A comprehensive land surface vegetation model for multi-stream data assimilation, D&B v1.0

Wolfgang Knorr[1], Matthew Williams[2], Tea Thum[3], Thomas Kaminski[1], Michael Voßbeck[1], Marko Scholze[4], Tristan Quaife[5], T. Luke Smallman[2], Susan C. Steele-Dunne[6], Mariette Vreugdenhil[7], Tim Green[2], Sönke Zaehle[8], Mika Aurela[3], Alexandre Bouvet[9], Emanuel Bueechi[7], Wouter Dorigo[7], Tarek S. El-Madany[8], Mirco Migliavacca[8,9], Marika Honkanen[3], Yann H. Kerr[10], Anna Kontu[3], Juha Lemmetyinen[3], Hannakaisa Lindqvist[3], Arnaud Mialon[10], Tuuli Miinalainen[3], Gaétan Pique[10], Amanda Ojasalo[3], Shaun Quegan[11], Peter. J. Rayner[1], Pablo Reyez-Muñoz[12], Nemesio Rodríguez-Fernández[9], Mike Schwank[13], Jochem Verrelst[12], Songyan Zhu[2], Dirk Schüttemeyer[14], and Matthias Drusch[14]

[1]The Inversion Lab, Tewessteg 4, D-20249 Hamburg, Germany
[2]University of Edinburgh, Edinburgh, UK
[3]Finnish Meteorological Institute, Helsinki, Finland
[4]Lund University, Lund, Sweden
[5]University of Reading, Reading, UK
[6]Department of Geosciences and Remote Sensing, Delft University of Technology, The Netherlands
[7]Vienna University of Technology, Wien, Austria
[8]Max-Planck Institute for Biogeochemistry, Jena, Germany
[9]DG Joint Research Centre, European Commission, Ispra, Italy
[10]Centre d'études Spatiales de la Biosphère (CESBIO), Université de Toulouse 3 Paul Sabatier, Centre National de la Recherche Scientifique, Centre National d'Etudes Spatiales, Institut de Recherche pour le Développement, Institut National de Recherches pour l'Agriculture, l'Alimentation et l'Environnement, Toulouse, France
[11]University of Sheffield, Sheffield, UK
[12]University of Valencia, Valencia, Spain
[13]Swiss Federal Institute for Forest, Snow and Landscape Research, Birmensdorf, Switzerland
[14]European Space Agency, ESTEC, Noordwijk, The Netherlands

Correspondence: Wolfgang Knorr (wolfgang.knorr@inversion-lab.com)

Abstract. Advances in Earth Observation capabilities mean that there is now a multitude of spatially resolved data sets available that can support the quantification of water and carbon pools and fluxes at the land surface. However, such quantification ideally requires efficient synergistic exploitation of those data, which in turn requires carbon and water land-surface models with the capability to simultaneously assimilate several of such data streams. The present article discusses the requirements for such a model and presents one such model based on the combination of the existing DALEC land vegetation carbon cycle model with the BETHY land-surface and terrestrial vegetation scheme. The resulting D&B model, made available as a community model, is presented together with a comprehensive evaluation for two selected study sites of widely varying climate. We then demonstrate the concept of land surface modelling aided by data streams that are available from satellite remote sensing. Here we present D&B with four observation operators that translate model-derived variables into measurements available from such data streams, namely: fraction of photosynthetically active radiation (FAPAR), solar-induced chlorophyll fluorescence

(SIF), vegetation optical depth (VOD) at microwave frequencies, and near-surface soil moisture, also available from microwave measurements. As a first step, we evaluate the combined model system using local observations, and finally discuss the potential of the system presented for multi-stream data assimilation in the context of Earth Observation systems.

1 Introduction

Monitoring the status of land surface carbon pools has gained significant attention following various climate related pledges to balance carbon sources and sinks (Heiskanen et al., 2022). Indeed, even though anthropogenic carbon fluxes are responsible for creating a large imbalance of the global carbon cycle that has led to sustained and accelerating greenhouse-gas forcing, the largest CO_2 fluxes globally are related to plant photosynthesis, plant respiration and the decay of dead plant matter (Friedlingstein et al., 2022). These carbon fluxes are determined by climatic factors, the presence and amount of photosynthesising vegetation, and soil water availability, the latter due to the intrinsic water limitation of biological processes (Gerten et al., 2005).

A reliable characterisation of both carbon and water fluxes and pools at a range of spatial scales is therefore of paramount importance, as we currently lack a robust, spatially and temporally explicit knowledge of the sources and sinks of CO_2 within the terrestrial biosphere, or of the drivers of those variations. Current climate predictions and climate policy scenarios crucially depend on assumptions about the future fate of the terrestrial carbon pools and their interaction with future climate variations, but how variations in carbon fluxes interact with various forcing factors (such as climate, land use, CO_2 fertilisation) is still only partially understood (Arora et al., 2020). This makes policies that rely on future climate scenarios intrinsically unreliable.

The lack of knowledge exists despite of the availability of products of net or gross carbon uptake by terrestrial vegetation, such as those from MODIS with daily and up to 250 m resolution (Zhao et al., 2005), or from the Copernicus Global Land Service with a 300 m spatial and 10-day temporal resolution (Swinnen et al., 2021). One issue is that those products are no direct observations of carbon fluxes, but rather a combination of remotely sensed information and a set of assumptions. They thus do not necessarily agree with each other or with the results of ecosystem models (Turner et al., 2006; Sun et al., 2021). Another issue is that these only refer to CO_2 sinks, while we lack spatially distributed data sets of terrestrial biosphere CO_2 sources.

However, in order to identify the drivers of terrestrial carbon sources and sinks, such as vegetation state, soil carbon content of different qualities, temperature, soil moisture, atmospheric humidity, or light availability, we need models that are both internally consistent – i.e. can be run without remotely sensed input — and at the time thoroughly evaluated against reliable observations. Those observations should be as independent of specific model assumptions as possible, so that it is possible to clearly distinguish between model predictions by themselves (when we run without using those observations), and predictions resulting from the combination of observations and model assumptions.

Furthermore, if we also want to identify existing carbon sources and sinks and attribute those to certain drivers and processes, we also need to be able to run and evaluate those models at the spatial and temporal resolution of interest. Running models at high spatial and temporal resolution is not an issue in principle. The problem lies in finding suitable observations at high

temporal and spatial resolution for terrestrial ecosystem model evaluation and in finding out which model formulations, initial conditions and paramerisations can reproduce those observations.

Earth observation technology offers a powerful tool for observing the land vegetation and soil water status in multiple, complementary ways across time and space. However, there remain serious challenges for their exploitation, in particular a lack of a direct link between the variable of interest and remotely sensed information. In other words, remote sensing regularly provides only an indirect measure (Disney et al., 2016; Gao et al., 2021) – with widely varying accuracy – of some underlying processes, such as photosynthetic activity, rather than a quantification of carbon pools and fluxes. Or if carbon fluxes or pools are estimated, those estimates heavily rely on models and auxiliary inputs (Running and Zhao, 2015), or have so far been validated only locally (Liu et al., 2022).

Data assimilation is a valuable method for automatically finding the optimal combination of model initial values, parameters and even input quantities given the observations assimilated, pertinent to certain assumptions about prior values and uncertainties of models and data within a Bayesian framework (Tarantola, 2005). While not providing a ready made answer – it always needs to be assured that the thus optimised model simulations "make sense" – data assimilation can be used to find the most reliable model and data based estimates of quantities of interest, e.g. carbon fluxes, and serve as a tool for evaluating assumptions about the inherent processes driving changes in those fluxes. What we want to avoid, however, is to assimilate data streams that themselves rely substantially on model assumptions, such as the global GPP products mentioned above, because this would make the results depend on model assumptions outside of the core model used for assimilation. Thus, we expect significant added value if earth observation data are used within a data assimilation framework, allowing the synergistic use of multiple data streams (Berger et al., 2012; Ciais et al., 2014; Scholze et al., 2017). This is particularly relevant given that remote sensing offers unparalleled data coverage over large regional to global scales at high temporal frequency.

In this study we present a process-based modelling system that is suitable for the assimilation of a wide range of such data streams, enabling a synergistic multi-data stream land surface carbon monitoring and prediction system. So far, there have been a number of relevant studies, mostly using single data streams. For example, the Biosphere Energy-Transfer HYdrology scheme (BETHY, Knorr, 2000) has a long record of data assimilation studies using land surface temperature (Knorr and Lakshmi, 2001), atmospheric CO_2 (Rayner et al., 2005), fraction of photosynthetically active radiation (FAPAR, Knorr et al., 2010; Kaminski et al., 2012), eddy-flux measurements (Knorr and Kattge, 2005; Kato et al., 2013), solar-induced fluorescence (SIF, Norton et al., 2017, 2018, 2019), the combination of CO_2 and passive-microwave-derived soil moisture (Scholze et al., 2016, VOD,) as well as the combination of CO_2, L-band passive-microwave soil moisture and vegetation optical depth (Scholze et al., 2019). BETHY is a combined carbon and water land surface model that focuses on faster processes, such as energy and water exchanges, and carbon fluxes at time scales from hours up to several years, coinciding with the typical time span of satellite missions. BETHY is also the core of the first Carbon Cycle Data Assimilation System (CCDAS, Rayner et al., 2005; Kaminski et al., 2013). BETHY has been developed specifically for the purpose of assimilating both satellite and locally measured carbon and energy flux data. The main limitation of the above studies, however, is that BETHY does not account for plant growth and allocation, and therefore cannot capture slow increases in living-plant biomass over time.

Following up from the experiences gained from the studies previously cited, we identify the following essential requirements for a process-model at the centre of the envisaged land surface carbon monitoring system:

- representation of internal processes affecting carbon, water and energy fluxes at time scales of hours to several years, to permit spatial and temporal scaling;
- representation of specific variables that directly link to remotely sensed information and, if needed, related 'observation operators', i.e. modules that simulate the same variable as provided by the assimilated data stream (e.g. FAPAR, VOD);
- computational efficiency, high degree of simplicity while retaining sufficient realism, and ideally the availability of the adjoint model code to enable the use of efficient variational assimilation approaches.

There are a number of models that could potentially fulfill those requirements. They range from carbon models incorporated into routine weather forecasting, such as C-TESSEL (Boussetta et al., 2013), to highly complex land surface and ecosystem models that can be operated both within earth system models or independently, such as JULES (Best et al., 2011; Harper et al., 2016) or ORCHIDEE (Traore et al., 2014). Of these, C-TESSEL has probably the strongest track-record for assimilation of satellite data, mainly for the purpose of constraining soil moisture (Scipal et al., 2008). However, it does not simulate the mass balance of carbon, despite simulating photosynthesis and respiration, nor can it predict leaf area index, which it requires as input data. C-TESSEL is therefore of limited use when assimilating FAPAR, or variables related to biomass.

By contrast, JULES includes a full set of carbon fluxes and pools (Clark et al., 2011). An adjoint version of JULES has been developed to optimise parameters at site level using eddy flux data (Raoult et al., 2016). ORCHIDEE includes not only carbon but also nitrogen cycling (Vuichard et al., 2019). A data-assimilation framework for ORCHIDEE also exists, which has been successfully employed at site level for the step-wise optimisation of model parameters using remote sensing data (e.g. FAPAR), as well as water and carbon flux observations from the eddy covariance flux networks (Peylin et al., 2016).

We note that less complex models such as C-TESSEL are often much better suited for data assimilation than complex models, because a simpler structure with fewer parameters, omitting processes not relevant at the time scales of interest, makes optimisation both computationally and mathematically much more feasible. For example, C-TESSEL and BETHY lack representation of carbon pools (except for leaf area in the case of BETHY) due to a focus on short time scales of up to a few years. This is contrasted by another model, DALEC (Data Assimilation Linked Ecosystem Carbon, Williams et al., 2005), which focuses on carbon pools and longer-term processes, but is structurally also simple. DALEC has been developed specifically for assimilating information on C fluxes and pools from satellite observations (Bloom and Williams, 2015), eddy covariance systems (Bloom and Williams, 2015; Famiglietti et al., 2021), and biometric data including biomass (Smallman et al., 2017; Quegan et al., 2019).

In this study, we therefore address the above requirements by the development of a new process model, which combines the BETHY and DALEC models, both of which have been specifically designed with data assimilation in mind and have a corresponding track record. BETHY provides a detailed representation of fast and intermediate-time scale (hours to months) processes while DALEC provides a focus on slower processes of carbon allocation and turnover (months to years). This

combination opens up the possibility of retrieving variables such as biomass carbon stocks, that were not the focus of the assimilation studies using BETHY. DALEC is a mass-balance model that simulates the dynamics of live and dead carbon pools and associated fluxes (Williams et al., 2005; Bloom and Williams, 2015). Data assimilation is used to assign parameter values and their uncertainty ranges, as well as model initial conditions, at the pixel-scale across the modelled domain. DALEC requires input in the form of either gross or net primary productivity (NPP) from a separate model, in this case BETHY. In return, DALEC provides information on leaf area back to BETHY.

In this contribution we present the newly developed D&B (DALEC & BETHY) model together with original measurements from two study sites of widely varying climate and vegetation. Both, the model development and the field campaign, were carried out within the ESA-funded Land surface Carbon Cycle (LCC) study. Due to the considerable computational demands of data assimilation, D&B avoids any complexity that cannot be justified by the need to improve the realistic simulation of target variables. Coupling of the two models together with their respective components, state and driving variables is shown in Section 2 (Figure 1). D&B has been developed with the specific purpose of providing a modular and flexible modelling scheme for the assimilation of multiple data streams. We present the various components of the core model, and a detailed evaluation of the *a priori*, uncalibrated model. We also present observation operators for FAPAR, SIF, VOD and near-surface soil moisture, and a further evaluation of the model combined with each observation operator against locally measured data, as a precursor to the use of satellite-derived Earth observation data.

2 Model description

The D&B model is comprised of three interconnected components: (i) photosynthesis and autotrophic respiration, (ii) energy and water balance, and (iii) carbon allocation and cycling, including heterotrophic respiration (Figure 1). The first comprises processes that lead to the uptake of CO_2 via plant photosynthetic activity (gross primary production, GPP), influenced by temperature, light absorption across the canopy, and stomatal control, as well as carbon loss from the respiration of live vegetation (R_A, autotrophic respiration). The remaining carbon flux is then passed as net primary production (NPP = GPP - R_A) into the Carbon Allocation and Cycling component. The Energy and Water Balance determines the energy input to and output from the canopy in the form of radiative heat, latent and sensible heat transport, taking into account the water balance of the canopy and soil, as well as the rate of water uptake from the roots. Components (i) and (ii) are based on BETHY, and component (iii) on DALEC.

Depending on the domain for which the model is set up, D&B distinguishes up to 13 Plant Functional Types (PFTs) as shown in Table 1. Each PFT is characterised by a unique set of parameter values. All PFTs use the C3 photosynthetic pathway, except for PFT 10, for which a separate module for C4 photosynthesis is used (see the Supplementary Information, SI, Section S1.1.1). Management of arable crops is represented by appropriate parameters for leaf onset and fall, as well as assumptions about a minimum level to which soil moisture is allowed to fall, as an approximation of irrigation (Section S1.2.7).

The fundamental model time step is 1 hour. The following components are, however, simulated at a daily time step in order to decrease the computational effort:

Table 1. Parameter combinations are available for the following plant functional types in D&B:

PFT#	Short name	Description
1	TrEV	Tropical broadleaf evergreen tree
2	TrDD	Tropical broadleaf deciduous tree
3	TmEv	Temperate broadleaf evergreen tree
4	TmSg	Temperate broadleaf deciduous tree
5	EvCn	Evergreen coniferous tree
6	SgCn	Deciduous coniferous tree
7	EShr	Evergreen shrub
8	DShr	Deciduous shrub
9	C3Gr	C3 grass
10	C4Gr	C4 grass
11	TunV	Tundra
12	WetV	Wetland
13	ArbC	Arable crop

- soil water balance
- canopy water balance
- snow module
- the observation operators for VOD and surface soil moisture

The model simulates several PFTs in sub-grid tiles. Each PFT is simulated separately as if it would cover the full grid cell, with the results re-scaled by multiplying them with the grid-cell fraction occupied by the specific tile. Inter-PFT competition for light or water are neglected. A given grid cell can thus comprise several PFTs each with its specific cover fraction.

2.1 Photosynthesis and Autotrophic Respiration

The C3 photosynthesis module (Section S1.1.1) is based on the biochemical model of photosynthesis by Farquhar et al. (1980). It determines light absorption, light-limited electron transport, CO_2-limited carboxylation rate and the resulting gas exchange of CO_2. Light absorption in the photosynthetically active spectrum is calculated within a two-flux approximation (Section S1.1.1), following Sellers (1985). D&B divides the canopy into several vertical layers of equal LAI, the sum of which constitutes the total canopy LAI. In the standard configuration, the number of layers is three. The amount of light absorbed and thus available for photosynthesis is dependent on LAI, statistical leaf orientation (assumed to be isotropic) and leaf single-scattering albedo. Photosynthetic capacity decreases from top to bottom of the canopy, assuming that decreasing levels of daily-average solar radiation drive decreases in leaf nitrogen content and maximum rates of light-limited photosynthesis.

The photosynthesis module further divides GPP into NPP and R_A (Section S1.1.2). R_A is modelled as the sum of maintenance and growth respiration (Knorr, 1997). While maintenance respiration is proportional to photosynthetic capacity, growth respiration is proportional to NPP, and zero when NPP is negative. Both continually increase with temperature. Negative NPP is also passed on to the C Allocation and Cycling component, where it leads to the depletion of the labile C pool.

The rate of photosynthesis is first computed under standard conditions without limitation by water availability. This potential photosynthesis rate is translated into an equivalent stomatal conductance, i.e. the stomatal conductance necessary to provide the flow of CO_2 to the leaf interior. This value for stomatal conductance without water limitation is reduced depending on the vapour pressure deficit of the surrounding air, and available soil moisture. This modified stomatal conductance, or "actual stomatal conductance", then determines actual photosynthesis and, using information from the Energy and Water Balance component, the rate of transpiration.

2.2 Energy and Water Balance

The Energy and Water Balance component requires the rate of transpiration from the photosynthesis module, due to the tight coupling between water loss through transpiration and CO_2 uptake by leaves. Transpiration (Section S1.2.4) is subsequently combined with other evaporative fluxes, namely of intercepted water (canopy evaporation, Figure 2, Section S1.2.3), and from the soil surface (soil evaporation, Section S1.2.5), including snow sublimation (Section S1.2.7), to derive total evapotranspiration and latent heat flux. Latent heat flux is constrained by the available net radiative energy input, which the model computes separately for the vegetation canopy and the soil (Section S1.2.6). Sensible heat flux is computed from the assumption of energy closure from net radiation, latent and soil heat flux. The model uses incoming shortwave (solar) and longwave (thermal) radiation as input, but simulates both outgoing radiation components internally, using information on the albedo of the soil background and vegetation (Section S1.2.8).

Soil evaporation proceeds at the equilibrium rate driven by the soil net radiation from a thin surface layer. This corresponds to a typical depth for which microwave remote sensing can provide soil moisture estimates (Babaeian et al., 2019). Evapotranspiration from the canopy happens as either canopy evaporation from leaf surfaces at the equilibrium rate (Jarvis and McNaughton, 1986), or as transpiration through leaf pores. Precipitation enters either the leaf interception pool, or the soil pool (Section S1.2.1). Precipitation happens as either snow (Section S1.2.7), or rainfall, partitioned into a canopy-interception part, soil infiltration, and surface runoff. Soil water drains as sub-surface runoff, or base flow. Infiltration into the soil (Section S1.2.2), runoff, drainage and baseflow (Section S1.2.5) are simulated following a new implementation of the variable infiltration capacity approach (Wood et al., 1992), where a thin surface layer has been added to a single root-zone layer, (Scholze et al., 2016). The surface soil moisture layer overlaps with the root zone layer, so that the near-surface soil water pool forms part of the root-zone soil water pool (Figure 2). The former has a nominal depth of 4 cm, the latter a depth equal to a PFT-specific root depth, d_r (Table S1 in Section S1.1.2). Both depths are limited by depth to bedrock. Soil water exiting the root zone downwards is considered subsurface drainage, while there is no upward water movement from below the root zone. The root zone soil moisture pool contains all simulated soil water, while the surface layer is added in order to be able to calculate

soil evaporation, as well as for diagnostic purposes, accounting for the impact of surface soil moisture on microwave remote sensing.

Once per day around the time of maximum evaporative demand, assumed to be at the hourly time step closest to 13:00 hours local solar time (Knorr, 1997), the parameters determining actual stomatal conductance are reset to reflect soil water status. To do this, transpiration is simulated as the minimum of a root water supply rate, which increases linearly as soon as soil water exceeds the permanent wilting point, reaching a maximum with soil water at field capacity, and the demand for transpiration. This rate of demand is determined by the potential rate of photosynthesis without water stress computed previously at each time step. Potential photosynthesis is assumed as the rate at a fixed ratio of leaf to atmospheric CO_2 content (0.87 for C3 and 0.67 for C4 photosynthesis). Actual photosynthesis and stomatal conductance are then set such that transpiration is downregulated from its potential rate to the rate of maximum root water supply. A supply-demand calculation then determines the rate at which leaf stomata close in response to the water vapour deficit of the surrounding air. (See Section S1.2.4, Equ. S67).

Finally, the surface reflectance, or background albedo (ρ_S), is affected by soil brightness, surface soil water content, and the presence of snow. Vegetation albedo as a function of absorption in the photosynthetically active spectrum, computed in the photosynthesis module, and snow albedo is modelled depending on snow age (Loth and Graf, 1996; Knorr, 1997).

2.3 Carbon Allocation and Cycling

The carbon cycle in D&B is expressed as a series of six equations describing the dynamics of six carbon pools. Other than the original DALEC model, D&B employs an hourly time step for allocation, the same as the time step used by the photosynthesis module. There are four live C pools, for foliage (*fol*), wood (*wd*) and fine roots (*fr*), a labile (*lab*) pool which supports foliage expansion, and two dead organic matter pools, namely litter (*lit*) and soil organic matter (*SOM*). The state equations describe the change over time in pool sizes on the basis of C fluxes in and out of the pool. Carbon inputs are all derived originally from NPP. NPP is allocated to each of the four live biomass pools based on fixed site or PFT-specific fractions.

The labile C pool in D&B represents the stored C used to initiate accelerated leaf development at the start of the growing season (Section S1.3.1). The phenology scheme parameterises timing of local bud burst via allocation to leaves from the labile pool based on calibrated climate sensitivity. Leaf development thus depends on the allocation of labile carbon, replenished from NPP, to the leaf carbon pool in addition to direct allocation from NPP. The leaf area index is determined by the conversion of leaf carbon pool size to leaf area by way of fixed values of leaf mass per area.

Losses from fine root (*fr*) and wood pools (*wd*) are determined by first-order differential equations, using a decay constant. Biomass dynamics of plant pools are the outcome of NPP allocation and these mortality losses (Section S1.3.2). Parameters for the C cycle in D&B use PFT calibrations derived for DALEC using the CARDAMOM model-data fusion approach (Bloom and Williams 2015). CARDAMOM uses ecological and dynamical constraints to ensure that allometric relationships arising from parameter selection (like emergent root:shoot ratios) are kept within ecologically realistic bounds. By calibrating DALEC using both LAI and woody biomass data, a constraint is placed on relevant model parameters to match the measured biomass of these plant organs.

Losses from the fine root pool replenish the litter pool, added by strongly periodic inputs linked to leaf senescence, while wood directly feeds SOM. The litter pool decays either to CO_2 via heterotrophic respiration, or is decomposed to the SOM pool. Mineralisation of both SOM and litter C pools by heterotrophic respiration thus results in further CO_2 fluxes. Total ecosystem respiration (TER) is determined by the sum of autotrophic growth and maintenance respiration, and mineralisation of dead organic matter (*lit* or *som*), creating a flux of heterotrophic respiration. Following the procedure used for DALEC, the prior parameter values of the Carbon Allocation and Cycling are set through a regional-scale calibration procedure, as described in Section S1.3.3.

3 Observation Operators

The task of an observation operator is to simulate the equivalent of an observation from the model's state variables. This includes the simulation of the variable that is retrieved at the time when it was observed and over the footprint of the observations, i.e. the source area of the signal measured by the respective instrument. (Kaminski and Mathieu, 2017). In this manuscript, we will present the simulation of four data streams, namely FAPAR, SIF, L-band VOD and near-surface soil moisture, and then confront model simulations with local measurements. Of the four data streams, FAPAR and surface soil moisture are internally calculated.

3.1 Fraction of absorbed photosynthetically active radiation (FAPAR)

FAPAR is a measure of the capacity of terrestrial vegetation to absorb sunlight in the visible spectrum, i.e. that part that can be utilised for photosynthesis. It is defined as the amount of photosynthetically active radiation (PAR) absorbed by green, functioning leaves divided by the total incoming PAR. FAPAR is calculated within the two-flux canopy radiative transfer scheme (Section S1.1.1) required for the calculation of GPP (Section 2.1). However, due to the dependence of FAPAR on solar zenith angle, it is necessary to take into account the solar zenith angle at time of observation. Therefore, the observation operator for FAPAR utilizes FAPAR calculations performed within the model's photosynthesis component at the times and dates where model and observations solar zenith angles match.

3.2 Solar Induced Fluorescence (SIF)

Strictly speaking, the canopy level solar-induced chlorophyll fluorescence, or SIF, is a measure not of the photosynthetic rate as such, but of the amount of radiation absorbed by the leaf and not used for the purpose of photosynthesis. Some of that surplus radiation is re-emitted as fluorescent light as part of a copying mechanism of the photosynthetic system. Under normal field conditions, however, SIF can often be used as an indication of photosynthetic activity, as opposed to FAPAR, which only characterises the photosynthetically active light that is potentially available (Porcar-Castell et al., 2014; Mohammed et al., 2019).

To calculate SIF, we use the formulation of Gu et al. (2019). This choice is motivated by the direct link to the photosynthesis routines and the relatively parsimonious implementation, which fits with the modelling strategy adopted here. The canopy layer SIF, S_n, is given by:

$$S_n = s_{SIF} J_n \frac{1 - \psi_{PSIImax}}{q_L \psi_{PSIImax}(1 + k_{DF})}, \quad (1)$$

where J_n is the electron transport in canopy layer n (Equ. S8), $\psi_{PSIImax}$ is the maximum photochemical quantum yield of photosystem II, q_L is the fraction of open photosystem II reaction centres and k_{DF} the ratio of the first order rate constants for heat dissipation and fluorescence. We take the values prescribed by Gu et al. (2019). Note that the original equation in that paper also has a term for the photon escape probability from the canopy. In D&B, this is calculated explicitly by the layered 2-stream model (Section S2) and hence is not required here. As an extension to the model by Gu et al. (2019) in view of the anticipated calibration in a data assimilation scheme, we further introduce the scaling factor s_{SIF}, which compensates for large uncertainties in (1) the values of the three constants $\psi_{PSIImax}$, q_L, and k_{DF} and (2) the spectral conversion that is described below. We set the prior value of s_{SIF} to 1.

SIF produced by the D&B model via the layered 2-stream model described in Section S2 has native units of $mol.m^{-2}s^{-1}$. It represents the total flux of photons into the hemisphere above the canopy for all wavelengths. Satellite measurements and *in situ* observations, however, are typically recorded in energy flux units per steradian, per nano-metre of the SIF spectra, e.g. $Wm^{-2}s^{-1}nm^{-1}sr^{-1}$. To convert from molar to energy units, we apply the molar form of the Planck equation, providing energy per mole of photons, e. ($e = ahc/\lambda_\phi$, where a is Avogadro's number, 6.023×10^{23}, h is the Planck constant, $6.626 \times 10^{-34}\ m^2kg.s^{-1}$), c is the speed of light, $3.0 \times 10^8\ m.s^{-1}$ and λ_ϕ is the wavelength of the SIF photons in metres.

We convert to steradians by using a constant factor of $\frac{1}{2\pi}$, which assumes that the emittance of SIF from the top of the canopy is isotropic, and finally weight by the relative strength of emissions at λ_ϕ compared to a reference SIF spectrum, i.e.:

$$w = \frac{E(\lambda_\phi)}{\sum_i E(\lambda_{\phi,i})}, \quad (2)$$

where E is the SIF emission spectrum of arbitrary units. Hence:

$$SIF' = SIF \frac{ew}{2\pi}, \quad (3)$$

Where SIF has units of $mol.m^{-2}s^{-1}$ and SIF' has units of $Wm^{-2}s^{-1}sr^{-1}nm^{-1}$.

For the present study we use a SIF emission spectrum observed at the Hyytiälä site in Finland (Magney et al., 2019). The SIF spectrum was measured for four Scots pine trees at light level of 1200 $\mu molm^{-2}s^{-1}$ and then averaged.

3.3 Vegetation optical depth (VOD)

D is essentially a parameter describing the attenuation of microwave radiation at some wavelength due to the presence of vegetation. It depends on the dielectric properties (due to water content, temperature and chemical composition) as well as

the structure and geometry of the vegetation, and sensor properties (e.g. wavelength, polarization). Due to the relatively static nature of structure, dynamics of VOD are generally attributed to changes in above ground biomass and water content (Ulaby and Wilson, 1985; Konings et al., 2019). It is measured within the microwave spectrum with passive instruments, using the black body radiation of the surface in the microwave domain), or active instruments such as scatterometers or synthetic aperture radars.

Common retrieval methods may extract both the surface soil moisture and VOD simultaneously from satellite remote sensing data, provided enough measurements are performed. For example the SMOS (Kerr et al., 2010) retrieval algorithm (Kerr et al., 2012) is based on the so-called $\tau - \omega$ formulation for the vegetation contribution (Kirdiashev et al., 1979; Mo et al., 1982) of the microwave signal, where VOD is denoted by τ, the perpendicular vegetation optical depth (Wigneron et al., 2007, 2010).

Here, however, we compare simulations to locally measured L-band passive VOD measurements. Due to the local setup where separate sensors are placed above and below the canopy (see Section S3.6), it is possible to measure VOD directly without having to solve for soil moisture simultaneously.

We use a semi-empirical formulation for L-band VOD, expressed as:

$$\tau_\lambda = f(T)(l_{wd}C_{wd} + l_{fol}C_{fol})(l_s f_{soil} + l_f f_E + l_0), \tag{4}$$

where the subscript λ denotes its wavelength dependence, C_{fol} and C_{wd} are the leaf and woody biomass pools, respectively (see Section S1.3.2), f_{soil} (Equ. S92) is fractional plant-available soil water content, and

$$f_E = E_t / E_{t,max} \tag{5}$$

i.e. the ratio of actual to potential transpiration (see Equs. S59 and S64). f_{soil} describes slow changes in the plant's hydrological status, hence multiplied by parameter l_s, and f_E fast changes, multiplied by parameter l_f. The other empirical parameters are l_0, l_{wd} for dependence on woody biomass, and l_{fol} for dependence on leaf biomass. We note that the five empirical parameters are wavelength-dependent; for convenience we refrain from adding an additional subscript τ.

Following Schwank et al. (2021), we include an explicit temperature dependency in the form of:

$$f(T) = 0.25 + 0.75/(1 + e^{-0.5(T+3)}), \tag{6}$$

which approximates theoretically derived behaviour around the freezing point, with T being 2-m air temperature. This formulation can be used across a range of microwave wavelengths, using different parameter values in each case. The second multiplicative factor in Equ. 4 is an empirical, linear expression using both woody and foliar biomass with the assumption that VOD will be zero if no biomass is present. The third multiplicative factor describes how the water status of the vegetation modifies this expression. This last one also contains a constant factor, $l_0 > 0$, because we expect positive VOD even if vegetation water stress is at its maximum.

In this contribution, we apply it to passive L-band microwave measurements. The values of the parameters for the empirical VOD observation operator, shown in Table 2, were selected such that the model reproduces a reasonable fit to L-band observations from SMOS over the Sodankylä and the Majadas del Tietar sites.

Table 2. Parameters for empirical observation operator for L-VOD.

Parameter	Unit	Value
l_{wd}	m^2/gC	2.0×10^{-4}
l_{fol}	m^2/gC	2.0×10^{-4}
l_s	-	1.20
l_f	-	4.0
l_0	-	0.4

3.4 Near-surface soil moisture

In the D&B model, near-surface soil moisture is represented by an explicitly modelled thin surface soil moisture layer, with a depth of 4 cm, unless depth-to-bedrock indicates a lower value. It is therefore a state variable in the model's soil water component, and is described in detail in Section 2.2. This surface layer therefore here serves a dual purpose. First, to simulate soil evaporation. Second, to diagnose a variable that can be retrieved from satellite observations. Near-surface soil moisture is usually available from passive microwave measurements when retrieved simultaneously with VOD Section 5.3. These retrieval algorithms explicitly separate the contributions to the microwave signal that come either from the vegetation (VOD) or from the soil (surface soil moisture).

4 Model evaluation

We first present an evaluation of the D&B model on its own, followed by an evaluation of the model together with the observation operators for the four data streams FAPAR, SIF, L-VOD and surface soil moisture. The methods used to derive the driving data for the model as well as those of the measurements undertaken for driving and evaluating the model are described in Section S3.

4.1 Study sites

The D&B model is run for two study sites with widely varying climate, one representing a boreal forest – Sodankylä in Finland, a Class 1 site of the ICOS network (Rebmann et al., 2018) – the other representing a temperate savanna ecosystem – Majadas de Tietar in Spain, also an ICOS network site. The Sodankylä Scots pine forest site (67^{o}21'44.6"N, 26^{o}38'18.9"E) is situated 100 km north of the Arctic Circle (Thum et al., 2007; Honkanen et al., 2023). It also has an understory of evergreen ericacecous shrub (mostly blueberry) as well as lichen and mosses. The soil is characterised as predominantly sand (0-10 cm: 0.5/6.0/88.1/5.4% clay/silt/sand and stone, respectively; 10-20 cm: 0.3/4.1/93.5/2.1%; 20-40 cm: 0.2/2.8/91.9/5.1%.

The Majadas de Tietar site is a Mediterranean open woodland of evergreen Holm oak in western Spain (39^{o}56'24.68"N, 5^{o}45'50.27"E, Wang et al., 2016; El-Madany et al., 2018). The soil (Nair et al., 2020) contains an upper sandy layer (5% Clay, 20% silt, 75% sand, 20 cm deep) underlain by a clay layer (30 to 60 cm depth, no information for 20 to 30 cm).

4.2 Model setup

We use locally observed data to drive the model (see Section S3.1). C_a is set to a uniform value of 405 ppm, i.e $C_a = 405 \times 10^{-6}$mol(CO_2)/mol(air), which is approximately the annual mean value at Mauna Loa, Hawaii, centered around the beginning of 2017 (NOAA, 2024). Model runs are with *a priori* values of the parameters for all modules and observation operators. Initial water content of the soil was set to 50% of field capacity. Simulations for the first two calendar years were discarded to avoid model biases due to initial conditions of the water balance and short-lived carbon pools. The fractional vegetation cover is set to $f_c = 1$ for both sites.

For Sodankylä, the model simulation was run for the period 1 Jan 2009 to 31 Dec 2021, with two PFTs, evergreen coniferous forest (PFT 5, 67% of ground area) and evergreen shrub (PFT 7, 33%). Measured soil temperature as model input is for 1 m depth. The soil texture class is "medium/coarse" (cf. Section S1.2.5, Table 4), following the classification of the global soil texture data set by Zobler (1986). Parameter and initial values related to carbon turnover are set according to Table S6 (Section S1.4).

For Majadas de Tietar, the simulation is for 1 Apr 2014 to 31 Dec 2021. We assume the site area to comprise C3 grass (PFT 9, 80% of ground area) and temperate evergreen trees (PFT 3, 20%). The model is driven by soil temperature measured at 80 cm depth averaged between four two locations, two in open grassland and two under a tree canopy. The soil texture class is "medium". Parameters and initial values related to carbon turnover are set according Table S7.

4.3 Evaluation approach

The D&B model is compared against eddy covariance data of carbon and energy fluxes, locally observed radiation balance, and, in the case of the boreal site, snow depth. This is a first evaluation of the model with its *a priori* parameterisation and its purpose is to assess whether the model is able to reproduce measurements with a reasonable degree of realism. The role of the in situ observations is to serve as an independent evaluation data set.

We compare multi-year time series by showing the following values for both observations and model simulations: $f_j, f_j^{min}, f_j^{max}$, where

$$f_j = \frac{1}{n}\sum_{i=1}^{n} f(i,j), \tag{7}$$

and f is the flux of interest, i counts the n simulation years used for this analysis, j the day within the year (Jan 1 to Dec 31). f_j^{min} and f_j^{max} denote the minimum or maximum across the n values $\{f(1,j),\ldots,f(n,j)\}$, respectively. We also show the multi-year mean for both observations and models as

$$\bar{f} = \frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m} f(i,j), \tag{8}$$

where m is the number of days per year. In addition, we provide the following metrics: root-mean square error (RMSE) of daily

$$\text{RMSE}_{\text{daily}} = \frac{1}{nm}\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{m}\left[f_{\text{mod}}(i,j) - f_{\text{obs}}(i,j)\right]^2} \tag{9}$$

and annual values:

$$\text{RMSE}_{\text{annual}} = \frac{1}{n}\sqrt{\sum_{i=1}^{n}\left(\bar{f}_{i,\text{mod}} - \bar{f}_{i,\text{obs}}\right)^2}, \tag{10}$$

with $\bar{f}_i$ denoting annual average fluxes for year i for either simulations (mod) or observations (obs); as well as explained variance (r^2) at daily and annual time scales:

$$r^2_{\text{daily}} = \frac{\sum_{i=1}^{n}\sum_{i=j}^{m}\left(f_{\text{mod}}(i,j) - \bar{f}_{\text{obs}}\right)^2}{\sum_{i=1}^{n}\sum_{i=j}^{m}\left(f_{\text{obs}}(i,j) - \bar{f}_{\text{obs}}\right)^2} \tag{11}$$

and

$$r^2_{\text{annual}} = \frac{\sum_{i=1}^{n}\left(\bar{f}_{i,\text{mod}} - \bar{f}_{\text{obs}}\right)^2}{\sum_{i=1}^{n}\left(\bar{f}_{i,\text{obs}} - \bar{f}_{\text{obs}}\right)^2} \tag{12}$$

Carbon fluxes are gross primary productivity (GPP), total ecosystem respiration (TER), and net ecosystem exchange (NEE=TER-GPP). NEE is defined as going from the vegetation to the atmosphere, i.e. positive values denote a flux of CO_2 towards the atmosphere. Energy fluxes are latent heat flux (LHF) and sensible heat flux (SHF), with the addition of net radiation, which is the balance of incoming minus outgoing solar and thermal radiation fluxes. Both carbon and energy fluxes, and net radiation are measured over a representative area of each ecosystem, consisting of different PFTs.

For the purpose of calculating the above statistics, we used the 6 year period from 2016 to 2021 for both sites. We additionally use snow depth measurements from the period 2011 to 2021 for validation at the Sodankylä site. Biomass and soil carbon measurements, also at the Sodankylä site, were taken in 2011, and are compared to mean values of the simulations from 2011 to 2021 (Sodankylä) or 2016 to 2021 (Majadas del Tietar). See Section S3 for details of measurement methods.

4.4 Evaluation at Boreal forest site: Sodankylä

At the boreal-forest site (Figure 3), measurements show a markedly smaller amplitude of the annual cycle of carbon fluxes (NEE) than the model. While in the spring time, there is a reasonable agreement with an initial rapid increase in the magnitude of NEE, carbon loss during the winter, and carbon uptake during summer and early autumn are clearly larger for the model. Between day-of-year (DOY) 200 and 260 (mid July to mid September), the discrepancy in NEE is associated with an underestimation by the model of respiration (TER) and an overestimate of GPP. Not surprisingly, GPP agrees well during the winter as it is well constrained due to the lack of light and low temperatures, but TER is generally higher for the model. There is also

a conspicuous phase shift of TER between the two curves, with measurements showing TER peaking much later, while the phases of GPP agree reasonably.

For the energy flux evaluation (Figure 4), what stands out is the good agreement between modelled and simulated SHF, except for the spring (ca. DOY 50 to 100), where observations exceed simulations. LHF is also well matched during the summer, (ca. DOY 120–260). Since in the model, energy balance is exactly fulfilled, we would expect an equally good match for the net energy input (i.e. net radiation minus ground heat flux, cf. Equ. S124, S127) if the energy balance is also fulfilled for the observations. However, observations during the summer period are systematically lower than simulations. Therefore, we attribute the mismatch in net radiation during the summer to a lack of energy closure of the eddy covariance measurements. However, for the winter months, SHF is in good agreement, but both net energy input and LHF show systematically higher values for the model, and hence there is no evidence of lack of energy closure for the measurements. The difference might thus be mostly due to an overestimate by the model.

Another noteworthy period is the winter time, where the model produces slightly negative SHF and at the same time overestimates LHF compared to observations. Both deviations about cancel each other, and there does not seem to be an issue with energy closure for the observations.

Snow depth observations taken from within the forest and simulated snow depth for the evergreen conifer tree PFT agree generally very well with each other (Figure 5). The model tends to somewhat underestimate the observations, especially at the time of snow melt, where snow depth is receding, but the differences are small and the comparison favourable, in particular when noting the good agreement in interannual variations. The peak winters with highest values (e.g. the winter 2019/2020) are also well reproduced.

On an annual average basis, modelled NEE shows a small carbon sink with NEE equal to -197 $\mathrm{gCm^{-2}yr^{-1}}$, against a smaller source in magnitude for the measurements of +34 $\mathrm{gCm^{-2}yr^{-1}}$ (values given in Table 3 converted from molar units). GPP is 534 $\mathrm{gCm^{-2}yr^{-1}}$ observation-derived against 927 $\mathrm{gCm^{-2}yr^{-1}}$ for the model. In contrast to the mean, the explained variance, r^2, is not sensitive to the absolute magnitude of the fluxes, and since the phases agree well (Figure 3) it is not surprising that it shows a very high value of 0.87 for GPP at the daily time step. For TER, however, due to the phase shift previously discussed, we find a lower value ($r^2 = 0.69$). The value of r^2 for NEE is smaller than for GPP and TER, as we would expect, because NEE is the difference of two larger fluxes and has therefore a much smaller magnitude. (Let us assume that the true temporal average of NEE is zero, while each, the model and the measurements, reproduce temporal average of GPP equal to temporal average of TER, but with some temporal "noise" added. This noise might be due to measurement or model error, but it is uncorrelated between model and measurements. In this case, we would expect the model-measurement correlation for NEE to be zero. However, correlation between modelled and measured GPP or TER could still be substantial due to coinciding temporal variations shorter than the averaging period.)

The value of the annual r^2 in Table 3 captures exclusively interannual variations and the values are much smaller than derived on the basis of daily averages. It appears that the model only partially reproduces the observation-derived interannual variability, especially for NEE. Note, however, that the number of data points is only 6. RMSE for GPP and NEE on a daily

Table 3. Metrics for the different variables simulated at Sodankylä for the period 1/2016 to 12/2021. RMSE: root-mean square. Units: μmol m^{-2} s^{-1} (Gross primary productivity, GPP; total ecosystem respiration, TER; net ecosystem exchange, NEE), W m^{-2} (latent heat flux, LHF; sensitive heat flux, SHF), cm (snow depth).

Metric	GPP	TER	NEE	LHF	SHF	Snow depth
r^2 daily	0.87	0.69	0.51	0.39	0.50	0.78
r^2 annual	0.49	0.26	0.12	0.36	0.55	0.84
RMSE daily	1.86	0.89	1.55	26.97	41.57	14.57
RMSE annual	1.09	0.52	0.63	16.00	9.63	6.88
observed mean	1.38	1.46	0.09	22.74	28.87	21.82
model mean	2.45	1.92	-0.52	34.64	19.23	15.22

basis is similar to the annual mean GPP, likely due to day-to-day variations in the measurements not being captured by the model.

Over all seasons, the model shows much higher LHF than the measurements, but much lower SHF (Table 3). The difference comes almost entirely from the winter and spring seasons, as noted when discussing Figure 4. r^2 are also significantly smaller for energy than for carbon fluxes (Table 3) due to the seasonally varying model-observation differences, which create differing seasonal cycles between the two. By contrast, snow depth shows a very high r^2 at both the daily and the annual time scale, as it was apparent from Figure 5.

Modelled carbon pools differ substantially from locally derived values: The mean and standard deviation of total soil organic carbon found was 3.70$\pm$0.16 kgC m^{-2}, against a model-based estimate of 38.7 kgC m^{-2}. It appears that the model underestimates the turnover time of the slowest soil carbon pool. The observed above-ground biomass at the site was 37.3 t/ha and against a model estimate of 62.5 t/ha, assuming 50% carbon content of dry mass and 85% of woody biomass above ground (Helmisaari et al., 2002).

4.5 Evaluation at Temperate savannah site: Majadas de Tietar

The seasonal course of carbon exchanges at the temperate-savannah site (Figure 6) is characterised by a pronounced spring-time net carbon uptake and a prolonged period of carbon loss during the summer and autumn. However, the strength of the spring draw-down (ca. DOY 50 to 150) derived from the observations is much lower than the model-derived one. For the remaining seasons, model and observed NEE largely agree in terms of magnitude and timing, except for pronounced fluctuations in the measured NEE flux during summer and autumn that are not reproduced by the model. Such fluctuations are also found in the observation-derived TER flux. The discrepancy in the spring draw-down appears to be the result of a model over-estimate of GPP combined with an under-estimate of TER.

Table 4. Metrics for the different variables simulated at Majadas de Tietar for the period 1/2016 to 12/2021. RMSE: root-mean square. Units: μmol m^{-2} s^{-1} (Gross primary productivity, GPP; total ecosystem respiration, TER; net ecosystem exchange, NEE), W m^{-2} (latent heat flux, LHF; sensitive heat flux, SHF)

Metric	GPP	TER	NEE	LHF	SHF
r^2 daily	0.61	0.47	0.27	0.31	0.81
r^2 annual	0.16	0.26	0.54	0.56	0.66
RMSE daily	2.11	1.56	1.88	30.12	54.37
RMSE annual	1.19	1.23	0.16	7.35	40.49
observed mean	3.39	3.34	-0.05	39.87	40.10
model mean	2.25	2.16	-0.09	32.87	80.49

If we consider the climate of the site, with hot-dry summers, cool winters, and a winter rainfall maximum, we can assume that the most favourable growth conditions happen in the spring, where we indeed find the largest net CO_2 uptake rate in both model and observations. Under those spring-time conditions, however, the model predicts a higher GPP than the observation-based value, but a lower GPP value for the remaining seasons where growth is either temperature and light (winter/autumn) or water (summer/autumn) limited. In other words, compared to the observations, the model over-predicts GPP under favourable conditions, but under-predicts GPP under conditions of GPP limitations — by way of dry conditions in the summer, low light levels in the autumn and temperature during winter. We thus find that the model likely overestimates moisture limitation, and also other GPP limiting factors, while overestimating photosynthetic capacity.

On an annual basis (Table 4), simulated GPP (844 gC m^{-2} $year^{-1}$) is generally smaller compared to the observation derived estimate (1283 gC m^{-2} $year^{-1}$). The same also applies to TER (814 gC m^{-2} $year^{-1}$ vs. 1264 gC m^{-2} $year^{-1}$ derived from eddy covariance measurements). Net flux (mean NEE) is small and agrees well (-30 vs. -19 gC m^{-2} $year^{-1}$ for model vs. observations). In contrast to Sodankylä, r^2 for the annual values shows that the interannual variability of NEE is reproduced well, in fact better than that for the components GPP and TER. However, daily r^2 is much lower for NEE than for GPP or TER, due to the different shape of the seasonal cycle of the model, showing a pronounced spring draw-down, as already discussed. RMSE of GPP and TER on a daily basis are similar in magnitude to the modelled, but less than the observed mean, while annual RMSE for NEE is remarkably low. While the high r^2 suggests that the model reproduces the interannual variability of the net carbon fluxes well for this site, the combination of rather high RMSE and similar observed means suggests that day-to-day variations are less well captured.

As far as the energy balance is concerned, we find a similar result for Majadas de Tietar (Figure 7) compared to the boreal-forest site: net energy input (net radiation radiation minus ground heat flux) agrees very well between model and observations, but there is a rather large over-estimate by the model of the sensible heat flux, albeit with a very similar shape of the seasonal cycle ($r^2 = 0.81$, SHF daily, Table 4). For most of the year, except for a pronounced summer decline for the model but not

for the observations, LHF agrees well. This is likely related to the model's pronounced under-prediction, compared to the observations, of GPP, resulting in a lower transpiration flux through more pronounced stomatal closure, and thus also lower LHF. Given strict energy closure for the model, if net energy input and LHF agree between model and measurements, then SHF should also agree. However, while the model has exact energy closure, the data apparently do not. For instance, at the start of the years until ca. DOY 130, net radiation and LHF agreement suggests an imbalance in the observations starting close to zero at the start of the year and increasing to around 40 Wm^{-2}. During the summer season, the model overestimates SHF by around 100 Wm^{-2}, but underestimates LHF by only around 60 Wm^{-2}, while net radiation agrees, which also suggests a deviation from energy closure of around 40 Wm^{-2}. This has again to be taken into account when evaluating the model.

On average (Table 4), the model slightly underestimates LHF, but overestimates SHF by close to a factor of two. For both LHF and SHF, we find high values for r^2 based on annual averages, and a very small value for annual RMSE for LHF, which suggest that the model, apart from a general overestimate of LHF, simulates interannual variability of energy fluxes reasonably well, with the caveat that only six full years are being considered here.

5 Evaluation of observation operators

5.1 Evaluation of FAPAR simulations

The simulations at the Sodankylä site showed larger FAPAR values during the summer than the observations (Figure 8), with a pronounced seasonal cycle. We find this to be a robust feature of the simulations (not shown). By contrast, observed values stay at approximately the same level during the observation period, with some larger values during the autumn. The values of the observed FAPAR match the expected behaviour of the largely aseasonal evergreen canopies of the PFTs for the boreal region. The pronounced seasonal cycle of FAPAR in the model runs corresponds to a seasonal cycle in the LAI of the model. The modelled LAI behaviour results from calibration using Copernicus LAI time series which have a strong (and unexpected) seasonality. By contrast, measured FAPAR shows only weak signs of seasonality, such as a very slight increase between DOY 170 and 200. There is, however, a cluster of elevated measured FAPAR values towards late summer/ autumn, alternating with lower values in line with those measured earlier. Here we must take into account that maximum solar elevation towards the end of the measurement period (22 October) did not exceed 12^o. Therefore, rays of direct sunshine have a longer path through the canopy, increasing FAPAR. The effect is seen to a lesser extent also in the simulations, but with an LAI driven seasonality dominating the time course of the data.

Extensive LAI sampling during the summer 2022 from hemispherical photographs give an average value of 1.37, and measurements using Li-Cor LAI-2200 an average value 1.32. These agree rather well with a simulated annual-average LAI of 1.3 for the tree PFT, however, there is a pronounced seasonality of simulated LAI corresponding to the seasonality of FAPAR seen in Figure 8, with significantly lower values for September (DOY 244 to 273: 1.37) than for mid June to the end of August (from DOY 166 to 243: 2.96). The across-plot average at different dates from the hemispheric photographs show no such seasonality, with a June to August average (measured on DOY 166, 192, 207, 212 and 217) of 1.39 vs. a September average of 1.34 (measured on DOY 254 and 271).

5.2 Evaluation of SIF simulations

SIF measurements provide an opportunity to document the presence of photosynthetically active plant material, and are therefore an interesting quantity for model validation. At the Sodankylä site, the observations started in spring 2021 as part of the LCC campaign activities. The measurement angle was adjusted in early June and therefore we show comparisons to the simulations starting only from June 3 onwards.

Simulated SIF values are shown here (Figures 9 to 11) with a multiplication factor of 10, i.e. with the scaling factor s_{SIF} in the SIF source term, Equ. 1, set to 10. While the prior value of s_{SIF} was 1, this change reflects the high uncertainty regarding the absolute magnitude of the measured SIF. Observations are shown for two methods of retrieving SIF from the actual measurements, namely Frauenhofer line discrimination and spectral fitting (see Section S3.5).

The difference in magnitude between the modelled and observed SIF is likely due to the choice of prior parameters for the SIF model, taken from Gu et al. (2019), and the specific spectral conversion used (Equ. 2). Although it has not been done here, there is scope within D&B to adjust these parameters in the assimilation. This is achieved through calibration of the scaling factor s_{SIF} in the observation operator for SIF (Equ. 1). Given that the model with its prior parameter set can already track the seasonal and diurnal cycle of the observations reasonably well, we expect that SIF measurements will be able to provide constraints on processes that affect the temporal evolution of photosynthetic rates, such as leaf phenology, or timing of stomatal closure.

At the Sodankylä site, the simulations are able to track both the diurnal and seasonal cycles of the observations reasonably well (Figure 9). However, there are indications of water stress in the measured diurnal cycles in June, July and August. These are shown in as a dip in far-red SIF during midday (Figure 9a), and for June also in red SIF (Figure 9b). The decline in SIF is likely due to midday depression of photosynthesis (Lin et al., 2024). The model reproduces this behaviour only for June and to a much lesser extent. Also, the model shows larger SIF signals for June compared to July, but not the measurements. Since mid-day depression is observed as a response to stomatal closure due to water stress, the comparison indicates that the model underestimates water stress at the boreal site. The simulations also show an earlier increase and later decrease during the day during the summer months. This may partly be attributed to retrieval problems for high sun zenith angles.

The measured far-red SIF (760 nm) of the trees at Majadas de Tietar site (PFT 3, Figure 10a) shows a clear seasonal cycle of SIF peaking in July. For red SIF (687 nm, Figure 10b), there is no clear seasonal maximum. This is independent of the retrieval method. The model by contrast shows a clear peak in May. The diurnal cycle of SIF in the model peaks later, usually around 2 pm, and extends further into the afternoon compared to the measurements, which peak around 11am to 12pm (central vertical line).

By contrast, the SIF measurements on grass (PFT 9) show almost complete senescence of the grass during June and July when using the spectral fitting method, but some remaining activity when using Frauenhofer line discrimination of the red-spectrum signal (Figure 11). For this combination, red SIF with Frauenhofer line discrimination, model simulations are in good agreement with the measurements, with suitable scaling factor s_{SIF} in the SIF source term (Equ. 1). However, judging from

the other spectral bands or retrieval methods, the results suggest that the model may under-estimate the water stress of the grasses.

5.3 Evaluation of VOD simulations

Figure 12 shows the comparison between observed and simulated L-VOD for the period after the first change in measurement geometry, for all three elevation angles, for the Sodankylä site. Observations only include the trees, and therefore simulated L-band VOD is for the tree PFT only. The temporal variations of the measurements are well captured by the simulation, in particular after the second change in viewing geometry after DOY 280.

The increase in biomass in the field of view through the first change in measurement geometry on 17 September (DOY 260) was estimated to be a factor of three (see Section S3.6). The revised field of view was also found to better represent typical conditions of the wider area, with the initial field of view capturing the signal from much sparser vegetation. To simulate L-VOD for the period before DOY 260, we therefore reduce assumed biomass entering the VOD observation operator (C_{wd}, C_{fol} in Equ. 4) to one third of their default modelled values.

With this provision, the simulations match both temporal variations and magnitude of the locally measured L-VOD rather well (Figure 13, see Section S3.6 for the default simulations). This includes the rise in spring, including a peak around DOY 60, and also temporal variations between DOY 90 and 130. Only ca. DOY 130 to 180 shows a systematic overestimate compared to measurements. A slow decline after DOY 220 is also reproduced by the model. We thus find a very satisfactory performance of the empirical L-VOD observation operator together with D&B.

5.4 Evaluation of surface soil moisture

Measured soil moisture at Sodankylä (Figure 14) shows very similar temporal variations between different depths. The temporal variations of the D&B simulations are also similar, only that the overall magnitude differs, even though the magnitude of the shallowest measured depth (5 cm) is closest to the model. We point out that the depth of the surface layer in D&B is 4 cm. Both measurements and simulations also indicate significant interannual variability, with some years (e.g. 2019, 2020) exhibiting some pronounced summer drying, of which only some is captured by the measurements due to data gaps.

At Majadas de Tietar, variations in soil moisture measured between different depths are again relatively small, showing that the exact depth for which these are simulated is of lesser importance (Figure 15). In fact, the two depths closest to the surface (5 and 10 cm depth) show an almost identical temporal profile, including the maximum soil moisture depletion during the summer (July, August). The main characteristics of the observed seasonal cycle are also well reproduced by the D&B model. The timing of individual rain events can be well traced in the measurements, and is well reproduced by the simulations, including the lack of such events during the summer months. However, in the simulations, soil moisture decreases to near zero, whereas according to the measurements some soil moisture remains even at the peak of the summer.

6 Discussion

6.1 Implications of study results

The comparison of the model simulations at the two sites against local data indicates that D&B does a reasonable job at representing energy and carbon fluxes between the atmosphere and terrestrial vegetation, albeit with the seasonal amplitude of the net carbon exchange overestimated at the boreal site. The comparison shows that carbon fluxes in particular are simulated reasonably well, with lesser agreement for energy fluxes but also significant imbalances between the measured energy fluxes and the net radiation available to the canopy, i.e. there is a significant deviation from energy closure. We conclude that there is a need for multi data sources to be used for evaluating carbon and water flux models of terrestrial ecosystems, as opposed to relying chiefly on eddy covariance data.

The addition of dedicated observation operators led to further insights regarding model performance. In particular, local SIF measurements further revealed the power of those measurements to detect limitations on photosynthesis, such as water stress, beyond the capability of FAPAR measurements, and this despite remaining uncertainties regarding the absolute magnitude of the simulated SIF signal (see Figure 9). We were able to identify a possible under-estimate of soil water limitation of the Scots Pine forest at Sodankylä during the summer, which may partly explain why the model overestimates GPP at this site.

At the Majadas de Tietar temperate savannah site, we clearly identified that the model underestimates latent heat flux during the summer months, while it also underestimated the site's overall photosynthetic uptake (GPP). This underestimation appears to be a result in particular of the model over-estimating moisture limitations of the savanna ecosystems during the summer, possibly due to non-matching parameterisation of the stomatal model. This matched the insights provided by the SIF measurements that the trees of the ecosystem continue transpiring and photosynthesising across the summer without major limitations due to water stress. Surface soil moisture data also indicated too much soil drying during the summer months. Possibly, the model fails to represent the strongly heterogeneous soil texture at Majadas del Tietar, with a sandy top and deeper clay soil, underestimating the soil water holding capacity of the deeper soil layers to which only the trees have access. These considerations demonstrate the added value of the dedicated observation operators for the evaluation of D&B at the local scale.

6.2 Potential for further applications

The process model in combination with its observation operators presented here has been designed to be used within a variational data assimilation framework, planned to be set up following the existing CCDAS (Rayner et al., 2005). This means that D&B will be complemented with tangent and adjoint versions, which efficiently provide derivative information for variational assimilation. The anticipated default setup in data assimilation mode is for combined calibration and initialisation, i.e. adjustment of parameters of the process model, its observation operators and of the initial state of the carbon pools. Assimilated data streams are planned to come chiefly from Earth Observation sources. This setup will provide both capabilities for assimilating more data streams than previous studies (e.g. Scholze et al. (2019)), while also including a full description of biomass pools as so far provided by other, more complex process models, e.g. LPJ-GUESS (Smith et al., 2001), albeit with lesser data assimilation capabilities.

In anticipation of such an application, we have in this contribution refrained from adjusting individual parameters "by hand" in order to improve the match to any of the validation data sets used in Sections 4 and 5. However, we can already assess, to an initial degree, the potential of the system to obtain a superior fit to measurements by way of optimising its parameters. As an example, comparing measured and simulated surface soil moisture (Section 5.4) and taking into account model's functional dependencies, we can infer that changing the assumed texture of the soil near its surface will immediately change the absolute magnitude of the simulated signal, but have only a negligible impact (via soil evaporation) on its temporal course.

In particular the good match between simulated and locally measured L-VOD, which includes details of most temporal variations, offers considerable opportunities for the assimilation of widely available satellite-derived L-VOD over larger regions.

6.3 Limitations

While the initial task to match and compare modelled and observed data streams was successfully demonstrated, the results of this study also point at the need to further investigate the representation of the seasonal cycle of LAI in northern evergreen conifer forests and shrubs. Earth observation products for the boreal region show seasonality in LAI that is not consistent with ecological expectation and FAPAR data. The phenology scheme of D&B has the flexibility to simulate vegetation with a small amount of seasonal variation in LAI, if corresponding information is provided for the prior calibration of the parameters in the phenology scheme. Such information could come from field observations of LAI time series in boreal regions or improved satellite products.

Another issue that occurred is that the scaling factor s_{SIF} in the SIF source term (Equ. 1) is highly uncertain. In a data assimilation mode, it would be included (possibly in PFT-specific form) into the list of parameters to be adjusted. This would effectively allow scaling the simulated SIF time series shown in Figs. 9, 10 and 11. Similarly, the parameters in the empirical observation operator for VOD would also be included in the set of parameters to be adjusted in assimilation mode. We also note that many of the model's parameters are not very well constrained, and could therefore change substantially. For example, an adjustment of the turnover times for the litter and soil organic matter pools will change heterotrophic respiration, and according to Equ. S153 the fit to simulated TER shown in Figure 3 and Figure 6. This could happen in the framework of either a local-scale assimilation of eddy flux measurements as used here for evaluation, or on a regional to global scale with the assimilation of atmospheric CO_2 data, including those from space-based remote sensing (Buchwitz et al., 2017).

A further limitation we found is that the model overestimates soil water limitation at the savannah site. This may be linked to the parameterisation of soil hydrological properties, or to the parameterisation of rooting depth and root penetration of the soil, both of which warrant further investigation. We also find that the model may overestimate soil evaporation for very dry soils.

A principal advantage of the process-based modelling approach presented here is that the system can be used to identify, better investigate and quantify specific processes – a fundamental and often decisive advantage over machine learning, or complex statistical modelling systems (Thessen, 2016; Lary et al., 2018). The advantage translates into a principal limitation in that if a given process contributes to the measured signal, it has to be represented. Otherwise, missing process representation can lead to misleading parameter choices that use processes included within the system to compensate for the missing process

– also known as "matching observations for the wrong reasons". Therefore, process modelling requires significant expert knowledge on ecosystem functioning as well as experience with or direct contact to experimental teams, compared to statistical interference methods, including machine learning – which by definition can never be "right for the wrong reasons", as they are used essentially as black boxes. The potential advantage of D&B coupled to multiple observation operators is that it allows model testing via multiple data streams, thus providing are more comprehensive model evaluation which makes it less likely the model matches observations while misrepresenting important processes.

6.4 Outlook

In this study, we have shown the value of the four data streams (FAPAR, SIF, VOD and surface soil moisture), as opposed to the intrinsically local measurements used for the initial evaluation, which lies in their availability over large spatial scales. Therefore, such data streams derived from Earth Observation sources will make it possible to evaluate the model across larger, regional scales. The immediate next step would therefore be to evaluate D&B with regional rather than local observations and see if in such a set-up the noted model-observation differences are reduced. The advantage of regional comparisons is that substantial uncertainties arising from small-scale conditions are averaged out, and the scale of comparison may by more appropriate for a typical application of the model.

At such a regional scale, it will further be possible to assimilate those data streams using the principle setup described above. Here, it will be possible to either adjust parameters spatially grouped by PFT, following CCDAS (Kaminski et al., 2013), or independently at each pixel following DALEC (Quaife et al., 2007). A further approach that has not yet been tested would be a combination of the two, where parameters are adjusted at every grid cell independently, but with a partial constraint on parameter values assuming that those values co-vary depending on closeness of geographical location, altitude, land use, PFT, or soil type.

A significant advantage of such a data assimilation system will be the possibility to investigate if the process model is capable of matching the observations not only for a specific parameter set, but within reasonable bounds of the entire model-parameter space. Only if that is not possible can we rigorously conclude that the remaining model-observation mismatch is caused by missing or unsuitable process representation. We consider such investigations the next logical step of development of the D&B modelling system, besides any inclusion of further Earth Observation data streams. The further goal would then be to apply it to the task of routinely producing data products on carbon and energy fluxes.

Code and data availability. The D&B code in Fortran90 is hosted, with simulation results, at the Zenodo repository under the AGPL license, available through https://zenodo.org/doi/10.5281/zenodo.11243753 and is also available, with updates, from its repository via https://gitlab.gwdg.de/tccas-team/TCCAS.git. The observations are available on the TCCAS home page https://tccas.inversion-lab.com/database.html and have been permanently archived at https://zenodo.org/doi/10.5281/zenodo.12725764

Author contributions. WK drafted and finalized manuscript. MD developed the layout of the activity combining modelling, field work and EO. WK and MW provided original model design and description of D&B. WK, MW, TK, MV, TLS, MS, TG, TT, TLQ, SZ further designed and validated D&B. TK, MV, and WK implemented D&B and integrated module for carbon balance and phenology. MV and TK performed model simulations. AK, JL, MA, MH, TG, TSeM provided field data. MV, AO, and TM prepared data and plots. TLS and MW drafted sections on carbon balance and phenology and provided module for carbon balance and phenology. TLQ drafted sections on SIF observation operator. All authors contributed to the discussions on analysis of results or commented on or edited manuscript.

Competing interests. At least one of the (co-)authors is a member of the editorial board of Geoscientific Model Development.

Acknowledgements. This work was funded by ESA through contracts 4000131497 (LCC study) and 4000141232 (TCCAS study) within the Carbon Science Cluster and by the European Union's Horizon Europe Programme through Grant Agreement 101082194. T. Thum and T. Miinalainen were funded by the Research Council of Finland (grant number 330165). Y. H. Kerr was funded by the Centre National d'Etudes Spatiales through the Terre Ocean Surface Continentales Atmosphere (TOSCA) programme. T. S. El-Madany thanks the Alexander von Humboldt Stiftung for financial support of the MaNiP project and Arnaud Carrara, Ramon Lopez-Jiminez, Gerardo Moreno, Victor Rolo, and M. Pilar Martín and her SpecLab team for support of fieldwork. J.Verelst and P. Reyez-Muñoz were funded by the European Union (ERC, FLEXINEL, 101086622). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them. All authors thank the two anonymous referees for there helpful comments and suggestions.

References

Arora, V. K., Katavouta, A., Williams, R. G., Jones, C. D., Brovkin, V., Friedlingstein, P., Schwinger, J., Bopp, L., Boucher, O., Cadule, P., et al.: Carbon–concentration and carbon–climate feedbacks in CMIP6 models and their comparison to CMIP5 models, Biogeosciences, 17, 4173–4222, 2020.

Babaeian, E., Sadeghi, M., Jones, S. B., Montzka, C., Vereecken, H., and Tuller, M.: Ground, proximal, and satellite remote sensing of soil moisture, Reviews of Geophysics, 57, 530–616, 2019.

Berger, M., Moreno, J., Johannessen, J. A., Levelt, P. F., and Hanssen, R. F.: ESA's sentinel missions in support of Earth system science, Remote Sensing of Environment, 120, 84–90, 2012.

Best, M. J., Pryor, M., Clark, D., Rooney, G. G., Essery, R., Ménard, C., Edwards, J., Hendry, M., Porson, A., Gedney, N., et al.: The Joint UK Land Environment Simulator (JULES), model description–Part 1: energy and water fluxes, Geoscientific Model Development, 4, 677–699, 2011.

Bloom, A. A. and Williams, M.: Constraining ecosystem carbon dynamics in a data-limited world: Integrating ecological "common sense" in a model-data fusion framework, Biogeosciences, 12, 1299–1315, https://doi.org/10.5194/bg-12-1299-2015, 2015.

Boussetta, S., Balsamo, G., Beljaars, A., Panareda, A.-A., Calvet, J.-C., Jacobs, C., van den Hurk, B., Viterbo, P., Lafont, S., Dutra, E., et al.: Natural land carbon dioxide exchanges in the ECMWF integrated forecasting system: Implementation and offline validation, Journal of Geophysical Research: Atmospheres, 118, 5923–5946, 2013.

Buchwitz, M., Reuter, M., Schneising, O., Hewson, W., Detmers, R. G., Boesch, H., Hasekamp, O. P., Aben, I., Bovensmann, H., Burrows, J. P., et al.: Global satellite observations of column-averaged carbon dioxide and methane: The GHG-CCI XCO2 and XCH4 CRDP3 data set, Remote Sensing of Environment, 203, 276–295, 2017.

Ciais, P., Dolman, A. J., Bombelli, A., Duren, R., Peregon, A., Rayner, P. J., Miller, C., Gobron, N., Kinderman, G., Marland, G., et al.: Current systematic carbon-cycle observations and the need for implementing a policy-relevant carbon observing system, Biogeosciences, 11, 3547–3602, 2014.

Clark, D., Mercado, L., Sitch, S., Jones, C., Gedney, N., Best, M., Pryor, M., Rooney, G., Essery, R., Blyth, E., et al.: The Joint UK Land Environment Simulator (JULES), model description–Part 2: carbon fluxes and vegetation dynamics, Geoscientific Model Development, 4, 701–722, 2011.

Disney, M., Muller, J.-P., Kharbouche, S., Kaminski, T., Voßbeck, M., Lewis, P., and Pinty, B.: A new global fAPAR and LAI dataset derived from optimal albedo estimates: Comparison with MODIS products, Remote Sensing, 8, 275, 2016.

El-Madany, T. S., Reichstein, M., Perez-Priego, O., Carrara, A., Moreno, G., Martín, M. P., Pacheco-Labrador, J., Wohlfahrt, G., Nieto, H., Weber, U., et al.: Drivers of spatio-temporal variability of carbon dioxide and energy fluxes in a Mediterranean savanna ecosystem, Agricultural and Forest Meteorology, 262, 258–278, 2018.

Famiglietti, C. A., Smallman, T. L., Levine, P. A., Flack-Prain, S., Quetin, G. R., Meyer, V., and et al.: Optimal model complexity for terrestrial carbon cycle prediction, Journal of Advances in Modeling Earth Systems, 18, 2727–2754, https://doi.org/10.5194/bg-18-2727-2021, 2021.

Farquhar, G., von Caemmerer, S. v., and Berry, J.: A biochemical model of photosynthetic CO2 assimilation in leaves of C3 species, Planta, 149, 78–90, 1980.

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., et al.: Global carbon budget 2022, Earth System Science Data, 14, 4811–4900, 2022.

Gao, H., Liu, S., Lu, W., Smith, A. R., Valbuena, R., Yan, W., Wang, Z., Xiao, L., Peng, X., Li, Q., et al.: Global analysis of the relationship between reconstructed solar-induced chlorophyll fluorescence (SIF) and gross primary production (GPP), Remote Sensing, 13, 2824, 2021.

Gerten, D., Hoff, H., Bondeau, A., Lucht, W., Smith, P., and Zaehle, S.: Contemporary "green" water flows: Simulations with a dynamic global vegetation and water balance model, Physics and Chemistry of the Earth, Parts A/B/C, 30, 334–338, 2005.

Gu, L., Han, J., Wood, J. D., Chang, C. Y.-Y., and Sun, Y.: Sun-induced Chl fluorescence and its importance for biophysical modeling of photosynthesis based on light reactions, New Phytologist, 223, 1179–1191, 2019.

Harper, A. B., Cox, P. M., Friedlingstein, P., Wiltshire, A. J., Jones, C. D., Sitch, S., Mercado, L. M., Groenendijk, M., Robertson, E., Kattge, J., et al.: Improved representation of plant functional types and physiology in the Joint UK Land Environment Simulator (JULES v4. 2) using plant trait information, Geoscientific Model Development, 9, 2415–2440, 2016.

Heiskanen, J., Brümmer, C., Buchmann, N., Calfapietra, C., Chen, H., Gielen, B., Gkritzalis, T., Hammer, S., Hartman, S., Herbst, M., et al.: The integrated carbon observation system in Europe, Bulletin of the American Meteorological Society, 103, E855–E872, 2022.

Helmisaari, H.-S., Makkonen, K., Kellomäki, S., Valtonen, E., and Mälkönen, E.: Below-and above-ground biomass, production and nitrogen use in Scots pine stands in eastern Finland, Forest ecology and management, 165, 317–326, 2002.

Honkanen, M., Heikkinen, P., MacArthur, A., Thum, T., Kivi, R., and Lindqvist, H.: UAV-Borne Measurements of Solar-Induced Chlorophyll Fluorescence (SIF) at a Boreal Site, in: International conference on FinDrones, pp. 115–135, Springer, 2023.

Jarvis, P. G. and McNaughton, K.: Stomatal control of transpiration: scaling up from leaf to region, Advances in Ecological Research, 15, 1–49, 1986.

Kaminski, T. and Mathieu, P.-P.: Reviews and syntheses: Flying the satellite into your model: on the role of observation operators in constraining models of the Earth system and the carbon cycle, Biogeosciences, 14, 2343–2357, https://doi.org/10.5194/bg-14-2343-2017, 2017.

Kaminski, T., Knorr, W., Scholze, M., Gobron, N., Pinty, B., Giering, R., and Mathieu, P.-P.: Consistent assimilation of MERIS FAPAR and atmospheric CO_2 into a terrestrial vegetation model and interactive mission benefit analysis, Biogeosciences, 9, 3173–3184, https://doi.org/10.5194/bg-9-3173-2012, 2012.

Kaminski, T., Knorr, W., Schürmann, G., Scholze, M., Rayner, P. J., Zaehle, S., Blessing, S., Dorigo, W., Gayler, V., Giering, R., Gobron, N., Grant, J. P., Heimann, M., Hooker-Strout, A., Houweling, S., Kato, T., Kattge, J., Kelley, D., Kemp, S., Koffi, E. N., Köstler, C., Mathieu, P., Pinty, B., Reick, C. H., Rödenbeck, C., Schnur, R., Scipal, K., Sebald, C., Stacke, T., van Scheltinga, A. T., Vossbeck, M., Widmann, H., and Ziehn, T.: The BETHY/JSBACH Carbon Cycle Data Assimilation System: experiences and challenges, J. Geophys. Res., 118, doi:10.1002/jgrg.20 118, http://onlinelibrary.wiley.com/doi/10.1002/jgrg.20118/abstract, 2013.

Kato, T., Knorr, W., Scholze, M., Veenendaal, E., Kaminski, T., Kattge, J., and Gobron, N.: Simultaneous assimilation of satellite and eddy covariance data for improving terrestrial water and carbon simulations at a semi-arid woodland site in Botswana, Biogeosciences, 10, 789–802, https://doi.org/10.5194/bg-10-789-2013, 2013.

Kerr, Y., Waldteufel, P., Wigneron, J.-P., Delwart, S., Cabot, F., Boutin, J., Escorihuela, M.-J., Font, J., Reul, N., Gruhier, C., Juglea, S., Drinkwater, M., Hahne, A., Martin-Neira, M., and Mecklenburg, S.: The SMOS Mission: New Tool for Monitoring Key Elements ofthe Global Water Cycle, Proceedings of the IEEE, 98, 666–687, https://doi.org/10.1109/JPROC.2010.2043032, 2010.

Kerr, Y. H., Waldteufel, P., Richaume, P., Wigneron, J. P., Ferrazzoli, P., Mahmoodi, A., Al Bitar, A., Cabot, F., Gruhier, C., Juglea, S. E., et al.: The SMOS soil moisture retrieval algorithm, IEEE transactions on geoscience and remote sensing, 50, 1384–1403, 2012.

Kirdiashev, K., Chukhlantsev, A., and Shutko, A.: Microwave radiation of the earth's surface in the presence of vegetation cover, Radiotekhnika i Elektronika, 24, 256–264, 1979.

Knorr, W.: Satellitengestützte Fernerkundung und Modellierung des Globalen CO_2 -Austauschs der Landvegetation: Eine Synthese, Ph.D. thesis, Max-Planck-Institut für Meteorologie, Hamburg, Germany, 1997.

Knorr, W.: Annual And Interannual CO_2 Exchanges Of The Terrestrial Biosphere: Process-Based Simulations And Uncertainties, Glob. Ecol. Biogeogr., 9, 225–252, 2000.

Knorr, W. and Kattge, J.: Inversion of terrestrial biosphere model parameter values against eddy covariance measurements using Monte Carlo sampling, Global Change Biology, 11, 1333–1351, 2005.

Knorr, W. and Lakshmi, V.: Assimilationof fAPAR and Surface Temperature Into a Land Surface and Vegetation Model, in: Land Surface Hydrology, Meteorology, and Climate: Observations and Modeling, edited by Lakshmi, V., Albertson, J., and Schaake, J., pp. 177 – 200, American Geophysical Union, ISBN 0875903525, 2001.

Knorr, W., Kaminski, T., Scholze, M., Gobron, N., Pinty, B., Giering, R., and Mathieu, P.-P.: Carbon cycle data assimilation with a generic phenology model, J. Geophys. Res., 115, https://doi.org/10.1029/2009JG001119, 2010.

Konings, A. G., Rao, K., and Steele-Dunne, S. C.: Macro to micro: microwave remote sensing of plant water content for physiology and ecology, New Phytologist, 223, 1166–1172, https://doi.org/https://doi.org/10.1111/nph.15808, _eprint: https://nph.onlinelibrary.wiley.com/doi/pdf/10.1111/nph.15808, 2019.

Lary, D. J., Zewdie, G. K., Liu, X., Wu, D., Levetin, E., Allee, R. J., Malakar, N., Walker, A., Mussa, H., Mannino, A., et al.: Machine learning applications for earth observation, Earth observation open science and innovation, 165, 2018.

Lin, J., Zhou, L., Wu, J., Han, X., Zhao, B., Chen, M., and Liu, L.: Water stress significantly affects the diurnal variation of solar-induced chlorophyll fluorescence (SIF): A case study for winter wheat, Science of The Total Environment, 908, 168 256, 2024.

Liu, Z., Zhao, F., Liu, X., Yu, Q., Wang, Y., Peng, X., Cai, H., and Lu, X.: Direct estimation of photosynthetic CO2 assimilation from solar-induced chlorophyll fluorescence (SIF), Remote Sensing of Environment, 271, 112 893, 2022.

Loth, B. and Graf, H.: Modelling the snow cover for climate studies, Tech. Rep. 190, Max-Planck-Institut für Meteorologie, Hamburg, Germany, 1996.

Magney, T. S., Frankenberg, C., Köhler, P., North, G., Davis, T. S., Dold, C., Dutta, D., Fisher, J. B., Grossmann, K., Harrington, A., Hatfield, J., Stutz, J., Sun, Y., and Porcar-Castell, A.: Disentangling Changes in the Spectral Shape of Chlorophyll Fluorescence: Implications for Remote Sensing of Photosynthesis, Journal of Geophysical Research: Biogeosciences, 124, 1491–1507, https://doi.org/https://doi.org/10.1029/2019JG005029, 2019.

Mo, T., Choudhury, B. J., Schmugge, T. J., Wang, J. R., and Jackson, T. J.: A model for microwave emission from vegetation-covered fields, Journal of Geophysical Research: Oceans, 87, 11 229–11 237, https://doi.org/https://doi.org/10.1029/JC087iC13p11229, 1982.

Mohammed, G. H., Colombo, R., Middleton, E. M., Rascher, U., van der Tol, C., Nedbal, L., Goulas, Y., Pérez-Priego, O., Damm, A., Meroni, M., et al.: Remote sensing of solar-induced chlorophyll fluorescence (SIF) in vegetation: 50 years of progress, Remote sensing of environment, 231, 111 177, 2019.

Nair, R. K., Morris, K. A., Migliavacca, M., Moreno, G., and Schrumpf, M.: Plant-available N: P alters root litter N recycling in a Mediterranean tree–grass ecosystem, Journal of Plant Nutrition and Soil Science, 183, 517–529, 2020.

NOAA: Trends in Atmospheric Carbon Dioxide, Global Monitoring Laboratory, https://gml.noaa.gov/webdata/ccgg/trends/co2/co2_annmean_mlo.txt, accessed 1 Feb 2024, 2024.

Norton, A. J., Rayner, P. J., Koffi, E. N., and Scholze, M.: Assimilating solar-induced chlorophyll fluorescence into the terrestrial biosphere model BETHY-SCOPE: Model description and information content, Geoscientific Model Development Discussions, 2017, 1–26, https://doi.org/10.5194/gmd-2017-34, 2017.

Norton, A. J., Rayner, P. J., Koffi, E. N., and Scholze, M.: Assimilating solar-induced chlorophyll fluorescence into the terrestrial biosphere model BETHY-SCOPE v1.0: model description and information content, Geoscientific Model Development, 11, 1517–1536, https://doi.org/10.5194/gmd-11-1517-2018, 2018.

Norton, A. J., Rayner, P. J., Koffi, E. N., Scholze, M., Silver, J. D., and Wang, Y.-P.: Estimating global gross primary productivity using chlorophyll fluorescence and a data assimilation system with the BETHY-SCOPE model, Biogeosciences, 16, 3069–3093, https://doi.org/10.5194/bg-16-3069-2019, 2019.

Peylin, P., Bacour, C., MacBean, N., Leonard, S., Rayner, P., Kuppel, S., Koffi, E., Kane, A., Maignan, F., Chevallier, F., et al.: A new stepwise carbon cycle data assimilation system using multiple data streams to constrain the simulated land surface carbon cycle, Geoscientific Model Development, 9, 3321–3346, 2016.

Porcar-Castell, A., Tyystjärvi, E., Atherton, J., Van der Tol, C., Flexas, J., Pfündel, E. E., Moreno, J., Frankenberg, C., and Berry, J. A.: Linking chlorophyll a fluorescence to photosynthesis for remote sensing applications: mechanisms and challenges, Journal of experimental botany, 65, 4065–4095, 2014.

Quaife, T., Lewis, P., Disney, M., De Kauwe, M., Williams, M., and Law, B.: Assimilating reflectance data into a ecosystem model to improve estimates of terrestrial carbon flux, in: 10th International Symposium on Physical Measurements and Signatures in Remote Sensing, publisher Copyright: © 2007 International Society for Photogrammetry and Remote Sensing. All rights reserved., 2007.

Quegan, S., Le Toan, T., Chave, J., Dall, J., Exbrayat, J.-F., Minh, D. H. T., Lomas, M., D'Alessandro, M. M., Paillou, P., Papathanassiou, K., Rocca, F., Saatchi, S., Scipal, K., Shugart, H., Smallman, T. L., Soja, M. J., Tebaldini, S., Ulander, L., Villard, L., and Williams, M.: The European Space Agency BIOMASS mission: Measuring forest above-ground biomass from space, Remote Sensing of Environment, 227, 44–60, https://doi.org/https://doi.org/10.1016/j.rse.2019.03.032, 2019.

Raoult, N. M., Jupp, T. E., Cox, P. M., and Luke, C. M.: Land-surface parameter optimisation using data assimilation techniques: the adJULES system V1. 0, Geoscientific Model Development, 9, 2833–2852, 2016.

Rayner, P., Scholze, M., Knorr, W., Kaminski, T., Giering, R., and Widmann, H.: Two decades of terrestrial Carbon fluxes from a Carbon Cycle Data Assimilation System (CCDAS), Global Biogeochemical Cycles, 19, doi:10.1029/2004GB002 254, 2005.

Rebmann, C., Aubinet, M., Schmid, H., Arriga, N., Aurela, M., Burba, G., Clement, R., De Ligne, A., Fratini, G., Gielen, B., et al.: ICOS eddy covariance flux-station site setup: a review, International Agrophysics, 32, 471–494, 2018.

Running, S. W. and Zhao, M.: Daily GPP and annual NPP (MOD17A2/A3) products NASA Earth Observing System MODIS land algorithm, MOD17 User's Guide, 2015, 1–28, 2015.

Scholze, M., Kaminski, T., Knorr, W., Blessing, S., Vossbeck, M., Grant, J., and Scipal, K.: Simultaneous assimilation of SMOS soil moisture and atmospheric CO_2 in-situ observations to constrain the global terrestrial carbon cycle, Remote Sensing of Environment, 180, 334–345, 2016.

Scholze, M., Buchwitz, M., Dorigo, W., Guanter, L., and Quegan, S.: Reviews and syntheses: Systematic Earth observations for use in terrestrial carbon cycle data assimilation systems, Biogeosciences, 14, 3401–3429, 2017.

Scholze, M., Kaminski, T., Knorr, W., Voßbeck, M., Wu, M., Ferrazzoli, P., Kerr, Y., Mialon, A., Richaume, P., Rodríguez-Fernández, N., Vittucci, C., Wigneron, J.-P., Mecklenburg, S., and Drusch, M.: Mean European Carbon Sink Over 2010-2015 Estimated by Simultane-

ous Assimilation of Atmospheric CO2, Soil Moisture, and Vegetation Optical Depth, Geophysical Research Letters, 46, 13 796–13 803, https://doi.org/10.1029/2019GL085725, 2019.

Schwank, M., Kontu, A., Mialon, A., Naderpour, R., Houtz, D., Lemmetyinen, J., Rautiainen, K., Li, Q., Richaume, P., Kerr, Y., and Mätzler, C.: Temperature effects on L-band vegetation optical depth of a boreal forest, Remote Sensing of Environment, 263, 112 542, https://doi.org/https://doi.org/10.1016/j.rse.2021.112542, 2021.

Scipal, K., Drusch, M., and Wagner, W.: Assimilation of a ERS scatterometer derived soil moisture index in the ECMWF numerical weather prediction system, Advances in water resources, 31, 1101–1112, 2008.

Sellers, P. J.: Canopy reflectance, photosynthesis and transpiration, International Journal of Remote Sensing, 6, 1335–1372, https://doi.org/10.1080/01431168508948283, 1985.

Smallman, T. L., Exbrayat, J.-F., Mencuccini, M., Bloom, A. A., and Williams, M.: Assimilation of repeated woody biomass observations constrains decadal ecosystem carbon cycle uncertainty in aggrading forests, Journal of Geophysical Research: Biogeosciences, 122, 528–545, https://doi.org/https://doi.org/10.1002/2016JG003520, 2017.

Smith, B., Prentice, I. C., and Sykes, M. T.: Representation of vegetation dynamics in the modelling of terrestrial ecosystems: comparing two contrasting approaches within European climate space, Global ecology and biogeography, pp. 621–637, 2001.

Sun, J., Yue, Y., and Niu, H.: Evaluation of NPP using three models compared with MODIS-NPP data over China, PLoS One, 16, e0252 149, 2021.

Swinnen, E., Van Hoolst, R., and Toté, C.: Copernicus Global Land Operations" Vegetation and Energy", Issue I1, 12, 2021.

Tarantola, A.: Inverse Problem Theory and methods for model parameter estimation, SIAM, Philadelphia, 2005.

Thessen, A.: Adoption of machine learning techniques in ecology and earth science, One Ecosystem, 1, e8621, 2016.

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: Chemical and Physical Meteorology, 59, 874–890, 2007.

Traore, A. K., Ciais, P., Vuichard, N., Poulter, B., Viovy, N., Guimberteau, M., Jung, M., Myneni, R., and Fisher, J. B.: Evaluation of the ORCHIDEE ecosystem model over Africa against 25 years of satellite-based water and carbon measurements, Journal of Geophysical Research: Biogeosciences, 119, 1554–1575, 2014.

Turner, D. P., Ritts, W. D., Cohen, W. B., Gower, S. T., Running, S. W., Zhao, M., Costa, M. H., Kirschbaum, A. A., Ham, J. M., Saleska, S. R., et al.: Evaluation of MODIS NPP and GPP products across multiple biomes, Remote sensing of environment, 102, 282–292, 2006.

Ulaby, F. T. and Wilson, E. A.: Microwave attenuation properties of vegetation canopies, IEEE Transactions on Geoscience and Remote Sensing, pp. 746–753, 1985.

Vuichard, N., Messina, P., Luyssaert, S., Guenet, B., Zaehle, S., Ghattas, J., Bastrikov, V., and Peylin, P.: Accounting for carbon and nitrogen interactions in the global terrestrial ecosystem model ORCHIDEE (trunk version, rev 4999): Multi-scale evaluation of gross primary production, Geoscientific Model Development, 12, 4751–4779, 2019.

Wang, J., Xiao, X., Wagle, P., Ma, S., Baldocchi, D., Carrara, A., Zhang, Y., Dong, J., and Qin, Y.: Canopy and climate controls of gross primary production of Mediterranean-type deciduous and evergreen oak savannas, Agricultural and forest meteorology, 226, 132–147, 2016.

Wigneron, J.-P., Kerr, Y., Waldteufel, P., Saleh, K., Escorihuela, M.-J., Richaume, P., Ferrazzoli, P., de Rosnay, P., Gurney, R., Calvet, J.-C., Grant, J., Guglielmetti, M., Hornbuckle, B., Mätzler, C., Pellarin, T., and Schwank, M.: L-band Microwave Emission of the Biosphere (L-MEB) Model: Description and calibration against experimental data sets over crop fields, Remote Sensing of Environment, 107, 639–655, https://doi.org/https://doi.org/10.1016/j.rse.2006.10.014, 2007.

Wigneron, J.-P., Chanzy, A., Kerr, Y. H., Lawrence, H., Shi, J., Escorihuela, M. J., Mironov, V., Mialon, A., Demontoux, F., De Rosnay, P., et al.: Evaluating an improved parameterization of the soil emission in L-MEB, IEEE Transactions on Geoscience and Remote Sensing, 49, 1177–1189, 2010.

Williams, M., Schwarz, P. A., Law, B. E., Irvine, J., and Kurpius, M. R.: An improved analysis of forest carbon dynamics using data assimilation, Global Change Biology, 11, 89–105, 2005.

Wood, E., Lettenmaier, D., and Zartarian, V.: A land-surface hydrology parameterization with subgrid variability for general circulation models, J. Geophys. Res., 97, 2717–2728, 1992.

Zhao, M., Heinsch, F. A., Nemani, R. R., and Running, S. W.: Improvements of the MODIS terrestrial gross and net primary production global data set, Remote sensing of Environment, 95, 164–176, 2005.

Zobler, L.: A World Soil File for Global Climate Modelling, NASA Tech. memo 87802, 32, 1986.

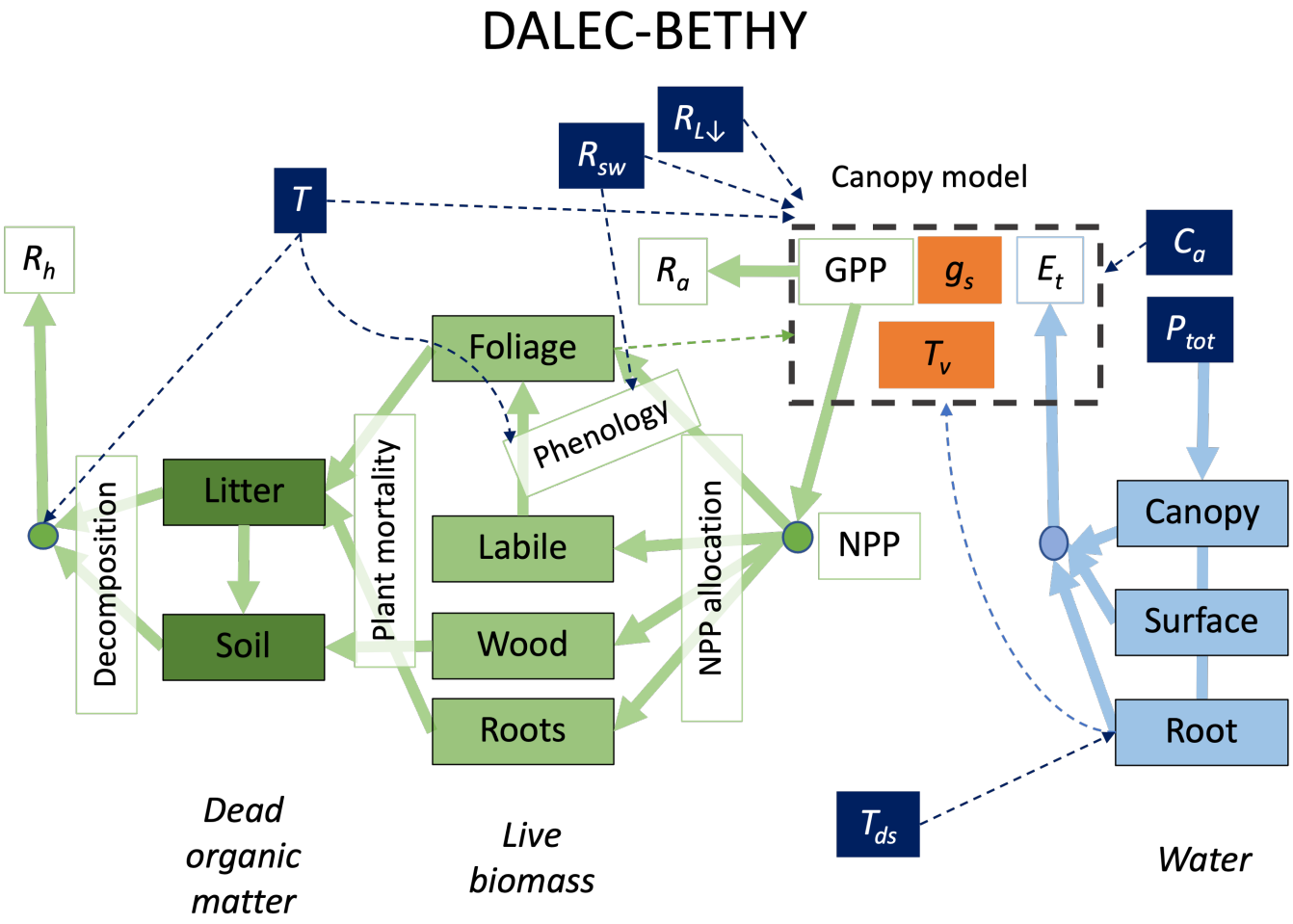

Figure 1. Structure of the DALEC-BETHY coupled model. State variables are in filled boxes; green for carbon, pale blue for water, orange for canopy states of stomatal conductance (g_s) and leaf temperature (T_v). Drivers are shown with white text in dark blue boxes (T: air temperature, T_{ds}: deep-soil temperature, R_{sw}: downwelling shortwave radiation, $R_{L\downarrow}$: downwelling longwave radiation; P_{tot}: total precipitation, C_a: CO_2 concentration in air). Fluxes are shown as solid arrows and annotated by open boxes; coloured green for C fluxes and pale blue for water. GPP is gross primary production; NPP is net primary production; R_A is autotrophic respiration; R_H is heterotrophic respiration; E_t is evapotranspiration. Dashed arrows show influences – for example T and R_{sw} influence the modelling of phenology.

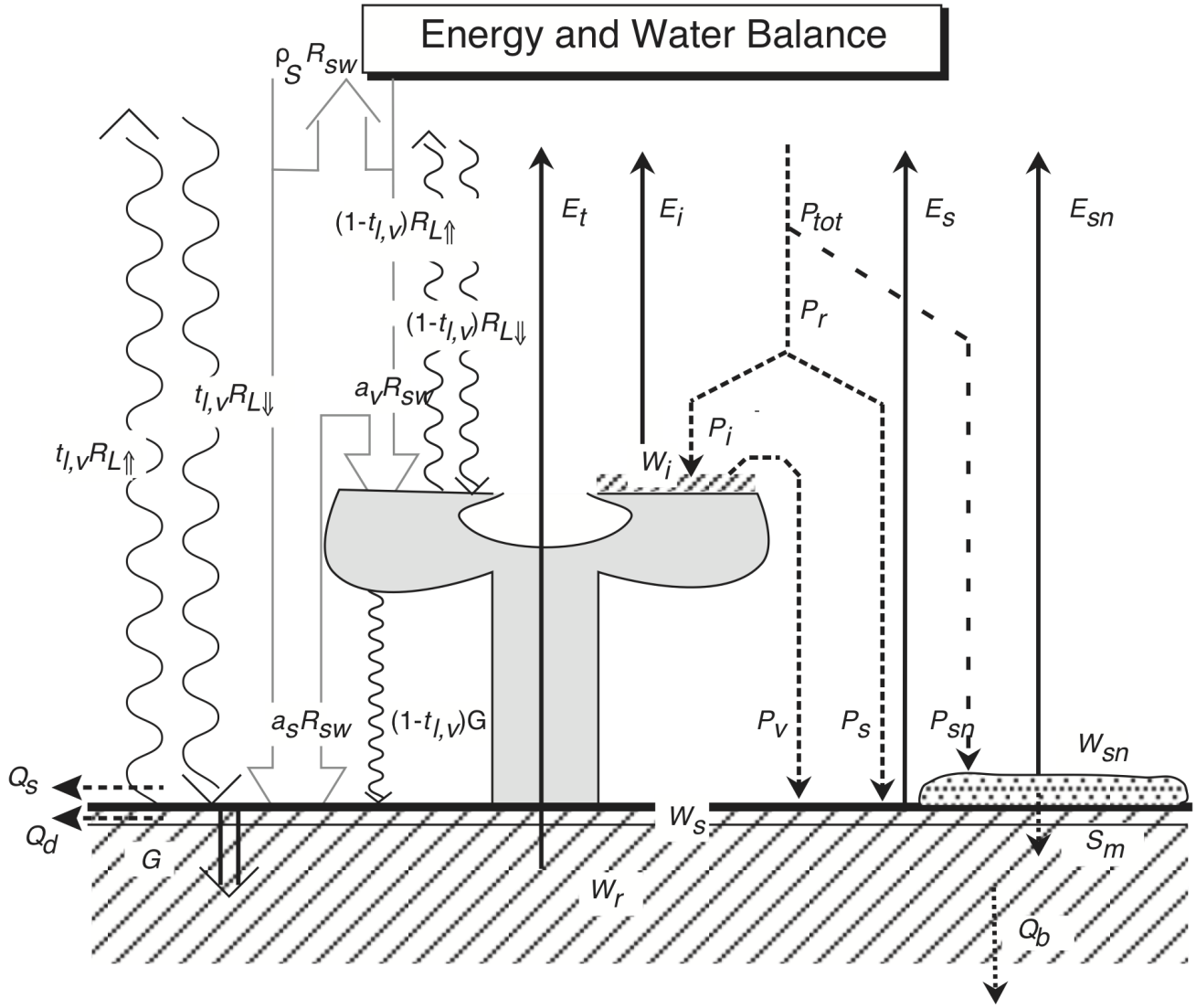

Figure 2. Energy and water balance of the D&B model with symbols for fluxes (normal) and reservoirs (italics). a_s: soil absorption of shortwave radiation, a_v: canopy absorption of shortwave radiation, E_i: intercepted-water (canopy) evaporation, E_s: soil evaporation; E_{sn}: snow evaporation, E_t: transpiration, G: ground heat flux, P_i: intercepted rainfall; P_r: rainfall; P_s rainfall on soil; P_{sn}: snowfall, P_{tot}: total precipitation; P_v: throughfall, Q_b: baseflow, Q_d: horizontal drainage, Q_s: surface runoff, $R_{L,\uparrow}$: upwelling longwave radiation, $R_{L,\downarrow}$: downwelling longwave radiation, R_{sw}: downwelling shortwave radiation, S_m: snow melt, $t_{l,v}$: longwave canopy transmission, W_i: intercepted water amount, W_r: root-zone soil moisture, W_s: surface-layer soil moisture, W_{sn}: snow amount, ρ_S: surface reflectance.

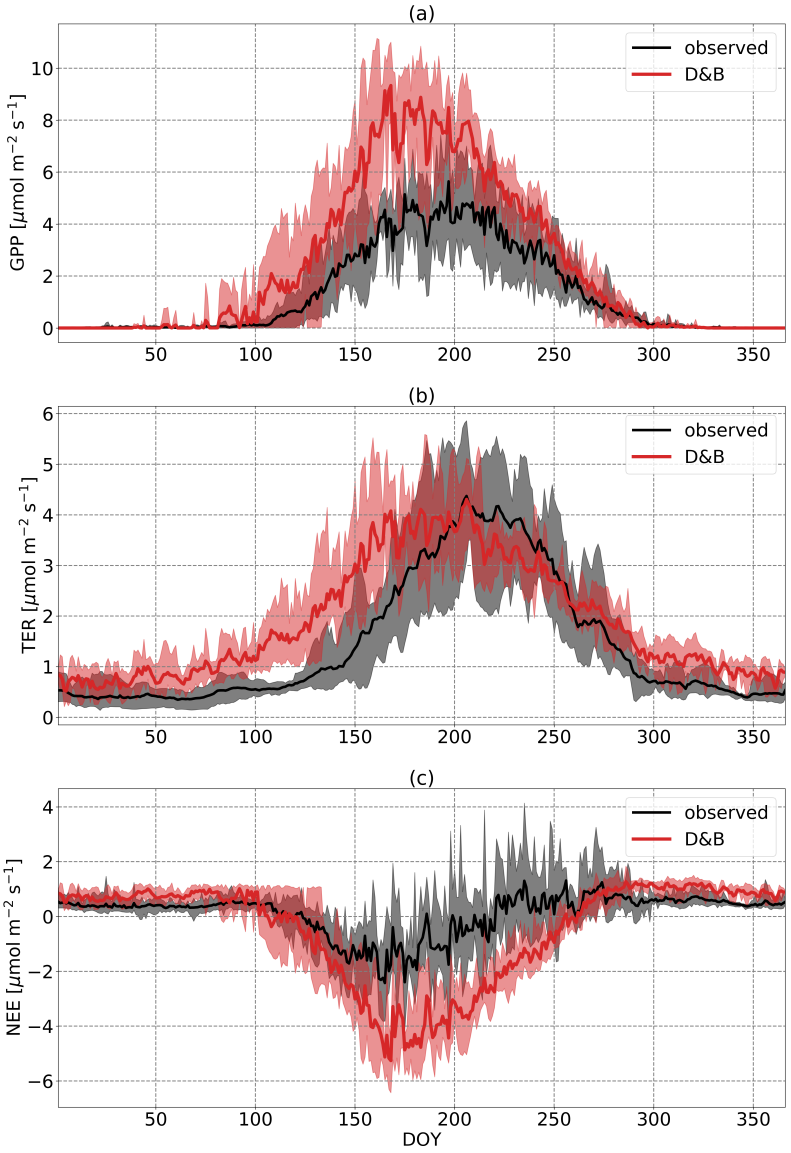

Figure 3. Annual cycles of daily (a) GPP, (b) TER, and (c) NEE at Sodankylä, averaged over the years 2016 to 2021. Black line is observation based on eddy-covariance data, the red line D&B. The shaded areas represent the ranges of the observed and simulated daily cycles over the period.

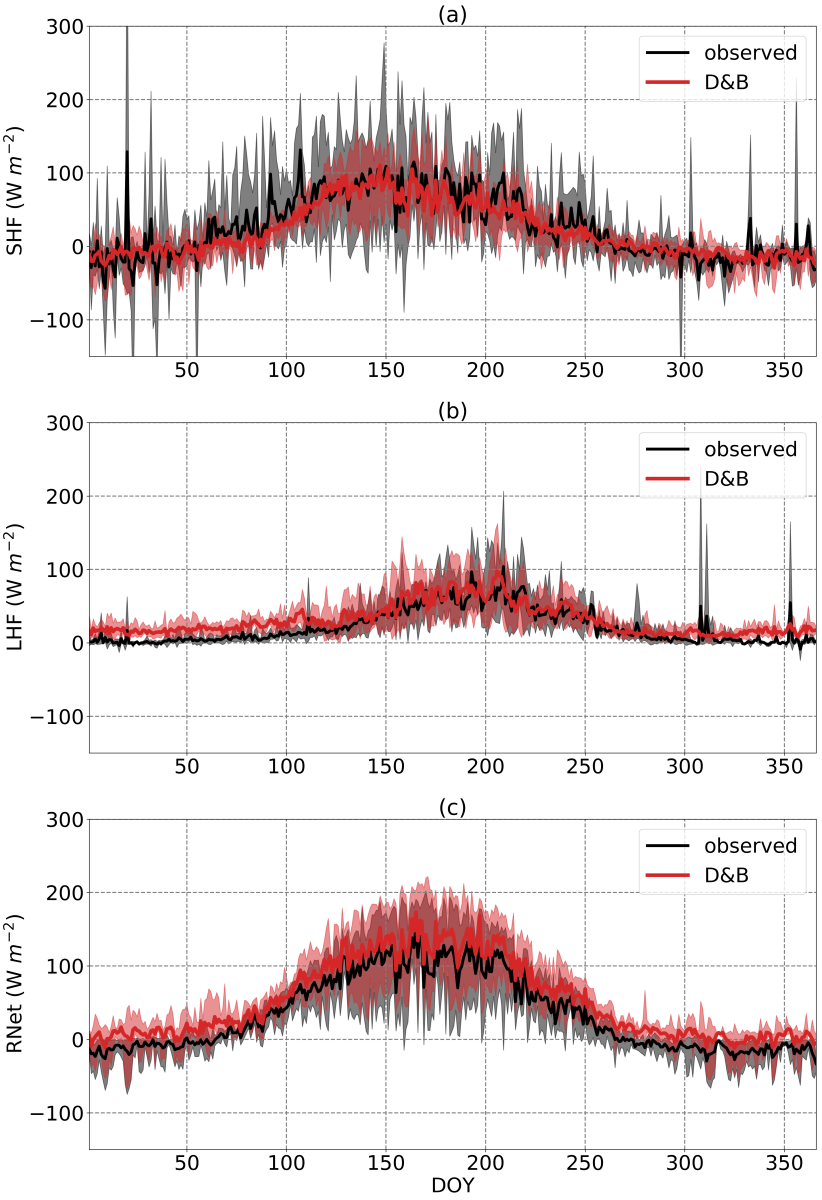

Figure 4. Annual cycles of daily (a) sensible heat flux, (b) latent heat flux, and (c) net radiation minus ground heat flux at Sodankylä, averaged over the years 2016 to 2021. The shaded areas represent the ranges of the observed and simulated daily cycles over the period.

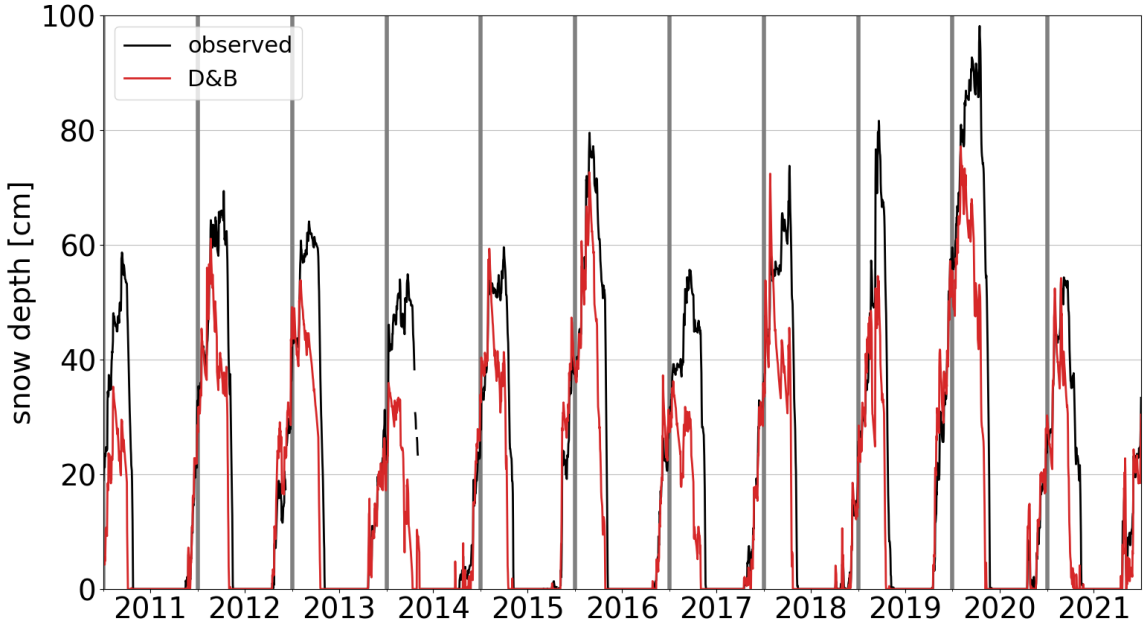

Figure 5. Daily observed (black) and simulated (red) snow depth at Sodankylä for the years 2011-2021. Simulated snow depth is for the evergreen conifer PFT only.

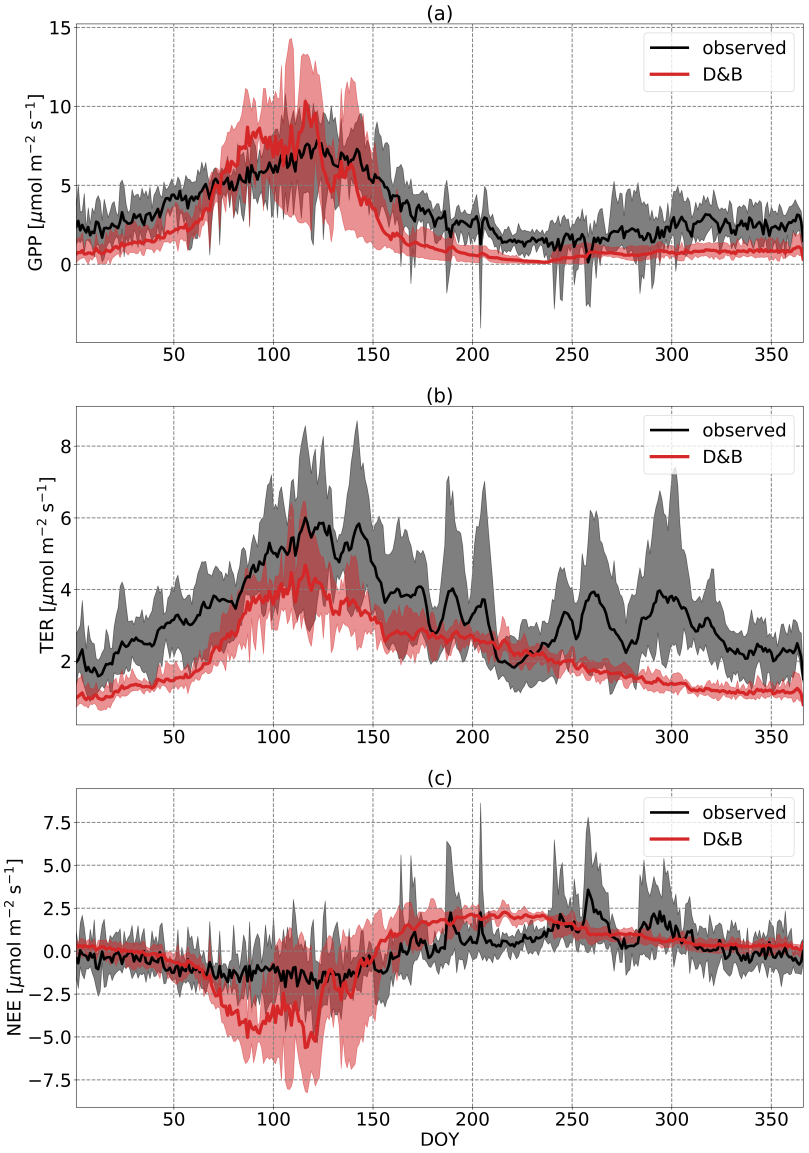

Figure 6. Annual cycles of daily (a) GPP, (b) TER, and (c) NEE at Majadas de Tietar, averaged over the years 2016 to 2021. Black line is observation based on eddy-covariance data, the red line D&B. The shaded areas represent the ranges of the observed and simulated daily cycles over the period.

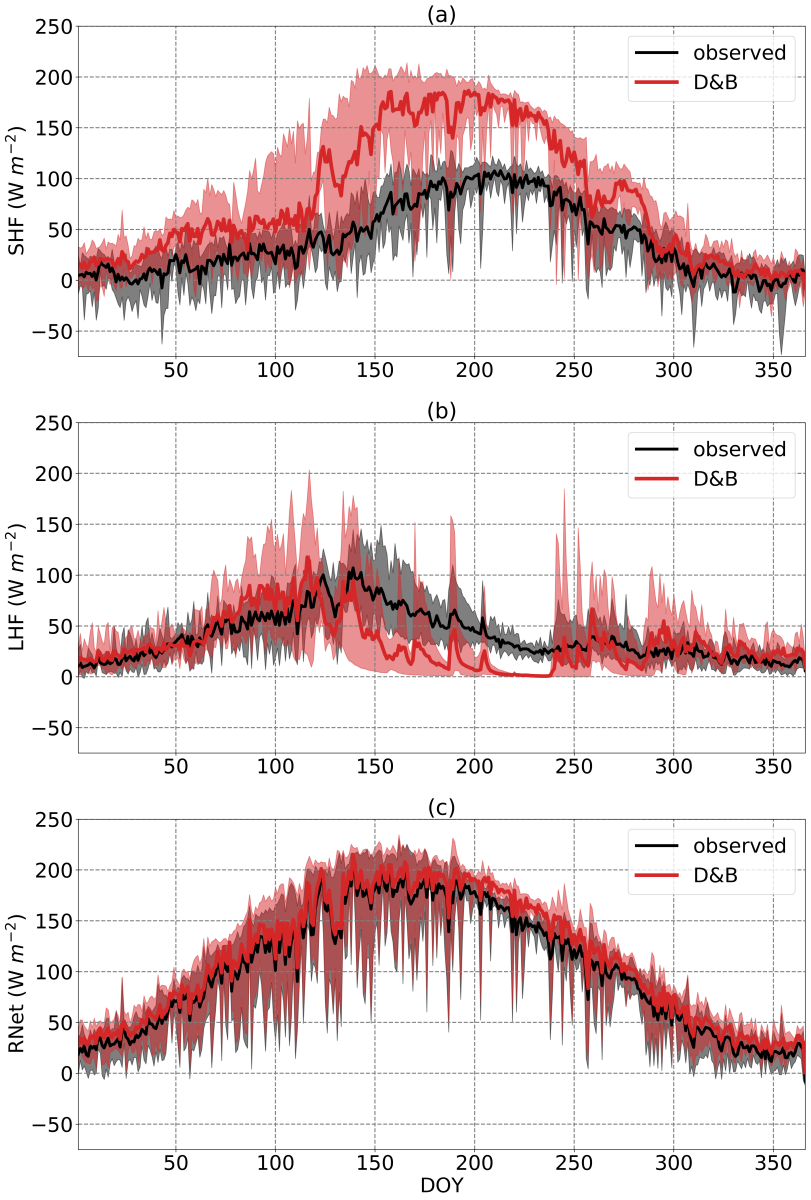

Figure 7. Annual cycles of daily (a) sensible heat flux, (b) latent heat flux, and (c) net radiation minus ground heat flux at Majadas de Tietar, averaged over the years 2016 to 2021. The shaded areas represent the ranges of the observed and simulated daily cycles over the period.

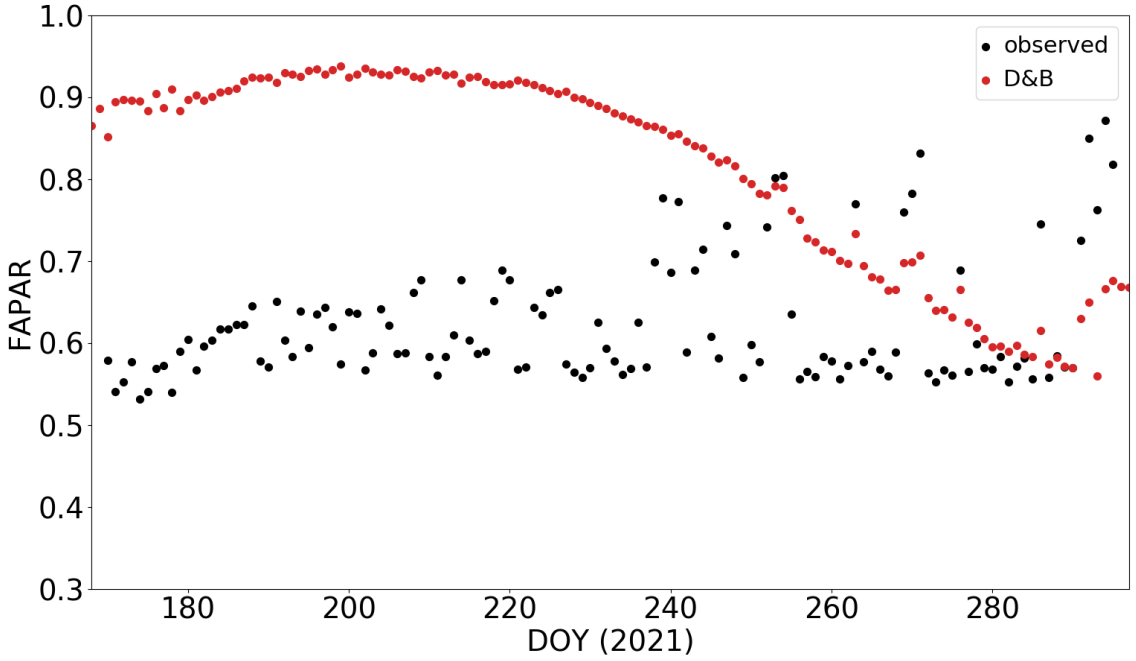

Figure 8. Observed (black) and simulated (red) FAPAR at Sodankylä between 19 June and 22 October 2021. Simulated FAPAR is for the evergreen conifer PFT only, in accordance with the vegetation within the field of view of the FAPAR sensor.

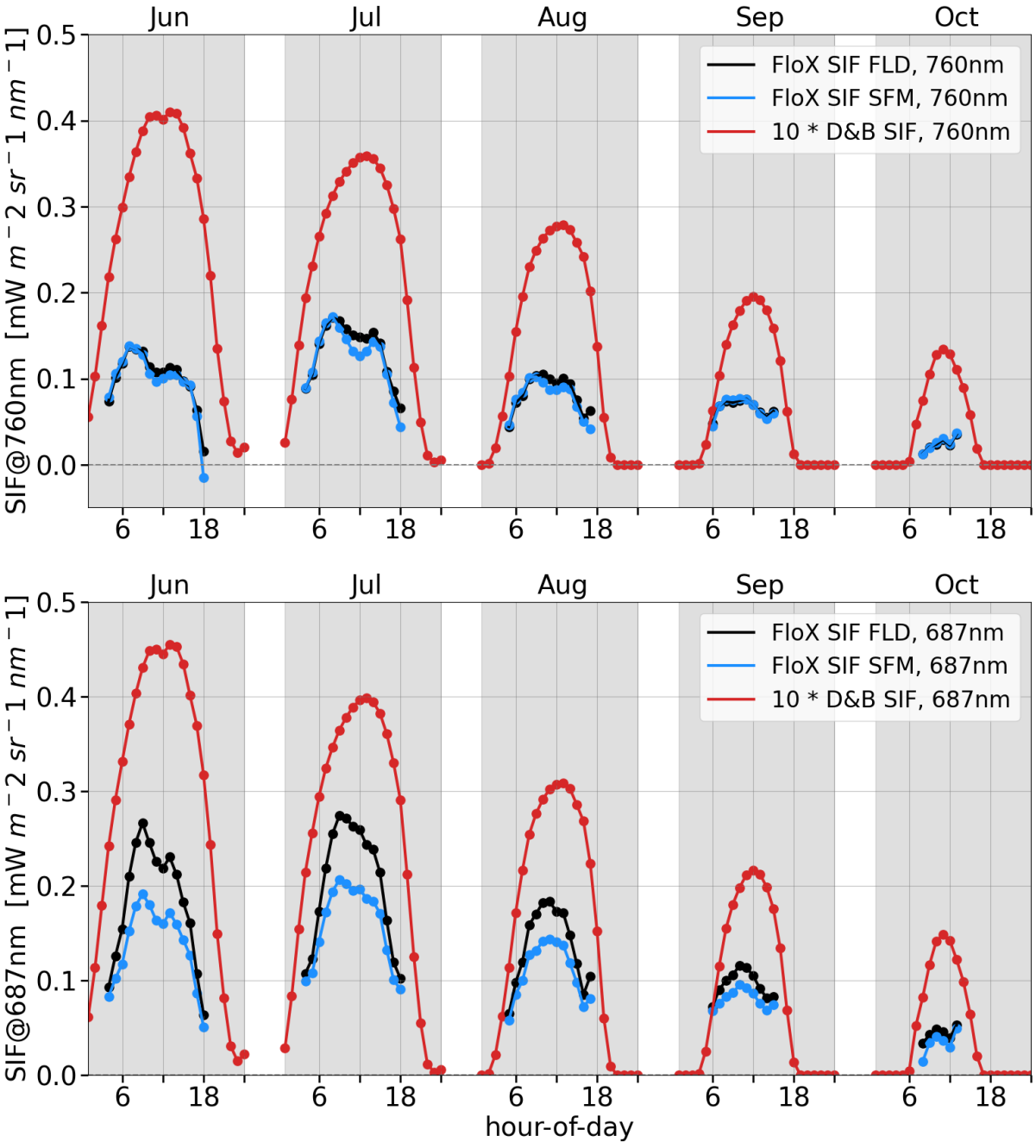

Figure 9. Average diurnal cycle by month of far-red (upper panel) and red SIF (lower panel) for pine forest (PFT 5) at Sodankylä for months June to October in 2021. D&B simulations (red) against measurements with the Fluorescence Box (FloX): retrievals made with the Frauenhofer line discrimination (black) and retrieval made with the spectral fitting method (blue).

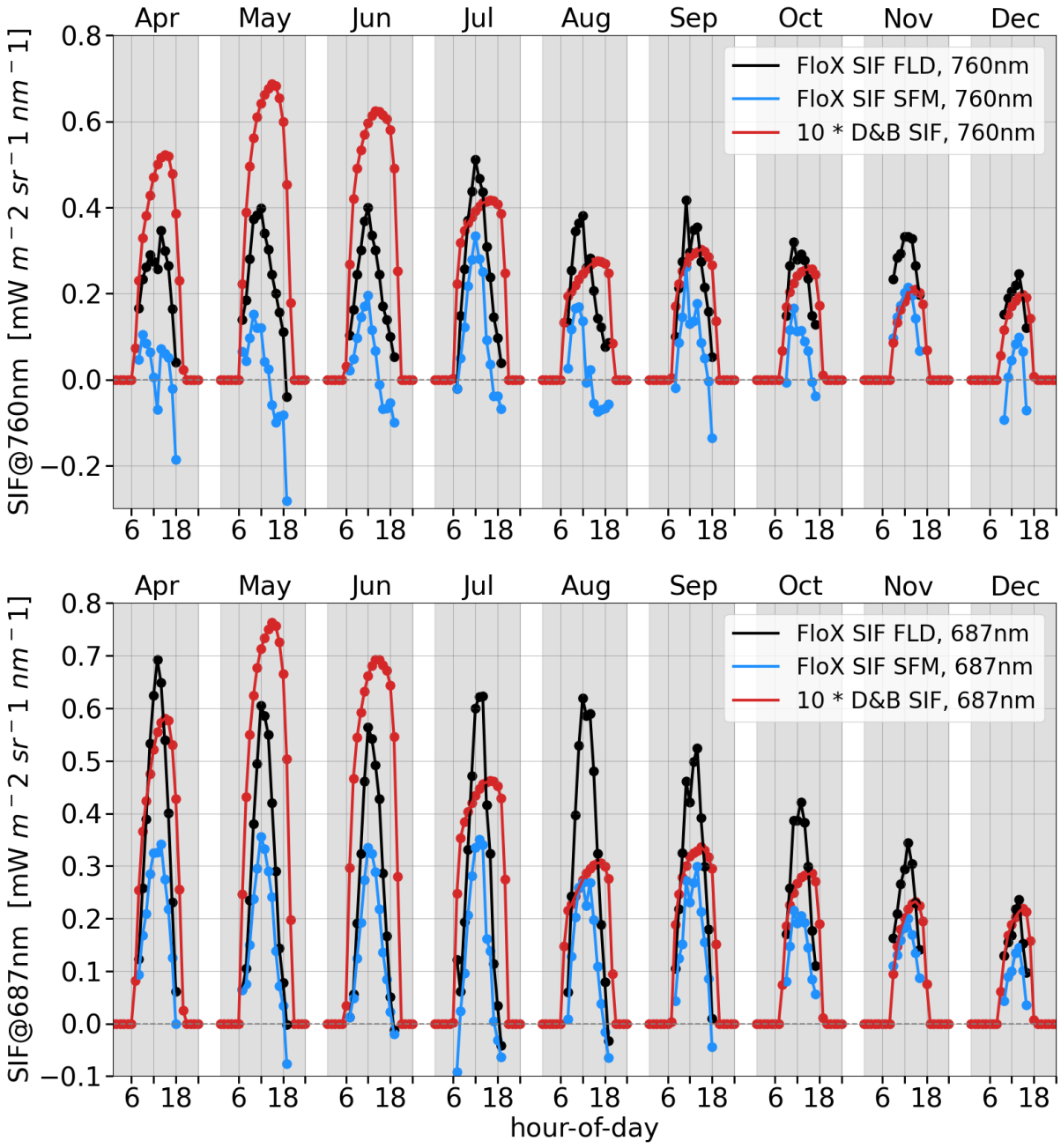

Figure 10. Average hourly diurnal cycle by month of SIF in the far-red (upper panel) and red (lower panel) for evergreen trees (PFT 3) at Majadas de Tietar for months April to December in 2021. D&B simulations (red) against measurements: retrievals made with Frauenhofer line discrimination (black) and spectral fitting method (blue).

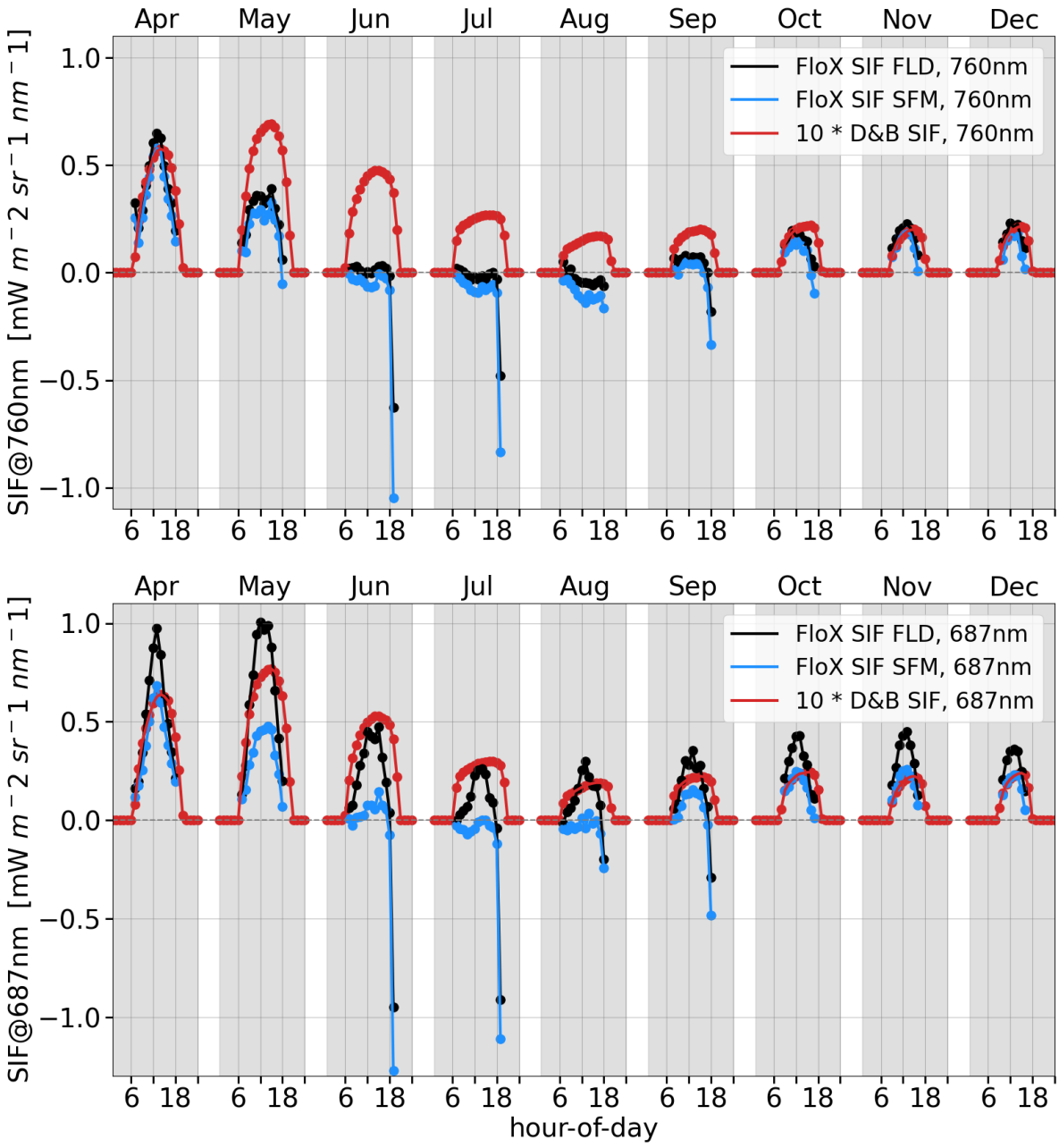

Figure 11. Average hourly diurnal cycle by month of SIF in the far-red (upper panel) and red (lower panel) for C3 grass (PFT 9) at Majadas de Tietar for months April to December in 2021. D&B simulations (red) against measurements: retrievals made with Frauenhofer line discrimination (black) and spectral fitting method (blue).

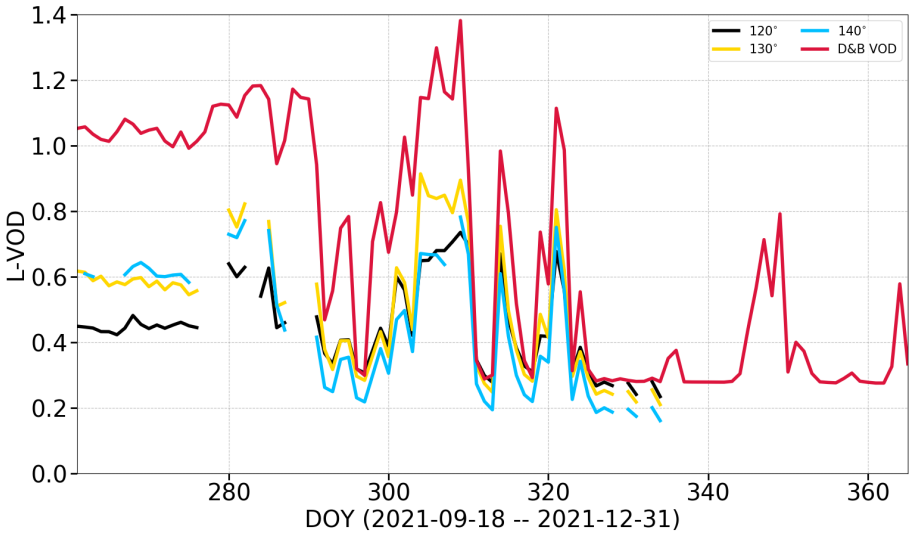

Figure 12. L-band VOD from Elbara II over a pine stand (PFT 5) for different elevation angles compared to D&B simulated L-band VOD, for PFT 5 only. Time axis starts on 18 September 2021 when the azimuth angle of the Elbara II instrument was changed for the first time.

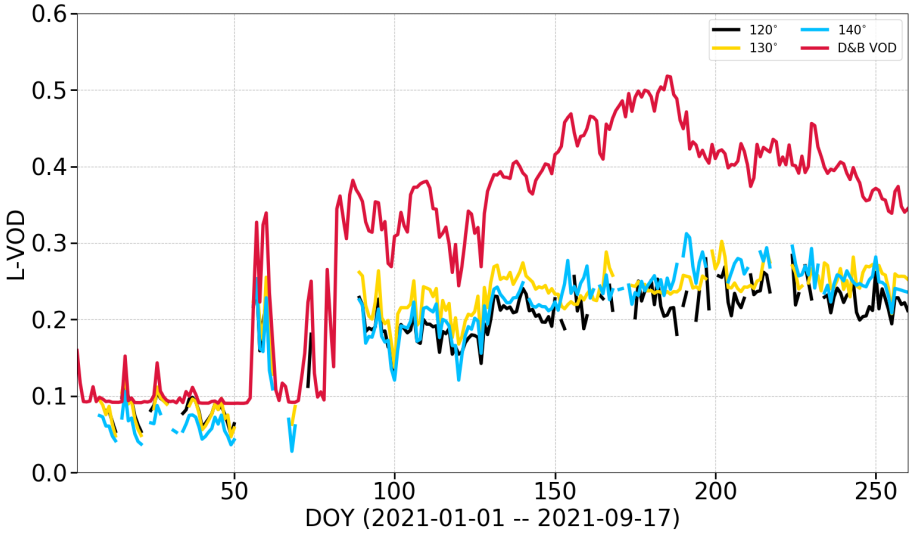

Figure 13. L-band VOD from Elbara II over a pine stand (PFT 5) for different elevation angles compared to D&B simulated L-band VOD, for PFT 5 only, before first change in view geometry, with biomass reduced to 1/3 of default simulated values.

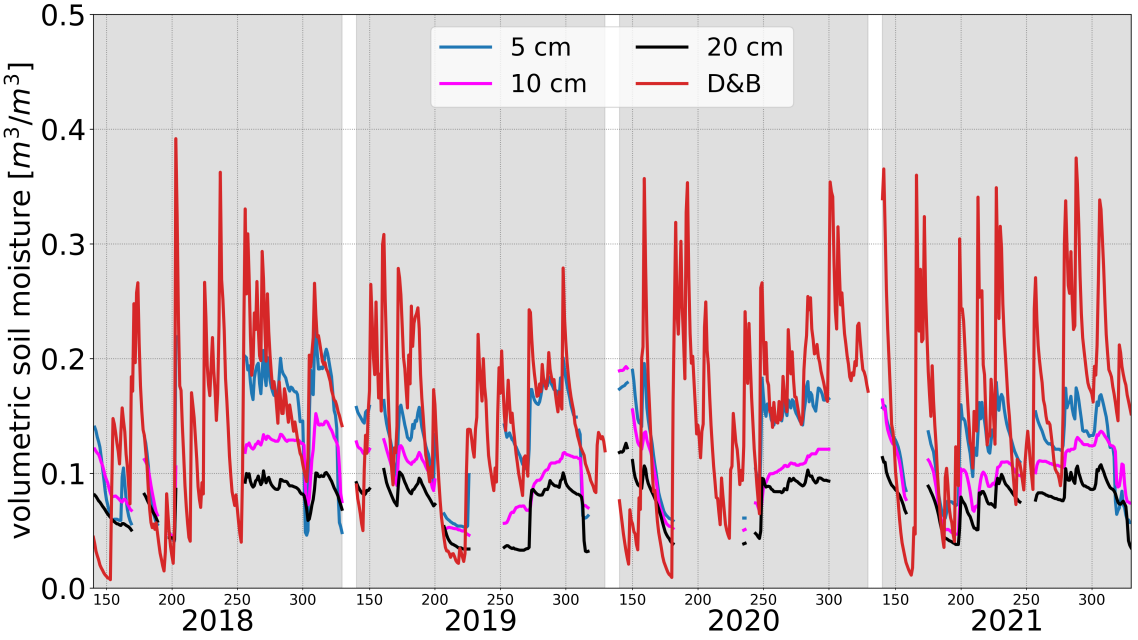

Figure 14. Daily *in situ* observed soil moisture at 5cm, 10cm and 20cm depth, and simulated near-surface soil moisture, W_s, at Sodankylä for Day-Of-Year 140—330 in the years 2018–2021. Winter months not shown to avoid impact of frost on soil moisture sensors.

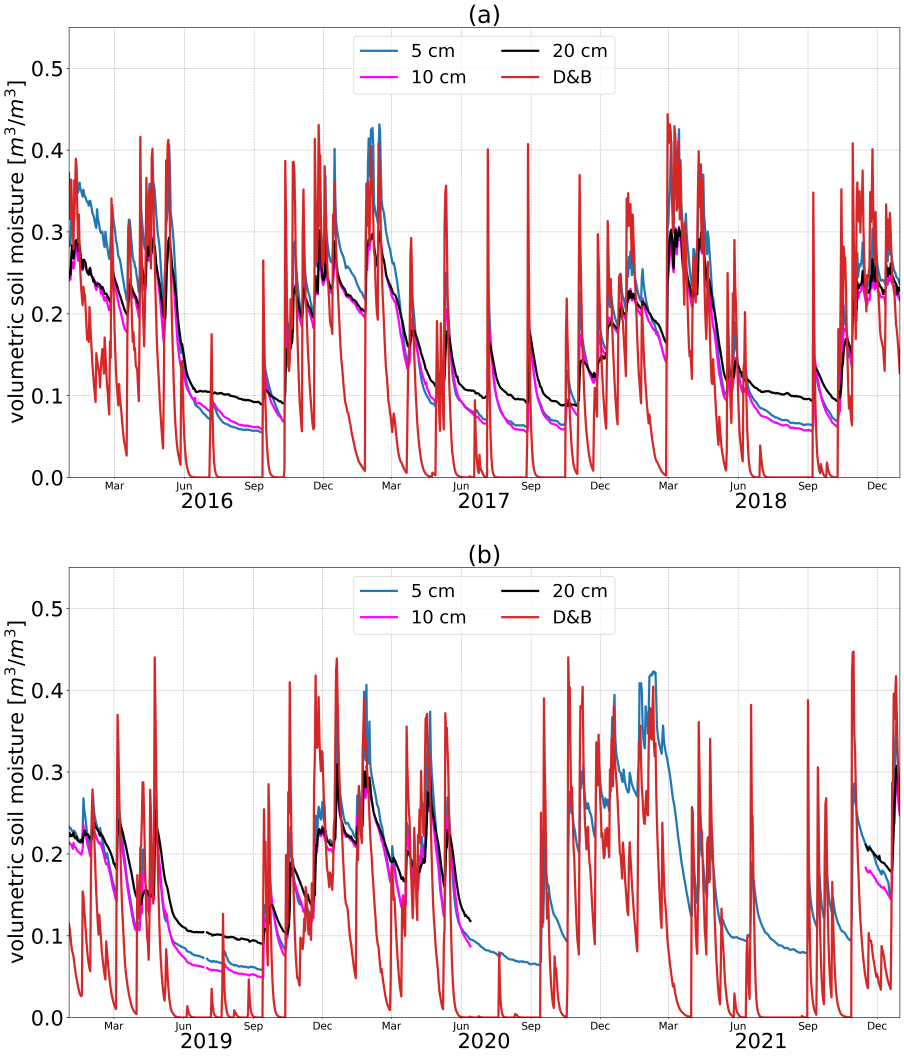

Figure 15. Daily *in situ* observed soil moisture at 5cm, 10cm and 20cm depth, and simulated near-surface soil moisture, W_s, at Majadas de Tietar for years 2016–2018 (a) and 2019–2021 (b).