Sense4Fire CCN2

Sentinel-based fuel, fire and emissions products to constrain the changing role of vegetation fires in the global carbon cycle

ESA Contract Number: 4000134840/21/I-NB

Algorithm Theoretical Baseline Document Version 4.1 (ATBDv4.1)

A revised full description of the **GFA-S4F-v0.3** and **TUD-S4F-v0.3** approaches with the near real-time (**TUD-S4F-v0.3-NRT**) and high resolution (**TUD-S4F-v0.3-HR**) setups, and of the **KNMI-S5p-v0.2** evaluation approach

6 November 2026

Prepared by:

Matthias Forkel[1], Daniel Kinalczyk[1], Niels Andela[2], Xiao Liu, Christopher Marrs[1], Dave van Wees[2], Jos de Laat[3]

With contributions from:

Martin Böttcher[4], Gunnar Brandt[4], Johanna Kranz[1], Christine Wessollek[1]

(1) TUD Dresden University of Technology, Faculty of Environmental Sciences, Dresden, Germany
(2) BeZero Carbon Ltd., London, UK
(3) Royal Netherlands Meteorological Institute (KNMI), De Bilt, The Netherlands
(4) Brockmann Consult GmbH, Hamburg, Germany

Contents

Contents 2
Figures 4
Tables 6
1 Introduction 8
1.1 Previous developments in Sense4Fire 8
1.2 Scope of this ATBD 9
2 GFA-S4F-v0.3 12
2.1 Overview and updates 12
2.2 Input datasets 12
2.2.1 VIIRS active fire observations 12
2.2.2 ESA FireCCISFD20 burnt area 13
2.2.3 ESA WorldCover v200 13
2.3 Fire event monitoring 13
2.4 Fire types, burned area and emissions 15
2.4.1 Burned area scaling factors 15
2.4.2 Emissions 17
2.4.3 Benchmarking against GFED5 18
3 TUD-S4F-v0.3 20
3.1 Overview 20
3.1.1 TUD-S4F approach in a nutshell 20
3.1.2 Summary of updates in version TUD-S4F-v0.3 21
3.1.3 Latency and maturity setups of the TUD-S4F-v0.3 approach 23
3.2 Input datasets and data pre-processing 24
3.2.1 Data pre-processing for all datasets 24
3.2.2 CGLS Proba-V/Sentinel-3 leaf area index (LAI) 26
3.2.3 LAI time series filtering, gap-filling and forecasting 26
3.2.4 MODIS and Sentinel-3 burnt area (ESA FireCCI51 and FireCCIS311) 32
3.2.5 Sentinel-2 Burnt Area (S4F-FireCCIS211/ FireCCISFD20 / FireCCIS2v1) 33
3.2.6 ESA WorldCover 34
3.2.7 ESA CCI HRLC land cover 34
3.2.8 ESA CCI above-ground biomass 35
3.2.9 Canopy height 35
3.2.10 ALOS-2 PALSAR-2 mosaic datasets 36

3.2.11 CGLS ASCAT soil water index 36
3.2.12 VOD2LFMC live fuel moisture content 36
3.2.13 Databases of ground observations 37
3.3 TUD-S4F-v0.3 model description and data-model fusion steps 37
3.3.1 Fractional cover and LAI of vegetation types 38
3.3.2 Canopy height 43
3.3.3 Biomass and allometry of trees 44
3.3.4 Biomass and allometry of surface live woody vegetation 45
3.3.5 Herbaceous biomass 47
3.3.6 Biomass turnover and harvest 48
3.3.7 Fuel moisture content 49
3.3.8 Combustion completeness 50
3.3.9 Surface fuel dynamics 51
3.3.10 Dry matter burnt and fire behaviour 52
3.3.11 Combustion, emission factors, and fire radiative energy 54
3.3.12 Local calibration and spatial prediction of model parameters 57
3.4 TUD-S4F-v0.3-NRT near-real time setup 59
3.5 TUD-S4F-v0.3-HR high resolution setup 60
3.5.1 Downscaling of CGLS LAI (TUD-S4F-0.3-HR-S3) 60
3.5.2 Retrieval of LAI from Sentinel-2 (TUD-S4F-0.3-HR-S2) 61
3.6 Downscaling of ESA CCI above-ground biomass 63
4 KNMI-S5p validation approach 65
4.1 Overview 65
4.1.1 Summary of updates 65
4.1.2 Sentinel-5p/TROPOMI and wildfire emissions 65
4.1.3 Wildfire emissions evaluation with IFS-COMPO and TROPOMI 66
4.1.4 Extending the evaluation of Sense4Fire wildfire emissions 67
4.2 Input data 68
4.3 IFS model setup 69
4.4 Preliminary 2019-2023 GFAS IFS-COMPO results 70
5 Appendix: KNMI-S5p β-method 74
5.1 Introduction 74
5.2 Inverse modelling of trace gas emissions 74
5.3 Alternative approach: β-method short lived trace gases NO_2/NO_x 75

5.4 Limitations of the β-method for long lived trace gases and possible modifications 75
5.5 The standard β-Method 76
5.6 Limitations of the β-method for CO 76
5.7 Proposed Modification: Incorporating a Pseudo-Influx Term 77
5.7.1 Conceptual Basis 77
5.7.2 Extracting the Local Contribution 77
5.7.3 Corrected Local β-Factor 78
5.7.4 Emission Update Formula 78
5.8 Discussion of Limitations and Caveats 78
5.8.1 Transport Anisotropy 78
5.8.2 Chemical Coupling 78
5.8.3 Zero Emission Cells Are Not Chemically Passive 78
5.8.4 A.5.4 Large Spatial Influence 79
5.8.5 Practical Usefulness 79
5.9 Conclusion 79
References 80

Figures

Figure 1: Overview about the study regions in Sense4Fire. 8
Figure 2: Example of the spatial and temporal evolution of VIIRS active fire detections for Amazon deforestation and forest fires provides an estimate of daily burned area needed for the Global Fire Atlas algorithm. Day of active fire detections (circles) from both VIIRS instruments for (a) a large forest fire and (c) a deforestation fire in Brazil. Panels (b) and (d) show the estimated daily fire progression on a 0.005° (~550 m) grid. The background depicts fractional tree cover (Hansen et al., 2013). Black lines indicate the fire perimeters derived from active fire detections and red lines show the corresponding burned area estimates from the MCD64A1 C6 product (Giglio et al., 2018) in (a) and (c) and Landsat-derived burned area estimates from MapBiomas in (b) and (d). 14
Figure 3: Scaling grid for 2019 southern Africa using an iterative region-growing algorithm, based on the ratio of FireCCISFD20 and VIIRS SNPP and NOAA-20 active fire detections. 16
Figure 4: Comparison of FireCCISFD11 burned area for 2016 and scaled VIIRS SNPP burned area using a 2019-based scaling factor for the southern African domain. 17
Figure 5: Comparison of FireCCISFD20 burned area for 2019 and scaled VIIRS SNPP burned area using a 2016-based scaling factor for the southern African domain. 17

Figure 6: A comparison of burned area and carbon emissions estimates derived using the GFA-S4F-v0.3 and those of GFED5. Note that burned area and fire emissions estimates from GFED5 were not available for the full period. 19
Figure 7: Schematic structure of the TUD-S4F-v0.3 approach to estimate fuel loads, fuel moisture, fuel consumption and fire emissions. Various satellite datasets are used a forcing (top) or for training of machine learning models and the calibration of model parameters (right). 20
Figure 8: Example time series of LAI and vegetation cover for a grid cell at 14.59625°E and 6.7987°S from the TUD-S4F-v0.3-S311 setup. The observed gap-filled and filtered CGLS LAI (red, left) dataset is used together with class frequencies from ESA WorldCover to compute multi-annual dynamics in fractional vegetation cover of trees, herbaceous and SLW vegetation (right) and to split LAI in a woody and herbaceous component (left). 40
Figure 9: Regional example of canopy height from (a) GEDI L3 data and (b) TUD-S4F-v0.3-S311. (c and d) Spatial patterns and histogram of the residuals, and scatter plot (e). 44
Figure 10: Example time series of above-ground biomass (left) and biomass components (right) for a grid cell at 14.59625°E and 6.7987°S from the TUD-S4F-v0.3-S311 setup. (left) AGB from ESA CCI (obs) and from TUD-S4F prior to the calibration of allometry parameters and with optimised allometry parameters (opt). 46
Figure 11: Regional comparison of above-ground biomass from ESA CCI data and TUD-S4F-v0.3-S311. 47
Figure 12: Calibration of herbaceous biomass. Random grid cells from the study region are sampled in a buffer of 200 km around locations for which observations are available from the H14 and vW22 databases (left). Herbaceous biomass is available for 40 sites and the statistical distribution estimated in TUD-S4F is fitted to the observed distribution (lower right) by calibrating the specific leaf area *sla* (upper right). 48
Figure 13: Results of the calibration of parameters for combustion completeness for the TUD-S4F-v0.3-S311 setup against regional distributions of CC from van Wees et al. (2022). 51
Figure 14: Boxplots of dry matter burnt (kg/m^2) for different fuel components as estimated with TUD-S4F-v0.3-S311 for 220 randomly sampled grid cells within a 200 km buffer of field observations from the vW22 dataset. The location of grid cells correspond to the one shown in Section 3.3.5. 54
Figure 15: Overview about the calculation of fire emissions in the TUD-S4F approach. .. 55
Figure 16: Scatterplot and correlation matrix of optimised model parameters and the corresponding RMSE for canopy height, above ground biomass and LFMC from 1000 sampled grid cells from TUD-S4F-v0.3-S311. Orange and blue colours indicate negative and positive correlations, respectively. 58
Figure 17: Comparison of estimated and reference canopy height, woody biomass, and mean LFMC for 1000 randomly sampled grid cells after calibration of each grid cell. Black are all grid cells and blue and red are grid cells for which the optimisation resulted in good and no acceptable performance, respectively. 59

Figure 18: An example of CGLS LAI (left, 0.0025° resolution) and HR S3 LAI (right, 0.0002° resolution). The temporal median values of 280 scenes (from 2014-01-10 to 2021-10-20) are presented. 61
Figure 19: Comparison of the temporal median value (280 scenes from 2014-01-10 to 2021-10-20) of CGLS LAI and resampled HR S3 LAI at a 5° × 5° test site. 61
Figure 20: An example of CGLS LAI (left, 0.0025° resolution) and HR S2 LAI (right, 0.0002° resolution). The temporal median values of 93 scenes (from 2016-04-16 to 2022-10-02) are presented. 62
Figure 21: Comparison of the temporal median value (93 scenes from 2016-04-16 to 2022-10-02) of CGLS LAI and resampled HR S2 LAI at a 0.5° × 0.5° test site. 62
Figure 22: Comparison of the temporal median value (280 scenes from 2014-01-10 to 2021-10-20) of CGLS LAI and resampled HR S3 LAI at the same test site as HR S2 LAI. ... 63
Figure 23: An example of CCI AGB (left, 100 m resolution) and downscaled AGB (right, 20 m resolution). 64
Figure 24: Comparison of CCI AGB and resampled HR AGB at a 1° × 1° test site. 64
Figure 25: Time series of the daily total carbon monoxide mass over Amazonia for the region 80°W-40°W and 20°S-10°N (displayed in figure X02) for the period 2019-2023. The blue line denotes the total CO mass of IFS-COMPO GFAS simulation, collocated with TROPOMI CO total column observations. The red line displays the total CO mass from the TROPOMI CO total column observations themselves. As there can be model grids without data due to cloudiness, hence the area total daily CO mass values are weighted with the ratio of model grid cells with data and the total number of model grid cells. 71
Figure 26: Box region over Amazonia (80°W-40°W and 20°S-10°N) for which the time series in Figure 25 are calculated. 72
Figure 27: As Figure 26 but for the central Africa region in Figure 28. 72
Figure 28: Box region over central Africa (12.2°E-39.7° E and 16.3°S-14.0°N) for which the time series in Figure 27 are calculated. 73

Tables

Table 1: Overview of the study regions and test areas used in Sense4Fire. The African study regions highlighted in bold are used for the approaches described in this ATBD and the accompanying PVRv4. 9
Table 2: Development of the Sense4Fire approaches with the related versions of Algorithm Theoretical Baseline Documents (ATBD), Product Validation Reports (PVR), Impact Assessment Reports (IAR), Databases (DB) and publications. 11
Table 3: Overview of emissions factors used in the GFA-S4F approach. 18
Table 4: Overview of the setups of the TUD-S4F-v0.3 approach. 24
Table 5: Overview of pre-processing steps applied to all datasets. 25

Table 6: Overview of parameters for the TUD-S4F-v0.3 model. “Prior” refers to the initial default value of a parameter that was taken from the literature or estimated from databases. The “Optimised” value was calibrated and used in model runs. 41
Table 7: Overview about the performance of several regression models to predict canopy height from dynamic mean LAI (LAI_{mean}) and fractional vegetation cover of trees and SLW (f_{Tree} and f_{SLW}). Performance metrics were computed between estimated canopy height and the canopy height from GEDI L3 based on a sample of thousands of testing grid cells across the S-Africa study region. AIC is computed based on the training data. 43
Table 8: Example of the calculation of EF_{CO} for the three used fuel compounds and for three different types of combustion efficiency. 56
Table 9: Planned IFS-COMPO model simulations with different fire emission databases and the time period for which the IFS-COMPO model simulations will be performed. ... 69

1 Introduction

1.1 Previous developments in Sense4Fire

The Sense4Fire project, funded by the European Space Agency (ESA), aims to increase the scientific understanding of fire dynamics and their role in the carbon cycle by integrating observations from the Sentinels into new fire products. Fire dynamics encompass a broad range of processes, including pre-fire conditions of the land surface (i.e. fuel loads and fuel moisture), fire behaviour (fire ignitions, spread, speed, size, burnt area, fire type, and radiative power), combustion and production of fire emissions (combustion completeness, dry matter burnt, combustion efficiency, and composition of emissions) and the effect of fire emissions on atmospheric composition.

Previously, the Sense4Fire project provided estimates of fuel loads, live fuel moisture content, fire types, dry matter burnt, emission factors and fire emissions from three complementary approaches in three versions of Databases for four study regions, namely South America (Amazon and Cerrado region), southern Africa, a region in eastern Siberia and southern Europe (Mediterranean) (Figure 1, Table 1). The main focus of the scientific analyses was the Amazon and Cerrado region, which resulted in a series of high-level scientific publications (Andela et al., 2022; Forkel et al., 2025; de Laat et al., 2026).

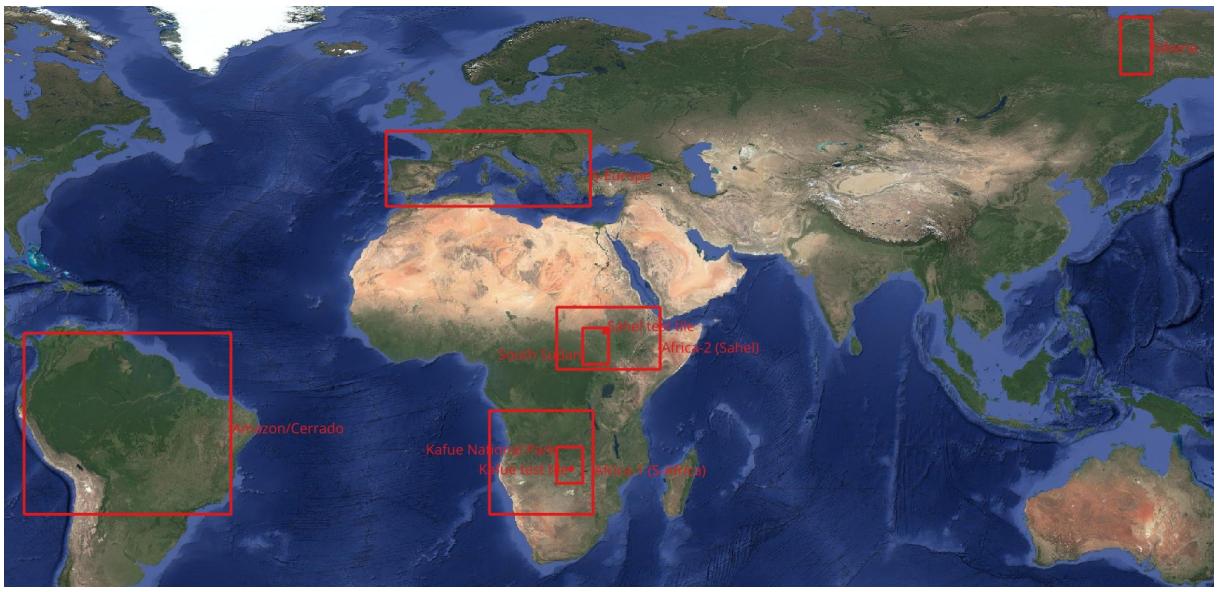

Figure 1: Overview about the study regions in Sense4Fire.

Table 1: Overview of the study regions and test areas used in Sense4Fire. The African study regions highlighted in bold are used for the approaches described in this ATBD and the accompanying PVRv4.

Study region or test area	East/West extent	North/South extent	Included in Database Version
S-Africa: southern Africa study region (Africa-1)	10°E – 30°E	5°S – 25°S	DBv1, DBv2, DBv4
• **Kafue National Park**	23°E – 28°E	12°S – 19°S	DBv4 (high resolution)
Sahel study region (Africa-2)	23°E – 43°E	3°N – 15°N	DBv4
• **South Sudan**	28°E – 33°E	4°N – 11°N	DBv4 (high resolution)
Amazon/Cerrado study region	40°W - 80°W	25°S - 10°N	DBv2, DBv3
Europe study region	10°W – 29.5°E	34.5°N - 49°N	DBv1, DBv2
Siberia test area	132°E - 138°E	60°N - 71°N	DBv1, DBv2

Sense4Fire provides two bottom-up fire emission approaches (GFA-S4F and TUD-S4F) and a set of approaches (summarised as KNMI-S5p) to provide a top-down benchmark of the estimated fire emissions:

- The **GFA-S4F** approach is based on the Global Fire Atlas (GFA) algorithm and uses observations of active fires from the VIIRS instrument and Sentinel-2-derived burnt area maps to provide near-real time estimates of burnt area and performs a mapping of fire types to estimate fire emissions (Andela et al., 2019, 2022).
- The **TUD-S4F** approach is a data-model fusion approach that combines several Earth observation products from the Sentinels, from ESA's Climate Change Initiative (CCI) and from the Copernicus Global Land Service (CGLS) to estimate fuel loads, fuel moisture, fuel consumption, and fire emissions (Forkel et al., 2025).
- **KNMI-S5p** is a set of approaches that uses observations of atmospheric CO and NOx with the Integrated Forecasting Model (IFS) to evaluate bottom-up fire emission estimates. The approaches have been used in Forkel et al. (2025) and in de Laat et al. (2026) to evaluate emissions estimates from GFA-S4F and TUD-S4F, among others.

The development of the three Sense4Fire approaches along with their versioning, application to different study regions, and the related deliverables and publications are shown in Table 2.

1.2 Scope of this ATBD

The fourth version of the Algorithm Theoretical Baseline Document (ATBDv4) provides an updated and complete description of the most recent versions (GFA-S4F-v0.3, TUD-S4F-v0.3, KNMI-S5p-v0.2) of the three approaches used in Sense4Fire.

The three approaches were previously described in ATBDv2.1 and ATBDv3 (Forkel et al., 2023b, 2024) as well as in the corresponding scientific publications (Andela et al., 2022; Forkel et al., 2025; de Laat et al., 2026). ATBDv4 documents advancements since ATBDv3, reflecting the integration of new Earth observation datasets and methodological developments in the approaches. Specifically, it describes the setups of GFA-S4F and TUD-S4F to provide fire emission estimates with low latency, i.e. as **provisional (PRV)** analysis up to the previous year or as **near-real time (NRT)**, here defined as providing fire emission estimates with a latency of <1 month. In addition, ATBDv4 documents the approach to apply the TUD-S4F approach at **high spatial resolution** (**HR**, 20 m).

The ATBDv4 accompanies the fourth version of the Sense4Fire Database (**DBv4**), which applies the approaches to study regions in **southern Africa and the Sahel** (Table 1). ATBDv4 also serves as the methodological foundation for the datasets that will be evaluated and inter-compared in the fourth version of the Product Validation Report (**PVRv4**), including the near-real time and high-resolution setups.

The Sense4Fire project's deliverables, including datasets and scientific publications, are accessible through the project's website (https://sense4fire.eu/).

Table 2: Development of the Sense4Fire approaches with the related versions of Algorithm Theoretical Baseline Documents (ATBD), Product Validation Reports (PVR), Impact Assessment Reports (IAR), Databases (DB) and publications.

Approach and version	Short description	Related ATBDs, PVRs and IARs	Study region and temporal coverage Publication	Related DB	Data publi-cation date
GFA-S4F-v0.3 (this ATBD)	Based on GFA-S4F-v0.1 but with updated estimation of NRT burnt area based on Sentinel-2 for Africa	ATBDv4, PVRv4	S-Africa 2019-2025	DBv4	Nov 2026
			Sahel 2019-2025	DBv4	Nov 2026
GFA-S4F-v0.2	Based on GFA-S4F-v0.1 but with improved parametrisation for NO_x emissions and prolonged time series (2019-2024)	ATBDv2.1, ATBDv3, IARv2.1	Amazon/Cerrado 2019-2023 (Forkel et al., 2025)	DBv2	Oct 2023
			Amazon/Cerrado 2024 (de Laat et al., 2026)	DBv3	Dec 2024
GFA-S4F-v0.1	Original description of the GFA-S4F approach based on Andela et al. (2022)	ATBDv1, PVRv3	Amazon/Cerrado, Europe, S-Africa, and Siberia 2020	DBv1	Nov 2021
TUD-S4F-v0.3 (this ATBD)	Based on TUD-S4F-v0.2 but with new input datasets and representation of surface live woody vegetation	ATBDv4, PVRv4	S-Africa 2014-2024/2025	DBv4	Apr 2026
			Sahel 2014-2024/2025	DBv4	Sep 2026
TUD-S4F-vNRT01	Initial near-real time setup based on a machine learning model trained against TUD-S4F-v0.2	IARv2.1	Amazon/Cerrado 2024 (de Laat et al., 2026)	DBv3	Dec 2024
TUD-S4F-v0.2	Default setup of TUD-S4F for all study regions	ATBDv2.1, ATBDv3, PVRv3, IARv2.1	Amazon/Cerrado (Forkel et al., 2025), Europe, S-Africa, and Siberia 2014-2021	DBv2	Oct 2023
TUD-S4F-v0.1 *(deprecated)*	Initial version of the TUD-S4F approach, replaced by TUD-S4F-v0.2	ATBDv1, PVRv2.1	Amazon/Cerrado, Europe, S-Africa, and Siberia 2014-2021	DBv1	Nov 2021
KNMI-S5p-v0.2 (this ATBD)	Revised setup of the IFS/S5p-based validation approach; no emission estimates	ATBDv4, PVRv4	--	--	--
KNMI-S5p-v0.1	Initial description of the approach, including top-down emission estimates of NO_x and CO	ATBDv2.1, ATBDv3, PVRv3	Amazon/Cerrado (Forkel et al., 2025; de Laat et al., 2026), Europe, S-Africa, and Siberia 2020	DBv1/v2	Nov 2021

2 GFA-S4F-v0.3

2.1 Overview and updates

The Global Fire Atlas provides an object-based approach to estimate global fire emissions. The approach builds upon the recently developed Global Fire Atlas (Andela et al., 2019) and "Amazon dashboard" (Andela et al., 2022), that tracks individual fires and their behaviour (e.g., perimeters, speed, duration, radiative power) based on near real time detection of active fires from the VIIRS sensors on-board NOAA20 and Suomi-NPP. Emissions are modelled through the identification of fire types, and development of fire type specific scaling factors for burned area and fuel consumption.

Here we present a revised model for southern hemisphere Africa that introduces burned area scaling based on the ESA Sentinel-2 burned area product (FireCCISFD20) (Ramo et al., 2021) and subsequent calibration of emissions scaling factors that translate fire radiative energy per area burned (MJ m^{-2}) to fuel consumption (g m^{-2}). We generate results for 2019 – 2025 and compare results against GFED5.

2.2 Input datasets

This section provides an overview of the datasets used. Active fire detections from the VIIRS instruments onboard S-NPP and NOAA20 are used to track individual fire events and characterise their behaviour (2.2.1). Because active fire detections provide only information about actively burning fires at cloud free satellite overpass, we use burned area products from Sentinel-2 to develop scaling factors and estimate the burned area associated with each fire event (2.2.2). Finally, we use land cover information from ESA WorldCover (v200) to classify fires into different fire types (2.2.3).

2.2.1 VIIRS active fire observations

We use VIIRS VNP14IMG 375-m resolution active fire detections from the Suomi-NPP (SNPP) and NOAA-20 satellites (Schroeder et al., 2014). SNPP VIIRS active fire detections are available from February 2012 onwards and NOAA-20 active fire detections are available only from January 2018. The two satellites are separated by half an orbit, allowing them to observe the same fire event approximately 50 minutes apart from different viewing angles, thereby improving the delineation of individual events. The active fire detections were gridded onto an 0.005° (approximately 500 metre resolution) annual global grid, binning all active fire detections falling within each grid cell. We selected this spatial resolution to accommodate typical under-story forest fire spread rates of 100 to 400 m day^{-1} (Balch et al., 2008, 2011), reduce potential effects of geolocation error from the VIIRS Sensor, and for consistency with the TUD-S4F models produced at 0.0025° spatial resolution. Finally, we create a Boolean (values 0 or 1) grid indicating the occurrence of at least one active fire detection per year within each 0.005° grid cell. This Boolean active fire grid was produced for 2019 to 2025 by combining the active fire detections from VIIRS SNPP and NOAA-20. Two additional Boolean active fire grids were produced for the evaluation of burned area scaling factors, the first is based on SNPP VIIRS only for 2016 and 2019, and the second on NOAA-20 VIIRS only for 2019.

2.2.2 ESA FireCCISFD20 burnt area

The FireCCISFD20 (Small Fire Dataset v2.0) product is a burnt area dataset for the year 2019 at 20 m spatial resolution based on the Sentinel-2 MSI (Multi-Spectral Instrument) sensor, developed within the ESA Fire Climate Change Initiative (Fire CCI) project (Chuvieco et al., 2022). This product was specifically designed to exploit the relatively high spatial resolution of the Sentinel-2 MSI sensor, representing a significant improvement in spatial detail compared to the coarser resolution FireCCI51 (~250 m, MODIS-based) and FireCCIS311 (Sentinel-3-based) products. This 20 m spatial resolution enables the detection of small fires that are often omitted by coarse-resolution burnt area products, which is particularly relevant in regions where small-scale agricultural burning and fragmented fire regimes dominate. Here we use the FireCCISFD20 data to develop burned area scaling factors. In addition, we use 2016 data from FireCCISFD11, the successor to FireCCISFD20, to evaluate the predictive performance of our scaling factors for an independent year (as explained in 2.4.1).

2.2.3 ESA WorldCover v200

For the African study region, we separated three fire types based on the WorldCover V200 dataset (Zanaga et al., 2022). If tree cover fraction (WorldCover V200) within the fire perimeter exceeded 50% a fire was classified as woodland fire. Any fire not meeting this criterion, was classified as cropland or grassland and savannah fire, depending on the dominant land cover type within the fire perimeter.

2.3 Fire event monitoring

This section provides an overview of how we identify and track individual fire events. The VIIRS active fire detections provide a point cloud of active fire detections and associated detection dates at a daily temporal resolution. Here we cluster these active fire detections into fire events based on the Fire Atlas algorithm (Andela et al., 2019). The Global Fire Atlas algorithm was initially applied to the MCD64A1 collection 6 burned area product, a daily estimate of global burned area at 500 m resolution (Giglio et al., 2018).

Burned area data have two important limitations compared to active fire detections for the development of an algorithm to track fires and associated emissions. First, burned area data are not available in near-real time, as burned area algorithms rely on time series of data before and after the fire, leading to a typical two to three-month delay in the production of global products. Second, coarse resolution (e.g., 500 m) satellite observations are often unable to accurately map the burned area associated with small fire types or low-intensity forest fires burning beneath a dense forest canopy (Morton et al., 2011).

Previous studies have demonstrated that active fire data can provide a robust estimate of burned area in forested ecosystems based on a persistent thermal signal within each larger grid cell from the slow spread rates and residual smouldering (Oliva and Schroeder, 2015; Veraverbeke et al., 2014). To map individual fire perimeters from VIIRS active fire detections, we first gridded active fire detections at a 0.005° (~550 m) resolution based on

the centre of each fire pixel (Figure 2). We selected this spatial resolution to accommodate typical forest fire spread rates of 300 to 500 m day-1 and reduce potential effects of geolocation error from the VIIRS Sensor.

To prepare the data for the Fire Atlas algorithm, we select the earliest active fire detection within each 550 m grid cell (Figure 2). Following the Global Fire Atlas approach, we then apply a spatial filter to remove inconsistencies in the estimated burn date within each fire to identify the ignition location (Andela et al., 2019). Inconsistencies can originate from gaps in the observation record, e.g., due to cloud cover. An additional threshold is required to separate adjacent fires that burned at different times during the fire season. This threshold sets the maximum number of days for a fire to spread into an adjacent grid cell; here, we set this threshold as ≤5 days after the last active fire detection within any given 550 m grid cell.

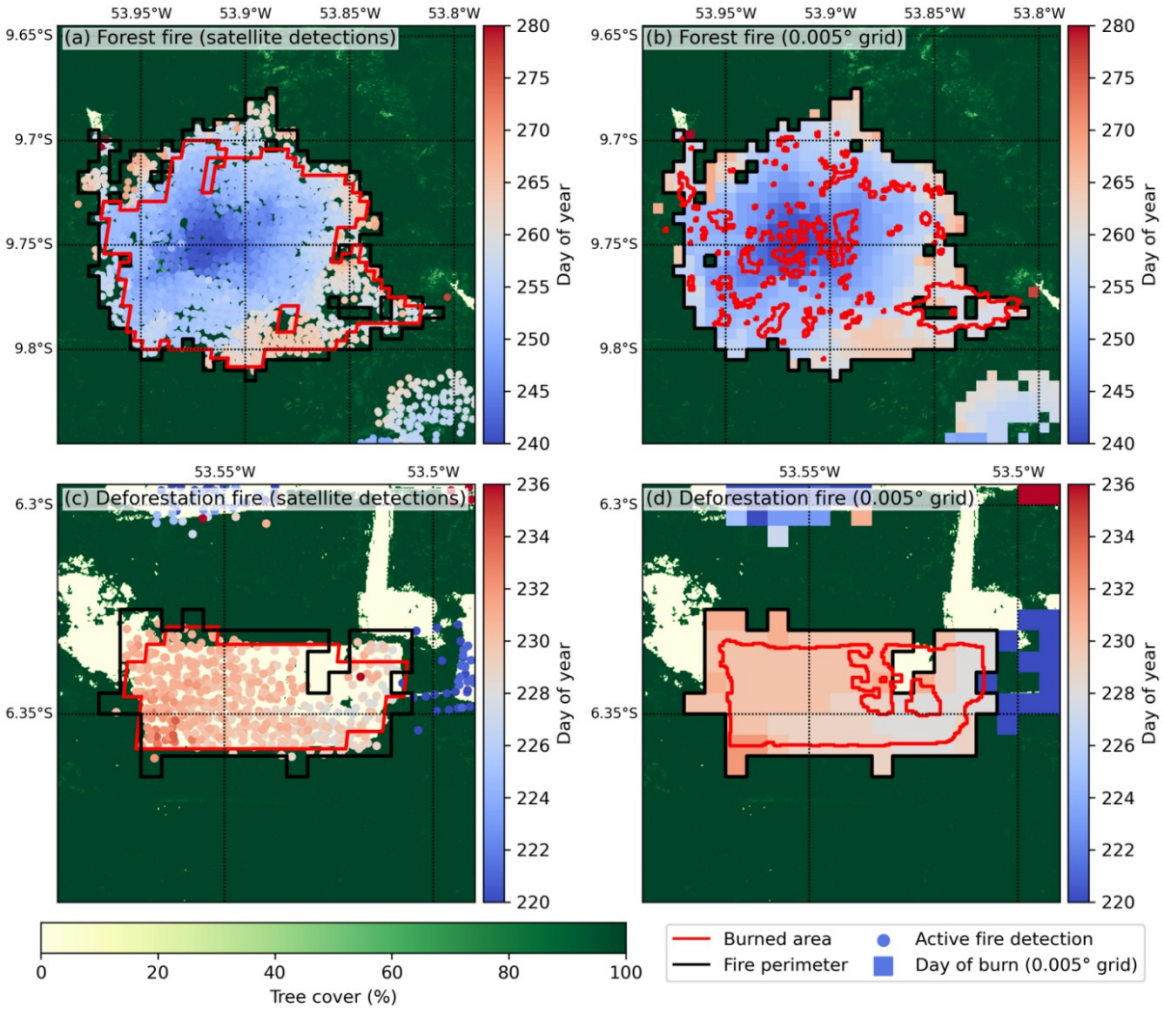

Figure 2: Example of the spatial and temporal evolution of VIIRS active fire detections for Amazon deforestation and forest fires provides an estimate of daily burned area needed for the Global Fire Atlas algorithm. Day of active fire detections (circles) from both VIIRS instruments for (a) a large forest fire and (c) a deforestation fire in Brazil. Panels (b) and (d) show the estimated daily fire progression on a 0.005° (~550 m) grid. The background depicts fractional tree cover (Hansen et al., 2013). Black lines indicate the fire perimeters derived from active fire detections and red lines show the corresponding burned area estimates from the MCD64A1 C6 product (Giglio et al., 2018) in (a) and (c) and Landsat-derived burned area estimates from MapBiomas in (b) and (d).

In Figure 2c and Figure 2d, a deforestation fire in an adjacent field burned into the understorey of neighbouring forest area, resulting in fire detections adjacent to our example fire on day 220, eight days before the example fire was ignited. In this case, the algorithm successfully classified these as two separate fire events despite their spatial proximity. In contrast to forest fires, savannah and grassland fires typically spread several kilometres per day, resulting in artificial fragmentation of individual events in our finer 550 m grid. We therefore use land cover data to distinguish grassland, forest and cropland fires and develop a burned area scaling factor to estimate associated burned areas for each fire type.

2.4 Fire types, burned area and emissions

2.4.1 Burned area scaling factors

Approach: A new approach was developed to estimate historic and near-real time burnt areas using scaling factors and VIIRS active fire detections. Scaling factors were computed based on the ratio between FireCCISFD20 Sentinel-2 burnt area and the Boolean active fire grid based on VIIRS SNPP and NOAA-20 for the year 2019. For this purpose, each Boolean grid cell with value one (i.e. at least on active fire detection within a year) is considered to be fully burned, representing roughly 500 x 500 m = 0.25 km^2 of burnt area at the equator. The scaling factors were based on an iterative region-growing algorithm (like Andela et al. 2022), initially creating scaling factors for all 0.25° grid cells that contained at least 50 km^2 burnt area for both FireCCISFD20 (aggregated to 0.005°) and the Boolean VIIRS active fire grid. Scaling factors were not derived at 0.005° resolution because individual grid cells may contain only a single active fire detection, providing insufficient information to reliably estimate the relationship between active fire detections and burned area. Scaling factors at this resolution would therefore be expected to vary substantially between years, making factors calibrated using a single year less reliable when applied to other years. Next, we clustered the remaining 0.25° grid cells that did not meet the criteria to 0.5° and evaluated the larger regions using the same threshold of 50 km^2. These grid cell aggregation and evaluation steps were repeated at 1°, 2°, and 4° resolutions. Finally, all remaining grid cells were combined into a single continental-scale region for the final evaluation step. Separate scaling factors were computed for three different fire types: forest fires, grassland and savannah fires, and cropland fires. Figure 3shows the resulting scaling grid for 2019, as a composite of the three fire types.

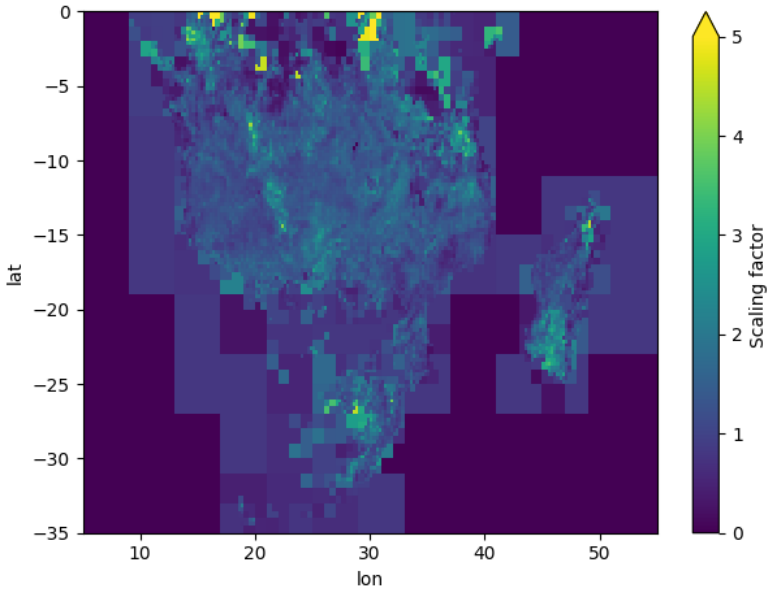

Figure 3: Scaling grid for 2019 southern Africa using an iterative region-growing algorithm, based on the ratio of FireCCISFD20 and VIIRS SNPP and NOAA-20 active fire detections.

Evaluation of scaling factors: The scaling grid is created only for 2019, the only year for which both FireCCISFD20 and VIIRS active fire data from the SNPP and NOAA-20 sensors are available. Therefore, scaling for more recent years, including near-real-time estimates, assumes that the 2019-based scaling grid remains representative. In order to test this, we used FireCCISFD11 to compare how well our scaling method represents spatial burnt area patterns in 2016 versus 2019. This exercise is limited to VIIRS SNPP for creating the Boolean active fire grid, since the NOAA-20 sensor had not yet been launched in 2016. As a result, scaling factors are roughly a factor two higher, omitting about half of the active fires counts due to the absence of NOAA-20 detections. For this test, VIIRS SNPP active fires for 2016 are scaled using the 2019 scaling factors and compared to 2016 FireCCISFD11 burned area. Secondly, active fires for 2019 are scaled using the 2016 scaling factors and compared to 2019 FireCCISFD20 burnt area (Figure 4, Figure 5).

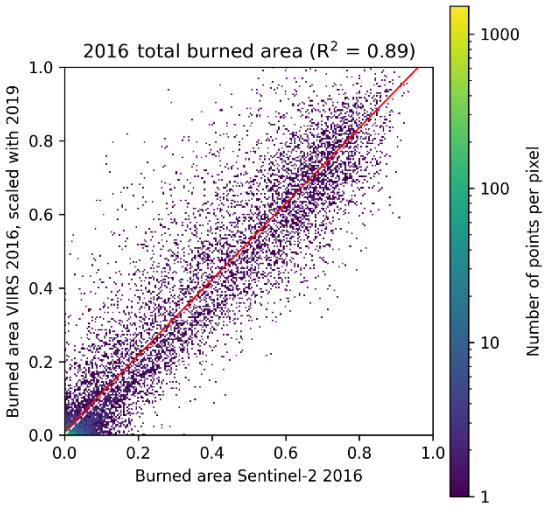

Figure 4: Comparison of FireCCISFD11 burned area for 2016 and scaled VIIRS SNPP burned area using a 2019-based scaling factor for the southern African domain..

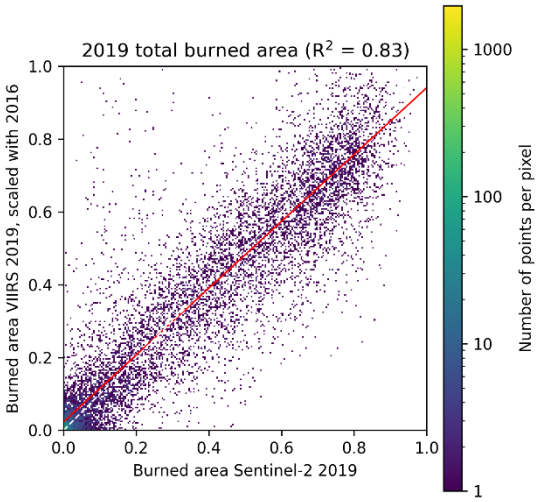

Figure 5: Comparison of FireCCISFD20 burned area for 2019 and scaled VIIRS SNPP burned area using a 2016-based scaling factor for the southern African domain.

2.4.2 Emissions

We estimate fuel consumption for each fire event based on fire type specific conversion factors. Conversion factors translate Fire Radiative Energy (FRE) per area burned (MJ m^{-2}) to fuel consumption per area burned (g C m^{-2}).

For each fire event, we derive FRE (MJ) from VIIRS fire radiative power (FRP, MW) by summing all FRP values of a single fire and multiplying this by the average duration between satellite overpasses: FRE (MJ) = (FRP in MW * 43,200 seconds). Subsequently, we derive FRE per area burned (MJ m-2) by dividing by the burned area associated with the fire event. Active fire data in this study are not corrected for cloud cover, since statistical models used in other studies to account for cloud cover cannot be easily attributed to specific fire events (e.g., (Kaiser et al., 2012)).

For fuel consumption, we rely on the dataset of fuel consumption observations compiled by van Wees et al. (2022). For each fire type, we reviewed the database to include the most relevant field observations of fuel consumption (g C m^{-2}) and match these to the distribution fire FRE per area burned (MJ m^{-2}) using the method explained in Andela et al. (2022). Finally, we convert carbon to trace gas emissions. We review emissions factors compiled by Andrea et al. (2019) to derive fire type specific emissions factors of CO, summarised in Table 3.

For the fire types considered, CO emission factors have coefficients of variation ranging from approximately 10% to 30%, indicating modest variability within individual fire types. In comparison, conversion factors show greater variability, with coefficients of variation of 55% for savannas and grasslands, 73% for tropical forests and woodlands, and 89% for croplands. This suggests that uncertainty associated with converting fire radiative power into fuel consumption may be greater than that arising from the subsequent application of emission factors to estimate trace gas emissions.

Table 3: Overview of emissions factors used in the GFA-S4F approach.

Fire type	**CO2 g kg^{-1}**	**CO g kg^{-1}**	**NOx (NO) g kg^{-1}**
Deforestation fires	1641	95.5	1.7
Tropical forest fires	1510	104.0	2.0
Small clearing and agricultural fires	1431	76.2	2.4
Temperate forest fires	1647	88.0	1.9
Boreal forest fires	1489	127.0	0.9
Savannah and grassland fires	1656	69.2	2.5
Cropland fires	1585	102	3.1

2.4.3 Benchmarking against GFED5

We compare burned area and carbon emissions estimates derived here (GFA-S4F-v0.3) against those from GFED5 (Figure 6) (Chen et al., 2023; van der Werf et al., 2025). Both approaches are calibrated against Sentinel-2, and therefore our average annual burned area estimate of 228 Mha compares closely to the 217 Mha estimated by GFED5 between 2019 to 2024. Both methods also show reasonable agreement (r-squared of 0.62) in the inter-annual variability of burned area over the same period. The emissions scaling implemented here deviates from the emissions model of GFED5, resulting in substantially higher (35%) emissions of 1211 Tg C annually, compared to 894 Tg C repored by GFED (2019 – 2022). While the aim of our model is not to replicate GFED5 emissions, we are planning to re-calibrate our emissions model before September 2026 based on a greater number of field observations and development of scaling factors (MJ m-2 to g C m-2) that are more representative of the regions field observations are located in. Currently, a single calibration is carried out including all of southern Africa, despite field observations being predominately located in part of the study region.

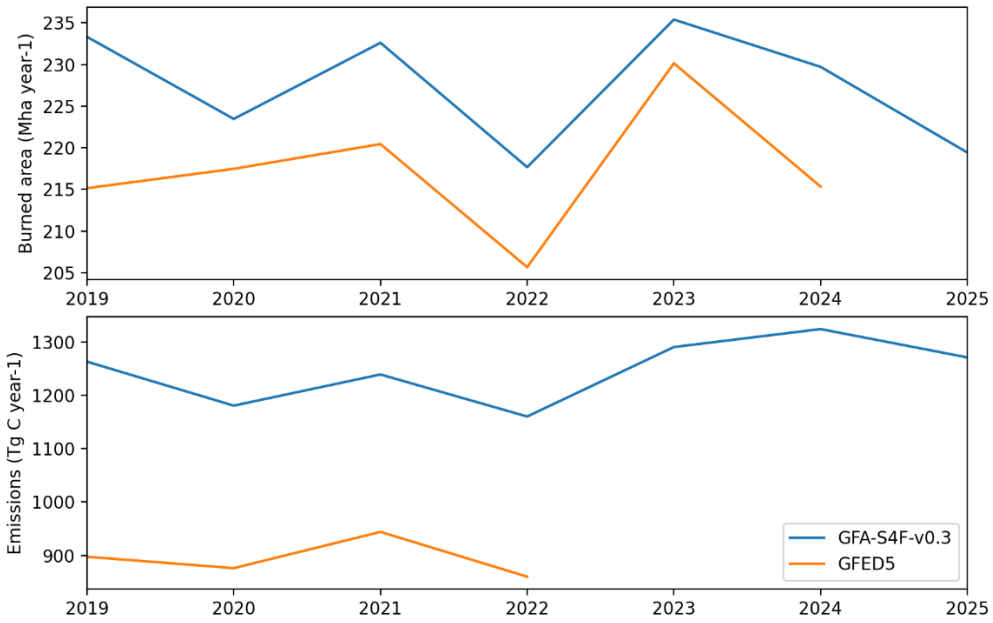

Figure 6: A comparison of burned area and carbon emissions estimates derived using the GFA-S4F-v0.3 and those of GFED5. Note that burned area and fire emissions estimates from GFED5 were not available for the full period.

3 TUD-S4F-v0.3

3.1 Overview

3.1.1 TUD-S4F approach in a nutshell

TUD-S4F is a data-model fusion approach that integrates different Earth observation products such as land cover, leaf area index (LAI), soil water index (SWI), and burnt area within a simple ecosystem model to estimate the spatial patterns and temporal dynamics of biomass, fuel loads and fuel moisture, combustion completeness, fuel consumption, combustion efficiency, emission factors, and fire emissions (Figure 7). The following description is largely based on text from the Supplementary Information in Forkel et al. (2025), with updates described in Section 3.1.2 and in the text.

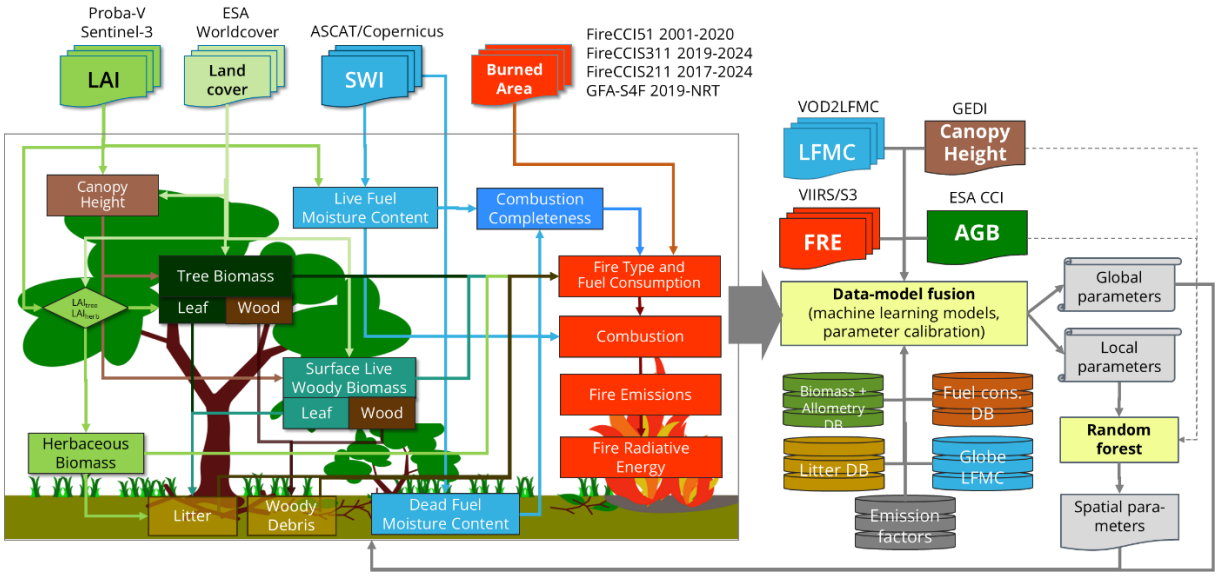

Figure 7: Schematic structure of the TUD-S4F-v0.3 approach to estimate fuel loads, fuel moisture, fuel consumption and fire emissions. Various satellite datasets are used a forcing (top) or for training of machine learning models and the calibration of model parameters (right).

The TUD-S4F core model represents different biomass and fuel components such as tree leaves, branches and stems, herbaceous vegetation, surface live woody biomass (SLW, i.e. shrubs and understorey), surface litter, and fine and coarse woody debris (FWD, CWD). Biomass in trees and SLW vegetation is estimated from canopy height, LAI and land cover using allometric equations that are calibrated based on satellite products of canopy height and above-ground biomass. The accumulation of surface fuels is then estimated from changes in land cover and from seasonal and long-term temporal changes in LAI, which affect the turnover of herbaceous and leaf biomass to litter, and of branches and stems to FWD and CWD, respectively. Fuel moisture content (FMC) is estimated from SWI and LAI for leaves and herbaceous biomass (i.e., live-fuel moisture content, LFMC) and from SWI for live woody biomass. Additionally, SWI is used as a proxy for the dead fuel moisture content (DFMC) of surface fuels. The estimated FMC is then used to estimate combustion completeness.

Unlike in other fire emission inventories, TUD-S4F estimates emission factors dynamically dependent on the chemical composition of biomass and fuel components and fuel moisture by using a chemical-based combustion model. The combustion model is calibrated against field and laboratory databases of emission factors and against satellite observations of fire radiative energy.

3.1.2 Summary of updates in version TUD-S4F-v0.3

Here we describe the most recent version TUD-S4F-v0.3. The previous version TUD-S4F-v0.2 has been described in ATBDv2.1 and ATBDv3 (Forkel et al., 2023b, 2024) and in the Supplementary Information of Forkel et al. (2025). In summary, TUD-S4F-v0.3 includes the following major updates and details can be found in the indicated Sections:

Use of new input datasets: All used input datasets for TUD-S4F are described in Section 3.2. In TUD-S4F-v0.3, we replaced the previously used Plant Functional Types (PFTs) from the ESA CCI Land Cover map with the ESA CCI WorldCover land cover map. In addition, for the Sahel region, we also evaluate the use of the ESA CCI High-Resolution Land Cover map (ESA HRLC) to test the effect of different land cover datasets. Both WorldCover and HRLC have a higher spatial resolution (10 m and 20 m) than the previously used ESA CCI Land Cover dataset.

For burnt area we produced burnt area time series at 20 m spatial resolution for the African study regions for the period June 2017 to December 2025 by making use of the FireCCI algorithm for Sentinel-2 data (Roteta et al., 2019). This product is hereinafter called Sense4Fire FireCCIS211.

In addition, we use the FireCCIS311 burnt product derived from Sentinel-3 observations (Lizundia-Loiola et al., 2022). By stacking the FireCCIS211 (2017-2025) or FireCCIS311 (2019-2024) products with the already previously used MODIS-based FireCCI51 product (used from 2014 onwards until 2017 or 2019) we achieve to run TUD-S4F for a period from 2014 to 2024/2025 (Section 3.2). Given the stacking of the datasets, TUD-S4F results should not be analysed for trends or inter-annual variability across the 2017 (FireCCIS211) or 2019 (FireCCIS311) limits.

Pre-processing and forecasting of LAI time series: LAI time series from the Copernicus Global Land Service (CGLS) product based on Proba-V and Sentinel-3 satellite observations are one of the main input datasets in TUD-S4F. Although this dataset includes a simple filtering of noise and interpolation of missing values, it can still contain gaps and short-term changes in LAI between 10-daily periods that appear as artefacts or noise. Those short-term changes in LAI can propagate in the TUD-S4F model to unrealistic changes in leaf and herbaceous biomass and hence in surface litter. Therefore, we apply a harmonic reconstruction approach to fill remaining gaps and to smooth the CGLS LAI time series (Section 3.2.3). The method also supports temporal resampling (e.g., decadal to daily) and forecasting of LAI, which is required for the near-real-time application of TUD-S4F. Uncertainties are propagated analytically through the linear model, with additional

block bootstrap and Monte Carlo approaches used to account for temporally correlated residuals and nonlinear functionals.

Implementation of surface live woody (SLW) vegetation: SLW vegetation such as shrubs in open shrublands or understorey woody vegetation in forests represent a major surface fuel in forest fires. Previously, SLW was not considered in TUD-S4F and shrublands were assigned to the forest land cover and hence modelled as "small" forests. In TUD-S4F-v0.3, we implemented a specific allometry model for SLW vegetation to estimate the biomass and fuel consumption of SLW woody and leaf biomass (Section 3.3.4).

Improved calculation of canopy height: Previously, we tested several alternative linear regression models to estimate canopy height from tree cover and long-term changes in mean LAI. This has been revised by using generalised linear models with a Gamma distribution, which is a better suited to model strict positive skewed data such as canopy height. In addition, temporal changes in the fractional cover of trees and SLW vegetation were included as additional predictor variables for canopy height (Section 3.3.1-3.3.2).

Improved calibration of emission factors: Vernooij et al. (2023) compiled biomass burning emission factors of several gas species and related ancillary information for Savannahs in a database. This database (hereinafter called "V23") contains more than 4600 measurements of emission factors for CO_2, CH_4, CO, N_2O and the Modified Combustion Efficiency (MCE) from savannahs in Australia, Brazil and southern African countries, many from within the southern Africa study region used in Sense4Fire. We used this database to calibrate parameters that affect the computation of emission factors in TUD-S4F-v0.3. This is a major advancement over the previous use of the emission database from Andreae et al. (2019) (A19), which contained only around 110 measurements for savannahs and grasslands. However, we still use also the A19 database as this includes emission factors for particulate matter (PM) and NO_x.

No use of VOD: Please note that we previously used satellite datasets of Vegetation Optical Depth (VOD) to calibrate parameters in TUD-S4F-v0.2 (Forkel et al., 2025), which we did not include in TUD-S4F-v0.3. While the calibration of parameters against VOD time series was helpful to constrain estimates of vegetation water content (VWC) in TUD-S4F, it did not strongly influence the estimated biomass, fuel dynamics and fire emissions. A critical aspect thereby was also the mismatch in resolution between the 0.25° resolution of VOD datasets and the approx. 300 m resolution of TUD-S4F. The assessment of the actual benefit of using VOD data to constrain TUD-S4F requires more detailed analyses and data-model fusion experiments, including testing the effect of the resolution mismatch and the formulation of the observation operator to integrate VOD data, which is currently out of scope in the development for TUD-S4F-v0.3. Therefore, we did not use VOD data in TUD-S4F-v0.3, although VWC and VOD are still computed internally in TUD-S4F.

3.1.3 Latency and maturity setups of the TUD-S4F-v0.3 approach

TUD-S4F-v0.3 is implemented so that it can provide estimates of fuels and fire emissions in various setups (Table 4). We define four product latency and maturity levels:

- **DEF: The default setup** represents a fully reprocessed, newly calibrated model setup with state-of-the-art input datasets (e.g. with burnt area from FireCCI). DEF setups can be only provided with 2-3 years delay, depending on the availability and latency of the input datasets. Results from this setup are recommended for statistical, seasonal and multi-annual analyses.
- **PRV: The provisional setup** provides temporally complete annual coverage of the previous year and has moderate latency of approx. 4 months after year end. PRV setups rely on burnt area estimates (like from GFA-S4F) which is available a few days to weeks after a fire but make use of observed LAI data, which has currently a time lag of 60 days for the CGLS LAI product. Results from this setup can be used for an assessment of the previous wildfire seasons.
- **NRT: The near-real time setup** provides estimates of fuels and fire emissions of recent fire events, available within 2–4 weeks after a fire event. This setup relies on incomplete input data and specifically requires NRT estimates of burnt area (from GFA-S4F, section 2.4) and LAI (based on a NRT forecast, section 3.4).
- **EXP: Experimental setup** are model runs that reflect sensitivity experiments with TUD-S4F, for example to test the effect of different input datasets or to assess the effect of new model developments.

In summary, the different latency levels and setups differ by the used input datasets. For example, the default setup TUD-S4F-v0.3-**S311** uses coarse resolution (0.0025°, approx. 250 m) burnt area from MODIS and Sentinel-3 from the FireCCI51 (2014-2018) and FireCCIS311 (2019-2024) products, respectively, while the TUD-S4F-v0.3-**S211** setup uses burnt area from Sentinel-2 from the FireCCIS211 product, produced under Sense4Fire, (2017-2025) at 0.0025° resolution. Note that the datasets are only temporally stacked and not harmonised, which prevents the use of the entire time series analyses of trends or inter-annual variability. In order to use the S4F-FireCCIS211 burnt area dataset (20 m spatial resolution) at coarse resolution, the dataset was aggregated to 0.0025° by computing the fractional burnt area per grid cell.

The provisional TUD-S4F-v0.3-**GFA** setup provides emission up to the end of the previous year by using the NRT GFA-S4F burnt area estimate. TUD-S4F-v0.3-**NRT** is the near-real time setup at 0.0025° resolution that makes use of a NRT estimate of leaf area index (LAI) and NRT burnt area from GFA-S4F-v0.3. TUD-S4F-v0.3-**HR** are experimental setups at 0.0002° (approx. 20 m) spatial resolution that differ by the used input data for LAI or land cover. Please note that although we refer to the resolution of 0.0025° as 'coarse' resolution that this resolution represents one of the large-scale fire emission datasets with the highest spatial resolution available.

Table 4: Overview of the setups of the TUD-S4F-v0.3 approach.

Approach, version and setup	Spatial resolution, temporal coverage	Main input data	Product latency level and description
TUD-S4F-v0.3-**S311**	0.0025° 2014/2019-2024	ESA WorldCover 2020 CGLS LAI FireCCI51 BA 2014-2018 + **FireCCIS311** BA 2019-2024	**DEF:** Default setup with observed LAI, new representation of surface live woody vegetation, land cover from ESA WorldCover, and prolonged burnt area time series (FireCCIS311 from 2019 onwards)
TUD-S4F-v0.3-**S211**	0.0025° 2017-2025	ESA WorldCover 2020 CGLS LAI **FireCCIS211** BA 2017-2025	**DEF:** Default setup in addition to TUD-S4F-v0.3-S311 by using burnt area from Sentinel-2 (S4F-FireCCIS211 product)
TUD-S4F-v0.3-**S211-HRLC**	0.0025° 2017-2025	ESA CCI **HRLC30** land cover CGLS LAI **FireCCIS211** BA 2017-2025	**EXP:** Experimental setup like TUD-S4F-v0.3-S211 but using HRLC30 land cover for the year 2019
TUD-S4F-v0.3-**GFA**	0.0025° 2014/2019-2025	ESA WorldCover 2020 CGLS LAI FireCCI51 2014-2018 + **GFA-S4F NRT** BA 2019-2025	**PRV:** Provisional setup of TUD-S4F-v0.3 with observed LAI and near-real time burnt area from GFA-S4F
TUD-S4F-v0.3-**NRT**	0.0025° 2019-NRT	ESA WorldCover 2020 CGLS LAI + **NRT LAI** **GFA-S4F NRT** BA 2019-NRT	**NRT:** Near-real time setup case for selected fire events of TUD-S4F-v0.3 with NRT LAI estimate and NRT burnt area from GFA-S4F
TUD-S4F-0.3-**HR-S3**	**0.0002°** 2017-2025	ESA WorldCover 2020 **Downscaled Sentinel-3 LAI** **FireCCIS211** BA 2017-2025	**EXP:** Experimental high-resolution setup of TUD-S4F-v0.3 with LAI downscaled from Sentinel-3
TUD-S4F-0.3-**HR-S2**	**0.0002°** 2017-2025	ESA WorldCover 2020 **Sentinel-2 LAI** **FireCCIS211** BA 2017-2025	**EXP: Experimental high-resolution** setup of TUD-S4F-v0.3 with LAI derived from Sentinel-2

3.2 Input datasets and data pre-processing

3.2.1 Data pre-processing for all datasets

Input datasets are either used as static or dynamic datasets. Static datasets neither show seasonal variations throughout the year nor contain more than one time step per year; they provide a single annual snapshot (e.g. land cover, canopy height, above-ground biomass). In contrast, dynamic datasets capture processes that change over time, feature

multiple observations each year, and exhibit distinct sub-annual temporal variability (e.g. LAI, SWI, burnt area, LFMC).

All input datasets are resampled onto a common geographical grid for spatial consistency. The grid uses the WGS84 coordinate system (EPSG:4326) and covers the entire areas of interest. It is divided into regular tiles measuring 0.5° longitude × 0.5° latitude. For the coarse resolution setup, each tile has a spatial resolution of 0.0025°, forming a 200 rows × 200 column grid. This setup ensures seamless spatial alignment within and across tiles, preventing resampling artefacts. For the HR setup, tiles have a spatial resolution of 0.0002°.

Table 5: Overview of pre-processing steps applied to all datasets.

Dataset	Category	Spatial Resampling	Temporal Resampling
CGLS LAI	dynamic	Bilinear interpolation	None required — native 10-daily temporal resolution grid aligned with target grid (10th, 20th, and last day of each month) but filtering and gap-filling applied
FireCCI51	dynamic	Average resampling to derive burnt area fractions	Resampled to 10-daily LAI target grid; binary classification indicating fire or no fire per time-step
FireCCIS311	dynamic	Average resampling to derive burnt area fractions	Resampled to 10-daily LAI target grid; binary classification indicating fire or no fire per time-step
S4F-FireCCIS211	dynamic	Average resampling to derive burnt area fractions	Resampled to 10-daily LAI target grid; binary classification indicating fire or no fire per time-step
ESA WorldCover	Static	Average resampling to derive landcover fractions	N/A
ESA CCI HRLC Land Cover	static	Average resampling to derive landcover fractions	N/A
ESA CCI Above-Ground Biomass	static	Bilinear interpolation	N/A

Temporal resampling applies only to dynamic datasets, whose resolutions range from daily to 10-daily steps. The target temporal grid uses the 10-daily resolution of the CGLS LAI product. Thereby, the days in date definitions always correspond to conditions in the 10 days before. For example, the date 2020-01-10 correspond to average state variables (e.g. LAI, LFMC) or biomass pools (e.g. tree biomass) or to the total accumulated area (for burnt area) or flux (e.g. dry matter burnt or fire emissions) over the period 2020-01-01 to

2020-01-10. The spatial and temporal resampling method used for each dataset depends on its characteristics (Table 5). Water bodies are masked using the water class from the ESA WorldCover 2021 dataset. Grid cells identified as water are set to NaN.

3.2.2 CGLS Proba-V/Sentinel-3 leaf area index (LAI)

The Copernicus Global Land Service (CGLS) Leaf Area Index (LAI) product, version 2, provides global time series of LAI, defined as half the total area of green elements of the canopy per unit horizontal ground area. The dataset is available at a spatial resolution of approximately 300 m and a temporal frequency of 10 days. From January 2014 to June 2020, the product was derived from PROBA-V satellite observations. Following the decommissioning of PROBA-V, the dataset continuity has been ensured by the Sentinel-3 satellite mission, specifically using data from the Ocean and Land Colour Instrument (OLCI) and the Sea and Land Surface Temperature Radiometer (SLSTR). This transition from PROBA-V to Sentinel-3 was designed to guarantee long-term, sustained global land surface monitoring at moderate resolution. The LAI estimates are produced in near real time at global scale, making the dataset well suited for operational vegetation monitoring and land surface modelling applications. The product is also commonly referred to as GEOV3 in the scientific literature (Baret et al., 2016; Fuster et al., 2020; de Muck et al., 2025).

The CGLS LAI product is one of the main inputs to TUD-S4F and defines the period and temporal resolution of model applicability (starting 2014, 10-daily).

3.2.3 LAI time series filtering, gap-filling and forecasting

We apply a harmonic reconstruction approach to fill remaining gaps, to smooth and to forecast the CGLS LAI time series. The algorithm reconstructs gap-affected Leaf Area Index (LAI) time series using a continuous, differentiable parametric model. This approach assumes that LAI follows quasi-periodic variation, repeating an approximately similar annual cycle each year, driven by seasonal cycles of insolation, temperature, and precipitation (Jakubauskas et al., 2002; Jönsson and Eklundh, 2004; Roerink et al., 2000). The LAI signal is modelled as the superposition of (i) a repeating annual waveform, (ii) higher harmonics that encode the overall waveform shape, particularly the asymmetry between growing and dormant seasons, and (iii) additive measurement noise. A harmonic Fourier representation is employed, as any periodic function can be expressed as a sum of sinusoids at integer multiples of the fundamental frequency. This framework enables reconstruction by leveraging information from all observed years to constrain the model in data-sparse intervals, allowing the recurring cycle to bridge gaps rather than relying on local interpolation between neighbouring samples (Das, 2015).

The periodicity assumption is applied in a relaxed manner; a strictly periodic model would require identical annual cycles, which is inconsistent with observed interannual variability. The model maintains fixed annual periodicity for the harmonic carriers while permitting their amplitudes and baseline to vary smoothly over time (Jiang et al., 2010). Periodicity constrains the cycle timing, while slowly varying coefficients capture interannual variation.

For weakly seasonal or aseasonal LAI (e.g., evergreen tropical forests), the oscillatory coefficients are set to zero, allowing the constant term to represent the signal without switching models.

Processing consists of four stages: (i) an admissibility test to determine whether a pixel's observation record supports reconstruction at the required spectral bandwidth, (ii) derivation of the number of annual harmonics from the data, (iii) fitting of an amplitude-varying harmonic model using weighted least squares, and (iv) analytical propagation of uncertainties from the per-observation uncertainty layer to all derived quantities. The time axis is expressed in decimal years, with a nominal sampling rate of 36 samples per year (decadal compositing), so that the fundamental annual cycle corresponds to a frequency of exactly 1 yr^{-1}. Each observation y_i is accompanied by a per-step uncertainty obtained from the product's RMSE layer.

Because the model is defined continuously, gap filling requires no separate interpolator; observed values are retained and missing steps are replaced by the model prediction. The method also supports temporal resampling (e.g., decadal to daily) and forecasting of LAI, which is required for the near-real-time application of TUD-S4F. Uncertainties are propagated analytically through the linear model, with additional block bootstrap and Monte Carlo approaches used to account for temporally correlated residuals and nonlinear functionals.

3.2.3.1 Determination of reconstructability

Reconstruction is considered admissible only when the observation record sufficiently constrains the model at the bandwidth determined by the selected harmonic order. The gate evaluates four independent criteria on the binary presence mask $m_i = 1[y_i < \infty]$, all of which must be satisfied.

According to the Landau rate, a band-limited signal containing harmonics up to order $W = K - 1$ requires a minimum average sampling density equal to twice the one-sided bandwidth (Gröchenig and Razafinjatovo, 1996).

Over a record of Y years, the required number of valid observations is:

$N_{min} = \lceil 2, W, Y, \kappa \rceil$

where $\kappa = 1.3$ is a safety factor accommodating non-uniform sampling. The criterion is $N_{valid} \geq N_{min}$.

Global density is insufficient if the missing observations are contiguous: a gap spanning a substantial fraction of the shortest resolved period leaves that period locally unconstrained. With shortest period $T_{min} = \frac{f_s}{W}$ samples, the admissible gap length is:

$$G_{max} = \left\lfloor \gamma, \frac{f_s}{W} \right\rfloor, \gamma = 0.5$$

and the criterion is $G_{\text{longest}} \leq G_{max}$ or $f_s = 36$ and $W = 3$ this permits gaps of up to six decads (≈60 days). Basis functions are least constrained at time-series' ends, where the fit has no two-sided support. At least two valid observations must fall within each of the

leading and trailing $[0.06, n]$ samples. An absolute floor $N_{\text{valid}} \geq N_{\text{floor}}$ guards against broken time-series'. The gate returns a binary verdict along with diagnostic quantities, including G_{longest}, G_{max} the number of gap segments, and the number of missing steps. If the verdict is negative, processing for that pixel terminates, and the pixel is set to NaN while retaining the verdict flag. No fitted values, uncertainties, or forecasts are generated. The gate is applied identically in all fitting and extrapolation cases, ensuring that admissibility rules remain consistent across processing paths.

3.2.3.2 Derivation of the harmonic order

The number of annual harmonics is derived from the data rather than assumed (Draper and Smith, 1998). Harmonics are added sequentially, $h = 1, 2, \ldots$, each contributing a sine–cosine pair to the design matrix. At each step, the reduction in residual sum of squares is tested by a nested F-test:

$$F_h = \frac{(\text{RSS}_{\text{h-1}} - \text{RSS}_{\text{h}})/2}{\text{RSS}_{\text{h}}/(n - p_h)}$$

with two numerator degrees of freedom (the sine and cosine coefficients) and $n - p_h$ denominator degrees of freedom. We retain a harmonic when it is both statistically significant ($p < \alpha, \alpha = 0.05$) and materially important, as quantified by the partial coefficient of determination:

$$R^2_{\text{partial},h} = \frac{\text{RSS}_{\text{h-1}} - \text{RSS}_{\text{h}}}{\text{TSS}} \geq \tau, \qquad \tau = 0.02$$

The selected order is $K = 1 + h_{max}$ here h_{max} the highest harmonic satisfying both conditions in an unbroken sequence from $h = 1$. The significance test optionally includes a linear trend column, tested separately with one numerator degree of freedom. The trend is a zero-frequency effect and does not enter K, which governs the spectral bandwidth $W = K - 1$ used by the gate. The model order is selected on the denoised series to ensure that the test responds to phenological structure rather than measurement noise. Model order is cross-checked by the Bayesian Information Criterion evaluated with known per-point variances (Schwarz, 1978):

$$\text{BIC} = p \ln n - 2 \ln \mathcal{L}, \qquad \ln \mathcal{L} = -\frac{1}{2} \sum_i \left[\frac{r_i^2}{\sigma_i^2} + \ln\left(2\pi\sigma_i^2\right) \right]$$

The adequacy of the selected order is verified a posteriori using the residual periodogram. Under a correct representation of the seasonal waveform, residual power is spectrally flat. Residual energy concentrated at integer multiples of $1\ \text{yr}^{-1}$ indicates an unmodelled harmonic and the need to increase K. Peak significance is assessed against a block-bootstrap null that preserves residual autocorrelation while destroying periodicity, and peak location is additionally checked for stability across alternative model configurations (Bühlmann and Künsch, 1999). A persistent peak is physical, whereas a migrating one is related to noise. For the reference semi-arid savanna site, $K = 4$ ($H = 3$ harmonics) is selected; three harmonics are required to represent the asymmetric green-up and senescence limbs of the annual cycle.

3.2.3.3 Derivation of the harmonic fit

Model formulation: A fixed-amplitude harmonic model cannot represent interannual variation in seasonal magnitude and underfits at all orders K when tested against the per-point uncertainty (Harvey, 1990; Young et al., 1999). Therefore, the operational model allows the baseline and each harmonic amplitude to vary in time:

$$\widehat{\mathrm{LAI}}(t) = c(t) + \sum_{h=1}^{H} [A_h(t)\sin(2\pi ht) + B_h(t)\cos(2\pi ht)]$$

Each time-varying coefficient is expanded in a cosine basis on normalised time $\tilde{t} = (t - t_{min})/L \in [0,1]$:

$$\phi_m(\tilde{t}) = \cos(\pi m\tilde{t}), \qquad m = 0, \ldots, n_{\mathrm{amp}},$$

$$c(t) = \sum_m \beta_m^c \phi_m(\tilde{t}), \qquad A_h(t) = \sum_m \beta_m^{A_h} \phi_m(\tilde{t}), \qquad B_h(t) = \sum_m \beta_m^{B_h} \phi_m(\tilde{t}).$$

The model is linear in the parameter vector β, so it can be written $\hat{y} = \boldsymbol{X}\beta$ with design matrix X of dimension $n \times P$, where $P = (n_{\mathrm{amp}} + 1)(1 + 2H)$. Linearity in β is essential because it permits closed-form covariance propagation, analytic differentiation, and extrapolation using the same algebra as the fit. The constant mode $\phi_0 \equiv 1$ is part of the design, so a flat series is represented exactly by the null state $\beta_0^c = \bar{y}$ with all remaining coefficients zero. Therefore, no model switching is required for weak or non-seasonal vegetation.

Estimation: Coefficients are estimated by weighted least squares with weights derived from the per-step uncertainty, $W = \mathrm{diag}(\sigma_i^{-2})$: $\hat{\beta} = \arg\min \beta\ (y - X\beta)^T W (y - X\beta) = (X^{\mathrm{T}} W X)^{-1} X^{\mathrm{T}} W y$.

Numerically, the system is solved in the following form via least-squares factorisation: $X_w = X/\sigma, y_w = y/\sigma$. The fit uses only observations. The equations exclude interpolated steps, which carry no independent information.

Gap Filling and Resampling: Because the model is defined continuously, gap filling requires no separate interpolator; interpolation is a by-product of the method. The design is rebuilt on the full record grid and evaluated with the fitted coefficients:

$$\widehat{\mathrm{LAI}}(t) = x(t)^{\mathrm{T}}\hat{\beta}$$

Observed values are retained and missing steps are replaced by the model prediction, with each step labelled as observed, model-filled, or missing. The normalisation constants t_{min} and L are fixed at their fitted values whenever the design is rebuilt on a different time-series length. This step is mandatory because recomputing $\tilde{t}$ from a time-series that extends beyond the fitting record would alter the basis and invalidate $\hat{\beta}$. Temporal resampling, such as decadal to daily, is performed by evaluating the model on the target time-series length using the same fixed normalisation, rather than by interpolating the fitted decadal values. This approach reproduces the model exactly between nodes and yields the uncertainty at the target time-series length analytically, which interpolation of a variance array cannot provide (Olsson and Eklundh, 1994).

Forecast: Forecast steps $t_* > t_n$ are evaluated identically, $\widehat{\mathrm{LAI}}(t_*) = x(t_*)^{\top}\hat{\beta}$, with $\tilde{t_*} > 1$. A forecast of $n_{\mathrm{ahead}} = 7$ decadal steps ($\approx$70 days) covers the maximum admissible gap defined in Section 2 and reaches $\tilde{t_*} = 1 + (n_{\mathrm{ahead}}/f_s)/L$. For a 12.1-year record, this is $\tilde{t_*} \approx 1.016$, a 1.6 % extension of the amplitude basis (Olsson and Eklundh, 1994).

3.2.3.4 Limitations

Sideband Collision: The amplitude basis modulates each harmonic carrier, producing sidebands. Mode m of the amplitude basis oscillates at frequency $f_m = \frac{m}{2L}$ cycles yr^{-1}, and multiplying it with the fundamental carrier at $1\ \mathrm{yr}^{-1}$ places energy at $1 \pm f_m$. The lower sideband of the fundamental therefore lies at $f_{\mathrm{sb}} = 1 - \frac{n_{\mathrm{amp}}}{2L}$. The periodogram resolution of a record of length L is $\Delta f = 1/L$. Expressed in resolution cells, the separation between the fundamental and its highest sideband is the sideband margin: $\mathcal{M} = L - n_{\mathrm{amp}}$. Collision occurs at $\mathcal{M} = 0$, i.e. $n_{\mathrm{amp}} = L$. The amplitude modulation becomes spectrally indistinguishable from the seasonal cycle itself, and interannual amplitude variation can no longer be separated from the seasonal signal. The basis order is therefore set as an offset below the record length:

$$n_{\mathrm{amp}} = \lfloor L \rfloor - \delta, \qquad \mathcal{M} = \mathrm{frac}(L) + \delta,$$

with $\delta = 1$ as the minimum-clearance setting and $\delta = 2$ required for products that depend on record-end behaviour such as derivatives and forecasts, where the basis is least constrained. Deriving n_{amp} from L rather than fixing it ensures the guarantee travels with the data when shorter archives are processed (Amiri-Simkooei, 2007).

Structural Limitations: The model represents the persistence of learned seasonal and interannual structure. It cannot anticipate abrupt disturbances such as fires or storms occurring within a bridged gap or forecast horizon. Such events appear as unmodelled mean shifts not captured by the parameter covariance. Additionally, the phase $\varphi_h(t)$ and its uncertainty are undefined in the limit $R_h \to 0$ and are masked for aseasonal pixels (Verbesselt et al., 2010, 2012).

3.2.3.5 Error propagation methodology

Coefficient Covariance: Since $\hat{\beta}$ is linear in y and $\mathrm{Cov}(y) = W^{-1}$:

$$\mathrm{Cov}(\hat{\beta}) = (X^T W X)^{-1} X^T W W^{-1} W X (X^T W X)^{-1} = (X^T W X)^{-1} \equiv \Sigma_{\beta}$$

The input uncertainty layer is trusted in its relative structure but rescaled in magnitude to the empirical residual dispersion via the reduced χ^2_{ν},

$$\chi^2_{\nu} = \frac{1}{n-P}\sum_i \frac{r_i^2}{\sigma_i^2}, \qquad \hat{\Sigma}_{\beta} = \chi^2_{\nu}\,(X^{\top}\boldsymbol{W}\boldsymbol{X})^{-1}..$$

Values $\chi^2_{\nu} < 1$ indicate the stated RMSE overstates the true dispersion, and rescaling prevents systematically inflated output uncertainties (Bevington and Robinson, 2003).

Derived Quantities: For any evaluation time, observed time-series, resampled time-series, or forecast step or steps, the variance is the quadratic form: $\mathrm{Var}[\widehat{\mathrm{LAI}}(t)] = x(t)^{\top}\hat{\boldsymbol{\Sigma}}_{\beta}x(t)$, evaluated efficiently as $\mathrm{diag}(X\hat{\boldsymbol{\Sigma}}_{\beta}\boldsymbol{X}^T)$ with the 95 % confidence band being $\widehat{\mathrm{LAI}} \pm 1.96\,\mathrm{SE}$. The per-harmonic amplitude and phase are nonlinear functions of the

coefficients, $R_h = \sqrt{A_h^2 + B_h^2}$ and $\varphi_h = \arctan 2\,(B_h, A_h)$, and are propagated by the delta method. With Φ the amplitude-basis matrix and v_{AA}, v_{BB}, v_{AB} the propagated (co)variances of $A_h(t)$ and $B_h(t)$:

$$\mathrm{Var}(R_h) = \frac{A_h^2 v_{AA} + B_h^2 v_{BB} + 2A_h B_h v_{AB}}{R_h^2}, \qquad \mathrm{Var}(\varphi_h) = \frac{B_h^2 v_{AA} + A_h^2 v_{BB} - 2A_h B_h v_{AB}}{R_h^4}$$

Both expressions are singular as $R_h \to 0$ and are masked below a configured amplitude threshold. Amplitude is additionally masked within the boundary fraction used by the gate, where it is poorly constrained. Because $d\widehat{LAI}/dt = \dot{X}\beta$ remains linear in β, derivative uncertainty follows the same quadratic form with $\dot{X}$ in place of X (Politis and Romano, 1994; Tellinghuisen, 2001).

Forecast Uncertainty: The forecast error decomposes into independent estimation and observation components: $\mathrm{LAI}(t_*) - \widehat{\mathrm{LAI}}(t_*) = x(t_*)^T(\beta - \hat{\beta}) + \varepsilon(t_*)$, with $\boldsymbol{x}(t_*)^T(\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}})$being the estimation and $\varepsilon(t_*)$ the observation, giving a confidence variance $x^{\top}\hat{\boldsymbol{\Sigma}}_{\beta}x$ for the mean curve and a prediction variance:

$$\mathrm{Var}_{\mathrm{pred}}(t*) = x(t_*)^{\top}\hat{\Sigma}_{\beta}, x(t_*) + \sigma_*^2,$$

where σ_* is the anticipated observation uncertainty, estimated from the median of the most recent valid σ_i. The prediction band is the appropriate reference for comparison against future observations. The harmonic carriers $\sin(2\pi ht)$, $\cos(2\pi ht)$ are bounded and contribute no horizon-dependent growth. The amplitude basis, however, is evaluated at $\tilde{t_*} > 1$, so x (Carroll and Ruppert, 2017; Seber and Lee, 2003).

Block Bootstrap for Correlated Errors: The breakdown of the formula in the coefficient covariances section is predicated on $\mathrm{Cov}(y)$ being diagonal. The decadal compositing windows do overlap between time steps so the residuals are temporally correlated and the analytical covariance is therefore optimistic. This is quantified via resampling rather than by some alternative closed form. Residuals are first standardized $\tilde{r_i} = r_i/\sigma_i$, which removes the heteroscedasticity already accounted for by W, so that the comparison is purely a function of correlation. The standardised residuals are then resampled in contiguous blocks of length b, chosen to match the compositing correlation scale, and re-coloured by the local uncertainty at the destination index: $y_i^{*(j)} = \widehat{\mathrm{LAI}}_i + \tilde{r}_i^{*(j)}\sigma_i$ (Kunsch, 1989).

Each surrogate series $y^{*(j)}$ is refitted with the identical model configuration, and the ensemble $\widehat{\mathrm{LAI}}^{(j)}(t)_{j=1}^{B}$ yields a percentile band of form $\left[Q_{2.5}\widehat{\mathrm{LAI}}^{(j)}(t),\ Q_{97.5}\widehat{\mathrm{LAI}}^{(j)}(t)\right]$. Because block resampling preserves within-block dependence while destroying longer-range structure, the ratio of bootstrap to analytical band width gives: $\rho(t) = \frac{W_{\mathrm{boot}}(t)}{W_{\mathrm{analytical}}(t)}$ (Freedman, 1981).

Estimates the inflation due to correlated errors. An effective sample size then follows as $N_{eff} \cong n/\bar{\rho}^2$. The analytical covariance must be χ_ν^2-rescaled for the comparison to be valid, so that both bands are on the same empirical residual scale (Theiler et al., 1992).

Monte Carlo propagation: Any quantities which are non-linear functionals of β are not amenable to closed form propagation. For these we use parametric Monte Carlo: $\boldsymbol{\beta}^{(j)} \sim \mathcal{N}(\hat{\beta}, \hat{\Sigma}_\beta), j = 1, \ldots, M$. The covariance is propagated through the full nonlinear chain and percentiles are taken across the ensemble. It is symmetrised and given a trace-scaled jitter prior to sampling in order to ensure positive definiteness. This is different to the block bootstrapped outlined in the paragraph before, where the input distribution is the fitted Gaussian covariance and only the propagation is stochastic. The resulting band therefore inherits the independence assumption of $\hat{\Sigma}_\beta$ and should be scaled by $\bar{\rho}$ where correlation is important (Efron and Tibshirani, 1994).

Gap-specific uncertainty: Model-filled steps carry an error component absent at observed steps, which neither $\hat{\Sigma}_\beta$ nor the block represents, since both are conditioned on the realised sampling pattern. It is characterised empirically by injecting synthetic gaps into gap-free reference series, with block length and number of segments matched to that actually encountered, reconstructing, and compiling the error distribution at the filled positions:

$$e_i = \widehat{\mathrm{LAI}}_i - \mathrm{LAI}_i^{\mathrm{true}}, \qquad i \in \text{filled}.$$

The standard deviation and mean absolute error of e are reported as the interpolation uncertainty. Injected block lengths are capped at the gate's admissible maximum G_{max} and trials whose realised geometry fails the gate. For instance, randomly placed blocks become adjacent and exceed G_{max} are discarded and counted, keeping the characterisation consistent with the admissibility rules. A complementary Monte Carlo variant perturbs the observed values by their stated uncertainties, $y_i^{(j)} = y_i + \epsilon_i^{(j)}$ with $\epsilon_i^{(j)} \sim \mathcal{N}(0, \sigma_i^2)$ at observed positions only, and reconstructs each realisation. The resulting percentile envelope expresses how measurement noise at the observed steps propagates into the filled steps, which is the uncertainty most relevant to a bridged gap (Falge et al., 2001).

3.2.4 MODIS and Sentinel-3 burnt area (ESA FireCCI51 and FireCCIS311)

The FireCCI51 and FireCCIS311 datasets are global burnt area (BA) products developed within the ESA Fire Climate Change Initiative (FireCCI) programme, which aims to produce long-term, consistent datasets of burnt area information from satellite observations.

FireCCI51 provides monthly global burned area information at an approximate spatial resolution of 250 m, covering the period from 2001 to 2020. The product is derived from MODIS (Moderate Resolution Imaging Spectroradiometer) satellite data (Chuvieco et al., 2018; Lizundia-Loiola et al., 2020; Pettinari et al., 2021).

The FireCCIS311 dataset is the successor burnt area product, extending the temporal record with observations from the Sentinel-3 SYN (Synergy) processor (Chuvieco et al., 2024; Lizundia-Loiola et al., 2022). It provides global burnt area information covering the period from 2019 to 2024, thereby ensuring continuity with the earlier FireCCI51 record.

Together, FireCCI51 and FireCCIS311 form a complementary pair of burnt area datasets that span over two decades of global fire observations, transitioning from MODIS-based to Sentinel-3-based detection while maintaining methodological consistency within the Fire CCI framework (Khairoun et al., 2023; Pettinari, 2024). However, although the two datasets are defined as complementary, they still show differences as the two datasets are not harmonised or bias-adjusted against each other. For example, the Sentinel-3 based product shows up to 12% larger burnt area than the MODIS-based product in tropical savannahs (Lizundia-Loiola et al., 2022).

In the **TUD-S4F-v0.3-S311** default setup, we use burnt area from FireCCI51 for the period 2014-2018 and from FireCCIS311 for 2019-2024. This was done to include a fire and disturbance 'history' for the period 2014-2018 but, however, implies that the TUD-S4F-v0.3 datasets should not be used for analyses of trends or interannual variability for the full period from 2014-2024 because of the inconsistency between the two burnt area datasets. Hence, we recommend using the TUD-S4F-v0.3 datasets only from the years 2019 (TUD-S4F-v0.3) and 2017 (TUD-S4F-S211) onwards.

3.2.5 Sentinel-2 Burnt Area (S4F-FireCCIS211/ FireCCISFD20 / FireCCIS2v1)

A burnt area dataset has been produced from Sentinel-2 imagery for the Sense4Fire project (named S4F-FireCCIS211) by Brockmann Consult (Martin Böttcher, Gunnar Brandt) and is based on the method for the FireCCIS2v1 dataset, which is currently developed within the FireCCI project. FireCCIS2v1 is a high-resolution global burnt area product at 20 m spatial resolution for the year 2023. It follows the same algorithmic approach as the FireCCISFD20 (Small Fire Database 20m) product for Africa (Chuvieco et al., 2022), which was specifically designed to exploit the high spatial resolution capabilities of the Sentinel-2 MSI (Multi-Spectral Instrument) sensor. The products provide burnt area information at a spatial resolution of 20 m, representing a significant improvement in spatial detail compared to the coarser resolution FireCCI51 (~250 m, MODIS-based) and FireCCIS311 (Sentinel-3-based) products. This high-resolution capability enables the detection of small fires that are often omitted by moderate-resolution burnt area products (Ramo et al., 2021), which is particularly relevant in regions where small-scale agricultural burning and fragmented fire regimes dominate.

The S4F-FireCCIS211 product produced here is an extension of the FireCCIS2v1 product, for which the method is applied to the S-Africa and Sahel study regions for the period June 2017 to December 2025 to provide a regional multi-annual high resolution burnt area time series. The FireCCIS211 product was used both in the coarse resolution (0.0025°) setup in **TUD-S4F-v0.3-S211** as an alternative to the FireCCI51/FireCCIS311 burnt area datasets as well as in the high resolution (0.0002°) setups in **TUD-S4F-v0.3-HR**.

The burnt area datasets S4F-FireCCIS211, the predecessor FireCCISFD20, and the FireCCIS311 are compared for the study regions in PVRv4.

3.2.6 ESA WorldCover

The ESA WorldCover product provides global land cover types at 10 m resolution for the years 2020 (version 100) and 2021 (version 200) (Zanaga et al., 2021, 2022). It is derived based on synthetic aperture radar (Sentinel-1) and optical sensor (Sentinel-2) observations using a gradient boosting decision tree algorithm. Different algorithm versions are used for the two different versions and years and hence any changes between the two maps represent both real land cover changes and artefacts from classification errors. The land cover map has a global overall accuracy of 74.4% for the year 2020 and 76.7% for the year 2021 (Tsendbazar et al., 2021, 2022). Xu et al. (2024) report a higher global accuracy for WorldCover 2021 because they applied a different validation method. The original eleven land cover types are reclassified into seven types.

WorldCover is used as land cover input in all setups of TUD-S4F-v0.3. Therefore, the tree, shrubland, grassland, cropland, and herbaceous wetland classes were aggregated to 0.0025° (default) and 0.0002° (HR) resolutions by calculating the relative frequency of each land cover class per grid cell. The difference in relative frequency to unity caused by the remaining classes (built-up, bare, snow and ice, permanent water bodies, mangroves and moss and lichen) were treated as 'non-burnable' and hence were not considered in TUD-S4F.

In addition, ESA WorldCover product also provides the annual composite of terrain corrected Sentinel-1 backscatter (VV, VH and VH/VV) and Sentinel-2 normalised differential vegetation index (NDVI) percentiles (10th, 50th and 90th) at 10 m resolution (Van De Kerchove et al., 2020, 2022). The NDVI percentile and Sentinel-1 backscatter composites from WorldCover were used in the high-resolution setup of TUD-S4F to calculate high resolution LAI (Section 3.5.1) and AGB datasets (Section 3.6), respectively.

3.2.7 ESA CCI HRLC land cover

The ESA CCI High Resolution Land Cover (HRLC) dataset is a regional land cover product developed within the ESA Climate Change Initiative (CCI) framework for Amazonia, Northern Siberia and the Eastern Sahel region (Bruzzone et al., 2024b). The primary objective of the HRLC project is to study and investigate the role of the spatial resolution of land cover (LC) and land cover change (LCC) in supporting climate modelling research. The HRLC product involves the accurate description and analysis of land cover and land cover change using Earth Observation (EO) data at high spatial resolution, leveraging the capabilities of the Sentinel-2 satellite constellation (Bruzzone et al., 2024a). This represents a significant improvement over the standard ESA CCI Land Cover maps, which provide global annual land cover classifications at a coarser spatial resolution of 300 m. By providing land cover information at higher spatial resolution, the HRLC dataset enables a more detailed characterisation of heterogeneous landscapes and fine-scale land cover transitions that are not adequately captured by moderate-resolution products. This is particularly relevant for climate modelling applications, where sub-pixel land cover

heterogeneity can have significant impacts on surface energy and water balance parameterisations.

In TUD-S4F-v0.3, the ESA CCI HRLC10 dataset for the year 2019 is used as an alternative to the WorldCover dataset. ESA CCI HRLC is spatially resampled using nearest neighbour interpolation to preserve the integrity of the thematic land cover classifications on the common base grid and then aggregated to target resolution as the fractional coverage of each land cover class. The various tree cover and shrub classes are then summed to an overall tree cover and shrubland cover class, respectively.

3.2.8 ESA CCI above-ground biomass

The ESA CCI Above-Ground Biomass (AGB) dataset is a global product developed within the ESA Climate Change Initiative (CCI) Biomass project (Santoro and Cartus, 2025a). The primary science objective of the project is to provide global maps of above-ground biomass (in Mg ha^{-1}) at hectare scale for multiple epochs. Vegetation biomass is a crucial ecological variable for understanding the evolution and potential future changes of the climate system. The dataset provides spatially explicit estimates of forest above-ground biomass along with related standard deviations as separate map products. The most recent version (v6.0) covers the years 2007, 2010, 2015, 2016, 2017, 2018, 2019, 2020, 2021, and 2022, offering a multi-temporal perspective on global biomass dynamics. The AGB maps are derived from a combination of satellite observations, including data from ALOS-2 PALSAR-2 synthetic aperture radar. Mapping above-ground biomass is a priority for several current and upcoming space agency missions, including GEDI, ICESat-2, BIOMASS, ALOS-4, and NISAR, underscoring the importance of this variable within the broader Earth observation community (Santoro and Cartus, 2024, 2025b).

Within TUD-S4F-v0.3, the ESA CCI AGB dataset is treated as a static multi-annual input and is spatially resampled using bilinear interpolation onto the common base grid, as the product provides annual estimates without sub-annual temporal variability. The dataset is then used to calibrate parameters of the allometry and above-ground biomass module in TUD-S4F and as explanatory variable to predict those parameters in space (Sections 3.3.3 and 3.3.10).

3.2.9 Canopy height

Canopy height is another important input dataset in TUD-S4F, which is used to calibrate the sub-module for the temporal dynamics in canopy height and to subsequently estimate changes in above-ground biomass of trees and SLW vegetation (Section 3.3.2). We currently use different datasets of canopy height in TUD-S4F for different purposes. All of those datasets are based on space-borne Lidar observation from the GEDI sensor onboard the International Space Station (ISS). We have not tested the impact of canopy height on our results.

GEDI L3 1 km gridded: GEDI Level 3 (L3) Gridded Land Surface Metrics product (version 2) provides mean canopy height (RH100) and standard deviation of canopy height at 1 km

resolution (Dubayah et al., 2021b). Inside each 1 km × 1 km grid cell, mean canopy top height (i.e., relative height 100, RH100) and its standard deviation are calculated using all valid GEDI Level 2A footprints. For the use in TUD-S4F-v0.3, this dataset from period from 2019-04-09 to 2021-08-04 was bilinearly interpolated to the 0.0025° resolution and then used as target variable to train regression models for the estimation of canopy height in TUD-S4F (Section 3.3.2).

Pauls et al. 2024 canopy height: Pauls canopy height is a 10 m global canopy height product for the year 2020 (Pauls et al., 2024) (called P24). It is generated using neural network which include Sentinel-1 backscatter (median value in summer leaf-on season) and Sentinel-2 multispectral bands (median value of cloud-masked images) as predictor variables for canopy height (RH100) from GEDI footprints. The P24 dataset has a RMSE of 4.73 m overall and 6.72 m for trees taller than 5 m. P24 data is used to estimate high resolution AGB data for the TUD-S4F-v0.3-HR setup (Section 3.6).

3.2.10 ALOS-2 PALSAR-2 mosaic datasets

ALOS-2 PALSAR-2 mosaic datasets (version 2.6.0) provides global 25 m resolution L-band SAR backscatter produced by Japan Aerospace Exploration Agency (Japan Aerospace Exploration Agency, 2026). The backscatter includes HH and HV polarisation and has been geometrically and radiometrically calibrated. PALSAR-2 mosaics is annually updated and covers the period 2015 to 2022. This dataset is used to estimate high resolution AGB data for the TUD-S4F-v0.3-HR setup (Section 3.6).

3.2.11 CGLS ASCAT soil water index

The soil water index from Metop/ASCAT is a proxy for soil moisture at various depths. The dataset is available at a resolution of 0.1° for the period since 2007 (Bauer-Marschallinger et al., 2018; Copernicus Land Monitoring Service and Copernicus Land Monitoring Service helpdesk, 2025). The SWI dataset is already available as a 10 daily product, which uses the same 10 daily interval as the LAI dataset (Raml, 2025). Therefore, temporal resampling is not necessary. We resampled the SWI dataset to the spatial resolution of the LAI data set using bilinear interpolation.

In TUD-S4F, the SWI is used as a proxy for surface dead fuel moisture content (DFMC = SWI) and to estimate live woody fuel moisture content. Although the SWI might not be fully reliable as a proxy for soil moisture in dense tropical forests because of the attenuation of microwaves in the vegetation layer, we assume that soil and vegetation moisture are correlated, which allows us to use the SWI as a predictor for LFMC and as proxy for DFMC.

3.2.12 VOD2LFMC live fuel moisture content

The VOD2LFMC live-fuel moisture content (LFMC) dataset (Forkel et al., 2022, 2023a) provides daily estimates of LFMC at a spatial resolution of 0.25° for the period February 2000 to July 2017. The dataset has been created by calibrating an empirical model that predicts LFMC from the Globe-LFMC database (Yebra et al., 2019) of in situ measurements

by using Ku-band VOD as a predictor. We temporally resampled the dataset to the 10-daily time steps of the CGLS LAI dataset by using the mean LFMC for each 10-daily interval, and spatially resampled the dataset to the 0.0025 grid using bilinear interpolation. The VOD2LFMC dataset was used to calibrate the computation of LFMC in the TUD-S4F model (Section 3.3.6).

3.2.13 Databases of ground observations

Several databases of ground observations were used to constrain parameters or to validate the TUD-S4F model. The Biomass and Allometry Database (BAAD) provides information on biomass stocks in different tree components such as stems, branches, leaves, and roots (Falster et al., 2015). The database includes 259,634 measurements from 176 studies and 678 tree species. We use BAAD to constrain parameters of the allometry module of the TUD-S4F model to represent the relationships between tree height, stems, branches, and leaves biomass. For calibration and/or validation of further model components, we used the Global Database of Litterfall Mass and Litter Pool Carbon (Holland et al., 2014) (H14), the database of fuel loads, combustion completeness and fuel consumption (van Wees et al., 2022) (vW22), and the databases of emission factors for savannahs by Vernooij et al. (2023) (V23) and globally by Andreae (2019) (A19).

3.3 TUD-S4F-v0.3 model description and data-model fusion steps

In TUD-S4F, the dynamics of biomass, fuel loads, fuel moisture, fuel consumption and fire emissions are estimated in a series of sub-modules (Figure 7):

(a) Temporal dynamics in the fractional cover of vegetation types are estimated from long-term dynamic mean LAI and land cover (Section 3.3.1).
(b) The temporal dynamics of canopy height (CH) are estimated from long-term dynamic mean LAI and fractional cover. Therefore, different alternative regression models are trained against GEDI canopy height. Canopy height is not used as direct input to the model because it does not provide any information on temporal changes in canopy height (Section 3.3.2).
(c) The estimated canopy height is used in the allometry modules to estimate biomass compartments of trees (Section 3.3.3) and of surface live woody (SLW) vegetation (Section 3.3.4), respectively. Thereby, temporal changes in canopy height, as estimated in Step *(b)* directly translate into changes in stem, branches and leaf biomass of trees and of woody and leaf biomass of SLW vegetation, respectively.
(d) Herbaceous biomass of grasslands and croplands is directly estimated from the 10-daily LAI time series (Section 3.3.5).
(e) Biomass turnover from living vegetation to the surface (e.g., leaf fall, transfer of woody biomass) is computed as the temporal reduction in herbaceous, tree and SLW biomass, which is mainly driven by temporal changes in the gap-filled and filtered CGLS LAI dataset. A reduction in 10-daily LAI directly causes a reduction in the corresponding amount of leaf and herbaceous biomass, and hence causes turnover. Woody turnover occurs when a reduction in canopy height occurs, which itself is caused by a reduction in the fractional cover of tree and SLW vegetation and long-

term mean LAI. A fraction of the woody turnover is assumed to be extracted through harvest or through the collection of small diameter firewood and hence does not enter the woody debris pools. The same applies to turnover of herbaceous vegetation in croplands, where a fraction is assumed to be extracted through harvest (Section 3.3.6).

(f) Fuel moisture content (FMC) of herbaceous and leaf live fuels is computed from LAI and SWI. Moisture content of live wood is estimated from SWI. Dead fuel moisture content is directly taken from SWI (Section 3.3.7).

(g) Combustion completeness (CC) is then computed as a linear function of FMC for each respective fuel compartment and the functions are calibrated against the field-based CC data from vW22 (Section 3.3.8).

(h) Dead surface fuel loads are represented in terms of litter, fine and coarse woody debris (FWD and CWD). They are estimated from the biomass turnover from herbaceous vegetation and leaves, and from woody biomass, respectively. The decomposition of litter, FWD, and CWD is represented by fixed decomposition rates (Section 3.3.9).

(i) Dry matter burnt is then estimated during fire occurrence from burnt area, combustion completeness, and fuel loads by taking into account the estimated fire type (surface or canopy fire) (Section 3.3.10).

(j) Dry matter burnt of each fuel component is then provided to a chemical-based fire combustion model to estimate heat yields, fire radiative energy (FRE), and emission factors. The dynamic emission factors depend on parameters for the relative composition of a fuel component by lignin, cellulose, and volatiles, its fuel moisture content, and a parametrisation of the equivalent oxygen-to-fuel ratio. The emission factors are multiplied by dry matter burnt to receive total emissions of a trace gas species (Section 3.3.11).

(k) Parameters of several modules of the TUD-S4F model are calibrated in various steps, which include calibration of the parameters of individual modules (see respective sections) as well as joint calibration of the full model for individual grid cells against ESA CCI AGB, GEDI canopy height, VOD2LFMC, FRE and a spatial prediction of the locally calibrated parameters (Section 3.3.12).

3.3.1 Fractional cover and LAI of vegetation types

Satellite products of land cover and canopy height provide mostly static maps, which do not account for vegetation cover and height changes that would, however, strongly affect the accumulation of surface fuels. In order to compute temporal changes in vegetation cover (and subsequently in canopy height, biomass and surface fuels), we assume that the long-term dynamic mean LAI is a realistic proxy for changes in the fractional cover of woody vegetation. Long-term dynamic mean LAI (LAI_{mean}) is computed as a Locally Weighted Scatterplot Smoothing (LOWESS) of the 10-daily filtered CGLS LAI time series (LAI_{Obs}) and hence represents the long-term non-linear trend component of the LAI time series:

$$LAI_{mean} = LOWESS(LAI_{Obs})$$

LOWESS is applied using 40% of a time-series data (frac), residual-based reweighting of 1 (it) and always use linear-interpolation instead of weighted regression (delta=0).

LAI_{mean} is then used to compute dynamic mean fractional vegetation cover ($fCover_{mean}$) based on the Beer-Lambert law for light extinction:

$$fCover_{mean} = 1 - e^{(-k \cdot LAI_{mean})}$$

whereby *k* is the light extinction coefficient with a default value of *k = 0.5* (Monsi and Saeki, 1953, 2005) (Table 6). Subsequently, the fractional cover of trees (f_{Tree}) and of SLW vegetation (f_{SLW}) is computed by multiplying $fCover_{mean}$ with the relative frequency of the respective land cover (LC) classes per grid cell:

$$f_{Tree} = LC_{Tree} \cdot fCover_{mean}$$

$$f_{SLW} = LC_{SLW} \cdot fCover_{mean}$$

$$f_{Woody} = f_{Tree} + f_{SLW}$$

whereby LC_{Tree} *and* LC_{SLW} equal the relative frequency of forest and shrubland cover in a grid cell, respectively. f_{Woody} is the total fractional cover of woody vegetation. The fractional cover of herbaceous vegetation is then computed based on the relative frequency of herbaceous land cover (LC_{Herb}), which is the sum of the relative frequencies of grasslands, croplands and herbaceous wetlands from the WorldCover dataset, and based on the fractional space that is not occupied by woody vegetation cover within the tree and shrubland land cover classes:

$$f_{Herb} = LC_{Herb} + ((LC_{Tree} + LC_{SLW}) - f_{Woody})$$

The fractional cover of herbaceous and woody vegetation is then used to split the 10-daily LAI data into an herbaceous and woody component:

$$LAI_{Herb} = \max\left(3, \frac{f_{Herb}}{f_{Herb} + f_{Woody}} \cdot LAI_{Obs}\right)$$

$$LAI_{Woody} = LAI_{Obs} - LAI_{Herb}$$

Thereby we assume that the maximum LAI of herbaceous vegetation is 3. While the original LC input data is a static map (e.g. for the year 2020 when using WorldCover v100), the use of the long-term dynamic mean LAI introduces long-term dynamics in the computation of the fractional cover and LAI of the different vegetation types. Specifically, this calculation implies a dynamic interaction between the fractional cover of herbaceous and woody vegetation, where an increase or reduction in woody vegetation results in a decline or increase of herbaceous cover and LAI, respectively (Figure 8).

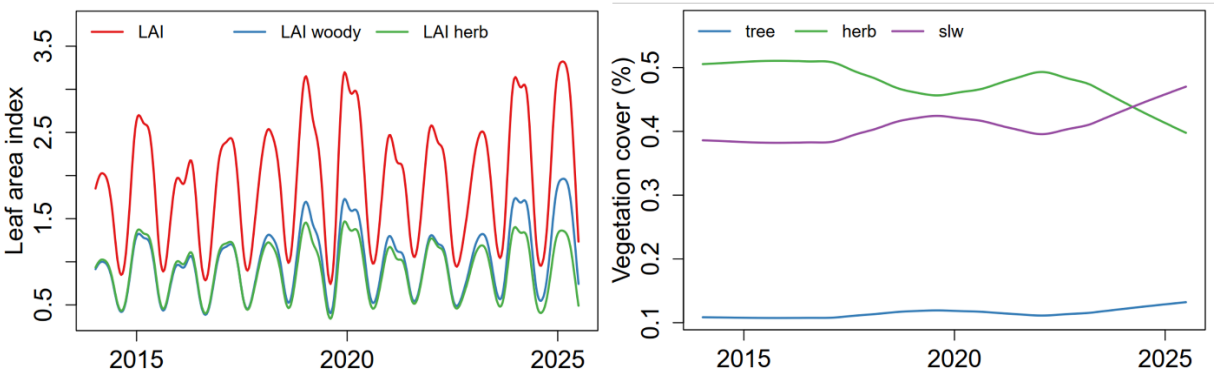

Figure 8: Example time series of LAI and vegetation cover for a grid cell at 14.59625°E and 6.7987°S from the TUD-S4F-v0.3-S311 setup. The observed gap-filled and filtered CGLS LAI (red, left) dataset is used together with class frequencies from ESA WorldCover to compute multi-annual dynamics in fractional vegetation cover of trees, herbaceous and SLW vegetation (right) and to split LAI in a woody and herbaceous component (left).

Table 6: Overview of parameters for the TUD-S4F-v0.3 model. "Prior" refers to the initial default value of a parameter that was taken from the literature or estimated from databases. The "Optimised" value was calibrated and used in model runs.

Name (unit)	Prior	Min	Max	Optimized	Description, reference and calibration
k	0.5	0.3	0.9	--	Light extinction coefficient to compute fCover
h1	1	0.7	1.3	Spatially varying	Local scalar and offset for canopy height, calibrated locally against GEDI canopy height
h2	0.001	-10	10	Spatially varying	
a0	1	0.2	3	Spatially varying	Local scalar for stem biomass, calibrated locally against above-ground biomass from ESA CCI
a1 (kg m-2)	0.0199	0.001	0.7	0.4915	Parameters for allometric relations between canopy height and tree stem, branch and leaf biomass, Calibration against BAAD (Falster et al., 2015)
a2	0.4666	0.1	1.5	0.7751	
a3	0.2095	0.05	2.5	0.4611	
a4	0.9244	0.3	2	1.5286	
a5	0.1172	0.01	2.5	0.0533	
a6	2.3018	0.5	3.5	1.2904	
s1	1.403	--	--	--	Parameters for allometric relations between SLW height, canopy diameter and biomass, taken from Conti et al. 2019
s2	-0.370	--	--	--	
s3	1.903	--	--	--	
s4	0.652	--	--	--	
CD (m)	2.5	0.5	7	1.8255	Crown diameter of shrubs, calibrated against above-ground biomass from ESA CCI
sla (m^2/kg)	6	5	70	6.4028	Specific leaf area to compute herbaceous biomass from LAI (Reich et al., 1998), calibrated against fuel load data (van Wees et al., 2022)
t_{leaf} (yr^{-1})	0.5	0.2	1	--	Constant baseline turnover rate for leaves (Kikuzawa et al., 2013)
t_{wood} (yr^{-1})	0.0319	0.01	0.5	--	Constant baseline turnover rate for stem and branches (Johnson et al., 2016)
t_{slw} (yr^{-1})	0.01	0.003	0.1	--	Constant turnover rate for SLW biomass (Li et al., 2020)
f_{sb}	0.5	0.05	0.95	--	Fraction of small branches (diameter < 3" = 7.62 cm) relative to all branches for turnover of branches biomass to FWD or CWD
HI_{herb}	0.6	0	1	--	Maximum harvest index of herbaceous biomass at 60% cropland cover
$HI_{branches}$	0.8	0	1	--	Harvest index for small branches biomass
HI_{stem}	0.9	0	1	--	Harvest index for stem and large branches
k_l (yr^{-1})	0.98	0.2	1.1	--	Decomposition rate for annual herbaceous litter (Zhao et al., 2025)
k_{fwd} (yr^{-1})	0.344	0.0336	2.433	--	FWD decomposition rate, estimated from (Harmon et al., 2020) for tropical deciduous forests
k_{cwd} (yr^{-1})	0.172	0.0112	0.811	--	CWD decomposition rate, estimated from (Harmon et al., 2020) for tropical deciduous forests

f0_leaf	0.8	0.005	1.699	Spatially varying	Parameters for the relationships between LFMC, LAI and SWI. Calibrated against the VOD2LFMC dataset
f0_herb	0.964	0.005	1.939	Spatially varying	
f1_leaf	1.012	-0.457	2.461	Spatially varying	
f1_herb	1.403	0	3.577	Spatially varying	
f2_leaf	1	-371.4	211.871	Spatially varying	
f2_herb	1	-250.774	291.860	Spatially varying	
m1 (%)	20	5	50	20	Minimum (m1) and maximum (m2) moisture content of living wood (Glass and Zelinka, 2010)
m2 (%)	200	100	300	Spatially varying	
c1_herb (%)	250	46	420	234.57	c1: Maximum FMC of the respective fuel component at which combustion completeness = 0 c2: Maximum combustion completeness of the respective fuel component Calibration against statistical distributions of combustion completeness data (van Wees et al., 2022)
c2_herb	0.95	0.2	1	0.99	
c1_leaf (%)	250	21	190	105.97	
c2_leaf	0.95	0.2	1	0.99	
c1_wood (%)	150	30	400	118.5	
c2_wood	0.3	0.2	0.7	0.59	
c1_litter (%)	100	6	62	34.54	
c2_litter	0.9	0.18	1	0.93	
c1_fwd (%)	100	30	300	34.54	
c2_fwd	0.7	0.5	1	0.69	
c1_cwd (%)	100	5	50	27.65	
c2_cwd	0.5	0.01	0.35	0.18	
fvo_leaf	0.27	0.0	0.27	0.00082	Fraction of volatiles (fvo) and cellulose (fce) of the respective fuel components; priors from (Rego et al., 2021; Waliszewska et al., 2021); Calibration against regional distributions of emissions factors and modified combustion efficiency (Andreae, 2019; Vernooij et al., 2023)
fce_leaf	0.54	0.40	0.9	0.799	
fvo_herb	0.09	0.0	0.27	0.0	
fce_herb	0.51	0.47	0.9	0.799	
fce_wood	0.70	0.40	0.8	0.489	
eofr_leaf	0.95	0.84	0.98	0.893	Equivalent oxygen to fuel ratio for combustion of the respective fuel components Calibration against regional distributions of emissions factors and modified combustion efficiency (Andreae, 2019; Vernooij et al., 2023)
eofr_herb	0.95	0.84	0.98	0.959	
eofr_wood	0.95	0.84	0.98	0.849	
eofr_litter	0.95	0.84	0.98	0.959	
eofr_fwd	0.95	0.84	0.98	0.956	
eofr_cwd	0.89	0.84	0.99	0.846	
frad	0.189	0.09	0.25	0.189	Fraction of FRE to total heat release
ef_nox_tree	2.31	--	--	--	Emission factor for NOx emissions (g NO/kg) for forested pixels ($f_{tree} > 0.5$) and herbaceous pixels ($f_{tree} \leq 0.5$), Median value for topical forests and savannahs from (Andreae, 2019)
ef_nox_herb	2.39	--	--	--	

3.3.2 Canopy height

The long-term dynamic mean LAI and fractional vegetation cover are then used as predictors for canopy height *H* based on a regression model *f()*:

$$H = h1 \cdot f(LAI_{mean}, f_{tree}, f_{SLW}) + h2$$

We tested several regression models for *f()*, such as Ordinary Least Squares (OLS) linear regression, Generalised Linear Models (GLM) with a Gamma distribution, Generalised Additive Models (GAM), and Random Forest regression (RF). The performance of the four regression models was similar for the training and test datasets (Table 7). However, Model 1 (OLS) showed heteroscedasticity in the residuals and Model 4 (RF) resulted in unrealistic step-jumps in the estimated canopy height time series. Although Model 3 (GAM) performed slightly better than Model 2 (GLM), it took more processing time in application. Hence, we decided to use Model 2, the GLM, which is suited to estimate positive-only values of canopy height by using the Gamma link function.

The parameters *h1* and *h2* in the equation are used as scalars to fit the estimated canopy height from the regionally calibrated regression model to the canopy height data of individual grid cells, i.e., to eliminate local biases from the regression model (Figure 9). Spatial fields of the parameters *h1* and *h2* are estimated in the later model-data fusion exercises (Section 3.3.12).

Table 7: Overview about the performance of several regression models to predict canopy height from dynamic mean LAI (LAI_{mean}) and fractional vegetation cover of trees and SLW (f_{Tree} and f_{SLW}). Performance metrics were computed between estimated canopy height and the canopy height from GEDI L3 based on a sample of thousands of testing grid cells across the S-Africa study region. AIC is computed based on the training data.

Model	**Correlation (r)**	**Root-mean squared error (RMSE)**	**Kling-Gupta efficiency (KGE)**	**Akaike's Information Criterion (AIC)**
M1 (OLS): H = a LAI_{mean}^2 + b LAI_{mean} + c f_{Tree} + d f_{SLW}	0.74	3.12 m	0.73	74293
M2 (GLM with Gamma link): ln(H) = a LAI_{mean}^2 + b LAI_{mean} + c f_{Tree} + d f_{SLW}	0.78	2.73 m	0.68	66384
M3 (GAM): H = s(LAI_{mean}) + s(f_{tree}) + s(f_{SLW})	0.79	2.66 m	0.71	69557
M4 (RF): H = RF(LAI_{mean}, f_{tree}, f_{SLW})	0.78	2.72 m	0.71	--

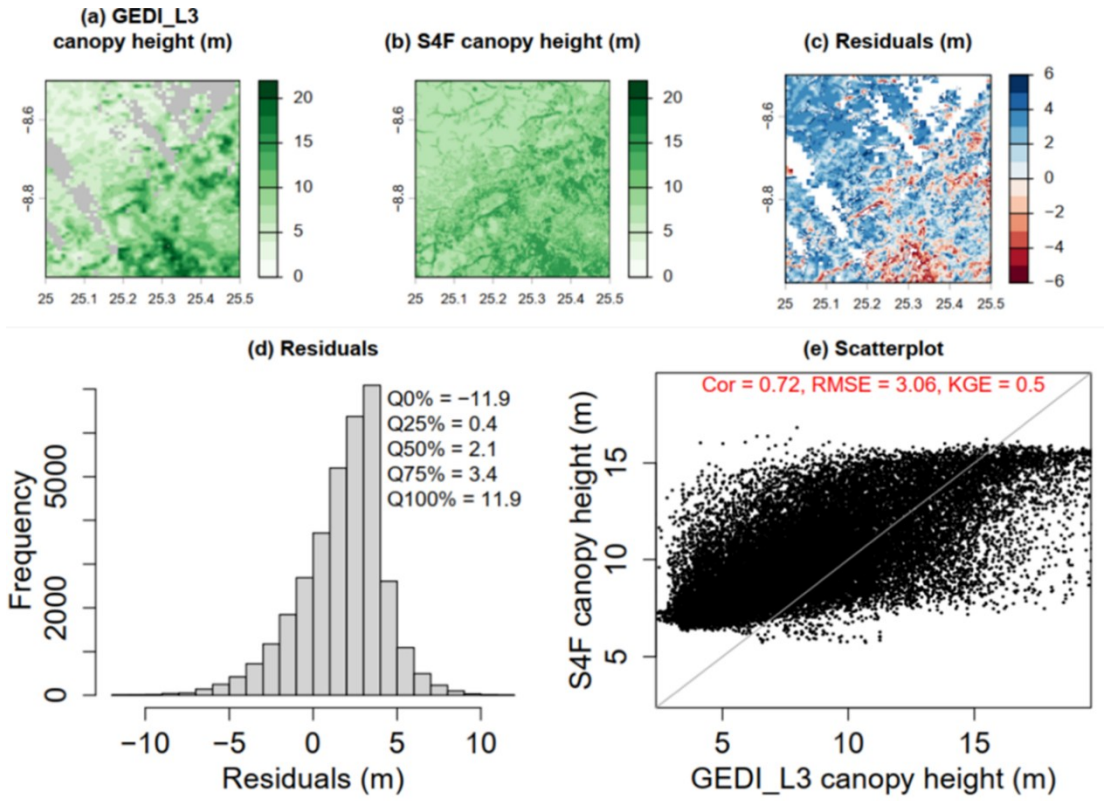

Figure 9: Regional example of canopy height from (a) GEDI L3 data and (b) TUD-S4F-v0.3-S311. (c and d) Spatial patterns and histogram of the residuals, and scatter plot (e).

3.3.3 Biomass and allometry of trees

After the computation of canopy height, allometric equations are used to compute the biomass in stems, branches, and leaves of trees. Stem biomass BM_{stem} is computed from canopy height based on an exponential allometric function:

$$BM_{stem} = f_{Tree} \cdot a0 \cdot \left(a1 \cdot H^{\frac{1}{a2}}\right)$$

whereby *a1* and *a2* are the parameters of the allometric relationship. The parameter a0 is used as a local scaling factor to fit the estimated woody biomass of individual grid cells based on the woody biomass of the ESA CCI biomass map, which is done during the local calibration and prediction of parameters (Section 3.3.12). Following the approach used by Thurner et al. (2014) to estimate different biomass compartments from satellite datasets, we then use BM_{stem} to compute the biomass of tree branches $BM_{branches}$ and leaves BM_{leaves} as:

$$BM_{branches} = a3 \cdot BM_{stem}^{\frac{1}{a4}}$$

$$BM_{leaf} = a5 \cdot BM_{stem}^{\frac{1}{a6}} \cdot P$$

whereby *a3* to *a6* are the allometric parameters and *P* is the phenology status of tree canopies. P is zero in the case of no leaf cover and one in the case of full leaf cover. The

phenology status is computed from the seasonal fractional cover of woody vegetation ($fCover_{Woody}$) as:

$$fCover_{Woody} = 1 - e^{(-k \cdot LAI_{Woody})}$$

$$P = \frac{fCover_{Woody} - fCover_{Woody,min}}{fCover_{Woody,max} - fCover_{Woody,min}}$$

Whereby $fCover_{Woody,min}$ and $fCover_{Woody,max}$ are the minimum and maximum seasonal fractional cover of woody vegetation values of a grid cell, respectively. Please note that here $fCover_{Woody}$ is computed from the seasonally changing woody LAI and hence represents seasonal dynamics, unlike the mean dynamic $fCover_{mean}$ used in the calculation of fractional vegetation cover.

The parameters *a1* to *a6* for the tree biomass module were calibrated against the Biomass and Allometry Database (BAAD) (Falster et al., 2015) and against total above-ground woody biomass from the ESA CCI biomass map (Santoro et al., 2021). Therefore, the reported values for tree height in BAAD and tree height from GEDI from sampled grid cells in the study region were used jointly to estimate leaf, branches, stem, total wood, and total above-ground biomass and to calibrate the allometry module against the respective biomass components from BAAD and against total above-ground woody biomass from ESA CCI). From BAAD, only values from tropical or savannah angiosperm tree species were considered.

3.3.4 Biomass and allometry of surface live woody vegetation

Surface live woody (SLW) vegetation has been newly implemented in TUD-S4F-v0.3. We here define SLW vegetation as small woody vegetation, specifically shrubs in shrublands or understorey woody vegetation in forests. However, we currently only use the implementation for shrublands as we currently lack any reliable dataset about the coverage of understorey vegetation in forests. This will be however considered in future developments of TUD-S4F.

Different allometric models have been used for the estimation of shrub biomass depending on the physiognomy of the shrubs and the available measurements (Conti et al., 2019). While shrubs with a single main stem are already represented in TUD-S4F through the allometry module for trees (previous section), the area-based biomass density of multi-stemmed shrubs can be estimated from shrub height and canopy diameter (*CD*):

$$BM_{SLWwood} = \mathrm{s1} \cdot e^{\mathrm{s2}+s3 \cdot \ln(CD)+\mathrm{s4} \cdot \ln(H)} \cdot \frac{N}{A}$$

The model and the parameters *s1* to *s4* are directly taken from (Conti et al., 2019). The ratio of the number of shrubs *N* to the area *A* of a grid cell is used to scale individual-level biomass to the grid cell. This model can be relatively easily applied to remote sensing data as *CD*, *H* and *N* can be estimated from high-resolution satellite imagery. However, in the current setup in TUD-S4F-v0.3, we treat *CD* as a fixed model parameter that we include in

model calibration Table 6. *N* can be estimated from the canopy area (*CA*) and the fractional coverage of SLW in a grid cell as:

$$CA = \frac{\pi}{4} CD^2$$

$$N = \frac{A \cdot f_{SLW}}{CA}$$

The leaf biomass of SLW vegetation is then estimated based on the ratio of leaf to woody biomass r_{LW} in shrubs and the woody phenology status *P*:

$$BM_{SLW\text{leaf}} = r_{LW} \cdot BM_{SLW\text{wood}} \cdot P$$

According to Hiernaux et al. (2023), the ratio of leaf to woody biomass in woody Savannah vegetation of the Sahel can be estimated from canopy area (CA) as:

$$r_{LW} = 0.028 + 0.651 \cdot CA^{-0.719}$$

We kept all parameters *s1* to *s4* and the parameters to compute r_{LW} constant at the values provided by Conti et al. (2019) and Hiernaux et al. (2023), respectively. Only the parameter *CD* was included in the calibration of both tree and SLW allometry modules against ESA CCI above-ground biomass (AGB). Total AGB is defined as the total biomass of tree and SLW biomass:

$$BM_{agb} = BM_{stem} + BM_{branches} + BM_{leaf} + BM_{SLWwood} + BM_{SLWleaf}$$

An example of total AGB and the different biomass components is given for one grid cell in Figure 10 and a regional comparison of AGB from ESA CCI and TUD-S4F is provided in Figure 11.

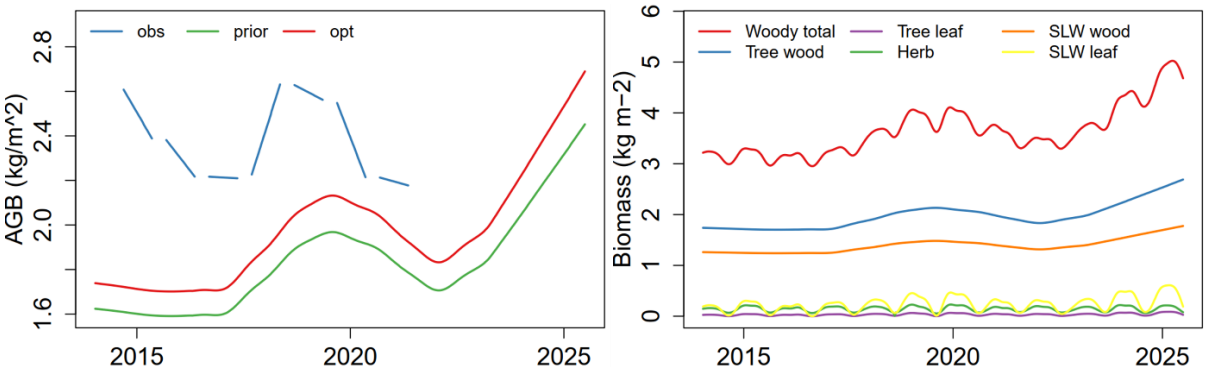

Figure 10: Example time series of above-ground biomass (left) and biomass components (right) for a grid cell at 14.59625°E and 6.7987°S from the TUD-S4F-v0.3-S311 setup. (left) AGB from ESA CCI (obs) and from TUD-S4F prior to the calibration of allometry parameters and with optimised allometry parameters (opt).

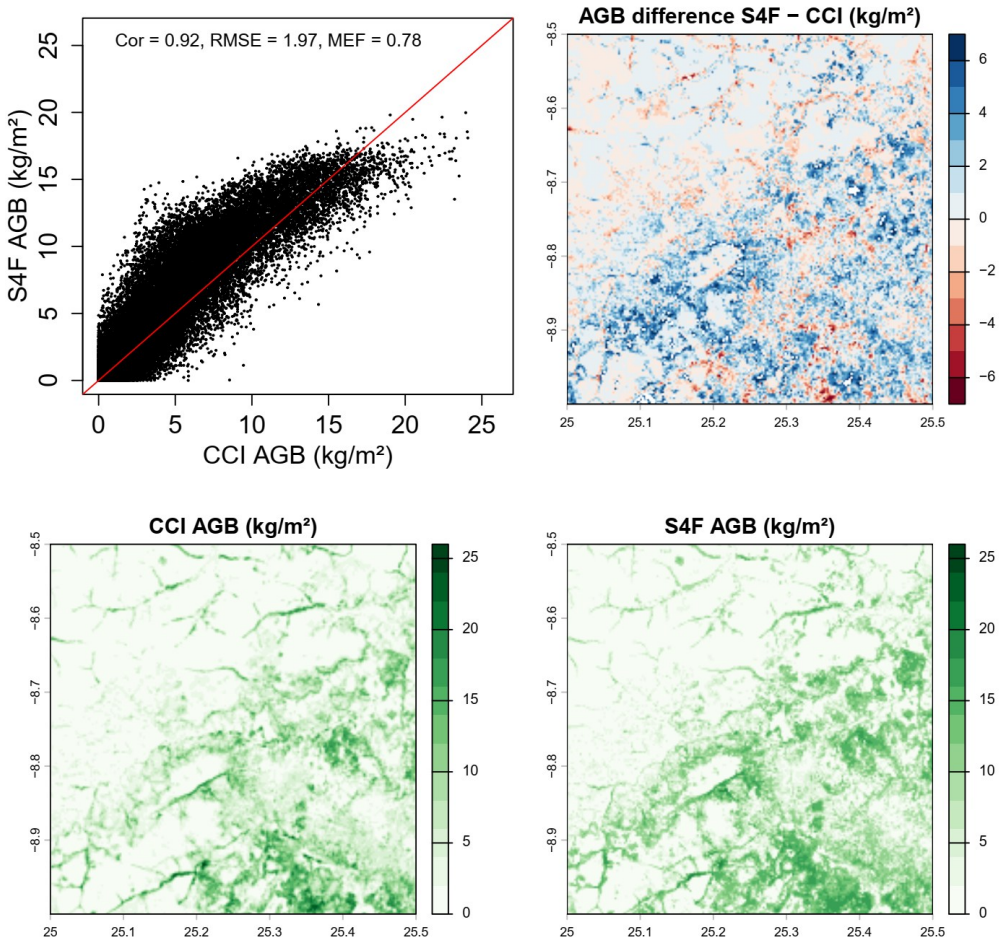

Figure 11: Regional comparison of above-ground biomass from ESA CCI data and TUD-S4F-v0.3-S311.

3.3.5 Herbaceous biomass

Herbaceous biomass is estimated from herbaceous LAI, as several studies report linear relationships between LAI and grass biomass (Makuma-Massa et al., 2017; Punalekar et al., 2018; Schwieder et al., 2020). Thereby, we estimate herbaceous biomass (BM_{herb}), from herbaceous LAI (LAI_{Herb}) and the specific leaf area *sla*:

$$BM_{herb} = \frac{LAI_{Herb}}{sla}$$

The parameter *sla* has been optimised using a quantile-based cost function so that the statistical distribution of BM_{herb} agrees with the observed regional statistical distribution of herbaceous biomass from in situ observations as reported in the updated fuel consumption database (van Wees et al., 2022) (Figure 12). Therefore, the Nelder-Mead (downhill simplex) optimiser is used to minimise the following cost function:

$$J_Q = \sqrt{\frac{1}{n}\sum_{q=0.05}^{q=0.95}\left(\frac{sim_q + 1}{obs_q + 1} - 1\right)^2}$$

whereby sim_q and obs_q are the quantiles 0.05 to 0.95 (in steps of 0.05, i.e. n = 19) of the simulated and observed herbaceous biomass, respectively.

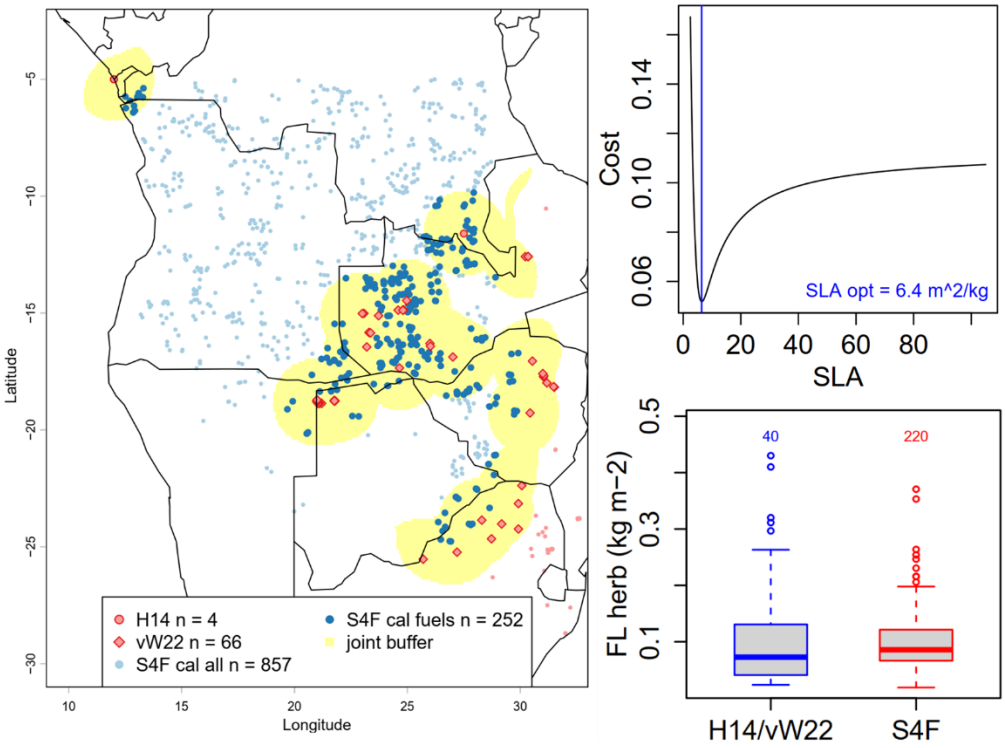

Figure 12: Calibration of herbaceous biomass. Random grid cells from the study region are sampled in a buffer of 200 km around locations for which observations are available from the H14 and vW22 databases (left). Herbaceous biomass is available for 40 sites and the statistical distribution estimated in TUD-S4F is fitted to the observed distribution (lower right) by calibrating the specific leaf area *sla* (upper right).

3.3.6 Biomass turnover and harvest

The turnover T of living vegetation biomass to surface fuels is estimated as a constant baseline turnover T_B and from the temporal changes in the estimated biomass of tree leaves, branches, and stems; of SLW woody and leaf biomass; and of herbaceous vegetation. The baseline turnover is:

$$T_{B,i} = BM_i \cdot (1 - e^{-t_i})$$

Whereby the index i refers to different biomass components (herbaceous, tree leaf, tree branches, tree stem, SLW wood and SLW leaf biomass) and the parameter t_i is a constant turnover rate, which is defined for leaves and wood separately (Table 6). Values for the turnover rates t were taken for leaf biomass from Kikuzawa et al. (2013) and for woody biomass from Johnson et al. (2016), who report stem mortality rates for Amazon forests.

The total turnover of a biomass component is then:

$$T_i = (1 - HI_i) \cdot T_{B,i} + \begin{cases} \Delta BM_i = (BM_{i,t-1} - BM_{i,t}) \; if \; \Delta BM_i > 0 \\ 0 \; if \; \Delta BM_i < 0 \end{cases}$$

Whereby *HI* is the harvest index, i.e. a fraction of the turnover that is removed from the ecosystem through harvest and hence does not enter the woody debris and litter pools. The turnover of branches biomass is split into small (diameter < 7.62 cm) and large branches (> 7.62 cm), which are assigned to the turnover of stems:

$$T_{branches,small} = T_{branches} \cdot f_{SB}$$

$$T_{stem} = T_{stem} + T_{branches} \cdot (1 - f_{SB})$$

The turnover of leaves and herbaceous biomass will enter the litter surface fuels. Turnover of small branches will enter fine woody debris (FWD) and turnover of large branches and stems will enter coarse woody debris (CWD) (Section 3.3.9).

For tree trunks and large branches (diameter > 7.62 cm), we assume HI_{stem} = 0.9, i.e. a strong extraction of woody material. For small branches, we assume $HI_{branches}$ = 0.8, i.e. a strong extraction of small woody materials, e.g. for the use as firewood, which is a common energy source in Africa (Agoundedemba et al., 2023). For herbaceous biomass at croplands, we assume HI_{herb} = 60% at 100% cropland cover and HI_{herb} = 0% at 0% cropland cover (i.e., no harvest at grid cells that are purely classified as grasslands).

Please note that the *HI* values differ substantially from the values used for TUD-S4F-v0.2 in Forkel et al. (2025) for the Amazon and Cerrado. Although the *HI* parameters account for a real biomass flux, they cannot be estimated reliable regionally or spatially as they depend on local human activities. However, the *HI* parameters can be treated as effective parameters to scale regional distributions of surface litter, FWD and CWD.

3.3.7 Fuel moisture content

Fuel moisture content in forest fire research is commonly defined as the amount of water over the dry biomass of a vegetation sample (Yebra et al., 2013) and hence is related to wet biomass (BMwet) and dry biomass (BMdry) as:

$$FMC = \frac{BM_{wet} - BM_{dry}}{BM_{dry}} \cdot 100\,\%$$

FMC is commonly distinguished between living and dead vegetation components (LFMC and DFMC). In TUD-S4F model, FMC is represented separately for live fuels (e.g., leaves $LFMC_{leaf}$, herbaceous vegetation $LFMC_{herb}$, and live wood $LFMC_{wood}$) and for dead fuels (DFMC).

Please note that DFMC is not distinguished between litter, FWD, and CWD, which typically would react as 1, 10, or 100 hours fuels increasingly slower to changes in ambient relative humidity (e.g. Cawson et al., 2020; Matthews, 2014; Rodrigues et al., 2024; Zhao et al., 2021). DFMC is not specifically modelled in TUD-S4F but taken from the soil water index (SWI). SWI represents the percentage of soil saturation (0-100%) of the surface layer which is however comparable to ranges that are commonly reported for dead fuels (Rodrigues et al., 2024). We acknowledge this current limitation of TUD-S4F and see this as a priority area for improvement in future work.

LFMC is estimated from LAI and SWI as those variables are highly correlated in many ecosystems (Forkel et al., 2022; Lu and Wei, 2021; Yebra et al., 2013). We estimate the LFMC of leaves and herbaceous vegetation based on a single-layer perceptron (as used in neural networks) with a sigmoidal activation function by using LAI and SWI as input:

$$LFMC_i = \frac{400\,\%}{1+e^{-(z-f_{0,i})}}$$

$$z = \left(f1_i \cdot \frac{LAI-LAI_{min}}{LAI_{max}-LAI_{min}}\right) + \left(f2_i \cdot \frac{SWI-SWI_{min}}{SWI_{max}-SWI_{min}}\right)$$

Whereby the sigmoidal activation function ensures that LFMC is scaled to plausible ranges between 0% and 400%. LAI and SWI are normalised based on the temporal minimum and maximum LAI and SWI in each grid cell. The parameters *f0*, *f1*, and *f2* are calibrated for different vegetation types *i* (leaves and herbaceous) against the VOD2LFMC dataset (Forkel et al., 2022).

Woody FMC (FMC_{wood}) is assumed to increase linearly with SWI, whereby FMC_{wood} equals to a minimum woody moisture content *m1* if SWI ≤ 10% and increases up to a maximum woody moisture content *m2* at the 99%-ile of SWI (Table 6).

3.3.8 Combustion completeness

We combustion completeness (CC) is estimated as a function of FMC following the approach which is used in GFED (van der Werf et al., 2006, 2010, 2017). In GFED, CC depends on modelled soil moisture. In TUD-S4F, the CC of leaves and wood depends on leaf and woody FMC, respectively, and the CC of litter, FWD, and CWD depends on SWI. CC declines linearly with FMC from a maximum CC (parameter *c2*) at FMC = 0% to CC = 0 at a maximum FMC (parameter *c1*). Parameters c1 and c2 are defined for leaves, live wood, herbaceous vegetation, litter, FWD and CWD.

The parameters *c1* and *c2* are calibrated to fit observed regional statistical distributions of CC from the database by (van Wees et al., 2022). Therefore, the same approach and quantile-based cost function like in the calibration for herbaceous biomass was used (Section 3.3.5). As a result, the estimated parameters *c1* and *c2* are able to reproduce the regional statistical distributions of CC, although the reference data is limited specifically for leaves, litter and woody debris (Figure 13).

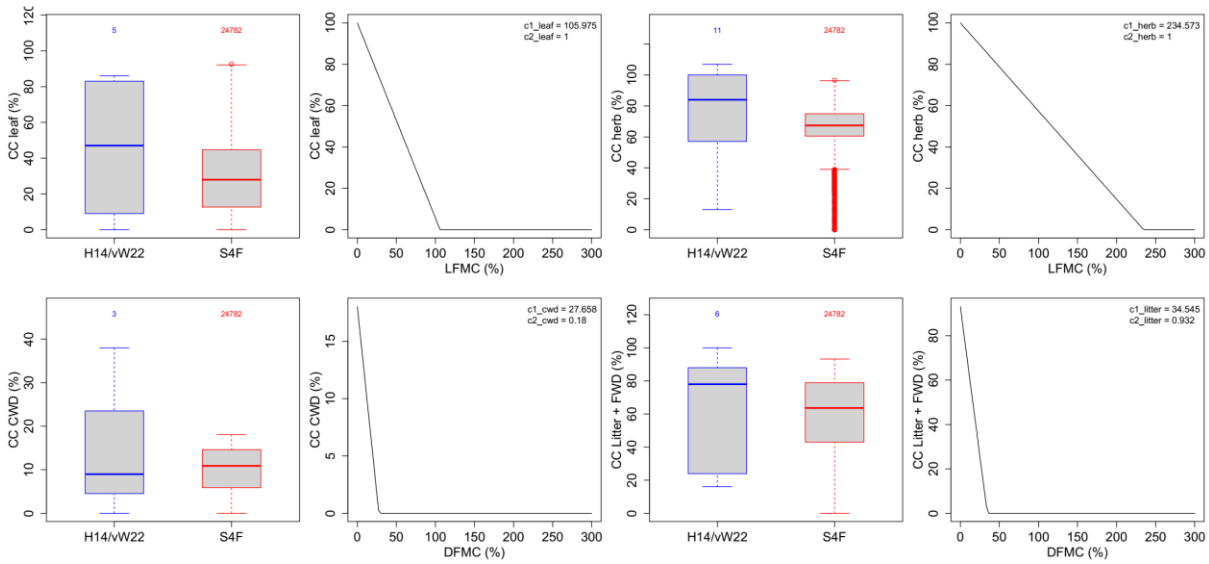

Figure 13: Results of the calibration of parameters for combustion completeness for the TUD-S4F-v0.3-S311 setup against regional distributions of CC from van Wees et al. (2022).

3.3.9 Surface fuel dynamics

Based on the turnover fluxes of leaves and herbaceous vegetation and of branches and stems, we estimate the dynamics of litter (L), FWD, and CWD, respectively. Following the definitions in the North American Wildland Fuel Database (Prichard et al., 2019), L is dead leaf or herbaceous material, and FWD and CWD are dead woody material with a diameter threshold below and above 3" = 7.62 cm, respectively. Hence, CWD also includes dead woody trunks.

To estimate initial values of L, FWD, and CWD fuel loads, we assume that those fuel load components are in a steady state, i.e., the input of biomass from turnover equals the release of carbon through decomposition (heterotrophic respiration). The steady state assumption was previously applied to remotely sensed estimates of gross primary production, above-ground biomass, and soil carbon stocks to estimate total ecosystem carbon turnover times globally (Carvalhais et al., 2014; Fan et al., 2020) and can also be used to estimate surface fuels (Rego et al., 2021, pp. 370–371). Following these approaches, the carbon turnover time τ of a system is related to the carbon stock S and the carbon flux F as follows:

$$\tau = \frac{S}{F}[yr]$$

The annual decomposition rate k (yr^{-1}) is related to the annual turnover time (yr) as $k = 1/\tau$. Hence, we estimate the initial fuel load as the ratio of the turnover and the decomposition rate. The initial load of litter $L_{t=0}$ is computed from total turnover from leaves and herbaceous biomass over the first one third of all time steps t, but excluding the time steps when a fire occurred:

$$L_{t=0} = \frac{\sum_{t=1}^{n} T_{leaf} + T_{herb}}{nyears \cdot k_{litter}}$$

Whereby *nyears* is the number of years over which the summation is applied (here 3.6 years, June 2014 – July 2017 in TUD-S4F-v0.3-S311) and k_{litter} is the annual litter decomposition rate. The initial loads of FWD and CWD are computed accordingly from the turnover of small branches and stems (including large branches):

$$FWD_{t=0} = \frac{\sum_{t=1}^{n} T_{branches,small}}{nyears \cdot k_{fwd}}$$

$$CWD_{t=0} = \frac{\sum_{t=1}^{n} T_{stem}}{nyears \cdot k_{cwd}}$$

Values for the decomposition rate for CWD k_{cwd} = 0.172 were taken for tropical deciduous forests from Harmon et al. (2020). For FWD, we assume that the decomposition rates are by a factor 2 higher than for CWD. For litter, the value of k_{litter} = 0.98 was taken for annual herbaceous litter from Zhao et al. (2025).

Please note that the amount of CWD in an ecosystem depends especially on the time since the last disturbance, the disturbance type, and the stand age (Harmon et al., 2020; Pedlar et al., 2002; Sturtevant et al., 1997). As such information is not directly available from Earth observation data over large areas, we intend to approximate the involved dynamics by using multi-annual time series of LAI and of f_{tree} to potentially capture past disturbance events, although we cannot account for the accumulation of woody debris that occurred before 2014 (start of the LAI dataset).

Starting from the initial estimates, we then compute the temporal dynamics of L, FWD, and CWD from the turnover *T*, the daily decomposition *D* and the daily dry matter burnt *DMB* of the corresponding fuel components as:

$$L_t = \left(L_{t-1} + T_{leaf,t} - DMB_{leaf,t} + T_{herb,t} - DMB_{leaf,t}\right) \cdot d_L - DMB_{litter,t}$$

$$FWD_t = \left(FWD_{t-1} + T_{branches,t} - DMB_{branches,t} + T_{SLWwood,t} - DMB_{SLWwood,t}\right) \cdot d_{fwd} - DMB_{fwd,t}$$

$$CWD_t = \left(CWD_{t-1} + T_{stem,t} - DMB_{stem,t}\right) \cdot d_{cwd} - DMB_{cwd,t}$$

Whereby the daily decomposition rate *d* for the three fuel components simply depends on the annual decomposition rate k:

$$d_{i,t} = e^{\frac{-k_i}{tsy}}$$

The parameter *tsy* defines the number of time steps per year (e.g., tsy = 36 in the case of 10 daily time steps) and distributes the annual decomposition rate k_i to an estimate for each time step. Note that this is a strong simplification, as the daily decomposition rate depends on daily variations in temperature and soil moisture conditions.

3.3.10 Dry matter burnt and fire behaviour

Dry matter burnt of the different fuel components depends on fuel loads, burnt area, combustion completeness and the behaviour of a fire, i.e. surface or crown fires. Thereby,

we assume that each observed fire (i.e. fractional burnt area fBA > 0) is in the first place a surface fire that always affects the surface fuel components (i.e., herbaceous vegetation, litter, FWD, and CWD). Combustion of tree leaves and SLW leaf and woody biomass in crown fires is only estimated if we observe during a fire also a reduction in woody LAI from the previous to the current time step. Combustion of woody biomass in tree stems and branches is only estimated if we estimate a reduction in tree height between the previous and current time steps.

Dry matter burnt is computed for live fuel components *i* (i.e. tree leaves, tree branches, tree stems, SLW leaves, and SLW wood) as:

$$DMB_{i,t} = CC_{i,t} \cdot fBA_t \cdot T_{i,t}$$

This implies that dry matter burnt of living biomass components is only applied to the turnover fluxes (i.e., reduction of biomass pools) that are estimated at the time of fire occurrence.

The dry matter burnt of herbaceous biomass and dead surface fuels (i.e. litter, FWD, and CWD) depends on the available surface fuel loads *FL* at a time step:

$$DMB_{i,t} = CC_{i,t} \cdot fBA_t \cdot FL_{i,t}, with\ FL_i \in \{L, FWD, CWD\}$$

Total dry matter burnt is the sum across the different fuel components:

$$DMB_{total} = DMB_{leaf} + DMB_{branches} + DMB_{stem} + DMB_{SLWwood} + DMB_{SLWleaf} + DMB_{herb} + DMB_L + DMB_{FWD} + DMB_{CWD}$$

Preliminary comparisons of the estimated dry matter burnt from TUD-S4F-v0.3-S311 show the TUD-S4F can represent regional distributions of observed DMB, although it seems to overestimate DMB from litter and FWD (Figure 14).

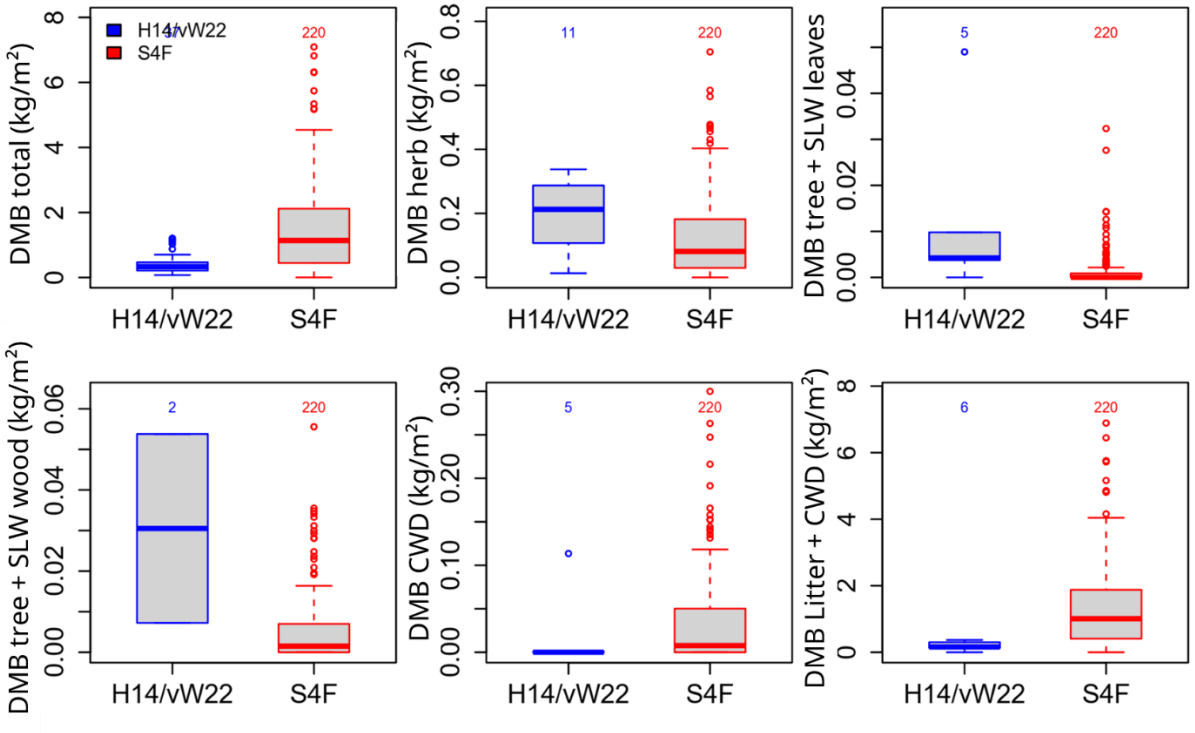

Figure 14: Boxplots of dry matter burnt (kg/m^2) for different fuel components as estimated with TUD-S4F-v0.3-S311 for 220 randomly sampled grid cells within a 200 km buffer of field observations from the vW22 dataset. The location of grid cells correspond to the one shown in Section 3.3.5.

3.3.11 Combustion, emission factors, and fire radiative energy

The composition of fire emissions and the produced heat and fire radiative energy depend on the dry matter burnt, the chemical composition of the combusted material (e.g., lignin, cellulose, volatiles), the fuel moisture content, and the type of combustion (i.e., flaming or smouldering combustion). In order to represent and quantify those factors in the TUD-S4F model, we adopted the chemical combustion model described and implemented in Chapter 2 of Rego et al. (2021) (Figure 15). The combustion model requires information about FMC, the chemical composition of the fuel (i.e., cellulose, lignin, and volatiles), the type of combustion, air temperature, and relative humidity. The combustion model then quantifies the modified combustion efficiency (MCE), the heat yield, and emission factors for H_2O, CO_2, CO, CH_4 and particulate matter (PM). In the following, we summarize the calculation of the emission factor for CO (EF_{CO}), but we also implemented the calculation for emission factors for CO_2, CH_4, H_2O, total PM, and $PM_{2.5}$. For details of the fuel combustion model, we refer to Chapter 2 of Rego et al. (2021).

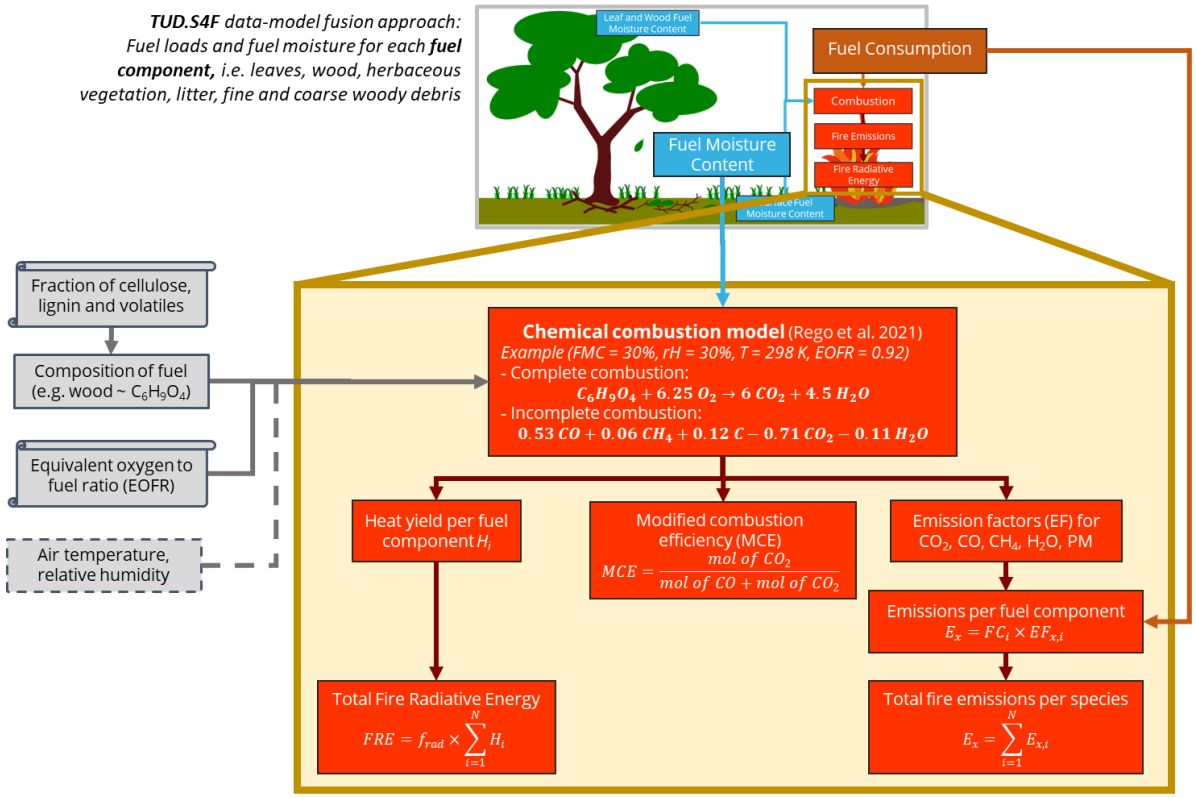

Figure 15: Overview about the calculation of fire emissions in the TUD-S4F approach.

The EF_{CO} (g/kg) for one fuel compound is calculated as follows:

$$EF_{CO,compound} = \frac{N_{CO,inc} \cdot (M_C + M_O)}{M_{compound} \cdot 1000}$$

Whereby M_C = 12 and M_O = 16 are the atomic weights of carbon and oxygen, respectively, and $M_{compound}$ is the molar mass of the fuel molecule (i.e., lignin, cellulose, or a volatile). $N_{CO,inc}$ is the number of CO molecules produced in incomplete combustion. Following the calculation in Rego et al. (2021, pp. 25–36), $N_{CO,inc}$ is computed as:

$$N_{CO,inc} = -0.75 \cdot \frac{N_O \cdot (EOFR - 1)}{1.41} \cdot 2$$

Whereby N_O is the number of oxygen atoms consumed in complete combustion and EOFR is the equivalent oxygen to fuel ratio. The EOFR is defined as the ratio of oxygen consumed during the actual combustion and the oxygen consumed in complete combustion (Rego et al., 2021). Fires with complete combustion have an EOFR = 1, flaming combustion has around EOFR = 0.96 and smouldering combustion has an EOFR < 0.9. We took the typical values provided by Rego et al. (2021) as a priori EOFR values for leaves, wood, herbaceous vegetation, litter, FWD, and CWD (Table 6).

In TUD-S4F, we assume that each fuel component consists of the three compounds cellulose, lignin, and volatiles (assuming that volatiles are monoterpenes) and un-burnable minerals, and then we compute the combustion and emission factors separately for each compound. An example for this is shown in Table 8.

Table 8: Example of the calculation of EF_{CO} for the three used fuel compounds and for three different types of combustion efficiency.

Fuel compound	Formula	Molar mass (g/mol)	EF_{CO} (g/kg) Smouldering fire (EOFR = 0.8)	EF_{CO} (g/kg) Flaming fire (EOFR = 0.93)	EF_{CO} (g/kg) Complete combustion (EOFR = 1)
Cellulose	$C_6H_{10}O_5$	162	221	77	0
Lignin	$C_{10}H_{12}O_3$	180	381	133	0
Volatiles (here: monoterpenes)	$C_{10}H_{16}$	136	613	215	0

The emission factor for the entire fuel component (e.g., leaf biomass) is then computed as the weighted average of emission factors for cellulose, lignin, and volatiles:

$$EF_{CO,fuel} = fce_{fuel} \cdot EF_{CO,cellulose} + fli_{fuel} \cdot EF_{CO,lignin} + fvo_{fuel} \cdot EF_{CO,volatiles}$$

whereby *fce*, *fli*, and *fvo* are the fractions of cellulose, lignin, and volatiles in a fuel component (Table 6). CO emissions from a fuel component are finally computed as the product of dry matter burnt and emission factor; for example, CO emissions from CWD are computed as:

$$E_{CO,CWD} = DMB_{CWD} \cdot EF_{CO,CWD}$$

Total CO emissions of the grid cell are the sum of emissions from all fuel components (i.e., leaf and herbaceous biomass, wood, FWD, CWD, litter).

The composition of each fuel by the fuel compounds cellulose, lignin, volatiles (e.g., monoterpenes), and minerals is an important control on the composition of the fire emissions. We here parametrise this composition by the parameters *fce*, *fli*, and *fvo.* Those fractions are known for some plants or fuel components but show a large variability (e.g. Rego et al., 2021; Waliszewska et al., 2021) and are not available for a large-scale spatial application in a satellite data-driven fire combustion model. Hence, we made a couple of assumptions and treated them in TUD-S4F as calibrate-able model parameters. Specifically, we assume that the fractional content of lignin (f_{li}) in leaves, live wood and herbaceous vegetation is:

$$f_{li} = 1 - f_{ce} - f_{vo} - 0.01$$

Which assumes a mineral content of 1%. We assume that the volatiles are monoterpenes ($C_{10}H_{16}$) because most other volatiles have higher carbon contents and hence would cause high emissions of CO and CH_4, which seemed implausible in a priori sensitivity experiments. We treat the fraction of volatiles as a model parameter that we calibrate against observed statistical distributions of emission factors provided by the V23 and A19 datasets.

Decaying plant material has a higher content of lignin than cellulose because cellulose decomposes faster. Hence, we assume that f_{ce} in litter, FWD, and CWD is 50% of f_{ce} for leaves and wood, respectively, and adjust the fraction of lignin accordingly.

As the production of nitrous oxides (NO_x) during fire is highly complex and not considered in the used combustion model, NO_x emissions are estimated with fixed emission factors for tropical forests and savannahs (Table 6).

Fire radiative energy (FRE) is estimated from the heat release *Q* of each fuel component *i:*

$$FRE = f_{rad} \cdot \sum_i^{\square} Q_i$$

Whereby f_{rad} is the fraction of heat released in the form of radiation. The heat release of each fuel component is computed from dry matter burnt and from the heat yield *q* (i.e., kJ of heat per mass of combusted fuel) of the combustion of the cellulose, lignin, and volatile compounds:

$$Q_i = f_{ce} \cdot DMB_i \cdot q_{ce} + f_{li} \cdot DMB_i \cdot q_{li} + f_{vo} \cdot DMB_i \cdot q_{vo}$$

The heat yield is computed according to Chapter 3 in Rego et al. (2021) as the difference between the heat content of the fuel, the heat required for pre-ignition of the fuel, and the heat content of emitted products from incomplete combustion (i.e., CO, CH_4, C):

$$q = q_{fuel} - q_{preign} - q_{products}$$

The computation of these heat components are based according to the combustion model described in Chapters 2 and 3 of Rego et al. (2021).

3.3.12 Local calibration and spatial prediction of model parameters

The modules for canopy height, tree biomass, herbaceous biomass and combustion completeness have been first calibrated as stand-alone modules as described before. Afterwards, the parameters of the fully coupled TUD-S4F model are calibrated against satellite datasets of AGB, canopy height, LFMC for several individual fire events for single grid cells. Therefore, we sampled 1000 grid cells from the entire study region. For each grid cell, we calibrated the parameters h1, h2, a0, $f0_{leaf}$, $f1_{leaf}$, $f2_{leaf}$, $f0_{herb}$, $f1_{herb}$ and $f2_{herb}$, by using the Genetic Optimisation Using Derivatives (GENOUD) algorithm (Mebane and Sekhon, 2011). The used cost function *J* directly compares spatial-temporal data pairs is based on the Kling-Gupta efficiency (Gupta et al., 2009) summed over all datasets *d*:

$$J_d = \sqrt{\left(\frac{\bar{s}+1}{\bar{o}+1} - 1\right)^2 + \left(\frac{\sigma_s+1}{\sigma_o+1}\right)^2 + (r-1)^2}$$

$$J = \sum_{d=1}^{N} J_d$$

Whereby *s* and *o* refer to the mean values and standard deviations of simulations (from TUD-S4F) and observations (from satellite datasets), respectively, and *r* is the Pearson correlation coefficient.

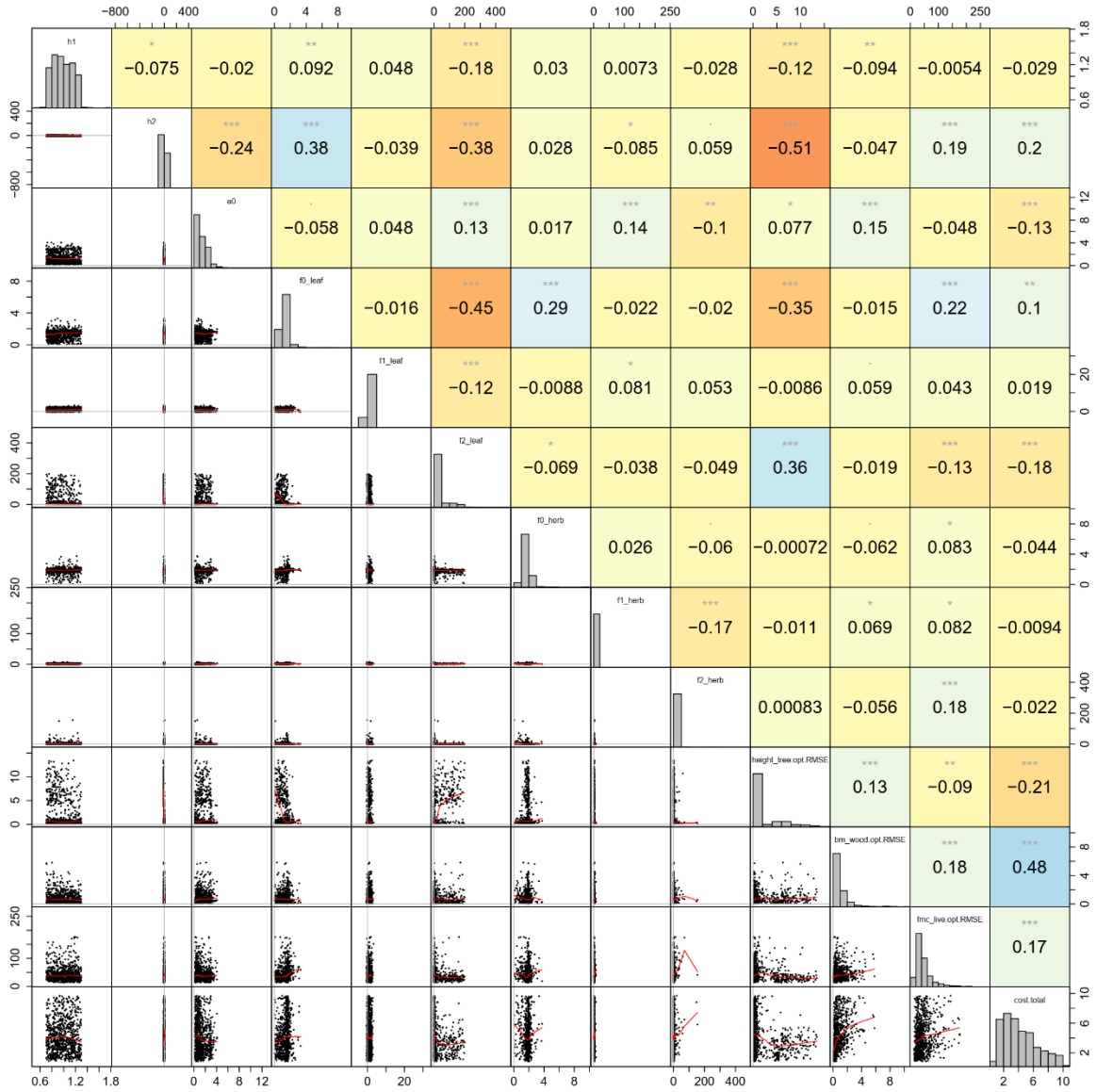

Figure 16: Scatterplot and correlation matrix of optimised model parameters and the corresponding RMSE for canopy height, above ground biomass and LFMC from 1000 sampled grid cells from TUD-S4F-v0.3-S311. Orange and blue colours indicate negative and positive correlations, respectively.

In the calibration of TUD-S4F-v0.3-S311 based on 1000 randomly sampled grid cells, the total cost is mostly correlated with the error for above-ground biomass and is mostly correlated to the parameters *h2* and *a0* (Figure 16). The local calibration results in high agreement between canopy height, AGB and LFMC from TUD-S4F and the respective reference datasets, although some grid cells show low performance (blue in Figure 17), mostly because the reference datasets appear to be inconsistent in such grid cells.

Finally, the locally calibrated parameters from the 1000 sampled grid cells are then used to estimate spatial fields of those parameters in order to apply the TUD-S4F model spatially. Therefore, we used the average and standard deviation of LAI (CGLS) and SWI (CGLS), the above-ground biomass (ESA CCI), and the fractional land cover of trees and herbaceous land cover (ESA Worldcover), and canopy height (from TUD-S4F predictions

using the GLM) of the reference year 2020 in a random forest regression model as predictors for the locally calibrated parameters a0, h1, h2, b0, $f0_{leaf}$, $f0_{herb}$, $f1_{leaf}$, $f1_{herb}$, $f2_{herb}$, and $f2_{leaf}$. This random forest model was then applied to all grid cells to predict the parameters and to apply the TUD-S4F model.

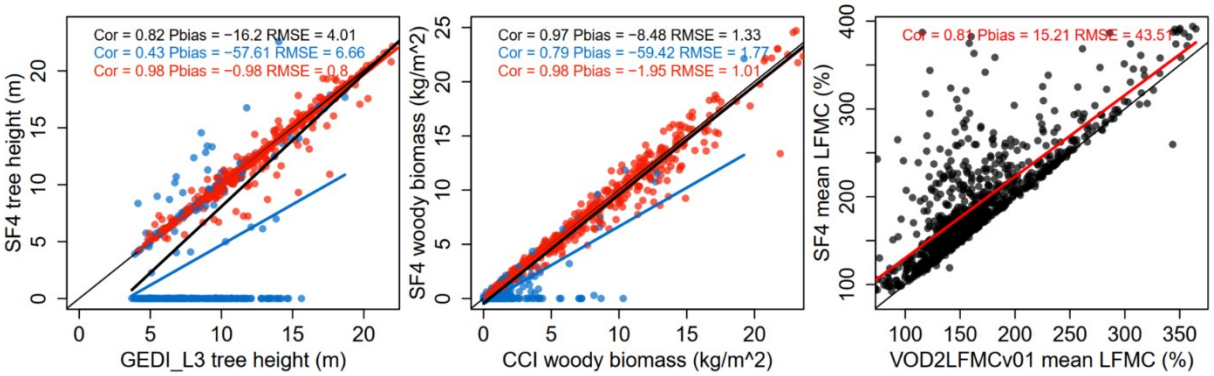

Figure 17: Comparison of estimated and reference canopy height, woody biomass, and mean LFMC for 1000 randomly sampled grid cells after calibration of each grid cell. Black are all grid cells and blue and red are grid cells for which the optimisation resulted in good and no acceptable performance, respectively.

3.4 TUD-S4F-v0.3-NRT near-real time setup

In order to apply TUD-S4F in near real time, all input datasets are required until the final date of model application. Therefore, we assume that the static input datasets are still representative and hence only require NRT values for burnt area and LAI. For burnt area, we use the NRT estimates from GFA-S4F. For LAI, we use the CGLS LAI time series until the last available RT6 estimate (i.e. with a time lag of 60 days) and apply a LAI forecast until the final date.

This method employs a harmonic reconstruction approach to fill gaps, smooth, and forecast the CGLS LAI time series, a key input to TUD-S4F. It assumes that LAI exhibits a quasi-periodic annual cycle influenced by insolation, temperature, and precipitation. The signal is represented as a repeating annual waveform, with higher harmonics capturing the asymmetry between growing and dormant seasons, alongside measurement noise. Periodicity is applied flexibly: harmonic timing remains fixed, while amplitudes and baseline vary smoothly to account for interannual variability.

The processing workflow consists of four stages: an admissibility gate that evaluates whether a pixel's observation record supports reconstruction; selection of the number of harmonics using nested F-tests, a partial R^2 threshold, and the Bayesian Information Criterion (BIC). Coefficients are estimated by weighted least squares based on per-observation uncertainty, and uncertainties are propagated analytically. Due to the model's continuity, gap filling does not require a separate interpolator. Observed values are retained, while missing steps are replaced with model predictions. The same fitted coefficients enable temporal resampling and forecasting. Uncertainty is further quantified using block bootstrapping for temporally correlated errors, Monte Carlo methods for nonlinear quantities, and gap-specific error characterisation.

This method has several limitations. It cannot anticipate abrupt disturbances, such as fires or storms, that occur within a gap or forecast horizon. Phase and amplitude are undefined or masked in regions where seasonality is weak. Additionally, the number of amplitude basis functions is constrained by record length to prevent spectral sideband collision.

3.5 TUD-S4F-v0.3-HR high resolution setup

TUD-S4F at high resolution is applied in a similar setup like at the coarse resolution but by using higher resolution input data at 0.0002° (approx. 20 m). While datasets for burnt area (FireCCIS211), land cover (Worldcover or CCI HRLC) and canopy height (Pauls et al., 2024) are already available at such resolutions, datasets for LAI, above-ground biomass and SWI are not. For SWI, we follow the same approach like in the coarse resolution setup and bilinearly interpolate the data to the high-resolution grid. However, LAI and biomass should represent local vegetation patterns in order to accurately estimate fuel loads. Therefore, we apply several approaches to estimate LAI and AGB at higher resolution. For LAI, we either apply a downscaling approach to estimate LAI at high resolution from the coarse resolution CGLS LAI data (Section 3.5.1) or we directly retrieve LAI from high resolution Sentinel-2 imagery (Section 3.5.2). For AGB, we downscale the ESA CCI AGB product to higher resolution (Section 3.5.3).

3.5.1 Downscaling of CGLS LAI (TUD-S4F-0.3-HR-S3)

LAI at coarse resolution (LAI_{CR}, 0.0025°) is the average of LAI high resolution (LAI_{HR}, 0.0002°) resolution weighted by the fractional cover of each land cover type:

$$LAI_{CR} = \frac{1}{N}\sum_{i=1}^{N}\sum_{type} LAI_{type,i} \cdot f_{type,i}$$

Whereby $LAI_{type,i}$ and $f_{type,i}$ are the LAI values and fractional coverage of tree, shrub, grass land, herbaceous wetland and crop (5 types) in each i-th HR grid cell within the CR grid cell. $f_{type,i}$ is derived from ESA WorldCover by calculating the percentage of 10 m pixels for each land cover type within the larger 20 m grid cell.

$LAI_{type,i}$ is calculated from the NDVI percentiles 10% and 90% of the WorldCover dataset by making use of the relationship between NDVI and LAI that can be represented by the Beer-Lambert law of light extinction:

$$LAI_{type,i} = \frac{\log\left(1 - \left(NDVIp10_i + P_{CR} \cdot \left(NDVIp90_i - NDVIp10_i\right)\right)\right)}{-k_{type}}$$

Whereby P_{CR} is the phenology status for a CR grid cell and ranges between 0 (no seasonal leaf cover) and 1 (full seasonal leaf cover). It is computed from scaling LAI_{CR} to [0,1] using the maximum and minimum values of the whole time series per pixel. $NDVIp10_i$ and $NDVIp90_i$ are the 10th and 90th percentiles of NDVI at HR resolution. They are computed from ESA WorldCover NDVI percentile using bilinear interpolation. k_{type} is a land-cover-specific light extinction coefficient. It is computed using the genetic optimisation calibration approach. The cost function for this optimisation is the root mean square error between estimated and observed CGLS LAI_{CR} time series.

The output LAI at HR is calculated by:

$$LAI_{HR} = \sum_{type} LAI_{type,i} \cdot f_{type,i}$$

For this ATBD, the algorithm was first tested in a 5° × 5° test site (11°S – 16°S, 23°E – 28°E) in southern Africa. The input data includes CGLS LAI from 2014-01-10 to 2021-10-20 (10 daily, 280 scenes in total), ESA WorldCover and NDVI percentiles (year 2021). 5000 randomly sampled CGLS LAI cells are used to estimate k_{type}. The optimised k_{type} for tree, shrub, grass, crop and herb is 0.35, 0.33, 0.28, 0.23 and 0.50, respectively. In comparison with CGLS LAI, the downscaled high resolution (HR S3) LAI presents much more spatial details and has a similar range of LAI value (Figure 18). To validate the HR S3 LAI, temporal median values of CGLS LAI and HR S3 LAI are firstly calculated. The median value of HR S3 LAI is then resampled to same resolution as CGLS LAI using average resampling and is compared with CGLS LAI (Figure 19). The resampled HR S3 LAI has a RMSE of 0.33.

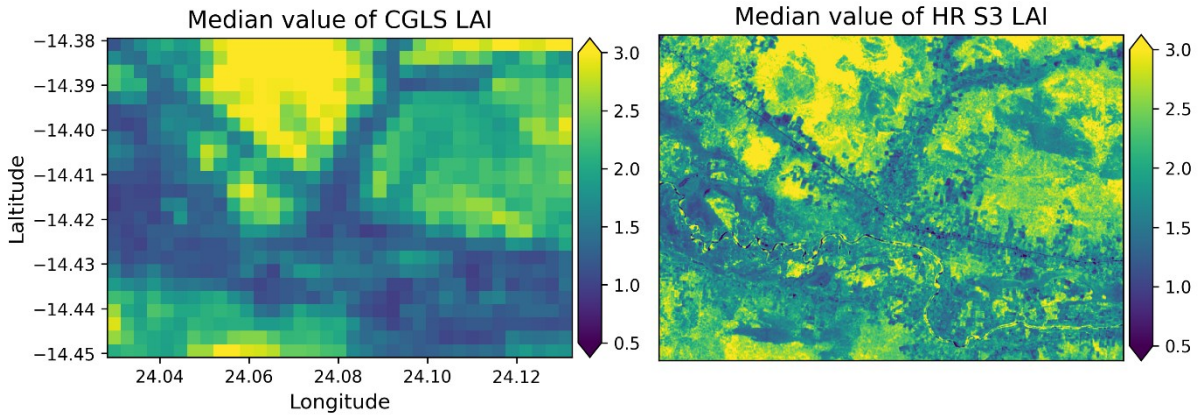

Figure 18: An example of CGLS LAI (left, 0.0025° resolution) and HR S3 LAI (right, 0.0002° resolution). The temporal median values of 280 scenes (from 2014-01-10 to 2021-10-20) are presented.

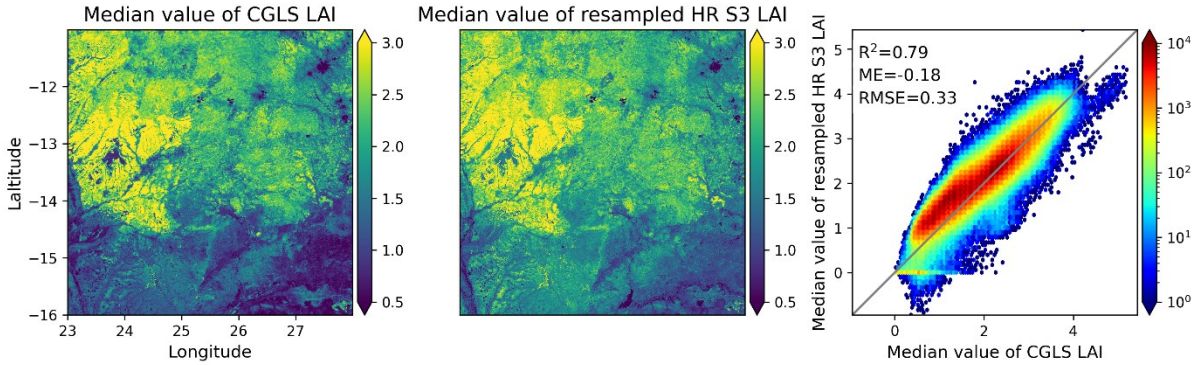

Figure 19: Comparison of the temporal median value (280 scenes from 2014-01-10 to 2021-10-20) of CGLS LAI and resampled HR S3 LAI at a 5° × 5° test site.

3.5.2 Retrieval of LAI from Sentinel-2 (TUD-S4F-0.3-HR-S2)

As an alternative LAI dataset, we retrieved LAI from Sentinel-2 imagery using the SNAP Biophysical Processor S2 (SNAP version 13) at one 0.5° × 0.5° test site (14°S – 14.5°S, 24°E

– 24.5°E). The input data is Sentinel-2A Level 2A Surface Reflectance product from 2016-04-16 to 2022-10-02 (93 scenes in total). The output LAI (HR S2 LAI) has 20 m resolution.

In comparison with CGLS LAI, S2 LAI has more spatial details but much lower LAI values than the CGLS S3 product (Figure 20). The same validation method for downscaled CGLS LAI is applied to validate S2 LAI. Although S2 is systematically underestimated, it shows strong linear correlation with CGLS LAI (Figure 21). A scaling factor (around 0.5) is needed to correct S2 LAI. For the same test site, HR S3 LAI has a better performance than HR S2 LAI (Figure 22).

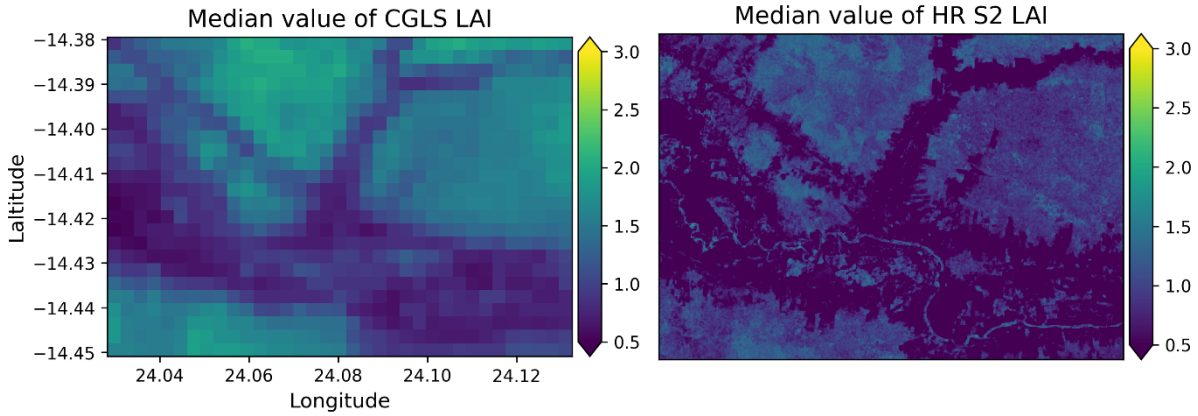

Figure 20: An example of CGLS LAI (left, 0.0025° resolution) and HR S2 LAI (right, 0.0002° resolution). The temporal median values of 93 scenes (from 2016-04-16 to 2022-10-02) are presented.

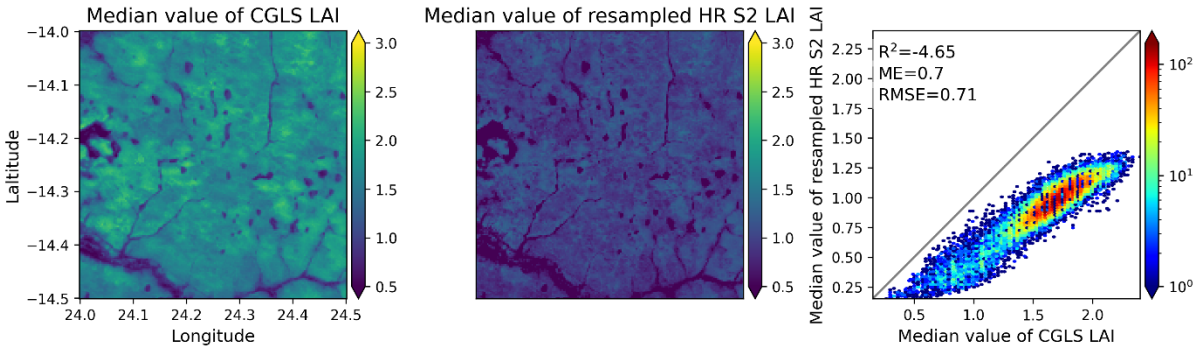

Figure 21: Comparison of the temporal median value (93 scenes from 2016-04-16 to 2022-10-02) of CGLS LAI and resampled HR S2 LAI at a 0.5° × 0.5° test site.

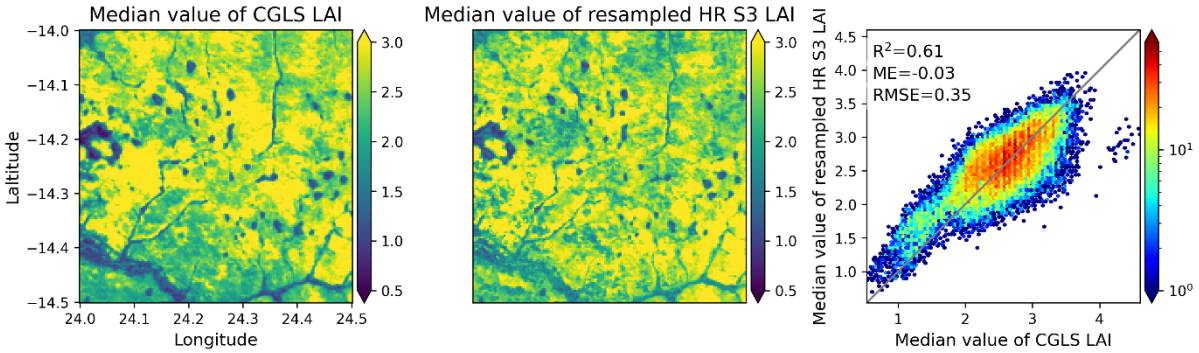

Figure 22: Comparison of the temporal median value (280 scenes from 2014-01-10 to 2021-10-20) of CGLS LAI and resampled HR S3 LAI at the same test site as HR S2 LAI.

3.6 Downscaling of ESA CCI above-ground biomass

To downscale ESA CCI AGB to HR, we applied a multiple linear regression trained using Sentinel-1 backscatter, ALOS backscatter and canopy height as predictor variables and ESA CCI AGB as response variable:

$$\log_{10} AGB_{100m} = a \cdot \gamma_{S1,VH,100m} + b \cdot \gamma_{ALOS2,HV,100m} + c \cdot \log_{10} H_{100m} + d$$

Whereby $\gamma_{S1,VH,100m}$ is the resampled ESA WorldCover annual composite of Sentinel-1 VH Gamma-0 backscatter in decibel scale. $\gamma_{ALOS,HV,100m}$ is the resampled global mosaic of ALOS-2 HV Gamma-0 backscatter in decibel scale. Both Sentinel-1 and ALOS2 backscatter have been geometrically and radiometrically calibrated. H_{100m} is the resampled global canopy height from Pauls et al. (2024). Median resampling is used to resample all three predictor variables from their original resolution to 100 m resolution. Ordinary Least Squares regression is used to estimate regression coefficients a, b, c and d.

The regression function is then applied to the resampled Sentinel-1 backscatter, ALOS2 backscatter and canopy height to estimate AGB at HR (HR AGB):

$$\log_{10} AGB_{HR} = a \cdot \gamma_{S1,VH,HR} + b \cdot \gamma_{ALOS2,HV,HR} + c \cdot \log_{10} H_{HR} + d$$

For the ATBD, the model is tested at one 1° × 1° tile (14°S – 15°S, 24°E – 25°E). The regression coefficients for this test site are a = 0.177, b = 0.068, c = 0.66, and d = 4.609. Spatial resolution is improved in the downscaled AGB (Figure 23). HR AGB is resampled to 100 m resolution using average resampling and is compared with CCI AGB, which results in a RMSE of 17.05 Mg/ha (Figure 24).

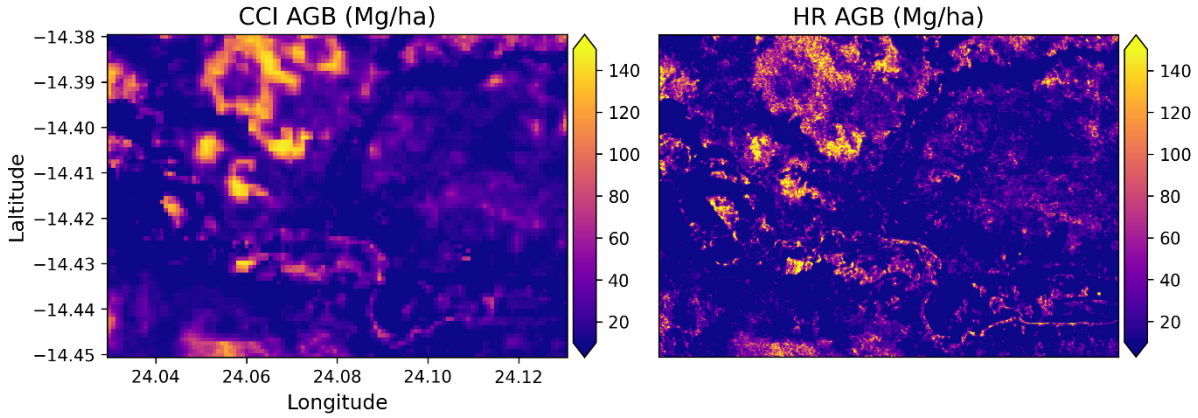

Figure 23: An example of CCI AGB (left, 100 m resolution) and downscaled AGB (right, 20 m resolution).

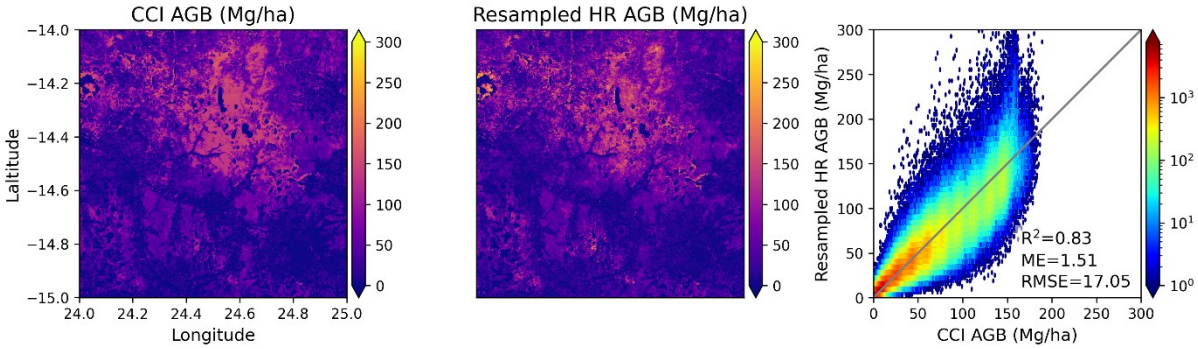

Figure 24: Comparison of CCI AGB and resampled HR AGB at a 1° × 1° test site.

4 KNMI-S5p validation approach

4.1 Overview

4.1.1 Summary of updates

As outlined in the introduction and the scoping of this ATBDv4, the KNMI-S5p validation approach is based on the use of Sentinel-5p satellite observations from the TROPOspheric Monitoring Instrument (TROPOMI) of atmospheric composition and air pollution.

The previous Sense4Fire ATBDs (ATBD v2.1 and ATBD v3) primarily focussed on describing possible approaches and parameters that could be used for evaluating and validation the wildfire emission databases developed and used in Sense4Fire.

The Sense4Fire results have since then narrowed down the strategy how to evaluate wildfire emissions using IFS-COMPO atmospheric composition modelling and intercomparison with Sentinel-5p measurements. The context and strategy that has emerged will be described in more detail later in this section.

ATBDv3 also briefly explained the strategy to test the IFS-COMPO modelling suite for a near-real-time application (NRT) which in the context of wildfire emissions means being able to evaluate emissions per month and within approximately a maximum delay of one month. This has been proven achievable, also discussed in de Laat et al. (de Laat et al., 2026).

4.1.2 Sentinel-5p/TROPOMI and wildfire emissions

The Sentinel-5p satellite was launched in October 2017 and has been providing data operationally from the summer of 2018 onwards. The TROPOMI instrument on sentinel-5p was developed based on the legacy of satellite instruments such as SCIAMACHY and OMI who demonstrated that air quality and air pollution could be monitored using satellite observations of earth reflected solar radiation at ultraviolet and visible wavelengths with high spectral resolution (also known as hyperspectral).

TROPOMI was therefore specifically designed to improve such air quality measurements to allow for daily mapping of emission plumes at the so-called city scale, which in practice has been in the region of 7x7 km to 5.5x3.5 km depending on the type of air pollution.

The TROPOMI instrument characteristics therefore opened a whole new landscape for monitoring wildfire emission plumes on a daily basis. Prior to TROPOMI, using satellite observations assessing wildfire emission databases was limited. Satellites like SCIAMACHY, OMI, but also GOME-2, OMPS, MOPITT, IASI and AIRS were characterised by either coarse spatial resolution, complex vertical sensitivity or a need to average multiple observations over time to meet the required measurement precision.

That also complicated the translation of wildfire emissions estimates to satellite observations of air pollution. Some effort had been put into using numerical model simulations of atmospheric composition as an interface, but required averaging limited the usefulness of these model simulations.

In addition, methods for more formally inverting satellite observations of air pollution back to fire emissions have historically always been similarly hampered by the lower quality of satellite observations of air pollution of satellite instruments prior to TROPOMI. Such formal wildfire emission inversions have remained underdeveloped.

Furthermore, possibilities to independently verify wildfire emission estimates with *in situ* observations – either from permanent surface stations or field campaigns – are extremely limited. The most important wildfire regions worldwide generally are remote making it difficult to reach them while infrastructure for setting up a monitoring network is lacking.

Hence why the advance of earth observation technology – and TROPOMI in particular – has been so important for wildfire emission monitoring, evaluation and validation. TROPOMI for the first time allows for spatiotemporally detailed a continuous monitoring of wildfire emissions.

4.1.3 Wildfire emissions evaluation with IFS-COMPO and TROPOMI

Although in principle TROPOMI measurements allow for formal wildfire emission inversions, the approach chosen here is to use an atmospheric chemistry-transport model as an interface between the wildfire emission estimates on the one hand and the satellite measurements of air pollution on the other hand. Wildfire emissions are fed into the model whose results are collocated with the satellite observations of air pollution – taking satellite instrument operators into account (Douros et al., 2023). This provides an extensive dataset that can be used to explore how well the model reproduces the observations. Model simulations can be tweaked with other fire emission databases or modifications of those emission databases to bring model results and observations in closer agreement.

Furthermore, for the short live trace gas nitrogen dioxide (NO_2) the model simulations also allow for applying a post-hoc NO_2 emission correction based on the modelled local sensitivity of tropospheric NO_2 columns to local NO_2 emissions (Castellanos et al., 2014; Forkel et al., 2025; Lamsal et al., 2011). Assuming linearity between both, that sensitivity – denoted as β-factor – can be multiplied with the differences between model simulated and observed tropospheric NO_2 to a local multiplication factor which to first order brings model simulations and observations of tropospheric NO_2 in better agreement.

In the first original phase of the Sense4Fire project the TROPOMI measurements were explored and evaluated on how to use them for wildfire emission monitoring. Several air pollutants were explored for their usefulness: aside from NO_2 also Carbon Monoxide (CO), aerosols via the absorbing aerosol index (AAI) and formaldehyde (HCHO) were evaluated.

Results indicated that CO was a useful addition to NO_2. Both traces are emitted wildfires but have different sensitivities to fire temperatures. Whereas NO_2 emissions increase with increasing fire temperature – and thus reflect combustion completeness, CO emissions increase with decreasing fire temperatures. CO thus reflects incomplete combustion.

It was also shown that the AAI and CO are spatiotemporally well correlated. Since the AAI is only a qualitative index it has limited added value on top of CO measurements. HCHO turned out to be less useful due to the lower quality of TROPOMI measurements, requiring the need for averaging over space and/or time. The TROPOMI HCHO measurements thus were hampered by some of the same limitations as satellite instruments prior to TROPOMI. Hence why for the Sense4Fire project the primary focus has been on CO and NO_2.

4.1.4 Extending the evaluation of Sense4Fire wildfire emissions

The first results of the exploration of the use of TROPOMI measurements for wildfire emission monitoring have focused on short time periods of several months in single year over a few small regions worldwide with a focus Amazonia (Forkel et al., 2025; de Laat et al., 2026; other regions were sub-equatorial Africa, Mediterranean Europe and in Siberia).

Although limited in their spatiotemporal coverage, results nevertheless clearly indicated that CO emissions were underestimated over the primary wildfire emissions regions Amazonia and (sub)Equatorial Africa. Depending on the wildfire emission dataset used, wildfire CO emissions over Amazonia were underestimated by a factor of 1.5-3. The main cause of this large discrepancy appears to be small scale fires and limitations to their detection with currently satellites.

NO_2 was generally better represented albeit that for very large fires their emissions were significantly overestimated by up to a factor of 10. Although for the total NO_2 emissions the contribution of this “large-fire-bias” was limited, this result nevertheless signals emissions from high temperature wildfires is not well understood, warranting additional research.

This next phase of the Sense4Fire project will rely on the same approach for evaluating and validating wildfire emissions. Wildfire emissions estimates are used as a boundary condition for atmospheric composition modelling. Model results are then collocated with TROPOMI measurements while applying the appropriate satellite operators for comparison.

However, whereas until now the modelling time periods have been small, only a couple of months, for this project phase (Contract Change #2; CCN02) the model simulations will be performed for the multi-year period 2019-2025 covering mostly the entire TROPOMI dataset. In addition, the primary focus over this period will be (sub)equatorial Africa rather than Amazonia, although analyses for Amazia will be made.

Furthermore, it will be at least briefly explored whether HCHO could be more useful for the multi-year model simulation period. As outlined before, TROPOMI HCHO is not as accurate as CO and NO_2. However, monthly averages of TROPOMI HCHO are more than sufficiently accurate to use for evaluating the model simulations (Anderson et al., 2023; Fu et al., 2025; Sfendla et al., 2026; Shen et al., 2026)

Finally, although not the scope of the Sense4Fire CCN02, the possibility of developing a β-approach for CO will be explored in the EU HORIZON project CoFlame (Copernicus Fire and Land Monitoring of Emissions) that will commence during the second half of 2026 and in which KNMI participates using the model-as-interface approach as developed within S4F. Because the basic idea that the β-approach could be modified for CO also emerged during the S4F project, the conceptual outline of this idea will be documented here in the appendix (A1), thereby allowing for it to be used in future documents and projects if indeed achievable.

4.2 Input data

The evaluation of the wildfire emission databases developed within S4F (see this ATBD v4 sections 2 and 3 and Table KNMI-1 below) will continue to use the same approach as applied before. The IFS-COMPO model will be run using one of the fire emission databases as input. Model output will be collocated with TROPOMI measurements while TROPOM instrument operators will be applied. The most important operators are the averaging kernel, using relatively cloud free TROPOMI measurements, *i.e.* measurements with a cloud radiance fraction < 0.5, and integrating/averaging TROPOMI measurements within an IFS-COMPO model grid of 0.5°x0.5° provided that at least 50% of the model grid is covered by cloud free TROPOMI measurements. Note that a TROPOMI pixel is 7x7 km for CO and 3.5x5.5 km for NO_2, so effectively approximately 60 (CO) to 160 (NO_2) TROPOMI pixels can fit within the 0.5°x0.5° model grid.

The IFS-COMPO model setup and model simulations that will be performed are summarised in table X1. There will be three model setups covering the entire 2019-2025 period: GFAS, GFA-SAF-v0.3 and TUD-S4F-S311. For the last two setups the emissions are provided for Southern Hemisphere Africa while applying GFAS emissions elsewhere. For the other emission databases – which will have been scheduled to be delivered later during the project– only a one-year simulation will be performed. The year itself will be decided later. If emissions for the Sahel region are also provided, then these will be incorporated in the model simulation.

The TROPOMI data – CO, NO_2, possibly HCHO - will be used for the period 2019-2025. Note that currently TROPOMI data over that period consists of two retrieval approaches. For the period up until 17 July 2022 the reprocessing (RPRO) algorithm data will be used. The RPRO algorithms consist of the most complete retrieval algorithms. For the period after 17 July 2022 – for which no reprocessing has been performed yet – TROPOMI data from the OFFL algorithms will be used. OFFL algorithms contain a few simplifications that allow faster run time. Note that there are also Near-Real-Time algorithms containing even more simplifications that meet the requirement of providing TROPOMI data within 3 hours after observation. The OFFL data generally becomes available within a few days after the observations.

Table 9: Planned IFS-COMPO model simulations with different fire emission databases and the time period for which the IFS-COMPO model simulations will be performed.

Emission database	Region	Time period and comment
GFAS	Global	2019-2025
TUD-SAF-v0.3-S311	Africa-1 (S-Africa) + GFAS elsewhere	2019-2025
TUD-S4F-v0.3-S211	Africa-1 (S-Africa) Africa-1 (Sahel) + GFAS elsewhere	2024
TUD-S4F-v0.3-NRT	S-Africa + Sahel region is optional + GFAS elsewhere	one year (TBD[#]) Model simulation is optional if GFA-S4F-S311 differs significantly from TUD-S4F-v0.3-NRT
TUD-S4F-v0.3-HR-S3	S-Africa + Sahel region is optional + GFAS elsewhere	one year (TBD[#])
GFA-S4F-v0.4	S-Africa + Sahel region is optional + GFAS elsewhere	one year (TBD[#])
β-method	*see notes*	*one year (TBD[#])* *Simulation with a 20% emissions increase. Decision on which emission database to use will be decided later*

[#] *at the time of writing it has yet to be decided which year will be used for performing these IFS-COMPO model simulations. The criteria based on which the year will be selected have yet to be discussed within the Sense4Fire consortium.*

4.3 IFS model setup

The IFS-COMPO global atmospheric composition modelling system (IFS-COMPO) is used to simulate the atmospheric conditions during the periods. As IFS-COMPO is a global model it automatically covers any region under consideration. The model results serve to bridge the gap between bottom-up emission estimates (Table 9) and Sentinel-5p observations. The model results can be compared with Sentinel-5p observations after applying satellite observation operators.

The IFS-COMPO model currently used is CY50R1 tropospheric chemistry and aerosol modules. It is driven by ECMWF operational meteorology, and using emissions (including anthropogenic and biogenic) as prepared for the operational model configuration. It is setup to run at an approximately 40 km spatial resolution (T511 reduced Gaussian grid), with 137 vertical model levels (approximately 60 layers in the lowest 10 km of the

atmosphere). Model output is then generated at 0.5° resolution or approximately 50 km spatial resolution. This horizontal model resolution is much less refined than the Sentinel 1-2-3 data (sub-km scale), emission estimates (km scale and finer) and the Sentinel-5p spatial resolution (3.5x5.5 km at best). This impact the model-observation comparison.

Note that in the Sense4Fire project potential biases due atmospheric processes with spatial scales smaller than the model grid have been explore. For example, localised sub-model grid scale non-linear emission plume chemistry may result in changes in atmospheric composition that differ from direct injection of fire emissions at the native model resolution, particularly for short-lived trace gases. This is a well-known phenomenon from for example transport (aviation, shipping, road traffic) but also for stationary sources such as power plants. Such impacts traditionally have been mitigated in modelling by applying resolution-dependent correction factors to emission values. IFS-COMPO model simulations using parameterised sub-grid scale effects that to first order correct for such processes indicated that for fires in the Amazon region sub-grid processes had negligible effects on modelled CO and NO_2values.

To compare Sentinel-5p observations with IFS/CAMS output we consider all relevant aspects that are required when matching observation data to model data. Only observations with quality assurance threshold above 0.75 are used, as recommended by the NO_2 product user manual. This concerns observations with cloud radiance fraction of less than 0.5 and excludes retrievals with ground pixels covered with snow/ice, as well as problematic retrievals. Note that a cloud radiance fraction is not the same as cloud fraction but depends on the type of cloud (optically thick or thin) and the surface reflectance (light or dark). To illustrate: for a cloud radiance fraction of 0.5 approximately translates to a cloud fraction of 50% for optically thick clouds like cumulus and a relatively dark surface like vegetation or the oceans.

The model fields are interpolated in time to match with local overpass time of TROPOMI. The collocated model-observation pairs are gridded on a common 0.5x0.5-degree output field (or different resolution, which is configuration setting), and only written to NetCDF if a threshold coverage of 50% of the grid cell is reached.

The averaging is done by an area-weighted approach, hence considering the area of the TROPOMI-pixel that falls within the model grid box (Douros et al., 2023).

4.4 Preliminary 2019-2023 GFAS IFS-COMPO results

The IFS-COMPO model benchmark simulations using GFAS wildfire emissions have already been performed for the period 2019-2023. Here we show some first preliminary results.

Figure 25 shows the time series of the daily total CO mass over Amazonia for the region 80°W-40°W and 20°S-10°N (see Figure 26) for both TROPOMI observations and the corresponding IFS model simulation using GFAS wildfire emissions.

The results show that IFS does a very credible job in reproducing CO variability on a daily basis as well as the intra-seasonal variability and the seasonal cycle. There clearly is a bias

during the peak Amazon fire season from August to October with a peak during September. This is consistent with Forkel et al. (2025) and de Laat et al. (2026), as expected.

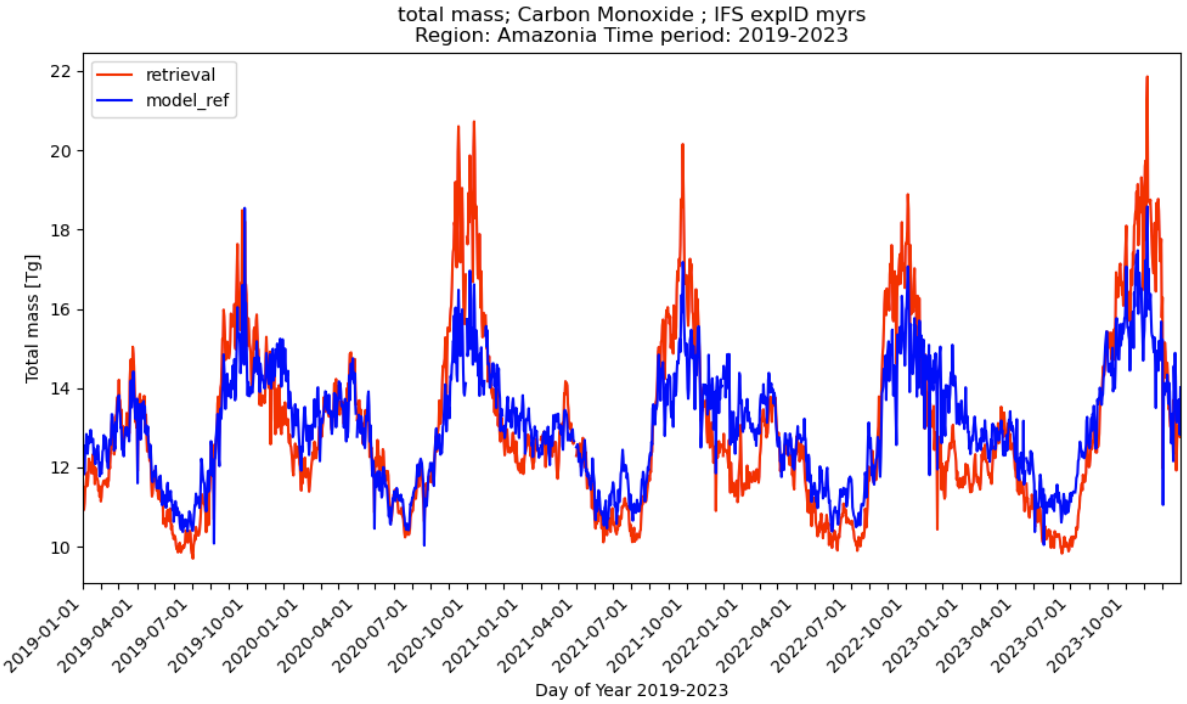

Figure 25: Time series of the daily total carbon monoxide mass over Amazonia for the region 80°W-40°W and 20°S-10°N (displayed in figure X02) for the period 2019-2023. The blue line denotes the total CO mass of IFS-COMPO GFAS simulation, collocated with TROPOMI CO total column observations. The red line displays the total CO mass from the TROPOMI CO total column observations themselves. As there can be model grids

without data due to cloudiness, hence the area total daily CO mass values are weighted with the ratio of model grid cells with data and the total number of model grid cells.

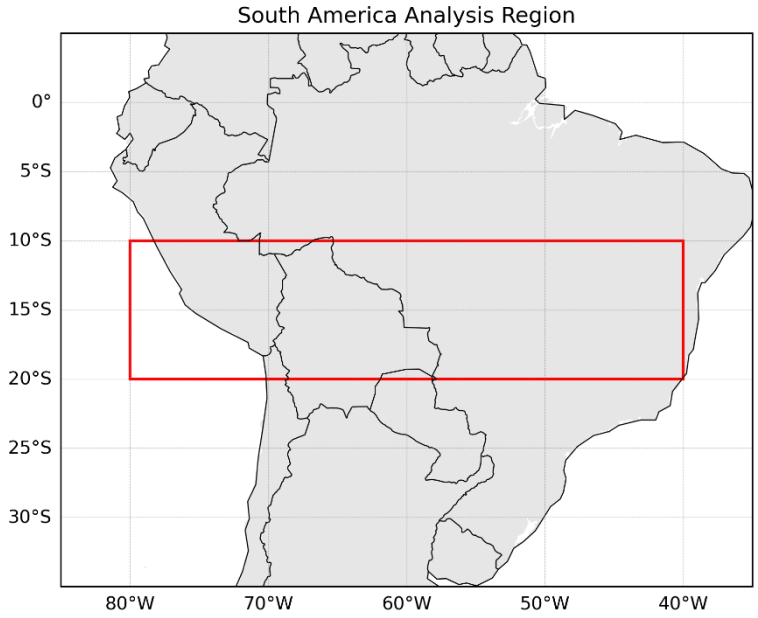

Figure 26: Box region over Amazonia (80°W-40°W and 20°S-10°N) for which the time series in Figure 25 are calculated.

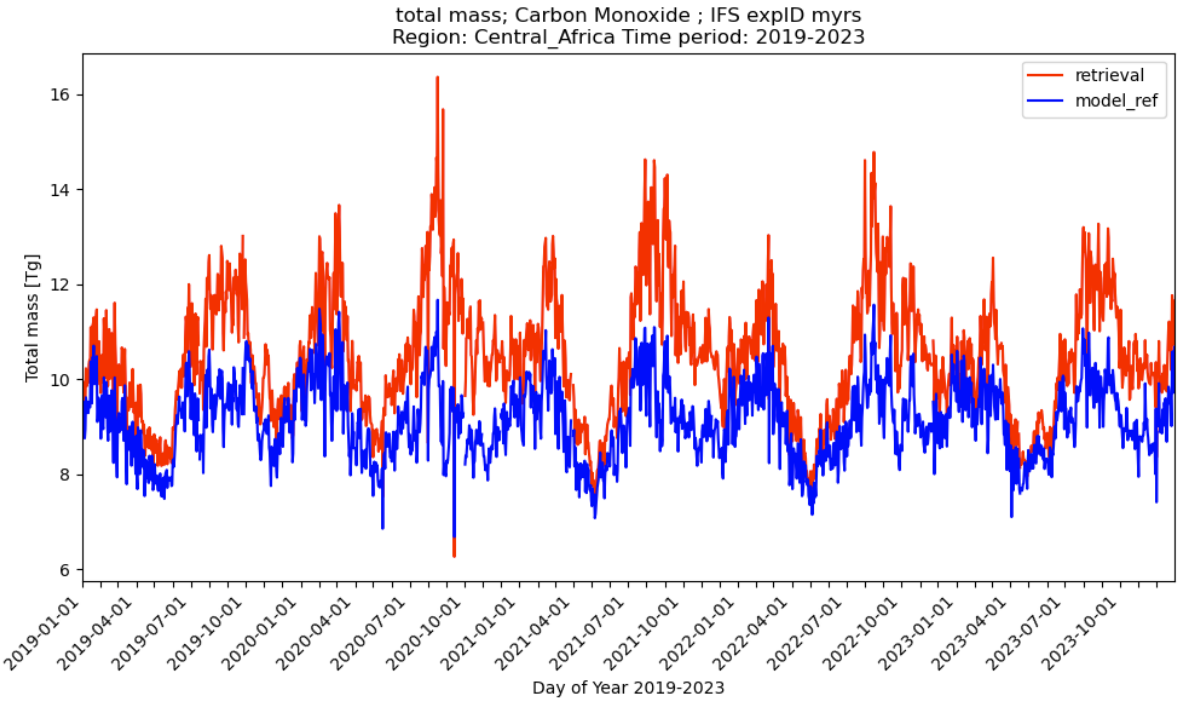

Figure 27: As Figure 26 but for the central Africa region in Figure 28.

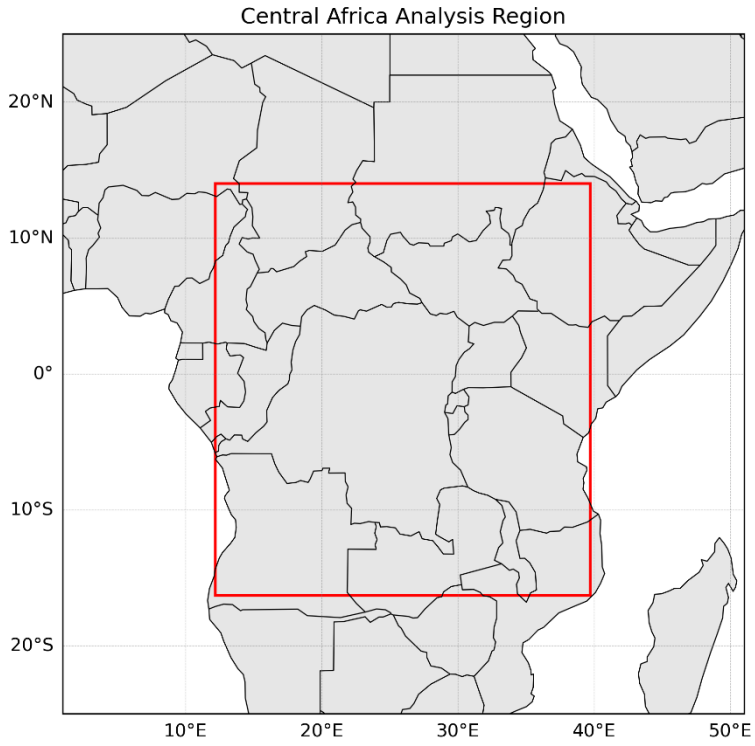

Figure 28: Box region over central Africa (12.2°E-39.7° E and 16.3°S-14.0°N) for which the time series in Figure 27 are calculated.

For the Amazon region displayed in Figure 26 the seasonal cycle in total CO mass has a particular evolution.

During January-April the CO total mass is enhanced, related to wildfires to the north that peak during these months (Velásquez-García et al., 2026). Some of those emissions are advected into the domain displayed here. To the south of the displayed domain there are no CO wildfire emissions during January-April.

During April-June wildfire emissions are limited throughout entire South America, causing the CO total mass to decrease to its seasonal minimum.

During July wildfire emissions rapidly increase towards their seasonal maximum in August-September in this region but also south of the region. The CO total mass correspondingly increases rapidly with a delay due to the CO lifetime and time needed for CO to accumulate so also in October the CO total mass can still be high.

During November-December the CO total mass decreases again as wildfire activity rapidly ceases during October.

On a monthly basis the Amazonian fire season biases in observed CO total columns (not shown) typically are 50-100% larger than modelled with IFS-COMPO but only for the main fire months August, September and October. There is also a significant interannual variability in the observation-model biases.

Figure 27 similarly show the total CO mass over the central African region 12.2°E-39.7° E and 16.3°S-14.0°N. This region also has a particular seasonal evolution of the total CO mass with essentially two active seasons and two quiescent periods (Jiang et al., 2020;

Petetin et al., 2018). This strong seasonality in wildfire activity is driven by rainfall patterns, with opposite fire seasons north and south of the equator. Northern Central Africa burns mainly in the dry months of November–February, while southern Central Africa burns mostly in June–September, reflecting the shift of the Inter-Tropical Convergence Zone (ITCZ).

Due to a time-lag delay of atmospheric CO in response to fire emissions this pattern shifts in time. The result for the entire box region in Figure 28 is that the CO total mass is enhanced from January to March (fires north of the equator), suppressed during April-June (limited to no fires), then is enhanced again from July to October (fires south of the equator) and is suppressed again during November-December (limited to no fires).

The IFS-COMPO model simulations follow this pattern but throughout the two main fire seasons the CO total mass is clearly biased low in GFAS. The two shorter low fire activity periods are well represented by IFS-COMPO.

On a monthly basis the central African fire season biases in observed CO total columns (not shown) typically are a factor of 2-3 larger than modelled with IFS-COMPO for the wildfires south of the equator but for an occasional month the observations can be a factor 4 larger. North of the equator observed CO total columns are typically 40-60% larger on a monthly basis. These biases appear to be more consistent between the years compared to the Amazon region.

5 Appendix: KNMI-S5p β-method

5.1 Introduction

This section introduces a proposal for modifying the β-method designed to address the “locality” criterion when applied to long-lived trace gases such as carbon monoxide. This locality criterion reflects the implicit assumption that variations in the trace-gas burden within a given model grid cell arise primarily from variations in emissions within that same cell, with only limited – thus negligible - influence from upwind regions. For long-lived species, this assumption is violated because transport will introduce substantial non-local contributions. However, in certain cases it may be possible to mitigate this limitation by explicitly estimating a “pseudo-influx” term that represents the contribution of neighbouring grid cells without local emissions. CO variations in nearby grid cells that have no local emissions should reflect non-local effects on CO. Such an approach thus would account for non-local transport effects within the β-framework.

5.2 Inverse modelling of trace gas emissions

Inverse modelling of atmospheric trace gas emissions often apply 4DVar or Ensemble Kalman Filter assimilation systems. A numerical model simulation of atmospheric composition is used to implicitly or explicitly determine the complex relationship between emissions of the trace gas and measurements of that trace gas. Assuming that the

numerical model simulation takes all relevant process into account, this then allows to determine what emissions should have been to explain the observations.

Although these approaches are powerful and mathematically sound, they face challenges related to high computational demands, sensitivity to model imperfections, and difficulties representing the true complexity of atmospheric processes. These approaches can struggle with nonlinear chemistry, sparse or indirect observations, and uncertainties in transport, all of which can distort the inferred emission patterns. They also rely on assumptions about error statistics and system behaviour that may not hold in real atmospheric conditions, leading to solutions that are sensitive to tuning choices or prone to compensating for model biases rather than revealing true emission signals.

5.3 Alternative approach: β-method short lived trace gases NO_2/NO_x

An alternative approach that has been developed for short-lived nitrogen dioxide (NO_2), tropospheric NO_2 columns and NO_x emissions, is the so-called β-method. In that approach an atmospheric chemistry model is run with baseline emissions and with perturbed emissions (typically +20%) to estimate the local sensitivity of the modelled tropospheric column to changes in emissions. It offers a set of practical advantages because it avoids the heavy machinery of full inverse-modelling systems and instead uses a simple, observation-driven sensitivity relationship between tropospheric NO_2 columns and NO_x emissions. It reduces the uncertainty by aggregating many fires and avoids the need for adjoints or ensembles. It is computationally lighter, easy to implement, and robust in regions where localised single source types dominate emissions – like wildfires. It thus works well for wildfire NO_x emissions as the short lifetime of atmospheric NO_2 ensures that the sensitivity relationship between NO_2 columns and NO_x emissions is primarily a local relationship that holds within the bounds of a model grid cell. This is the so-called “locality” criterion.

5.4 Limitations of the β-method for long lived trace gases and possible modifications

For long-lived species such as carbon monoxide (CO), whose atmospheric lifetime ranges from one to several months, the assumption of locality breaks down. CO at any location is also strongly influenced by long-range and even intercontinental transport and regional background concentrations. This means that local CO levels are determined by local emissions but also other non-local processes. Hence why application of the β-method has been limited to satellite measurements of NO_2 and NO_x emissions.

However, it might be possible to modify the β-method such that it may be applicable for CO emissions from wildfires. TROPOMI observations of wildfire CO clearly reveal emission plumes extending away from wildfires, which morphologically are to first order rather similar to NO_2 emission plumes extending away from wildfires. Although the evaluation of the wildfire emission plumes of both differ - NO_2 dilutes and is photochemically

destroyed whereas CO is only diluted – the similarities suggest that to first order also for CO there is a local relationship between the wildfire emissions and the measured CO total column. CO enhancements due to fires are clearly identifiable above the local background CO level. The complexity comes from accounting for whatever determines that local background CO level.

Although as described above the locality assumption needed for applying the β-method breaks down for CO, it is a trace gas with simple atmospheric chemistry whose concentrations are determined either by direct emissions, atmospheric reaction with OH to form CO_2, long range transport and oxidation of methane and non-methane hydrocarbons. Most of these processes can be (reasonably) well simulated with atmospheric chemistry modelling.

5.5 The standard β-Method

The classical β-method relies on establishing a local sensitivity the modelled tropospheric column C_x to local emissions E_x. The term "local" in a modelling context refers to the model grid: how does tropospheric column C_x in a model grid cell change due to emission E_x in the same model grid cell. In practice this is established by performing two model simulations:

- Baseline emissions: E_x^0
- Perturbed emissions: $E_x^1 = \alpha \cdot E_x^0$, typically with *α=1.2*

The local sensitivity factor β is then approximated as:

$$\beta = \frac{(C_x^1 - C_x^0)}{(E_x^1 - E_x^0)}$$

This formulation assumes:

- Linear chemistry
- Locality of response
- Short atmospheric lifetime

This works well under the assumption that all atmospheric chemistry processes are linear and that the local amount of a trace gas column value only depends on the local emissions. Typically, short lived trace gases like NO_2 meet this criterion.

5.6 Limitations of the β-method for CO

For CO, several assumptions underlying the β-method are violated:

- CO has a long lifetime (1–3 months), allowing long-range transport
- CO chemistry interacts with OH, introducing nonlinearities
- Columns at a given location reflect emissions from a wide region

- Columns at a given location reflect production from atmospheric oxidation processes

Thus, the local model column can be expressed as:

$$C_x = C_x(\mathrm{loc}) + C_x(\mathrm{infl})$$

where:

- $C_x(\mathrm{loc})$ is the contribution from local emissions
- $C_x(\mathrm{infl})$ is the non-local contribution from transported CO of atmospheric oxidation processes ("pseudo influx")

The local CO tropospheric columns thus not only depend on local emissions but also on emissions - or other processes leading to CO production like atmospheric oxidation and fire emissions - from a (large) surrounding area that flow into a model grid cell, the so-called "pseudo influx").

5.7 Proposed Modification: Incorporating a Pseudo-Influx Term

5.7.1 Conceptual Basis

The proposed method estimates the transported component $C_x(\mathrm{infl})$ by examining nearby grid cells that have zero local CO emissions. Any CO present in such cells must originate from transport and/or non-local atmospheric oxidation Processes.

Let:

- N_x be the set of neighbouring grid cells around grid cell x
- N_x^0 be the subset with zero emissions

Then the pseudo influx for grid cell x is estimated as:

$$C_x(\mathrm{infl}) = \mathrm{stat}\left(C_y \,|\, y \in N_x^0\right)$$

where "stat" may be the mean, median, or another robust statistic from neighbouring grid cells that have no emissions (N_x^0). What passes as a "neighbouring" grid cell is to be determined, but preferably results should to first order be insensitive ("robust") to the exact definition of "neighbouring".

5.7.2 Extracting the Local Contribution

For each simulation (baseline = 0 and perturbed = 1), the local contribution is:

$$C_x(\mathrm{loc})^0 = C_x^0 - C_x(\mathrm{infl})^0$$

$$C_x(\mathrm{loc})^1 = C_x^1 - C_x(\mathrm{infl})^1$$

5.7.3 Corrected Local β-Factor

A corrected sensitivity factor is then defined as:

$$\beta\,(\mathrm{loc}) = \frac{(C_x(\mathrm{loc})^1 - C_x(\mathrm{loc})^0)}{(E_x{}^1 - E_x{}^0)}$$

5.7.4 Emission Update Formula

Given a satellite-observed column $C_x(\mathrm{sat})$, emissions may be updated as:

$$E_x(\mathrm{new}) = E_x{}^0 + \frac{(C_x(\mathrm{sat}) - C_x{}^0)}{\beta\,(\mathrm{loc})}$$

or, if correcting only the local component:

$$E_x(\mathrm{new}) = E_x{}^0 + \frac{(C_x(\mathrm{sat}) - C_x(\mathrm{infl})^0 - C_x(\mathrm{loc})^0)}{\beta\,(\mathrm{loc})}$$

5.8 Discussion of Limitations and Caveats

5.8.1 Transport Anisotropy

Transport into a zero-emission cell may not reflect transport into a nearby emitting cell due to:

- Wind direction
- Synoptic variability
- Vertical mixing
- Convection

Thus, the pseudo influx estimate may not be spatially transferable.

5.8.2 Chemical Coupling

CO influences OH, and OH controls CO lifetime. Therefore:

- Local emission perturbations can alter regional chemistry.
- The transported component may not scale linearly with emissions.

This violates the linearity assumption of the β-method.

5.8.3 Zero Emission Cells Are Not Chemically Passive

Even without emissions, such cells may differ in:

- OH concentration
- Boundary layer height
- Vertical mixing
- CO production due to oxidation of methane or non-methane hydrocarbons

Thus, CO levels in zero emission grid cells may not be representative of only transported CO into emitting cells.

5.8.4 A.5.4 Large Spatial Influence

CO columns integrate emissions over:

- Hundreds to thousands of kilometres
- Weeks to months

A local pseudo influx estimate may miss contributions from distant sources.

5.8.5 Practical Usefulness

The proposed modified β-method may function only as a heuristic:

- It may reduce overcorrections.
- It may partially account for regional background.
- It cannot fully resolve nonlocal sensitivities.

More robust alternatives include adjoint models, ensemble perturbations, and 4DVar or Ensemble Kalman Filter assimilation systems.

5.9 Conclusion

The proposed modification of the β-method introduces a pseudo influx term to approximate non-local contributions to CO columns. This approach is conceptually appealing and computationally inexpensive, relying only on information already present in the model fields. However, significant limitations may arise from transport anisotropy, chemical coupling, the long lifetime of CO and local CO production due to oxidation of methane or non-methane hydrocarbons. As such, the method should be primarily viewed and be used as a first order approximate correction of the CO emissions that have been used. Validation against more rigorous sensitivity analysis frameworks could provide more insight in the realism of the method.

References

Agoundedemba, M., Kim, C. K., and Kim, H.-G.: Energy Status in Africa: Challenges, Progress and Sustainable Pathways, Energies, 16, 7708, https://doi.org/10.3390/en16237708, 2023.

Amiri-Simkooei, A.: Least-Squares Variance Component Estimation. Theory and GPS Applications, Nederlandse Commissie voor Geodesie, https://doi.org/10.54419/fz6c1c, 2007.

Andela, N., Morton, D. C., Giglio, L., Paugam, R., Chen, Y., Hantson, S., Van Der Werf, G. R., and Randerson, J. T.: The Global Fire Atlas of individual fire size, duration, speed and direction, Earth Syst. Sci. Data, 11, 529–552, 2019.

Andela, N., Morton, D. C., Schroeder, W., Chen, Y., Brando, P. M., and Randerson, J. T.: Tracking and classifying Amazon fire events in near real time, , Sci. Adv.8eabd2713https://doi.org/doi:10.1126/sciadv.abd2713, 2022.

Anderson, L. D., Dix, B., Schnell, J., Yokelson, R., Veefkind, J. P., Ahmadov, R., and de Gouw, J.: Analyzing the Impact of Evolving Combustion Conditions on the Composition of Wildfire Emissions Using Satellite Data, Geophys. Res. Lett., 50, e2023GL105811, https://doi.org/10.1029/2023GL105811, 2023.

Andreae, M. O.: Emission of trace gases and aerosols from biomass burning–an updated assessment, Atmospheric Chem. Phys., 19, 8523–8546, https://doi.org/https://doi.org/10.5194/acp-19-8523-2019, 2019.

Balch, J. K., Nepstad, D. C., Brando, P. M., Curran, L. M., PorTELa, O., De CARVALHO JR, O., and Lefebvre, P.: Negative fire feedback in a transitional forest of southeastern Amazonia, Glob. Change Biol., 14, 2276–2287, https://doi.org/10.1111/j.1365-2486.2008.01655.x, 2008.

Balch, J. K., Nepstad, D. C., Curran, L. M., Brando, P. M., Portela, O., Guilherme, P., Reuning-Scherer, J. D., and de Carvalho Jr, O.: Size, species, and fire behavior predict tree and liana mortality from experimental burns in the Brazilian Amazon, For. Ecol. Manag., 261, 68–77, 2011.

Baret, F., Weiss, M., Verger, A., and Smets, B.: Algorithm Theoretical Basis Document (ATBD) for LAI, fAPAR, and fCover PROBA-V products at 300m resolution (GEOV3), 1.73, Copernicus Global Land Services, 2016.

Bauer-Marschallinger, B., Paulik, C., Hochstöger, S., Mistelbauer, T., Modanesi, S., Ciabatta, L., Massari, C., Brocca, L., and Wagner, W.: Soil Moisture from Fusion of Scatterometer and SAR: Closing the Scale Gap with Temporal Filtering, Remote Sens., 10, 1030, https://doi.org/10.3390/rs10071030, 2018.

Bevington, P. R. and Robinson, D. K.: Data reduction and error analysis for the physical sciences, 3rd ed., McGraw-Hill, Boston, 2003.

Box, G. E. P., Jenkins, G. M., and Reinsel, G. C.: Time series analysis: forecasting and control, Fourth edition., J. Wiley & Sons, Hoboken, N.J, 1 p., https://doi.org/10.1002/9781118619193, 2008.

Bruzzone, L., Bovolo, F., Amodio, A., Brovelli, M. A., Corsi, M., Defourny, P., Domingo, C., Gamba, P., Kolitzus, D., Lamarche, C., Moser, G., Ottlé, C., Perantoni, G., Pesquer, L., and Zanetti, M.: ESA High Resolution Land Cover Climate Change Initiative - Product

User Manual, 2.0, https://climate.esa.int/media/documents/CCI_HRLC_Ph1-D4.3_PUG_v2.2.pdf, 2024a.

Bruzzone, L., Bovolo, F., Amodio, A., Brovelli, M. A., Corsi, M., Defourny, P., Domingo, C., Gamba, P., Kolitzus, D., Lamarche, C., Moser, G., Ottlé, C., Perantoni, G., Pesquer, L., and Zanetti, M.: ESA High Resolution Land Cover Climate Change Initiative (High_Resolution_Land_Cover_cci): High Resolution Land Cover Maps in Africa (Eastern Sahel region) at 10m spatial resolution for 2019 in Geotiff format, v1.2, https://doi.org/10.5285/F107A4CE186844BB8ADF8CD1F2F6D552, 2024b.

Bühlmann, P. and Künsch, H. R.: Block length selection in the bootstrap for time series, Comput. Stat. Data Anal., 31, 295–310, https://doi.org/10.1016/S0167-9473(99)00014-6, 1999.

Carroll, R. J. and Ruppert, D.: Transformation and Weighting in Regression, Chapman and Hall/CRC, New York, 264 pp., https://doi.org/10.1201/9780203735268, 2017.

Carvalhais, N., Forkel, M., Khomik, M., Bellarby, J., Jung, M., Migliavacca, M., Mu, M., Saatchi, S., Santoro, M., Thurner, M., Weber, U., Ahrens, B., Beer, C., Cescatti, A., Randerson, J. T., and Reichstein, M.: Global covariation of carbon turnover times with climate in terrestrial ecosystems, Nature, 514, 213–217, https://doi.org/10.1038/nature13731, 2014.

Castellanos, P., Boersma, K. F., and van der Werf, G. R.: Satellite observations indicate substantial spatiotemporal variability in biomass burning NO_x emission factors for South America, Atmospheric Chem. Phys., 14, 3929–3943, https://doi.org/10.5194/acp-14-3929-2014, 2014.

Cawson, J. G., Nyman, P., Schunk, C., Sheridan, G. J., Duff, T. J., Gibos, K., Bovill, W. D., Conedera, M., Pezzatti, G. B., Menzel, A., Cawson, J. G., Nyman, P., Schunk, C., Sheridan, G. J., Duff, T. J., Gibos, K., Bovill, W. D., Conedera, M., Pezzatti, G. B., and Menzel, A.: Estimation of surface dead fine fuel moisture using automated fuel moisture sticks across a range of forests worldwide, Int. J. Wildland Fire, 29, 548–559, https://doi.org/10.1071/WF19061, 2020.

Chen, Y., Hall, J., van Wees, D., Andela, N., Hantson, S., Giglio, L., van der Werf, G. R., Morton, D. C., and Randerson, J. T.: Multi-decadal trends and variability in burned area from the fifth version of the Global Fire Emissions Database (GFED5), Earth Syst. Sci. Data, 15, 5227–5259, https://doi.org/10.5194/essd-15-5227-2023, 2023.

Chuvieco, E., Pettinari, M. L., Lizundia-Loiola, J., Storm, T., and Padilla Parellada, M.: ESA Fire Climate Change Initiative (Fire_cci): MODIS Fire_cci Burned Area Pixel product, version 5.1 (3.1), https://doi.org/10.5285/58F00D8814064B79A0C49662AD3AF537, 2018.

Chuvieco, E., Roteta, E., Sali, M., Stroppiana, D., Boettcher, M., Kirches, G., Storm, T., Khairoun, A., Pettinari, M. L., Franquesa, M., and Albergel, C.: Building a small fire database for Sub-Saharan Africa from Sentinel-2 high-resolution images, Sci. Total Environ., 845, 157139, https://doi.org/10.1016/j.scitotenv.2022.157139, 2022.

Chuvieco, E., Pettinari, M. L., Lizundia-Loiola, J., Khairoun, A., Danne, O., Boettcher, M., and Storm, T.: ESA Fire Climate Change Initiative (Fire_cci): Sentinel-3 SYN Burned Area Pixel product, version 1.1, https://doi.org/10.5285/D441079FC77F49FABEB41330612B252F, 2024.

Conti, G., Gorné, L. D., Zeballos, S. R., Lipoma, M. L., Gatica, G., Kowaljow, E., Whitworth-Hulse, J. I., Cuchietti, A., Poca, M., Pestoni, S., and Fernandes, P. M.: Developing allometric models to predict the individual aboveground biomass of shrubs worldwide, Glob. Ecol. Biogeogr., 28, 961–975, https://doi.org/10.1111/geb.12907, 2019.

Copernicus Land Monitoring Service and Copernicus Land Monitoring Service helpdesk: Soil Water Index 2007-present (raster 12.5 km), global, 10-daily - version 4 (Version 4), https://doi.org/10.2909/FF1E948D-36CA-4C70-9E34-393EA3D9C785, 2025.

Das, J. C.: Fourier Analysis, in: Power System Harmonics and Passive Filter Designs, John Wiley & Sons, Ltd, 31–69, https://doi.org/10.1002/9781118887059.ch2, 2015.

Douros, J., Eskes, H., van Geffen, J., Boersma, K. F., Compernolle, S., Pinardi, G., Blechschmidt, A.-M., Peuch, V.-H., Colette, A., and Veefkind, P.: Comparing Sentinel-5P TROPOMI NO_2 column observations with the CAMS regional air quality ensemble, Geosci. Model Dev., 16, 509–534, https://doi.org/10.5194/gmd-16-509-2023, 2023.

Draper, N. R. and Smith, H.: Applied Regression Analysis, 1st ed., Wiley, https://doi.org/10.1002/9781118625590, 1998.

Dubayah, R., Luthcke, S. B., Sabaka, T. J., Nicholas, J. B., Preaux, S., and Hofton, M. A.: GEDI L3 Gridded Land Surface Metrics, Version 2, https://doi.org/10.3334/ORNLDAAC/1952, 2021.

Efron, B. and Tibshirani, R. J.: An Introduction to the Bootstrap, 0 ed., Chapman and Hall/CRC, https://doi.org/10.1201/9780429246593, 1994.

Falge, E., Baldocchi, D., Olson, R., Anthoni, P., Aubinet, M., Bernhofer, C., Burba, G., Ceulemans, R., Clement, R., Dolman, H., Granier, A., Gross, P., Grünwald, T., Hollinger, D., Jensen, N.-O., Katul, G., Keronen, P., Kowalski, A., Ta Lai, C., Law, B. E., Meyers, T., Moncrieff, J., Moors, E., William Munger, J., Pilegaard, K., Rannik, Ü., Rebmann, C., Suyker, A., Tenhunen, J., Tu, K., Verma, S., Vesala, T., Wilson, K., and Wofsy, S.: Gap filling strategies for long term energy flux data sets, Agric. For. Meteorol., 107, 71–77, https://doi.org/10.1016/S0168-1923(00)00235-5, 2001.

Falster, D. S., Duursma, R. A., Ishihara, M. I., Barneche, D. R., FitzJohn, R. G., Vårhammar, A., Aiba, M., Ando, M., Anten, N., Aspinwall, M. J., Baltzer, J. L., Baraloto, C., Battaglia, M., Battles, J. J., Bond-Lamberty, B., van Breugel, M., Camac, J., Claveau, Y., Coll, L., Dannoura, M., Delagrange, S., Domec, J.-C., Fatemi, F., Feng, W., Gargaglione, V., Goto, Y., Hagihara, A., Hall, J. S., Hamilton, S., Harja, D., Hiura, T., Holdaway, R., Hutley, L. S., Ichie, T., Jokela, E. J., Kantola, A., Kelly, J. W. G., Kenzo, T., King, D., Kloeppel, B. D., Kohyama, T., Komiyama, A., Laclau, J.-P., Lusk, C. H., Maguire, D. A., le Maire, G., Mäkelä, A., Markesteijn, L., Marshall, J., McCulloh, K., Miyata, I., Mokany, K., Mori, S., Myster, R. W., Nagano, M., Naidu, S. L., Nouvellon, Y., O'Grady, A. P., O'Hara, K. L., Ohtsuka, T., Osada, N., Osunkoya, O. O., Peri, P. L., Petritan, A. M., Poorter, L., Portsmuth, A., Potvin, C., Ransijn, J., Reid, D., Ribeiro, S. C., Roberts, S. D., Rodríguez, R., Saldaña-Acosta, A., Santa-Regina, I., Sasa, K., Selaya, N. G., Sillett, S. C., Sterck, F., Takagi, K., Tange, T., Tanouchi, H., Tissue, D., Umehara, T., Utsugi, H., Vadeboncoeur, M. A., Valladares, F., Vanninen, P., Wang, J. R., Wenk, E., Williams, R., de Aquino Ximenes, F., Yamaba, A., Yamada, T., Yamakura, T., Yanai, R. D., and York, R. A.: BAAD: a Biomass And Allometry Database for woody plants, Ecology, 96, 1445–1445, https://doi.org/10.1890/14-1889.1, 2015.

Fan, N., Koirala, S., Reichstein, M., Thurner, M., Avitabile, V., Santoro, M., Ahrens, B., Weber, U., and Carvalhais, N.: Apparent ecosystem carbon turnover time: uncertainties and robust features, Earth Syst. Sci. Data, 12, 2517–2536, https://doi.org/10.5194/essd-12-2517-2020, 2020.

Forkel, M., Schmidt, L., Zotta, R.-M., Dorigo, W., and Yebra, M.: Leaf moisture content (live-fuel moisture content) at global scale from passive microwave satellite observations of vegetation optical depth (VOD2LFMC) (01), https://doi.org/10.5281/zenodo.6545571, 2022.

Forkel, M., Schmidt, L., Zotta, R.-M., Dorigo, W., and Yebra, M.: Estimating leaf moisture content at global scale from passive microwave satellite observations of vegetation optical depth, Hydrol. Earth Syst. Sci., 27, 39–68, https://doi.org/10.5194/hess-27-39-2023, 2023a.

Forkel, M., Andela, N., de Laat, J., Huijnen, V., Awotwi, A., Kinalczyk, D., Marrs, C., and Wessollek, C.: Sense4Fire ATBDv2.1. Sentinel-based fuel, fire and emissions products to constrain the changing role of vegetation fires in the global carbon cycle. Algorithm Theoretical Baseline Document Version 2.1, 2023b.

Forkel, M., Andela, N., de Laat, J., Huijnen, V., van Wees, D., Kinalczyk, D., Marrs, C., and Wessollek, C.: Sense4Fire ATBDv3. Sentinel-based fuel, fire and emissions products to constrain the changing role of vegetation fires in the global carbon cycle. Algorithm Theoretical Baseline Document Version 3, 2024.

Forkel, M., Wessollek, C., Huijnen, V., Andela, N., de Laat, A., Kinalczyk, D., Marrs, C., van Wees, D., Bastos, A., Ciais, P., Fawcett, D., Kaiser, J. W., Klauberg, C., Kutchartt, E., Leite, R., Li, W., Silva, C., Sitch, S., Goncalves De Souza, J., Zaehle, S., and Plummer, S.: Burning of woody debris dominates fire emissions in the Amazon and Cerrado, Nat. Geosci., 18, 140–147, https://doi.org/10.1038/s41561-024-01637-5, 2025.

Freedman, D. A.: Bootstrapping Regression Models, Ann. Stat., 9, 1218–1228, https://doi.org/10.1214/aos/1176345638, 1981.

Fu, W., Zhu, L., Kwon, H.-A., Park, R. J., Lee, G. T., De Smedt, I., Liu, S., Li, X., Chen, Y., Pu, D., Li, J., Zuo, X., Zhang, P., Li, Y., Yan, Z., Zhang, X., Zhang, J., Wu, X., Shen, H., Ye, J., Wang, C., Fu, T.-M., and Yang, X.: Evaluating GEMS HCHO Retrievals With TROPOMI Product, Pandora Observations, and GEOS-Chem Simulations, Earth Space Sci., 12, e2024EA003894, https://doi.org/10.1029/2024EA003894, 2025.

Fuster, B., Sánchez-Zapero, J., Camacho, F., García-Santos, V., Verger, A., Lacaze, R., Weiss, M., Baret, F., and Smets, B.: Quality Assessment of PROBA-V LAI, fAPAR and fCOVER Collection 300 m Products of Copernicus Global Land Service, Remote Sens., 12, 1017, https://doi.org/10.3390/rs12061017, 2020.

Giglio, L., Boschetti, L., Roy, D. P., Humber, M. L., and Justice, C. O.: The Collection 6 MODIS burned area mapping algorithm and product, Remote Sens. Environ., 217, 72–85, https://doi.org/10.1016/j.rse.2018.08.005, 2018.

Gröchenig, K. and Razafinjatovo, H.: On Landau's Necessary Density Conditions for Sampling and Interpolation of Band-Limited Functions, J. Lond. Math. Soc., 54, 557–565, https://doi.org/10.1112/jlms/54.3.557, 1996.

Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, J. Hydrol., 377, 80–91, https://doi.org/10.1016/j.jhydrol.2009.08.003, 2009.

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J., Loveland, T. R., Kommareddy, A., Egorov, A., Chini, L., Justice, C. O., and Townshend, J. R. G.: High-Resolution Global Maps of 21st-Century Forest Cover Change, Science, 342, 850–853, https://doi.org/10.1126/science.1244693, 2013.

Harmon, M. E., Fasth, B. G., Yatskov, M., Kastendick, D., Rock, J., and Woodall, C. W.: Release of coarse woody detritus-related carbon: a synthesis across forest biomes, Carbon Balance Manag., 15, 1, https://doi.org/10.1186/s13021-019-0136-6, 2020.

Harvey, A. C.: Forecasting, Structural Time Series Models and the Kalman Filter, Cambridge University Press, Cambridge, https://doi.org/10.1017/CBO9781107049994, 1990.

Hiernaux, P., Issoufou, H. B.-A., Igel, C., Kariryaa, A., Kourouma, M., Chave, J., Mougin, E., and Savadogo, P.: Allometric equations to estimate the dry mass of Sahel woody plants mapped with very-high resolution satellite imagery, For. Ecol. Manag., 529, 120653, https://doi.org/10.1016/j.foreco.2022.120653, 2023.

Holland, E. A., Post, W. M., Matthews, E. G., Sulzman, J. M., Staufer, R., and Krankina, O. N.: A Global Database of Litterfall Mass and Litter Pool Carbon and Nutrients, , https://doi.org/10.3334/ORNLDAAC/1244, 2014.

Jakubauskas, M. E., Legates, D. R., and Kastens, J. H.: Crop identification using harmonic analysis of time-series AVHRR NDVI data, Comput. Electron. Agric., 37, 127–139, https://doi.org/10.1016/S0168-1699(02)00116-3, 2002.

Japan Aerospace Exploration Agency: Global 25 m Resolution PALSAR-2/PALSAR Mosaic (Ver.2.6.0), https://www.eorc.jaxa.jp/ALOS/index_e.htm (last access: 20 March 2026), 2026.

Jiang, Y., Zhou, L., and Raghavendra, A.: Observed changes in fire patterns and possible drivers over Central Africa, Environ. Res. Lett., 15, 0940b8, https://doi.org/10.1088/1748-9326/ab9db2, 2020.

Johnson, M. O., Galbraith, D., Gloor, M., De Deurwaerder, H., Guimberteau, M., Rammig, A., Thonicke, K., Verbeeck, H., von Randow, C., Monteagudo, A., Phillips, O. L., Brienen, R. J. W., Feldpausch, T. R., Lopez Gonzalez, G., Fauset, S., Quesada, C. A., Christoffersen, B., Ciais, P., Sampaio, G., Kruijt, B., Meir, P., Moorcroft, P., Zhang, K., Alvarez-Davila, E., Alves de Oliveira, A., Amaral, I., Andrade, A., Aragao, L. E. O. C., Araujo-Murakami, A., Arets, E. J. M. M., Arroyo, L., Aymard, G. A., Baraloto, C., Barroso, J., Bonal, D., Boot, R., Camargo, J., Chave, J., Cogollo, A., Cornejo Valverde, F., Lola da Costa, A. C., Di Fiore, A., Ferreira, L., Higuchi, N., Honorio, E. N., Killeen, T. J., Laurance, S. G., Laurance, W. F., Licona, J., Lovejoy, T., Malhi, Y., Marimon, B., Marimon Junior, B. H., Matos, D. C. L., Mendoza, C., Neill, D. A., Pardo, G., Peña-Claros, M., Pitman, N. C. A., Poorter, L., Prieto, A., Ramirez-Angulo, H., Roopsind, A., Rudas, A., Salomao, R. P., Silveira, M., Stropp, J., ter Steege, H., Terborgh, J., Thomas, R., Toledo, M., Torres-Lezama, A., van der Heijden, G. M. F., Vasquez, R., Guimarães Vieira, I. C., Vilanova, E., Vos, V. A., and Baker, T. R.: Variation in stem mortality rates determines patterns of above-ground biomass in Amazonian forests: implications for dynamic global vegetation models, Glob. Change Biol., 22, 3996–4013, https://doi.org/10.1111/gcb.13315, 2016.

Jönsson, P. and Eklundh, L.: TIMESAT—a program for analyzing time-series of satellite sensor data, Comput. Geosci., 30, 833–845, https://doi.org/10.1016/j.cageo.2004.05.006, 2004.

Kaiser, J. W., Heil, A., Andreae, M. O., Benedetti, A., Chubarova, N., Jones, L., Morcrette, J.-J., Razinger, M., Schultz, M. G., Suttie, M., and van der Werf, G. R.: Biomass burning emissions estimated with a global fire assimilation system based on observed fire radiative power, Biogeosciences, 9, 527–554, https://doi.org/10.5194/bg-9-527-2012, 2012.

Khairoun, A., Loiola, J. L., Chuvieco, E., Pettinari, M. L., Storm, T., Boettcher, M., and Kirches, G.: ESA CCI ECV Fire Disturbance: D2.2 Algorithm Theoretical Basis Document-SYN, 1.1, ESA Climate Change Initiative, https://climate.esa.int/media/documents/Fire_cci_D2.2_ATBD-SYN_v1.1.pdf, 2023.

Kikuzawa, K., Onoda, Y., Wright, I. J., and Reich, P. B.: Mechanisms underlying global temperature-related patterns in leaf longevity, Glob. Ecol. Biogeogr., 22, 982–993, https://doi.org/10.1111/geb.12042, 2013.

Kunsch, H. R.: The Jackknife and the Bootstrap for General Stationary Observations, Ann. Stat., 17, 1217–1241, https://doi.org/10.1214/aos/1176347265, 1989.

de Laat, A. T. J., Andela, N., Forkel, M., Huijnen, V., Kinalczyk, D., and van Wees, D.: Sentinel-5p Reveals Unexplained Large Wildfire Carbon Emissions in the Amazon in 2024, Geophys. Res. Lett., 53, e2025GL115123, https://doi.org/10.1029/2025GL115123, 2026.

Lamsal, L. N., Martin, R. V., Padmanabhan, A., van Donkelaar, A., Zhang, Q., Sioris, C. E., Chance, K., Kurosu, T. P., and Newchurch, M. J.: Application of satellite observations for timely updates to global anthropogenic NOx emission inventories, Geophys. Res. Lett., 38, https://doi.org/10.1029/2010GL046476, 2011.

Li, Z., Shi, H., Vogelmann, J. E., Hawbaker, T. J., and Peterson, B.: Assessment of Fire Fuel Load Dynamics in Shrubland Ecosystems in the Western United States Using MODIS Products, Remote Sens., 12, 1911, https://doi.org/10.3390/rs12121911, 2020.

Lizundia-Loiola, J., Otón, G., Ramo, R., and Chuvieco, E.: A spatio-temporal active-fire clustering approach for global burned area mapping at 250 m from MODIS data, Remote Sens. Environ., 236, 111493, https://doi.org/10.1016/j.rse.2019.111493, 2020.

Lizundia-Loiola, J., Franquesa, M., Khairoun, A., and Chuvieco, E.: Global burned area mapping from Sentinel-3 Synergy and VIIRS active fires, Remote Sens. Environ., 282, 113298, https://doi.org/10.1016/j.rse.2022.113298, 2022.

Lu, Y. and Wei, C.: Evaluation of microwave soil moisture data for monitoring live fuel moisture content (LFMC) over the coterminous United States, Sci. Total Environ., 771, 145410, https://doi.org/10.1016/j.scitotenv.2021.145410, 2021.

Makuma-Massa, H., Bemigisha, J., Kyasimire, B., Nyiramahoro, E., Begumana, J., Mugerwa, S., Egeru, A., and Cho, M.: Mapping the Potential for Hay Making in Rangelands: A Methodological Proposition, Rangelands, 39, 152–162, https://doi.org/10.1016/j.rala.2017.07.002, 2017.

Matthews, S.: Dead fuel moisture research: 1991–2012, Int. J. Wildland Fire, 23, 78–92, 2014.

Mebane, W. R. and Sekhon, J. S.: Genetic Optimization Using Derivatives: The rgenoud Package for R, J. Stat. Softw., 42, 1–26, https://doi.org/10.18637/jss.v042.i11, 2011.

Monsi, M. and Saeki, T.: Über den Lichtfaktor in den Pflanzengesellschaften und seine Bedeutung für die Stoffproduktion, Jpn. J. Bot., 14, 22–52, 1953.

Monsi, M. and Saeki, T.: On the Factor Light in Plant Communities and its Importance for Matter Production, Ann. Bot., 95, 549–567, https://doi.org/10.1093/aob/mci052, 2005.

Morton, D. C., DeFries, R. S., Nagol, J., Souza, C. M., Kasischke, E. S., Hurtt, G. C., and Dubayah, R.: Mapping canopy damage from understory fires in Amazon forests using annual time series of Landsat and MODIS data, Remote Sens. Environ., 115, 1706–1720, https://doi.org/10.1016/j.rse.2011.03.002, 2011.

de Muck, D., Verger, A., Van Der Goten, R., and Sánchez-Zapero, J.: Product User Manual - CGLS Leaf Area Index (LAI), Fraction of Absorbed Photosynthetically Active Radiation (FAPAR) and Fraction of green Vegetation Cover (FCover), 2.0, Copernicus Global Land Operations, 2025.

Oliva, P. and Schroeder, W.: Assessment of VIIRS 375 m active fire detection product for direct burned area mapping, Remote Sens. Environ., 160, 144–155, https://doi.org/10.1016/j.rse.2015.01.010, 2015.

Olsson, L. and Eklundh, L.: Fourier Series for analysis of temporal sequences of satellite sensor imagery, Int. J. Remote Sens., 15, 3735–3741, https://doi.org/10.1080/01431169408954355, 1994.

Pauls, J., Zimmer, M., Kelly, U. M., Schwartz, M., Saatchi, S., Ciais, P., Pokutta, S., Brandt, M., and Gieseke, F.: Estimating Canopy Height at Scale, https://doi.org/10.48550/arXiv.2406.01076, 3 June 2024., 2024.

Pedlar, J. H., Pearce, J. L., Venier, L. A., and McKenney, D. W.: Coarse woody debris in relation to disturbance and forest type in boreal Canada, For. Ecol. Manag., 158, 189–194, https://doi.org/10.1016/S0378-1127(00)00711-8, 2002.

Petetin, H., Sauvage, B., Parrington, M., Clark, H., Fontaine, A., Athier, G., Blot, R., Boulanger, D., Cousin, J.-M., Nédélec, P., and Thouret, V.: The role of biomass burning as derived from the tropospheric CO vertical profiles measured by IAGOS aircraft in 2002–2017, Atmospheric Chem. Phys., 18, 17277–17306, https://doi.org/10.5194/acp-18-17277-2018, 2018.

Pettinari, M. L.: ESA CCI ECV Fire Disturbance: D4.2.5 Product User Guide – Sentinel-3 SYN, Zenodo, https://doi.org/10.5281/ZENODO.10721951, 2024.

Pettinari, M. L., Loiola, J. L., and Chuvieco, E.: ESA CCI ECV Fire Disturbance: D4.2.1 Product User Guide - MODIS, 1.1, ESA Climate Change Initiative, https://dap.ceda.ac.uk/neodc/esacci/fire/docs/v5.1/Fire_cci_D4.2.1_PUG-MODIS_v1.1.pdf, 2021.

Politis, D. N. and Romano, J. P.: The Stationary Bootstrap, J. Am. Stat. Assoc., 89, 1303–1313, https://doi.org/10.1080/01621459.1994.10476870, 1994.

Prichard, S. J., Kennedy, M. C., Andreu, A. G., Eagle, P. C., French, N. H., and Billmire, M.: Next-Generation Biomass Mapping for Regional Emissions and Carbon Inventories: Incorporating Uncertainty in Wildland Fuel Characterization, J. Geophys. Res. Biogeosciences, 124, 3699–3716, https://doi.org/https://doi.org/10.1029/2019JG005083, 2019.

Punalekar, S. M., Verhoef, A., Quaife, T. L., Humphries, D., Bermingham, L., and Reynolds, C. K.: Application of Sentinel-2A data for pasture biomass monitoring using a physically based radiative transfer model, Remote Sens. Environ., 218, 207–220, https://doi.org/10.1016/j.rse.2018.09.028, 2018.

Raml, B.: Algorithm Theoretical Basis Document - Soil Water Index (SWI) 0.1°, Soil Water Index 10 Days (SWI10) 0.1°, SWI Time Series (SWI_TS) 0.1°, 1.40, TU Wien, https://land.copernicus.eu/en/technical-library/algorithm-theoretical-basis-document-soil-water-index-0.1deg/@@download/file, 2025.

Ramo, R., Roteta, E., Bistinas, I., Wees, D. van, Bastarrika, A., Chuvieco, E., and Werf, G. R. van der: African burned area and fire carbon emissions are strongly impacted by small fires undetected by coarse resolution satellite data, Proc. Natl. Acad. Sci., 118, e2011160118, https://doi.org/10.1073/pnas.2011160118, 2021.

Rasmussen, C. E. and Williams, C. K. I.: Gaussian processes for machine learning, MIT Press, Cambridge, Mass, 248 pp., 2006.

Rego, F. C., Morgan, P., Fernandes, P., and Hoffman, C.: Fire Science: From Chemistry to Landscape Management, Springer International Publishing, Cham, https://doi.org/10.1007/978-3-030-69815-7, 2021.

Reich, P. B., Walters, M. B., Ellsworth, D. S., Vose, J. M., Volin, J. C., Gresham, C., and Bowman, W. D.: Relationships of leaf dark respiration to leaf nitrogen, specific leaf area and leaf life-span: a test across biomes and functional groups, Oecologia, 114, 471–482, https://doi.org/10.1007/s004420050471, 1998.

Rodrigues, M., Resco de Dios, V., Sil, Â., Cunill Camprubí, À., and Fernandes, P. M.: VPD-based models of dead fine fuel moisture provide best estimates in a global dataset, Agric. For. Meteorol., 346, 109868, https://doi.org/10.1016/j.agrformet.2023.109868, 2024.

Roerink, G. J., Menenti, M., and Verhoef, W.: Reconstructing cloudfree NDVI composites using Fourier analysis of time series, Int. J. Remote Sens., 21, 1911–1917, https://doi.org/10.1080/014311600209814, 2000.

Roteta, E., Bastarrika, A., Padilla, M., Storm, T., and Chuvieco, E.: Development of a Sentinel-2 burned area algorithm: Generation of a small fire database for sub-Saharan Africa, Remote Sens. Environ., 222, 1–17, 2019.

Santoro, M. and Cartus, O.: ESA CCI Biomass - Algorithm Theoretical Basis Document v6.0, 6.0, https://climate.esa.int/media/documents/D2_2_Algorithm_Theoretical_Basis_Document_ATBD_V6.0_20241219_DmDjgI9.pdf, 2024.

Santoro, M. and Cartus, O.: ESA Biomass Climate Change Initiative (Biomass_cci): Global datasets of forest above-ground biomass for the years 2007, 2010, 2015, 2016, 2017, 2018, 2019, 2020, 2021 and 2022, v6.0, https://doi.org/10.5285/95913FFB6467447CA72C4E9D8CF30501, 2025a.

Santoro, M. and Cartus, O.: ESA CCI Biomass - Product User Guidev6.0, 6.0, https://climate.esa.int/media/documents/D4.3_CCI_PUG_V6.0_20250606.pdf, 2025b.

Santoro, M., Cartus, O., Carvalhais, N., Rozendaal, D. M. A., Avitabile, V., Araza, A., de Bruin, S., Herold, M., Quegan, S., Rodríguez-Veiga, P., Balzter, H., Carreiras, J., Schepaschenko, D., Korets, M., Shimada, M., Itoh, T., Moreno Martínez, Á., Cavlovic, J.,

Cazzolla Gatti, R., da Conceição Bispo, P., Dewnath, N., Labrière, N., Liang, J., Lindsell, J., Mitchard, E. T. A., Morel, A., Pacheco Pascagaza, A. M., Ryan, C. M., Slik, F., Vaglio Laurin, G., Verbeeck, H., Wijaya, A., and Willcock, S.: The global forest above-ground biomass pool for 2010 estimated from high-resolution satellite observations, Earth Syst. Sci. Data, 13, 3927–3950, https://doi.org/10.5194/essd-13-3927-2021, 2021.

Schroeder, W., Oliva, P., Giglio, L., and Csiszar, I. A.: The New VIIRS 375 m active fire detection data product: Algorithm description and initial assessment, Remote Sens. Environ., 143, 85–96, https://doi.org/10.1016/j.rse.2013.12.008, 2014.

Schwarz, G.: Estimating the Dimension of a Model, Ann. Stat., 6, 461–464, https://doi.org/10.1214/aos/1176344136, 1978.

Schwieder, M., Buddeberg, M., Kowalski, K., Pfoch, K., Bartsch, J., Bach, H., Pickert, J., and Hostert, P.: Estimating Grassland Parameters from Sentinel-2: A Model Comparison Study, PFG – J. Photogramm. Remote Sens. Geoinformation Sci., 88, 379–390, https://doi.org/10.1007/s41064-020-00120-1, 2020.

Seber, G. A. F. and Lee, A. J.: Linear Regression Analysis, 1st ed., Wiley, https://doi.org/10.1002/9780471722199, 2003.

Sfendla, Y., Stavrakou, T., Müller, J.-F., Oomen, G.-M., Opacka, B., Danckaert, T., De Smedt, I., and Lerot, C.: Global VOC emissions quantified from inversion of TROPOMI spaceborne formaldehyde and glyoxal data, Atmospheric Chem. Phys., 26, 733–767, https://doi.org/10.5194/acp-26-733-2026, 2026.

Shen, J., Kaiser, J., and Jin, X.: Fire-driven formaldehyde enhancement and population health burden revealed by TROPOMI in the contiguous U.S., Environ. Res. Lett., 21, 034012, https://doi.org/10.1088/1748-9326/ae3788, 2026.

Sturtevant, B. R., Bissonette, J. A., Long, J. N., and Roberts, D. W.: Coarse Woody Debris as a Function of Age, Stand Structure, and Disturbance in Boreal Newfoundland, Ecol. Appl., 7, 702–712, https://doi.org/10.2307/2269532, 1997.

Tellinghuisen, J.: Statistical Error Propagation, J. Phys. Chem. A, 105, 3917–3921, https://doi.org/10.1021/jp003484u, 2001.

Theiler, J., Eubank, S., Longtin, A., Galdrikian, B., and Doyne Farmer, J.: Testing for nonlinearity in time series: the method of surrogate data, Phys. Nonlinear Phenom., 58, 77–94, https://doi.org/10.1016/0167-2789(92)90102-S, 1992.

Thurner, M., Beer, C., Santoro, M., Carvalhais, N., Wutzler, T., Schepaschenko, D., Shvidenko, A., Kompter, E., Ahrens, B., Levick, S. R., and Schmullius, C.: Carbon stock and density of northern boreal and temperate forests, Glob. Ecol. Biogeogr., 23, 297–310, https://doi.org/10.1111/geb.12125, 2014.

Tsendbazar, N., Li, L., Koopman, M., Carter, S., Herold, M., Georgieva, I., and Lesiv, M.: WorldCover - Product Validation Report v1.1, https://esa-worldcover.s3.eu-central-1.amazonaws.com/v100/2020/docs/WorldCover_PVR_V1.1.pdf, 2021.

Tsendbazar, N., Xu, P., Herold, M., Lesiv, M., and Duerauer, M.: WorldCover - Product Validation Report v2.0, 2.0, https://esa-worldcover.s3.eu-central-1.amazonaws.com/v200/2021/docs/WorldCover_PVR_V2.0.pdf, 2022.

Van De Kerchove, R., Zanaga, D., De Keersmaecker, W., Li, L., Tsendbazar, N., and Lesiv, M.: WorldCover - Product User Manual v1.0, 1.0, https://esa-worldcover.s3.eu-central-1.amazonaws.com/v100/2020/docs/WorldCover_PUM_V1.0.pdf, 2020.

Van De Kerchove, R., Zanaga, D., Xu, P., Tsendbazar, N., and Lesiv, M.: WorldCover - Product User Manual v2.0, https://esa-worldcover.s3.eu-central-1.amazonaws.com/v200/2021/docs/WorldCover_PUM_V2.0.pdf, 2022.

Velásquez-García, M. P., Pope, R. J., Turnock, S. T., Deva, C., Moore, D. P., Mataveli, G., Arnold, S. R., Doherty, R. M., and Chipperfield, M. P.: Understanding drivers and biases of simulated CO emissions from the INFERNO fire model over South America, Biogeosciences, 23, 1341–1364, https://doi.org/10.5194/bg-23-1341-2026, 2026.

Veraverbeke, S., Sedano, F., Hook, S. J., Randerson, J. T., Jin, Y., and Rogers, B. M.: Mapping the daily progression of large wildland fires using MODIS active fire data, Int. J. Wildland Fire, 23, 655–667, https://doi.org/10.1071/WF13015, 2014.

Verbesselt, J., Hyndman, R., Zeileis, A., and Culvenor, D.: Phenological change detection while accounting for abrupt and gradual trends in satellite image time series, Remote Sens. Environ., 114, 2970–2980, https://doi.org/10.1016/j.rse.2010.08.003, 2010.

Verbesselt, J., Zeileis, A., and Herold, M.: Near real-time disturbance detection using satellite image time series, Remote Sens. Environ., 123, 98–108, https://doi.org/10.1016/j.rse.2012.02.022, 2012.

Vernooij, R., Eames, T., Russell-Smith, J., Yates, C., Beatty, R., Evans, J., Edwards, A., Ribeiro, N., Wooster, M., Strydom, T., Giongo, M. V., Borges, M. A., Menezes Costa, M., Barradas, A. C. S., van Wees, D., and Van der Werf, G. R.: Dynamic savanna burning emission factors based on satellite data using a machine learning approach, Earth Syst. Dyn., 14, 1039–1064, https://doi.org/10.5194/esd-14-1039-2023, 2023.

Waliszewska, B., Grzelak, M., Gaweł, E., Spek-Dźwigała, A., Sieradzka, A., and Czekała, W.: Chemical Characteristics of Selected Grass Species from Polish Meadows and Their Potential Utilization for Energy Generation Purposes, Energies, 14, 1669, https://doi.org/10.3390/en14061669, 2021.

van Wees, D., van der Werf, G. R., Randerson, J. T., Rogers, B. M., Chen, Y., Veraverbeke, S., Giglio, L., and Morton, D. C.: Global biomass burning fuel consumption and emissions at 500 m spatial resolution based on the Global Fire Emissions Database (GFED), Geosci. Model Dev., 15, 8411–8437, https://doi.org/10.5194/gmd-15-8411-2022, 2022.

van der Werf, G. R., Randerson, J. T., Giglio, L., Collatz, G. J., Kasibhatla, P. S., and Arellano Jr, A. F.: Interannual variability in global biomass burning emissions from 1997 to 2004, Atmospheric Chem. Phys., 6, 3423–3441, https://doi.org/https://doi.org/10.5194/acp-6-3423-2006, 2006.

van der Werf, G. R., Randerson, J. T., Giglio, L., Collatz, G. J., Mu, M., Kasibhatla, P. S., Morton, D. C., DeFries, R. S., Jin, Y., and van Leeuwen, T. T.: Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997–2009), Atmos Chem Phys, 10, 11707–11735, https://doi.org/10.5194/acp-10-11707-2010, 2010.

van der Werf, G. R., Randerson, J. T., Giglio, L., Van Leeuwen, T. T., Chen, Y., Rogers, B. M., Mu, M., Van Marle, M. J., Morton, D. C., Collatz, G. J., Yokelson, Robert J., and Kasibhatla, Prasad S.: Global fire emissions estimates during 1997-2016, Earth Syst. Sci. Data, 9, 697–720, https://doi.org/https://doi.org/10.5194/essd-9-697-2017, 2017.

van der Werf, G. R., Randerson, J. T., van Wees, D., Chen, Y., Giglio, L., Hall, J., Vernooij, R., Mu, M., Binte Shahid, S., Barsanti, K. C., Yokelson, R., and Morton, D. C.: Landscape fire

emissions from the 5th version of the Global Fire Emissions Database (GFED5), Sci. Data, 12, 1870, https://doi.org/10.1038/s41597-025-06127-w, 2025.

Xu, P., Tsendbazar, N.-E., Herold, M., de Bruin, S., Koopmans, M., Birch, T., Carter, S., Fritz, S., Lesiv, M., Mazur, E., Pickens, A., Potapov, P., Stolle, F., Tyukavina, A., Van De Kerchove, R., and Zanaga, D.: Comparative validation of recent 10 m-resolution global land cover maps, Remote Sens. Environ., 311, 114316, https://doi.org/10.1016/j.rse.2024.114316, 2024.

Yebra, M., Dennison, P. E., Chuvieco, E., Riaño, D., Zylstra, P., Hunt Jr., E. R., Danson, F. M., Qi, Y., and Jurdao, S.: A global review of remote sensing of live fuel moisture content for fire danger assessment: Moving towards operational products, Remote Sens. Environ., 136, 455–468, https://doi.org/10.1016/j.rse.2013.05.029, 2013.

Yebra, M., Scortechini, G., Badi, A., Beget, M. E., Boer, M. M., Bradstock, R., Chuvieco, E., Danson, F. M., Dennison, P., Dios, V. R. de, Bella, C. M. D., Forsyth, G., Frost, P., Garcia, M., Hamdi, A., He, B., Jolly, M., Kraaij, T., Martín, M. P., Mouillot, F., Newnham, G., Nolan, R. H., Pellizzaro, G., Qi, Y., Quan, X., Riaño, D., Roberts, D., Sow, M., and Ustin, S.: Globe-LFMC, a global plant water status database for vegetation ecophysiology and wildfire applications, Sci. Data, 6, 1–8, https://doi.org/10.1038/s41597-019-0164-9, 2019.

Young, P. C., Pedregal, D. J., and Tych, W.: Dynamic harmonic regression, J. Forecast., 18, 369–394, https://doi.org/10.1002/(SICI)1099-131X(199911)18:6%3C369::AID-FOR748%3E3.0.CO;2-K, 1999.

Zanaga, D., Van De Kerchove, R., De Keersmaecker, W., Souverijns, N., Brockmann, C., Quast, R., Wevers, J., Grosu, A., Paccini, A., Vergnaud, S., Cartus, O., Santoro, M., Fritz, S., Georgieva, I., Lesiv, M., Carter, S., Herold, M., Li, L., Tsendbazar, N.-E., Ramoino, F., and Arino, O.: ESA WorldCover 10 m 2020 v100 (v100), https://doi.org/10.5281/ZENODO.5571936, 2021.

Zanaga, D., Van De Kerchove, R., Daems, D., De Keersmaecker, W., Brockmann, C., Kirches, G., Wevers, J., Cartus, O., Santoro, M., Fritz, S., Lesiv, M., Herold, M., Tsendbazar, N.-E., Xu, P., Ramoino, F., and Arino, O.: ESA WorldCover 10 m 2021 v200, https://doi.org/10.5281/zenodo.7254221, 2022.

Zhao, L., Yebra, M., van Dijk, A. I. J. M., Cary, G. J., Matthews, S., and Sheridan, G.: The influence of soil moisture on surface and sub-surface litter fuel moisture simulation at five Australian sites, Agric. For. Meteorol., 298–299, 108282, https://doi.org/10.1016/j.agrformet.2020.108282, 2021.

Zhao, X., Tian, Q., Michelsen, A., Ren, B., Feng, Z., Chen, L., Jiang, Q., Zhao, R., and Liu, F.: Global pattern in terrestrial leaf litter decomposition: The effects of climate, litter chemistry, life form, growth form and mycorrhizal association, Agric. For. Meteorol., 362, 110368, https://doi.org/10.1016/j.agrformet.2024.110368, 2025.