CO2MVS RESEARCH ON SUPPLEMENTARY OBSERVATIONS

D4.5: Recommendations on required evolution of IFS and ECLand for biomass and city scale

Due date of deliverable	December 2025
Submission date	December 2025
File Name	D4.5
Work Package /Task	WP4
Organisation Responsible of Deliverable	Meteo-France
Author name(s)	Jean-Christophe Calvet, Marko Scholze, Thomas Kaminski, Anna Agusti-Panareda, Cédric Bacour, Wolfgang Knorr, Erik Koene, Gerrit Kuhlmann, Fabienne Maignan, Philippe Peylin, Vincent Tartaglione, Pierre Vanderbecken, Michael Voßbeck
Revision number	V2.0
Status	Final
Dissemination Level / location	PUBLIC www.corso-project.eu

The CORSO project (grant agreement No 101082194) is funded by the European Union.
Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the Commission. Neither the European Union nor the granting authority can be held responsible for them.

1 Executive Summary

The required evolution of the IFS in CO2MVS involves the calculation of carbon fluxes over global forests and at city scale. This work aims to investigate how to constrain biomass and biospheric fluxes over forests and estimate point-source emissions at city scale using satellite images and mass-balance methods. Recommendations are made on possible developments in the ECMWF IFS. This work is supported by developments conducted in ORCHIDEE (CEA), ISBA (Meteo-France), D&B (iLab/U Lund) and ICON-ART (EMPA, ECMWF), which allow different methodologies with varying levels of complexity to be explored. This report presents the possible methodologies and preliminary results obtained in each system to guide further developments for potential CO2MVS applications. Modelling and data assimilation over forests show that simple methods can be used to monitor carbon pools and fluxes. These approaches may enable the sequential assimilation of above-ground biomass observations using Kalman filtering without the need for long assimilation windows. Using multiple data streams simultaneously provides a stronger constraint on the terrestrial carbon cycle. In this approach, a long assimilation window is preferable for constraining slowly changing quantities in the cycle. However, a deeper understanding of disturbances and the history of carbon pools would require a more complex modelling approach that accounts for stand age and mortality. At the city level, it has been demonstrated that biospheric fluxes are the main source of background CO_2 heterogeneity.

Table of Contents

2 Introduction

2.1 Background

To enable the European Union (EU) to move towards a low-carbon economy and implement its commitments under the Paris Agreement, a binding target was set to cut emissions in the EU by at least 40% below 1990 levels by 2030. European Commission (EC) President von der Leyen committed to deepen this target to at least 55% reduction by 2030. This was further consolidated with the release of the Commission's European Green Deal on the 11th of December 2019, setting the targets for the European environment, economy, and society to reach zero net emissions of greenhouse gases in 2050, outlining all needed technological and societal transformations that are aiming at combining prosperity and sustainability. To support EU countries in achieving the targets, the EU and European Commission (EC) recognised the need for an objective way to monitor anthropogenic CO_2 emissions and their evolution over time.

Such a monitoring capacity will deliver consistent and reliable information to support informed policy- and decision-making processes, both at national and European level. To maintain independence in this domain, it is seen as critical that the EU establishes an observation-based operational anthropogenic CO_2 emissions Monitoring and Verification Support (MVS) (CO2MVS) capacity as part of its Copernicus Earth Observation programme.

The CORSO research and innovation project will build on and complement the work of previous projects such as CHE (the CO_2 Human Emissions), and CoCO2 (Copernicus CO_2 service) projects, both led by ECMWF. These projects have already started the ramping-up of the CO2MVS prototype systems, so it can be implemented within the Copernicus Atmosphere Monitoring Service (CAMS) with the aim to be operational by 2026. The CORSO project will further support establishing the new CO2MVS addressing specific research & development questions.

The main objectives of CORSO are to deliver further research activities and outcomes with a focus on the use of supplementary observations, i.e., of co-emitted species as well as the use of auxiliary observations to better separate fossil fuel emissions from the other sources of atmospheric CO_2. CORSO will deliver improved estimates of emission factors/ratios and their uncertainties as well as the capabilities at global and local scale to optimally use observations of co-emitted species to better estimate anthropogenic CO_2 emissions. CORSO will also provide clear recommendations to CAMS, ICOS, and WMO about the potential added-value of high-temporal resolution $^{14}CO_2$ and APO observations as tracers for anthropogenic emissions in both global and regional scale inversions and develop coupled land-atmosphere data assimilation in the global CO2MVS system constraining carbon cycle variables with satellite observations of soil moisture, Leaf Area Index (LAI), Solar Induced Fluorescence (SIF), and vegetation biomass. Finally, CORSO will provide specific recommendations for the topics above for the operational implementation of the CO2MVS within the Copernicus programme.

2.2 Scope of this deliverable

2.2.1 Objectives of this deliverables

The aim of this deliverable is to summarise the results of Task 4.4, which focuses on exploring the potential integration of new satellite products in CO2MVS, including at the city scale. Recommendations are provided for the evolution of ECLand and IFS in order to constrain biomass and biospheric fluxes over forests. The required evolution of the IFS in CO2MVS at city scale is based on ICON-ART simulations.

2.2.2 Work performed in this deliverable

For this task, we investigated and compared assimilation methods to constrain biomass variables using ESA Climate Change Initiative (CCI) biomass products, with a view to their potential integration into the future CO2MVS. We also used offline numerical experiments with ISBA, ORCHIDEE, and D&B models to identify the requirements for model complexity and future ECLand evolution in order to monitor carbon pools. Additionally, in close collaboration with ICOS-Cities, we explored the use of high-resolution satellite products for constraining biospheric fluxes at city scale. We carried out simulations with ICON-ART at city scale, using biospheric fluxes computed from MODIS observations (i.e. Enhanced Vegetation Index (EVI) and Land Surface Water Index (LSWI)) and the Vegetation Photosynthesis and Respiration Model (VPRM), to constrain the biospheric signal in the emission plumes of cities. This document presents the methods and results for ISBA, ORCHIDEE, D&B and ICON-ART. Recommendations are provided in the final section.

2.2.3 Deviations and counter measures

No deviations.

2.3 Task 4.4 partners contributing to the deliverable

Partners	
LUNDS UNIVERSITET	ULund
KAMINSKI THOMAS HERBERT	iLab
COMMISSARIAT A L ENERGIE ATOMIQUE ET AUX ENERGIES ALTERNATIVES	CEA
METEO-FRANCE	MF
EIDGENOSSISCHE MATERIALPRUFUNGS- UND FORSCHUNGSANSTALT	EMPA
EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS	ECMWF

3 Data

3.1 Biomass

The ESA CCI Above Ground Biomass (AGB) project provides annual, global maps that show the amount of above-ground biomass in forests, expressed in megagrams per hectare (Mg/ha). These data are crucial for monitoring the health of vegetation and estimating carbon stocks. To estimate the latter, we used the ESA-CCI AGB v5.0 product, which covers the years 2010 and 2015–2021. The yearly maps are computed at a resolution of 100 metres and then aggregated at various resolutions ranging from 1 to 50 kilometres. In this case, the maps have a spatial resolution of 25 km x 25 km (Santoro et al., 2024). AGB is provided for forests, as defined by the ESA-CCI land use map. Shrublands are excluded from this definition, as shown in Figure 1.

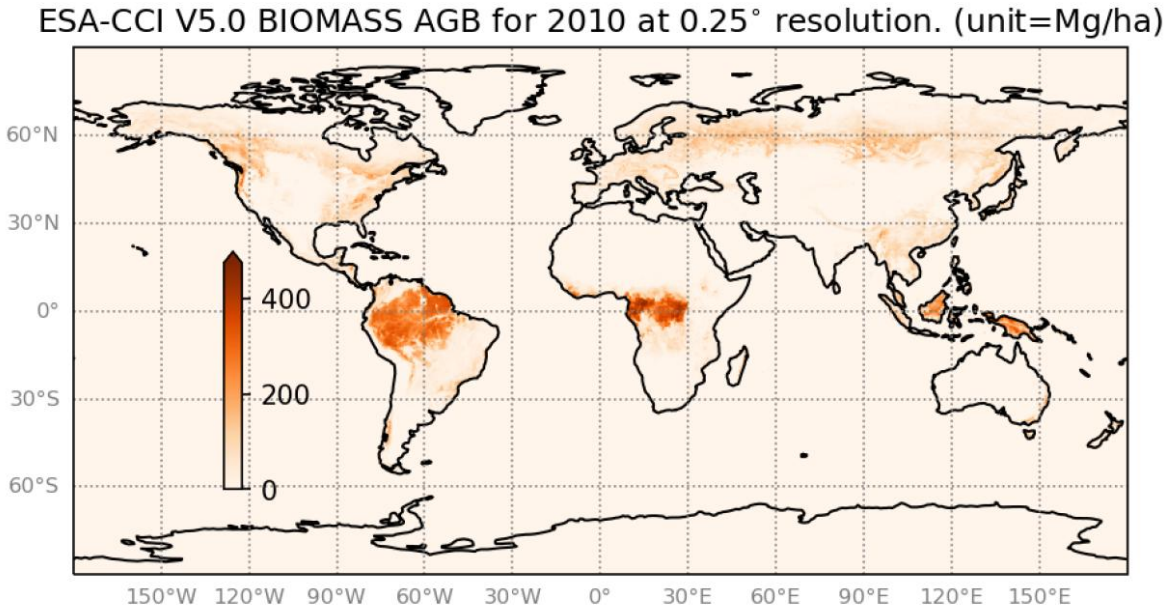

Figure 1: ESA-CCI above ground biomass global map for the year 2010 at 0.25° of resolution

Equatorial forests have the highest maximum AGB. The maximum AGB for boreal and temperate forests is about half that of equatorial forests.

3.2 MODIS vegetation indices

The following MODIS vegetation index observations were used: Enhanced Vegetation Index (EVI) and Land Surface Water Index (LSWI)).

4 Methods

4.1 ORCHIDEE

ORCHIDEE is a mechanistic land surface model (LSM) designed to simulate the fluxes of carbon, water, and energy between the biosphere and atmosphere (Krinner et al., 2005). It is a component of the Earth System Model developed by Institut Pierre-Simon Laplace IPSL-CM. The model operates from local to global scale, representing the spatial distribution of vegetation using fractions of plant functional types (PFTs) for each grid cell. Currently 15 PFTs are used: https://orchidas.lsce.ipsl.fr/dev/lccci/orchidee_pfts.php (last access: December 2025).

In this study, we use the latest version of the model (ORCHIDEE-V4), that will be included in the upcoming CMIP-7-FastTrack climate experiment. Since the last version that was used for CMIP-6, the following major functionalities were added:

- Since 2020, ORCHIDEE simulates a dynamic nitrogen cycle that includes the effects of nitrogen on photosynthesis, phenology, allocation, litter and soil decomposition, as well as plant nitrogen uptake, nitrogen losses to leaching, and nitrogen gas emissions (Vuichard et al., 2019). The nitrogen cycle implemented in ORCHIDEE required rewriting the allocation scheme (Naudts et al., 2015) following the principles of the pipe-model (Shinozaki et al., 1964).
- Canopies are simulated as three-dimensional heterogeneous canopies (Naudts et al., 2015). This representation is used in the calculation of radiative transfer through the canopy (Ardaneh et al., 2025).
- Canopy structure is now driven by demographic processes (Eckes-Shephard et al., 2025) and disturbances including forest management (Naudts et al., 2015), wind storms (Chen et al., 2022), and insect outbreaks (Marie et al., 2023).
- The physical properties of the soil (thermal conductivity and thermal capacity) were modified to account for cryoturbation, insulation from soil organic matter, and insulation from a surface organic layer. For these processes, soil organic matter has been discretized along soil depth.
- A large number of parameters have been optimized using various observational products, from in-situ observations to remote sensed Earth observations, following standard Bayesian optimization (Tarantola, 2005). Among these we mention, the optimization of the main parameters controlling the vegetation, soil background and snow albedos using satellite observations from MODIS instrument, and the optimization of the dependence of stomatal conductance on vapor pressure deficit (VPD).

For the data assimilation, we will use the ORCHIDAS Data Assimilation tool (https://orchidas.lsce.ipsl.fr/) (MacBean et al., 2022; Bacour et al., 2023). The assimilation relies on a Bayesian framework with a global misfit function between model simulations and observational data, considering error covariance matrices and prior information. We use a Genetic Algorithm optimization approach (Goldberg, 1989), to iteratively minimise the misfit function (Bastrikov et al., 2018).

In this study we focus on the African continent, trying to simulate the AGB gradients that are observed from satellite products (ESA-CCI) or from in-situ measurements. We choose Africa as it is currently less investigated than other tropical regions and this will set the base for a global pan-tropical assimilation of AGB data. The figure below summarizes the key processes that control AGB of forest ecosystems, from recruitment and photosynthesis to plant mortality processes, through carbon allocation and autotrophic respiration.

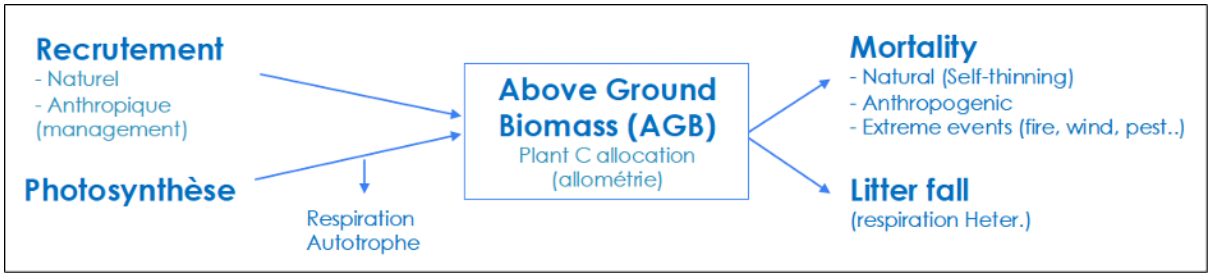

Figure 2: Main drivers of the Above Ground Biomass (AGB) simulated by the ORCHIDEE land surface model.

In order to calibrate the AGB gradient, focussing on mortality and growth parameters, we decided to proceed in several steps:

1. Preliminary step: we first calibrated the model with in-situ AGB data that are collected over the tropics choosing primarily to calibrate the regrowth curves of the model, i.e. the increase of AGB as a function of forest age. This first step was done in order to define prior parameter for the growth and mortality that are compatible with a few in-situ measurements. This was a preliminary calibration to ensure model realism.
2. In a second step (ongoing) we decided to assimilate the ESA-CCI AGB data to improve the simulated spatial gradients of AGB as a function of climate (mainly precipitation). For this step, we choose to first focus on the potential AGB that could be realistically simulated by the ORCHIDEE-V4 model. Indeed, ORCHIDEE-V4 does not yet include the effect of fire (upcoming feature), nor the impact of herbivores or very local human perturbations and we thus decided to calibrate the model with respect to potential maximum AGB that can be sustain given local climate and soil texture conditions. For that, we derived “potential AGB” at coarse resolution (0.5°) using the high resolution CCI data (300 m) and selecting in each 0.5° pixel only the upper quartile of the 300 m CCI-AGB data. The hypothesis is that the upper quartile represents mostly low-disturbance plots and thus potential AGB. The model will then be run and compared and calibrated with these potential AGB at 0.5° using ERA5 climate forcing. In addition, we also include in the calibration data-driven estimates of Gross Primary Productivity (GPP) taking the FluxComXbase estimates (Nelson et al., 2024).
3. In a final step, we will use the full ESA-CCI AGB data to compare with the optimized ORCHIDEE output with respect to potential biomass. Such comparison will allow us to derive on a regional basis the additional mortality (fire, herbivores, insects, …) or uptake by humans that are not currently represented in the selected version ORCHIDEE-V4. In this final step we will define an additional mortality flux, spatially varying across Africa.

We choose this approach of using first the “potential AGB” to calibrate specific processes of the ORCHIDEE model, without combining all effects not currently modelled in a generic mortality function. The results of the potential biomass map derived from ESA-CCI data and the gradients of these potential biomass as a function of annual mean precipitation are illustrated in the figure below. We note that the potential biomass increases with precipitation up to 1200 mm/yr and then becomes more scattered. It is likely that factors other than precipitation determine AGB density above this value.

This first step of the task has been completed at the time of the writing. It was mainly done with a trial-and-error approach for the calibration of a few parameters only (see results). A more comprehensive approach would involve using a formal Bayesian calibration to first select an ensemble of sites representing a gradient of annual mean precipitation.

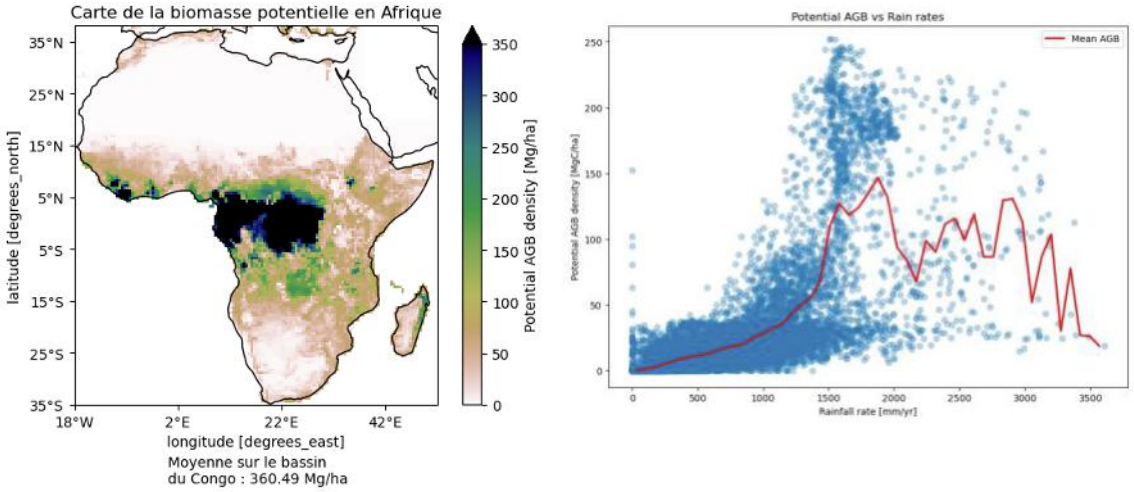

Figure 3: Left: map of potential AGB data over Africa derived from the ESA-CCI product; Potential biomass is derived from the high resolution CCI data (300 m) at 0.5° taking only the upper quartile of the AGB distribution in each grid at 0.5°. Right: gradient of potential AGB data as a function of annual mean precipitation.

4.2 ISBA

This study uses an updated version of the ISBA model (Gibelin et al., 2006), which can simulate soil moisture and temperature, as well as photosynthesis, plant growth, senescence, and wood biomass, both above and below ground. This version is denoted as ISBA-CC. Phenology is driven entirely by photosynthesis via a straightforward allocation scheme. Net leaf CO_2 assimilation represents the incoming carbon flux for leaf biomass growth. A photosynthesis-dependent leaf mortality rate is then calculated. The balance between leaf carbon uptake and leaf mortality results in an increase or decrease in leaf biomass. Leaf biomass is converted to LAI using a fixed surface leaf area (SLA) value per plant functional type. Simulations were performed using the SURFEX v9 modelling platform. Vegetation was categorised into 19 plant functional types based on the ECOCLIMAP-II land use map (Faroux et al., 2013). Deviating from the standard 12 PFTs used in other project deliverables was motivated by the need to distinguish shrublands from forest vegetation to align with the ESA-CCI description of forest above-ground biomass, which excludes low vegetation. The 19 PFTs are: bare soil (NO), bare rock (ROCK), snow and ice (S&I), C3 crops (C3), C4 crops (C4), grassland (GRASS), boreal grassland and tropical grassland (BOGR), irrigated areas (IRR), parks (PARK), temperate broadleaf deciduous forest (TEBD), tropical broadleaf deciduous forest (TRBD), boreal broadleaf deciduous forest (BOBD), temperate broadleaf evergreen forest (TEBE), shrublands (SHRB), tropical broadleaf evergreen forest (TRBE), boreal needleleaf evergreen forest (BONE), and temperate evergreen needleleaf forest (TENE).

The first step in checking the feasibility and relevance of assimilating ESA-CCI above-ground biomass to improve biomass representation in ISBA was to run two experiments starting from 2010 to 2021: one based solely on ISBA physics and one initialised with biomass from ESA-CCI. This required us to address the challenge of allocating ESA-CCI biomass to different PFTs within the ISBA model. This was achieved in two steps. First, areas where the forest fraction agreed between the model and the ESA-CCI were selected. Then, these biomasses were distributed conservatively among the different PFTs. The first step in the assimilation process is an offline screening step, while the second step corresponds to the observation operator.

To derive the forest fraction associated with the 25 km resolution global ESA-CCI AGB map, we start with the 1 km resolution ESA-CCI AGB map, which is the highest resolution our

computer can load. We then define an indicator ranging between 0 and 1 to emulate the forest fraction at 1 km resolution, according to the following equation:

$FF_{1km} = \min(AGB/150, 1)$.

The 150 Mg ha^{-1} threshold was established by adjusting the parameters in order to maximise the area of forest considered. The forest fraction (FF1km) was then averaged over the 25 km resolution map to produce the map shown in Figure 4(b). The forest fraction for the model is straightforward: it is obtained by summing the fractions of the eight forest types (TEBD, TRBD, BOBD, TEBE, TRBE, BONE, TENE and BOND) within the 25 km grid cell (see Figure 4(a)). To select grid cells with consistent forest fractions in both the model and the CCI AGB maps, we calculate the absolute difference between the two forest fraction estimations, as shown in Figure 4(c). We then filter out all cells where these estimates differ by more than 0.2 and all cells where the model forest fraction is lower than 0.2.

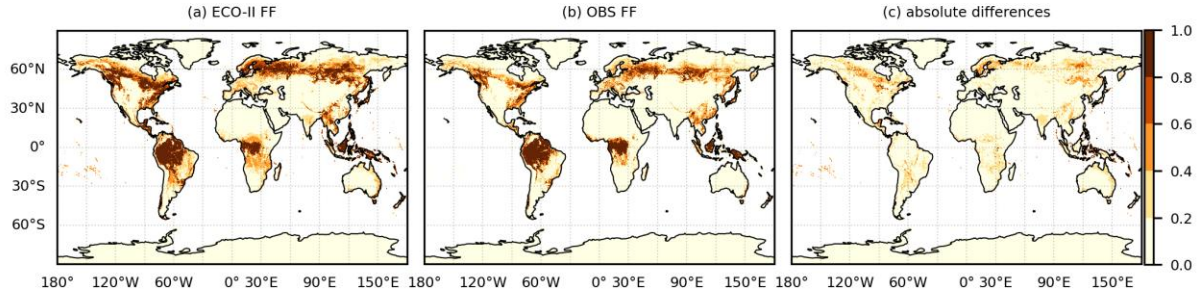

Figure 4: Comparison between the forest fraction (FF) derived from (a) ECOCLIMAP-II (ECO-II) and (b) ESA-CCI biomass observations (OBS). The ECO-II minus OBS absolute difference is shown (c) at 25km of resolution.

In the remaining grid cells with significant AGB values, ESA-CCI biomass is distributed according to the predicted PFT fraction, ensuring that this fraction represents more than a quarter of the total forest fraction. This final filtering step prevents the value stored for a given PFT from increasing significantly due to that PFT having a very low fraction. Figure 5 shows the resulting initial distribution for the different PFTs.

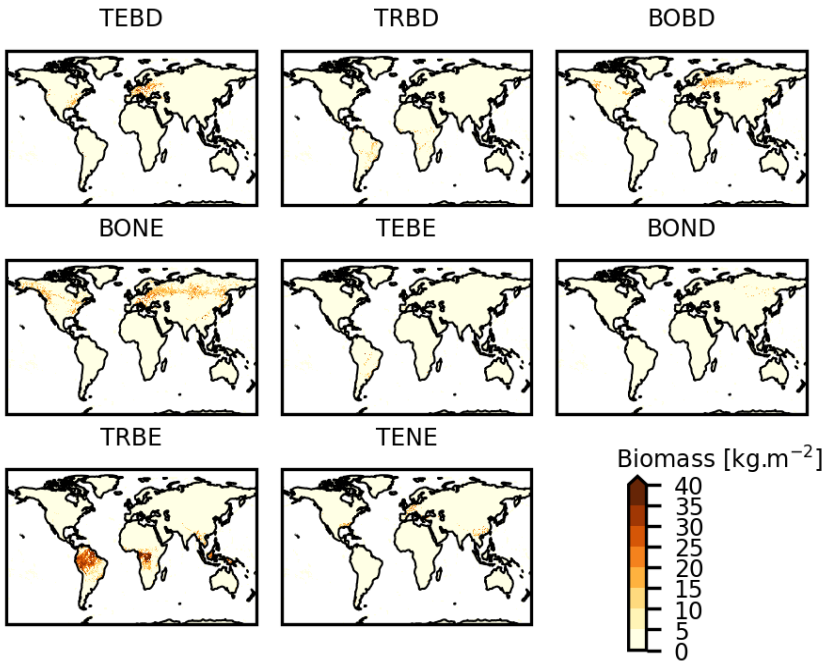

Figure 5: 2010 initial distribution of above ground woody biomass set in ISBA for the initialised experiment per PFT.

One disadvantage of not considering the PFT when its fraction is insignificant is that the final value of the above-ground biomass model may be lower than that provided by the ESA-CCI. This means that the initial conditions tend to underestimate the actual AGB.

Figure 6 shows a comparison between the initial ESA-CCI biomass provided for 2010 and the biomass used for initialisation. We proportionately scaled the below-ground woody biomass (BGB) to the above-ground woody biomass for each plant functional type (PFT), using factors inspired by those proposed by Huang et al. (2021) and presented in Table 1.

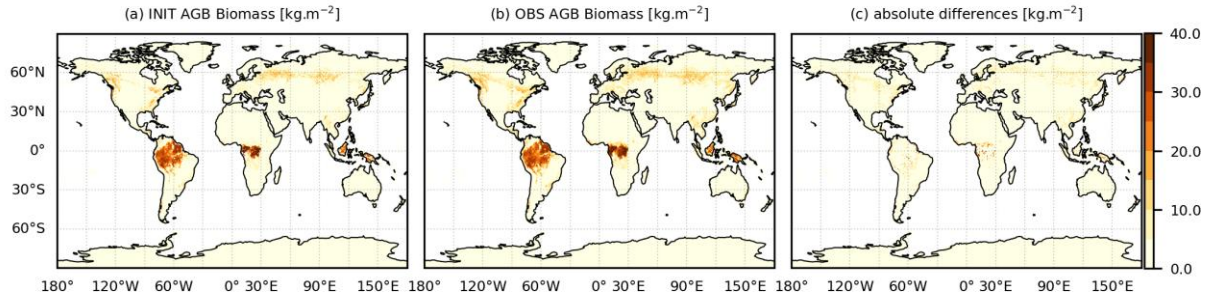

Figure 6: Comparison between the initial 2010 values of above ground woody biomass (a) in ISBA and (b) observed by ESA-CCI (OBS). The ISBA minus OBS absolute difference is shown (c).

Table 1: Reference ratio between the woody below ground biomass and the woody above ground biomass for the different forest PFTs used in ISBA.

PFT	TRBE	TEBD, BONE, TEBE, TENE	TRBD, BOBD, BOND
Below- to above-ground ratio	0.24	0.25	0.28

For both experiments, the initial LAI is set to ECOCLIMAP-II values. The soil temperature is set to 293.16 K for all soil layers. The soil water index is set to 1.3 in all layers.

4.3 D&B

ULUND and iLab used the D&B terrestrial ecosystem community model (Knorr et al. 2025). The D&B model is based upon three interconnected sub-model components: (i) photosynthesis and autotrophic respiration, (ii) energy and water balance, and (iii) carbon allocation and cycling, including heterotrophic respiration. The first component includes a process-based description of uptake of CO_2 via plant photosynthetic activity (gross primary production, GPP), regulated by temperature, light absorption across the canopy, and stomatal control, and of carbon loss from the respiration of live vegetation (RA, autotrophic respiration). The remainder, net primary production (NPP = GPP - RA), is then passed over to the Carbon Allocation and Cycling component and distributed among the various carbon pools. The Energy and Water Balance component regulates the energy input to and output from the canopy in the form of radiative, latent and sensible heat exchange with the atmosphere, taking into account the hydrological status of the canopy and soil, as well as the plant transpiration. Components (i) and (ii) are based on BETHY (Knorr, 2000), and component (iii) on DALEC (Williams et al., 2005).

D&B has observation operators for the following EO data streams:

- FAPAR (fraction of absorbed photosynthetically active radiation) using the two-stream radiative transfer scheme according to Sellers (1985),
- SIF (solar-induced fluorescence) offering the formulations of Gu et al. (2019), van der Tol et al. (2014) and Li et al. (2022) for the SIF source terms with calculation of canopy-level SIF via the layered 2-stream model (L2SM) by Quaife (2025),
- L-band VOD (vegetation optical depth) using an empirical formulation (Knorr et al., 2025) that combines impacts of biomass and plant water, and
- near-surface soil moisture following the VIC-1L approach (Wood et al., 1992).

D&B and its observation operators are embedded into the variational Terrestrial Carbon Community Assimilation System (TCCAS). TCCAS assimilates all data streams simultaneously in a single long assimilation window. TCCAS includes tangent and adjoint versions of D&B and its observation operators. TCCAS is an open-source development; its code is available from TCCAS.Inversion-Lab.com.

4.4 ICON-ART and IFS plume simulations and processing

To accurately estimate point source emissions from satellite images using mass balance methods, the mass enhancement due to the point source (i.e., the mass exclusively related to the point source of interest) needs to be well known. It is typically computed by subtracting a background estimated from the image itself, but could also be estimated using a low-resolution modelling system. This is done, for example, by using values upstream, or by the values surrounding the plume, as a proxy for the background within the plume. This proxy method can fail if the estimated background is biased or if the true background shows strong spatial

variability from sources and sinks (e.g., biospheric fluxes) that are not sufficiently approximated by a constant or spatially smooth background estimate.

EMPA and ECMWF study the effects of such a proxy method in two cities (Zurich and Paris) using the high-resolution modelling systems ICON-ART (0.5 km resolution for Zurich, 1 km resolution for Paris) and the Integrated Forecasting System (IFS, with three resolutions: 28 km, 9 km, and 4.5 km). These models simulated XCO_2 fields over Zurich and Paris during June and July 2023 and produced two output fields: one including all CO_2 components (referred to as 'total'), and another excluding the respective city anthropogenic emissions (referred to as 'background'). The goal is to determine whether high-resolution models yield a good estimate of the background. The ICON-ART model computes biospheric fluxes using the Vegetation Photosynthesis and Respiration Model (VPRM, Mahadevan et al., 2008). VPRM requires meteorological parameters, which were provided online and satellite measurements of two indices, the enhanced vegetation index (EVI) and the land surface water index (LSWI), which were obtained from the MODIS instrument at the resolution of the ICON-ART simulations. The ICON-ART simulations were conducted in collaboration with the ICOS-Cities project (Ponomarev et al., 2025).

The IFS simulations were performed using the CAMS forecasting system (Agusti-Panareda et al., 2019) with CY49R2. The biogenic fluxes are simulated with ECLand (Boussetta et al., 2021) and they are bias corrected using the CO2 fluxes from the CAMS inversion dataset (CAMS, 2025) and the Biogenic Flux Adjustment Scheme (Agusti-Panareda et al.,2014). The fossil fuel emissions are from the CAMS-GLOB-ANT v6.2 (Soulie et al., 2024).

To estimate the background from the models, we developed an algorithm that automatically masks a city plume based on the ICON-ART 'city plume' image. This image is formed by taking the difference between the 'total' and 'background' XCO_2 images, which thus corresponds solely to the city plume. The mask is created by initiating a 'seed' centered at the city, selecting all cells within 4 km for Zurich or 10 km for Paris, and then flood filling a mask in the direction downwind using the modeled ICON-ART 10 m wind speeds to within ±90°, as long as the XCO_2 values lie above a threshold. The algorithm tests 13 thresholds (ranging from 0.8 to 2.0 times the mean value at the city seed), and selects the one that scores best on a cost function defined by three criteria:

- Contrast (the difference between the mean of pixels inside the mask and the mean of the pixels outside the mask)
- Elongation (the plume is expected to not twist and turn much, tested by comparing the mean angle of the plume to the mean wind speed)
- Fraction (the fraction of the image covered by the mask, i.e., a number between [0,1])

The weighted scoring function favours plumes with clear contrast and reasonable (i.e. not too large) elongation and extent. The threshold that maximises this score is used to generate the plume mask, which is then expanded by 1.5 km in all directions for Zurich and by 4 km for Paris to ensure the margins of the plume are included. We take the 'total' pseudo-XCO_2 image for each of the models (ICON-ART and IFS), mask out the city plume using the estimated mask, and fill in the data through an operation like a normalized convolution (i.e., a convolution operation aware of missing data). As ICON-ART operates on an unstructured grid, we could not implement normalized convolution using a 2-D kernel – instead, we implement it by iteratively averaging each node's value with those of its neighbors, while skipping invalid (NaN) entries. Interpreted differently, we solved the heat equation on the unstructured grid, with initial and boundary conditions defined by the unmasked area to fill the masked area with a diffuse field based on its edges. As we resorted to this method for ICON-ART, we implemented it

similarly for IFS, even though the results are essentially identical to those obtained using normalized convolution with a 2D Gaussian kernel.

The reconstructed background was then compared with the "true" background derived from the ICON-ART model. This comparison was conducted for the model runs at 10 and 11 AM, close to the expected CO2M satellite overpass time.

5 Results

5.1 ORCHIDEE

Step 1: pre-calibration of the ORCHIDEE model with in-situ data

In this first step a collection of in-situ AGB data has been made, grouping them as a function of forest age. This allows the reconstruction of a regrowth curve (AGB as a function of forest age) for a given Plant Functional Type of ORCHIDEE. Several parameters of the ORCHIDEE model were manually tuned based on our understanding of the model equations and a priori knowledge of the most uncertain parametrisations. Five parameters controlling self-thinning, two parameters controlling recruitment and two for the carbon allocation were adjusted. The figure below illustrates the results for the deciduous broadleaf tropical trees (rain-green PFT). The ORCHIDE-V4 represents reasonably well the increase of AGB as a function of age, with slightly lower biomass that the observation before 50 years and then a relatively good match with the in-situ observations. This pre-calibration showed the importance of calibrating the self-thinning curves (i.e., Tree size = f(density)) as well as the recruitment in the model.

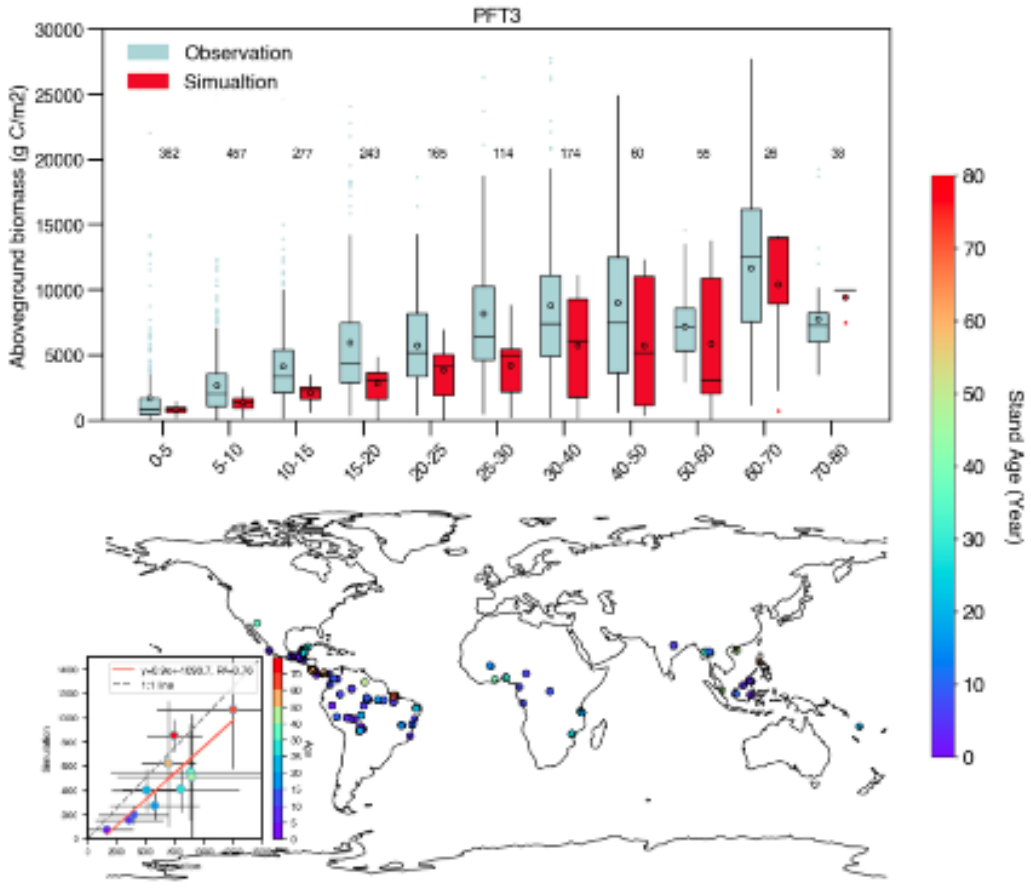

Figure 7: Top: Above Ground Biomass (AGB) as a function of age (regrowth curve) simulated by the ORCHIDEE-V4 model (red) for the tropical deciduous (raingreen) plant function type (PFT3) and from in-situ observations; the box plots correspond to the spread of the observations or model AGB for the different sites. Bottom: a map showing the location of the in-situ plots over the tropics that were used to create the observation regrowth curve and scatter plot of the model versus the observations.

Step 2: Assimilation of the ESA-CCI AGB data, using only potential biomass at 0.5°

We ran the pre-calibrated ORCHIDEE-V4 version over Africa, using ERA5 climate forcing to drive the simulations. We started the procedure over the southern part of Africa to decrease the computing time and step up the overall data assimilation procedure. Figure 8 illustrates the spatial gradient of potential AGB derived from ESA-CCI data (see method) and the one simulated by ORCHIDEE, averaged over the last decade (2020-2020). We clearly see some similarities with decreasing AGB as we go from the equatorial region to the southern part of Africa, following the decrease of annual mean precipitation. However, the regional gradients show large discrepancies between the model and the data-driven product, highlighting potentially significant improvements of the model after a formal Bayesian data assimilation (DA) step. To prepare such DA step, we have first started with an exhaustive sensitivity analysis including around 20 parameters. Such analysis allowed us to consider only a restricted set (around 10) of most influential parameters with respect to AGB. We also selected a set of parameters controlling primarily the GPP in order to assimilate also the GPP from FluxComxbase at the same time than potential AGB. This calibration step was initiated and will be finalised at a later stage. The calibration is performed with a gradient-based minimisation approach (using the BFGS standard algorithm). Note that, the sensitivity of the cost function (function minimizing the model – data mismatch) to the different parameters is estimated using a finite difference approach. Instead of using the whole African region (all pixels at 0.5°) we first made a selected set of points to test the approach. Preliminary results of the calibration highlight the importance of the DA setup. We indeed made the hypothesis that the selected pixels are in equilibrium with respect to AGB and thus we have to select a long enough time window over which the model needs to be run with updated parameters (still using a common spinup). We are currently testing different time window from 100 years to only 50 years. From these preliminary results we anticipate that optimisation of the AGB gradients will be substantial and can already draw some recommendation (see next section).

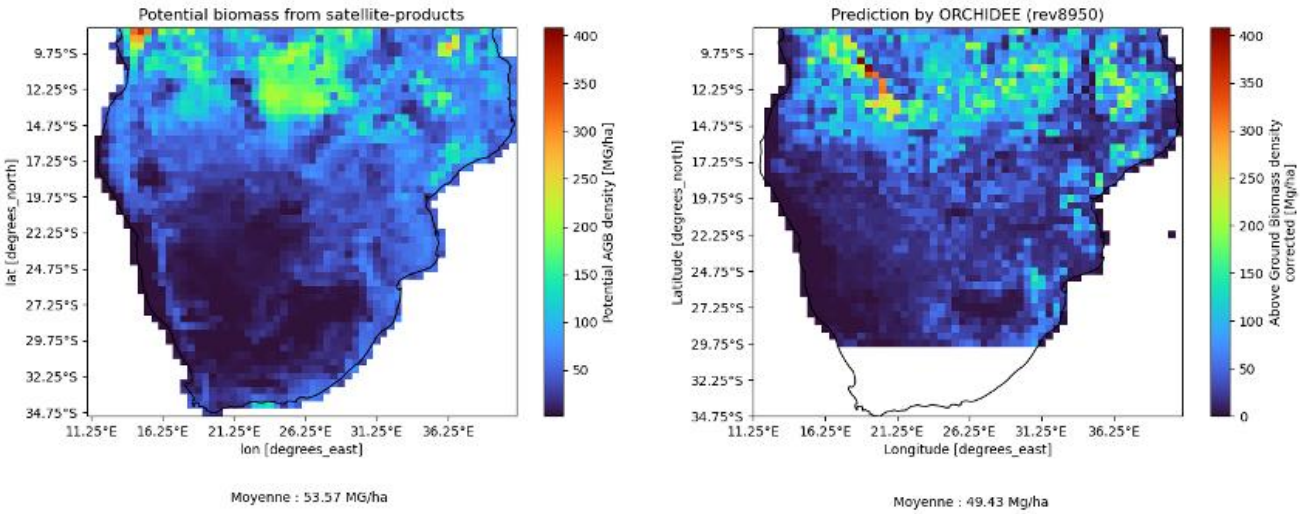

Figure 8: Potential Above Ground Biomass (see method section) for the southern part of Africa estimated from ESA-CCI data (left) and simulated by the ORCHIDEE-V4 model, for the 2010-2020 period.

5.2 ISBA

An increase in biomass storage does not imply a change in leaf area index (LAI) or other vegetation parameters. Here, we focus on representing woody aboveground biomass (AGB). After five years of simulation, we compared the final AGB for the two experiments with the 2015 ESA-CCI biomass data. Figure 9 shows that the baseline simulation does not reach the observed amount of biomass, since five years is probably shorter than the lifespan of most forests. The experiments with initialisation maintain the same spatial distribution, and after five years of growth the initial underestimation has decreased, resulting in a global mean difference of -0.21 kg m^{-2} compared to the ESA-CCI biomass data.

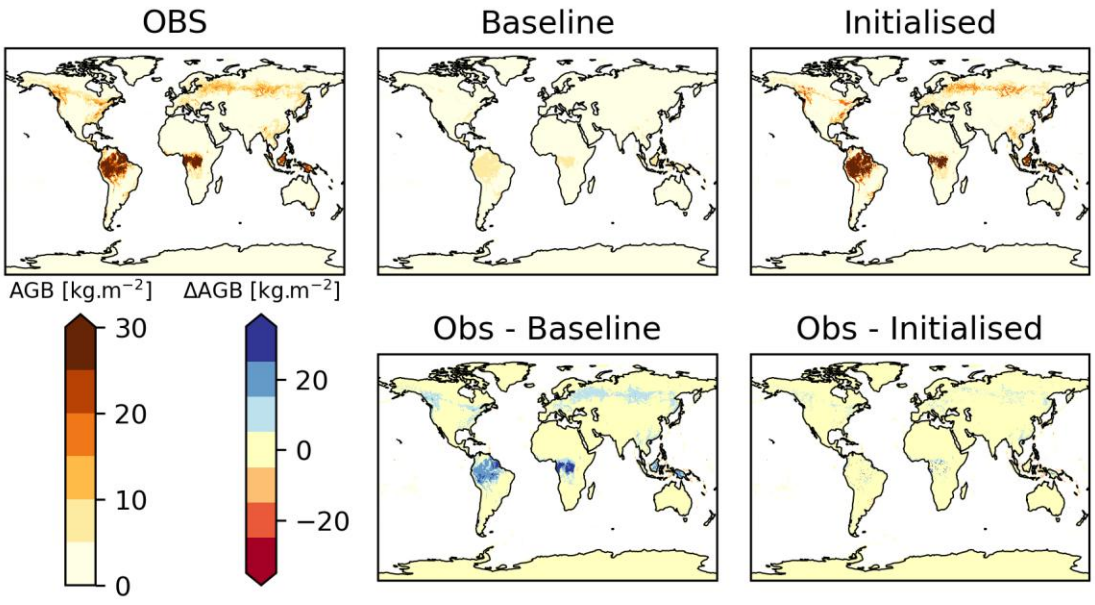

Figure 9: Comparison between the woody above ground biomass simulated by both the baseline and the initialised experiments for the year 2015.

Figure 10 shows the increase in woody biomass over the simulation period for both experiments. After 10 years, the mean values produced by the baseline experiment were similar to those of the ESA-CCI biomass. However, a large range of differences was observed.. For the initialised simulation, the best agreement in terms of mean squared error (MSE) was achieved between years 6 and 7. After that, the results remained better than the baseline, but the woody biomass was overestimated compared to the ESA-CCI data. The distribution of differences is more centered and has less variance for the initialised simulation than for the baseline experiment.

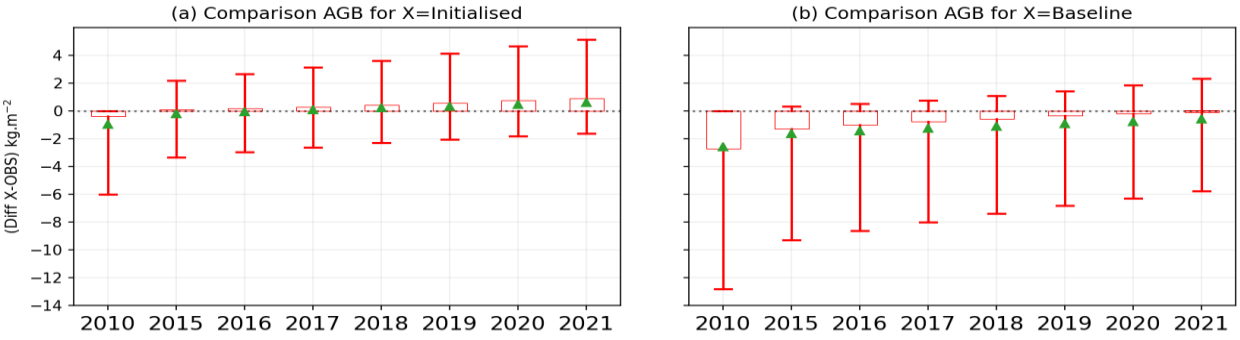

Figure 10: Boxplots of the differences between the woody above ground biomass simulated by respectively the initialised (a), the baseline experiments (b) and the ESA-CCI observed biomass for each year available. Box represents the 25th and the 75th percentiles, the whiskers the 5th and 95th percentiles and the green triangle the mean.

By the end of the simulation period, the expected biomass value had been reached in the baseline experiment. However, as Figure 11 shows, the AGB values in the equatorial regions were still underestimated, and the AGB distribution in the northern hemisphere was too homogeneous. While the initialised experiment produced greater heterogeneity, as the boxplots indicate, it overestimated the above-ground biomass. Focusing on the Amazon area reveals that the observed biomass is more heterogeneous than the simulated biomass. The model's biomass growth appears uniform and fails to capture these observed heterogeneities. As the model does not account for forest management and deforestation, the simulated AGB tends to smooth out these heterogeneities.

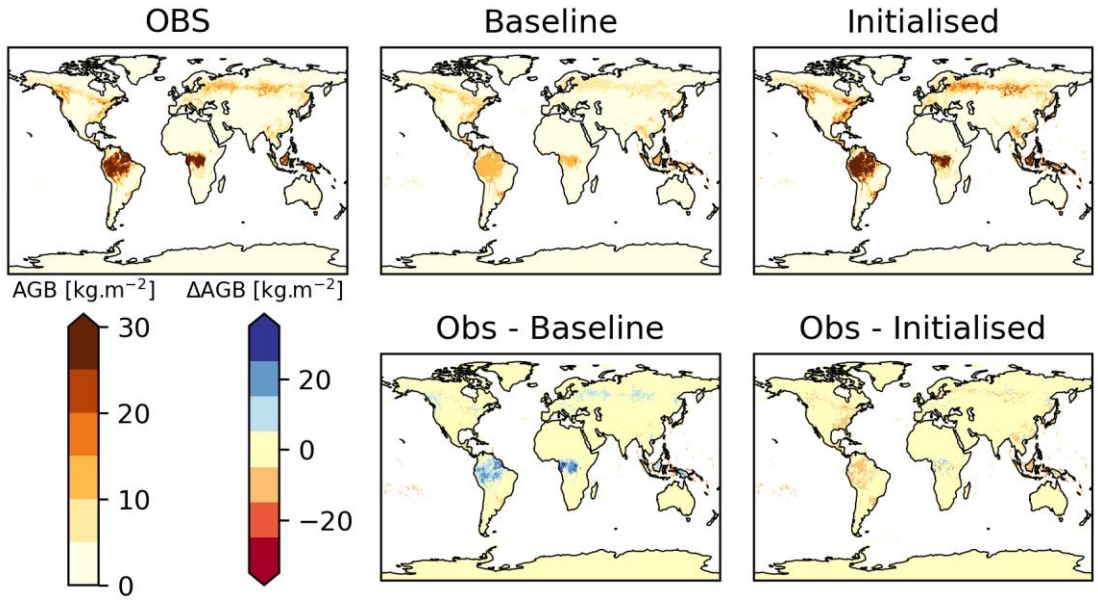

Figure 11: Comparison between the woody above ground biomass simulated by both the baseline and the initialised experiments for the year 2021.

Below-ground woody biomass was also initialised. As the Figure 12 shows, the above-to-below-ground biomass ratios differ from the reference values provided by Huang (Table 1). For both experiments, all ratios stabilised at values lower than the reference values for the PFTs, except for temperate broadleaf evergreen forests (TEBEs). Initialising the below-ground biomass resulted in a slightly higher ratio than without initialisation, but still lower than the reference value.

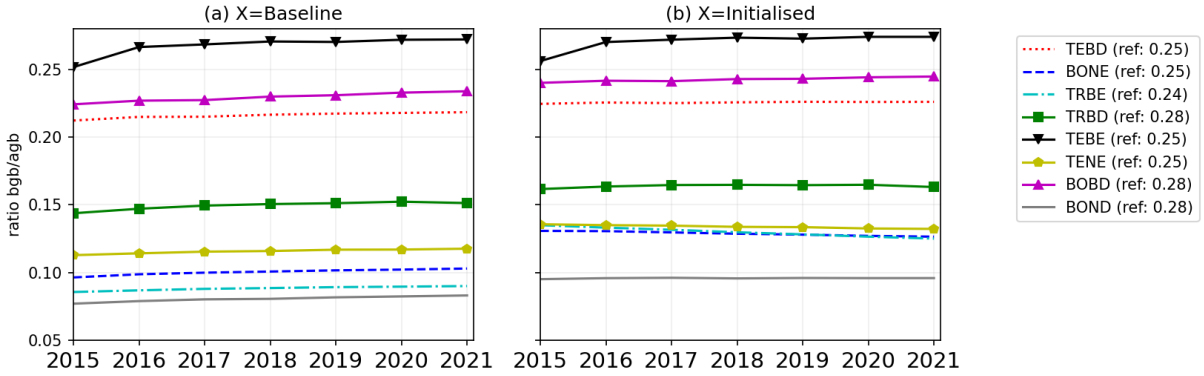

Figure 12: Comparison of the woody above-to-below-ground biomass ratio for the 8 forest PFT of the ISBA model from 2015 to 2021. (a) Baseline experiment (b) Initialised experiment.

The AGB resulting from simulations with no initialisation takes decades to grow from zero to a value closer to the observed figures. This growth rate appears consistent across different vegetation types and regions. Therefore, a spin-up step is irrelevant for AGB in ISBA; otherwise, AGB values in temperate forests would be similar to those in equatorial forests. This is due to the lack of representation of forest management and wildfires in ISBA, which results in a mortality rate that is too low for biomass and exhibits less heterogeneity in biomass distribution. While initialising the AGB addresses the issue of biomass relevance without the need for decades of simulation, it also results in a growth rate that is too high given the lack of mortality, leading to excessive AGB by the end of the simulation. This demonstrates the need to assimilate AGB observations yearly in ISBA.

The first step in this process demonstrates the feasibility of sequentially assimilating such products into ISBA using the method described in Section 4.2. However, before implementing sequential assimilation using a Kalman filtering approach, for example, the model and observation errors need to be assessed. An initial estimate of the observation error can be obtained from the boxplots in Figure 11, i.e. 2.15 kg m^{-2}. While it is possible to assimilate below-ground biomass simultaneously, the model appears to require modification to align with the reference values for the ratio of below-ground to above-ground biomass.

5.3 D&B

We operated D&B/TCCAS at the boreal ICOS site of Sodankylä for a seven-year period from January 1, 2015 to December 31, 2021, with a two-year spin-up phase followed by a five year assimilation window. The system was used to jointly estimate the sizes of D&B's initial carbon pools along with process parameters in the core model and the observation operators. We can regard the online calibration of the model and the observation operators as a form of online bias correction. We performed three experiments:

- Our default experiment assimilates only L-band VOD derived from SMOS (Kerr et al, 2010) over the full five-year assimilation window.
- The second experiment assimilates L-Band VOD derived from SMOS together with FAPAR provided by JRC-TIP (Pinty et al., 2007) from MODIS broadband albedos (Pinty et al., 2011), surface soil moisture provided by SMOS (Kerr et al, 2010), and SIF retrieved from TROPOMI (Guanter et al., 2021) over the full five-year assimilation window.
- The third experiment repeats Experiment 2 but with a reduced assimilation window of only one year.

For validation we employ independent eddy covariance measurements (The Inversion Lab, 2024) of NEE and a GPP data set derived from the NEE data set. We further use EO data streams not assimilated for validation.

Figure 13 shows the comparison of the assimilated data stream and a range of validation data streams against the corresponding simulation with the posterior (i.e. calibrated) model after assimilation of L-VOD only (Experiment 1). As expected, the assimilation improved the fit to L-VOD compared to the prior, while it degraded the fits to SIF and surface soil moisture. It also improved the fit to NEE, GPP, and FAPAR. However, the time series of the latter two variables do indicate that the assimilation of L-VOD drove the model into an unrealistic state of too low productivity.

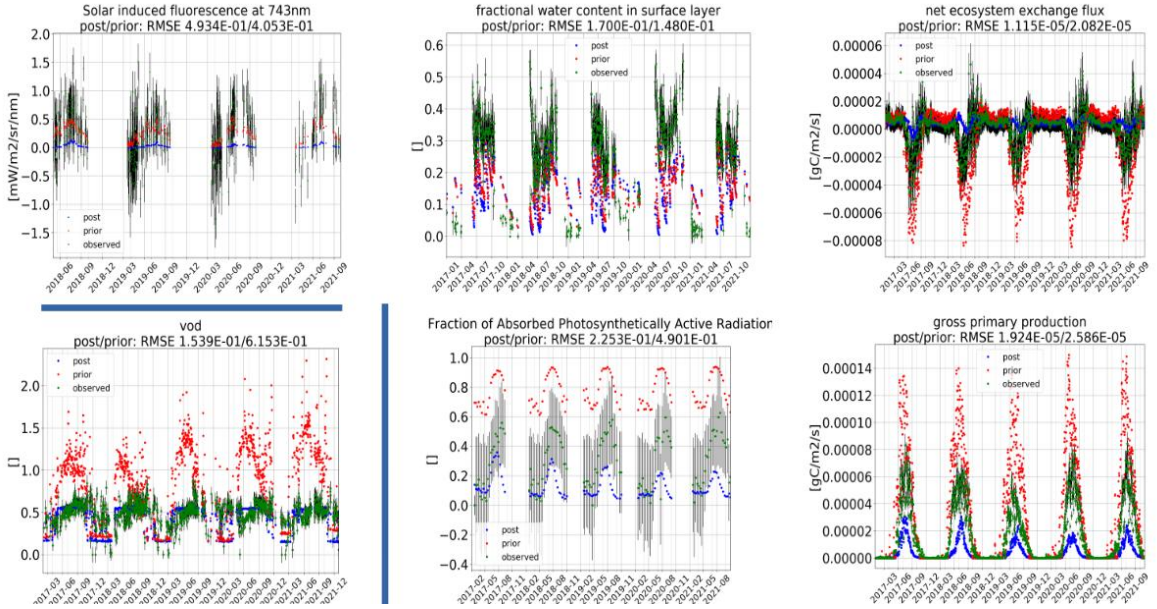

Figure 13: Comparison of assimilation data (left and below blue bar) and independent data (right and above of blue bar) against prior and posterior model output after assimilation L-VOD only (experiment 1). Prior and posterior RMSE values are for each data stream on top of the graphs.

By contrast, Experiment 2, which assimilates all four EO data streams, achieves a good fit to the independent eddy covariance measurement (Figure 14). In particular, the posterior simulation exhibits a reasonable productivity. The fit to all assimilated data streams has improved compared to the prior.

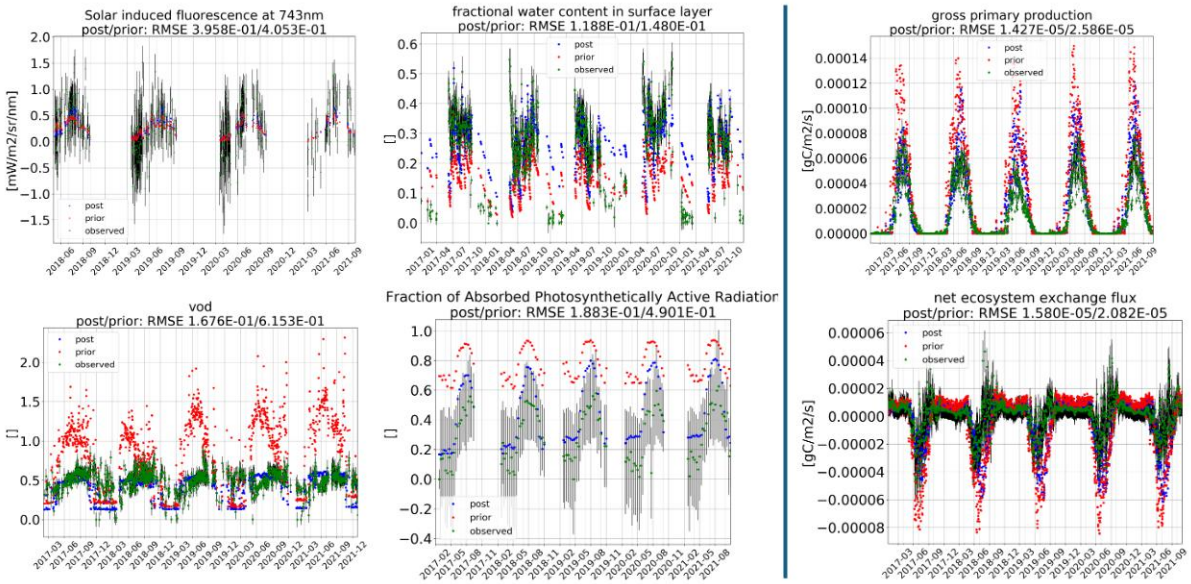

Figure 14: Comparison of assimilation data (left of blue bar) and independent data (right of blue bar) against prior and posterior model output after simultaneous assimilation of L-VOD, SIF, FAPAR and soil moisture (experiment 2). Prior and posterior RMSE values are for each data stream on top of the graphs.

Experiment 3, assimilating all data streams but with an assimilation window of only one year, improves the fit to all assimilated data streams and both validation data sets compared to the prior (as shown in Figure 15). Note that the SIF data set only became available from 2018.

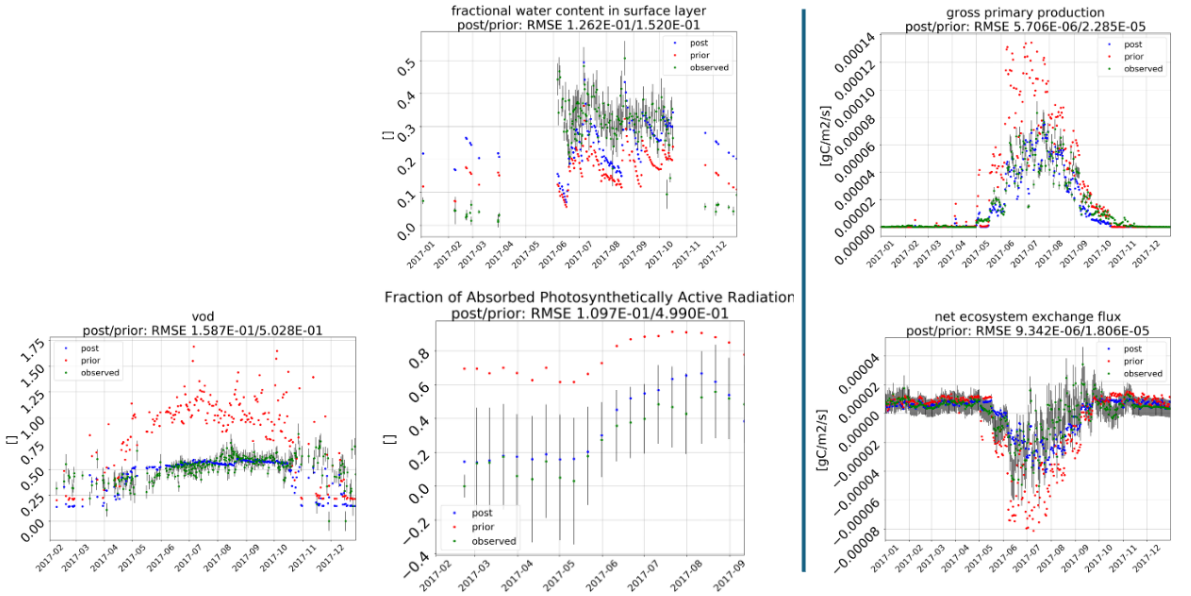

Figure 15: Comparison of assimilation data (left of blue bar) and independent data (right of blue bar) against prior and posterior model output after simultaneous assimilation of L-VOD, SIF, FAPAR and soil moisture but with an assimilation window of one year (experiment 3). Prior and posterior RMSE values are for each data stream on top of the graphs.

The effect of the length of the assimilation window on the posterior carbon pools is illustrated by Figure 16, which shows the posterior evolution of the slow carbon pools (woody biomass and soil organic matter) of the evergreen trees for Experiments 2 and 3. While posterior initial woody biomass in Experiment 2 is around 4050 gc/m^2 and, except for a seasonal cycle, remains constant, the posterior initial woody biomass in Experiment 3 is around 2400 gc/m^2 and increases to roughly 2600 gc/m^2 over three years. The posterior initial soil organic matter pool has roughly the same size in both experiment but decreases faster in Experiment 3.

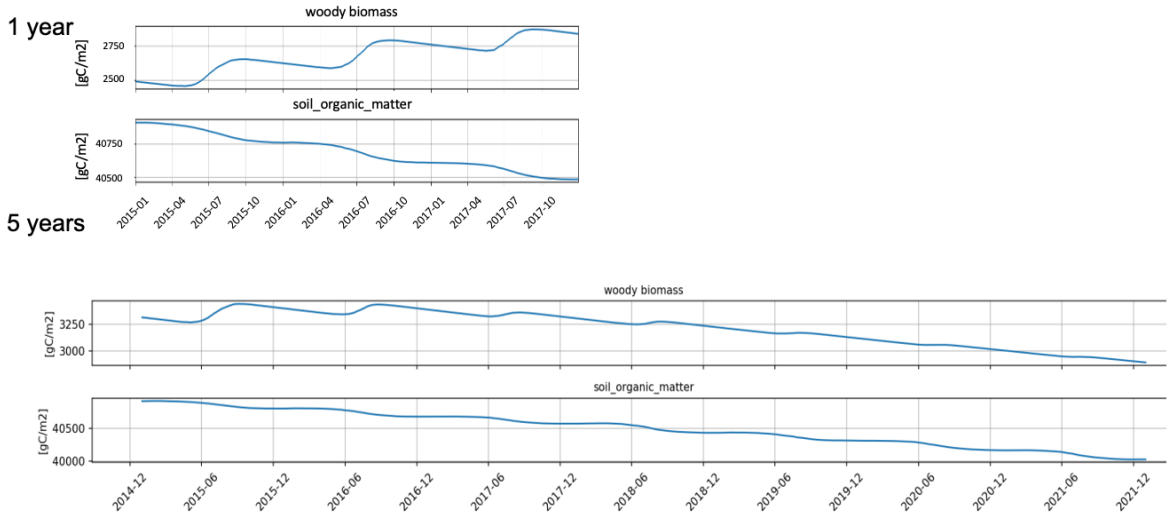

Figure 16: Evolution of the size of the slow carbon pools (woody biomass and soil organic matter) after assimilation (posterior) of 1 year (top) and 5 years of data (bottom). Note that the top graph is shrunk along the x-axis to correspond to the same time period as in the bottom graph.

5.4 ICON-ART and IFS plume simulations

Figure 17 shows an example of Paris city plume simulated with ICON-ART and IFS at different resolutions. We can see that the identified city plume mask is representative of the city, indicating our automated city plume algorithm performs well. Furthermore, for ICON-ART, the reconstructed background shows amplitudes slightly larger than the true background. The range of amplitudes in the 'true' city plume and its reconstruction is, however, relatively narrow. Conversely, the IFS results show a much broader range of XCO_2 values; and all reconstructions approach the 'true' background as simulated with IFS, albeit with a smaller spread. To derive statistically relevant parameters, we view the ICON-ART results as ground truths, as the simulated fields performed well against in-situ observations in and around Zurich and Paris (Ponomarev et al., 2025). We compute the mean bias (i.e., the mean of the true ICON-ART city plume background XCO_2 is compared to the mean of the reconstructed city plume background) for all simulated cases at 10 and 11 AM during June and July 2023.

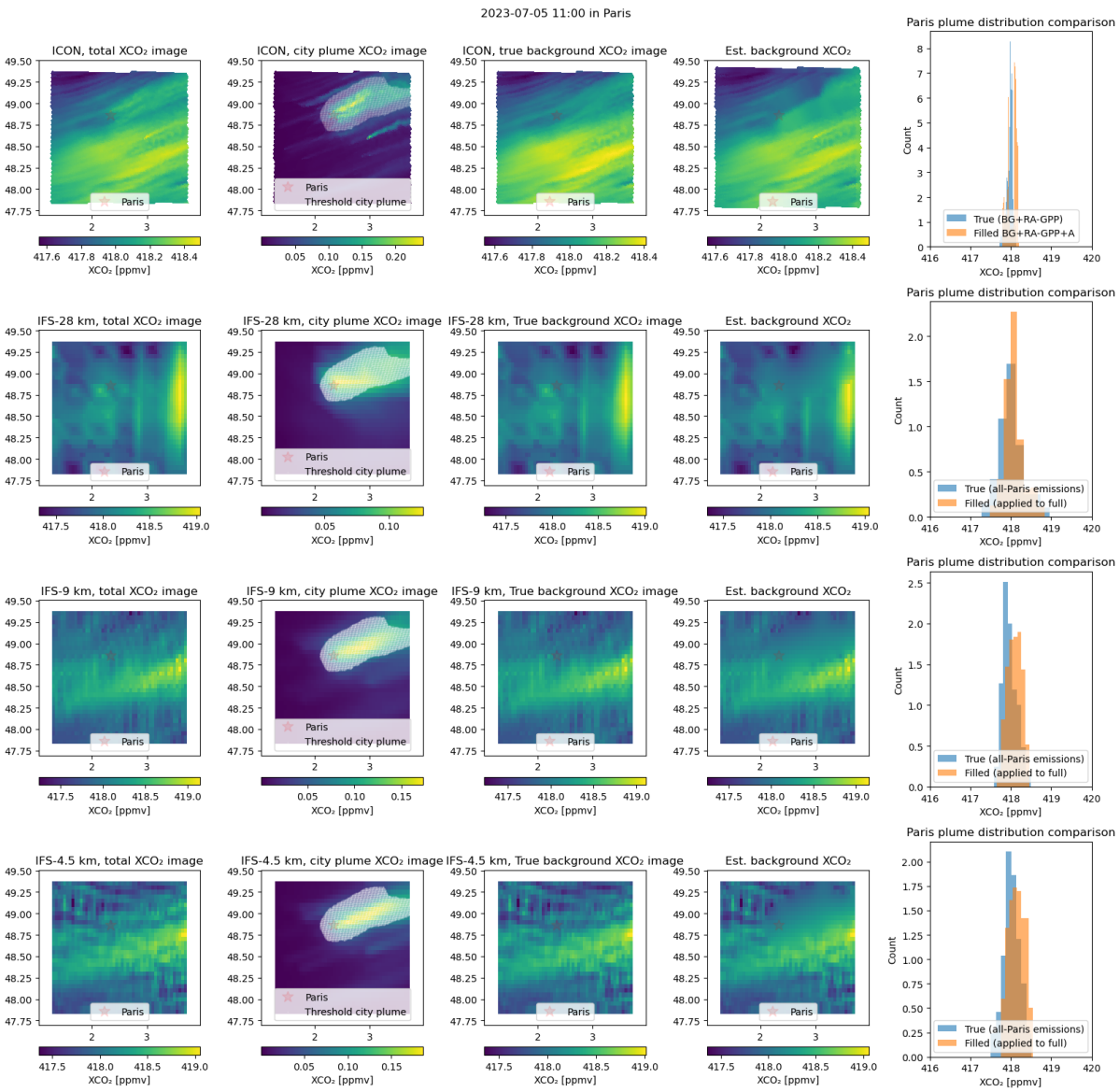

Figure 17: An example of the Paris city plume. The top row corresponds to ICON-ART output, while the other rows correspond to IFS at 28 km, 9 km, and 4.5 km, respectively. The first column presents the 'total' XCO_2 images around the source. The second column presents the derived XCO_2 images of the city plume. The third column presents the true (simulated) 'background' XCO_2 images. The fourth column presents reconstructed background XCO_2 images derived from the 'total' XCO_2 image. The fifth column displays the distribution of the true (simulated, for that method) 'background' XCO_2 image and the filled (i.e., reconstructed) section. Note that the color scales are not unified across the various images.

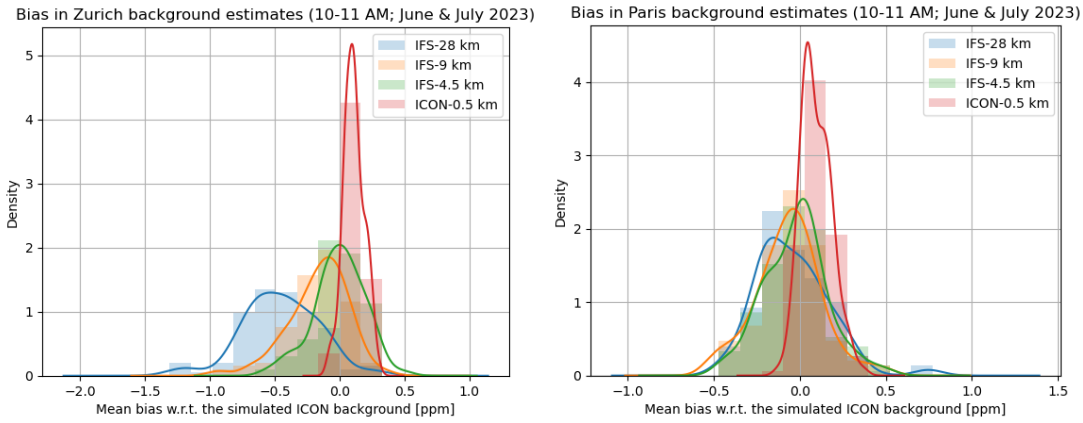

Figure 18: Mean bias relative to the simulated ICON background during 10 and 11 AM cases for Zurich (left) and Paris (right).

The results reveal differences between the cities. For both Zurich and Paris, the IFS simulations at 9 km and 4.5 km resolution produce a smaller bias in the reconstructed background, indicating better performance at higher resolutions. In contrast, for Zurich the lower resolution IFS simulation of 28 km resulted in a larger mean bias and a broader spread can be observed. This suggests we can accurately estimate the size of the bias, indicating that a correction of this bias might be possible in the future. Thus, the effectiveness of the background estimation methods may vary significantly per city (the urban setting and local atmospheric dynamics, as well as how well these factors are represented by the various models). A clear observation is that the 9 km and 4.5 km resolution IFS simulations yield nearly identical results, and their bias distribution peak near zero; therefore, a 9 km model simulation is preferred over a 4.5 km model simulation when considering model efficiency. The smallest spread in the bias is observed for the 0.5-1 km ICON-ART simulations. However, we have to keep in mind that for the latter, we are comparing ICON-ART against ICON-ART, whereas previously we compared ICON-ART against IFS – this means there is a remaining aspect of inter-model differences that complicate this analysis. Despite the inter-model differences, it is reassuring that the bias of the background reconstructions from the 9 km and 4.5 km resolution IFS simulations are much closer to the ICON-ART reconstruction than the 28 km IFS simulation.

An additional metric we can consider is the Pearson's correlation coefficient. By interpolating the reconstructed backgrounds of IFS onto the ICON grid, we can compute the correlation between the reconstructed background and the ICON-simulated 'true' background within the plume. What is clear here, again, is that the ICON results are much correlated with the 'true' background (averaging at about r≈60%) compared to the IFS computed fields (which have a peak at about r≈0% for Zurich and one between 40% and 50% for Paris). For Zurich, the IFS-9 km and IFS-4.5 km results have an average Pearson's correlation coefficient larger than 0. For Paris, the IFS reconstructions are much better correlated with the ICON output, probably due to the simpler topography compared to Zurich. In a way, this is of little surprise that IFS cannot reproduce the fine features associated with the background in Zurich, as the ICON output may contain a lot finer spatially resolved features which cannot be reconstructed using a normalized convolution at all. Hence, although the IFS fields generally have a relatively small mean bias, they do not manage to resolve the spatial heterogeneity of the (ICON-simulated) 'true' background for a small city with complex topography.

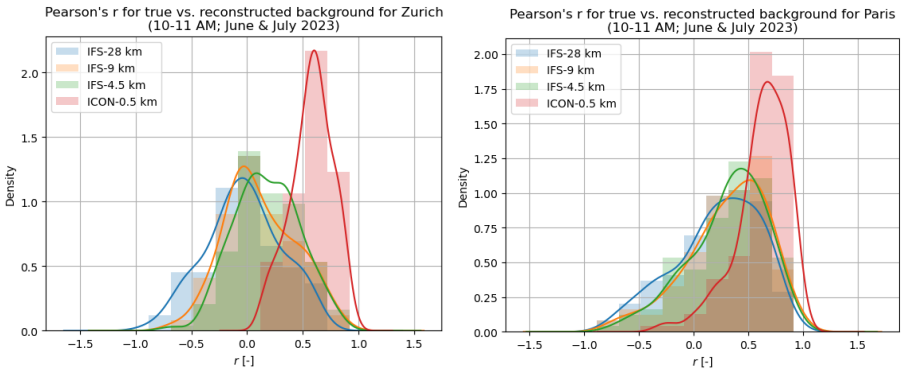

Figure 19. Pearson's correlation coefficients comparing the reconstructed background using normalized convolution to the 'true' background as computed with ICON-ART for all simulations.

The ICON-ART simulations enable an inspection of the VPRM outputs within the selected plume region as they are modelled by a separate tracer. What was referred to as "background" in other sections is more appropriately decomposed here into "boundary + VPRM". The "boundary" is the XCO_2 transported from outside the domain, while VPRM is the net ecosystem exchange within the domain. Hence, while a particular distribution of standard deviations emerges when collating information for just the VPRM component, a slightly larger standard deviation arises when the boundary is included. This indicates that the boundary is heterogeneous, leading to a modest increase in the XCO_2 standard deviation. Clearly, for the two cities considered here, the primary contribution to variations in the background is due to vegetation factors. For both cities, the signal due to the vegetation is of a magnitude comparable to the city plume – with reverse signs, as in the summer months we have net uptake of CO_2 due to vegetation.

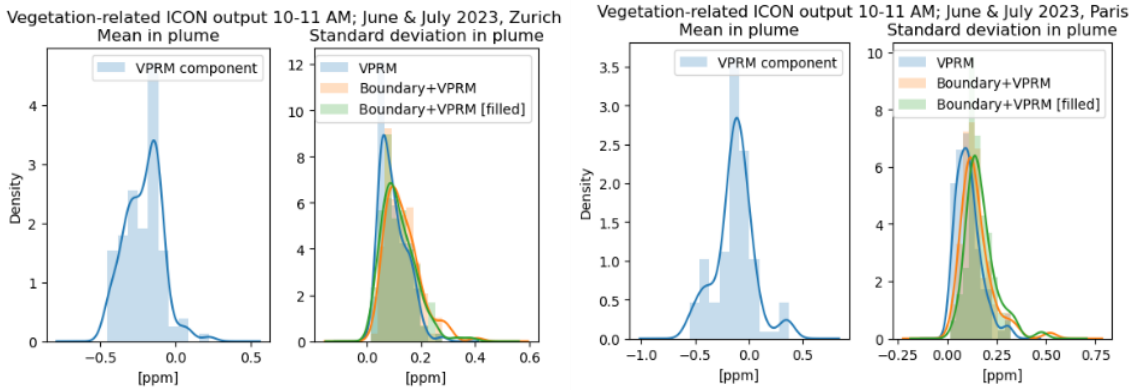

Figure 20. The mean and standard deviation of the simulated ICON VPRM component in the plume, along with the full background, are presented for 10 and 11 AM cases in Zurich (left) and Paris (right).

An example of the VPRM component is given in Figure 21. It is clear that the VPRM component shows considerable small-scale heterogeneities, and that it has a considerable effect on the resulting net effect when summing the city plume (as seen on the third panel). We can compute the estimated VPRM from the background estimation by taking the difference (estimated VPRM = estimated background – (true background – true VPRM)). Hence, we would perfectly recover the VPRM component if our estimated background would equal the true background. As we did not recover the background perfectly, it is no surprise that the

reconstructed VPRM component shows (spatially varying) deviations from the true VPRM, with regions of both strong over- and underestimation. Hence, even if the net mean bias in the plume is positive or small (in the case depicted here, the mean of BG_reconstructed – BG_true = 0.05 ppmv), we may still have considerable spatial error patterns.

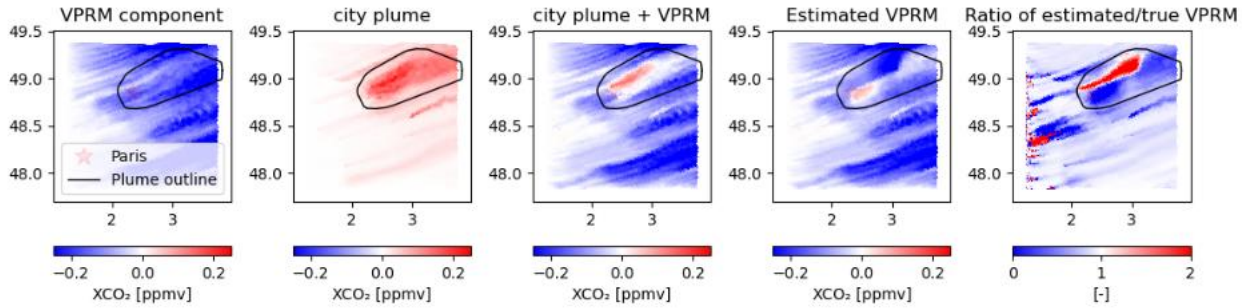

Figure 21. Map of the ICON-ART simulated VPRM component and city plume for the same case as shown in Figure 17. The third panel shows the sum of the first two panels. The fourth panel shows the estimated VPRM due to the background estimation, and the fifth panel shows the ratio between the estimated and the true VPRM for this case.

We can make a scatter plot combining the various bits of information shown before, looking at merely the ICON-ART data, and see that there is relatively little correlation between (1) the mean bias, (2) the Pearson's correlation coefficient, and (3) the standard deviation in the VPRM field within the plume. There is a small apparent correlation between the VPRM standard deviation (color) and the mean bias (vertical axis), where the standard deviation is smaller for a mean bias around 0 ppmv – meaning that for a more homogenous background (as implied by the small standard deviation in the VPRM component), it is easier to reconstruct the background using normalized convolution.

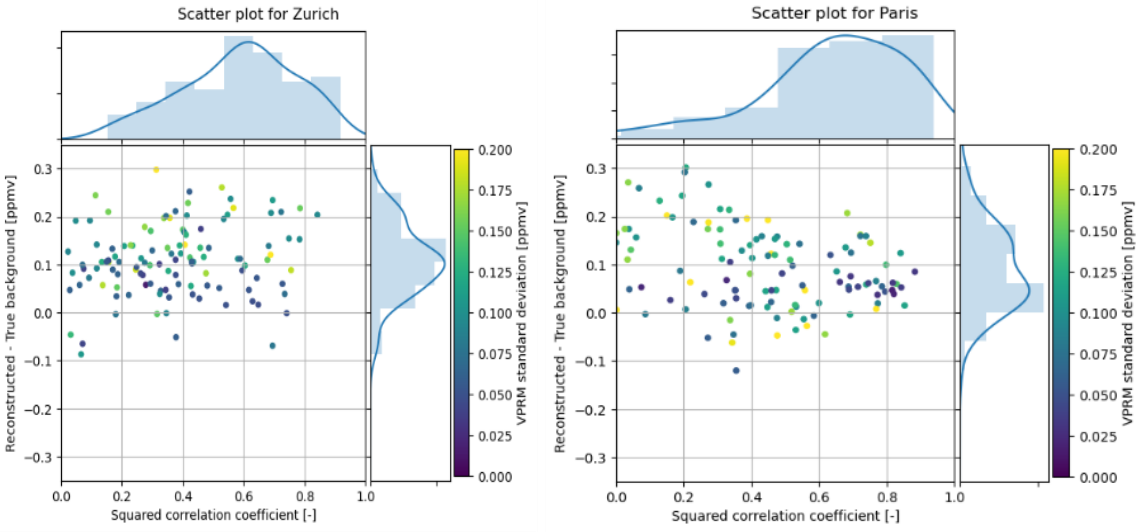

Figure 22. Scatter plot of the ICON-ART simulated fields for Zurich (left) and Paris (right). The vertical axis represents the bias between the reconstructed and true background, while the horizontal axis represents the squared correlation coefficient, and the colors represent the mean VPRM strength within the plume region.

Based on the above observations (i.e., that the VPRM is the main contributor to the heterogeneity of the background of the city plume, and its spatial distribution being hard to recover with a normalized convolution approach), it would be great if there was a data product with a reliable high-resolution VPRM simulation instead, as this would greatly improve the possibility of estimating the true background.

6 Recommendations

From the work done with the ORCHIDEE land surface model, presented in Section 5.1, we may recommend the following developments of the ECLand model with respect to the assimilation of AGB data derived from satellite instruments:

- simulate separately at least: **above- vs. below-ground** biomass and leaf mass with a C allocation scheme that is flexible enough to simulate realistic growth of AGB as a function of the age of the stand.
- simulate **forest age** evolution and recruitment which is critical to explain current forest C sink (especially over Europe where forest are recovering from past disturbances).
- simulate **mortality** either statistically or using a process-based approach. In this later case, it would allow to simulate the impact of extreme events (wildfires, high-speed wind, pests, etc.) on AGB rapid changes.

The results of the ISBA above- and below-ground biomass simulations presented in Section 5.2 show that the simple approach used in ISBA may enable the sequential assimilation of above-ground biomass observations using Kalman filtering. This data-driven approach would not require long assimilation windows and could be easily implemented.

From the D&B/TCCAS experiments presented in Section 5.3 we can recommend that using multiple data streams simultaneously provides a stronger constrain on the terrestrial carbon cycle as illustrated by the comparison against independent in-situ data. In this approach, the use of one long assimilation window is preferable for constraining slowly changing quantities in the terrestrial carbon cycle such as the size of the slow carbon pools (woody biomass and soil organic matter) along with rapidly changing quantities.

The comparison of IFS and ICON ART simulations suggests that a resolution of 9 km provides an advantage in resolving urban CO_2 plumes, producing smaller biases than 28 km resolution. In contrast, the benefits of a resolution of 4.5 km seem to be less significant. As biospheric fluxes appear to be the main source of background CO_2 heterogeneity, developing and comparing different approaches for modelling biospheric fluxes at CO2M spatial resolution (2 km) is recommended, as these could substantially reduce uncertainties in city-scale CO_2 emission estimates.

7 Conclusion

The objective of this report was to summarise the results of Task 4.4, which focused on exploring the potential integration of new satellite products in CO2MVS. Recommendations were provided for the development of ECLand and the IFS to enhance biomass and biospheric fluxes in forest areas. The required evolution of the IFS in CO2MVS at city scale was based on ICON-ART and IFS simulations. Modelling and data assimilation over forests demonstrate that simple approaches can be implemented for monitoring carbon pools and fluxes. However, a deeper understanding of disturbances and carbon pool history would require a more complex modelling approach. At the city level, it has been demonstrated that an accurate representation of the vegetation within and around cities is required to represent the background CO_2 heterogeneity.

8 References

Agustí-Panareda, A., Diamantakis, M., Massart, S., Chevallier, F., Muñoz-Sabater, J., Barré, J., Curcoll, R., Engelen, R., Langerock, B., Law, R. M., Loh, Z., Morguí, J. A., Parrington, M., Peuch, V.-H., Ramonet, M., Roehl, C., Vermeulen, A. T., Warneke, T., and Wunch, D.: Modelling CO2 weather – why horizontal resolution matters, Atmos. Chem. Phys., 19, 7347–7376, https://doi.org/10.5194/acp-19-7347-2019, 2019.

Agustí-Panareda, A., Massart, S., Chevallier, F., Balsamo, G., Boussetta, S., Dutra, E., and Beljaars, A.: A biogenic CO2 flux adjustment scheme for the mitigation of large-scale biases in global atmospheric CO2 analyses and forecasts, Atmos. Chem. Phys., 16, 10399–10418, https://doi.org/10.5194/acp-16-10399-2016, 2016.

Ardaneh, K., Maignan, F., Luyssaert, S., Peylin, P., and Boucher, O.: Fast, and Accurate Radiative Transfer for Land Surface Models, Procedia Computer Science, 255, 3–12, 2025.

Bacour, C., MacBean, N., Chevallier, F., Léonard, S., Koffi, E. N., and Peylin, P.: Assimilation of multiple datasets results in large differences in regional- to global-scale NEE and GPP budgets simulated by a terrestrial biosphere model, Biogeosciences, 20, 1089–1111, doi: 10.5194/bg-20-1089-2023

Bastrikov, V., MacBean, N., Bacour, C., Santaren, D., Kuppel, S., and Peylin, P.: Land surface model parameter optimisation using in situ flux data: comparison of gradient-based versus random search algorithms (a case study using ORCHIDEE v1.9.5.2), Geosci. Model Dev., 11, 4739-4754, doi: 10.5194/gmd-11-4739-2018d

Boussetta, S.; Balsamo, G.; Arduini, G.; Dutra, E.; McNorton, J.; Choulga, M.; Agustí-Panareda, A.; Beljaars, A.; Wedi, N.; Munõz-Sabater, J.; et al. ECLand: The ECMWF Land Surface Modelling System. Atmosphere 2021, 12, 723. https://doi.org/10.3390/atmos12060723

CAMS: Global inversion-optimised greenhouse gas fluxes and concentrations [dataset], https://ads.atmosphere.copernicus.eu/datasets/cams-global-greenhouse-gas-inversion (last access: December 2025), 2025.

Chen, Y.-Y., Huang, W., Cheng, C.-T., Hong, J.-S., Yeh, F.-L., and Luyssaert, S.: Simulation of the impact of environmental disturbances on forest biomass in Taiwan, Journal of Geophysical Research: Biogeosciences, 127, e2021JG006 519, 2022.

Eckes-Shephard, A. H., Argles, A. P., Brzeziecki, B., Cox, P. M., De Kauwe, M. G., Esquivel-Muelbert, A., Fisher, R. A., Hurtt, G. C., Knauer, J., Koven, C. D., et al.: Demography, dynamics and data: building confidence for simulating changes in the world's forests, New Phytologist, 2025.

Faroux, S. and Kaptué Tchuenté, A. T. and Roujean, J.-L. and Masson, V. and Martin, E. and Le Moigne, P., ECOCLIMAP-II/Europe: a twofold database of ecosystems and surface parameters at 1 km resolution based on satellite information for use in land surface, meteorological and climate models, Geoscientific Model Development, volume 6, 2013, 2, 563–582, https://gmd.copernicus.org/articles/6/563/2013/, doi: 10.5194/gmd-6-563-2013.

Gibelin, A. L., Calvet, J. C., Roujean, J. L., Jarlan, L., and Los, S. O.: Ability of the land surface model ISBA-A-gs to simulate leaf area index at the global scale: Comparison with satellites products. Journal of Geophysical Research: Atmospheres, 111(D18), 2006. https://doi.org/10.1029/2005JD006691

Goldberg, D. (1989). *Genetic algorithms in search optimization and machine learning.* Reading, MA: Addison-Wesley.

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.

Guanter, L., Bacour, C., Schneider, A., Aben, I., Van Kempen, T. A., Maignan, F., Retscher, C., Köhler, P., Frankenberg, C., Joiner, J. and Zhang, Y.: The TROPOSIF global sun-induced fluorescence dataset from the Sentinel-5P TROPOMI mission, Earth Syst. Sci. Data, 13(11), 5423–5440, https://doi.org/10.5194/essd-13-5423-2021, 2021

Huang, Y. and Ciais, P. and Santoro, M. and Makowski, D. and Chave, J. and Schepaschenko, D. and Abramoff, R. Z. and Goll, D. S. and Yang, H. and Chen, Y. and Wei, W. and Piao, S., A global map of root biomass across the world's forests, Earth System Science Data volume 13, 2021, 9, 4263–4274, https://essd.copernicus.org/articles/13/4263/2021/, doi:10.5194/essd-13-4263-2021

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, doi:10.1109/JPROC.2010.2043032, 2010.

Knorr, W.: Annual and interannual CO_2 exchanges of the terrestrial biosphere: process-based simulations and uncertainties, Glob. Ecol. Biogeogr., 9, 225–252, 2000. https://doi.org/10.1046/j.1365-2699.2000.00159.x

Knorr, W., Williams, M., Thum, T., Kaminski, T., Voßbeck, M., Scholze, M., Quaife, T., Smallman, T. L., Steele-Dunne, S. C., Vreugdenhil, M., Green, T., Zaehle, S., Aurela, M., Bouvet, A., Bueechi, E., Dorigo, W., El-Madany, T. S., Migliavacca, M., Honkanen, M., Kerr, Y. H., Kontu, A., Lemmetyinen, J., Lindqvist, H., Mialon, A., Miinalainen, T., Pique, G., Ojasalo, A., Quegan, S., Rayner, P. J., Reyes-Muñoz, P., Rodríguez-Fernández, N., Schwank, M., Verrelst, J., Zhu, S., Schüttemeyer, D., and Drusch, M.: A comprehensive land-surface vegetation model for multi-stream data assimilation, D&B v1.0, Geosci. Model Dev., 18, 2137–2159, 2025. https://doi.org/10.5194/gmd-18-2137-2025

Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S. and Prentice, I. C.: A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system, Global Biogeochem. Cycles, 19(1), 1–33, https://doi.org/10.1145/2939672.293978510.1029/2003GB002199, 2005.

Li, R., Lombardozzi, D., Shi, M., Frankenberg, C., Parazoo, N. C., K¨ohler, P., Yi, K., Guan, K., and Yang, X.: Representation of leaf-to-canopy radiative transfer processes improves simulation of far-red solar-induced chlorophyll fluorescence in the community land model version 5, Journal of advances in modeling earth systems, 14, e2021MS002 747, 2022

Marie, G., Jeong, J., Jactel, H., Petter, G., Cailleret, M., McGrath, M., Bastrikov, V., Ghattas, J., Guenet, B., Lansø, A.-S., et al.: Simulating Bark Beetle Outbreak Dynamics and their Influence on Carbon Balance Estimates with ORCHIDEE r7791, EGUsphere, 2023, 1–35,15 2023.

MacBean, N., Bacour, C., Raoult, N., Bastrikov, V., Koffi, E. N., Kuppel, S., Maignan, F., Ottlé, C., Peaucelle, M., Santaren, D., and Peylin, P.: Quantifying and Reducing Uncertainty in Global Carbon Cycle Predictions: Lessons and Perspectives From 15 Years of Data Assimilation Studies with the ORCHIDEE Terrestrial Biosphere Model. Global Biogeochemical Cycles, 36, e2021GB007177, doi: 10.1029/2021GB007177

Naudts, K., Ryder, J., McGrath, M. J., Otto, J., Chen, Y., Valade, A., Bellasen, V., Berhongaray, G., Bönisch, G., Campioli, M., Ghattas, J., De Groote, T., Haverd, V., Kattge, J., MacBean, N., Maignan, F., Merilä, P., Penuelas, J., Peylin, P., Pinty, B., Pretzsch, H., Schulze, E. D., Solyga, D., Vuichard, N., Yan, Y., and Luyssaert, S.: A vertically discretised canopy description for ORCHIDEE (SVN r2290) and the modifications to the energy, water

and carbon fluxes, Geoscientific Model Development, 8, 2035–2065, https://doi.org/10.5194/gmd-8-202035-2015, http://www.geosci-model-dev.net/8/2035/2015/, 2015.

Nelson, J. A., Walther, S., Gans, F., Kraft, B., Weber, U., Novick, K., ... & Jung, M. (2024). X-BASE: the first terrestrial carbon and water flux products from an extended data-driven scaling framework, FLUXCOM-X. *Biogeosciences*, *21*(22), 5079-5115.

Pinty, B., Lavergne, T., Voßbeck, M., Kaminski, T., Aussedat, O., Giering, R., Gobron, N., Taberner, M., Verstraete, M. M., and Widlowski, J.-L.: Retrieving Surface Parameters for Climate Models from MODIS-MISR Albedo Products, J. Geophys. Res.-Atmos., 112, D10116, doi:10.1029/2006JD008105, 2007.

Pinty, B., Clerici, M., Andredakis, I., Kaminski, T., Taberner, M., Verstraete, M. M., Gobron, N., Plummer, S., and Widlowski, J.-L.: Exploiting the MODIS albedos with the Two-stream Inversion Package (JRC-TIP): 2. Fractions of transmitted and absorbed fluxes in the vegetation and soil layers, J. Geophys. Res.-Atmos., 116, D09106, doi:10.1029/2010JD015373, 2011.

Ponomarev, N., Steiner, M., Koene, E., Rubli, P., Grange, S., Constantin, L., Ramonet, M., David, L., Emmenegger, L., and Brunner, D.: Estimation of CO2 fluxes in the cities of Zurich and Paris using the ICON-ART CTDAS inverse modelling framework, EGUsphere [preprint], https://doi.org/10.5194/egusphere-2025-3668, 2025.

Quaife, T. L. (2025). A two stream radiative transfer model for vertically inhomogeneous vegetation canopies including internal emission. Journal of Advances in Modeling Earth Systems, 17, e2024MS004712.https://doi.org/10.1029/2024MS004712

Santoro M., Cartus, O., Quegan, S., Kay H., Lucas, R. M., Araza, A., Herold, M., Labrière, N., Chave, J., Rosenqvist, Å., Tadono, T., Kobayashi, K., Kellndorfer, J., Avitabile, V., Brown, H., Carreiras, J., Campbell, M. J., Cavlovic, J., Conceição Bispo, P., Gilani, H., Khan, M. L., Kumar, A., Lewis, S. L., Jingjing Liang, J., Mitchard, E. T. A., Pacheco-Pascagaza, A. M., Phillips, O. L., Ryan, C. M., Saikia, P., Schepaschenko, D., Sukhdeo, H., Verbeeck, H., Vieilledent, G., Wijaya, A., Willcock, S. and Seifert, F.M.: Design and performance of the Climate Change Initiative Biomass global retrieval algorithm. Science of Remote Sensing, 2024. https://doi.org/10.1016/j.srs.2024.100169

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

Shinozaki, K., Yoda, K., Hozumi, K., and Kira, T.: A quantitative Analysis of Plant Form-The Pipe Model Theory, The Ecological Society of Japan, 14, 97–105, 1964.

Soulie, A., Granier, C., Darras, S., Zilbermann, N., Doumbia, T., Guevara, M., Jalkanen, J.-P., Keita, S., Liousse, C., Crippa, M., Guizzardi, D., Hoesly, R., and Smith, S. J.: Global anthropogenic emissions (CAMS-GLOB-ANT) for the Copernicus Atmosphere Monitoring Service simulations of air quality forecasts and reanalyses, Earth Syst. Sci. Data, 16, 2261–2279, https://doi.org/10.5194/essd-16-2261-2024, 2024.

The Inversion Lab. “In Situ Observations Used in GMD Manuscript "A Comprehensive Land Surface Vegetation Model for Multi-stream Data Assimilation, D&B V1.0"”. Zenodo, July 11, 2024. https://doi.org/10.5281/zenodo.12725765.

van der Tol, C., Berry, J. A., Campbell, P. K. E., and Rascher, U.: Models of fluorescence and photosynthesis for interpreting measurements of solar-induced chlorophyll fluorescence, Journal of Geophysical Research: Biogeosciences, 119, 2312–2327, doi:10.1002/2014JG002713, URL https://onlinelibrary.wiley.com/doi/abs/10.1002/2014JG002713, 2014.

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, https://doi.org/10.5194/gmd-12-4751-2019, https://gmd.copernicus.org/articles/12/4751/2019/, 2019.

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. https://doi.org/10.1111/j.1365-2486.2004.00891.x

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.

Document History

Version	Author(s)	Date	Changes
1.1	Jean-Christophe Calvet	07.10.2025	
1.2	Thomas Kaminski and Marko Scholze Gerrit Kuhlmann and Erik Koene Pierre Vanderbecken	24.11.2025	D&B-related text ICON-ART text ISBA text
1.3	Philippe Peylin	01.12.2025	ORCHIDEE text
1.4	Anna Agusti-Panareda	04.12.2025	IFS text
2.0	All	17.12.2025	Final revisions

Internal Review History

Internal Reviewers	Date	Comments
Matt Rigby	11/12/2025	Comment on document text