rate of electron transport or the potential maximum rate of photosynthesis under optimal conditions. In the default CLM5 configuration, *jmaxb0* is set to 0.03 and ranges between 0.01 – 0.05. However, our research findings suggest this parameter can be constrained to a narrower range. Specifically, our results indicate that *jmaxb0* can be constrained within a range of 0.012 to 0.018. This narrower range may accurately represent photosynthesis capacity under similar conditions explored in the present study.

Hydrological Processes: The *n_baseflow* and *baseflow_scalar* parameters are crucial to determining the movement of water through the soil and subsurface runoff. The *n_baseflow,* also known as the drainage power exponent, regulates the rate of water movement from the soil into deeper layers. It represents the water flow rate through the soil as a function of the hydraulic gradient. Within the CLM5, *n_baseflow* has a default value of 1, and some existing literature defines its range as between 1 and 2. Our study reveals that 75% of the top 10% simulation values for *n_baseflow* are greater than 1.2. This suggests that under the specific conditions of our research, a stronger drainage power exponent could better align with NEON observations. Therefore, water movement through the soil and deeper layers may occur faster than anticipated under default settings within specific contexts. The *baseflow_scalar* is a scalar multiplier for the baseflow rate in CLM5. It is applied when the lower boundary condition is not an aquifer or a water table, as it allows for the adjustment of subsurface runoff. In the default configuration, *baseflow_scalar* has

a default value of 0.001 and typically ranges from 0.0005 to 0.1; our study findings suggest a narrower and more constrained *baseflow_scalar* parameter value between 0.02 and 0.07. Like *n_baseflow* parameter results, higher values of the *baseflow_scalar* parameter could potentially improve the accuracy of capturing the baseflow rate in scenarios where alternative lower boundary conditions are prevalent in CLM5.

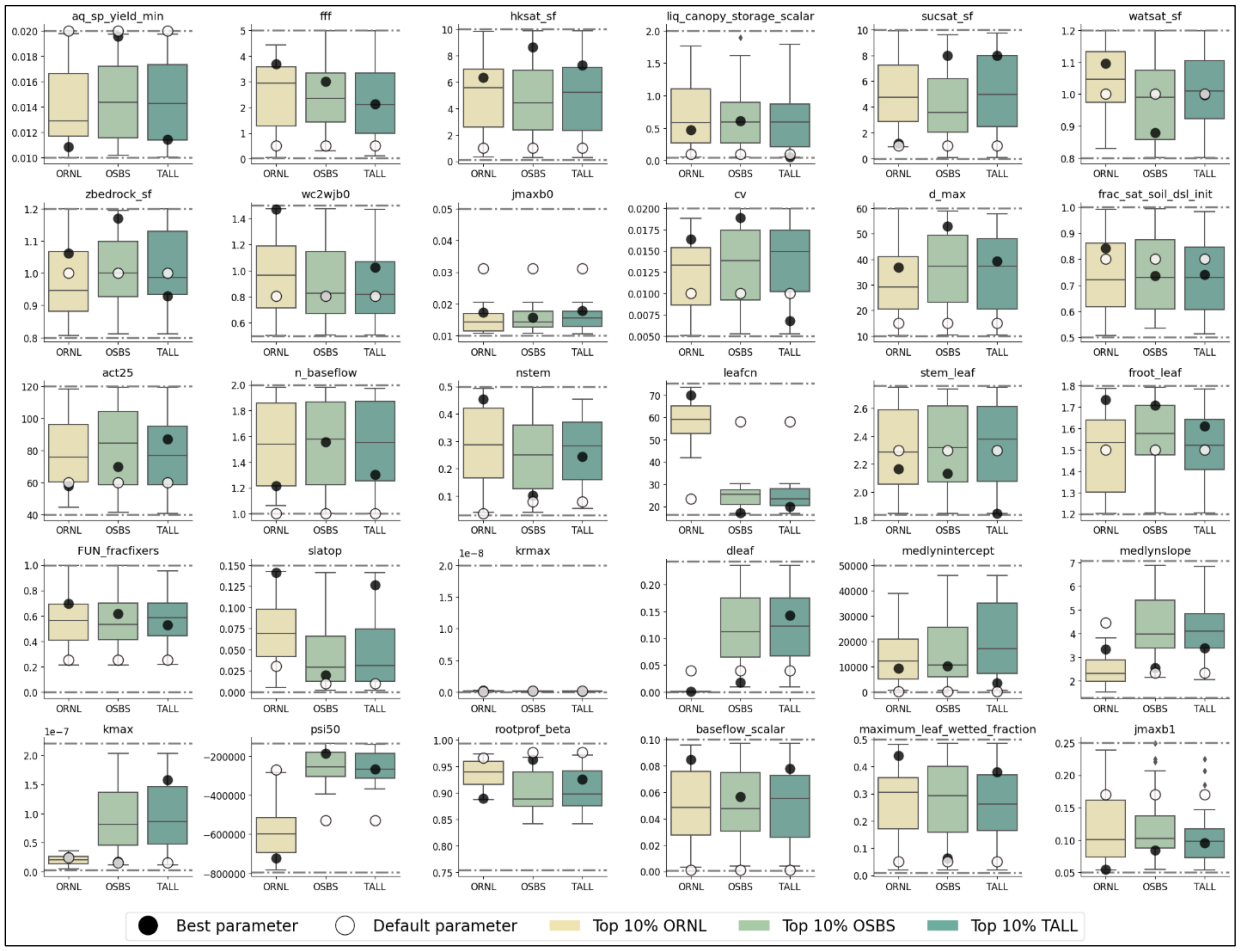

Figure 2.5: NCAR-NEON System parameterization and optimization at three NEON sites, using the parameter perturbation experiment and selected 30 parameters using monthly data. The plot includes these parameters' default values, optimized parameter value, and box-whisker plot of the top 10% best-performing models, which was done by combining three performance metrics: RMSE, MBE, and R^2. The y-axis shows the full range of parameter values based on maximum and minimum values.

The dynamics of soil water movement in the CLM5 are significantly influenced by the saturated hydraulic conductivity, which is controlled by the parameter *hksat_sf*. This parameter serves as a scalar adjustment factor for saturated hydraulic conductivity. It is essential in shaping

soil water movement, plant water uptake, and soil moisture dynamics within CLM5, which is usually assigned a default value of 1. Typically, this parameter is assigned values between 0.1 and 10. Approximately 75% of the top 10% of best simulations have hksat_sf parameter values constrained within a narrower range between 3 and 7, significantly different from the default value of 1. The parameter determines how water can move through the soil, impacting plant water uptake and soil moisture dynamics. Our findings suggest that this parameter can be tuned within a narrower range (3 to 7) in certain environmental conditions, deviating significantly from the default value of 1. This emphasizes the importance of considering specific soil and environmental characteristics when configuring CLM5 to accurately represent soil water movement and associated ecosystem processes.

Fixation and uptake of Nitrogen: The nitrogen fixation sensitivity to temperature changes has significant implications for vital ecosystem processes as it influences nutrient availability, thus affecting ecosystem productivity. For example, a study by Bytnerowicz et al. (2022) discovered that nitrogen fixation adjusts to growing temperatures, with the optimal temperature for nitrogen fixation being 5.2°C higher than that for photosynthesis. The *FUN_fracfixers* represent the sensitivity of nitrogen fixation to temperature variations in CLM5. The *FUN_fracfixers* parameter is essential in nitrogen fixation and uptake processes in CLM5 as it controls the slope of the response curve of nitrogen fixation to temperature changes. Typically, the *FUN_fracfixers* parameter value ranges between 0 and 1, with 0.25 in the CLM5 default configuration. Our results show that the FUN_fracfixers values obtained for the top 10% best simulations cluster within a narrower range of between 0.4 and 0.7. The narrower range observed implies that elevated *FUN_fracfixers* parameter settings could provide a more accurate representation of the temperature-dependent dynamics of nitrogen fixation in CLM5.

Carbon and Nitrogen allocation: The *leafcn* parameter is crucial in determining plant leaves stoichiometry in CLM5. The *leafcn* parameter has significant implications for plant growth and physiology and is influenced by several factors, including plant functional type, atmospheric CO2 concentration, soil nutrient availability, and water availability. It is essential to emphasize that leaf nitrogen is a vital nutrient for plant growth, limiting primary productivity. Our research focused on two distinct plant functional types (PFTs): needleleaf evergreen trees - temperate and broadleaf deciduous trees – temperate. With a default value of 23.4, our results show that the top

10% best-performing simulations for broadleaf deciduous temperate trees had *leafcn* parameter values ranging between 40 and 73. This means that the deciduous temperate trees exhibited a significantly higher carbon-to-nitrogen ratio in their leaves compared to the default setting. Our findings align with expectations as deciduous plants are known to have higher *leafcn* ratios compared to other plant types reflecting a higher carbon content relative to nitrogen. The needleleaf evergreen temperate trees have a value of 58 in the default CLM5 configuration. According to our results, the top 10% best-performing simulations for the needleleaf temperate tree ecosystem have *leafcn* values ranging between 16 and 30. This indicates a lower carbon-to-nitrogen ratio in the leaves of the needleleaf temperate tree compared to the default setting. Again, the results align with needleleaf characteristics, which typically have lower *leafcn* ratios reflecting a balanced carbon and nitrogen content.

Plant Hydraulics: During water stress, plants react by closing stomata, preventing further water loss, which leads to reduced stomatal conductance. The *psi50* parameter indicates a plant's drought tolerance or sensitivity. Different pfts have different *psi50* values based on their abilities to tolerate water stress before experiencing significant reductions in their physiological activity. The *psi50* parameter, which characterizes the water stress threshold at which a plant's stomatal conductance significantly declines, is a vital indicator of a plant's drought tolerance or sensitivity. Stomatal conductance represents the rate at which water vapor is exchanged between leaves and the atmosphere through stomata. As water stress increases, plants close their stomata to reduce water loss, reducing stomatal conductance. Our study is based on two distinct PFTs: needleleaf evergreen temperate (NET) trees and broadleaf deciduous temperate (BET) ecosystems. The NET ecosystem has an assigned default *psi50* of -530,000 mm H2O in CLM5. The 10% best simulations for NET ecosystems have *psi50* parameter values constrained within a narrower range, between -135,546- and -394,790-mm H2O, significantly different from the default value. The narrower range suggests that NET trees exhibit a different water stress threshold than the default setting. The BET ecosystem has a default *psi50* value of -270,000 in CLM5. Our results show that BET ecosystems in the top 10% of best simulations exhibit different *psi50* values from default, with values between -783,713- and -284,831-mm H2O. This indicates that BET ecosystems in the top-performing simulations have a distinct water stress threshold compared to the default setting, allowing them to withstand more severe water stress conditions before experiencing significant reductions in stomatal conductance.

2.3.2 Parameter Sensitivity Analysis

The variability of carbon-water processes consists of two components: the variance of each parameter individually (referred to as the main effect) and the combined impact of individual parameters and interactions among parameters (known as the total effect). A higher interaction effect implies that the combined effect of multiple parameters is not additive, and these interactions significantly impact model behavior. To evaluate the impact of each parameter, we measure their effects as follows: the main effect – individual parameter contribution to overall uncertainty; total effect: the combined effect of individual and interaction effects. If the sum of the "Total Effects" exceeds 100%, it indicates the presence of a significant interaction (synergistic) effect caused by overlapping variance contributions between parameters. However, This is expected due to the complex nature of land surface models, especially CLM5, in representing land surface processes.

Individual parameter and interaction uncertainty contribute to the overall variability of carbon cycle processes (for example, GPP) in CLM5. There is a more robust interaction effect across all three sites, with a combined sum of the main effect and interactions effect explaining about 131% to 163% of the total variance. The interaction effect is more substantial at the OSBS site, with a total effect of 163%, while the main effect only explains 61% of the total variance. The combined sum of the main and interaction effects at the TALL site explains 152%, while the sum of the main effects explains about 63% of the total variance. The only exception is the ORNL site, where the main effect becomes more robust, explaining about 78%. The combined sum of the main and interaction effects accounts for only 132% of the total variance. The individual parameter effects for both carbon and water cycle processes are similar at all three sites (Fig. 2.6). Parameters such as slatop, jmaxb1, medlynintercept, wc2wjbo, leafcn, and jmaxb0 are consistently influential to both water and carbon cycle processes and consistently across the three sites.

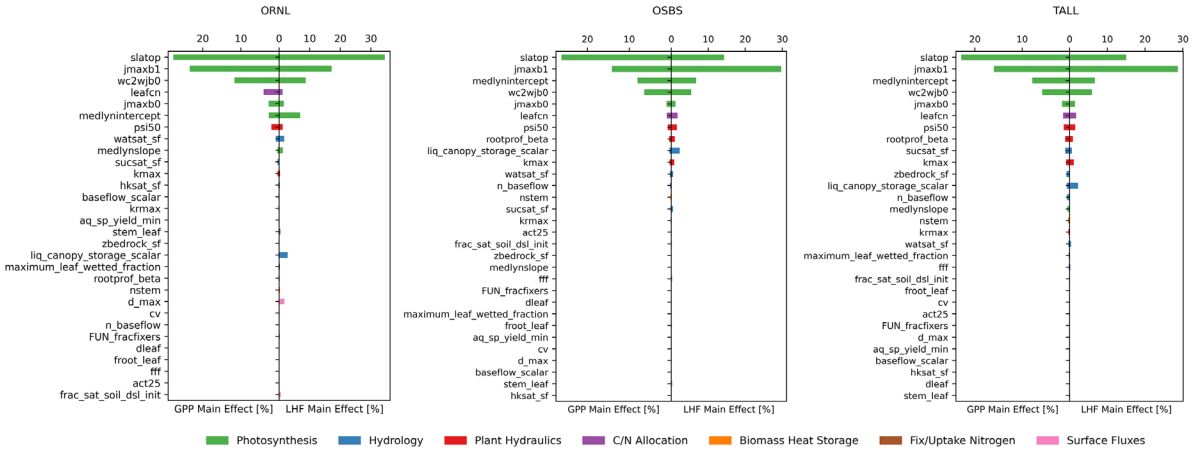

Figure 2.6: The parameter sensitivity analysis of carbon-water cycle processes variance (%) explained by individual parameter [main effect] for all the three selected sites. The color code of the legend presents the various processes influenced by each of these parameters.

The parameters *jmaxb0, jmaxb1, leafcn, medlynintercept, psi50, slatop,* and *wc2wjb0* explain most gross primary productivity uncertainties in CLM5. Our results from the three sites show that the main effects of these parameters explain between 57% and 74% of the total observed variance. The combined sum of the main and interaction effects accounts for 110% to 141% of the total variance. The remaining twenty-three parameters contribute to less uncertainty overall, with a combined sum of main and interaction effects accounting for about 22% to 32% of the total variance.

The relative importance of these seven key parameters varies significantly across the three sites (Fig 2.7). For instance, the *slatop, jmaxb1,* and *wc2wjb0* parameters contribute to the most considerable uncertainty in gross primary productivity. The *slatop* explains 23% to 28%, *jmaxb1* explains 14% to 23%, and the *wc2wjb0* parameter explains 6% to 12% of the observed variance. The *leafcn* explains 1% to 4%, *medlynintercept* 3% to 8%, *jmaxb0* 1% to 3%, and *psi50* 1% to 2% of the observed variance across the three sites.

Carbon-water cycle processes have similar parametric sensitivity but with some differences. The main effect is more substantial for water cycle processes, explaining about 72% of the total variance, while the main effect accounts for only 68%. Interactions between model parameters play a significant role in explaining overall evapotranspiration uncertainty. The combined sum of the main effect and interaction effects explains between 126% and 150% of the total variance. Specifically, the OSBS site has a more substantial interaction effect, with the main

effect explaining 67% of the total variance and a combined sum of the main effect and interactions explaining 150%.

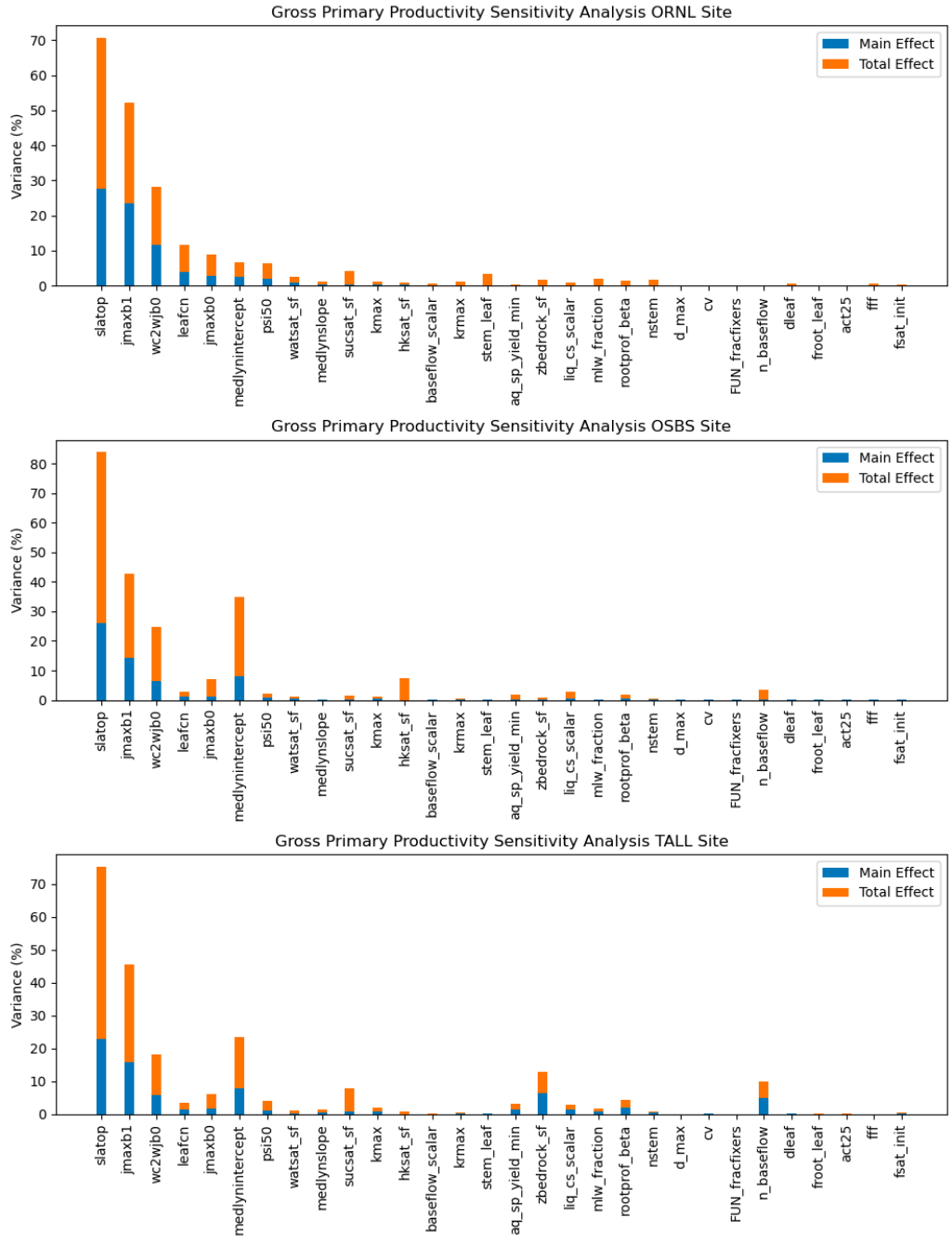

Figure 2.7: The main and total effect for each individual parameter on gross primary productivity across the three sites.

Similarly, the interaction effects account for more ET uncertainty at the TALL site. At the TALL site, the summed variance contribution by the individual parameters explains 68%, while

the combined sum of the main effect and interaction effects accounts for 147% of the total variance. The main effects are, however, more significant than interaction effects at the ORNL site. The summed variance contribution by the individual parameters and the combined sum of the main effect and interaction effects accounts for 82% and 126% of the total variance, respectively.

The main factors driving ET uncertainty in CLM5 are *slatop, jmaxb1, wc2wjb0, medlynintercept*, and *liq_canopy_storage_scalar* (Fig 2.8). Each parameter's summed variance contribution ranges from 58% to 70% of the observed variance. Our results at the OSBS and TALL sites show that the uncertainty significantly influences ET uncertainty in the *jmaxb1* parameter—the *jmaxb1* parameter accounts for 28% to 30% of the variance. However, at the ORNL site, the *slatop* parameter plays a significant role in explaining ET uncertainty, accounting for 35% of the total variance. The explained variance by *wc2wjb0* parameter ranges from 5% - 9%, *medlynintercept* from 6% - 7%, and *liq_canopy_storage_scalar* from 2% - 3%. The remaining twenty-five parameters have a minor effect on ET uncertainty in CLM5, with a summed variance contribution ranging from 9% to 12% across all three sites.

The relative importance of the five key parameters varies across different sites. At the OSBS and TALL sites, the *jmaxb1* parameter contributes the most significant uncertainty to ET, accounting for 28% to 30% of the total variance. The *slatop* parameter explains 14% to 15% of the total variance at these two sites. At the ORNL site, the *slatop* parameter plays a more substantial role and accounts for significant uncertainty in ET, accounting for about 35% of the total variance. Meanwhile, the *jmaxb1* parameter only explains around 17% of the total variance. Across all three sites, the *wc2wjb0* parameter explains about 5% to 9% of the total variance. The *medlynintercept* parameter explains around 6% to 7% of the variance, while the *liq_canopy_storage_scalar* parameter accounts for 7% to 8% of the total variance.

Plant functional type affects parameter sensitivity in CLM5. The OSBS and TALL sites have similar plant functional types (PFT) - needleleaf evergreen temperate tree, while the ORNL site has the broadleaf deciduous temperate tree. The pft-related parameters show some exciting behavior, especially when comparing the two distinct pfts examined in the present study. For example, the *leafcn* parameter explains the fourth largest uncertainty (4%), followed by Medlynintercept (sixth (3%) in GPP at the ORNL site. However, the *medlynintercept* explains the fourth largest uncertainty in GPP (8%) for both OSBS and TALL sites. The *medlynintercept* and

the *jmaxb0* explain more considerable variance in GPP than the *leafcn* parameter at these two sites. The *medlynintercept* parameter contributes to more uncertainty than the *leafcn* parameter at both the OSBS and TALL sites. In contrast, the *leafcn* parameter contributes to more uncertainty than the *medlynintercept* at the ORNL site.

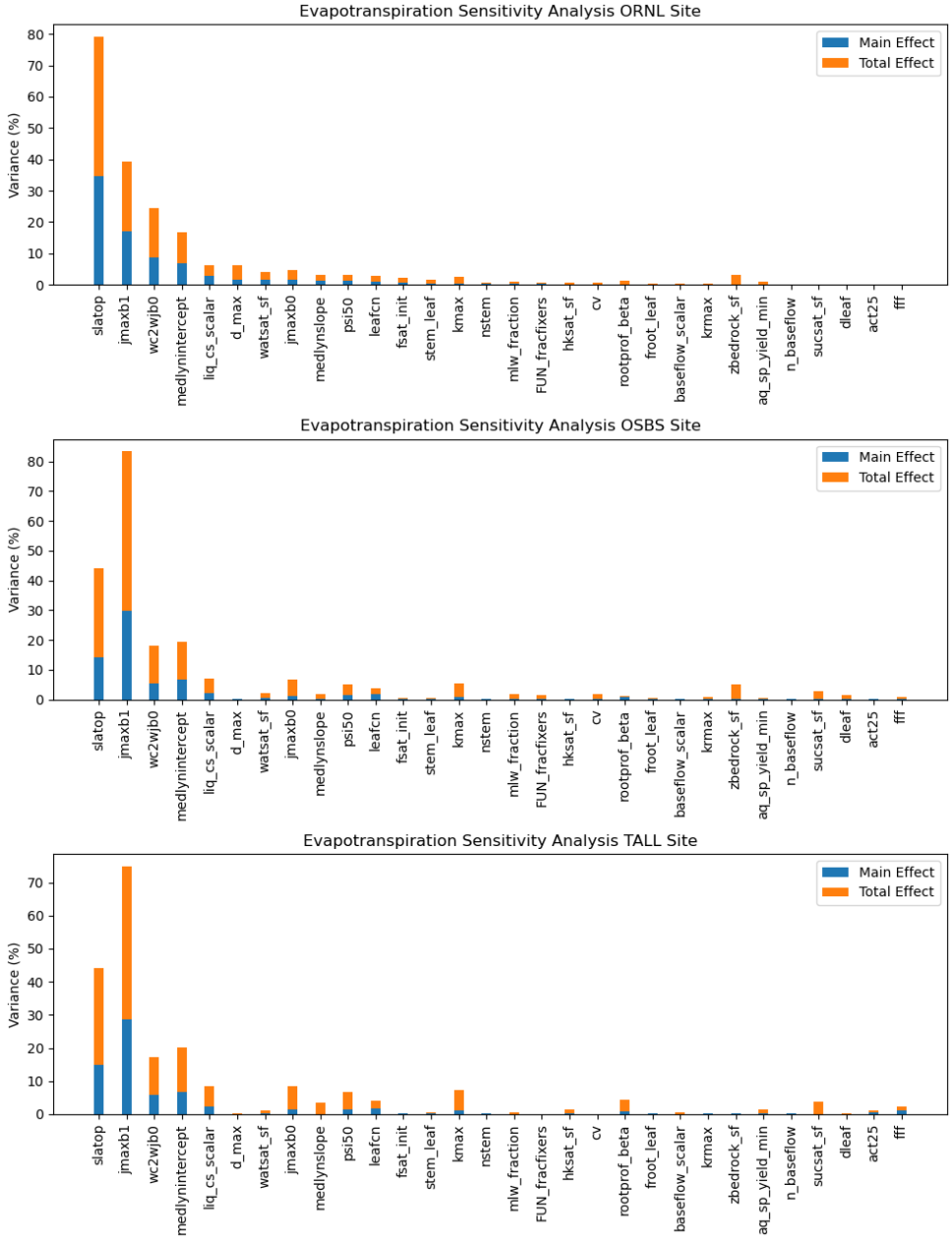

Figure 2.8: The main and total effect for each individual parameter on gross primary productivity across the three sites.

However, ET uncertainty is not only significantly highlighted by pft parameters. We observed that the *wc2wjb0* parameter (not a pft) contributes to more significant uncertainty at the ORNL site, accounting for 9% of the total variance, compared to the *medlynintercept* (pft) parameter, which explains about 7% of the total variance. It is also important to note the difference at the other two sites, TALL and OSBS, where the *medlynintercept* parameter contributes to more considerable uncertainty than the *wc2wjb0* parameter, explaining approximately 7% and 6% of the total variance.

The sensitivity analysis results on water use efficiency show that individual parameter contributions and interactions between parameters contribute to carbon-water-cycle projection uncertainties in CLM5. The combined sum of the total effects across all three sites exceeds 100% (ORNL: 150%, OSBS: 165%, TALL: 161%), indicating strong interactions between parameters. This suggests that individual parameters' combined effects and interactions explain more variability in the model output than individual parameters alone.

The main factors driving water use efficiency uncertainty include *slatop, medlynintercept, jmaxb1, wc2wjb0, leafcn, medlynslope,* and *jmaxb0*. At ORNL, the *slatop* parameter contributes the most significant uncertainty (Main Effect: 23.29%, Total Effect: 49.34%). Other significant parameters contributing to large uncertainties include *medlynintercept* (Main Effect: 13.48%, Total Effect: 19.78%) and *jmaxb1* (Main Effect: 11.56%, Total Effect: 19.23%). Similarly, at OSBS, *slatop* (Main Effect: 37.48%, Total Effect: 65.61%), *medlynintercept* (Main Effect: 15.53%, Total Effect: 39.2%), and wc2wjb0 (Main Effect: 2.04%, Total Effect: 13.22%) are the most influential parameters. At the TALL site, *slatop* (Main Effect: 35.43%, Total Effect: 62.38%), *medlynintercept* (Main Effect: 18.27%, Total Effect: 39.46%), and *jmaxb1* (Main Effect: 3.58%, Total Effect: 19.49%) contribute significantly to uncertainties in water use efficiency uncertainties.

Notably, the *slatop, medlynintercept,* and *jmaxb1* parameters significantly drive carbon-water cycle projection uncertainty in CLM5. According to our results, these parameters consistently contribute to more significant variances for the examined processes across all three sites. All three are photosynthetic related, contribute to more significant individual effects, and strongly interact with carbon-water processes in CLM5. These findings suggest that other factors

may also significantly contribute to the uncertainty in carbon-water cycle processes. Therefore, conducting parametric sensitivity analyses at additional sites with different PFTs would be valuable in generalizing these conclusions.

2.3.3 Regional Simulation

The best parameter values from each site from point scale parameter perturbation experiments were extracted and manually keyed to update the default CLM5 parameter file for a high-resolution CLM5 regional simulation. The updated parameter files were then used to run a regional simulation (upscaling). Upscaling our point scale calibrations to the regional scale shows mixed results.

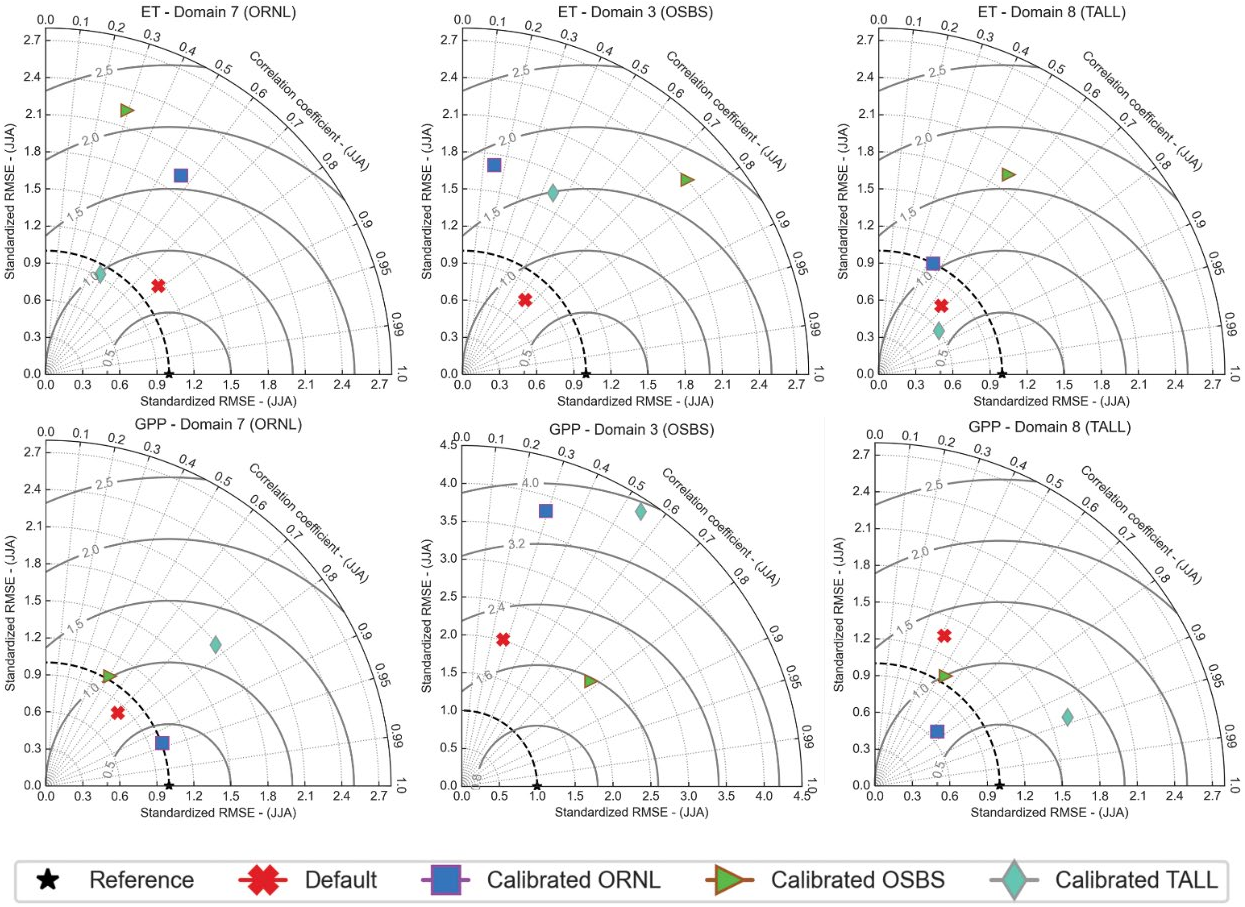

Figure 2.9: Taylor Plots Visualizing Standardized Root mean square error and Correlation Coefficients for Carbon-water Cycle Processes (ET and GPP) in the Southeastern United States region. The diagram presents a comprehensive assessment of the agreement and variability

between reference and simulated ET and GPP highlighting the performance and consistency of the CLM5 model.

It is important to note that point scale simulations use a single pft and do not factor in land use heterogeneity. Therefore, our point scale calibration efforts only considered two pfts represented at the three NEON flux tower sites: Needleleaf evergreen–temperate (TALL and OSBS) and broadleaf deciduous-temperate (ORNL). Our analysis focuses on three NEON ecoclimatic domains (D07, D03, and D08) represented by our three study sites. The simulated gross primary productivity (gC/m2/day) and evapotranspiration (mm/month) for each of these ecoclimatic domains are evaluated against reference datasets (ERA5 – Evapotranspiration and FLUXCOM – gross primary productivity).

Site-specific calibrations improve CLM5's ability to simulate carbon-water cycle processes (GPP and ET). We extracted the peak season for GPP and ET in June, July, and August (JJA). Looking at the correlation results, the site-specific calibration shows relative improvements for ET, which improves from default 0.66 to 0.76 (D03) and 0.68 – 0.81 (D08). At D07, the default configuration shows a better ET performance, with correlation declining from 0.78 (default) to 0.56 (calibrated). On the contrary, the standardized root mean squared error shows improved performance at D08 from 0.75 (default) to 0.56 (calibrated). However, the simulated ET at D03 shows declined model performance with increased standardized root mean squared error from 0.79 (default) to 2.41 (calibrated). The same is observed at D07, where the standardized root mean squared error increases slightly from 1.16 to 1.94.

Calibrated GPP at the regional level shows improvements in CLM5 model performance. Our results show that point-specific calibrations reduce CLM5 uncertainties with improved correlation at the ecoclimatic domains (D07 from 0.70 to 0.94, D03 from 0.27 to 0.78, and D08 from 0.41 to 0.94). However, despite improvements in correlation, the standardized root mean squared error increases minimally across all three ecoclimatic domains (D07 from 0.83 to 1.01, D03 from 2.01 to 2.4, and D08 from 1.35 to 1.64).

The upscaling of calibrated results for a single plant functional type to heterogeneous land use impacts the growth and presence of other pfts not considered during calibration. This effect is demonstrated by the living area plot where significant data gaps are evident (explained by significant differences between the default and the calibrated configurations). The magnitude of

this challenge is particularly pronounced for the needleleaf ecosystem. The needleleaf ecosystem suffers a significant loss of vegetation, mainly when evaluated against default model output. However, the deciduous broadleaf ecosystem demonstrates comparatively better performance when upscaled, with results closely resembling the default simulation result. This observation is further supported by the Taylor plot (Fig. 2.9), which demonstrates that D07 calibration outperforms D03 and D08 sites. It is important to note that the Southeastern United States is dominated by the broadleaf deciduous forest, which further supports the better performance at D07, which is calibrated for the deciduous broadleaf ecosystem.

Pft-specific calibration could potentially address land use heterogeneity. The parameter subset is perturbed in our PPE setup using the Latin Hypercube sampling approach. The approach perturbs parameter values, and the changes apply to every parameter represented. This means that if we have a pft parameter, we do not only perturb the specific pft we are interested in representing but all the other pfts are also perturbed during the sampling. This means that the optimal parameter setting for one pft is different for another. We set up another high-resolution regional experiment using the best simulation parameter file to test this.

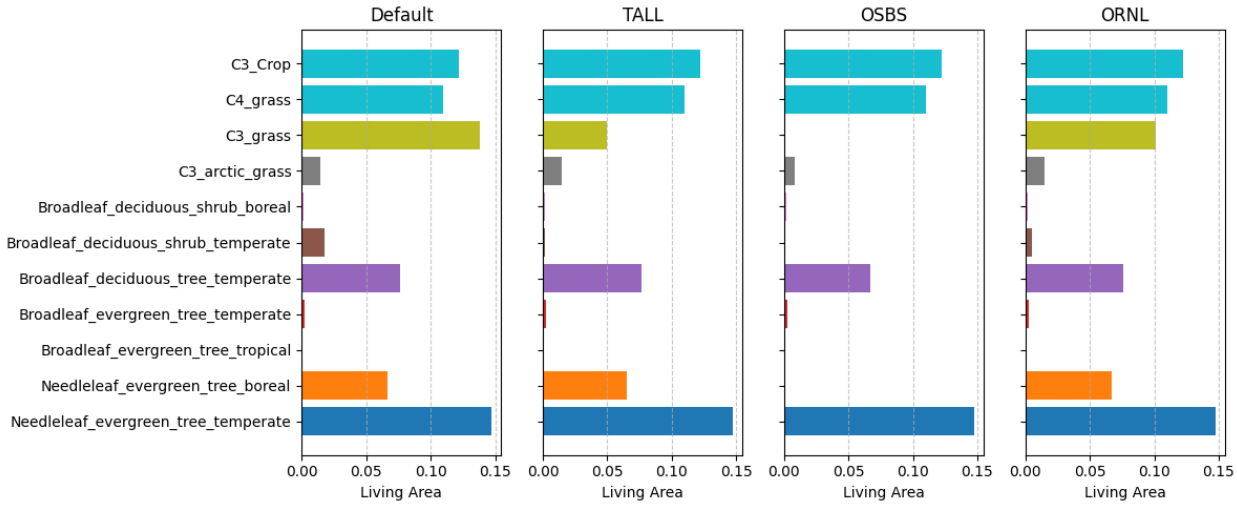

Figure 2.10: Plotting the living area as per pft types. The figure shows how single site-specific calibrations affect other plant functional types when applied at a larger scale.

The results demonstrate significant vegetation loss primarily to the uncalibrated/unoptimized pfts. It is important to note that our case study only optimizes the pft-related parameters for two distinct pfts. For instance, the ORNL calibration optimized for the

broadleaf deciduous tree–temperate shows a similar living area percentage compared to the default simulation pft living area (Fig. 2.10). The only difference between the two is on c3_grass pft, which loses about one-third of its living area and the broadleaf deciduous shrub temperate which loses its entire living area. This suggests that, contrary to expectations, the broadleaf deciduous tree temperate optimization representatively captures other pfts in CLM5 simulations.

Even though we only optimized parameters for two pfts, we see an effect on the living area for other pfts, including *c3_actic_grass, broadleaf_deciduous_shrub_temperate,* and *c3_grass.* The reduction in living area for these pfts could be attributed to the influence of other non-pft parameters. According to our sensitivity experiment, we identified photosynthesis-related parameters primarily driving gross primary productivity, and their influence could explain why we have a reduction in the living of uncalibrated pfts. This is further supported (Fig. 2.11) by significantly reducing the living area for many other pfts using calibrated configurations when the best simulation parameter file containing uncalibrated pfts is used to run CLM5 carbon-water cycle simulations.

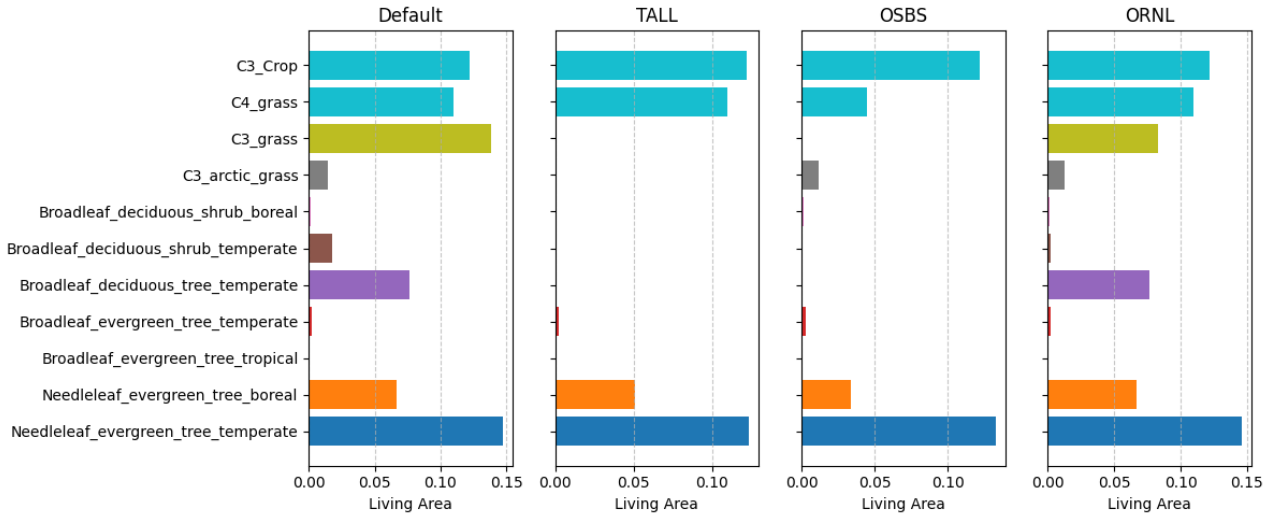

Figure 2.11: Plotting the living area as per pft types. The figure illustrates how single site-specific calibrations affect other plant functional types when applied at a larger scale.

The default plot shows the living area of all the pft represented, but these sizes shrink with TALL and OSBS plots showing significant loss in pfts. Emphasis on significant differences between default simulation and calibrated results without considering pft-specific calibration shows the importance of pft-specific calibrations. Therefore, pft-specific parameter

calibration/optimization can potentially guide carbon-water cycle process simulations in CLM5. These findings highlight the nuanced relationships between pfts and the potential for pft-specific calibrations to yield unexpectedly positive outcomes.

2.4 Discussions

The assessment of land surface models using in situ land surface data is restricted due to the limited spatial coverage of land surface process observations and the significant degree of heterogeneity in land surfaces. Understanding the heterogeneity of land surfaces and their processes is crucial as it informs model development, improving carbon-water cycle predictions. However, the high computational demands of global simulations limit their practicality, making point-scale simulations a more feasible option. In this regard, parameter perturbation experiments conducted at point scale have proven efficient and effective in improving model performance. This highlights the potential of such experiments to improve our understanding of land surface processes and develop more accurate models.

Implementing the parameter perturbation experiment significantly improves carbon-water cycle simulations in CLM5. Notably, our findings align with similar studies in the field. For instance, Dagon et al. (2020) employed the PPE approach to improve plant phenology simulation in a land surface model, resulting in notable performance improvements. Similarly, Fanourgakis et al. (2019) and Li et al. (2019) performed PPE-based investigations to assess land surface model sensitivity to different parameter configurations. Their work further substantiates the PPE approach's effectiveness in minimizing CLM5 model output uncertainties.

A central motivation for this study revolves around comprehending and minimizing carbon-water cycle prediction uncertainties in CLM5. The goal is to identify the primary drivers of carbon-water cycle projection uncertainties in land surface models. Our results emphasize the importance of specific parameters in shaping carbon-water projection uncertainties. Photosynthesis-related parameters, including *Jmaxb1, slatop, froot_leaf, leafcn,* and *Medlynintercept*, emerged as primary contributors to carbon-water cycle projection uncertainties. These parameters affect plant growth, physiology, and the relationship between photosynthesis and absorbed light (De Kauwe et al., 2017; Rogers et al., 2017).

Hydrological parameters regulating evapotranspiration rates influence water partitioning between ground, plant transpiration, canopy evaporation, and runoff. Partitioning affects water and

energy availability during photosynthesis, affecting vegetation carbon uptake. Our results identify *Liq_canopy_storage_scalar* as an essential parameter driving carbon-water cycle projection uncertainties. Our findings agree with existing research that has highlighted the significance of canopy water on ecosystem water balance and productivity (Dagang Wang et al., 2007; Zhang et al., 2014). The parameter determines the amount of water stored in the canopy, which affects the interception of precipitation and subsequent evapotranspiration, thus crucial to regulating ecosystem water balance.

Plant hydraulic conductivity is essential in determining the rates of water movement within plants and significantly affects water use efficiency, a critical component of the terrestrial carbon cycle (Kennedy et al., 2019). As demonstrated by previous studies, the psi50 parameter affects plant water relations and photosynthetic rates (Flo et al., 2021; Gul et al., 2023; Xiong & Nadal, 2020). Moreover, the Kmax parameter, which represents the maximum segmental conductance, is crucial in determining the overall hydraulic conductivity of plants (Kennedy et al., 2019). They identified Kmax as a primary source of uncertainty in CLM5, particularly in predicting gross primary productivity. Their findings reveal a robust and nonlinear relationship between Kmax and the skill of predicting gross primary productivity. This suggests that a parameter range may be extensive and constrained by reducing the maximum values or tailoring values to specific plant functional types (PFTs). Our results underscore this issue by showing that the top 10% of the best-performing models deviate significantly from the default values and are constrained within a narrower parameter range. It is worth noting that the challenge of constraining plant hydraulic stress models is exacerbated by the scarcity of relevant observational data, a concern echoed by Kennedy et al. (2019)

Understanding the Medlyn process is essential as it helps understand how vegetation responds to changes in atmospheric CO2 concentrations. The Medlynintercept and medlynslope parameters have garnered considerable attention due to their role in stomatal conductance sensitivity to atmospheric CO2 concentration and the slope of the stomatal conductance-photosynthesis relationship in CLM5. Existing research highlights the ecological significance of these parameters, particularly the influence of medlynslope on ecosystem water status (Dombrowski et al., 2022) and its more significant impact on carbon-water cycle projections (Dagon, 2020; Luo et al., 2022). Our findings show that the medlynintercept contributes more to

a considerable carbon-water cycle projection uncertainty than the medlynslope. This suggests that the role of atmospheric CO2 concentration in regulating stomatal conductance may be more critical for carbon and water cycle simulations than the stomatal conductance-photosynthesis relationship. Further, our sensitivity analysis shows a lesser influence of the medlynslope parameter in carbon-water cycle projections, even though it plays a significant role in stomatal responses to CO2. This aligns with an earlier study by Fisher et al. (2019), which also failed to establish a significant contribution of the medlynslope parameter to stomatal responses to CO2 changes. According to Fisher et al. (2019), the linear response of stomatal conductance to CO2 suggests that the direct impact of medlynslope primarily relates to changes in pre-existing CO2 concentrations.

Upscaling experiments using high-resolution CLM5 at the ecoregional level using the best parameter values show mixed results. These results are attributed to the limitations of upscaling (Xiao et al., 2012), as the calibrated parameters depend on site-specific characteristics and may not adequately capture and represent vegetation dynamics. However, the results provide important insights into the effectiveness of upscaling from point scale to regional scale. The performance of the High-resolution CLM5 simulation at the ecoregional level yields negative results, contrary to the point scale simulations, as they improved the model performance significantly.

Previous studies have reported limitations of upscaling due to the dependence of calibrated parameters on site-specific characteristics (Xiao et al., 2012; Chaney et al., 2016). Our study confirms these findings and provides valuable insights into improving CLM5 simulations at a point scale. Additionally, our study highlights the limitations of upscaling and guides how to upscale point-scale calibrated parameters to a regional or global scale. Further research is necessary to address these limitations and to fully understand the potential for point-scale calibration using NEON observations to improve water cycle projections in land surface models at a regional scale.

2.5 Conclusions

The current work was designed to constrain carbon-water cycle prediction uncertainties in CLM5 using NEON observation data for NEON ecoclimatic domains within the Southeastern United States. We used a subset of 30 parameters affecting various water carbon cycle processes in land surface models. These parameters are resampled using the Latin Hypercube sampling technique to create samples that are then used to set up and run parameter perturbation experiments in CLM5. We performed a parameter perturbation experiment using CLM5 and evaluated

parameter sensitivities at the three flux sites. The study realizes about 244,800 years of CLM5 simulations, including 200 years of model initialization at each site for each parameter configuration and 4-production years, which produce the evaluated data in this study. The simulation outputs are compared with NEON observations using a composite performance evaluation metric developed by combined mean bias error, root mean squared error, and R. The PPE results are then used to analyze global sensitivity to identify the most influential parameters driving carbon-water cycle projections in CLM5. Also, the best parameter configurations are applied to run CLM configuration at a regional scale to evaluate the consequence of transferring point-scale specific calibrations in CLM5 at a regional scale.

The PPEs improved CLM5 performance at a point scale from 0.37 to 0.78. The sensitivity analysis shows that photosynthesis-related parameters significantly influence carbon-water cycle projections. Evaluated at a larger scale, the parameter calibration efforts improve CLM5 model performance at the ecoclimatic domain calibrated for. However, these improvements were not as significant as observed at single-site calibration, attributed to different configurations at both point-scale and regional-scale CLM5 simulations. In our calibration (PPEs), we worked with a single pft and did not factor in land use heterogeneity at the point scale. There is ecosystem diversity at a regional scale, and surface characteristics are quite different. From the current study, we can draw the following conclusions:

The top 10% best-performing simulations across all three sites sufficiently capture NEON observations for latent heat flux and gross primary productivity. We can conclude that PPEs can potentially constrain carbon-water cycle projection in CLM5. The best simulation results closely match NEON observations across all three sites, which could potentially address the bias between default model output and NEON observations.

Land surface model parameter space can be further constrained. Parameters such as *jmaxb0, n_baseflow, baseflow_scalar, leafcn, psi50,* and *hksat_sf* parameters generate a better model fit against NEON observations within a constrained range as compared to the entire parameter space prescribed. We extracted and plotted the top 10% simulations parameter configurations, and most of the parameters mentioned above in the top 10% best simulations are far from the prescribed default value. Hence, considering values within the constrained ranges shown in this study could potentially reduce carbon-water cycle projection uncertainties in CLM5.

Photosynthesis-related processes are the primary drivers of carbon-water cycle prediction uncertainties in CLM5. The most influential group of parameters contributing significantly to the overall observed variance in CLM5 output is related to photosynthesis. Sensitivity analysis shows that parameters: *Jmaxb1* (response of electron transport rate to light absorption), *slatop* (Relationship between photosynthesis and absorbed light), *froot-leaf* (influences carbon partitioning between above and below-ground biomass), *leafcn* (the ratio of leaf carbon to nitrogen, affects plant growth and physiology), and *Medlynintercept* (sensitivity of stomatal conductance to atmospheric CO2) have the most decisive influence on carbon-water cycle processes. These parameters contributed significantly to the overall individual uncertainty and interaction uncertainty of latent heat flux and gross primary productivity.

Upscaling point scale calibrations to regional scale shows mixed results. It is important to note that point scale simulations use a single pft and do not factor in land use heterogeneity. Pft parameter values show a considerable influence on carbon-water cycle predictions. We attribute the inferior performance of the optimized parameter values at a regional scale to the limitations of upscaling, as the calibrated parameters depend on site-specific characteristics and may not adequately capture and represent vegetation dynamics. However, our results can potentially inform and guide point-scale and regional simulations in CLM5.

Chapter Three: Advancing Land Surface Modeling With Gaussian Process Emulators

3.1 Introduction

Representing land surface processes adequately in land surface models is a challenging process that leads to increased complexity of these models. Efforts to better represent land surface processes introduce additional uncertainties into the model projections, adding to the overall model bias (Yuan et al., 2022). These biases emanate from insufficient observational data and incomplete numerical representations of these complex processes (Dagon, 2020). Improving these models and their predictive skills is vital for future projections and climate-carbon feedback. The realization of these goals necessitates a concerted effort to minimize sources of model error while enhancing the capabilities of land surface models (Bonan and Doney 2018). The dual challenge of enhancing capabilities while mitigating uncertainties constitutes a pivotal endeavor in advancing land modeling and its broader scientific applications.

Increasing comprehensiveness increases the need for the computational processing power of these models. This presents a challenge to experiment with different parameter configurations in global analysis. The balance between computational cost and practical relevance hinges significantly on the resolution at which a model is executed. While there is a prevailing, though not always justified, trend towards increased resolution in climate modeling, the application of Land Surface Models (LSMs) to inform ecosystem management decisions demands a careful consideration of resolution, mainly when supporting decisions at the local or regional level (Ref: Baker 2021. As a result, performing higher-resolution simulations using CLM5 becomes computationally expensive.

To address this challenge, researchers have introduced emulators (statistical surrogate models) (Kennedy & O'Hagan, 2001). An emulator is a statistical framework that rapidly predicts a computer model's output, including uncertainties, without additional simulations (Baker et al., 2021). The resulting statistical model can help understand and improve the underlying process-based model, e.g., CLM5. Fast land surface model predictions can enhance decision-making, advance scientific understanding, and facilitate the effective integration of multiple models, all while quantifying uncertainties.

Earth system science analysis employs two main types of emulators: Statistical and Deep Learning (Chaure & Bhattacharya, 2023). Statistical surrogates are data-driven approaches that rely heavily on data to emulate land surface processes, making them suitable for understanding environmental systems with limited mathematical descriptions (Field et al., 2011). Land surface modeling involves training simplified models to predict land surface processes for new scenarios efficiently (Lu & Ricciuto, 2019). However, their accuracy may be lower than LSMs, necessitating consideration of uncertainties. Deep Learning emulators utilize artificial neural networks to model complex system behavior learning patterns in large datasets and include multiple layers of interconnected neurons; they are trained to predict system behavior under varying input conditions (Weber et al., 2020). Both statistical and deep learning emulators are valuable resources, especially in land surface modeling, with statistical emulators reducing computational resources and deep learning emulators capturing complex non-linear relationships.

3.1.1 Gaussian Processes Regression and Earth System Modeling

Gaussian Process regression is a data-driven machine-learning technique that models the input-output relationship as a probabilistic distribution, distinguishing itself from traditional fixed mathematical functions (Rasmussen & Williams, 2006). The Gaussian Process (GP) concept refers to a statistical framework with implications for modeling complex relationships in data, and its fundamental characteristic is that any finite subset of a random variable adheres to a multivariate Gaussian distribution, providing GPR with its inherent flexibility (Rasmussen & Williams, 2006).

Earth system modeling is primarily driven by the need to understand the complex interactions between various earth system components, including land, atmosphere, and oceans. These interactions are described by dynamic and non-linear relationships that often confront simplistic parameter representation. A non-parametric modeling technique such as the Gaussian processes regression could address these challenges as it can capture input-output relationships without making prior assumptions about the functional forms (Rasmussen & Williams, 2006). The adaptability of Gaussian processes contrasts conventional parametric models often used in machine learning, which assume a prior fixed functional form for the relationship between input and output variables (Gramacy, 2020). This adaptability resonates well with the complex dynamics of Earth systems and the multifaceted interactions governing climate processes.

Gaussian process regression application has enhanced land-atmosphere interactions by providing a more realistic representation of earth system processes, improving accuracy in climate projections (Yadav et al., 2020), and making it possible to quantify climate prediction uncertainties (Gong et al., 2015; Lovenduski & Bonan, 2017; Prihodko et al., 2008). Using observational data to improve land surface model accuracy is a powerful data assimilation tool in Earth system modeling (Gómez-Dans et al., 2016; Sawada, 2020). By capturing spatial and temporal patterns, they can reconcile predictions with real-world observations (Baker et al., 2022). A problem with Gaussian process prediction methods is that they become very slow and computationally demanding (O(n3)) as the amount of data grows (Rasmussen & Williams, 2006).

The current study is directed to quantify land surface process uncertainty by systematically exploring the parameter space of the Community Land Model version 5 (CLM5). We aim to identify the main uncertainty contributors to carbon-water cycle projections. We leverage the power of Gaussian Processes Regression (GPR) to develop an emulator using output from a concluded Perturbation parameter experiment. Our emulator is designed to capture the seasonal variations in water fluxes.

3.2 Material And Methods.

3.2.1 Research Design

The design and implementation of the emulator come after parameter perturbation experiments. You need sufficient data to develop, execute, and implement a machine learning algorithm. The parameter perturbation experiments serve as the input data source into our emulator, including CLM5 simulation outputs and inputs. Although parameter perturbation experiments improved CLM5 model performance, the final result did not adequately capture observations. Therefore, we build our emulators (Fig. 3.1) based on the perturbed parameter experiment results.

To train and evaluate the emulator, we will use a subset of seven parameters (input parameters) affecting carbon-water cycle processes in CLM5 and seven climatological variables (forcing datasets) and CLM5 outputs (evapotranspiration).

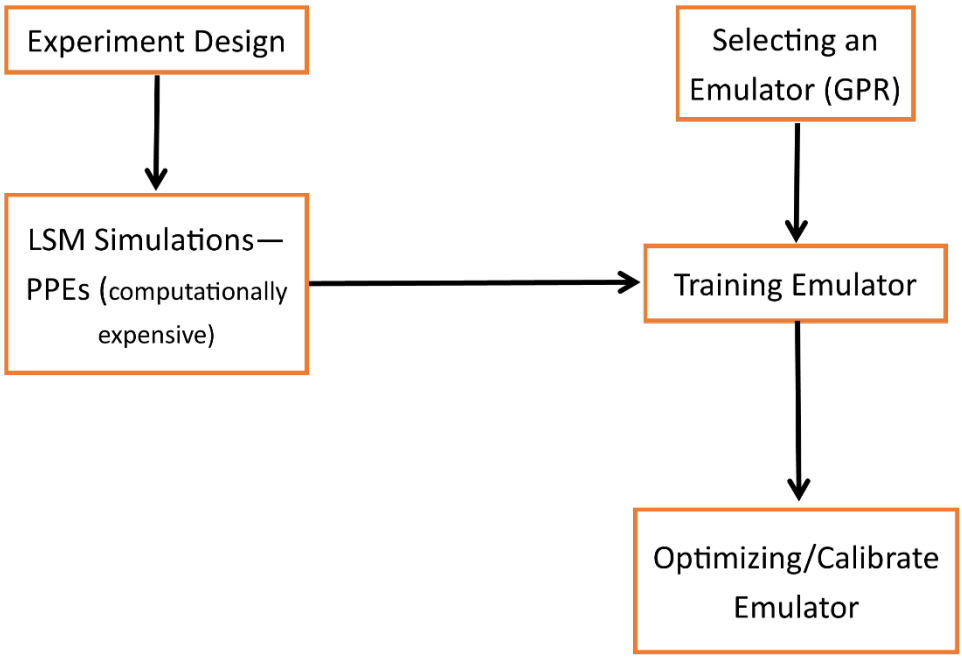

Figure 3.1: The graphical representation of the Gaussian Process Regression model.

The input parameters, forcing datasets, and the CLM5 outputs were used to train and evaluate our emulator. The training and testing dataset was split at 85% and 15%, respectively. Often, emulators in earth system science train to capture a statistic that might not fully represent the datasets. Notably, carbon-water cycle processes are affected significantly by seasons with high activity during summer and minimal to no activity during winter. Using the sufficiently trained and tested emulator, we employ a Markov Chain Monte-Carlo (MCMC) sampling approach to generate input samples by exploring the plausible regions of our parameters driven by the observations data. These generated samples are parsed through our emulator while holding the forcing dataset constant. The emulator is then subjected to monthly evapotranspiration predictions, which are compared against NEON observations. This works in a loop and is directed to identifying the input values that yield minimum bias between emulator output and NEON observations.

3.2.2 Gaussian Processes Regression

Gaussian processes (GPs) are infinite-dimensional normal distributions used for system emulation due to their simplicity and reliable uncertainty estimates, especially in scenarios with limited training data(Rasmussen & Williams, 2006). We developed a Gaussian Processes regression emulator to emulate CLM5 processes. Gaussian process emulators serve as statistical

models for functions, treating the simulator as a function with inherent uncertainty. Emulators determine probability distribution across these functions using Gaussian processes, allowing for quick sampling or evaluation. The formulation of a Gaussian process emulator can be expressed as equation 3.1 below:

$$y(\theta) = GP(m(\theta), k(\theta, \theta) \tag{3.1}$$

Here, y represents the land surface model output, θ refers to the vector of input parameters, m is the mean function. It represents prior beliefs about the model, and k is the covariance function. A vital feature of a Gaussian process is that any finite set of realizations, $y(\theta_1), \ldots, y(\theta_n)$, follows a multivariate normal distribution with mean vector $(m(\theta_1), \ldots, m(\theta_n))$, and covariance matrix K, where $K_{ij} = k(\theta_i, \theta_j)$. After observing model outputs (y), the posterior $(y(\theta)|y)$ process also follows a Gaussian process with a known mean and covariance functions (Rasmussen and Williams, 2006). The common practice of re-scaling input parameters to (0,1) and standardizing model outputs is applied.

The mean function provides information about the expected value m (θ)for any given value of θ, while the covariance function defines the shape of the Gaussian process. The shape regulates the mean function's smoothness. It provides a correlation between $m(\theta_i)$ and $m(\theta_j)$.' After initializing the prior distribution, observation data (training data) is used to refine the mean function estimates using the Bayes theorem. The likelihood function [eq. 3.2] is expressed below.

$$p(y|\theta, \mu) = Normal(y|m(\theta), \sigma^2) \tag{3.2}$$

This function quantifies the likelihood of the observed data y, considering our estimate of f(x) and noise level (σ^2). Using Bayes' rule, we derive the posterior distribution [eq. 3.3] over m (θ) after observing data:

$$p(m(\theta)|\theta, y) = Normal(m(\theta)|\mu_{post}, \sigma_{post}) \tag{3.3}$$

Where, μ_{post} and σ_{post} represent the mean and covariance functions of the posterior distribution, respectively. These are computed based on the prior mean and covariance functions, the likelihood function, and the observed data. We employ the posterior distribution over m (θ) to predict new θ values. For a new input θ*, we compute the predictive distribution over y* expressed in [eq. 3.4] below.

$$p(y^*|\theta^*, \theta, y) = Normal(y^*|\mu^*_{pred}, \sigma^*_{pred}) \quad (3.4)$$

Here, μ^*_{pred} and σ^*_{pred} refers to the mean and variance of the predictive distribution, respectively.

Our current work employs the ESEm GP emulator developed by Watson-Parris et al. (2021), built on the GPflow implementation (Matthews et al., 2017).

A key consideration in GPR is the kernel selection. Common kernel types used in GPR include constant, linear, radial basis function (RBF or squared exponential), and Matern family, each with varying degrees of differentiability. Kernels can also be customized to capture specific data characteristics, such as non-stationarity or periodicity. The ability to select a specific covariance function is dependent on the knowledge of the data and allows for more control and interpretability in the resulting model compared to alternatives like neural networks.

3.2.3 Kernel Description

The current work combines two kernels (the radial basis function and the Matern32 kernels) to model the covariance structure. The Radial Basis Function (squared exponential) [eq. 3.5] is a stationary kernel in that its properties remain constant with changes in input space (Rasmussen & Williams, 2006). It captures smooth, continuous functions and is applied when the expected exhibits such behavior.

$$k_{exp}(\theta, \theta') = \exp\left(-\frac{1}{2l^2}\left|\left|\theta - \theta'\right|\right|^2\right) \quad (3.5)$$

The Matern32 kernel [eq. 3.6] is an extended form of the RBF kernel but with an additional parameter v that controls the smoothness of the output function. A smaller parameter v value is likely to generate a less smooth approximated function, and as v goes to infinity, the Matern kernel behaves more like the RBF kernel. Parameter (v) can either be a single value for uniform behavior in all directions (isotropic) or a vector of matching input dimensions for varied behavior in different directions (anisotropic) (Rasmussen & Williams, 2006).

$$k_{mat}(\theta, \theta') = \frac{1}{\Gamma(v)2^{v-1}}\left(\frac{\sqrt{2v}}{l}d(\theta, \theta')\right)^v k_v\left(\frac{\sqrt{2v}}{l}d(\theta, \theta')\right) \quad (3.6)$$

$$k(\theta, \theta') = k_{exp}(\theta, \theta') + k_{mat}(\theta, \theta') \quad (3.7)$$

k (θ, θ') is the composite covariance function designed for this study, and l is the lengthscale (hyperparameter) and controls the covariance function width. The $||\theta - \theta'|| = d$ refers to the Euclidean distance between θ and θ,' k_v refers to a modified Bessel function and Γ is the gamma function.

The RBF kernel works under the assumption that nearby data points share a similar function. In contrast, the Matern kernel offers greater flexibility by accommodating various degrees of smoothness within the function. By combining these two kernels [eq. 3.7], we can capture our function's smooth and non-smooth components.

3.2.4 Model and Parameters Selection

CLM5 was used to initialize and execute the parameter perturbation experiments at three NEON sites. Using the Latin Hypercube sampling technique, these experiments used 400x30 different parameter configurations generated from a subset of 1 x 30 parameters. The sensitivity analysis from the perturbation experiments identifies the seven most influential parameters, which we further explore using the emulator.

3.2.5 Emulating land Surface model

We consider the case of emulating Latent Heat Flux (ET) at a single flux tower site, which provides point-scale measurements of ET over time. Land surface models generate outputs of varied dimensions, including spatiotemporal attributes. Land surface models derive the spatiotemporal correlation from the input forcing datasets. To illustrate, CLM5 independently solves the same differential equations in each grid cell without the solver passing any horizontal information exchange (Lawrence et al., 2019). Our experiment is designed for a single point to capture the temporal variability at a specific location. Specifically, our emulator is designed for the TALL NEON site, providing access to high-frequency observational data, including latent heat flux. Mathematically, we can represent this as follows:

$$y_t(W, \theta) = f(W_t, \theta) \tag{3.8}$$

Where $y_t(W, \theta)$ refers to CLM5 output (latent heat flux) at time t with forcing datasets W and input parameters θ, while W_t is the forcing dataset at time t and f is the land surface model (CLM5). Based on this information and considering the land surface model as having a 1D output, we develop a conventional emulator:

$$f(W_t, \theta) \sim GP(W_t, \theta) \tag{3.9}$$

The assumption of spatial independence, conditioned on the available forcing data, is often a reasonable simplification, but temporal dependence is less straightforward in land surface modeling (Baker et al., 2022). In CLM5, the model maintains and updates various internal state variables at each time step, introducing some inherent temporal structure. For instance, precipitation and prior time steps influence the soil moisture modeled in CLM5. Similarly, the leaf area index exhibits temporal dependencies, as future carbon assimilation and respiration depend on the leaf area index from the preceding time step.

The emulation of such time structures could be realized by adopting a "dynamic emulator," as proposed by Mohammadi et al. in (2019). However, it is essential to note that the computational expense associated with this approach is typically too high for land surface models. The current research addresses this issue by adopting a monthly averaging approach to smooth out short-term variability.

3.2.6 Obtaining an Ensemble and Data Abundance

Developing an emulator requires multiple simulations of the land surface model. The conventional approach involves running the land surface model at the region of interest (e.g., NEON flux tower site) and repeating this process multiple times with varying parameter configurations to establish an initial ensemble (Booth et al., 2012; Williamson, 2015). The emulator experiment comes after the parameter perturbation experiments, which generated 400 CLM simulations with different parameter configurations.

The PPEs were set up using a subset of 30 parameters affecting various carbon-water cycle processes in CLM5, including hydrology, photosynthesis, and flux energy exchange. The 400 simulations were evaluated against NEON observations (GPP and ET), and the best-performing simulations were identified. We extracted the top 10% of the best simulations for the emulator development. A global sensitivity analysis identified a subset of 7 parameters as the most influential parameters driving carbon-water projection uncertainties in CLM5.

Suppose we use all the 400 simulations, the 30 parameters, and the forcing dataset and extract for higher frequency time-step. In that case, the result will be a substantial amount of data with hundreds of thousands of data points per simulation. While having abundant data should, in

theory, enhance predictive capabilities, Gaussian processes are not well-suited for handling large datasets. In computer experiments, it is typically recommended to have around 10 data points per input dimension (Loeppky et al., 2009). While alternative statistical models like standard linear regression or neural networks can be employed, the preference for the Gaussian process lies in its flexibility and ability to provide valuable uncertainty estimates. The PPE simulations dataset forms the foundation for our emulator, which aims to capture the seasonality of Evapotranspiration. The emulator's function is contingent on the forcing dataset, a critical factor in the modeling process.

3.2.7 Emulating CLM5

To illustrate the framework described above, we developed an emulator for the Community Land Model (CLM5) to investigate evapotranspiration. In climate and land use changes, CLM5 is commonly employed globally, typically with a coarse spatial resolution of at least 0.5° x 0.5°. Our research is motivated by the growing demand for detailed process models to inform land use and management decisions at a much finer resolution. We execute and emulate CLM5 [eq. 3.13] at a point scale that allows us to capture localized nuances. The general design of our emulator for simulated ET is presented here:

$$ET(W_t, \theta) \sim GP(W_t, \theta) \tag{3.10}$$

Where ET is the emulated ET, θ denotes input parameters, and W_t denotes forcing inputs at time t. we experimented with many different emulator designs, as summarized (Table 3.1) below:

Table 3.1: Summary of different emulator designs

Design	Equation	Notes
Using mean CLM5 output (e.g., ET)	$ET(\theta) \sim GP(\theta)$	- Total observation points 400 x 7 - No forcing dataset used ~ constant. - Includes non-plausible regions of the parameter space
Using performance metrics (e.g., composite, rmse)	$Metric(\theta) \sim GP(\theta)$	- Total observation points 400 x 7 - No forcing dataset used ~ constant. - Includes non-plausible regions of the parameter space.

		- Training the model with bad data ~ when the metric is poor, the ML learns and considers that when making predictions. - Difficult to achieve an accurate model
Using monthly CLM5 output (e.g., ET) ~ best performing PPE simulations	$ET(W_t, \theta) \sim GP(W_t, \theta)$	- Total observation points 1920 x 14 - Includes forcing dataset. - Includes only plausible regions of the parameter space

Here, we experiment with different settings to train different emulators and with different observation points and variables. As shown in the table, if we use a mean statistic or a performance metric, we omit the observation as *t=0.* The forcing dataset becomes constant, does not affect the emulator training, and thus drops out. The mean statistics cannot adequately capture the observations as they represent the ensemble mean, and carbon-water cycle processes are significantly affected by seasons. Therefore, training an emulator with the mean statistic cannot capture this seasonality as it eliminates the forcing data that brings in the time variable. Training the emulator using a performance metric presented many challenges. The performance metric is the result of comparing the PPE results against NEON observations. The metric identifies which simulations closely match the observations. That means we have mixed metrics; in other words, some metrics suggest good correlations between simulations and observations, while others suggest poor correlations between simulations and observations. As a result, when we use all the PPE metrics to train the emulator, we have a mixture of good and poorly performing models as inputs to the emulator. The flawed metric is a good indicator of non-plausible regions of the parameter space and becomes difficult even to achieve an accurate emulator. Therefore, using the mean and performance metric means we are still experimenting with non-plausible parameter space regions and eliminating the forcing dataset that brings in the time function. Developing an emulator that only considers the plausible region of the parameter space is meaningful, especially for calibration purposes.

An overview of the input parameters and forcing variables is presented below (Table 3.2). The table includes their default values, the minimum and maximum ranges, and a brief description of their roles within CLM5.

Table 3.2: This table details the parameters, their default values, ranges, units, forcing variables, and concise descriptions.

Parameters	**Default**	**min**	**max**	**Description**
wc2wjb0	0.81	0.5	1.5	*The baseline ratio of rubisco limited rate vs. light limited photosynthetic rate (Wc: Wj)*
jmaxb1	0.17	0.05	0.25	*Determines response of electron transport rate to light availability*
jmaxb0	0.03	0.01	0.05	*Baseline proportion of nitrogen allocated for electron transport (J)*
slatop	0.01	0	0.15	*Specific leaf area at the canopy top (m2/gC)*
psi50	-530000	-50%	50%	*Water potential at 50% loss of conductance (mm H2O)*
leafcn	58	-30%	30%	*Leaf carbon to nitrogen ratio (gCgN−1)*
medlynintercept	100	1	50000	*Medlyn intercept of conductance-photosynthesis relationship (µmol H2O/ m2s))*
Forcing Variables	Units			Description
FLDS	W/m^2			*Incident longwave*
FSDS	W/m^2			*Incident shortwave*
PRECTmms	mm/s			*Precipitation*
RH	%			*Relative humidity*
PSRF	pa			*Pressure at the lowest atmospheric level*
TBOT	K			*Temperature at the lowest atm level*
WIND	m/s			*Wind speed at the lowest atm level*

3.2.8 Emulator Calibration/Tuning

The formal adjustment, or calibration, of input parameters within a land surface model can be realized using a well-designed and efficient emulator. Calibration is adjusting input parameters to align the resulting outputs with real-world observations. Exploring various parameter configurations without an emulator exhaustively becomes prohibitively expensive, especially in land surface models. Calibration involves a degree of subjectivity, heavily relying on personal experience and intuition. On the other hand, an optimization procedure can be pursued (Raoult et al., 2021). However, optimization procedures can be computationally demanding, not fully quantify uncertainties, and lack alternative options if the results do not align with established scientific knowledge.

An emulator enables direct testing of numerous parameter configurations, streamlining the exploration of the parameter space for greater efficiency. A fast, robust emulator is trained to calibrate our model with available observations. The calibration follows the outlined GP calibration procedure using ESEm developed by Watson-Parris et al. (2021). The posterior is proportional to the observations Y^0 given parameters (θ) multiplied by the prior which presents the knowledge before observing the data. Bayes' theorem can be summarized:

$$p(\theta|Y^0) \propto p(Y^0|\theta)p(\theta) \tag{3.11}$$

Here $p(\theta|Y^0)$is the posterior probability, $p(Y^0|\theta)$ is the likelihood function and represents the joint probability of the data (Y^0) given parameters (θ), and $p(\theta)$ refers to the prior probability distribution of parameters (θ).

Since the full likelihood distribution is often unknown and intractable, we must approximate it. The choice of approximation method depends on the calibration purpose and assumptions about $p(Y^0|\theta)$. In our study, the goal is to find the optimal input parameters to our developed emulator to match NEON observed evapotranspiration, and Markov-Chain Monte-Carlo (MCMC) techniques proposed by Watson-Parris et al. (2021) is implemented.

When using the MCMC method, the acceptance or rejection of a set of parameters provides a better understanding of our posterior probability $p(\theta|Y^0)$. Specifically, rejecting a set of parameters guides the sampler to focus on plausible parameter regions, thereby improving the approximations. The observation (target) drives the ESEm MCMC sampler and accepts samples generated if they closely match our target. The acceptance probability *r* for generated samples is determined as expressed below:

$$r = \frac{p\left(Y^0\middle|Y'\right)p\left(\theta'\middle|\theta\right)p(\theta')}{p\left(Y^0\middle|Y\right)p\left(\theta\middle|\theta'\right)p(\theta)} \tag{3.12}$$

The sampler uses TensorFlow-Probability's Hamiltonian-Monte Carlo (HMC) with gradient information assuming an even proposal distribution: $p(\theta'|\theta) = p(\theta|\theta')$, which simplifies the log probability as:

$$\log(r) = \log\left(p(Y^0|Y')\right) + \log\left(p(\theta')\right) - \log\left(p(Y^0|Y)\right) - \log\left(p(\theta)\right) \tag{3.13}$$

According to Watson-Parris (2021), the distance metric ρ is given as a probability distribution with values between 0 and 1. Therefore, we assume that this discrepancy can be estimated effectively by modeling it as a normal distribution centered around the emulator mean and a standard deviation equal to the sum of the squares of the variances [eq. 3.14]. Since the simulator proposal deviates from the observations, ESEm includes an implausibility metric which defines this discrepancy as the standardized Cartesian distance between observations (Y^0) and emulator mean (μE) (Watson-Parris et al., 2021).

$$\rho\left(Y^0, \mu E(\theta)\right) = \frac{|Y^0 - \mu E|}{\sqrt{\sigma^2{}_E + \sigma^2{}_Y + \sigma^2{}_R + \sigma^2{}_S}} = \rho(Y^0, \theta) \tag{3.14}$$

Here, the total standard deviation (σ_t) is emulator variance squared sum ($\sigma^2{}_E$), observation uncertainties ($\sigma^2{}_Y$), model structure ($\sigma^2{}_S$), and representation (observations expected to match model output) uncertainties ($\sigma^2{}_R$).

$$p(Y^0|\mu E) \approx \frac{1}{\sigma_t\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{Y^0 - \mu E}{\sigma_t}\right)^2}, \sigma_t = \sqrt{\sigma^2{}_E + +\sigma^2{}_Y + \sigma^2{}_R + \sigma^2{}_S} \tag{3.15}$$

The sampler returns requested samples together with an acceptance rate. The ESEm MCMC sampler uses a Hamiltonian Montecarlo kernel, enabling us to sample from the posterior distribution. To run the sampler, one needs to define the objective function. We need to obtain emulator inputs that yield ET estimates closer to observations. In our three emulator designs, the objective function was defined differently: Mean guided emulator: observation mean metric, performance metric guided emulator design, the max value, and the PPE guided emulator, NEON ET monthly observations. For each emulator design, we generated 80,000 samples.

The generated samples were parsed through the emulator for different purposes. The mean emulator-generated samples were used to predict ET using the emulator, and the generated outputs were filtered to identify the outputs that closely match the mean observations. The input samples that correlate with the predicted outputs that match the objective function (observed mean) are identified as the optimized parameter and are used to run CLM5 simulations. The metric-guided emulator samples were used to make predictions. The outputs were also filtered to remain with outputs above 0.75 only. These outputs are then matched with their respective input parameters, and the highest predicted metric input parameters are identified as the optimized parameters that are again used to run CLM5.

The PPE-guided emulator samples were dealt with differently. We developed an algorithm that takes these samples, filters out the forcing datasets, and concatenates the input samples with the actual forcing datasets. The new inputs are then subjected to monthly four-year predictions (48 months) and compared against NEON observations. The algorithm compares the predicted ET and the NEON observations by computing the root mean squared error. We set our objective function in the algorithm targeting a minimum rmse value. All the samples are passed through the emulator. If the target rmse is realized, the algorithm breaks and returns the corresponding input parameters together with the root mean square realized. Otherwise, it continues until all the samples are passed through, and then the inputs with the minimum rmse are returned and considered the optimized new set of parameters. It is important to note that the forcing dataset does not change and remains constant while the input parameters are obtained from the generated 80,000 samples. We did not change/calibrate the forcing dataset as we wanted to replicate conditions similar to those in the CLM5 environment.

3.3 Results

3.3.1 Emulator Performance

Evaluating the accuracy of an emulator is a crucial step, just as it is when assessing the performance of the land surface model itself. While training our emulator, we randomly set aside 15% of the data points as test datasets. These held-out data points are then used to verify the accuracy of the emulator. After training our emulator to a satisfactory level, we conduct predictions using the 15% test data to assess the emulator's performance against the ET outputs (y_test) that were not used during the training. We calculate the correlation coefficient (R^2) between the emulator predictions and the y_test datasets to evaluate performance. The emulators (Fig. 3.2) are summarized below.

The mean and metric-based emulators were challenging to train for higher accuracy. However, the top 10% PPE emulator, which included larger dimensions and more observation points, gave a better prediction accuracy. Using the developed emulator, we generated monthly evapotranspiration predictions from January 2018 to December 2021 and compared them with NEON ET observations. The results show a mean bias error of -6.21 mm/month. The best overall PPE simulation yielded a mean bias error of -2.28 mm/month using the same parameter set.

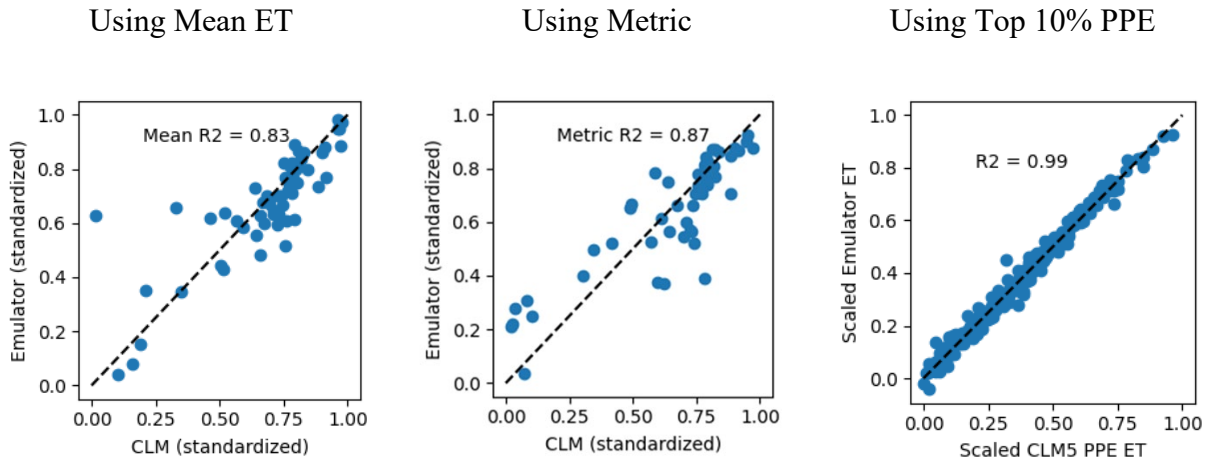

Figure 3.2: Evaluating the efficiency of the emulator to capture ET dynamics using test datasets (train_x) and comparing the predicted ET against the PPE (test_y) ET.

Comparing our results against the PPE results (Fig. 3.3), we notice that the emulator without calibration further underestimates ET, leading to a more significant mean bias error. However, this result is from the PPE-optimized parameters and forcing datasets only without any adjustments, demonstrating our emulator's ability to capture identical dynamics within the dataset.

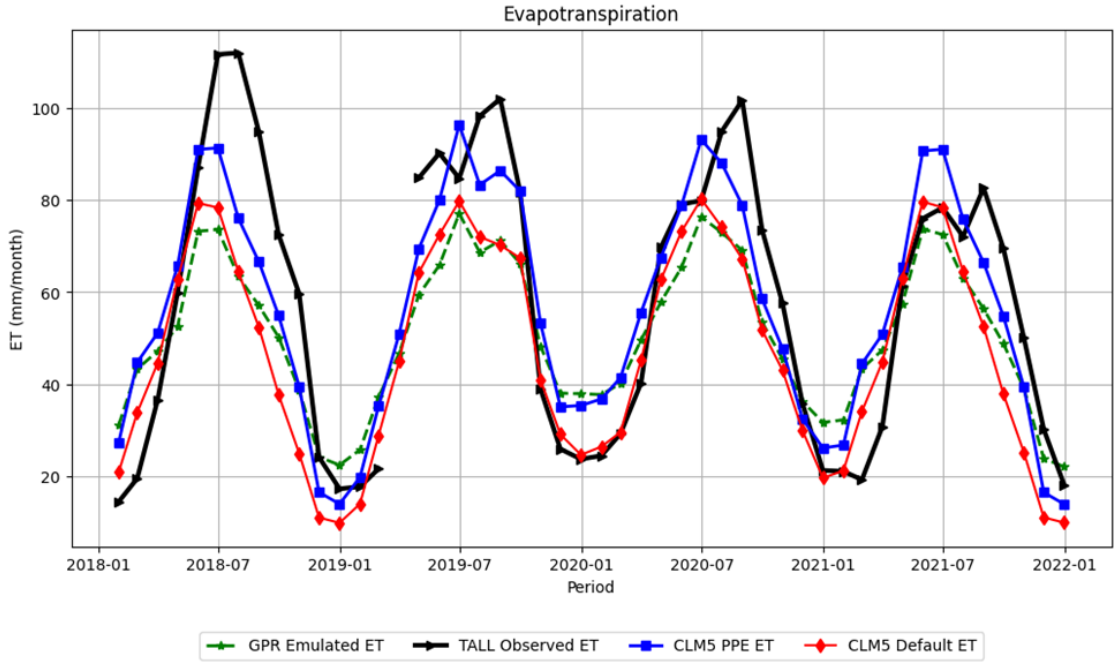

Figure 3.3: Determining the efficiency of the emulator in capturing ET seasonality. The figure demonstrates the ability of the emulator to predict monthly ET using optimized parameter values from the PPE experiments and monthly forcing dataset for the period 2018 - 2021.

3.3.2 Sensitivity analysis

A Gaussian process emulator can perform an automatic preliminary sensitivity analysis. This is achieved when the covariance function, denoted as 'k' (see section 3.2.3 [Eq. 3.7]), is selected to be a "non-isotropic squared exponential covariance function," sometimes referred to as an "automatic relevance determination kernel." When coupled with a constant mean function, the emulator automatically provides estimates of the relative importance of various inputs. In other words, it assesses how sensitive the outputs are to individual inputs.

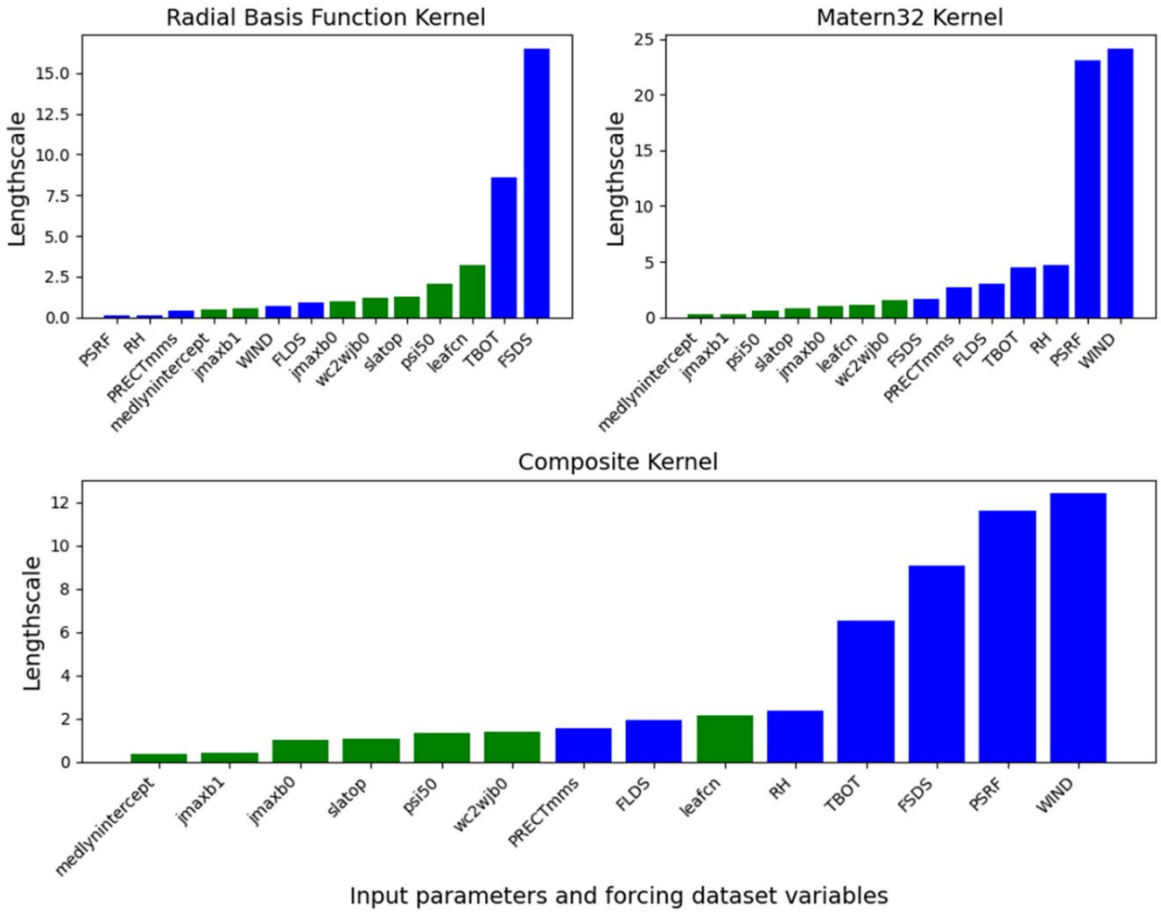

Figure 3.4: Estimated lengthscale values from the developed Gaussian process emulator; a small value indicates a stronger relationship (more sensitivity) between input parameters and output (latent heat flux). The green bars denote tuning parameters, while the blue bars correspond to climatological forcing datasets.

This choice of covariance function is available as an option in most Gaussian process software. The emulator training processes estimates lengthscale parameters, which measure the

distance two points must be in each dimension before they become uncorrelated. A smaller lengthscale implies a stronger relationship between the parameter and the output, signifying greater importance. The sensitivity analysis illustrates these length scale estimates for each input, providing an understanding of input influences on the emulator output.

In Gaussian Process modeling, a non-isotropic kernel allows the model to have different relationships between input parameters and model output independently for each dimension. As a result, the shape and strength of the relationship between the input dimensions and model output can vary from one dimension to another, clearly suggesting that the input parameters may impact the output differently depending on the dimension. A non-isotropic covariance function provides more flexibility in capturing complex relationships. We use a composite covariance function determined by taking the sum of the RBF and Matern32 kernels.

In our experiment, we set the initial conditions for kernel hyperparameters by providing an initial length scale value of 1 and a variance of 1. This gives the emulator a starting point and a prior belief about the underlying relationships within the datasets. During training, the model learns from the provided data to understand better the relationship between the inputs and the outputs, and it dynamically updates these hyperparameters based on the likelihood of the data. This adaptive process allows the model to refine its understanding of the sensitivity of input parameters relative to the output. In the process, the model fine-tunes itself, adjusting its lengthscales, variance, and other hyperparameters better to capture the complex structures and dependencies within the data.

Since combining two kernels, we are introducing two lengthscale parameters for our model output. As a result, we now have the model output sensitivity influenced by both lengthscale parameters. The reason is that each lengthscale parameter represents the characteristic distance in its respective kernel, indicating how far two points need to be within a specific dimension before they become uncorrelated. The composite kernel allows for more flexibility and adaptability in modeling the sensitivity of our outputs to the input parameters. It reflects the recognition that different aspects of our data may have different characteristic lengthscales, and the emulator can account for these variations.

The sensitivity analysis is summarized (Fig. 3.4), where each covariance function influence is individually evaluated, and an average of both is also computed to estimate the overall uncertainty contribution by each input parameter. The idea of evaluating each lengthscale independently and then a combination presents a better overview of how changes in input parameters affect the output predictions, considering the influence of both lengthscale parameters. Smaller lengthscales indicate greater sensitivity to variations in that input dimension. The RBF lengthscales identify the forcing dataset as more sensitive to emulator output, with PSRF, RH, PRECTmms, Medlynintercept, and jmaxb1 being the most sensitive. On the contrary, the matern32 lengthscale parameters identify the input variables (medlynintercept, jmaxb1, psi50, slatop, and jmaxb0) most sensitive to the emulator. When combined, the input parameters are sensitive to emulator outputs, highlighting that medlynintercept, jmaxb1, jmaxb0, slatop, and psi50 are more sensitive.

3.3.3 Emulator Calibration/Tuning

The emulator calibration significantly improved its performance (Fig. 3.5), resulting in better alignment with observations than the previous Parameter Perturbation Ensemble (PPE) results. The mean bias error improved from -6.21 W/m² to 5.13 W/m² while the RMSE improved from 18.81 W/m² to 12.5 W/m². The R-squared value improved from 0.77 to 0.90.

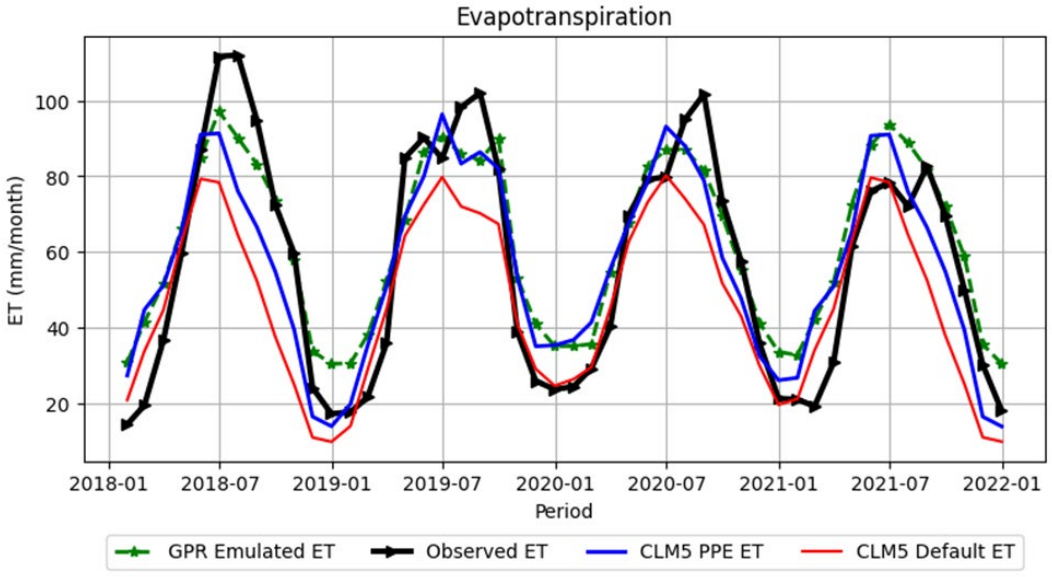

Figure 3.5: Comparison between NEON observations, CLM5 output and Emulator simulated ET

The design of the mean-based emulator and metric-based emulator does not allow monthly simulations like the top 10% PPE-based emulator. However, these two can generate a statistic that can guide optimization. Using the metric emulator, we generated samples using the MCMC sampler by setting our objective to zero. The reason was to generate input samples that yield the minimum rmse values. We then use the generated samples to simulate rmse values and identify the input values corresponding to the minimum rmse value generated. A similar approach was used for the mean-based emulator but with an objective set based on the mean observed ET. To evaluate the efficiency of the generated optimized parameters, we initialized a CLM5 model to run but with the updated new parameter sets. The CLM5 results (Fig. 3.6) for the rmse optimized parameters give rmse = 19.51 mm/month, mbe = 13.83 mm/month, and R_squared = 0.80. The mean emulator results are rmse = 53.32 mm/month, mbe = 47.97 mm/month, and R_squared = 0.77. In both cases, the optimized parameters result in an overestimation of ET.

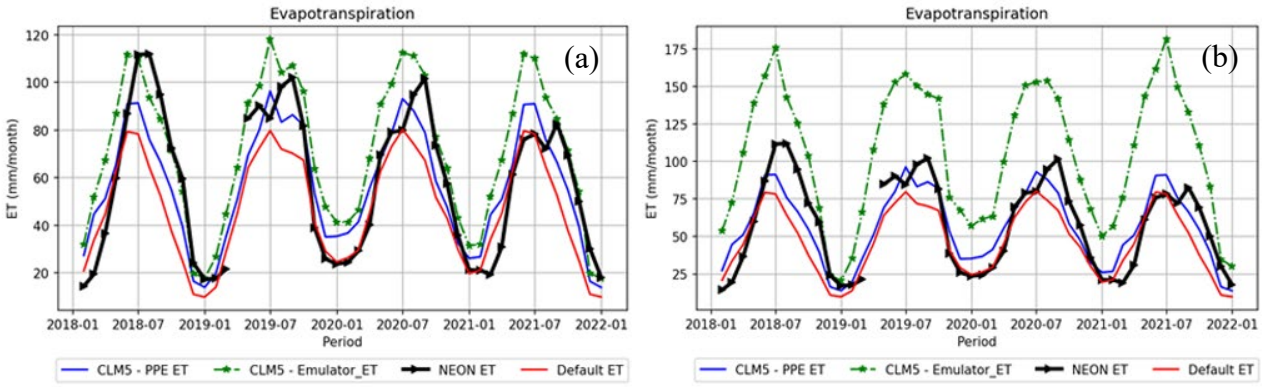

Figure 3.6: The CLM5 model output using the emulator optimized parameters. Panel (a) shows CLM5 results using rmse-based optimization, while panel (b) shows results using mean ET-optimized parameters.

The top 10% PPE emulator optimization generates similar results to mean ET when applied to run CLM5. The CLM5 model's performance is contrary to emulator predictions as the top 10% PPE-guided emulator shows more significant improvements to ET predictions than observations.

3.4 Discussion

We have developed a framework for emulating land surface models using Gaussian Process Regression, which incorporates external forcing data, allowing us to expand our initial simulation ensembles without additional effort. This approach has gained popularity in the land

surface modeling community, making it easier to perform tasks like fast prediction, parameter tuning, and sensitivity analysis (Baker et al., 2022; Chaure & Bhattacharya, 2023; Gómez-Dans et al., 2016; Gramacy, 2020; Sun et al., 2014).

The current work is motivated by creating a CLM5 emulator, especially for estimating evapotranspiration, but the procedure can be adapted for other land surface model processes. Each process and model have unique characteristics, so adjustments may be necessary to align with the modeler's goals and priorities.

Optimization is an essential application of land surface emulators, where we aim to maximize or minimize specific outputs. Many land surface models have large uncertainties with significant discrepancies between model outputs and observation data or reference datasets (Yuan et al., 2022). For instance, CLM5 underestimates and overestimates ET in different regions within the globe (Dayang Wang et al., 2021). Therefore, developing a method that minimizes the discrepancies between simulated and observed ET is fundamental. Efficient and accurate land surface emulators on any scale have the potential to inform decision-making, making this an exciting research direction. We employ the MCMC approach to sample new parameter sets driven by a target, learning from each iteration and adapting the parameters accordingly (Watson-Parris et al., 2021).

Accurate ET estimates are crucial for understanding and managing the Earth's water and energy cycles. ET is critical in linking the water, energy, carbon, and nitrogen cycles, making it a vital component of the Earth's system. Accurate ET data is essential for hydrological modeling, climate studies, and agricultural planning. Wang et al. (2021) argue that significant efforts have been made to improve the performance of LSMs by upgrading parameterization schemes and utilizing more accurate forcing and land surface data. However, substantial differences exist in quantifying the magnitude of terrestrial ET and its temporal and spatial patterns among different LSMs.

Our sensitivity analysis underscores the significant impact of specific forcing datasets on simulated Evapotranspiration. The accuracy of these forcing datasets plays a pivotal role in mitigating uncertainties within land surface models (Fyfe et al., 2021). Our prior analyses did not assess the uncertainties within the Community Land Model version 5 (CLM5) associated with these forcing datasets. However, by examining the lengthscale parameters, the emulator

development process effectively quantifies uncertainty for all input variables (Dunbar et al., 2021). This approach helps to identify significant drivers quickly and easily by assigning a lengthscale value to each input variable. The smaller the lengthscale, the greater the influence on the simulated output.

The sensitivity analysis shows *jmaxb1, jmaxb0, slatop,* and *medlynintercept* parameters as the most influential to ET uncertainties in the designed emulator. The *jmaxb1* parameter is essential in determining the response of the electron transport rate to light availability. The *jmaxb0* parameter represents the baseline proportion of nitrogen allocated for electron transport. Fisher et al. (2019) emphasize nitrogen allocation's importance in determining plant photosynthetic capacity. The response of the electron transport rate to light availability is a critical aspect of plant photosynthesis. As light availability increases, the electron transport rate rises linearly until it reaches its maximum capacity (Jmax) (Ali et al., 2016). The baseline proportion of nitrogen allocated for electron transport (Jmaxb0) significantly influences the baseline electron transport rate. A higher Jmaxb0 corresponds to an increased allocation of nitrogen for electron transport, thereby enhancing the photosynthetic capacity of plants. These relationships emphasize the link between light availability and nitrogen allocation in shaping photosynthetic processes. Understanding these dynamics is vital for ecosystem process modeling, as they shed light on how plants respond to varying light conditions and optimize their photosynthetic efficiency.

The Specific Leaf Area at the canopy top *(slatop)* (m2/gC) is an essential parameter for the canopy's ability to intercept light and perform photosynthesis. The slatop parameter influences the area-based leaf nitrogen concentration, which, in turn, determines the vcmax25—representing the maximum carboxylation rate at 25°C (Luo et al., 2020). A higher slatop value leads to a lower leaf nitrogen concentration and subsequently reduced vcmax25, directly impacting the rate of photosynthesis (Luo et al., 2020). Identifying the *slatop* as an influential parameter of ET agrees with existing research, thereby confirming its importance in ET estimations regardless of the method employed.

The *medlynintercept* parameter is equally significant in photosynthesis due to its impact on stomatal conductance, which governs carbon dioxide movement through leaf stomata. Working with the *Medlynslope* parameter, the *medlynintercept* parameter establishes the stomatal conductance model used in the CLM (Medlyn et al., 2011). A higher *medlynintercept* value

promotes increased stomatal conductance, facilitating greater carbon dioxide uptake during photosynthesis (Luo et al., 2020). A higher *medlynintercept* parameter value signifies that plants have a more sensitive stomatal response to atmospheric conditions (especially CO2 concentrations), affecting ET. Our findings are consistent with existing research, emphasizing the importance of these parameters to the projection of water cycle processes in land surface models.

Various criteria can further constrain input parameters regarding the tuning procedure. Seasonality, especially in the southeastern United States, presents challenges in identifying the target for the sampler. Our observation distribution provides a more extensive spread, and when generating samples using MCMC, the algorithm is designed to generate samples from the higher distribution. This might not be effective as it is essential to capture all regions, and carbon-water cycle processes are significantly affected by seasonal variations. Determining the most effective sampling criteria remains an open question. One of the key advantages of the emulator is its computational speed compared to running the land surface model directly (Chaure & Bhattacharya, 2023). For instance, performing parameter perturbation experiments on a supercomputer can take hours (point scale simulations take approximately 6 hours for a 200-year spin-up when performing parameter perturbation experiments). However, our emulator can be trained in about 3 minutes and run in less than 5 seconds. Parallelization can further enhance speed, and the emulator does not require spin-up time, simplifying the process of obtaining the desired information.

Despite the emulator significantly decreasing the bias between observed and simulated evapotranspiration, we noted a decline in performance for the CLM5 when evaluated using the optimized parameter values. The result shows that CLM5 significantly overestimates ET. We have demonstrated the accessibility and value of modern statistical techniques like Gaussian process emulators. While they may be sophisticated, they are also intuitive and competent. They provide substantial benefits with essential expertise compared to the complexity of developing a land surface model.

3.5 Conclusion

The current research used an emulator to minimize and quantify uncertainties in water cycle projections in land surface models. The developed emulator is robust in emulating evapotranspiration with an R2 value of 0.99 when evaluated against test datasets. The robustness and ability to capture ET dynamics, including seasonality, highlights the potential efficiently and

effectively for advancing land surface modeling and offers new avenues for research in this field. Besides, the emulator performance against NEON observations (ET) is better than the previous experiment, where we minimize ET uncertainties even further using the developed emulator.

The sensitivity analysis demonstrated the significant impact of specific forcing datasets and input parameters on ET predictions. The findings highlighted the importance of parameters such as *jmaxb1, jmaxb0, slatop,* and *medlynintercept* in driving ET uncertainties. This emphasizes the importance of these parameters in water cycle projections as they are emphasized in both approaches, including PPEs and emulator simulations. The flexibility of the designed emulator can further be assessed to explore these parameters, presenting an opportunity to improve our understanding of ecosystem processes and their response to changing environmental conditions.

The calibration process, using a Markov Chain Monte Carlo technique, improved the accuracy of the emulator by aligning its predictions with NEON observed ET. This calibration process resulted in better performance than the Parameter Perturbation Ensemble (PPE) results, further emphasizing the potential of the emulator to refine and optimize land surface model predictions. Even though the optimized parameters did not yield improvements when applied to CLM5 simulations, the result can potentially guide best practices to optimize parameters for CLM5 simulations.

The developed framework is practical and fast and can be utilized for prediction, parameter tuning, and sensitivity analysis. The approach can be adapted for different land surface models and processes, allowing researchers to address specific modeling challenges and uncertainties. This represents a promising approach for enhancing land surface modeling capabilities and addressing the challenges associated with parameterization, sensitivity, and calibration in climate model applications. The current work contributes to a better understanding of land-atmosphere interactions but reveals the power of statistical techniques in advancing Earth system modeling.

Chapter Four: Summary and Conclusions

Carbon-water cycle predictions are considerably uncertain in climate models, which overestimate vegetation growth with increasing CO2 concentrations, reducing water availability. The present work is designed to constrain carbon-water cycle prediction uncertainties in land surface models. Parameter sensitivity studies and parameter optimization techniques are extensively utilized in hydrological literature to minimize uncertainties in water projections. While ideally, parameter sensitivity and optimization techniques could potentially constrain output uncertainties in earth system models, the approach is challenging to realize due to the nature of these models. Land surface models represent many processes, making them computationally resource-intensive, even when using supercomputers. As a result, parameter sensitivity analysis and optimization techniques that require many simulations have made limited inroads into climate modeling applications. However, developments and advances in earth system science have realized possible ways of adequately exploring model input and output relationships.

The present research explores and implements a systematic methodology integrating parameter perturbation experiments, machine learning approaches, and sensitivity analysis to understand carbon-water cycle projection uncertainties in land surface models. In our initial study, we configured the Community land model to run at point scale using 400 different parameter combinations. These parameter configurations were generated by sampling identified parameters that affect carbon-water cycle processes in CLM5. The Latin Hypercube, an efficient way to adequately sample a parameter space, was used. The CLM simulations show how different parameter configurations affect the model processes by comparing the simulation results against NEON observations. As we are keen on exploring the plausible regions of the parameter yielding better alignments with observations, we extract the top 10% best-performing simulations, together with the associated parameters. Together with the forcing data, we develop a machine learning-based Gaussian processes regression emulator using these top 10% best simulations.

The parameter perturbation experiment improves CLM5 model performance. The model performance improved from 0.3 to about 0.78, measured using a composite metric that combines root mean squared error, mean bias error, and R^2. The emulator also showed better performance and improved the projections, further constraining carbon-water cycle prediction uncertainties. Sensitivity analysis from the parameter perturbation experiment and emulator identifies

photosynthesis-related parameters as critical drivers to carbon-water cycle projection uncertainties.

The present study is motivated by addressing carbon-water cycle projection uncertainties in land surface models. Variables like evapotranspiration and gross primary productivity are measured at point scale using flux tower sites and estimated using various techniques, including land surface models over larger scales. Land surface models tend to underestimate water cycle processes and overestimate carbon cycle processes. It is, therefore, essential to reduce the uncertainties between estimated products and observations, which has the potential to better guide resource management, policy development, and implementation.

The current work calibrates and improves land surface model performance at point-scale. To assess the effectiveness at a regional scale, the model performance shows mixed results, suggesting (1) the need for plant functional type calibration and (2) challenges associated with upscaling. According to the study area description, the southeastern region covered by the three ecoclimatic domains is dominated by broadleaf deciduous forests (21%). Two of our calibrations were for the needleleaf ecosystem, which covers a relatively small percentage (13%) of the region. Results suggest that the deciduous forest calibration performed relatively better than the other two calibrations, supported by the study area description.

The emulator performs better with a better prediction, which aligns with NEON observations compared to parameter perturbation experiments. As the emulator is robust and fast, we experimented with thousands of different parameter configurations. That means we could further explore the plausible parameter regions identified in the parameter perturbation experiments, leading to better predictions. However, when the optimized parameters from the emulator calibration are input into the CLM, the model does not react well. We noted that even though the emulator can be fast and efficient and generate results better aligning with observations, the method employed in its development is not ideal for generating optimized parameters for land surface models.

Understanding land surface processes is essential to develop accurate models for carbon-water cycle projections, especially in future climate scenarios. While the limited spatial coverage of surface process observations and the significant degree of heterogeneity in land surfaces restrict the assessment of land surface models, the parameter perturbation experiments approach can

effectively improve model performance. Our study demonstrates the potential of these experiments in improving land surface model performance. We also identify the most influential parameters driving water and carbon cycle projection uncertainties, including photosynthetic, hydrology, and plant hydraulic conductivity parameters, which can be constrained to optimize model performance.

Our study highlights the potential of PPEs to be an invaluable tool in improving our understanding of land surface processes and ultimately developing more accurate and efficient models. Importantly, we evaluated the applicability of point-scale calibration efforts at a larger scale and concluded that surface characteristics are a major determining factor in the success and failure of these attempts. One main observation is the need to calibrate for every pft available within the region where you want to conduct CLM5 simulations. We believe calibrating for each specific pft could potentially address the vegetation loss experienced by transferring a single calibrated pft to a heterogeneous region. The current work addresses the ongoing challenge of carbon-water cycle projection uncertainties in land surface models. It serves as a basis for future related work, providing guidelines on what works and does not yield significant results.

Future research should aim to expand on these efforts and include a broader range of parameters that influence various other processes, including processes such as vegetation dynamics, surface energy fluxes, and biogeochemical cycling. Researchers can enhance the land surface model calibration process by incorporating observational data from the National Ecological Observatory Network (NEON) and potentially minimize model output uncertainties. NEON provides access to high-quality and high-frequency data on climate, vegetation, and hydrology, which can be used to refine and validate model simulations. Leveraging these datasets could improve model performance by addressing the challenges associated with global-scale parameter calibration and sensitivity experiments in earth system models.

Chapter Five: References

Ainsworth, E. A., & Long, S. P. (2005). What have we learned from 15 years of free-air CO2 enrichment (FACE)? A meta-analytic review of the responses of photosynthesis, canopy properties and plant production to rising CO2. *New Phytologist*, *165*(2), 351–372. https://doi.org/10.1111/j.1469-8137.2004.01224.x

Ali, A. A., Xu, C., Rogers, A., Fisher, R. A., Wullschleger, S. D., Massoud, E. C., et al. (2016). A global scale mechanistic model of photosynthetic capacity (LUNA V1.0). *Geoscientific Model Development*, *9*(2), 587–606. https://doi.org/10.5194/gmd-9-587-2016

Andrew, D. F., Wolfgang, L., Tim, T. R., Rozenn, K., Richard, B., Patricia, C., et al. (2014). Carbon residence time dominates uncertainty in terrestrial vegetation responses to future climate and atmospheric CO2 _ Enhanced Reader. *Proceedings of the National Academy of Sciences (PNAS)*, *111*(9), 3280–3285. https://doi.org/10.1073/pnas.1222477110

Arora, V. K., Boer, G. J., Friedlingstein, P., Eby, M., Jones, C. D., Christian, J. R., et al. (2013). Carbon-concentration and carbon-climate feedbacks in CMIP5 earth system models. *Journal of Climate*, *26*(15), 5289–5314. https://doi.org/10.1175/JCLI-D-12-00494.1

Bai, Z., Wei, H., Xiao, Y., Song, S., & Kucherenko, S. (2021). A vine copula-based global sensitivity analysis method for structures with multidimensional dependent variables. *Mathematics*, *9*(19), 1–20. https://doi.org/10.3390/math9192489

Baker, E., Harper, A., Williamson, D., & Challenor, P. (2021). Emulation of high-resolution land surface models with sparse Gaussian processes with application to JULES. *Geoscientific Model Development*. https://doi.org/10.5194/gmd-2021-205

Baker, E., Harper, A. B., Williamson, D., & Challenor, P. (2022). Emulation of high-resolution land surface models using sparse Gaussian processes with application to JULES. *Geoscientific Model Development*, *15*(5), 1913–1929. https://doi.org/10.5194/gmd-15-1913-2022

Boit, A., Sakschewsk, B., Boysen, L., Crespo, A. C.-, Clement, J., Garcia-Alaniz, N., et al. (2016). Large-scale impact of climate change vs land-use change on future biome shifts in Latin America. *Global Change Biology*, *22*, 3689–3701.

Bonan, G. B., Lombardozzi, D. L., & Wieder, W. R. (2021). The signature of internal variability in the terrestrial carbon cycle. *Environmental Research Letters*, *16*(3). https://doi.org/10.1088/1748-9326/abd6a9

Booth, B. B. B., Jones, C. D., Collins, M., Totterdell, I. J., Cox, P. M., Sitch, S., et al. (2012). High sensitivity of future global warming to land carbon cycle processes. *Environmental Research Letters*, *7*(2). https://doi.org/10.1088/1748-9326/7/2/024002

Brovkin, V., Boysen, L., Raddatz, T., Gayler, V., Loew, A., & Claussen, M. (2013). Evaluation of vegetation cover and land-surface albedo in MPI-ESM CMIP5 simulations. *Journal of Advances in Modeling Earth Systems*, *5*(1), 48–57. https://doi.org/10.1029/2012MS000169

Bytnerowicz, T. A., Akana, P. R., Griffin, K. L., & Menge, D. N. L. (2022). Temperature sensitivity of woody nitrogen fixation across species and growing temperatures. *Nature Plants*, *8*(3), 209–216. https://doi.org/10.1038/s41477-021-01090-x

Chaure, A., & Bhattacharya, S. (2023). Gaussian Process Regression for Climate Modeling : Potentials , Limitations , and Advances in Emulation Techniques, (May). https://doi.org/10.1729/Journal.34007

Collier, N., Hoffman, F. M., Lawrence, D. M., Keppel-Aleks, G., Koven, C. D., Riley, W. J., et al. (2018). The International Land Model Benchmarking (ILAMB) System: Design, Theory, and Implementation. *Journal of Advances in Modeling Earth Systems*, *10*(11), 2731–2754. https://doi.org/10.1029/2018MS001354

Cox, P. M., Betts, R. A., Jones, C. D., & Spall, S. A. (2000). Acceleration of global warming due to carbon-cycle feedbacks in acoupled climate model. *Nature*, *408*(November), 184–187.

Dagon, K. (2020). Supplement of A machine learning approach to emulation and biophysical parameter estimation with the Community Land Model, version 5 The copyright of individual parts of the supplement might differ from the CC BY 4.0 License. *Adv. Stat. Clim. Meteorol. Oceanogr*, *6*, 223–244. Retrieved from https://doi.org/10.5194/ascmo-6-223-2020-supplement

Dagon, K., Sanderson, B. M., Fisher, R. A., & Lawrence, D. M. (2020). A machine learning approach to emulation and biophysical parameter estimation with the Community Land

Model, version 5. *Advances in Statistical Climatology, Meteorology and Oceanography*, *6*(2), 223–244. https://doi.org/10.5194/ascmo-6-223-2020

Deb, J. C., Phinn, S., Butt, N., & McAlpine, C. A. (2018). Climate change impacts on tropical forests: Identifying risks for tropical Asia. *Journal of Tropical Forest Science*, *30*(2), 182–194. https://doi.org/10.26525/jtfs2018.30.2.182194

Dewitz, J., & U.S. Geological Survey. (2021). *National Land Cover Database (NLCD) 2019 Products (ver. 2.0, June 2021)*.

Dombrowski, O., Brogi, C., Hendricks Franssen, H. J., Zanotelli, D., & Bogena, H. (2022). CLM5-FruitTree: A new sub-model for deciduous fruit trees in the Community Land Model (CLM5). *Geoscientific Model Development*, *15*(13), 5167–5193. https://doi.org/10.5194/gmd-15-5167-2022

Dunbar, O. R. A., Garbuno-Inigo, A., Schneider, T., & Stuart, A. M. (2021). Calibration and Uncertainty Quantification of Convective Parameters in an Idealized GCM. *Journal of Advances in Modeling Earth Systems*, *13*(9). https://doi.org/10.1029/2020MS002454

Fang, H., Baret, F., Plummer, S., & Schaepman-Strub, G. (2019). An Overview of Global Leaf Area Index (LAI): Methods, Products, Validation, and Applications. *Reviews of Geophysics*, *57*(3), 739–799. https://doi.org/10.1029/2018RG000608

Fanourgakis, G. S., Kanakidou, M., Nenes, A., Bauer, S. E., Bergman, T., Carslaw, K. S., et al. (2019). Evaluation of global simulations of aerosol particle and cloud condensation nuclei number, with implications for cloud droplet formation. *Atmospheric Chemistry and Physics*, *19*(13), 8591–8617. https://doi.org/10.5194/acp-19-8591-2019

Fisher, R. A., & Koven, C. D. (2020). Perspectives on the Future of Land Surface Models and the Challenges of Representing Complex Terrestrial Systems. *Journal of Advances in Modeling Earth Systems*, *12*(4). https://doi.org/10.1029/2018MS001453

Fisher, R. A., Wieder, W. R., Sanderson, B. M., Koven, C. D., Oleson, K. W., Xu, C., et al. (2019). Parametric Controls on Vegetation Responses to Biogeochemical Forcing in the CLM5. *Journal of Advances in Modeling Earth Systems*, *11*(9), 2879–2895. https://doi.org/10.1029/2019MS001609

Flo, V., Martínez-Vilalta, J., Mencuccini, M., Granda, V., Anderegg, W. R. L., & Poyatos, R. (2021). Climate and functional traits jointly mediate tree water-use strategies. *New Phytologist*, *231*(2), 617–630. https://doi.org/10.1111/nph.17404

Fyfe, J. C., Kharin, V. V., Santer, B. D., Cole, J. N. S., & Gillett, N. P. (2021). Significant impact of forcing uncertainty in a large ensemble of climate model simulations. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(23), 1–6. https://doi.org/10.1073/pnas.2016549118

Gao, X., Avramov, A., Saikawa, E., & Adam Schlosser, C. (2021). Emulation of community land model version 5 (Clm5) to quantify sensitivity of soil moisture to uncertain parameters. *Journal of Hydrometeorology*, *22*(2), 259–278. https://doi.org/10.1175/JHM-D-20-0043.1

Gentine, P., Green, J. K., Guérin, M., Humphrey, V., Seneviratne, S. I., Zhang, Y., & Zhou, S. (2019). Coupling between the terrestrial carbon and water cycles - A review. *Environmental Research Letters*, *14*(8). https://doi.org/10.1088/1748-9326/ab22d6

Gómez-Dans, J. L., Lewis, P. E., & Disney, M. (2016). Efficient emulation of radiative transfer codes using gaussian processes and application to land surface parameter inferences. *Remote Sensing*, *8*(2), 1–32. https://doi.org/10.3390/rs8020119

Gong, W., Duan, Q., Li, J., Wang, C., Di, Z., Dai, Y., et al. (2015). Multi-objective parameter optimization of common land model using adaptive surrogate modeling. *Hydrology and Earth System Sciences*, *19*(5), 2409–2425. https://doi.org/10.5194/hess-19-2409-2015

Gramacy, R. B. (2020). *Surrogates: Gaussian process modeling, design, and optimization for the applied sciences*. *Journal of Quality Technology* (Vol. 53). Taylor & Francis Group, LLC. https://doi.org/10.1080/00224065.2020.1764416

Gray, C. L., Hill, S. L. L., Newbold, T., Hudson, L. N., Boïrger, L., Contu, S., et al. (2016). Local biodiversity is higher inside than outside terrestrial protected areas worldwide. *Nature Communications*, *7*. https://doi.org/10.1038/ncomms12306

Gul, M. U., Paul, A., Manimurugan, S., & Chehri, A. (2023). Hydrotropism: Understanding the Impact of Water on Plant Movement and Adaptation. *Water (Switzerland)*, *15*(3), 1–16. https://doi.org/10.3390/w15030567

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, *377*(1–2), 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003

Heimann, M., & Reichstein, M. (2008). Terrestrial ecosystem carbon dynamics and climate feedbacks. *Nature*, *451*(7176), 289–292. https://doi.org/10.1038/nature06591

Jacovides, C. P., & Kontoyiannis, H. (1995). *Agricultural water management Statistical procedures for the evaluation of evapotranspiration computing models*. *Agricultural Water Management* (Vol. 27).

Jennifer, M. (2017). *NEON Site-Level Plot Summary Oak Ridge National Laboratory (ORNL) Document Information Site Background*.

Jin, H., Li, A., Bian, J., Nan, X., Zhao, W., Zhang, Z., & Yin, G. (2017). Intercomparison and validation of MODIS and GLASS leaf area index (LAI) products over mountain areas: A case study in southwestern China. *International Journal of Applied Earth Observation and Geoinformation*, *55*, 52–67. https://doi.org/10.1016/j.jag.2016.10.008

Jung, C., Lee, Y., Cho, Y., & Kim, S. (2017). A study of spatial soil moisture estimation using a multiple linear regression model and MODIS land surface temperature data corrected by conditional merging. *Remote Sensing*, *9*(8). https://doi.org/10.3390/rs9080870

De Kauwe, M. G., Medlyn, B. E., Knauer, J., & Williams, C. A. (2017). Ideas and perspectives: How coupled is the vegetation to the boundary layer? *Biogeosciences*, *14*(19), 4435–4453. https://doi.org/10.5194/bg-14-4435-2017

Kennedy, D., Swenson, S., Oleson, K. W., Lawrence, D. M., Fisher, R., Lola da Costa, A. C., & Gentine, P. (2019). Implementing Plant Hydraulics in the Community Land Model, Version 5. *Journal of Advances in Modeling Earth Systems*, *11*(2), 485–513. https://doi.org/10.1029/2018MS001500

Koven, C. D., Knox, R. G., Fisher, R. A., Fisher, R. A., Chambers, J. Q., Chambers, J. Q., et al. (2020). Benchmarking and parameter sensitivity of physiological and vegetation dynamics using the Functionally Assembled Terrestrial Ecosystem Simulator (FATES) at Barro

Colorado Island, Panama. *Biogeosciences*, *17*(11), 3017–3044. https://doi.org/10.5194/bg-17-3017-2020

Lawrence, D. M., Fisher, R. A., Koven, C. D., Oleson, K. W., Swenson, S. C., Bonan, G., et al. (2019). The Community Land Model Version 5: Description of New Features, Benchmarking, and Impact of Forcing Uncertainty. *Journal of Advances in Modeling Earth Systems*, *11*(12), 4245–4287. https://doi.org/10.1029/2018MS001583

Lewis, S. L. (2006). Tropical forests and the changing earth system. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *361*(1465), 195–210. https://doi.org/10.1098/rstb.2005.1711

Li, J., Chen, F., Zhang, G., Barlage, M., Gan, Y., Xin, Y., & Wang, C. (2018). Impacts of Land Cover and Soil Texture Uncertainty on Land Model Simulations Over the Central Tibetan Plateau. *Journal of Advances in Modeling Earth Systems*, *10*(9), 2121–2146. https://doi.org/10.1029/2018MS001377

Li, S., Rupp, D. E., Hawkins, L., Mote, P. W., McNeall, D., Sparrow, S. N., et al. (2019). Reducing climate model biases by exploring parameter space with large ensembles of climate model simulations and statistical emulation. *Geoscientific Model Development*, *12*(7), 3017–3043. https://doi.org/10.5194/gmd-12-3017-2019

Loeppky, J. L., Sacks, J., & Welch, W. J. (2009). Choosing the sample size of a computer experiment: A practical guide. *Technometrics*, *51*(4), 366–376. https://doi.org/10.1198/TECH.2009.08040

Lombardozzi, D., Levis, S., Bonan, G., & Sparks, J. P. (2012). Predicting photosynthesis and transpiration responses to ozone: Decoupling modeled photosynthesis and stomatal conductance. *Biogeosciences*, *9*(8), 3113–3130. https://doi.org/10.5194/bg-9-3113-2012

Lombardozzi, D. L., Wieder, W. R., Sobhani, N., Bonan, G. B., & Durden, D. (2023). Overcoming barriers to enable convergence research by integrating ecological and climate sciences : The NCAR-NEON system Version 1 Short Summary, (April), 1–37.

Lovenduski, N. S., & Bonan, G. B. (2017). Reducing uncertainty in projections of terrestrial carbon uptake. *Environmental Research Letters*, *12*(4). https://doi.org/10.1088/1748-9326/aa66b8

Lu, D., & Ricciuto, D. (2019). Efficient surrogate modeling methods for large-scale Earth system models based on machine-learning techniques. *Geoscientific Model Development*, *12*(5), 1791–1807. https://doi.org/10.5194/gmd-12-1791-2019

Lucas, D. D., Klein, R., Tannahill, J., Ivanova, D., Brandon, S., Domyancic, D., & Zhang, Y. (2013). Failure analysis of parameter-induced simulation crashes in climate models. *Geoscientific Model Development*, *6*(4), 1157–1171. https://doi.org/10.5194/gmd-6-1157-2013

Luo, Q., Wen, J., Hu, Z., Lu, Y., & Yang, X. (2020). Parameter Sensitivities of the Community Land Model at Two Alpine Sites in the Three-River Source Region. *Journal of Meteorological Research*, *34*(4), 851–864. https://doi.org/10.1007/s13351-020-9205-8

Luo, Q., Zhang, T., & Li, Z. (2022). Numerical simulation of the surface flux of an alpine grassland in the source region of the Yellow River by the land surface model. *Arabian Journal of Geosciences*, *15*(22). https://doi.org/10.1007/s12517-022-10993-8

Mara, T. A., Tarantola, S., & Annoni, P. (2015). Non-parametric methods for global sensitivity analysis of model output with dependent inputs. *Environmental Modelling and Software*, *72*, 173–183. https://doi.org/10.1016/j.envsoft.2015.07.010

Marck C. Kennedy, & Anthony O'Hagan. (2001). Bayesian calibration of computer models. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, *63*(3), 425–464. https://doi.org/10.1111/1467-9868.00294

Matthews, A. G. de G., van der Wilk, M., Nickson, T., Fujii, K., Boukouvalas, A., Pablo, L.-V., et al. (2017). GPflow: A Gaussian process library using TensorFlow. *Journal of Machine Learning Research*, *18*(40), 1–6.

McKay, M. D., Beckman, R. J., & Conover, W. J. (2000). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, *42*(1), 55–61. https://doi.org/10.1080/00401706.2000.10485979

Medlyn, B. E., Duursma, R. A., Eamus, D., Ellsworth, D. S., Prentice, I. C., Barton, C. V. M., et al. (2011). Reconciling the optimal and empirical approaches to modelling stomatal conductance. *Global Change Biology*, *17*(6), 2134–2144. https://doi.org/10.1111/j.1365-2486.2010.02375.x

Murphy, C., Fealy, R., Charlton, R., & Sweeney, J. (2006). The reliability of an "off-the-shelf" conceptual rainfall runoff model for use in climate impact assessment: Uncertainty quantification using Latin hypercube sampling. *Area*. https://doi.org/10.1111/j.1475-4762.2006.00656.x

NEON (National Ecological Observatory Network). (n.d.-a). Oak Ridge NEON | NSF NEON | Open Data to Understand our Ecosystems. Retrieved October 19, 2023, from https://www.neonscience.org/field-sites/ornl

NEON (National Ecological Observatory Network). (n.d.-b). Ordway-Swisher Biological Station NEON | NSF NEON | Open Data to Understand our Ecosystems. Retrieved October 19, 2023, from https://www.neonscience.org/field-sites/osbs

NEON (National Ecological Observatory Network). (n.d.-c). Talladega National Forest NEON | NSF NEON | Open Data to Understand our Ecosystems. Retrieved October 19, 2023, from https://www.neonscience.org/field-sites/tall

Pang, M., Xu, R., Hu, Z., Wang, J., & Wang, Y. (2021). Uncertainty and sensitivity analysis of input conditions in a large shallow lake based on the latin hypercube sampling and morris methods. *Water (Switzerland)*, *13*(13). https://doi.org/10.3390/w13131861

Powell, T. L., Koven, C. D., Johnson, D. J., Faybishenko, B., Fisher, R. A., Knox, R. G., et al. (2018). Variation in hydroclimate sustains tropical forest biomass and promotes functional diversity. *New Phytologist*, *219*(3), 932–946. https://doi.org/10.1111/nph.15271

Prihodko, L., Denning, A. S., Hanan, N. P., Baker, I., & Davis, K. (2008). Sensitivity, uncertainty and time dependence of parameters in a complex land surface model. *Agricultural and Forest Meteorology*, *148*(2), 268–287. https://doi.org/10.1016/j.agrformet.2007.08.006

Le Quéré, C., Moriarty, R., Andrew, R. M., Canadell, J. G., Sitch, S., Korsbakken, J. I., et al. (2015). Global Carbon Budget 2015. *Earth System Science Data*, *7*(2), 349–396. https://doi.org/10.5194/essd-7-349-2015

Raoult, N., Ottlé, C., Peylin, P., Bastrikov, V., & Maugis, P. (2021). Evaluating and Optimizing Surface Soil Moisture Drydowns in the ORCHIDEE Land Surface Model at In Situ Locations. *Journal of Hydrometeorology*, *22*(4). https://doi.org/https://doi.org/10.1175/JHM-D-20-0115.1

Rasmussen, C. E., & Williams, C. K. I. (2006). *Gaussian Processes for Machine Learning*. *Journal fur Urologie und Urogynakologie* (Vol. 7). the Massachusetts Institute of Technology (MIT) Press. Retrieved from www.GaussianProcess.org/gpml

Razi, S., Shervan, G., Simon, Michael, P., & Martyn, P. C. (2021). VISCOUS A Variance-Based Sensitivity Analysis Using Copulas for Efficient Identification of Dominant Hydrological Processes. *Water Resource Research*, *57*.

Richard V. Field, Paul Constantine, & Mark Boslough. (2011). *SANDIA REPORT: Statistical surrogate models for prediction of high-consequence climate change*. Retrieved from http://www.ntis.gov/help/ordermethods.asp?loc=7-4-0#online

Rogers, A., Medlyn, B. E., Dukes, J. S., Bonan, G., Caemmerer, S. von, Michael Dietze, C., et al. (2017). Viewpoints A roadmap for improving the representation of photosynthesis in Earth system models.

Ryu, Y., Verfaillie, J., Macfarlane, C., Kobayashi, H., Sonnentag, O., Vargas, R., et al. (2012). Continuous observation of tree leaf area index at ecosystem scale using upward-pointing digital cameras. *Remote Sensing of Environment*, *126*, 116–125. https://doi.org/10.1016/j.rse.2012.08.027

Sabine, C., Bala, G., Bopp, L., Brovkin, V., Canadell, J., Chhabra, A., et al. (2013). Evaluation of Climate Models. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. In T. F. , Stocker, D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, et al.

(Eds.), *Climate Change 2013: The Physical Science Basis*. Cambridge, United Kingdom and New York, NY, USA: Cambridge University press.

van der Sande, M. T., Poorter, L., Kooistra, L., Balvanera, P., Thonicke, K., Thompson, J., et al. (2017). Biodiversity in species, traits, and structure determines carbon stocks and uptake in tropical forests. *Biotropica*, *49*(5), 593–603. https://doi.org/10.1111/btp.12453

Sawada, Y. (2020). Machine Learning Accelerates Parameter Optimization and Uncertainty Assessment of a Land Surface Model. *Journal of Geophysical Research: Atmospheres*, *125*(20). https://doi.org/10.1029/2020JD032688

Schuur, E. A. G., McGuire, A. D., Schädel, C., Grosse, G., Harden, J. W., Hayes, D. J., et al. (2015). Climate change and the permafrost carbon feedback. *Nature*, *520*(7546), 171–179. https://doi.org/10.1038/nature14338

Singh, A., Kumar, S., Akula, S., Lawrence, D. M., & Lombardozzi, D. L. (2020). Plant Growth Nullifies the Effect of Increased Water-Use Efficiency on Streamflow Under Elevated CO2 in the Southeastern United States. *Geophysical Research Letters*, *47*(4). https://doi.org/10.1029/2019GL086940

Sun, A. Y., Wang, D., & Xu, X. (2014). Monthly streamflow forecasting using Gaussian process regression. *Journal of Hydrology*, *511*, 72–81. https://doi.org/10.1016/j.jhydrol.2014.01.023

Tewari, A., Giering, M. J., & Raghunathan, A. (2011). Parametric characterization of multimodal distributions with non-Gaussian modes. *Proceedings - IEEE International Conference on Data Mining, ICDM*, 286–292. https://doi.org/10.1109/ICDMW.2011.135

Vicente-Serrano, S. M., Miralles, D. G., McDowell, N., Brodribb, T., Domínguez-Castro, F., Leung, R., & Koppa, A. (2022). The uncertain role of rising atmospheric CO2 on global plant transpiration. *Earth-Science Reviews*, *230*(May), 104055. https://doi.org/10.1016/j.earscirev.2022.104055

Wang, Dagang, Wang, G., & Anagnostou, E. N. (2007). Evaluation of canopy interception schemes in land surface models. *Journal of Hydrology*, *347*(3–4), 308–318. https://doi.org/10.1016/j.jhydrol.2007.09.041

Wang, Dayang, Wang, D., & Mo, C. (2021). The use of remote sensing-based et estimates to improve global hydrological simulations in the community land model version 5.0. *Remote Sensing*, *13*(21). https://doi.org/10.3390/rs13214460

Watson-Parris, D., Williams, A., Deaconu, L., & Stier, P. (2021). Model calibration using ESEm v1.1.0-an open, scalable Earth system emulator. *Geoscientific Model Development*, *14*(12), 7659–7672. https://doi.org/10.5194/gmd-14-7659-2021

Weber, T., Corotan, A., Hutchinson, B., Kravitz, B., & Link, R. (2020). Technical note: Deep learning for creating surrogate models of precipitation in Earth system models. *Atmospheric Chemistry and Physics*, *20*(4), 2303–2317. https://doi.org/10.5194/acp-20-2303-2020

Wieder, W. R., Knowles, J. F., Blanken, P. D., Swenson, S. C., & Suding, K. N. (2017). Ecosystem function in complex mountain terrain: Combining models and long-term observations to advance process-based understanding. *Journal of Geophysical Research: Biogeosciences*, *122*(4), 825–845. https://doi.org/10.1002/2016JG003704

Wieder, W. R., Lawrence, D. M., Fisher, R. A., Bonan, G. B., Cheng, S. J., Goodale, C. L., et al. (2019). Beyond Static Benchmarking: Using Experimental Manipulations to Evaluate Land Model Assumptions. *Global Biogeochemical Cycles*, *33*(10), 1289–1309. https://doi.org/10.1029/2018GB006141

Williamson, D. (2015). Exploratory ensemble designs for environmental models using k-extended Latin Hypercubes. *Environmetrics*. https://doi.org/10.1002/env.2335

Winkler, A. J., Myneni, R. B., Alexandrov, G. A., & Brovkin, V. (2019). Earth system models underestimate carbon fixation by plants in the high latitudes. *Nature Communications*, *10*(1). https://doi.org/10.1038/s41467-019-08633-z

Wu, Y., Miao, C., Fan, X., Gou, J., Zhang, Q., & Zheng, H. (2022). Quantifying the Uncertainty Sources of Future Climate Projections and Narrowing Uncertainties With Bias Correction Techniques. *Earth's Future*, *10*(11), 1–16. https://doi.org/10.1029/2022EF002963

Xiao, J., Chen, J., Davis, K. J., & Reichstein, M. (2012). Advances in upscaling of eddy covariance measurements of carbon and water fluxes. *Journal of Geophysical Research: Biogeosciences*, *117*(1), 2–5. https://doi.org/10.1029/2011JG001889

Xiong, D., & Nadal, M. (2020). Linking water relations and hydraulics with photosynthesis. *Plant Journal*, *101*(4), 800–815. https://doi.org/10.1111/tpj.14595

Yadav, N., Ravela, S., & Ganguly, A. R. (2020). Machine learning for robust identification of complex nonlinear dynamical systems: Applications to earth systems modeling. *ArXiv*.

Yan, H., Sun, N., Eldardiry, H., Thurber, T. B., Reed, P. M., Malek, K., et al. (2023). Characterizing uncertainty in Community Land Model version 5 hydrological applications in the United States. *Scientific Data*, *10*(1), 187. https://doi.org/10.1038/s41597-023-02049-7

Yuan, K., Zhu, Q., Riley, W. J., Li, F., & Wu, H. (2022). Understanding and reducing the uncertainties of land surface energy flux partitioning within CMIP6 land models. *Agricultural and Forest Meteorology*, *319*(September 2021). https://doi.org/10.1016/j.agrformet.2022.108920

Zhang, C., Wang, D., Wang, G., Yang, W., & Liu, X. (2014). Regional differences in hydrological response to canopy interception schemes in a land surface model. *Hydrological Processes*, *28*(4), 2499–2508. https://doi.org/10.1002/hyp.9762

Chapter Six: Appendices

6.1 Appendix A: Parameter Description

6.1.1 Hydrology

The baseflow_scalar is a scalar multiplier for the baseflow rate and is only applied when the lower boundary condition is not an aquifer or a water table. At the same time, n_baseflow is the drainage power exponent. The baseflow_scalar is a name list parameter in CLM-5, which can be modified to increase or decrease subsurface runoff. Lateral sub-surface runoff occurs when saturated soil moisture conditions exist within the soil column.

$$q_{drai} = \Theta_{ice} K_{baseflow} fan(\beta) \Delta z_{sat}{}^{N_{baseflow}}$$

Where $K_{baseflow}$ is a calibration parameter, β is the topographic slope, the exponent $N_{baseflow}$ = 1, and Δz_{sat} is the thickness of the saturated portion of the soil column and Θ_{ice} is an ice impedance factor. The drainage power exponent, the "n_baseflow" parameter in the CLM5, controls water movement from the soil into deeper layers. It denotes the water flow rate through the soil as a function of the hydraulic gradient. The *n_baseflow* parameter can help regulate the quantity of water available, especially during the dry period for stream flow and other ecosystem processes.

The "zbedrock_sf" parameter refers to the bedrock depth scalar adjustment factor and represents a scaling factor applied to the standard bedrock depth values in CLM5, usually determined according to geological characteristics. The *zbedrock* parameter in CLM5 regulates the depth at which roots can penetrate, thus affecting plant water uptake and nutrient cycling.

$$\Delta z_{sat} = z_{bedrock} - z_{\nabla}$$

Where z_{∇} is the water table depth determined by ascertaining the first soil layer above the bedrock depth where the soil volumetric water content falls below a specific threshold, typically 0.9.

The specific yield (s_y) depends on water table depth and soil properties. It is obtained by deriving the difference between two soil moisture profiles at equilibrium with a water table depth of infinitesimal amount.

$$S_y = \theta_{sat}\left(1 - \left(1 + \frac{z_\nabla}{\psi_{sat}}\right)^{-\frac{1}{B}}\right)$$

Where B ("aq_sp_yield_min" parameter) is the Clapp-Hornberger exponent and describes the relationship between soil water content and its ability to transmit water.

The CLM5 applies a TOPMODEL-based runoff model known as SIMTOP (Niu et al., 2005) to parameterize runoff. The fractional saturated area (fsat) is a critical component of this approach and is determined by the grid-cell soil moisture state and topography. The fractional saturated area is a function of soil moisture.

$$f_{sat} = f_{max} \exp\left(-0.5 f_{over} z_\nabla\right)$$

Where f_{max} refers to the potential or maximum fractional saturated area value, f_{over} refers to the decay factor (fff m^{-1}) and z_∇ is the water table depth in meters. The decay factor ("*fff* ") parameter explains how the proportion of the saturated area of a grid cell decreases over time due to evaporation and drainage and determines the quantity of water available in the soil for plants, runoff, and groundwater recharge as it affects water movement within the soil and has a significant effect on vegetation growth and ecosystem dynamics.

The hydraulic conductivity and soil matric potential in CLM5 for specific soil layers varies according to soil texture and the volumetric soil water. Like soil thermal properties, soil hydraulic properties could be a weighted mixture of soil mineral properties established according to the soil clay and sand and its organic matter content. The hydraulic conductivity is determined at two adjacent layers of interface depth in CLM5. It is a function of saturated hydraulic conductivity (k_{sat}), soil moisture content, and the ice impedance factor (Lawrence et al., 2019). The *hksat_sf* parameter refers to the scalar adjustment factor for saturated hydraulic conductivity in CLM5. It determines the rate at which water can move through the soil, affecting plant water uptake and soil moisture dynamics.

$$k[z_h, i] = \begin{cases} \Theta_{ice} k_{sat}[z_h, i]\left[\dfrac{0.5(\theta_i + \theta_{i+1})}{0.5(\theta_{sat}, i + \theta_{sat,i+1})}\right]^{2B_i+3}, & 1 \le i \le N_{levsoi} - 1 \\ \Theta_{ice} k_{sat}[z_h, i]\left(\dfrac{\theta_i}{\theta_{sat}, i}\right)^{2B_i+3}, & i = N_{levsoi} \end{cases}$$

Where $k[z_h, i]$ is the saturated hydraulic conductivity, θ_i & $\theta_{i+1}$is the liquid soil moisture volumetric content of the two adjacent soil layers, Θ_{ice} is the impedance ice factor, and θ_{sat} is the volumetric water content in the soil at saturation (porosity).

The *sucsat_sf* parameter refers to the saturated soil matric potential scalar adjustment factor. It measures the required energy to extract water from the soil, which is closely linked to the capillary and adsorptive forces in the soil matrix. Soil matric potential (ϕ_i mm) is defined at node depth z_i for each soil layer "i" in CLM5.

$$\psi_i = \psi_{sat,i}\left(\frac{\theta_i}{\theta_{sat,i}}\right)^{-B_i} \geq -1 \times 10^8 \quad 0.01 \leq \frac{\theta_i}{\theta_{sat,i}} \leq 1$$

Where ψ_i is the soil matric potential, and ψ_{sat} refers to the saturated soil matric potential. The sucsat_sf parameter is based on specific conditions such as the soil organic matter content, soil structure, and texture and affects the soil's ability to transmit and hold water. The scalar adjustment factor for the soil water holding capacity, also known as the "watsat_sf, " is based on soil characteristics, including soil organic matter content and texture, and regulates the maximum amount of water that can be held within the soil.

The "liq_canopy_storage_scalar" determines the scalar adjustment factor for the liquid canopy water storage capacity in CLM5 based on leaf area index and vegetation type. The parameter determines the maximum amount of water that can be stored on plant canopies intercepted from precipitation, affecting the amount of water reaching the soil and evaporating from the plant canopies. It is worth noting that evaporation and interception of precipitation on vegetation surfaces have significant consequences on any ecosystem's water and energy balances.

$$W_{can,liq}^{max} = p_{liq}(L + S)$$

Where $W_{can,liq}^{max}$ (mm or Kg/m of water) is the maximum amounts of liquid water a canopy can hold, L is the exposed leaf area index and S is the exposed stem area index.

The "maximum_leaf_wetted_fraction" refers to the maximum fraction of leaf area that can be moistened by any form of precipitation and affects the quantity of water intercepted by plant canopies. This parameter affects the plant canopy interception process and implies plant water use efficiency and ecosystem water balance.

$$more\ significant f_{wet} = \begin{cases} \left[\frac{W_{can}}{p_{liq}(L+S)}\right]^{2/3} \geq 1, & L+S>0 \\ 0, & L+S=0 \end{cases}$$

Where f_{wet} is wetted canopy fraction, W_{can} is canopy water and p_{liq} refers to the maximum storage of liquid water.

6.1.2 Photosynthesis Capacity (LUNA)

The J_{maxb0} parameter indicates the allocated nitrogen proportion for electron transport in vegetation in CLM5. The parameter represents the maximum rate of electron transport or the potential maximum rate of photosynthesis under optimal conditions. The J{maxb1} parameter regulates the response of the electron transport rate to light availability when the level of Ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) enzyme activity is lower than its maximum value. The *Wc2Wjb0* parameter denotes the baseline ratio or Rubisco-limited rate to light-limited photosynthetic rate (Wc: Wj) at a reference temperature of 250C. The "WcWjb0" simulates the effects of temperature on the balance between Rubisco-limited and light-limited photosynthesis. The Wc: Wj ratio is fundamental to determining overall vegetation photosynthesis capacity and dramatically varies based on the light intensity, temperature, and carbon dioxide concentrations. The Jmaxb0, Jmaxb1, and Wc2Wjb0 parameters are estimated by fitting a LUNA model to a global compilation of more than 800 observations located at various canopy locations, biomes, and time of the year between 1993 and 2013 (Ali et al., 2015)

The maximum electron transport rate (j_{max}: electron / m^2/s) is simulated in CLM5 to obtain reference nitrogen allocation and additional nitrogen allocation to vary based on mean daytime photosynthesis active radiation, day length, and air humidity.

$$j_{max} = j_{max0} + j_{maxb1} f(day\ length) f(humidity) \alpha PAR$$

$$j_{max0} = j_{maxb0} FNC_a NUE j_{max}$$

Where j_{max0} refers to the reference nitrogen allocation rate, j_{maxb1} is a coefficient determining electron transport rate response to light absorption, j_{maxb0} is the reference proportion of nitrogen allocated for electron transport, FNC_a refers to a functional nitrogen content, αPAR refers to the amount of absorbed light and FNCNUEj_{max} refers to the nitrogen use efficiency of j_{max}

6.1.3 Stomatal Resistance and Photosynthesis

The default CLM5 stomatal conductance computations use the Medlyn conductance model. In CLM5, the leaf stomatal resistance required for water vapor flux is linked to leaf photosynthesis. The Medlyn model computes stomatal conductance based on the net leaf photosynthesis, leaf surface CO2 concentrations, and vapor pressure deficit. The medlynslope refers to the Medlyn slope of conductance to photosynthesis relationship, while the medlynintercept refers to the Medlyn intercept of conductance to photosynthesis relationship. The equation where these two parameters affect is presented below.

$$One\frac{1}{r_s} = g_s = g_0 + 1.6\left(1 + \frac{g_1}{\sqrt{D}}\right)\frac{A_n}{C_s/P_{atm}}$$

Where r_s refers to the leaf stomatal resistance, g_0 is the minimum stomatal conductance, A_n refers to the leaf net photosynthesis, c_s refers to the CO2 partial pressure at the leaf surface, P_{atm} is the atmospheric pressure, D refers to the vapor pressure deficit at the leaf surface, and g_1 refers to the plant functional type dependent parameter.

The maximum rate of carboxylation at 25°C, represented as V_{cmax25} depends on nitrogen foliage concentration and specific leaf area, and its computation is outlined in (). At 25ºC, V_{cmax25} can be expressed as:

$$V_{cmax25} = N_{cb}NUE_{Vcmax25}$$

Here, N_{cb} represents the content of nitrogen required for carboxylation (in grams of nitrogen per square meter of leaf), and $NUE_{Vcmax25}$ = 47.3 x 6.25 represents the nitrogen use efficiency for V_{cmax25}. The constant 47.3 used in $NUE_{Vcmax25}$ computation refers to the specific activity of Rubisco ("act25" with units' micromoles of CO2 per gram of Rubisco per second) measured at 25^0C. In contrast, the 6.25 constant refers to the Rubisco nitrogen binding factor (grams of Rubisco per gram of nitrogen).

6.1.4 Carbon and Nitrogen Allocation

The $f_{root\text{-}leaf}$ parameter denotes the root-to-shoot ratio, which refers to the ratio between below-ground and above-ground biomass in plants. The $f_{root\text{-}leaf}$ parameter influences carbon partitioning between above and below ground and affects the root system size and structure. The

$f_{root\text{-}leaf}$ parameter is a significant determinant of plant growth, productivity, and carbon and nutrient cycling within the ecosystem in CLM5.

$$RAI = (LAI + SAI).r_i.f_{root-leaf}$$

The leafcn parameter is the leaf carbon to leaf nitrogen ratio and determines stoichiometry. Leafcn affects plant growth and physiology and is influenced by factors such as plant functional type, atmospheric CO2 concentration, soil nutrient availability, and water availability. It is important to note that leaf nitrogen is an essential plant growth nutrient often considered a limiting factor in primary productivity.

$$CN_{leaf} = -C{:}N\ for\ leaf$$

All the C: N parameters are typically expressed as constants for a given plant functional type.

The stem_leaf ratio parameter is the plants' stem and leaf biomass ratio. The stem_leaf parameter affects carbon partitioning between above-ground plant tissues and can affect plant growth, resource allocation, and productivity. It is important to note that a high stem_leaf ratio negatively affects vegetation growing in stressful conditions since the cost of building and maintaining leaves might be too high. However, a low stem_leaf ratio favors vegetation growing in resource-rich conditions where leaves are easy to develop and maintain.

The allocation parameter (flivewd) refers to the allocation of new wood biomass between live (phloem and ray parenchyma) and dead (heartwood and sapwood) tissues. The flivewd parameter affects plant carbon and nutrient partitioning, affecting growth, productivity, and carbon storage. A more considerable flivewd parameter value translates to a more significant portion of new wood biomass allocated to live tissues. It can have negative consequences as it benefits plant growth but can lead to high maintenance costs as live tissues require more energy to manage than dead tissues.

$$CN_{lw} = -C{:}N\ for\ live\ wood\ (in\ stem\ and\ coarse\ root)$$

6.1.5 Plant Hydraulics

The Psi_{50} parameter is the water potential at 50% conductance loss. In CLM-5, water from soil to root, root to stem -i.e., water fluxes are modeled according to Darcy's law.

$$q = kA(\psi_1 - \psi_2)$$

Where, q is water fluxes (mmH_2O/s) between the segment $\psi_1\ and\ \psi_2$, k is the hydraulic conductance per second, A is the area basis (m^2/m^2) relating to the conducting area basis to the ground area and $\psi_1 - \psi_2$ is the gradient in water potential (mmH^2O) across the segment.

As water potential decreases, hydraulic conductance also decreases. This is achieved by multiplying the maximum segment conductance by a sigmoidal function, capturing the percent loss of conductivity. This function uses two parameters to fit the experimental vulnerability curves, i.e., the water potential at 50% loss of conductivity ($p50$) and the shape fitting parameter (c_k)

$$k = k_{max}.\, 2^{-(\frac{\psi_1}{p50})^{c_k}}$$

k_{max} is the maximum segment conductance (s^{-1}), P50 is the water potential at 50% loss of conductivity (mmH_2O/s), and ψ_1 is the water potential of the lower segment terminus (mmH_2O/s)

The krmax ($K_{3,\,max}$) parameter, also known as the Root Segment Maximum Conductance, represents the maximum rate of nitrate ions uptake per unit of root biomass and determines nitrates uptake by plant roots. Krmax is influenced by soil properties and plant functional type (pft) and regulates nitrate availability for plant growth and productivity. The "Root Segment Maximum Conductance" denotes the maximum rate at which water is obtained by roots and transported through the plant and is a function of the root system's cross-sectional area and soil and roots' hydraulic properties.

$$k_{r,i} = \frac{k_{3,max}}{z_{3,i}}.\, RAI.\, 2^{-\left(\frac{\psi_{soil,i}}{p50_3}\right)^{c_k}}$$

$$RAI = (LAI + SAI).\, r_i.\, f_{root-leaf}$$

Where $z_{3,i}$ refers to the conducting path of the root tissue, $k_{3,\,max}$ refers to the root segment maximum conductance, r_i refers to the root fraction within the soil layer "i", $k_{soil,\,i}$ refers to the Brooks-Corey soil conductivity within layer "i", LAI refers to the total leaf area, SAI is the stem area index, $f_{root-leaf}$ refers to the root-to-shoot ratio and RAI refers to the root area index. The root fraction r_i is each layer within the soil varies and depends on the plant functional type.

$r_i = (\beta^{z_{h,1-1\cdot 100}} - \beta^{z_{h,i\cdot 100}})$ for $1 \leq i \leq N_{levsoi}$

Where $z_{h,i}$ refers to the soil surface to interface between layers *i* and i+1($z_{h,0}$), and βrefers to the plant-dependent root distribution parameter (rootprof_beta)

6.1.6 Sensible, Latent Heat & Momentum Fluxes

The aerodynamic resistances to heat (i.e., moisture) transfer between ground height and canopy air.

$$r_{ah} = r_{aw} = \frac{1}{C_s U_{av}}$$

$$U_{av} = V_a \sqrt{\frac{1}{r_{am} V_a}} = u_*$$

Where U_{av} refers to the wind velocity incident magnitude on leaves while C_s is the turbulent transfer coefficient between underlying soil and canopy air, which is interpolated between values for dense canopy and bare soil.

$$C_s = C_{s,base} W + C_{s,dense}(1 - W)$$

Where $C_{s,\,dense}$ is the dense canopy turbulent transfer coefficient and $C_{s,\,bare}$ is the bare soil turbulent transfer coefficient. The leaf boundary resistance r_b is a function of the turbulent transfer coefficient between the canopy surface and canopy air (C_v = 0.01 $ms^{-1/2}$), the characteristic dimension of the leaves in wind flow direction (d_{leaf}), and the magnitude of wind velocity incident on leaves (U_{av})

$$r_b = \frac{1}{C_v} \left(\frac{U_{av}}{d_{leaf}}\right)^{-\frac{1}{2}}$$

The Frac_sat_soil_dsl_init parameter refers to the fraction of saturated soil for moisture value at which DSL initiates. The Dry Surface Layer (DSL) refers to the thickness of the dry surface layer (m) and is given as shown below. $DSL = \begin{cases} D_{max} \frac{\theta_{init} - \theta_1}{\theta_{init} - \theta_{air}} & When\ \theta_1 < \theta_{init} \\ 0 & \theta_1 \geq \theta_{init} \end{cases}$

Where D_{max} is a parameter specifying the length scale of the maximum DSL thickness (default = 15mm) θ_{init} (mm^3/mm^3) is the moisture value at which the DSL initiates, θ_1 (mm^3/mm^3) is the moisture value at the top of the model soil layer and θ_{air} (mm^3/mm^3) is the air dry soil moisture value

6.1.7 Stomatal resistance and photosynthesis

The maximum carboxylation rate at 25^0C varies with the foliage nitrogen concentration and specific leaf area.

$$V_{c\,max25} = N_{cb} NUE_{V_{c\,max25}}$$

Where N_{cb} refers to the nitrogen for carboxylation (g N m^{-2} leaf) and $NUE_{V_{c\,max25}} = 47.3\ X\ 6.25)$ and refers to the nitrogen use efficiency for $V_{c\,max25}$. The constant 47.3is the specific Rubisco activity (Act25 μ mol CO_2 g^{-1} Rubisco s^{-1}) and is measured at 25^0C. In contrast, the 6.25 constant refers to the binding factor for Rubisco (g Rubisco g^{-1}N).

6.1.8 Fixation and Uptake of Nitrogen (FUN)

The "FUN_fracfixers" is the parameter 'a' of temperature response of nitrogen fixation and represents the sensitivity of nitrogen fixation to temperature changes. The FUN_fracfixers parameter describes the dependence of nitrogen fixation on temperature and includes several parameters that control the shape and slope of the response curve. The parameter 'a' precisely controls the slope of the response curve, and a higher value of 'a' indicates a steeper response of nitrogen fixation to temperature changes.

$$N_{cost,fix} = -s_{fix} / \left(1.25 e^{a_{fix} + b_{fix} \cdot t_{soil} \left(1 - \frac{0.5 t_{soil}}{c_{fix}}\right)} \right)$$

Where a_{fix}, b_{fix}, and c_{fix} all refers to the temperature response function of the fixation, and t_{soil} refers to the soil temperature.

6.1.9 Biomass heat storage

The "nstem" refers to the stem number or several individual stems per square meter. The nstem parameter determines the level of detail in the representation of plant structure, biomass allocation, and the simulation computational cost. A higher value of nstem calls for a more detailed representation of plant structure and biomass allocation, potentially enhancing model prediction accuracy significantly. However, a higher nstem increases the simulation computational cost. The nstem parameter varies depending on the specific model application and the detail required in the output. It is important to note that individual tree simulations require higher nstem value than only simulating for the entire forest.

6.2 Appendix B: Model Performance Evaluation Metrics Comparison

6.2.1 Comparing composite metric and t-statistic result.

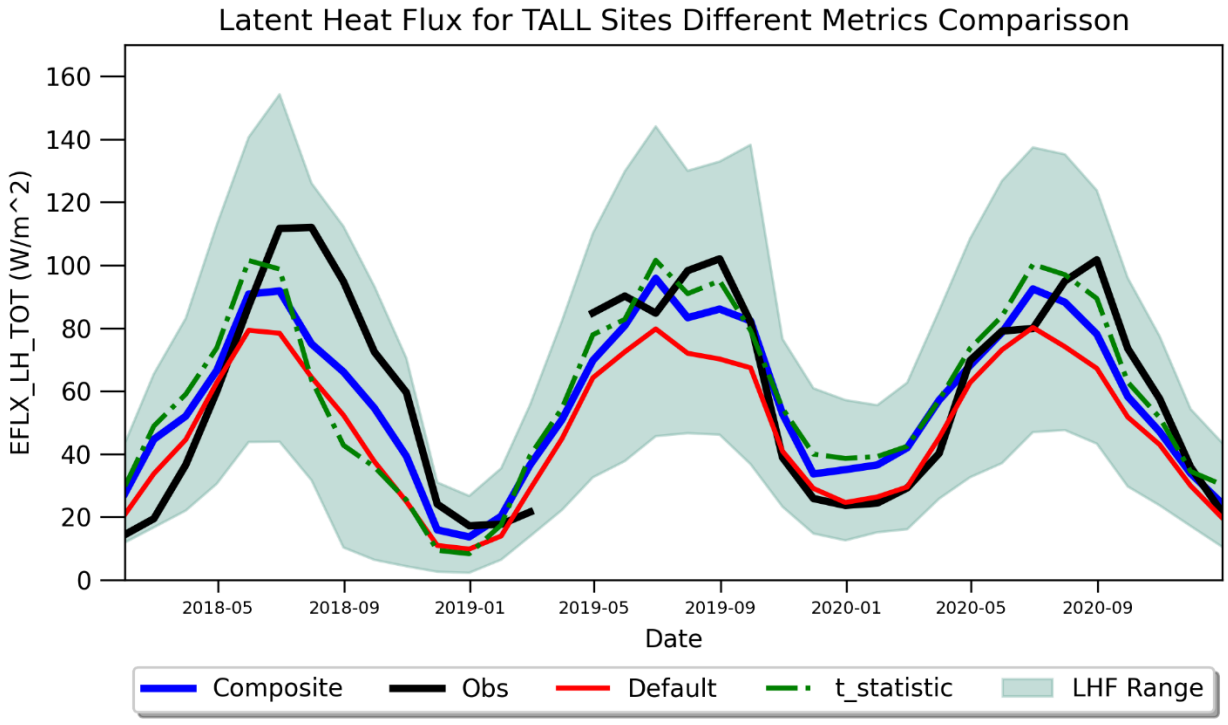

Figure 6.1: Comparing the metrics composite (combines RMSE, MBE, and R^2) and t-statistics combining MBE and RMSE.

The t-statistic metric computed using the proposed method by Jacovides et al. (1995) shows similar results but only for latent heat flux, with our composite result giving even a better constrain with a result closer to observations.

$$t = \left[\frac{(N-1)MBE^2}{RMSE^2 - MBE^2}\right]^{1/2}$$

Here, (N-1) refers to the degrees of freedom where we used daily time steps with 1461 observations, therefore having 1460 degrees of freedom. The KGE method also shows a better constraint for latent heat flux, but when the same model is evaluated for GPP, the result is significantly off (Fig. 6.3) and does not capture observation data.

The Kling-Gupta Efficiency (KGE) is calculated using the following equation:

$$KGE = 1 - \sqrt{((r-1)^2 + (s-1)^2 + (b-1)^2)}$$

Here, r refers to correlation coefficient, s is the ratio of the standard deviation of simulated values to the standard deviation of observed values and b is the ratio of the mean of simulated values to the mean of observed values.

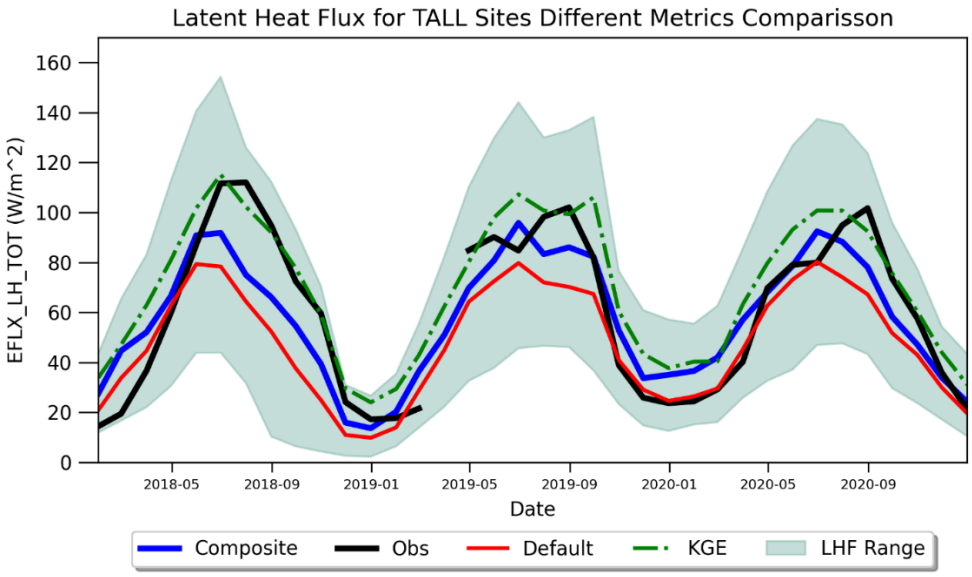

Figure 6.2: Comparing the different metrics composite (combines RMSE, MBE, and R^2) and KGE.

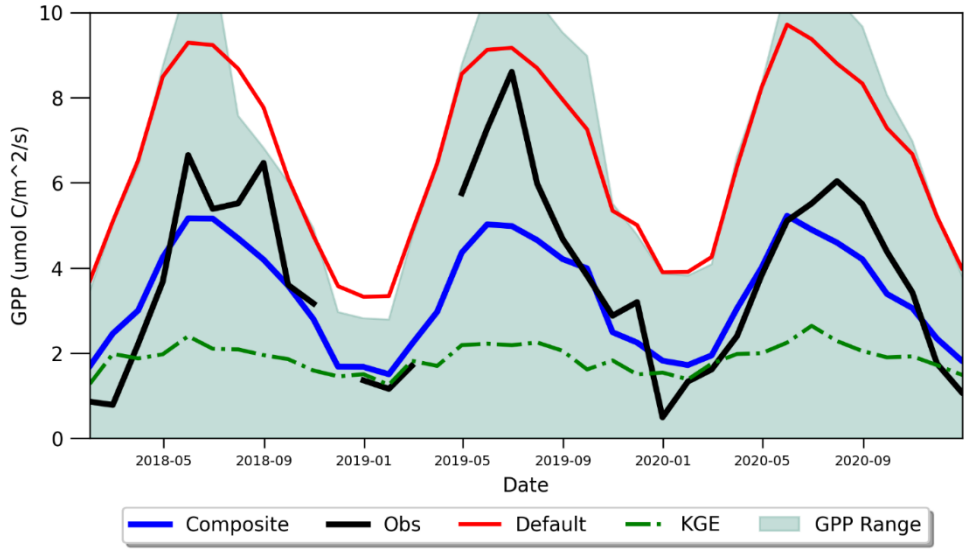

Figure 6.3: Comparing the different metrics composite (combines RMSE, MBE, and R^2) and KGE.

6.3 Appendices C: Supplementary Plots

6.3.1 Dotty Plots

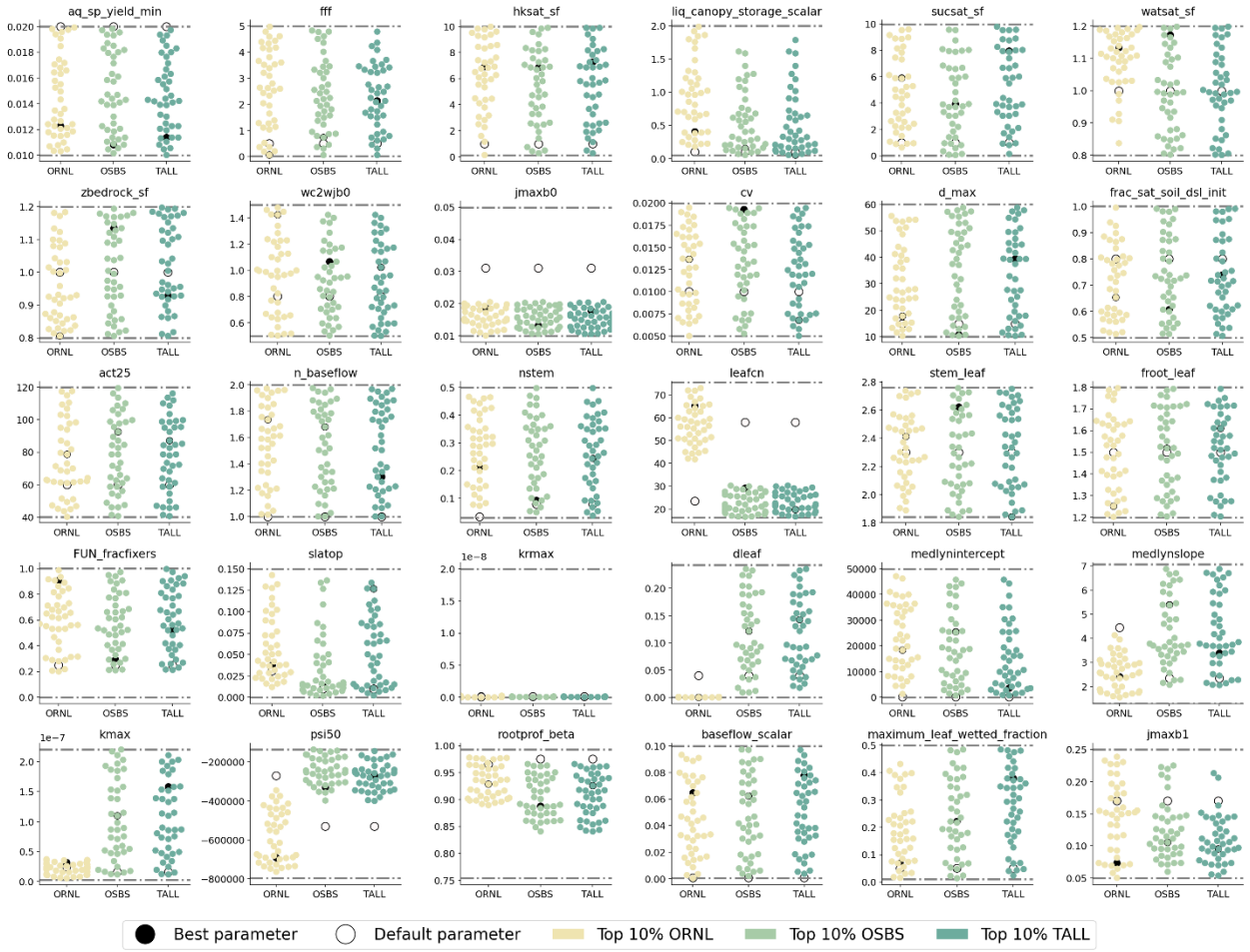

Figure 6.4: Dotty plots for the selected 10% best performing simulations from the perturbed parameter experiments.