Title

TEMulation: A Forest Productivity Emulator of the Terrestrial Ecosystem Model with Extreme Gradient Boosting

Author

Rauschenbach, Sonya

Publication Date

2024

Peer reviewed|Thesis/dissertation

TEMulation: A Forest Productivity Emulator of the Terrestrial Ecosystem Model with Extreme Gradient Boosting

By

SONYA RAUSCHENBACH

THESIS

Submitted in partial satisfaction of the requirements for the degree of

MASTER OF SCIENCE

in

Atmospheric Science

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Erwan Monier, Chair

Kyaw Tha Paw U

Yufang Jin

Committee in Charge

2024

Abstract

Forests play a critical role in the global carbon cycle, and the age of the forest is a key driver of a forest's carbon productivity. However, the environmental conditions the forest experiences growing also drive global variation in global carbon stocks and fluxes. Integrated Assessment Models would benefit from a representation of forest productivity that is impacted by local climate conditions. The Terrestrial Ecosystem Model (TEM) represents the forest carbon cycle response to age and environmental conditions for different forest types, however it is a complex, process-based model and cannot run fast enough to reasonably be part of a coupled, human-Earth model system. Therefore we have developed a statistical emulator of TEM using XGBoost, which can emulate TEM with high skill in a fraction of the run time.

Acknowledgements

I would like to thank my advisors, Dr. Erwan Monier and Dr. Kyaw Tha Paw U for supporting me and guiding me through the masters' thesis process and helping me grow as a researcher in my first years of grad school. I am extremely appreciative of Dr. Ben Felzer, the developer of TEM-hydro, for his patience in teaching me how to run the model and tenacious support in trouble-shooting any issues I encountered. I am grateful for the encouragement and feedback from Dr. Abigail Snyder throughout the project development. I am also very thankful for the valuable suggestions and support from my committee member, Dr. Yufang Jin. I am so grateful for all the students in Hoagland Hall who were there through the ups and downs of this research journey, I wouldn't have been able to do this without all the laughs and the support. I am also very thankful for all the fun times with my labmate Yanyan Lu while working on the TEM simulations. And finally, a huge thank you to all my friends, my parents Kurt and Kristin, my sister Kyra, my partner Pancho and my cat Permian for all of your love - I wouldn't be the person I am today without you all.

This research was supported by the U.S. Department of Energy, Office of Science, as part of research in MultiSector Dynamics, Earth and Environmental System Modeling Program. The views and opinions expressed in this paper are those of the authors alone.

Contents

1 Introduction

Forests cover 31% of the landmass on Earth and play a crucial role in the global carbon cycle. Collectively, they store an estimated 80% of global terrestrial biomass, with 50% in Tropical forests and 30% in Temperate and Boreal forests (Chapin et al., 2011; Unep, 2020). Globally, around 7% of forests are managed (H. Xu et al., 2024) and an estimated 86 million people are formally or informally employed through the forest sector, with the 33 million people directly employed making up 1% of the global population (Unep, 2020). In 2015, the forest sector contributed over 663 billion USD to the global economy through harvesting and logging as well as the production of solid wood products, pulp and paper.

Integrated Assessment Models (IAMs) are critical tools to understand the consequences of climate change policies through simulating complex interactions in the human-Earth system (Calvin et al., 2019). Forests are a key component of this system because of their role in sequestering carbon and their impact on the socio-economic system (Rose, 2014; Ohrel, 2019). When a forest is harvested, there is major disturbance to its carbon cycle. This disturbance has such an impact on the carbon dynamics of forests that the forest stand age is typically assumed to be the primary driver of variation in forest carbon stocks (Chapin et al., 2011; Pregitzer et al., 2004). IAMs often model the quantity of forest carbon stocks by stand age and use these estimations to make harvesting and land allocation decisions within the model (Narayan et al., 2024). However, while stand age is a primary driver of carbon dynamics, climate, atmospheric chemistry and soil structure (environmental) conditions also influence stocks and have the potential to impact stand dynamics as climate change becomes more severe. The global forest cover impacts climate as well, and both climate and prices of forest products are in turn impacted by socio-economic factors that IAMs represent (Calvin et al., 2019). Therefore it is important to account for how climate and atmospheric chemistry might impact the forest carbon response to stand age.

The Terrestrial Ecosystem Model (TEM) is a complex biogeochemistry model that is well-suited to represent complex interactions between stand age and the environment (B. Felzer et al., 2004; B. S. Felzer, Cronin, Melillo, Kicklighter, and Schlosser, 2009; B. S. Felzer, Cronin, Melillo, Kicklighter, Schlosser, and Dangal, 2011; H. Tian et al., 1999; McGuire et al., 1995). TEM is a cohort-based model where the land use history of each grid cell affects the carbon, nitrogen and water pools for that cohort. The advantage of the cohort-based approach is that TEM is able to track the carbon, nitrogen and water pools over time to simulate post-disturbance growth as well as how the growth is impacted by environmental conditions. The model is designed to run at the global or regional scale, which makes it a good candidate to be coupled to an IAM, and it has been applied to analyze the impact of global policy and future climate on forest productivity before (Reilly et al., 2007; Rose, 2014). However, like other process-based Land Surface Models

(LSMs) (Baker et al., 2022), TEM is computationally intensive and cannot feasibly be coupled with an IAM, as one global simulation from 2000 to 2100 requires five days when run on 300 parallel cores. We have developed an emulator of TEM that runs the same simulation in minutes and therefore can be reasonably coupled with an IAM.

Baker et al., 2022 lays out a framework for emulating LSMs, and recommends using Gaussian Process (GP) for emulation. However, because we are emulating the carbon cycle response to stand age as well as environmental variables, our study requires a larger training set than emulating a grid cells' response to the environment alone, because there are multiple forest ages within one grid cell. Additionally, because our goal is to have an emulator that represents how forest growth will respond to future climates, our emulator training data requires additional simulations due to climate uncertainty. Therefore, we build upon the Baker et al., 2022 framework by demonstrating how Extreme Gradient Boosting (XGBoost) can be used to emulate process-based models of the terrestrial carbon cycle. We also apply statistical tools to understand if the emulators are learning physically-sound relationships between emulator input and output. In this paper we discuss our methods, specifically about the processes TEM represents, our global simulation, and the emulation method we employed. We then discuss the results of our emulators, as well as tools we utilize to understand how the emulators are estimating TEM's processes. We apply the emulator to see what the model predicts for the stand age response of our key carbon variables to warmer and wetter climates. Finally, we discuss our results as well as future work.

2 Background

2.1 Forest Productivity Response to Stand Age

The relationship of stand age to the forest carbon cycle has been established primarily from experimental studies of managed forests as well as process-based models and a handful of studies of unmanaged forests. It has been established that forests grow fastest for the first few decades and then growth slows to a steady state, and can even decline in some cases (Chapin et al., 2011; Zaehle et al., 2006; Pregitzer et al., 2004; Gower et al., 1996). The hypothesis behind the decreasing growth rate with age are 1) nutrient supply limitations, 2) a decline in growth after the canopy closes (Ryan, Gower, et al., 1995; Gower et al., 1996) and 3) that as trees age, more carbon goes into structural tissue (e.g. bark, stem and roots), than leaves which are the primary photosynthesis drivers (Ryan, Gower, et al., 1995; Chapin et al., 2011). Many trees do not keep growing taller and producing more leaves, which would increase photosynthesis, because there is increased hydraulic resistance due to tree height (Zaehle et al., 2006; Ryan and Yoder, 1997).

The rate of increase of the growth of young forests depends on available nutrients and the general production measure being considered. Direct growth is measured through Gross Primary Production (GPP) which represents the aggregated ecosystem photosynthesis. The factors that affect GPP are related specifically to photosynthesis and accumulation of photosynthate, while Net Primary Production (NPP) includes autotrophic from vegetation (Spurr, S.H et al., 1973). Net Ecosystem Productivity (NEP) includes heterotrophic respiration from the soil in addition to autotrophic respiration which is included in NPP (WOODWELL et al., 1968; Lovett et al., 2006). While biomass accumulation, and therefore NPP and GPP, can be low initially in the case of primary succession of forests due to a lack of organic matter, forests can start to regrow quickly in the case of secondary succession, which occurs after harvesting or fires and other natural disturbances, providing there are not other nutrient limitations (Chapin et al., 2011). Key sources of variation in carbon accumulation by stand age include the initial carbon and nitrogen stocks of the regrown forest as well as variation in environmental conditions. Traditional ecosystem dynamics predicts GPP will rise to a maximum and then decline slowly with ecosystem age, while NEP will similarly rise to a maximum at a similar time to the GPP maximum, but then decline gradually to approach zero as the respiration increases to match the GPP (Kira et al., 1967; Odum, 1969).

The slope of the forest growth curve as well as the growth peak varies by forest type and biome. A study that looked at managed forests in Europe found that the fastest rate of increase in carbon occurred in the period before 50-80 years (Zaehle et al., 2006). A meta study that analyzed data from managed and un-managed forests for the Temperate, Tropical and Boreal biomes found that NPP started to decline after 11-30 years for Temperate forests and at around 71-120 years for Boreal (Tropical forests did not have enough data for a statistically significant trend)(Pregitzer et al., 2004). For Tropical and Temperate biomes, living biomass continued to increase over time but growth of Boreal forests leveled off after 120-200 years. A simulation that looked at Temperate regions found that for Deciduous trees, NPP increases up to 30 years, whereas for Temperate Coniferous forests, NPP reaches a max at around 20 years, then decreases until 35 years and reaches a lower, but non-zero, equilibrium. The study also found that for both Temperate Deciduous and Coniferous, vegetation carbon increased monotonically with age, though coniferous at a faster pace (Pan et al., 2002). However, this study only used one site per forest type to test the modeled stand age dynamics.

2.2 Statistical Emulation of Process-based Models

Physics-based models that prescribe complex system processes are computationally costly and oftentimes it is unfeasible to run the multitude of simulations that are necessary for policy making. The computational

efficiency of emulators have also made them popular in coupled systems. Emulators are widely used in IAMs for fast coupling of sub-models like agricultural yield (Snyder et al., 2019), hydrology (Liu et al., 2018) and even climate (Tebaldi et al., 2022). These models sometimes replace physical models that are coupled to IAMs ((Snyder et al., 2019; Liu et al., 2018)), and other times are swapped in place of the physical model when a large number of scenarios need to be run (Tebaldi et al., 2022).

The traditional approach to emulation is to assume that the relationship between the input of the process-based model and its output is an unknown function which can be represented as a Gaussian Process (GP) (Sacks et al., 1989; Kennedy et al., 2001). GP is a stochastic process that, given a specific mean and covariance function, gives a probability distribution over all functions that fit the input to the output, estimated using Bayesian inference. GP has been used to emulate process-based models, for ex. (Baker et al., 2022) (Li et al., 2019; O'Hagan, 2006; Kennedy et al., 2001) and its strength the method's robust uncertainty estimation. A similar approach to emulation specifies the basis functions that can explain the relationship between input and output variables and similarly uses Bayesian inference to solve for the optimal parameter set to fit to these functions (Snyder et al., 2019; Chakraborty et al., 2017). These types of statistical emulators are often used in calibration and tuning of the physical model or to explore and identify areas of improvement from the physical model, and so do not require large numbers of simulations for training (Baker et al., 2022; Kennedy et al., 2001; Marmin et al., 2022; Sacks et al., 1989). Therefore, for these applications, it is not a limitation that the computational time of the GP models scales with $O(n^3)$ (Rasmussen et al., 2006). While there are techniques to make this method more computationally efficient ((Li et al., 2019; Baker et al., 2022)), for our simulation training set, which includes all stand age groups for each grid cell, this approach would require sub-sampling on a very small percent of the available training data.

Other approaches to statistical emulation include conventional statistical learning and deep learning (Sudakow et al., 2022). There are different benefits to each approach. Decision tree regression, which falls under conventional statistical learning techniques, have the benefit that the predicted variable will stay within the range of values it was trained on (Limon et al., 2023; Yuval et al., 2020). This makes it less likely that the emulator will predict non-physical quantities, but limits the validity of emulator prediction to the input training data ranges, as opposed to the traditional GP and basis function approach, as well as deep learning techniques which can generalize the response function outside of the training ranges (Limon et al., 2023; Yuval et al., 2020; Snyder et al., 2019). Deep learning methods allow for reliable prediction when the input is high dimensional, whereas conventional statistical methods are more prone to the "curse of dimensionality", where high dimensional input can lead to optimal models that have spurious input-output relationships and so do not generalize well. A drawback of deep learning techniques is that they are known

for having a more difficult time constraining the output to physical quantities (Limon et al., 2023). For our application, we do not have high dimensional input and we are prioritizing emulation approaches whose predictions remain within physically feasible ranges. Gradient Boosting Machines (GBMs), implemented with XGBoost, are the conventional statistical learning approach we have chosen because they represent TEM with high model fidelity and XGBoost is designed to be run on very large training datasets(Friedman, 2001; Chen et al., 2016).

3 Methods

3.1 The Terrestrial Ecosystem Model

We are emulating TEM, a process-based biogeochemical model, specifically the version TEM-hydro, which represents carbon-nitrogen-water fluxes in terrestrial ecosystems between the soil, detritus, roots, stems, leaves, seeds and atmosphere (B. Felzer et al., 2004; B. S. Felzer, Cronin, Melillo, Kicklighter, and Schlosser, 2009; B. S. Felzer, Cronin, Melillo, Kicklighter, Schlosser, and Dangal, 2011; B. S. Felzer, 2023). The processes that drive these fluxes include photosynthesis, respiration, nitrogen mineralization and uptake, litterfall, and plant growth (allocation) for the different plant tissue pools. TEM is a monthly model typically run at 0.5 x 0.5 degree spatial resolution over land. Each grid cell is run independently and so is no interaction between grid cells. TEM uses an adaptive Runge-Kutta integration process to solve it's differential equations over time (B. S. Felzer, Cronin, Melillo, Kicklighter, and Schlosser, 2009). See Figure 2 for a schematic of TEM-hydro.

TEM takes a cohort-based approach to land use and land cover change (i.e. disturbance). After a disturbance, a new cohort is formed from one of the existing cohorts within the grid cell. Disturbances are assumed to occur annually and in the first month of the year. At the time of the disturbance, the carbon and nitrogen stocks within the soil are equal to the carbon and nitrogen pools for the prior cohort, plus the detritus left over from the harvest. After harvest, the vegetation carbon stock is reset, and initially there is little leaf area (parameterized through Leaf Area Index - LAI) for photosynthesis. Once there is enough carbon allocated to the leaves, there is a nonlinear increase in vegetation carbon stocks and carbon assimilation fluxes, though the rate of the increase is limited by the available nutrients (water, carbon and nitrogen) for that cohort plus what is input from the current time step's precipitation and nitrogen deposition. Within this framework, stand age is equivalent to the age of the cohort. The growth of vegetation is not prescribed as a function of stand age, but rather related to LAI and available nutrients, as well as the impact of climate and atmospheric chemistry forcing (H. Tian et al., 1999). Therefore the impact of stand age is not dictated

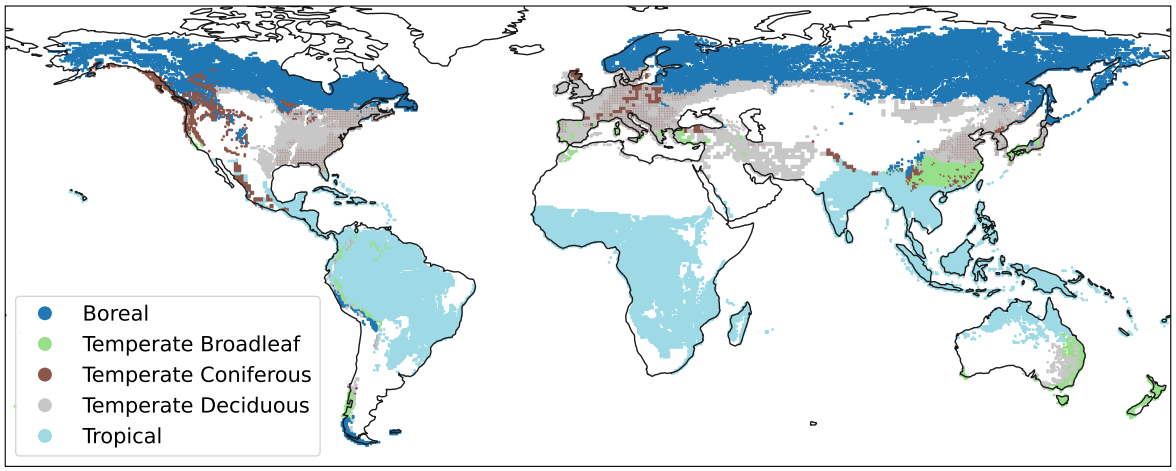

Figure 1: This map shows the forest PFT locations used in TEM.

Parameter	Description	Tropical	Temperate Broad.	Temperate Conif.	Temperate Decid.	Boreal
Tmin	(°C) where GPP becomes 0	10	5	1	0	-5
Tmax	(°C) where GPP becomes 0	37	35	38	40	32
$Tmin_{opt}$	min (°C) where GPP can peak	22	20	22	23	17
$Tmax_{opt}$	max (°C) where GPP can peak	33	32	30	31	25
rootz	rooting depth	8	3	2	2	2
kc	half saturation constant of CO_2	400	200	200	200	200

Table 1: This table shows key parameters in the functional response of GPP to climate.

by the model but is instead simulated through the process-based equations. This makes TEM a suitable model to study the complex relationship of how carbon fluxes and stocks respond to the environment and stand age.

TEM uses Plant Functional Types (PFTs) to represent different ecosystem types. PFTs group plant species that respond similarly to the environment and available resources (Duckworth et al., 2000). We ran TEM with 5 unique sets of parameters for its forest PFTs, Temperate Coniferous, Temperate Deciduous, Temperate Broadleaf Evergreen, Tropical and Boreal. Figure 1 shows where on the globe these PFTs are located, according to TEM's potential vegetation (McGuire et al., 1995). In total there are around 38,000 grid cells where the potential vegetation is forested. There is no competition between PFTs in TEM, and each grid cell only has one potential vegetation PFT, which can then be harvested and regrown or converted to crop, pasture or urban land use. Each PFT has a unique set of parameters that impact TEM's GPP equation and influences how the carbon fluxes and stocks respond to climate and atmospheric chemistry forcing. See B. S. Felzer, Cronin, Melillo, Kicklighter, and Schlosser, 2009 and B. S. Felzer, Cronin, Melillo, Kicklighter, Schlosser, and Dangal, 2011 for a description of PFT dependent parameters and their source. Table 1 lists the key parameters that will be discussed in this study.

3.1.1 Functional Forms of Response of TEM Emulation Variables to TEM Climate Forcing

TEM's equation for GPP factors in the impact from the environment (the climate, atmospheric chemistry, soil properties and elevation) directly. GPP is a core variable because it represents the gross photosynthesis, from which NPP is calculated after subtracting maintenance and growth (autotrophic) respiration. The net increase in carbon, from NPP, is partitioned into different parts of the plant structure based on a marginal benefit calculation (B. S. Felzer, Cronin, Melillo, Kicklighter, Schlosser, and Dangal, 2011). The total of carbon in the different plant structures is equal to the vegetation carbon. Raich et al., 1991 showed the response of GPP to temperature, atmospheric CO_2, and downwelling shortwave radiation (rsds) to a prior version of TEM and demonstrated the hyperbolic relationship CO_2 and rsds have to GPP, and the parabolic relationship temperature has to GPP (Raich et al., 1991). TEM-hydro utilizes similar functional forms, but has modified GPP calculations to represent leaf level photosynthesis which is then aggregated to canopy level, as opposed to calculating canopy level GPP. Precipitation impacts GPP through its direct relationship to soil moisture, which is assumed to have a hyperbolic relationship to GPP, where after a level of soil moisture, additional moisture no longer has an incremental positive benefit. These equations are shown in Appendix A.

While climate variables directly impact GPP through the prescribed non-linear functional forms, there are many complex interactions and indirect effects. For example, while temperature impacts GPP according to the relationship described in f_T, temperature will also impact f_{H_2O} through its impact on evapotranspiration, $f_{C_a,D}$ through its impact on vapor pressure deficit (VPD), and will also have a feedback affect on $\int_0^{LAI} f_{PAR} dL$, as optimal temperatures allow for more leaf growth and therefore a higher LAI. Precipitation rates, and its impact on moisture stress, will also impact LAI through its impact on GPP, and so will have a feedback effect. The effects of CO_2 and precipitation on GPP interact, because CO_2 has a larger positive benefit on GPP when there is moisture stress (Raich et al., 1991). Furthermore, GPP is limited by nitrogen uptake, which is affected by temperature in the same relationship as maintenance respiration.

3.2 Simulation Set Up

In order for the emulator to learn TEM's response to a wide range of plausible climate scenarios, we run a global TEM simulation on historical and future climate input. The Climate Model Intercomparison Project (CMIP6) represents the state of the art of climate modeling at the time of conducting this research (Eyring et al., 2016). We train the emulator on realistic climate scenarios, as opposed to simulated ranges of the required TEM climate input, because climate models have already factored in the complex physical relationships between climate variables. We emulate GPP, NPP, which are gross and net annual rates of

carbon assimilation, as well as vegetation and soil carbon stocks, VEGC and SOILORGC, which together represent the total biomass of these ecosystems. While many CMIP6 models also output these core ecosystem variables, they do not include forest age and so cannot be used for our application.

We train the emulator on transient climate and atmospheric chemistry input (see Table 2) as well as land use and land cover change, in order to simulate the most realistic environment TEM can represent. For historical climate forcing, we follow the TRENDY v11 protocol, which provides the land carbon cycle components for the Global Carbon Budget (Friedlingstein et al., 2023). The climate data are primarily from the Climate Research Unit (CRU) Time Series 4.05 (Harris et al., 2020), with the exception of surface downwelling shortwave radiation flux, where we use the data product from O'Sullivan et al., 2021, which better accounts for aerosols' impact on downwelling radiation. For future climate, we use the CMIP6 model CanESM5-1, specifically the Shared Socioeconomic Pathway (SSP) scenario 245 from initial conditions r1i1p1f1 (Sigmond et al., 2023). We bias correct the future climate data using the Quantile Delta Mapping method in order to preserve variation within the biased corrected data. This method is described in Cannon et al., 2015. We use a rolling 30-year window over which to calculate the quantiles, and based the bias correction on the overlapping 1984-2014 time period between CRU and the historical CanESM5-1 simulation. Additionally we downscaled the climate data to TEM's 0.5° x 0.5° resolution using a nearest neighbor approach (Gangopadhyay et al., 2005).

For the atmospheric chemistry inputs, we also use the TRENDY provided gridded datasets for nitrogen deposition (Lamarque et al., 2013) and the nitrogen fertilization on crop and pastureland, from Hanqin Tian et al., 2018. We apply the nitrogen deposition to every grid cell, but only apply the nitrogen fertilization where there was active cropland or pastureland in a given cohort. We keep future fertilization steady from 2018 - 2100 for all scenarios as we are not aware of any dataset with nitrogen fertilization projects for crop or pasture lands. Our ozone and future nitrogen deposition data are calculated from the multi-model mean of the Atmospheric Chemistry Modeling Improvement Project (ACCMIP) following the methodology in Lamarque et al., 2013, which linearly interpolates values between missing years. We then calculate AOT40, , following the calculation described in B. Felzer et al., 2004, which is most relevant for plant life because it only looks at the ozone levels during the daytime periods when stomata are open. At the time of running the ensemble, ACCMIP was last run on CMIP5. Therefore there is a scenario mismatch with CMIP6. We use the future scenario with the closest greenhouse gas forcing, so SSP245 is matched with RCP45. Historical and future CO_2 are from the historical and SSP245 forcing input for CMIP6.

We use the Land Use Harmonization (LUH2) database for historical and future land use land cover change (Hurtt et al., 2020). This database describes transitions for the percent of land harvested or converted to crop, pasture, and urban land use. LUH2 has different scenarios for each SSP, which are the land use land

Input Variable	Shorthand	Units	Emulator Input
Diurnal Temperature Range	trange	°C	X
Near-surface Wind Speed at 10m	wind	m s-1	X
Vapor Pressure	vpr	hPa	X
Near Surface Air Temperature	tair	°C	X
Precipitation Flux	precip	mm month-1	X
Surface Downwelling Shortwave Flux	rsds	W m-2	X
Percent of Sand	s1	%	X
Percent of Silt	s2	%	
Percent of Clay	s3	%	X
Carbon Dioxide	CO_2	ppm	X
Accumulated Ozone	AOT40	ppb	
Nitrogen Deposition	Ndep	mgN m-2 yr-1	

Table 2: TEM input forcing and emulator input variables.

cover input for the CMIP6. We use the historical scenario and the SSP245 scenario for our two time periods. LUH2 does not have a land cover map with detail about vegetation, therefore we use the primary vegetation map from (McGuire et al., 1995). TEM is a cohort-based model, so we convert the transitions to cohort land area. LUH2 also includes land states for given years as well as their transition. Our ensemble begins in 1700, so we take the land states from LUH2, which are primary, secondary, urban, pasture or crop, and apply this to our primary vegetation map and treat each separate land type as a distinct cohort. So if in 1700, at the beginning of our simulation, LUH2 has 5% of the land as crop, 5% secondary and 90% primary, the grid cell would have three cohorts. In 1701, the transitions from LUH2 are applied. So, if in 1701 1% of the grid cell was transitioned from primary to secondary, indicating that the forest was harvested, there would be four cohorts and the two cohorts of secondary land would have different stand ages.

In order to ensure that each TEM grid cell was in a steady equilibrium, we run a 200 year spin up period after the initial equilibration. For the spin up, we repeat the last 20 years available for each forcing dataset. For climate, this is from 1900-1920 for atmospheric chemistry, this is from 1850-1870. For the period 1900 to 2014, we use the historical forcing for all datasets, and then for 2015-2100 we run the future simulations according to the methodology described above. TEM also has static datasets that provide soil texture and elevation for each grid cell, which are described in (B. S. Felzer, 2023). Table 2 includes the TEM input variables, as well as a flag for if these variables were included in the emulator input.

3.2.1 Sources of Uncertainty from Simulation Set Up

There are various sources of uncertainty from our simulation, some of which we focus on accounting for, others we leave open for future improvements. This study is focused on the uncertainty in emulator performance (i.e. how well the emulator reproduces TEM output). We account for this through cross validation of our emulator, using 50 different train test splits, as well as comparing model performance when trained and tested

on different time periods. However, this study does not look at the impact of different initial conditions as our future climate data is from a single initial condition. Additionally, there are sources of uncertainty from the bias correction of CanESM5-1, as well as with the observational datasets we used. Furthermore, we were only interested in emulating one version of TEM, and therefore there is uncertainty related to the sets of parameters we are using with the five forest PFTs.

3.3 Emulation Input-Output Structure and Temporal and Spatial Scale

IAMs typically run at wide time steps and spatial scales, for example the Global Change Analysis Model operates on 5-year time steps and breaks the globe into 256 land areas (Calvin et al., 2019). TEM is a monthly model and has over 60,000 0.5° x 0.5° grid cells over land. We have developed the emulator at an annual time scale, to smooth out the seasonal influence of the climate input that IAMs do not capture and instead focus on capturing inter-annual variability. With an annual time scale, the emulator output could be aggregated into as many years as required to couple with an IAM. We do not aggregate the native 0.5° x 0.5° spatial resolution of TEM when building the emulator, in order to train the emulator on as wide a range of climate scenarios as possible.

When describing their methods of emulating land surface models (LSMs), the authors of Baker et al., 2022 assumed temporal and spatial independence for their emulator. In TEM, each grid cell is run separately, and so it is reasonable to assume spatial independence for our emulator, and that any spatial dependence will be captured by the emulator input. However, the assumption of temporal independence is less reasonable given that the carbon cycle output of TEM, as well as other LSMs, depends on the amount of carbon accumulated in past time steps, which in turn depends on past climate output. Therefore the assumption of temporal independence is a simplification to the emulation structure (Baker et al., 2022). While there are emulators that can include time series structures (Mohammadi et al., 2019), for our purposes this simplification is necessary as most IAMs do not store information from past time steps to conserve memory. Therefore in our emulator we only include input from the current time step in order to facilitate IAM coupling.

We are interested in emulating TEM output for GPP, NPP, VEGC and SOILORGC. These variables are key components of the carbon cycle that indicate forest health and carbon storage. GPP is the total CO_2 assimilation, excluding photorespiration, modeled in $gCm^{-2}month^{-1}$. NPP is a measure of CO_2 assimilation into plant organic matter, and so is GPP minus maintenance and growth respiration, also modeled in $gCm^{-2}month^{-1}$. VEGC is modeled in gCm^{-2} and includes all plant organic carbon, which in TEM is partitioned into roots, stem tissue and leaves. SOILORGC is the total of all the organic carbon in the soil, modeled in gCm^{-2}. We convert these monthly TEM output variables into annual output by summing

the monthly GPP and NPP, so that the units are now $gCm^{-2}year^{-1}$, and taking the annual average of VEGC and SOILORGC. We expect stocks to be easier to predict than fluxes, since the fluxes are a change in stocks and are more variable. See Figure 2 for a schematic of the TEM simulation and emulator set-up.

For the climate input to the emulator, we use the average, minimum and maximum monthly values of all annual climate variables, in order to preserve the climate signal from the extremes of the monthly values that might impact the annual value. We convert vapor pressure to vapor pressure deficit using the average monthly air temperature. We do this to make the response easier to interpret than vapor pressure alone, because the same vapor pressure value can have a very different impact on vegetation depending on the ambient temperature. Because this relationship between vapor pressure and VPD is well-established and easy to calculate with vapor pressure and temperature, another input variable, there was no need for the emulator to learn it. We recognize that the non-linearity of the saturation vapor curve with respect to temperature causes longer-term averages of VPD based shorter-term values to deviate from VPD calculated from longer-term averages of air temperature and vapor pressure, but for practical purposes we ignore this. It should be noted that TEM includes assumed diurnal patterns that partially mitigate the effects of the non-linearity.

In TEM there is a new cohort each time there is a land disturbance, so an individual grid cell can contain thousands of cohorts and often contains hundreds. In order to couple with IAMs, we want the disturbance variable to be the forest stand age. We parameterize the cohorts in TEM by using five year stand age bins for the first 100 years, and then including a bin for 100-125 years and 125-150 years and 150+ years, for a total of 23 stand age bins. In order to create a more continuous input variable, we convert the stand age bins to ordinal values ranging from 0-22. We drop all cohorts that were primary vegetation and have never been disturbed, so our maximum stand age is 300 years, which represents a cohort that is disturbed in the first year of the simulation. We also include soil texture, specifically the percent of clay and sand as static predictors.

To facilitate coupling with IAMs, we only included the CO_2 atmospheric forcing data, as nitrogen deposition and AOT40 are not typically modeled in IAMs. Because we are only using TEM input from the current time step, we could not include the nitrogen fertilizer amount to agricultural land and pasture land because we are emulating forested cohorts, and fertilization is only applied to crop and pasture. However, fertilization impacts the nitrogen stocks when the forest was converted from crop or pasture land, and so is an additional unaccounted for source of variability along with AOT40 and nitrogen deposition.

3.4 Emulation Statistical Methods

We use a statistical model that is efficient for very large datasets, due to the size of the TEM simulations. Because our input data is relatively low dimensional (22 input features versus over 50 million samples), we implement Gradient Boosting Machines (GBMs) (Friedman, 2001). GBMs are a non parametric decision tree method consisting of piecewise constant functions. See Figure 3 for a stylized example of these piecewise constant functions. We also test a method that approximates piecewise linear functions, Multivariate Adaptive Regression Splines, however GBMs perform significantly better for all PFTs and variables (see Appendix B).

GBMs are an ensemble tree boosting machine learning method that learn from the additive output of many weak models. Each new tree learns from minimizing the loss of the residuals from the previous trees. We minimize the loss with respect to root mean squared error (RMSE). The weak models are added together with weights, so better performing trees have higher weights, and the weights of worse performing trees are shrunk (Friedman, 2001). GBMs are a regularized model with a penalty for more complex regression tree functions to avoid overfitting. We implement GBMs with the eXtreme Gradient Boosting algorithm (XGBoost) (Chen et al., 2016). XGBoost is able to scale to large datasets because it has an algorithm to approximate the split function to find the best tree based on different quantiles of the input features. Additionally XGBoost is the only sparsity aware gradient boosting machine, which has advantages in coming to an optimal solution given the concentration of zeros in input variables like VPD min and precipitation min.

Beyond the statistical metrics that show how the emulated variables compare to the simulated variables, we are interested in whether the emulators have learned the physically relevant relationships between the climate input and stand age and the output variables. As GBMs grow in complexity, they lose interpretability and we must rely on additional tools to interpret the model, for example variable importance and partial dependence plots (PDPs) recommended in Friedman, 2001. We check the soundness of our XGBoost emulators by looking at the relative importance of input variables on the learned function. We estimate variable importance (VI) by the method described in Friedman, 2001, which approximates VI to be the sum of the squared error improvement as a result of the split from each variable, averaged, with weight, over all trees. Another method that Friedman, 2001 recommended to gain clarity on the GBMs functions is to plot the partial dependence of the input variables on the output. The partial dependence shows the relationship between each input variable to the output, holding all else constant, which is akin to estimating the partial differential equation. For the variable importance, we are looking to see if the identified variables for each emulator are in line with what we would expect given the TEM equations and each PFTs parameterized climate limits. For the partial dependence plots, we are testing if the functional form of the input variables

match what TEM's equations imply for GPP. We will only analyze the partial dependence plots for GPP, where the equations for the climate input to GPP are easiest to interpret (see Appendix A).

Along with the Partial Dependence Plots (PDP), we plot the Individual Conditional Expectation (ICE), which show the effect of changing one input variable for each individual GPP sample, holding all other input variables constant (Goldstein et al., 2015). As is typical with ICE plots, we center the response so that for each observation, the impact of the minimum of the relevant input variable (the first value on the x axis) on GPP is zero. This facilitates comparisons of the impacts for observations that have a wide range of GPP values. A PDP is the average of all ICE, however plotting the ICE shows the variation of the average response. We show ICE plots for a subset of 1,200 runs, where we randomly sampled 200 GPP values from the 0-5th percentile, 5th - 25th percentile, 25th-50th percentile, 50th to 75th percentile, 75th to 95th percentile, and then 95th to 100th percentile of GPP for each PFT. We divide the quantiles unevenly to highlight the responses of outlier GPP values below the 5th and above the 95th percentile.

While XGBoost allows for multi-model output, the GBM method does not allow for the emulator to learn from the relationship between the multi-output (i.e. the relationship between NPP and GPP or VEGC and SOILORGC). As there is no benefit to the emulator performance for a multi-model output, we fit a separate emulator for each variable. Additionally, because each PFT has a different parameter set, we train a different emulator for each PFT. Figure 2 shows a schematic of the TEM simulation and how it relates to the emulators' construction. Figure 3 shows a stylized example of the XGBoost emulators.

3.4.1 Parameter Tuning

When tuning the emulator, we focus on three key parameters: the learning rate, which shrinks the weights on model features to prevent overfitting, the maximum depth, which determines the maximum decision levels each weak tree learner can have, and the minimum child weight (minchildweight), which reduces node split complexity. We optimize the model based on the root mean squared error. In order to avoid overfitting we choose a parameter set where the training performance is similar to the testing performance. We find the best combination of parameters to be a learning rate of 0.1, a depth of 14 and a minchildweight of 100. This proves to be an optimal combination for all emulated variables. While the difference between the training and testing R2 is 0.02, there is considerable improvement in overall R2 (0.88 versus 0.76) and we feel that improvement is enough to warrant a slight difference in training versus testing. Note that while more complex models with higher maximum depths and lower minchildweights improves performance in the testing set, they do not give enough of a performance boost to warrant adding model complexity. We test other parameters to reduce overfitting, including alpha, which controls the L1 penalization, but they have little impact on performance, only increasing it by 0.1-0.2%. See Appendix C for comparisons of the

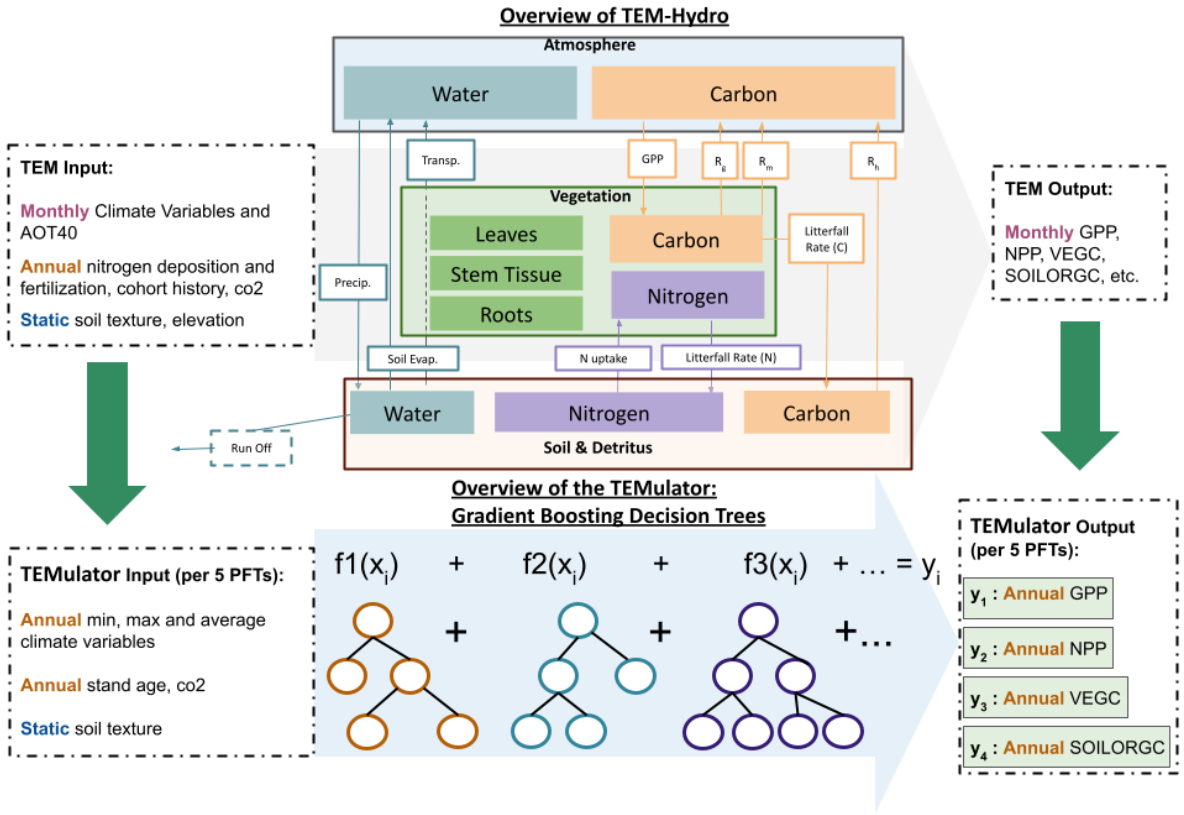

Figure 2: This figure shows how the input and output of TEM relate to the emulator's (TEMulator's) construction. There are a total set of 20 emulators, with one emulator per each output variable and each PFT. The emulator runs at an annual time scale, whereas TEM is a monthly model. Not all of the TEM input variables were included in the emulator in order for the emulator to be easily coupled with IAMs. The schematic of TEM-Hydro is adapted from B. S. Felzer, Cronin, Melillo, Kicklighter, Schlosser, and Dangal, 2011.

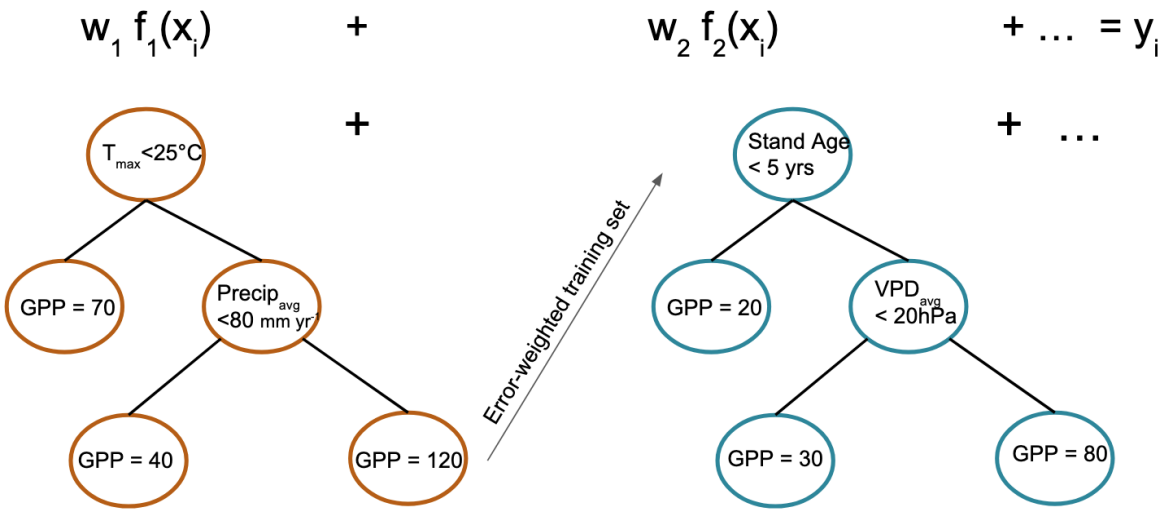

Figure 3: This figure shows a stylized GBM example, where w_j is the weight of each decision tree, f_j is the decision tree function and y_j is the predicted output variable. GPP is in units gC/m2/month. The arrow for the error-weighted training set represents how each additional tree learns from the residuals of the prior decision tree by giving a higher weight to training data with a higher residual, in order to fix the errors in the next tree.

performance and difference between training and testing for different parameter sets. We show the results for all variables of Temperate Coniferous PFT, which had overall the worst performance of all PFTs.

3.5 Criteria for Model Assessment

We compare the tuned best models for their skill in emulating TEM based on the following criteria:

1. How well does the emulator represent TEM, as quantified with the statistical metrics of R^2, RMSE, and bias?

2. How do the emulated 25th, 50th and 75th percentiles, and standard deviation compare to the simulated metrics?

3. How well does the emulator capture the response of the output variables to stand age, as quantified by comparing the 25th, 50th and 75th percentiles of the response per stand age bin?

4. How well do the emulators capture the functional form of the underlying TEM equations for GPP?

3.6 Training Data Comparison

As explained above, our global TEM simulation provides an abundance of training data for the emulator. We want to understand the incremental benefit to performance metrics by increasing the amount of data in the training set. Therefore, we analyze the emulator performance with a 75/25, 50/25, and 25/25 train test split, where the testing data was the same but the training data increased. We use Monte Carlo cross validation

techniques in order to build a confidence interval for our performance metrics (Q.-S. Xu et al., 2001). We randomly subsample the simulation dataset 50 times to get a distribution of the model performance metrics. For each of the 50 random subsamples, the 25% testing data is kept the same, so that pseudo-random differences in sampling do not impact the comparison of the percent of training data included.

Additionally, we test the emulator performance for different time periods. Throughout the 2015 to 2100 time period, many grid cells experience a drastic change in climate, and we test how well the emulator is able to simulate TEM with and without including this time period in the training dataset. We conduct three tests: 1) we train the emulators on the entire time period of 1900-2100, with a 75/25 percent train - test split (47 million samples), 2) we train the emulators only on historical data from 1900 - 2014 (22 million samples) and test on the entire future training set (25 million samples), 3) we train and test the emulators only on future data, similarly with a 75/25 train-test split (25 million samples). We again use Monte Carlo cross validation and randomly sub-sampled each test 50 times.

We also analyzed the impact of adding, cumulatively, 20 years of training data, starting at 1900 - 1920 until we get to the full time period of 1900-2100 to understand how additional decades of climate trends impact the predictive power of the emulator. For this experiment, we hold out 15% of the data for the entire time period for testing, and then we include in the training dataset the cumulative time period between 1900 and the 20 year period. The first period is 1900 - 1920, the second 1900-1940 until we had the entire 85% of the simulated data for the time period between 1900-2100. These results will be discussed in Appendix D as the conclusions are not thematically different from the time period experiment.

4 Results

4.1 Emulation Performance Among Train-Test Splits

The results of our cross validation train-test splits show little uncertainty in performance metrics for different subsamples, but a sizable difference in model performance between the 25-75% of training data. Figure 4 shows the performance metrics for all emulators for each of the 50 subsamples. The colors distinguish between the training size inclusion tests, the scatter plot shows the result for each cross validation run, and the box plot shows the distribution of the performance metrics across the cross validation subsamples for the 25% of testing data. Note that there is little variation in performance across the subsamples, which gives high confidence in the performance metrics, in comparison to the variation between the different training data inclusion percentages. For the three different tests, the incremental increase in performance between including 25% versus 50% of the training data is larger than the increase between 50% and 75%, indicating

that the incremental benefit is beginning to level off.

Overall, the performance metrics demonstrate that the emulator represents TEM simulations well. As shown in Figure 4, the R^2 is highest for the Temperate Broadleaf PFT for most of the carbon variables, with the exception of VEGC, where Temperate Deciduous is highest performing. The worst performing PFTs were Boreal and Temperate Coniferous forests. While Temperate Broadleaf has the highest R^2, that PFT has the largest range of bias across the subsample iterations, whereas Boreal has the lowest. The emulators capture well the 50th, 25th and 75th percentiles of all carbon output variables and again there is little variation among the cross validation. For GPP, NPP and VEGC the bias in the 50th percentile is positive, whereas for SOILORGC for the 50th percentile the bias is primarily negative with the exception of temp_c. At the 25th percentile, the bias is positive for all variables, and at the 75th percentile the bias is primarily negative, indicating that the emulator typically predicts closer to the center of the distribution. This can also be seen in the comparison on kernel density estimation (KDE) for the simulated versus emulated variables shown in Appendix E.

4.2 Emulation Performance for Different Training Time Periods

In general across all PFTs and output variables, training the emulators on the whole time period has comparable performance to training the emulator on future data, which indicates that the emulator is able to learn functions across time periods. The emulator that is trained on the entire time period generally outperformed the future, which is expected as the entire time period has a larger training dataset than the future only. This is in line with the results in the prior section, where larger training sizes increased the performance.

As shown in Figure 5, the emulator trained on historical data and tested on future data does not perform well as the emulator that has never learned the response to future climates which deviate from historical norms. This shows the importance of including future climate in statistical models that are meant to predict future periods. This also demonstrates why training statistical models on present day climate can miss important nonlinear effects and tipping points. For all variables, the bias is negative indicating that the emulator trained only in the present day predicted that the carbon cycling would accumulate more carbon than TEM simulated. This negative bias is carried through in all quantiles. Appendix D shows the impact of cumulatively adding 20 decades to the training dataset, and the conclusions are similar.

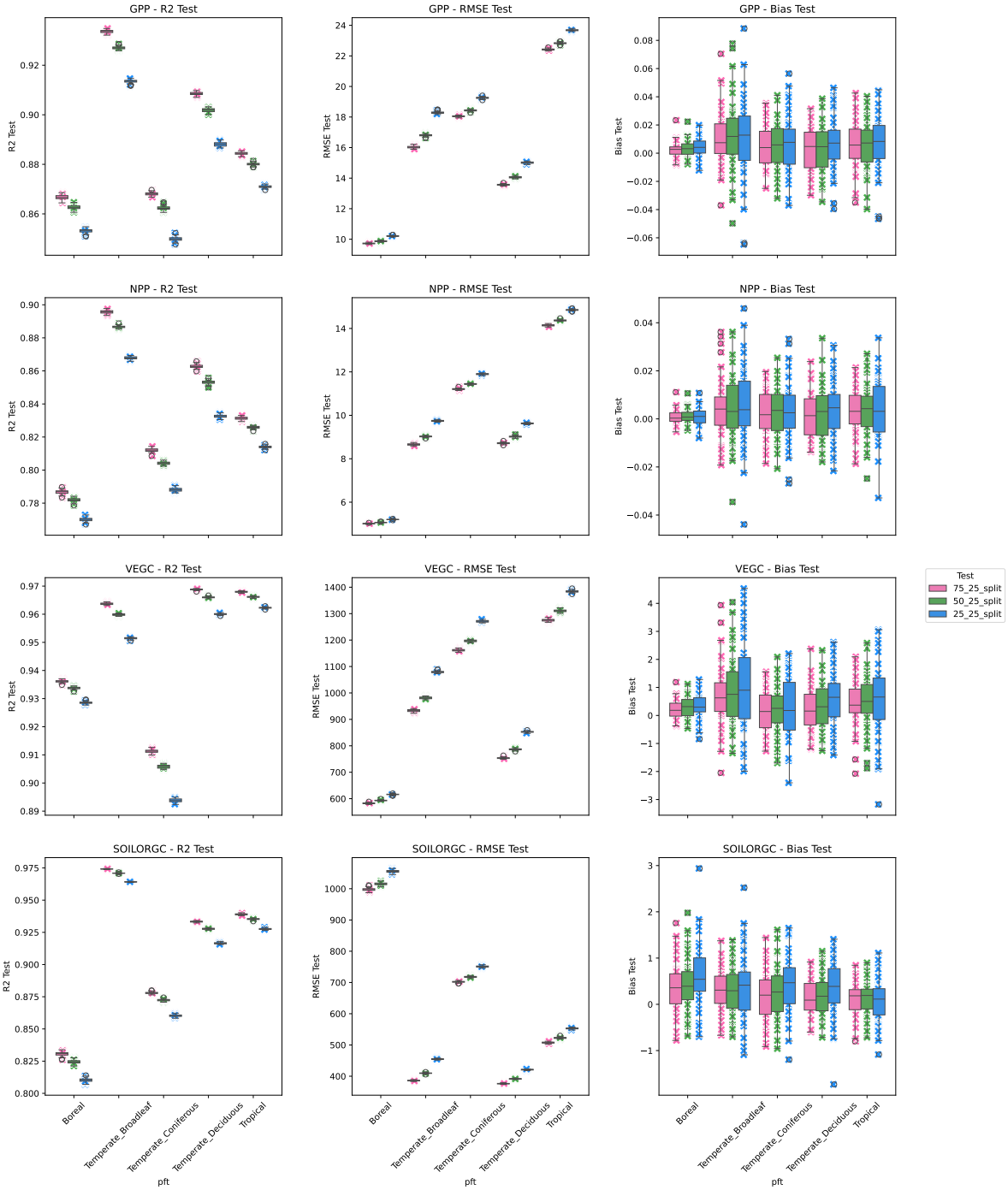

Figure 4: This shows the box plots for the R^2 and RMSE and Bias for the 50x Cross Validation for all three training split tests.

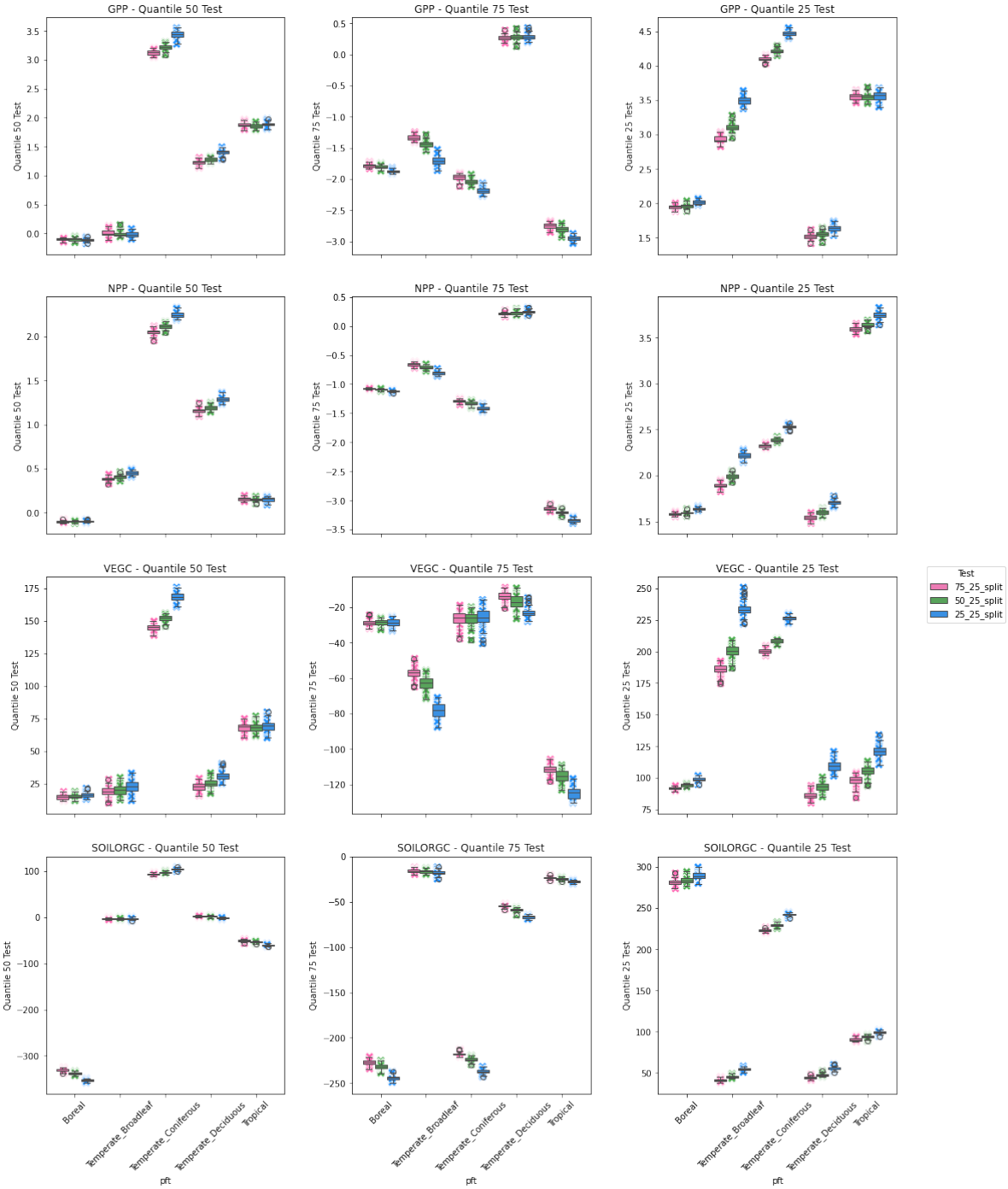

Figure 5: This shows the box plots for the 25th, 50th and 75th quantiles for the 50x Cross Validation for all three training split tests.

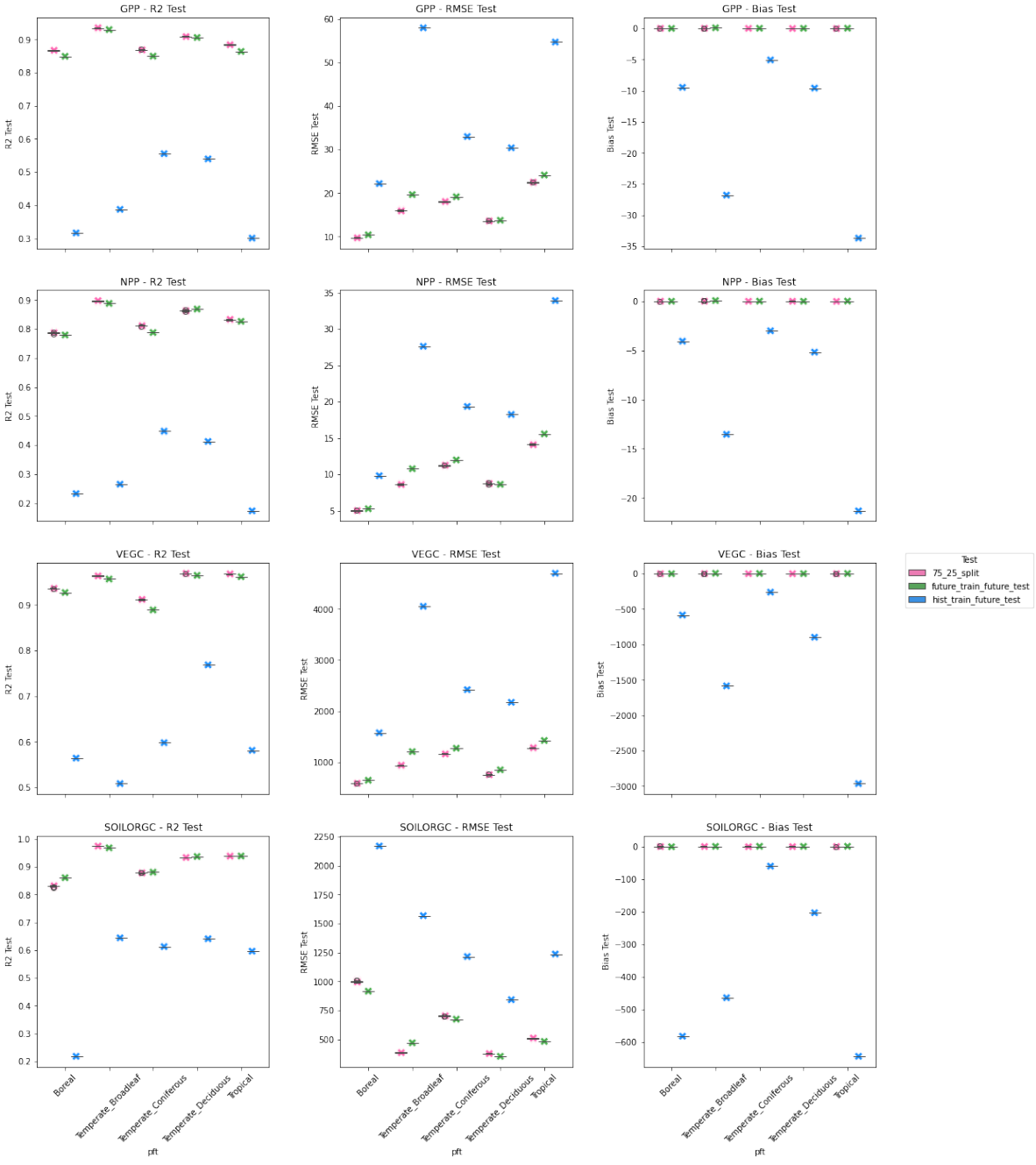

Figure 6: This shows the box plots for the R^2 and RMSE and Bias for the 50x Cross Validation for all three training time period tests.

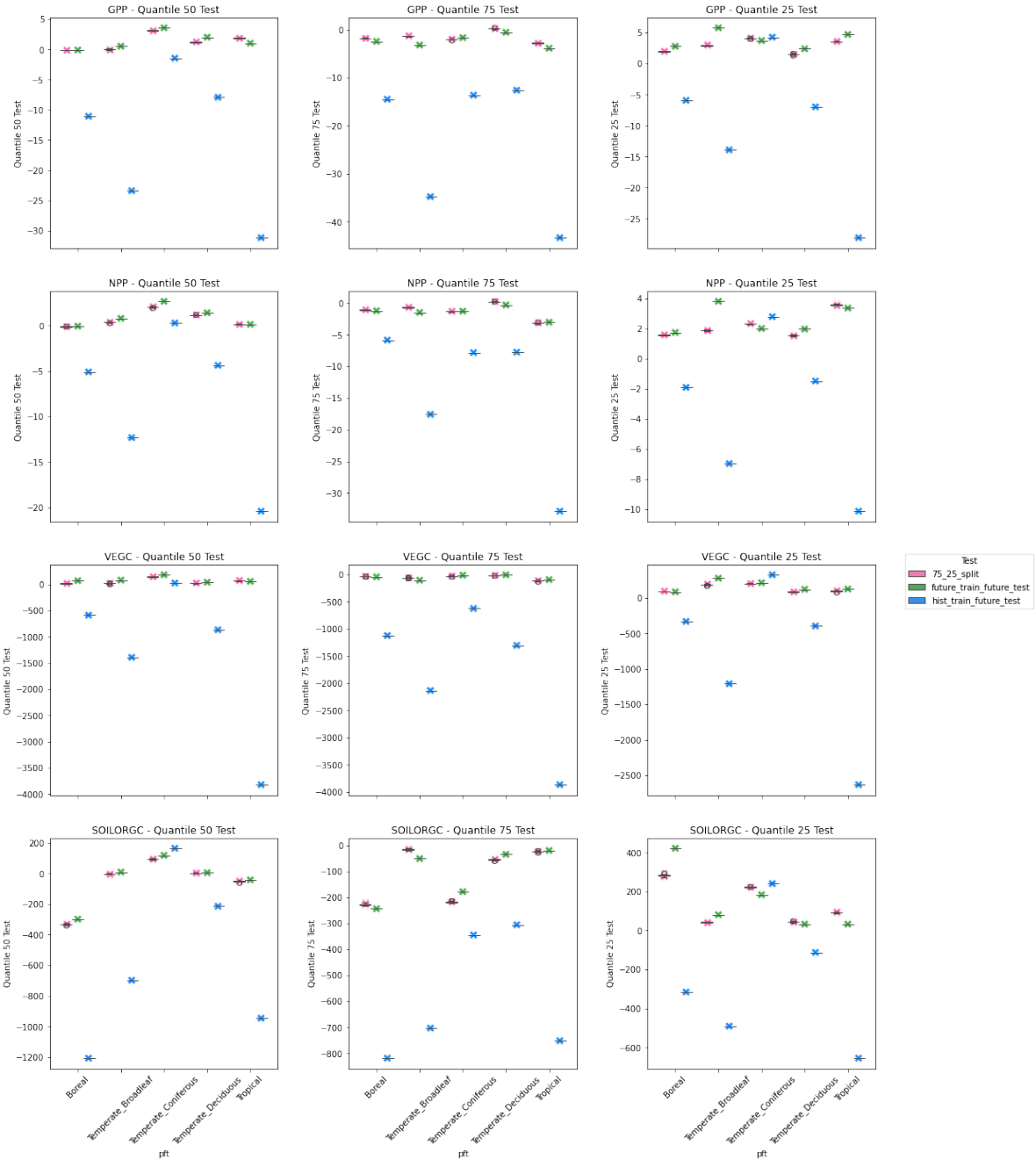

Figure 7: This shows the box plots for the 25th, 50th and 75th quantiles for the 50x Cross Validation for all three training time period tests.

4.3 Response to Stand Age

Because stand age is an important variable for harvesting decisions and considered the key driver of variation in carbon stocks and fluxes, we test how well the emulator can represent the range of stand age responses in the TEM simulation. As you can see in Figure 8, the emulator captures the average response to stand age but tends to show a smaller range of variation, as indicated by the overestimation of the 25th percentile and the underestimation of the 75th percentile. This is particularly evident for GPP and NPP in the first 50 year stand age bins. For SOILORGC, it is the initial carbon stock where the simulated response showed more variability than the emulated. For Temperate Deciduous, the emulated response actually shows a slightly wider variation for the 75th percentile of GPP, NPP and VEGC.

4.4 Emulator Variable Importance

As indicated in Figure 9, the most important variables for emulation are temperature, precipitation, stand age and vapor pressure deficit. Among PFT groups, there is consistency for the most important emulator variables for GPP and NPP. VEGC often follow a similar top 5 to NPP and GPP, however stand age is typically the most important variable, with Temperature Coniferous a notable exception. Because GPP, NPP and VEGC are all related to carbon assimilation in plant matter, we expect the most important variables to be similar. SOILORGC has distinct important variables, which is unsurprising as the processes that govern SOILORGC are different from those which govern photosynthesis. We refer to the carbon output variables GPP, NPP and VEGC as the "photosynthesis and autotrophic respiration governed variables" and to SOILORGC as the "heterotrophic respiration governed variable".

For the photosynthesis and autotrophic respiration governed variables, the average annual precipitation is important to almost all PFTs, except the Boreal PFT. While temperature is a top variable, whether the annual min, max or average (or some combination) were most important is PFT dependent. For the Tropical PFT, minimum temperature is important, whereas for Boreal the maximum temperature is more important. Maximum temperature is important for all the Temperate PFTs, though additionally average temperature is the most significant variable for Temerate Coniferous, and minimum temperature is important for Temperate Broadleaf. Ambient CO_2 is important only for Tropical and Boreal forests, and it does not show up in the top 5 for any Temperate PFTs.

Besides stand age, temperature, precipitation and vapor pressure deficit, certain PFTs have other variables that held top 5 importance. Average rsds is important for Temperate Coniferous. Average wind temperature is important for the Tropical PFT, suggesting that evapotranspiration and therefore soil moisture is more strongly limited by wind speeds for this PFT than others. Soil texture is typically an important

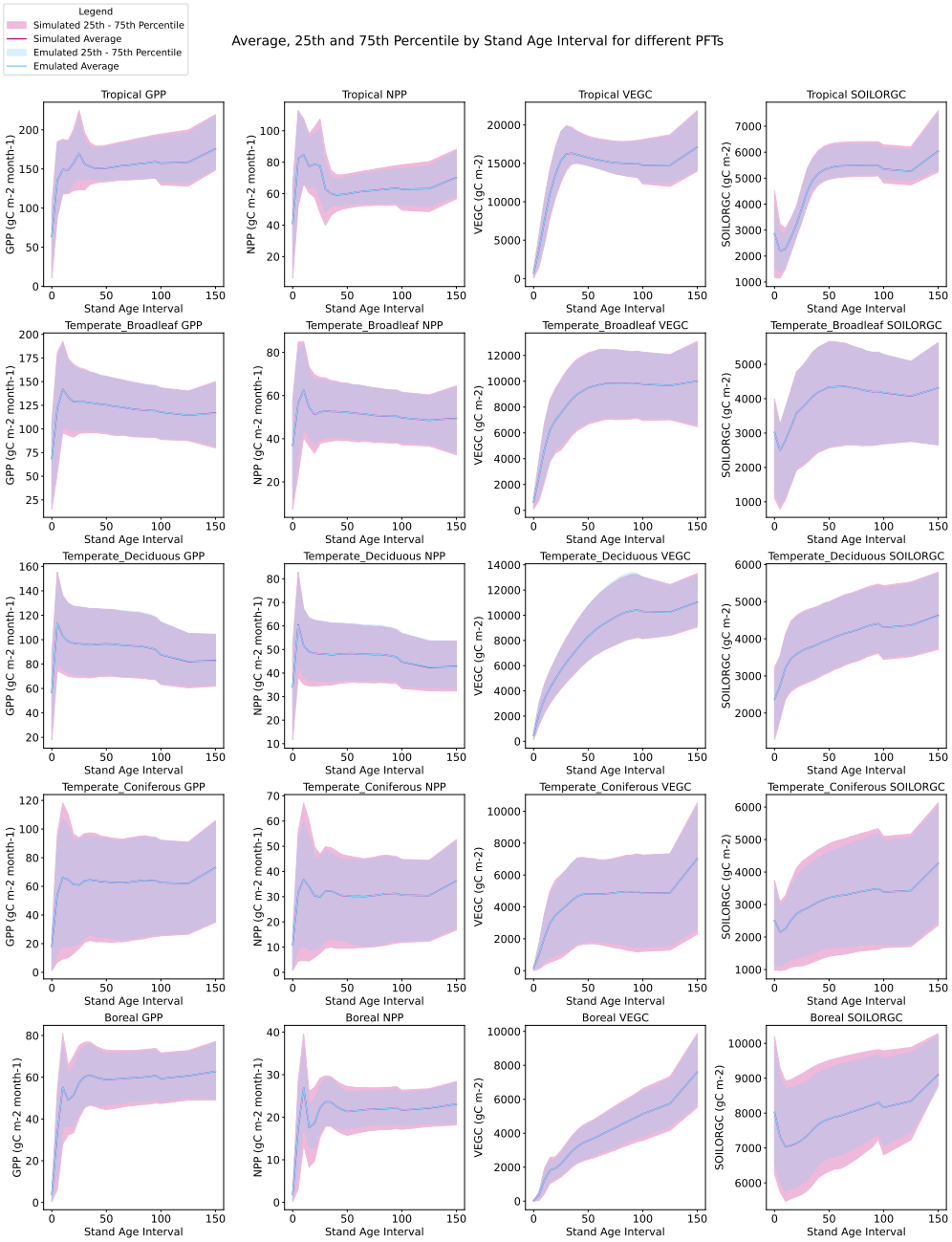

Figure 8: This figure shows the average response of the output variables to stand age as well as the shaded range from the 25th to 75th percentile for the emulated and simulated output variable.

variable across PFTs, particularly the percent of clay for Temperate forests, and the percent of sand for Boreal.

For SOILORGC, the heterotrophic governed variable, the top five variables typically include a few different climate inputs than the photosynthesis and autotrophic respiration governed variables. For the Tropical PFT, maximum temperature is more important and minimum temperature most important for Boreal, a reversal from the important annual temperature statistic from the photosynthesis and autotrophic respiration governed variables. Soil texture is important for SOILORGC in Temperate Coniferous and very important for Boreal as both the percent of sand and of clay are in the top, and the percent of clay is the most important variable. Stand age is most important to the Tropical PFT, and it is in the top five for Temperate Broadleaf and Temperate Deciduous as well. There is a notable difference in magnitude of the top variables for Temperate Coniferous and Boreal PFTs and no variable sticks out as most important, in comparison to Tropical, Temperate Broadleaf and Temperate Deciduous where the first or first and second ranked variables are 3-4 times as important as the next top five variable.

4.5 Partial Dependence Plots

In order to determine if the emulators are learning the functional forms between climate input and TEM, we show the PDPs and ICE plots for the top five important variables for each PFT for the GPP output variable, which has the simplest functional forms to visualize in the TEM equations (see Appendix A). As explained in the methodology section, the PDPs represent the average response, and the ICE plots show the variation in individual responses. ICE plots are centered on zero for the first value of the dependent variable to normalize the response curve for different values of GPP. Based on the functional form of climate input to GPP, holding all else constant, we expect the relationship with precipitation to follow the relationship between GPP and soil moisture, which is a hyperbolic relationship that varies with rooting depth parameter and soil texture. We expect the relationship with temperature to be parabolic. We expect the relationship with CO_2 to be hyperbolic. We expect that as the vapor pressure deficit increases, GPP decreases. See Table 1 for more details on TEM GPP equations and PFT dependent parameters. And, because stand age is not explicitly related to GPP, we expect that the relationship with stand age to follow the average relationship seen in the TEM simulation output. See Appendix F for the average relationship of all emulated variables to all the input variables.

For the Temperate Broadleaf GPP emulator, shown in Figure 10, we see approximately the expected relationships for all variables in the PDPs, though there is considerable variation in the direction of the response in the ICE plots for average precipitation and maximum air temperature. The precipitation-GPP

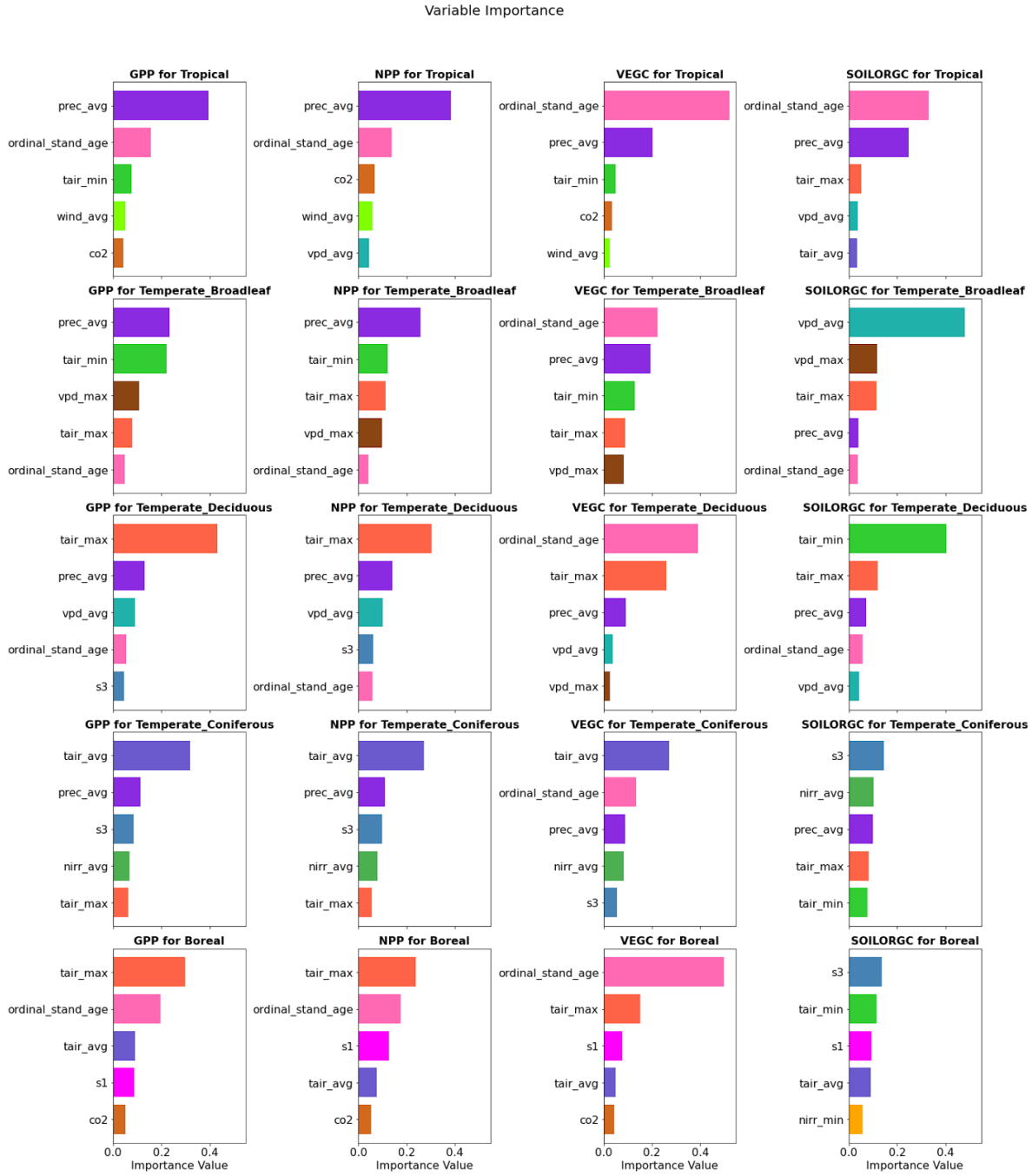

Figure 9: This figure shows the top five variable importance from each emulator. The different colors represent different input variables.

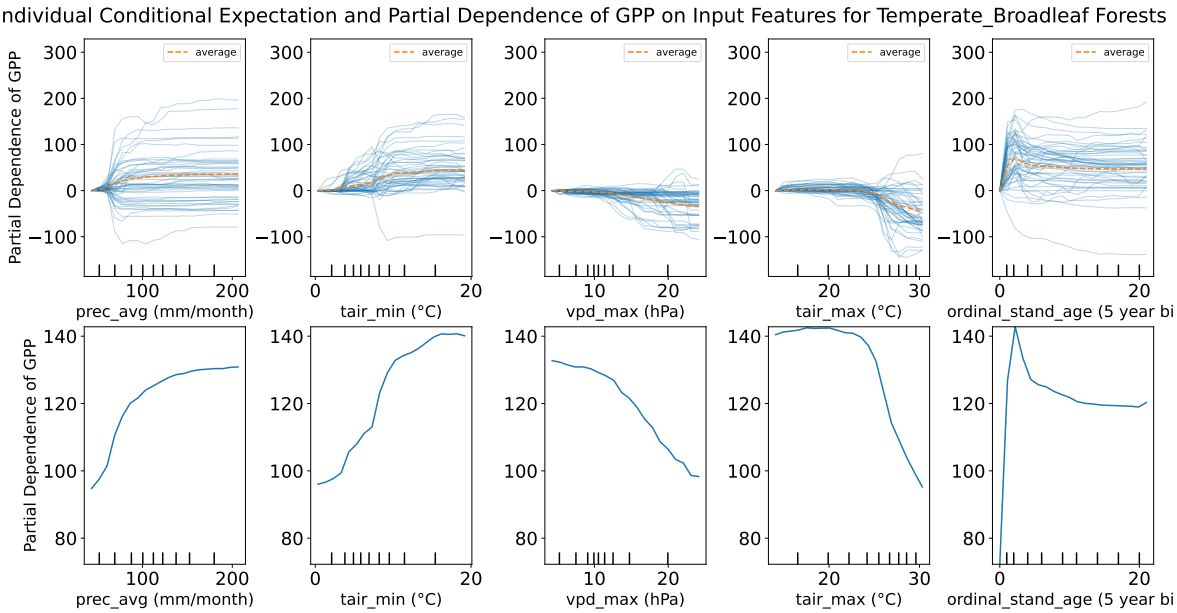

Figure 10: This figure shows the individual contribution expectation and partial dependence plots for the top five variables of the Temperate Broadleaf-GPP emulator. Note that the ICE plot's response dependence is centered on zero for the first value on the x-axis.

curve seems to follow a sigmoid relationship to GPP, where small amounts of precipitation lead to an exponential relationship, but overall we see that after a certain amount of monthly average precipitation, around 150 mm/month, precipitation stops having a significant effect on GPP, so there is a saturation, similar to the expected hyperbolic relationship. However, the ICE plot shows how variable the functional response to precipitation is, and there are times when precipitation has a negative effect on GPP even though the overall effect is positive. We see that precipitation does not have a negative effect until after 60mm/month for any of the plotted ICE, which gives confidence to the emulator as precipitation at such low rates would be beneficial to GPP. However, it is likely not feasible that precipitation has a negative effect on GPP starting at around 75 mm/month, though at least the negative effect does not increase with more precipitation. The relationships between VPD max and stand age are as expected for both the PDP and ICE, as they follow the expected functional form for VPD max, with variation that increases as VPD becomes more negative. Interestingly, when put together, the relationships with minimum and maximum temperature show a parabolic function form, where GPP increases until around 15°C, then begins to decrease after 25°C.

For the Tropical GPP emulator, shown in Figure 11, we see the expected response of GPP to precipitation, stand age and wind. Precipitation increases with GPP until saturation at 160 mm/month and the majority of ICE relationships show this increase, with only one example of precipitation having a negative effect on GPP. The relationship between stand age and GPP follows the double peak at 10 years and 25 years that we

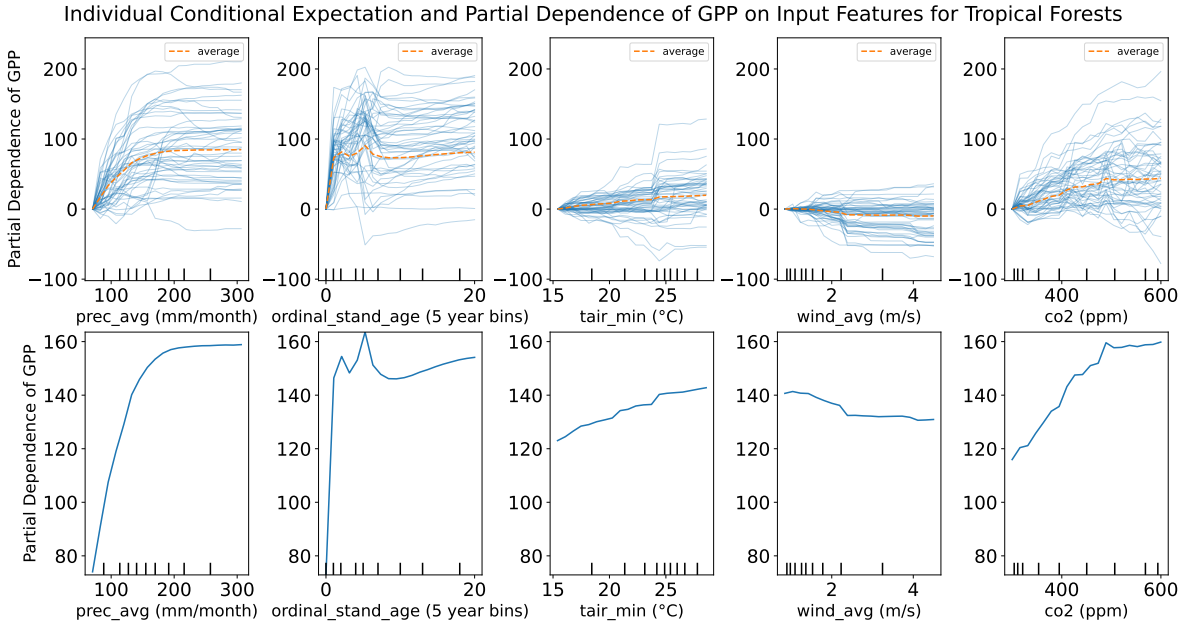

Figure 11: This figure shows the individual contribution expectation and partial dependence plots for top five variables of the Tropical-GPP emulator. Note that the ICE plot's response dependence is centered on zero for the first value on the x-axis.

see in the simulated output, though there is considerable variation in the ICE plot that again is seen in the simulated data. Wind speed starts to have a negative impact on GPP after 2 m/s, likely through an increase in ET leading to a decrease in soil moisture, though there are many instances where wind speed has little effect and this could be when soil moisture is sufficient. While not very smooth, the relationship between CO_2 and GPP shows the expected increase in GPP with CO_2 up until around 500 ppm, where there is an apparent saturation of the CO_2 fertilization effect. Note that the half saturation constant is around 400 ppm for Tropical forests, double what is the parameter for other PFTs (see Table 1). However, as the ICE plots make apparent, there are some sharp valleys and peaks in CO_2's impact on GPP. The relationship of GPP to minimum temperature does not show a parabolic relationship, but instead shows a slight linear increase with an increase in temperature on average. While this increase is physically sound, as GPP would increase with minimum temperature, it does not follow the function form of TEM's GPP equations. The ICE plots with minimum temperature show cases of sharp increases and decreases that occur at 23°C, likely this is a split point in the model which relates to minimum temperature as well as other input variable values.

For the Temperate Deciduous GPP emulator, shown in Figure 12, the PDP relationships for maximum air temperature, average precipitation, clay %, and stand age to GPP follow the expected functional form, however the relationship between VPD and GPP shows the opposite relationship from what is expected. For maximum air temperature, we see the first half of the parabolic relationship ends at 30°C. Based on the

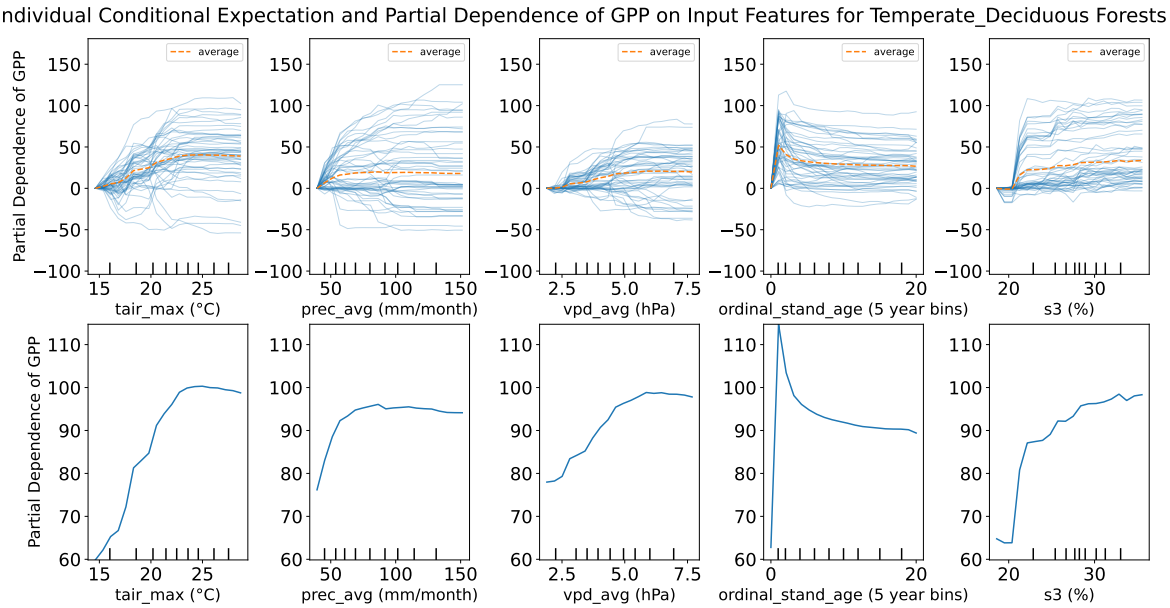

Figure 12: This figure shows the individual contribution expectation and partial dependence plots for top five variables of the Temperate Deciduous-GPP emulator. Note that the ICE plot's response dependence is centered on zero for the first value on the x-axis.

parameters for this PFT, the parabolic peak for Temperate Deciduous PFT is around 31°C, and the emulator has not been trained on temperatures higher than the peak. This warrants caution when using the emulator and the need for more training data, as the emulator has not learned that at higher maximum temperatures, GPP will decrease when the parabolic relationship in TEM suggests a decrease in GPP, all other factors constant. The relationship between average precipitation and the percent of clay in the soil both show similar functional forms to the relationship of soil moisture to GPP. For average precipitation, there is a saturation at around 95 mm/month. However, looking at the ICE there are instances of precipitation having negative effects on GPP. For the percent of clay, there is a sharp increase in GPP between 20 and 21 percent of soil. An increase in VPD, however, increases GPP on average and then its impact remains steady after a deficit of $\tilde{5}$ hPa. This is the opposite of the relationship to be expected, though it is notable that this relationship is seen in the simulation data as well, see Appendix F, but nonetheless suggests that the emulator learned a spurious relationship not supported by TEM's equations. The ICE plot shows that there are examples where GPP decreases with VPD, however you can also find these examples with maximum air temperature and precipitation.

For the Temperate Coniferous GPP emulator, shown in Figure 13, we see the expected functional form for max temperature and precipitation. Similar to the Temperate Deciduous training data, the Temperate Coniferous simulation does not have the temperature range needed for the second half of the parabolic relationship, however we do see the downward trend after around 30°C, the maximum optimal temperature.

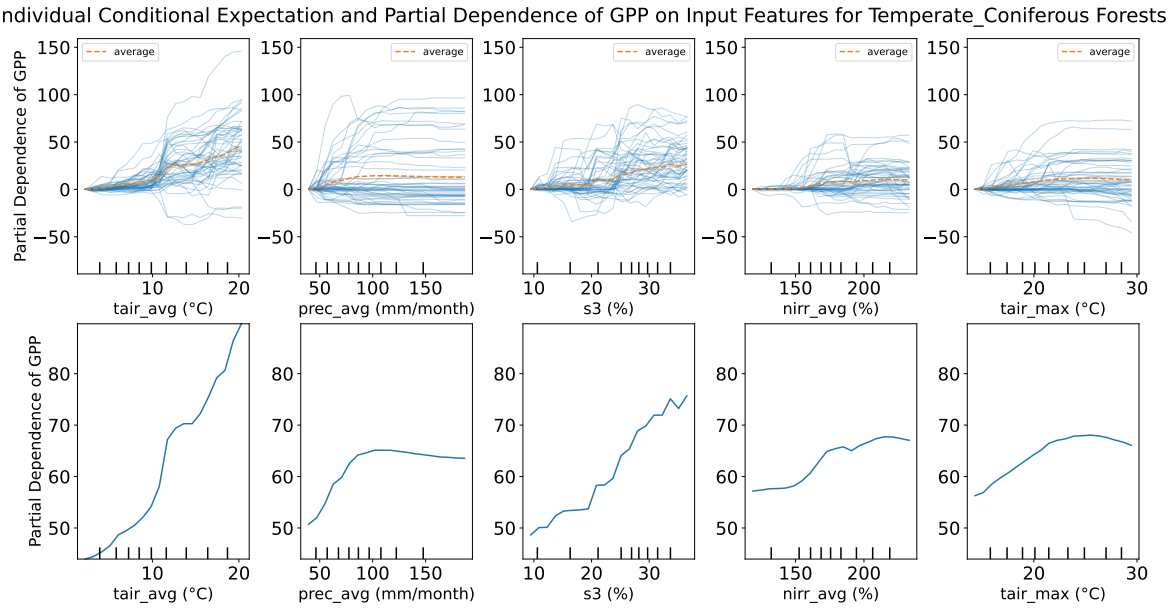

Figure 13: This figure shows the individual contribution expectation and partial dependence plots for top five variables of the Temperate Coniferous-GPP emulator. Note that the ICE plot's response dependence is centered on zero for the first value on the x-axis.

For the percent of clay, we see a general increase in GPP, as we would expect from the increase in soil moisture. Average air temperature does not follow the expected functional form, instead there are two inflection points and a sharp increase in GPP after 11°C, though as expected GPP increases with average air temperature. We would expect downwelling shortwave radiation to have a hyperbolic relationship to GPP, as PAR is assumed to be 48% of rsds. Instead, we see a sigmoid relationship with an inflection point at 150 W/m2.

The emulator for Boreal Forests, shown in Figure 14, shows increasing relationships with temperature, which is generally supported by the ICE plots, and also an increasing relationship with CO_2, though there is more variation in the ICE plots response. GPP should peak around 25°C, and so again the simulation training data does not include enough observations of high temperatures. GPP is highly dependent on stand age, with a relationship that follows the average in the TEM output. The percent of sand follows the expected relationship, as sand does not retain as much water as clay or silt.

As the emulator's goal is to effectively represent the stand age response to variation in climate, we show contour plots of the PDPs for the top four variables along with stand age in Figure 15, to look at interactions between stand age and climate variables and determine if they are physically sound. For the Tropical emulator, the contour plots show that the GPP is most sensitive to interactions between precipitation and CO_2 with stand age. Forests under 30 years do not have an apparent limit to the CO_2 fertilization effect, though for forests between 40 and 50 years high CO_2 levels can cause a decrease in GPP, which

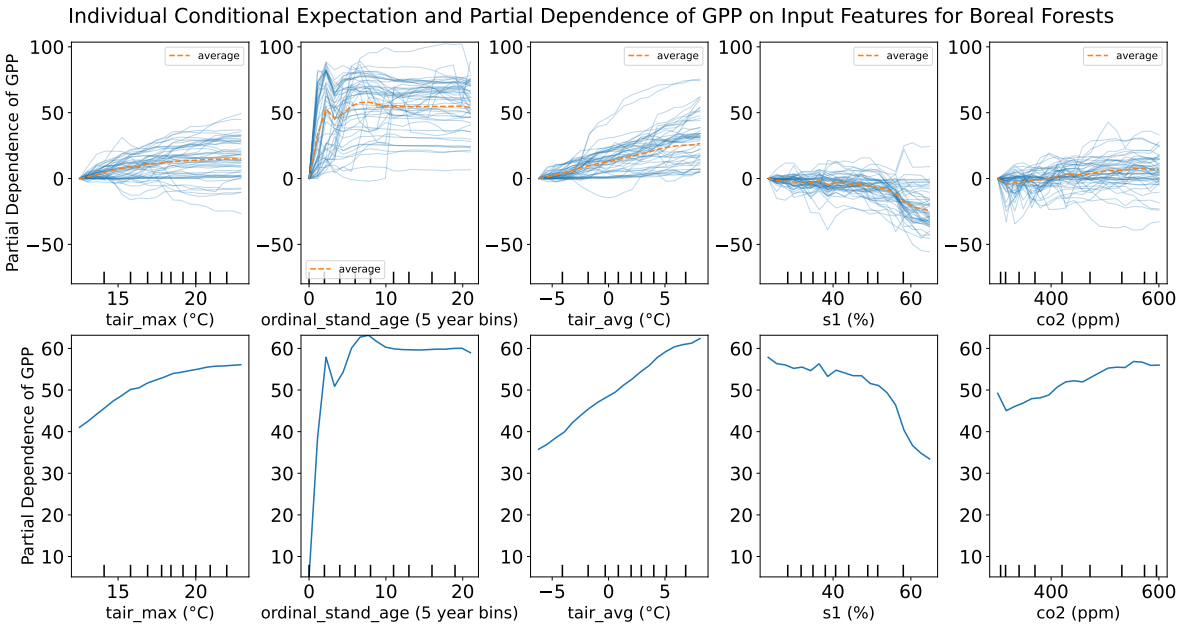

Figure 14: This figure shows the individual contribution expectation and partial dependence plots for top five variables of the Boreal-GPP emulator. Note that the ICE plot's response dependence is centered on zero for the first value on the x-axis.

might explain some of the decreasing relationships seen in the CO_2 ICE plot. Forests under 20 years do not have any upper bound on the amount of precipitation needed for growth, though forests between 25 and 30 years have an ideal precipitation range between around 130-240 mm/month. The sensitivity of GPP to minimum temperature is not very strong, though there are some age ranges where a higher minimum air temperature increases GPP. The relationship between stand age and wind is similarly not strong. We would have expected wind speed to be more important in the early growth phase, when there is less canopy cover, but instead lower wind speeds are beneficial in the late growth phase between 20 and 35 years, and again in more mature forests. For the Boreal emulator, GPP reaches its maximum at warmer maximum and average temperatures for the peak growth period of 30-40 years, as well as with lower percentages of sand and higher levels of CO_2. GPP remains high for stand ages over 25 years when the average annual temperature is over 3 °C.

All of the Temperate PFTs showed that air temperature is particularly important for young forests. For the Temperate Broadleaf emulator, the contour plots of the interaction of climate variables to stand age show expected relationships. GPP starts at low values uniformly, and then the peak growth rate between 5 and 20 years depends on favorable climate conditions like a low vapor pressure deficit, sufficient precipitation, and annual temperature between 10°C and 25°C. These plots also show that the temperature range between 10°C and 25°C is the climate interaction that leads to the largest increase in GPP for young Temperate

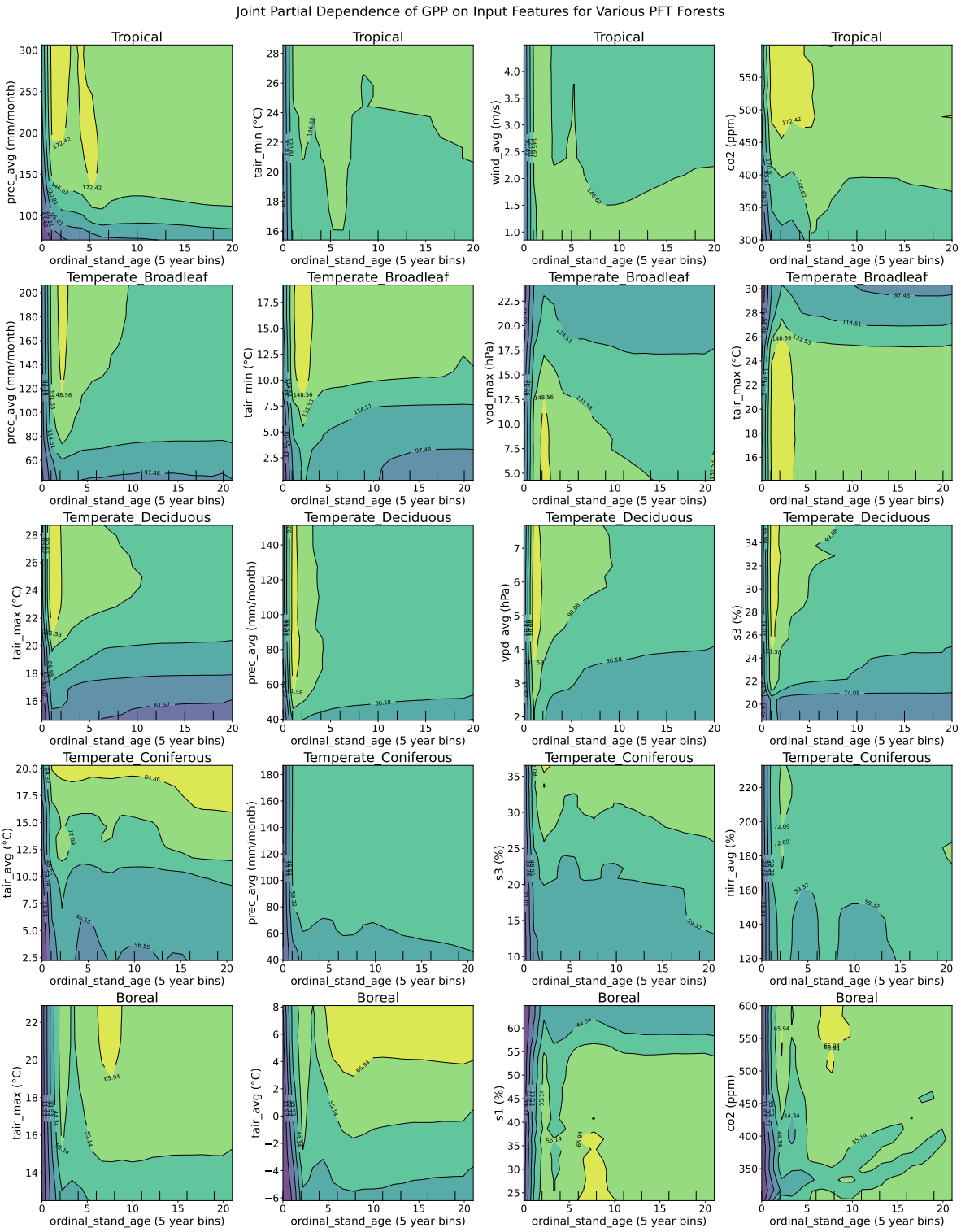

Figure 15: This figure shows contour partial dependence plots for the GPP emulators of each PFT, with stand age on the x-axis and the other top four variables on the y-axis. Note that stand age is shown here as an ordinal value, ranging from 0-22 for the 23 stand age bins.

Broadleaf forests. For the Temperate Deciduous emulator, the contour plot with stand age and VPD shows that a higher VPD increases GPP for young forests, which is not a feasible relationship. However, the other climate variables show more established relationships with stand age, as GPP is fairly sensitive to higher maximum temperatures, particularly for young forests. For precipitation, young forests have a high required amount of precipitation, around 60 mm/month they achieve high GPP but for forests over 25 years, they are not sensitive to an increase in precipitation over 30 mm/month. There is a similar relationship with the percent of clay, which makes sense because clay increases the amount of water storage. For the Temperate Coniferous emulator, stand age is not a top five variable for GPP for temperate coniferous, so we compared the interaction between stand age and the top 4 variables. Average air temperature and the % of clay are most sensitive in the interaction with stand age and rsds and precipitation are most important for young forests of around 10 years, and then old forests over 120 years.

5 Emulator Application - Stand Age Response to Annual Climate Perturbations

We apply the emulator to understand how the productivity of forests would be impacted under different global degrees of warming or rates of precipitation. We calculate the average annual temperature and precipitation for all forest grid cells over a control period of 1950 - 2000, then we take the average annual climate from 2000-2100 and compare when a given year is between 1 - 2 degrees (+1), 2 - 3 (+2), 3 - 4 (+3) and 4 - 5 (+4) degrees warmer. We also look at years where the precipitation rate is between 5-10%, 11-15%, and 16-20% higher than the control period. Table 3 shows the different scenarios and the number of years in each scenario group. The CanESM5-1 simulation does not have any years where the bias-corrected annual average precipitation is less than 5% from the baseline, so we do not include any drier scenarios. We take the annual climate input for all grid cells in each PFT, and create stand age bins for each grid cell. We run this input data through each emulator to look at the impact of these warmer and wetter worlds on the growth response functions for different PFTs. Note that while on average these forest grid cells are warmer and had more precipitation, there is significant variation among grid cells. Therefore, in the +4 degree world, it is possible some grid cells do not have any change in temperature, or even are cooler than the baseline. Our goal is to understand the impact on global GPP from changes in global averages, as opposed to the impact of a given grid cell.

As shown in Figures 16 and 17, in general, there is enhanced photosynthesis activity among all PFTs with the increase in temperature and precipitation, particularly for younger forests, which represent the peak

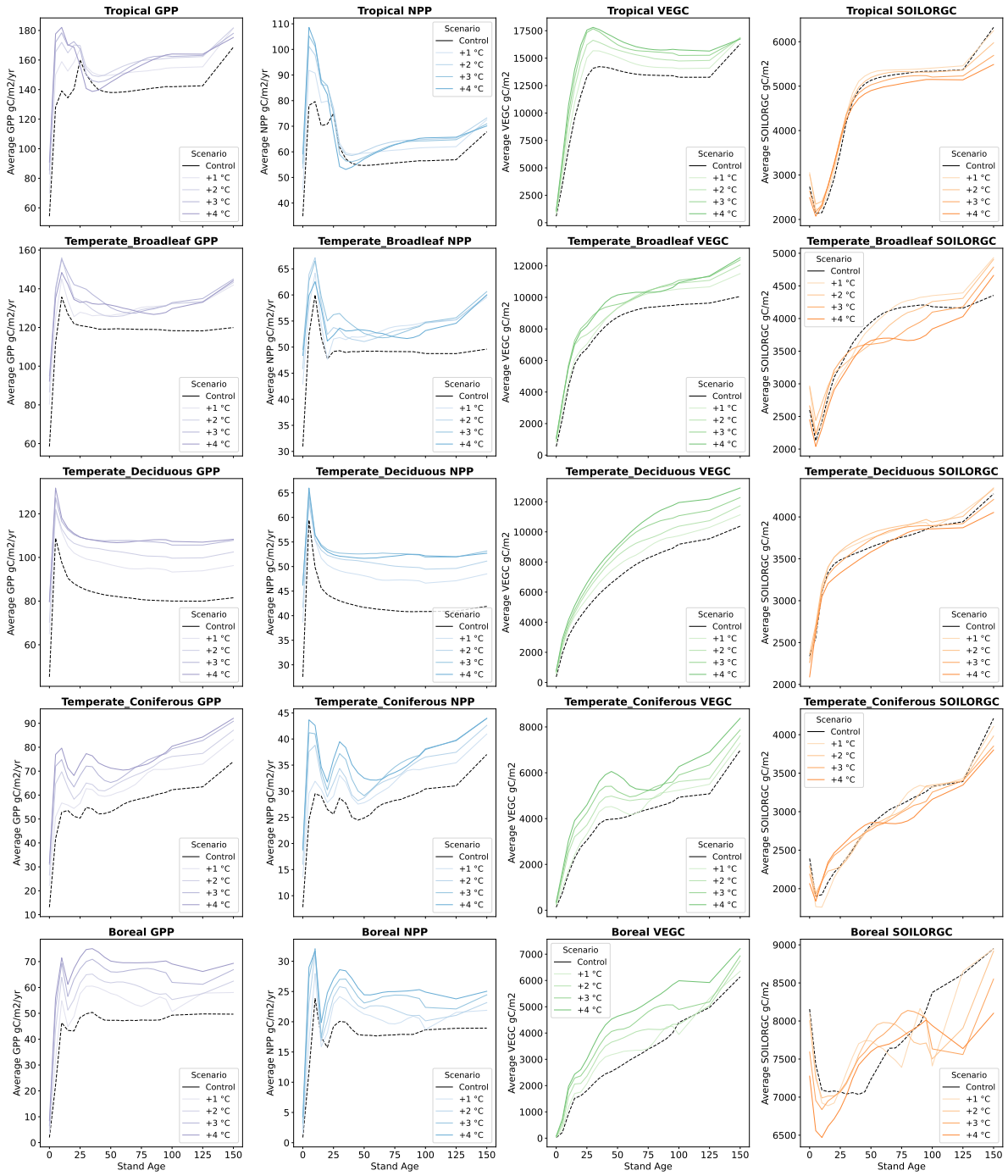

Figure 16: This figure shows each PFT's response of each output variable to stand age for different degrees of warming.

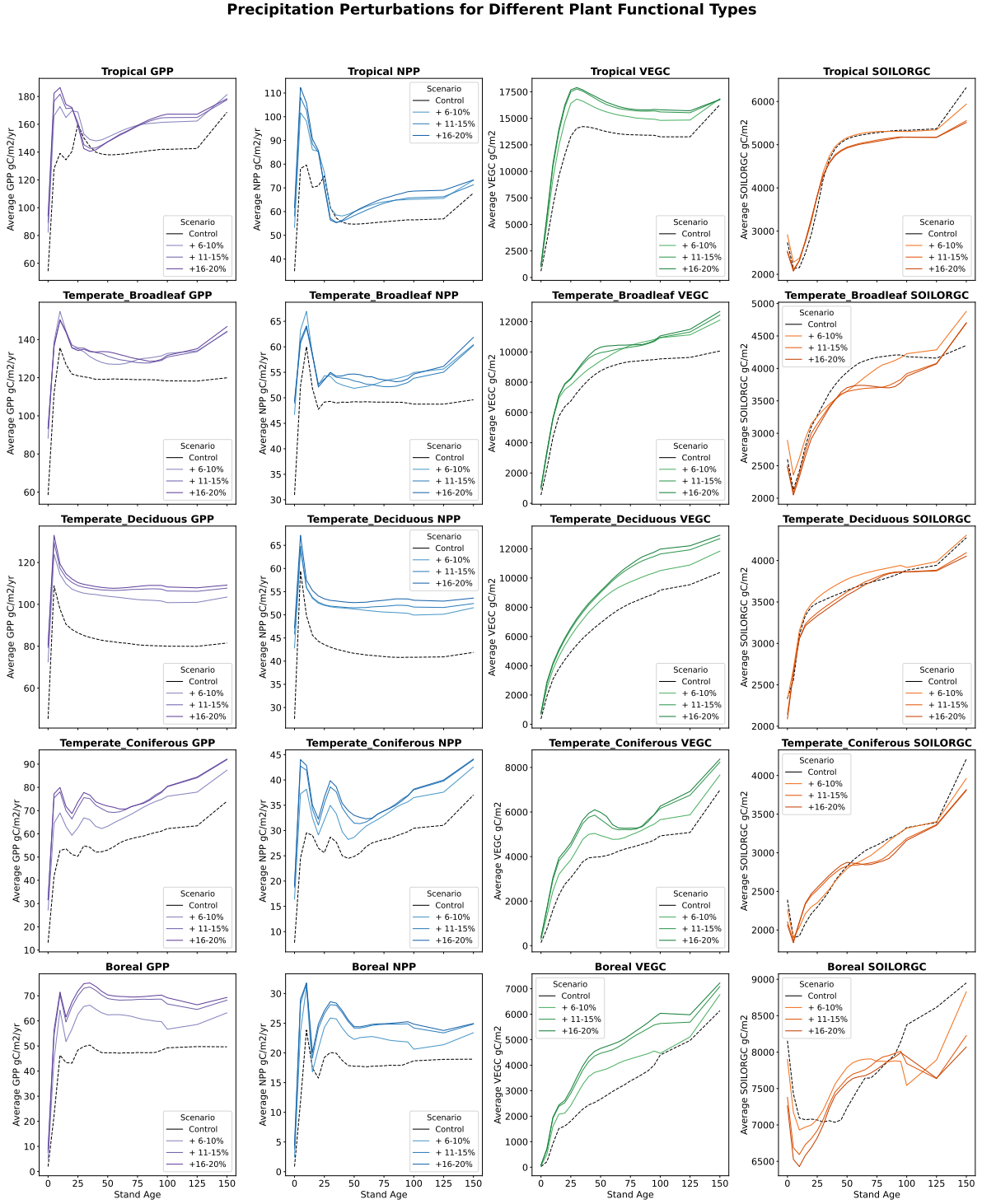

Figure 17: This figure shows each PFT's response of each output variable to stand age for different percent increases in precipitation.

Scenario	Range	N Years
1 °C	[1, 2) °C	16
2 °C	[2, 3) °C	21
3 °C	[3, 4) °C	18
4 °C	[4, 5) °C	31
5-10%	[5-10] %	30
11-15%	(10-15] %	26
16-20%	(15-20] %	12

Table 3: Number of years in each scenario group.

growth period. However most PFTs saw a decrease in soil carbon. The emulator results show that younger forests benefit most for the Tropical PFT, older forests of Temperate Broadleaf and middle age forests for Boreal and uniformly across all ages for Temperate Deciduous and Temperate Coniferous. For each incremental temperature degree, there are incremental benefits to GPP, NPP and VEGC for the Tropical, Temperate Coniferous, Temperate Deciduous and Boreal PFTs, but the Temperate Broadleaf PFT only has incremental benefits for young forests for GPP and NPP, and then for only young and old forests for VEGC, but not forests aged 60-100 years. However, for the +4 degree scenario, the incremental benefit decreases from the +3 degree scenario for Tropical, Temperate Broadleaf and Temperate Deciduous. For percent increases in precipitation, many PFTs only have an incremental benefit in GPP, NPP, and VEGC up to the + 10-15% increase in precipitation, after which the benefits from higher precipitation appear to saturate. We see this pattern for the Tropical, Boreal, Temperate Deciduous and Temperate Coniferous PFTs, whereas the Temperate Broadleaf PFT saturated at a lower percent increase in precipitation (+ 5-10%). For the photosynthesis-driven output variables of GPP, NPP and VEGC, there are positive benefits for the warmer and wetter scenarios, however for SOILORGC there are positive or negative impacts, depending on the stand age and the scenario. Temperature increases have little impact on the response for Tropical, Temperate Coniferous and Temperate Deciduous forests, but for Temperate Broadleaf there is a negative response for higher temperature for middle age forests and positive response for old forests, whereas Boreal shows the opposite relationship. Higher rates of precipitation create a negative response for Tropical, Temperate Broadleaf and most of Temperate Coniferous and Deciduous forest ages. Boreal forests have substantial impacts to the SOILORGC that change the functional form with age as young forests are negatively impacted, middle age forests see a positive impact, whose peak depends on the degree of temperature and precipitation increase, and older forests are negatively impacted. Carbon fluxes are positively impacted as is VEGC, though older forests are less positively impacted.

Metric	Unit	Tropical	Temperate Broadleaf	Temperate Coniferous	Temperate Deciduous	Boreal
GPP $RMSE$	gC/m2/month	22	16	17	13	10
NPP $RMSE$	gC/m2/month	14	9	11	8	5
VEGC $RMSE$	gC/m2	1277	934	1097	698	583
SOILORGC $RMSE$	gC/m2	507	389	673	351	990
GPP R^2		.88	.93	.87	.91	.87
NPP R^2		.83	.90	.81	.86	.79
VEGC R^2		.97	.96	.91	.97	.94
SOILORGC R^2		.94	.97	.88	.93	.83

Table 4: $RMSE$ and R^2 Performance Metrics for all Emulators.

6 Discussion

6.1 Emulator Performance

Emulating TEM with GBMs produce key carbon cycle output with high model fidelity. See Table 4 for an overview of key performance metrics R^2 and $RMSE$ for all emulators. The emulators are able to represent the 25th 50th and 75th percentiles of carbon cycle response with minimal bias, at maximum +/- 4 gC/m2/month for GPP and NPP, +/-200 gC/m2 for VEGC and +/-300 gC/m2 for SOILORGC. While GBMs do not have the advantage of robust confidence intervals and sensitivity testing that emulation methods like GP have, we are able to get a confidence interval of emulation performance variability using Monte Carlo cross validation, which shows that model performance is robust to differences in subsampled training data. Additionally, we demonstrate tools to understand how the input features are driving emulation prediction, for example VI and PDPs, which generally show the expected important variable as well as feasible relationships between the input variables and the predicted response. Importantly for our use case, which requires large training datasets because we are including stand age level input data, GBMs implemented through XGBoost are highly computationally efficient, able to train a 6 GB data set in 10 minutes.

While the PDPs are a powerful way to visualize the emulator's learned relationships, it is important to note that they provide the most accurate representations of relationships when the input features are independent. The correlation charts in Appendix F show that most input features are not correlated. While some climate variables showed high correlations between their maximum, average, and minimum annual values, the emulators do not choose highly correlated features as top variables. For example, for Temperate Coniferous, average and maximum temperature are both chosen as important features, which have a correlation of 0.6, but minimum temperature, which has a 0.9 correlation with average temperature, is not chosen as a top variable. Also note that the L1 penalty on additional features limits the inclusion of correlated features as they do not contribute to improving the overall R^2. However, while the correlation charts show that the majority of features are linearly independent through their low correlation, it is unreasonable to

assume that the features are truly independent given that they are output from a climate model. So results should be interpreted with caution.

The GBMs emulation method is able to be applied across PFTs and variables, all with comparable performance (compared to the MARS emulator, discussed in Appendix B). The relative performance of the GPP and NPP emulators by PFT were similar, with Temperate Broadleaf performing the best, followed by Temperate Deciduous, Tropical, Temperate Coniferous and Boreal. NPP emulation performance is lower than GPP, which makes sense given that NPP has the additional complexity of being related to the maintenance and growth respiration, and while respiration processes are closely linked to temperature, they are also dependent on the allocation percents of leaves, stem, roots, sapwood which relate to all input variables as well as the past carbon stocks. VEGC emulators have high performance metrics, with all R^2 metrics above .9. SOILORGC have the largest variation in R^2 among the PFTs, with a low of 0.83 for Boreal forests and a high of 0.97 for Temperate Broadleaf. While Temperate Broadleaf overall have the best emulation results, it is also the smallest PFT in terms of number of grid cells, and therefore sample size. However, the two PFTs that covered the largest land area, Boreal and Tropical forests, do not show similar performance results with Tropical typically being the second or third best emulated and Boreal being the worst or second worse.

Another key finding from this research is the importance of including the maximum and minimum of climate input when emulating at a different temporal resolution than the simulated output. This is particularly true for temperature and VPD as typically the maximum or minimum of temperature, and occasionally VPD, are chosen as the most important variables. This shows that the average annual temperature is not the most informative to the annual carbon stocks and fluxes, as can be expected from the nonlinear GPP equations. For VPD, the maximum and average are both selected a comparable amount of times, with the maximum being important for Temperate Broadleaf and the average for Temperate Deciduous. We would not expect the minimum of VPD to be informative, since it would not be a limiter on GPP and since it is oftentimes 0. For precipitation, the average is always chosen to be the most important variable, signifying its importance over the monthly extremes. In the few times that wind or rsds are selected as top variables, the average is important, with the one exception of the Boreal SOILORGC emulator. While we only included the min, max and average as annual statistics for the climate variables, seasonal climate variables like growing season minimum, maximum and average temperature could be important to include. Future versions of the emulator should test the importance of including growing season metrics as predictor variables.

There is a lot of potential for these emulators to be coupled with IAMs, as they efficiently reproduce the stand age carbon response of TEM. The average stand age response is captured well with little bias. The 25th and 75h percentile response by stand age of the emulator reduces the variance in the response

compared to the simulated output, however this is a workable limitation for an IAM that is typically run at the regional level and so will have reduced variance from the grid cell level. We have demonstrated how this emulator can apply the stand age response of all emulated carbon variables in climates whose annual average is +1-5° C hotter, or 5-20% wetter. These response curves could be coupled with an IAM to show how different forest PFTs might respond on aggregate, which can facilitate decisions around land use and harvesting. Additionally, IAMs can couple the full emulator, since the emulators can produce output for one carbon cycle variable at a rate of around 150,000 observations per second.

6.2 TEM Simulation Stand Age Response - Comparison to Existing Research

There have been a few other studies that looked at the carbon cycle impact from CMIP6 climate scenarios, and their results support key findings from our emulator application. Lu et al., 2024 looked at GPP output from CMIP6 models and found that for the CMIP6 models they chose, there was an increase in GPP everywhere except the tropics (Lu et al., 2024). The researchers trained a statistical model to predict satellite derived GPP from climate variables and then used this to run simulations with CMIP6 climate input. A study looking at the NPP output (derived from GPP-Ra) of the abruptly quadrupled CO_2 simulation of 23 CMIP6 models found an increase in NPP in the extra-tropics, which they attributed to warmer, wetter average climate, and a decrease in NPP in the tropics, which they attributed to warmer, drier climates (J. Zhu et al., 2022). Another study looked at peak vegetation greening found that 70% of global vegetative areas would increase, however all forest types besides evergreen needle-leaf forests were predicted to have browning compared to the baseline period, though this does not have a direct link to GPP and NPP (Teng et al., 2023). Our emulator results imply that warmer or wetter climates lead to an increase in carbon fluxes and vegetation stocks compared to the 1950-2000 baseline. However, we only included one CMIP6 model, CanESM5-1, and the inclusion of more CMIP6 models would allow for a look into drier climates to confirm that this would lead to a decrease in carbon fluxes and vegetation stocks.

While we establish that the emulator is able to reproduce TEM's carbon flux and stock response to stand age, it is also important to establish that TEM's stand age response is supported by the literature. It is well established that for secondary forests, which have enough carbon and nitrogen stores in the soil from previous forest generations, VEGC increases at a logarithmic or sigmoid rate, with the fastest growth rates typically while the forest is less than 50 years old (Zaehle et al., 2006; Pan et al., 2002; C. Zhu et al., 2020), though some studies have found evidence of linear increases in VEGC for temperate and tropical biomes (Pregitzer et al., 2004). The sigmoid shape of the curve comes from the regeneration phase, as defined by NFBI (Kurz et al., 1999), however when there is sufficient carbon and nitrogen in the soil forests can

begin to regrow within 5-10 years, leading to the logarithmic relationship seen in many experimental and theoretical studies (Chapin et al., 2011; Pan et al., 2002; Peichl et al., 2010). The logarithmic response is also found in many LSMs that are coupled with the climate models that run CMIP simulations. Zhu and Xia 2020 demonstrated the implicit CMIP5 and CMIP6 response of VEGC to forest age and found that the majority of models fit the nonlinear, logarithmic growth for VEGC to stand age (C. Zhu et al., 2020). For NPP, meta-studies of experimental data have found that NPP increases hyperbolically with stand age, peaking around 20-30 years, or parabolically, sometimes decreasing down to zero (Zaehle et al., 2006), other times then decreasing but remaining elevated above zero as the forest ages. Peichl et al 2010 demonstrated how forest productivity metrics, like GPP and NPP, can be shown to decrease towards zero, as opposed to remaining elevated after peaking due to failure to account for site differences in species specific relationships (Peichl et al., 2010).

In TEM and its emulators, the VEGC response for all PFTs follow forest age structures typically seen in the literature for the control experiment. In warmer and wetter worlds Temperate Coniferous and Tropical forests have a more pronounced peak, between 20 and 30 years for tropical and at 40 years for Temperate Coniferous. The TEM NPP stand age response match the responses found in Pan et al 2002, where there is a peak in NPP, followed by a decline to an elevated steady state (Pan et al., 2002). We see this pattern for Temperate Deciduous and Temperate Broadleaf, however for Temperate Coniferous, Boreal, and Tropical forests we see a double peak in GPP and NPP that is not supported by the literature. Further research is needed to understand this response. Soil carbon responses are complicated by the varying depths at which soil carbon is measured over, and so difficult to directly compare to TEM output.

7 Conclusions and Future Work

We develop a collection of emulators that can represent TEM's response of key carbon cycle variables to stand age and climate impact with high model fidelity. We use the emulator to demonstrate the impact, by stand age, of climates that are globally warmer or wetter. Future work could look at the impact of climates that are warmer AND wetter or warmer and drier. We show that XGBoost can be used for emulation tasks that require large amounts of simulation data, and we demonstrate that PDPs and VI metrics give insight into how the emulator is approximating the functional relationship between input features and output carbon variables. We use the model insight recommendations provided in Friedman's original paper describing GBMs, however future research could also apply the SHAP (SHapley Additive exPlanations) method to understand model relationships (Lundberg et al., 2017).

One key shortcoming of the emulator is that despite the training data abundance, the emulator has

not learned that with higher temperatures there will be a negative response in GPP according to the PFT dependent parameters shown in Table 1. This could be improved by adding warmer climate scenarios, like SSP585, to the training data. Another area to explore is the inclusion of temporal lags. For LSMs, TEM included, each gridcell is run independently and so it is reasonable to assume spatial independence in emulation of LSMs (Baker et al., 2022). However, temporal independence is a simplification, as the numeric schemes of LSMs take into account past time steps and furthermore stocks are a variable that holds memory of cohort history. For the purpose of an emulator to couple with an IAM, the inclusion of only the current time step as input features will make for easier coupling. But if one wanted a more accurate emulator, future research could explore adding in time lags to input features.

This emulator also does not include any of TEM's atmospheric chemistry input of ozone, nitrogen deposition nor nitrogen fertilization. Ozone directly impacts GPP though a damage function, which has been shown to be an important determinant of how forest productivity will respond to different pollution scenarios (Reilly et al., 2007), and nitrogen deposition is important to the amount of available nitrogen for a grid cell. We do not include these variables to allow for coupling with IAMs which might not have these input variables available. However, further work could look into the impact of these variables. While nitrogen fertilization is only applied to crop and pasture, if a future emulator were to look at temporal lags, fertilization would be a variable to include.

These emulators are developed to be coupled with IAMs, which need to include stand age level predictions of carbon stocks or fluxes. Many IAMs do not factor in the impact of future climate on forest growth assumptions. These emulators provide a computationally efficient solution that can represent complex biogeochemical responses. Future work could apply these emulators to an IAM in order to determine the impact on harvesting and land use decisions.

Book Sources

Chapin, F. Stuart, Pamela A. Matson, and Peter M. Vitousek (2011). *Principles of Terrestrial Ecosystem Ecology*. New York, NY: Springer. ISBN: 978-1-4419-9503-2 978-1-4419-9504-9. DOI: 10.1007/978-1-4419-9504-9.

Rasmussen, Carl Edward and Christopher K. I. Williams (2006). *Gaussian processes for machine learning*. Adaptive computation and machine learning. Cambridge, Mass: MIT Press. 248 pp. ISBN: 978-0-262-18253-9.

Spurr, S.H and Barnes, B.V (1973). *Forest Ecology*. New York, NY: The Ronald Press Co. 571pp.

Unep, FAO and (2020). *The State of the World's Forests 2020*. FAO and UNEP ; ISBN: 978-92-5-132419-6.

Other Sources

Baker, Evan et al. (Mar. 8, 2022). "Emulation of high-resolution land surface models using sparse Gaussian processes with application to JULES". In: *Geoscientific Model Development* 15.5, pp. 1913–1929. ISSN: 1991-959X. DOI: 10.5194/gmd-15-1913-2022.

Calvin, Katherine et al. (Feb. 15, 2019). "GCAM v5.1: representing the linkages between energy, water, land, climate, and economic systems". In: *Geoscientific Model Development* 12.2, pp. 677–698. ISSN: 1991-959X. DOI: 10.5194/gmd-12-677-2019.

Cannon, Alex J., Stephen R. Sobie, and Trevor Q. Murdock (Sept. 1, 2015). "Bias Correction of GCM Precipitation by Quantile Mapping: How Well Do Methods Preserve Changes in Quantiles and Extremes?" In: *Journal of Climate* 28.17, pp. 6938–6959. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-14-00754.1.

Chakraborty, Avishek et al. (Apr. 3, 2017). "Emulation of Numerical Models With Over-Specified Basis Functions". In: *Technometrics* 59.2, pp. 153–164. ISSN: 0040-1706. DOI: 10.1080/00401706.2016.1164078.

Chen, Tianqi and Carlos Guestrin (Aug. 13, 2016). "XGBoost: A Scalable Tree Boosting System". In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–794. DOI: 10.1145/2939672.2939785.

Duckworth, Jennifer C, Martin Kent, and Paul M Ramsay (2000). "Plant functional types: an alternative to taxonomic plant community description in biogeography?" In: *Progress in Physical Geography* 24.4, pp. 515–542.

Eyring, Veronika et al. (May 26, 2016). "Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization". In: *Geoscientific Model Development* 9.5, pp. 1937–1958. ISSN: 1991-959X. DOI: 10.5194/gmd-9-1937-2016.

Felzer, B. et al. (Jan. 1, 2004). "Effects of ozone on net primary production and carbon sequestration in the conterminous United States using a biogeochemistry model". In: *Tellus B: Chemical and Physical Meteorology* 56.3, pp. 230–248. ISSN: null. DOI: 10.3402/tellusb.v56i3.16415.

Felzer, Benjamin S. (Feb. 6, 2023). "Effect of land-use legacy on the future carbon sink for the conterminous US". In: *Biogeosciences* 20.3, pp. 573–587. ISSN: 1726-4170. DOI: 10.5194/bg-20-573-2023.

Felzer, Benjamin S., Timothy W. Cronin, Jerry M. Melillo, David W. Kicklighter, and C. Adam Schlosser (2009). "Importance of carbon-nitrogen interactions and ozone on ecosystem hydrology during the 21st century". In: *Journal of Geophysical Research: Biogeosciences* 114 (G1). ISSN: 2156-2202. DOI: 10.1029/2008JG000826.

Felzer, Benjamin S., Timothy W. Cronin, Jerry M. Melillo, David W. Kicklighter, C. Adam Schlosser, and Shree R. S. Dangal (2011). "Nitrogen effect on carbon-water coupling in forests, grasslands, and shrublands in the arid western United States". In: *Journal of Geophysical Research: Biogeosciences* 116 (G3). ISSN: 2156-2202. DOI: 10.1029/2010JG001621.

Friedlingstein, Pierre et al. (Dec. 5, 2023). "Global Carbon Budget 2023". In: *Earth System Science Data* 15.12, pp. 5301–5369. ISSN: 1866-3508. DOI: 10.5194/essd-15-5301-2023.

Friedman, Jerome H. (2001). "Greedy Function Approximation: A Gradient Boosting Machine". In: *The Annals of Statistics* 29.5, pp. 1189–1232. ISSN: 0090-5364.

Gangopadhyay, Subhrendu, Martyn Clark, and Balaji Rajagopalan (2005). "Statistical downscaling using K-nearest neighbors". In: *Water Resources Research* 41.2. ISSN: 1944-7973. DOI: 10.1029/2004WR003444.

Goldstein, Alex et al. (Jan. 2, 2015). "Peeking Inside the Black Box: Visualizing Statistical Learning With Plots of Individual Conditional Expectation". In: *Journal of Computational and Graphical Statistics* 24.1, pp. 44–65. ISSN: 1061-8600. DOI: 10.1080/10618600.2014.907095.

Gower, Stith T., Ross E. McMurtrie, and Danuse Murty (Sept. 1, 1996). "Aboveground net primary production decline with stand age: potential causes". In: *Trends in Ecology & Evolution* 11.9, pp. 378–382. ISSN: 0169-5347. DOI: 10.1016/0169-5347(96)10042-2.

Harris, Ian et al. (Apr. 3, 2020). "Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset". In: *Scientific Data* 7.1, p. 109. ISSN: 2052-4463. DOI: 10.1038/s41597-020-0453-3.

Hurtt, George C. et al. (Nov. 10, 2020). "Harmonization of global land use change and management for the period 850–2100 (LUH2) for CMIP6". In: *Geoscientific Model Development* 13.11, pp. 5425–5464. ISSN: 1991-9603. DOI: 10.5194/gmd-13-5425-2020.

Kennedy, Marc C. and Anthony O'Hagan (2001). "Bayesian calibration of computer models". In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 63.3, pp. 425–464. ISSN: 1467-9868. DOI: 10.1111/1467-9868.00294.

Kira, Tatuo and Tsunahide Shidei (1967). "Primary Production and Turnover of Organic Matter in Different Forest Ecosystems of the Western Pacific". In: 17.2, pp. 70–87. DOI: 10.18960/seitai.17.2_70.

Kurz, Werner A. and Michael J. Apps (1999). "A 70-Year Retrospective Analysis of Carbon Fluxes in the Canadian Forest Sector". In: *Ecological Applications* 9.2, pp. 526–547. ISSN: 1939-5582. DOI: 10.1890/1051-0761(1999)009[0526:AYRAOC]2.0.CO;2.

Lamarque, J.-F. et al. (Aug. 20, 2013). "Multi-model mean nitrogen and sulfur deposition from the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP): evaluation of historical and projected future changes". In: *Atmospheric Chemistry and Physics* 13.16, pp. 7997–8018. ISSN: 1680-7316. DOI: 10.5194/acp-13-7997-2013.

Li, Yuxiao and Ying Sun (2019). "Efficient Estimation of Nonstationary Spatial Covariance Functions with Application to High-Resolution Climate Model Emulation". In: *Statistica Sinica* 29.3, pp. 1209–1231. ISSN: 1017-0405.

Limon, Garrett C. and Christiane Jablonowski (2023). "Probing the Skill of Random Forest Emulators for Physical Parameterizations Via a Hierarchy of Simple CAM6 Configurations". In: *Journal of Advances in Modeling Earth Systems* 15.6, e2022MS003395. ISSN: 1942-2466. DOI: 10.1029/2022MS003395.

Liu, Yaling et al. (Mar. 23, 2018). "A hydrological emulator for global applications – HE v1.0.0". In: *Geoscientific Model Development* 11.3, pp. 1077–1092. ISSN: 1991-959X. DOI: 10.5194/gmd-11-1077-2018.

Lovett, Gary M., Jonathan J. Cole, and Michael L. Pace (Feb. 1, 2006). "Is Net Ecosystem Production Equal to Ecosystem Carbon Accumulation?" In: *Ecosystems* 9.1, pp. 152–155. ISSN: 1435-0629. DOI: 10.1007/s10021-005-0036-3.

Lu, Qikai et al. (Feb. 20, 2024). "Global prediction of gross primary productivity under future climate change". In: *Science of The Total Environment* 912, p. 169239. ISSN: 0048-9697. DOI: 10.1016/j.scitotenv.2023.169239.

Lundberg, Scott and Su-In Lee (Nov. 24, 2017). *A Unified Approach to Interpreting Model Predictions*. DOI: 10.48550/arXiv.1705.07874.

Marmin, Sébastien and Maurizio Filippone (Dec. 2022). "Deep Gaussian Processes for Calibration of Computer Models (with Discussion)". In: *Bayesian Analysis* 17.4, pp. 1301–1350. ISSN: 1936-0975, 1931-6690. DOI: 10.1214/21-BA1293.

McGuire, A. David et al. (1995). "Equilibrium Responses of Soil Carbon to Climate Change: Empirical and Process-Based Estimates". In: *Journal of Biogeography* 22.4, pp. 785–796. ISSN: 0305-0270. DOI: 10.2307/2845980.

Mohammadi, Hossein, Peter Challenor, and Marc Goodfellow (Nov. 1, 2019). "Emulating dynamic non-linear simulators using Gaussian processes". In: *Computational Statistics & Data Analysis* 139, pp. 178–196. ISSN: 0167-9473. DOI: 10.1016/j.csda.2019.05.006.

Narayan, Kanishka B et al. (2024). "Seeing the forest for the trees: Implementing dynamic representation of forest management and forest carbon in a long-term global multisector model". In: *Environmental Research Letters*. ISSN: 1748-9326. DOI: 10.1088/1748-9326/ad6ea3.

O'Hagan, A. (Oct. 1, 2006). "Bayesian analysis of computer code outputs: A tutorial". In: *Reliability Engineering & System Safety*. The Fourth International Conference on Sensitivity Analysis of Model Output (SAMO 2004) 91.10, pp. 1290–1300. ISSN: 0951-8320. DOI: 10.1016/j.ress.2005.11.025.

O'Sullivan, Michael et al. (Dec. 2021). "Aerosol–light interactions reduce the carbon budget imbalance". In: *Environmental Research Letters* 16.12, p. 124072. ISSN: 1748-9326. DOI: 10.1088/1748-9326/ac3b77.

Odum, Eugene P. (Apr. 18, 1969). "The Strategy of Ecosystem Development". In: *Science* 164.3877, pp. 262–270. DOI: 10.1126/science.164.3877.262.

Ohrel, Sara Bushey (2019). "Policy Perspective on the Role of Forest Sector Modeling". In: *Journal of forest economics* 34.3, pp. 187–204. ISSN: 1104-6899. DOI: 10.1561/112.00000506.

Pan, Yude et al. (2002). "A biogeochemistry-based dynamic vegetation model and its application along a moisture gradient in the continental United States". In: *Journal of Vegetation Science* 13.3, pp. 369–382. ISSN: 1654-1103. DOI: 10.1111/j.1654-1103.2002.tb02061.x.

Peichl, Matthias, M. Altaf Arain, and Jason J. Brodeur (July 15, 2010). "Age effects on carbon fluxes in temperate pine forests". In: *Agricultural and Forest Meteorology* 150.7, pp. 1090–1101. ISSN: 0168-1923. DOI: 10.1016/j.agrformet.2010.04.008.

Pregitzer, Kurt S. and Eugénie S. Euskirchen (2004). "Carbon cycling and storage in world forests: biome patterns related to forest age". In: *Global Change Biology* 10.12, pp. 2052–2077. ISSN: 1365-2486. DOI: 10.1111/j.1365-2486.2004.00866.x.

Raich, J. W. et al. (1991). "Potential Net Primary Productivity in South America: Application of a Global Model". In: *Ecological Applications* 1.4, pp. 399–429. ISSN: 1939-5582. DOI: 10.2307/1941899.

Reilly, J. et al. (Nov. 1, 2007). "Global economic effects of changes in crops, pasture, and forests due to changing climate, carbon dioxide, and ozone". In: *Energy Policy* 35.11, pp. 5370–5383. ISSN: 0301-4215. DOI: 10.1016/j.enpol.2006.01.040.

Rose, Steven K. (Nov. 1, 2014). "Integrated assessment modeling of climate change adaptation in forestry and pasture land use: A review". In: *Energy Economics* 46, pp. 548–554. ISSN: 0140-9883. DOI: 10.1016/j.eneco.2014.09.018.

Ryan, Michael G., Stith T. Gower, et al. (1995). "Woody Tissue Maintenance Respiration of Four Conifers in Contrasting Climates". In: *Oecologia* 101.2, pp. 133–140. ISSN: 0029-8549.

Ryan, Michael G. and Barbara J. Yoder (1997). "Hydraulic Limits to Tree Height and Tree Growth". In: *BioScience* 47.4, pp. 235–242. ISSN: 0006-3568. DOI: 10.2307/1313077.

Sacks, Jerome et al. (Nov. 1989). "Design and Analysis of Computer Experiments". In: *Statistical Science* 4.4, pp. 409–423. ISSN: 0883-4237, 2168-8745. DOI: 10.1214/ss/1177012413.

Sigmond, Michael et al. (Nov. 15, 2023). "Improvements in the Canadian Earth System Model (CanESM) through systematic model analysis: CanESM5.0 and CanESM5.1". In: *Geoscientific Model Development* 16.22, pp. 6553–6591. ISSN: 1991-959X. DOI: 10.5194/gmd-16-6553-2023.

Snyder, Abigail et al. (Apr. 3, 2019). "A crop yield change emulator for use in GCAM and similar models: Persephone v1.0". In: *Geoscientific Model Development* 12.4, pp. 1319–1350. ISSN: 1991-959X. DOI: 10.5194/gmd-12-1319-2019.

Sudakow, Ivan, Michael Pokojovy, and Dmitry Lyakhov (Jan. 2022). "Statistical mechanics in climate emulation: Challenges and perspectives". In: *Environmental Data Science* 1, e16. ISSN: 2634-4602. DOI: 10.1017/eds.2022.15.

Tebaldi, Claudia, Abigail Snyder, and Kalyn Dorheim (Nov. 11, 2022). "STITCHES: creating new scenarios of climate model output by stitching together pieces of existing simulations". In: *Earth System Dynamics* 13.4, pp. 1557–1609. ISSN: 2190-4979. DOI: 10.5194/esd-13-1557-2022.

Teng, Hongfen et al. (July 1, 2023). "Future changes and driving factors of global peak vegetation growth based on CMIP6 simulations". In: *Ecological Informatics* 75, p. 102031. ISSN: 1574-9541. DOI: 10.1016/j.ecoinf.2023.102031.

Tian, H. et al. (Jan. 1, 1999). "The sensitivity of terrestrial carbon storage to historical climate variability and atmospheric CO2 in the United States". In: *Tellus B: Chemical and Physical Meteorology* 51.2, pp. 414–452. ISSN: null. DOI: 10.3402/tellusb.v51i2.16318.

Tian, Hanqin et al. (June 1, 2018). "The Global N2O Model Intercomparison Project". In: *Bulletin of the American Meteorological Society* 99.6, pp. 1231–1251. ISSN: 0003-0007, 1520-0477. DOI: 10.1175/BAMS-D-17-0212.1.

Vörösmarty, C. J., C. A. Federer, and A. L. Schloss (June 25, 1998). "Potential evaporation functions compared on US watersheds: Possible implications for global-scale water balance and terrestrial ecosystem modeling". In: *Journal of Hydrology* 207.3, pp. 147–169. ISSN: 0022-1694. DOI: 10.1016/S0022-1694(98)00109-7.

WOODWELL, GEORGE M. and ROBERT H. WHITTAKER (Feb. 1, 1968). "Primary Production in Terrestrial Ecosystems". In: *American Zoologist* 8.1, pp. 19–30. ISSN: 0003-1569. DOI: 10.1093/icb/8.1.19.

Xu, Hongtao et al. (2024). "Global Forest Plantations Mapping and Biomass Carbon Estimation". In: *Journal of Geophysical Research: Biogeosciences* 129.3, e2023JG007441. ISSN: 2169-8961. DOI: 10.1029/2023JG007441.

Xu, Qing-Song and Yi-Zeng Liang (Apr. 16, 2001). "Monte Carlo cross validation". In: *Chemometrics and Intelligent Laboratory Systems* 56.1, pp. 1–11. ISSN: 0169-7439. DOI: 10.1016/S0169-7439(00)00122-2.

Yuval, Janni and Paul A. O'Gorman (July 3, 2020). "Stable machine-learning parameterization of subgrid processes for climate modeling at a range of resolutions". In: *Nature Communications* 11.1, p. 3295. ISSN: 2041-1723. DOI: 10.1038/s41467-020-17142-3.

Zaehle, Sönke et al. (2006). "The Importance of Age-Related Decline in Forest Npp for Modeling Regional Carbon Balances". In: *Ecological Applications* 16.4, pp. 1555–1574. ISSN: 1939-5582. DOI: 10.1890/1051-0761(2006)016[1555:TIOADI]2.0.CO;2.

Zhu, Chen and Jianyang Xia (2020). "Nonlinear Increase of Vegetation Carbon Storage in Aging Forests and Its Implications for Earth System Models". In: *Journal of Advances in Modeling Earth Systems* 12.12, e2020MS002304. ISSN: 1942-2466. DOI: 10.1029/2020MS002304.

Zhu, Jiawen, Xiaofei Gao, and Xiaodong Zeng (2022). "Response of terrestrial net primary production to climate change associated with the quadrupling CO2 forcing in CMIP6 models". In: *Atmospheric Science Letters* 23.9, e1098. ISSN: 1530-261X. DOI: 10.1002/asl.1098.

A Appendix A

The following equation is for potential GPP (GPP_p), not limited by nitrogen interactions. All variables are at the current time step, t, and GPP_p is calculated in the units $gCm^{-2}month^{-1}$.

$$GPP_p = C_{max} \times f_T \times f_{H_2O} \times f_{C_a,D} \times f_{O_3} \times \int_0^{LAI} f_{PAR} dL \tag{1}$$

where F_T, f_{H_2O}, $f_{C_a,D}$, f_{O_3}, and $\int_0^{LAI} f_{PAR} dL$ are functions or temperature, soil moisture stress, CO_2 concentration and vapor pressure deficit, ozone, and photosynthetically active radiation, and C_{max} is a PFT defined parameter for the maximum rate of leaf-level photosynthesis.

For the emulator, we expect that temperature and precipitation will be the most important variables. We will look at the functional forms of the updated TEM F_T and f_{H_2O}. TEM-hydro has an updated temperature response function that includes a T_{opt}.

$$f_T = \frac{(t_{max} - t_{air}) \times (t_{air} - t_{min})^{\frac{t_{opt}-t_{min}}{t_{max}-t_{opt}}}}{(t_{max} - t_{opt}) \times (t_{opt} - t_{min})^{\frac{t_{opt}-t_{min}}{t_{max}-t_{opt}}}} \tag{2}$$

In this equation T_{air} is the average monthly air temperature, T_{max} and T_{min} are parameters that represent bounds on photosynthetic activity and unique to each PFT. T_{opt} a five year rolling average of the maximum temperature and is meant to factor in some adaptability to climate changes, however there are bounds on T_{opt} which are PFT dependent and are shown in Table 1. Figure 16 plots these and shows the potential peaks of the parabolic relationship, depending on the adapted T_{opt}.

The function for soil moisture impact is from the Water Balance Model in Vörösmarty et al., 1998.

$$f_{H_2O} = 1 - e^{\frac{-lcs}{lcs_{min}} \times \frac{availw}{awcap}} \tag{3}$$

Where $availw$ is the amount of available water in the soil and $awcap$ is the maximum possible plant extractable water in the soil profile, a function of the rooting zone depth, a PFT parameter, and the percent

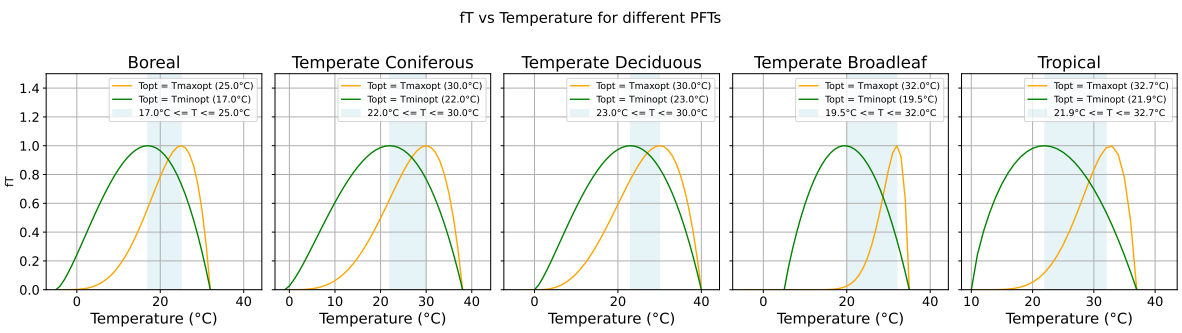

Figure 18: Potential fT for different PFT parameters.

of sand and clay in the soil, which is a static percent for each grid cell. The variable lsc is leaf specific conductance, a function of the LAI and the carbon in the stem and roots and their hydraulic conductance and lcs_{min} is a PFT parameter.

f_{PAR} is related to the amount of rsds, assumed to be 48%, integrated over the current timestep LAI. If we perform the integral function we find:

$$\int_0^{LAI} f_{PAR} dL = \frac{1}{k_{ext}} ln(\frac{k_I + k_{ext} \times PAR(L)}{k_I + k_{ext} \times PAR(L) \times e^{-k_{ext} LAI}}) \tag{4}$$

Where k_I is the half saturation constant, and k_{ext} is the extinction coefficient, and these are PFT dependent. In the following chart, we map the relationship between irradiance and GPP_p assuming that LAI is the average maximum annual LAI for each PFT, to account for seasonality in deciduous forests.

NPP will depend on GPP as well as autotrophic respiration. The temperature dependence on autotrophic respiration, f_{rmt} is as follows:

$$f_{rmt} = \frac{[Q_{ref} \times e^{-\alpha(T - T_{ref})}]^{\frac{T - T_{ref}}{10}} / [1 + e^{(\beta - T)} + e^{(T - \gamma)}]}{[Q_{ref} \times e^{-\alpha(T_{opt} - T_{ref})}]^{\frac{T_{opt} - T_{ref}}{10}} / [1 + e^{(\beta - T_{opt})} + e^{(T_{opt} - \gamma)}]} \tag{5}$$

B Appendix B

MARS, described in Friedman (1991), estimates the input - output function with additive linear splines. It utilizes recursive partitioning regression to split the data into various sub regions to estimate parameters for each spline. The method selects variable subsets locally, which makes it dynamic and adaptable because it can track interactions of complex functions which allows the model to approximate high dimensional functions. The method includes a forward step, which adds a basis function and deliberately overfits the model, and then a backward step removes splits that do not contribute to model performance. MARS tries to isolate outliers with a series of splits. While MARS is more robust to outliers than linear models, outliers

are particularly important around the partitioned data boundaries. For the MARS method it is important to choose variables without a high correlation. Collinearity in MARS leads to a suboptimal model because it can lead to spurious interaction effects as it is challenging for the function approximation to separate additive and interactive contributions. A large benefit of MARS is that the model is reasonably interpretable when the number of interactions between variables is less than 2, which is what we selected when testing MARS emulators. MARS allows for ANOVA decomposition to interpret the model, and so does not need to rely on methods like PDPs and ICE to understand how each input variable is affecting the output variable. As MARS grow in complexity they do lose interpretability as well however it is still possible to examine the piecewise linear relationships and compare them to the function form of TEM.

MARS is stepwise, versus GBM which is stagewise, this makes MARS more computationally intensive. Friedman (2001) compared MARS and GBM and found that there were many target functions by which each method substantially outperformed the other. However, MARS typically had a higher RMSE than average absolute error, indicating that the predictions were either close to or far away from the target. Friedman reasoned that the piecewise constant approach is more robust to outliers and wide tails than the piecewise continuous MARS function approximation. For emulating TEM for all PFTs and all variables, XGBoost outperformed MARS, as you can see in S1. Because of the MARS runtime and that there was little variation in mode performance over cross validation, likely due to the large training dataset, we only ran 5 folds for MARS, and therefore randomly selected 5 out of 50 of the iterations for the XGBoost comparison.

C Appendix C

This appendix shows the R^2 results for different tuning parameters for all output variables of the Temperate Coniferous PFT.

D Appendix D

This appendix shows the KDE plots for the simulated versus emulated output variables, as well as the mean and 25th and 75th percentiles.

E Appendix E

This appendix shows the results of adding incrementally 20 year periods of training data for NPP for the Temperate Broadleaf PFT.

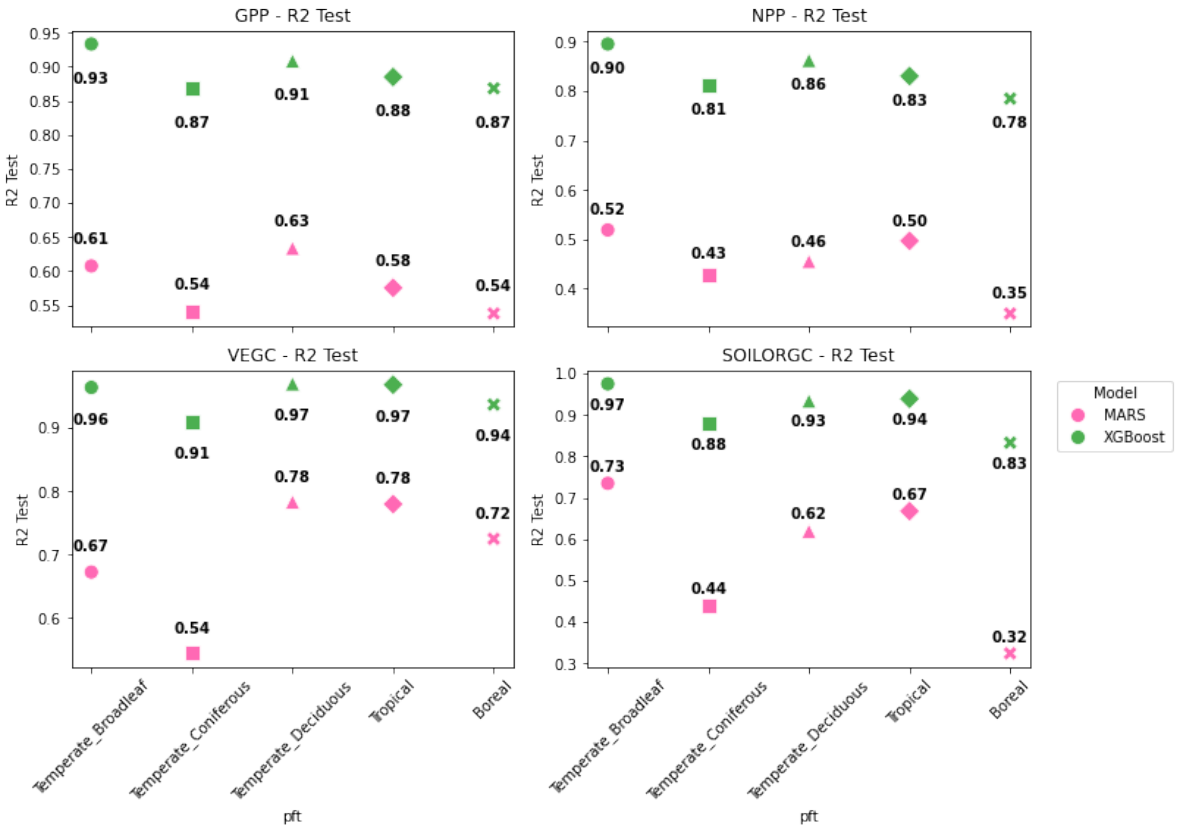

Figure 19: A comparison of MARS to XGBoost R^2 for all emulators.

F Appendix F

This appendix shows data summary plots for all input and output variables for all PFTs.

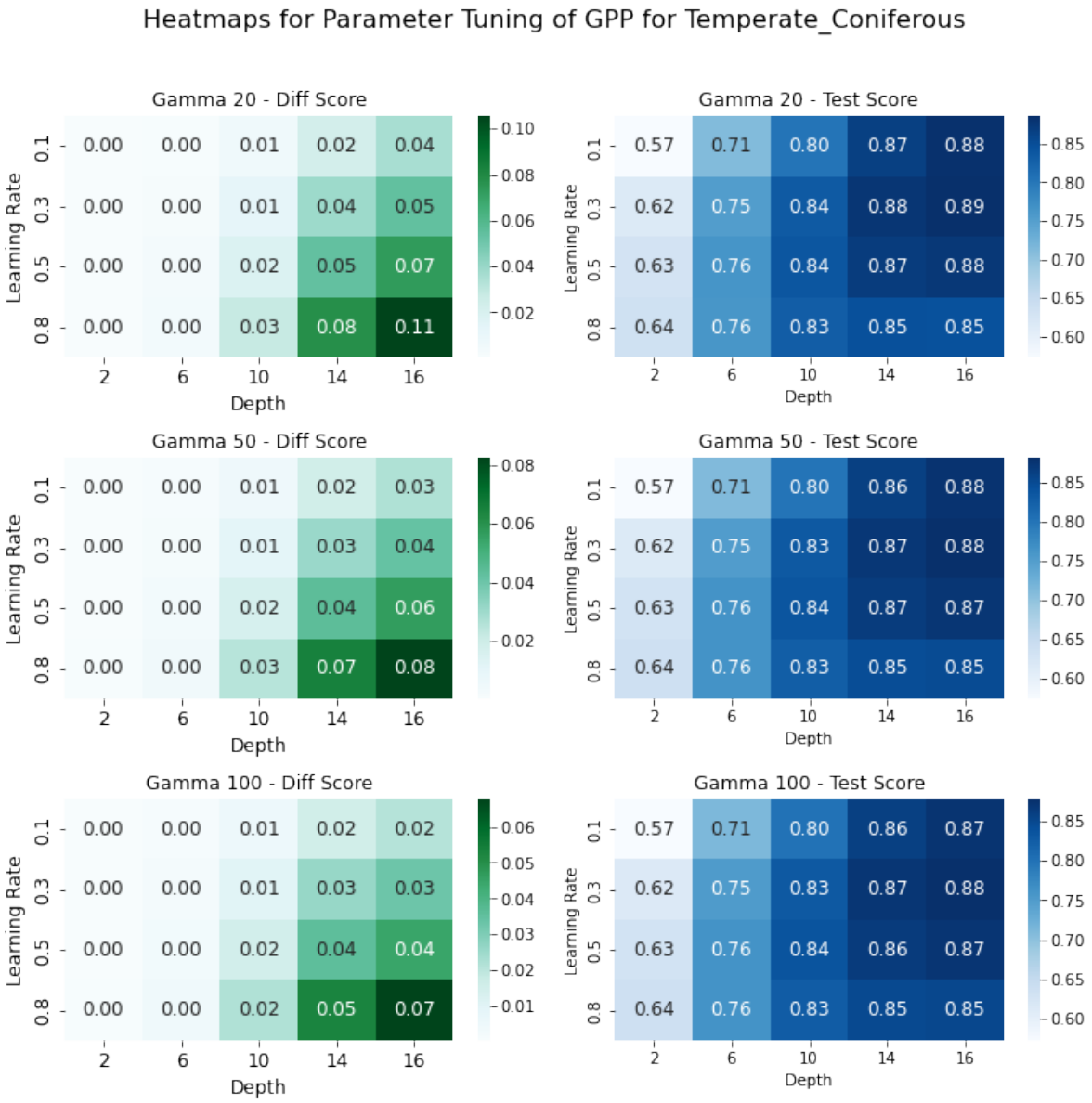
Heatmaps for Parameter Tuning of GPP for Temperate_Coniferous
Gamma 20 - Diff Score
Gamma 20 - Test Score
Gamma 50 - Diff Score
Gamma 50 - Test Score
Gamma 100 - Diff Score
Gamma 100 - Test Score
Learning Rate
Depth

Figure 20:

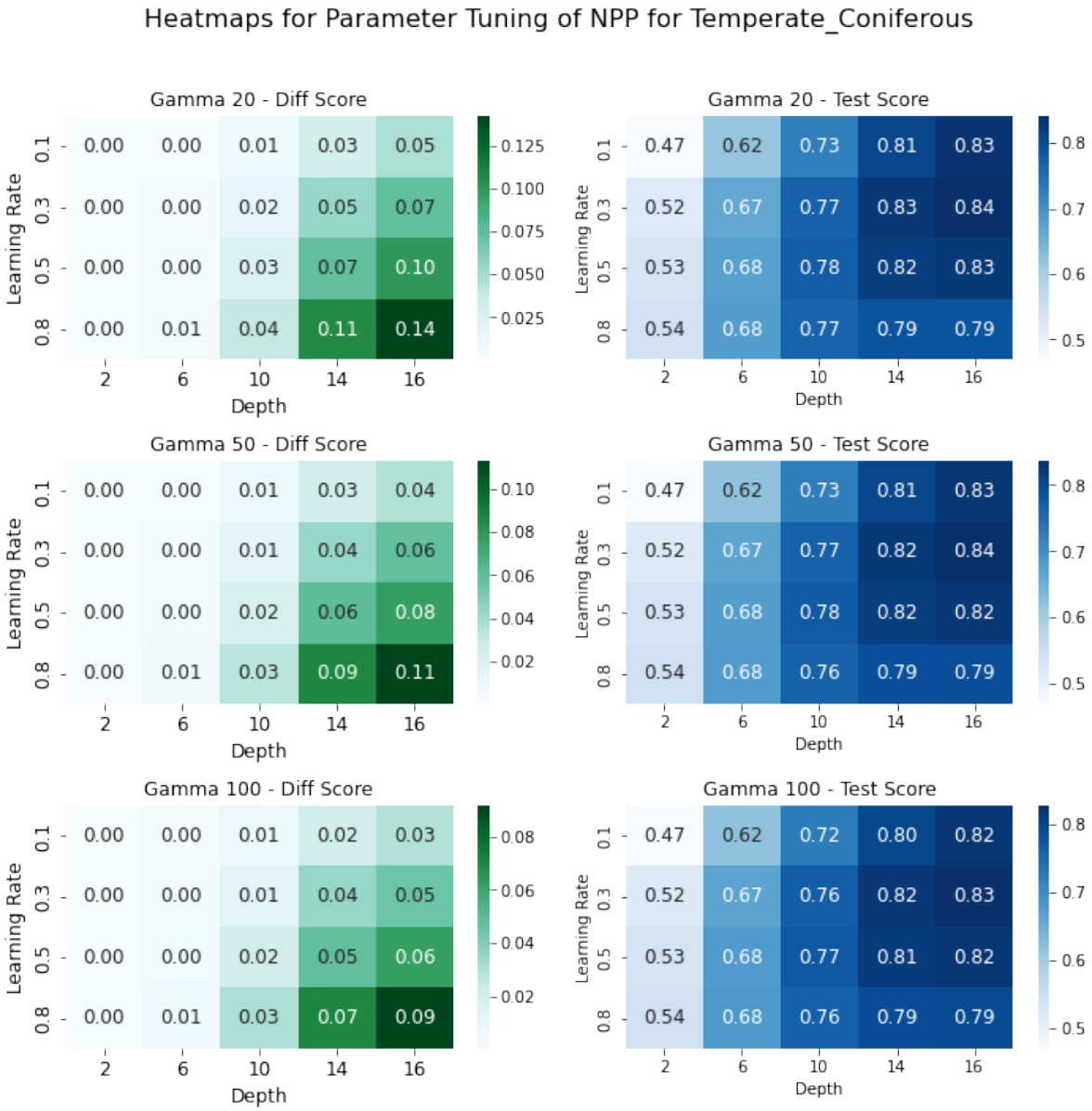
Heatmaps for Parameter Tuning of NPP for Temperate_Coniferous
Gamma 20 - Diff Score
Gamma 20 - Test Score
Gamma 50 - Diff Score
Gamma 50 - Test Score
Gamma 100 - Diff Score
Gamma 100 - Test Score
Learning Rate
Depth

Figure 21:

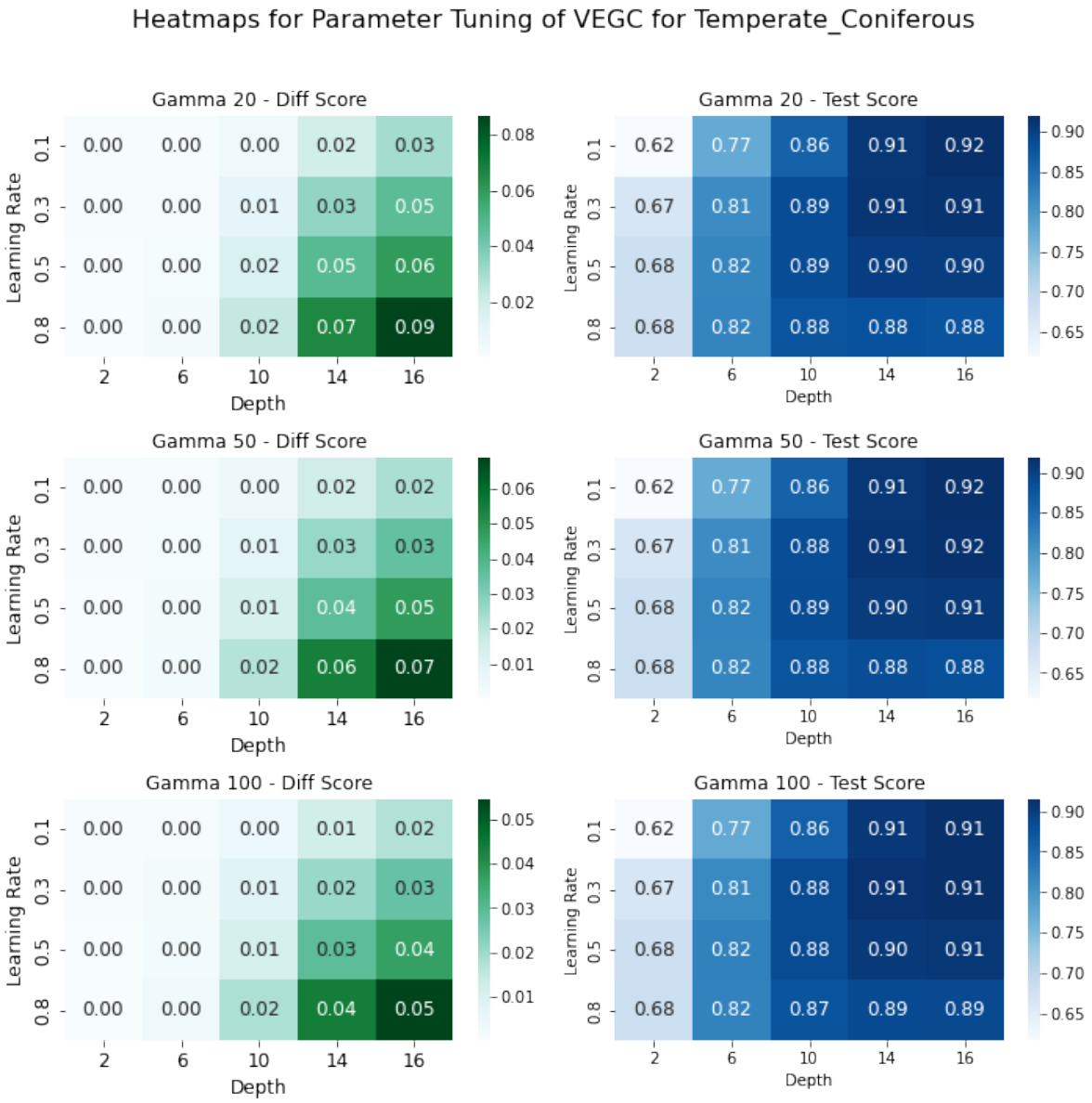
Heatmaps for Parameter Tuning of VEGC for Temperate_Coniferous
Gamma 20 - Diff Score
Learning Rate
0.1 0.00 0.00 0.00 0.02 0.03
0.3 0.00 0.00 0.01 0.03 0.05
0.5 0.00 0.00 0.02 0.05 0.06
0.8 0.00 0.00 0.02 0.07 0.09
2 6 10 14 16
Depth
0.08 0.06 0.04 0.02
Gamma 20 - Test Score
Learning Rate
0.1 0.62 0.77 0.86 0.91 0.92
0.3 0.67 0.81 0.89 0.91 0.91
0.5 0.68 0.82 0.89 0.90 0.90
0.8 0.68 0.82 0.88 0.88 0.88
2 6 10 14 16
Depth
0.90 0.85 0.80 0.75 0.70 0.65
Gamma 50 - Diff Score
Learning Rate
0.1 0.00 0.00 0.00 0.02 0.02
0.3 0.00 0.00 0.01 0.03 0.03
0.5 0.00 0.00 0.01 0.04 0.05
0.8 0.00 0.00 0.02 0.06 0.07
2 6 10 14 16
Depth
0.06 0.05 0.04 0.03 0.02 0.01
Gamma 50 - Test Score
Learning Rate
0.1 0.62 0.77 0.86 0.91 0.92
0.3 0.67 0.81 0.88 0.91 0.92
0.5 0.68 0.82 0.89 0.90 0.91
0.8 0.68 0.82 0.88 0.88 0.88
2 6 10 14 16
Depth
0.90 0.85 0.80 0.75 0.70 0.65
Gamma 100 - Diff Score
Learning Rate
0.1 0.00 0.00 0.00 0.01 0.02
0.3 0.00 0.00 0.01 0.02 0.03
0.5 0.00 0.00 0.01 0.03 0.04
0.8 0.00 0.00 0.02 0.04 0.05
2 6 10 14 16
Depth
0.05 0.04 0.03 0.02 0.01
Gamma 100 - Test Score
Learning Rate
0.1 0.62 0.77 0.86 0.91 0.91
0.3 0.67 0.81 0.88 0.91 0.91
0.5 0.68 0.82 0.88 0.90 0.91
0.8 0.68 0.82 0.87 0.89 0.89
2 6 10 14 16
Depth
0.90 0.85 0.80 0.75 0.70 0.65

Figure 22:

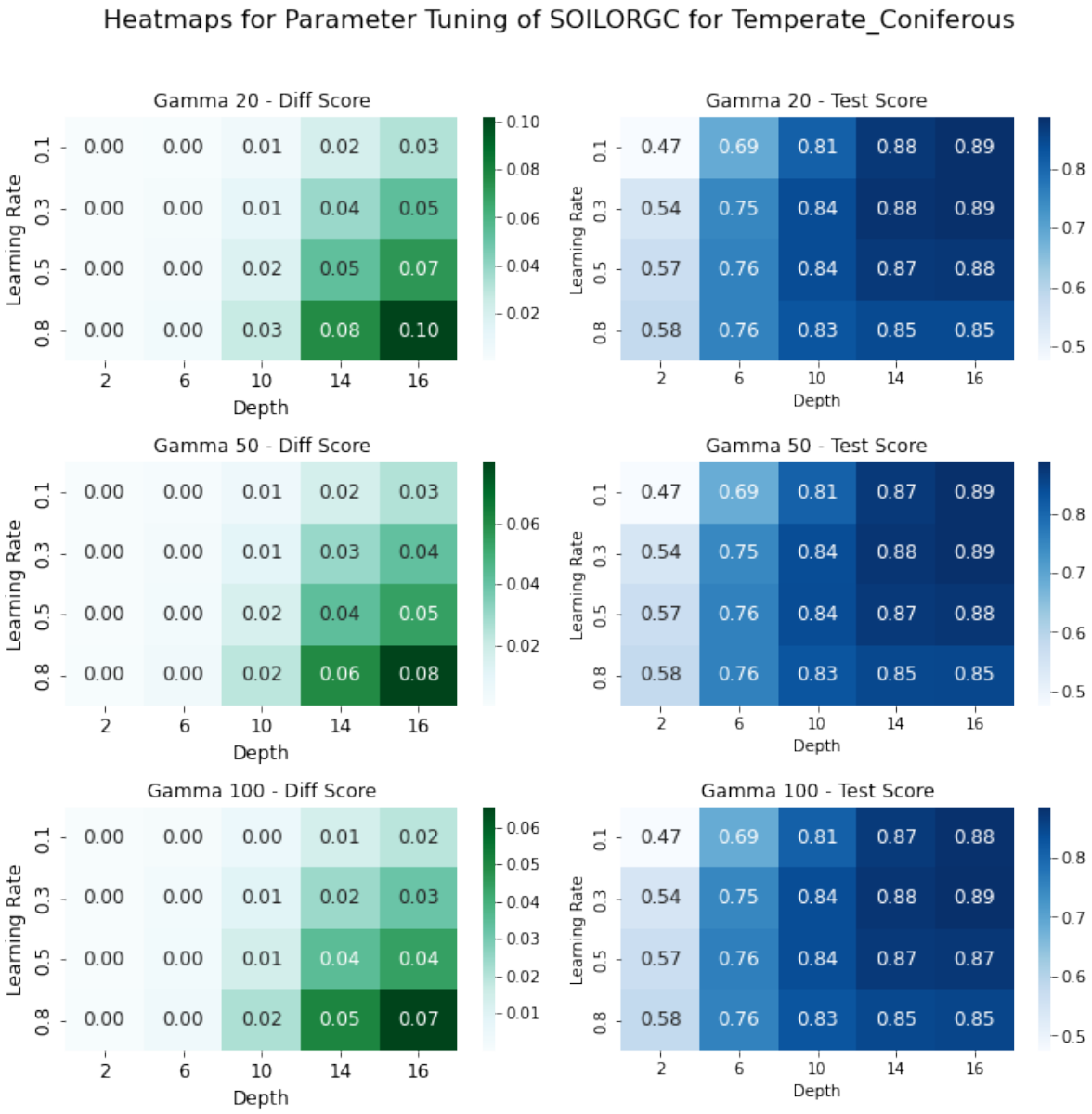
Heatmaps for Parameter Tuning of SOILORGC for Temperate_Coniferous
Gamma 20 - Diff Score
Gamma 20 - Test Score
Gamma 50 - Diff Score
Gamma 50 - Test Score
Gamma 100 - Diff Score
Gamma 100 - Test Score
Learning Rate
Depth

Figure 23:

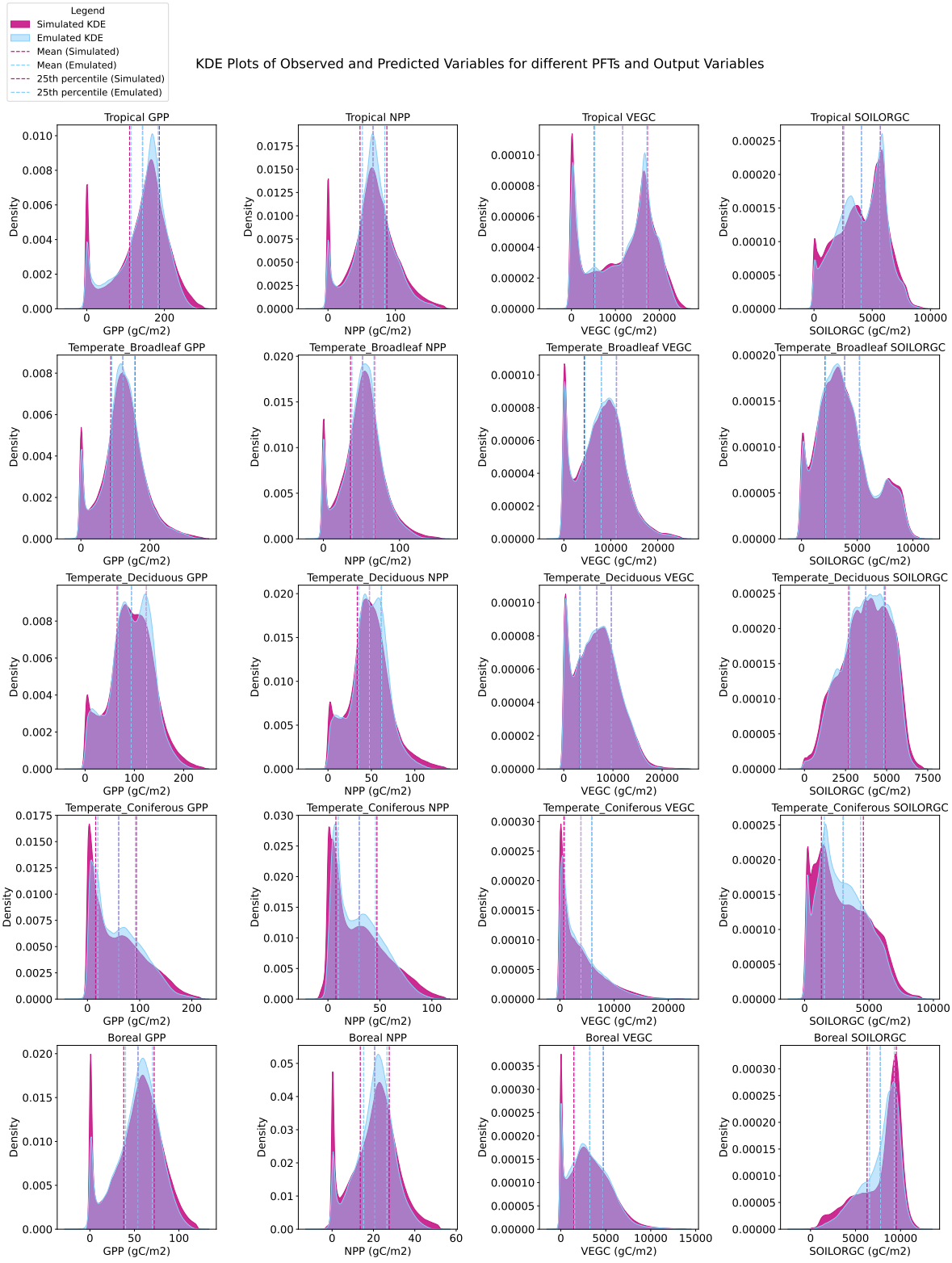
Legend
Simulated KDE
Emulated KDE
Mean (Simulated)
Mean (Emulated)
25th percentile (Simulated)
25th percentile (Emulated)
KDE Plots of Observed and Predicted Variables for different PFTs and Output Variables
Tropical GPP
Tropical NPP
Tropical VEGC
Tropical SOILORGC
Temperate_Broadleaf GPP
Temperate_Broadleaf NPP
Temperate_Broadleaf VEGC
Temperate_Broadleaf SOILORGC
Temperate_Deciduous GPP
Temperate_Deciduous NPP
Temperate_Deciduous VEGC
Temperate_Deciduous SOILORGC
Temperate_Coniferous GPP
Temperate_Coniferous NPP
Temperate_Coniferous VEGC
Temperate_Coniferous SOILORGC
Boreal GPP
Boreal NPP
Boreal VEGC
Boreal SOILORGC
Density
GPP (gC/m2)
NPP (gC/m2)
VEGC (gC/m2)
SOILORGC (gC/m2)

Figure 24:

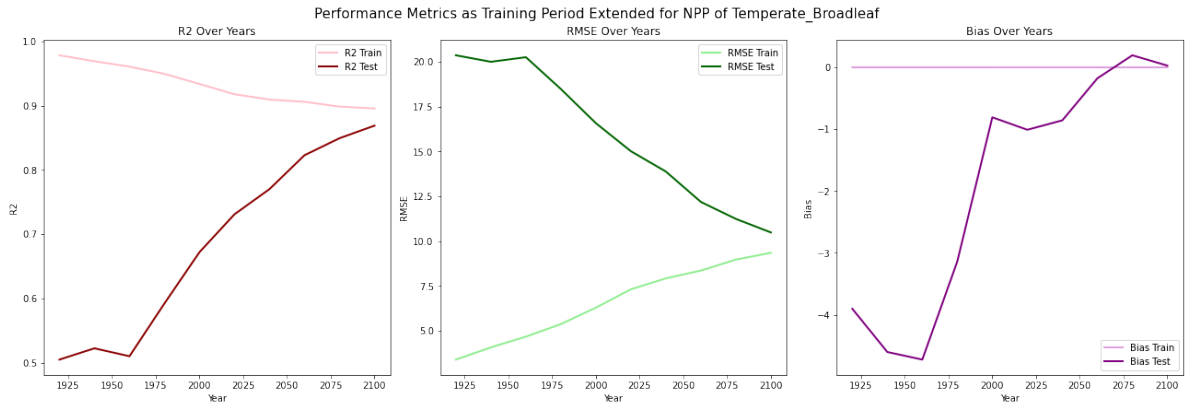

Figure 25:

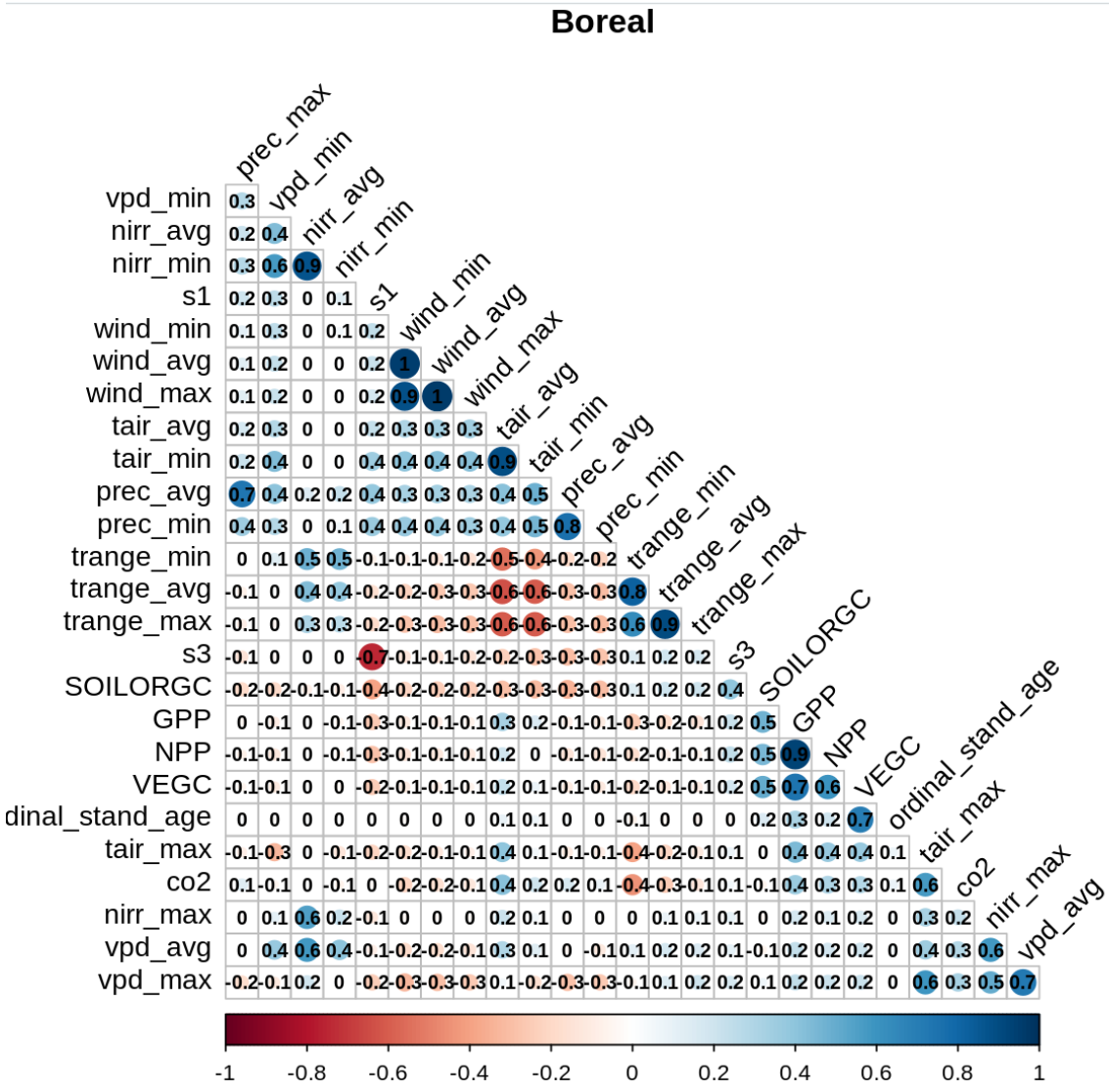

Figure 26:

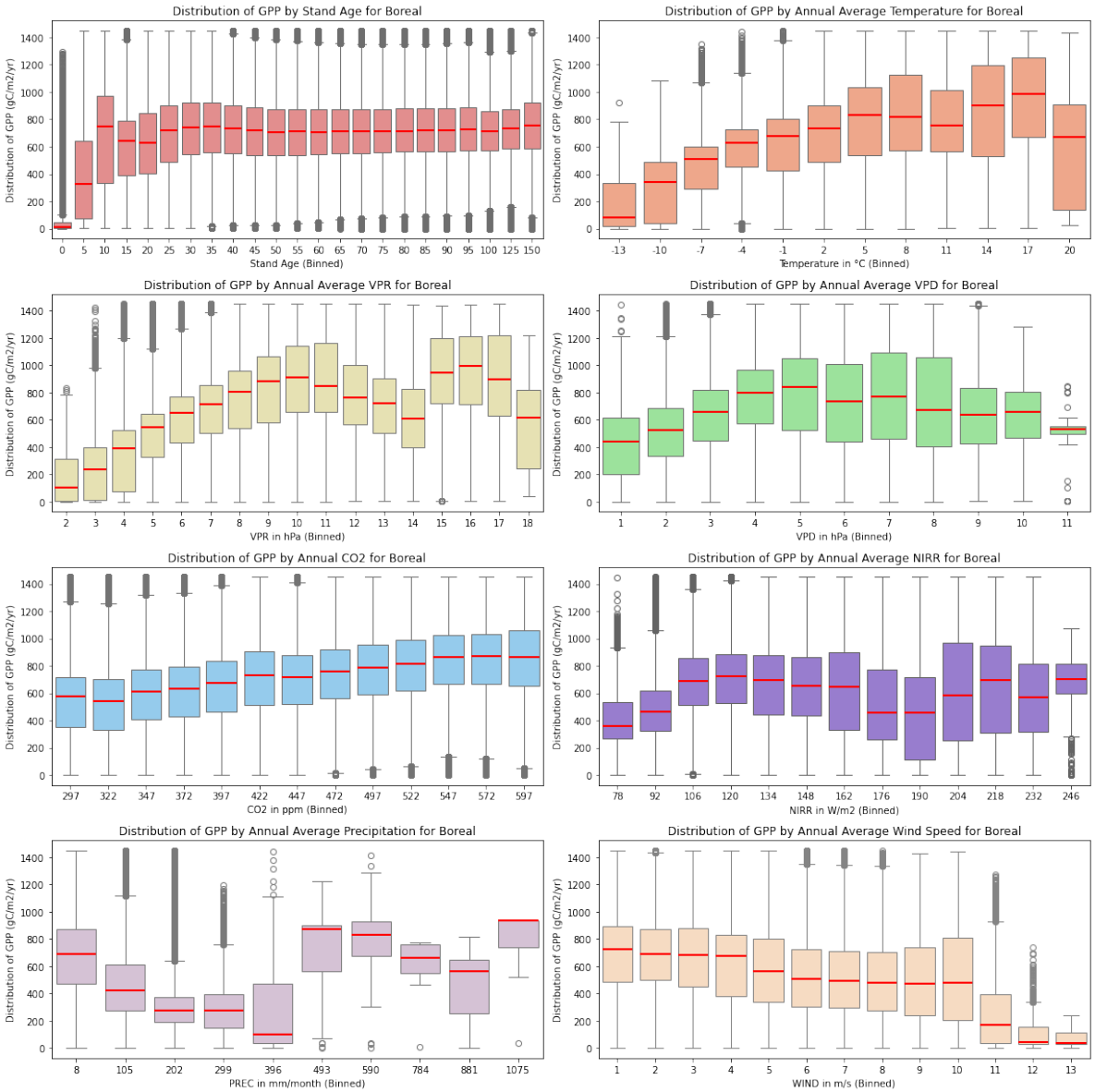
Distribution of GPP by Stand Age for Boreal
Distribution of GPP (gC/m2/yr)
Stand Age (Binned)
Distribution of GPP by Annual Average Temperature for Boreal
Temperature in °C (Binned)
Distribution of GPP by Annual Average VPR for Boreal
VPR in hPa (Binned)
Distribution of GPP by Annual Average VPD for Boreal
VPD in hPa (Binned)
Distribution of GPP by Annual CO2 for Boreal
CO2 in ppm (Binned)
Distribution of GPP by Annual Average NIRR for Boreal
NIRR in W/m2 (Binned)
Distribution of GPP by Annual Average Precipitation for Boreal
PREC in mm/month (Binned)
Distribution of GPP by Annual Average Wind Speed for Boreal
WIND in m/s (Binned)

Figure 27:

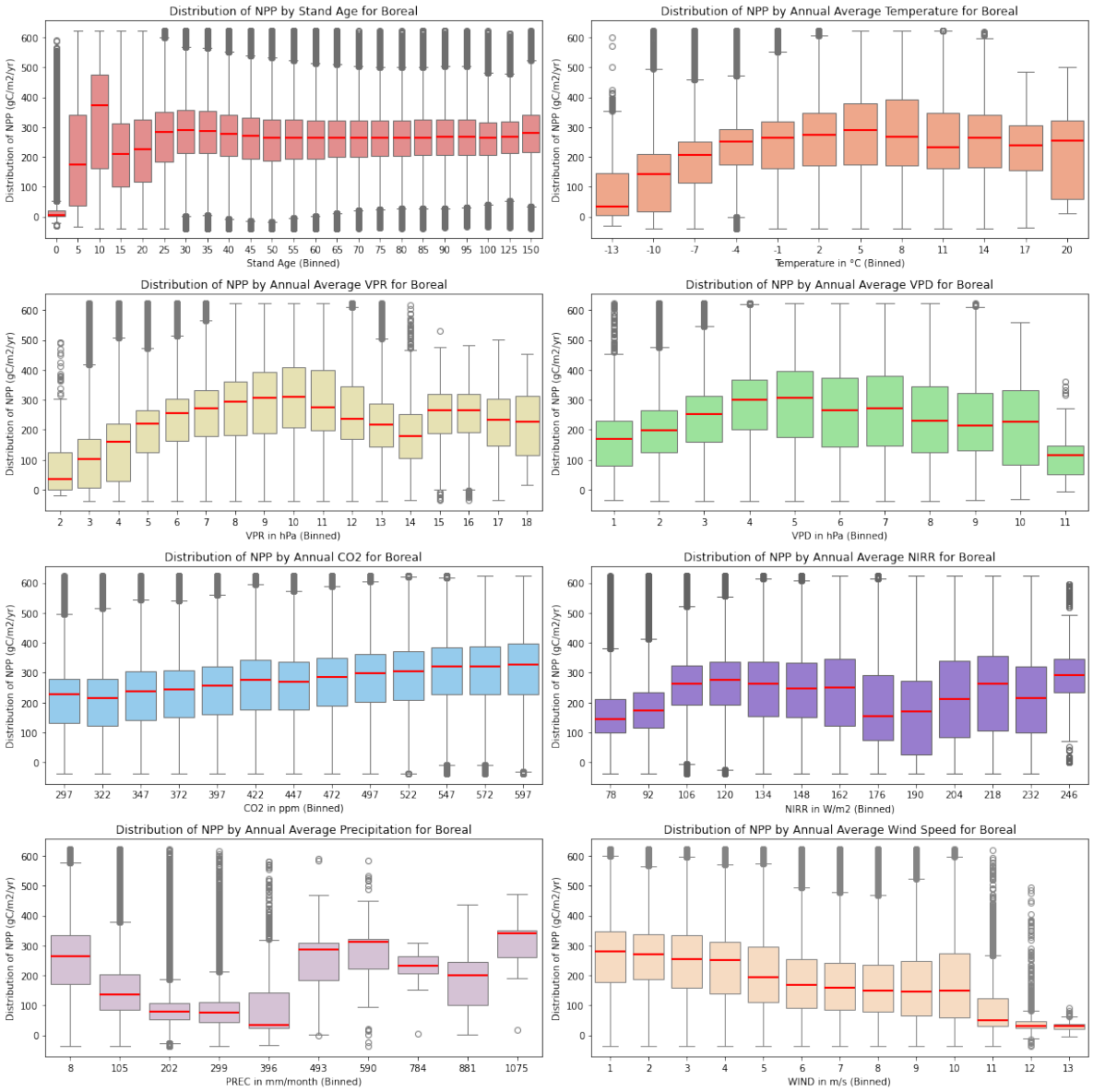
Distribution of NPP by Stand Age for Boreal
Distribution of NPP (gC/m2/yr)
Stand Age (Binned)
Distribution of NPP by Annual Average Temperature for Boreal
Temperature in °C (Binned)
Distribution of NPP by Annual Average VPR for Boreal
VPR in hPa (Binned)
Distribution of NPP by Annual Average VPD for Boreal
VPD in hPa (Binned)
Distribution of NPP by Annual CO2 for Boreal
CO2 in ppm (Binned)
Distribution of NPP by Annual Average NIRR for Boreal
NIRR in W/m2 (Binned)
Distribution of NPP by Annual Average Precipitation for Boreal
PREC in mm/month (Binned)
Distribution of NPP by Annual Average Wind Speed for Boreal
WIND in m/s (Binned)

Figure 28:

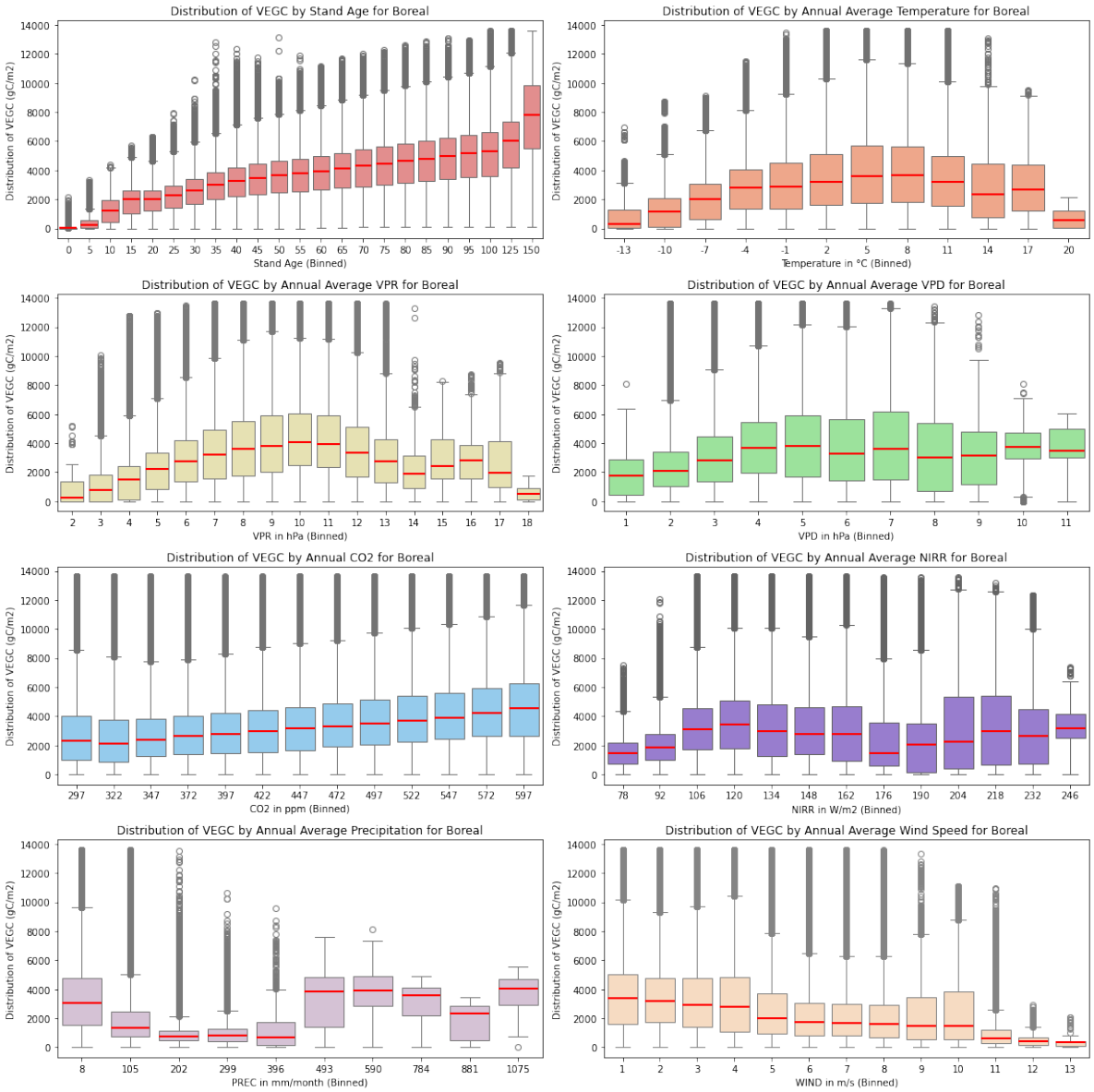
Distribution of VEGC by Stand Age for Boreal
Distribution of VEGC (gC/m2)
Stand Age (Binned)
Distribution of VEGC by Annual Average Temperature for Boreal
Temperature in °C (Binned)
Distribution of VEGC by Annual Average VPR for Boreal
VPR in hPa (Binned)
Distribution of VEGC by Annual Average VPD for Boreal
VPD in hPa (Binned)
Distribution of VEGC by Annual CO2 for Boreal
CO2 in ppm (Binned)
Distribution of VEGC by Annual Average NIRR for Boreal
NIRR in W/m2 (Binned)
Distribution of VEGC by Annual Average Precipitation for Boreal
PREC in mm/month (Binned)
Distribution of VEGC by Annual Average Wind Speed for Boreal
WIND in m/s (Binned)

Figure 29:

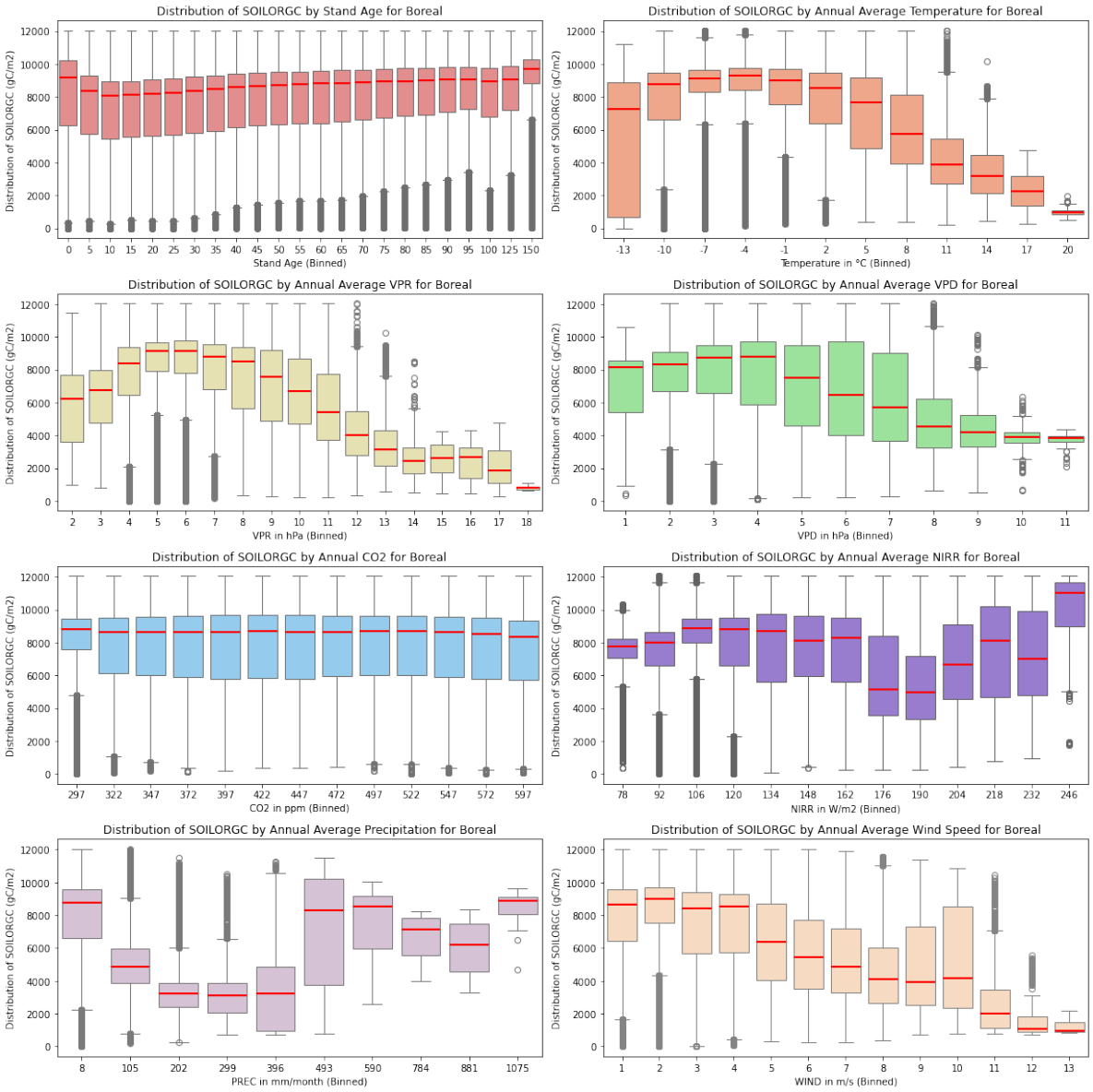
Distribution of SOILORGC by Stand Age for Boreal
Stand Age (Binned)
Distribution of SOILORGC by Annual Average Temperature for Boreal
Temperature in °C (Binned)
Distribution of SOILORGC by Annual Average VPR for Boreal
VPR in hPa (Binned)
Distribution of SOILORGC by Annual Average VPD for Boreal
VPD in hPa (Binned)
Distribution of SOILORGC by Annual CO2 for Boreal
CO2 in ppm (Binned)
Distribution of SOILORGC by Annual Average NIRR for Boreal
NIRR in W/m2 (Binned)
Distribution of SOILORGC by Annual Average Precipitation for Boreal
PREC in mm/month (Binned)
Distribution of SOILORGC by Annual Average Wind Speed for Boreal
WIND in m/s (Binned)
Distribution of SOILORGC (gC/m2)

Figure 30:

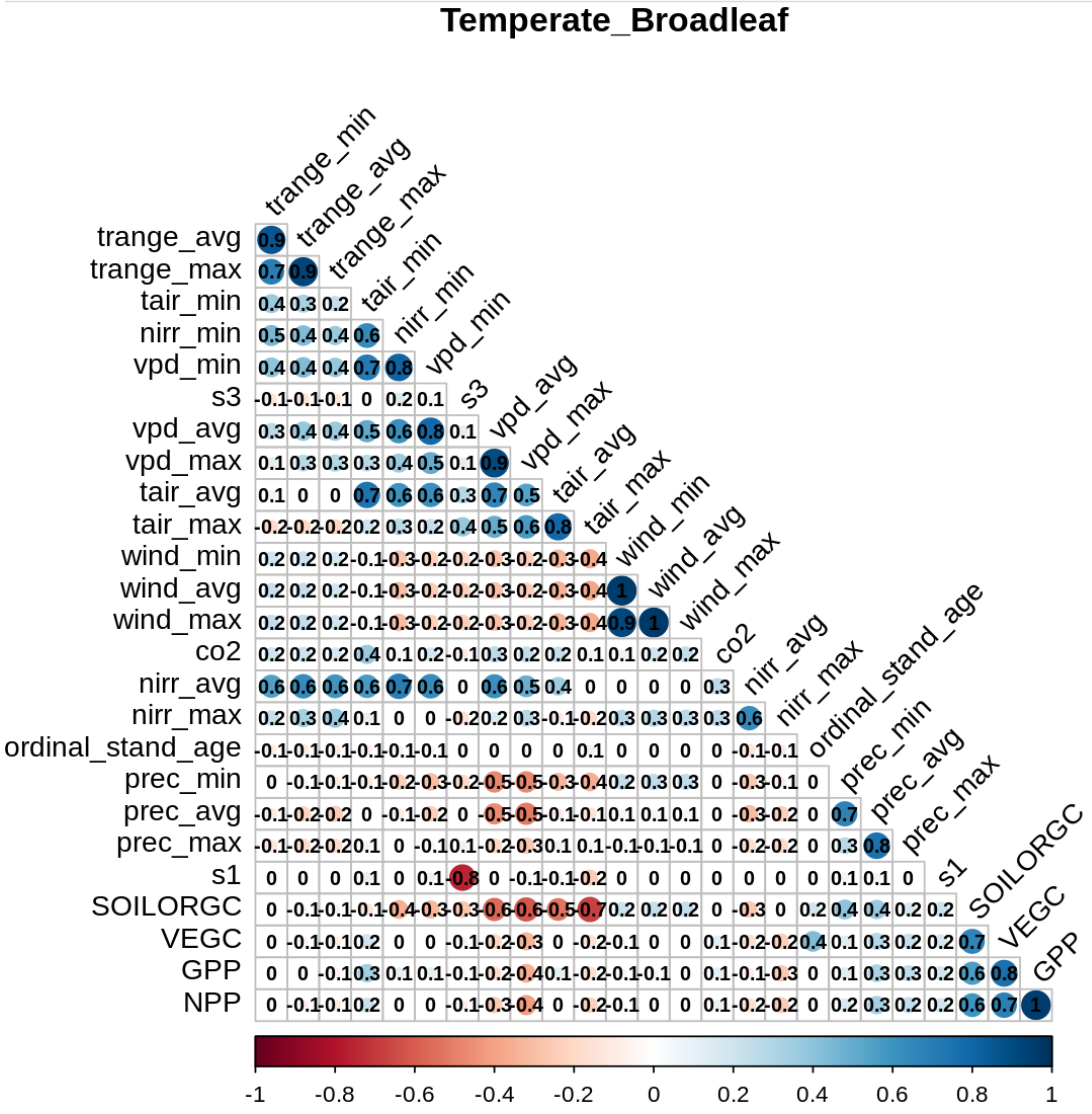
Temperate_Broadleaf
trange_min
trange_avg
trange_max
tair_min
nirr_min
vpd_min
s3
vpd_avg
vpd_max
tair_avg
tair_max
wind_min
wind_avg
wind_max
co2
nirr_avg
nirr_max
ordinal_stand_age
prec_min
prec_avg
prec_max
s1
SOILORGC
VEGC
GPP
NPP
-1
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8
1

Figure 31:

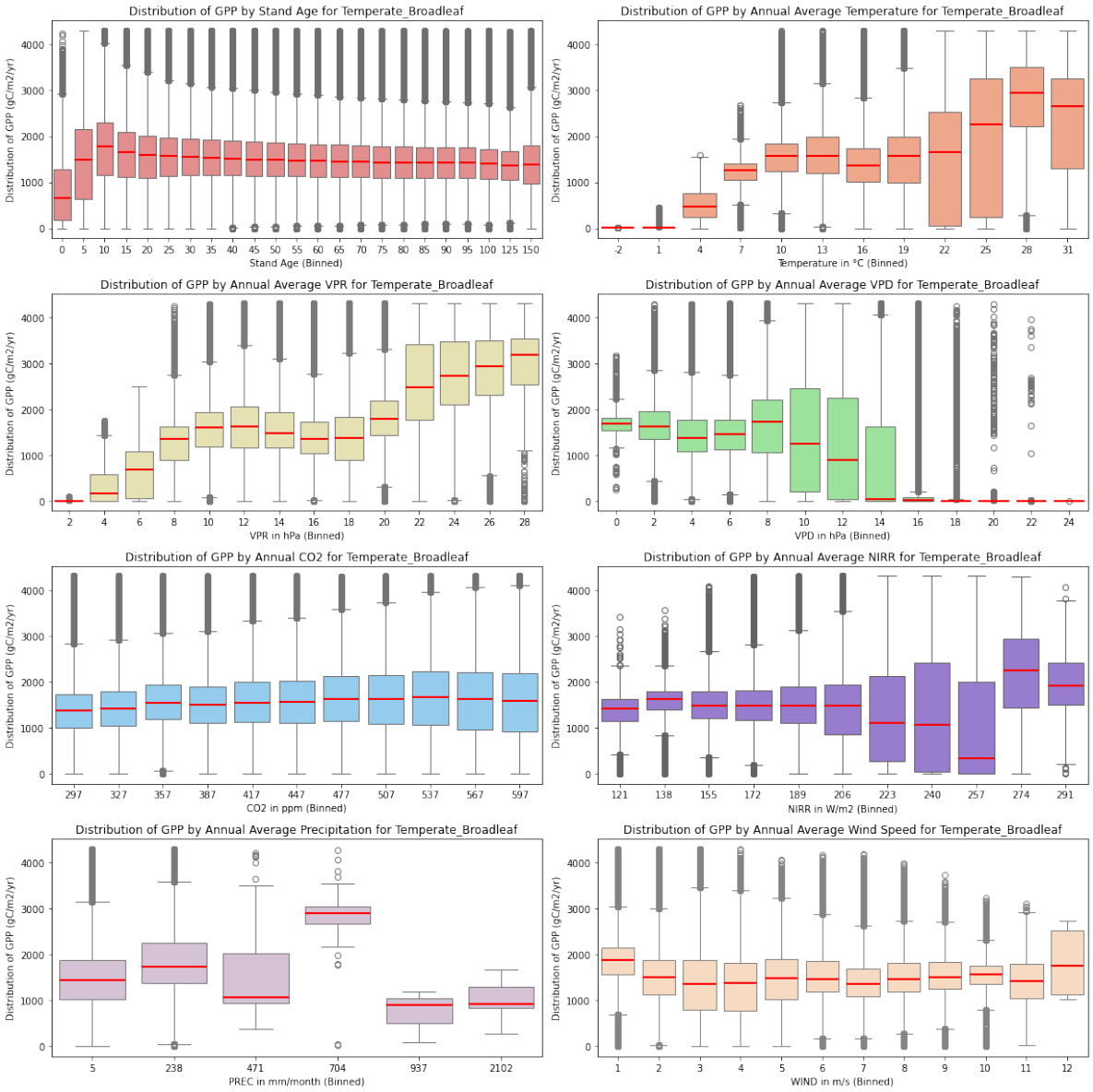
Distribution of GPP by Stand Age for Temperate_Broadleaf
Distribution of GPP (gC/m2/yr)
Stand Age (Binned)
Distribution of GPP by Annual Average Temperature for Temperate_Broadleaf
Temperature in °C (Binned)
Distribution of GPP by Annual Average VPR for Temperate_Broadleaf
VPR in hPa (Binned)
Distribution of GPP by Annual Average VPD for Temperate_Broadleaf
VPD in hPa (Binned)
Distribution of GPP by Annual CO2 for Temperate_Broadleaf
CO2 in ppm (Binned)
Distribution of GPP by Annual Average NIRR for Temperate_Broadleaf
NIRR in W/m2 (Binned)
Distribution of GPP by Annual Average Precipitation for Temperate_Broadleaf
PREC in mm/month (Binned)
Distribution of GPP by Annual Average Wind Speed for Temperate_Broadleaf
WIND in m/s (Binned)

Figure 32:

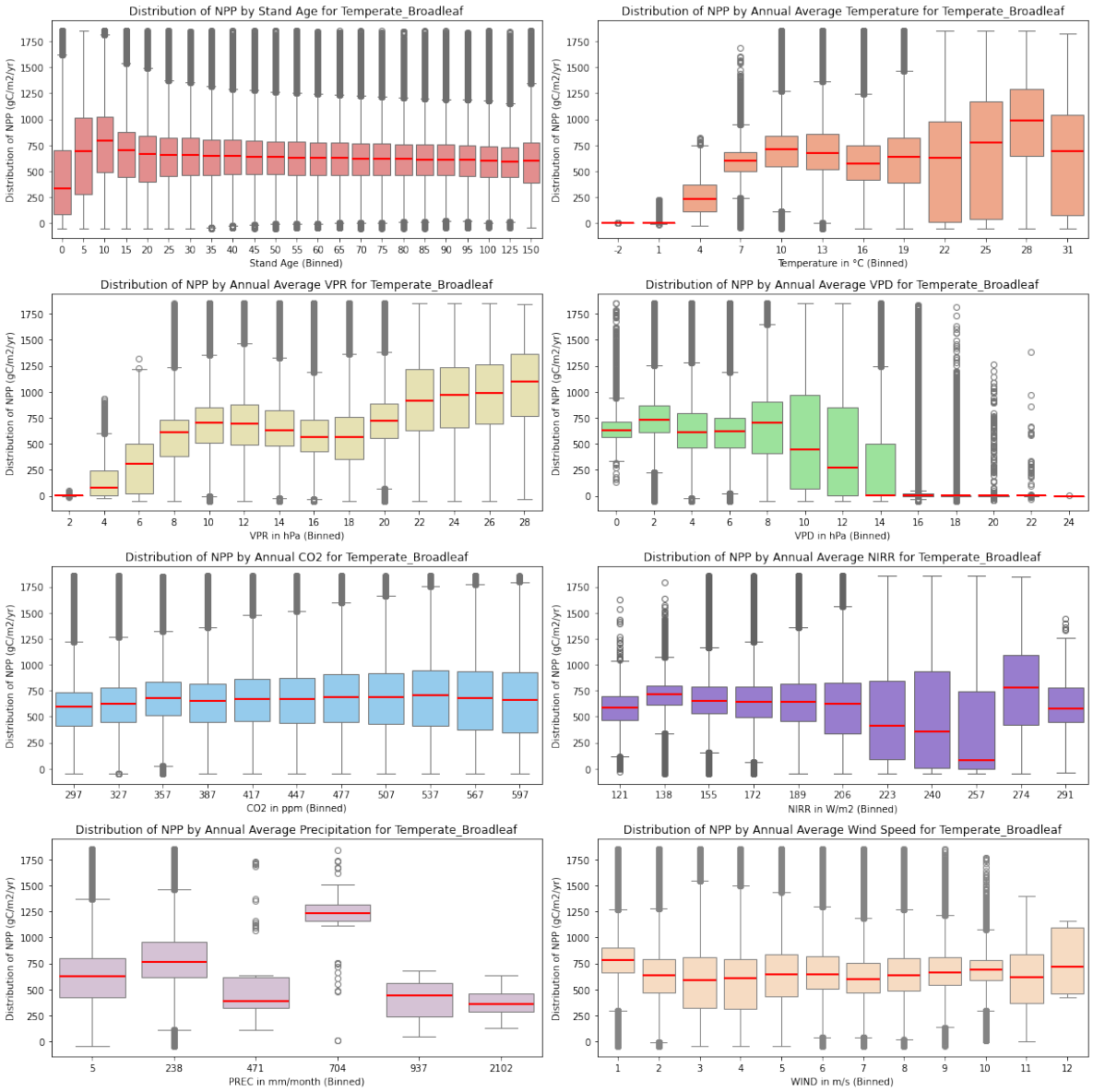
Distribution of NPP by Stand Age for Temperate_Broadleaf
Distribution of NPP (gC/m2/yr)
Stand Age (Binned)
Distribution of NPP by Annual Average Temperature for Temperate_Broadleaf
Distribution of NPP (gC/m2/yr)
Temperature in °C (Binned)
Distribution of NPP by Annual Average VPR for Temperate_Broadleaf
Distribution of NPP (gC/m2/yr)
VPR in hPa (Binned)
Distribution of NPP by Annual Average VPD for Temperate_Broadleaf
Distribution of NPP (gC/m2/yr)
VPD in hPa (Binned)
Distribution of NPP by Annual CO2 for Temperate_Broadleaf
Distribution of NPP (gC/m2/yr)
CO2 in ppm (Binned)
Distribution of NPP by Annual Average NIRR for Temperate_Broadleaf
Distribution of NPP (gC/m2/yr)
NIRR in W/m2 (Binned)
Distribution of NPP by Annual Average Precipitation for Temperate_Broadleaf
Distribution of NPP (gC/m2/yr)
PREC in mm/month (Binned)
Distribution of NPP by Annual Average Wind Speed for Temperate_Broadleaf
Distribution of NPP (gC/m2/yr)
WIND in m/s (Binned)

Figure 33:

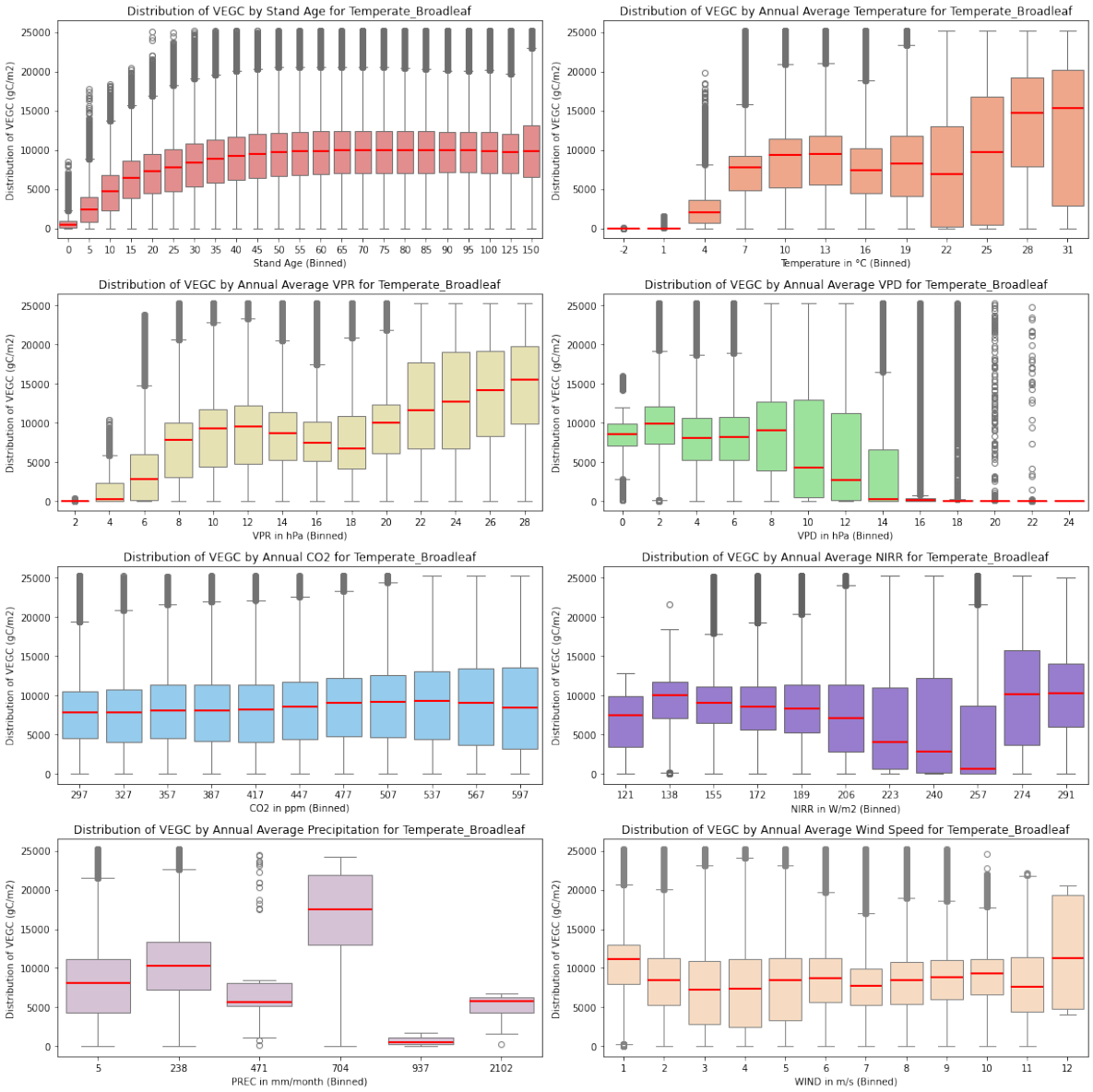
Distribution of VEGC by Stand Age for Temperate_Broadleaf
Distribution of VEGC (gC/m2)
Stand Age (Binned)
Distribution of VEGC by Annual Average Temperature for Temperate_Broadleaf
Distribution of VEGC (gC/m2)
Temperature in °C (Binned)
Distribution of VEGC by Annual Average VPR for Temperate_Broadleaf
Distribution of VEGC (gC/m2)
VPR in hPa (Binned)
Distribution of VEGC by Annual Average VPD for Temperate_Broadleaf
Distribution of VEGC (gC/m2)
VPD in hPa (Binned)
Distribution of VEGC by Annual CO2 for Temperate_Broadleaf
Distribution of VEGC (gC/m2)
CO2 in ppm (Binned)
Distribution of VEGC by Annual Average NIRR for Temperate_Broadleaf
Distribution of VEGC (gC/m2)
NIRR in W/m2 (Binned)
Distribution of VEGC by Annual Average Precipitation for Temperate_Broadleaf
Distribution of VEGC (gC/m2)
PREC in mm/month (Binned)
Distribution of VEGC by Annual Average Wind Speed for Temperate_Broadleaf
Distribution of VEGC (gC/m2)
WIND in m/s (Binned)

Figure 34:

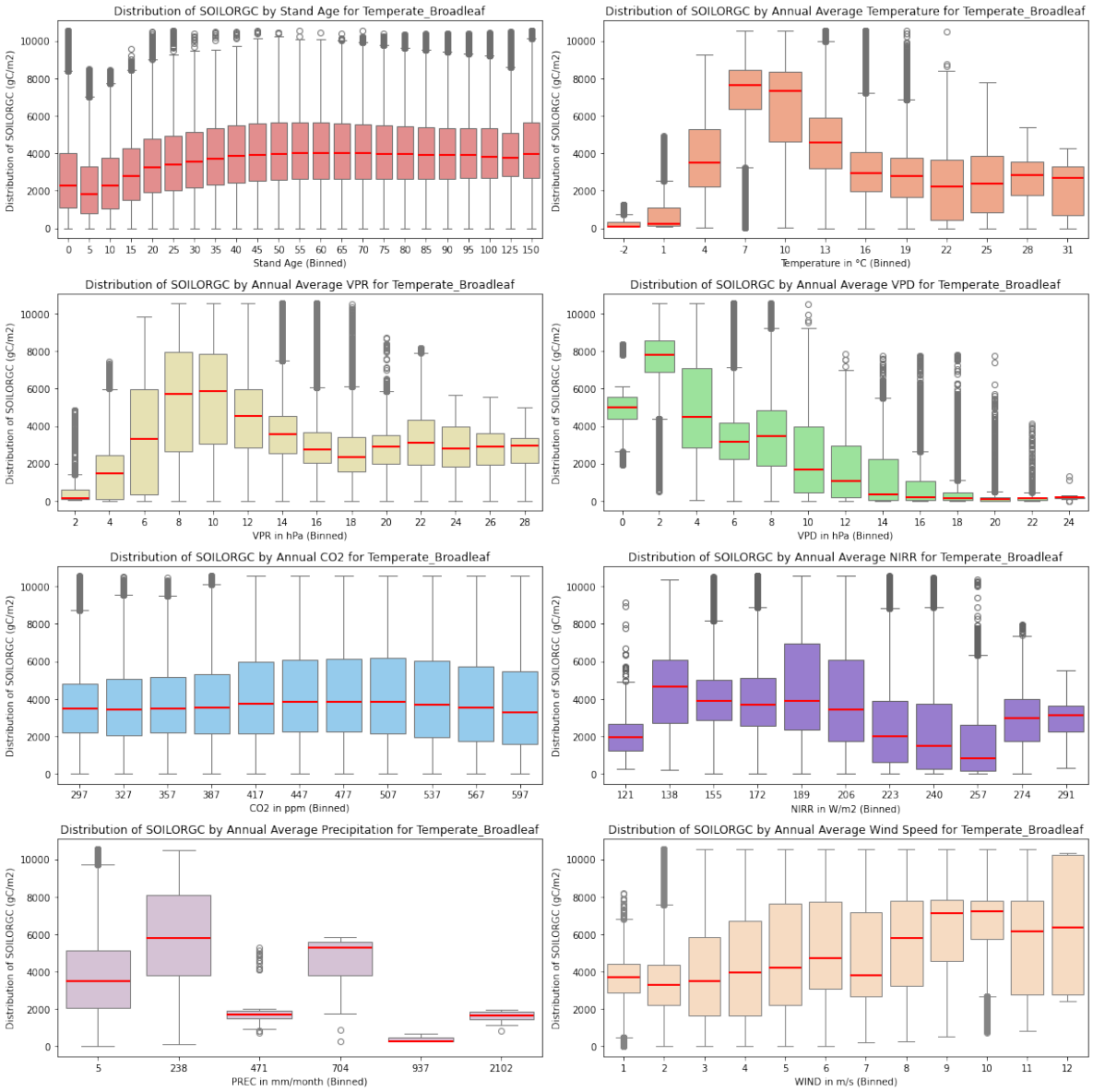
Distribution of SOILORGC by Stand Age for Temperate_Broadleaf
Distribution of SOILORGC (gC/m2)
Stand Age (Binned)
Distribution of SOILORGC by Annual Average Temperature for Temperate_Broadleaf
Temperature in °C (Binned)
Distribution of SOILORGC by Annual Average VPR for Temperate_Broadleaf
VPR in hPa (Binned)
Distribution of SOILORGC by Annual Average VPD for Temperate_Broadleaf
VPD in hPa (Binned)
Distribution of SOILORGC by Annual CO2 for Temperate_Broadleaf
CO2 in ppm (Binned)
Distribution of SOILORGC by Annual Average NIRR for Temperate_Broadleaf
NIRR in W/m2 (Binned)
Distribution of SOILORGC by Annual Average Precipitation for Temperate_Broadleaf
PREC in mm/month (Binned)
Distribution of SOILORGC by Annual Average Wind Speed for Temperate_Broadleaf
WIND in m/s (Binned)

Figure 35:

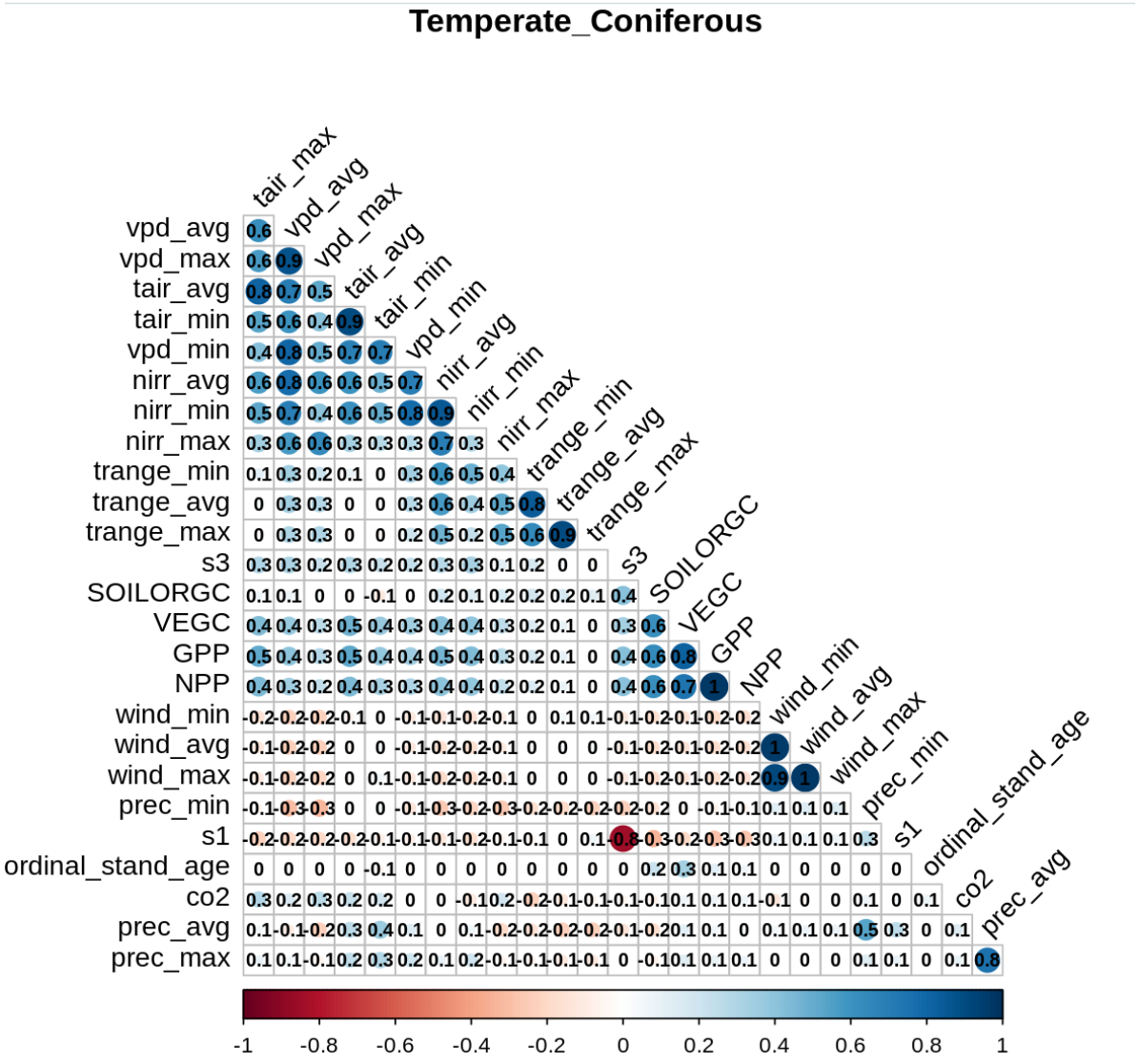
Temperate_Coniferous
tair_max
vpd_avg
vpd_max
tair_avg
tair_min
vpd_min
nirr_avg
nirr_min
nirr_max
trange_min
trange_avg
trange_max
s3
SOILORGC
VEGC
GPP
NPP
wind_min
wind_avg
wind_max
prec_min
s1
ordinal_stand_age
co2
prec_avg
prec_max
-1
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8
1

Figure 36:

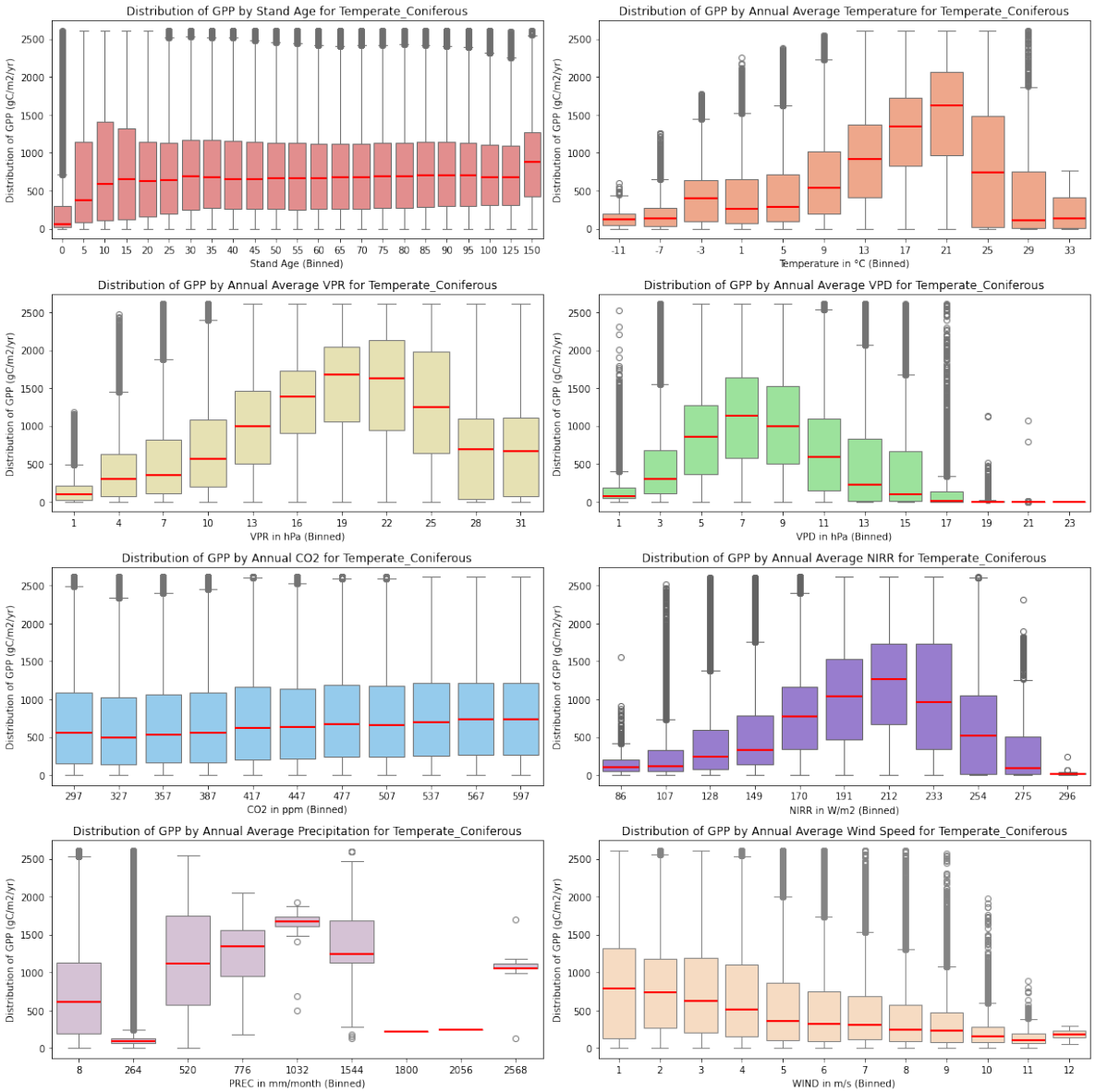
Distribution of GPP by Stand Age for Temperate_Coniferous
Distribution of GPP (gC/m2/yr)
Stand Age (Binned)
Distribution of GPP by Annual Average Temperature for Temperate_Coniferous
Temperature in °C (Binned)
Distribution of GPP by Annual Average VPR for Temperate_Coniferous
VPR in hPa (Binned)
Distribution of GPP by Annual Average VPD for Temperate_Coniferous
VPD in hPa (Binned)
Distribution of GPP by Annual CO2 for Temperate_Coniferous
CO2 in ppm (Binned)
Distribution of GPP by Annual Average NIRR for Temperate_Coniferous
NIRR in W/m2 (Binned)
Distribution of GPP by Annual Average Precipitation for Temperate_Coniferous
PREC in mm/month (Binned)
Distribution of GPP by Annual Average Wind Speed for Temperate_Coniferous
WIND in m/s (Binned)

Figure 37:

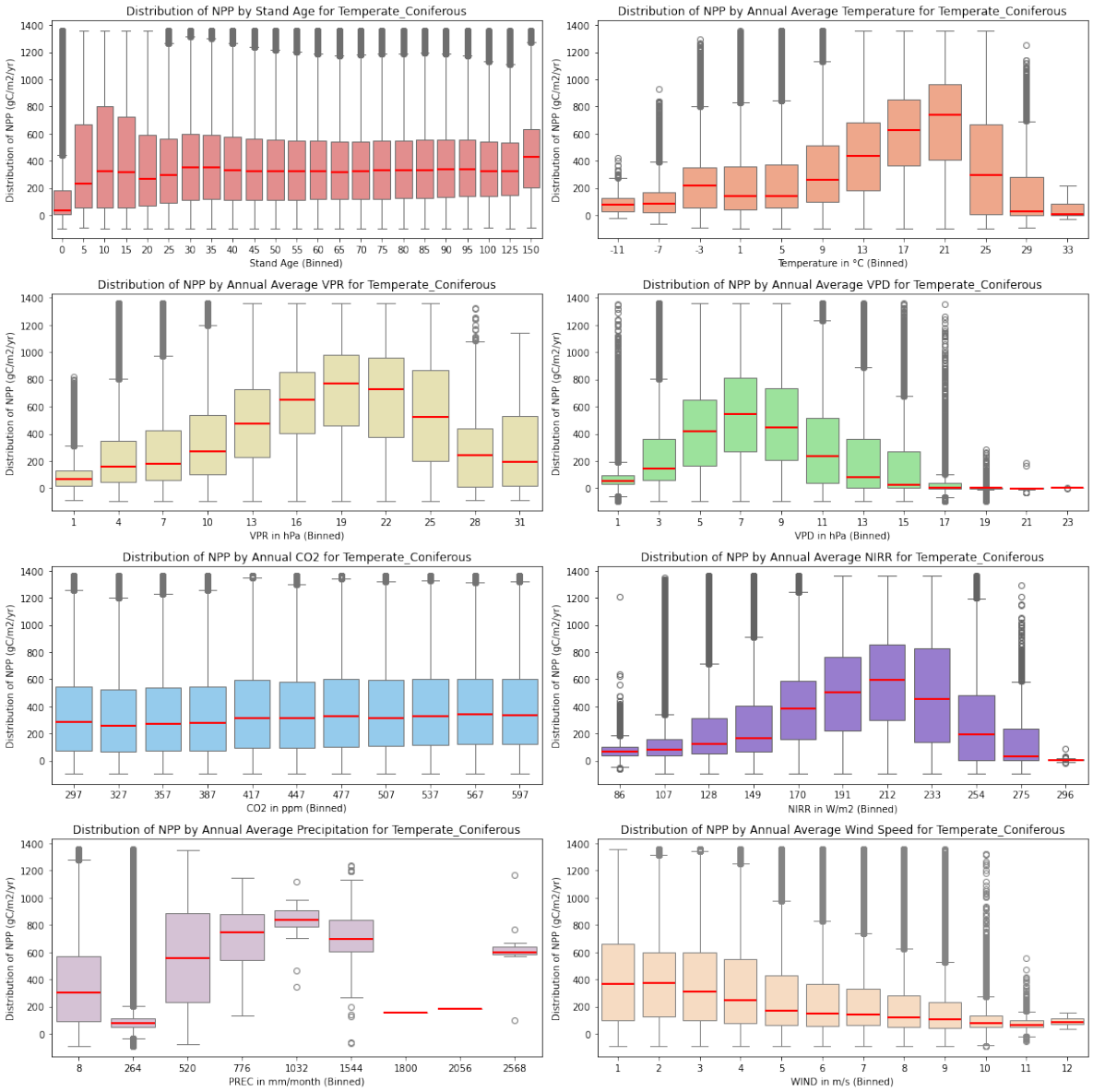
Distribution of NPP by Stand Age for Temperate_Coniferous
Distribution of NPP (gC/m2/yr)
Stand Age (Binned)
Distribution of NPP by Annual Average Temperature for Temperate_Coniferous
Distribution of NPP (gC/m2/yr)
Temperature in °C (Binned)
Distribution of NPP by Annual Average VPR for Temperate_Coniferous
Distribution of NPP (gC/m2/yr)
VPR in hPa (Binned)
Distribution of NPP by Annual Average VPD for Temperate_Coniferous
Distribution of NPP (gC/m2/yr)
VPD in hPa (Binned)
Distribution of NPP by Annual CO2 for Temperate_Coniferous
Distribution of NPP (gC/m2/yr)
CO2 in ppm (Binned)
Distribution of NPP by Annual Average NIRR for Temperate_Coniferous
Distribution of NPP (gC/m2/yr)
NIRR in W/m2 (Binned)
Distribution of NPP by Annual Average Precipitation for Temperate_Coniferous
Distribution of NPP (gC/m2/yr)
PREC in mm/month (Binned)
Distribution of NPP by Annual Average Wind Speed for Temperate_Coniferous
Distribution of NPP (gC/m2/yr)
WIND in m/s (Binned)

Figure 38:

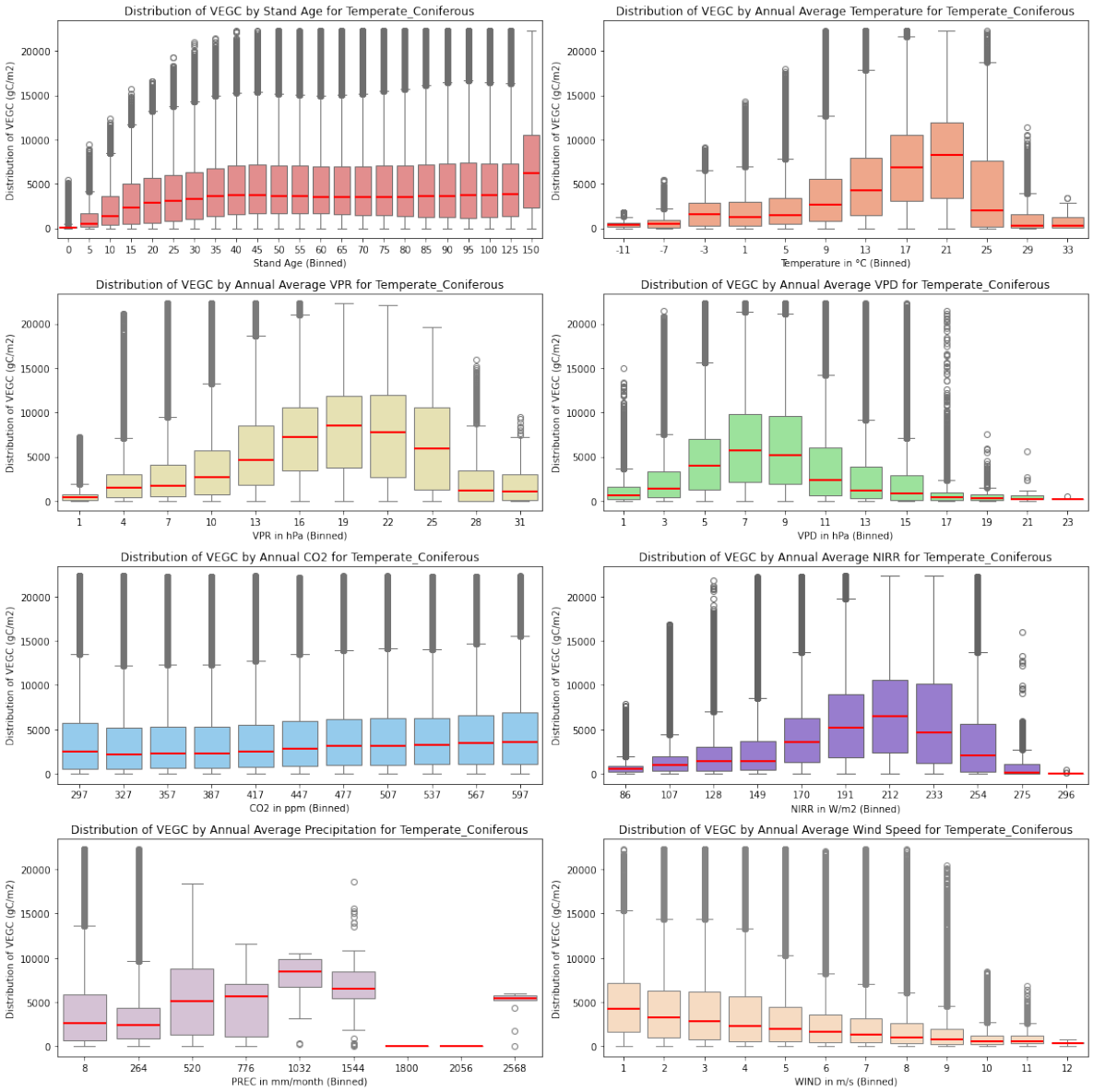
Distribution of VEGC by Stand Age for Temperate_Coniferous
Distribution of VEGC (gC/m2)
Stand Age (Binned)
Distribution of VEGC by Annual Average Temperature for Temperate_Coniferous
Temperature in °C (Binned)
Distribution of VEGC by Annual Average VPR for Temperate_Coniferous
VPR in hPa (Binned)
Distribution of VEGC by Annual Average VPD for Temperate_Coniferous
VPD in hPa (Binned)
Distribution of VEGC by Annual CO2 for Temperate_Coniferous
CO2 in ppm (Binned)
Distribution of VEGC by Annual Average NIRR for Temperate_Coniferous
NIRR in W/m2 (Binned)
Distribution of VEGC by Annual Average Precipitation for Temperate_Coniferous
PREC in mm/month (Binned)
Distribution of VEGC by Annual Average Wind Speed for Temperate_Coniferous
WIND in m/s (Binned)

Figure 39:

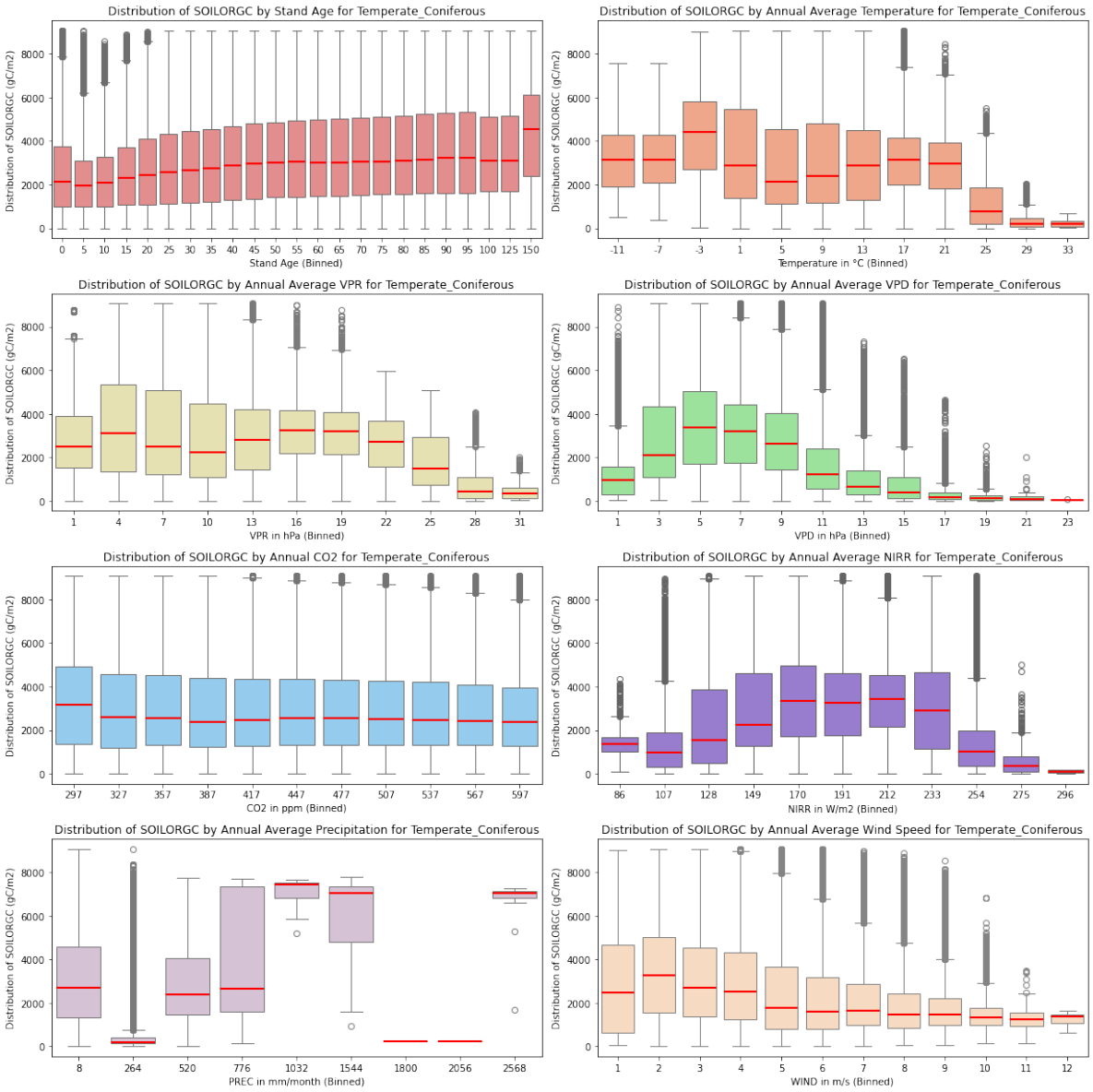
Distribution of SOILORGC by Stand Age for Temperate_Coniferous
Distribution of SOILORGC (gC/m2)
Stand Age (Binned)
Distribution of SOILORGC by Annual Average Temperature for Temperate_Coniferous
Temperature in °C (Binned)
Distribution of SOILORGC by Annual Average VPR for Temperate_Coniferous
VPR in hPa (Binned)
Distribution of SOILORGC by Annual Average VPD for Temperate_Coniferous
VPD in hPa (Binned)
Distribution of SOILORGC by Annual CO2 for Temperate_Coniferous
CO2 in ppm (Binned)
Distribution of SOILORGC by Annual Average NIRR for Temperate_Coniferous
NIRR in W/m2 (Binned)
Distribution of SOILORGC by Annual Average Precipitation for Temperate_Coniferous
PREC in mm/month (Binned)
Distribution of SOILORGC by Annual Average Wind Speed for Temperate_Coniferous
WIND in m/s (Binned)

Figure 40:

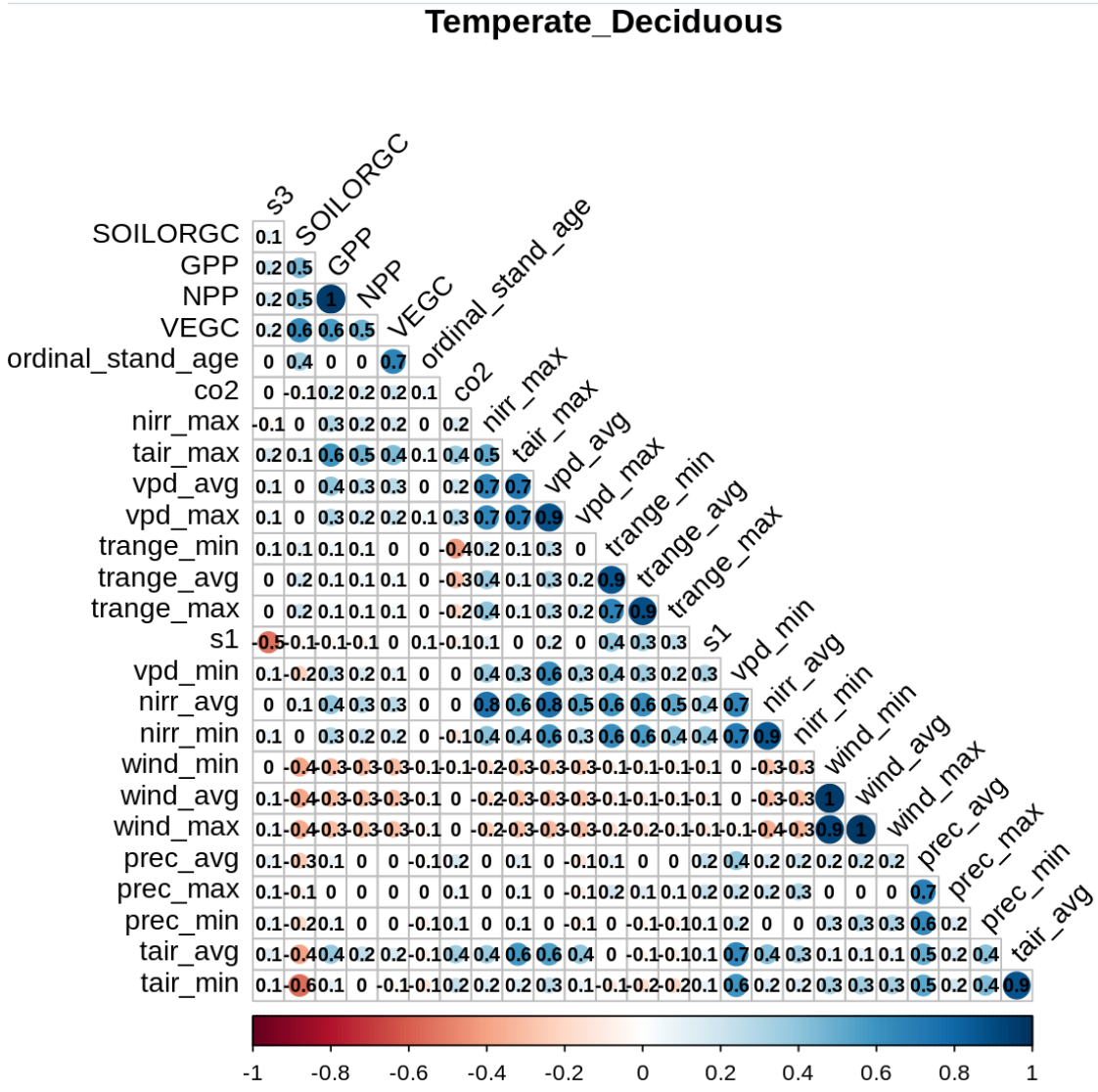
Temperate_Deciduous
s3
SOILORGC
GPP
NPP
VEGC
ordinal_stand_age
co2
nirr_max
tair_max
vpd_avg
vpd_max
trange_min
trange_avg
trange_max
s1
vpd_min
nirr_avg
nirr_min
wind_min
wind_avg
wind_max
prec_avg
prec_max
prec_min
tair_avg
tair_min
-1
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8
1

Figure 41:

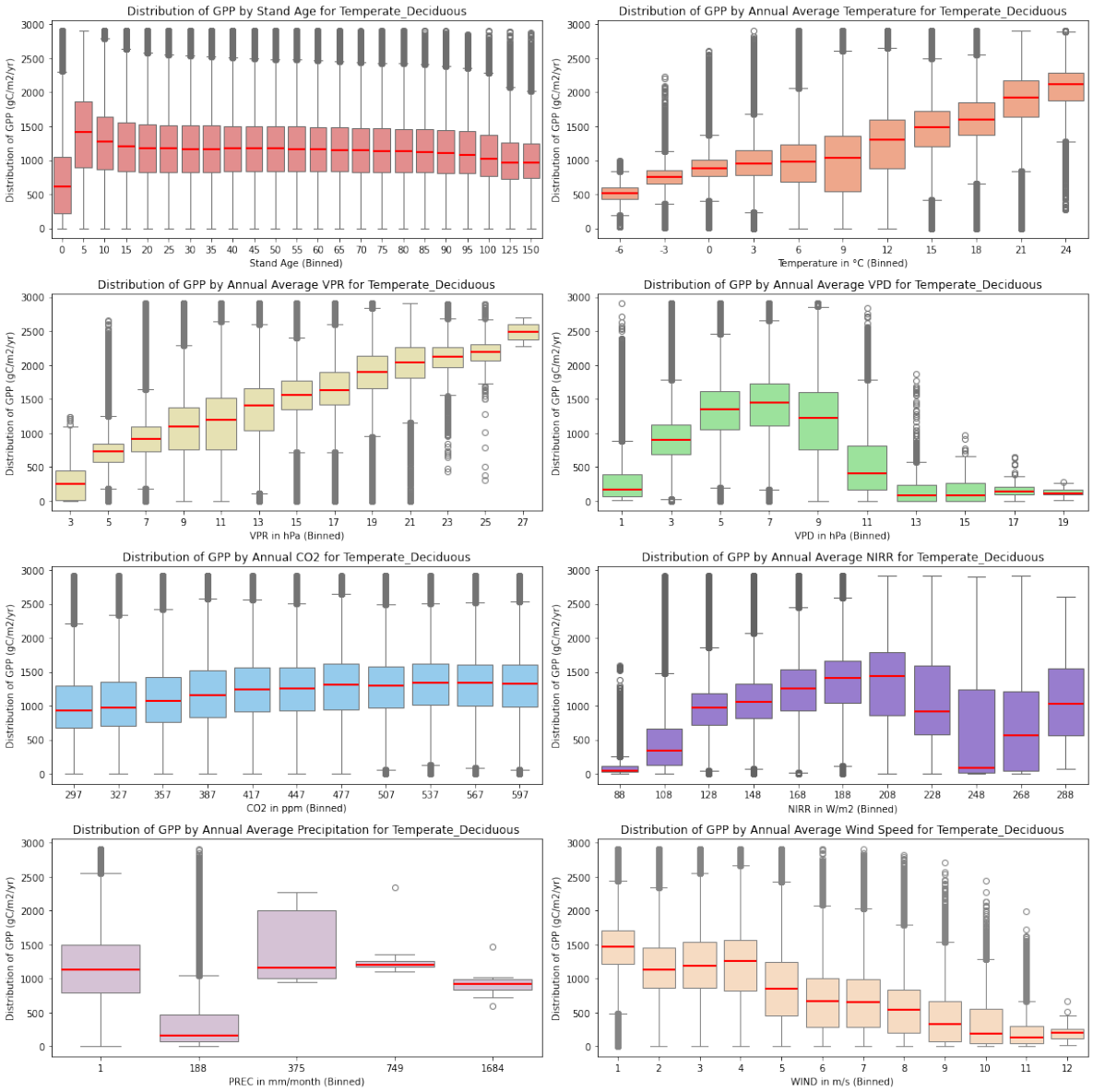
Distribution of GPP by Stand Age for Temperate_Deciduous
Distribution of GPP (gC/m2/yr)
Stand Age (Binned)
Distribution of GPP by Annual Average Temperature for Temperate_Deciduous
Temperature in °C (Binned)
Distribution of GPP by Annual Average VPR for Temperate_Deciduous
VPR in hPa (Binned)
Distribution of GPP by Annual Average VPD for Temperate_Deciduous
VPD in hPa (Binned)
Distribution of GPP by Annual CO2 for Temperate_Deciduous
CO2 in ppm (Binned)
Distribution of GPP by Annual Average NIRR for Temperate_Deciduous
NIRR in W/m2 (Binned)
Distribution of GPP by Annual Average Precipitation for Temperate_Deciduous
PREC in mm/month (Binned)
Distribution of GPP by Annual Average Wind Speed for Temperate_Deciduous
WIND in m/s (Binned)

Figure 42:

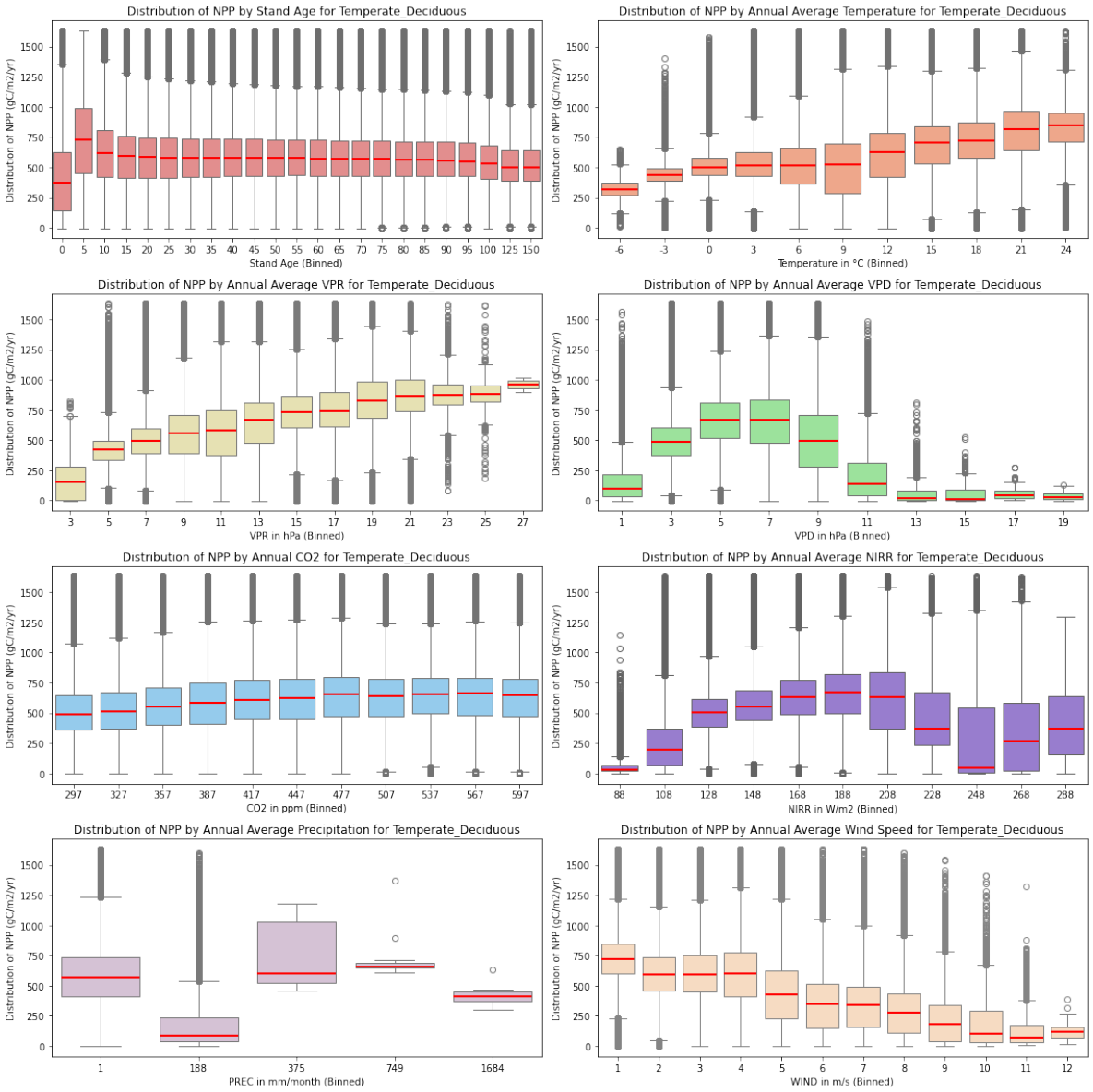
Distribution of NPP by Stand Age for Temperate_Deciduous
Distribution of NPP (gC/m2/yr)
Stand Age (Binned)
Distribution of NPP by Annual Average Temperature for Temperate_Deciduous
Temperature in °C (Binned)
Distribution of NPP by Annual Average VPR for Temperate_Deciduous
VPR in hPa (Binned)
Distribution of NPP by Annual Average VPD for Temperate_Deciduous
VPD in hPa (Binned)
Distribution of NPP by Annual CO2 for Temperate_Deciduous
CO2 in ppm (Binned)
Distribution of NPP by Annual Average NIRR for Temperate_Deciduous
NIRR in W/m2 (Binned)
Distribution of NPP by Annual Average Precipitation for Temperate_Deciduous
PREC in mm/month (Binned)
Distribution of NPP by Annual Average Wind Speed for Temperate_Deciduous
WIND in m/s (Binned)

Figure 43:

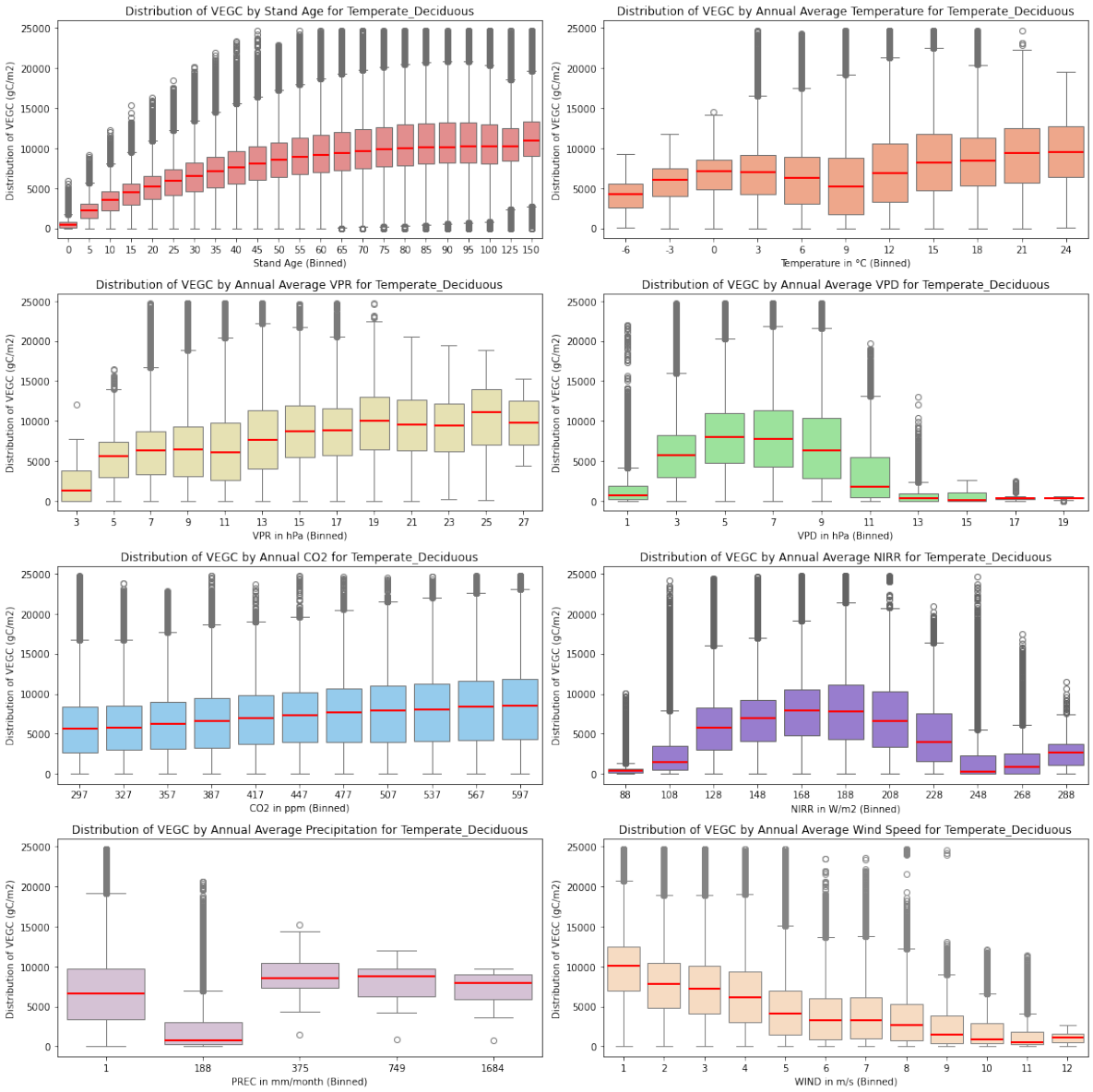
Distribution of VEGC by Stand Age for Temperate_Deciduous
Distribution of VEGC (gC/m2)
Stand Age (Binned)
Distribution of VEGC by Annual Average Temperature for Temperate_Deciduous
Temperature in °C (Binned)
Distribution of VEGC by Annual Average VPR for Temperate_Deciduous
VPR in hPa (Binned)
Distribution of VEGC by Annual Average VPD for Temperate_Deciduous
VPD in hPa (Binned)
Distribution of VEGC by Annual CO2 for Temperate_Deciduous
CO2 in ppm (Binned)
Distribution of VEGC by Annual Average NIRR for Temperate_Deciduous
NIRR in W/m2 (Binned)
Distribution of VEGC by Annual Average Precipitation for Temperate_Deciduous
PREC in mm/month (Binned)
Distribution of VEGC by Annual Average Wind Speed for Temperate_Deciduous
WIND in m/s (Binned)

Figure 44:

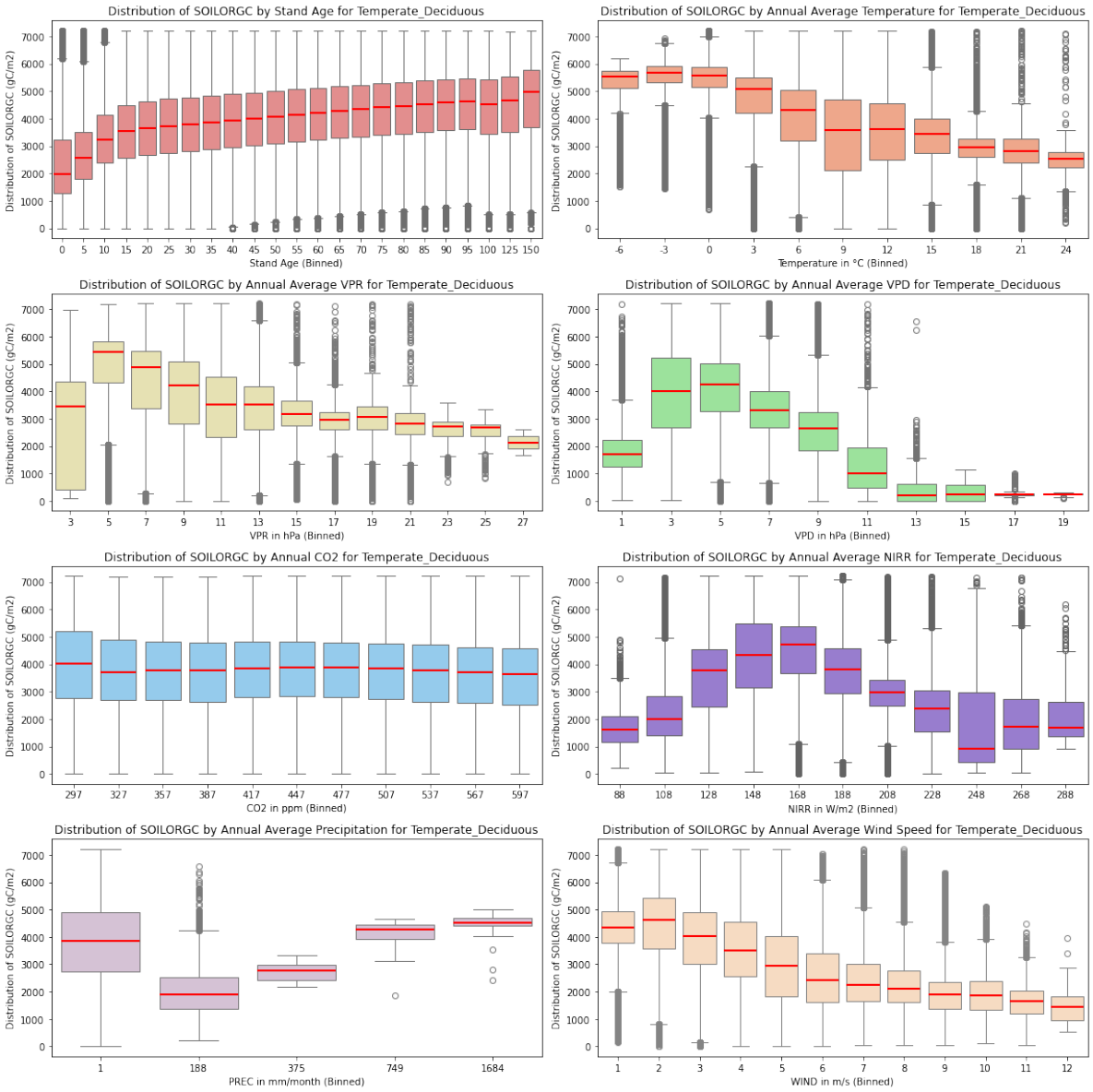
Distribution of SOILORGC by Stand Age for Temperate_Deciduous
Distribution of SOILORGC (gC/m2)
Stand Age (Binned)
Distribution of SOILORGC by Annual Average Temperature for Temperate_Deciduous
Temperature in °C (Binned)
Distribution of SOILORGC by Annual Average VPR for Temperate_Deciduous
VPR in hPa (Binned)
Distribution of SOILORGC by Annual Average VPD for Temperate_Deciduous
VPD in hPa (Binned)
Distribution of SOILORGC by Annual CO2 for Temperate_Deciduous
CO2 in ppm (Binned)
Distribution of SOILORGC by Annual Average NIRR for Temperate_Deciduous
NIRR in W/m2 (Binned)
Distribution of SOILORGC by Annual Average Precipitation for Temperate_Deciduous
PREC in mm/month (Binned)
Distribution of SOILORGC by Annual Average Wind Speed for Temperate_Deciduous
WIND in m/s (Binned)

Figure 45:

Tropical

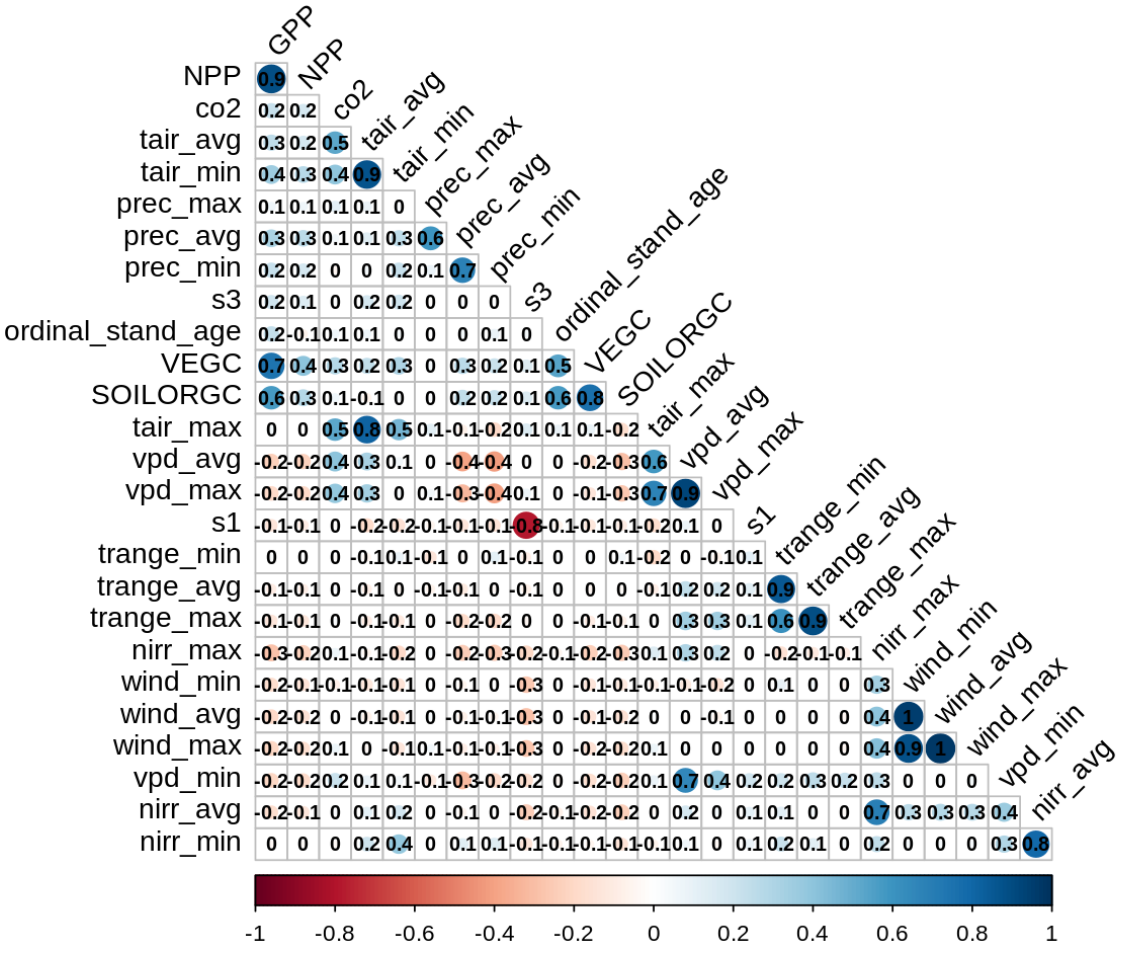

Figure 46:

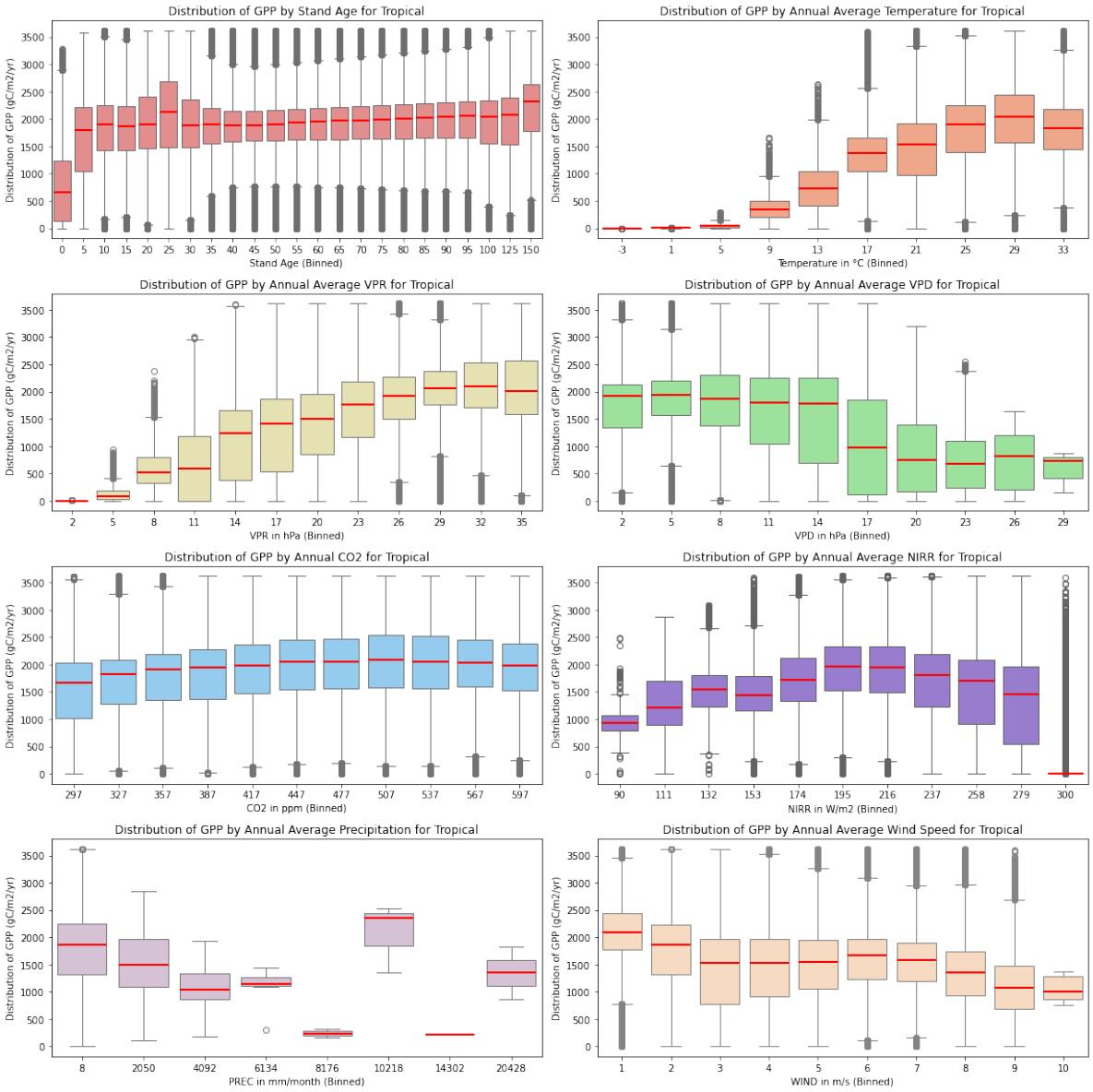
Distribution of GPP by Stand Age for Tropical
Distribution of GPP (gC/m2/yr)
Stand Age (Binned)
Distribution of GPP by Annual Average Temperature for Tropical
Temperature in °C (Binned)
Distribution of GPP by Annual Average VPR for Tropical
VPR in hPa (Binned)
Distribution of GPP by Annual Average VPD for Tropical
VPD in hPa (Binned)
Distribution of GPP by Annual CO2 for Tropical
CO2 in ppm (Binned)
Distribution of GPP by Annual Average NIRR for Tropical
NIRR in W/m2 (Binned)
Distribution of GPP by Annual Average Precipitation for Tropical
PREC in mm/month (Binned)
Distribution of GPP by Annual Average Wind Speed for Tropical
WIND in m/s (Binned)

Figure 47:

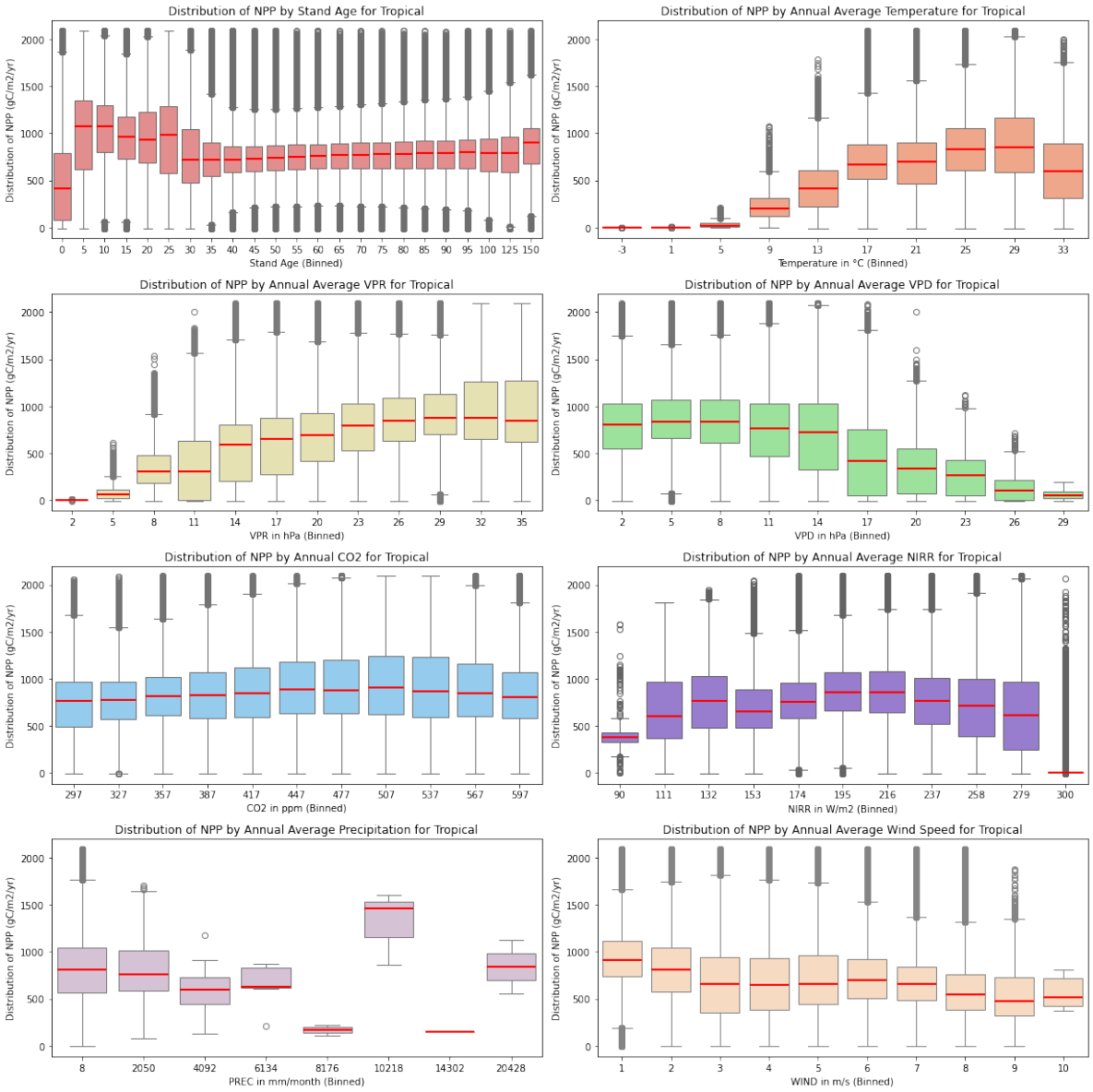
Distribution of NPP by Stand Age for Tropical
Distribution of NPP (gC/m2/yr)
Stand Age (Binned)
Distribution of NPP by Annual Average Temperature for Tropical
Temperature in °C (Binned)
Distribution of NPP by Annual Average VPR for Tropical
VPR in hPa (Binned)
Distribution of NPP by Annual Average VPD for Tropical
VPD in hPa (Binned)
Distribution of NPP by Annual CO2 for Tropical
CO2 in ppm (Binned)
Distribution of NPP by Annual Average NIRR for Tropical
NIRR in W/m2 (Binned)
Distribution of NPP by Annual Average Precipitation for Tropical
PREC in mm/month (Binned)
Distribution of NPP by Annual Average Wind Speed for Tropical
WIND in m/s (Binned)

Figure 48:

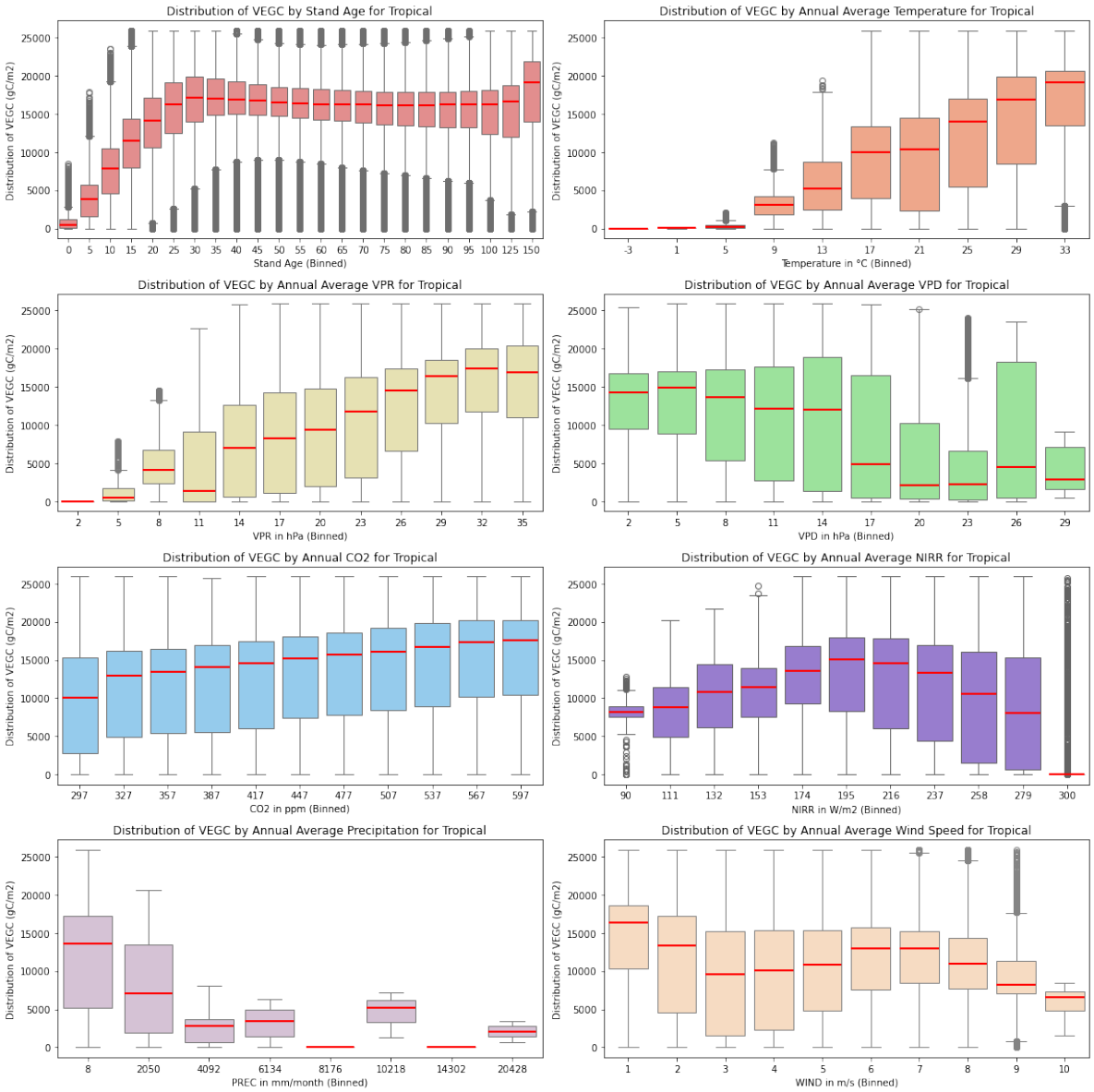
Distribution of VEGC by Stand Age for Tropical
Distribution of VEGC (gC/m2)
Stand Age (Binned)
Distribution of VEGC by Annual Average Temperature for Tropical
Temperature in °C (Binned)
Distribution of VEGC by Annual Average VPR for Tropical
VPR in hPa (Binned)
Distribution of VEGC by Annual Average VPD for Tropical
VPD in hPa (Binned)
Distribution of VEGC by Annual CO2 for Tropical
CO2 in ppm (Binned)
Distribution of VEGC by Annual Average NIRR for Tropical
NIRR in W/m2 (Binned)
Distribution of VEGC by Annual Average Precipitation for Tropical
PREC in mm/month (Binned)
Distribution of VEGC by Annual Average Wind Speed for Tropical
WIND in m/s (Binned)

Figure 49:

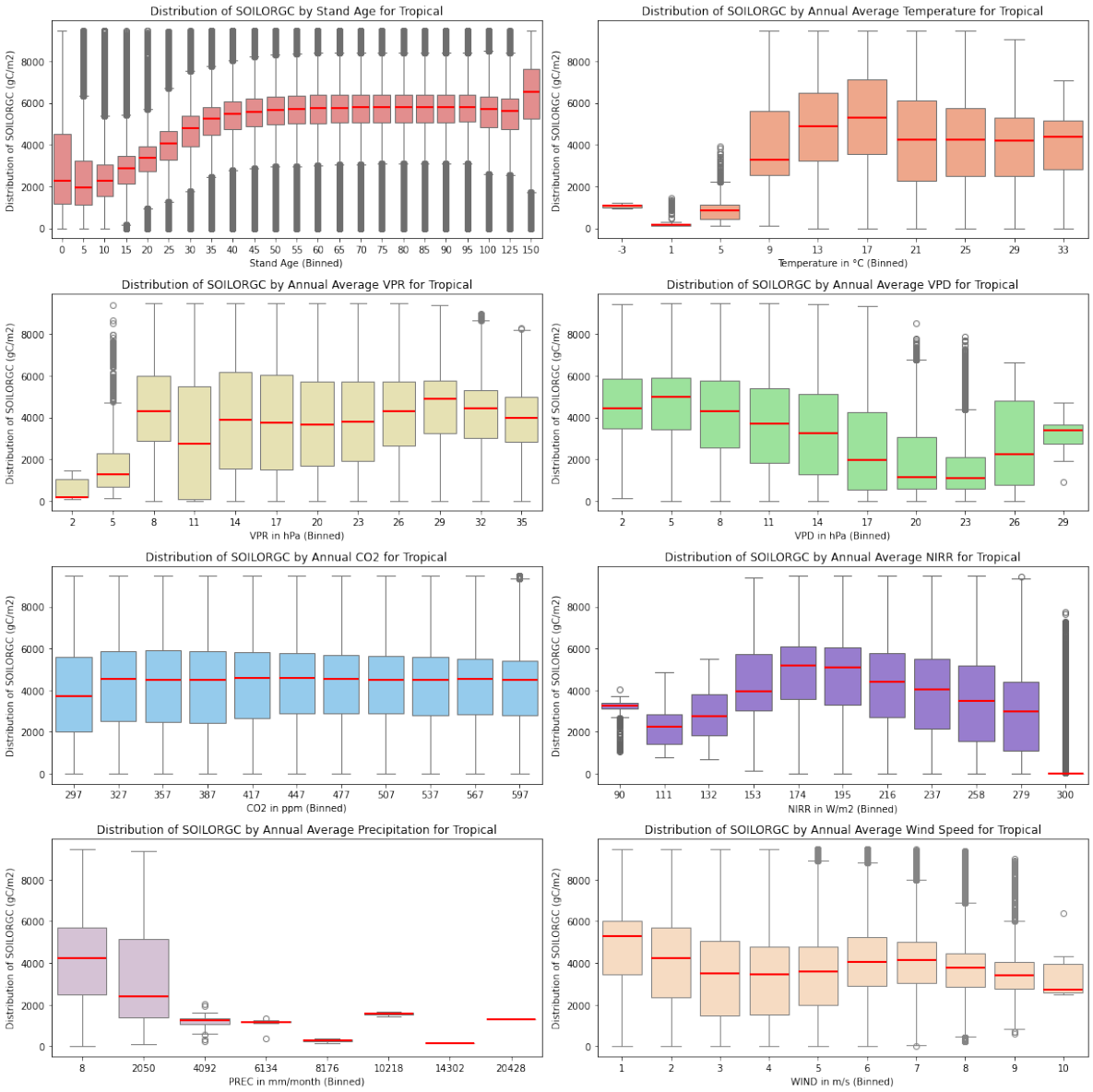
Distribution of SOILORGC by Stand Age for Tropical
Distribution of SOILORGC (gC/m2)
Stand Age (Binned)
Distribution of SOILORGC by Annual Average Temperature for Tropical
Temperature in °C (Binned)
Distribution of SOILORGC by Annual Average VPR for Tropical
VPR in hPa (Binned)
Distribution of SOILORGC by Annual Average VPD for Tropical
VPD in hPa (Binned)
Distribution of SOILORGC by Annual CO2 for Tropical
CO2 in ppm (Binned)
Distribution of SOILORGC by Annual Average NIRR for Tropical
NIRR in W/m2 (Binned)
Distribution of SOILORGC by Annual Average Precipitation for Tropical
PREC in mm/month (Binned)
Distribution of SOILORGC by Annual Average Wind Speed for Tropical
WIND in m/s (Binned)

Figure 50: