CAMS Service Evolution

D7.1 Report on the implementation of updated biogenic organic aerosol precursor emissions

Due date of deliverable	30.6.2025
Submission date	30.6.2025
File Name	CAMAERA-D7.1-V1.1
Work Package /Task	WP7, Task 7.1, Improved sources of biogenic VOCs
Organisation Responsible of Deliverable	MET Norway, KNMI
Author name(s)	David Simpson, Jhilik Majumdar, Vincent Huijnen
Revision number	V1.1
Status	
Dissemination Level	PU

The CAMAERA project (grant agreement No 101134927) is funded by the European Union.
Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the Commission. Neither the European Union nor the granting authority can be held responsible for them.

1 Executive Summary

The two main objectives of WP7, Task 7.1 were:

1. Update of the biogenic VOC (BVOC) emissions system of the EMEP-MSC-W model
2. Integrate the BVOC emissions system into the IFS

In the first part of this document (Sect. 4, dealing with item 1) we made use of databases and methods from EMEP BVOC system, and have outlined work towards an updated database of BVOC emissions databases, with a focus on the underlying land-cover and LAI data sets, and on emission potentials (EPs) for some key species. This work showed that there is a major need for improved land-cover maps, which provide well defined ecosystem and ideally species-level coverage, together with plant functional type (PFT)-specific LAI values. There are major problems with these land-cover issues for the purposes of both BVOC emissions and for deposition modelling.

The current EPs for isoprene and monoterpenes seemed reasonable for some key species, but the EPs used for sesquiterpenes (SQT) likely needs to be increased for some species, for example Norway spruce, and the temperature functions used in EMEP EPs for SQT may need modification. The EPs for semi-natural vegetation are highly uncertain, also because of the widely differing definitions used in different land-cover datasets.

The main product of this EMEP-based work is a python-based software system that has been constructed to modify and improve global landcover datasets for use in BVOC inventories. The system starts with a global land-cover dataset and modifies the PFTs to provide data more appropriate for both BVOC and deposition modelling. The modifications include replacement of: bare land with desert, some semi-natural regions with Mediterranean scrub, some forests with PFTs more appropriate to the Mediterranean. Although currently being tested with the ECOSG datasets, the aim is to also enable IFS and other global data to be used in due course.

Some recommendations for improved handling of these updated BVOC emissions to the IFS modelling can be made from the item (1) activities. The recommendations include:

- Use of PFT-specific LAI data as far as possible
- Increased consideration/evaluation of the LAI data in use.
- Application of European/“EMEP” emission potentials (EPs) for boreal forest conifers to Eurasian boreal regions (at least Russia) of the outer MEGAN domain, thus ensuring seamless EP maps across the region. This recommendation also follows from the fact that the species dominating this area (Norway spruce, Scots pine) are ubiquitous across much of Eurasia.

As for the global model component (item 2), we have addressed uncertainties related to the EP specification for isoprene, and implemented the option to use EP information for isoprene from an offline dataset. Additionally we assessed uncertainties related to activity factors as based on Guenther et al. (2006). We performed evaluations of our parameterisation at a local site in Belgium, and globally, in IFS-COMPO, against satellite observations of formaldehyde and isoprene. This led to the following conclusions:

- Based on local observations, we find an overall good match of the activity factors used in IFS-COMPO, although the peak values associated with high-temperature events are somewhat under-estimated. This gives confidence in the chosen parameterization for normal summertime conditions.

- When using EP for isoprene which was taken from CAMS-GLOB-BIO v3.1, instead of attributing EP for different vegetation types, then indeed our emission totals and its seasonal cycle, can closely follow the one produced by CAMS-GLOB-BIO v3.1. However, when evaluating this change in an IFS-COMPO simulation against satellite retrievals for isoprene and HCHO the results are more mixed, and no obvious improvement can be detected globally. This stresses the remaining uncertainties in assumed EP data as discussed above, but also the fact that the use of TROPOMI HCHO for model evaluation is challenging due to uncertainties in the data product (the need for a bias correction), and also consideration of the many contributing factors that govern simulated HCHO, apart from changes in isoprene emissions.

Table of Contents

2 Introduction

2.1 Background

The European Union's flagship Space programme Copernicus provides a key service to the European society, turning investments in space-infrastructure into high-quality information products. The Copernicus Atmosphere Monitoring Service (CAMS, https://atmosphere.copernicus.eu) exploits the information content of Earth-Observation data to monitor the composition of the atmosphere. By combining satellite observations with numerical modelling by means of data assimilation and inversion techniques, CAMS provides in near-real time a wealth of information to answer questions related to air quality, climate change and air pollution and its mitigation, energy, agriculture, etc. CAMS provides both global atmospheric composition products, using the Integrated Forecasting System (IFS) of ECMWF - hereafter denoted the global production system -, and regional European products, provided by an ensemble of eleven regional models - the regional production system.

The CAMS AERosol Advancement (CAMAERA) project will provide strong improvements of the aerosol modelling capabilities of the regional and global systems, on the assimilation of new sources of data, and on a better representation of secondary aerosols and their precursor gases. In this way CAMAERA will enhance the quality of key products of the CAMS service and therefore help CAMS to better respond to user needs such as air pollutant monitoring, along with the fulfilment of sustainable development goals. To achieve this purpose CAMAERA will develop new prototype service elements of CAMS, beyond the current state-of-art. It will do so in very close collaboration with the CAMS service providers, as well as other tier-3 projects. In particular CAMAERA will complement research topics addressed in CAMEO, which focuses on the preparation for novel satellite data, improvements of the data assimilation and inversion capabilities of the CAMS production system, and the provision of uncertainty information of CAMS products.

2.2 Scope of this deliverable

2.2.1 Objectives of this deliverables

The two main objectives of WP7, Task 7.1 were:

3. Update of the BVOC emissions system of the EMEP-MSC-W model
4. Integrate the BVOC emissions system into the IFS

2.2.2 Work performed in this deliverable

In this deliverable the work as planned in the Description of Action (DoA, WP7 T7.1) was performed. Details of this work are provided in Sections 3-5 below, with conclusions presented in Section 6.

2.2.3 Deviations and counter measures

No deviations have been encountered.

2.2.4 CAMAERA Project Partners:

HYGEOS	HYGEOS SARL
ECMWF	EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS
Met Norway	METEOROLOGISK INSTITUTT
RC.io	RESEARCHCONCEPTS IO
BSC	BARCELONA SUPERCOMPUTING CENTER-CENTRO NACIONAL DE SUPERCOMPUTACION
KNMI	KONINKLIJK NEDERLANDS METEOROLOGISCH INSTITUUT-KNMi
SMHI	SVERIGES METEOROLOGISKA OCH HYDROLOGISKA INSTITUT
FMI	ILMATIETEEN LAITOS
MF	METEO-FRANCE
TNO	NEDERLANDSE ORGANISATIE VOOR TOEGEPAST NATUURWETENSCHAPPELIJK ONDERZOEK TNO
INERIS	INSTITUT NATIONAL DE L ENVIRONNEMENT INDUSTRIEL ET DES RISQUES - INERIS
IOS-PIB	INSTYTUT OCHRONY SRODOWISKA - PANSTWOWY INSTYTUT BADAWCZY
FZJ	FORSCHUNGSZENTRUM JULICH GMBH
AU	AARHUS UNIVERSITET
ENEA	AGENZIA NAZIONALE PER LE NUOVE TECNOLOGIE, L'ENERGIA E LO SVILUPPO ECONOMICO SOSTENIBILE

3 Biogenic VOC emissions - overview of current status

Terrestrial vegetation is known to emit trace gases in many different structures of hydrocarbons. They are typically highly reactive, and form about 90% of the total atmospheric Volatile Organic Compounds (VOC) as found in the atmosphere.

Biogenic emissions of volatile organic compounds (BVOC) are an important player in atmospheric chemistry. Apart from affecting the gas-phase chemistry, in the scope of the CAMAERA project their relevance to the aerosol budget is highlighted as partial reaction products can condensate and form secondary organic aerosol (SOA) particles. Nevertheless, the gas-phase chemistry is also relevant. Production of formaldehyde (HCHO) and, in turn, CO from biogenic sources can be used to provide constraints on BVOC emissions, as these tracers can be observed from space.

The largest portion of BVOC emissions is known to be emitted in the form of isoprene (C_5H_8), with a value of about 440 Tg yr^{-1} (Sindelarova et al., 2022), followed by monoterpenes, methanol, and a range of other aromatic compounds. The emission magnitude and variability of these different BVOC's varies depending on environmental conditions (vegetation type, meteorology, atmospheric composition), which implies that these emissions vary significantly in space and time. Any of these trace gases can be characterized by different physical properties such as reactivity, volatility and solubility, and therefore different fate once emitted in the atmosphere.

The modelling of BVOC emissions can be treated through the more process-based parameterization (e.g. LPJ-GUESS; Hantson et al., 2017) or the more empirical algorithms (e.g. MEGAN, Guenther et al., 1995, 2006, 2012). In any of these parameterizations the effective emission of individual trace gases depends on a reference emission factor (EF) (or emission potential, EP) for a specific plant functional type (PFT) under standard conditions, which is modulated depending on environmental conditions. In the atmospheric composition modelling community the MEGAN module is amongst the most popular approaches to simulate BVOC emissions, although different research groups apply variations of the basic parameters depending on model specifics. Here, for brevity we refer to the Guenther et al (1993, 1995, 2006 and 2012) papers and algorithms as G93, G95, G06 and G12.

However, there are very large uncertainties in all estimates of biogenic VOC emissions. For example, Messina et al. (2016) illustrated that global isoprene emissions from different sources varied from about 250 – 600 Tg(C)/yr, and estimates of monoterpenes varied from ca. 10 – 140 Tg(C)/yr. In the USA, estimates made by the BEIS3 (Environmental Protection Agency Biogenic Emissions Inventory System 3) (http://www.epa.gov/asmdnerl/biogen.html) and MEGAN systems have shown factors of 2-3 differences (Warneke et al., 2010; Wang et al., 2017), despite the fact that both systems originated in the USA and have access to the same databases. In Europe, major differences have been found between MEGAN and national estimates (e.g. Rinne et al., 2009), which can largely be attributed to the fact that the species which underlie the MEGAN EPs do not reflect the species present in Europe. For example, Guenther et al. (1994) and Geron et al. (1994) assigned an emission potential of 14 ug/g/h (G93 standard definitions, i.e. 30 deg C, 1000 uE PAR) for N. American spruce, but this value is far higher than those measured for Norway spruce (Picea abies) which dominates most European spruce landscapes (Simpson et al., 1995, 1999; Rinne et al., 2009; Keenan et al., 2009; Hakola et al., 2023). Indeed, in the EMEP model Simpson et al. (1995) used 1.75 ug/g/h for Norway spruce, and this value was further reduced to 1 ug/g/h for the latest EMEP inventory (Simpson et al., 2012). As shown in Langner et al. (2012), a CTM model which used MEGAN for BVOC emissions predicted twice as much isoprene over Europe as the EMEP CTM model.

As well as uncertainties in EPs, other significant uncertainties are associated with the algorithms used. For example, issues with respect to the temperature response have been

recently reported by Emmerson et al. (2020) for Australian Eucalyptus trees, and also by DiMaria et al. (2023), who use flux observations and in-situ concentrations of isoprene as constraints. While the default temperature response parameterization works well for a site in the UK, it largely under-estimated this response for a site in the Amazon, suggesting that ecosystem-specific parameterizations are required. Furthermore, it was also indicated that biases in the meteorological drivers, such as surface temperature, can lead to biases in the resulting emissions. These can be compensated by alternative response-functions depending on the assumed meteorology.

An important, though complex question concerns the question how to constrain, and validate such modelling attempts. Measurements of direct fluxes of BVOC emissions are sparse, and when available they are difficult to interpret for model validation purposes, because of the inherent local nature: the observations have limited representativity compared to the generally coarse size of the grid boxes, due to the combination of large variations in vegetation types within a typical model grid box, and the short lifetime of the emitted trace gas.

Likewise, measurements of VOC concentrations of emitted trace gases such as isoprene and terpenes, can be used to infer emissions. In-situ, ground-based observations are less sparse than the emission flux measurements, but still suffer from similar shortcomings associated with representativity. Recently, satellite-based observations of isoprene have been presented (Fu et al., 2019; Wells et al., 2022). These have the potential to provide a representative distribution of C_5H_8 across large regions. A limitation to the use of these observations to infer their emissions is that this still requires assumptions on their lifetime, particularly their dependency on local OH abundance, which may be rather uncertain itself.

A larger amount of observations (both in-situ and satellite-based) have been published for trace gases which are produced from oxidation of emitted BVOCs: CHOCHO, HCHO and CO. For instance, from images of HCHO and CHOCHO satellite retrievals a clear footprint of biogenic emissions activity can be identified. In-situ observations suffer from similar limitations as described above, while both in-situ and remote-sensing observations of these trace gases are additionally depending on uncertainties due to emissions from other sources, and uncertainties associated to the chemistry mechanism which links the precursor emissions with these observed quantities.

In this report, in Section 4 we provide a brief review of some updates concerning EPs for European ecosystems, and focus on investigations of the land-cover and LAI data underlying the BVOC emission inventory. This work is being tested within the EMEP framework, but the results are intended for later use in IFS-COMPO and other CTMs. We will refer to the land-cover and BVOC data sets under development here as “CAMAERA/EMEP”. Then, in Section 5, we continue with reporting on efforts done in IFS-COMPO, with emphasis on the evaluation and improvement of the existing MEGAN-style implementation of BVOC emissions. We describe the default, and updated model configuration, and provide evaluation results from local measurements at a site in Belgium. Section 6 summarises these efforts and outlines ideas for next steps in the BVOC inventory improvement.

4 Towards updated emissions databases and algorithms for BVOC modelling.

4.1 BVOC inventories: EMEP and CAMS-GLOB-BIOv3.1

In the CAMS-GLOB-BIOv3.1 inventory of BVOC emissions (Sindelarova et al., 2022), monthly mean and monthly averaged daily profile emissions were calculated by the Model of Emission of Gases and Aerosols from Nature (MEGANv2.1, Guenther et al. 2012). This model was driven by meteorological reanalyses of the European Centre for Medium-Range Weather Forecasts over years from 2000 onwards. Furthermore, European isoprene emission potential data were updated using high-resolution land cover maps and detailed information of tree species composition and emission factors from the EMEP MSC-W model system (Simpson et al., 1999, 2012).

The motivation for using the EMEP emissions over the European area was that (i) the EMEP system made use of far more detailed vegetation data over the region than MEGAN, and (ii) that several studies have shown large discrepancies between emissions calculated using species-specific emission factors and those calculated by MEGAN-based inputs (Rinne et al., 2009; Langner et al., 2012; Jiang et al., 2019).

For Europe these BVOC emission potentials are built upon a complex system which makes use of procedures developed over many years for the EMEP model (Simpson et al., 1995, 1999, 2012). As well as species-specific estimates of emission potentials, this system utilises maps of 115 forest species from 30 countries (from Köble and Seufert, 2001). These data were merged with land-cover data from the European Environment Agency (EEA) Corine Land Cover 2000 maps, and further processed to the EMEP grid by the Stockholm Environment Institute at York in order to add data from other countries in the (2000 era) EMEP domain and for non-forested vegetation (Cinderby et al., 2007). In later years the EMEP domain was significantly expanded to the east, and data for the expanded area and indeed globally make use of a merger of the GLC_2000 dataset (https://forobs.jrc.ec.europa.eu/products/glc2000/products.php) and data from the Community Land Model (https://www.cesm.ucar.edu/models/clm/, Oleson et al. 2010; Lawrence et al. 2011) as described in Simpson et al. (2017).

In order to provide a manageable number of plant functional types (PFTs) for use in MEGAN, tree species were aggregated in six classes, as summarized in Table 4.1, and for each grid cell the average EP was derived from the area-weighted EPs of each forest-species (Simpson et al., 2012). Thus, the emission potential for a specific land-cover category (e.g. boreal/temperate coniferous forests, CF) can differ substantially from cell to cell, depending on the mix of tree species within the cells, and such potentials are read in as gridded maps for the model.

For non-forest vegetation types (e.g. grasslands, seminatural vegetation) in Europe, or for forest areas not covered by the Köble and Seufert maps (e.g. for eastern Russia), default emission factors taken from Simpson et al. (2012) were applied.

The crop category is the most difficult to deal with in terms of BVOC emissions, not least because the types of crops are not well known (and can change significantly over the years), and the growing seasons are very difficult to specify. In EMEP work to now we have used a simple system which defines the phenology and emission factors of crops using EMEP model definitions. For use with MEGAN, the EMEP model's temperate crop (e.g. wheat), Mediterranean crop (e.g. maize) and root crop (e.g. potato) were aggregated into one crop PFT.

Figure 4.1 illustrates the results of this CAMS-GLOB-BIO – EMEP merge over Eurasia and nearby areas. This figure also shows one of the issues that we hope to solve with the new CAMAERA/EMEP approach. The isoprene EPs shown here have an obvious discontinuity at about 65 degrees East, which is the edge of the EMEP domain. Within the EMEP domain the EMEP EPs and land-cover are used, and outside the domain EPs are directly from MEGAN. A less visible and irregular discontinuity is that the EMEP model made use of detailed forest species data for EU countries, but had to use much more aggregated data for non-EU countries within the domain.

Table 4.1. Generic PFTs used for European emission potential maps provided to CAMS-GLOB-BIO, based on EMEP. From Sindelarova et al. 2022.

PFT	Vegetation included	Examples	LAI variation	LAI max ($m^2\ m^{-2}$)
CF	Temperate and boreal coniferous forest	Norway spruce, Scots pine	Constant	5
DF	Temperate and boreal deciduous forest	European oak, beech, birch	Variable	4
NF	Mediterranean needleleaf forest	Cedars, eucalyptus, stone pine	Constant	4
BF	Mediterranean broadleaf forest	Holm oak, cork oak, arbutus	Constant	4
SNL	Seminatural	Moorland, tundra, shrub	Variable	3
CR	Crops	All crops	Variable	3.5

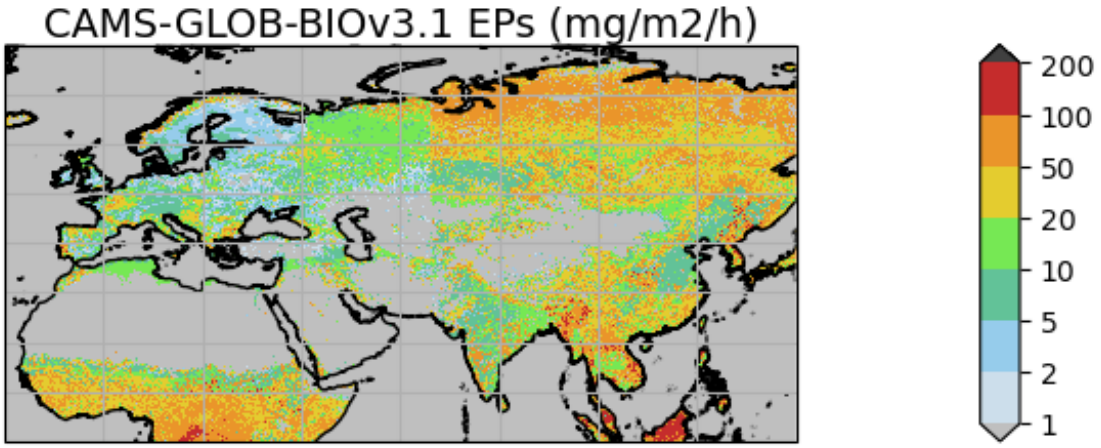

Fig. 4.1 Emissions potentials for isoprene in CAMS-GLOB-BIOv3.1 system (data from Sinderalova et al., 2022).

4.2 Emission potentials, definitions

There are several widely-used approaches to BVOC emission modelling, but the main approaches can be summarised as either using the G95 algorithms or the MEGAN algorithms. In both approaches emission potentials are defined at some specified standard conditions. With G95 (also used in the EMEP model), the standard conditions are defined as 30 degrees C and a photosynthetically active radiation (PAR) levels of 1000 µmol/m^2/s.The standard conditions for the MEGAN model (Guenther et al., 2006) include a leaf area index, LAI, of 5,

a canopy with 80% mature, 10% growing and 10% old foliage; current environmental conditions including a solar angle (degrees from horizon to sun) of 60 degrees, photosynthetic photon flux density (PPFD) transmission (ratio of PPFD at the top of the canopy to PPFD at the top of the atmosphere) of 0.6, air temperature=303 K, humidity=14 g/kg, wind speed=3 m/s and soil moisture=0.3 m^3/m^3; average canopy environmental conditions of the past 24 to 240 h include leaf temperature=297 K and PPFD=200 µmol/m^2/s for sun leaves and 50 µmol/m^2/s for shade leaves. A somewhat simpler big-leaf version of MEGAN (the PCEEA algorithm, see Guenther et al. (2006), also used in IFS-MEGAN, see Sect. 5) uses fewer parameters, only requiring LAI, solar transmission, and monthly temperature and PPFD.

Apart from the definition of standard conditions, another difference in these methods is the specification of the emission potential (EP), which gives the emissions at these standard conditions. In the G95 system, EPs are specified as mass based, gram(BVOC) per gram foliage (dw) per hour, i.e. $g/g_{dw}/h$. In the MEGAN system EPs are specified as gram(BVOC) per m^2 land-area per h, i.e. $g/m^2/h$. (Emissions are sometimes specified as gram species and sometimes as grams of carbon. Here we use simply grams. Other uncertainties are much larger.)

For any vegetation or PFT *i* emissions are calculated at each model time-step using:

$$E_i\ (\mu g/m^2/h) = EP_i^{G95}\ (\mu g/g_{dw}/h) \times D_i\ (g_{dw}/m^2(veg)) \times AF_i \times CCE_i^{G95} \quad (4.1)$$

for the G95/EMEP system, and:

$$E_i\ (\mu g/m^2/h) = EP_i^{MEG}\ (\mu g/m^2/h) \times AF_i \times CCE_i^{MEG} \quad (4.2)$$

for the IFS-MEGAN system. Here D_i is the foliar biomass density (grams of dry weight leaf), AF_i is the area fraction of vegetation *i*, and CCE_i is the canopy environment correction factor (accounting for the instantaneous temperature and PAR levels, as well as other factors for MEGAN). Because of the different standard conditions in the two approaches, EPs from one system cannot be converted consistently to EPs of the other, but Alex Guenther (pers. comm., 2025) has informed that a simple conversion factor between the G95 and MEGAN EPs is used for cases where data are insufficient for specifying the MEGAN-type EPs directly:

$$EP_i^{MEG} \approx EP_i^{G95} \times SLW_i \times LAI_{std} \quad (4.3)$$

Where SLW_i is the specific leaf weight (assumed 150 g/m2 for needleleaf, 100 g/m^2 for broadleaf), and LAI_{std} is 5 m^2/m^2.

Another major difference between the EMEP model BVOC calculations and those of MEGAN is that MEGAN uses the grid-average LAI for all PFTs within a grid-cell, whereas the EMEP model uses PFT-specific LAIs. The advantage of the MEGAN approach is simplicity, in that satellite-derived LAI can be used without further processing or assumptions, but the disadvantage is that deciduous trees will be assigned LAI values even when they should have zero foliage. In practice the MEGAN LAI simplification may not be too important for BVOC emissions as temperatures are low in the winter months, but on the other hand warm periods are possible (e.g. springtime), and LAI has important impacts on deposition modelling.

4.3 Emission potentials, update

The emission potentials used in the EMEP model and hence CAMS-GLOB-BIOv3.1 stem to a large extent from measurements and studies made in the 1990s or previously. For the CAMAERA project we have conducted a literature search for some of the key species in order to determine if and where changes in these EPs are needed. Some of the material is still being evaluated, but here we summarise findings for two of the key species of the boreal forest, and also for the important oak species. (The main issues identified for oaks actually relate to land-cover; efforts to resolve this are discussed below and in Sect. 4.4.2.)

Considering the boreal forests first, this accounts for ca. 30% of the world's forest cover, with a limited number of species (pines, spruces and larch for needleleaf, Betula sp., Alnus sp., and Populus sp. for broadleaved, Hakola et al., 2023). Although the EPs are usually low for both isoprene and monoterpenes, the spatial coverage is large, and biomass density can be much larger than for deciduous species. In the European part of the boreal forest by far the most common species are Norway spruce (Picea abies) and Scots pine (Pinus sylvestris), and especially Norway spruce may be responsible for a significant fraction of isoprene emissions (Simpson et al., 1999, Keenan et al., 2009).

1.1.1 Norway spruce

Hakola et al. (2023) collected all available emission rate data of Norway spruce in Europe that could be used for emission potential calculations and investigated if the growing location or age of a tree would control emission rates, among other things. Some important conclusions from this study were that (a) emissions of sesquiterpenes (SQT) exceeded those of isoprene and monoterpenes (MT), i.e. the EP for SQT might be far higher than we currently assume in models, (b) there is a large variability in EPs even within same tree species, (c) on average, older trees (>80 years) emitted about ten times more isoprene and MTs than younger ones (<80 years), but no clear difference was seen in SQT emissions. and (d) EPs for MTs seemed to be highest in spring, declining towards the autumn.

Figure 4.2 illustrates the summertime EPs for isoprene and monoterpenes documented by Hakola et al. (2023). The current EMEP EPs are also shown. The scatter seen in these plots is a good illustration of the major uncertainties underlying EP estimation for these BVOC, and, combined with the many confounding and not understood factors mentioned in e.g. Hakola et al. (2023), there seems to be little reason to change the current EP defaults in the EMEP model for either isoprene or monoterpenes.

One major discrepancy did arise though from these comparisons. In the default EMEP setup we set sesquiterpene emissions to be 10% of the monoterpene emissions. van Meeningen et al. (2017) suggested a ratio of 20%, and MEGANv2.1seems to suggest 16% for the Finnish locations (Hakola et al., 2023). The new measurements from Hakola et al. (2023) suggest much larger ratios, however, with SQT emissions far exceeding those of MT (SQT/MT = 800-1200%!). The latest data files for MEGANv3 (https://zenodo.org/records/10939297) also suggest a high ratio, 100%. It seems that this ratio cannot be specified with any certainty, but there are grounds to increase the SQT/MT ratio in forthcoming BVOC emission datasets.

Finally, one other issue arose when reviewing these Norway spruce EPs. The EMEP model assumes a foliar biomass density of D = 1400 g/m^2 for Norway spruce up to 60°N (from Veldt, 1989), whereas the MEGAN defaults would assign a value of SLW of 150 g/m^2 (for all needleleaf) and default LAI of 5 m2/m2, giving D = 750 g/m^2; almost 50% lower than EMEP. For comparison, the US EPA BEIS system uses SLW of 75 g/m^2 for broadleaf, 214 g/m^2 for spruce, and 233 g/m^2 for pines (Alex Guenther, pers.comm., 2025). Keenan et al. (2009)

suggested SLW of 235 g/m^2 for Norway spruce. Such higher values might help to explain this discrepancy between EMEP and MEGAN, though it should be noted that Hakola et al. (2023) used SLW of 150 g/m^2 also.

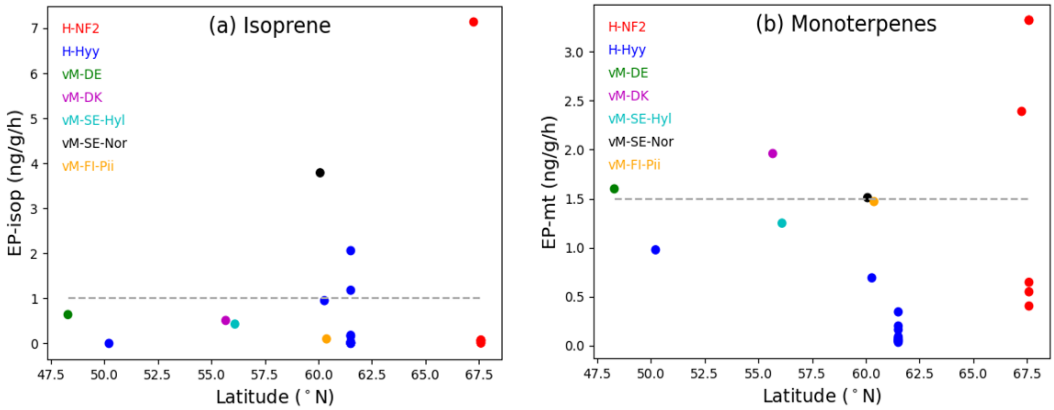

Fig. 4.2. Emission potentials (at G95 standard conditions) for (a) isoprene and (b) monoterpenes, from summer (JJA) values compiled by Hakola et al. (2023). Codes and colours represent sources (H= Hakola et al. 2023, Bo=Bourtsoukidis et al. 2014 vM=van Meeningen et al. 2017) and sites (NF2, Hyy and Fi-Pii are Finnish sites, DE, DK are German and Danish sites, SE-Hyl and SE-Nor are Swedish sites). The EMEP default EPs are shown as the grey dashed line. For more details see Hakola et al. (2023).

1.1.2 Scots pine

For Scots pine the EMEP model currently assumes EP values (G95-style) of 0.1 ug/g/h for isoprene, and 3 ug/g/h for monoterpenes (with a light dependent fraction, LDF = 0.33). Rantala et al. (2015) studied PTR-MS data for BVOC fluxes since 2010, using a tower over the Scots pine dominated Finnish Hyytiälä site, thus providing a very long time-series. Although dominated by Scots pine, some other species (Norway spruce, aspen and birch) were thought to be responsible for at least some of the isoprene fluxes. The data showed that the Guenther 1991/1993 algorithms worked quite well in predicting the measured fluxes, especially when a fraction of the emissions were assumed to be from synthesis and not just from the terpene pools. Although they found some variations in the standardised EPs from month to month, we can derive summertime's (JJA) EMEP-style EPs for monoterpenes of 1.8 ug/g/h (with LDF of 0.44).

These values are thus somewhat smaller than the EMEP model's current settings. However, other studies show a spread of values. Rinne et al. (2009) compiled data from several studies, and reported observed MT EPs of between 0-5–2.5 ug/g/h (high summer values) from observations. Tarvainen et al. (2007) suggested 2.39 ug/g/h in the April-June period, and 1.46 ug/g/h in July-October, though the LDF was assumed to be zero. (They also suggested an isoprene EP of 0.1 ug/g/h.) Karl et al. (2009) suggested 0.1 ug/g/h for isoprene, and 5 ug/g/h (with LDF=0.5) for monoterpenes. Emissions have also been shown to be extremely sensitive to assumptions concerning the impact of new shoots (with extremely high MT emissions) (Aalto et al., 2014, 2015; Taipale et al., 2011), though such effects have very limited experimental support and cannot yet be reliably quantified for CTM usage.

Thus, the EMEP values EPs for monoterpenes from Scots pine seem a little high compared to recent estimates from Finnish data, but not unreasonable when compared to data from other regions. Again, the main unresolved question is likely to what extent SQT emissions need to be modified.

1.1.3 Oaks

The oak genus displays a wide range of EP across the different species, with the EPs of some of the important European ones shown in Table 4.2. The EMEP PFTs in this table are DF = boreal/temperate deciduous broadleaf, and BF = evergreen broadleaf, which is common in southern Europe (Simpson et al., 1999). Despite sharing the same genus, the different classes of oaks have significantly different BVOC characteristics. The two DF species emit isoprene in significant quantities, but little MT. The two BF species emit little isoprene, but are significant emitters of MT. Indeed, both Simpson et al. (1999) and Keenan et al. (2009) have shown that the few species which comprise the DF oaks account for a very large fraction of European isoprene emissions. Conversely, Holm oak was found to be one of the most significant MT emitters in Europe in both studies.

Langford et al. (2017) reviewed five observational datasets of isoprene fluxes from a range of oak forests in the UK, Italy and France, and compared EPs estimated using a range of MEGAN-like definitions and averaging strategies. They were also able to account for the fraction and types of oaks at each site, providing a thorough comparison and evaluation on the range of Guenther/MEGAN type algorithms. The calculated bias of the different methods ranged between +29 and −4 % for the G93 algorithm and between +9 and −40 % for the MEGAN 2.1 approaches. The bias for the G93 algorithm is typically positive, which reflects the fact that the algorithm performs well at the reference conditions which represent typical daytime conditions but performs worse in the morning and afternoon, overestimating emission fluxes due to its inability to account for the attenuation of light and temperature through the canopy.

Using the G93 algorithms, they obtained isoprene EPs of 85 ± 75 and 78 ± 25 µg/g/h for Q. robur and Q. pubescens, respectively. These are essentially identical to the EPs used in EMEP (or Keenan et al., 2009), adding valuable new data in support of these EP values.

We can note that by far the biggest issue we found when investigating data for CAMAERA was that most land-cover databases do not account for the BF-type oaks. This is discussed further in Sect. 4.4.2 below.

Table 4.2 EMEP Emission potentials (G95-style) for Isoprene and monoterpenes (MT) from oak species, from Simpson et al., 2012.

EMEP PFT	Species	Common name	D	EP-Isop.	EP-MT[(a)]
			g/m2	ug/g/h	ug/g/h
DF	Q. robur	English oak	320	80	0.4(0.5)
DF	Q. pubescens	Downy oak	320	80	0.2(1.0)
DF	Q. petraea	Sessile oak	320	45	0.5(1.0)
BF	Q. ilex	Holm oak	500	0.1	30(1.0)
BF	Q. suber	Cork oak	500	0.2	20(1.0)

(a) Numbers in parenthesis give light-dependent fraction where appropriate.

1.1.4 Semi-natural low vegetation

Table 4.3 shows the EPs used in the current EMEP model. There are very large differences in EPs between these different semi-natural categories, which is a serious problem for BVOC inventories since different land-cover data sets differ considerably in their coverage and definitions of these categories. In fact, each PFT typically encompasses many different species and ecosystems. A large number of papers have been published on such ecosystems (especially from northern latitude sites) since the Simpson et al. 1999 review (Kramshøj et al., 2016; Langford et al., 2022; Rinnan et al., 2013, 2020; Simin et al., 2021; Staudt et al., 2017; Vedel-Petersen et al., 2015, e.g.), and these also often show both EPs and temperature responses that differ from the standard G95 and MEGAN approaches. These data are currently being compiled and compared, but the biggest difficulty is likely to decide if a particular land-area should be defined as tundra, shrub, moorland, or other ecosystem. For example, land defined as grass in one data set may well be defined as moorland or shrubs in another; Sect. 4.4.1 discusses this issue further.

Table 4.3 EMEP Emission potentials (G95-style) for Isoprene and monoterpenes (MT) from semi-natural vegetation.

EMEP PFT	Species	D	EP-Isop.	EP-MT[(a)]
		g/m2	ug/g/h	ug/g/h
SNL	Moorland	200	8.5	2.7(0.63)
GR	Grass	400	0.2	0.5(0.6)
MS	Medit. scrub	150	14	2.85(0.3)
WE	wetlands	150	3.4	1.35(0.63)
TU	Tundra	200	8.5	1.35(0.63)
C3_ARCT_GRSS	C3 arctic grass	300	5	0.2(0.5)
C3_NARC_GRSS	C3 non-arctic grass	300	5	0.2(0.5)
C4_GRSS	C4 grass	300	5	0.8(0.5)

(a) Numbers in parenthesis give light-dependent fraction where appropriate.

4.4 Land-cover data sets

An essential input to all BVOC emission data sets, and indeed to CTMs, is the land-cover. Typically CTMs require only a limited number of plant functional types (e.g. deciduous or coniferous forest, grass, etc.), but for BVOC emissions it is advantageous to have more details of the ecosystem and ideally species characteristics. We have investigated a large number of global land-cover (LC) databases, some of which have indeed been used in CTMs and for biogenic emission development. These include systems with 300m global resolution such as the ECLand system used in IFS (Boussetta et al., 2021), the European Space Agency (ESA) CCI (ESA, 2017) and ESA PFT (Harper et al., 2023) systems, and the ECOCLIMAP-Second generation (opensource.umr-cnrm.fr/projects/ecoclimap-sg/wiki, Munier et al. 2018; Druel et al. 2022), also known as ECOSG. The Olson land-cover map (Olson, 1992) provided 57

ecosystem types and was used for the global BVOC estimates of Guenther et al. (1995). The 2001 update of this map was processed for use in the GEOS-Chem CTM (http://wiki.seas.harvard.edu/geos-chem/index.php/Olson_land_map), and is used in the current EMEP modelling system to distinguish base-land from deserts (Simpson et al., 2023). The MEGAN-type inventories have typically used the land-cover data set of the Community Land Model (CLM, http://www.cgd.ucar.edu/models/clm/, Oleson et al. 2010; Lawrence et al. 2011). As described in Simpson et al. (2012), the standard EMEP model uses fine-scale (5 km resolution) European data which merges the CORINE land-cover maps with data from the Stockholm Environment Institute at York (SEIY) which had more detail on agricultural land-cover (Cinderby et al., 2007).

We identified several major issues when investigating how to update and improve the BVOC emissions system for IFS and EMEP:

1. None of the available land-cover databases provided the detail and/or appropriate land-cover classes we require for both BVOC and deposition modelling, either on a global or European basis.
2. Some of the databases provide dominant land-cover, rather than the fractional land-cover which is more suitable for BVOC emissions and CTM models. In these cases cross-walk tables are required to convert the assigned LC to CTM-suitable PFTs (Poulter et al., 2015).
3. Important categories of vegetation, especially evergreen broadleaf forests (e.g. olive, eucalyptus, Holm oak, Cork oak), are not visible in southern Europe, presumably being incorrectly classified as broadleaf deciduous. This is important as such species constitute a substantial fraction of the forest cover in some southern European countries (Simpson et al., 1999), and can have very different BVOC emission rates to the broadleaf species.
4. The various databases are often inconsistent with one another. For example, the ESA PFT database (Harper et al., 2023) defines shrublands as having heights of over 3 m, whereas most other LC datasets (e.g. IGBP classifications, https://fluxnet.org/data/badm-data-templates/igbp-classification/, CLM), define shrubs as below this height.
5. We identified serious problems with the satellite-derived leaf-area index (LAI) typically used with the available data-bases. Particular problems were found with needleleaf forests and croplands.

Table 4.3 provides a brief summary of some of the pros and cons of the main databases considered or used in this study. We will consider two of the main issues in sections 4.4.1 - 4.4.2, and also look at LAI in Sect. 4.5 below. The land cover data sets shown in Table 4.3 have all been converted to netcdf files on a regular lat/lon grid, which enables easier side-by-side comparisons.

Table 4.3 Pros and cons of some global land-cover (LC) datasets as a basis for biogenic emissions and deposition modelling in CTMs.

ESA CCI	**Pro:** Global 300m data-set. Used by IFS. Continuous update. **Con:** Requires cross-walk table. Misses wetlands, tundra, etc. No desert category. Issues with European Med. vegetation.
ESA PFT	**Pro:** Global 300m data-set. Continuous update. PFTs are represent "pure" land-cover types - no cross-walk table required. This simplifies usage in CTMs. **Con:** Few categories of low vegetation (and shrubs defined as >3m). Misses wetlands, tundra, etc. No distinction of boreal, etc.
IFS ECLand	**Pro:** Link to IFS model. Has bogs & marshes. **Con:** Many crops seem to be "Tall Grass". Unclear sources.
ECOCLIMAP-SG	**Pro:** Global 300m data-set, derived from ESA CCI. Distinction of boreal, etc. Distinction of C3 and C4 crops, also summer and winter crops. Has associated height and LAI of vegetation. **Con:** Issues with European Med. vegetation. Misses wetlands, tundra, etc.
CLM	**Pro:** Land-cover used by MEGAN BVOC system. **Con:** MEGAN version fixed. Only categories for vegetation (ie no urban, lakes, deserts , etc.)
Olson	**Pro:** Many LC categories. **Con:** Not updated.
MODIS	**Pro:** Associated with several biogenic inventory systems, e.g. Guenther et al. (2012) for BVOC and Hudman et al. (2012) for soil NOx. **Con:** Categories include "mixed forest". Definition of savannahs places them further north (e.g. Norway) than in other land cover classifications. Does not distinguish fresh from sea water.
SEI/CCE	**Pro:** Harmonised European land-cover map, combining CORINE with JRC-Forest data. Distinguishes several types of low vegetation: grass, moorland, tundra, Mediterranean scrub. Designed with EMEP CTM in mind. **Con:** Lack of global coverage. Discontinued (available through EMEP system).

4.4.1 Landcover - low seminatural categories

Figure 4.3 shows the various semi-natural low vegetation categories in the current EMEP model, and Figs. 4.4 and 4.5 give equivalent results from the ECLand and ECOSG data sets. Major differences are readily seen. For example, for Norway the EMEP data has an extensive coverage of Tundra (“TU”) and wetlands (“WE”), whereas for ECLand the tundra areas seem to be defined as short grass and semi-desert. The ECLand “bogs and marshes” seems roughly equivalent to EMEP’s wetlands. The ECOSG dataset has extensive areas of both temperate grassland and shrubs over Norway, but then the somewhat strange (in this area) classification of flooded grassland seems to be equivalent to EMEP wetlands or ECLand bogs and marshes. As noted in Sect 4.3 above and Table 4.3, these differing classifications can have substantial implications for BVOC emissions.

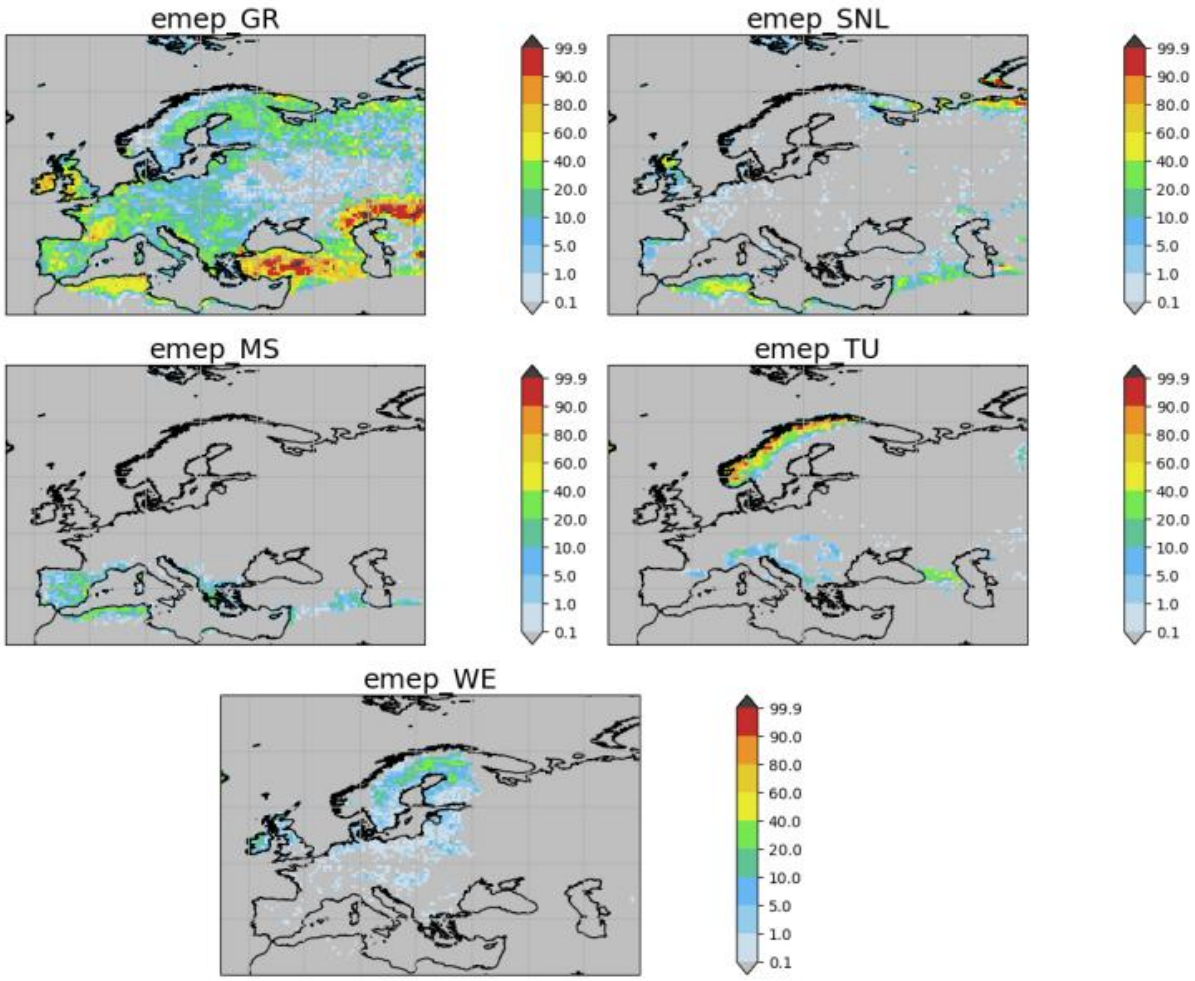

Fig. 4.3 Default EMEP land-cover categories (from Simpson et al., 2012) for low seminatural vegetation.

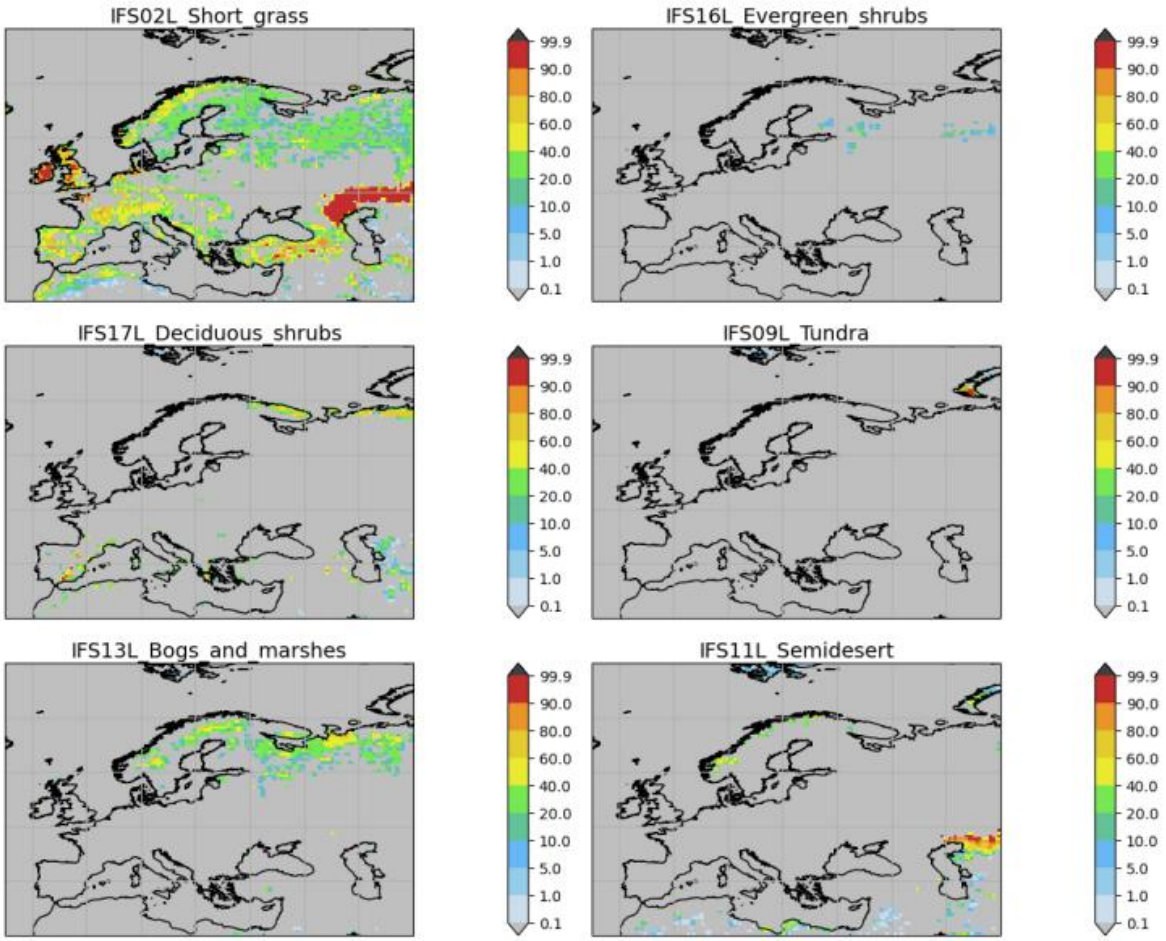

Fig. 4.4 IFS (ECland) land-cover categories for low semi-natural vegetation.

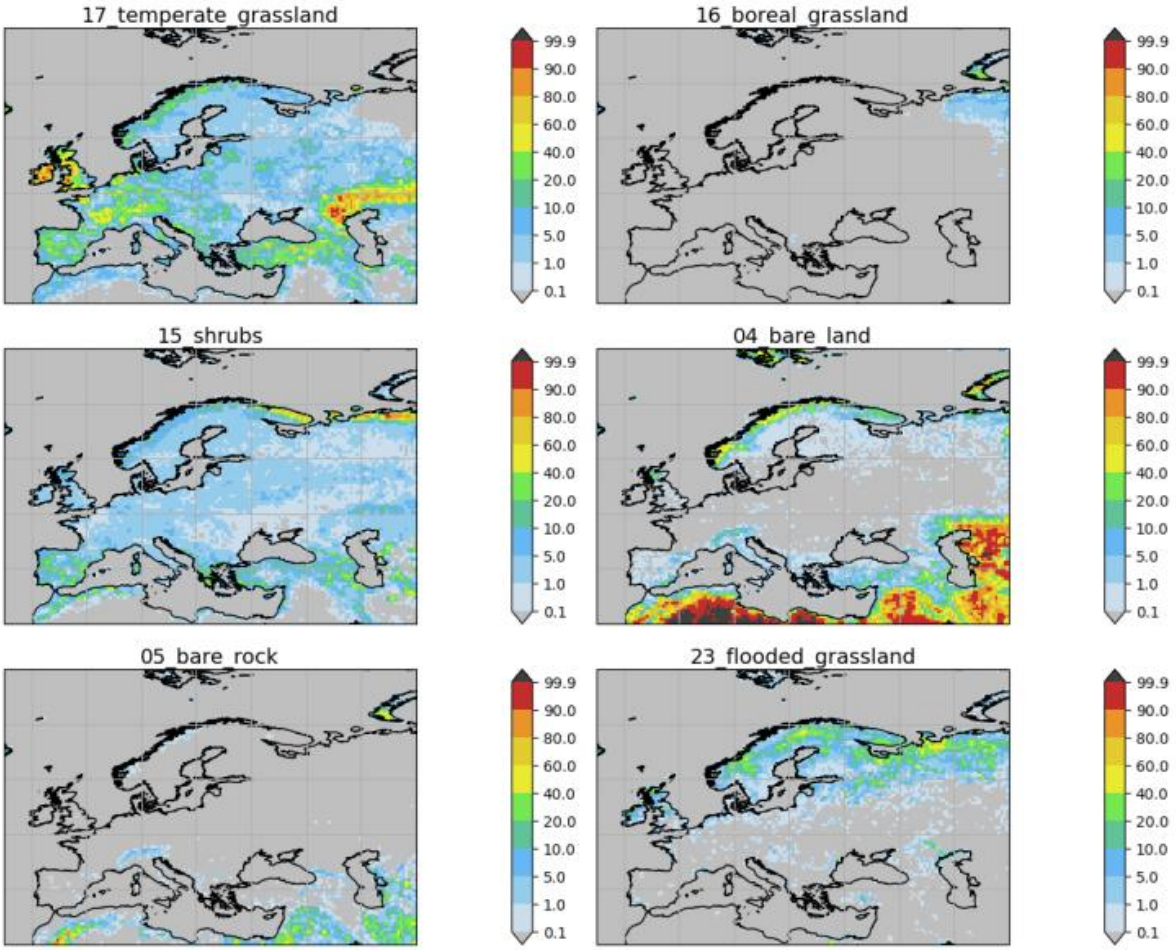

Fig. 4.5 ECO-SG land-cover categories for low seminatural vegetation.

4.4.2 Landcover - southern European forests (evergreen broadleaf etc.)

EMEP has four categories of forests: DF = boreal/temperate deciduous (e.g. beech, English oak), CF = boreal/temperate coniferous (e.g. Norway spruce, Scots pine), BF = Mediterranean evergreen broadleaf (e.g. Holm oak, Walnut), and NF = Mediterranean evergreen conifers (e.g. Maritime pine, juniper). Examples As noted inSect. 4.3 (Oaks) and Table 4.2 the emission potentials for the types of evergreen broadleaf oaks common in southern Europe (Holm oak, Cork oak, in the EMEP BF class) are very different to those of deciduous oaks more typical in central and northern Europe. There are several other evergreen species which have significant forest cover in southern Europe and also rather different BVOC rates to the temperate/boreal species, such as olive or eucalyptus (Simpson et al., 1999). However, Figure 4.6 shows that the ECLand model has essentially zero coverage of evergreen broadleaf species over Europe, and the same is true of both ESA CCI and ESA PFT, and also ESOSG and CLM. In contrast, Figure 4.7 shows the coverage of the EMEP PFTs, and this has a significant coverage of the Mediterranean BF and NF classes. This coverage also matches better the compilation of forest statistics presented in Simpson et al. (1999).

It seems likely that the satellite derived global maps are misinterpreting BF as deciduous broadleaf, and the particular species from EMEP's NF are simply part of the evergreen needleleaf categories. In order to preserve the different contributions of these BF and NF

forests to BVOC emissions (and also to deposition-related characteristics), even when starting with e.g. the IFS or ECOSG global datasets, we have constructed a system to re-allocate some of the global model deciduous or evergreen needleleaf forests to the EMEP categories. This re-allocation is only done within the Mediterranean area though (including Portugal).

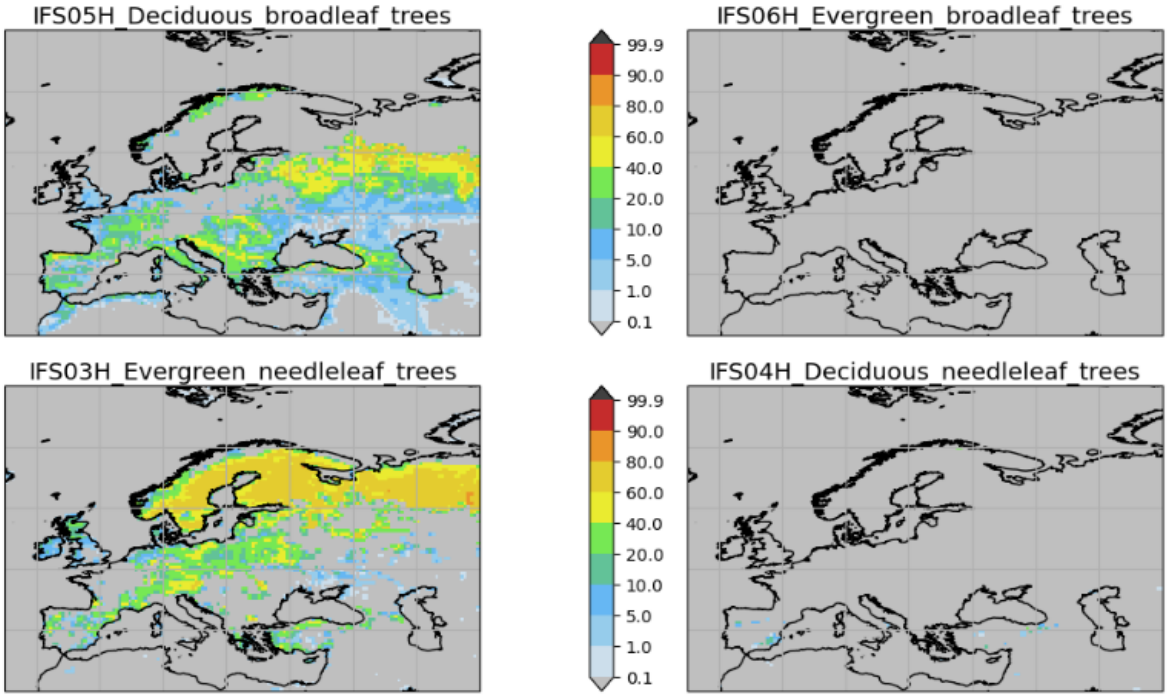

Figure 4.6: Forest categories in Europe from the IFS/ECLand system.

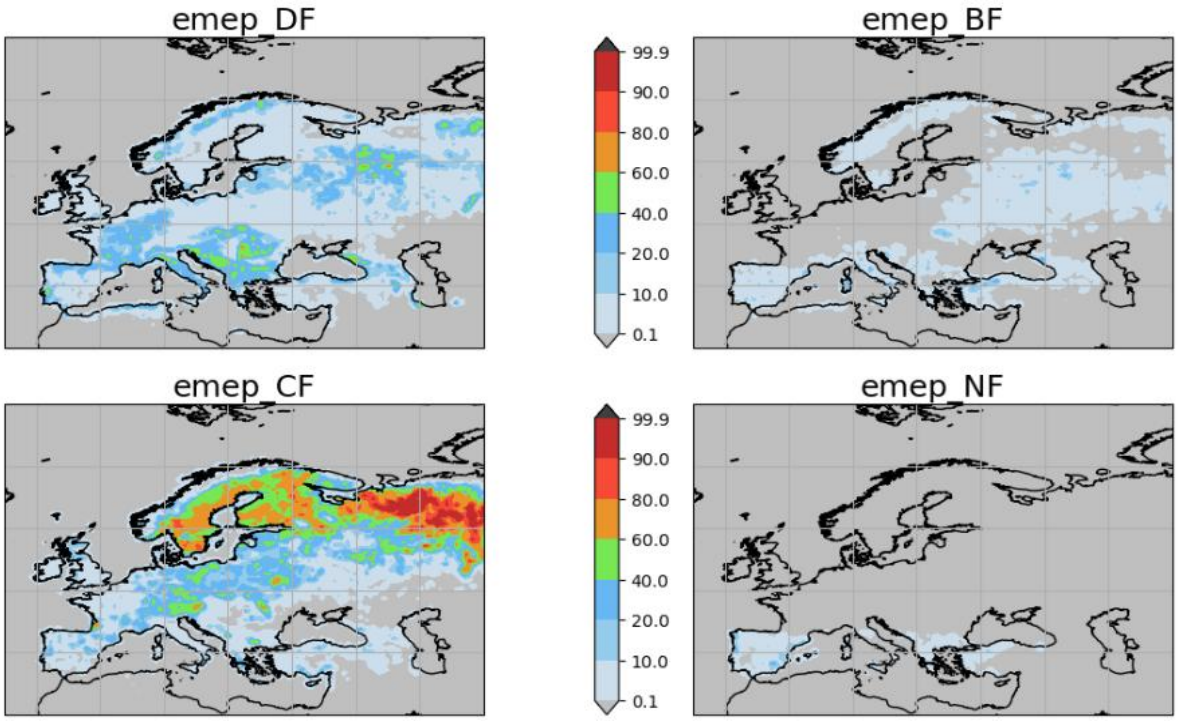

Fig. 4.7. Forest categories in the EMEP land-cover system.

4.5 Leaf area index (LAI)

In the current EMEP model (as used for the CAMS-GLOB-BIOv3.1 data set), each landcover (LC) is assigned a maximum value of LAI, LAImax . In "European" mode, the start and end of growing seasons (SGS,EGS), together with some shape parameters, are set for each LC and a simple polygon function determines LAI on a daily basis. For global runs we use the same LAImax values, but monthly variations in LAI are derived from the LPJ-GUESS ecosystem model (Smith et al., 2001; Schurgers et al., 2009). However, such global LAI values were only available for the forest PFTs, and C3 and C4 "grasslands".

For CAMAERA we initially planned to make use of satellite-derived LAI, or at least temporal variations in LAI, for the major PFTs. Given the increasing availability of such data, and improvements in retrieval algorithms, it was hoped that LAI could be specified for all the major PFT classes. Two datasets were investigated in detail, that of Yuan et al. (2011) (and a PFT-specific extract of this data, kindly provided by Dr. Yuan, 2025), and that of ECOSG Munier et al. (2018). (The Yuan et al. 2011 data were used in Sindelarova et al. (2022).)

Both data sets have advantages and disadvantages, and for deciduous forests both data sets look realistic. However, two particular categories stood out as problematic: (a) evergreen needleleaf forests, and (b) crops.

4.5.1 Forest LAI

Both the Yuan et al and ECOSG data sets showed seasonal variations for evergreen needleaf trees that are not fully realistic. As seen in Fig. 4.8, the LAI values in wintertime become very low in these retrievals. Similar results can be found for ECOSG, GEOV2 and other retrievals (Garrigues et al., 2008; Baret et al., 2006; Verger et al., 2023). Likely causes of the low LAI values in wintertime include issues due to snow cover, low zenith angles, and to mixtures of deciduous and evergreen vegetation within the areas designated as evergreen. Given these issues, however, we cannot recommend the use of satellite-derived LAI for evergreen forests, and the EMEP/CAMAERA land-inputs will instead assign constant LAI values.

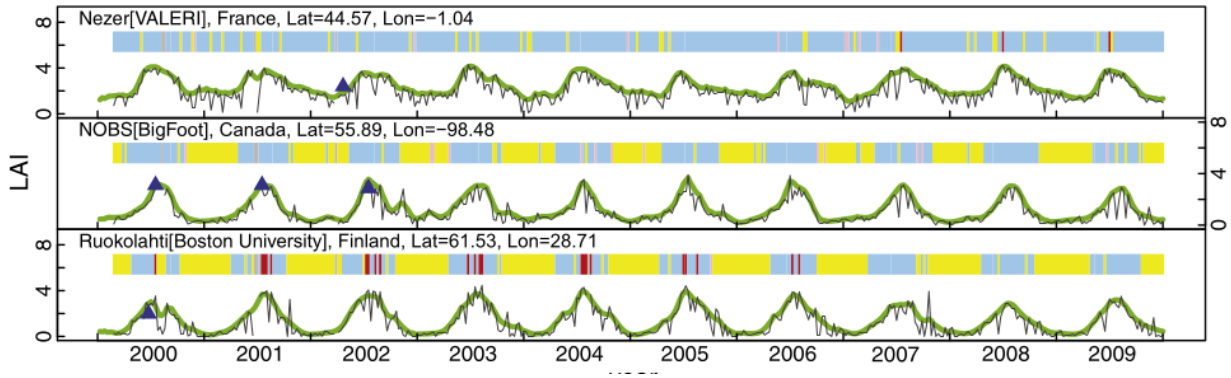

Fig. 4.8. Time-series plot of pixel mean LAI values at three locations. From Yuan et al. (2011). The black and green lines are from MODIS and 'improved MODIS'. The coloured bars represent quality control markers, e.g. blue = "main algorithm", yellow="back-up" - see Yuan et al. (2011) for further explanation.

4.5.2 Crops, LAI

Croplands dominate large areas of the globe, and represent a significant fraction of the LAI across populated land-areas. Although generally not large BVOC emitters, crops are essential for food, and in CTM modelling they also represent an important deposition sink for reactive gases. In the MEGAN system LAI contributions from crops are part of the grid.average LAI

used to estimate BVOC. In the EMEP model we have traditionally represented crops with three generic vegetation types: temperate/boreal crops (TC, representing wheat-like C3 crops), root-crops (RC, e.g. potato), and Mediterranean crops (MC, representing maize-like C4 crops). These three crop types were assigned growing seasons and deposition parameters based upon tabulations and functions detailed in Emberson et al. (2000) and Simpson et al. (2012). For regions outside of Europe, crops were simply treated as grassland, since we did not find appropriate data for the growing seasons or crop types in regions outside of Europe.

For the revised landcover inputs we wish to apply a uniform system across the globe, and to make use of the improvements in satellite and other data which have taken place over the last years. Here we show comparisons between four different methodologies for estimating LAI variations:

- EMEP2012 - uses the methods of Emberson et al. (2000) and Simpson et al. (2012). This methodology is only applicable in Europe.
- GGCMI-Spam - combines data on:
  (a) crop planting and maturity dates from the Agricultural Model Intercomparison and Improvement Project's Global Gridded Crop Model (GGCMI, Jägermeyr et al., 2021). For each of the crop types provided, we have fit a 2-parameter beta distribution between the planting day and the maturity day, assuming that the maximum LAI occurs at a point 80% in between these two dates. (This simple procedure was found to fit the data for wheat presented in Pande et al. 2024 rather well, but of course its applicability to other crops and regions is questionable.),
  (b) crop location and physical area from the 'Spatial Production Allocation Model' (MapSPAM), version SPAM 2020 V1r0 (Yu et al., 2020; International Food Policy Research Institute , IFPRI, 2024).
- Yuan - the LAI dataset of Yuan et al. (2011) reprocessed MODIS collection 5 data to provide time-varying maps at high spatial resolution, and over many years. The published dataset provides LAI per grid-cell, and thus can represent many PFTs. For this work Hua Yuan has kindly provided data which were calculated for specific PFTs, and here we consider C3-crops, and C4-grass as a surrogate for C4-crops.
- ECOSG - The ECOCLIMMAP Second Generation dataset (Munier et al., 2018), provides 300m global resolution land-cover and 10-daily (approx) LAI data for a number of PFTs, including winter C3 crops, summer C3 crops, and C4 crops. These data look very promising for EMEP model usage, since we can simply associate these LAI data with the ECOSG land-cover PFTs.

The comparisons have entailed a rather detailed look at these different LAI methodologies. We start with a look at Fig. 4.9 (Denmark). The caption explains the figure contents, but we can make some points:

1. The top row shows that rape (and wheat, sadly hidden) have a much earlier development compared to barley and maize.
2. The net LAI in this grid cell (middle row) starts to show a complex pattern for C3 crops, and the mean LAI is now significantly less than the crop-specific maximum value of 5 m^2/m^2.
3. The lowest row shows that Yuan and ECOSG have rather similar LAI developments, though differ in magnitude. (All values should be for one-sided LAI, but this question is a common source of confusion.)
4. The Yuan/ECOSG system suggests a simpler (single-peak) seasonal variation in LAI compared to the GGCMI-Spam estimate.
5. The EMEP TC (boreal/temperate crop) estimate looks very crude. This highlights that the EMEP model assumes a single crop which follows the phenology specified in the ICP-Vegetation suggestion for wheat (UNECE 2017), whereas the satellite and crop-calendar estimates allow for multiple crops within the same grid-cells.

Figures 4.10 and 4.11 provide further examples, from the USA and China respectively. The USA example is interesting in that the Yuan/ECOSG C3 LAI seems to peak earlier than the crop-calendar based estimates, and the C4 from ECOSG peaks much earlier. Such issues may be caused by differences in climate between years, or even by the use of different crops to those assumed. Fig. 4.11 for China shows an interesting double peak for C4 crops in the satellite data, but not from the crop-calendar method; likely due to some omitted crop in the calendar approach.

Many other examples have been investigated. In many cases (especially in areas with extensive agricultural coverage such as western Europe) the satellite-derived LAI looks reasonable as far as seasonality is concerned, though the absolute values of LAI can be puzzling (e.g. that LAI in Finland is greater than in France). In some other cases (e.g. some South American cells) almost no seasonality was seen (not shown). A likely confounding issue in these cases is that other vegetation categories (buses, trees) are contributing to the LAI assigned as cropland.

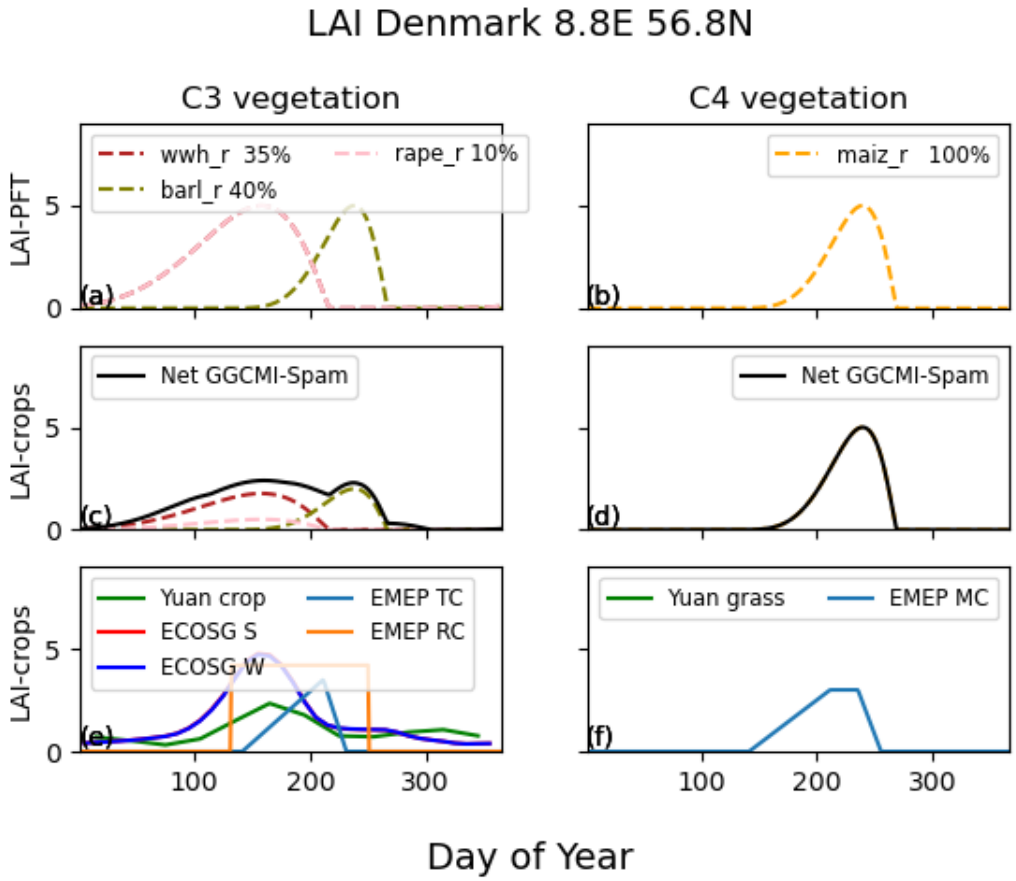

Fig. 4.9 LAI comparisons, for a site in Denmark. Left column is for C3 crops, right for C4 crops (or grass for Yuan), and we have assumed that all crops have a maximum LAI of 5 m^2/m^2. The percentage area covered by each crop (compared to C3/4 total area) is given in the legend. Only crops contributing more than 5% are shown. The top row shows the LAI development from GGCMI-Spam for each crop, regardless of crop coverage in grid cell. (Note that here the wheat (wwh_r) line is masked by that of rape (rape_r). The suffixes (i or r) refer to irrigated or rainfed vegetation from SPAM. The middle row shows the PFT-average development (over all C3 or C4 crops areas) as a black solid line, plus the area-weighted LAI from each of the crops shown in the top row. The bottom row shows the PFT-average LAI from Yuan and ECO-SG (S=summer crop, W=winter crop), as well as from the original EMEP model.

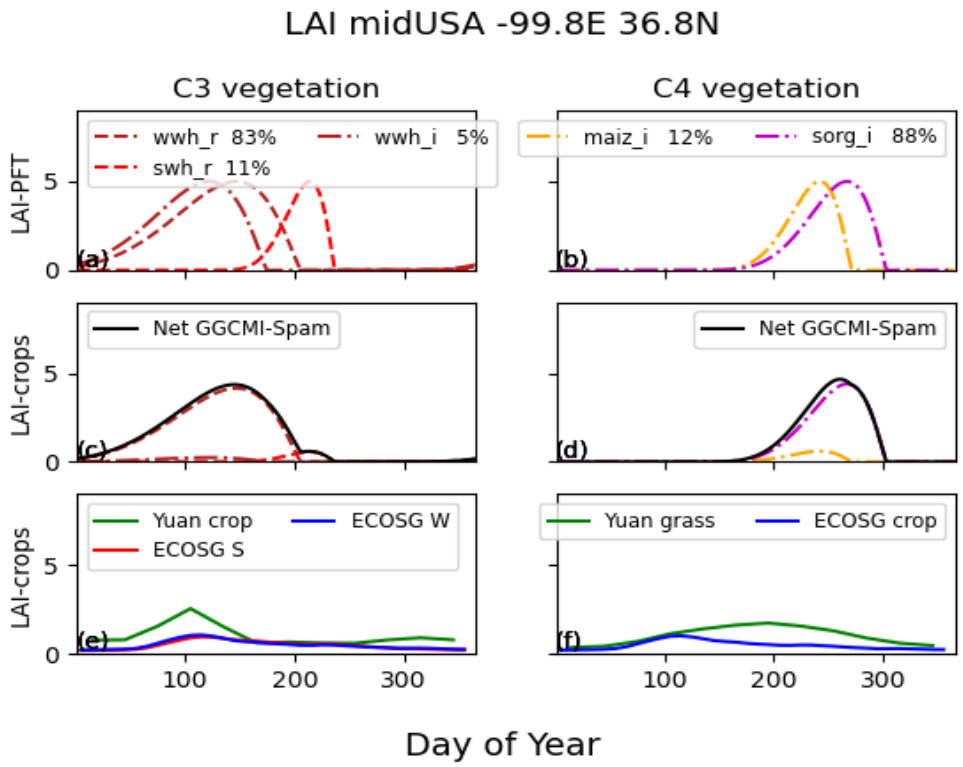

Fig. 4.10. As Fig. 4.9, but for a site in central USA.

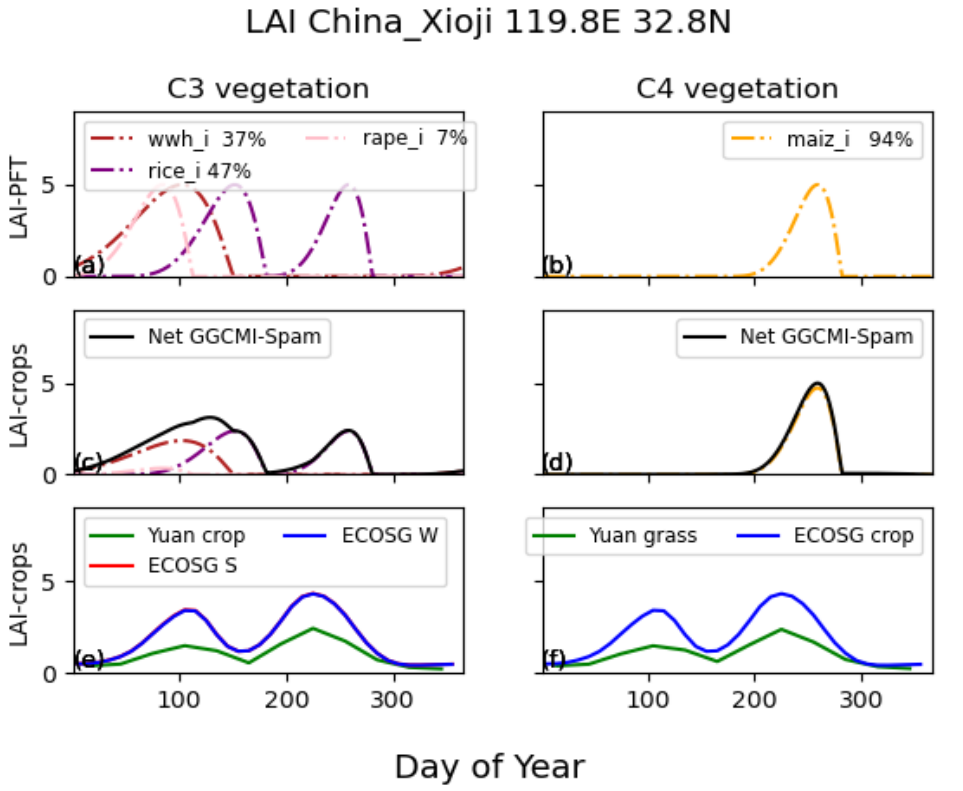

Fig. 4.11 As Fig. 4.9, but for a site in China.

For the updated EMEP/CAMAERA dataset, the preliminary decision is to make use of the satellite-derived LAI for crops, accepting that the values cannot be accurate everywhere, and that the LAI used may be an average over several crop types in the same grid-cell.

4.6 Use of HCHO?

Although most efforts have so far been aimed at improving the land-cover and EP system, some attempts have been made to investigate satellite-derived isoprene emissions. This work will continue in phase 2 of the project, but Fig. 4.12 provides an example of the importance of BVOC emissions for modelled HCHO in Europe for July 2019. When considering the surface HCHO only (left-hand fig.), the BVOC emissions do make substantial contributions (> 50%) to the modelled HCHO, at least in southern Europe, but in much of Europe the contribution is less. When considering the column HCHO (right-hand fig.), which is more relevant for comparison with satellite data, the BVOC contribution is everywhere less than 50%, and in many areas less than 20%. This lack of sensitivity will make it challenging to use HCHO as a tracer for BVOC emissions. In phase 2 we will follow up on these comparisons, but also make use of CrIS data (Wells et al., 2022) which provides isoprene columns directly - these data are discussed in relation to IFS COMPO in Sect. 5.4.

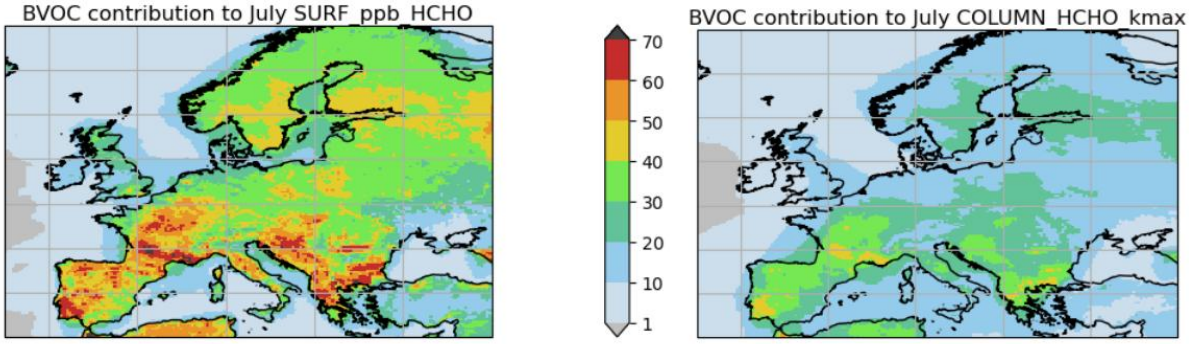

Fig. 4.12 BVOC contribution (in %) to modelled modelled HCHO for July 2019. Left figure shows results for surface concentrations; right figure shows results for column HCHO. Calculations with EMEP MSC-W model.

4.7 Putting it all together: revised BVOC system

As noted above, a large number of data sets have been examined as candidates for the revised BVOC emission system. It is clear that no one dataset fulfills all requirements, so a system (python based) has been constructed with the following features:

- The global dataset (here denoted GLOB-PFT) is first converted (if needed) to a netcdf format with lat/long projection.
- The set of desired PFTs is created, here denoted NEW-PFT.
- A mapping table is created between the PFTs from GLOB-PFT to NEW-PFT. Where no direct mapping is possible, a code is used to suggest which modification is needed. Currently these codes allow:
  - conversion of “bare” land to deserts, based upon Köppen-Geiger categories.
  - conversion of “flooded trees” (from ECOSG) to tropical, temperate or boreal trees, depending on the Köppen-Geiger categories.

    - conversion of seminatural GLOB-PFTs to EMEP's mediterranean scrub (MS) category where appropriate (from EMEP land-cover).
    - conversion of deciduous and coniferous forests to EMEP BF and NF categories (from EMEP land-cover).
- Use of satellite-derived LAI for deciduous PFTs, grassland and semi-natural.
- Use of satellite-derived LAI for crops.
- Retain constant LAI values for evergreen species.

This system will certainly be updated in the future, especially if issues with LAI can be resolved, but the current system works and produces new files readable by the EMEP model as intended. Although currently being tested with the ECOSG data sets, the aim is to also enable IFS and other global data to be used in due course.

Emission potentials for isoprene and terpene have so far remained unchanged since the EPs seem consistent with recent data, but further analysis of the EPs for seminatural ecosystems is warranted (however, again, the main problem is likely with land-cover definitions). For sesquiterpenes an increase in the EP is anticipated, but results from more PFTs will be assembled before making this change.

Given the difficulties found with especially the land-cover definitions, use of satellite data to constrain and improve the basic maps of EP would be extremely useful. This should be a priority for the next steps in developing improved BVOC inputs.

5 Development of online BVOC modeling in IFS-COMPO

In a previous effort, a technical implementation of online emission calculation in ECMWF's Integrated Forecasting System (IFS) has been achieved, as described in Huijnen et al. (2023). In this approach, the BVOC emission rate *E* of chemical compound *j* is described by:

$$E_j = \gamma_j \sum_i^{PFT} f_{PFTi} \varepsilon_{ij} \quad (5.1)$$

where γ is a dimensionless activity factor, f is a fraction of a grid box covered by PFT_i and ε_{ij} is the emission factor which represents amount of the emission under standard conditions in units [µg m^{-2} h^{-1}]. The values for ε_{ij} used here are based on MEGAN PFTs as specified in G12, except for α-pinene, which use a factor 2 lower values than reported there (Messina et al., 2016). This implementation is referred to as IFS-MEGAN, and is based on the G06 and G12 parameterizations. The net activity factor is computed based on an emission response to light, temperature, leaf area index leaf age and CO_2 inhibition, essentially following G06 with a Parameterized Canopy Environment Emission Activity (PCEEA) algorithm.

Specifically, the net activity factor is computed based on an emission response to light (γ_P), temperature (γ_T), leaf area index (γ_{LAI}) leaf age (γ_A) and CO_2 inhibition (γ_C) following:

$$\gamma = \gamma_P \gamma_T \gamma_{LAI} \gamma_A \gamma_C \quad (5.2)$$

The BVOC emissions parameterization is implemented as a separate module as part of ECLand (Boussetta et al., 2021), i.e. the land component of the IFS. Here we have addressed both uncertainties related to the emission factor description (Sec 5.1), as well as uncertainties related to the activity factor (Sec 5.2). Finally, we have tested the default, and updated model configuration in IFS-COMPO, and confronted these simulations against independent satellite observations (Sec. 5.3).

1.2 5.1 Vegetation description and emission factors

The MEGAN PFT categories are different from the ones that are available in ECLand, which implies that a set of conversions is required. The ECLand scheme uses a set of Biosphere-Atmosphere Transfer Scheme (BATS) classes. In CY49R1, these classes are converted from a version of ESA CCI data (ESA, 2017) using a dedicated cross-walking table (Boussetta et al., 2021). Here "residual" (non pure) vegetation types such as the interrupted forest (which covers about 25% of the land points in the current ECMWF maps) are removed compared to earlier distribution of BATS classes. This was done as an attempt to filter for more actual/pure plant types, to better allow for the characterization of the model parameters (Boussetta et al., 2021). Two BATS-type vegetation tiles are available in the IFS: "high vegetation" and "low vegetation". These two tiles come with specifications of the type of high/low vegetation.

For the purpose of estimating BVOC emission, the vegetation types in the BATS classes have to be converted to MEGAN-type PFTs, in order to identify the relevant ε_{ij}. Specifically, MEGAN employs 16 PFT categories derived from the Community Land Model (CLM4). These 16 categories capture the diversity of vegetation by grouping species with similar traits, such as broadleaf deciduous trees, needleleaf evergreen trees, grasses, and shrubs. For the conversion of ECLand BATS classes towards MEGAN PFTs a very basic set of rules was also defined, see Huijnen et al. (2023), which was based on a rough analysis of the distribution of ECLand vegetation categories compared to the distribution of PFT categories as used in

CAMS-GLOB-BIO. Particularly in IFS-MEGAN simplifications have been made: the Köppen-Geiger classification maps, which govern the climate zone, have been represented by very basic latitudinal bands. This implies particularly that both longitudinal variations in vegetation types, and vegetation in high-altitude (mountainous) regions are currently poorly represented.

In CAMAERA we have invested in the implementation of a better way to describe the distribution of isoprene emissions, which is one of the most important BVOC emissions. Instead of using an emission factor per land use category we make use of a gridded map of emission potential data. This means that the equation (5.1) reduces to:

$$E_{C5H8} = \gamma_{C5H8} \mathrm{EP}_{C5H8} f \tag{5.3}$$

Here EP_{C5H8} refers to the emission potential for isoprene and *f* the gridbox fraction covered by vegetation. Here we use a gridded dataset which is the same as the one used in CAMS-GLOB-BIO v3.1 (Sindelarova et al., 2022), except that also for Europe we adopt a constant value throughout the year. The implementation required technical investments in order to (1) convert the input netcdf dataset into grib, which is the default data format for input/output data of the IFS; (2) design and implement a structure for the IFS to ingest this new dataset; (3) implement the code infrastructure to make the dataset available to the BVOC emissions module and (4) validate and evaluate the new parameterization. After evaluation, it was found that an additional scaling factor of 1.5 is needed to match the regional total isoprene emissions as simulated in IFS-COMPO to those from CAMS-GLOB-BIO v3.1. The reasons for this discrepancy are so far not understood, and subject of further research.

1.3 5.2 Parameterisation of activation factors

The net activity factor is computed based on equation 2 given above, essentially following G06 with a Parameterized Canopy Environment Emission Activity (PCEEA) algorithm. The parameterization of the light response to emissions (γ_P) is simulated as:

$$\gamma_P = \left(1 - \mathrm{LDF}_j\right) + \mathrm{LDF}_j \gamma_{\mathrm{P\_LDF}} \tag{5.4}$$

with LDF_j the light-dependent fraction specific for each emission compound, and γ_{P_LDF} (following the corrigendum to G06):

$$\gamma_{\mathrm{P\_LDF}} = \sin(a)\left[2.46\,\Phi\left(1 + 0.0005\,(P_{daily} - 400)\right) - 0.9\Phi^2\right] \quad \text{for } \sin(a) > 0 \tag{5.5}$$

Here a is a solar angle and P_{daily} represents the daily average above canopy photosynthetic photon flux density (PPFD) representative of the simulation period (week or month). In IFS-MEGAN CY49R2 P_{daily} is approximated through a climatological value of the PPDF, which has a seasonal cycle only. Furthermore, is the above canopy PPFD transmission, which is estimated as

$$\Phi = \frac{P}{\sin(a)\, P_{toa}} \tag{5.6}$$

Here P is above-canopy PPFD, and P_{toa} is the PPFD at the top of the atmosphere. For this we adopt the expression given in G06, which depends on the day of the year (DOY) only:

$$P_{toa} = 3000 + 99\cos(6.28\,(\mathrm{DOY} - 10)/365) \quad (5.7)$$

For the temperature activity factor we follow G06:

$$\gamma_T = E_{opt}\left(\frac{C_{T2}\exp\,(C_{T1}x)}{C_{T2} - C_{T1}[1 - \exp\,(C_{T2}x)]}\right) \quad (5.8)$$

where

$$x = \left(\frac{1}{T_{opt}} - \frac{1}{T_{hr}}\right)/0.00831 \quad (5.9)$$

with $E_{opt} = 1.75\exp\,(0.08(T_{daily} - 297))$ and $T_{opt} = 313 + (0.6(T_{daily} - 297))$. T_{hr} represents hourly average air temperature (K), T_{daily} as daily average air temperature (K) representative of the simulation period, C_{T1} (=80 kJ mol^{-1}) and C_{T2} (=200 kJ mol^{-1}) empirical coefficients. In IFS-MEGAN, T_{hr} is taken the instantaneous temperature at the lowest model level in the atmosphere, while T_{daily} is taken from the 3rd level (28-100 cm) soil temperature. DiMaria et al. (2023) note that this temperature response function is actually highly uncertain.

For emission response to leaf area index variations, γ_{LAI}, we again follow G06 using the formulation

$$\gamma_{LAI} = 0.49\,\frac{\mathrm{LAI}}{\sqrt{1 + 0.2\mathrm{LAI}^2}} \quad (5.10)$$

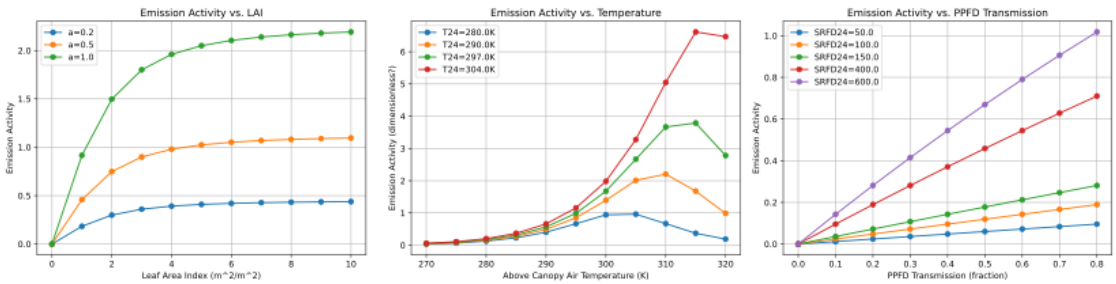

Figure 5.1. Illustration of instantaneous activation factors for LAI, γ_{LAI} (left), as function of LAI (horizontal axis), and pre-factor (color), Temperature γ_T (middle), as function of air temperature (horizontal axis) and soil temperature (colors) and radiation γ_P (right), as function of PPFD (horizontal axis).

Following G06, the dependency of the activity factors to the various input variables (LAI, temperature and radiation aspects) is given in Figure 5.1, showing the different types of sensitivity to these variables. In Figure 5.2 we also present maps with instantaneous values for these activation factors, giving an impression of the spatial variation of the given factors. While LAI is also high at the high latitudes, due to the lower temperatures (and radiation) there the effective emissions will still be lower. Conversely, over the Sahara the solar radiation and temperature are highest, but in absence of vegetation there are no BVOC emissions to be expected.

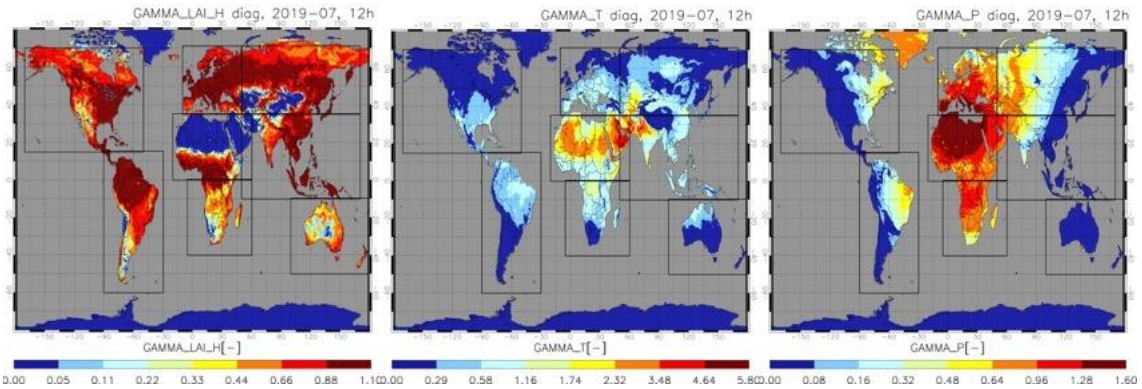

Figure 5.2. Illustration of instantaneous activation factors for LAI (left), Temperature (middle) and radiation (right), here for 1 July 2019, 12 CET, here for isoprene.

The emission response to leaf age, γ_A, also follow G06, while for γ_C, i.e. the CO_2 inhibition, equation 14 in G12 is adopted. However, both activity factors are very close to unity for default IFS-COMPO 'present day' simulations. γ_A is close to unity everywhere because it is parameterized as a function of temporal variation in LAI, while the LAI input data in the IFS shows only very weak variations as this is governed by climatological data. The γ_C is close to unity because the parameterization is normalized towards a CO_2 concentration of 400 ppmv, which is the present-day concentration. Therefore we do not consider these activity factors any further. Nevertheless, in the next subsection we test the validity of the activation factors that govern the isoprene emissions against in-situ observations, with emphasis on the temperature and light-response.

In Figure 5.3 we present the resulting isoprene emissions for the original, and updated model implementation, and compare them with CAMS-GLOB-BIO. Indeed, when using the emission potential map for C_5H_8, together with the 1.5 scaling, the match with CAMS-GLOB-BIO improves a lot, especially over Africa and South America. Still, over Australia we find a large discrepancy, which is likely associated to differences in the modeling of activation factors. Also the regional totals show improved consistency between CAMS-GLOB-BIO and IFS-MEGAN using the EP, Figure 5.4. Especially over Europe a large decrease is seen, consistent with the updated EP in CAMS-GLOB-BIO v3.1.

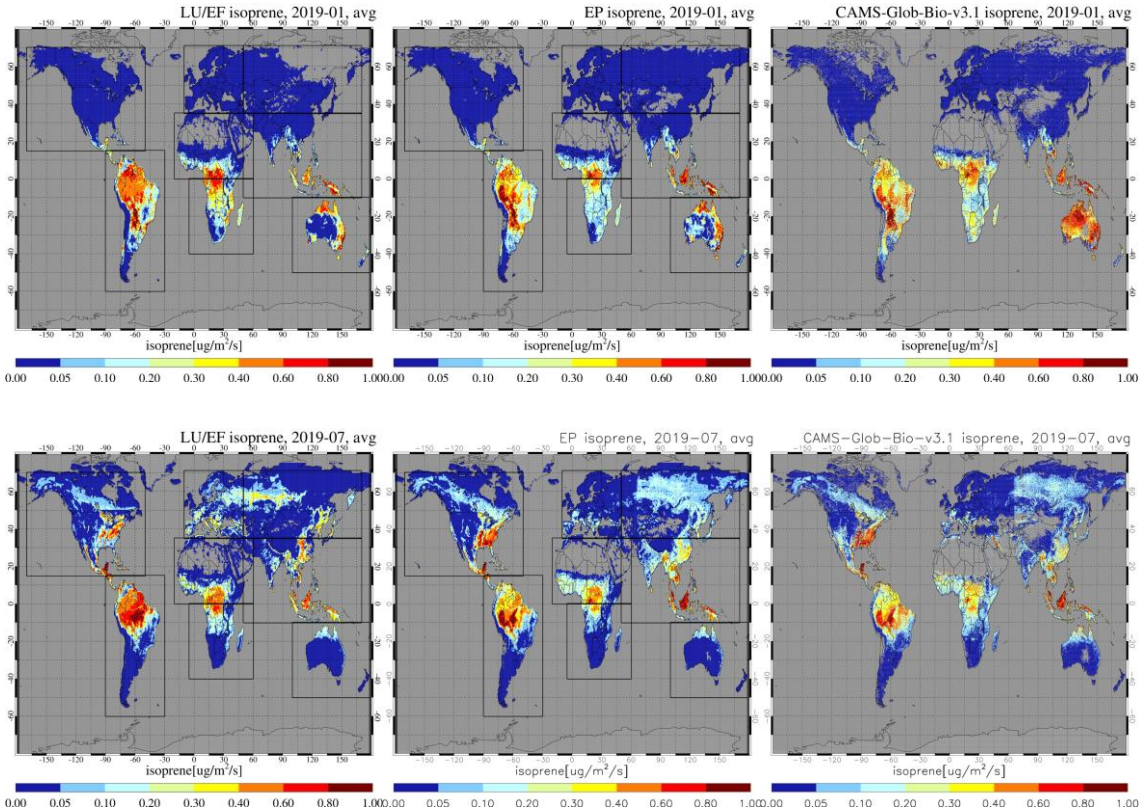

Figure 5.3. Comparison of monthly mean isoprene emissions for January (top) and July (bottom) estimated using fixed emission factor per land-use category (LU/EF, left), and when using an emission potential map (EP, middle), compared to CAMS-GLOB-BIO v3.1 emissions (right)

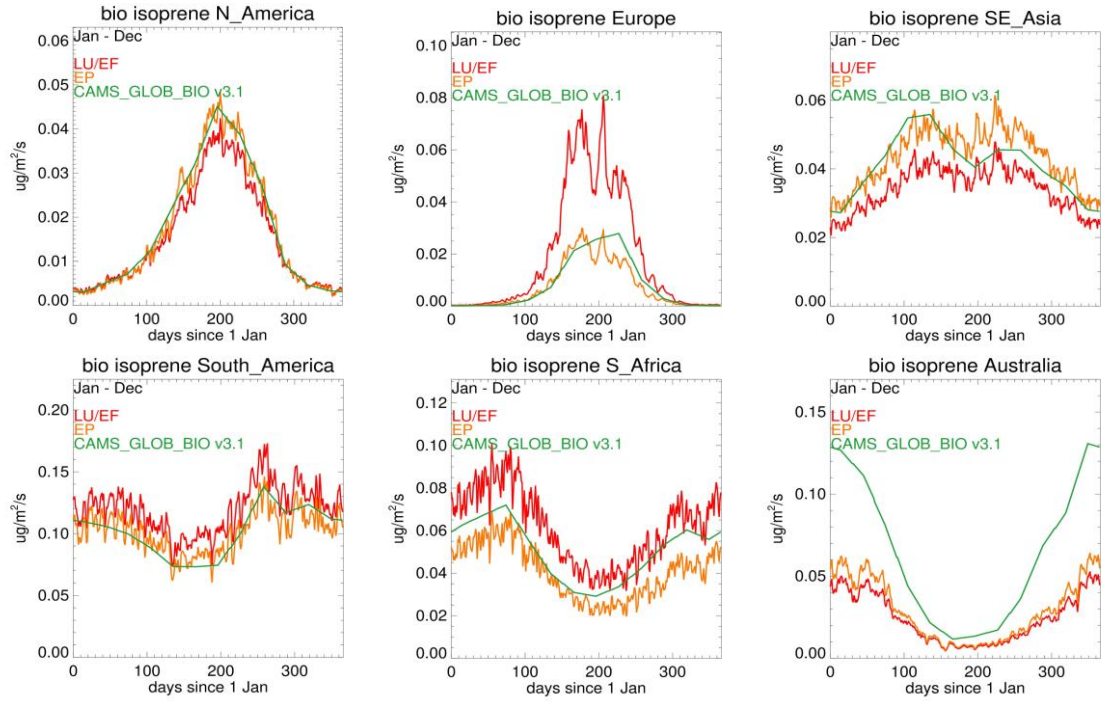

Figure 5.4. Comparison of daily mean isoprene emissions for the year 2019 estimated using fixed emission factor per land-use category (LU/EF, red), and when using an emission potential map (EP, orange), compared to CAMS-GLOB-BIO v3.1 (green), for various regions in the world

5.3 Evaluation of box model simulations against in-situ observations

The Vielsalm flux observations

The Vielsalm eddy-covariance site (50.30°N, 5.90°E) is located in the Belgian Ardennes amid a temperate forest canopy and serves as the reference for VOC flux measurements (Amelynck et al., 2024). This dataset contains volatile organic compound concentrations and fluxes obtained above the canopy of a mixed forest from a flux tower at 52 m a.g.l. at the corresponding ICOS ecosystem site, from May to October 2009 and from March to October 2010, i.e. during the growing season. The dataset contains both concentration measurements and flux data, as described in Laffineur et al. (2011, 2012; 2013). Here we use these observations to assess the performance of our parameterization.

The isoprene fluxes were measured above the canopy by an eddy-covariance system comprising a PTR-MS (Proton-Transfer-Reaction Mass Spectrometer) at ~29 m height and a 3D sonic anemometer. Data were recorded at half-hourly to hourly intervals, then subjected to u* threshold filtering and spike removal to ensure diurnal and daily-mean reliability. The observational flux datafiles used here ('vie_2009_voc_database.csv') includes hourly VOC flux observations in units μg m^{-2} s^{-1}. The data for isoprene fluxes have been extracted to compute both hourly time series and monthly averaged diurnal cycles.

Apart from composition and flux observations, meteorological data is also provided, including T2m (2 m air temperature in K) and SSRD (surface-downward shortwave radiation in J m^{-2} accumulated over 3 h), but these are not used here.

The flux observation data have been used previously to evaluate the MEGAN–MOHYCAN simulations (Bauwens et al., 2018), to assess the model's ability to simulate the diurnal and seasonal emission estimates. Observed daily-mean isoprene emissions were shown to peak near 5 mg $m^{-2}h^{-1}$, while the monthly average isoprene fluxes reach a value of about 1.2 mg $m^{-2}h^{-1}$. In their H3 simulation, adopting a locally observed emission factor of SEF = 2.88 mg $m^{-2}h^{-1}$, the mean bias was -8.5% (r = 0.92) in 2009 and –1.0% (r = 0.91) in 2010, confirming a strong agreement in the diurnal- cycle (Bauwens et al., 2018).

The MEGAN emissions module has been extracted from the IFS-COMPO code, and converted into a python code to allow assessment of the parameterization with local input data. Specifically the model is forced every hour with 2 m air temperature (K) and surface downward shortwave irradiance, converted to photosynthetic photon flux density (PPFD) via a factor of 4.6 μmol J^{-1}, taken from the Copernicus European Regional ReAnalysis (CERRA) dataset, Schimanke et al. (2021).

Within this box, canopy isoprene emissions are calculated online using the unmodified MEGAN v2.1 algorithm, which applies:

- Temperature response factors C_{T1} =80 kJ mol^{-1} and C_{T1} =200 kJ mol^{-1} used in the computation of γ_T
- A top-of-atmosphere PPFD constant value of 3000 μmol $m^{-2}s^{-1}$
- A fixed leaf- area index (LAI) of 3.0 m^2m^{-2}
- the observed EF set at 2.88 mg $m^{-2}h^{-1}$
- Ambient CO_2 set to 410 ppm

We note that the EF used in this analysis is to some extent different from the one used in the IFS-COMPO model simulations, which runs on about 40 km horizontal resolution. When

interpolated to the observation location the EF field used here is only 1.7±0.3 mg $m^{-2}h^{-1}$, which would cause an underestimate in total simulated emissions by about 40% locally. This underscores the importance of using locally observed EF to evaluate the parameterization against observations, and also the critical importance, of realistic (representative) maps of EF in the global model.

We compared modeled against observed isoprene emissions using four complementary diagnostics at Vielsalm, during May – October 2009. The daily time- series (Figure 5.5) demonstrates that the model reproduces synoptic- scale variability in emissions (e.g. week- to- week changes), although extreme peaks (e.g. heat- wave days) are muted. This is confirmed by the scatter plots of daily- mean values (Figure 5.6), which reveals a generally linear relationship but with the model systematically underestimating higher- flux days.

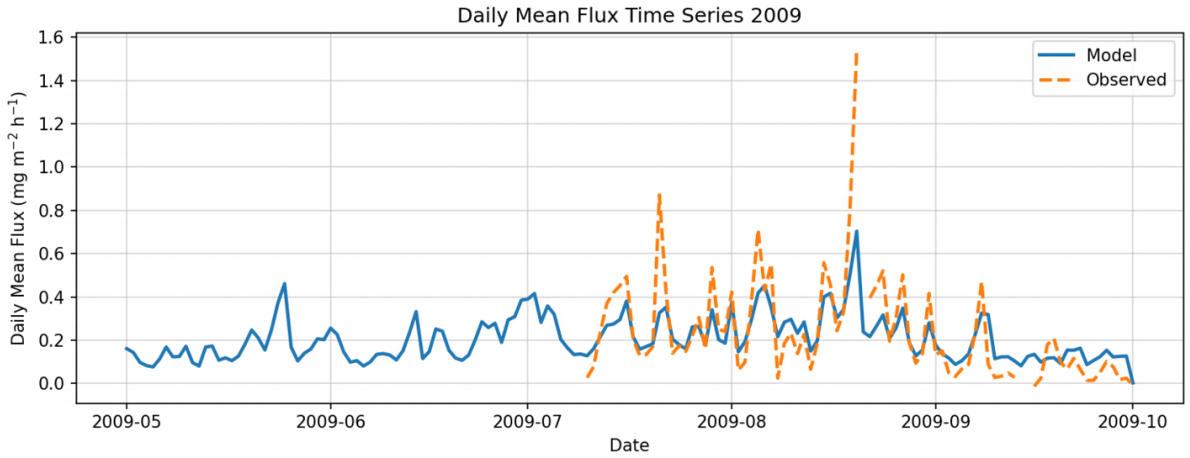

Figure 5.5. Time series of daily mean isoprene flux from May through October 2009. Solid blue line depicts model; dashed orange line depicts observations. Synoptic- scale variability is reproduced, but observed extreme high- emission events exceed model values.

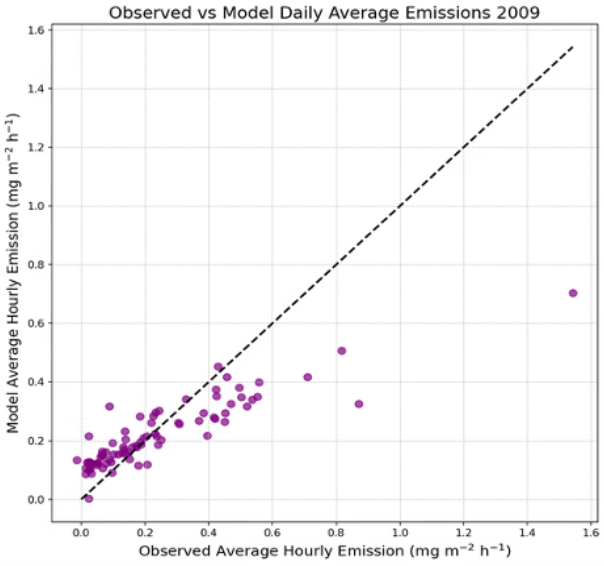

Figure 5.6. Scatter of observed vs. modeled daily average isoprene emissions for July – October, 2009. Each point represents one calendar- day mean.

Monthly averages (Figure 5.7) show that the model captures the seasonal decline from summer into autumn, yet underpredicts the emission magnitudes in July and August by ~10–15 %. Diurnal- cycle panels (Figure 5.8) confirm that the timing of the sunrise- to- noon ramp and afternoon decline is well represented, but that peak amplitudes remain too low, particularly in midsummer. Together, these plots indicate that the default MEGAN v2.1 configuration captures the broad diurnal and seasonal patterns at Vielsalm but underestimates absolute flux levels during high- emission periods.

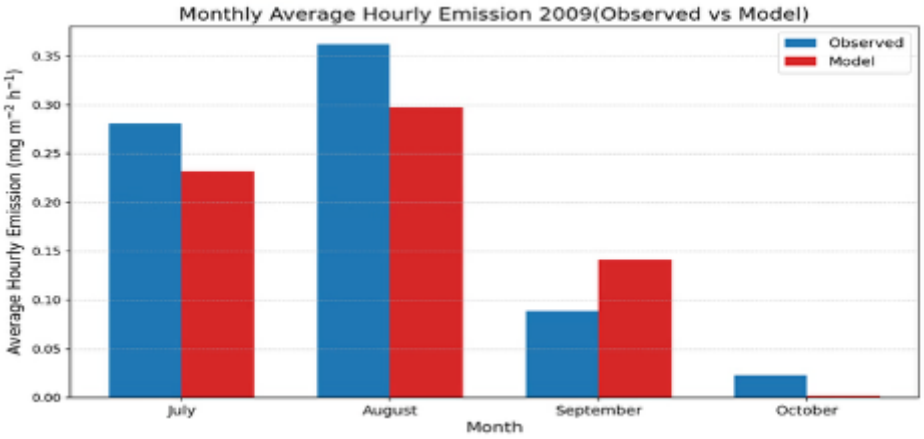

Figure 5.7. Monthly average emissions in 2009, comparing observed (blue bars) and modeled (red bars) values for July–October. Bars show the mean across all hours in each month.

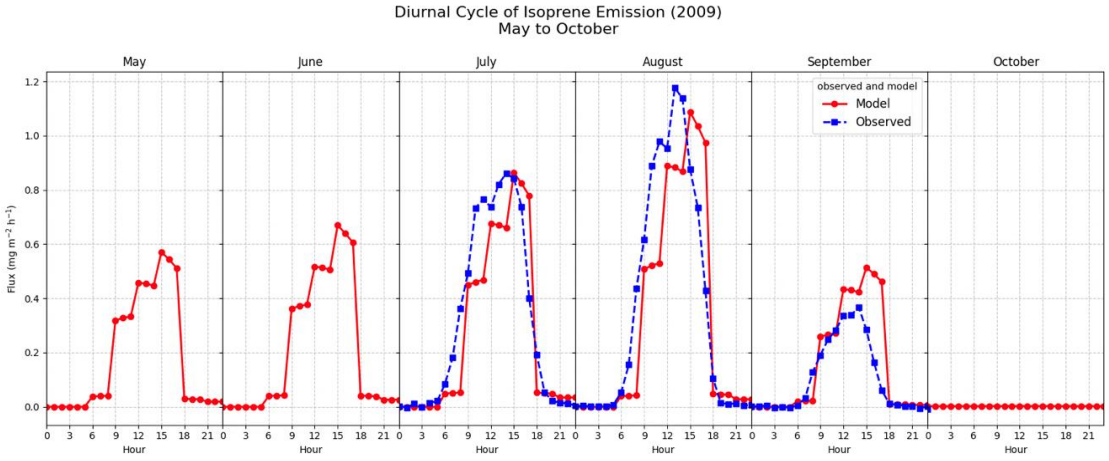

Figure 5.8. Diurnal cycle of isoprene flux (hourly means) for each month from May to October 2009. Red lines/circles are model results; blue dashed lines/squares are observations. While the timing of the daily peak is well captured, the model peak amplitude remains systematically lower.

The simulated results reveal that indeed radiation and temperature drive the bulk of the emission variability. Temperature directly influences the enzymatic processes responsible for BVOC synthesis; even slight temperature changes can exponentially alter isoprene production rates. Similarly, light - quantified as photosynthetically active radiation (PAR) - is a primary driver of photochemical reactions and closely controls the dynamic response of isoprene emissions. A realistic estimation of these factors is critical for the model's ability to reproduce the observed variability. Particularly during midday when both temperature and PAR peak, the

combined sensitivity emphasizes their importance for simulating realistic diurnal emission profiles, and their changes from day-to-day.

Parameter optimization

To assess the sensitivity of the model to the input parameters, and test for methods to better capture the high-emission conditions, we optimized the temperature- dependence factor γ_T, through a fitting procedure. The temperature activity factor is governed by eqn 5.8 above, with the input parameters described as:

$T_{opt}=T_{opt,0}+\beta(T_{daily}-T_0)$ and

$E_{opt}=D_0 \exp(\alpha(T_{daily}-T_0))$

Default values as proposed by Guenther et al. (2006) given in Table 5.1.

The fitting targeted daily- mean isoprene fluxes over July–October 2009 at the Vielsalm site. We performed a grid search, systematically varying:

- Reference optimum temperature ($T_{opt,0}$) from 310 K to 320 K in 1 K increments.
- Exponential amplitude sensitivity (α) from 0.07 K^{-1} to 0.09 K^{-1} in 0.002 K^{-1} increments.

For each parameter pair, the model computed hourly fluxes driven by observed air temperature, soil temperature, and surface solar radiation. Hourly outputs were averaged to daily means to match the observational dataset. The fit objective was to minimize the root-mean- square error (RMSE) defined as:

$$RMSE = \mathrm{sqrt}((1/N)\ \Sigma\ [F_{model} - F_{obs}]^2)$$

where, *N* is the number of days and *F* represents the daily- mean (modeled/ observed) emission flux. This selection emphasizes the diurnal cycle average rather than instantaneous hourly peaks, providing a robust evaluation of model performance at the daily scale.

For observation data for the year 2009 the optimization converged on values for $T_{opt,0}$ = 311.0 K and α = 0.07 K^{-1}, while all other parameters remained at their default values. This configuration achieved an RMSE of 0.14 mg $m^{-2}h^{-1}$, a ~15% improvement over the RMSE when using default values (~0.16 mg $m^{-2}h^{-1}$). By refining the γ_T parameterization, particularly lowering the reference optimum temperature and adjusting the exponential sensitivity, the model more accurately captures the interdaily emission peaks, see Figure 5.9.

Table 5.1 Definitions and default values of parameters in the γ_T formulation.

Parameter	Symbol	Default Value	optimized value (2009)
Reference optimal temperature	$T_{opt,0}$	313.0 K	311.0K
Linear sensitivity of T_{opt}	β	0.60	unchanged
Soil-temperature offset	T_0	297.0 K	unchanged
Exponential amplitude sensitivity	α	0.08 K^{-1}	0.07 K^{-1}
Base amplitude	D_0	1.75	unchanged

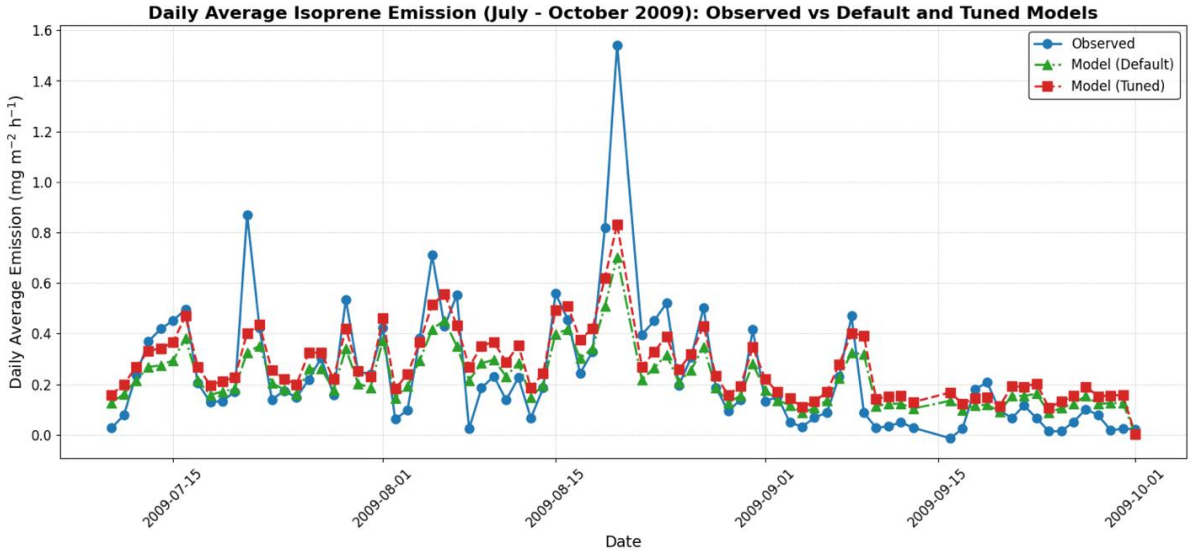

Figure 5.9. Daily- mean isoprene emission fluxes (July–October 2009) at Vielsalm: observed (blue circles), model with default temperature- sensitivity parameters (green triangles, dash- dot), and model with optimized parameters (red squares, dashed).

When adopting the tuned parameters for the 2010 data the emission peaks in mid- July and mid- August are also better reproduced than the default version, Figure 5.10, which are associated with the warmest days of the season. So the optimized temperature- response (γ_T) based on 2009 data remains effective for 2010.

This improved peak fitting is also evident in the scatter plot (Figure 5.11), where daily observed emissions are plotted against modeled values. High- emission days (upper- right cluster) lie noticeably closer to the 1:1 line with the optimized parameters than with the default settings, in particular, for high-emission events. At lower emission values, the optimized values result in slightly larger over-estimates, suggesting the need for an adaptation of the underlying default emission potential. In a follow-up activity the performance statistics, and its changes, will be quantified in more detail.

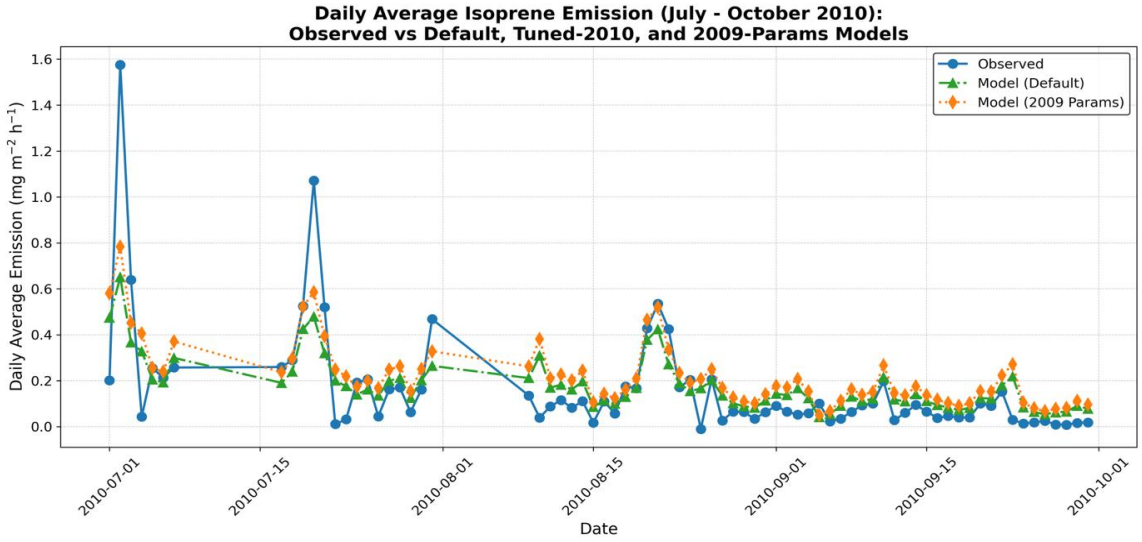

Figure 5.10. Daily average isoprene emission (July–October 2010) time series comparing observed values (blue circles) with the default model version (green triangles) and the optimized parameters based on 2009 data (orange diamonds).

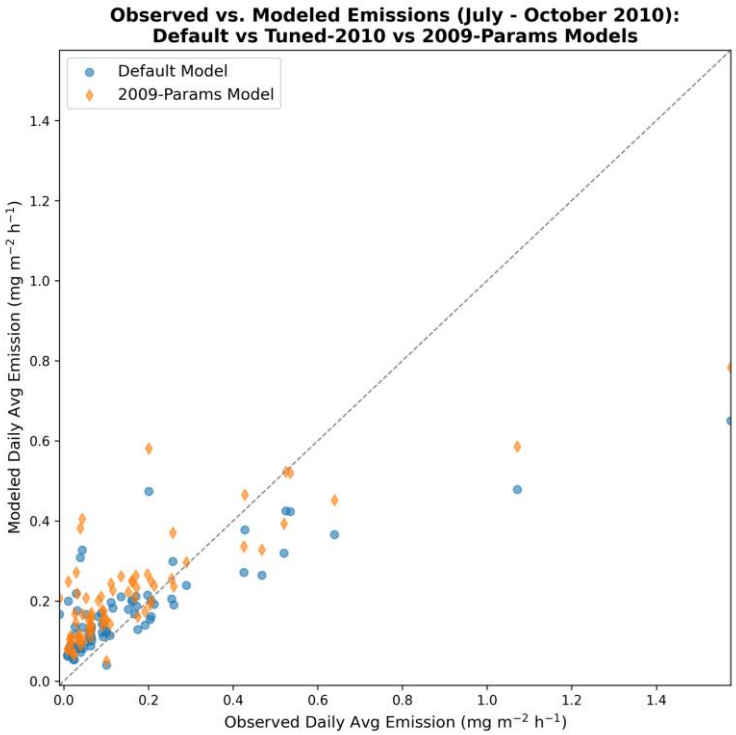

Figure 5.11. Scatter plot of observed vs. modeled daily average isoprene emission (July–October 2010) for the default model (blue circles), and the 2009 parameters applied to 2010 data (orange diamonds).

5.4 Evaluation against satellite observations

We test the use of the emission potential map originating from CAMS-GLOB-BIO v3.1, as compared to the original option, which uses fixed emission factors per land use class, in a set of IFS-COMPO global model simulations for the year 2019, and compare these to satellite observations of isoprene (C_5H_8) and formaldehyde (HCHO). The simulations have been setup to run with tropospheric chemistry only, using a CY49R1 branch that contains many of the composition modeling updates that are planned for CY50R1. Also, the isoprene chemistry has been revisited to include a more realistic recycling of OH, based on an intercomparison of the chemistry mechanism implemented in IFS-COMPO against MAGRITTE, which is planned to be activated in a future cycle as well. The ifs-source branch used for these experiments is nk9_CY49R1.M_COMPO_50R1_R4_C5H8chem.

Anthropogenic emissions are taken from CAMS-GLOB-ANT v6.2, which have additionally been scaled to better match the observed long-term trends in CO, NOx and SO_2 over the NH, US and China.

Two experiments are reported here (see Table 5.2), and are compared against a range of observation data. For this, we select isoprene total column data from the for the Cross-track Infrared Sounder (CrIS) instrument (Wells et al., 2022), complemented with HCHO observations from TROPOMI.

Table 5.2 Definition of experiments evaluated in this report

label	Configuration	expid
LU/EF	old default option, using EF per LU for all BVOC emissions	b2rj
EP	Updated configuration, where for isoprene an EP map is adopted	b2qp

Isoprene (C_5H_8)

We start the evaluation of C_5H_8 columns against a CrIS retrieval dataset, provided by Wells et al (2022). They reported on a machine-learning based isoprene retrieval product from the CrIS instrument at a comparatively high spatial resolution. In this comparison we use the CrIS climatological monthly mean C_5H_8 columns (2012-2020) and compare this to monthly mean columns at 13:30h Local Time (i.e. CrIS overpass time) for our simulations for the year 2019. The data product does not provide information on the averaging kernel, which implies that different profile shapes in our model and in the model (implicitly) used in the retrieval are not accounted for.

Figure 5.12 shows the evaluation for January 2019. While for simulation 'EP' the positive bias over South America and central Africa is reduced, there are still large differences visible on a global scale, suggesting persistent uncertainties that are remaining. For instance, over the Amazon basin, and southeast Asia, the model still shows a positive bias. Also over Yucatan, a region in Mexico known for its special vegetation, the CrIS observations suggest markedly lower C_5H_8 emissions than suggested by either of the model versions. Remarkably, over Australia the model simulations show an average positive bias wrt CrIS C_5H_8.

These evaluations suggest the potential of using CrIS observations to further constrain, and optimize, the prior isoprene emission fluxes on a global scale.

Figure 5.13 shows a similar evaluation, but for July 2019. Again, a marked reduction in positive bias against CrIS is seen over South America when using the EP, while over south-east Asia (mainly Indonesia) the model version with EP shows actually an increase in the positive bias. Over the NH mid-latitudes the C_5H_8 columns increase over the south-eastern part of the US, while they decrease over Eurasia. Both aspects can mostly be considered an improvement with respect to the CrIS observations.

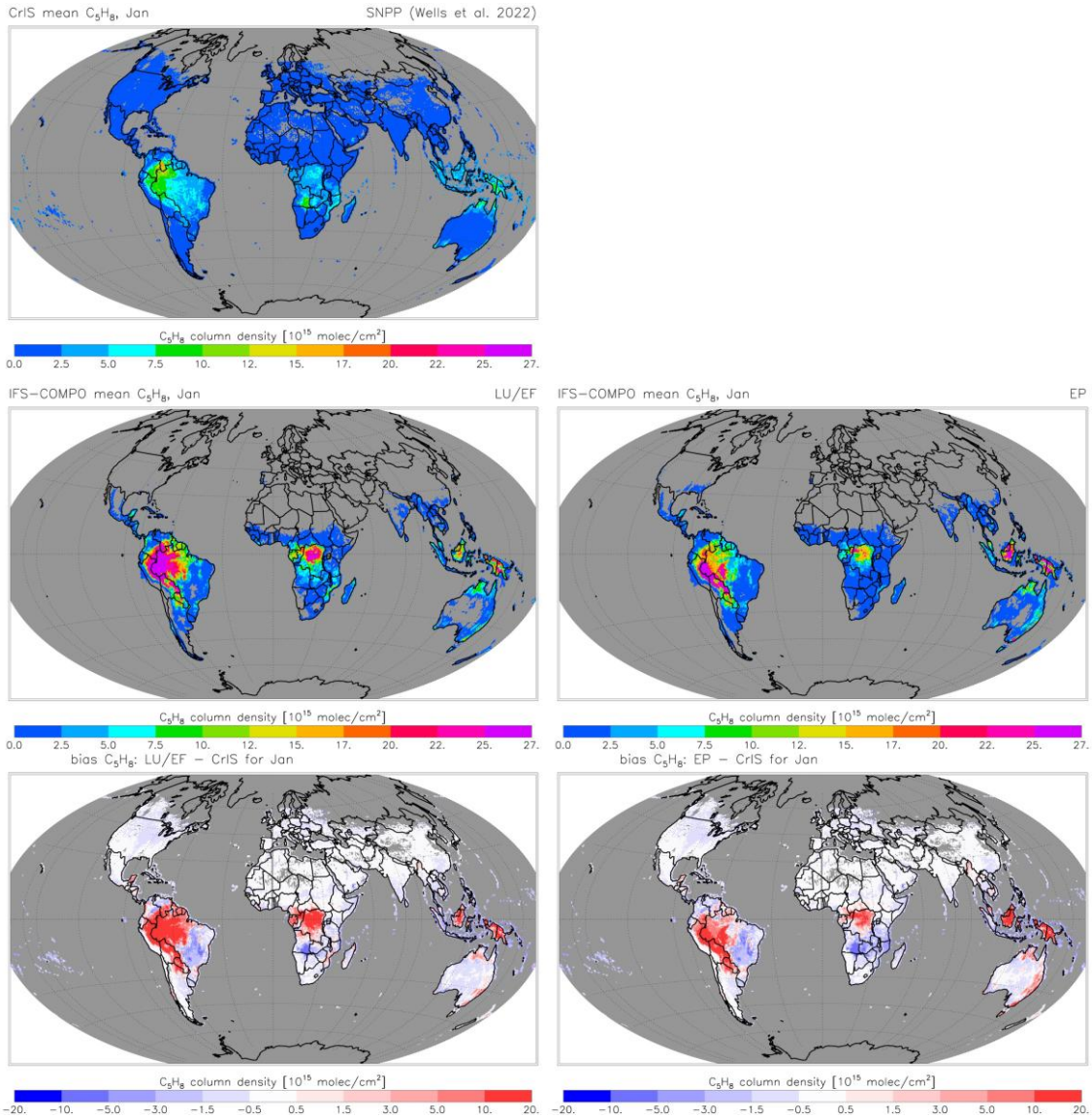

Figure 5.12. Evaluation of monthly mean C_5H_8 simulated columns against CrIS retrievals for January 2019. Top: CrIS climatological values; middle row, left: IFS-COMPO using LU/EF; right: IFS-COMPO using EP. Bottom: model bias against CrIS for the two simulations.

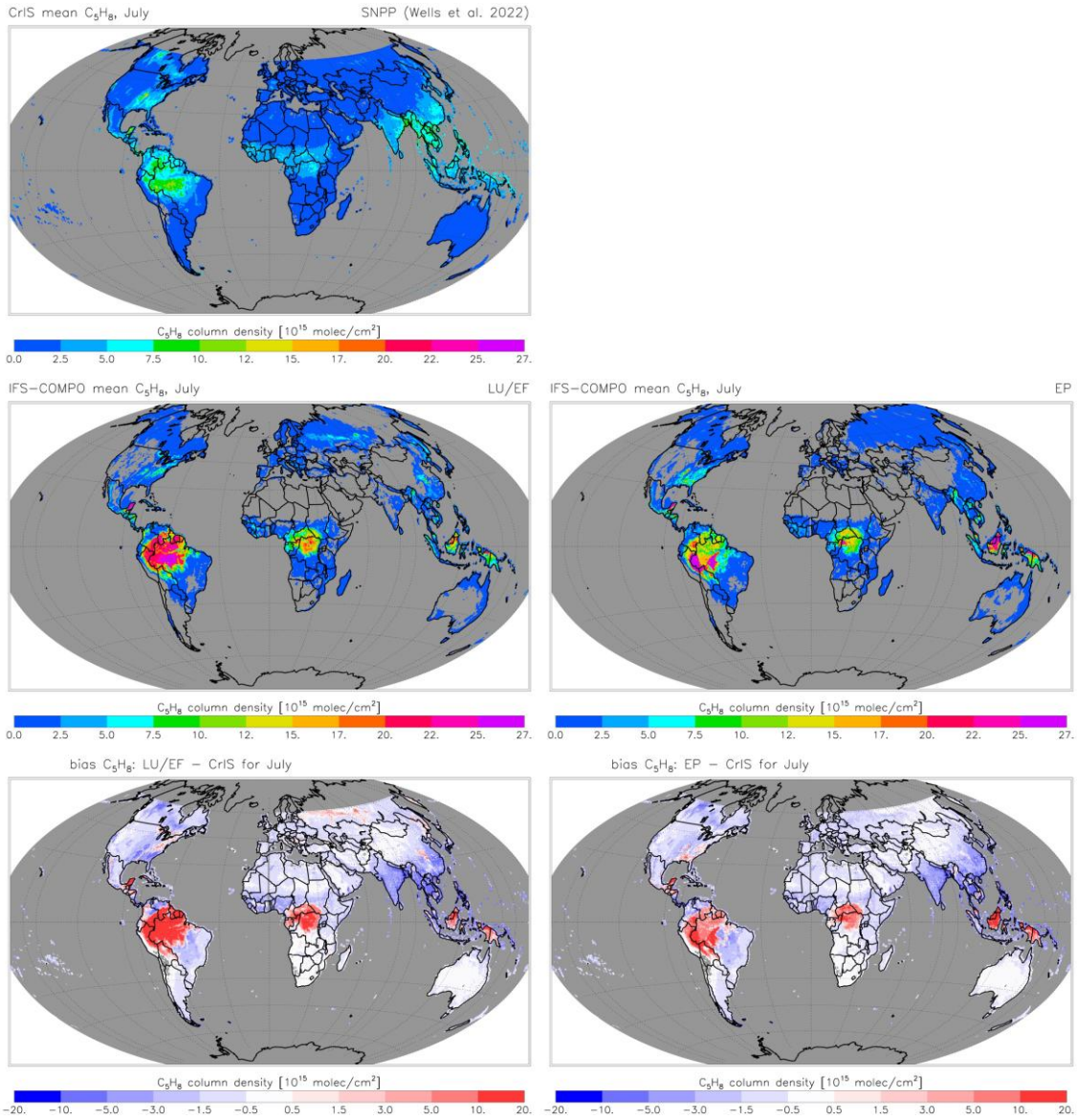

Figure 5.13. Evaluation of monthly mean C5H8 simulated columns against CrIS retrievals for July 2019. Top: CrIS climatological values; middle row, left: IFS-COMPO using LU/EF; right: IFS-COMPO using EP. Bottom: model bias against CrIS for the two simulations.

Formaldehyde (HCHO)

In this study we adopt the re-processed HCHO retrievals (v2.04) from TROPOMI, as reported in Vigouroux et al. (2020). Comparisons of satellite HCHO vertical columns with ground-based data have revealed that low HCHO columns are found to be generally too high whereas high HCHO columns are most often too low (Vigouroux et al. 2020; De Smedt et al. 2021; Zhu et al. 2020). The definite source of this bias is unknown, but it is generally assumed to be related to the retrieval procedure of the satellite HCHO column densities. For this reason we adopt a recommended bias correction prior to comparing the retrieval product to the model HCHO columns, following, for instance, Müller et al. (2024):

$\Omega_{BC} = 1.61 \times \Omega - 1.84 \times 10^{15}$ molec cm^{-2},

in which Ω and Ω_{BC} are the retrieved and bias-corrected TROPOMI HCHO columns in molec cm^{-2} , respectively. The application of this correction leads to strongly increased columns over source regions, and a reduction of the columns over remote, clean regions.

Figures 5.14 and 5.15 show an evaluation of the two IFS-COMPO model simulations against the TROPOMI retrievals for January 2019. Whereas the patterns with elevated HCHO columns are well captured, over the source regions there are considerable differences between the model simulations and TROPOMI. Over southern and western Africa and Australia the model is consistently biased low, while over the Amazon and south-east Asia it is biased high.

The differences between the two approaches for describing the emissions are relatively small, with some reduction in columns over the Amazon in IFS-COMPO experiment EP. This is further quantified in Figure 5.15, which shows a reduction in the mean bias from $1.5x10^{15}$ molec cm^{-2} towards $0.5x10^{15}$ molec cm^{-2} for this region, together with some improvement in the correlation. For Australia the model remains persistently biased low, particularly over the mainland and western part of the continent. This appears in contradiction with the evaluation against CrIS retrievals of C_5H_8, where the model was mostly biased high - but this mostly concerns the area towards the east-coast of Australia.

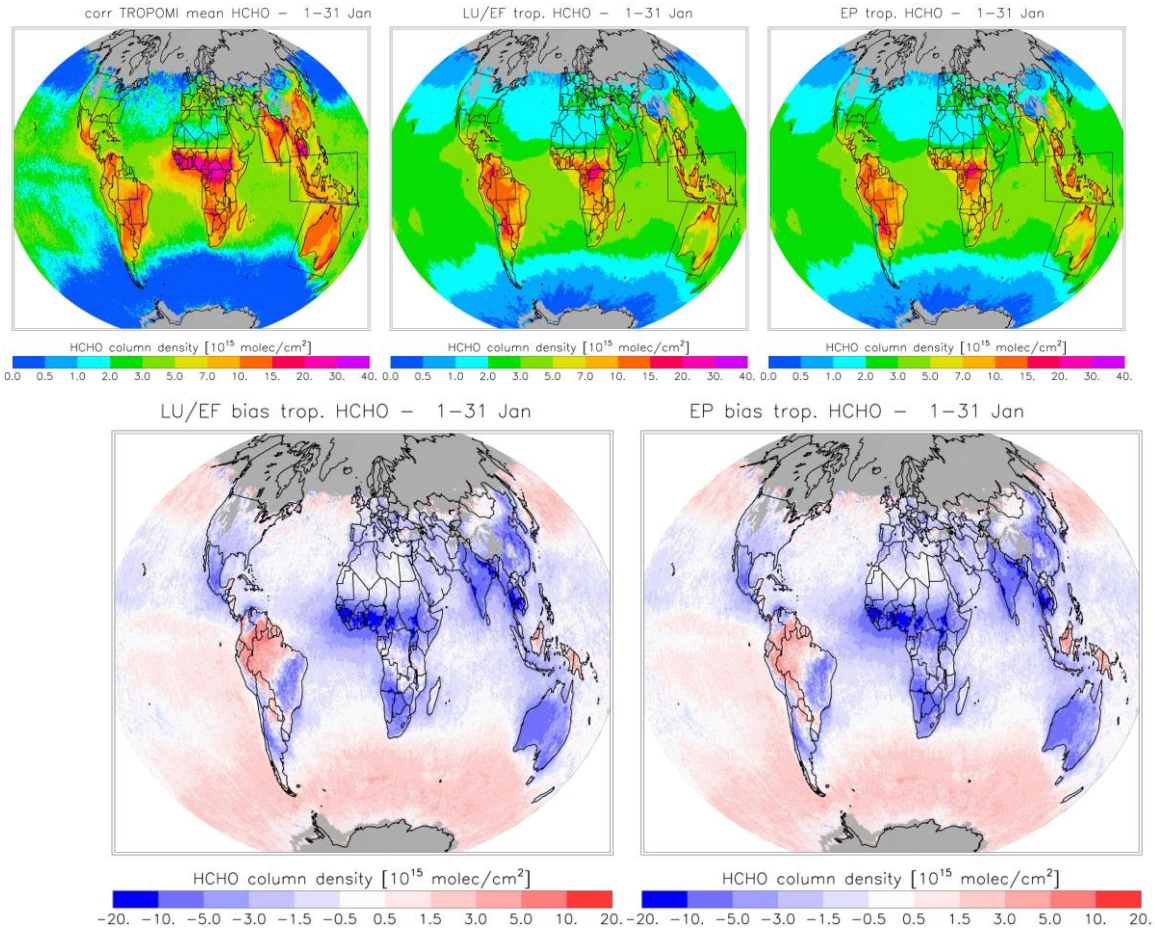

Figure 5.14. Evaluation of monthly mean HCHO simulated columns against bias-corrected TROPOMI HCHO retrievals for January 2019. Top left: TROPOMI HCHO columns; middle: IFS-COMPO using LU/EF; top right: IFS-COMPO using EP. Bottom row: corresponding model bias.

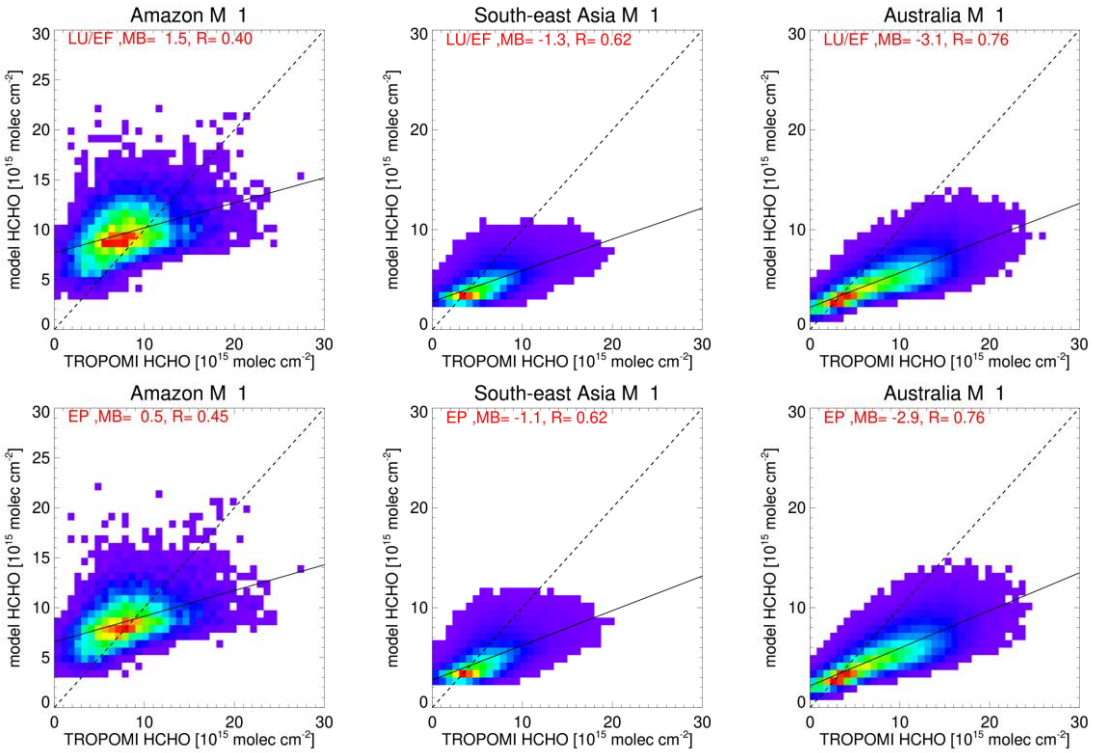

Figure 5.15. Scatter density plots of daily instantaneous HCHO columns simulated with IFS-COMPO, against bias-corrected TROPOMI HCHO retrievals for January 2019. Top row: IFS-COMPO LU/EF; bottom row: IFS-COMPO EP. Three regions are considered here, as indicated in Figure 5.14 above. From left to right: Amazon, south-east Asia and Australia.

Similarly, Figures 5.16 and 5.17 show an evaluation for July 2019, which allows to assess the performance during the NH summer season. Again, over Eastern US, southern Europe and eastern Asia the model is mostly biased low. The differences between the two approaches for describing the emissions are again relatively small. Over the south-eastern part of the US the columns show an increase, resulting in a reduction in the mean bias with respect to the TROPOMI observations. This is confirmed in Figure 5.17, which shows a reduction in the mean bias from -6.2x10^{15} molec cm^{-2} towards -5x10^{15} molec cm^{-2} for this region, although the correlation is not improved. For the other regions no significant change in model performance could be detected.

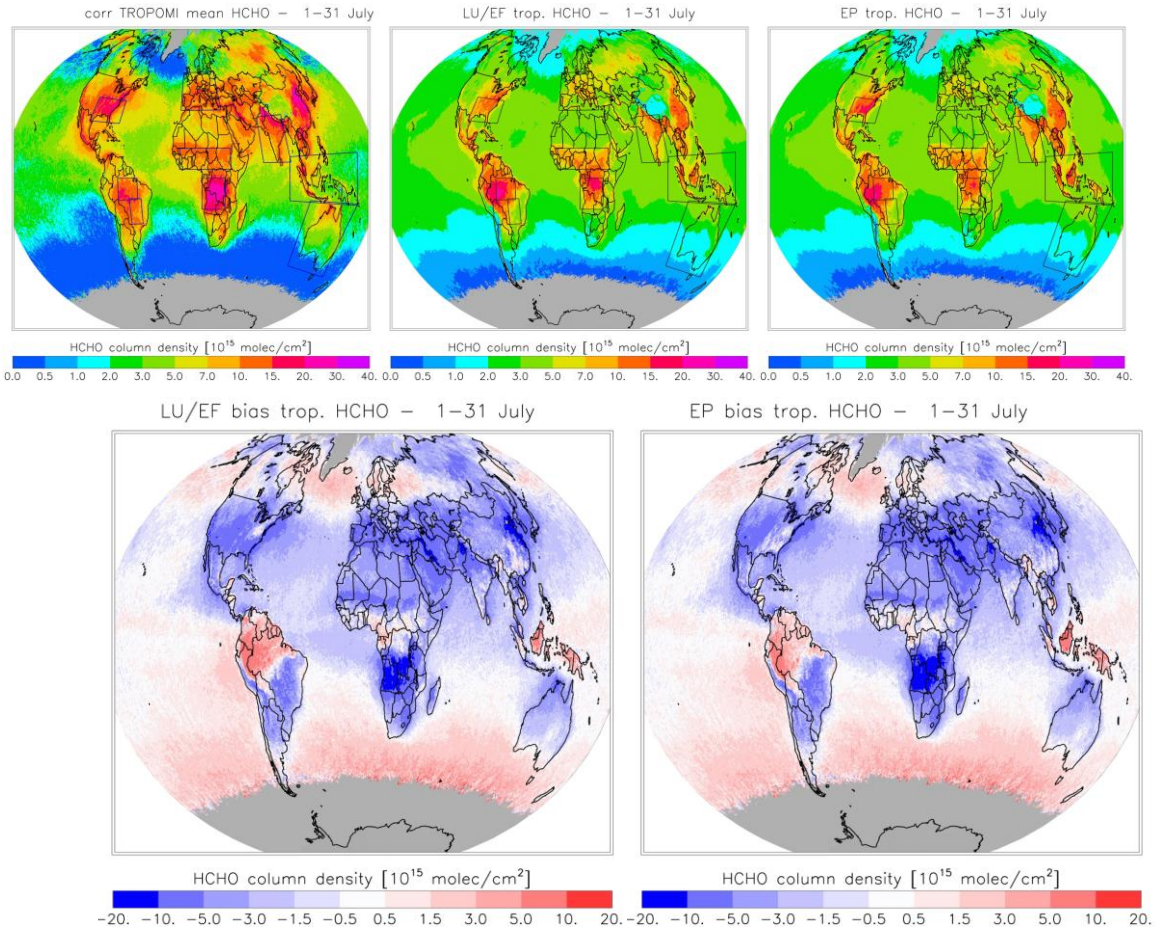

Figure 5.16. Evaluation of monthly mean HCHO simulated columns against bias-corrected TROPOMI HCHO retrievals for July 2019. Top left: TROPOMI HCHO columns; top middle: IFS-COMPO using LU/EF; top right: IFS-COMPO using EP. Bottom row: corresponding model bias.

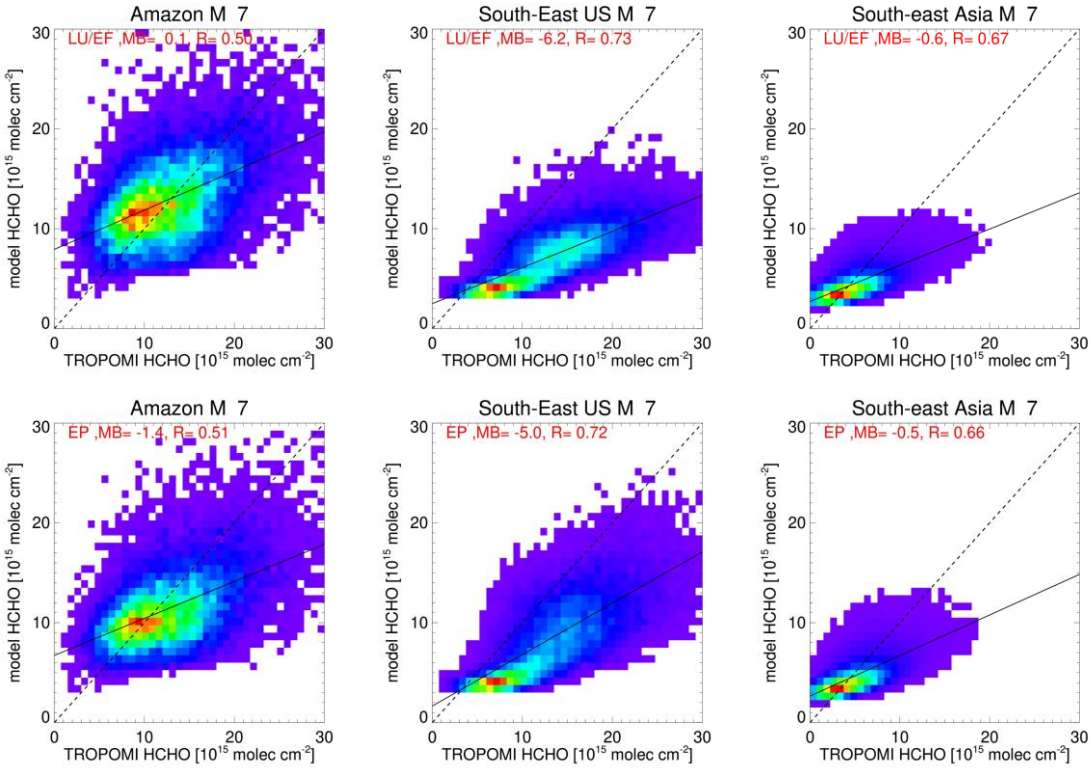

Figure 5.17. Scatter density plots of instantaneous HCHO columns simulated with IFS-COMPO, against bias-corrected TROPOMI HCHO retrievals for July 2019. Top row: IFS-COMPO LU/EF; bottom row: IFS-COMPO EP. Three regions are considered here, as indicated in Figure 5.16 above. From left to right: Amazon, south-east US and south-east Asia.

Impact on other trace gases and aerosol

The impact on air quality aspects, measured by simulated PM2.5, PM10, and ozone during the summer period of 2019, evaluated against air quality networks in the US and Europe, is marginally affected by these changes (not shown here). A review of secondary organic aerosol formation, and a closer evaluation of model parameter choices on the impact on simulated organic aerosol, will be subject of the follow-up activities in this project.

6 Conclusion

In this document we have reported progress from WP7 Task 7.1 of the CAMAERA project.

In Section 4 we have outlined work towards an updated database of BVOC emissions databases, with a focus on the underlying land-cover and LAI data sets, and on emission potentials for some key species. Some conclusions from this activity can be summarised:

- There is a major need for improved land-cover maps, which provide well defined ecosystem and ideally species-level coverage, together with PFT-specific LAI values. There are major problems with these land-cover issues for the purposes of both BVOC emissions and for deposition modelling.

- The current EPs for isoprene and monoterpenes seem reasonable for some key species, but:
    - The EPs used for sesquiterpenes (SQT) likely needs to be increased for some species, for example Norway spruce.
    - The temperature functions used in EMEP EPs for SQT may need modification
    - The EPs for semi-natural vegetation are highly uncertain, also because of the widely differing definitions used in different land-cover datasets.

- We have made a first test on the contribution of BVOC emissions to the total HCHO column. Generally this is below 20% over Europe, which makes it hard to use satellite-based retrievals to provide constraints on the emissions. Nevertheless, new data products, including those from CrIS, will be assessed for their usefulness on the European scale.
- The main product of this work is a python-based software package that has been constructed to modify and improve global landcover data sets for use in BVOC inventories. The system starts with a global land-cover data set, and modifies the PFTs to provide data more appropriate for both BVOC and deposition modelling. The modifications include replacement of: bare land with desert, some semi-natural regions with Mediterranean scrub, some forests with PFTs more appropriate to the Mediterranean.
- This system will certainly be updated in the future, especially if issues with LAI can be resolved, but the current system works and produces new files readable by the EMEP model as intended. Although currently being tested with the ECOSG data sets, the aim is to also enable IFS and other global data to be used in due course.
- Some recommendations for improved handling of these updated BVOC emissions to the IFS modelling can be made. The recommendations include:
    - Use of PFT-specific LAI data as far as possible
    - Increased consideration/evaluation of the LAI data in use.
    - Application of European/“EMEP” emission potentials (EPs) for boreal forest conifers to Eurasian boreal regions (at least Russia) of the outer MEGAN domain, thus ensuring seamless EP maps across the region. This recommendation also follows from the fact that the species dominating this area (Norway spruce, Scots pine) are ubiquitous across much of Eurasia.

It should also be stressed that updating BVOC emission factors is a complex enterprise, and much depends on work done by other groups - for example from the various land-cover

groups. Regarding emission potentials, the MEGANv3 system looks promising as a method of merging EMEP's species-specific EPs within a global system, but so far this system remains undocumented and so hard to evaluate.

As for the global model component discussed in Sec. 5, we have assessed various components which feed into the BVOC emissions model as implemented in ECLand, the land component of IFS. Specifically we have addressed uncertainties related to the emission potential (EP) specification, as well as uncertainties related to the activity factor associated with temperature variations. Finally, we have tested the default, and updated model configuration in IFS-COMPO with respect to an update of the isoprene emission potential, and confronted these simulations against independent satellite observations. We derive the following conclusions:

2 The main driver for temporal variability in the emissions is governed by the temperature activity factor. When using local observations at a site in Vielsalm (Belgium), we find that overall this is well simulated with the parameterization suggested by Guenther et al. (2006), although the peak values associated with high-temperature events are somewhat under-estimated. This could be slightly improved using alternative factors that drive the temperature response parameterisation.
3 The specification of a map that describes the EP for different trace gases is associated with very large uncertainties, which drive the overall uncertainty in BVOC emission estimates. When using EP for isoprene which was taken from CAMS-GLOB-BIO v3.1, then indeed the emission totals and its seasonal cycle can closely follow the one produced by CAMS-GLOB-BIO v3.1. Still a global upscaling factor of 1.5 for the isoprene emissions (not any of the other BVOC's) is needed to match the IFS-COMPO emissions, which is so far not understood.
4 However, when testing this change in EP in an IFS-COMPO simulation, and comparing the results to satellite retrievals for C_5H_8 and HCHO, the results are more mixed, and no obvious improvement can be detected globally. Regionally there are some changes, with improved performance over South America with respect to CrIS C_5H_8, associated with a reduction in performance over Indonesia. The use of TROPOMI HCHO for model evaluation is additionally challenging due to uncertainties in the data product (the need for a bias correction), and also the many contributing factors that govern simulated HCHO, apart from changes in isoprene emissions, as also highlighted above.

A first assessment on the impact of these changes towards secondary organic aerosol (SOA) formation, as well as for trace gases was done, and suggested minor changes so far. One reason for this is that sensitivities in terpenes emissions, which have a larger contribution to SOA, have not been considered so far. This will be the subject of the next phase in the CAMAERA project.

Acknowledgements

Thanks to Hua Yuan (Beijing Normal Univ.), Paul Hamer (NILU), and Katerina Sindelarova (CU) for providing land-cover-related data and useful discussions, and to Samuel Rémy & Roselyne Lacaze (Hygeos), and Simon Munier, Diane Tzanos and Jean-Christophe Calvet (Meteo-France) for help in interpretation of various land-cover and LAI products. We are grateful to Jean-Francois Müller, Jenny Stavrakou and Crist Amelynck (BIRA-IASB) for discussions on the use of satellite and in-situ observations. Also to Alex Guenther (University of California, Irvine) for help in interpreting the MEGAN emission system files.

7 References

Aalto, J., Kolari, P., Hari, P., Kerminen, V.-M., Schiestl-Aalto, P., Aaltonen, H., Levula, J., Siivola, E., Kulmala, M., and Bäck, J.: New foliage growth is a significant, unaccounted source for volatiles in boreal evergreen forests, Biogeosciences, 11, 1331–1344, https://doi.org/10.5194/bg-11-1331-2014, 2014.

Aalto, J., Porcar-Castell, A., Atherton, J., Kolari, P., Pohja, T., Hari, P., Nikinmaa, E., Petäjä, T., and Bäck, J.: Onset of photosynthesis in spring speeds up monoterpene synthesis and leads to emission bursts, Plant, Cell & Environ., 38, 2299–2312, https://doi.org/https://doi.org/10.1111/pce.12550, 2015.

Amelynck, C., Laffineur, Q., Schoon, N., & Heinesch, B. (2024). (O)VOC concentration and flux measurements above a mixed forest at the Vielsalm ICOS ecosystem station (Belgium) (Version 1) [Data set]. Royal Belgian Institute for Space Aeronomy. https://doi.org/10.18758/H659PDRV.

Baret, F., Morissette, J., Fernandes, R., Champeaux, J., Myneni, R., Chen, J., Plummer, S., Weiss, M., Bacour, C., Garrigues, S., and Nickeson, J.: Evaluation of the representativeness of networks of sites for the global validation and intercomparison of land biophysical products: proposition of the CEOS-BELMANIP, IEEE Transactions on Geoscience and Remote Sensing, 44, 1794–1803, https://doi.org/10.1109/TGRS.2006.876030, 2006.

Bauwens, M., Stavrakou, T., Müller, J.-F., Van Schaeybroeck, B., De Cruz, L., De Troch, R., Giot, O., Hamdi, R., Termonia, P., Laffineur, Q., Amelynck, C., Schoon, N., Heinesch, B., Holst, T., Arneth, A., Ceulemans, R., Sanchez-Lorenzo, A., and Guenther, A.: Recent past (1979–2014) and future (2070–2099) isoprene fluxes over Europe simulated with the MEGAN–MOHYCAN model, Biogeosciences, 15, 3673–3690, https://doi.org/10.5194/bg-15-3673-2018, 2018.

Bourtsoukidis, E., Williams, J., Kesselmeier, J., Jacobi, S., and Bonn, B.: From emissions to ambient mixing ratios: online seasonal field measurements of volatile organic compounds over a Norway spruce-dominated forest in central Germany, Atmos. Chem. Physics, 14, 6495–6510, https://doi.org/10.5194/acp-14-6495-2014, 2014.

Boussetta, S., Balsamo, G., Arduini, G., Dutra, E., McNorton, J., Choulga, M., Agustí-Panareda, A., Beljaars, A., Wedi, N., Munõz-Sabater, J., de Rosnay, P., Sandu, I., Hadade, I., Carver, G., Mazzetti, C., Prudhomme, C., Yamazaki, D. and Zsoter, E. Ecland: The ecmwf land surface modelling system. Atmosphere, 12(6), doi:10.3390/atmos12060723, 2021.

Cinderby, S., Emberson, L., Owen, A., and Ashmore, M.: LRTAP land cover map of Europe, in: CCE Progress Report 2007, pp. 59–70, Coordination Centre for Effects, RIVM, Bilthoven, The Netherlands, 2007.

De Smedt, I., Theys, N., Yu, H., Danckaert, T., Lerot, C., Compernolle, S., Van Roozendael, M., Richter, A., Hilboll, A., Peters, E., Pedergnana, M., Loyola, D., Beirle, S., Wagner, T., Eskes, H., van Geffen, J., Boersma, K. F., and Veefkind, P.: Algorithm theoretical baseline for formaldehyde retrievals from S5P TROPOMI and from the QA4ECV project, Atmos. Meas. Tech., 11, 2395–2426, https://doi.org/10.5194/amt-11-2395-2018, 2018.

Denier van der Gon, H., Gauss, M., Granier, C., Arellano, S., Benedictow, A., Darras, S., Dellaert, S., Guevara, M., Jalkanen, J.-P., Krueger, K., Kuenen, J., Liaskoni, M., Liousse, C., Markova, J., Perez, A. P., Quack, B., Simpson, D., Sindelarova, K., and Soulie, A.: Documentation of CAMS emission inventory products, Copernicus Atmosphere Monitoring Service, https://doi.org/10.24380/q2si-ti6i, 2023.

DiMaria, C. A., Jones, D. B. A., Worden, H., Bloom, A. A., Bowman, K., Stavrakou, T., et al. (2023). Optimizing the isoprene emission model MEGAN with satellite and ground-based observational constraints. Journal of Geophysical Research: Atmospheres, 128, e2022JD037822. https://doi.org/10.1029/2022JD037822.

Druel, A., Munier, S., Mucia, A., Albergel, C., and Calvet, J.-C.: Implementation of a new crop phenology and irrigation scheme in the ISBA land surface model using SURFEX_v8.1, Geoscientific Model Dev., 15, 8453–8471, https://doi.org/10.5194/gmd-15-8453-2022, 2022.

Emmerson, K. M., Possell, M., Aspinwall, M. J., Pfautsch, S., and Tjoelker, M. G.: Temperature response measurements from eucalypts give insight into the impact of Australian isoprene emissions on air quality in 2050, Atmos. Chem. Phys., 20, 6193–6206, https://doi.org/10.5194/acp-20-6193-2020, 2020.

ESA. Land Cover CCI Product User Guide Version 2. Tech. Rep. (2017). Available at: maps.elie.ucl.ac.be/CCI/viewer/download/ESACCI-LC-Ph2-PUGv2_2.0.pdf

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

Fu, D., Millet, D. B., Wells, K. C., Payne, V. H., Yu, S., Guenther, A., and Eldering, A.: Direct retrieval of isoprene from satellite-based infrared measurements, Nat. Commun., 10, 3811, https://doi.org/10.1038/s41467-019-11835-0, 2019.

Garrigues, S., Lacaze, R., Baret, F., Morisette, J. T., Weiss, M., Nickeson, J. E., Fernandes, R., Plummer, S., Shabanov, N. V., Myneni, R. B., Knyazikhin, Y., and Yang, W.: Validation and intercomparison of global Leaf Area Index products derived from remote sensing data, J. Geophys. Res., 113, https://doi.org/https://doi.org/10.1029/2007JG000635, 2008.

Geron, C., Guenther, A., and Pierce, T.: An improved model for estimating emissions of volatile organic compounds from forests in the eastern United States, J. Geophys. Res., 99, 12 773–12 792, https://doi.org/10.1029/94JD00246, 1994.

Guenther, A., Zimmerman, P., and Wildermuth, M.: Natural volatile organic compound emission rate estimates for U.S. woodland landscapes, Atmos. Environ., 28, 1197–1210, https://doi.org/https://doi.org/10.1016/1352-2310(94)90297-6, 1994.

Guenther, A., Hewitt, C. N., Erickson, D., Fall, R., Geron, C., Graedel, T., ... & Zimmerman, P. A global model of natural volatile organic compound emissions. *Journal of Geophysical Research: Atmospheres*, *100*(D5), 8873-8892, 1995.

Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P. I., and Geron, C.: Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature), Atmos. Chem. Phys., 6, 3181–3210, https://doi.org/10.5194/acp-6-3181-2006, 2006.

Guenther, A. B., Jiang, X., Heald, C. L., Sakulyanontvittaya, T., Duhl, T., Emmons, L. K., and Wang, X.: The Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGAN2.1): an extended and updated framework for modeling biogenic emissions, Geoscientific Model Dev., 5, 1471–1492, https://doi.org/10.5194/gmd-5-1471-2012, 2012.

Hakola, H., Taipale, D., Praplan, A., Schallhart, S., Thomas, S., Tykkä, T., Helin, A., Bäck, J., and Hellén, H.: Emissions of volatile organic compounds from Norway spruce and potential atmospheric impacts, Frontiers in Forests and Global Change, 6, https://doi.org/10.3389/ffgc.2023.1116414, 2023.

Hantson, S., Knorr, W., Schurgers, G., Pugh, T.A.M., Arneth, A., Global isoprene and monoterpene emissions under changing climate, vegetation, CO2 and land use, Atmospheric Environment, 155, pp35-45, https://doi.org/10.1016/j.atmosenv.2017.02.010 , 2017.

Harper, K. L., Lamarche, C., Hartley, A., Peylin, P., Ottlé, C., Bastrikov, V., San Martín, R., Bohnenstengel, S. I., Kirches, G., Boettcher, M., Shevchuk, R., Brockmann, C., and Defourny, P.: A 29-year time series of annual 300 m resolution plant-functional-type maps for climate models, Earth System Sci. Data., 15, 1465–1499, https://doi.org/10.5194/essd-15-1465-2023, 2023.

International Food Policy Research Institute (IFPRI): Global Spatially-Disaggregated Crop Production Statistics Data for 2020 Version 1.0.0, https://doi.org/10.7910/DVN/SWPENT, Harvard Dataverse, V2, harvard Dataverse, V2, 2024.

Jägermeyr, J., Müller, C., Ruane, A. C., Elliott, J., Balkovic, J., Castillo, O., Faye, B., Foster, I., Folberth, C., Franke, J. A., Fuchs, K., Guarin, J. R., Heinke, J., Hoogenboom, G., Iizumi, T., Jain, A. K., Kelly, D., Khabarov, N., Lange, S., Lin, T.-S., Liu, W., Mialyk, O., Minoli, S., Moyer, E. J., Okada, M., Phillips, M., Porter, C., Rabin, S. S., Scheer, C., Schneider, J. M., Schyns, J. F., Skalsky, R., Smerald, A., Stella, T., Stephens, H., Webber, H., Zabel, F., and Rosenzweig, C.: Climate impacts on global agriculture emerge earlier in new generation of climate and crop models, Nature Food, 2, 873–885, https://doi.org/10.1038/s43016-021-00400-y, 2021.

Jiang, J., Aksoyoglu, S., Ciarelli, G., Oikonomakis, E., El-Haddad, I., Canonaco, F., O'Dowd, C., Ovadnevaite, J., Minguillón, M. C., Baltensperger, U., and Prévôt, A. S. H.: Effects of two different biogenic emission models on modelled ozone and aerosol concentrations in Europe, Atmos. Chem. Physics, 19, 3747–3768, https://doi.org/10.5194/acp-19-3747-2019, 2019.

Karl, M., Guenther, A., KFIXble, R., Leip, A., and Seufert, G.: A new European plant-specific emission inventory of biogenic volatile organic compounds for use in atmospheric transport models, Biogeosciences, 6, 1059–1087, http://www.biogeosciences.net/6/1059/2009/, 2009.

Keenan, T., Niinemets, U., Sabate, S., Gracia, C., and Penuelas, J.: Process based inventory of isoprenoid emissions from European forests: model comparisons, current knowledge and uncertainties, Atmos. Chem. Physics, 9, 4053–4076, 2009.

Kramshøj, M., Vedel-Petersen, I., Schollert, M., Rinnan, Å., Nymand, J., Ro-Poulsen, H., and Rinnan, R.: Large increases in Arctic biogenic volatile emissions are a direct effect of warming, Nature Geoscience, 9, 349–352, https://doi.org/10.1038/ngeo2692, 2016.

Huijnen, V., Sindelarova, K., Remy, S.: CAMS2_35 D2.3.1: Implementation of online biogenic emissions in the IFS. CAMS2_35 Deliverable report D2.3.1, 2023.

Köble, R. and Seufert, G.: Novel Maps for Forest Tree Species in Europe., in: A Changing Atmosphere, 8th European Symposium on the Physico-Chemical Behaviour of Atmospheric Pollutants, Torino, Italy, 17-20 Sept., 2001., http://ies.jrc.ec.europa.eu/Units/cc/events/torino2001/torinocd/Documents/Terrestrial/TP35.htm, 2001.

Laffineur, Q., Aubinet, M., Schoon, N., Amelynck, C., Müller, J. F., Dewulf, J., ... & Heinesch, B. (2011). Isoprene and monoterpene emissions from a mixed temperate forest. Atmospheric Environment, 45(18), 3157-3168.

Laffineur, Q., Aubinet, M., Schoon, N., Amelynck, C., Müller, J. F., Dewulf, J., ... & Heinesch, B. (2012). Abiotic and biotic control of methanol exchanges in a temperate mixed forest. Atmospheric Chemistry and Physics, 12(1), 577-590.

Laffineur, Q., Aubinet, M., Schoon, N., Amelynck, C., Müller, J. F., Dewulf, J., ... & Heinesch, B. (2013). Impact of diffuse light on isoprene and monoterpene emissions from a mixed temperate forest. Atmospheric Environment, 74, 385-392.

Langford, B., Cash, J., Acton, W. J. F., Valach, A. C., Hewitt, C. N., Fares, S., Goded, I., Gruening, C., House, E., Kalogridis, A.-C., Gros, V., Schafers, R., Thomas, R., Broadmeadow, M., and Nemitz, E.: Isoprene emission potentials from European oak forests derived from canopy flux measurements: an assessment of uncertainties and inter-algorithm variability, Biogeosciences, 14, 5571–5594, https://doi.org/10.5194/bg-14-5571-2017, 2017.

Langford, B., Cash, J. M., Vieno, M., Heal, M. R., Drewer, J., Jones, M. R., Leeson, S. R., Simmons, I., Braban, C. F., and Nemitz, E.: Evaluation of isoprene light response curves for bryophyte-dominated ecosystems and implications for atmospheric composition, Environmental Research: Ecology, 2, 011 002, https://doi.org/10.1088/2752-664X/aca2ad, 2022.

Langner, J., Engardt, M., Baklanov, A., Christensen, J. H., Gauss, M., Geels, C., Hedegaard, G. B., Nuterman, R., Simpson, D., Soares, J., Sofiev, M., Wind, P., and Zakey, A.: A multi-model study of impacts of climate change on surface ozone in Europe, Atmos. Chem. Physics, 12, 10 423–10 440, https://doi.org/10.5194/acp-12-10423-2012, 2012.

Lawrence, D. M., Oleson, K. W., Flanner, M. G., Thornton, P. E., Swenson, S. C., Lawrence, P. J., Zeng, X., Yang, Z.-L., Levis, S., Sakaguchi, K., Bonan, G. B., and Slater, A. G.: Parameterization Improvements and Functional and Structural Advances in Version 4 of the Community Land Model, J. Adv. Modeling Earth Systems, 3, https://doi.org/10.1029/2011MS000045, 2011.

Messina, P., Lathiere, J., Sindelarova, K., Vuichard, N., Granier, C., Ghattas, J., Cozic, A., and Hauglustaine, D. A.: Global biogenic volatile organic compound emissions in the ORCHIDEE and MEGAN models and sensitivity to key parameters, Atmos. Chem. Physics, 16, 14 169–14 202, https://doi.org/10.5194/acp-16-14169-2016, 2016.

Müller, J.-F., Stavrakou, T., Oomen, G.-M., Opacka, B., De Smedt, I., Guenther, A., Vigouroux, C., Langerock, B., Aquino, C. A. B., Grutter, M., Hannigan, J., Hase, F., Kivi, R., Lutsch, E., Mahieu, E., Makarova, M., Metzger, J.-M., Morino, I., Murata, I., Nagahama, T., Notholt, J., Ortega, I., Palm, M., Röhling, A., Stremme, W., Strong, K., Sussmann, R., Té, Y., and Fried, A.: Bias correction of OMI HCHO columns based on FTIR and aircraft measurements and impact on top-down emission estimates, Atmos. Chem. Phys., 24, 2207–2237, https://doi.org/10.5194/acp-24-2207-2024, 2024.

Munier, S., Carrer, D., Planque, C., Camacho, F., Albergel, C., and Calvet, J.-C.: Satellite Leaf Area Index: Global Scale Analysis of the Tendencies Per Vegetation Type Over the Last 17 Years, Remote Sensing, 10, https://doi.org/10.3390/rs10030424, 2018.

Oleson, K., Lawrence, D., Bonan, G., Flanner, M., Kluzek, E., Lawrence, P., Levis, S., Swenson, S., Thornton, P., Dai, A., Decker, M., Dickinson, R., Feddema, J., Heald, C., Hoffman, F., Lamarque, J., Mahowald, N., Niu, G., Qian, T., Randerson, J., Running, S., Sakaguchi, K., Slater, A., Stockli, R., Wang, A., Yang, Z., Zeng, X., and Zeng, X.: Technical Description of version 4.0 of the Community Land Model (CLM), NCAR Technical Note NCAR/TN-478+STR, National Center for Atmospheric Research, Boulder, CO, https://doi.org/10.5065/D6FB50WZ, iSSN Electronic Edition 2153-2400, 2010.

Rantala, P., Aalto, J., Taipale, R., Ruuskanen, T. M., and Rinne, J.: Annual cycle of volatile organic compound exchange between a boreal pine forest and the atmosphere, Biogeosciences, 12, 5753–5770, https://doi.org/10.5194/bg-12-5753-2015, 2015.

Rinnan, R., Gierth, D., Bilde, M., Rosenørn, T., and Michelsen, A.: Off-season biogenic volatile organic compound emissions from heath mesocosms: responses to vegetation cutting, Frontiers in Microbiology, 4, 224, https://doi.org/10.3389/fmicb.2013.00224, 2013.

Rinnan, R., Iversen, L. L., Tang, J., Vedel-Petersen, I., Schollert, M., and Schurgers, G.: Separating direct and indirect effects of rising temperatures on biogenic volatile emissions in the Arctic, Proceedings of the National Academy of Sciences, 117, 32 476–32 483, https://doi.org/10.1073/pnas.2008901117, 2020.

Rinne, J., Back, J., and Hakola, H.: Biogenic volatile organic compound emissions from the Eurasian taiga: current knowledge and future directions, Boreal Env. Res., 14, 807–826, 2009.

Schimanke S., Ridal M., Le Moigne P., Berggren L., Undén P., Randriamampianina R., Andrea U., Bazile E., Bertelsen A., Brousseau P., Dahlgren P., Edvinsson L., El Said A., Glinton M., Hopsch S., Isaksson L., Mladek R., Olsson E., Verrelle A., Wang Z.Q., (2021): CERRA sub-daily regional reanalysis data for Europe on single levels from 1984 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS), DOI: 10.24381/cds.622a565a (Accessed on 01-05-2025)

Schurgers, G., Hickler, T., Miller, P. A., and Arneth, A.: European emissions of isoprene and monoterpenes from the Last Glacial Maximum to present, Biogeosciences, 6, 2779–2797, https://doi.org/10.5194/bg-6-2779-2009, 2009.

Simin, T., Tang, J., Holst, T., and Rinnan, R.: Volatile organic compound emission in tundra shrubs – Dependence on species characteristics and the near-surface environment, Environmental and Experimental Botany, 184, 104 387, https://doi.org/https://doi.org/10.1016/j.envexpbot.2021.104387, 2021.

Simpson, D., Guenther, A., Hewitt, C., and Steinbrecher, R.: Biogenic emissions in Europe 1. Estimates and uncertainties, J. Geophys. Res., 100, 22 875–22 890, 1995.

Simpson, D., Winiwarter, W., Börjesson, G., Cinderby, S., Ferreiro, A., Guenther, A., Hewitt, C. N., Janson, R., Khalil, M. A. K., Owen, S., Pierce, T. E., Puxbaum, H., Shearer, M., Skiba, U., Steinbrecher, R., Tarrasón, L., and Öquist, M. G.: Inventorying emissions from Nature in Europe, J. Geophys. Res., 104, 8113–8152, 1999.

Simpson, D., Benedictow, A., Berge, H., Bergström, R., Emberson, L. D., Fagerli, H., Flechard, C. R., Hayman, G. D., Gauss, M., Jonson, M. E., Jenkin, M. E., Nyíri, A., Richter, C., Semeena, V. S., Tsyro, S., Tuovinen, J.-P., Valdebenito, A., and Wind, P.: The EMEP MSC-W chemical transport model – technical description, Atmos. Chem. Physics, 12, 7825–7865, https://doi.org/10.5194/acp-12-7825-2012, 2012.

Sindelarova, K., Markova, J., Simpson, D., Huszar, P., Karlicky, J., Darras, S., and Granier, C.: High-resolution biogenic global emission inventory for the time period 2000–2019 for air quality modelling, Earth Syst. Sci. Data, 14, 251–270, https://doi.org/10.5194/essd-14-251-2022, 2022.

Smith, B., Prentice, I., and Sykes, M.: Representation of vegetation dynamics in the modelling of terrestrial ecosystems: comparing two contrasting approaches within European climate space, Glob. Ecol. Biogeogr., 10, 621–637, 2001.

Staudt, M., Bourgeois, I., Al Halabi, R., Song, W., and Williams, J.: New insights into the parametrization of temperature and light responses of mono - and sesquiterpene emissions from Aleppo pine and rosemary, Atmos. Environ., 152, 212–221, https://doi.org/https://doi.org/10.1016/j.atmosenv.2016.12.033, 2017.

Taipale, D., Aalto, J., Schiestl-Aalto, P., Kulmala, M., and Bäck, J.: The importance of accounting for enhanced emissions of monoterpenes from new Scots pine foliage in models - A Finnish case study, Atmos. Environ., 8, 100 097, https://doi.org/https://doi.org/10.1016/j.aeaoa.2020.100097, 2020.

Tarvainen, V., Hakola, H., Rinne, J., Hellen, H., and Haapanala, S.: Towards a comprehensive emission inventory of terpenoids from boreal ecosystems, Tellus, 59, 526–534, https://doi.org/10.1111/j.1600-0889.2007.00263.x, 2007.

UNECE: Mapping Critical Levels for Vegetation, in: Manual for modelling and mapping critical loads and levels, chap. 3, ICP vegetation, CEH, UK, icpvegetation.ceh.ac.uk/get-involved/manuals/mapping-manual, 2017.

Vedel-Petersen, I., Schollert, M., Nymand, J., and Rinnan, R.: Volatile organic compound emission profiles of four common arctic plants, Atmos. Environ., 120, 117 – 126, https://doi.org/https://doi.org/10.1016/j.atmosenv.2015.08.082, 2015.

Veldt, C.: Leaf biomass data for the estimation of biogenic VOC emissions, apeldoorn, The Netherlands, MT-TNO Report 89-306, 1989.

Verger, A., Sánchez-Zapero, J., Weiss, M., Descals, A., Camacho, F., Lacaze, R., and Baret, F.: GEOV2: Improved smoothed and gap filled time series of LAI, FAPAR and FCover 1 km Copernicus Global Land products, International Journal of Applied Earth Observation and Geoinformation, 123, 103 479, https://doi.org/10.1016/j.jag.2023.103479, 2023.

Vigouroux, C., Langerock, B., Bauer Aquino, C. A., Blumenstock, et al.: TROPOMI Sentinel-5 Precursor formaldehyde validation using an extensive network of ground-based Fourier-transform infrared stations, Atmos. Meas. Tech., 13, 3751–3767, https://doi.org/10.5194/amt-13-3751-2020, 2020

Warneke, C., de Gouw, J. A., Del Negro, L., Brioude, J., McKeen, S., Stark, H., Kuster, W. C., Goldan, P. D., Trainer, M., Fehsenfeld, F. C., Wiedinmyer, C., Guenther, A. B., Hansel, A., Wisthaler, A., Atlas, E., Holloway, J. S., Ryerson, T. B., Peischl, J., Huey, L. G., and Hanks, A. T. C.: Biogenic emission measurement and inventories determination of biogenic emissions in the eastern United States and Texas and comparison with biogenic emission inventories, J. Geophys. Res., 115, https://doi.org/10.1029/2009JD012445, 2010.

Wang, P., Schade, G., Estes, M., and Ying, Q.: Improved MEGAN predictions of biogenic isoprene in the contiguous United States, Atmospheric Environment, 148, 337–351, https://doi.org/https://doi.org/10.1016/j.atmosenv.2016.11.006, 2017.

Wells, K. C., Millet, D. B., Payne, V. H., Vigouroux, C., Aquino, C. A. B., De Mazière, M., et al. (2022). Next-generation isoprene measurements from space: Detecting daily variability at high resolution. Journal of Geophysical Research: Atmospheres, 127, e2021JD036181. https://doi.org/10.1029/2021JD036181

Yu, Q., You, L., Wood-Sichra, U., Ru, Y., Joglekar, A. K. B., Fritz, S., Xiong, W., Lu, M., Wu, W., and Yang, P.: A cultivated planet in 2010 – Part 2: The global gridded agricultural-production maps, Earth System Science Data, 12, 3545–3572, https://doi.org/10.5194/essd-12-3545-2020, 2020.

Yuan, H., Dai, Y., Xiao, Z., Ji, D., and Shangguan, W.: Reprocessing the MODIS Leaf Area Index products for land surface and climate modelling, Remote Sens. Environ., 115, 1171–1187, 2011.

Zhu, L., González Abad, G., Nowlan, C. R., Chan Miller, C., Chance, K., Apel, E. C., DiGangi, J. P., Fried, A., Hanisco, T. F., Hornbrook, R. S., Hu, L., Kaiser, J., Keutsch, F. N., Permar, W., St. Clair, J. M., and Wolfe, G. M.: Validation of satellite formaldehyde (HCHO) retrievals using observations from 12 aircraft campaigns, Atmos. Chem. Phys., 20, 12329–12345, https://doi.org/10.5194/acp-20-12329-2020, 2020.

Document History

Version	Author(s)	Date	Changes
1.0	D. Simpson, J. Majumdar, V. Huijnen	17-06-2025	Initial version
1.1	D. Simpson, J. Majumdar, V. Huijnen	24-06-2025	Revised version

Internal Review History

Internal Reviewers	Date	Comments
Jesper Christensen	23-06-2025	OK from me with few comments and suggestions
V. Guidard	23-06-2025	