Improving Carbon Cycle Uncertainty Through Ensemble Based Temporal Downscaling

Author

James Simkins

A thesis submitted in partial fulfillment of

the requirements for the degree of

MASTER OF SCIENCE

(Atmospheric and Oceanic Sciences)

at the

UNIVERSITY OF WISCONSIN-MADISON

2017

Thesis Declaration and Approval

I, James Simkins, declare that this thesis titled “Improving Carbon Cycle Uncertainty Through Ensemble Based Temporal Downscaling” and the work presented in it are my own.

James Simkins		
Author	Signature	Date

I hereby approve and recommend for acceptance this work in partial fulfillment of the requirements for the degree of Master of Science:

Dr. Ankur R. Desai		
Committee Chair	Signature	Date
Dr. Christopher Kucharik		
Faculty Member	Signature	Date
Dr. Daniel Vimont		
Faculty Member	Signature	Date

Abstract

The terrestrial biosphere assimilates nearly one fourth of anthropogenic carbon dioxide emissions, providing a significant ecosystem service. Anthropogenic climate changes that influence the distribution and frequency of subdaily weather phenomena can have a momentous impact on this useful function that ecosystems provide. However, most ecosystem-model based analyses of the impact of future extreme events on site to regional ecosystem carbon uptake do not consider subdaily meteorology. In order to improve these ecosystem model forecasts, we developed an ensemble based high resolution temporal downscaling routine designed to propagate uncertainty as part of the Predictive Ecosystem Analyzer (PEcAn, http://pecanproject.org) workflow and package. This routine uses a multi-variate regression based approach to downscale any daily or subdaily meteorological dataset to an hourly resolution. Temporally downscaling to an hourly resolution allows for the larger, more energetic eddies within the planetary boundary layer to overturn. Further, an hourly timestep allows for better model simulations of plant physiological response to rapidly changing weather (e.g., radiation, temperature, humidity) and microclimate. In order to test the strength of our downscaling algorithm, we aggregated a year of hourly resolution eddy covariance data from the Willow Creek Fluxnet tower to a daily resolution, used our downscaling procedure to downscale it back to an hourly resolution, and then compared our observations against our downscaled output. Following this test, we performed the same procedure for each year between 1999 and 2015 and compared ecosystem model output for each resolution. We find that the hourly downscaled meteorology improved our net ecosystem exchange values by 94%, bringing our modeled values closer to the actual observations. Next, we use the temporal downscaling algorithm to sample uncertainty in future projected meteorological drivers at site-level daily

resolution using Coupled Model Intercomparison Project 5 (CMIP5) forcing data from the Multivariate Adaptive Constructed Analogs (MACA) dataset. MACA offers spatially downscaled data for each CMIP5 model for two Representative Concentration Pathways (RCP) of 4.5 and 8.5 (W/m2 trapped by 2100, respectively) emissions scenarios at a resolution of 4-km from the present day to 2100 (Abatzoglou et al., 2012). We temporally downscaled each model in the MACA dataset from 2020-2030 so that the forcings were calibrated using observations at an hourly eddy covariance flux tower. After running the SIPNET ecosystem model using these temporally downscaled future climate data, we find a large divergence of cumulative net ecosystem exchange (~4000 g C m-2) across the CIMP5 models during the next decade. This large spread is due to differences in frequency of extreme temperatures, extreme precipitation events, and prolonged periods without precipitation.

Acknowledgements

I would like to express my sincere gratitude to the Desai lab and the PEcAn project members including Christy Rollinson, all of whom assisted in this research and offered support throughout my graduate career. I wish to express my appreciation for Dr. Ankur Desai, my adviser, who served as an incredible mentor and supported me and my research throughout my time in graduate school. I would also like to thank my other committee members, Dr. Christopher Kucharik and Dr. Daniel Vimont, who through classes and individual meetings helped cultivate the ideas I express in this thesis and always provided invaluable mentorship along the way.

I would like to acknowledge the generous support from the following institutions which provided me the financial and technical support allowing me to conduct this research: the University Of Wisconsin Department Of Atmospheric and Oceanic Sciences and the US National Science Foundation under PEcAn Awards DBI-1458021, DBI-1457897, and DBI-1457890.

Lastly, I would like to express my gratitude for my family and friends who supported me throughout this academic journey. I am incredibly fortunate to have such wonderful people in my life who are so encouraging and caring. Also, a special thanks to CJ for helping me out with getting this printed!

Table of Contents

List of Figures

List of Tables

List of Equations

Introduction

Human activity over the past 200 years has caused an increase in atmospheric concentration of carbon dioxide. Carbon dioxide emissions will continue to increase globally into the future (Raupach et al., 2008). Given our anticipated rates of emission, the International Panel on Climate Change (IPCC) Fifth Assessment Report (2014) shows that global-mean temperature is expected to increase between 2.6 and 4.8 degrees Celsius by 2100 relative to 1986-2005 measurements. One reason for the range in temperature increase is due to uncertainty of natural carbon sequestration. The Earth's oceanic ecosystems and terrestrial biosphere provide a significant service in reducing the impact of anthropogenic carbon dioxide emissions through the assimilation of this greenhouse gas. The terrestrial biosphere alone assimilates nearly one fourth of these carbon dioxide emissions which diminishes the full effect of human activity, but future rates range from significantly larger carbon sinks to global carbon sources (Friedlingstein et al., 2016; Pachauri, Rajendra K. et al., 2014). Understanding how this useful ecosystem function will change as the climate changes is important for increased confidence in prediction of further impacts to society and the health of planet Earth.

The reason future land uptake estimates diverge is that global warming causes temperature increases and changing precipitation patterns that will have varied regional impacts on ecosystems because of differences in soils, vegetation, and land management (Walther et al., 2002). For example, ecosystem response to phenomena such as daily temperature extremes, changing precipitation event duration and intensity, and an increase in extreme weather events each have implications for the global carbon budget. An extreme weather event is defined in the IPCC (2014) as an event that is as rare or rarer than the 10^{th} or 90^{th} percentile of a Gaussian

distribution. Frank et al. (2015) suggested that a relatively small change in the mean or variance of a climate variable inherently leads to a disproportional change in the frequency of extremes.

It is projected that longer and more frequent heat waves will be a consequence of global warming. These temperature extremes are usually coupled with drought which can induce water stress on forests in particular with an increase in tree mortality as a consequence (Breda et al., 2006; Bigler et al., 2007; Adams et al., 2010). These anomalously high temperature episodes coupled with a lack of precipitation leads to stomatal closure, higher vapor pressure deficit, and lower levels of leaf transpiration and evaporative cooling by plants (De Boeck et al., 2011). De Boeck et al. (2011) goes onto discuss how high temperatures and increased levels of sensible heat also increase the rate of soil moisture depletion, leading to even further stresses on plant hydraulics. More negative soil water potential and low soil hydraulic conductivity are consequences of the high temperatures projected due to anthropogenic climate change (Reichstein et al., 2013).

These examples of temperature impacts on ecosystems are not limited to long term events. Subdaily fluctuations in temperature can have significant implications for plant function and gross primary production (GPP) (Medvigy et al., 2010). Long (1991) showed that photosynthesis is optimized at a particular temperature and photosynthetic rates are diminished at extreme low and high temperatures. Thus, temperature variability throughout a day can significantly impact GPP values and ultimately control the carbon fluxes. When compared to using just daily mean, maximum, and minimum temperature data in model experiments, hourly temperature data can provide greater insight to diurnal photosynthetic response to temperature variation leading to improved accuracy of carbon flux estimates.

Changes in subdaily fluctuations in temperature and specifically diurnal variability due to climate change has consequences for agriculture as well. Peng et al. (2004) investigated diurnal temperature range impacts for rice crop yields using a 24-year observational dataset. They reported varied increases in annual mean, minimum and maximum temperatures and show a decrease in the diurnal temperature range. As a consequence, they discovered that increased nighttime temperature associated with global warming decreased rice yields.

While Peng et al. (2004) provided us with evidence showing how maximum and minimum temperature ranges can impact rice yields, climate change impacts on agriculture are not limited to rice. Lobell et al. (2012) showed that carbon dioxide trends are likely to increase global yields by roughly 1.8% per decade over the next few decades. Conversely, warming trends are likely to reduce global yields by 1.5% per decade should the plants fail to adapt effectively. Changes in local temperature and precipitation conditions will also play a role in determining crop yields. Hatfield et al. (2010) discusses how agronomists will need to consider spatial and temporal variations in temperature and precipitation as part of the production system to determine future yield responses to climate change. This review discusses how sustained periods when temperature exceeds thresholds for damage to plants can decrease crop production.

Changing precipitation patterns, event duration, and intensity are also altering hydrological systems that the terrestrial biosphere depends on. Extreme precipitation trends (greater than 95^{th} percentile of Gaussian distribution for location) have been observed since 1950 and show that the number of heavy precipitation events have increased over most land areas (IPCC, 2014). These changes will not be equally distributed across the globe, leaving specific regions to be affected differently. Mean precipitation in the mid-high latitudes and equatorial Pacific are likely to increase whereas many mid-latitude and subtropical dry regions show a

decrease under our current emission levels and socio-economic activities following the Representative Concentration Pathway (RCP) 8.5 (W/m2 increase in radiation) scenario. Precipitation events are very likely to become more intense and frequent over wet tropical regions and mid-latitude land masses (IPCC, 2014).

1.1 Evidence of Ecosystem Response to Subdaily Meteorology

Medvigy et al. (2010) compared the influences of hourly, daily, and monthly variance of carbon fluxes using the Ecosystem Demography model version 2 (ED2) and found that terrestrial ecosystems are highly sensitive to high-frequency meteorological variability. Through their modeling efforts, they discovered substantial differences in ecosystem functioning when ED2 was driven by hourly meteorological forcing data compared to coarser resolution forcing data with the same means. They found that projected changes in radiation and precipitation fluctuations have significant ramifications for carbon storage and ecosystem structure and composition. They concluded that the statistical variability at high temporal frequency is essential for robust forecasting of ecosystem structure and functioning, thus providing further evidence that subdaily meteorological data are required to accurately forecast carbon cycle changes.

Precipitation regimes in the mid-latitudes are more likely to have greater intensities but occur less frequently due to climate change (IPCC, 2014). Air has a greater capacity for water vapor at higher temperatures and the vapor pressure deficit (VPD) between plants and the atmosphere will grow, leading to increased evaporation and transpiration rates to restore the balance (Novick et al., 2016). This study showed higher levels of atmospheric water vapor can intensify precipitation events because of increased moisture availability and latent heat release

which fuel low pressure systems. Heavier precipitation events and stronger localized convective systems present a new problem for ecosystems on a local scale. Rain events initiated by convective forcing and local convergence in the United States typically occur on a spatial scale of 10 to 100 km and at a subdaily timescale where the timing of initiation varies region to region (Carbone et al., 2001). Essential plant nutrients such as nitrogen and phosphorus that are either found naturally in soils or applied as organic and inorganic amendments in agriculture are more likely to be lost to surface and groundwater with extreme rainfall events (Trenberth, 2011). Trenberth (2011) also showed how the irregularity of precipitation events can cause surface crusting during drought events which can increase runoff during subsequent heavier precipitation events. Flooding events such as the Boscastle, England flood in August 2004 was caused by 181 mm of rain that fell in a span of 5 hours (Wheater 2006), highlighting the need for subdaily resolution rainfall intensity data.

The shifts in precipitation patterns and intensity are also likely to impact soil moisture availability coupled with increased evapotranspiration (ET). ET is highly variable at both a daily and subdaily time step (Fisher et al., 2017). Solar radiation, humidity, air temperature, and wind speed regulate the transfer of water between plant and atmosphere and thus drive ET. Another significant driver examined in ecology but not mentioned with the traditional meteorological variables is vapor pressure deficit which tells us the relative magnitude of atmospheric water vapor demand. Novick et al. (2016) showed that increased atmospheric water vapor demand limits stomatal conductance, a measure of a plant's ability to exchange water vapor, CO2 and O2 with the atmosphere. Their results indicate that in the future, the terrestrial carbon cycle will be impacted by higher temperatures because the warmer conditions will increase the respective importance of vapor pressure deficit in limiting stomatal conductance and ET. Vapor pressure

deficit has a linear relationship with gross primary production and ET and considering these variables at a subdaily timescale is important for maintaining the carbon-water connection (Zhou et al., 2014). The studies each highlight the importance of subdaily meteorological information, as it will ultimately impact plant function and the carbon cycle.

Parmesan et al. (2000) provides evidence that extreme weather events are mechanistic drivers of broad ecological response and are essential in estimating carbon cycle feedbacks from ecosystem models. We have shown that extreme weather events occur at a local to regional spatial scale and can be subdaily events. In order to estimate extreme event impacts, a spatially and temporally downscaled dataset is required. Individual plants respond to meteorological drivers on a subdaily timescale and can vary from site to site across large spatial domains.

1.2 Current Approaches to Downscaling

Recent efforts in global climate modeling have been shown through the fifth Coupled Model Intercomparison Project (CMIP5) where 40 different climate centers around the world ran global climate models with multiple ensemble outputs for four likely scenarios and multiple physical realizations (Taylor et al., 2010). This is the most robust GCM effort to date, yet the average spatial resolution of these models is around 1.75 degrees latitude by 1.75 degrees longitude (ENES, 2011).

The coarse spatial resolution has motivated multiple spatial downscaling efforts to identify climate changes on the regional scale. Zhang et al. (2005) spatially downscaled GCM output from the native GCM grid scale to station-scale using a transfer function method and used this method to explore the site-specific impact assessment of climate change on water resources at Kingfisher, Oklahoma, USA. Vimont et al. (2010) used empirical orthogonal

function/principal component analysis with linear regression, maximum covariance analysis, and canonical correction analysis to downscale precipitation over Indonesia. They investigated the skill of the linear method employed and found that most of the skill is attributed to year to year El-Nino Southern Oscillation variation and long-term trend in precipitation and large scale fields. Other efforts for statistical downscaling of spatial resolution have specifically focused on precipitation (Schmidli et al., 2007; Gutmann et al., 2014) or temperature and precipitation (Clark et al., 2004; Vrac et al., 2016) where multiple approaches are used such as stepwise multiple linear regression, bias correction and spatial disaggregation to generate the statistics required to downscale. Each of these methods employ the use of statistical spatial downscaling which assumes a relationship between the large scale and local scale flow and that this relationship will exist into the future. These spatial downscaling efforts have increased our understanding of meteorological impacts on a much finer grid space and have enhanced our ability to predict local-scale climate change.

While there has been a breadth of research performed for spatially downscaling GCM output, little has been done to temporally downscale these coarse meteorological projections. The majority of the temporal downscaling efforts thus far have been geared towards downscaling precipitation or temperature. Mendes and Marengo (2010) used artificial neural networks to temporally disaggregate precipitation and showed that their method outperformed a statistical downscaling model, but they were only temporally downscaling to a daily resolution and this was only performed for precipitation. Kumar et al. (2012) statistically downscaled multiple meteorological variables using neural networks from monthly to sub-daily time steps while retaining consistent correlations between variables. They concluded that neural networks may not be a realistic approach if the time frame extends to a period with significantly different

weather variability. Kirchmeier-Young et al. (2016) provided evidence that a probabilistic approach to temporal downscaling is necessary to model extreme events and showed a variety of validation metrics for examining the results of the University of Wisconsin Probabilistic Downscaling dataset. While this dataset provides a robust representation of extreme events at a fine spatial resolution, the daily resolution data is not fine enough in a temporal sense to fully investigate ecosystem responses to subdaily meteorological variation.

1.3 Objectives

The present study is motivated by the lack of high temporal resolution GCM output available for testing subdaily extremes and their impact on the terrestrial carbon cycle at a local scale. Further, ecosystem models not only require hourly meteorology, they also require all variables relevant to land-atmosphere interaction including temperature, precipitation, incoming shortwave and longwave radiation, pressure, relative or absolute humidity, and wind speed. We have developed the first open-source ensemble based multi-variable temporal downscaling routine that can be used to disaggregate any resolution dataset in Climate and Forecast (CF) convention based on statistics from a subdaily observational dataset. This routine is available through the Predictive Ecosystem Analyzer (PEcAn) R package, which accelerates ecosystem modeling and makes the process reproducible. We use a linear regression based approach where we leverage statistics of the provided training data using a linear regression based approach. Our method is also unique as we generate multiple ensembles of downscaled output to propagate uncertainty into the future. By creating multiple ensembles, we are able to analyze our uncertainty that is inherent to any downscaling procedure (Bierkens et al., 2000; Jianguo, 2006).

The purpose of this manuscript is to: (i) describe the open source, ensemble based downscaling procedure we have developed, (ii) evaluate the temporal downscaling algorithm, (iii) and use the temporal downscaling procedure to estimate future net ecosystem exchange (NEE). This paper is directed toward the ecosystem modeling community and climate downscaling community, with the goals of: (i) to show evidence supporting the need for high resolution meteorological driver data, (ii) validate the temporal downscaling algorithm detailed above, and (iii) temporally downscale projected climate data to understand the effect future climate has on ecosystems. This paper is organized as follows. The methodology and data descriptions are detailed in Section 2. Results from our algorithm validation and future carbon cycle response are presented in Section 3. A discussion of the results and future goals are presented in Section 4, and a summary of results are presented in Section 5.

2. Methods

The focus of the present study is to account for subdaily meteorological variations and their impact on future ecosystem carbon cycling. We developed a unique temporal downscaling algorithm that we use to temporally disaggregate daily resolution CMIP5 output into hourly resolution forcing data and use the downscaled data product to drive the Sipnet ecosystem model within the PEcAn workflow. We describe the PEcAn ecoinformatics workflow in section 2.1 and the specifics of the temporal downscaling algorithm created within PEcAn in 2.2. Data and models are discussed in section 2.3 followed by the experimental design in section 2.4.

2.1 The Predictive Ecosystem Analyzer (PEcAn)

Ecosystem modeling plays a critical role in basic ecological research and the lack of conventional methods and standards has slowed the pace of improvement (Dietze et al., 2013; Moorcorft, 2006). Diversity in data collections without unified metadata guidelines makes reproducibility of a modeling effort difficult. Many models aren't available to the wider ecosystem community and collecting multiple data sources creates a large time sink. PEcAn attempts to solve these problems and accelerate the pace of model improvement by making data and models more accessible (Lebauer et al., 2013). PEcAn is an informatics project that acts as an accessible home for the ecosystem community by allowing a user to parameterize, run, analyze and benchmark a suite of ecosystem models in one place. PEcAn integrates multiple sources of meteorological data and trait data that all are converted to the CF metadata convention standards and follow a unified coding protocol. The tools that have been developed for PEcAn can be combined into customizable workflows and used for ecosystem modeling and synthesis along with decision support. The overarching goal of PEcAn aims to improve ecosystem models and reduce the uncertainty of climate change impacts on ecosystems and the carbon cycle. PEcAn has been developed in the R programming language (R Development Core Team 2008) and all code is open-source and is available at (https://github.com/PecanProject/pecan).

a. Ecosystem Modeling Example Using PEcAn

The first step of the PEcAn involves selecting a location anywhere on the globe and an ecosystem model to run at that location. Once that is declared, the next step is downloading the meteorological driver data. Within the PEcAn framework, we've developed a suite of data extraction functions that pull data from various meteorological data archives. These include reanalysis datasets spanning as far back as 1900, recent observational datasets, and climate

projected data reaching out to 2100. Point location observational datasets and gridded datasets at various resolutions allow a user to span across multiple spatial and temporal scales. These are designed to collect variables of interest for a specified range of years, rename them using CF convention, and save them as individual years in the Network Common Data Form (NetCDF) format. A user can simply select the function that downloads the dataset of their choice, enter the logistical arguments necessary, and download the data on the fly.

The downloaded dataset is then gapfilled to fill any missing data values using the Marginal Distribution Sampling (MDS) method described in Reichstein et al. (2005). The meteorological data is now configured for that specific model input format. These two steps are done by using *metgapfill()* and *met2model()* which are functions that have been developed in PEcAn and generalized for all of our data and models.

Next, a user selects the plant functional type (PFT) (availability varies by model, SIPNET offers Temperate Deciduous and Temperate Coniferous) for the location of interest which is used to parameterize the ecosystem model. These traits and priors are queried from the Biofuel Ecophysiological Traits and Yields Database (BETYdb) and a meta-analysis is performed. A user can select to run the model with multiple ensemble runs with varying trait data parameters for the specific model. This is important for quantifying our trait data uncertainty.

After the input stage, a PEcAn creates an XML file containing all of the settings required to run the chosen model. The workflow is now ready to run the model. Each model run can be found in the automatically generated run, PFT, and output folders. One can find the results of the model in the output folder for each year of meteorology that drove the model. The output files are where we can find out the model results for NEE, Net Primary Productivity, Gross Primary

Productivity, etc. An ensemble analysis of the model output immediately follows the model execution stage. Then, a sensitivity analysis and variance decomposition of the model output is performed. Parameter data assimilation comes next and performs a Bayesian Markov Chain Monte Carlo on the model parameters by suggesting parameter values, running the model, computing the likelihoods, and accepting or rejecting the proposed parameters.

State data assimilation then runs each ensemble member through a Kalman filter and generalized filter code. This function outputs each of the model run outputs, a PDF of diagnostics, and an Rdata (R's native data file type) object that includes the forecast, analysis, and prior and posterior covariance matrix for each time step. A benchmarking analysis can also be performed if selected by the user which evaluates the parameters by comparing to a collection of standards. A key advantage of this workflow is that it's coherent across multiple models and parameters with a uniform structure that is reproducible. Ecosystem modeling has traditionally been unsystematic with lack of a central parameter database, agreed unit conventions, or a reproducible open-sourced workflow. PEcAn solves these issues and increases the quality and quantity of science we can do.

This paper focuses on the recent addition of a temporal downscaling algorithm to PEcAn R package. The Temporal Downscale Meteorology (TDM) workflow has been synthesized within PEcAn and directly functions in between the driver data gap filling and the conversion of that data to model specific meteorological data syntax. The TDM functions will temporally downscale coarse meteorology downloaded before it is read into the model format. The development of this algorithm and the specific routine will be detailed in the following section.

2.2 Temporal Downscaling Algorithm

Our statistical downscaling scheme uses a linear regression based approach motivated by the plethora of high resolution observational and reanalysis meteorological datasets currently available. Linear regression is a popular statistical method used to investigate the relationships between predictor and predictand (Neter et al., 1996). An advantage of a linear regression approach is that it enables us to quantify and propagate uncertainty. Additionally, other methods described by Wilby and Wigley 1997 suggest that statistical downscaling techniques such as regression are not as computationally intensive as others. We chose the linear regression technique for its computational speed and to address the statistical uncertainties in the downscaling (Jianguo, 2006). The basic linear regression equation is displayed below.

Linear Regression Equation

$$y = a + bx + e$$

$y = dependent, response, or\ predictand\ variable$

$a = intercept$

$b = beta\ coefficient\ or\ slope$

$x = independent, explanatory, or\ predictor\ variable$

$e = residual\ error$

Equation 1. *Linear regression equation (Montgomery et al., 2015)*

In the present study, we are investigating future meteorology and its impact on the ecosystem. It's important to note that we incorporate the future climate change signal by temporally downscaling projected meteorological datasets. We are assuming that the relationships between the dependent and independent variables remains valid into the future, an

assumption often referred to as stationarity (Schmith 2008). Our temporal downscaling routine is the first open sourced ensemble based temporal downscaling function that downscales a suite of meteorological variables that are necessary to drive ecosystem models. Other open source temporal downscaling methods only downscale one or a few meteorological variables and don't quantify uncertainty through the generation of multiple ensembles (Wilby et al., 2004; Maraun et al., 2010; Mascaro et al., 2014). We will now discuss each step of the temporal downscaling workflow and explain how it was synthesized within the PEcAn workflow.

Meteorological Variables Downscaled	
Variables	**Units**
Air Temperature	Kelvin
Surface Downwelling Shortwave Flux In Air	Watts meter^{-2}
Precipitation Flux	Kilograms meter-2 second^{-1}
Air Pressure	Pascals
Surface Downwelling Longwave Flux In Air	Watts meter^{-2}
Specific Humidity	Kilogram Kilogram^{-1}
Wind Speed	Meters second^{-2}

Table 1. *Downscaled meteorological variables and units*

a. Extracting Meteorological Datasets

The TDM workflow in PEcAn requires two datasets for downscaling, a fine resolution training dataset and a coarse resolution dataset to downscale. The training data should ideally be at least a decade in length consisting of observations at a high temporal frequency (sub-daily,

ideally hourly or half-hourly). The coarse resolution dataset that we wish to downscale can span any range of dates but should be at a daily temporal resolution. We assume that future inter-variable relationships will be more robust during years closer to the observational dataset. A user can select to download a single year or multi-year dataset at any point in time that they wish to downscale.

b. Generate Linear Regression Models – gen.subdaily.models()

The first step of our TDM workflow is to generate linear regression models using provided training data. At the front end of the model generation function, multiple arguments must be specified. The number of beta coefficients desired is a parameter that must be defined and the larger number of betas leads to a larger pool to draw from for uncertainty. Beta coefficients measure the slope or steepness of the regression line and any intercepts. Most importantly, the beta coefficients preserve the covariance among coefficients and intercepts. The larger the number of betas set, the larger memory storage required and the slower the calculations. We do not recommend setting the number of betas below 30 due to the central tendency theorem. The central tendency theorem states that the mean of a sample of data will be closer to the mean of the entire dataset as the sample size grows (Rosenblatt 1956). As a general rule, sample sizes greater than 30 are considered sufficient for this theorem to hold. The routine generates the specified number of beta coefficients from the normal distribution by preserving covariance among predictors and the fine temporal resolution so the temporal autocorrelation is calculated.

The window day argument specifies the number of days around the current day to include in the normal distribution. For example, if we are analyzing the training data at 01:00 AM on

January 5th, we will pull values from 01:00 AM on January 3,4,5,6 and 7th given a window day length of 5 days. We use a small window day length in order to keep smooth transitions between days. This helps us avoid problems with lack of data in small datasets or zero inflated variables such as precipitation flux. Another argument important for this function is declaring whether or not we want to calculate residuals for each time step. Residuals are the difference between the predicted value and dependent variable. If we calculate the residuals and store them in the linear regression model output, we will use these to help us add noise to the downscaled product and propagate uncertainty. Logistical arguments such as the training data file path and a folder to save the models are also needed.

Once the initial arguments are specified, the training data is read into the function and additional variables are created that are used as predictors. For each meteorological variable in the training dataset, we calculate daily means for the previous day, current day, and next day which help us identify daily transitions. For air temperature, we additionally calculate daily minimums, minimum departures, maximums, and maximum departures for the previous, current, and next days. If it is the first day of the first year in the training dataset, we declare the previous day to be equal to the current day. The same protocol is done for the last day in the last year of the training dataset. We also leverage the previous and next time step values because the previous time step if not independent from the next time step, especially at fine temporal resolutions such as hourly. Failure to do this causes abrupt, illogical shifts in variables such as temperature. Once the additional variables have been created, the algorithm will model each meteorological variable and the time step of the training data.

Each variable in Table 1 undergoes a similar but unique linear regression calculation that gets applied on a daily basis. We perform the calculations for each variable independent of the others.

A multivariate linear regression technique was tested but returned results well outside the range of possibility for most variables, so specific humidity has the only multivariate approach. The covariances of predictors at the diurnal resolution are poorly constrained, but specific humidity has a clearer physical relationship with other predictors. Each linear regression calculation in R is performed using the “lm” function which fits a linear model to the data. The formula arguments includes a stated response variable and at least one explanatory variable. In our algorithm, the training data observations of each meteorological variable is the response variable. The explanatory variables differ between meteorological variables used.

I. Air Temperature

Air temperature has the most explanatory variables which include air temperature means, minimums, and maximums for each previous day, current day, next day, previous time step, and next time step. This gives us a smooth transition between time steps and helps us capture the diurnal signature of temperature.

II. Shortwave Radiation

For Shortwave radiation (formally known as surface downwelling shortwave flux in air), we initially identify the time steps where there is no sunlight and make sure to save those linear regression models as a dark time step. This ensures that random noise at those time steps won’t exist so we don’t get shortwave radiation values at night when we actually downscale some data. It also helps constrain an otherwise double-ended zero-inflation issue unique to shortwave radiation. The only explanatory variables for shortwave radiation are the daily means.

III. Precipitation

Precipitation needs to be a bit different than the other variables because at some locations on Earth, the number of observations without measureable precipitation outweighs those with measureable precipitation, especially at a fine temporal resolution. We saw it necessary to use a scheme where we first calculate the fraction of precipitation occurring at each time step and estimate the probability distribution of rain occurring in that given time step. In this case, the probability distribution function is now used as the response variable. Our explanatory variable is the fraction of precipitation flux multiplied by the total daily precipitation flux.

IV. Air Pressure

The linear regression calculation for air pressure is straight forward. As with every variable besides precipitation, the variable of interest is used as the response variable. The explanatory variables include the means of the previous day, current day, and next day as well as the values for the previous and next time step.

V. Longwave Radiation

Longwave radiation (formally known as surface downwelling longwave flux in air) undergoes the same approach as air pressure with the daily means of previous day, current day, and next day and the values of the previous and next time step used as the explanatory variables.

VI. Specific Humidity

Specific humidity is the only variable that has a multivariate approach for the linear regression model calculation. We log transform the specific humidity variable and use that as the

response variable because it's zero bounded and log distributed. The explanatory variables for specific humidity include the daily mean for the current day, the previous and next time step values, and the daily air temperature maximum. We experimented with other predictors but the downscaled values fell well outside the observational dataset Gaussian distribution, leading us to believe these values were unrealistic.

VII. Wind Speed

Wind Speed is the final variable that we generate linear regression models for. Our response variable is the square root of the observational value for wind speed because of its non-negative bound. The daily means of the current day and the previous and next time step values are used as our explanatory variables.

A linear regression model is created for each time step of the training data. The linear regression models for each variable are saved for each day of the year as an Rdata file within the output folder. We save models by day to decrease memory demand and query them as needed in the following step. We save each variable in its own folder for organization. Now that we have generated the linear regression models for the variables in the training dataset, we are ready to utilize the *predict_subdaily_met()* function.

c. Predicting Subdaily Meteorology – predict_subdaily_met()

The second step to the TDM workflow is where we apply the temporal downscaling functions the coarse data previously downloaded using the linear regression models calculated from the training data. Logistical arguments must be passed such as the file paths to the dataset we wish to temporally downscale, the training dataset, and the location where the linear

regression models are stored. Each year for the data we downloaded is saved to an individual NetCDF file. In order to efficiently loop over each file, the date range of the coarse dataset must be specified. If we calculated residuals in *gen.subdaily.models()* and want to use them to add noise and thus uncertainty to our downscaled product, we will set this logical argument to true.

An argument is required that specifies the number of ensemble members we wish to generate. In order to propagate the uncertainty when we transform coarse data to a fine resolution, we create multiple ensemble members which are sampled from the probability distribution functions contained within the linear regression models for each time step. We are able to better quantify our uncertainty by generating multiple ensembles and we are able to transfer this through the ecosystem model execution by driving the chosen ecosystem model using each ensemble member generated.

Once the initial parameters have been defined, we are able to run the function that first opens the training dataset, which we use to keep our output file organized. The final downscaled file will have the same resolution, latitude, and longitude as the training dataset and by opening it first we obtain this information. Next, the date range of the coarse dataset is sequenced and a loop is initiated to read in and downscale one year at a time beginning with the first year. For the given year in the coarse dataset, we read in the previous year (if it exists), the current year, and the next year. By reading in the previous and next years, our transitions are coherent and smooth from year to year when we have a cut off at December 31 – January 1. For years where we are unable to access previous or next years (first year or last year of the dataset), we duplicate the closest available value. Once the appropriate data files have been read in and values have been saved, we are ready to undergo the subdaily prediction and generate multiple ensembles. During the prediction phase of the downscaling procedure, each variable is modeled separately.

However, the general scheme is the same for all, so we will describe the common workflow and then discuss the differences across variables.

The first step to the prediction phase is creating column propagation list that randomly chooses which ensemble member value to propagate onto the equation for the next value. This list has a length equal to the length of the desired temporal resolution and randomly samples between one and the number of ensembles based on the normal distribution. Similarly, a beta propagation list is created that randomly chooses which betas that will be selected from the linear regression models for each ensemble member. This list also has a length equal to the length of the desired temporal resolution and randomly samples between one and the number of betas that we calculated for the linear regression model generation. These lists are created at the beginning of the prediction phase so that these random selections are consistent across variables.

For each variable to be downscaled, we loop over the length of the desired time step beginning with the minimum time step value. Inside the loop, we grab the values from the previous time step. From each time step in that particular variable, we read in the appropriate day (saved a day at a time with information at the specific time step) from the linear regression model output and pull the information from that particular time step. Additionally, we read in the beta file for that day that we have saved and reference the beta propagation list to select the appropriate value to use for that time step.

Each time step that we are downscaling is predicted by first reading in the model terms, coefficients, covariances, and residuals. We extract the model frame from the linear regression model output that was used when fitting the model for that time step. Then, we create a model matrix for the time step using the model frame, the closest coarse data values and the previous and next downscaled time step values for the variable we are downscaling. Finally, we use the

linear regression approach detailed in equation 2 to simulate each ensemble member by multiplying the model matrix by the transposition of the beta value selected for each ensemble member.

Linear Regression Approach to Ensemble Downscaling

$$downscaled\ ensemble\ value = model\ matrix * ensemble\ beta\ value + residual\ error$$

Equation 2. *Linear regression approach to ensemble downscaling.*

If we calculated residual errors in linear regression model generation step, the error is added at the end of this equation. The predictand solved for in this equation for each ensemble member is the downscaled value.

Air temperature is downscaled uniquely in that we include the maximum and minimum temperature for the coarse data time step and the previous downscaled values. We use these to help us downscale air temperature and constrain our values by set boundaries for the maximum and minimum temperatures possible for downscaling. For shortwave radiation, we only model the time steps in the training data where this value is positive for that particular day. Downwelling shortwave radiation is only non-zero while the sun is out so we make sure we are only modeling time steps where shortwave radiation is greater than zero.

After the downscaled values are predicted, some variables undergo quality control procedures. Shortwave radiation and precipitation flux are zero truncated variables so we ensure that the values generated are non-negative. For longwave radiation, we need to square these values because we took the square root when generating the linear regression models to prevent negative values. We also set boundaries for longwave radiation by declaring that the values must be greater than 100 W meter^{-2} and less than 600 W meter^{-2}. These values were chosen semi-arbitrarily based on values seen at Willow Creek. However, this is not a long term solution and

we will be implementing a new boundary value creation routine in the future. Specific humidity needs to be log transformed because we took the natural log of these values when generating the linear regression models. Specific humidity can sometimes end up with high or infinite values so we constrain this by ensuring that the values in the 99^{th} percentile are less than 0.03 kilogram kilogram^{-1}. Wind speed needs to be squared since we square rooted these values at the model generation stage to prevent negative values. The values for each ensemble member at each time step must undergo these quality checks.

Once the ensemble predictions have been complete and our downscaled members are finished, we save each ensemble member for that specific year as a NetCDF file in CF metadata conventions. Each ensemble member can now be used as the meteorological driver for an ecosystem model.

2.3 Data and Models Used

Our experimental design (explained in Section 2.4) requires the use of two large datasets and one ecosystem model. We employed the use of the Willow Creek Flux Tower eddy covariance observational dataset (Cook et al., 2004) as well as the Multivariate Adaptive Constructed Analogs dataset which is a spatially downscaled version of the CMIP5 global climate model output. The SIPNET model (Braswell et al., 2005) was chosen as our ecosystem model for its fast processing time.

d. Willow Creek FLUXNET Dataset

FLUXNET is a global collection of ecological research sites that measure land-atmosphere exchange of carbon cycle and related atmospheric constituents that help us monitor the earth

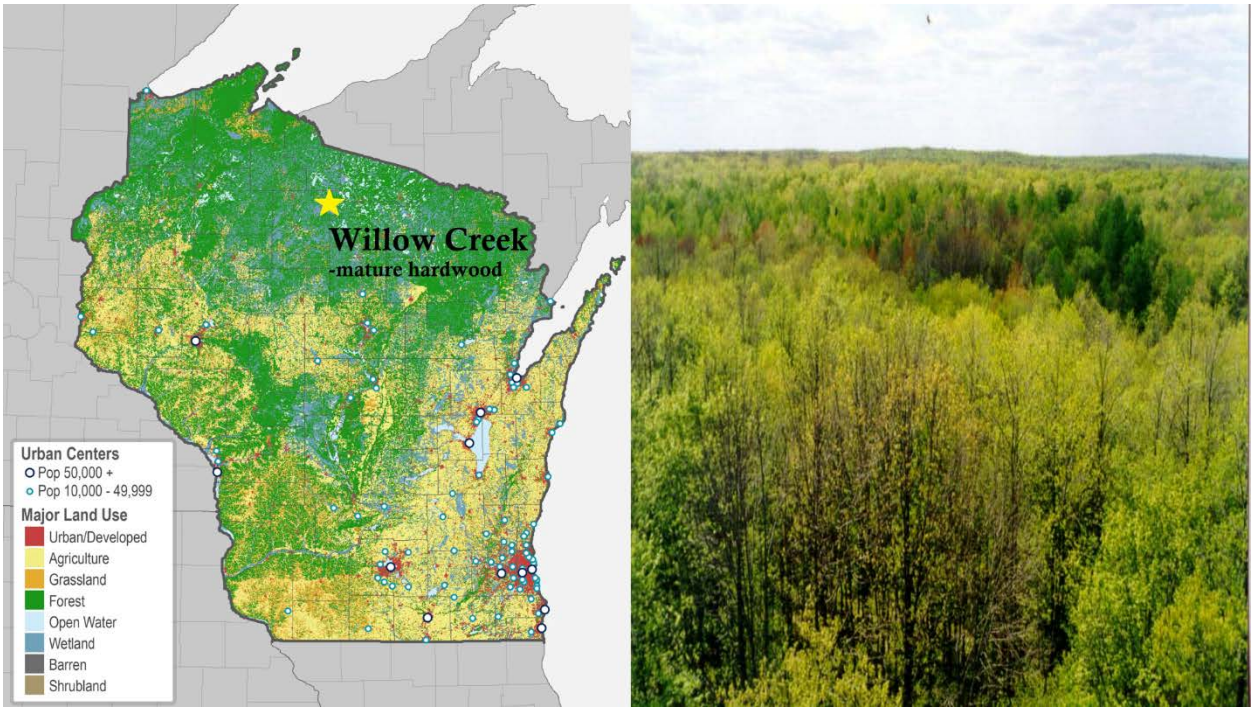

Figure 1. *Willow Creek location and view of phenology from the top of the flux tower. Land cover map of Wisconsin courtesy of Wisconsin Department of Natural Resources. View of forest canopy structure taken from http://cheas.psu.edu/.*

system (Baldocchi et al., 2001). The measurements obtained by the sensors in this network help us monitor meteorology, climate trends, the carbon cycle, and the water cycle. These sensors are most commonly placed on flux towers which are constructed by an individual party. The individual scientist or lab is responsible for ensuring the quality of the data through regular upkeep and calibration of instrumentation as well as rigorous post-collection quality controls. In 2016, the FLUXNET organization gathered all submissions for each tower in the network up to 2015. The data for each site went through an advanced gap filling algorithm to fill in points in the time series where data may have been unavailable. Then, the data was quality assured and published to the FLUXNET2015 data archive.

The Willow Creek Flux Tower stands in Northern Wisconsin and is located in a deciduous broadleaf forest. As described in Desai et al. (2005) and Cook et al. (2004), this mature hardwood forest is located in the Chequamegon-Nicolet National Forest, WI, USA (45°48′21″N,

90°04′47″W). The flux tower stands at 30 meters and has been measuring atmospheric state variables and eddy covariance metrics since 1999. We utilized PEcAn's extraction function, *download.Fluxnet2015(),* to attain 16 years (1999-2015) of hourly data from the Willow Creek flux tower. We will be using the variables listed in Table 1. These long standing flux tower observations at a high temporal resolution provide a rich dataset that we will use to test the TDM functions and use as training data for the downscaling.

e. MACAv2-METDATA Dataset

Multivariate Adaptive Constructed Analogs (MACA) is a statistical downscaling method used to spatially downscale CMIP5 global climate model output (Abatzoglou and Brown, 2011). The goal of MACA is to remove historical biases and match spatial patterns in climate model output by using a meteorological observation dataset. The GCM output from 20 GCMs of the CMIP5 experiment were spatially downscaled for the RCP4.5 and RCP8.5 scenarios to a spatial resolution of 4-KM.

Model Name	**Model Country**	**Model Agency**	**Atmosphere Resolution(Lon x Lat)**	**Ensemble Used**
bcc-csm1-1	China	Beijing Climate Center, China Meteorological Administration	2.8 deg x 2.8 deg	r1i1p1
bcc-csm1-1-m	China	Beijing Climate Center, China Meteorological Administration	1.12 deg x 1.12 deg	r1i1p1
BNU-ESM	China	College of Global Change and Earth System Science, Beijing Normal University, China	2.8 deg x 2.8 deg	r1i1p1
CanESM2	Canada	Canadian Centre for Climate Modeling and Analysis	2.8 deg x 2.8 deg	r1i1p1
CCSM4	USA	National Center of Atmospheric Research, USA	1.25 deg x 0.94 deg	r6i1p1

CNRM-CM5	France	National Centre of Meteorological Research, France	1.4 deg x 1.4 deg	r1i1p1
CSIRO-Mk3-6-0	Australia	Commonwealth Scientific and Industrial Research Organization/Queensland Climate Change Centre of Excellence, Australia	1.8 deg x 1.8 deg	r1i1p1
GFDL-ESM2M	USA	NOAA Geophysical Fluid Dynamics Laboratory, USA	2.5 deg x 2.0 deg	r1i1p1
GFDL-ESM2G	USA	NOAA Geophysical Fluid Dynamics Laboratory, USA	2.5 deg x 2.0 deg	r1i1p1
HadGEM2-ES	United Kingdom	Met Office Hadley Center, UK	1.88 deg x 1.25 deg	r1i1p1
HadGEM2-CC	United Kingdom	Met Office Hadley Center, UK	1.88 deg x 1.25 deg	r1i1p1
inmcm4	Russia	Institute for Numerical Mathematics, Russia	2.0 deg x 1.5 deg	r1i1p1
IPSL-CM5A-LR	France	Institut Pierre Simon Laplace, France	3.75 deg x 1.8 deg	r1i1p1
IPSL-CM5A-MR	France	Institut Pierre Simon Laplace, France	2.5 deg x 1.25 deg	r1i1p1
IPSL-CM5B-LR	France	Institut Pierre Simon Laplace, France	2.75 deg x 1.8 deg	r1i1p1
MIROC5	Japan	Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies,and Japan Agency for Marine-Earth Science and Technology	1.4 deg x 1.4 deg	r1i1p1
MIROC-ESM	Japan	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies	2.8 deg x 2.8 deg	r1i1p1
MIROC-ESM-CHEM	Japan	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies	2.8 deg x 2.8 deg	r1i1p1
MRI-CGCM3	Japan	Meteorological Research Institute, Japan	1.1 deg x 1.1 deg	r1i1p1

NorESM1-M	Norway	Norwegian Climate Center, Norway	2.5 deg x 1.9 deg	r1i1p1

Table 2*. CMIP5 Models Spatially Downscaled by MACA found at (http://maca.northwestknowledge.net/GCMs.php)*

The MACAv2-METDATA dataset is the result of the spatially downscaled CMIP5 output available over the entire coterminous United States at a daily temporal resolution from 2006 to 2100. All 20 GCMs in Table 2 were spatially downscaled and the meteorological variables in Table 3 are available for each GCM.

MACAv2-METDATA	
Variables	**Units**
Minimum Daily Temperature	Kelvin
Maximum Daily Temperature	Kelvin
Average Daily Precipitation	mm/day
Daily Average Surface Downwelling Shortwave Flux In Air	Watts meter^{-2}
Daily Average Specific Humidity	Kilogram Kilogram^{-1}
Daily Average Wind Speed	Meters second^{-2}

Table 3. *MACAv2-METDATA Variables and Units*

In this study, we downloaded a decade of data (2020-2030) for the RCP8.5 scenario for all 20 models available in the dataset using the *download.MACA()* function in PEcAn. The MACAv2-METDATA dataset is ideal for testing ecosystem response to future climate. The 20 spatially downscaled GCMs allow for a fine model intercomparison at 1/24th of a degree spatial

resolution. We will temporally downscale the MACAv2-METDATA to obtain a fine temporal and spatial resolution dataset.

f. SIPNET Ecosystem Model

The SIPNET model is a simple ecosystem model that can be used to explain the interaction between ecological biomes and the atmosphere (Braswell et al., 2005). This model has been built in a straightforward manner with the intention of being a basic, fast running ecosystem model. It stems from the PnET family of models (Aber and Federer 1992) and has been integrated into PEcAn. SIPNET is the fastest model PEcAn has and also requires the least amount of inputs. A user is able to select the meteorological driver and the PFT for a point location to run the model. PEcAn converts these selections to the climate and parameter input files in SIPNET format using the met2model() function. PEcAn also automatically generates the spatial parameter file and the SIPNET input file that holds information about the organization of the files created. The automated creation of these files make modeling with SIPNET fast, efficient, and straightforward.

SIPNET requires a number of meteorological variables to drive the model. These variables include air temperature, precipitation flux, shortwave radiation, humidity, and wind speed. Another convenient feature of SIPNET is that it doesn't require the meteorological data to be at any specific temporal resolution. This will allow us to compare SIPNET's output for varying temporal resolution meteorological datasets.

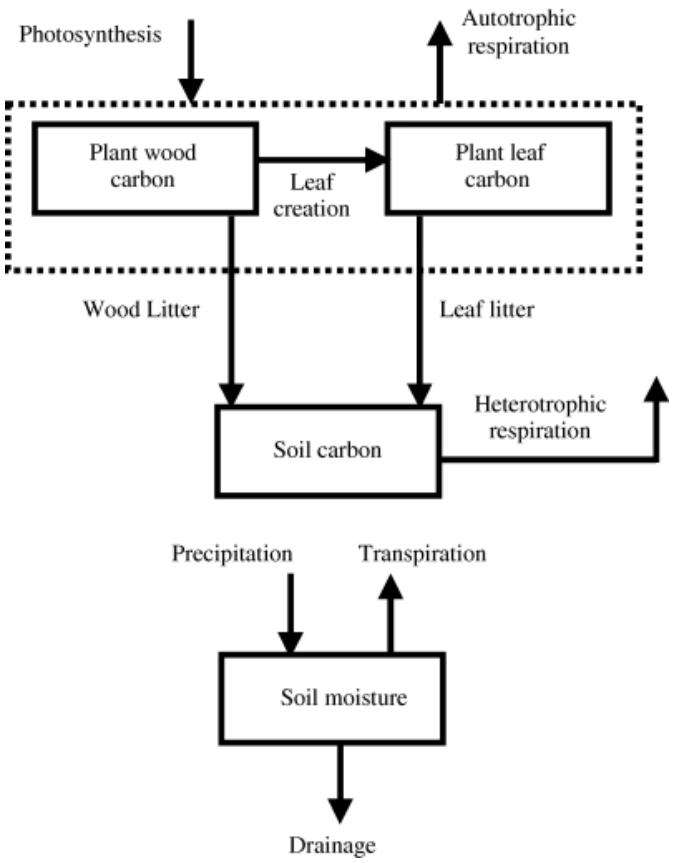

Figure 2*. Basic SIPNET workflow described in Braswell et al., 2005.*

2.4 Experimental Design

The goals of this study are: (i) to show evidence supporting the need for high resolution meteorological driver data, (ii) validate the temporal downscaling algorithm detailed above, and (iii) temporally downscale projected climate data to understand the effect future climate has on ecosystems. Each objective was approached in a specific manner and the methods employed for each goal will be detailed below.

g. The Importance of High Resolution Meteorological Data

The present study is centered on the argument that high resolution meteorological data is required to decrease carbon cycle uncertainty. In order to present evidence for this argument, we

ran the SIPNET ecosystem model in PEcAn using various resolutions of the same dataset. We took hourly resolution Willow Creek FLUXNET data for each year in our observational record (1999-2015) and aggregated the meteorology to 2 hour means, 3 hour means, 6 hour means, 12 hour means and daily means. We ran SIPNET with each of these resolutions for Willow Creek and accumulated daily NEE values for the entire year. The average yearly sum of NEE from 1999-2015 for each resolution are compared.

h. Temporal Downscaling Algorithm Validation

In order to test the validity of the temporal downscaling algorithm we wanted to see how well the routine performed with observational data. We aggregated 2006 hourly Willow Creek FLUXNET observations to a daily resolution and used the TDM workflow to downscale it back down to an hourly resolution. We chose to aggregate up to a daily resolution and downscale back to hourly because we will be temporally downscaling from a daily resolution to an hourly resolution in the next experiment. The year 2006 was chosen as the test year because it lies in the middle of our training dataset. A window day length of 5 days was selected and we generated the linear regression models with 100 betas. The window day length of 5 days allows us to provide the linear regression model calculations with sufficient data for precipitation. We chose 100 betas because of computational time restrictions. We compare the meteorology data between the downscaled ensemble outputs and the actual observations and evaluate the accuracy of the TDM procedure.

i. Future Ecosystem Responses at Willow Creek

Once we had validated the TDM procedure, we temporally downscaled climate projected data to understand the effect future climate has on ecosystems. We temporally downscaled each GCM in the MACAv2-METDATA dataset for the Willow Creek Flux Tower location. This dataset has a spatial resolution of 4-km, allowing us to extract spatially downscaled data near the location of the tower. We only temporally downscaled the RCP8.5 scenario, which has been coined the "business as usual" scenario. This scenario will be most accurate with current carbon emissions for the next decade given the low likelihood that immediate, drastic policy change on reducing greenhouse gas emissions would take place.

In order to temporally downscale this data, we trained it using the Willow Creek FLUXNET meteorological dataset. The 15 years of hourly resolution flux tower data offer a robust training dataset. We used the TDM workflow to temporally downscale each GCM in the MACAv2-METDATA dataset from a daily resolution to an hourly resolution for the years 2020-2030. We chose to perform this analysis for the next decade because SIPNET does not account for elevations in atmospheric carbon dioxide concentrations and the next decade will have similar levels compared to the end of the century. We generated 12 ensembles for each model and year downscaled to quantify our uncertainty. Consistent with our validation step, we chose a window day length of 5 days and chose to use 100 betas for the same reasons listed previously.

Once all of the data went through the TDM workflow and the ensemble meteorology was generated, the data was then debiased based on evaluations of the downscaling algorithm performance detailed above. Specifically, we debias the downscaled precipitation fluxes by multiplying by the percent error between the mean of the observed precipitation and the downscaled precipitation (17.75%). The other variable comparisons showed that we were within 1.4% of the mean for each and thus decided it was not worth debiasing. We ran each year of each

ensemble member through SIPNET to observe future NEE. We compare how differences in meteorological driver data affects the SIPNET model output and thus the future carbon cycle. By temporally downscaling spatially downscaled data, we are able to get our best estimate of NEE in the decade of 2020-2030. It is important to note that the MACAv2-METDATA dataset did not have longwave radiation or air pressure available. SIPNET does not require these meteorological variables but a more sophisticated model may need these variables to be specified.

3. Results

3.1 NEE Response to Varying Temporal Resolution Meteorology

We begin by comparing the SIPNET NEE output for varying temporal resolutions of Willow Creek flux tower observations. Figure 3 shows the SIPNET output for average yearly cumulative NEE for each temporal resolution aggregation and the actual observations of NEE.

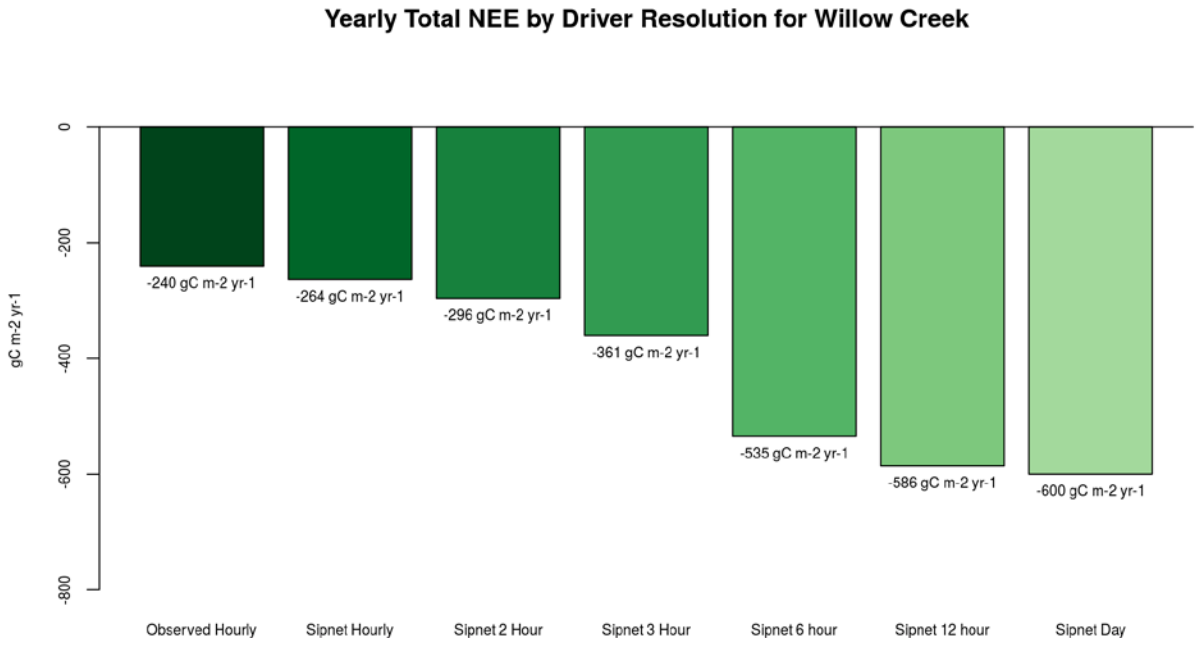

Figure 3*. Yearly total NEE by driver resolution for Willow Creek. Each resolution of Willow Creek observations from 1999-2015 was run using SIPNET. Average yearly total NEE from SIPNET output is compared to NEE observations.*

To ensure that each resolution of NEE output was being compared fairly, we computed the daily average sums for each SIPNET run to account for the models tendency to show high diurnal variability. The daily sums for each SIPNET run were then summed for each year of output and averaged across the date range. The same procedure was used for the hourly observations of NEE at Willow Creek. When we compare the observed hourly NEE with the SIPNET NEE from hourly meteorology, we find that SIPNET overestimates carbon uptake by about 24 gC m-2 yr-1, or 10%. When we compare observed hourly NEE with the SIPNET NEE from daily meteorology, we find that SIPNET overestimates carbon uptake by 360 gC m-2 yr-1, or 250%. A progression of carbon uptake overestimation is evident as we decrease the temporal frequency of the meteorological driver data.

Figure 4. *Boxplots of yearly total NEE for each year (1999-2015) in observed NEE and SIPNET output NEE for various aggregated Willow Creek observations.*

Figure 4 shows boxplots that highlight the range in the yearly sum of NEE for each year at each resolution. As we decrease our temporal frequency, the variance in yearly NEE output from SIPNET grows. A poor temporal resolution meteorological driver increases interannual variability. By running SIPNET at an hourly resolution, we decrease our uncertainty for interannual variability for yearly sums of NEE.

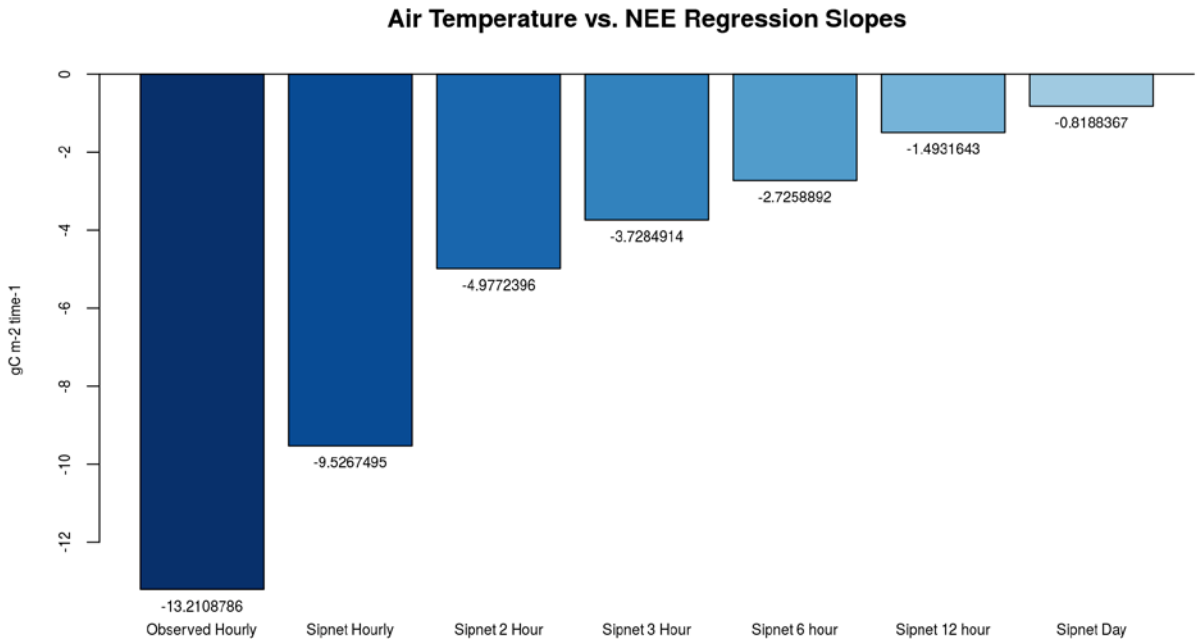

Figure 5*. Air Temperature vs. NEE regression slope magnitudes comparing actual observations and the various aggregations used for SIPNET.*

Figure 5 shows magnitudes of the linear regression slopes for observed NEE and temperature against the various temporal resolutions of SIPNET NEE and temperature. Temperature can be considered the most important driver of NEE because it controls plant photosynthetic efficiency and drives total soil respiration. As we decrease our temporal frequency, we decrease the variation in temperature by smoothing between the fine temporal variation and thus alter its photosynthetic efficiency and NEE. The magnitude of regression slopes drastically increases and moves further away from observations as we decrease our temporal frequency. Furthermore, Figure 5 shows the magnitude of covariances between air temperature and NEE for observations and the various SIPNET runs. The covariance structure becomes progressively disjointed and moves further from reality at lower temporal frequencies.

Since temperature is an important driver for photosynthetic efficiency, we need to preserve the covariance between it and NEE.

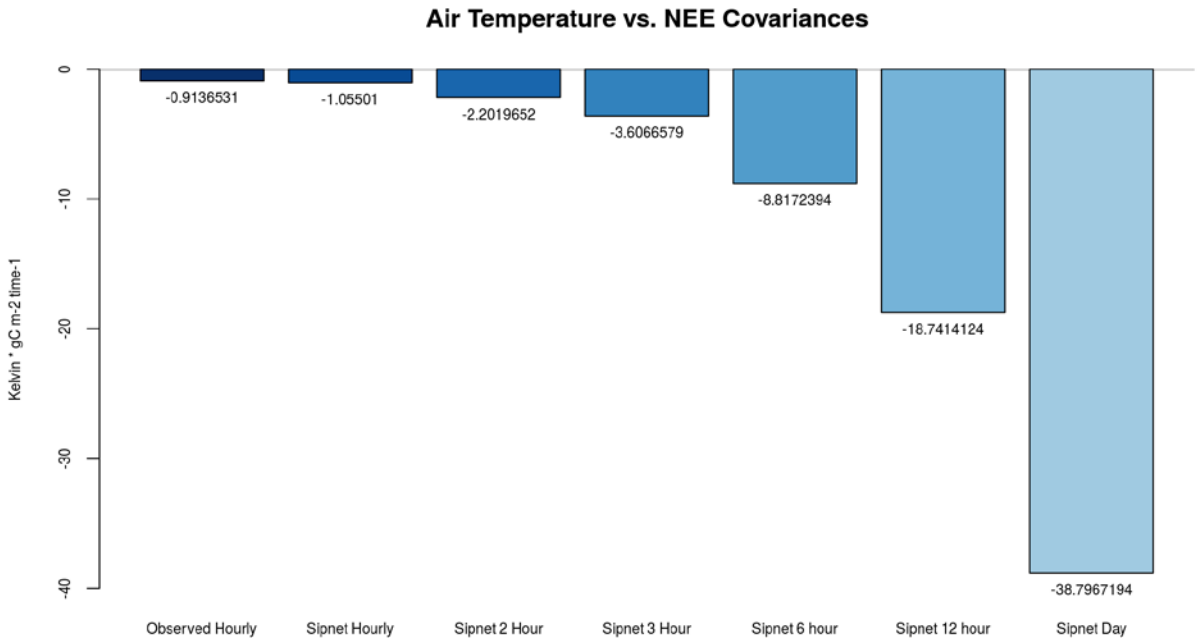

Figure 6*. Air Temperature vs. NEE covariance magnitudes comparing actual observations and the various aggregations used for SIPNET.*

In order to decrease our uncertainty for yearly sums of NEE, it's necessary to use high temporal frequency driver data for the ecosystem model. Figures 3-6 show evidence supporting the notion that finer temporal resolution meteorological data should be used to bring our SIPNET output closer to observed, both for means, for interannual variability, and for environmental sensitivity of the carbon cycle to climate.

3.2 Temporal Downscaling Algorithm Validation

Validating the temporal downscaling procedure was necessary to show that we trust the algorithm to reproduce meteorology. Each figure or table in this section will be comparing the

hourly observed meteorology and the ensemble averaged downscaled meteorology. Tables 4 and 5 show mean percent changes between the observations and the downscaled ensemble member average. A positive percentage value indicates a larger observational meteorology value and a negative value indicates a larger ensemble average downscaled value.

Mean Percent Change of Summary Statistics

	Temperature	Precipitation	Shortwave Radiation	Longwave Radiation	Air Pressure	Specific Humidity	Wind Speed
Minimum	−0.743	0	0	6.924	0.337	−93.646	0
1st Quartile	0.037	0	0	−0.724	−0.011	−2.507	−3.906
Median	0.072	0	−53.881	0.232	0.032	−5.955	1.532
Mean	−0.036	17.749	−0.526	−0.033	0	−1.395	0.597
3rd Quartile	−0.035	0	−5.091	0.469	0.021	−2.722	3.376
Maximum	−0.13	84.411	2.332	1.131	−0.186	−14.811	−34.036

(Observed − Ensemble Avg)

Table 4. *Mean Percent Change of the summary statistics for the year of 2006 highlighting the differences between the observed meteorology and the downscaled ensemble average.*

The summary statistics show small differences apart from precipitation flux and specific humidity. For precipitation, the mean value and maximum value for the observations are 17.75% and 84.41% higher, respectively. Precipitation is a difficult meteorological variable to downscale because at Willow Creek, like most mid-latitude locations, consistent patterns are rare at the hourly timescale. The timing and intensity of precipitation events is highly variable and the number of precipitable hours is largely outweighed by the number of dry hours. The lack of wet hours and in particular intense, convectively driven precipitation events causes the algorithm to distribute a daily value of precipitation throughout the day rather than capture a single, intense hourly value. Statistically speaking, the training data shows it to be more likely to have equal precipitation throughout the day rather than the short-lived but intense precipitation event. This is backed by our methodology which chooses to use a normal distribution to select the most

probable times of precipitation based on the observational dataset. Ultimately, this shifts our distribution of precipitation amounts towards smaller values that occur with a higher frequency.

The yearly total sums for each variable is not reported in Table 4 as this is only applicable to precipitation. The mean percent change of the sum of precipitation flux is 17.75% (consistent with the precipitation flux mean). The year of 2006 was only 2% below our average precipitation total values throughout the datatset, suggesting that our downscaling algorithm underestimates precipitation totals. When we take a look at the downscaled precipitation output and compare it to the observations for the entire 16 year period, a low precipitation bias of around 17.7% is consistent throughout each ensemble member in each year.

Specific humidity for the ensemble averages seem to have a negative bias at the tails, but the mean is within 1.4% of the observed meteorology. This gives us confidence that while the tails aren't captured, the overall mean is well modeled. Observed meteorology and the ensemble average of the downscaled data demonstrate reasonable agreement for all other variables. Most notably, the temperature downscaling procedure performs well in recapturing the means, minimums and maximums. This gives us confidence in the downscaling approach because temperature is an important driver for ecosystem response.

Mean Percent Change of Covariances

	Temperature	Precipitation	Shortwave Radiation	Longwave Radiation	Air Pressure	Specific Humidity	Wind Speed
Temperature	−0.05	0.2	1.58	1.17	−9.84	0.14	−3.81
Precipitation	0.2	91.29	85.28	35.74	33.7	5.26	79.25
Shortwave Radiation	1.58	85.28	5.45	−8.89	−1.71	1.61	36.04
Longwave Radiation	1.17	35.74	−8.89	9.92	14.34	0.21	−5.79
Air Pressure	−9.84	33.7	−1.71	14.34	0.75	21.17	0.46
Specific Humidity	0.14	5.26	1.61	0.21	21.17	−3.2	13.12
Wind Speed	−3.81	79.25	36.04	−5.79	0.46	13.12	10.22

(Observed − Ensemble Avg)

Table 5. *Mean percent change of the covariances for the year of 2006 highlighting the differences between the observed meteorology and the downscaled ensemble average.*

Propagating through the covariances through the downscaling procedure was important for realistic meteorological modeling. Despite our univariate linear regression modeling approach (aside from specific humidity), the mean percent change in covariances between observed meteorology and ensemble averaged downscaled meteorology are relatively small across all variables barring precipitation and its relationship with other variables. It is important to note that precipitation will NOT be debiased until we test the future carbon cycle response using the MACAv2-METDATA. This suggests that our approach propagates the covariances through and shows evidence that the relationships between variables is not lost during the procedure.

2006 Ensemble Averaged Forecast Accuracy

	Temperature	Precipitation	Shortwave Radiation	Longwave Radiation	Pressure	Specific Humidity	Wind Speed
Mean Error	0.055851	−3.8e−06	0.7762185	0.0048011	−0.0408408	8.95e−05	−0.0159113
Roor Mean Square Error	1.5928022	0.0001858	55.2400673	21.6842581	212.7199961	0.0012185	0.743945
Mean Absolute Error	1.2077908	3.04e−05	25.880918	16.905987	151.1461762	0.0008022	0.5642744
Mean Percentage Error	0.0182959	−32.4591481	−Inf	−0.1231403	−0.0002719	−0.9645232	−3.2951159
Mean Absolute Percentage Error	0.4339237	201.310305	Inf	5.9006034	0.1593654	14.7217113	22.8875422
t−test statistic	3.2836979	−1.9155156	1.3152246	0.0207216	−0.0179685	6.8957284	−2.0021233
p−value	0.0010286	0.0554593	0.1884689	0.9834682	0.9856644	0	0.0453022

Table 6. *2006 forecast accuracy metrics between the ensemble averaged downscaled meteorology and the observations. A paired t-test was used to calculate values for the t-test statistic and p-value. Results were rounded to the 7th decimal place.*

In Table 6, we evaluate the forecast accuracies for each meteorological variable we downscaled. We chose metrics described in Hyndman and Koehler 2006 to evaluate our forecast accuracies. We calculated mean error (ME), root mean square error (RMSE), mean absolute error (MAE), mean percentage error (MPE), mean absolute percentage error (MAPE), a paired t-test, and a p-value to help us determine the forecast accuracy of the downscaled meteorology. The values of ensemble averaged temperature show high accuracy to the observed temperature values. A popular metric for investigating forecasting accuracy is the p-value. P-values of below .05 indicate a statistical significance that rejects the null hypothesis which in this case is that the

models are disagreeing. Precipitation seems to be accurate and has a p-value suggesting high accuracy. However, this is probably due to the agreement of dry hours between the ensemble averaged values and the observations as we've seen from previous tables that the mean and maximum values are considerably different.

Shortwave radiation has a high p-value and this is most likely due to the fact that the downscaled values underestimate the maximum daily shortwave radiation values by 2.3%. Another caveat of the shortwave radiation downscaling is the smoothed diurnal cycle. TDM shows that it's unable to accurately predict abrupt variations in sub-daily shortwave radiation. This is an artifact of the training dataset which has 16 years of hourly values and doesn't have consistent variability of subdaily values for any variable.

The high p-values of longwave radiation and air pressure are contradicted by the small mean percentage error values. Specific humidity and wind speed values are accurately forecasted by the metrics shown above. Overall, forecasting hourly data from a daily resolution dataset can be done with relative accuracy using our methodology. Our results are consistent with Maurer et al. (2010) that show a linear regression approach abrupt hourly shifts per variable would be hard to capture given their inconsistent nature in the training data.

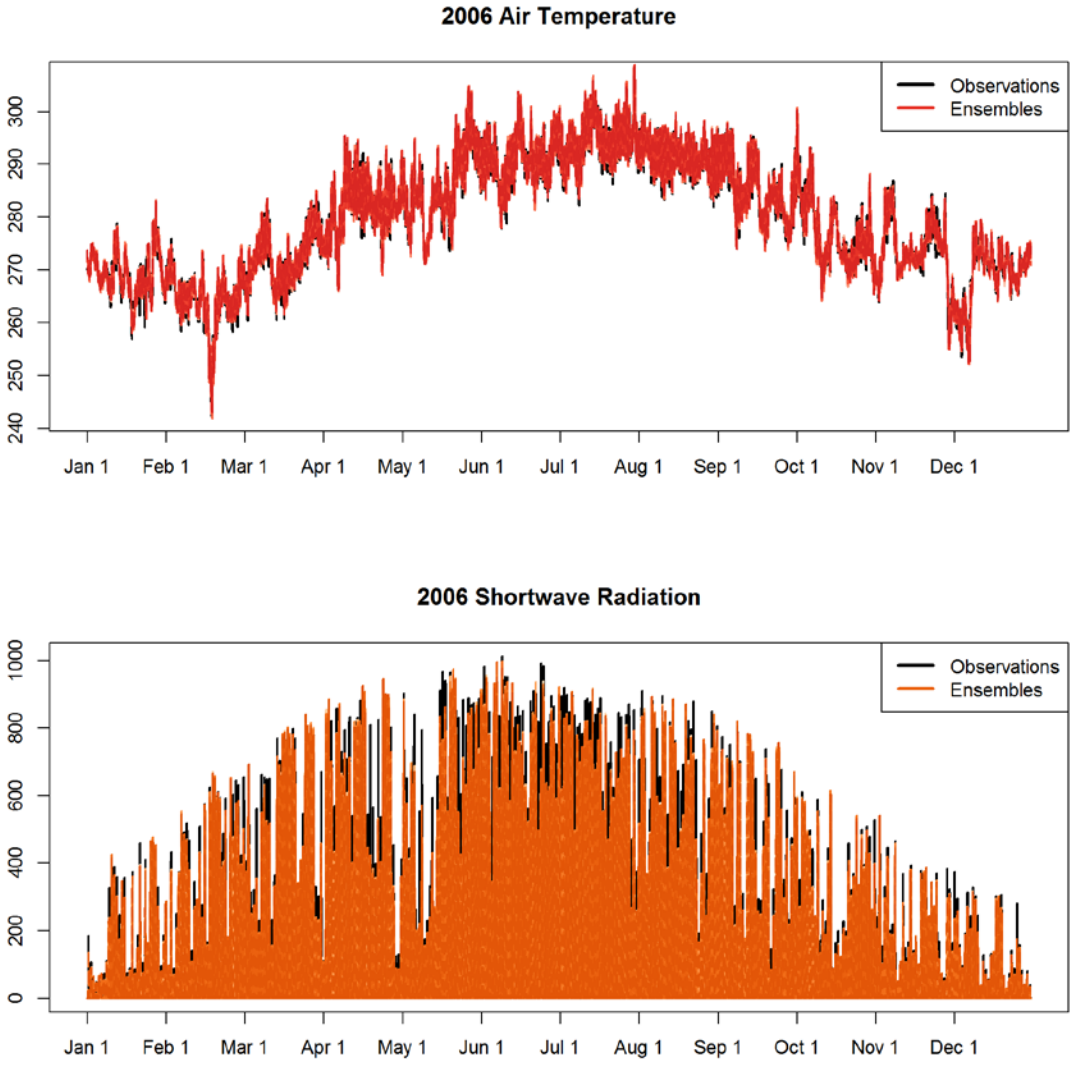

Figure 7*. Yearly time series of temperature and shortwave radiation for 2006. Observed meteorology is colored in black and the ensemble averages are colored in red for temperature and orange for shortwave radiation.*

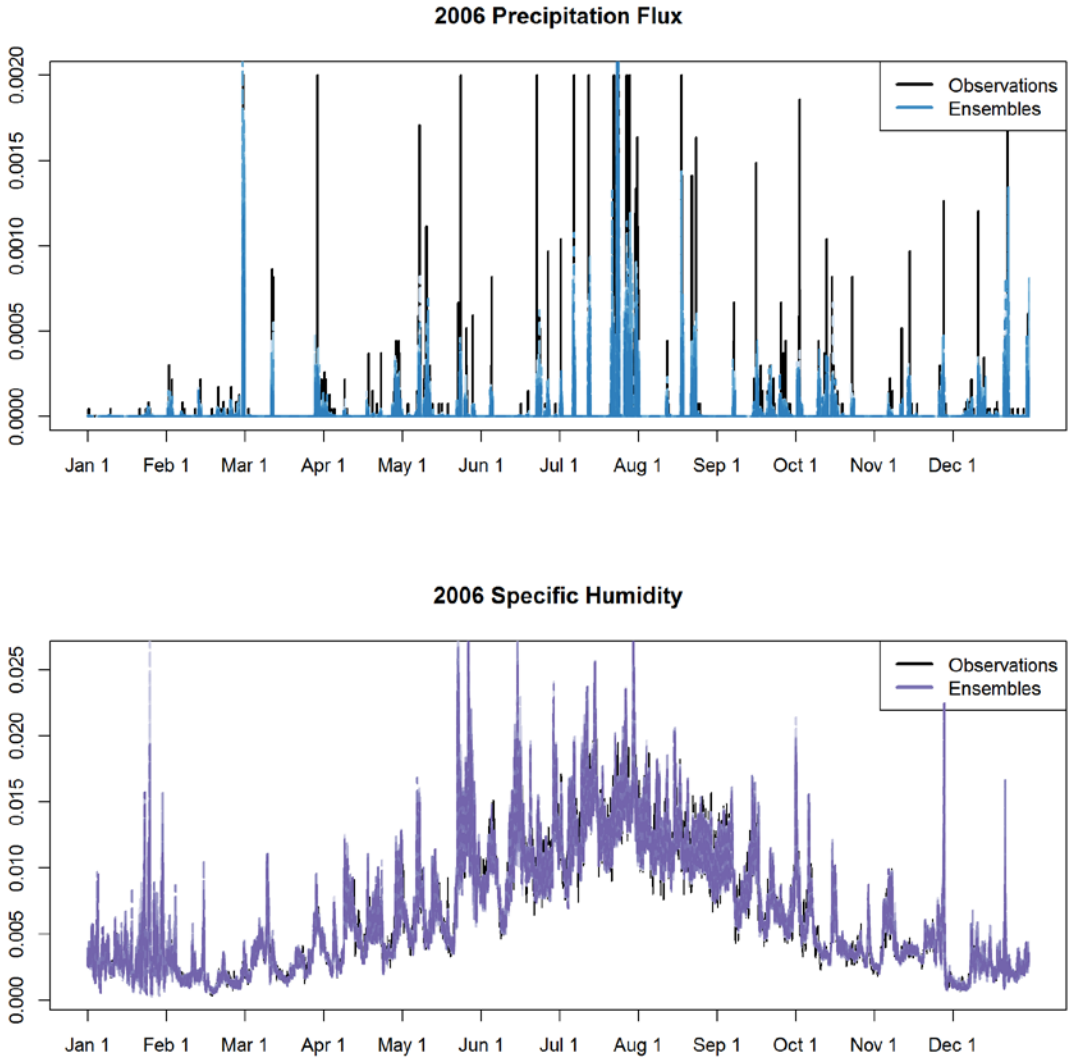

Figure 8*. Yearly time series of precipitation flux and specific humidity for 2006. Observed meteorology is colored in black and the ensemble averages are colored in blue for precipitation and purple for specific humidity.*

Figures 7 and 8 provide a visual representation of a yearly time series for temperature, shortwave radiation, precipitation flux, and specific humidity. The ensemble members do a fine job of modeling temperature, shortwave radiation, and specific humidity. For precipitation, however, one can see that even on a yearly scale, the magnitudes of the precipitation fluxes are not well captured for reasons described above.

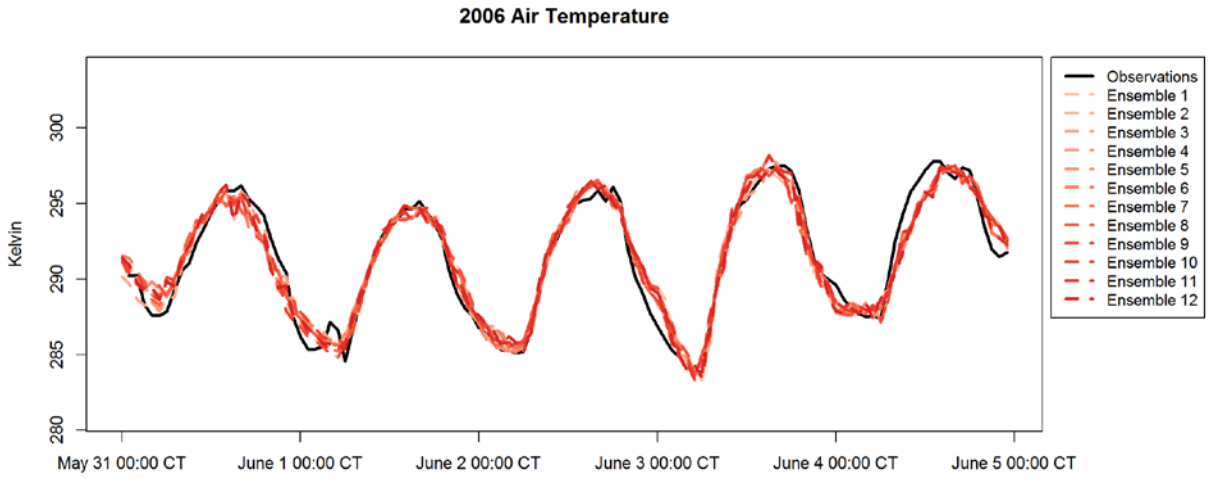

Figure 9*. May 31 - June 5 2006 air temperatures comparing the observed values to the downscaled ensemble member values.*

When we plot air temperature with a length of 5 days, we can show how the ensemble members are capturing the diurnal signature of temperature. We chose to show 5 days at the beginning of meteorological summer due to its importance for the growing season. The activity during the early stages of the growing season can have significant contributions to the yearly carbon cycle budget as shown in Wolf et al. (2016). Obtaining the diurnal cycle for temperature was a priority during our development of the algorithm. There is strong agreement between observations and ensemble members here, thus suggesting that the algorithm is doing a good job at capturing the diurnal cycle of temperature.

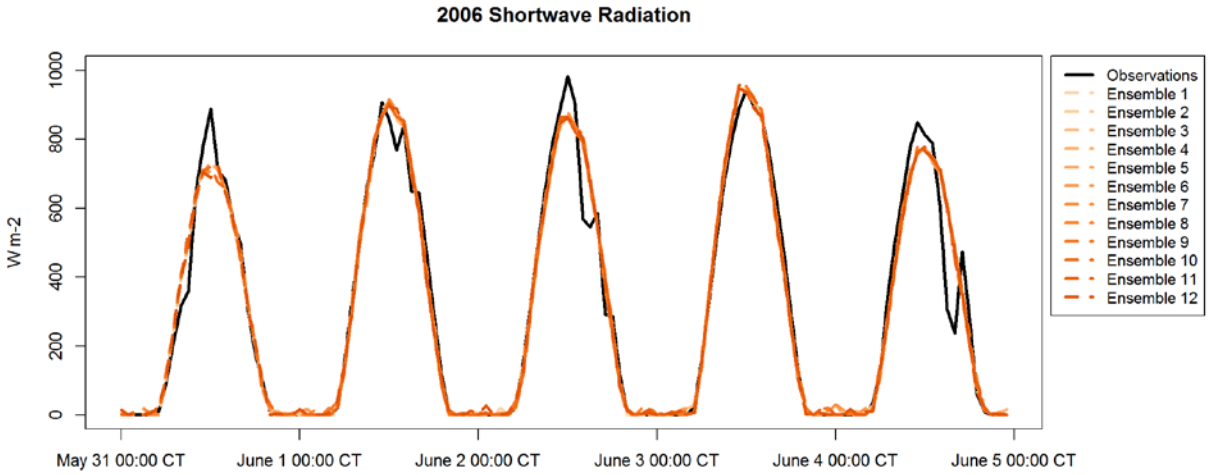

Figure 10. *May 31 - June 5 2006 shortwave radiation plot (surface downwelling shortwave flux in air) comparing the observed values to the downscaled ensemble member values.*

Figure 10 shows the same 5 day span as Figure 9 but for shortwave radiation. The diurnal cycle is evident in both the ensemble members and the observations. The ensemble members do a fine job at capturing the timing and magnitudes of the observations apart from the maximum value and abrupt shifts likely caused by clouds. The maximum values of the ensemble members falls short of the observed maximums for May 31, June 2, and June 4. Further, sudden diminished values of shortwave radiation in the observations are not well captured by the ensemble members.

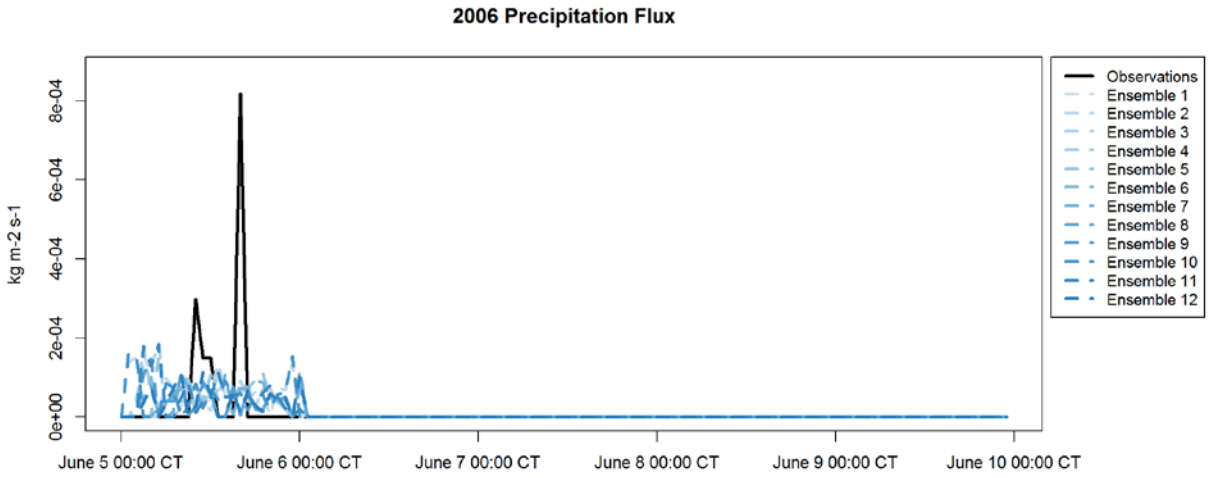

Figure 11. *June 5 – June 10 2006 precipitation fluxes comparing the observed values to the downscaled ensemble member values.*

Figure 11 shows another 5 day period in the early stages of the growing season. We had to push our 5 day time series forward another 5 days so we could show a day with precipitation. This figure provides a visual representation of the models inability to capture hourly precipitation spikes. Rather than model a few hourly values with higher precipitation values, the ensembles spread the precipitation out throughout the day as directed by the normal distribution modeling step for subdaily precipitation.

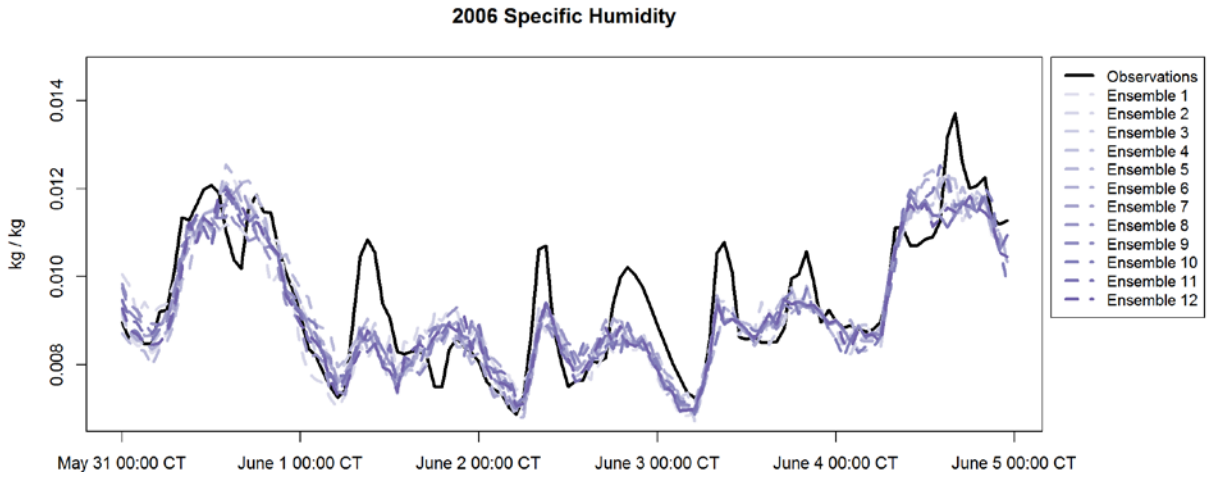

Figure 12. *May 31 - June 5 2006 specific humidity plot comparing the observed values to the downscaled ensemble member values.*

Figure 12 shows May 31 – June 5, 2006 hourly values of specific humidity for the observations and downscaled ensemble members. The ensemble members do well at following the general pattern of subdaily specific humidity but has trouble obtaining the magnitudes that the observed values have. Overall, it is encouraging to be able to capture these hourly patterns by downscaling a single daily value, even if the magnitudes are slightly suppressed.

The final step of our validation experiment was using the downscaled meteorology ensembles to drive SIPNET. Now that we are applying the downscaled meteorology to SIPNET, we decided to debias the downscaled precipitation by my increasing the values of each ensemble member by 17.75%. In order to decrease uncertainty for the carbon cycle, it's necessary that we correct for the precipitation mean disparity between the downscaled meteorology and the observations. While accounting for the low precipitation bias by raising the means will help our accuracy, further work should be done on improving the downscaled precipitation maximum values.

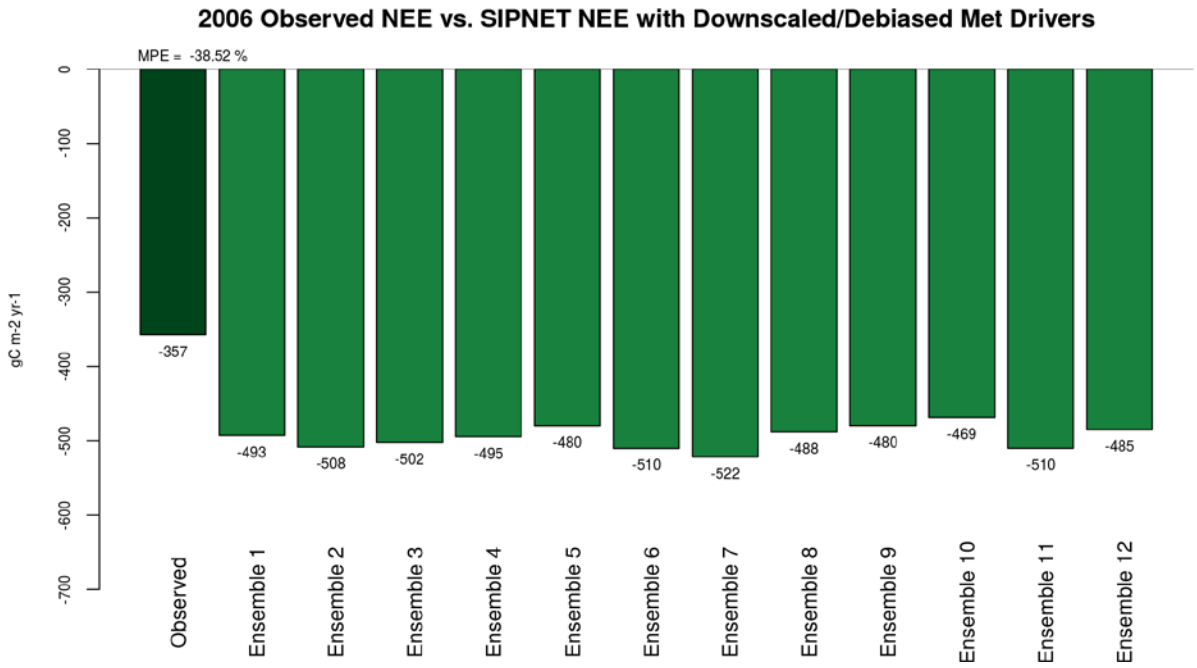

Figure 13. *2006 cumulative observed NEE vs. SIPNET cumulative NEE with downscaled/debiased met drivers. The average MPE for the ensemble members compared to the observed values is -38.52%.*

Figure 13 compares the observed NEE to the SIPNET NEE that was run using downscaled and debiased ensemble drivers for the year of 2006. The ensemble average mean percent error was -38.53% showing that our downscaled values seem to overestimate carbon sequestration. While this figure highlights the 2006 analysis, we need to perform this same analysis for every year in our dataset to quantify our uncertainty. Each year in our dataset followed the protocol detailed in this section of the results. We aggregated each year to a daily resolution and downscaled back to an hourly resolution. Running SIPNET with 12 downscaled and debiased ensemble members and averaging each ensemble member for each year between 1999 and 2015 will give us a more robust error estimate.

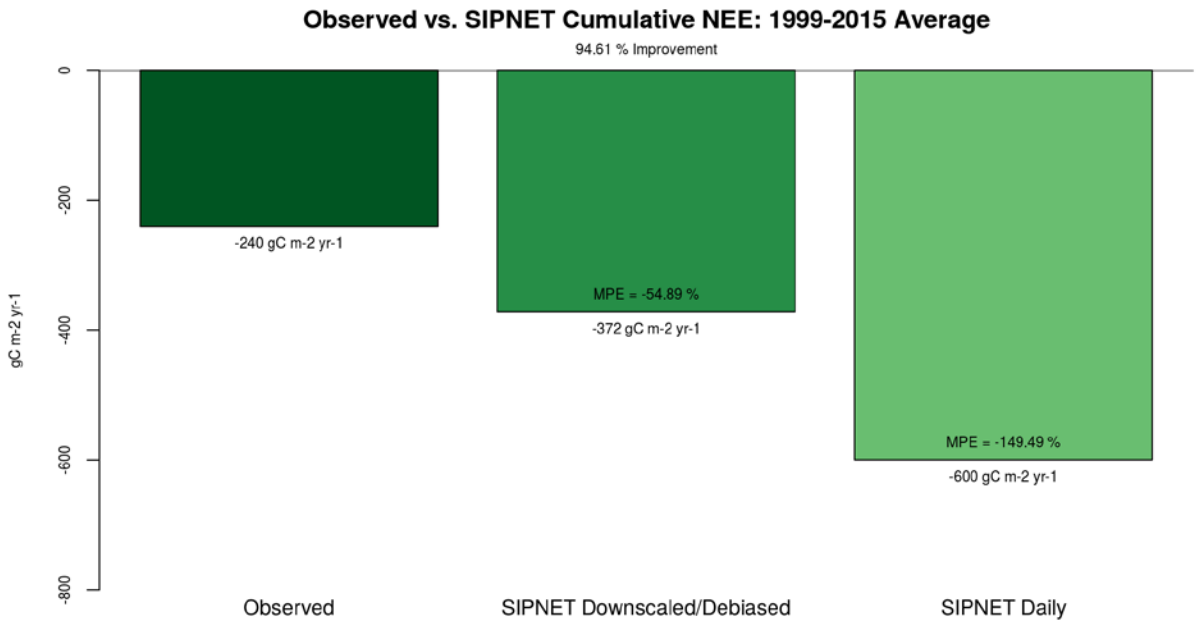

Figure 14. *Yearly cumulative NEE for observed data, SIPNET downscaled and debiased data, and SIPNET daily data.*

Figure 14 compares yearly cumulative NEE for the observed NEE, SIPNET NEE from downscaled and debiased hourly resolution data, and SIPNET NEE from the daily averaged resolution data. Our downscaled and debiased SIPNET NEE output shows that we overestimate carbon sequestration by 54.89%. However, this is significantly closer to the observed NEE compared to the daily resolution SIPNET NEE. By downscaling our data from a daily resolution to an hourly resolution and debiasing precipitation, we improve our accuracy by 94.61% and bring our values closer to observations.

3.3 Future NEE at Willow Creek

Future carbon cycle estimates come with high uncertainty. It's important to quantify this uncertainty and constrain it. After downscaling each CMIP5 model within the MACAV2-METDATA dataset, we debiased the precipitation fluxes by increasing the mean by 17.75%. The

results of SIPNET NEE output for each ensemble member of each year in each model will now be shown.

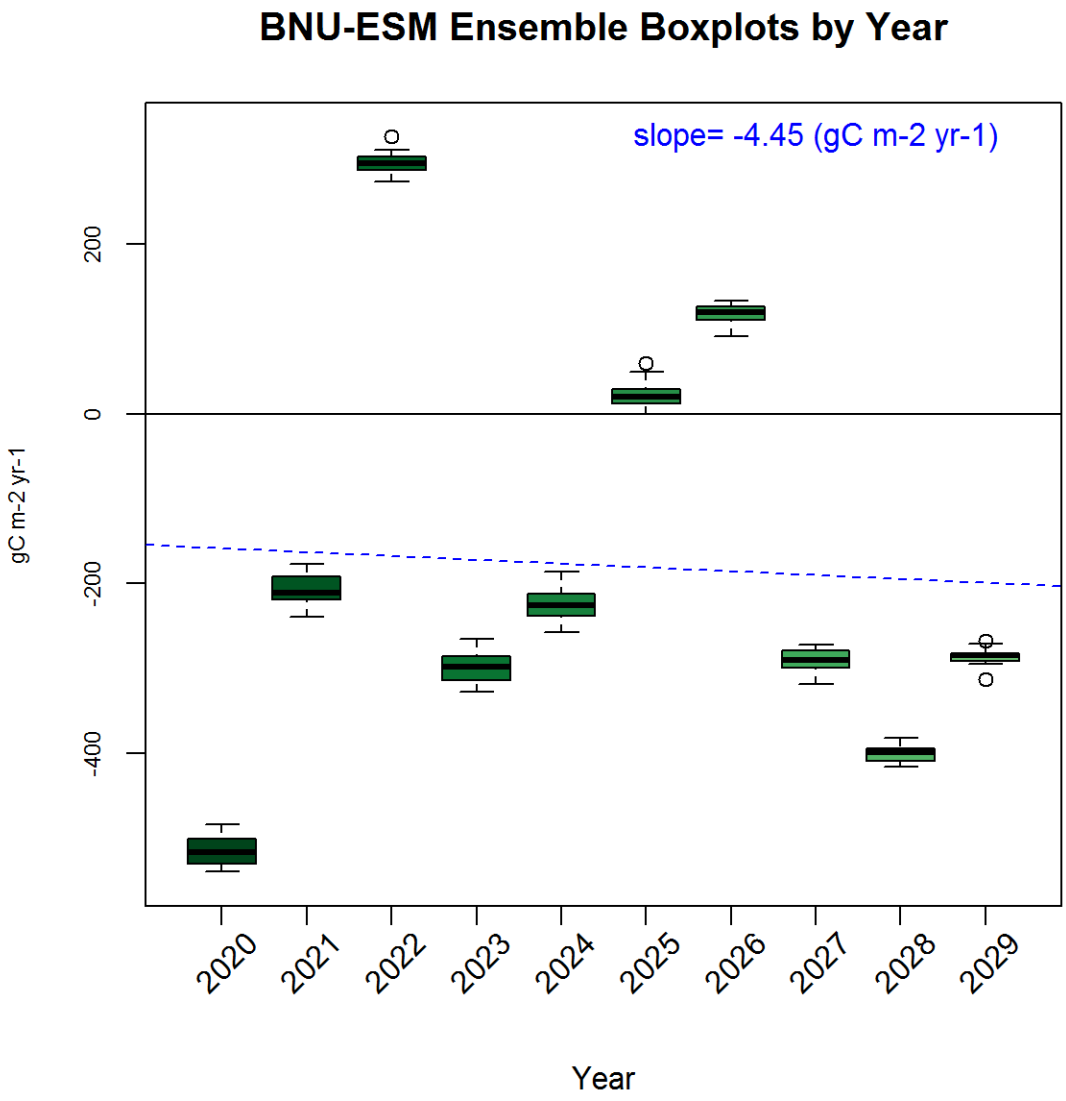

Figure 15*. Example of ensemble boxplots of yearly cumulative NEE from 2020-2029. This is for the MACAv2-METDATA BNU-ESM global climate model that was temporally downscaled using the TDM routine and then put through SIPNET.*

Figure 15 shows the temporally downscaled BNU-ESM ensemble boxplots for yearly cumulative NEE for each year from 2020 to 2029. This plot was made for each CMIP5 model so

that we could analyze interannual variability and the regression line of the decade. Figure 15 serves as an example for how we derived our cumulative NEE regression slopes for each model.

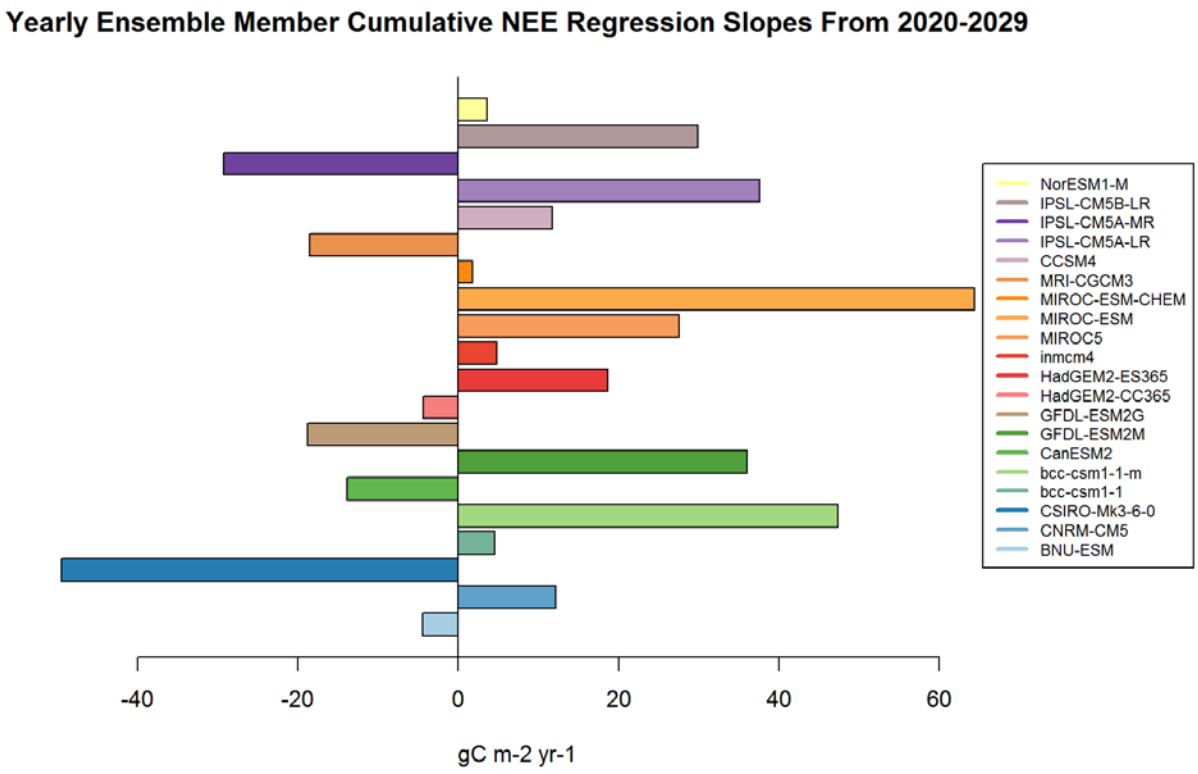

Figure 16. *Ensemble member regression slopes for yearly cumulative NEE. A positive slope indicates throughout the decade we are trending towards a carbon source. A negative slope indicates that we are trending towards a carbon sink.*

Figure 16 shows the divergence of ensemble member regression slopes for yearly cumulative NEE. Each CMIP5 model that was downscaled shows a wide range of regression slope magnitudes which ultimately tells us whether or not the model is trending Willow Creek towards a carbon source or carbon sink. A negative slope indicates that we are trending towards a carbon sink and a positive slope indicates that we are trending towards a carbon source. For this quantification, we will consider slopes between +-10 gC m-2 y-1 to be carbon neutral. Of the 20 CMIP5 models, 5 show we are trending towards a carbon sink, 6 are carbon neutral, and 9

show that we are trending towards a carbon source. The average model has a slope around of 5 gC m-2 y-1.

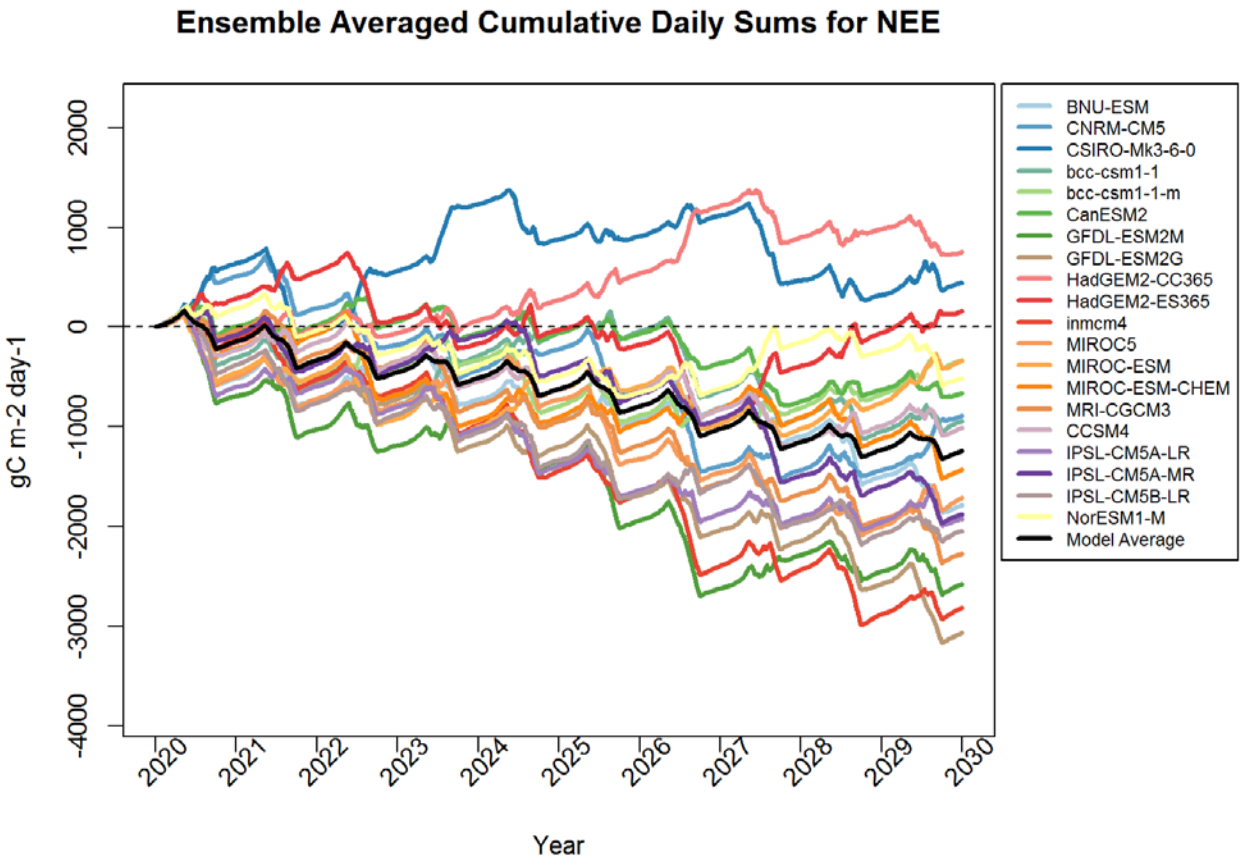

Figure 17. *2020-2030 ensemble averaged cumulative daily sums for NEE. Each year in the decade for each MACAv2-METDATA model was temporally downscaled from a daily to an hourly resolution. We generated 12 ensembles for each year and ran this data using SIPNET.*

Figure 17 shows the ensemble averaged cumulative daily sums of NEE for each CMIP5 model. This offers us insight into the total amount of carbon sequestered or respired during the decade. Despite the regression slopes showing that more models will become weaker carbon sinks throughout the decade, we see that we are still taking up carbon throughout that time period. The divergence of the cumulative daily sums shows ranges of nearly -3000 gC m-2 to 1000 gC m-2. Such a wide range indicates how sensitive even a basic ecosystem model like

SIPNET can be to different meteorology. This divergence also shows evidence that even over the relatively short time period of a decade, the model you choose to forecast ecological response to climate change can come with significant bias. In order to effectively quantify our uncertainty for future carbon cycle, multiple GCMs if not the entire suite of CMIP5 models is necessary for future ecosystem modeling. The closest model to the CMIP5 average cumulative daily sum of NEE was MIROC-ESM-CHEM which is an atmospheric chemistry coupled version of MIROC-ESM. The model showing the highest carbon source is HadGEM2-CC365 and the largest carbon sink is GFDL-ESM2G.

Comparing the differences between each CMIP5 models physics, parameters, and systematic biases is beyond the scope of this paper. Instead, we will investigate the differences between the downscaled hourly ensemble meteorology for HadGEM2-CC365 and GFDL-ESM2G, each tail of Figure 17s spread. As described in Dietze (2017), the factors contributing to predictive variance in ecosystem modeling are depicted in equation 3.

$$Var[Y_{t+1}] \approx \underbrace{\left(\frac{\partial f}{\partial Y}\right)^2}_{stability} \underbrace{Var[Y_t]}_{\substack{IC \\ uncert}} + \underbrace{\left(\frac{\partial f}{\partial X}\right)^2}_{\substack{driver \\ sens}} \underbrace{Var[X]}_{\substack{driver \\ uncert}} + \underbrace{\left(\frac{\partial f}{\partial \theta}\right)^2}_{\substack{param \\ sens}} \left(\underbrace{Var[\bar{\theta}]}_{\substack{param \\ uncert}} + \underbrace{Var[\alpha]}_{\substack{param \\ variability}} \right) + \underbrace{Var[\varepsilon]}_{\substack{process \\ error}}$$

$$= INTERNAL + EXTERNAL + \left(PARAMETERS + \frac{RANDOM}{EFFECTS}\right) + \frac{PROCESS}{ERROR}$$

Equation 3. *Factors contributing to predictive variance in ecosystem modeling (Dietze 2017).*

The focus of this paper is centered on driver sensitivity and uncertainty (external factors). By analyzing the differences between HadGEM2-CC365 and GFDL-ESM2G, an approach to exploring the driver sensitivity and uncertainty is illustrated.

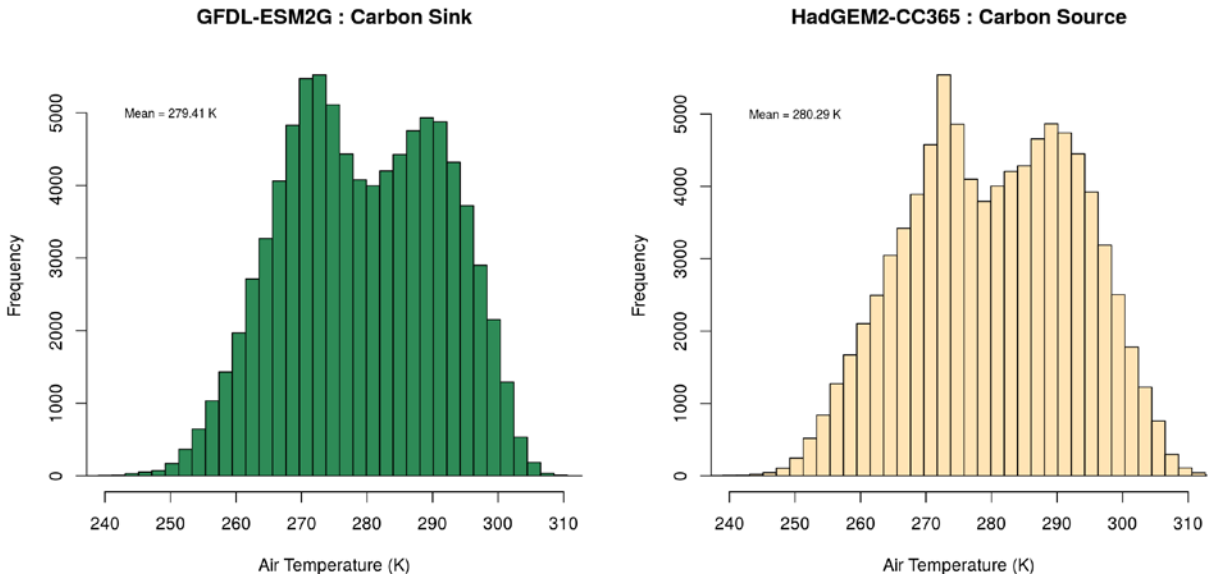

Figure 18. *Histogram of ensemble averaged temperature for GFDL-ESM2G (largest carbon sink) and HadGEM2-CC365 (largest carbon source) from 2020-2030.*

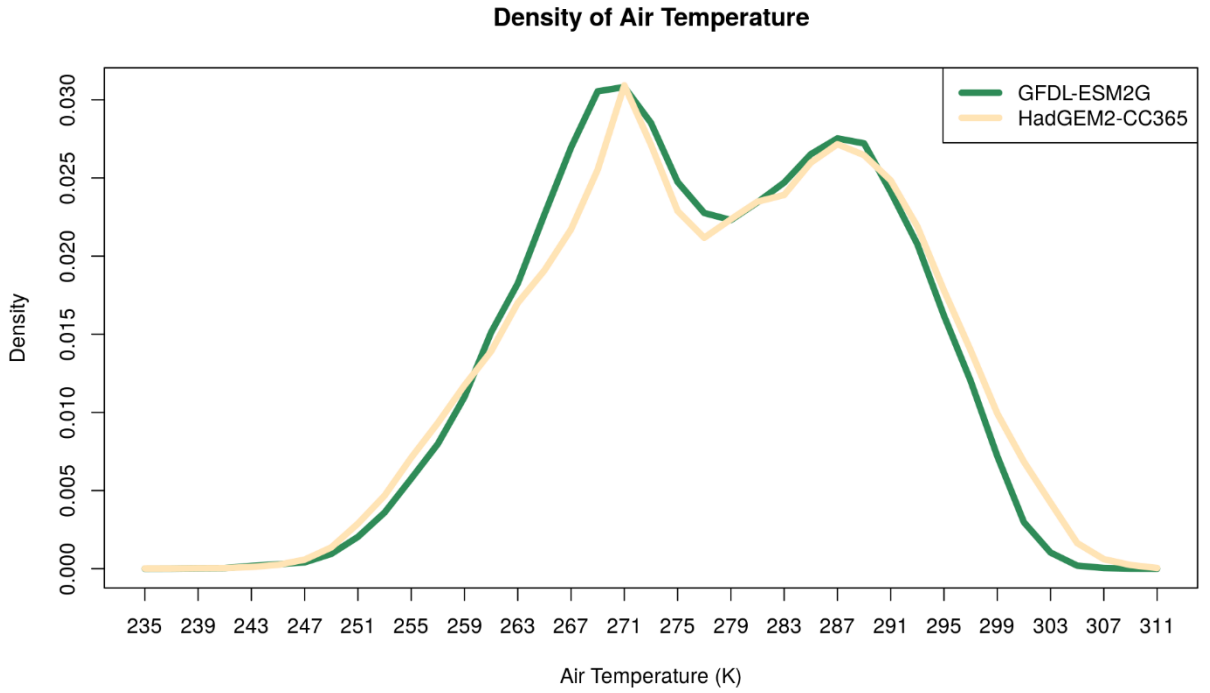

Figure 19. *Density of ensemble averaged, yearly averaged temperature for GFDL-ESM2G and HadGEM2-CC365 from 2020-2030.*

Temperature is an important driver for ecosystem function and the distribution of hourly temperatures could provide an explanation for the differences. Figures 18 & 19 show the ensemble averaged temperature distributions for the entire decade for GFDL-ESM2G (largest carbon sink) and HadGEM2-CC365 (largest carbon source). The models share a similar distribution shape but HadGEM2-CC365 mean decadal temperature 0.88 Kelvin higher. HadGEM2-CC365 appears to have a higher concentration of temperatures at the tails of the distribution, indicating that it has a more hourly temperature values that can induce ecological stress.

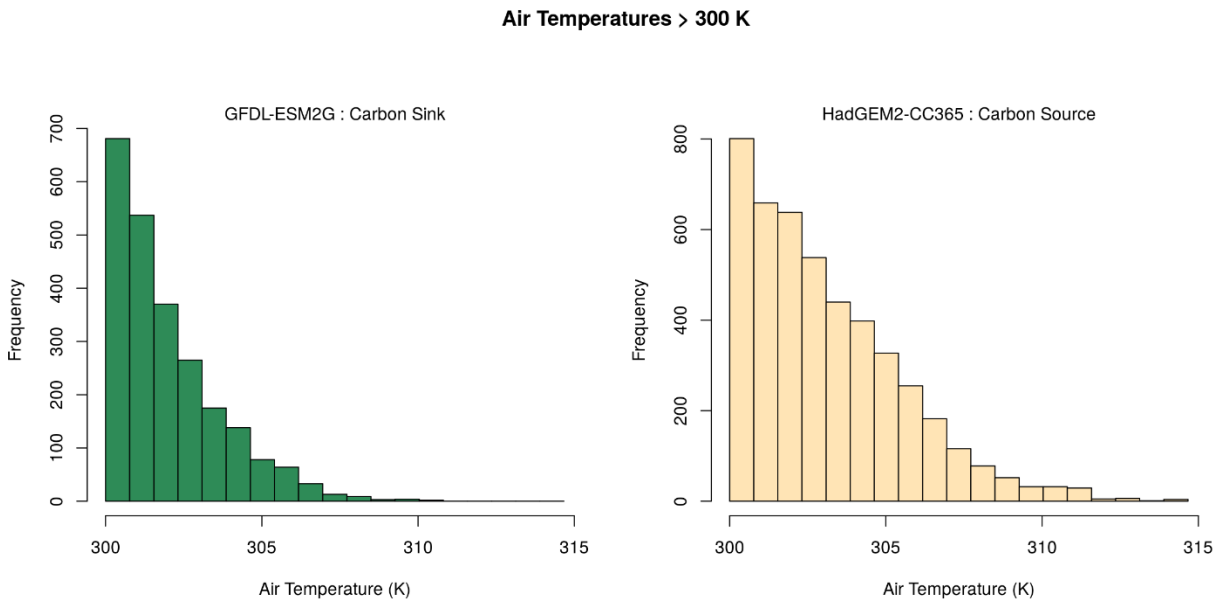

Figure 20. *Histogram of ensemble averaged, yearly averaged temperatures greater than 300 Kelvin for GFDL-ESM2G and HadGEM2-CC365 from 2020-2030.*

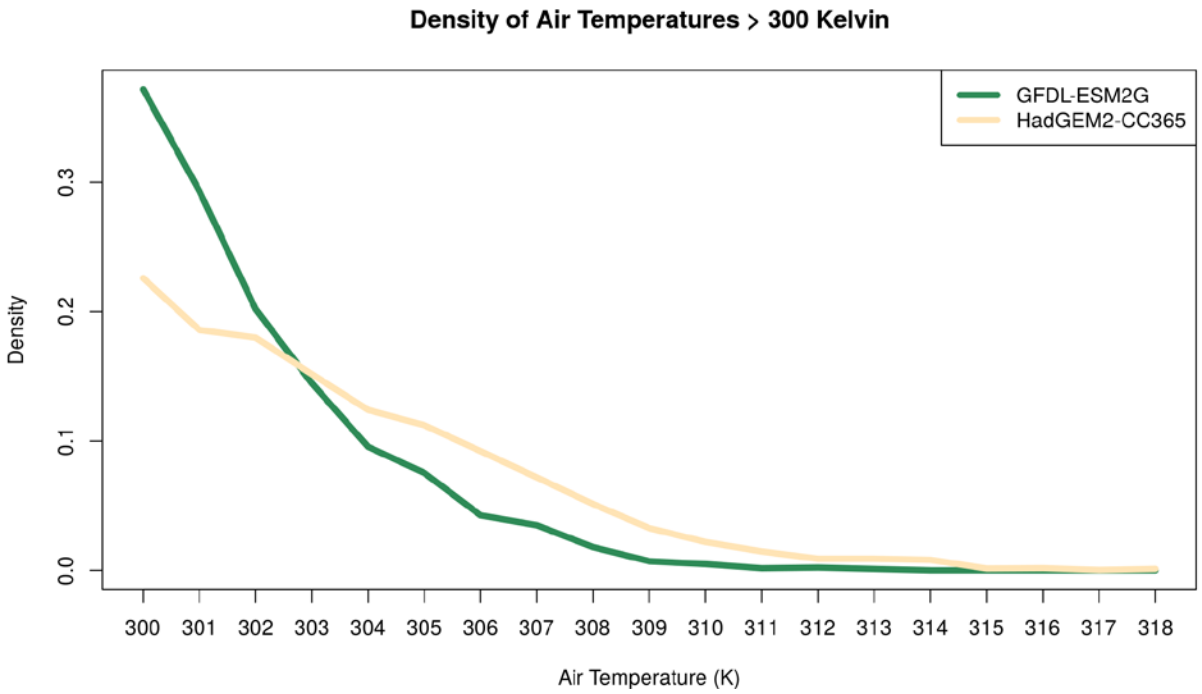

Figure 21. *Density of ensemble averaged, yearly averaged temperature greater than 300 Kelvin for GFDL-ESM2G and HadGEM2-CC365 from 2020-2030.*

One possible explanation is a disparity of hot temperatures in the models. Figures 20 and 21 show the ensemble averaged temperature distributions when temperature is greater than 300 Kelvin (80.33 Fahrenheit) for the entire decade for each model. This cutoff allows us to examine the concentration of temperatures that can threaten photosynthetic efficiency leading to decreased GPP. HadGEM2-CC365 has a higher concentration of warmer temperatures and has a higher density of temperatures beyond 303 Kelvin, suggesting higher plant stress and diminished productivity relative to GFDL-ESM2G.

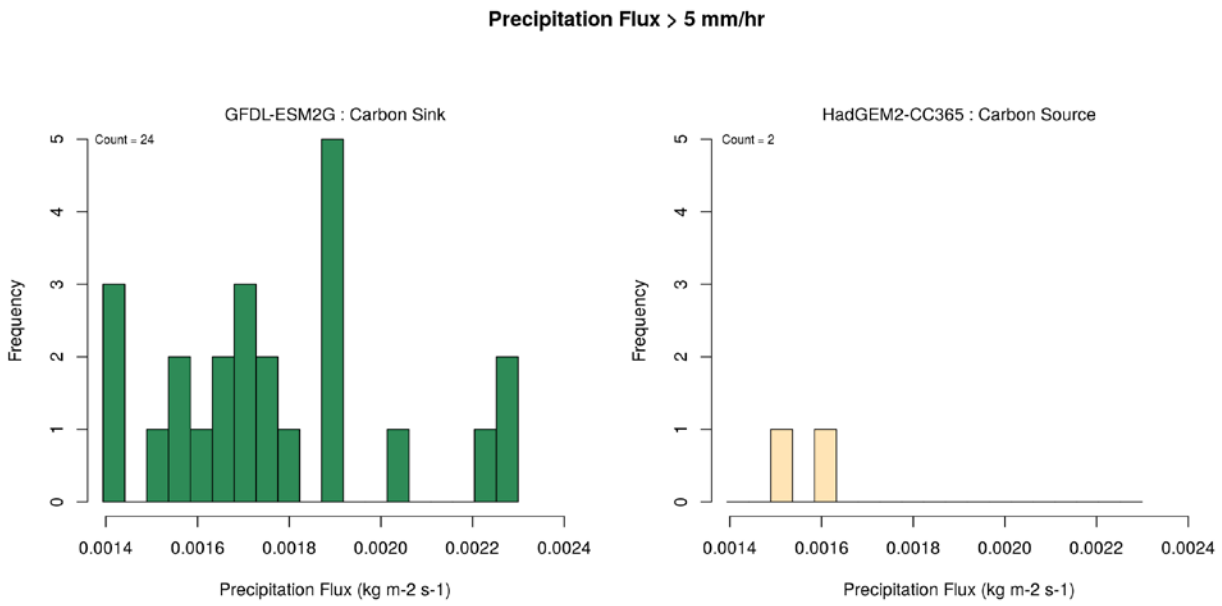

Figure 22. *Histogram of precipitation fluxes greater than 5 mm/hr (0.00139 kg m-2 s-1) for ensemble averaged, yearly averaged GFDL-ESM2G and HadGEM2-CC365 from 2020-2030.*

Another possible reason for that nearly 4000 gC m-2 decade1 cumulative NEE difference between GFDL-ESM2G and HadGEM2-CC365 is precipitation distribution; another key driver for ecosystem functions. Figure 22 shows the distribution of the concentrations of precipitation events greater than 5 mm/hr. GFDL-ESM2G has considerably more (22) precipitation events greater than 5 mm/hr than HadGEM2-CC365. The decadal sum of precipitation for GFDL-ESM2G is 852 mm higher than the decadal sum for HadGEM2-CC365. This suggests that the lack of higher precipitation events for HadGEM2-CC365 is not balanced by an increase in the number of smaller precipitation events.

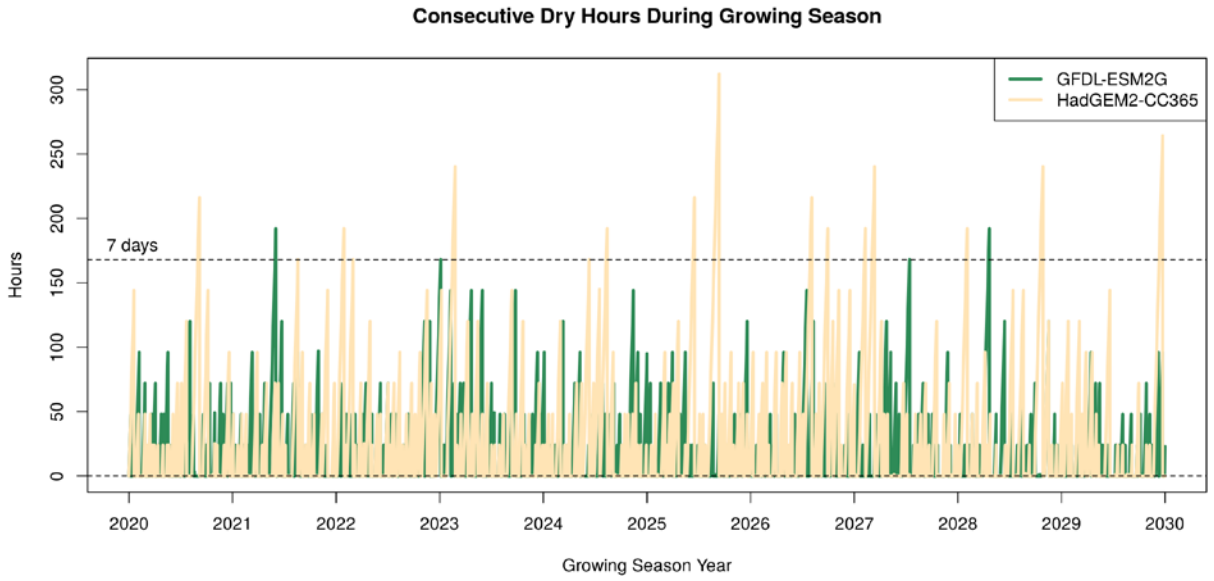

Figure 23. *Consecutive dry hours (no precipitation) during growing season for ensemble averaged, yearly averaged GFDL-ESM2G and HadGEM2-CC365 from 2020-2030. Over the decade, GFDL-ESM2G had 852.292 mm more precipitation than HadGEM2-CC365.*

A smaller decadal sum of precipitation and smaller number of heavy precipitation events offers us one piece of evidence for the high levels of decadal NEE for HadGEM2-CC365. A lower quantity of precipitation can be accounted for by smaller, more consistent precipitation events. Figure 23 investigates the frequency of precipitation events by showing the consecutive dry hours during the growing season. We chose to analyze the length of dry periods during the growing season because that is the period of time where GPP magnitudes are most susceptible. When we consider the number of dry periods lasting longer than 7 days, HadGEM2-CC365 has 11 more such periods than GFDL-ESM2G. Therefore, HadGEM2-CC365 has a lower quantity of precipitation and has a higher quantity of weeks without precipitation throughout the decade.

The greater density of temperatures greater than 300 K and drier conditions of the HadGEM2-CC365 meteorology likely increased the ecological stress and lead to decreased levels of carbon uptake. Alternatively, the cooler, wetter conditions of the GFDL-ESM2G

meteorology creates a more productive environment for photsynthetic efficiency and leads to greater carbon sequestration.

4. Discussion

4.1 Model performance

Generally, TDM routine had high accuracy for most meteorological variables. Temperature, a key driver for ecosystem response, performed well during our test case. The temporal downscaling routine also returned results with low error for shortwave radiation. The temporal downscaling for longwave radiation and air pressure captured the minimums, means, and maximums well but performed poorly during the accuracy analysis. This suggests that we aren't capturing the timing of these events correctly. TDM downscaled wind speed with high accuracy and recaptured the means. However, the maximum wind speed value is overestimated using our approach. Likewise, specific humidity downscaling had high accuracy regarding the timing of fluctuations but the magnitudes were suppressed (both maximums and minimums). Further tuning of the model is required to correct for these variables.

The lack of accuracy in the precipitation timing and intensity is a caveat in our methodology. Debiasing for precipitation fluxes using a scaling factor improved our accuracy by matching up precipitation means. This will decrease our mean percent error for the maximum in precipitation, though it won't correct the problem completely.

Precipitation has long been a difficult variable to temporally downscale due to its inconsistent nature (Gutmann et al., 2014; Clark et al., 2016). Other studies have used artificial networks for temporally downscaling precipitation (Coulibaly et al., 2005; Dibike et al., 2005; Anandhi et al., 2008; Mendes & Morengo, 2010). While these techniques are useful for

probabilistic prediction tasks, the computer code used to execute these the downscaling are openly available. Further, Kumar et al. (2010) shows that artificial neural networks not be useful for future time periods with increased variability. This is also true for our methodology as well which relies on an observational dataset. Further implementation of projected variability will need to be explored.

Considering that we derived hourly values from a daily measurement using the TDM workflow, the diurnal cycles were well captured for each variable. The inter-hourly variability is slightly constrained due to random variations due to temporary factors and this is a criticism of linear regression approaches (Wilby and Wigley, 1997).

4.2 Limitations of TDM approach

Although the linear regression based approach is a conservative method for reproducing extreme events (Wilby et al., 2004; Nguyen et al., 2007; Maraun et al., 2010), it's a straightforward method that employs a full range of available predictor variables. We were also driven to employ probabilistic temporal downscaling ideologies while also propagating future uncertainty that comes with any downscaling methodology. Linear regression also falls under scrutiny due to its poor representation of observed variance. However, we correct for this concern in our study by using high temporal resolution Fluxnet 2015 eddy covariance data to train our GCM output. By using high frequency flux tower data to train our GCM output, we will have the ability to better capture extreme subdaily events compared to previous linear regression efforts. The techniques developed for this study will be used to downscale CMIP5 output and used to estimate future carbon cycle responses.

The TDM functions have a univariate, independent approach, aside from specific humidity. A multivariate approach was tested, but took much longer to run and returned values well outside the realm of possibility. We hope to incorporate a multivariate regression technique in the future to more accurately account for covariances and interdependent variable relationships. Furthermore, while accounting for the residual errors allows for more uncertainty propagation, it is also more computationally intensive and for this reason we created that optional logical argument. Accounting for residual errors also poses problems for preserving covariances among drivers. We've built the function so that it could potentially be parallelizable, though that hasn't been fully implemented yet. This would greatly enhance the efficiency of the process but also requires a great deal of computational power.

TDM requires an hourly training dataset with a minimum of 6 years. Currently, generating a training dataset takes roughly 1 hour per year within the dataset. The more years present, the larger our pool of statistics to pull from. It takes roughly 1.5 hours to downscale a daily resolution dataset to an hourly resolution with 2 minutes per ensemble requested. The majority of the time is spent loading the stored linear regression models. Advancing the speed of the model is one of our future tasks.

Our TDM routine offers a unique way to temporally downscale any location on the globe given a high resolution training dataset. Our routine allows a user to also downscale for any time period, though we recommend staying close to the training dataset due to changes in variability that are currently unaccounted for. The functions above have been generalized to work with any meteorological dataset in CF convention and at any temporal resolution. These open-sourced functions are designed to work within the PEcAn framework and are available at (https://github.com/PecanProject/pecan/modules/data.atmosphere/R).

4.3 SIPNET Model Caveats

The SIPNET model was used for its computational efficiency and flexible meteorology variable requirements. It is not the best model available to quantify exact NEE and lacks the robust processes of other models such as the Ecosystem Demography model version 2 (Medvigy et al., 2009). SIPNET offers us estimates of the carbon cycle but a more robust model should be employed for finer estimates. Furthermore, a single set of parameters was used to initialize each SIPNET run. Willow Creek has semi-steady parameters (Cook et al., 2004), but these are likely to change slightly by 2020. SIPNET will also induce uncertainty via the propagation of those parameters over the time of the particular ecosystem model run. For example, SIPNET will build up a carbon pool over time and the parameters are changing as well. Typically, a parameter data assimilation should be employed for a range of parameter sets but that is beyond the scope of this study. As stated before, another caveat of SIPNET is that the model does not respond to changes in atmospheric CO2.

This study looks directly at the impact that climate change will have on ecosystems by temporally downscaling spatially downscaled CMIP5 model data. It's important to acknowledge that ecosystems can also impact physical properties within the boundary layer and alter meteorological variables. Ecosystems and meteorology represent a coupled system where each member influences the other. Estimating ecosystem influences on meteorology is beyond the scope of this study but should be considered in any carbon cycle quantification analysis.

4.3 Implications for carbon cycle

Our results are consistent with Medvigy et al. (2010) and demonstrate the importance of temporal resolution for ecosystem modeling. By feeding ecosystem models high frequency meteorological data at an hourly timestep, we can improve yearly NEE estimates and decrease

the interannual variability across the 16 year dataset at Willow Creek compared to coarser temporal resolution driver data. Further, we are better able to preserve the covariances and regression slope relationships between NEE and temperature. High temporal frequency data is imperative to accurate estimates of the carbon cycle.

4.4 Future work

Future work will primarily focus on improving the accuracy and speed of the TDM routine. Additionally, the precipitation bias and current TDM struggles associated with capturing precipitation maximum, specific humidity maximum and minimum, and wind speed maximum will be addressed. Future analysis will also focus on increasing the number of betas used for the linear regression models to quantify our meteorological uncertainty.

Conclusions

In this study, evidence is presented highlighting the importance that high temporal resolution meteorological plays in ecosystem modeling. We discuss a unique temporal downscaling algorithm and validate its accuracy. Future carbon cycle response to climate change at Willow Creek is tested using spatially and temporally downscaled CMIP5 data. Our results include:

1. Our temporal downscaling routine decreased our uncertainty by 94.61% when we downscaled from a daily resolution to an hourly resolution.
2. Covariances between NEE and temperature become decoupled at low frequency resolutions.

3. Our ensemble based temporal downscaling routine that returned multiple meteorological variables based on a high frequency training dataset quantifies our uncertainty and advances the science behind methodologies chosen for temporal downscaling routines.

4. Using our TDM routine, we are able to capture the diurnal cycles of a variety of meteorological variables, improve our NEE estimates, and decrease carbon cycle uncertainty.

5. The spatially and temporally downscaled CMIP5 models show varying regression slopes throughout the decade but more models point to Willow Creek trending towards a carbon source through the 2020-2030 period. The average model slope is 5 gC m-2 yr-1.

6. CMIP5 models show large divergence (~4000 gC m-2) in decadal cumulative NEE from 2020-2030. The model average has Willow Creek sequestering around 1200 gC m-2 by the end of the decade.

7. Differences in decadal carbon cycle between models due to strong differences in the amount of hourly temperatures above 300 K, amount of precipitation events greater than 5 mm/hr, and number of consecutive hours without precipitation. Ecological systems are highly sensitive to high-frequency variance in climate.

Citations

Aber, J. D., and C. A. Federer. 1992. A generalized, lumped-parameter model of photosynthesis, evapotranspiration and net primary production intemperate and boreal forest ecosystems. Oecologia92:463–474.

Abatzoglou, J. T. & Brown, T. J. A comparison of statistical downscaling methods suited for wildfire applications. Int. J. Climatol. 32, 772–780 (2012).

Adams, H. D. et al., Climate-induced tree mortality: Earth system consequences. Eos 91, 153–154 (2010)

Anandhi, Aavudai, et al. "Downscaling precipitation to river basin in India for IPCC SRES scenarios using support vector machine." International Journal of Climatology 28.3 (2008): 401-420.

Baldocchi DD, Falge E, Gu LH, Olson R, Hollinger D, Running S, Anthoni P, Bernhofer C, Davis K, Evans R, Fuentes J, Goldstein A, Katul G, Law B, Lee XH, Malhi Y, Meyers T, Munger W, Oechel W, Paw U KT, Pilegaard K, Schmid HP, Valentini R, Verma S, Vesala T, Wilson K, Wofsy S (2001a) FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. Bulletin of the American Meteorological Society 82, 2415-2434.

Bierkens, Marc F., Peter A. Finke, and Willigen. Upscaling and downscaling methods for environmental research. Dordrecht Boston: Kluwer Academic Publishers, 2000. Print.

Bigler, C., Gavin, D. G., Gunning, C. & Veblen, T. T. Drought induces lagged tree mortality in a subalpine forest in the Rocky Mountains. Oikos 116, 1983–1994 (2007).

Braswell, B. H., Sacks, W. J., Linder, E. & Schimel, D. S. Estimating diurnal to annual ecosystem parameters by synthesis of a carbon flux model with eddy covariance net ecosystem exchange observations. Glob. Chang. Biol. 11, 335–355 (2005).

Breda, N.,Huc, R., Granier, A. & Dreyer, E. Temperate forest trees and stands under severe drought: a review of ecophysiological responses, adaptation processes and long-term consequences. Ann. For. Sci. 63, 625–644 (2006).

Carbone, R. E., Tuttle, J. D., Ahijevych, D. A. & Trier, S. B. Inferences of Predictability Associated with Warm Season Precipitation Episodes. J. Atmos. Sci. 59, 2033–2056 (2002).

Clark, M., Gangopadhyay, S., Hay, L., Rajagopalan, B. & Wilby, R. The Schaake Shuffle: A Method for Reconstructing Space–Time Variability in Forecasted Precipitation and Temperature Fields. J. Hydrometeorol. 5, 243–262 (2004).

Cook, B.D., Davis, K.J., Wang, W., Desai, A.R., Berger, B.W., Teclaw, R.M., Martin, J.M., Bolstad, P., Bakwin, P., Yi, C., Heilman,W., 2004. Carbon exchange and venting anomalies in an upland deciduous forest in northern Wisconsin, USA. Agric. Forest

Coulibaly, Paulin, Yonas B. Dibike, and François Anctil. "Downscaling precipitation and temperature with temporal neural networks." Journal of Hydrometeorology 6.4 (2005): 483-496.

ENES, CMIP5 Models and Grid Resolution. European Network for Earth System Modeling, (2011).

De Boeck, H. & Verbeeck, H. Drought-associated changes in climate and their relevance for ecosystem experiments and models. Biogeosciences 8, 1121–1130 (2011).

Desai, A. R., Bolstad, P. V., Cook, B. D., Davis, K. J. & Carey, E. V. Comparing net ecosystem exchange of carbon dioxide between an old-growth and mature forest in the upper Midwest, USA. *Agric. For. Meteorol.* **128,** 33–55 (2005).

Dietze, Michael C., David S. Lebauer, and R. O. B. Kooper. "On improving the communication between models and data." Plant, Cell & Environment 36.9 (2013): 1575-1585.

Dietze, Michael C. "Prediction in ecology: a first-principles framework." Ecological Applications (2017).

Frank, D. et al., Effects of climate extremes on the terrestrial carbon cycle: Concepts, processes and potential future impacts. Glob. Chang. Biol. 21, 2861–2880 (2015).

Friedlingstein, Pierre, et al. "Climate–carbon cycle feedback analysis: results from the C4MIP model intercomparison." Journal of Climate 19.14 (2006): 3337-3353.

Fisher, J. B., et al., The future of evapotranspiration: Global requirements for ecosystem functioning, carbon and climate feedbacks, agricultural management, and water resources, Water Resour. Res., 53, (2017) doi:10.1002/2016WR020175.

Gutmann, D. A. et al., An Intercomparision of Statistical Downscaling Methods used for water resource assessments in the United States. 7167–7186 (2014). doi:10.1002/2014WR015559.

Hatfield, Jerry L., et al. "Climate impacts on agriculture: implications for crop production." Agronomy journal 103.2 (2011): 351-370.

Hyndman, R. J. & Koehler, A. B. Another look at measures of forecast accuracy. *Int. J. Forecast.* **22,** 679–688 (2006).

IPCC. Climate Change 2014 Synthesis Report Summary Chapter for Policymakers. *IPCC* 31 (2014). doi:10.1017/CBO9781107415324

Jianguo. Scaling and uncertainty analysis in ecology : methods and applications. Dordrecht, Netherlands: Springer, 2006. Print.

Kirchmeier-Young, M. C., Lorenz, D. J. & Vimont, D. J. Extreme event verification for probabilistic downscaling. J. Appl. Meteorol. Climatol. 55, 2411–2430 (2016).

Kumar, J., Brooks, B. J., Thornton, P. E. & Dietze, M. C. Sub-daily Statistical Downscaling of Meteorological Variables Using Neural Networks. Procedia Comput. Sci. 9, 887–896 (2012).

LeBauer, D. S., Wang, D., Richter, K. T., Davidson, C. C. and Dietze, M. C. (2013), Facilitating feedbacks between field measurements and ecosystem models. Ecological Monographs, 83: 133–154. doi:10.1890/12-0137.1

Lobell, David B., and Sharon M. Gourdji. "The influence of climate change on global crop productivity." Plant Physiology 160.4 (2012): 1686-1697.

Long, S. P. Modification of the response of photosynthetic productivity to rising temperature by atmospheric CO2 concentrations: Has its importance been underestimated? Plant. Cell Environ. 14, 729–739 (1991).

Maraun, D. et al., Precipitation downscaling under climate change: Recent developements to bridge the gap between dynamical models and the end user. Rev. Geophys. 48, 1–38 (2010).

Maurer, E. P., Hidalgo, H. G., Das, T., Dettinger, M. D. & Cayan, D. R. The utility of daily large-scale climate data in the assessment of climate change impacts on daily streamflow in California. Hydrol. Earth Syst. Sci. 14, 1125–1138 (2010).

Mascaro, G., E.R. Vivoni, D.J. Gochis, C.J. Watts, and J.C. Rodriguez, 2014: Temporal Downscaling and Statistical Analysis of Rainfall across a Topographic Transect in Northwest Mexico. J. Appl. Meteor. Climatol., 53, 910–927, https://doi.org/10.1175/JAMC-D-13-0330.1

Medvigy, D., Wofsy, S. C., Munger, J. W., Hollinger, D. Y. & Moorcroft, P. R. Mechanistic scaling of ecosystem function and dynamics in space and time: Ecosystem Demography model version 2. J. Geophys. Res. Biogeosciences 114, 1–21 (2009).

Medvigy, D., Wofsy, S. C., Munger, J. W. & Moorcroft, P. R. Responses of terrestrial ecosystems and carbon budgets to current and future environmental variability. Proc. Natl. Acad. Sci. U. S. A. 107, 8275–8280 (2010).

Mendes, D. & Marengo, J. A. Temporal downscaling: A comparison between artificial neural network and autocorrelation techniques over the Amazon Basin in present and future climate change scenarios. Theor. Appl. Climatol. 100, 413–421 (2010).

Montgomery, Douglas C., Elizabeth A. Peck, and G. Geoffrey Vining. Introduction to linear regression analysis. John Wiley & Sons, 2015.

Moorcroft, Paul R. "How close are we to a predictive science of the biosphere?." Trends in Ecology & Evolution 21.7 (2006): 400-407.

Neter, John, et al., Applied linear statistical models. Vol. 4. Chicago: Irwin, 1996.

Nguyen, V., Nguyen, T. & Cung, A. A statistical approach to downscaling of sub-daily extreme rainfall processes for climate-related impact studies in urban areas. Water Sci. Technol. Water Supply 7, 183–192 (2007).

Pachauri, Rajendra K., et al. Climate change 2014: synthesis report. Contribution of Working Groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change. IPCC, 2014.

Parmesan, C., Root, T. L. & Willig, M. R. Impacts of extreme weather and climate on terrestiral biota. Bull. Am. Meteorol. Soc. 81, 443–450 (2000).

Peng, S. et al., Rice yields decline with higher night temperature from global warming. Proc. Natl. Acad. Sci. 101, 9971–9975 (2004).

R Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing,Vienna, Austria. ISBN 3-900051-07-0, URL-http://www.R-project.org.

Raupach, M. R., Canadell, J. G. & Le Quéré, C. Anthropogenic and biophysical contributions to increasing atmospheric CO 2 growth rate and airborne fraction. *Biogeosciences* **5,** 1601–1613 (2008).

Reichstein, M. et al., On the separation of net ecosystem exchange into assimilation and ecosystem respiration: Review and improved algorithm. Glob. Chang. Biol. 11, 1424–1439 (2005).

Reichstein, M. et al., Climate extremes and the carbon cycle. Nature 500, 287–295 (2013).

Rosenblatt, Murray. "A central limit theorem and a strong mixing condition." Proceedings of the National Academy of Sciences 42.1 (1956): 43-47.

Schmidli, J., C. M. Goodess, C. Frei, M. R. Haylock, Y. Hundecha, J. Ribalaygua, and T. Schmith, 2007: Statistical and dynamical downscaling of precipitation: An evaluation and comparison of scenarios for the European Alps. J. Geophys. Res., 112, D04105, doi:10.1029/2005JD007026.

Schmith, Torben. "Stationarity of regression relationships: Application to empirical downscaling." Journal of Climate 21.17 (2008): 4529-4537.

Solomon S, Qin D, Manning M et al., (2007) Technical summary.In: Climate Change 2007: The Physical Science Basis. Contributionof Working Group I to the Fourth Assessment Report of theIntergovernmental Panel on Climate Change (eds Solomon S,Qin D, Manning M, Chen Z,

Marquis M, Averyt KB, TignorM, Miller HL). Cambridge University Press, Cambridge, UK/New York, NY, USA.

Taylor, K., R. Stouffer, and G. Meehl, 2012: An Overview of CMIP5 and the Experiment Design. Bull. Amer. Meteor. Soc., 93, 485–498, doi: 10.1175/BAMS-D-11-00094.1.

Trenberth, K. E. Changes in precipitation with climate change. Clim. Res. 47, 123–138 (2011).

Vrac, M. & Vaittinada Ayar, P. Influence of bias correcting predictors on statistical downscaling models. J. Appl. Meteorol. Climatol. JAMC-D-16-0079.1 (2016). doi:10.1175/JAMC-D-16-0079.1

Wheater, H. (2006), Flood hazard and management: A UK perspective, Philos. Trans. R. Soc. A, 364(1845), 2135–2145.

Wilby, R. L., and T. M. L. Wigley (1997), Downscaling general circulation model output: A review of methods and limitations, Prog. Phys. Geogr., 21, 530–548

Wilby RL, Charles SP, Zorita E, Timbal B, Whetton P, Mearns LO. *Guidelines for Use of Climate Scenarios Developed from Statistical Downscaling Methods*. IPCC Task Group on Data and Scenario Support for Impact and Climate analysis (TGICA). (2004). http://ipcc–ddc.cru.uea.ac.uk/guidelines/StatDown_Guide.pdf.

Wolf, S. *et al.,* Warm spring reduced carbon cycle impact of the 2012 US summer drought. *Proc. Natl. Acad. Sci. U. S. A.* 1519620113- (2016). doi:10.1073/pnas.1519620113

Zhang, X-C. "Spatial downscaling of global climate model output for site-specific assessment of crop production and soil erosion." Agricultural and Forest Meteorology 135.1 (2005): 215-229.

Zhou, S., Yu, B., Huang, Y. & Wang, G. The effect of vapor pressure deficit on water use efficiency at the subdaily time scale. Geophys. Res. … 5005–5013 (2014). doi:10.1002/2014GL060741.Received