



# Short-term forecasting of regional biospheric CO<sub>2</sub> fluxes in Europe using a light-use-efficiency model (VPRM, MPI-BGC version 1.2)

Jinxuan Chen<sup>1</sup>, Christoph Gerbig<sup>1</sup>, Julia Marshall<sup>1</sup>, and Kai Uwe Totsche<sup>2</sup>

<sup>1</sup>Department of Biogeochemical Signals, Max Planck Institute for Biogeochemistry, 07745 Jena, Germany

<sup>2</sup>Department of Hydrogeology, Institute of Geosciences, Friedrich Schiller University Jena, 07743 Jena, Germany

**Correspondence:** Jinxuan Chen (jichen@bgc-jena.mpg.de)

Received: 17 June 2019 – Discussion started: 23 July 2019

Revised: 11 June 2020 – Accepted: 14 July 2020 – Published: 7 September 2020

**Abstract.** Forecasting atmospheric CO<sub>2</sub> concentrations on synoptic timescales (~ days) can benefit the planning of field campaigns by better predicting the location of important gradients. One aspect of this, accurately predicting the day-to-day variation in biospheric fluxes, poses a major challenge. This study aims to investigate the feasibility of using a diagnostic light-use-efficiency model, the Vegetation Photosynthesis Respiration Model (VPRM), to forecast biospheric CO<sub>2</sub> fluxes on the timescale of a few days. As input, the VPRM model requires downward shortwave radiation, 2 m temperature, and enhanced vegetation index (EVI) and land surface water index (LSWI), both of which are calculated from MODIS reflectance measurements. Flux forecasts were performed by extrapolating the model input into the future, i.e., using downward shortwave radiation and temperature from a numerical weather prediction (NWP) model, as well as extrapolating the MODIS indices to calculate future biospheric CO<sub>2</sub> fluxes with VPRM. A hindcast for biospheric CO<sub>2</sub> fluxes in Europe in 2014 has been done and compared to eddy covariance flux measurements to assess the uncertainty from different aspects of the forecasting system. In total the range-normalized mean absolute error (normalized) of the 5 d flux forecast at daily timescales is 7.1 %, while the error for the model itself is 15.9 %. The largest forecast error source comes from the meteorological data, in which error from shortwave radiation contributes slightly more than the error from air temperature. The error contribution from all error sources is similar at each flux observation site and is not significantly dependent on vegetation type.

## 1 Introduction

Human activities have significantly influenced the carbon cycle of the earth system since industrialization, with the accumulation of greenhouse gases (GHG) in the atmosphere leading to radiative forcing and climate change (IPCC, 2014). The carbon exchange between the surface and the atmosphere still remains largely uncertain due to the complexity of processes and a lack of observations (Le Quéré et al., 2009). Therefore, more measurements are needed, especially over emission hotspots and regions lacking observations. Field campaigns to measure greenhouse gases, such as research flights and measurements in remote areas, can fill the observation gap in the troposphere and over regions not covered by existing networks, but they are often time limited. To make the best use of these limited measurements, field campaigns require careful planning. An atmospheric CO<sub>2</sub> forecast on synoptic timescales (~ days) can be helpful in such cases, for it provides an estimate of what signals are expected during the experiment and a physical explanation of the observations.

The research campaign CoMet (Carbon Dioxide and Methane Mission), organized by the Deutsches Zentrum für Luft- und Raumfahrt (DLR), made a series of airborne and ground-based measurements of greenhouse gases in Europe. The campaign took place from 15 May to 12 June 2018, during which four aircraft participated, including the High Altitude and Long Range Research Aircraft (HALO) and three light aircraft. During the campaign, HALO was equipped with an integrated path differential absorption (IPDA) lidar (CHARM-F) (Amediek et al., 2017) and carried out nine flights with a total of 65 flight hours. Continuous online in

situ CO<sub>2</sub>, CO, CH<sub>4</sub> and water vapor measurements were also made on board with the Jena Instrument for Greenhouse gas measurements (JIG), and air samples were collected with the Jena Air Sampler (JAS). The campaign performed measurements over different surfaces from northern Europe to northern Africa to assess and validate the new remote sensing instrument CHARM-F. Special attention was paid to two areas: Berlin (and nearby power plants) and the Upper Silesian basin, which are significant European point sources of CO<sub>2</sub> and CH<sub>4</sub>, respectively. Ground-based and light aircraft measurements were also made in the two regions with the remote sensing instrument Methane airborne MAPper (MAMAP) (Gerilowski et al., 2011) and portable ground-based Fourier transform infrared spectrometers (FTIRs) (Butz et al., 2017).

During the planning of the campaign, a CO<sub>2</sub> and CH<sub>4</sub> forecasting system was developed to support the mission; this paper focuses on the biogenic fluxes for the CO<sub>2</sub> component. The forecast provided 5 d CO<sub>2</sub> forecast fields at a fine spatial resolution (2 km × 2 km) within the observing area and a coarser resolution over the European domain (10 km × 10 km). The forecast product not only is helpful in terms of planning observations, offering meteorology and GHG fields to capture CO<sub>2</sub>/CH<sub>4</sub> plumes but also can provide a priori vertical information for the retrieval of remote sensing observations.

There are several existing models that can simulate atmospheric CO<sub>2</sub> on an appropriate scale, including Eulerian mesoscale models such as WRF-GHG (Beck et al., 2011; Pillai et al., 2016) and CHIMERE (Aulagnier et al., 2010). These models consist of an atmospheric tracer transport model coupled to fluxes representing the source and sink processes of CO<sub>2</sub>. By providing meteorological forecast fields and future fluxes of CO<sub>2</sub> to the model, the forecast CO<sub>2</sub> concentration fields can be obtained. The challenge of CO<sub>2</sub> forecasting comes with the provision of accurate CO<sub>2</sub> flux variations on subdaily timescales. A global atmospheric CO<sub>2</sub> forecast system has been developed as part of the Monitoring of Atmospheric Composition and Climate – Interim Implementation (MACC-II) service (Agusti-Panareda et al., 2014, 2016). These studies have shown that although transport plays a first-order role in synoptic CO<sub>2</sub> variability, the day-to-day variability in net ecosystem exchange (NEE) also plays an important role. Therefore, it is crucial for CO<sub>2</sub> forecasts to capture the day-to-day NEE variability in real-time, instead of using climatological values.

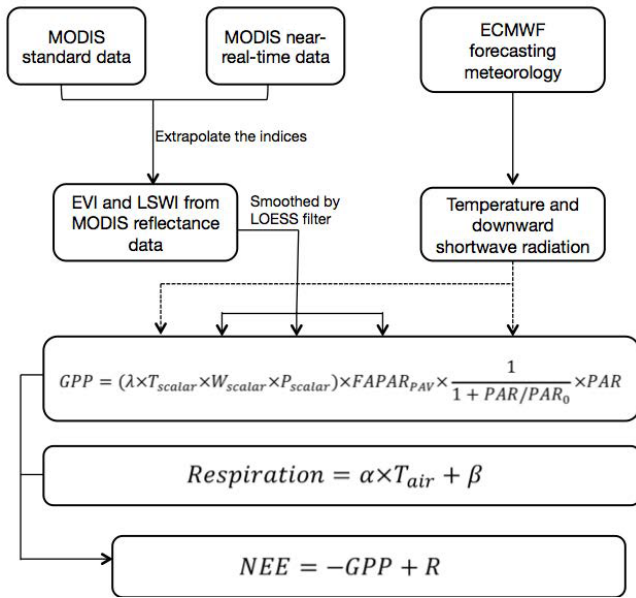
There are many models that can simulate biospheric CO<sub>2</sub> NEE on hourly timescales (Boussetta et al., 2013; Mahadevan et al., 2008). These models can be briefly grouped into two types: process-based models and light-use-efficiency (LUE) models. Process-based models use meteorological data as input and simulate the physiological processes of vegetation, for example BIOME-BGC (Running and Hunt Jr, 1993), TEM (Zhuang et al., 2003) or the Carbon Exchange in the Vegetation-Soil-Atmosphere model (CEVSA) (Woodward et al., 1995). Such models usually need a number of

parameters to describe the complex vegetation processes responding to meteorological drivers. The second type, LUE models, regard ecosystem gross primary production (GPP) as the product of photosynthetically active radiation (PAR), the fraction of photosynthetically active radiation absorbed by the photosynthetically active portion of the vegetation (FAPAR<sub>PAV</sub>) and the radiation use efficiency ( $\epsilon$ ). Such models include the Vegetation Photosynthesis and Respiration Model (VPRM) (Xiao et al., 2004; Mahadevan et al., 2008), the MODIS Daily Photosynthesis Model (Running et al., 2000) and the Carnegie-Ames-Stanford Approach (CASA) (Potter et al., 1993).

The CO<sub>2</sub> forecast in MACC-II uses the process-based model CTESSEL to compute biospheric CO<sub>2</sub> fluxes and evapotranspiration online (Boussetta et al., 2013; Agusti-Panareda et al., 2016), which makes the two variables consistent in the forecast system. However the challenge of providing accurate CO<sub>2</sub> fluxes is due to the complexity of vegetation processes and the lack of near-real-time (NRT) observations on vegetation state. Therefore, using a LUE model for CO<sub>2</sub> flux forecasting, which is a data-driven approach having less parameters compared to process-based models, is a possible way to improve the quality of CO<sub>2</sub> fluxes in forecasting. It should be noted that unlike the Copernicus Atmosphere Monitoring Service (CAMS) CO<sub>2</sub> forecasting, which is operational and global, we target to build a regional CO<sub>2</sub> forecast system and only operate the forecast within a shorter period (e.g., several months). Therefore, the issue of CO<sub>2</sub> budget conservation is less important comparing to an operational global forecast model.

In our case, we predict CO<sub>2</sub> fluxes based on the LUE model VPRM, which is driven by the enhanced vegetation index (EVI) and the land surface water index (LSWI) as well as the meteorological variables 2 m air temperature and downward shortwave radiation. The EVI and LSWI are derived from Moderate Resolution Imaging Spectroradiometer (MODIS) reflectance data, in which the MOD09A1N product provides NRT surface reflectance data, and thus the NRT observations on vegetation state can be used in flux forecasting. VPRM has a strong predictive ability for NEE while maintaining simplicity in having only four parameters for each of the seven vegetation types, which makes it suitable for our case. The flux forecast is then made by predicting the input of VPRM, for which different prediction methods were tested.

The model VPRM is one of the commonly used surface flux models in atmospheric CO<sub>2</sub> simulations and inversions (e.g., Ahmadov et al., 2007; Pillai et al., 2016; Wu et al., 2018). The uncertainty in the flux model is an essential question in inverse modeling (Lasslop et al., 2008), and the uncertainty in 3-hourly, monthly and annually integrated NEE simulated by VPRM has been well assessed by Lin et al. (2011). They established a general framework to attribute error to different sources of uncertainty (driving data, model parameter, and observation and model misrepresentation). In their



**Figure 1.** Diagram of the VPRM forecasting system. The top two levels show the drivers which are predicted into the future, while the bottom three boxes are based on the standard VPRM model (Mahadevan et al., 2008).

work the model’s sensitivity to each source of uncertainty is calculated. With an estimation of errors in each variable (input data, parameter, etc.), one can then attribute the total error to those uncertainty sources by multiplying the error in source with the model’s sensitivity.

Back to this study, our aim is to investigate the feasibility of using such a data-driven model to predict near-future carbon fluxes. Given the uncertainties in meteorological forecasts, the near-real-time MODIS product and all the necessary extrapolations, it is not clear if such a model can still predict realistic carbon fluxes.

This study describes the development and assessment of a biospheric CO<sub>2</sub> flux forecast based on the LUE model VPRM, with the goal of providing accurate hourly 5 d flux forecasts. By using a hindcast and comparing the results to flux tower sites across Europe, the error in the prediction is quantified, and the predictive ability of the CO<sub>2</sub> flux forecasts is assessed.

## 2 Methodology

The CO<sub>2</sub> flux forecast consists of two steps as shown in Fig. 1. Model inputs are first predicted 5 d into the future, then NEE is estimated based on the standard VPRM model, using parameters optimized in previous studies (Kountouris et al., 2018). Each input which must be forecast results in corresponding errors. We systematically evaluate the flux forecasting error associated with each of these predictands.

This section describes the framework of the VPRM forecasting model for biospheric CO<sub>2</sub> fluxes, as well as the method used to evaluate the error introduced by each element of the forecast.

For the meteorological input data, we use hourly ECMWF 5 d forecasts of temperature and shortwave radiation. The EVI and LSWI indices are derived from MODIS surface reflectance data. These provide the indices for an average of the past 8 d, and we forecast these indices for the next 5 d based on linear extrapolation or persistence. We then use these predicted input data to generate NEE using VPRM.

### 2.1 VPRM data processing

#### 2.1.1 Standard processing for past periods

The flux estimation is based on VPRM, a light use efficiency (LUE) model that calculates GPP with remote sensing data and meteorological data as inputs. The equation of GPP estimation is as follows:

$$GPP = \varepsilon \times FAPAR_{PAV} \times \frac{1}{1 + PAR/PAR_0} \times PAR. \quad (1)$$

The light use efficiency  $\varepsilon$  can be decomposed as

$$\varepsilon = \lambda \times T_{scalar} \times W_{scalar} \times P_{scalar}, \quad (2)$$

where  $T_{scalar}$ ,  $W_{scalar}$  and  $P_{scalar}$  represent the temperature sensitivity of photosynthesis, the water stress effect and the effects of leaf age on canopy photosynthesis, respectively, while  $\lambda$  is an adjustable parameter in the model.  $T_{scalar}$  is estimated from air temperature, and  $W_{scalar}$  and  $P_{scalar}$  are estimated from LSWI. See details in Mahadevan et al. (2008).

The  $FAPAR_{PAV}$  in the model is estimated as a linear function of EVI, and PAR is closely correlated with downward shortwave radiation. Therefore, the complete expression for GPP in VPRM is

$$GPP = (\lambda \times T_{scalar} \times W_{scalar} \times P_{scalar}) \times EVI \times \frac{1}{1 + \frac{PAR}{PAR_0}} \times PAR, \quad (3)$$

while the ecosystem respiration ( $R$ ) is estimated by a simple linear model:

$$R = \alpha \times T_{air} + \beta, \quad (4)$$

where  $T_{air}$  is the air temperature, and  $\alpha$  and  $\beta$  are vegetation-class-specific parameters.

The input of VPRM can be categorized into two groups: remote sensing data and meteorological data. The remote sensing data consist of EVI and LSWI at 10 km spatial resolution (the same resolution as the atmospheric transport model), where the EVI and LSWI are aggregated from MODIS surface reflectance 8 d L3 Global 500 m

(MOD09A1) version 6 data. It should be noted that in the forecasting model, the MODIS NRT surface reflectance data (MOD09A1N) would be used. A locally weighted least-squares (LOESS) filter ( $\alpha = 0.17$ ) is then applied to reduce the noise. The vegetation classification map that is used (SYNMAP) (Jung et al., 2006) is also a product originally derived from remote sensing. The meteorological data include air temperature at 2 m and downward shortwave radiation at the surface, which are obtained from a numerical weather prediction (NWP) model product, in our case the operational forecast archive from the European Centre for Medium-Range Weather Forecasts (ECMWF). In VPRM, there are four parameters ( $\lambda$ ,  $\text{PAR}_0$ ,  $\alpha$ ,  $\beta$ ) for each vegetation type. Model calibration for these parameters has been done using flux measurements in Europe in 2007 (Kountouris et al., 2018).

### 2.1.2 Processing for flux prediction

To use this diagnostic model in a predictive mode, we need to forecast all VPRM input variables 5 d into the future. Remote sensing data and meteorological data are predicted in different ways.

For the meteorological data, forecasts from a numerical weather prediction (NWP) model are needed. In this study, in order to assess the errors brought in by the meteorological forecasting, 5 d forecasts of 2 m temperature and downward shortwave radiation at the surface for each day of the year were used. The meteorological forecast is from the ECMWF operational forecast archive, with class “od” and type “fc”.

As for the remote sensing data, three sources of error had to be considered: the error induced by using the NRT version of the MODIS reflectances rather than the final product, the error of estimating the value of the indices into the future and the effect of the LOESS filter on the end value of the dataset.

We begin by describing the LOESS filter. This filter is usually applied to a full year of data, and when smoothing a truncated dataset there is an edge effect, meaning that when new data are added to the time series and the smoothing is repeated, the output at the former edge point will change slightly. In the following section we define the error caused by such an edge effect as “error due to data truncation”.

Following the filtering, the smoothed data are extrapolated 5 d into the future, either by linear extrapolation or by assuming persistence. The optimal extrapolation method was selected after testing the error contribution of each method.

The last error source comes from the difference between MODIS NRT and the standard product. The standard product is processed with the best available ancillary, calibration and geolocation information, while changes have been made in the NRT processing to expedite the data availability (see <https://earthdata.nasa.gov/earth-observation-data/near-real-time/near-real-time-versus-standard-products>, last access: 1 September 2020).

## 2.2 Uncertainty analysis

There are uncertainties in the model, in the forecast data as well as in the eddy covariance measurement, and each of these uncertainties has different impact on the final product of the flux forecast. Therefore, before getting into the error quantification and model evaluation, we will briefly discuss their roles in this study.

The uncertainty in the flux measurement has to be considered before being used as the “truth” in the model–data comparison. The uncertainty in flux measurement from eddy covariance tower and its impact on modeling have been well investigated by previous studies. Hollinger and Richardson (2005) attribute the random error in flux measurement to three reasons: the error associated with measurement system, the error associated with turbulence transport and the statistical error relating to footprint heterogeneity. They establish the method for flux measurement error estimation and analyze it on half-hourly timescale. Chevallier et al. (2012) calculate the flux measurement uncertainty on daily timescale based on hourly uncertainty estimation from Lasslop et al. (2008) and conclude that the daily uncertainty is small comparing to the daily NEE magnitude. A similar approach is used in Broquet et al. (2013), where the uncertainty in daily flux measurement is ignored in observation–model comparison. Therefore, in this study, where all comparisons are concerned with daily timescales, uncertainty from flux measurements can be neglected.

Estimating carbon fluxes with the data-driven model VPRM will result in additional uncertainties. These uncertainties are associated with uncertainties in the driving data, the misrepresentation of the LUE approach for vegetation processes and the spatial representation. We treat these uncertainties as an inherent part of the model, since they will exist despite whatever “good” data we are using to drive the model. We define these uncertainties as the VPRM “model error”, which can be quantified by comparing the flux estimation with best driving data available for VPRM to the flux measurement. This model error, as an inherent error in VPRM, is then chosen as a criterion for the evaluation of the forecasting result.

Lastly the error added by the flux forecasting needs to be considered. As described in Sect. 2.1.2, the flux forecast is made by predicting the driving data. Such a prediction has different impact on different variables, thus introducing different uncertainties. For meteorological data, they are from the Integrated Forecasting System (IFS) model of ECMWF, which will contain model error and representation error as any NWP model (Simmons et al., 1995; Simmons and Hollingsworth, 2002). Furthermore, the model error accumulates in weather forecasting, which means the further we predict into the future, the larger the error will be. As for the MODIS data, the use of NRT data and the extrapolation we apply will surely introduce uncertainties. In addition, VPRM applies LOESS filter in the MODIS data processing

**Table 1.** The selected FLUXNET2015 sites used for data–model comparison in this research.

Site ID	Latitude	Longitude	Vegetation types in VPRM	Data DOI	Reference
BE-Bra	51.3092	4.5206	Mixfrst	<a href="https://doi.org/10.18140/FLX/1440128">https://doi.org/10.18140/FLX/1440128</a>	Janssens et al. (2016)
BE-Lon	50.5516	4.7461	Crop	<a href="https://doi.org/10.18140/FLX/1440129">https://doi.org/10.18140/FLX/1440129</a>	Moureaux et al. (2006)
BE-Vie	50.3051	5.9981	Mixfrst	<a href="https://doi.org/10.18140/FLX/1440130">https://doi.org/10.18140/FLX/1440130</a>	Aubinet et al. (2001)
CH-Cha	47.2102	8.4104	Grass	<a href="https://doi.org/10.18140/FLX/1440131">https://doi.org/10.18140/FLX/1440131</a>	Merbold et al. (2014)
CH-Dav	46.8153	9.8559	Evergreen	<a href="https://doi.org/10.18140/FLX/1440132">https://doi.org/10.18140/FLX/1440132</a>	Zielis et al. (2014)
CH-Fru	47.1158	8.5378	Grass	<a href="https://doi.org/10.18140/FLX/1440133">https://doi.org/10.18140/FLX/1440133</a>	Imer et al. (2013)
CH-Lae	47.4781	8.365	Mixfrst	<a href="https://doi.org/10.18140/FLX/1440134">https://doi.org/10.18140/FLX/1440134</a>	Etzold et al. (2011)
CH-Oe2	47.2863	7.7343	Crop	<a href="https://doi.org/10.18140/FLX/1440136">https://doi.org/10.18140/FLX/1440136</a>	Dietiker et al. (2010)
CZ-wet	49.0247	14.7704	Grass	<a href="https://doi.org/10.18140/FLX/1440145">https://doi.org/10.18140/FLX/1440145</a>	Dušek et al. (2012)
DE-Akm	53.8662	13.6834	Grass	<a href="https://doi.org/10.18140/FLX/1440213">https://doi.org/10.18140/FLX/1440213</a>	Bernhofer et al. (2009–2014)
DE-Geb	51.1001	10.9143	Crop	<a href="https://doi.org/10.18140/FLX/1440146">https://doi.org/10.18140/FLX/1440146</a>	Anthoni et al. (2004)
DE-Gri	50.9495	13.5125	Grass	<a href="https://doi.org/10.18140/FLX/1440147">https://doi.org/10.18140/FLX/1440147</a>	Prescher et al. (2010)
DE-Kli	50.8929	13.5225	Crop	<a href="https://doi.org/10.18140/FLX/1440149">https://doi.org/10.18140/FLX/1440149</a>	Prescher et al. (2010)
DE-Obe	50.7836	13.7196	Evergreen	<a href="https://doi.org/10.18140/FLX/1440151">https://doi.org/10.18140/FLX/1440151</a>	Bernhofer et al. (2008–2014)
DE-RuR	50.6219	6.3041	Grass	<a href="https://doi.org/10.18140/FLX/1440215">https://doi.org/10.18140/FLX/1440215</a>	Post et al. (2015)
DE-RuS	50.8659	6.4472	Crop	<a href="https://doi.org/10.18140/FLX/1440216">https://doi.org/10.18140/FLX/1440216</a>	Mauder et al. (2013)
DE-SfN	47.8064	11.3275	Grass	<a href="https://doi.org/10.18140/FLX/1440219">https://doi.org/10.18140/FLX/1440219</a>	Hommeltenberg et al. (2014)
DE-Spw	51.8923	14.0337	Grass	<a href="https://doi.org/10.18140/FLX/1440220">https://doi.org/10.18140/FLX/1440220</a>	Bernhofer et al. (2010–2014)
DE-Tha	50.9636	13.5669	Evergreen	<a href="https://doi.org/10.18140/FLX/1440152">https://doi.org/10.18140/FLX/1440152</a>	Grünwald and Bernhofer (2007)
DK-Sor	55.4859	11.6446	Decid	<a href="https://doi.org/10.18140/FLX/1440155">https://doi.org/10.18140/FLX/1440155</a>	Pilegaard et al. (2011)
FI-Hyy	61.8475	24.295	Evergreen	<a href="https://doi.org/10.18140/FLX/1440158">https://doi.org/10.18140/FLX/1440158</a>	Suni et al. (2003)
FI-Sod	67.3619	26.6378	Evergreen	<a href="https://doi.org/10.18140/FLX/1440160">https://doi.org/10.18140/FLX/1440160</a>	Thum et al. (2007)
FR-Fon	48.4764	2.7801	Decid	<a href="https://doi.org/10.18140/FLX/1440161">https://doi.org/10.18140/FLX/1440161</a>	Delpierre et al. (2016)
FR-Pue	43.7414	3.5958	Evergreen	<a href="https://doi.org/10.18140/FLX/1440164">https://doi.org/10.18140/FLX/1440164</a>	Rambal et al. (2004)
IT-BCi	40.5238	14.9574	Crop	<a href="https://doi.org/10.18140/FLX/1440166">https://doi.org/10.18140/FLX/1440166</a>	Vitale et al. (2016)
IT-CA1	42.3804	12.0266	Decid	<a href="https://doi.org/10.18140/FLX/1440230">https://doi.org/10.18140/FLX/1440230</a>	Sabbatini et al. (2016)
IT-CA2	42.3772	12.026	Crop	<a href="https://doi.org/10.18140/FLX/1440231">https://doi.org/10.18140/FLX/1440231</a>	Sabbatini et al. (2016)
IT-CA3	42.38	12.0222	Decid	<a href="https://doi.org/10.18140/FLX/1440232">https://doi.org/10.18140/FLX/1440232</a>	Sabbatini et al. (2016)
IT-Col	41.8494	13.5881	Decid	<a href="https://doi.org/10.18140/FLX/1440167">https://doi.org/10.18140/FLX/1440167</a>	Valentini et al. (1996)
IT-Cp2	41.7043	12.3573	Evergreen	<a href="https://doi.org/10.18140/FLX/1440233">https://doi.org/10.18140/FLX/1440233</a>	Fares et al. (2014)
IT-Isp	45.8126	8.6336	Decid	<a href="https://doi.org/10.18140/FLX/1440234">https://doi.org/10.18140/FLX/1440234</a>	Ferréa et al. (2012)
IT-Lav	45.9562	11.2813	Evergreen	<a href="https://doi.org/10.18140/FLX/1440169">https://doi.org/10.18140/FLX/1440169</a>	Marcolla et al. (2003)
IT-Tor	45.8444	7.5781	Grass	<a href="https://doi.org/10.18140/FLX/1440237">https://doi.org/10.18140/FLX/1440237</a>	Galvagno et al. (2013)

to reduce noise, which means the data are constrained by the neighboring information. However, when forecasting, the data can only be constrained by the past, leading to another potential error source.

Altogether, the potential error sources of this flux forecasting system are as follows: (1) the VPRM model error (2) using meteorological model data rather than site-level meteorological data; (3) using ECMWF 5 d forecast meteorology, which accumulates extra error to its initial field; (4) using NRT MODIS data; (5) using LOESS filtering to smooth the MODIS data; and (6) the prediction of MODIS data. Error (6) contains two parts: (6a) EVI prediction and (6b) LSWI prediction. In the following discussion we use the numbering (1) to (6) to denote these error sources. The model error (1) defined above is regarded as a criterion for the forecast evaluation. We define (2) to (6) as the “forecast errors”, since they are introduced by the flux forecasting. In this study, we aim to quantify the forecast error and the error contribution

from each of the error sources, then evaluate the sum of forecast errors against the model error.

In order to quantify both the model error and the forecast error, a hindcast using the CO<sub>2</sub> flux forecast model has been done for the year 2014 for Europe. The evaluation and comparison were done at two spatial levels: at the flux observation site level and at the European domain level (1/8° longitude × 1/12° latitude). The comparison at site level aims to evaluate both the model error and the forecast error at locations with different vegetation types, while over the European domain, the aim is to investigate the spatial pattern of each forecast error term.

The surface CO<sub>2</sub> flux observation data come from eddy covariance tower measurements from the FLUXNET2015 tier one (open data) dataset (Baldocchi et al., 2001). Thirty-three European observation sites for which both MODIS data and flux measurements for 2014 are available were selected

**Table 2.** The experiment setup and the error sources addressed in each simulation. The numbering in the last column corresponds to the error from (1) the VPRM model, (2) the meteorological analysis, (3) the meteorological forecast, (4) the MODIS NRT data, (5) data truncation and (6) the prediction of MODIS indices.

	MODIS indices	Meteorology data	Error sources
Reference simulation	Standard MODIS products	Flux site observation	(1)
Simulation a	Standard MODIS products	ECMWF 12 h forecasting	(1) + (2)
Simulation b	Standard MODIS products	ECMWF fifth day forecasting	(1) + (2) + (3)
Simulation c	Truncated MODIS indices	Flux site observation	(1) + (5)
Simulation d	MODIS prediction based on fully filtered data	Flux site observation	(1) + (6)
Simulation e	NRT MODIS indices	Flux site observation	(1) + (4)
Simulation f	MODIS prediction based on truncated data	ECMWF fifth day forecasting	(1) + (2) + (3) + (5) + (6)

for data–model comparison. The selected sites' ID, location, vegetation type and data DOI are listed in Table 1.

To test the error contribution of the model and the 5 d flux forecast, experiments using the VPRM forecast model were carried out to evaluate the error contribution from different sources separately, as shown in Table 2. A control simulation and six experimental simulations (simulations a to f) were conducted. Although the CO<sub>2</sub> flux forecast targets hourly flux prediction for the next 5 d, model error and forecast error were analyzed on a daily timescale, as this scale is more relevant for synoptic CO<sub>2</sub> variability in the atmosphere.

The control simulation uses standard VPRM as a reference model with “perfect” input, meaning the MODIS EVI and LSWI standard products as well as shortwave radiation and temperature observed at the flux site. By comparing the modeled NEE to flux measurements, we can estimate the VPRM model error (1).

The experimental simulations (a) to (f) then included the error sources (2) to (6) in the VPRM model input data separately, and these are compared to the reference simulation in order to isolate the individual error contributions. The experiments aim to estimate the upper limit of forecast error, and therefore in simulations (b) and (f), 96 to 120 h meteorological forecasts, i.e., the last day (fifth) of a 5 d forecast, were used for each day of the year. For simulations (d) and (f), since the MODIS EVI and LSWI products have an 8 d period, MODIS data were first linearly interpolated to a daily scale. Then for each day of the year, MODIS data on the  $n$ th day were predicted from data on the  $n - 5$ th day.

There is a challenge in simulation (e) in that there are no archived NRT data for 2014, and thus it is impossible to have a comparison on the same basis as the other simulations. Instead we look at the model's sensitivity of NEE to EVI and LSWI bias, and we also compare the NRT EVI and LSWI, which we archived from February to June in 2018 for 120 d, to the standard MODIS product over the same period. In this way we were able to estimate the magnitude of the NRT indices' error and its impact on the model's output NEE.

In order to make the 33 different site results comparable, the simulation output NEE was first aggregated to daily averages and then normalized by the range (i.e., the difference

between maximum and minimum) of annual NEE at each site. The Bias<sub>NEE</sub>, which is defined as the output NEE from the experimental simulation minus the same variable from the reference model, was then calculated and normalized by the same scalar at each site. By applying such a normalization, positive and negative NEE keep their sign, and the normalized Bias<sub>NEE</sub> represents a fractional bias compared to the range of annual variation. (For example a normalized Bias<sub>NEE</sub> of 0.1 means that the magnitude of the bias equals 10 % of the annual variation.) Similar to Bias<sub>NEE</sub>, Bias<sub>-GPP</sub> and Bias<sub>R</sub> are also calculated as a measure for error in simulated GPP and  $R$ . Bias<sub>-GPP</sub> (or Bias<sub>R</sub>) is the difference of  $-GPP$  (or  $R$ ) in experimental and reference simulation, normalized by the annual range of observed NEE at each site (note that the sign of GPP is reversed). Bias<sub>-GPP</sub> and Bias<sub>R</sub> use the same normalization scalar so that they are additive and comparable to Bias<sub>NEE</sub>. Based on these definitions, we have the following:

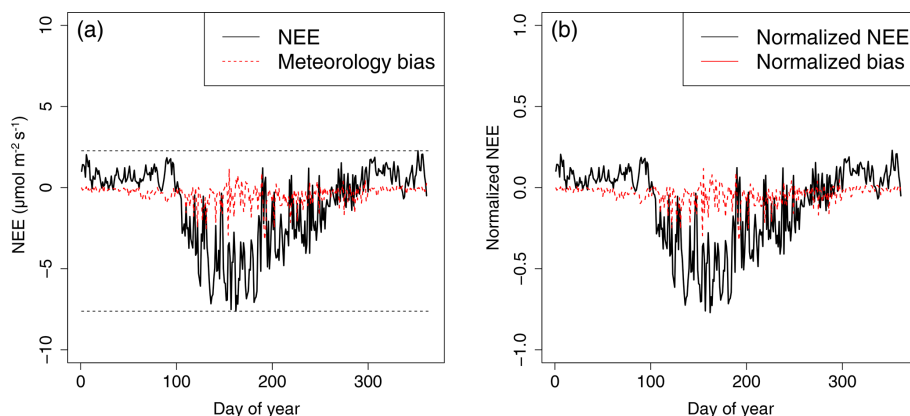
$$\text{Bias}_{\text{NEE}} = \text{Bias}_{-\text{GPP}} + \text{Bias}_R. \quad (5)$$

Thus the metrics Bias<sub>-GPP</sub> and Bias<sub>R</sub> represent the fractional bias of photosynthetic and nonphotosynthetic part in NEE. The mean of the absolute Bias<sub>NEE</sub> will be the mean absolute error (MAE), which is also used as a measure for error in this research. An example of such a normalization is shown for the station BE-Bra in Fig. 2.

### 3 Results and discussion

#### 3.1 Error attribution on site level

By comparing the NEE output from each experimental simulation, the impact of each error source on flux forecasting can be isolated and evaluated. The normalized mean absolute error (MAE) of NEE at all 33 sites is presented in Table 3. The MAE of the total forecast error is 0.071, which is smaller than the VPRM model error of 0.159. This indicates that the forecast model is reasonably capable of predicting fluxes on diurnal timescales.



**Figure 2.** Example of the data normalization at station BE-Bra: **(a)** NEE output from simulation (a) and the corresponding Bias<sub>NEE</sub>. The dashed black lines show the range of annual NEE. **(b)** NEE and bias after normalization by the range, conserving the physical meaning (release and uptake) of the sign.

**Table 3.** Normalized mean absolute error (MAE) of NEE for each error source. The compared objects are simulations (a) to (f), the reference simulation (ref.) and FLUXNET observation (obs.). Error sources (1) to (6) described in Sect. 2.2 can be isolated by calculating the MAE between different simulations.

Normalized mean absolute error (MAE) for each error source		
Compared objects	Error sources	MAE
a–ref.	(2) Meteorological analysis	0.046
b–a	(3) Meteorological forecast	0.040
b–ref.	(2) + (3) meteorological error	0.065
c–ref.	(5) Data truncation	0.015
d–ref.	(6ai) Linear EVI	0.016
d–ref.	(6aii) Persistence EVI	0.013
d–ref.	(6bi) Linear LSWI	0.012
d–ref.	(6bii) Persistence LSWI	0.010
f–ref.	(2) + (3) + (5) + (6aii) + (6bii) forecast error	0.071
ref.–obs.	(1) Model error	0.159

### 3.1.1 Meteorological error

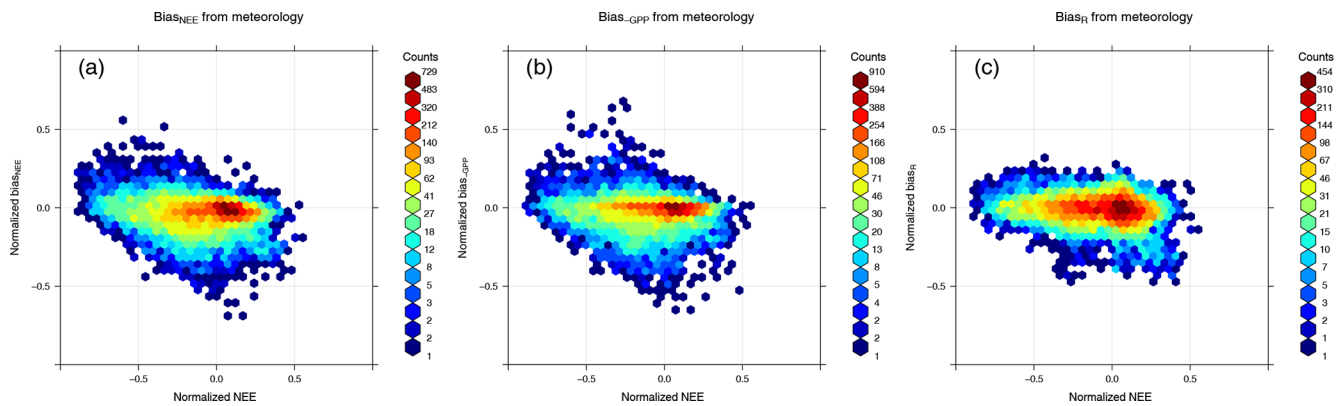
Among all forecast errors, the meteorological error accounts for the largest contribution. The meteorological error can be decomposed into (2) analysis error and (3) meteorological forecast error. The former corresponds to using meteorological analysis rather than observational data, while the latter comes from the numerical meteorological forecasting and can be estimated by comparing simulations (b) and (a). The analysis error and meteorological forecast error are of the same order of magnitude, namely 0.046 and 0.065, respectively.

The meteorological error is then analyzed further by dividing it into the photosynthetic part (Bias<sub>GPP</sub>) and the nonphotosynthetic respiration part (Bias<sub>R</sub>) as described in Sect. 2.2. The bias distributions of 33 × 365 data points are shown in Fig. 3.

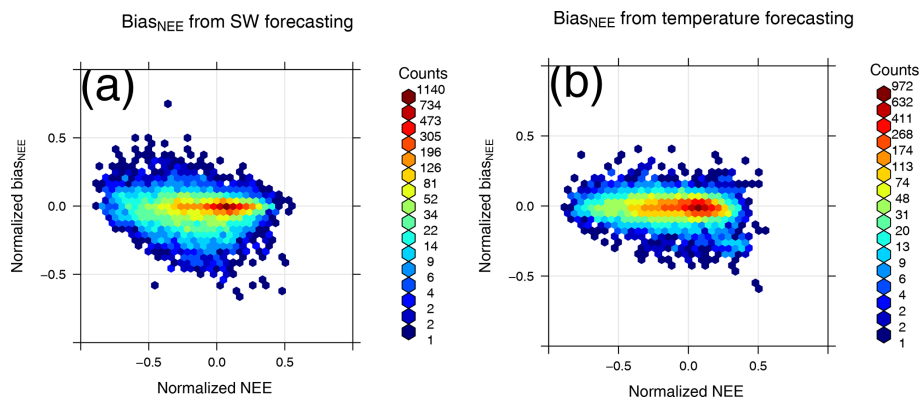
Figure 3a–c share the same *x* axis, and the bias in the *y* axes can be combined as Bias<sub>NEE</sub> = Bias<sub>GPP</sub> + Bias<sub>R</sub>.

Because a positive GPP bias will lead to a negative NEE bias, −GPP is used here to show its contribution to NEE. Bias<sub>GPP</sub> has a larger vertical spread towards negative values, which means a systematic bias in GPP. In contrast Bias<sub>R</sub> is basically symmetric about zero, which implies that the errors in temperature are random.

This seems to suggest that Bias<sub>NEE</sub> has a larger contribution from the photosynthetic part than the nonphotosynthetic part. Knowing that Bias<sub>NEE</sub> is the result of biases in the two meteorological variables used in the simulation, air temperature and downward shortwave radiation (SW), we conduct two further experiments, (b1) and (b2), to quantify the error contribution from these variables separately. In (b1) only the shortwave radiation is taken from the 5 d forecast, while all other variables are from the control simulation. (b2) is similar to (b1), but in this case the forecast value is used only for the 2 m air temperature. Figure 4 shows the bias distribution of the two experiments, in which the vertical spread of bias in (b1) (panel a) is slightly larger than (b2) (panel b). The



**Figure 3.** (a) Distribution of normalized Bias<sub>NEE</sub> due to meteorological error. The x axis refers to the normalized NEE, and the y axis refers to the corresponding Bias<sub>NEE</sub> defined in Sect. 2.2. Panels (b) and (c) share the same x axis with (a) but have Bias<sub>GPP</sub> and Bias<sub>R</sub> in y axis instead. The three biases combine as Bias<sub>NEE</sub> = Bias<sub>GPP</sub> + Bias<sub>R</sub>, suggesting larger contribution from photosynthetic part Bias<sub>GPP</sub>, which is controlled by the radiation parameter rather than temperature.



**Figure 4.** The Bias<sub>NEE</sub> distribution of experiment (b1) (a) and (b2) (b). In experiment (b1) only SW is from 5 d forecast, while other variables are the same with the reference simulation; in experiment (b2), it is air temperature that only comes from 5 d forecast. The MAEs to the reference experiment are 0.053 and 0.042, respectively.

overall normalized MAE compared to the control simulation is 0.053 when using forecast SW (b1), while it is 0.042 when using forecast 2 m air temperature (b2). Thus the error contribution resulting from forecast errors in downward shortwave radiation at the surface is found to be slightly larger than the error from 2 m air temperature.

### 3.1.2 MODIS error

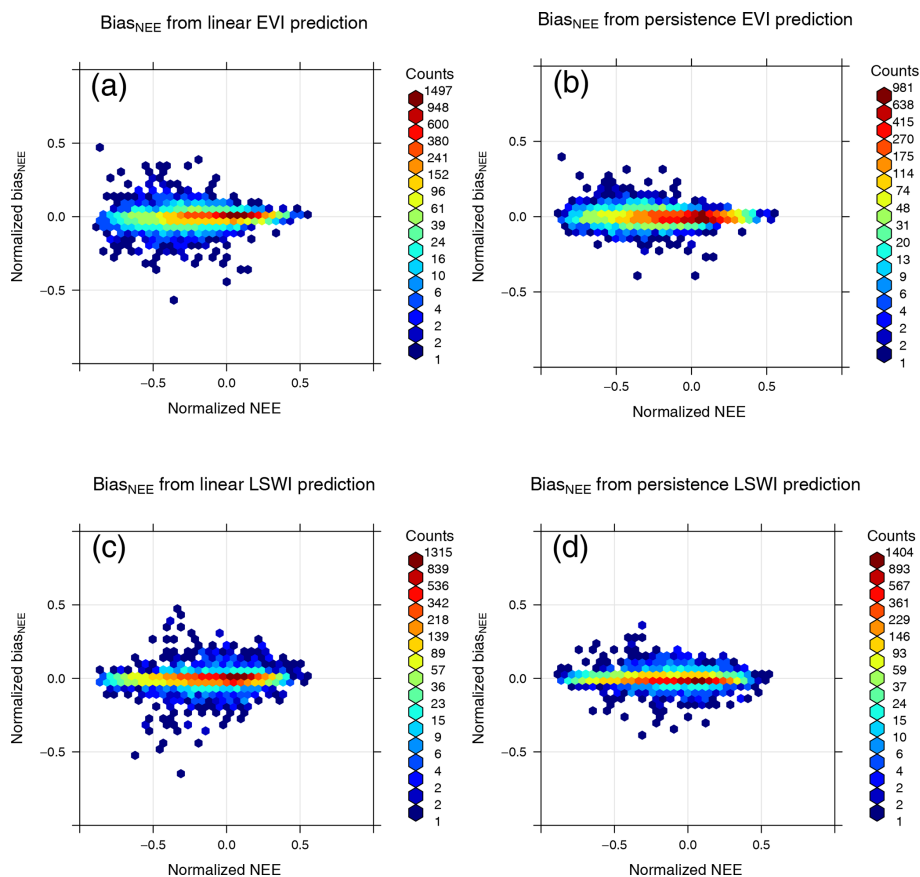
The MODIS error consists of three parts: using NRT products, using extrapolated indices and using truncated time series. These are tested in simulations (c), (d) and (e), respectively. In general, the MODIS error is less important than the meteorological error, and the errors due to data truncation, EVI extrapolation and LSWI extrapolation result in errors of similar magnitude: 0.015, 0.013 and 0.010, respectively.

As described in Sect. 2.1.1, the MODIS input data first need to be smoothed by a LOESS filter to reduce the noise. LOESS performs a local regression on the time series. Be-

cause the point at the end of the time series lacks a constraint from future data, it results in an error when the data are truncated. This error source is evaluated in simulation (c), where for each 8 d value, only data before this time are filtered. Thus the only difference between simulation (c) and the reference simulation is whether each MODIS-derived index is constrained by all local data or only constrained by preceding data. Comparing simulation (c) and the reference simulation finds that the error due to lack of constraint from future MODIS data introduces a MAE of 0.015.

For MODIS data extrapolation, different methods were tested in an attempt to minimize forecast error. Climatological values of EVI and LSWI were considered, but they lack the advantage of a data-driven approach for realistic estimation. After testing various alternatives, two simple methods were considered: linear extrapolation based on the last three data points and persistence (assuming the indices stay the same for the next 5 d). Figure 5 shows the NEE bias distribution by using linear extrapolation or persistence to predict





**Figure 5.** Bias<sub>NEE</sub> distribution of using linear extrapolation or persistence to predict EVI and LSWI. The persistence prediction introduces less bias than linear extrapolation for both EVI and LSWI. Therefore, persistence is used in the final forecast.

EVI and LSWI. For both indices, using the assumption of persistence results in a smaller error. The biases for the two extrapolation methods have similar distributions, but there are more outliers for linear extrapolation. This is due to the fact that linear extrapolation results in larger errors when the data are fluctuating.

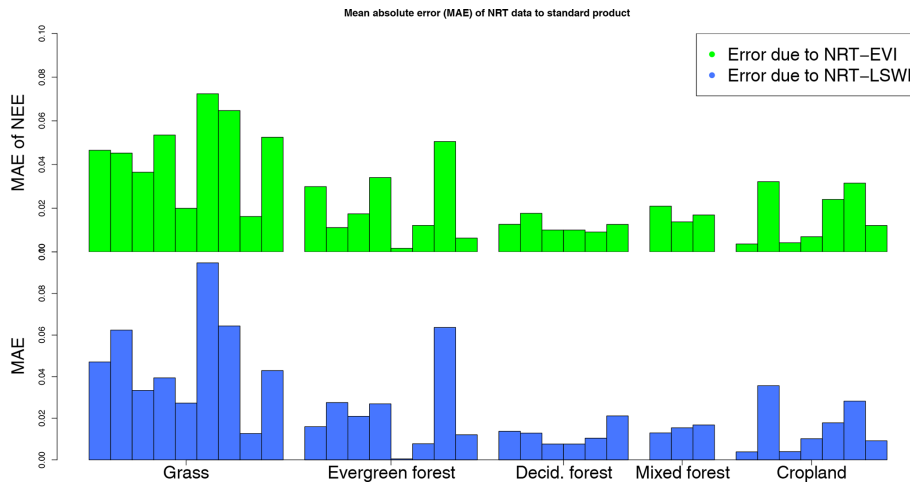
Finally, the difference between using MODIS NRT data and standard data has to be considered. This includes the effect of using different attitude and ephemeris data in processing, as well as using different ancillary data products for the Level 2 processing. For L2 Land Surface Reflectance data, National Oceanic and Atmospheric Administration Global Forecast System (GFS) ancillary product are used instead of the Global Data Assimilation System (GDAS) products used in the standard processing – this is described at NASA’s Land, Atmosphere Near real-time Capability for EOS (LANCE) website <https://earthdata.nasa.gov/earth-observation-data/near-real-time/near-real-time-versus-standard-products>, last access: 1 September 2020).

This presented a challenge, as no MODIS NRT data were archived for the test year 2014. Thus it was impossible to carry out a similar error evaluation as was done for other er-

ror sources. Therefore, we first use NRT EVI and LSWI that we archived for 120 d from February to June 2018 to calculate the MAE of the two indices to standard products at all flux sites. The MAE of NRT EVI and LSWI for all sites are 0.018 and 0.026, respectively. Considering the mean EVI and LSWI, which are 0.21 and 0.11 during this period, the magnitude of NRT EVI error is less than 10 % of EVI’s magnitude, while the number is 24 % for the magnitude of NRT LSWI error.

The impact of the errors in these NRT indices on the model is determined by the model’s sensitivity to EVI and LSWI. To investigate this sensitivity, we use the result from simulation (d) and the reference simulation, and we look at the difference in input EVI and LSWI and the corresponding difference in output NEE. The model’s sensitivity is different during the growing and the non-growing seasons, as in the non-growing season there would be no vegetation production anyway from a slight change of EVI and LSWI.

Therefore, the model sensitivity is analyzed for each season separately, as shown in Table 4. Difference in indices and the corresponding difference in daily NEE are applied with linear regression, and the rate of the linear function is regarded as model sensitivity. The maximum sensitiv-



**Figure 6.** The normalized error of NEE as a result of MODIS NRT error at 33 sites. Overall, 120 d from February to June in the year 2018 of MODIS NRT data are used to first calculate the EVI/LSWI differences, then multiplied by the sensitivities in Table 4 and normalized by the same scalar in the previous research. The flux sites in *x* axis are sorted by vegetation type and FLUXNET site ID (from left to right: CH-Cha, CH-Fru, CZ-wet, DE-Akm, DE-Gri, DE-RuR, DE-Sfn, DE-Spw, IT-Tor, CH-Dav, DE-Obe, DE-Tha, FI-Hyy, FI-Sod, FR-Pue, IT-Cp2, IT-Lav, DK-Sor, FR-Fon, IT-CA1, IT-CA3, IT-Col, IT-Isp, BE-Bra, BE-Vie, CH-Lae, BE-Lon, CH-Oe2, DE-Geb, DE-Kli, DE-Rus, IT-BCi, IT-CA2).

**Table 4.** The model's sensitivity of NEE to EVI/LSWI for four seasons. The result of simulation (d) is used in the sensitivity calculation. Linear regression is applied to the change in EVI and the change in corresponding NEE, the maximum sensitivity appears in summer, with a slope of  $-10.73$  ( $\mu\text{mol m}^{-2} \text{s}^{-1} \text{EVI}^{-1}$ ) for EVI and  $-6.29$  ( $\mu\text{mol m}^{-2} \text{s}^{-1} \text{LSWI}^{-1}$ ) for LSWI, respectively.

NEE sensitivity to EVI			NEE sensitivity to LSWI		
Seasons	Sensitivity ( $\mu\text{mol m}^{-2}$ $\text{s}^{-1} \text{EVI}^{-1}$ )	$R^2$	Seasons	Sensitivity ( $\mu\text{mol m}^{-2}$ $\text{s}^{-1} \text{LSWI}^{-1}$ )	$R^2$
Dec–Feb	−0.90	0.27	Dec–Feb	−0.57	0.28
Mar–May	−7.96	0.64	Mar–May	−3.41	0.51
Jun–Aug	−9.11	0.74	Jun–Aug	−6.29	0.58
Sep–Jan	−2.70	0.35	Sep–Jan	−1.16	0.29

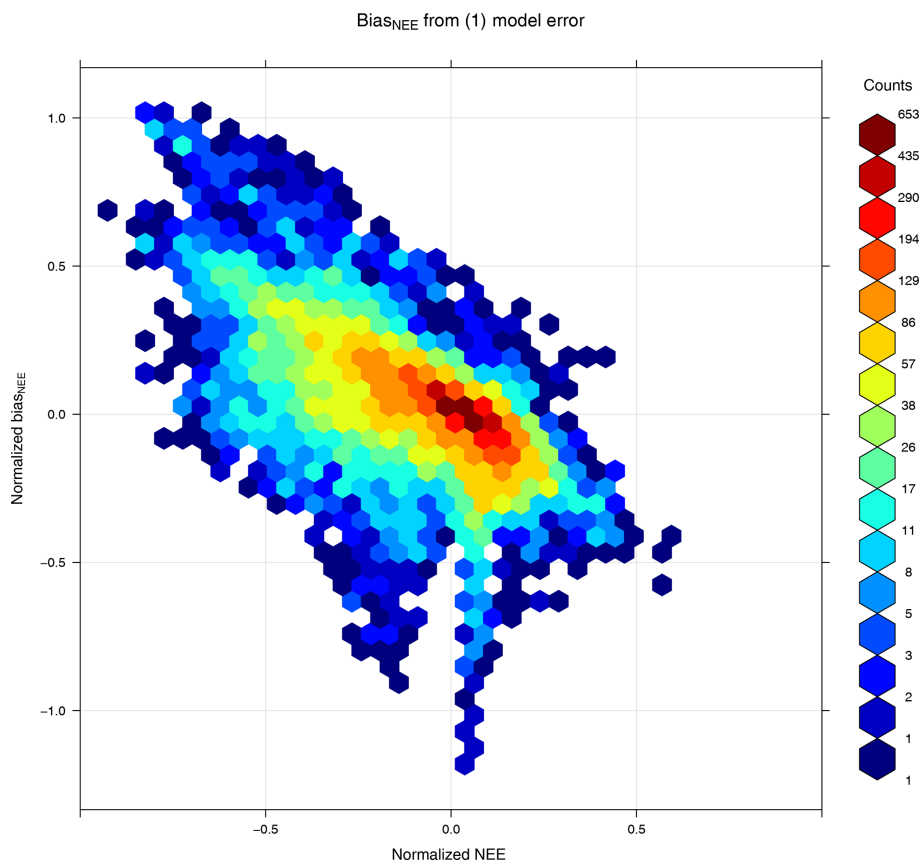
ity for both EVI and LSWI is in summer, with  $-9.11$  ( $\mu\text{mol m}^{-2} \text{s}^{-1} \text{EVI}^{-1}$ ) and  $-6.29$  ( $\mu\text{mol m}^{-2} \text{s}^{-1} \text{LSWI}^{-1}$ ), respectively. By assuming that the 120 d of archived NRT data is representative for MODIS NRT error, we can estimate the upper limit of forecasting error (4), as it is shown in Fig. 6. The normalized NEE error in Fig. 6 is calculated by using MODIS NRT error times the model sensitivity and then normalized by the same scalar used in previous analysis at each site. Therefore, the error here is comparable to the MAE in Table 3 if we assume the MODIS NRT data in the year 2014 and 2018 have similar error structure. The NEE error for all sites due to NRT-EVI and NRT-LSWI is 0.024 and 0.025, respectively, which is still small compared to the meteorology error in Table 3.

### 3.1.3 VPRM model error

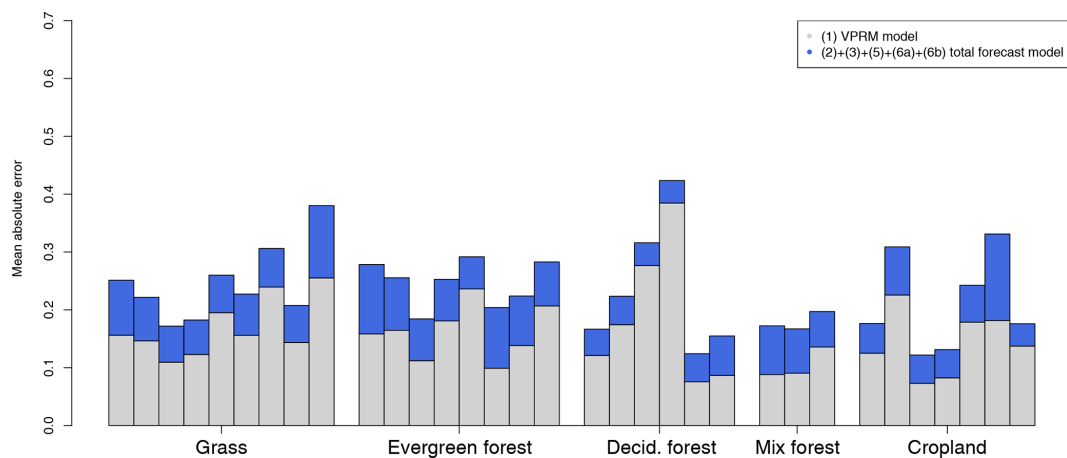
Unlike the forecast error discussed above, the  $\text{Bias}_{\text{NEE}}$  of (1) model error (reference model minus observation) distribution of the VPRM model error is asymmetric, as shown in Fig. 7. The model bias shows a negative correlation, which means a weaker uptake during the growing season and a weaker respiration during the non-growing season. Data with negative normalized NEE also correspond to a larger bias, which refers to larger model uncertainty during the growing season. The MAE of the model error is 0.166.

### 3.1.4 Errors at each flux observation site

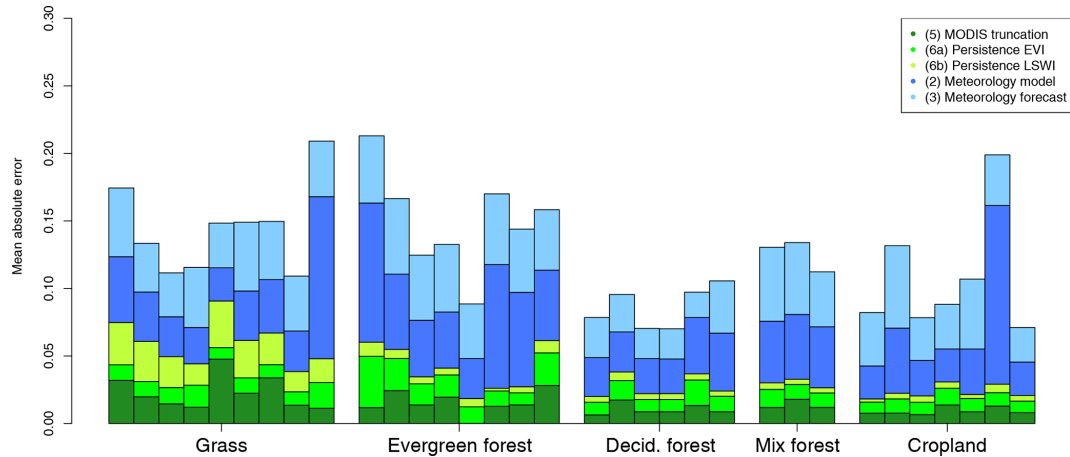
The MAE is also calculated at each flux measurement site and clustered according to vegetation types, as shown in Fig. 8. Generally the VPRM model error (grey) is larger or similar to the forecast error (blue), consistent with Table 3.



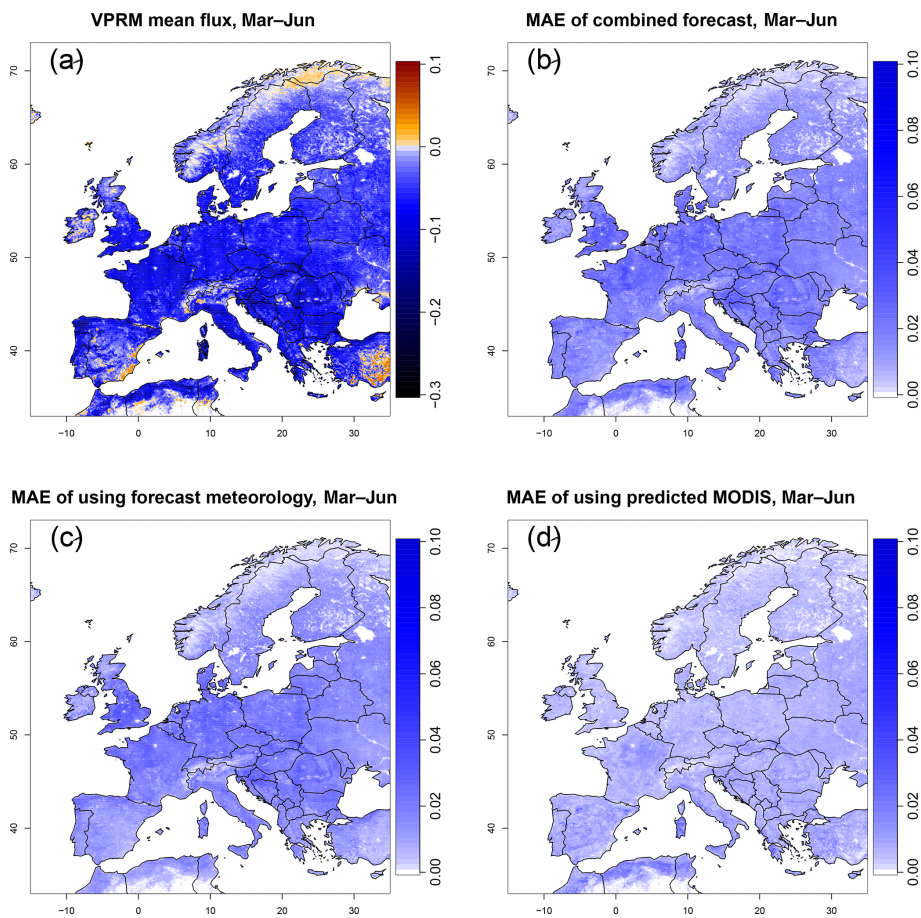
**Figure 7.** The Bias<sub>NEE</sub> distribution of the VPRM model error.



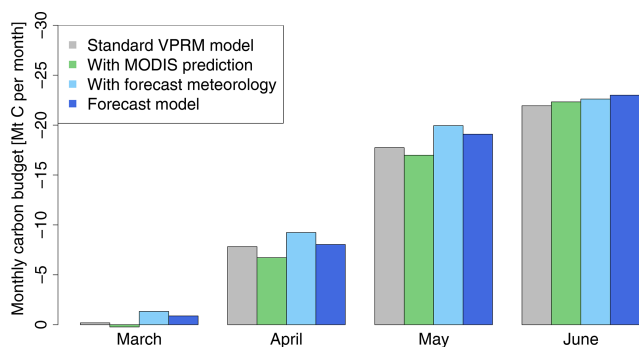
**Figure 8.** Mean absolute error of the forecast error compared to the VPRM model error at each flux observation site. The model error (1) is generally larger than the total forecast error (2) to (6), and the forecast error does not differ significantly across vegetation types. The order of the flux site is the same as in Fig. 6.



**Figure 9.** Mean absolute error for different error sources at each flux observation site. The meteorological error (2, meteorological model, + 3, meteorological forecast) is the dominant contributor at each site and has a similar contribution for different vegetation types. The data truncation error (4) has a stronger influence on some grass sites, likely due to the highly EVI variability resulting from mowing and regrowth during the growing season. The order of the flux site is the same as in Fig. 6.



**Figure 10.** (a) Mean VPRM NEE, during March to June 2014. (b) Spatial distribution of MAE for forecast error; (c) spatial distribution of MAE for meteorological error; (d) spatial distribution of MAE for MODIS error. The MAE of total forecast error in (b) has strong spatial relationship with the VPRM mean flux in (a), which indicates that the forecast error has a similar impact in all places. Panels (c) and (d) are consistent with Table 3, in that the forecast error is larger than the error from MODIS prediction.



**Figure 11.** Monthly carbon budget from March to June for original and forecast model for the European domain. The overall forecast flux budget is close to the original model, indicating the forecast flux model is appropriate for use in the GHG concentration forecasting system.

Moreover the forecast error does not differ significantly over different vegetation types. Figure 9 shows the error contribution from each source; the meteorological error (error 2 in dark blue and error 3 in light blue) is the dominant contributor at each site and has a similar contribution for different vegetation types. The data truncation error (4) has a stronger influence on some grass sites, because EVI at these sites is highly variable, possibly due to mowing and regrowing during the growing season. Overall, except for the data truncation error, all forecast error sources have a similar impact on each flux observation site. This shows that the forecast ability does not vary over different vegetation types.

### 3.2 Spatial pattern of forecast error

The forecast errors are also tested on the European domain from March to June (the season over which the CoMet campaign took place) in 2014, to analyze its spatial patterns. Three experiments have been done to represent the meteorological error (including analysis error and meteorological forecast error), the MODIS error (including extrapolation error and data truncation error) and the total forecast error (a combination of meteorological error and MODIS error). Figure 10 shows the mean VPRM NEE during the period and the corresponding spatial distribution of each error (in MAE).

By comparing Fig. 10a and b, it can be seen that the MAE of the total forecast error has a strong spatial relationship with the mean NEE, which indicates that the forecast error has a similar impact in all places. On a spatial level, the meteorological component still dominates compared to the MODIS error.

In the context of atmospheric CO<sub>2</sub> forecasting, the forecast CO<sub>2</sub> concentrations that are influenced by fluxes from larger-MAE areas (northern France, Germany and the Balkans) may have a larger bias due to poorer flux prediction in these areas.

The flux budget over the European domain was also calculated and is shown in Fig. 11. The carbon budget of the flux

forecast model (in dark blue) is close to the original VPRM model (in grey); thus we are able to confidently use this flux forecast model in the atmospheric GHG concentration forecasting system and predict reasonable CO<sub>2</sub> concentrations on synoptic timescales.

As mentioned in the introduction, we are aiming for not only a flux forecast but finally an atmospheric GHG concentration forecasting system. While this study has quantified how each error source affects the predicted biospheric fluxes, the next step is to use such predicted fluxes in an atmospheric transport model run in forecast mode and to assess the prediction error from each source in concentration space.

## 4 Conclusions

Based on the VPRM model, we developed a forecasting model that can predict biospheric NEE for the next 5 d and assess the error contribution from each aspect of forecasting. This CO<sub>2</sub> flux forecast model is a crucial component in an atmospheric CO<sub>2</sub> forecasting system, in which hourly to day-to-day CO<sub>2</sub> flux variability plays an important role. The forecast model inputs are MODIS near-real-time EVI and LSWI, as well as shortwave radiation and temperature from a meteorological forecast model. The error attribution shows that the dominant error is related to the meteorological data. We further attribute this error to the uncertainties in forecast shortwave radiation and temperature, and we found that the forecast shortwave radiation contributes slightly more to the meteorological error. Error from MODIS inputs is less important, and using a persistence assumption to predict MODIS indices resulted in smaller errors than a linear extrapolation. Overall the forecasting system error has a MAE of 0.071, which makes the model capable of forecasting CO<sub>2</sub> fluxes on the target timescale. The error contribution is insensitive to vegetation type and consistent over the whole EU domain. The error of the forecasting system is less than the VPRM model error at flux observation site level, which means that the system performs sufficiently well for its predictive task. From the spatial distribution of the error, the absolute flux errors are larger in northern France, Germany and the Balkans, which will lead to larger bias in atmospheric CO<sub>2</sub> forecasting system. The assessment of these (and other) errors in concentration space, using measurements from the CoMet mission as reference data, is foreseen as a followup study.

*Code and data availability.* The code for forecast VPRM model and the model outputs are available from <https://doi.org/10.17617/3.2d> (Chen and Gerbig, 2019). The code used for model assessment and figure plotting in this paper is also included in the same repository. The flux measurement data can be acquired from the FLUXNET2015 database (see DOIs in Table 1). The MODIS reflectance data can be acquired from NASA's Earth Science Data Systems (<https://doi.org/10.5067/MODIS/MOD09A1.006>, Vermote,

2015). The ECMWF meteorology data can be retrieved using ECMWF's Meteorological Archival and Retrieval System (MARS; <https://confluence.ecmwf.int/display/UDOC/MARS+user+documentation>, last access: 1 September 2020).

*Author contributions.* The experiments were planned by CG, JM, KUT and JC. CG prepared the standard VPRM model. JC made the forecast model and performed the model simulation and assessment. JM extensively commented and revised the paper. JC prepared the paper with contribution from all coauthors.

*Competing interests.* The authors declare that they have no conflict of interest.

*Acknowledgements.* We acknowledge the use of data products from the Land, Atmosphere Near real-time Capability for EOS (LANCE) system operated by NASA's Earth Science Data and Information System (ESDIS) with funding provided by NASA Headquarters. We acknowledge ECMWF for providing access to the ECMWF's archived data. This work used eddy covariance data acquired and shared by the FLUXNET community, including these networks: AmeriFlux, AfriFlux, AsiaFlux, CarboAfrica, CarboEuropeIP, CarboItaly, CarboMont, ChinaFlux, Fluxnet-Canada, GreenGrass, ICOS, KoFlux, LBA, NECC, OzFlux-TERN, TCOS-Siberia and USCCC. The ERA-Interim reanalysis data are provided by ECMWF and processed by LSCE. The FLUXNET eddy covariance data processing and harmonization was carried out by the European Fluxes Database Cluster, AmeriFlux Management Project and Fluxdata project of FLUXNET, with the support of CDIAC and ICOS Ecosystem Thematic Center as well as the OzFlux, ChinaFlux and AsiaFlux offices.

*Financial support.* The research is funded by the MPG (Max Planck Society) through the CoMet campaign and by the BMBF (German Federal Ministry of Education and Research) through AIRSPACE (FK 01LK1701C). Jinxuan Chen is funded by a PhD project from IMPRS-gBGC (the International Max Planck Research School for Global Biogeochemical Cycles).

The article processing charges for this open-access publication were covered by the Max Planck Society.

*Review statement.* This paper was edited by Tomomichi Kato and reviewed by two anonymous referees.

## References

Agustí-Panareda, A., Massart, S., Chevallier, F., Boussetta, S., Balsamo, G., Beljaars, A., Ciais, P., Deutscher, N. M., Engelen, R., Jones, L., Kivi, R., Paris, J.-D., Peuch, V.-H., Sherlock, V., Vermeulen, A. T., Wennberg, P. O., and Wunch, D.: Forecast-

- ing global atmospheric CO<sub>2</sub>, *Atmos. Chem. Phys.*, 14, 11959–11983, <https://doi.org/10.5194/acp-14-11959-2014>, 2014.
- Agustí-Panareda, A., Massart, S., Chevallier, F., Balsamo, G., Boussetta, S., Dutra, E., and Beljaars, A.: A biogenic CO<sub>2</sub> flux adjustment scheme for the mitigation of large-scale biases in global atmospheric CO<sub>2</sub> analyses and forecasts, *Atmos. Chem. Phys.*, 16, 10399–10418, <https://doi.org/10.5194/acp-16-10399-2016>, 2016.
- Ahmadov, R., Gerbig, C., Kretschmer, R., Koerner, S., Neining, B., Dolman, A., and Sarrat, C.: Mesoscale covariance of transport and CO<sub>2</sub> fluxes: Evidence from observations and simulations using the WRF-VPRM coupled atmosphere-biosphere model, *J. Geophys. Res.-Atmos.*, 112, D22107, <https://doi.org/10.1029/2007JD008552>, 2007.
- Amediek, A., Ehret, G., Fix, A., Wirth, M., Budenbender, C., Quatrevalet, M., Kiemle, C., and Gerbig, C.: CHARM-F-a new airborne integrated-path differential-absorption lidar for carbon dioxide and methane observations: measurement performance and quantification of strong point source emissions, *Appl. Opt.*, 56, 5182–5197, [10.1364/Ao.56.005182](https://doi.org/10.1364/Ao.56.005182), 2017.
- Anthoni, P., Knohl, A., Rebmann, C., Freibauer, A., Mund, M., Ziegler, W., Kolle, O., and Schulze, E. D.: Forest and agricultural land-use-dependent CO<sub>2</sub> exchange in Thuringia, Germany, *Glob. Change Biol.*, 10, 2005–2019, 2004.
- Aubinet, M., Chermanne, B., Vandenhaute, M., Longdoz, B., Yernaux, M., and Laitat, E.: Long term carbon dioxide exchange above a mixed forest in the Belgian Ardennes, *Agr. Forest Meteorol.*, 108, 293–315, 2001.
- Aulagnier, C., Rayner, P., Ciais, P., Vautard, R., Rivier, L., and Ramonet, M.: Is the recent build-up of atmospheric CO<sub>2</sub> over Europe reproduced by models. Part 2: an overview with the atmospheric mesoscale transport model CHIMERE, *Tellus B*, 62, 14–25, <https://doi.org/10.1111/j.1600-0889.2009.00443.x>, 2010.
- Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., and Evans, R.: FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities, *B. Am. Meteorol. Soc.*, 82, 2415–2434, 2001.
- Beck, V., Koch, T., Kretschmer, R., Marshall, J., Ahmadov, R., Gerbig, C., Pillai, D., and Heimann, M.: The WRF Greenhouse Gas Model (WRF-GHG), Technical Report No. 25, Max Planck Institute for Biogeochemistry, Jena, Germany, available at: <http://www.bgc-jena.mpg.de/bgc-systems/index.shtml> (last access: 1 September 2020), 2011.
- Bernhofer, C., Grünwald, T., Moderow, U., Hehn, M., Eichelmann, U., and Prasse, H.: FLUXNET2015 DE-Oberbärenburg, [10.18140/FLX/1440151](https://doi.org/10.18140/FLX/1440151), 2008–2014.
- Bernhofer, C., Grünwald, T., Moderow, U., Hehn, M., Eichelmann, U., and Prasse, H.: FLUXNET2015 DE-Akm Anklam, [10.18140/FLX/1440213](https://doi.org/10.18140/FLX/1440213), 2009–2014.
- Bernhofer, C., Grünwald, T., Moderow, U., Hehn, M., Eichelmann, U., and Prasse, H.: FLUXNET2015 DE-Spw Spreewald, [10.18140/FLX/1440220](https://doi.org/10.18140/FLX/1440220), 2010–2014.
- Boussetta, S., Balsamo, G., Beljaars, A., Panareda, A. A., Calvet, J. C., Jacobs, C., van den Hurk, B., Viterbo, P., Lafont, S., Dutra, E., Jarlan, L., Balzarolo, M., Papale, D., and van der Werf, G.: Natural land carbon dioxide exchanges in the ECMWF integrated forecasting system: Implementation and

- offline validation, *J. Geophys. Res.-Atmos.*, 118, 5923–5946, <https://doi.org/10.1002/jgrd.50488>, 2013.
- Broquet, G., Chevallier, F., Bréon, F.-M., Kadyrov, N., Alemanno, M., Apadula, F., Hammer, S., Haszpra, L., Meinhardt, F., Morguá, J. A., Necki, J., Piacentino, S., Ramonet, M., Schmidt, M., Thompson, R. L., Vermeulen, A. T., Yver, C., and Ciais, P.: Regional inversion of CO<sub>2</sub> ecosystem fluxes from atmospheric measurements: reliability of the uncertainty estimates, *Atmos. Chem. Phys.*, 13, 9039–9056, <https://doi.org/10.5194/acp-13-9039-2013>, 2013.
- Butz, A., Dinger, A. S., Bobrowski, N., Kostinek, J., Fieber, L., Fischerkeller, C., Giuffrida, G. B., Hase, F., Klappenbach, F., Kuhn, J., Lübcke, P., Tirpitz, L., and Tu, Q.: Remote sensing of volcanic CO<sub>2</sub>, HF, HCl, SO<sub>2</sub>, and BrO in the downwind plume of Mt. Etna, *Atmos. Meas. Tech.*, 10, 1–14, <https://doi.org/10.5194/amt-10-1-2017>, 2017.
- Chen, J. and Gerbig, C.: Short-term forecasting of regional biospheric CO<sub>2</sub> fluxes in Europe using a light-use-efficiency model – Model code and output, Max Planck Society, <https://doi.org/10.17617/3.2d>, 2019.
- Chevallier, F., Wang, T., Ciais, P., Maignan, F., Bocquet, M., Altaf Arain, M., Cescatti, A., Chen, J., Dolman, A. J., and Law, B. E.: What eddy-covariance measurements tell us about prior land flux errors in CO<sub>2</sub>-flux inversion schemes, *Global Biogeochem. Cy.*, 26, GB1021, <https://doi.org/10.1029/2010GB003974>, 2012.
- Delpierre, N., Berveiller, D., Granda, E., and Dufrêne, E.: Wood phenology, not carbon input, controls the interannual variability of wood growth in a temperate oak forest, *New Phytol.*, 210, 459–470, 2016.
- Dietiker, D., Buchmann, N., and Eugster, W.: Testing the ability of the DNDC model to predict CO<sub>2</sub> and water vapour fluxes of a Swiss cropland site, *Agr. Ecosyst. Environ.*, 139, 396–401, 2010.
- Dušek, J., Čížková, H., Stellner, S., Czerný, R., and Květ, J.: Fluctuating water table affects gross ecosystem production and gross radiation use efficiency in a sedge-grass marsh, *Hydrobiologia*, 692, 57–66, 2012.
- Etzold, S., Ruehr, N. K., Zweifel, R., Dobbertin, M., Zingg, A., Pluess, P., Häslar, R., Eugster, W., and Buchmann, N.: The carbon balance of two contrasting mountain forest ecosystems in Switzerland: similar annual trends, but seasonal differences, *Ecosystems*, 14, 1289–1309, 2011.
- Fares, S., Savi, F., Muller, J., Matteucci, G., and Paoletti, E.: Simultaneous measurements of above and below canopy ozone fluxes help partitioning ozone deposition between its various sinks in a Mediterranean Oak Forest, *Agr. Forest Meteorol.*, 198, 181–191, 2014.
- Ferréa, C., Zenone, T., Comolli, R., and Seufert, G.: Estimating heterotrophic and autotrophic soil respiration in a semi-natural forest of Lombardy, Italy, *Pedobiologia*, 55, 285–294, 2012.
- Galvagno, M., Wohlfahrt, G., Cremonese, E., Rossini, M., Colombo, R., Filippa, G., Julitta, T., Manca, G., Siniscalco, C., and di Cella, U. M.: Phenology and carbon dioxide source/sink strength of a subalpine grassland in response to an exceptionally short snow season, *Environ. Res. Lett.*, 8, 025008, <https://doi.org/10.1088/1748-9326/8/2/025008>, 2013.
- Gerilowski, K., Tretner, A., Krings, T., Buchwitz, M., Bertagnolio, P. P., Belemezov, F., Erzinger, J., Burrows, J. P., and Bovensmann, H.: MAMAP – a new spectrometer system for column-averaged methane and carbon dioxide observations from aircraft: instrument description and performance analysis, *Atmos. Meas. Tech.*, 4, 215–243, <https://doi.org/10.5194/amt-4-215-2011>, 2011.
- Grünwald, T. and Bernhofer, C.: A decade of carbon, water and energy flux measurements of an old spruce forest at the Anchor Station Tharandt, *Tellus B*, 59, 387–396, 2007.
- Hollinger, D. and Richardson, A.: Uncertainty in eddy covariance measurements and its application to physiological models, *Tree Physiol.*, 25, 873–885, 2005.
- Hommeltenberg, J., Schmid, H. P., Drösler, M., and Werle, P.: Can a bog drained for forestry be a stronger carbon sink than a natural bog forest?, *Biogeosciences*, 11, 3477–3493, <https://doi.org/10.5194/bg-11-3477-2014>, 2014.
- Imer, D., Merbold, L., Eugster, W., and Buchmann, N.: Temporal and spatial variations of soil CO<sub>2</sub>, CH<sub>4</sub> and N<sub>2</sub>O fluxes at three differently managed grasslands, *Biogeosciences*, 10, 5931–5945, <https://doi.org/10.5194/bg-10-5931-2013>, 2013.
- IPCC: 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, edited by: Core Writing Team, Pachauri, R. K., and Meyer, L. A., IPCC, Geneva, Switzerland, 151 pp., 2014.
- Janssens, I., Segers, J., Roland, M., and Arriga, N.: FLUXNET2015 BE-Bra Brasschaat, <https://doi.org/10.18140/FLX/1440128>, 2016.
- Jung, M., Henkel, K., Herold, M., and Churkina, G.: Exploiting synergies of global land cover products for carbon cycle modeling, *Remote Sens. Environ.*, 101, 534–553, 2006.
- Kountouris, P., Gerbig, C., Rödenbeck, C., Karstens, U., Koch, T. F., and Heimann, M.: Technical Note: Atmospheric CO<sub>2</sub> inversions on the mesoscale using data-driven prior uncertainties: methodology and system evaluation, *Atmos. Chem. Phys.*, 18, 3027–3045, <https://doi.org/10.5194/acp-18-3027-2018>, 2018.
- Lasslop, G., Reichstein, M., Kattge, J., and Papale, D.: Influences of observation errors in eddy flux data on inverse model parameter estimation, *Biogeosciences*, 5, 1311–1324, <https://doi.org/10.5194/bg-5-1311-2008>, 2008.
- Le Quéré, C., Raupach, M. R., Canadell, J. G., Marland, G., Bopp, L., Ciais, P., Conway, T. J., Doney, S. C., Feely, R. A., Foster, P., Friedlingstein, P., Gurney, K., Houghton, R. A., House, J. I., Huntingford, C., Levy, P. E., Lomas, M. R., Majkut, J., Metzl, N., Ometto, J. P., Peters, G. P., Prentice, I. C., Randerson, J. T., Running, S. W., Sarmiento, J. L., Schuster, U., Sitch, S., Takahashi, T., Viogy, N., van der Werf, G. R., and Woodward, F. I.: Trends in the sources and sinks of carbon dioxide, *Nat. Geosci.*, 2, 831–836, <https://doi.org/10.1038/ngeo689>, 2009.
- Lin, J. C., Pejam, M. R., Chan, E., Wofsy, S. C., Gottlieb, E. W., Margolis, H. A., and McCaughey, J. H.: Attributing uncertainties in simulated biospheric carbon fluxes to different error sources, *Global Biogeochem. Cy.*, 25, GB2018, <https://doi.org/10.1029/2010gb003884>, 2011.
- Mahadevan, P., Wofsy, S. C., Matross, D. M., Xiao, X. M., Dunn, A. L., Lin, J. C., Gerbig, C., Munger, J. W., Chow, V. Y., and Gottlieb, E. W.: A satellite-based biosphere parameterization for net ecosystem CO<sub>2</sub> exchange: Vegetation Photosynthesis and Respiration Model (VPRM), *Global Biogeochem. Cy.*, 22, GB2005, <https://doi.org/10.1029/2006gb002735>, 2008.

- Marcolla, B., Pitacco, A., and Cescatti, A.: Canopy architecture and turbulence structure in a coniferous forest, *Bound.-Lay. Meteorol.*, 108, 39–59, 2003.
- Mauder, M., Cuntz, M., Drüe, C., Graf, A., Rebmann, C., Schmid, H. P., Schmidt, M., and Steinbrecher, R.: A strategy for quality and uncertainty assessment of long-term eddy-covariance measurements, *Agr. Forest Meteorol.*, 169, 122–135, 2013.
- Merbold, L., Eugster, W., Stieger, J., Zahniser, M., Nelson, D., and Buchmann, N.: Greenhouse gas budget (CO<sub>2</sub>, CH<sub>4</sub> and N<sub>2</sub>O) of intensively managed grassland following restoration, *Glob. Change Biol.*, 20, 1913–1928, 2014.
- Moureaux, C., Debacq, A., Bodson, B., Heinesch, B., and Aubinet, M.: Annual net ecosystem carbon exchange by a sugar beet crop, *Agr. Forest Meteorol.*, 139, 25–39, 2006.
- Pilegaard, K., Ibrom, A., Courtney, M. S., Hummelshøj, P., and Jensen, N. O.: Increasing net CO<sub>2</sub> uptake by a Danish beech forest during the period from 1996 to 2009, *Agr. Forest Meteorol.*, 151, 934–946, 2011.
- Pillai, D., Buchwitz, M., Gerbig, C., Koch, T., Reuter, M., Bovensmann, H., Marshall, J., and Burrows, J. P.: Tracking city CO<sub>2</sub> emissions from space using a high-resolution inverse modelling approach: a case study for Berlin, Germany, *Atmos. Chem. Phys.*, 16, 9591–9610, <https://doi.org/10.5194/acp-16-9591-2016>, 2016.
- Post, H., Hendricks Franssen, H. J., Graf, A., Schmidt, M., and Vereecken, H.: Uncertainty analysis of eddy covariance CO<sub>2</sub> flux measurements for different EC tower distances using an extended two-tower approach, *Biogeosciences*, 12, 1205–1221, <https://doi.org/10.5194/bg-12-1205-2015>, 2015.
- Potter, C. S., Randerson, J. T., Field, C. B., Matson, P. A., Vitousek, P. M., Mooney, H. A., and Klooster, S. A.: Terrestrial Ecosystem Production – a Process Model-Based on Global Satellite and Surface Data, *Global Biogeochem. Cy.*, 7, 811–841, <https://doi.org/10.1029/93gb02725>, 1993.
- Prescher, A.-K., Grünwald, T., and Bernhofer, C.: Land use regulates carbon budgets in eastern Germany: From NEE to NBP, *Agr. Forest Meteorol.*, 150, 1016–1025, 2010.
- Rambal, S., Joffre, R., Ourcival, J., Cavender-Bares, J., and Rocheteau, A.: The growth respiration component in eddy CO<sub>2</sub> flux from a *Quercus ilex* mediterranean forest, *Glob. Change Biol.*, 10, 1460–1469, 2004.
- Running, S. W. and Hunt Jr., E. R.: Generalization of a forest ecosystem process model for other biomes, BIOME-BCG, and an application for global-scale models, Academic Press, 141–158, <https://doi.org/10.1016/B978-0-12-233440-5.50014-2>, 1993.
- Running, S. W., Thornton, P. E., Nemani, R., and Glassy, J. M.: Global terrestrial gross and net primary productivity from the Earth Observing System, in: *Methods in ecosystem science*, Springer, 44–57, 2000.
- Sabbatini, S., Arriga, N., Bertolini, T., Castaldi, S., Chiti, T., Consalvo, C., Njakou Djomo, S., Gioli, B., Matteucci, G., and Papale, D.: Greenhouse gas balance of cropland conversion to bioenergy poplar short-rotation coppice, *Biogeosciences*, 13, 95–113, <https://doi.org/10.5194/bg-13-95-2016>, 2016.
- Simmons, A., Mureau, R., and Petroliaigis, T.: Error growth and estimates of predictability from the ECMWF forecasting system, *Q. J. Roy. Meteor. Soc.*, 121, 1739–1771, 1995.
- Simmons, A. J. and Hollingsworth, A.: Some aspects of the improvement in skill of numerical weather prediction, *Q. J. Roy. Meteor. Soc.*, 128, 647–677, 2002.
- Suni, T., Rinne, J., Reissell, A., Altimir, N., Keronen, P., Rannik, U., Maso, M., Kulmala, M., and Vesala, T.: Long-term measurements of surface fluxes above a Scots pine forest in Hyttiala, southern Finland, 1996–2001, *Boreal Environ. Res.*, 8, 287–302, 2003.
- Thum, T., Aalto, T., Laurila, T., Aurela, M., Kolari, P., and Hari, P.: Parametrization of two photosynthesis models at the canopy scale in a northern boreal Scots pine forest, *Tellus B*, 59, 874–890, 2007.
- Valentini, R., De Angelis, P., Matteucci, G., Monaco, R., Dore, S., and Mucnozza, G. S.: Seasonal net carbon dioxide exchange of a beech forest with the atmosphere, *Glob. Change Biol.*, 2, 199–207, 1996.
- Vermote, E.: MOD09A1 MODIS/Terra Surface Reflectance 8-Day L3 Global 500m SIN Grid V006, Data set, NASA EOSDIS Land Processes DAAC, <https://doi.org/10.5067/MODIS/MOD09A1.006>, 2015.
- Vitale, L., Di Tommasi, P., D’Urso, G., and Magliulo, V.: The response of ecosystem carbon fluxes to LAI and environmental drivers in a maize crop grown in two contrasting seasons, *Int. J. Biometeorol.*, 60, 411–420, 2016.
- Woodward, F. I., Smith, T. M., and Emanuel, W. R.: A Global Land Primary Productivity and Phytogeography Model, *Global Biogeochem. Cy.*, 9, 471–490, <https://doi.org/10.1029/95gb02432>, 1995.
- Wu, K., Lauvaux, T., Davis, K. J., Deng, A., Coto, I. L., Gurney, K. R., and Patarasuk, R.: Joint inverse estimation of fossil fuel and biogenic CO<sub>2</sub> fluxes in an urban environment: An observing system simulation experiment to assess the impact of multiple uncertainties, *Elem. Sci. Anth.*, 6, 17, <https://doi.org/10.1525/elementa.138>, 2018.
- Xiao, X. M., Hollinger, D., Aber, J., Goltz, M., Davidson, E. A., Zhang, Q. Y., and Moore, B.: Satellite-based modeling of gross primary production in an evergreen needleleaf forest, *Remote Sens. Environ.*, 89, 519–534, <https://doi.org/10.1016/j.rse.2003.11.008>, 2004.
- Zhuang, Q., McGuire, A. D., Melillo, J. M., Clein, J. S., Dargaville, R. J., Kicklighter, D. W., Myneni, R. B., Dong, J., Romanovsky, V. E., Harden, J., and Hobbie, J. E.: Carbon cycling in extratropical terrestrial ecosystems of the Northern Hemisphere during the 20th century: a modeling analysis of the influences of soil thermal dynamics, *Tellus B*, 55, 751–776, <https://doi.org/10.1034/j.1600-0889.2003.00060.x>, 2003.
- Zielis, S., Etzold, S., Zweifel, R., Eugster, W., Haeni, M., and Buchmann, N.: NEP of a Swiss subalpine forest is significantly driven not only by current but also by previous year’s weather, *Biogeosciences*, 11, 1627–1635, <https://doi.org/10.5194/bg-11-1627-2014>, 2014.