



Universität Hamburg

DER FORSCHUNG | DER LEHRE | DER BILDUNG

Bachelor Thesis

Interpretable Machine Learning to understand Meteorological Impacts on Ecosystem Carbon Uptake

David Hafezi Racht

Meteorology
Matr.-Nr.: 7406480

First Supervisor: Dr. Alexander J. Winkler
Second Supervisor: Professor Dr. Felix Ament

Interpretable Machine Learning to understand Meteorological Impacts on Ecosystem Carbon Uptake

Bachelor Thesis

by

David Hafezi Rachtì

Student Name	Student Number
David Hafezi Rachtì	7406480

Supervisors: Dr. Alexander J. Winkler
Professor Dr. Felix Ament



Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Abstract

The land carbon uptake is the main driver of interannual variations in atmospheric CO₂ and one of the least understood parts of the global carbon cycle. Meteorological factors, such as variations in weather patterns and extreme climate events, are one of the main drivers of interannual variations in carbon uptake by terrestrial ecosystems. However, the impact of multi-scale meteorological events, their timing and duration on the carbon balance remains uncertain.

In this thesis, I adapt interpretable machine learning (ML) techniques from computer vision to investigate the role of multi-scale meteorological variability on the carbon balance of forest ecosystems. Specifically, I use a modelling framework based on a convolutional neural network trained on wavelet-transformed key meteorological variables to predict ecosystem carbon and water fluxes. The eddy covariance data comes from 15 deciduous broadleaf forest sites with a total of 112 site years. Input data for the model are meteorological measurements filled with reanalysis data, a remote sensing variable for the state of the ecosystem and a random walk variable for validation. The Integrated Gradients explanatory technique provides insights into the importance of different meteorological factors, as well as the length and timing of atmospheric events, for the anomalies in the annual carbon balance.

I can confirm that water availability is the dominant factor for local variations in carbon balance. In addition, I show that 20-40 day long vapour pressure deficit events during the summer are one of the most important drivers for the model to predict less annual carbon storage. In a case study of the 2003 heatwave, the model is able to capture the legacy effects of the heatwave on the 2004 carbon balance. Such experiments are important for understanding the role of meteorology on ecosystem carbon uptake, and for demonstrating the potential of interpretable machine learning methods to improve understanding of land-atmosphere interactions.

Contents

Abstract	i
Nomenclature	iii
1 Introduction	1
2 Fluxnet: a global network of flux measurements	4
2.1 Sites	4
2.2 Enhanced Vegetation Index	6
2.3 Gap Filling	8
3 Main Methods	9
3.1 Wavelet Transformation	9
3.2 Convolutional Neural Networks	10
3.3 Model architecture	10
3.4 Integrated Gradients	12
3.5 Model training and validation	12
4 Model Performance	15
5 Main Drivers of Positive NEE Anomalies	19
5.1 Methods	19
5.2 Overall importance of input variables	20
5.3 Temporal patterns of VPD importance	23
6 Case Study: Heatwave 2003 in Hainich Forest Germany	26
6.1 Methods	27
6.2 Atmospheric water demand caused by the 2003 heatwave has a legacy effect on 2004 NEE_{yr} anomaly	27
7 Conclusion	30
8 Outlook	31
References	39

Nomenclature

Abbreviations

Abbreviation	Definition
ML	Machine Learning
IAV	Interannual Variability
PFT	Plant Functional Type
DBF	Deciduous Broadleaf Forest
NA	North America
EU	Europe
QC	Quality Flags
WT	Wavelet Transform
CNN	Convolutional Neural Network
ResNet-18	Residual Neural Network of 18 layers
IG	Integrated Gradients
WCon	Wavelet Transform + ResNet-18 Model
WLR	Wavelet Transform + Linear Regression Model
SC	Mean Seasonal Cycle
CC	Correlation Coefficient
C	Carbon
Gt	Gigatons
yr	year

Variable Names

Symbol	Definition	Unit
NEE	Net ecosystem exchange	$\text{gC m}^{-2} \text{d}^{-1}$
GPP	Gross primary production	$\text{gC m}^{-2} \text{d}^{-1}$
TER	Total ecosystem respiration	$\text{gC m}^{-2} \text{d}^{-1}$
NEE_{yr}	Annual cumulative net ecosystem exchange	$\text{gC m}^{-2} \text{y}^{-1}$
LE	Latent heat flux	W m^{-2}
PA	Atmospheric pressure	kPa
TA	Air temperature	$^{\circ}\text{C}$

Symbol	Definition	Unit
VPD	Vapor pressure deficit	hPa
P	Precipitation	mm d ⁻¹
RH	Relative humidity	%
SW_IN	Shortwave radiation, incoming	W m ⁻²
SW_IN_POT	Shortwave radiation, incoming, potential (top of atmosphere)	W m ⁻²
D_SW_IN_POT	Derivative of SW_IN_POT	W m ⁻²
NETRAD	Net radiation	W m ⁻²
WS	Wind speed	m s ⁻¹
EVI	Enhanced vegetation index	
MAP	Mean annual precipitation	mm
MAT	Mean annual temperature	°C

1

Introduction

Atmospheric CO₂ continues to rise and in 2021 total anthropogenic CO₂ emissions reached $10.9 \pm 0.8 \text{ GtC yr}^{-1}$ (Friedlingstein et al., 2022). More than 50% of these emissions are absorbed by land (32.11%) and ocean (26.60%). In particular, the land carbon sink has significantly slowed global warming (Shevliakova et al., 2013). As atmospheric CO₂ increased over the past few decades, terrestrial ecosystems began to store more carbon. This is due to an extension of the growing season and increased vegetation growth, including CO₂ and nitrogen fertilisation. It is crucial to know how long this negative feedback will continue, but future projections are highly uncertain (Friedlingstein et al., 2006). A more complete understanding of the global carbon cycle is needed to reduce these uncertainties.

Terrestrial carbon uptake is the sum of carbon uptake through photosynthesis (Gross Primary Production, GPP) and carbon release through various respiration processes (Total Ecosystem Respiration, TER). The net balance is defined as net ecosystem exchange (NEE):

$$NEE = TER - GPP \quad (1.1)$$

The interannual variability (IAV) of the carbon cycle, driven primarily by land sink dynamics (Piao et al., 2020), is one of the most uncertain parts of the global carbon cycle (Friedlingstein et al., 2019). During the 1990s, the NEE IAV ranged from -4.0 GtC yr^{-1} (net uptake) to $+0.3 \text{ GtC yr}^{-1}$ (net emission) (Piao et al., 2020). A complex combination of climatic, ecological and disturbance variables can explain the IAV in the NEE (Baldocchi et al., 2018). Shao et al. (2015) showed in a large meta-analysis that biotic factors (ecological and disturbance) contributed to 57% of the variability in net ecosystem exchange, while climatic factors were associated with the remaining 43%.

There is increasing agreement on the influence of individual climatic factors, such as temperature or precipitation. For example, Jung et al. (2017) reported that water availability is the dominant driver of the local IAV of GPP and TER. For the IAV of NEE, these effects are partially compensated. However, the interactive effects of multiple climate factors were often neglected, contributing to confusion about the dominant climate factor driving the IAV of NEE (Piao et al., 2020).

An important factor in climate variability is the occurrence of extreme climate events. Extreme events can have a major impact on the carbon cycle by reducing carbon sinks or even causing net losses in carbon stocks, thereby releasing carbon into the atmosphere (Ciais et al., 2005). Ecosystem responses can be concurrent or lagged and often involve thresholds, with non-linear effects on carbon fluxes and stocks. Therefore, small shifts in the frequency or severity of climate extremes could significantly reduce carbon sinks, which could lead to positive feedback on climate warming (Reichstein et al., 2013). Reichstein et al. (2013) defined climate extremes in relation to the biosphere as:

"Conditions where an ecosystem function (such as carbon uptake) is higher or lower than a defined extreme percentile during a defined time period and over a certain area, traceable to single or multivariate anomalous meteorological variables."

To identify an extreme event, the focus is first on the impact of the event on ecosystems. Then the immediate and delayed effects of the meteorological variables have to be assigned. Frank et al. (2015) classified the impacts of climate extremes into four categories:

- Direct, concurrent impact: windthrow caused by storm; reduced productivity or increased mortality during drought
- Indirect, concurrent impact: loss of biomass or soil organic matter due to fire caused by lightning or human ignition
- Direct, lagged impact: reduced productivity or growth in the year(s) following the year of an extreme drought
- Indirect, lagged impact: increased pest- or pathogen-caused mortality following a climate extreme

Although extreme climate events could have far-reaching consequences for terrestrial ecosystems, the complex interactions are poorly understood (Frank et al., 2015). In particular, relatively little is known about the effects of extreme events other than droughts and heat-waves, such as heavy rainfall events, 'false springs' or exposure to high levels of solar radiation (Mahecha et al., 2022).

Different ecosystems respond differently to climate extremes and climate change. Forests have large stocks of above-ground biomass carbon per square metre, making them particularly vulnerable to extreme events. In addition, forests have a long recovery time due to the slow regrowth of trees. Forests are potentially vulnerable to a wide range of extreme events, such as storms, drought, heat, fire, ice storms and frost. However, on a global scale, droughts have the greatest impact on the carbon balance of forests (Reichstein et al., 2013).

A more accurate understanding of the key meteorological drivers and the effects of the timing and duration of meteorological events, such as climate extremes, on the carbon balance of forest ecosystems can help to reduce uncertainties in the carbon cycle. It will also improve the understanding of ecosystem responses to climate change.

Machine learning (ML) and deep learning (DL) are becoming increasingly popular in

Earth system science, providing promising tools for building new data-driven models to advance our understanding of the Earth (Reichstein et al., 2019). However, one of the major challenges is their lack of interpretability, which is crucial for understanding the results, improving the model or assessing physical plausibility.

In this thesis, I investigate the impact of multi-scale meteorological events on the IAV of NEE of deciduous broadleaf forests in Europe and North America. To do this, I train an ML model framework without any assumptions on eddy covariance data, with the goal of learning the main features of the meteorological input data internally. I then use an explanatory method from the field of Explainable Artificial Intelligence to identify the relevant factors learned by the model.

Specifically, I apply a machine learning framework using deep convolutional neural networks trained on wavelet-transformed key meteorological, remote sensing and random walk variables. I use Integrated Gradients in combination with the wavelet images to gain insight into the importance of various meteorological factors, as well as the length and timing of meteorological events, on the IAV of NEE.

First, I describe the dataset and explain the main methods. I then examine the performance of the model. The results are divided into two chapters. In the first chapter, I quantitatively investigate the main drivers of positive NEE anomalies. The second chapter is a case study on the legacy effects of the 2003 heat wave in the Hainich forest in Germany. Finally, I summarise the results and give an outlook on further possibilities to improve the model framework and further experiments.

2

Fluxnet: a global network of flux measurements

FLUXNET is a global network of eddy covariance towers. These measurement towers monitor carbon, water and energy cycles between the biosphere and the atmosphere, as well as other meteorological variables at the ecosystem scale. The network consists of 212 sites around the globe (a total of 1500 site years of data up to 2014). Some eddy covariance sites have been collecting data for several decades, making it possible to study the response of ecosystems to climate change.

Eddy covariance is a direct (in situ) flux measurement technique. It is based on the physical understanding of eddies moving air parcels. At time t , an eddy moves an air parcel containing different gases downwards at a certain speed. Then, at time $t+1$, another eddy moves an air parcel upwards at a certain speed. If the speed and gas concentrations of the parcel are known, the vertical flux can be estimated. An eddy covariance system consists of an ultrasonic anemometer and an infrared gas analyser. It measures high frequency (10-20 Hz) wind and scalar data (e.g. CO_2 , H_2O , temperature). Eddy covariance provides estimates of the net exchange of a scalar source footprint area extending up to hundreds of metres around the point (Pastorello et al., 2020).

All data are quality controlled and processed through the same data pipeline, ensuring consistency and comparability between sites (Pastorello et al., 2020). Each site is assigned to a specific ecosystem type, also known as a plant functional type (PFT), based on the International Geosphere-Biosphere Programme (IGBP) definition. PFTs are a system for classifying plants with similar responses to the environment.

2.1. Sites

In addition to FLUXNET2015 (Pastorello et al., 2020), I use the Warm Winter 2020 dataset from the ICOS network (Warm Winter 2020 Team, & ICOS Ecosystem Thematic Centre) and the AmeriFlux FLUXNET dataset from the AmeriFlux Management Project. ICOS is the European and AmeriFlux is the American eddy covariance network. Both datasets provide

additional data for the period 2015 - 2020. I select 15 PFT deciduous broadleaf forest (DBF) sites (Tab. 2.1) with a total of 112 site years between 2000 and 2020. Eight sites are located in North America (NA) and seven are in Europe (EU). For each site I select net ecosystem exchange (NEE), latent heat (LE) and several meteorological variables (2.2) at daily resolution.

Table 2.1: List of selected deciduous broadleaf forest sites with their climatic variables, number of siteyears and citations. The number of siteyears describes the number of siteyears included in the dataset for this thesis.

Site_id	MAT (°C)	MAP (mm)	Siteyears	Citation
CA-Oas	0.34	428.53	10	https://doi.org/10.18140/FLX/1440043
DE-Hai	8.3	720.0	11	https://doi.org/10.18160/2G60-ZHAK
DE-Lnf	6.96	894.6	2	https://doi.org/10.18140/FLX/1440150
DK-Sor	8.2	660.0	4	https://doi.org/10.18160/2G60-ZHAK
FR-Fon	10.2	720.0	9	https://doi.org/10.18160/2G60-ZHAK
FR-Hes	9.2	820.0	5	https://doi.org/10.18160/2G60-ZHAK
IT-Ro1	15.15	876.2	2	https://doi.org/10.18140/FLX/1440174
IT-Ro2	15.15	876.2	2	https://doi.org/10.18140/FLX/1440175
US-Ha1	6.62	1071.0	5	https://doi.org/10.18140/FLX/1440071
US-MMS	10.85	1032.0	19	https://doi.org/10.17190/AMF/1854369
US-MOz	12.11	986.0	13	https://doi.org/10.17190/AMF/1854370
US-Oho	10.1	849.0	4	https://doi.org/10.18140/FLX/1440088
US-UMB	5.83	803.0	14	https://doi.org/10.18140/FLX/1440093
US-UMd	5.83	803.0	6	https://doi.org/10.18140/FLX/1440101
US-WCr	4.02	787.0	6	https://doi.org/10.18140/FLX/1440095

I choose the input meteorological variables to describe the main meteorological characteristics as water availability (P, VPD and RH), temperature (TA, VPD) and radiation (SW_IN, NETRAD, SW_IN_P). Soil moisture, as an additional variable to describe soil water availability, could not be included, since it is not available for most of the FLUXNET sites. To allow the model to distinguish between spring and autumn, I also calculate the derivative of SW_IN_POT. The target variable for the model is LE as well as NEE. NEE is the difference between the total ecosystem respiration and the gross primary production. It describes the carbon flux between the ecosystem and the atmosphere. LE is the heat absorbed or released by water during a phase transition and is therefore a proxy for the water flux between the ecosystem and the atmosphere. The water and carbon cycles are strongly coupled through the stomata of plants. Thus, the prediction of both variables should add a physical constraint to the model.

Despite the same PFT, there are significant differences in climatic or static variables between the sites (Figure 2.1). For example, CA-Oas in Canada is the driest and coldest site with a mean annual precipitation (MAP) of 428.53 mm and a mean annual temperature (MAT) of 0.34 °C. The Harvard Forest EMS Tower site in NA has the highest MAP (1071.00 mm) and the Italian sites Roccarespampani one and two are the warmest with MAT of 15.15 °C. The climatic variables are also fed into the model to inform it of the climatic conditions of the site.

Table 2.2: List of the selected FLUXNET variable basenames, descriptions and units. PA to WS are meteorological variables. NEE and LE are the carbon and water fluxes between the atmosphere and the ecosystem.

Basename	Description	Unit
PA	Atmospheric pressure	kPa
TA	Air temperature	°C
VPD	Vapor pressure deficit	hPa
P	Precipitation	mm d ⁻¹
RH	Relative humidity	%
SW_IN	Shortwave radiation, incoming	W m ⁻²
SW_IN_POT	Shortwave radiation, incoming, potential (top of atmosphere)	W m ⁻²
D_SW_IN_POT	Derivative of SW_IN_POT	W m ⁻²
NETRAD	Net radiation	W m ⁻²
WS	Wind speed	m s ⁻¹
NEE	Net Ecosystem Exchange	gC m ⁻² d ⁻¹
LE	Latent heat flux	W m ⁻²

Forests absorb carbon from the atmosphere through photosynthesis. Carbon is released through autotrophic plant respiration and respiration by heterotrophic microorganisms (e.g. decomposition). A typical annual cycle of NEE is shown in Figure 2.2. In winter, from December to February, photosynthesis is dormant and thus NEE is positive. The forest releases more carbon into the atmosphere than it absorbs. With the start of the growing season in spring and the development of leaves, photosynthesis begins. When carbon uptake exceeds respiration, NEE becomes negative and the ecosystem stores more carbon than it emits. NEE reaches its minimum in midsummer and then increases again.

Figure 2.3 shows the distribution of the annual NEE sum (NEE_{yr} , annual sum over daily NEE values). Almost all siteyears exhibit a negative sum of NEE fluxes indicating a carbon sink function of the ecosystem with a mean value of about $-375 \text{ gC m}^{-2} \text{ y}^{-1}$. Ecosystems are growing.

2.2. Enhanced Vegetation Index

I use the enhanced vegetation index (EVI, Huete et al. (2002)) remote sensing variable as an additional feature to meteorology. The EVI is an optimised vegetation index designed to have improved sensitivity in high biomass regions such as forests. Other vegetation indices such as the normalised difference vegetation index (NDVI) are close to saturation in these regions. EVI is a state variable of the ecosystem. It provides the model information about phenology and the sensitivity of ecosystems to meteorological events. I use a dataset that provides daily EVI values for 338 flux sites between 2000 and 2020 (Walther et al., 2022).

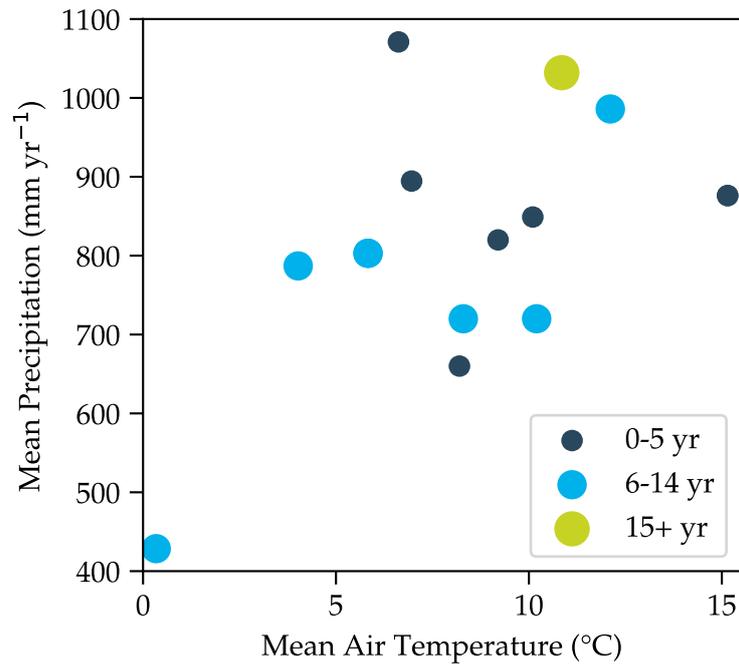


Figure 2.1: Distribution of the selected DBF sites across mean temperature and precipitation. The colours and size of the marker indicate the number of siteyears.

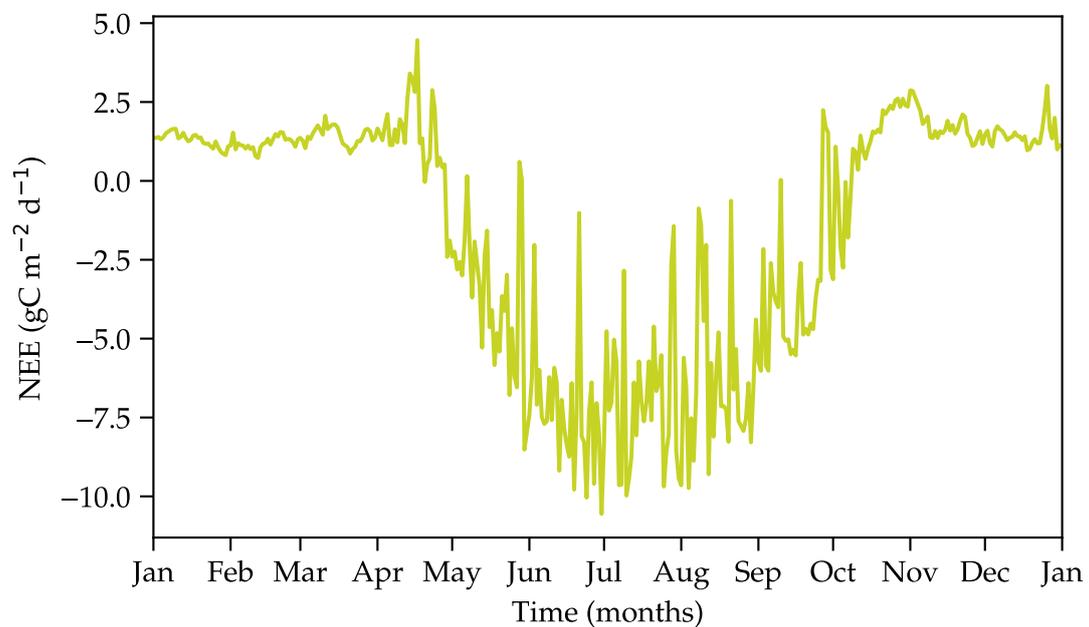


Figure 2.2: Annual cycle of NEE in 2007 in Hainich Forest, Germany.

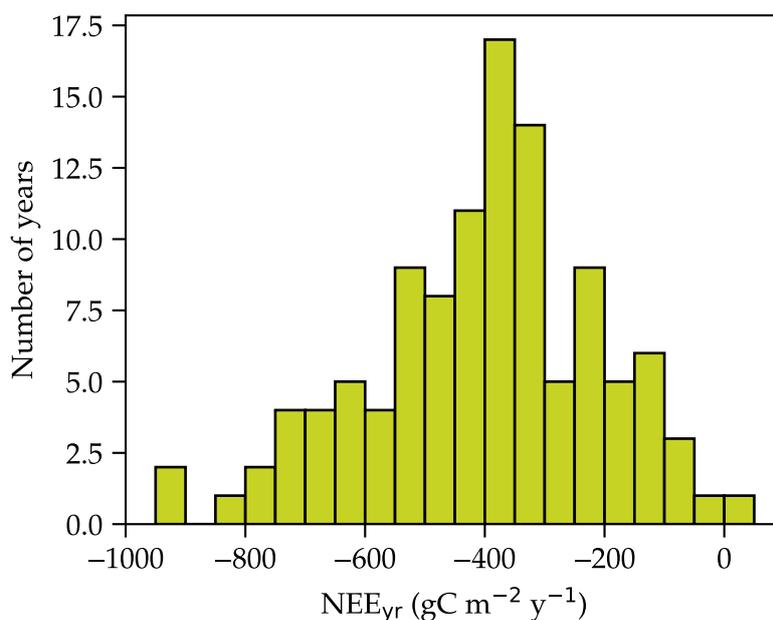


Figure 2.3: NEE_{yr} distribution of all 112 selected siteyears.

2.3. Gap Filling

The quality and completeness of the feature and target data is very important for data-driven learning. To ensure the quality for the target data (NEE and LE) I use the NEE Quality Flags (QC) from Jung et al. (2023). QC describes the quality of daily NEE values based on several factors. A QC of one corresponds to high quality data, while a QC of zero corresponds to low quality data. I have removed all years with more than ten consecutive days with a QC of zero from the data. The meteorological measurements from the FLUXNET sites contain gaps of various sizes. Precipitation in particular is a major problem. Precipitation time series are incomplete for many site years and difficult to interpolate due to the stochastic characteristics of precipitation. For the NA sites I use reanalysis data from the DayMet package (Thornton et al. (2022)) to fill the precipitation gaps. DayMet provides meteorological data with a spatial resolution of one kilometre. For all EU sites, I removed all siteyears with precipitation gaps longer than five days from the dataset. Smaller gaps were filled by piecewise linear interpolation. The remaining meteorological variables are filled with downscaled ERA5 data (Hersbach et al., 2020).

3

Main Methods

For this project, I am using a machine learning modelling framework based on a convolutional neural network trained on wavelet-transformed key variables to predict carbon and water fluxes. The modelling framework was first used in Hafezi Rachti et al. (2023) to predict phenological states. For a forward run, the ML model uses two years of daily values of ten different meteorological, one remote sensing and random walk variables to predict one year of daily values of net ecosystem exchange (NEE) and latent heat (LE). The explanation method Integrated Gradients (Sundararajan et al., 2017) combined with the wavelet images provides insight into the importance of different meteorological factors and the length and timing of meteorological events for NEE and LE.

3.1. Wavelet Transformation

The Wavelet Transform (WT, Daubechies (1990)) is the mathematical optimum between detecting frequencies and their position in a time series. It is similar to the Fourier transform. However, instead of just detecting the frequencies in the data series as the Fourier transform does, the WT also gives the position of the frequency. The continuous wavelet transform is given by the equation 3.1 . For a given scale (a) corresponding to a frequency, the convolution between the data series ($x(t)$) and a wavelet function $\psi(t)$ is computed. The resulting function shows how well $x(t)$ corresponds to a particular frequency at position t . To get a more general picture, this process is repeated for a range of scales. The final output is a two-dimensional wavelet spectrum. The WT transforms a one dimensional data series into two dimensional data that can also be interpreted as an image.

$$CWT_f(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \cdot \overline{\psi\left(\frac{t-b}{a}\right)} dt \quad (3.1)$$

- $CWT_f(a, b)$: Continuous wavelet transform
- $x(t)$: Function to be transformed, e.g. time series
- $\psi(t)$: Wavelet function
- a : Scaling parameter
- b : Translation parameter, for sampling the data

In this project I use the WT to transform a two year time series into a two-dimensional image showing the wavelet spectrum. The Nyquist frequency for a daily time series is 0.5/day. For better interpretability, I speak of periods and not frequencies ($T = 1/f$). The scales corresponding to the periods are logarithmically spaced from 2 days to 730 days, with a total of 64 scales. The wavelet function I use is the Ricker wavelet (Ricker, 1943). The Ricker wavelet is a second derivative of the Gaussian function and is one of the standard wavelet functions. It is useful for identifying both peaks and troughs in the data. Therefore, the Ricker wavelet is suitable for extracting features and detecting local extrema in time series. To implement the WT in Python I use the PyWavelets package (Lee et al., 2019). The boundaries of the signal are filled with zeros before the WT is applied. Figure 3.1a shows a two year temperature time series and Figure 3.1b the corresponding wavelet spectrum. The wavelet transformed timeseries (Fig. 3.1b) shows the seasonal change in the high periods and the alternation of weather in the low periods.

The WT has two major advantages. First, it provides an additional level of interpretability by decomposing the time series into periods. Second, images as input for neural network enables to make use of established convolutional neural networks from computer vision.

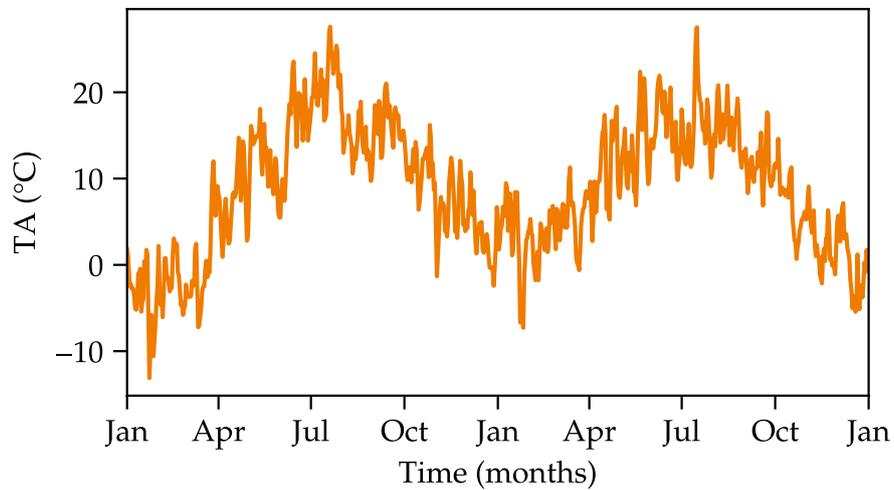
3.2. Convolutional Neural Networks

The main structure of a Convolutional Neural Network (CNN, LeCun et al. (1989)) consists of convolutional layers. Convolutional layers perform a convolution of the convolutional kernel with the layer's input matrix. As the convolutional kernel slides along the layer's input matrix, each value depends only on neighbouring inputs. The same convolution kernel is used for a single model layer. Depending on the kernel structure, convolution enhances features such as edges, trends or regions. It also reduces the number of free parameters and allows the network to be deeper. CNNs are very good at detecting features in images.

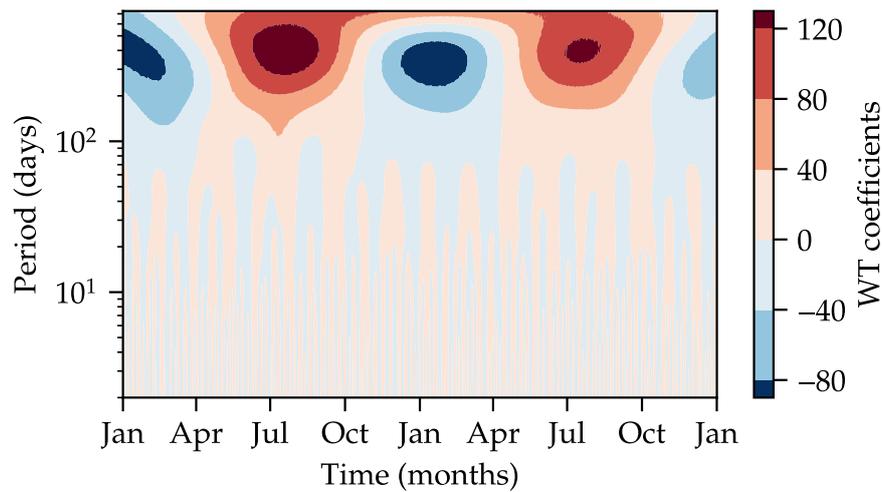
The CNN I use is a Residual Neural Network of 18 layers (ResNet-18, He et al. (2016)) pre-trained on ImageNet (Deng et al., 2009). ResNet is a deep learning model, i.e. it consists of several layers, and has shown very high performance in correctly classifying and recognising objects and scenes in images. To fit the framework, I change the input layers from fitting an image (224x224) with three channels (red, green, blue) to fitting an image of size 730x64 (730 days, 64 WT scales) with 12 channels. A channel is an input variable. I also change the output of the neural network to 730 values to predict 365 values each for NEE and LE.

3.3. Model architecture

Figure 3.2 shows the architecture of the model framework. The length of the input data is two years of daily values for each variable. The second year is the target year, for which the models predict daily values of NEE and LE. The first year is the memory year, which gives the model information about past meteorological events that may have a legacy effect on the ecosystem. In addition to the nine meteorological variables and one remote sensing variable, I add a random walk variable with autocorrelation. The random walk variable, in



(a) Example of a two-year air temperature time series.



(b) Wavelet transform of the two year air temperature time series from (a). It shows how well the time series at a given time corresponds to a given period in days.

Figure 3.1: Example of a two-year air temperature time series in (a) with its wavelet transform in (b).

combination with Integrated Gradients, helps to interpret the model. All two-year time series of the input variables are minimum-maximum normalised before and after the WT. All values are between zero and one. The resulting vector of 12 images or one image with 12 channels is fed into the ResNet model and the static climate variables (MAP, MAT) are added to the last layer of the ResNet model (late fusion). The model then directly predicts 365 daily values for NEE and LE for the target year. The ResNet model weights are updated with the loss calculated from the eddy covariance observations and the model prediction. For the results, I also calculate the total annual NEE from the 365 predicted NEE values.

3.4. Integrated Gradients

I use Integrated Gradients (IG, Sundararajan et al. (2017)) to interpret the machine learning model. IG is an explanation method from the field of Explainable Artificial Intelligence for neural networks and characterises the importance of the input features that contribute to the prediction. One of the essential points of IG is to set the baseline. The baseline, which can be zero or the mean state, is the starting point for calculating the gradient. For each input value, i.e. meteorological prediction in time and time scale, IG returns an importance score for the predicted output, i.e. annual carbon exchange, with respect to a specific baseline. IG calculates the gradients of the model's prediction along a straight path in the input space between the baseline and the actual input. The importance score is assigned to the integral over the gradients.

Applying IG to a model input gives an importance score for each value of the input vector and the two static variables. So for a given variable, the IG output resembles an image. I use the 99.9th percentile of the random variable IG output to set a minimum IG importance score for the other parameters. This reduces the uncertainty that an IG importance score was generated by chance and increases confidence in the results.

3.5. Model training and validation

The 14 sites were divided into training (nine sites with ~69.6% of siteyears), validation (two sites with ~10.7% of siteyears) and test data (three sites with ~19.6% of siteyears) to have a 70%/10%/20% split between training, validation and test data. The model framework based on wavelet transform and ResNet-18 is trained across sites on 78 siteyears (~69.6% of total siteyears) with the hyperparameter configuration of Table 3.1 and validated during training on the validation data (~10.7% of total siteyears). The model with the lowest mean square error on the validation data during training is selected.

To compare model performance, I include the mean seasonal cycle as a baseline and a linear regression framework version for comparison. I calculate the seasonal cycle by taking the average per day of the year across all training and validation site years. The linear regression framework is the same modelling framework based on the wavelet transform. However, the ResNet-18 is replaced by a linear regression model consisting of two linear layers. One linear layer applies a linear transformation to the incoming data. This model is

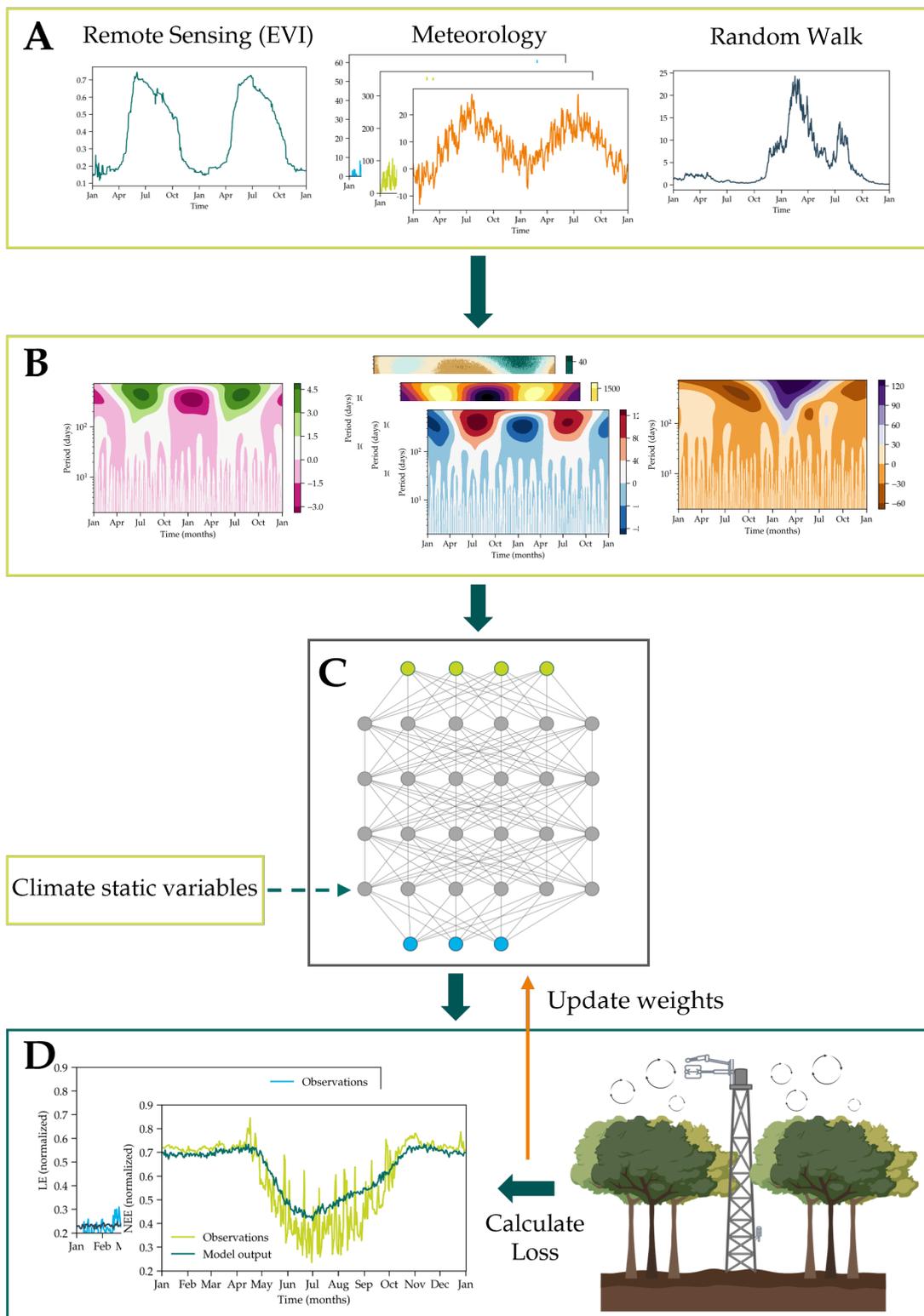


Figure 3.2: Model framework to predict carbon and water fluxes. **A** Example time series are shown for EVI, meteorological variables and the random walk. **B** The wavelet-transformed example time series are shown. **C** Static variables are attached with late fusion. Pre-trained Convolutional Neural Network (ResNet-18) predicts NEE and LE. **D** The loss is calculated by comparing observations with the model prediction. Then the weights of the model are updated. (Parts of this figure were created with BioRender.com)

also trained using the hyperparameters of Table 3.1.

Table 3.1: List of hyperparameters selected for training the ML framework. The parameters are set according to best practice.

Hyperparameter	Description	Defined Parameter
Learning rate	Step size at which a model's weights are updated	$\pi * 10^{-2}$
Learning rate scheduler	Technique to dynamically adjust the learning rate	CosineAnnealing-WarmRestarts (Loshchilov and Hutter, 2017)
Optimizer	Algorithm to update the model's parameter	Stochastic gradient descent
Batch size	Number of training samples in each iteration	4
Loss function	Measures the difference between predicted values and target values	Mean squared error
Epoch number	An epoch is one complete iteration through the entire training dataset	1000

4

Model Performance

To evaluate the model framework, I choose three different validation scores. First, I calculate the coefficient of determination (R^2) between observations and model predictions for daily values of NEE (and LE). I also calculate R^2 for anomalies by first subtracting the mean seasonal cycle per site. Additionally, I use the correlation coefficient of the annual sum of NEE (CC NEE_{yr}) as a simple measure to test the interannual variability (IAV). It is calculated per site between the predicted and observed annual total NEE (NEE_{yr}) for each siteyear. All validation scores are calculated per site and then averaged. The models are only tested on the three test sites that were not seen during training with 22 siteyears (~20% of the data). For the analysis, I only consider NEE as this is the target variable for the interpretation.

The model framework based on wavelet transformation and ResNet-18 (WCon) predicts the winter values and the slope of the seasonal cycle very well (Fig. 4.1). However, the amplitude of the seasonal cycle is not well captured and fluctuations during the growing season and winter are not reproduced. The same picture emerges for several years (Fig. 4.2). In addition, the model shows only minor variations between years. The validation scores reveal the same results.

The WCon model performs better than the baseline seasonal cycle (SC), but slightly worse than the linear version of the framework (WLR) for R^2 (Fig. 4.3). For the R^2 of anomalies, both models perform very poorly and are unable to capture any of the variability of the anomalies. For the correlation coefficient of NEE_{yr} , WCon performs better than WLR. WCon is able to capture 0.34 of the correlation of the annual sum of NEE between years. The WCon model is good at predicting high values, which are most likely winter NEE values (Fig. 4.4). It overestimates low NEE values. It appears that the model is not able to predict values below 0.40 and above ~0.75.

The WCon model captures the seasonal cycle quite well, but fails to describe anomalies in the NEE data. Furthermore, it can only partially capture the CC in NEE_{yr} . Most of the inter-annual variability is not captured. The slightly better performance of the WLR model for R^2 is surprising, as the CNN is expected to be better at extracting the main features in the wavelet-transformed input variables than the WLR. This may indicate that a deep CNN such as the ResNet-18 model is not necessary for predicting water and carbon fluxes. Overall, the WCon model is slightly better than the WLR and the baseline SC. The expectation of

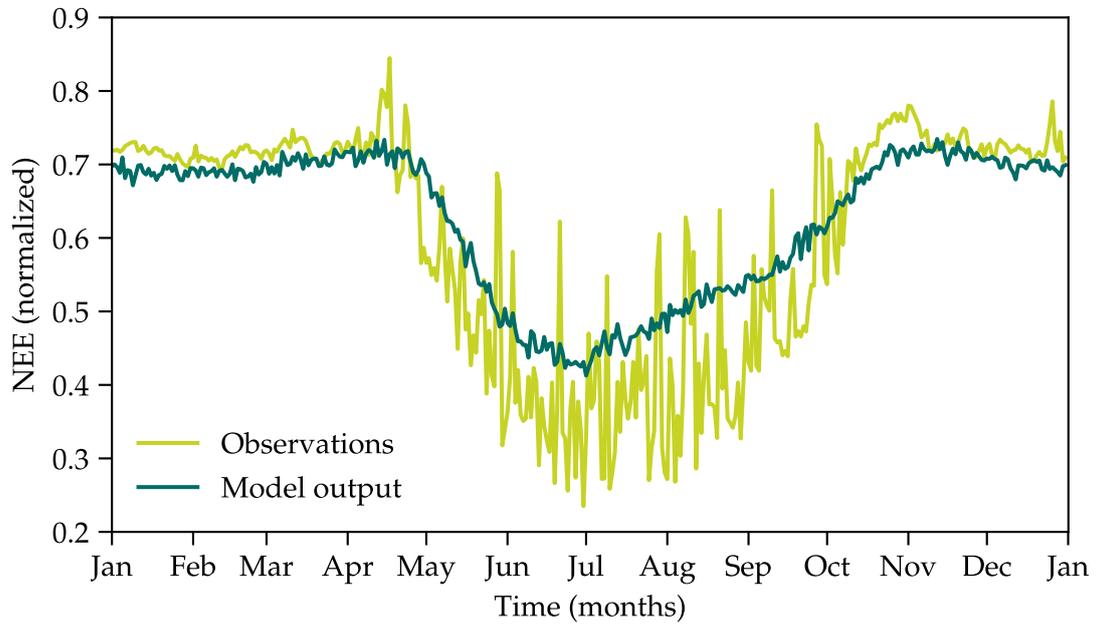


Figure 4.1: NEE Observations and WCon prediction for 2007 at the DE-Hai test site.

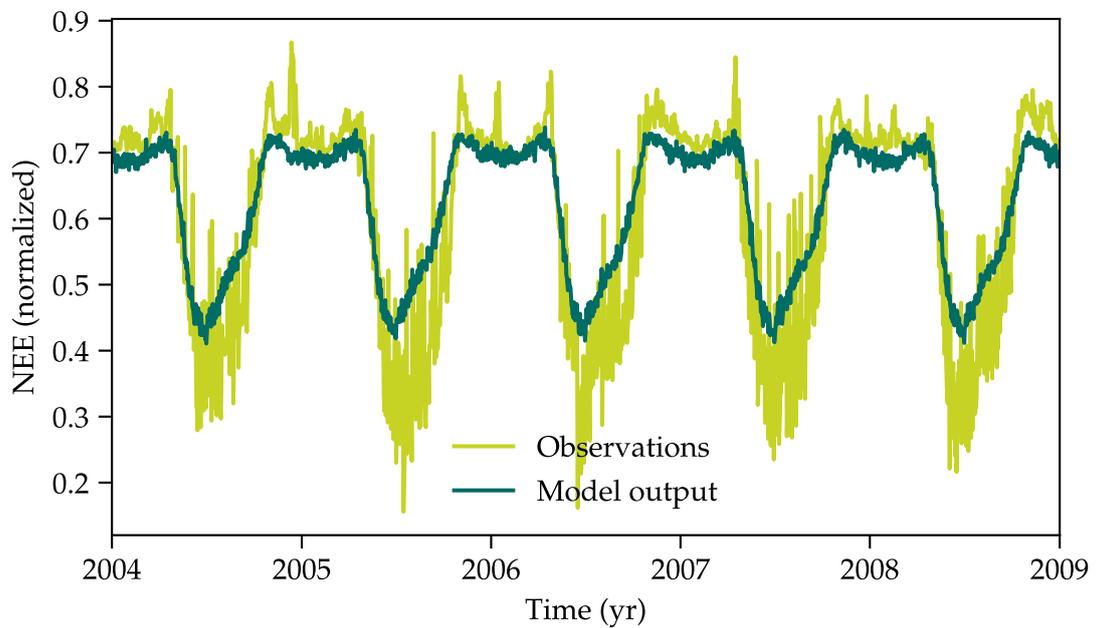


Figure 4.2: Same as 4.1, but for multiple years from 2004 to 2008.

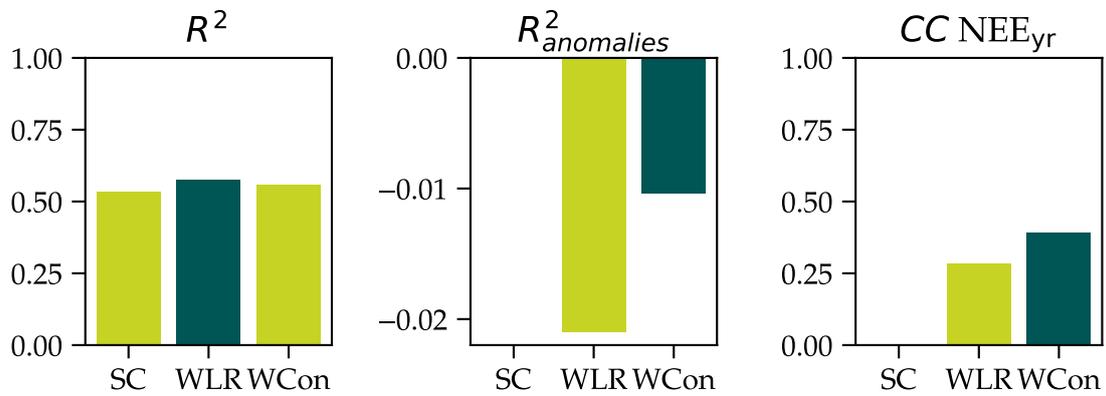


Figure 4.3: R^2 , R^2 of anomalies and CC of NEE_{yr} for the wavelet transform (WT) based models WCon (WT+ResNet-18) and WLR (WT+Linear Regression) as for the baseline seasonal cycle (SC). The dark green colour indicates the best performing model.

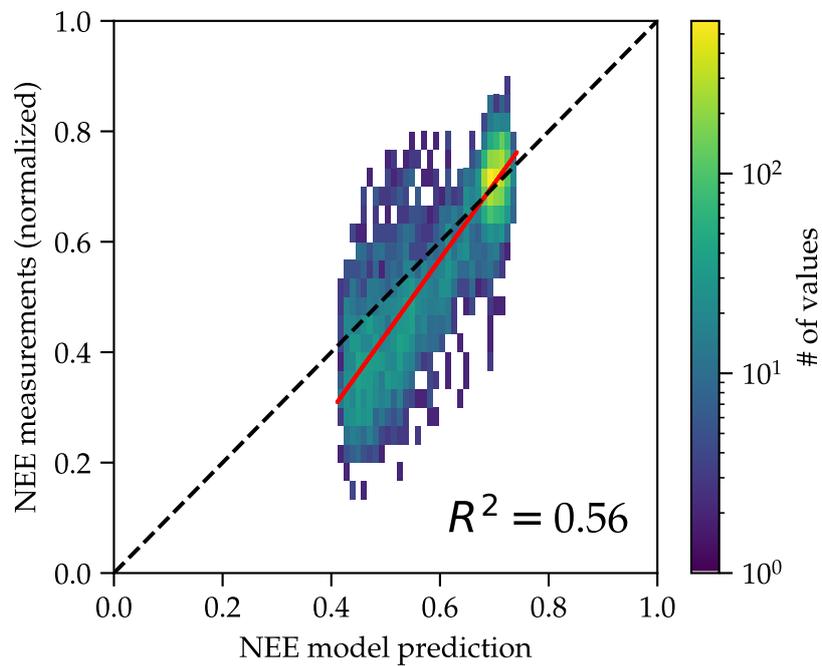


Figure 4.4: 2D histogram of WCon for the test data.

the deep convolutional network to extract more features from the data is not fulfilled under these training conditions. Nevertheless, the interpretive capacity of the framework is still a key advantage to interpret WCon prediction and investigate the meteorological drivers of the IAV in NEE_{yr} . Due to the poor performance of the model, caution should be exercised when applying the interpretation results to reality. In the following text, model or ML model always refers to the WCon model.

5

Main Drivers of Positive NEE Anomalies

Positive NEE anomalies are defined as periods with a higher NEE, i.e. the ecosystem stores less carbon, than the average. To identify positive NEE anomalies, I follow the definition of climate extremes in Reichstein et al. (2013):

“Conditions where an ecosystem function (such as carbon uptake) is higher or lower than a defined extreme percentile during a defined time period and over a certain area, traceable to single or multivariate anomalous meteorological variables.”

As ecosystem condition, I choose annual total NEE (sum of all daily values of one year, NEE_{yr}) and as extreme percentile the 90th. Therefore, for each site, all years with NEE_{yr} above the 90th NEE_{yr} percentile of that site are selected as positive anomalies. The 90th percentile is estimated using the linear method of the `numpy.percentile` function (NumPy, 2023). A total of 14 anomaly years are selected. Only sites with more than four years of data are considered. The NEE anomalies are site specific and are determined using the observational data. Looking at the local anomalies gives a more accurate picture of the meteorological factors. There are large differences in climate and mean NEE between sites. In a global view, effects such as site climate or forest condition could dominate the NEE. Looking at local anomalies should filter out most of the non-meteorological factors.

5.1. Methods

To determine the main drivers of the NEE_{yr} anomalies, I use the model framework and Integrated Gradients (IG, Sundararajan et al. (2017)). The target of IG is the annual nee total NEE_{yr} . Therefore, I add to the model framework an additional function that calculates NEE_{yr} from the predicted 365 daily values. IG has two inputs, firstly the target input data. This is the input data (meteorological data, EVI and the random walk variable) of the selected anomaly NEE_{yr} . The second input is for the baseline. In this case, these are the input variables averaged over all siteyears of the site. IG predicts with the model for the target and baseline inputs NEE_{yr} and returns the importance of each input value for the difference between the predicted target NEE_{yr} and the baseline NEE_{yr} . The baseline NEE_{yr} is a proxy for the mean NEE_{yr} of the site.

In a nutshell, I select the NEE_{yr} anomalies based on the NEE observations per site. I

then predict the same NEE_{yr} sum with the ML model. Finally, IG gives me the importance of each input value on predicting the anomaly (target NEE_{yr} - baseline NEE_{yr}). This process is repeated for all 14 of the selected anomaly years.

Due to the use of a proxy for the mean NEE_{yr} of a site and the poor model performance for anomalies, the difference between the predicted NEE_{yr} and the mean NEE_{yr} is significantly smaller than in the observations. Furthermore, in this analysis I am interpreting the predictions of the model, not the observations.

First, I am interested in the overall importance of each input variable. Therefore, for each of the selected anomaly examples, I calculate the total sum for each input variable. Once I take the absolute amount of IG importance scores (absolute IG importance) and once I take the actual IG importance scores before taking the sum. The absolute IG importance shows which input variable was overall most important for the model to predict NEE_{yr} anomaly. The actual IG importance shows the resulting effect of the input variable. A positive IG importance score means that the variable increases NEE_{yr} and therefore contributes to the anomaly. A negative IG importance means that the variable counteracts the anomaly and therefore provides more carbon storage. For the absolute IG importance scores, only those values above the 99.9th percentile of the absolute random walk importance scores are considered. To also reduce the impact of random IG importance values on the actual IG importance score, I set an upper and lower threshold for the IG values considered for the sum. The upper threshold is the same threshold as for the absolute importance score and the lower threshold is the negative upper threshold. Only IG values above and below the thresholds are considered. Using bootstrapping by varying the anomaly years in the sample, I create a box and whisker plot for the total absolute and actual importance of each input variable (Fig. 5.1 and 5.2).

I also calculate the sum for each IG importance score over all anomaly examples (See Figure 5.3 for VPD). Additionally, I want to identify the most important timings and scales of meteorological events. So for each day I calculate the sum over all periods (Fig. 5.4) and for each period I calculate the sum over all days (Fig. 5.5) for each input variable. Again, I only consider values above or below the threshold based on the 99.9th percentile of the random walk variable. The different scales of IG importance scores in Figures 5.3 - 5.5 result from the different methods of calculation.

5.2. Overall importance of input variables

Precipitation (P) has the highest absolute importance score, followed by vapour pressure deficit (VPD) and wind speed (WS) (Fig. 5.1). The large upper whisker for precipitation indicates that precipitation is highly important for a number of anomalies. For the actual importance values, the order changes (Fig. 5.2). Here VPD has the largest positive influence and therefore contributes to the anomaly. VPD is a major driver of the anomaly. Wind speed is third in absolute importance and has the highest negative score. Wind speed therefore has a counteracting effect and reduces the anomaly. The radiation variables incoming shortwave radiation (SW) and net radiation (NETRAD) play a less important role overall, but both variables drive anomalies as their importance scores are slightly positive. Air pressure (PA)

has an importance score close to zero and plays no role at all. The other radiation variables potential incoming shortwave radiation (SW_P) and its derivation (DV_SW_P) also have no influence, because for both variables the target input and the baseline input are the same.

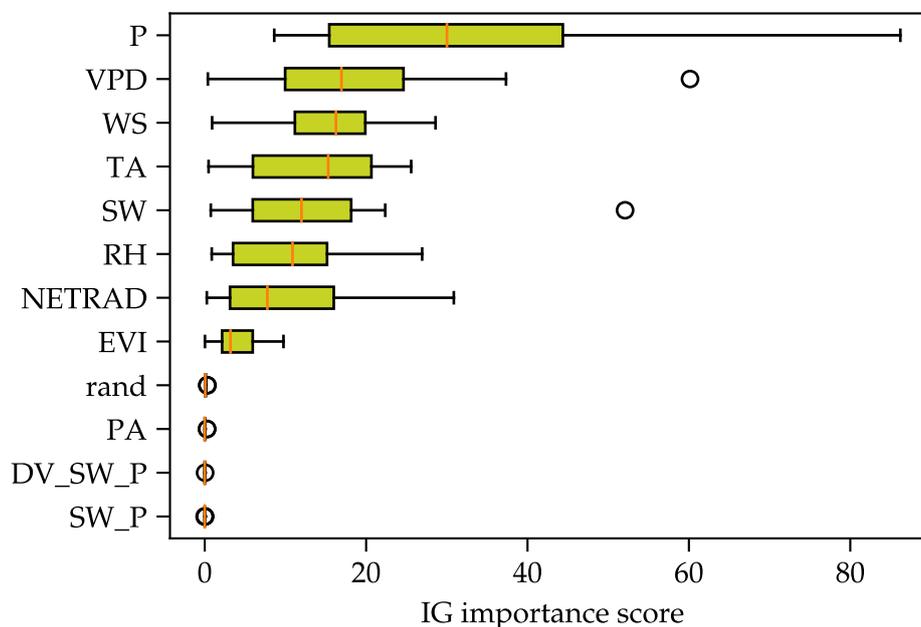


Figure 5.1: Box and whisker plot for the absolute total IG importance for each input variable.

The results show that water variables such as precipitation and VPD have the highest importance for the model and are, according to the model, the main drivers of the anomalies. In particular, VPD has a very positive effect on the NEE_{yr} anomaly. The vapour pressure deficit (VPD) is the difference between the saturated vapour pressure (e_s) calculated from the air temperature and the observed vapour pressure (e) (Eq. 5.1).

$$VPD = e_s(T) - e \quad (5.1)$$

VPD increases as the temperature rises or as the moisture in the air decreases. VPD is a combined measure of atmospheric heat and water. A high VPD is an indicator of high atmospheric water demand of atmospheric drought. As a result, plants evaporate more water and draw more water from their roots.

The high positive importance of VPD in combination with precipitation suggests that water availability for a site is the main factor for positive anomalies in NEE_{yr} . However, it is difficult to give a clear indication of ecosystem drought using VPD alone. This requires more information on soil water content. For example, a bucket model based on precipitation observations can estimate soil moisture.

In addition, VPD contains information from air temperature and has a similar meaning to relative humidity. Therefore, VPD could be the reason for a lower importance of temperature.

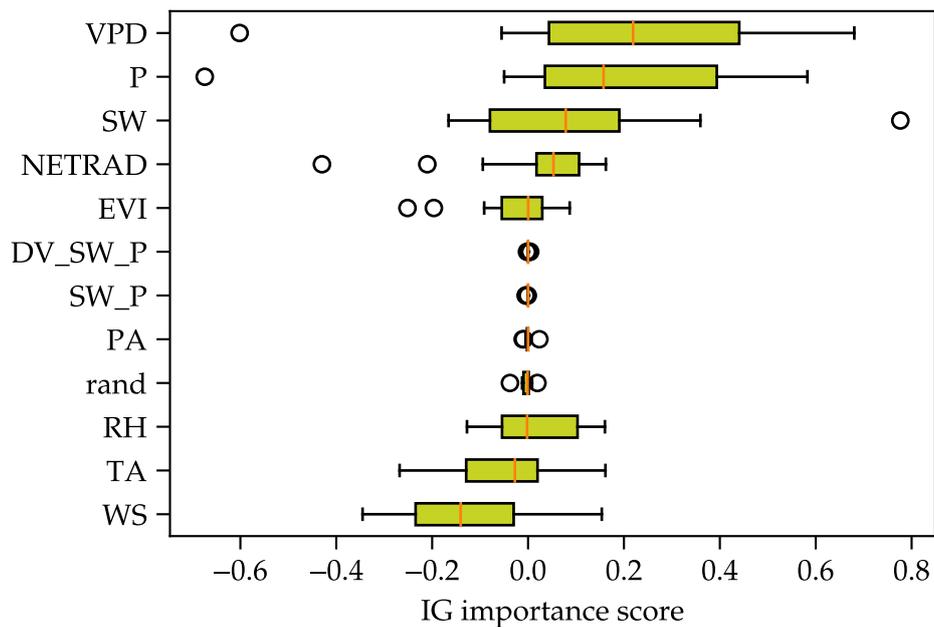


Figure 5.2: Box and whisker plot of actual total IG importance for each input variable.

Especially for the actual IG values (Fig. 5.2), VPD could hold a part of the positive temperature IG scores. Also, VPD may have a clearer relationship with NEE than RH due to its seasonal cycle, which causes the lower RH importance.

The interpretation of the model is in line with the results of Jung et al., who also report water availability as the main driver of interannual variability of NEE at the local scale. Even without clear information on soil moisture, the importance of VPD and precipitation may point to the effects of droughts on forest ecosystems carbon balance reported by Reichstein et al. (2013). However, this must be treated with caution, as it is not clear whether the high VPD importance scores are related to high VPD values.

The high importance of wind speed (WS) is surprising, especially the negative influence on positive NEE_{yr} anomalies. Frank et al. (2015) report on the vulnerability of forests to wind throw and its direct impact on the carbon balance. However, the interpretation results indicate that wind speed has an overall negative influence and causes more carbon storage. A clear explanation is difficult. Flux measurements depend on wind speed measurements. The high importance of WS could be an artefact of the use of wind speed for NEE measurements. Another reason could be that wind mixes the air through turbulence and transports carbon dioxide to plants, which increases GPP. Further research into the relationship between wind speed scales and their importance scores is needed to provide a clearer picture.

5.3. Temporal patterns of VPD importance

The previous results show that VPD is a dominant driver of NEE_{yr} . I use the interpretative power of the framework to look more closely at the importance of the timing and length of VPD events. Figure 5.3 shows that VPD events with a period of less than ten days do not play a significant role for the NEE_{yr} anomaly. VPD importance shows a seasonality. During the growing season, and especially in the summer months, VPD is more important than in the winter. The summer months of the target year are the most important, with the most important periods between 20 and 40 days. However, the summer months of the previous year also play a role. And it seems that the importance lies more in the higher periods, above 100 days.

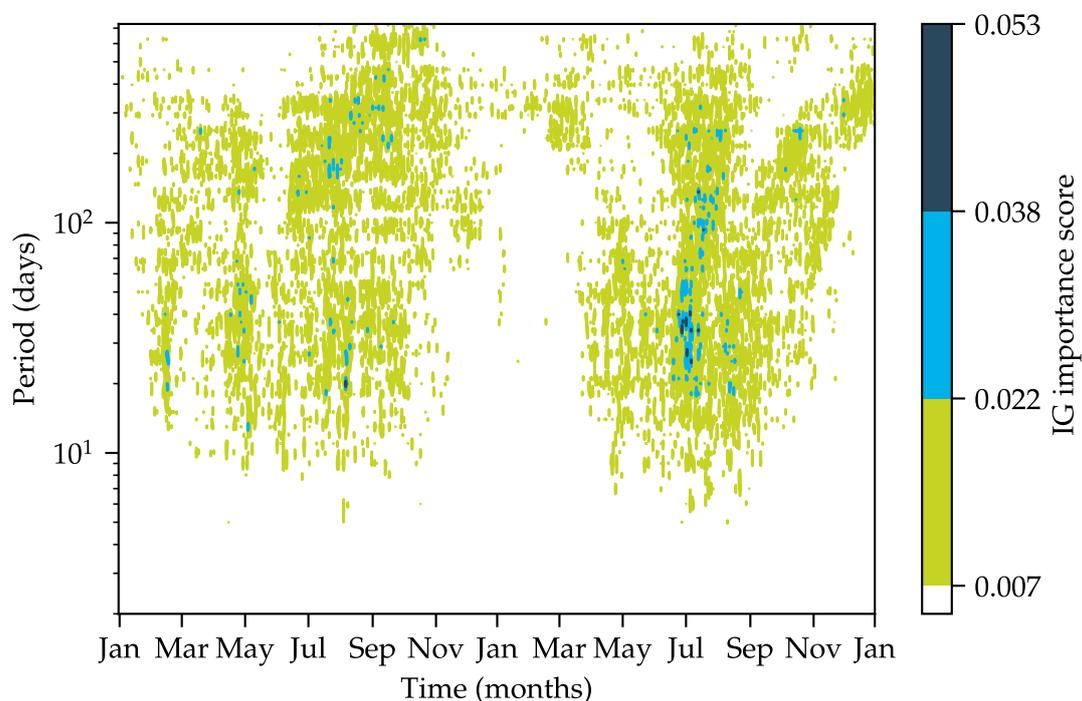


Figure 5.3: IG importance image for VPD. IG importance scores are summed over all anomaly examples.

Figure 5.4 highlights the importance of the summer of the target year. It also shows that the VPD is positive during these months, thus driving the anomalies. The memory year does not give a clear picture of the direction of the VPD importance values. Figure 5.5 compares the IG importance between the different periods. It shows two modes for VPD. One peak is around 25-40 days and the second peak is around 300-400 days.

The Integrated Gradients combined with the wavelet-transformed input reveal several things. Firstly, the model incorporates some physical plausibility. It is physically plausible that VPD has a higher importance during the growing season and especially in the summer months of

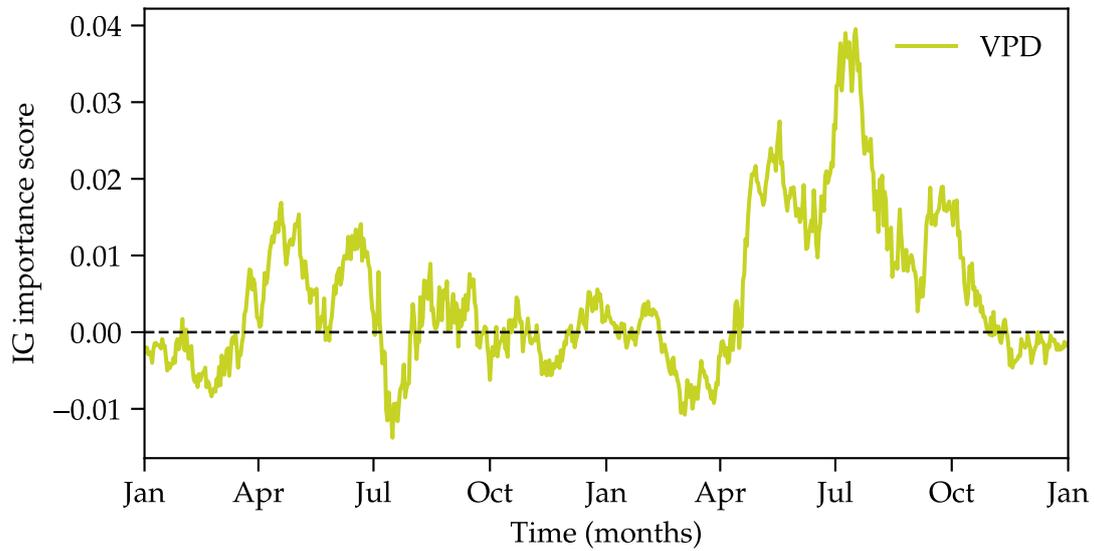


Figure 5.4: Daily actual IG scores summed over all periods for VPD. For better clarity, the data is smoothed with a 30-day moving average.

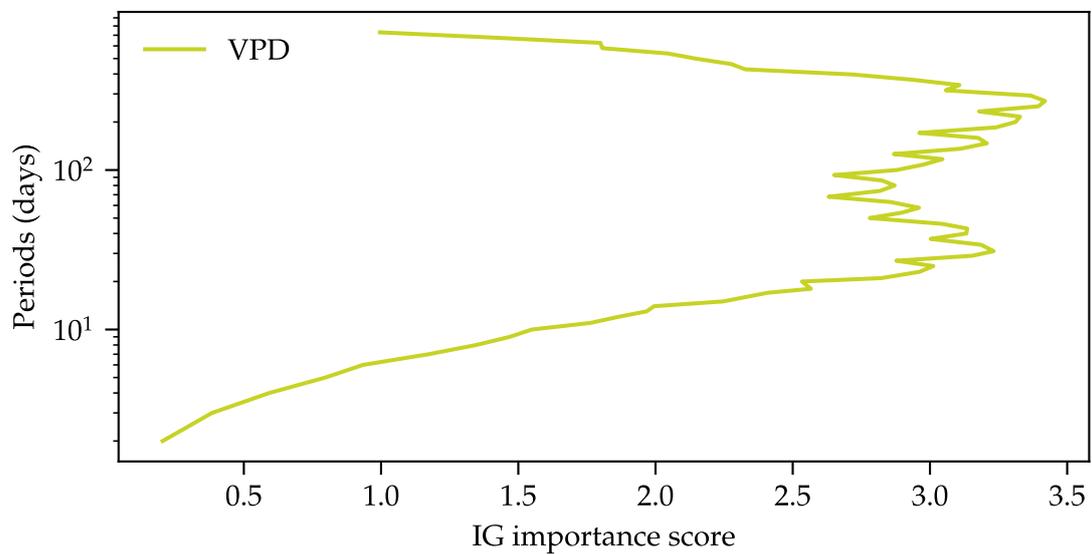


Figure 5.5: Absolute IG scores summed over all days for VPD. For better clarity, the data is smoothed with a five-period moving average.

the target year. In addition, the memory year also plays a role, with a slight shift in importance to higher periods, although to a lesser extent than in the target year.

Secondly, the interpretation methods are able to reveal temporal patterns of VPD importance and extract important periods of meteorological events. This provides a more accurate picture than reported in the literature. However, the limitations of IG are also apparent here. It is still unclear whether the high VPD importance scores are related to high or low VPD values and thus to possible droughts or heat waves or to other events such as storms or high precipitation events. With regard to the reported effects of summer droughts and heatwaves on the carbon balance of temperate forests (Ciais et al., 2005; Reichstein et al., 2013; Frank et al., 2015; Yu et al., 2022), the positive VPD importance is more likely to correlate with heatwaves or droughts.

6

Case Study: Heatwave 2003 in Hainich Forest Germany

The 2003 European heatwave with July temperatures up to 6 °C above long-term means and annual precipitation deficits up to 300 mm yr⁻¹ (Ciais et al., 2005) had far-reaching consequences for ecosystems and society. Ciais et al. (2005) reported a strong anomalous net source of carbon dioxide to the atmosphere, reversing the effect of four years of net ecosystem carbon uptake. Figure 6.1 shows the annual history of NEE_{yr} at the Hainich forest site in Germany (DE-Hai). The year 2003 shows a high NEE_{yr} value in the Hainich forest. The following year, 2004, is even higher and one of the highest NEE_{yr} measured at the site. This suggests strong legacy effects from the 2003 heatwave, as also reported in a study by Yu et al. (2022). I use the model framework and an interpretation scheme to investigate the drivers of the 2004 NEE_{yr} anomaly in more detail.

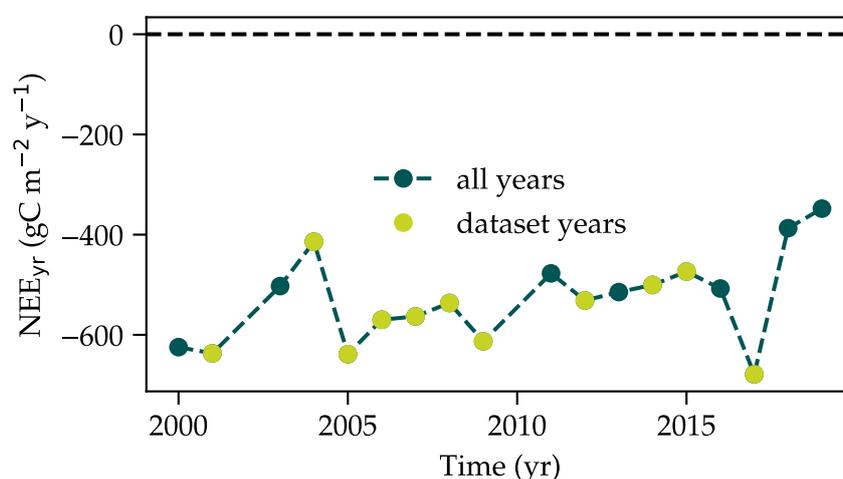


Figure 6.1: Interannual variation in NEE_{yr} for the Hainich forest site in Germany. All dots indicate observed siteyears. Light green dots indicate the siteyears that are also included in the test dataset.

6.1. Methods

To determine the overall importance of each variable I use a similar principle to that used in chapter 5. In this case study, the target input for IG are the input variables for the time period 2003 (memory year) till 2004 (target year) from the Hainich forest site. The ML model predicts daily values and calculates the NEE_{yr} for 2004. The baseline input are input variables averaged over the site. For each input variable I calculate the sum over all absolute IG importance scores. For the absolute total importance only values above the 99.9th percentile of the random variable IG scores are considered.

The interpretation scheme (Fig. 6.3) traces the model prediction back to a specific meteorological event. The first step is the absolute IG output for a given variable. Only IG importance scores above the 99.9th percentile of the random variable IG scores are shown. I find the most important point by smoothing the data with a Gaussian kernel and taking the maximum. This point in the two-dimensional space has a corresponding day and a corresponding time scale. The black dotted lines in all three plots of Figure 6.3 symbolise this length and time of the meteorological event detected by IG. The second plot shows the target input time series of the variable before the wavelet transform, compared to the baseline time series, which is the site mean. Both time series are smoothed with a seven-day moving average to improve clarity. The third plot shows the difference between the target wavelet transformed input and the baseline wavelet transformed input at the corresponding time period.

6.2. Atmospheric water demand caused by the 2003 heatwave has a legacy effect on 2004 NEE_{yr} anomaly

The total absolute IG sum in Figure 6.2 shows a similar order to the quantitative analysis in Figure 5.1. However, in this case the order of vapour pressure deficit (VPD) and precipitation (P) changes. VPD is clearly the most important variable for the NEE_{yr} 2004 anomaly. I take a closer look at the most important event for VPD using the interpretation scheme (Fig. 6.3).

The top plot in Figure 6.3 shows the IG output for VPD. The most important values are in the memory year and also in the higher periods. The Gaussian kernel also finds the most important point here (black box). With a timing in July and a length of 316 days it covers almost the whole of 2003. The second plot shows high VPD anomalies in the memory year 2003, especially during the summer. There is a clear positive anomaly for VPD when looking at the wavelet-transformed difference of the two time series in the corresponding period in the third plot. In 2003 the VPD was significantly higher than the site mean. Especially during the summer months there was a high atmospheric dryness. This had a significant impact on the model to predict the NEE_{yr} in 2004.

The model framework is able to show that the 2003 heat wave had a legacy effect on 2004. Most important were not only the high temperature anomalies, but the atmospheric dryness which is a consequence of the lack of precipitation and high temperatures in 2003. Reichstein et al. (2007) also report that precipitation and soil moisture deficit rather than

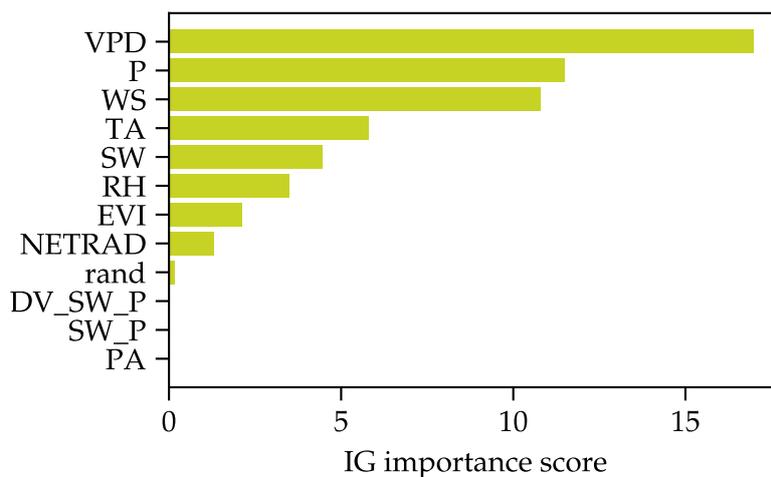


Figure 6.2: Total absolute IG sum for each input variable. Only values above the 99.9th percentile of the random variable IG scores are considered.

high temperatures were the main factor to reduce water and carbon fluxes. According to Frank et al. (2015), drought and high temperatures induce drought stress in vegetation, which negatively affects plant growth and plant health, and in turn increases plant mortality. The damage to plants caused by the 2003 heatwave could have consequences for carbon uptake in subsequent years. According to Yu et al. (2022) the legacy effect of the 2003 heatwave can be partly explained by reduced leaf development.

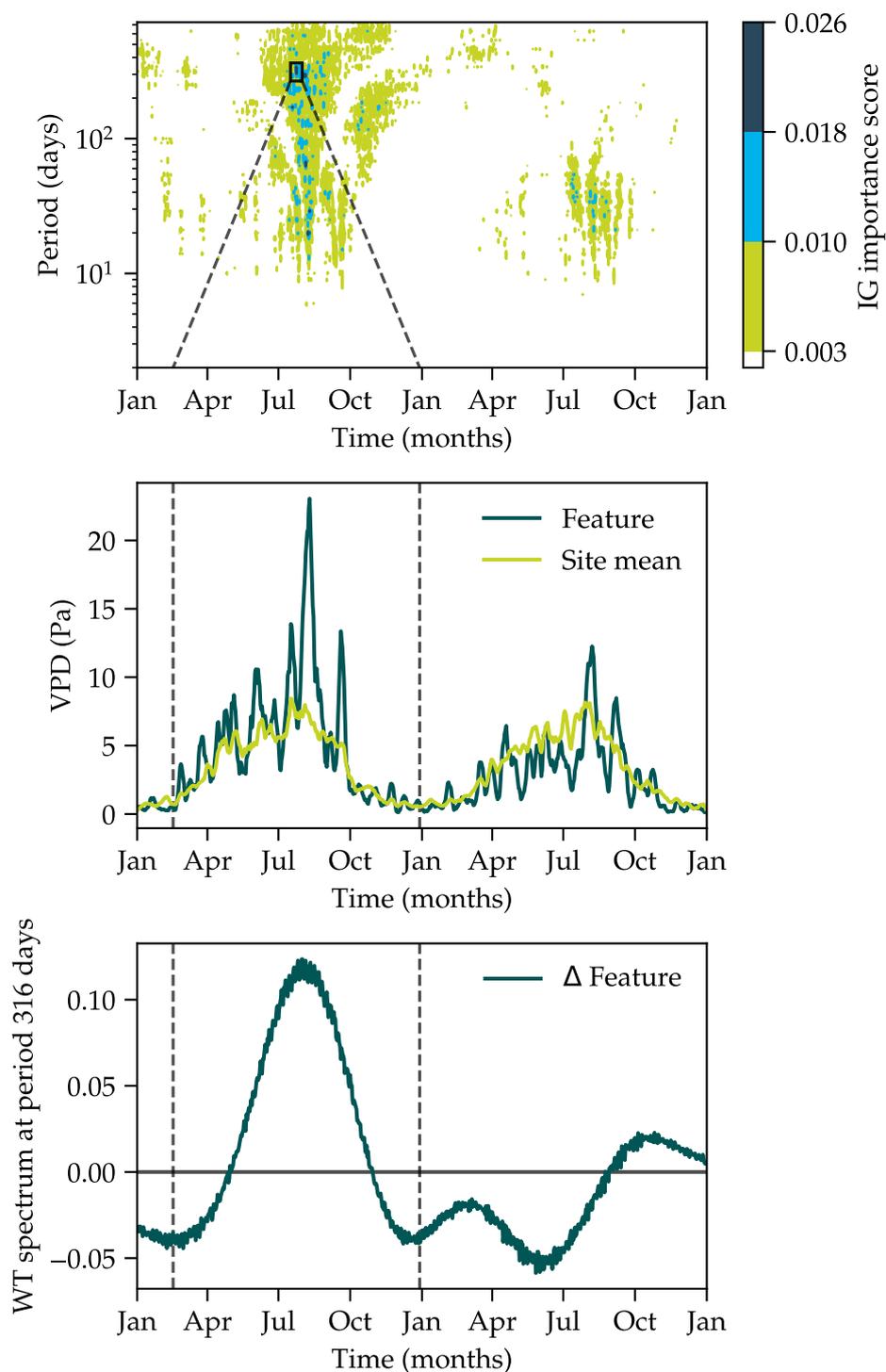


Figure 6.3: Interpretation scheme of VPD for 2004 of Hainich forest site in Germany. First plot shows the absolute IG importance scores for VPD. The black dotted lines in all three plots symbolise the length and time of the most important meteorological event detected by the Gaussian kernel (black box). The second plot shows the target input time series of the variable before the wavelet transform, compared to the baseline time series, which is the site mean. Both time series are smoothed with a seven day moving average. The third plot shows the difference between the target WT input and the baseline WT input at the corresponding time period.

7

Conclusion

The model framework captures the annual cycle of NEE at unseen sites well, but fails to reproduce anomalies. This reduces confidence in the results. Integrated gradients in combination with the wavelet-transformed time series can determine the importance of different input variables on NEE anomalies. Furthermore, the interpretation method is able to quantify the effects of multi-scale meteorological events on NEE_{yr} . It relates the model prediction to past meteorological events in terms of their position in time and scale. The interpretive capability demonstrates the potential of interpretable machine learning methods to accurately identify drivers of interannual variability in carbon fluxes and in Earth system science.

Using the model framework, I confirm the importance of water availability for the interannual variability of NEE_{yr} at site level. In particular, I show that the vapour pressure deficit is the most important driver of positive NEE_{yr} anomalies. The most important VPD events are those with a duration of about 20-40 days in the summer months of the target year. VPD events that occur in the memory year are quantitatively also of high importance and can lead to legacy effects. Furthermore, I analyse the drivers of the 2004 NEE_{yr} anomaly in the Hainich Forest, Germany, in a case study. My results show that the high VPD anomaly caused by the 2003 European heatwave was the dominant driver of the 2004 NEE_{yr} anomaly and thus had legacy effects.

The poor model performance on anomalies and the linkage of IG importance to the magnitude of input variables in the quantitative analysis are the main limitations of the modelling framework. However, both limitations can be improved with further studies. These studies are needed to increase confidence in the results.

8

Outlook

The poor performance of the model on anomalies is a major limitation in gaining more confidence in the results. There are several ways to improve the performance. One important point is to improve the performance by optimising the hyperparameters. For this thesis, I choose only a basic set of hyperparameters. Other possibilities are to change the model architecture. The loss function could be changed to include the loss for anomalies. This could shift the focus of the model from the seasonal cycle to anomalies. A different wavelet function may be better suited to meteorological data, and therefore lead to improved results.

Figure 4.1 shows that the model is not able to capture small variations in NEE on a daily basis. The time resolution of the daily input values is likely too coarse to represent short variations in NEE. Fluxnet and the downscaled Era5 products I use for gap filling provide hourly data. Reducing the time resolution to hourly data, with smaller scales for the wavelet transform, can help the model capture the daily variations in NEE and improve the overall model performance.

To overcome the interpretation limit in chapter 5, I can compare the IG output with the denormalised wavelet-transformed input. This could show quantitatively whether, for example, positive IG importance scores for VPD correspond to high VPD values.

The performance of the model and its interpretability can be improved by changing the input variables. It would be helpful to include information on soil moisture. Soil moisture is crucial for vegetation and therefore for NEE. Soil water content or soil moisture is only measured at a few Fluxnet sites. However, a bucket model based on precipitation could estimate soil moisture and provide useful information for the model. In addition, soil moisture information can be used in combination with VPD to accurately characterise droughts.

Different input variables contain similar information, e.g., RH and VPD. Reducing repetitive information, e.g., by removing RH as an input variable, reduces the complexity of the model and can lead to better interpretation results because the meaning is not shared between variables and the attribution is clearer. This effect could also be achieved by combining different water variables into a water availability index.

In addition to the experiments in Chapters 5 and 6, there are several other ways to assess meteorological effects on ecosystem carbon uptake. In chapter 5, I have only looked at

VPD in detail. The same methods can be used to investigate the importance of other variables and to further test the physical plausibility of the model. I want to repeat the experiment in chapter 5 to identify the driver of NEE_{yr} anomalies with a different baseline. Instead of using the mean input variables per site, the mean input variables across all sites can be used as the baseline. This reveals the drivers of NEE_{yr} anomalies as opposed to the mean meteorology across sites, providing a more global view. According to Jung et al. (2017), the importance of water availability should balance out between sites, leaving a dominant temperature signal in the annual variation of ecosystem carbon uptake. The interpretation scheme introduced in chapter 6 can be used to study already defined extreme events such as the 2003 heat wave in more detail. The interpretation method is able to attribute the NEE_{yr} anomaly to specific meteorological events. However, it remains unclear whether these meteorological events are extreme weather or climate events. Therefore, a further step is necessary to compare the identified event with data.

Another interesting experiment would be, instead of inferring meteorological factors from NEE anomalies, to identify extreme events in the input data and then investigate the consequences for NEE_{yr} . These extreme events could be droughts, extreme precipitation or 'false spring' events. Mahecha et al. (2022) report that the last two events in particular are understudied.

Keenan et al. (2014) show that a general trend towards an earlier start and later end of the growing season leads to increased carbon uptake through photosynthesis at US sites. The authors identify temperature as the main driver of these changes in phenology. In my opinion it is very interesting if changes in phenology due to meteorology and the resulting NEE consequences play a significant role for the IAV of NEE_{yr} . Looking at the main timing of temperature events could answer this question.

In this thesis I focus mainly on the IAV of NEE_{yr} . Looking at the annual sum of NEE_{yr} could lead to compensating effects between different events in a year. Perhaps a clearer signal could be obtained by looking at shorter-term NEE totals, such as monthly or seasonal totals.

Overall, there is a wide range of possibilities. If the model is able to capture NEE variation well at the site level, the possibility of spatially scaling up the model could be explored. Another possibility would be to include more forest plant functional types, such as evergreen or mixed forest, or to combine different models trained on these plant functional types to provide a more complete view of meteorological effects on ecosystem carbon uptake.

Acknowledgement

First of all I would like to thank my first supervisor Alexander J. Winkler for the great supervision and the opportunity to do a bachelor thesis on this exciting topic in cooperation with the Max Planck Institute for Biogeochemistry in Jena. I would also like to thank my second supervisor Felix Ament for his very helpful advice and for making this bachelor thesis possible.

Many thanks to Christian Reimers for the great co-supervision of my thesis and for always having an open door for questions.

I thank Jacob A. Nelson for providing the data for my thesis.

A very big thanks goes to my sister Schirin for her great support and advice as well as my parents for always being there for me.

Many thanks to Tim Schneider for proofreading my thesis and giving me very helpful comments.

And last but not least, I would like to thank my friends for their support and especially my fellow students in Hamburg for making my bachelor a wonderful time.

References

- Baldocchi, D., Chu, H., and Reichstein, M.: Inter-annual variability of net and gross ecosystem carbon fluxes: A review, *Agricultural and Forest Meteorology*, 249, 520–533, <https://doi.org/10.1016/j.agrformet.2017.05.015>, 2018.
- Ciais, P., Reichstein, M., Viovy, N., Granier, A., Ogée, J., Allard, V., Aubinet, M., Buchmann, N., Bernhofer, C., Carrara, A., Chevallier, F., De Noblet, N., Friend, A. D., Friedlingstein, P., Grünwald, T., Heinesch, B., Keronen, P., Knohl, A., Krinner, G., Loustau, D., Manca, G., Matteucci, G., Miglietta, F., Ourcival, J. M., Papale, D., Pilegaard, K., Rambal, S., Seufert, G., Soussana, J. F., Sanz, M. J., Schulze, E. D., Vesala, T., and Valentini, R.: Europe-wide reduction in primary productivity caused by the heat and drought in 2003, *Nature*, 437, 529–533, <https://doi.org/10.1038/nature03972>, number: 7058 Publisher: Nature Publishing Group, 2005.
- Daubechies, I.: The wavelet transform, time-frequency localization and signal analysis, *IEEE Transactions on Information Theory*, 36, 961–1005, <https://doi.org/10.1109/18.57199>, conference Name: IEEE Transactions on Information Theory, 1990.
- Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., and Fei-Fei, L.: ImageNet: A large-scale hierarchical image database, in: 2009 IEEE Conference on Computer Vision and Pattern Recognition, pp. 248–255, <https://doi.org/10.1109/CVPR.2009.5206848>, iISSN: 1063-6919, 2009.
- Frank, D., Reichstein, M., Bahn, M., Thonicke, K., Frank, D., Mahecha, M. D., Smith, P., van der Velde, M., Vicca, S., Babst, F., Beer, C., Buchmann, N., Canadell, J. G., Ciais, P., Cramer, W., Ibrom, A., Miglietta, F., Poulter, B., Rammig, A., Seneviratne, S. I., Walz, A., Wattenbach, M., Zavala, M. A., and Zscheischler, J.: Effects of climate extremes on the terrestrial carbon cycle: concepts, processes and potential future impacts, *Global Change Biology*, 21, 2861–2880, <https://doi.org/10.1111/gcb.12916>, <https://onlinelibrary.wiley.com/doi/pdf/10.1111/gcb.12916>, 2015.
- Friedlingstein, P., Cox, P., Betts, R., Bopp, L., Bloh, W. v., Brovkin, V., Cadule, P., Doney, S., Eby, M., Fung, I., Bala, G., John, J., Jones, C., Joos, F., Kato, T., Kawamiya, M., Knorr, W., Lindsay, K., Matthews, H. D., Raddatz, T., Rayner, P., Reick, C., Roeckner, E., Schnitzler, K.-G., Schnur, R., Strassmann, K., Weaver, A. J., Yoshikawa, C., and Zeng, N.: Climate–Carbon Cycle Feedback Analysis: Results from the C4MIP Model Intercomparison, *Journal of Climate*, 19, 3337–3353, <https://doi.org/10.1175/JCLI3800.1>, publisher: American Meteorological Society Section: Journal of Climate, 2006.
- Friedlingstein, P., Jones, M. W., O’Sullivan, M., Andrew, R. M., Hauck, J., Peters, G. P., Peters, W., Pongratz, J., Sitch, S., Le Quéré, C., Bakker, D. C. E., Canadell, J. G., Ciais, P., Jackson, R. B., Anthoni, P., Barbero, L., Bastos, A., Bastrikov, V., Becker, M., Bopp, L., Buitenhuis, E., Chandra, N., Chevallier, F., Chini, L. P., Currie, K. I., Feely, R. A., Gehlen, M., Gilfillan,

- D., Gkritzalis, T., Goll, D. S., Gruber, N., Gutekunst, S., Harris, I., Haverd, V., Houghton, R. A., Hurtt, G., Ilyina, T., Jain, A. K., Joetzjer, E., Kaplan, J. O., Kato, E., Klein Goldewijk, K., Korsbakken, J. I., Landschützer, P., Lauvset, S. K., Lefèvre, N., Lenton, A., Lienert, S., Lombardozzi, D., Marland, G., McGuire, P. C., Melton, J. R., Metzl, N., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S.-I., Neill, C., Omar, A. M., Ono, T., Pregon, A., Pierrot, D., Poulter, B., Rehder, G., Resplandy, L., Robertson, E., Rödenbeck, C., Séférian, R., Schwinger, J., Smith, N., Tans, P. P., Tian, H., Tilbrook, B., Tubiello, F. N., van der Werf, G. R., Wiltshire, A. J., and Zaehle, S.: Global Carbon Budget 2019, *Earth System Science Data*, 11, 1783–1838, <https://doi.org/10.5194/essd-11-1783-2019>, publisher: Copernicus GmbH, 2019.
- Friedlingstein, P., O’Sullivan, M., Jones, M. W., Andrew, R. M., Gregor, L., Hauck, J., Le Quéré, C., Luijkx, I. T., Olsen, A., Peters, G. P., Peters, W., Pongratz, J., Schwingshackl, C., Sitch, S., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S. R., Alkama, R., Arneeth, A., Arora, V. K., Bates, N. R., Becker, M., Bellouin, N., Bittig, H. C., Bopp, L., Chevallier, F., Chini, L. P., Cronin, M., Evans, W., Falk, S., Feely, R. A., Gasser, T., Gehlen, M., Gkritzalis, T., Gloege, L., Grassi, G., Gruber, N., Gürses, , Harris, I., Hefner, M., Houghton, R. A., Hurtt, G. C., Iida, Y., Ilyina, T., Jain, A. K., Jersild, A., Kadono, K., Kato, E., Kennedy, D., Klein Goldewijk, K., Knauer, J., Korsbakken, J. I., Landschützer, P., Lefèvre, N., Lindsay, K., Liu, J., Liu, Z., Marland, G., Mayot, N., McGrath, M. J., Metzl, N., Monacci, N. M., Munro, D. R., Nakaoka, S.-I., Niwa, Y., O’Brien, K., Ono, T., Palmer, P. I., Pan, N., Pierrot, D., Pockock, K., Poulter, B., Resplandy, L., Robertson, E., Rödenbeck, C., Rodriguez, C., Rosan, T. M., Schwinger, J., Séférian, R., Shutler, J. D., Skjelvan, I., Steinhoff, T., Sun, Q., Sutton, A. J., Sweeney, C., Takao, S., Tanhua, T., Tans, P. P., Tian, X., Tian, H., Tilbrook, B., Tsujino, H., Tubiello, F., van der Werf, G. R., Walker, A. P., Wanninkhof, R., Whitehead, C., Willstrand Wranne, A., Wright, R., Yuan, W., Yue, C., Yue, X., Zaehle, S., Zeng, J., and Zheng, B.: Global Carbon Budget 2022, *Earth System Science Data*, 14, 4811–4900, <https://doi.org/10.5194/essd-14-4811-2022>, publisher: Copernicus GmbH, 2022.
- Hafezi Rachti, D., Reimers, C., Liu, G., and Winkler, A. J.: Predicting Vegetation Phenology using Machine Learning based on Wavelet Transform of Meteorological Drivers, Tech. Rep. EGU23-10141, Copernicus Meetings, <https://doi.org/10.5194/egusphere-egu23-10141>, conference Name: EGU23, 2023.
- He, K., Zhang, X., Ren, S., and Sun, J.: Deep Residual Learning for Image Recognition, pp. 770–778, URL https://openaccess.thecvf.com/content_cvpr_2016/html/He_Deep_Residual_Learning_CVPR_2016_paper.html, 2016.
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, *Quarterly Journal of the Royal Meteorological Society*, 146, 1999–2049, <https://doi.org/10.1002/qj.3803>, <https://onlinelibrary.wiley.com/doi/pdf/10.1002/qj.3803>, 2020.
- Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., and Ferreira, L. G.: Overview of

- the radiometric and biophysical performance of the MODIS vegetation indices, *Remote sensing of environment*, 83, 195–213, 2002.
- Jung, M., Reichstein, M., Schwalm, C. R., Huntingford, C., Sitch, S., Ahlström, A., Arneeth, A., Camps-Valls, G., Ciais, P., Friedlingstein, P., Gans, F., Ichii, K., Jain, A. K., Kato, E., Papale, D., Poulter, B., Raduly, B., Rödenbeck, C., Tramontana, G., Viovy, N., Wang, Y.-P., Weber, U., Zaehle, S., and Zeng, N.: Compensatory water effects link yearly global land CO₂ sink changes to temperature, *Nature*, 541, 516–520, <https://doi.org/10.1038/nature20780>, number: 7638 Publisher: Nature Publishing Group, 2017.
- Jung, M., Nelson, J., Migliavacca, M., El-Madany, T., Papale, D., Reichstein, M., Walther, S., and Wutzler, T.: Technical Note: Flagging inconsistencies in flux tower data, *Biogeosciences Discussions*, pp. 1–45, <https://doi.org/10.5194/bg-2023-110>, publisher: Copernicus GmbH, 2023.
- Keenan, T. F., Gray, J., Friedl, M. A., Toomey, M., Bohrer, G., Hollinger, D. Y., Munger, J. W., O’Keefe, J., Schmid, H. P., Wing, I. S., Yang, B., and Richardson, A. D.: Net carbon uptake has increased through warming-induced changes in temperate forest phenology, *Nature Climate Change*, 4, 598–604, <https://doi.org/10.1038/nclimate2253>, number: 7 Publisher: Nature Publishing Group, 2014.
- LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., and Jackel, L. D.: Backpropagation Applied to Handwritten Zip Code Recognition, *Neural Computation*, 1, 541–551, <https://doi.org/10.1162/neco.1989.1.4.541>, conference Name: Neural Computation, 1989.
- Lee, G. R., Gommers, R., Waselewski, F., Wohlfahrt, K., and O’Leary, A.: Py-Wavelets: A Python package for wavelet analysis, *Journal of Open Source Software*, 4, 1237, <https://doi.org/10.21105/joss.01237>, 2019.
- Loshchilov, I. and Hutter, F.: SGDR: Stochastic Gradient Descent with Warm Restarts, <https://doi.org/10.48550/arXiv.1608.03983>, arXiv:1608.03983 [cs, math], 2017.
- Mahecha, M. D., Bastos, A., Bohn, F. J., Eisenhauer, N., Feilhauer, H., Hartmann, H., Hickler, T., Kalesse-Los, H., Migliavacca, M., Otto, F. E. L., Peng, J., Quaas, J., Tegen, I., Weigelt, A., Wendisch, M., and Wirth, C.: Biodiversity loss and climate extremes — study the feedbacks, *Nature*, 612, 30–32, <https://doi.org/10.1038/d41586-022-04152-y>, bandiera_abtest: a Cg_type: Comment Number: 7938 Publisher: Nature Publishing Group Subject_term: Biodiversity, Climate change, Solid Earth sciences, 2022.
- NumPy: NumPy - NumPy percentile, URL <https://numpy.org/doc/stable/reference/generated/numpy.percentile.html>, accessed: September 22, 2023, 2023.
- Pastorello, G., Trotta, C., Canfora, E., Chu, H., Christianson, D., Cheah, Y.-W., Poindexter, C., Chen, J., Elbashandy, A., Humphrey, M., Isaac, P., Polidori, D., Reichstein, M., Ribeca, A., van Ingen, C., Vuichard, N., Zhang, L., Amiro, B., Ammann, C., Arain, M. A., Ardö, J., Arkebauer, T., Arndt, S. K., Arriga, N., Aubinet, M., Aurela, M., Baldocchi, D., Barr, A., Beamesderfer, E., Marchesini, L. B., Bergeron, O., Beringer, J., Bernhofer, C., Berveiller, D., Billesbach, D., Black, T. A., Blanken, P. D., Bohrer, G., Boike, J., Bolstad, P. V., Bonal, D.,

- Bonnefond, J.-M., Bowling, D. R., Bracho, R., Brodeur, J., Brümmer, C., Buchmann, N., Burban, B., Burns, S. P., Buysse, P., Cale, P., Cavagna, M., Cellier, P., Chen, S., Chini, I., Christensen, T. R., Cleverly, J., Collalti, A., Consalvo, C., Cook, B. D., Cook, D., Coursolle, C., Cremonese, E., Curtis, P. S., D'Andrea, E., da Rocha, H., Dai, X., Davis, K. J., Cinti, B. D., Grandcourt, A. d., Ligne, A. D., De Oliveira, R. C., Delpierre, N., Desai, A. R., Di Bella, C. M., Tommasi, P. d., Dolman, H., Domingo, F., Dong, G., Dore, S., Duce, P., Dufrière, E., Dunn, A., Dušek, J., Eamus, D., Eichelmann, U., ElKhidir, H. A. M., Eugster, W., Ewenz, C. M., Ewers, B., Famulari, D., Fares, S., Feigenwinter, I., Feitz, A., Fensholt, R., Filippa, G., Fischer, M., Frank, J., Galvagno, M., Gharun, M., Gianelle, D., Gielen, B., Gioli, B., Gitelson, A., Goded, I., Goeckede, M., Goldstein, A. H., Gough, C. M., Goulden, M. L., Graf, A., Griebel, A., Gruening, C., Grünwald, T., Hammerle, A., Han, S., Han, X., Hansen, B. U., Hanson, C., Hatakka, J., He, Y., Hehn, M., Heinesch, B., Hinko-Najera, N., Hörtnagl, L., Hutley, L., Ibrom, A., Ikawa, H., Jackowicz-Korczynski, M., Janouš, D., Jans, W., Jassal, R., Jiang, S., Kato, T., Khomik, M., Klatt, J., Knohl, A., Knox, S., Kobayashi, H., Koerber, G., Kolle, O., Kosugi, Y., Kotani, A., Kowalski, A., Kruijt, B., Kurbatova, J., Kutsch, W. L., Kwon, H., Launiainen, S., Laurila, T., Law, B., Leuning, R., Li, Y., Liddell, M., Limousin, J.-M., Lion, M., Liska, A. J., Lohila, A., López-Ballesteros, A., López-Blanco, E., Loubet, B., Loustau, D., Lucas-Moffat, A., Lüers, J., Ma, S., Macfarlane, C., Magliulo, V., Maier, R., Mammarella, I., Manca, G., Marcolla, B., Margolis, H. A., Marras, S., Massman, W., Mastepanov, M., Matamala, R., Matthes, J. H., Mazzenga, F., McCaughey, H., McHugh, I., McMillan, A. M. S., Merbold, L., Meyer, W., Meyers, T., Miller, S. D., Minerbi, S., Moderow, U., Monson, R. K., Montagnani, L., Moore, C. E., Moors, E., Moreaux, V., Moureaux, C., Munger, J. W., Nakai, T., Neiryneck, J., Nestic, Z., Nicolini, G., Noormets, A., Northwood, M., Noretto, M., Nouvellon, Y., Novick, K., Oechel, W., Olesen, J. E., Ourcival, J.-M., Papuga, S. A., Parmentier, F.-J., Paul-Limoges, E., Pavelka, M., Peichl, M., Pendall, E., Phillips, R. P., Pilegaard, K., Pirk, N., Posse, G., Powell, T., Prasse, H., Prober, S. M., Rambal, S., Rannik, , Raz-Yaseef, N., Reibmann, C., Reed, D., Dios, V. R. d., Restrepo-Coupe, N., Reverter, B. R., Roland, M., Sabbatini, S., Sachs, T., Saleska, S. R., Sánchez-Cañete, E. P., Sanchez-Mejia, Z. M., Schmid, H. P., Schmidt, M., Schneider, K., Schrader, F., Schroder, I., Scott, R. L., Sedláč, P., Serrano-Ortíz, P., Shao, C., Shi, P., Shironya, I., Siebicke, L., Šigut, L., Silberstein, R., Sirca, C., Spano, D., Steinbrecher, R., Stevens, R. M., Sturtevant, C., Suyker, A., Tagesson, T., Takanashi, S., Tang, Y., Tapper, N., Thom, J., Tomassucci, M., Tuovinen, J.-P., Urbanski, S., Valentini, R., van der Molen, M., van Gorsel, E., van Huissteden, K., Varlagin, A., Verfaillie, J., Vesala, T., Vincke, C., Vitale, D., Vygodskaya, N., Walker, J. P., Walter-Shea, E., Wang, H., Weber, R., Westermann, S., Wille, C., Wofsy, S., Wohlfahrt, G., Wolf, S., Woodgate, W., Li, Y., Zampedri, R., Zhang, J., Zhou, G., Zona, D., Agarwal, D., Biraud, S., Torn, M., and Papale, D.: The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data, *Scientific Data*, 7, 225, <https://doi.org/10.1038/s41597-020-0534-3>, number: 1 Publisher: Nature Publishing Group, 2020.
- Piao, S., Wang, X., Wang, K., Li, X., Bastos, A., Canadell, J. G., Ciais, P., Friedlingstein, P., and Sitch, S.: Interannual variation of terrestrial carbon cycle: Issues and perspectives, *Global Change Biology*, 26, 300–318, <https://doi.org/10.1111/gcb.14884>, [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/gcb.14884](https://onlinelibrary.wiley.com/doi/pdf/10.1111/gcb.14884), 2020.
- Reichstein, M., Ciais, P., Papale, D., Valentini, R., Running, S., Viovy, N., Cramer, W., Granier, A., Ogée, J., Allard, V., Aubinet, M., Bernhofer, C., Buchmann, N., Carrara, A.,

- Grünwald, T., Heimann, M., Heinesch, B., Knohl, A., Kutsch, W., Loustau, D., Manca, G., Matteucci, G., Miglietta, F., Ourcival, J., Pilegaard, K., Pumpanen, J., Rambal, S., Schaphoff, S., Seufert, G., Soussana, J.-F., Sanz, M.-J., Vesala, T., and Zhao, M.: Reduction of ecosystem productivity and respiration during the European summer 2003 climate anomaly: a joint flux tower, remote sensing and modelling analysis, *Global Change Biology*, 13, 634–651, <https://doi.org/10.1111/j.1365-2486.2006.01224.x>, eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-2486.2006.01224.x>, 2007.
- Reichstein, M., Bahn, M., Ciais, P., Frank, D., Mahecha, M. D., Seneviratne, S. I., Zscheischler, J., Beer, C., Buchmann, N., Frank, D. C., Papale, D., Rammig, A., Smith, P., Thonicke, K., van der Velde, M., Vicca, S., Walz, A., and Wattenbach, M.: Climate extremes and the carbon cycle, *Nature*, 500, 287–295, <https://doi.org/10.1038/nature12350>, number: 7462 Publisher: Nature Publishing Group, 2013.
- Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., and Prabhat: Deep learning and process understanding for data-driven Earth system science, *Nature*, 566, 195–204, <https://doi.org/10.1038/s41586-019-0912-1>, number: 7743 Publisher: Nature Publishing Group, 2019.
- Ricker, N.: Further developments in the wavelet theory of seismogram structure*, *Bulletin of the Seismological Society of America*, 33, 197–228, <https://doi.org/10.1785/BSSA0330030197>, 1943.
- Shao, J., Zhou, X., Luo, Y., Li, B., Aurela, M., Billesbach, D., Blanken, P. D., Bracho, R., Chen, J., Fischer, M., Fu, Y., Gu, L., Han, S., He, Y., Kolb, T., Li, Y., Nagy, Z., Niu, S., Oechel, W. C., Pinter, K., Shi, P., Suyker, A., Torn, M., Varlagin, A., Wang, H., Yan, J., Yu, G., and Zhang, J.: Biotic and climatic controls on interannual variability in carbon fluxes across terrestrial ecosystems, *Agricultural and Forest Meteorology*, 205, 11–22, <https://doi.org/10.1016/j.agrformet.2015.02.007>, 2015.
- Shevliakova, E., Stouffer, R. J., Malyshev, S., Krasting, J. P., Hurtt, G. C., and Pacala, S. W.: Historical warming reduced due to enhanced land carbon uptake, *Proceedings of the National Academy of Sciences*, 110, 16730–16735, <https://doi.org/10.1073/pnas.1314047110>, publisher: Proceedings of the National Academy of Sciences, 2013.
- Sundararajan, M., Taly, A., and Yan, Q.: Axiomatic Attribution for Deep Networks, URL <http://arxiv.org/abs/1703.01365>, arXiv:1703.01365 [cs], 2017.
- Thornton, M., Shrestha, R., Wei, Y., Thornton, P., Kao, S.-C., and Wilson, B.: Daymet: Daily Surface Weather Data on a 1-km Grid for North America, Version 4 R1, <https://doi.org/10.3334/ORNLDAAC/2129>, 2022.
- Walther, S., Besnard, S., Nelson, J. A., El-Madany, T. S., Migliavacca, M., Weber, U., Carvalhais, N., Ermida, S. L., Brümmer, C., Schrader, F., Prokushkin, A. S., Panov, A. V., and Jung, M.: Technical note: A view from space on global flux towers by MODIS and Landsat: the FluxnetEO data set, *Biogeosciences*, 19, 2805–2840, <https://doi.org/10.5194/bg-19-2805-2022>, publisher: Copernicus GmbH, 2022.

Yu, X., Orth, R., Reichstein, M., Bahn, M., Klosterhalfen, A., Knohl, A., Koepsch, F., Migliavacca, M., Mund, M., Nelson, J. A., Stocker, B. D., Walther, S., and Bastos, A.: Contrasting drought legacy effects on gross primary productivity in a mixed versus pure beech forest, *Biogeosciences*, 19, 4315–4329, <https://doi.org/10.5194/bg-19-4315-2022>, publisher: Copernicus GmbH, 2022.

Eidesstattliche Versicherung

Hiermit versichere ich an Eides statt, dass ich die vorliegende Arbeit im Studiengang B.Sc. Meteorologie selbstständig verfasst und keine anderen als die angegebenen Hilfsmittel – insbesondere keine im Quellenverzeichnis nicht benannten Internet-Quellen – benutzt habe. Alle Stellen, die wörtlich oder sinngemäß aus Veröffentlichungen entnommen wurden, sind als solche kenntlich gemacht. Ich versichere weiterhin, dass ich die Arbeit vorher nicht in einem anderen Prüfungsverfahren eingereicht habe und dass die gedruckte der elektronischen Fassung entspricht.

Einer Veröffentlichung der vorliegenden Arbeit in der zuständigen Fachbibliothek des Fachbereichs stimme ich zu.

Hamburg, den 26.09.2023 Unterschrift: David Hefesi