



CoCo2

Prototype system for a  
Copernicus CO<sub>2</sub> service

# Validation of online global vegetation carbon fluxes in ECLand for prototype modelling developments

Emanuel Dutra, Francisco Lopes, Jean-Christophe Calvet,

Bertrand Bonan, Anna Agusti-Panareda, Gianpaolo Balsamo,

Souhail Boussetta and Martin Jung



Co-ordinated by  
 ECMWF





# CoCO2

Prototype system for a  
Copernicus CO<sub>2</sub> service

## D3.3 Validation of online global vegetation carbon fluxes in ECLand for prototype modelling developments

**Dissemination Level:** Public

**Author(s):** Emanuel Dutra & Francisco Lopes (FC.ID);  
Jean-Christophe Calvet & Bertrand Bonan (Meteo-France);  
Anna Augusti-Panareda, Souhail Boussetta & Gianpaolo Balsamo  
(ECMWF); Martin Jung (MPG)

**Date:** 05/07/2022

**Version:** 2.0

**Contractual Delivery Date:** 30/06/2022

**Work Package/ Task:** WP3/ T3.3a

**Document Owner:** FC.ID

**Contributors:** ECMWF/Meteo-France/MPG

**Status:** Final



# CoCO2: Prototype system for a Copernicus CO<sub>2</sub> service

Coordination and Support Action (CSA)  
H2020-IBA-SPACE-CHE2-2019 Copernicus evolution –  
Research activities in support of a European operational  
monitoring support capacity for fossil CO<sub>2</sub> emissions

**Project Coordinator:** Dr Richard Engelen (ECMWF)  
**Project Start Date:** 01/01/2021  
**Project Duration:** 36 months

**Published by the CoCO2 Consortium**

**Contact:**  
ECMWF, Shinfield Park, Reading, RG2 9AX,  
[richard.engelen@ecmwf.int](mailto:richard.engelen@ecmwf.int)



The CoCO2 project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958927.



# Table of Contents

1	Executive Summary .....	7
2	Introduction .....	8
2.1	Background.....	8
2.2	Scope of this deliverable .....	9
2.2.1	Objectives of this deliverable .....	9
2.2.2	Work performed in this deliverable .....	9
2.2.3	Deviations and counter measures .....	10
3	Processing of high-resolution land cover and Leaf Area Index .....	10
3.1	Data description.....	10
3.1.1	ESA-CCI land cover and Copernicus Global Land Service Leaf Area Index ..	10
3.2	Data processing .....	10
3.2.1	Land cover.....	10
3.2.2	Leaf Area index.....	11
3.2.3	Dominant c3/c4 and climate classification.....	12
4	Offline simulations and evaluation of biogenic fluxes .....	13
4.1	Data and methods.....	13
4.1.1	Model simulations .....	13
4.1.2	Land cover and LAI changes in ECLand.....	15
4.1.3	FLUXCOM and evaluation metrics.....	17
4.1.4	Satellite land surface temperature.....	18
4.2	Results.....	19
4.2.1	Surface water/energy/carbon fluxes evaluation.....	19
4.2.2	Land Surface Temperature .....	23
4.3	Discussion .....	24
5	Online simulations and evaluation of Numerical Weather Prediction metrics.....	26
5.1	Data and Methods.....	26
5.1.1	Observations.....	26
5.1.2	Model simulations .....	27
5.2	Results.....	28
5.2.1	Meteorological impact .....	28
5.2.2	Impact on near-surface temperature .....	32
5.3	Discussion .....	37
6	Conclusion .....	38
7	Acknowledgements.....	39
8	References .....	39
9	Appendix I: Supplementary Figures .....	44

## Figures

Figure 1: Dominant Köppen-Geiger climate classification. ....	12
Figure 2: Dominant mapping of the distribution of C <sub>3</sub> and C <sub>4</sub> for grass and crops plant types. ....	13
Figure 3: Dominant high vegetation (a, b) and low vegetation (c, d) in the default (CTR) ECLand configuration (a, c) and revised land cover based on ESA-CCI (b, d). ....	16
Figure 4: Differences between CLIM and control (CTR) simulations effective low vegetation cover (a), effective high vegetation cover (b), and effective bare ground cover (c). ....	16
Figure 5: Mean annual cycle of low vegetation LAI (a, d), high vegetation LAI (b, e), and total LAI (c, f), for two regions: (a-c) East North America [30° to 50° N, 90° to 70° W]; (d-f) South Africa [20° to 10° S, 12° to 34° E]. These figures compare the ECLand default LAI as in CTR (black curves) with the new revised CLIM LAI (solid red lines, v0). The total original LAI (dotted magenta in c, f panels), the intermediate LAI aggregation (eq. 7) (red dotted red lines), and an independent dataset based on MODIS data (blue lines, v0M) are also included. ....	17
Figure 6: Overall score $S_{all}$ for the 6 considered variables (horizontal axis) shown for each simulation (bars) considering (a) FLUXCOM RS as reference and (b) FLUXCOM RS_METEO as reference. In panel (b) an extra bar was included showing the scores of FLUXCOM RS (grey bar) considering RS_METEO as reference. ....	20
Figure 7: Overview of control (CTR) simulation metrics for Gross Primary Production (GPP) using FLUXCOM RS_METEO as reference. (top left) mean FLUXCOM field, (top centre) mean CTR field; (top right) mean difference CTR-RS_METEO; (bottom left) $S_{bias}$ ; (bottom centre) $S_{rmse}$ and (bottom right) $S_{phase}$ . The values between brackets on the title of each map indicate the spatial average of the metric. ....	20
Figure 8: Same as Figure 7 but for the CLIM_FvCB simulation. ....	20
Figure 9: Same as Figure 7 but for the SF_AS simulation. ....	21
Figure 10: Global budgets considering only vegetated regions for (a) R <sub>n</sub> , (b) LE, (c) GPP, (d) TER, and (e) NEE displayed for all 7 simulations and the two FLUXCOM products. ....	22
Figure 11: Case study for July 2010 Gross Primary Production (GPP) anomaly in respect to the 2002-2010 July mean in (a) FLUXCOM RS, (b) FLUXCOM RS_METEO, (c) SF_AS, (d) SF_OL, (e) CLIM_FvCB, (f) CLIM, and (g) VAR. The anomaly is computed as July 2010 GPP minus the 2002-2010 July mean GPP, so that the blue colour indicates a reduction of GPP in 2010 in respect to the climatology. ....	23
Figure 12: Evaluation of Land Surface temperature (LST) daily maximum (left figure) and daily minimum (right figure) values, displaying the mean bias of the control simulation (CTR) LST for each season, which include (a) DJF, (b) MAM, (c) JJA, and (d) SON, as well as the difference of the mean absolute error between CLIM – CTR also for each season (panels from b to h). ....	24
Figure 13: Location of each Global Historical Climatology Network (GHCN)-Daily ground station (marked by red dots) used for the evaluation of the online simulations with the IFS new land cover model. A total of 14625 stations are considered after the application of proper data screening. ....	27
Figure 14: Different domains with respective Global Historical Climatology Network (GHCN)-Daily ground stations (marked by red dots) used for the evaluation of the online simulations with the IFS revised land cover model. These include the United States of America (USA), Europe (EU), Australia (AUS), Russia (RUS), Indochina (ICH), India (IND), South America (SA), Central Africa (CAF), and South Africa (SAF). ....	27
Figure 15: Aggregated means for the 2-metre temperature (T2M) normalised RMSE differences between CCI0 and CTR0 forecasts (depicted by the green lines) for winter (JF), spring (MAM), and summer (JJA) seasons in 2019 with lead times up to 5 days.	

Regions considered: Southern Hemisphere (SH), between  $-90^{\circ}$  and  $-20^{\circ}$ ; Tropics, between  $-20^{\circ}$  and  $20^{\circ}$ ; and Northern Hemisphere (NH), between  $20^{\circ}$  and  $90^{\circ}$ . Confidence range of 95% with AR(1) inflation and Sidak correction for 16 independent tests. Additional experimental forecasts from previous versions are also depicted (black and red with a previous version of the land-cover mapping, and blue with low vegetation clumping). ..... 29

Figure 16: As Figure 15 but for temperature (T) during spring in the Northern Hemisphere (NH) (a-c) and summer in the Tropics (d-f) at different pressure levels: 1000hPa (a, d), 850 hPa (b, e), and 500 hPa (c, f). ..... 30

Figure 17: As Figure 15 but for the geopotential height (Z) at 500hPa during spring (a, b) and summer (c, d) for the Tropics (a, c) and Northern Hemisphere (NH) (b, d). ..... 31

Figure 18: Normalised Root Mean Square Error (RMSE) differences between CCI0 and CTR0 forecasts for 2-days lead time forecasts of 2-metre air temperature (T2M a, b), 2-metre dewpoint temperature (TD2M, c, d) and geopotential height (Z) at 1000hPa (e, f), for spring (MAM, left panels) and summer (JJA, right panels) periods in 2019. No statistical significance testing was applied..... 32

Figure 19: Daily maximum temperature (TMAX) for 2 days lead-time forecasts mean bias differences: CTR0-CHCN (left panels), CCI0-CTR0 (centre panels) and absolute bias changes between CCI0 and CTR0:  $\text{abs}(\text{CCI0-GHCN})-\text{abs}(\text{CTR0-GHCN})$  (left panels) for winter (JF a, b, c), spring (MAM d, e, f), summer (JJA g, h, i), and autumn (SON j, k, l) of 2019..... 34

Figure 20: Boxplots for the daily maximum temperature (TMAX) Root Mean Square Error (RMSE) differences (in K) between CCI0 and CTR0 forecasts for (a) winter (JF), (c) spring (MAM), (b) summer (JJA), and (d) autumn (SON) of 2019. The location of mean and median values is depicted inside each boxplot by a blue cross and a red line (respectively), while at top a reference is made to the median and sampling values used for each region: all data (Global), Europe (EU), Russia (RU), United States of America (USA), South America (SA), Central Africa (CAF), South Africa (SAF), India (IND), Indochina (ICH), and Australia (AUS). The forecast lead time of 2 days is considered. 35

Figure 21: (left panels) scatter plots for the daily maximum temperature (TMAX) bias magnitude change between CCI0-CTR0 (vertical axis) as function of the CTR0 bias (horizontal axis); (right panels) regional maps (USA) for TMAX bias differences between CCI0 and CTR0. Results are displayed for winter (JF a, b), spring (MAM c, d), summer (JJA e, f), and autumn (SON g, h) of 2019, with the forecast lead of 2 days. All units are in K. .... 37

**Figure A1: Scores for each of the considered metrics (a) S\_bias, (b) S\_rmse, (c) S\_dist, and (d) S\_phase for the 6 considered variables (horizontal axis) shown for each simulation (bars) considering FLUXCOM R\_METEO as reference and including the metrics of FLUXCOM RS using FLUXCOM RS\_METEO as reference (grey bars). ..... 44**

Figure A2: Distribution of S\_bias (left panel), S\_rmse (centre panels) and S\_phase (right panels) for each simulation CTR (a, b, c), CLIM (d, e, f), VAR (g, h, i), CTR\_FvCB (j, k, l), CLM\_FvCB (m, n, o), SF\_OL (p, q, r), and SF\_AS (s, t, u) for the 4 climate regions Tropical, Arid, Temperate, and Cold (horizontal axis). Each violin distribution displays the extremes and median as a horizontal line, with the value of the median shown at the base of the plot. .... 45

Figure A3: Same as Figure 19 but for the daily minimum temperature (TMIN) forecasts. .... 46

Figure A4: Same as Figure 20 but for the daily minimum temperature (TMIN) forecasts. .... 46

Figure A5: Same as Figure 21 but for daily minimum temperature (TMIN) forecasts. .... 47

## Tables

Table 1: Offline simulations configurations.....	15
Table 2: Weight factors ( $\beta$ ) for each variable used for the calculation of the overall score ( $S_{all}$ ).....	18
Table 3: Weather forecasts simulations configuration.....	28
Table 4: Summary for the number of samples (#) found within each performance condition (Q1, Q2, Q3, and Q4) for the IFS land cover model defined between control and revised simulations of TMAX, including the respective percentage in reference to the total sampling (i.e., total number of stations found in the USA region) for each season in 2019.....	36

## 1 Executive Summary

Following the CO<sub>2</sub> Human Emissions (CHE) project, and the Observation-based system for monitoring and verification of greenhouse gases (VERIFY) project, CoCO<sub>2</sub> aims to create a prototype of anthropogenic CO<sub>2</sub> emission Monitoring and Verification Support (i.e. the CO<sub>2</sub>MVS) systems at global, regional, and local scales. The global modelling and data assimilation work package 3 (WP3) of CoCO<sub>2</sub> is focused on the development of the CO<sub>2</sub>MVS at global scale, with the European Centre for Medium-range Weather Forecasts (ECMWF) Integrated Forecasting System (IFS) at the core, with several planned developments, among them the community land-surface modelling for vegetation carbon exchange fluxes based on ECLand. In this context, biogenic fluxes play a fundamental role in the carbon cycle with a key component centred on the capacity of plants to absorb carbon via photosynthesis, with land use and land cover being crucial for its correct representation in the model.

This deliverable is integrated in Task 3.3 of WP3, aiming to improve global simulations of biogenic fluxes and to integrate land-surface remote sensing observations and ancillary data related to vegetation carbon exchanges in the IFS prototype. The objectives are focused on land-surface mapping and modelling improvements by exploring the role of high-resolution land cover change products on biogenic fluxes in the ECLand system, as well as its impact on Numerical Weather Prediction (NWP) metrics.

The work included ECLand offline and online model simulations, data processing, analysis, and evaluation. Simulations with the surface modelling platform developed by Météo-France (SURFEX) were also performed and analysed. The results not only show model improvements but also highlight some of the weaknesses found in the proposed revision of land cover and Leaf Area Index (LAI) to ECLand, allowing to identify areas where the model requires additional developments. The work was organized in three main components:

- Data pre-processing of the high-resolution land cover from the European Space Agency Climate Change Initiative (ESA-CCI) for 1992-2015 and from the Copernicus Climate Change Service (C3S) for 2016-present, as well as the Copernicus Global Land Service (CGLS) LAI.
- Evaluation of the energy, water, and carbon fluxes of ECLand and SURFEX offline simulations, using FLUXCOM as reference, and the evaluation of ECLand Land Surface Temperature (LST) using satellite data.
- Assessment of the meteorological impact in the NWP context of the proposed revised land cover and LAI.

In addition to the evaluation of the model developments, the work carried out in this report sets the foundations for an integrated NWP and biogenic fluxes assessment of the system, which will be used throughout the project and beyond. The technical infrastructure has been developed under a version control system and is available on request by any interested project partner.

The surface offline simulations evaluation identified the added value of the revised land cover and LAI in term of Gross Primary Production (GPP) when used in conjunction with a model

configuration using the Farquhar photosynthesis model (CLM\_FvCB). Results also suggest that time-varying LAI conditions are relevant to the GPP estimates during large-scale extreme events. Limitations in the evaluation of Net Ecosystem Exchange (NEE) and terrestrial respiration (TER) arise from model uncertainties in ECLand, SURFEX, and in the reference data used (FLUXCOM), suggesting that further developments are required and that flux adjustments are paramount to mitigate biases in global CO<sub>2</sub> analysis. The evaluation of the energy fluxes indicated a very good agreement of surface net radiation between ECLand, SURFEX, and FLUXCOM, followed by latent and sensible heat flux.

The online simulations of weather forecasts for the year of 2019 focused on the evaluation of the meteorological impact of the revised land cover and LAI. The results identified a clear improvement of 2-metre temperature in Eurasia during spring, while during summer the results were mostly negative in the tropics, particularly over North-Eastern Brazil and central-south Africa. Similar results were found when considering 1000 hPa geopotential height. The geopotential height at 500 hPa indicated an error growth with the revised land cover between forecast lead times of 1 to 3 days. A more detailed analysis over the USA, using a high density of network of ground stations from the Global Historical Climatology Network (GHCN), indicated that the pre-existing bias in the control simulation explains the impact of the revised land cover. There is a general increase of temperature in the latter, mostly induced by the reduction of surface roughness (due to the reduction of high vegetation) that is positive in regions where the control simulation had a negative bias, while negative in regions with already a pre-existing warm bias.

Ongoing developments in ECLand for the next operational IFS cycle include a revision of the handling of roughness lengths disaggregation for the tiles and post-processing of 2-metre temperature. Preliminary results (not included in the report) indicate a positive impact on LST but require further testing with the revised land cover and LAI. Similarly, ongoing tests (performed during the preparation of this report) with revised model parameters for TER show the potential to improve the global NEE budgets. Therefore, the work in progress will continue to provide improvements, with an analysis towards the latest version of the system, being proposed to be carried out in a second version of this report, which is expected to take place in about one year time (i.e., mid-2023).

## 2 Introduction

### 2.1 Background

Following the CO<sub>2</sub> Human Emissions (CHE) project, and the Observation-based system for monitoring and verification of greenhouse gases (VERIFY) project, CoCO<sub>2</sub> aims to create a prototype anthropogenic CO<sub>2</sub> emission Monitoring and Verification Support (i.e., the CO<sub>2</sub>MVS) systems at global, regional, and local scales in the framework of low-carbon economy and implemented commitments under the Paris Agreement (Delbeke et al., 2019) and the European Green Deal (EU Commission, 2021). This effort will support countries with an independent assessment of their emissions and reduction targets. Moreover, the resulting information, being consistent and reliable, will provide support to policy- and decision-making processes both at national and European level.

The global modelling and data assimilation work package 3 (WP3) of CoCO<sub>2</sub> is focused on the development of the CO<sub>2</sub>MVS at a global scale, building on existing capacity, and will deliver a system that can provide information of CO<sub>2</sub> emissions and other surface exchanges in near-real time, as well as in past reanalysis mode. The European Centre for Medium-range Weather Forecasts (ECMWF) Integrated Forecasting System (IFS) is at the core of the CO<sub>2</sub>MVS with planned developments in several areas, such as data assimilation, tracer advection, representation of anthropogenic impact on land-surface exchanges, simulations of biogenic fluxes, among others.

Biogenic fluxes play a fundamental role in the carbon cycle with a key component centred on the capacity of plants to absorb carbon via photosynthesis. This process is also crucial for the land-surface water and energy cycles. These three cycles, water/energy/carbon, are coupled and interact on time-scales ranging from minutes to centuries. Among different aspects of the processes involved, Land Use and Land Cover (LULC) is extremely relevant in the estimation of biogenic carbon fluxes (Quaife et al., 2008) and their changes are extensively studied in the context of climate change. Moreover, the errors found in the models' representation of LULC effects on the lower troposphere have also been shown to limit the progress in weather and climate predictability (e.g. Guo et al., 2011; Orth et al., 2016).

In the current ECMWF land surface model (ECLand) and data assimilation system the vegetation cover is derived from the Global Land Cover Characteristics data set (GLCC, Loveland et al., 2000) and the Leaf Area Index (LAI) is based on a 2000-2008 climatology derived from Moderate Resolution Imaging Spectroradiometer (MODIS) collection 5 (MOD15A2) data and rescaled using a static LAI dataset to guarantee neutral impact on the ECMWF model (Boussetta et al., 2013). Therefore, these crucial input datasets do not fully benefit from the developments of remote sensing land cover and vegetation data sets during the past 20 years. Recent studies have identified limitations of the current land cover and LAI datasets used in ECLand (Johannsen et al., 2019), but proposed updates (Nogueira et al., 2020, 2021a), based on recent remote sensing products, have not been evaluated in terms of biogenic fluxes nor in a Numerical Weather Prediction (NWP) context.

## 2.2 Scope of this deliverable

### 2.2.1 Objectives of this deliverable

This deliverable is integrated in Task 3.3 aiming to improve global simulations of biogenic fluxes and to integrate land-surface remote sensing observations and ancillary data related to vegetation carbon exchanges in the IFS prototype. The objectives are focused on land-surface mapping and modelling improvements by exploring the role of high-resolution land cover change products on biogenic fluxes in the ECLand system. The key objectives are:

- Update the vegetation classification and land cover based on Copernicus products.
- Updated vegetation description to create quasi-prognostic LAI in IFS, preparing the integration of process description with remote-sensing data.
- Test the impact of high-resolution annual land use description updated in IFS biogenic fluxes, and benchmarking against data-driven biogenic CO<sub>2</sub> flux products (FLUXCOM).
- Benchmarking with near-surface parameters derived from the operational IFS system, scoring specifically on NWP metrics.

### 2.2.2 Work performed in this deliverable

This deliverable was produced through the collaboration between all authors from the different institutions over the initial course of the project. This included ECLand offline and online model simulations, SURFEX simulations, data processing, analysis, evaluation, as well as the writing and final review of the current report. The analyses described further on not only show model improvements but also highlight some of the weaknesses found in the proposed revision of land cover and LAI to ECLand, allowing to identify areas where the model requires additional developments.

The work performed is divided into three sections:

- Section 3 presents the data pre-processing of the high-resolution land cover from the European Space Agency Climate Change Initiative (ESA-CCI) for 1992-2015 and from the Copernicus Climate Change Service (C3S) for 2016-present (available from 1992-

present in the Copernicus Climate Data Store), as well as the Copernicus Global Land Service (CGLS) LAI developed for the IFS land cover.

- Section 4 describes the ECLand and SURFEX offline simulations and the impact of these on the energy/water and carbon fluxes, as well as on Land Surface Temperature (LST).
- Section 5 presents the ECLand online weather forecasts simulations and explores the impact of the revised land cover and LAI in the context of NWP metrics.

### **2.2.3 Deviations and counter measures**

There were no major deviations from the workplan. However, it was identified the need for an updated version of this report due to ongoing model developments. Therefore, we propose to prepare an updated evaluation of the developed system in terms of biogenic fluxes and meteorological impact by mid-2023.

## **3 Processing of high-resolution land cover and Leaf Area Index**

### **3.1 Data description**

#### **3.1.1 ESA-CCI land cover and Copernicus Global Land Service Leaf Area Index**

The land cover revision for ECLand (i.e., vegetation types and fractional coverage) was based on the ESA-CCI dataset. This dataset provides consistent maps of land cover, based on the 22 classes from the land cover classification system developed by the United Nations Food and Agriculture Organization. The latter are derived by combining remotely sensed surface reflectance and in situ observations (Defourny et al., 2014). The datasets used for this analysis are available at the 300 m spatial resolution, on an annual basis, in which ESA-CCI datasets are provided for 1992-2015, while C3S datasets are delivered for 2016–present.

The LAI was obtained from the CGLS version 2 at the 1 km resolution, covering the entire globe. The LAI product is derived from the SPOT-VEGETATION and PROBA-V satellite observations using the algorithm described by (Verger et al., 2014). It should be noted that this version was discontinued in June 2020, following the end of PROBA-V mission. For an operational implementation, the new version of CGLS 300 m based on Sentinel-3 could be explored, with a particular attention given to the stability of the LAI from different sensors.

### **3.2 Data processing**

#### **3.2.1 Land cover**

The land cover data processing included (i) the spatial aggregation from the original 300 m and (ii) the transformation of the ESA-CCI land cover classes to the ECLand land cover types. The 300 m spatial resolution ESA-CCI land cover grid points are spatially aggregated to the target model resolution by counting the number of 300 m pixels of each class occurring within each grid-cell of the target grid. This step creates maps of fraction cover for each ESA-CCI land cover type at the target grid. These fractional covers of ESA-CCI classes are then converted into the ECLand land cover classes using a cross-walking table. The latter identifies, for each of ESA-CCI class, the fraction of ECLand land cover type, being based on previous works (Nogueira et al., 2020, 2021b; Poulter et al., 2015), despite demonstrating a certain degree of uncertainty, particularly over sparse vegetated regions. Finally, the new ECLand land cover types fractions are processed to compute the fractional cover of low and high

vegetation, as well as the dominant types of low and high vegetation attributed to each grid point.

### 3.2.2 Leaf Area index

The total LAI data is processed in 3 steps:

1. A climatology for the period 2010-2019 was computed at the original resolution of 1 km.
2. The climatology (or the original data) total LAI is aggregated to the target grid and disaggregated into low and high vegetation.
3. The disaggregation of low/high LAI is further processed for consistency with the land cover map and to conserve the total LAI.

For the spatial aggregation and the low/high disaggregation, each 1 km pixel of CGLS LAI ( $lai_{hr}$ ) is aggregated to the target grid total LAI ( $lai$ ), while the low and high LAI ( $lai_{lv}$ ,  $lai_{hv}$ ) are derived considering the low vegetation fraction ( $cvl$ ), high vegetation fraction ( $cvh$ ), and land sea mask ( $ism$ ). The  $cvl$ ,  $cvh$ , and  $ism$ , are derived from aggregating the 300 m ESA-CCI land cover to the CGLS 1 km grid, by applying the same cross-walking table (as described in the previous section). For a particular target grid-cell, the aggregated  $lai$ ,  $lai_{lv}$ , and  $lai_{hv}$ , consider all the 1 km pixels that fall inside the grid-cell (i.e.,  $i = 1, n$  pixels, where  $ism > 0.5$ ). Each 1 km pixel is then assigned to high or low vegetation, depending on  $cvl$  and  $cvh$ , and average weighted on  $cvl$  or  $cvh$  to derive the low and high vegetation  $lai$ :

$$lai = \frac{1}{n} \sum_{i=1}^n lai_{hr_i} \quad (1)$$

$$lai_{lv} = \frac{1}{W} \sum_{i=1}^n w_i \times lai_{hr_i}, W = \sum_{i=1}^n w_i; w_i = \begin{cases} cvl_i, cvl_i > cvh_i \\ 0, cvl_i \leq cvh_i \end{cases} \quad (2)$$

$$lai_{hv} = \frac{1}{W} \sum_{i=1}^n w_i \times lai_{hr_i}, W = \sum_{i=1}^n w_i; w_i = \begin{cases} cvh_i, cvh_i > cvl_i \\ 0, cvh_i \leq cvl_i \end{cases} \quad (3)$$

A further processing is applied to reinforce land cover consistency and to guarantee total LAI conservation. The low and high vegetation  $lai$  for each vegetation type are scaled to guarantee consistency in term of annual maximum. For each vegetation type and grid-point, the maximum annual LAI is computed, and the median of the distribution is taken as representative for that vegetation type  $lai$ . A new  $lai$  is then computed by scaling the maximum  $lai$  in each grid-cell by that median value, weighted by the vegetation fraction:

$$lai_{vty\_max} = \max(lai_{vty}) \quad (4)$$

$$lai_{vty\_max\_med} = \text{median}(lai_{vty\_max}) \quad (5)$$

$$\alpha = \min(5, \max(0.2, \frac{lai_{vty\_max\_med}}{lai_{vty\_max}})) \quad (6)$$

$$lai_{vty} = (1 - cv) \times \alpha \times lai_{vty} + cv \times lai_{vty} \quad (7)$$

Finally, the updated  $lai_{lv}$  and  $lai_{hv}$  for each calendar month is scaled to conserve the total LAI with a multiplicative factor  $\beta$ , limited between 0.2 and 5 to avoid large unrealistic corrections:

$$lai^* = \frac{cvl \times lai_{lv} + cvh \times lai_{hv}}{cvl + cvh} \quad (8)$$

$$\beta = \min(5, \max(0.2, \frac{lai}{lai^*})) \quad (9)$$

$$lai_{lv} = \beta \times lai_{lv} \quad (10)$$

$$lai_{hv} = \beta \times lai_{hv} \quad (11)$$

### 3.2.3 Dominant c3/c4 and climate classification

In addition to the land cover and LAI datasets, two other datasets with the flexibility to adapt to any model resolution were also pre-processed, being used as auxiliary data for model diagnostics, namely: (i) the Köppen-Geiger climate classification; and (ii) the distribution of  $C_3$  and  $C_4$  low vegetation, which are relevant for photosynthesis pathways. The Köppen-Geiger climate classification was based on Beck et al. (2018) by performing an interpolation to the model grid using the nearest-neighbour approach (see **Figure 1**). The global distribution of  $C_3$  and  $C_4$  herbaceous cover is based on Still et al. (2003) for  $C_3$  by default and on the land-use harmonization dataset (LUH <https://luh.umd.edu/data.shtml>, in 2019) for crops using the following procedure:

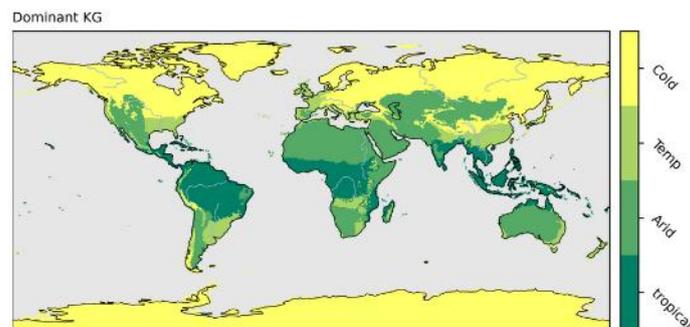
$$C_{3S} = 1 - \left( \frac{C_{4l}}{C_{3l} + C_{4l}} \right) \quad , \text{ from (Still et al., 2003),} \quad (12)$$

$$C_{4C} = \frac{C_{4ann} + C_{4per}}{C_{3ann} + C_{4ann} + C_{3per} + C_{4per} + C_{3nfx}} \quad \text{from LUH,} \quad (13)$$

where  $C_{4l}$  and  $C_{3l}$  are the  $C_4$  and  $C_3$  distributions from Still et al. (2003) and  $C_{4ann}$ ,  $C_{4per}$ ,  $C_{3ann}$ ,  $C_{3per}$  refer to  $C_3$  or  $C_4$  annual (*ann*) and perennial crops (*per*) and  $C_{3nfx}$  to nitrogen-fixing crops from the land-use harmonization data.

Since ECLand only considers the dominant low and high vegetation type, a dominant  $C_3/C_4$  map is generated based on the previous fields, assuming that  $C_{3S}$  is the baseline, being replaced by  $C_{4C}$  when crops are defined as the dominant low vegetation type. This procedure requires as input the land sea mask (*lsm*), low vegetation cover (*cvl*), and type of low vegetation (*tv*). The final  $C_3/C_4$  map is shown in **Figure 2**, and is given as:

- *map* = 3 in all land points by default;
- *map* = 4 if *cvl* > 0 & *tv* = 1 &  $C_{4C} > 0.5$  (from LUH);
- *map* = 3 if *cvl* > 0 & *tv* = 1 &  $C_{4C} \leq 0.5$  (from LUH);
- *map* = 3 if *cvl* > 0 & *tv* ≠ 1 &  $C_{3S} > 0.5$  (from Still);
- *map* = 3 if *cvl* > 0 & *tv* ≠ 1 &  $C_{3S} \leq 0.5$  (from Still);



**Figure 1: Dominant Köppen-Geiger climate classification.**

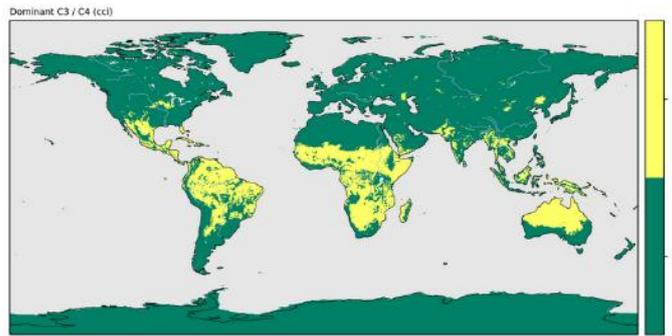


Figure 2: Dominant mapping of the distribution of C<sub>3</sub> and C<sub>4</sub> for grass and crops plant types.

## 4 Offline simulations and evaluation of biogenic fluxes

### 4.1 Data and methods

#### 4.1.1 Model simulations

Offline simulations with the ECLand and Interaction between Soil Biosphere and Atmosphere (ISBA) models were carried out driven by near-surface meteorological fields from the ECMWF ERA5 reanalysis (Hersbach et al., 2020). Details regarding model configurations are presented in the following sub-sections, along with a respective summary in [Table 1](#).

##### 4.1.1.1 SURFEX simulations

The ISBA land surface model is embedded in the SURFEX modelling platform (Masson et al., 2013). The latter contains data assimilation routine for land analysis. A python environment was built to manage the data flow, from satellite-derived observations to be assimilated to the display of open-loop (SF\_OL) and analysis simulation (SF\_AS) results. The combination of these elements is called LDAS-Monde (Albergel et al., 2017). This open-source tool has now reached a high technical readiness level and is currently being transferred to operational services of Météo-France. The current version of LDAS-Monde is able to assimilate true LAI values from the CGLS with or without the joint assimilation of surface soil moisture (SSM) products disseminated by CGLS, as well as to be adapted to ingest other sources of data. A unique capability of LDAS-Monde is the possibility to analyse the root-zone soil moisture (RZSM) at several depths by assimilating LAI and SSM, or both parameters individually.

In this report, ESA-CCI COMBINED soil moisture v6.1 and CGLS LAI v2 data are assimilated into the ISBA model (with the "NIT" option used), while analysis simulation (SF\_AS) results are obtained in addition to the SF\_OL (no assimilation). The original plan was to use the updates of SURFEX V9 of the "NCB" option (Gibelin, et al. 2008; Delire *et al.*, 2020), however SURFEX V9 is currently unavailable. The NCB option simulates carbon storage, all respiration terms, and spin-up procedures. On the other hand, the "NIT" option has a simple uncalibrated parametrization of the terrestrial sink (TER) (Albergel et al., 2010) based on near-surface soil temperature (Q10 function) and soil moisture (linear function). In order to fit TER to Gross Primary Production (GPP) we made the assumption that accumulated TER is equal to 70 % of the accumulated GPP, as observed over many plant functional types (PFTs) (Wang et al. 2008; Yuan *et al.*, 2011). This correction is performed only at a post-processing stage, therefore not providing any feedback during the model simulation. This is tantamount to tuning a TER value over a near-surface soil temperature of 25 °C at field capacity soil moisture conditions, grid-cell by grid-cell, in order to rescale TER to GPP. For this analysis, a rescaling to TER is applied so that the respective mean values over 2002-2019 can match the 0.7 x

GPP mean values over the same period. Simulations using the NCB option will be performed at a later stage, including a comparison with the ones made using the NIT option. In ISBA, GPP is determined using an uncalibrated A-gs approach (plant parameters are derived from the literature or from the TRY database).

#### **4.1.1.2 ECLand simulations**

The land-surface developments are based on the Carbon-Hydrology Tiled ECMWF Scheme for Surface Exchanges over Land (CHTESSEL) and form an integral part of the IFS model, supporting a wide range of global weather, climate and environmental applications. ECLand is a flexible system created to facilitate modular extensions in support of NWP and society-relevant operational services, for example, Copernicus (Boussetta et al., 2021). The ECLand system computes the land surface response to atmospheric forcing, estimating the surface energy and water fluxes, as well as the evolution of soil temperature, moisture, and snowpack. The surface-atmosphere exchanges take place in a surface skin layer separating the sub-soil from the lower troposphere. The skin layer has no heat capacity, and its temperature is used to compute the upwelling longwave radiation, making it directly comparable with LST. Each grid-box can represent different types of land cover using multiple tiles, including bare ground, dominant high vegetation type, dominant low vegetation type, intercepted water (on the canopy), and shaded and exposed snow. As an important component of the global carbon budget, a land surface CO<sub>2</sub> exchange module has been added to ECLand (Boussetta, Balsamo, Beljaars, Agusti-Panareda, et al., 2013), enabling environmental forecasting applications, which also involves interaction with atmospheric CO<sub>2</sub> concentration. The release of land biogenic CO<sub>2</sub> and the photosynthesis processes fixing carbon dioxide into biomass are parametrized in ECLand, enabling a response to meteorological and climate forcing and to the natural biomes distribution, including respective stress conditions (A-gs model). The soil respiration is also parametrized in an NWP adapted way as a function of land-use, similarly to the scheme used in ISBA. The A-gs model and soil respiration implementations are very similar to those in SURFEX, except for some model parameters that were tuned/calibrated differently. The photosynthesis and transpiration parametrizations are made modular within the land surface scheme, allowing an independent interaction with the atmospheric CO<sub>2</sub> concentrations for global monitoring and prediction purposes (Agusti-Panareda et al., 2014, 2016). A detailed description of the model can be found on the ECMWF IFS documentation (ECMWF, 2021).

In this analysis three configurations of ECLand with the default A-gs photosynthesis model are assessed:

- CTR: with the current operational model configuration.
- CLIM: with an updated land cover description, including low and high vegetation cover and types, as well as new LAI climatology (see section 3.1.1 for more details).
- VAR: using the new LULC data as CLIM (section 3.1.1) and considering the original data inter-annual variability (IAV). This configuration serves also as a proxy to an idealized model configuration with prognostic evolution of LAI.

In addition to the abovementioned configurations, another two configurations were also included following a recent implementation in ECLand of a photosynthesis model based on Farquhar et al. (1980) and Collatz et al. (1992), as implemented by Yin and Struik (2009) in the organising Carbon and hydrology In Dynamic Ecosystems (ORCHIDEE) land surface model (Krinner et al., 2005). Therefore, a control simulation using the default ECLand land cover and LAI data (CTR\_FvCB) and a simulation with the revised land cover and LAI climatology (CLIM\_FvCB) were also included in this report. All five ECLand simulations have applied the same TER parametrization and the same rescaling as in SURFEX to match 0.7 x GPP. Note that the model parameters of A-gs and Farquhar in the ECLand model have been optimized with in situ CO<sub>2</sub> flux-tower data. Since FLUXCOM also uses flux-tower data for

calibration purposes, the ECLand and FLUXCOM GPP products are not completely independent.

**Table 1: Offline simulations configurations.**

<b>Simulation</b>	<b>Setup Details</b>
CTR	ECLand-A-gs control simulation, with default ECLand configuration.
CLIM	ECLand-A-gs with new land cover and new LAI climatology.
VAR	ECLand-A-gs with yearly varying new land cover and LAI.
CTR_FvCB	ECLand-Farquhar control simulation.
CLIM_FvCB	ECLand-Farquhar with new land cover and new LAI climatology.
SF_OL	SURFEX control simulation with open loop.
SF_AS	SURFEX data assimilation, including ESA-CCI SSA and GEOV1 LAI.

#### 4.1.2 Land cover and LAI changes in ECLand

The revised land cover based on the ESA-CCI products for ECLand includes the update of the low and high vegetation types and cover. The current and revised dominant low and high vegetation types are shown in **Figure 3**. The revised high vegetation types do not include the types “interrupted forest” and “mixed forest”, which were mostly replaced by deciduous broad leaf trees (**Figure 3a** and **Figure 3b**). Additionally, the revised low vegetation types do not include the “semi-desert” type, which was replaced in most regions by shrubs. Despite these changes, the main dominant vegetation patterns remain similar. The largest changes came from the vegetation cover, which is shown in **Figure 4**, in terms of effective vegetation cover. In ECLand, the fraction of the vegetated tiles (i.e., effective vegetation) is computed as the product of the input vegetation cover by the vegetation density, which is given as a look-up table for each vegetation type (see Table 8.1 in Chapter 8 ECMWF (2021)). In the revised land cover, there is a general reduction of high vegetation, which is replaced by low vegetation and bare-ground. There is also a decrease in the bare ground cover in some semi-arid regions (e.g. Rock mountains, East Australia), which is associated with the removal of the “semi-desert” type, that had a low vegetation density. There are two key differences in the revised land cover: (i) the removal of two types of high vegetation, which include “interrupted forest” and “mixed forest”; and (ii) the reduction of high vegetation cover, which is compensated by an increase of low vegetation.

The revised LAI based on the CGLS dataset has a higher annual amplitude, when compared with the current climatology used in ECLand. This behaviour is exemplified in **Figure 5** for two regions (East North America and South Africa), where the increased annual amplitude results from both a reduction of the annual minimum and increase of annual maximum. The disaggregation of total LAI into the low and high vegetation components produces a consistent seasonal evolution of the two vegetation components, while guaranteeing the conservation of the total LAI given by the satellite data.

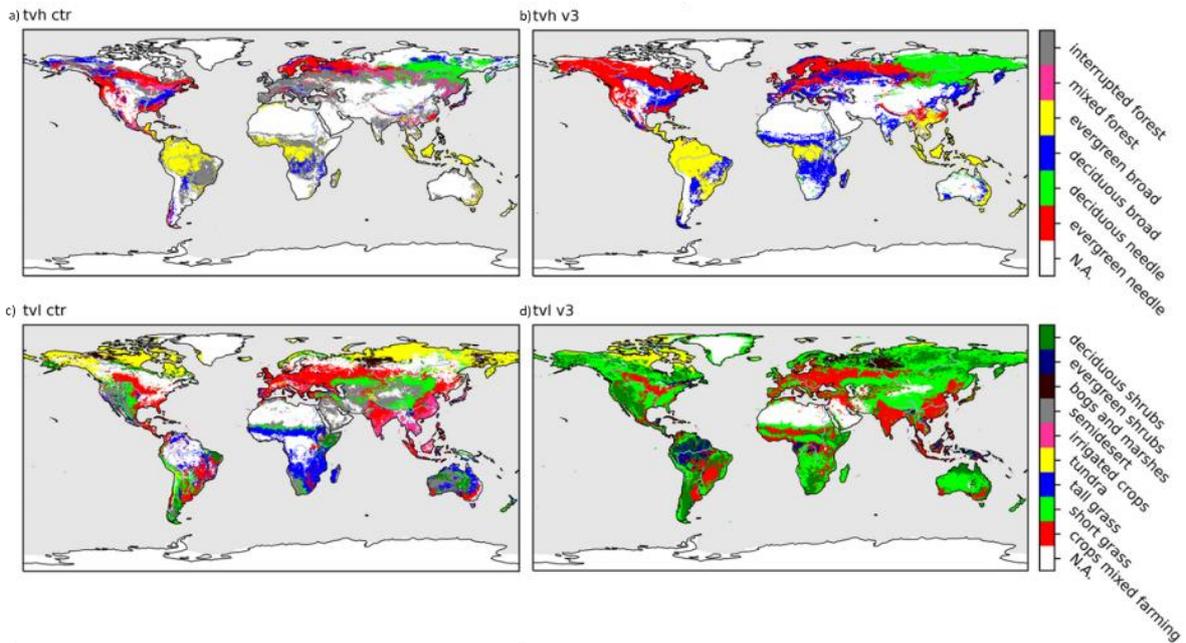


Figure 3: Dominant high vegetation (a, b) and low vegetation (c, d) in the default (CTR) ECLand configuration (a, c) and revised land cover based on ESA-CCI (b, d).

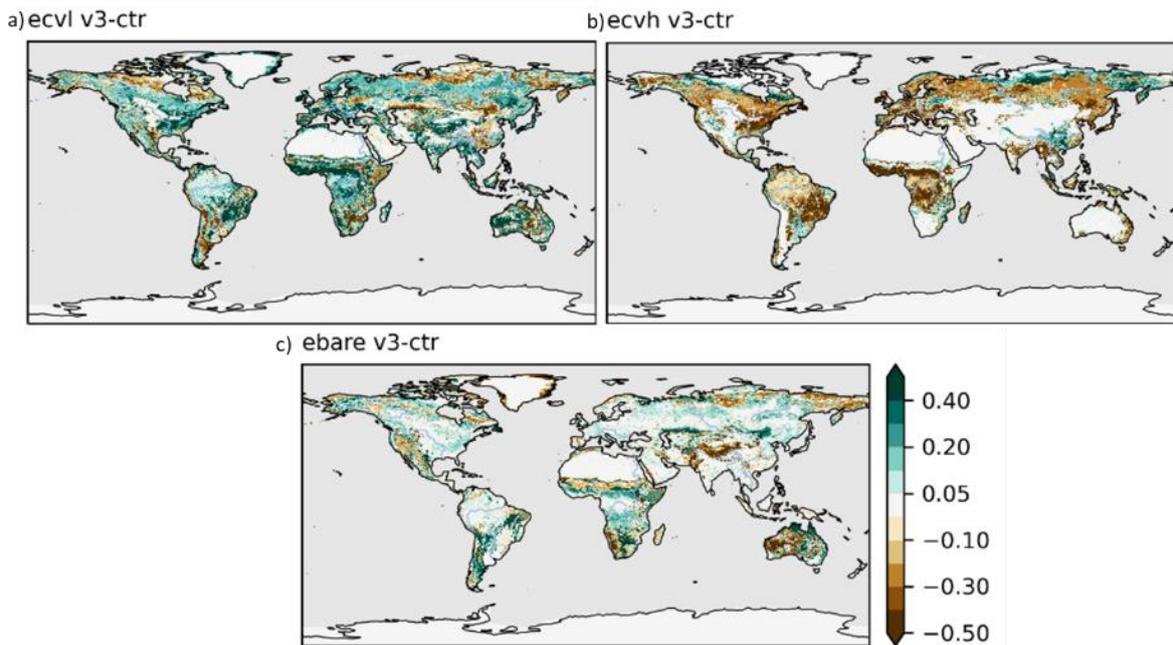
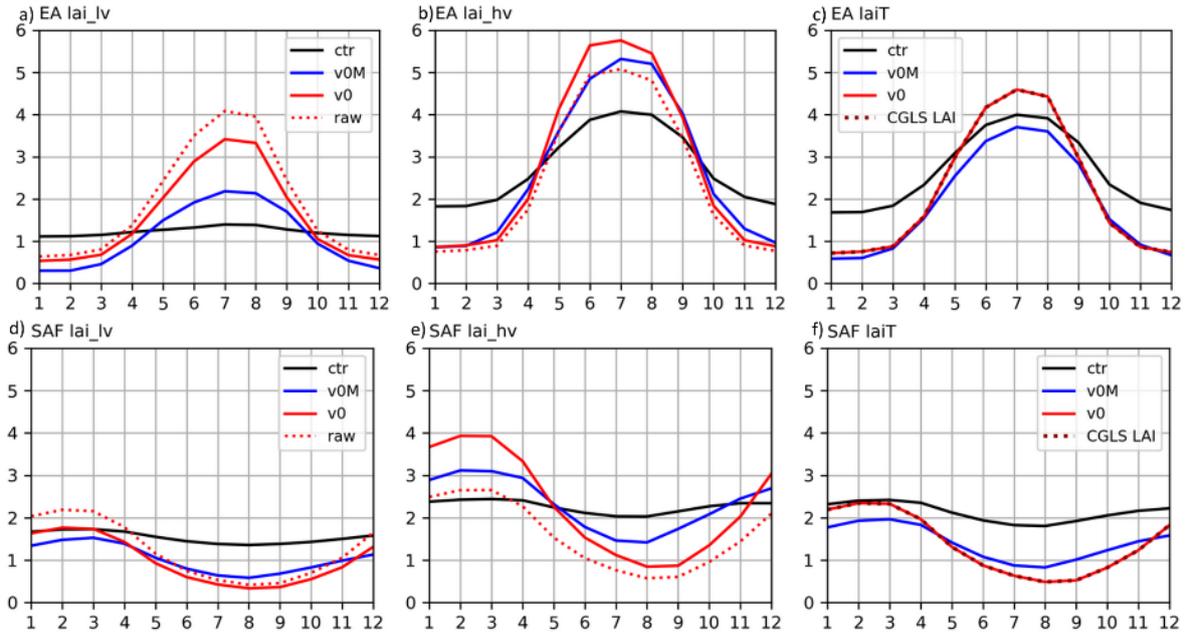


Figure 4: Differences between CLIM and control (CTR) simulations effective low vegetation cover (a), effective high vegetation cover (b), and effective bare ground cover (c).



**Figure 5: Mean annual cycle of low vegetation LAI (a, d), high vegetation LAI (b, e), and total LAI (c, f), for two regions: (a-c) East North America [30° to 50° N, 90° to 70° W]; (d-f) South Africa [20° to 10° S, 12° to 34° E]. These figures compare the ECLand default LAI as in CTR (black curves) with the new revised CLIM LAI (solid red lines, v0). The total original LAI (dotted magenta in c, f panels), the intermediate LAI aggregation (eq. 7) (red dotted red lines), and an independent dataset based on MODIS data (blue lines, v0M) are also included.**

#### 4.1.3 FLUXCOM and evaluation metrics

Surface water, energy, and carbon, fluxes of the offline simulations were compared against the FLUXCOM ensemble of global land-atmosphere energy and carbon fluxes (Jung et al., 2019, 2020). FLUXCOM uses machine learning to merge energy and carbon flux measurements from FLUXNET eddy covariance towers with remote sensing and meteorological data to estimate several global gridded quantities: net radiation (Rn), latent (LE) and sensible (H) heat, GPP, Net Ecosystem Exchange (NEE), and terrestrial ecosystem respiration (TER), and their uncertainties. In this analysis the ensemble mean of two FLUXCOM products are used: (i) the remote sensing (RS) and (ii) the remote sensing and meteorological data (RS\_METEO).

The comparison between simulations and FLUXCOM products followed the International Land Model Benchmarking (ILAMB) System protocol (Collier et al., 2018). ILAMB provides a comprehensive model assessment based on a set of normalized scores. Errors are transformed into normalized score ( $s$ ) on the unit interval  $[0,1]$  by passing the normalized error, generically represented here as  $\epsilon$ , and passing it through the exponential function:

$$s = e^{-\alpha\epsilon}, \quad (14)$$

where  $\alpha$  is a parameter that can be used to tune the mapping of error to score, being set to 1 in this study. A score of 1 is perfect and it tends to zero as the deviates from the reference dataset. The following scores were used:

- $S_{bias}$ : bias normalized by the centralized root mean square of the reference data (equations 11 to 15 in Collier et al. (2018));

- $S_{rmse}$ : centralized root mean square error (i.e., removing the mean) normalized by the root mean square of the reference data (equations 16 to 19 in Collier et al. (2018));
- $S_{phase}$ : phase shift of the annual cycle considering the timing of the annual maximum (equations 21 to 23 in Collier et al. (2018));
- $S_{dist}$ : spatial distribution of the time averaged variable (equations 29 and 30 in Collier et al. (2018));
- $S_{all}$ : the overall score for a given variable, computed as the composite of the four 4 metrics previously defined:

$$S_{all} = \frac{\beta_{bias}S_{bias} + \beta_{rmse}S_{rmse} + \beta_{phase}S_{phase} + \beta_{dist}S_{dist}}{\beta_{bias} + \beta_{rmse} + \beta_{phase} + \beta_{dist}}, \quad (15)$$

where the weight  $\beta$  for each metric depends on the variable considered, as shown in

**Table 2.** The weights were based on a qualitative interpretation of the strengths (seasonality and spatial pattern) and weakness (inter-annual variability and absolute mean fields of NEE and TER).

The interannual variability is also presented as a different metric in ILAMB, although it is not considered in this report due to limitations of FLUXCOM in representing interannual variability. The weight factors presented in

**Table 2** are based on a qualitative interpretation of the strengths and weaknesses of the FLUXCOM products, giving more weight to the seasonal cycle phase and spatial distribution, while the bias of NEE and TER is not considered.

**Table 2: Weight factors ( $\beta$ ) for each variable used for the calculation of the overall score ( $S_{all}$ ).**

	Rn	LE	H	GPP	NEE	TER
$\beta_{bias}$	2	1	1	1	0	0
$\beta_{rmse}$	1	1	1	1	1	1
$\beta_{phase}$	2	2	2	2	2	2
$\beta_{dist}$	2	2	2	2	2	2

#### 4.1.4 Satellite land surface temperature

Estimates of LST from the Satellite Application Facility on Land Surface Analysis (LSA-SAF) are available every 15 minutes (between 2004-present) over land pixels within the Meteosat Second Generation (MSG) disk. These comprise satellite zenith view angles between 0° and 80°, with a 3 km resolution at the nadir and are derived from the outgoing Thermal Infrared Radiation (TIR) measured at top-of-atmosphere by the Spinning Enhanced Visible and InfraRed Imager (SEVIRI) onboard the MSG satellite, employing a generalized “split-window” technique (Freitas et al., 2010). The TIR spectral band (8–13  $\mu\text{m}$ ) is particularly appropriated, since it presents relatively weak atmospheric attenuation under clear sky conditions and includes the peak of the Earth's spectral radiance. On the negative side, LSA-SAF LST estimates derived from TIR are mostly limited to clear-sky observations, which poses a significant limitation to their coverage (Trigo et al., 2011). Despite this limitation, several studies have used LSA-SAF LST datasets to evaluate LST of land surface models (e.g Johannsen et al., 2019; Nogueira et al., 2021b).

## 4.2 Results

### 4.2.1 Surface water/energy/carbon fluxes evaluation

In the following results, the comparison between the simulations and FLUXCOM is performed considering the period between 2002-2015 (14 years) when comparing against the RS FLUXCOM product, while the period between 2002-2010 (9 years) is used when comparing against the RS\_METEO product, following the datasets availability.

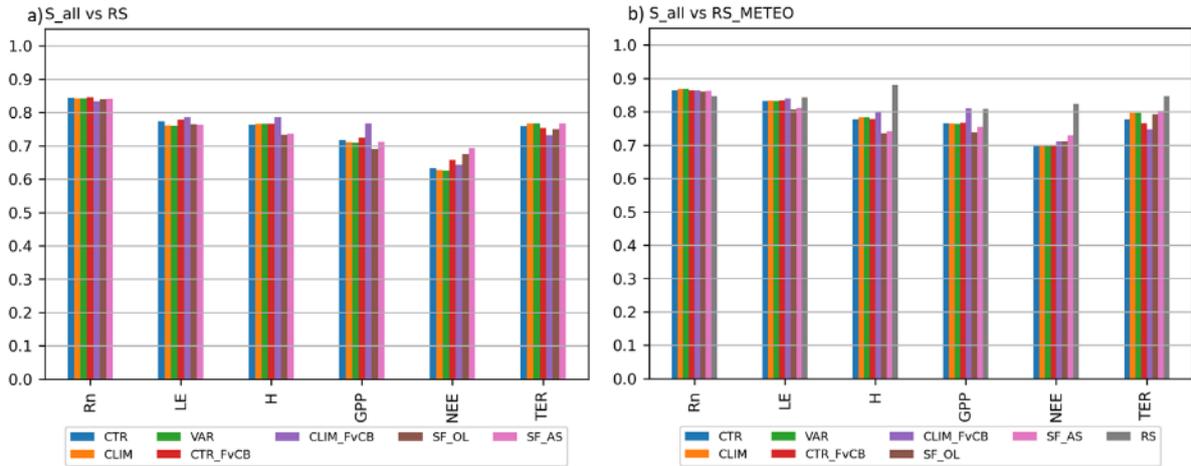
The overall scores  $S_{all}$  for the different simulations and variables are displayed in **Figure 6**, considering both FLUXCOM RS and RS\_METEO as reference datasets. Scores are generally higher when RS\_METEO is taken as reference, however the ranking among variables and simulations is similar when considering RS as reference (**Figure 6a** versus **Figure 6b**). Net radiation has the best scores, followed by LE, H, GPP and TER, with NEE presenting the poorest scores. Although all simulations score similarly per variable, there are two points worth noting: (i) CLIM\_FvCB shows higher scores than the other experiments for LE, H and GPP; (ii) SF\_AS scores are generally better than SF\_OL. The scores for RS taking RS\_METEO as reference (see grey bar in **Figure 6b**) can be interpreted as a benchmark score, estimating how the differences in the two products propagate into the score. This comparison shows very close results to the simulations for Rn and LE, and to some extent GPP and TER, while for H, and NEE there is higher agreement between RS and RS\_METEO than between the different simulations and RS\_METEO.

The contribution of the different metrics  $S_{bias}$ ,  $S_{rmse}$ ,  $S_{dist}$ , and  $S_{phase}$ , to  $S_{all}$  is shown in **Figure A1**, taking RS\_METEO as reference (results of considering RS as reference are similar). The  $S_{dist}$  and  $S_{phase}$  metrics, used to assess the spatial pattern of the mean field and timing of the annual maximum, respectively, have higher values for all variables, indicating a close agreement between the simulations and the reference data. The  $S_{bias}$  and  $S_{rmse}$  have lower values, which indicates less agreement, being also associated to larger uncertainties in the reference dataset, which was considered in the weights used to compute the overall score (see

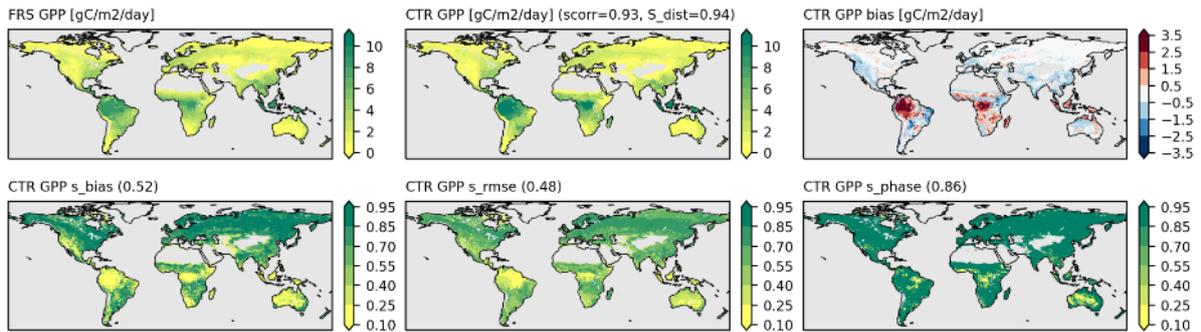
**Table 2**).

The spatial maps of GPP and the simulations, their differences, and of the metrics  $S_{bias}$ ,  $S_{rmse}$  and  $S_{phase}$  are shown in **Figure 7** for CTR, while for CLIM\_FvCB and for SF\_AS are shown in **Figure 8** and **Figure 9**, respectively. The results indicate that the tropical regions are the most problematic, with a clear improvement in terms of  $S_{rmse}$  in CLIM\_FvCB (0.60) when compared with CTR (0.48). SF\_AS presents the lowest biases (0.64), however some discrepancies are found in the annual cycle maximum over the tropical region.

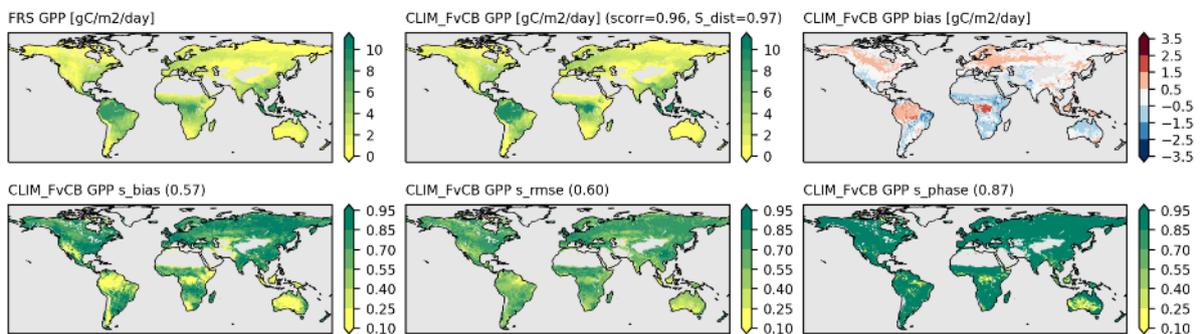
The distribution of GPP  $S_{bias}$ ,  $S_{rmse}$ , and  $S_{phase}$ , for the different simulations (see **Figure A2**) considering the 4 climate regions (see **Figure 1**), extends the spatial maps discussed previously. The results indicate a good agreement between the simulations and RS\_METEO in terms of the annual cycle maximum in all regions. For  $S_{bias}$  and  $S_{rmse}$ , there is a clear difference between Tropical / Arid and Temperate / Cold regions with the latter having the highest scores. These results can be partially attributed to the scores calculation, since both scores are normalized by the root mean square of the reference data, which is higher in the temperate and cold climates, with a marked seasonality, when compared with the tropical regions.



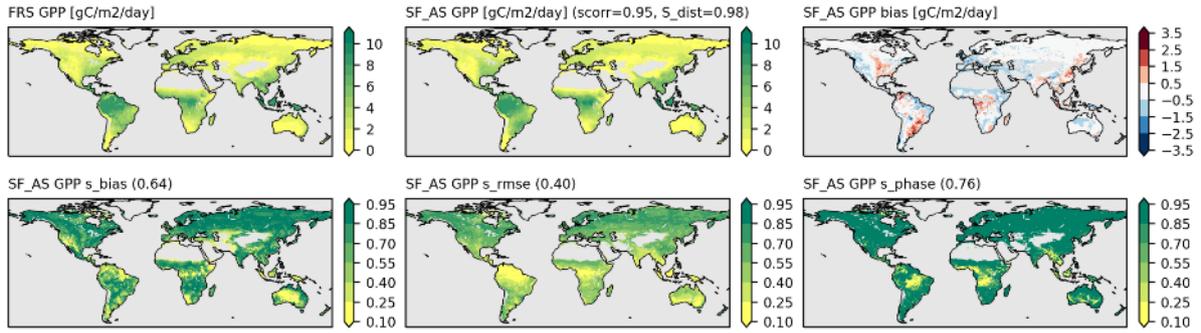
**Figure 6: Overall score  $S_{all}$  for the 6 considered variables (horizontal axis) shown for each simulation (bars) considering (a) FLUXCOM RS as reference and (b) FLUXCOM RS\_METEO as reference. In panel (b) an extra bar was included showing the scores of FLUXCOM RS (grey bar) considering RS\_METEO as reference.**



**Figure 7: Overview of control (CTR) simulation metrics for Gross Primary Production (GPP) using FLUXCOM RS\_METEO as reference. (top left) mean FLUXCOM field, (top centre) mean CTR field; (top right) mean difference CTR-RS\_METEO; (bottom left)  $S_{bias}$ ; (bottom centre)  $S_{rmse}$  and (bottom right)  $S_{phase}$ . The values between brackets on the title of each map indicate the spatial average of the metric.**



**Figure 8: Same as Figure 7 but for the CLIM\_FvCB simulation.**



**Figure 9: Same as Figure 7 but for the SF\_AS simulation.**

The annual global budgets computed only over vegetated areas, for the different variables and simulations are shown in **Figure 10**. Rn shows a variability of about 6 % (86 to 92 W m<sup>-2</sup>) between the different simulations, increasing to 16 % (46 to 54 W m<sup>-2</sup>) in LE, 32 % in GPP (90 to 125 Pg C year<sup>-1</sup>) and TER (65 to 90 Pg C year<sup>-1</sup>) and 84 % in NEE (-37 to -15 Pg C year<sup>-1</sup>). This inter-model variability of the budgets, which is increasing from the energy budgets to the final net carbon budget, reflects the large uncertainties and difficulties in constraining the global biogenic carbon budget, justifying flux adjustment to mitigate biases in global CO<sub>2</sub> analysis (Agustí-Panareda et al., 2016). The inter-model variability in TER is very close to GPP due to the correction applied to both ECLand and SURFEX simulations to constrain the long-term mean of TER to be 70 % of the long-term mean of GPP. Without this constrain the inter-model variability would have been much higher in TER and in NEE, and the comparison with FLUXCOM products would be significantly affected (not shown). However, this correction of TER results is a significant NEE sink in all model simulations, ranging between -37 to -25 Pg C year<sup>-1</sup>, which is higher than FLUXCOM estimates ranging between -15 to -20 Pg C year<sup>-1</sup>. It should be noted that FLUXCOM estimates are considered to be very high and beyond reasonable estimates given the current knowledge of the global carbon cycle, e.g. the land sink of the global carbon budget (Friedlingstein et al., 2022). Therefore, the current TER and NEE budgets from the models and FLUXCOM are considered unreliable, demanding further alternatives, as discussed at the end of this section.

Finally, the summer of 2010 was selected as a case study to investigate the impact of time-varying land cover and LAI conditions on GPP. This period was exceptionally warm in Eastern Europe and large parts of Russia (Barriopedro et al., 2011) with significant hydrometeorological and biospheric impacts (Flach et al., 2018). **Figure 11** shows the July 2010 anomaly (in respect to the 2002-2010 mean) for the different simulations. While there is a good agreement in terms of the spatial pattern of the anomalies, with a negative/positive dipole in Eastern Russia, the amplitude of the anomalies vary significantly. Both FLUXCOM products RS and RS\_METEO indicate smaller anomalies, followed by CLIM and CLIM\_FvCB, with SF\_AS, SF\_OL, and VAR, simulations showing the highest anomalies. The small anomalies found in the FLUXCOM product were expected due to its known limitations in terms of inter-annual variability (Jung et al., 2020). CLIM and CLIM\_FvCB differ from SF\_AS, SF\_OL, and VAR, in several ways, however there is a particularity in the fixed versus varying LAI conditions within the former. Although it is not possible to exactly quantify the GPP anomaly of this event, these results suggest that varying LAI conditions are relevant to the GPP estimates during large-scale extreme events, and that ECLand reproduces SURFEX anomalies when driven by time-varying LAI conditions, as it was the case in the VAR simulations with prescribed satellite LAI data.

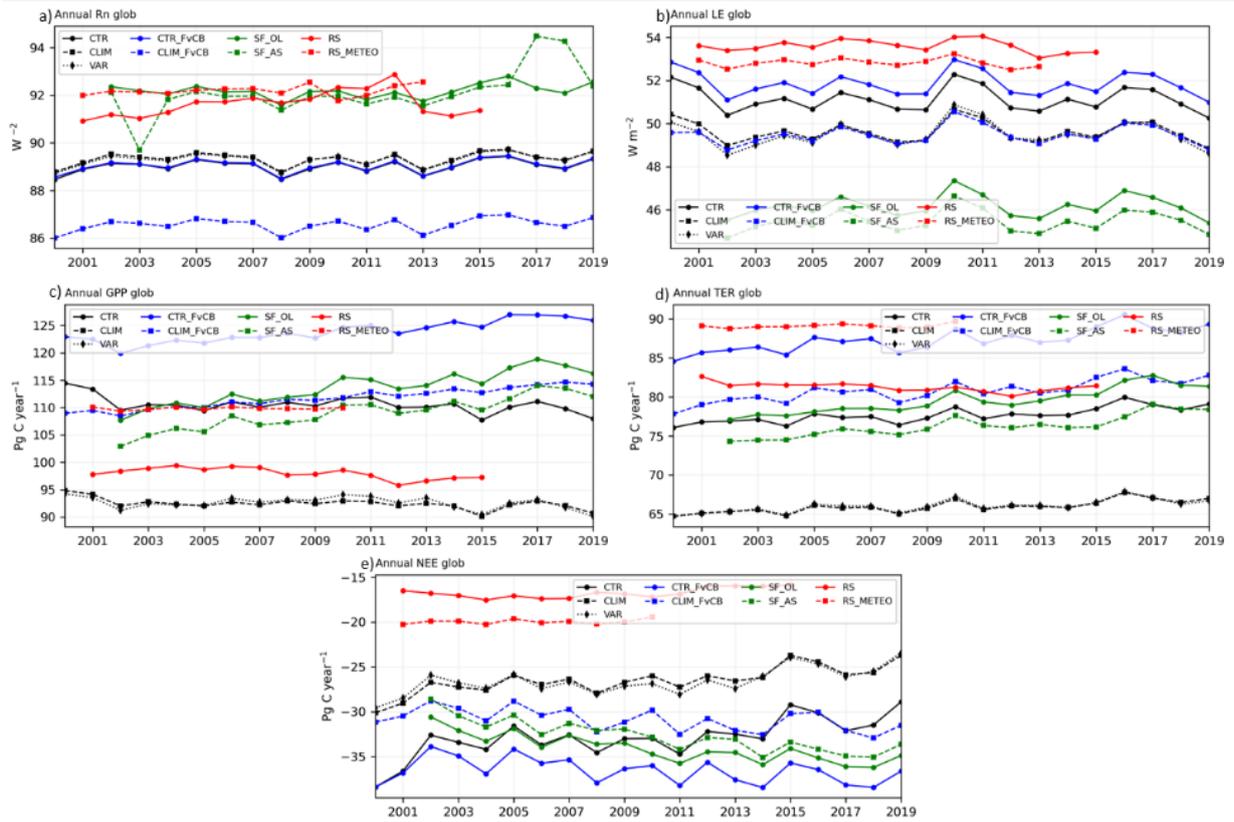
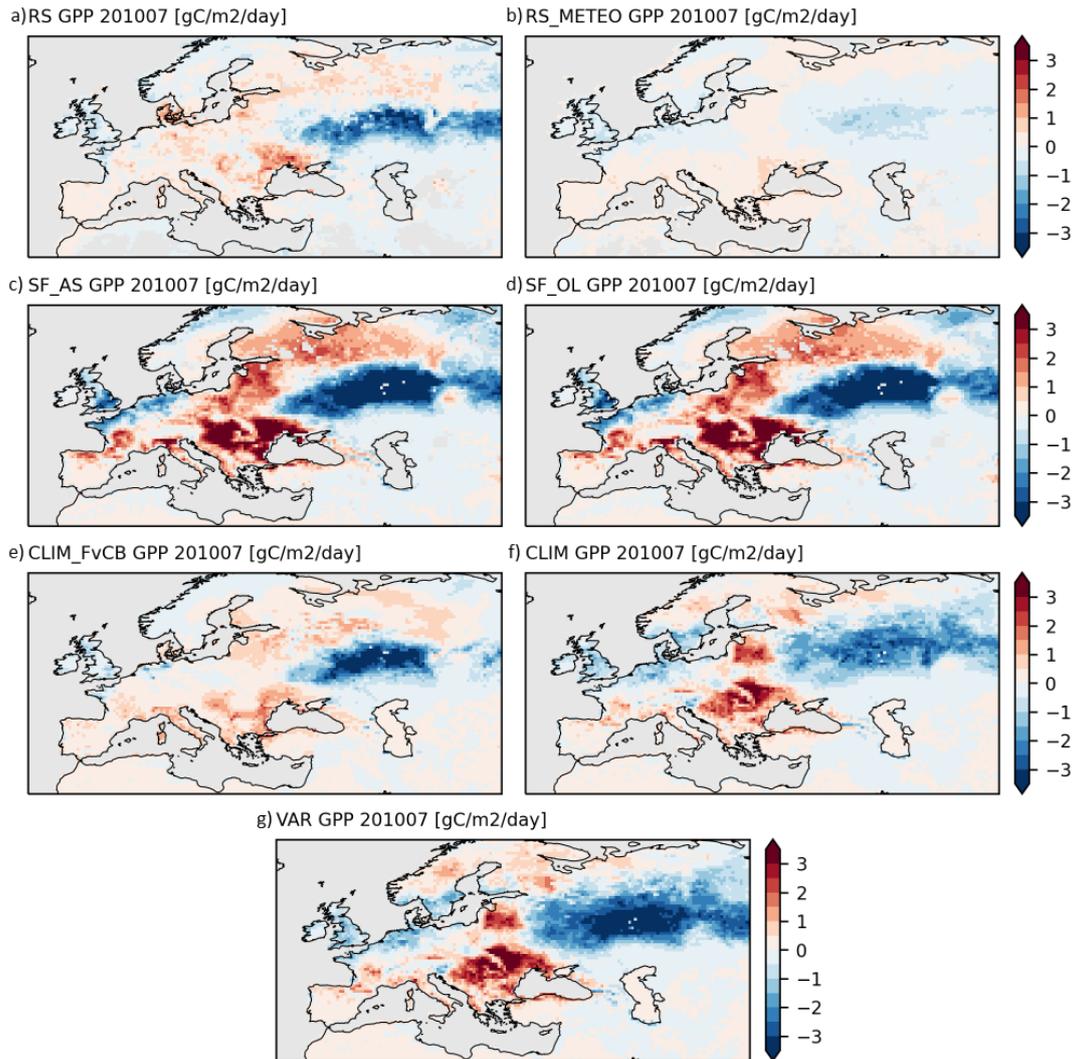


Figure 10: Global budgets considering only vegetated regions for (a) Rn, (b) LE, (c) GPP, (d) TER, and (e) NEE displayed for all 7 simulations and the two FLUXCOM products.



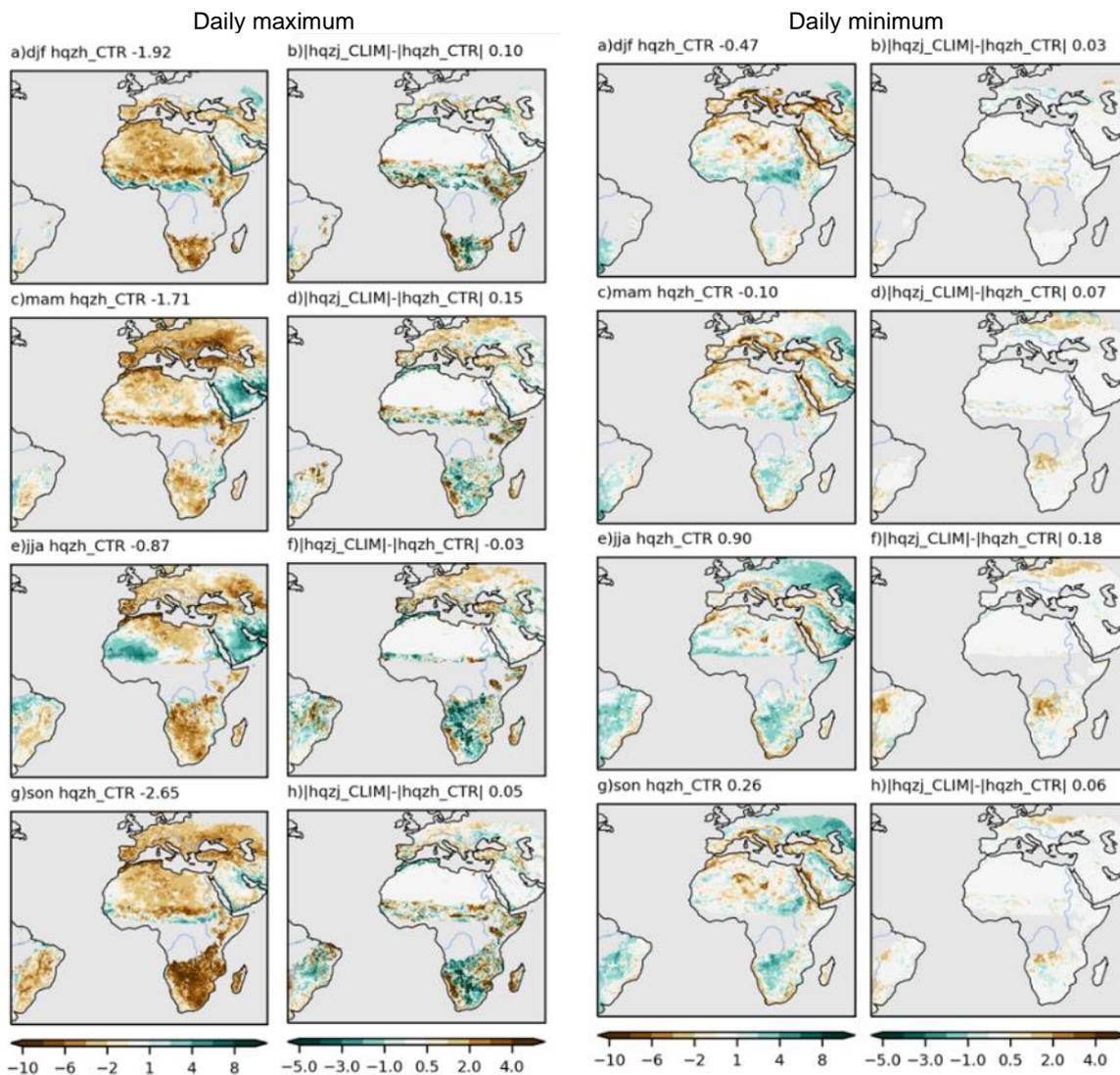
**Figure 11: Case study for July 2010 Gross Primary Production (GPP) anomaly in respect to the 2002-2010 July mean in (a) FLUXCOM RS, (b) FLUXCOM RS\_METEO, (c) SF\_AS, (d) SF\_OL, (e) CLIM\_FvCB, (f) CLIM, and (g) VAR. The anomaly is computed as July 2010 GPP minus the 2002-2010 July mean GPP, so that the blue colour indicates a reduction of GPP in 2010 in respect to the climatology.**

#### 4.2.2 Land Surface Temperature

The LST of the CTR and CLIM simulations was evaluated against the LSA SAF LST product over the period 2004-2019, which is available over the MSG disk considering clear-sky conditions. For each month the mean diurnal cycle was computed for LSA SAF and the simulations considering: (i) clear-sky conditions from the satellite data; and (ii) ERA5 total cloud cover conditions below 0.3 (Johannsen et al., 2019; Nogueira et al., 2021b). After this processing, a maximum and minimum daily LST for each month was selected for each grid-point when there is at least a minimum of 10 % of valid data. Due to this data screening, some regions (e.g. tropical forests) are not included in the analysis in the following results (depicted by the shaded grey contour within the presented Figures).

The daily maximum and minimum biases of the CTR simulations for each season is shown in **Figure 12**, along with the differences of the absolute bias between CLIM-CTR simulations. The results indicate that daily maximum LST in ECLand displays a large absolute bias over most land areas, in all seasons, particularly over Southern Europe, large portions of northern,

central and Southern Africa, and Eastern Brazil. In contrast, during JJA the Arabian Peninsula, Tropical Western Africa, and Western Brazil, display a warm bias, which is related to aerosols affecting both satellite and ERA5 radiation over the Arabian Peninsula and Tropical Western Africa. The impact of the revised land cover and LAI in the CLIM simulation (**Figure 12**) reveals mostly neutral impact in all seasons, except for spring in Europe with a marked deterioration. Despite this neutral impact, there is a coherent large bias reduction over Southern Africa.



**Figure 12: Evaluation of Land Surface temperature (LST) daily maximum (left figure) and daily minimum (right figure) values, displaying the mean bias of the control simulation (CTR) LST for each season, which include (a) DJF, (b) MAM, (c) JJA, and (d) SON, as well as the difference of the mean absolute error between CLIM – CTR also for each season (panels from b to h).**

### 4.3 Discussion

In this section offline simulations with ECLand and SURFEX model were evaluated in terms of the surface water/energy/carbon fluxes taking FLUXCOM as reference and the LST of ECLand was compared with satellite data.

The results indicate that CLIM\_FvCB has higher scores for H and GPP, when compared with the remaining simulations, and that SF\_AS scores are general higher than SF\_OL. The better

performance of CLIM\_FvCB can be partially attributed to the use of FLUXNET data in the optimization of the model as well as to the importance of using improved representation of LAI in the modelling of GPP. Lowest scores for the bias and rmse are associated with larger uncertainties in the FLUXCOM reference datasets, as well as with lower weights for the overall score calculation. A good agreement was found in terms of the annual cycle maximum in all climate regions for GPP. For S\_bias and S\_rmse there is a clear difference between Tropical / Arid and Temperate / Cold regions, with the latter having the highest scores. These results can be partially attributed to the scores calculation, since both scores are normalized by the root mean square of the reference data, which is higher in the temperate and cold climates, with a marked seasonality, when compared with the tropical regions.

There is a rise of inter-model variability in the annual budgets, increasing from the energy components to NEE. This reflects the large uncertainties and difficulties in constraining the global biogenic carbon budget, justifying flux adjustment to mitigate biases in global CO<sub>2</sub> analysis (Agustí-Panareda et al., 2016). The TER correction applied to ECLand and SURFEX simulations improved the comparison with FLUXCOM. However, this procedure resulted in a large NEE sink, compared to estimates from inversions, suggesting the need for further exploration of the TER parametrization and a better quantification of the uncertainties related to the different budget components (including multi-model/ensemble range). Finally, the analysis of the summer 2010 anomalies suggests that: (i) varying LAI conditions are relevant to the GPP estimates during large-scale extreme events: and (ii) that ECLand reproduces SURFEX anomalies when driven by time-varying LAI conditions, as it was the case in the VAR simulations with prescribed satellite LAI data.

The evaluation of the revised land cover and LAI in ECLand in terms of LST showed mostly neutral results, except for spring in Europe with a marked deterioration and a coherent large bias reduction over Southern Africa. These results are consistent with the findings of Nogueira et al. (2021b) that also suggested that the inclusion of a vegetation clumping parameterization based on LAI for low vegetation in addition to the revision of the land cover and LAI was beneficial for the LST simulations. However, in the most recent tests in coupled atmosphere forecast simulations, such change resulted in a negative impact in atmospheric scores and was not included in the presented analysis for this report. Uncertainties in the satellite data mapping of both the land cover and LAI might be also a source of error in transition regions and areas with sparse high vegetation (e.g. savannah, taiga and crops and small woodlands). Moreover, the relation between the satellite land cover and the actual model parameters used (e.g. effective vegetation cover) are also a potential source of errors and require further attention.

Ongoing developments in ECLand for the next operational IFS cycle include a revision of the handling of roughness lengths disaggregation for the tiles and post-processing of 2-metre temperature. Preliminary results (not included in the report) indicate a positive impact on LST, however, these findings require further testing with the revised land cover and LAI. Similarly, ongoing tests (performed during the preparation of the current report) with revised model parameters for TER show the potential to improve the global NEE budgets. The current scaling of TER to GPP is problematic as it implies that the sink is proportional to GPP. Adding an offset to the TER scaling is an option as well as non-linear relationship with GPP. These will be further explored to improve the TER (and NEE) budgets in the second version of this report.

## 5 Online simulations and evaluation of Numerical Weather Prediction metrics

### 5.1 Data and Methods

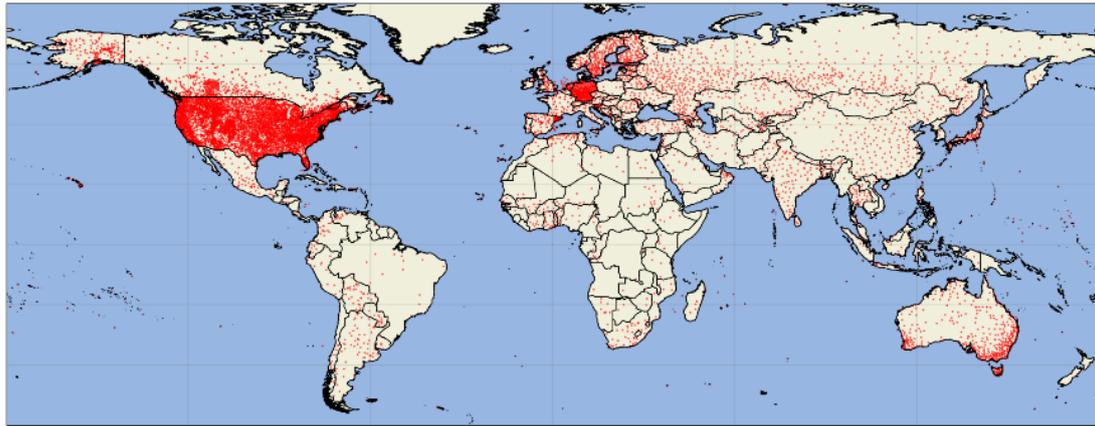
#### 5.1.1 Observations

A correct characterization of the land surface processes is fundamental for the accurate representation of physical and dynamical processes that take place near and above the Earth's surface. In this context, to evaluate the effects and impact of the land cover and LAI revision in ECLand towards its meteorological impact in short-range weather forecasts, particularly near-surface temperature, model forecasts are assessed against a set of Global Historical Climatology Network (GHCN) ground observations. To this end, GHCN daily data is extracted from the NOAA online repository<sup>1</sup>, allowing the access to in-situ records of more than 100,000 ground stations globally, with a record length ranging from less than one year to more than 175 years (depending on the station). Regarding the available atmospheric parameters, in addition to the total precipitation, snowfall, and snow depth, GHCN daily records also include maximum and minimum air temperatures values, which are updated daily and reprocessed once per week, as part of a quality control procedure. The GHCN records have been widely used for model validation (e.g. Mutti *et al.*, 2020; Kumar *et al.*, 2020; Durre *et al.*, 2010), covering a variety of locations, vegetation types and climates.

In this report, to evaluate the weather forecasts against GHCN observations, a nearest neighbour mapping of forecast data to the station location was adopted, and data screening was applied covering a period of one year (2019), including: i) only stations with more than 80 % of available data; ii) a minimum of 10 days with valid data; iii) a maximum altitude difference of 100 m between station and model altitudes; and iv) only land areas with a minimum fraction of 0.9 within each model pixel. As the result of applied filters, a total of 14625 stations were found suitable for the present analysis (**Figure 13**), where most of these are located within the United States of America (USA), being followed by Europe (EU), Australia (AUS), Russia (RUS), Indochina (ICH), India (IND), South America (SA), Central Africa (CAF), and South Africa (SAF), as depicted separately in **Figure 14**. When evaluating the weather forecast against GHCN, seasonal mean values for the daily temperature extremes are used by averaging the corresponding daily values within each season, being then subjected to an analysis where standard NWP metrics (e.g. bias and RMSE) are considered. More details concerning the GHCN and provided data are available in Menne *et al.*(2000). For comparison reasons, the seasonal means in the weather forecasts are computed only when there are available observations, matching in time, to guarantee a consistent comparison. The evaluation metrics are then computed independently for each station, considering the nearest model grid-cell, being spatially aggregated to each model grid-cell to partially homogenise the stations density.

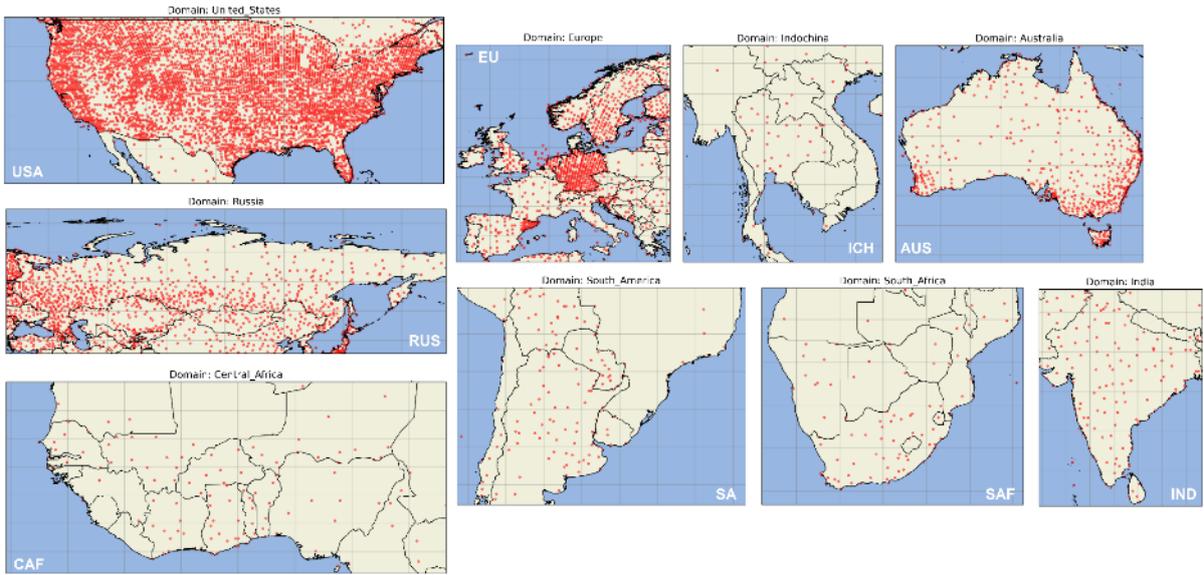
---

<sup>1</sup> <https://www.ncei.noaa.gov/products/land-based-station/global-historical-climatology-network-daily>



• GHCN stations

**Figure 13: Location of each Global Historical Climatology Network (GHCN)-Daily ground station (marked by red dots) used for the evaluation of the online simulations with the IFS new land cover model. A total of 14625 stations are considered after the application of proper data screening.**



• GHCN stations

**Figure 14: Different domains with respective Global Historical Climatology Network (GHCN)-Daily ground stations (marked by red dots) used for the evaluation of the online simulations with the IFS revised land cover model. These include the United States of America (USA), Europe (EU), Australia (AUS), Russia (RUS), Indochina (ICH), India (IND), South America (SA), Central Africa (CAF), and South Africa (SAF).**

**5.1.2 Model simulations**

As previously described in subsection 4.1.1.2, the land cover and LAI was revised for ECLand and tested in coupled IFS weather forecasts. For this report two configurations (see **Table 3**) were tested: (i) a control configuration (*CTR0*), which is used as reference for the entire evaluation procedure considering the current default configuration of the model, and (ii) a configuration using the revised land cover and the LAI climatology (*CC10*). It is worth noting that several extra configurations of land cover and LAI were also tested but not documented

in this report. The simulations consist in a set of medium-range weather forecasts with a 5-day lead time initialized at 00UTC on all days of 2019. Atmospheric initial conditions were taken from ERA5 and land initial conditions from the surface simulations with the same configurations, i.e., CTR0 was initialized with land conditions from CTR (see [Table 1](#)) and CCI0 from land conditions from CLIM (see [Table 1](#)).

Regarding the evaluation of the forecasts to assess the meteorological impact of CCI0, the ECMWF software package IVER (Geer, 2015) was used to compute the NWP scores and compare CCI0 to CTR0. Specifically, an assessment of several meteorological variables, affected by changes made in CCI0, is performed by using the normalised difference in RMSE (DRMSE), where blue (negative) areas indicate lower RMSEs (i.e., better scores) in CCI0 in comparison with CTR0. The change in the error is normalized by the RMS error of the control (CTR0) forecasts, representing the fractional change in the error, for example a value of -0.1 means a reduction of 10 % of the error in CCI0. This assessment is carried out considering different regions of the globe, including: the Southern Hemisphere (SH), between  $-90^{\circ}$  and  $-20^{\circ}$ ; the Tropics, between  $-20^{\circ}$  and  $20^{\circ}$ ; and the Northern Hemisphere (NH), between  $20^{\circ}$  and  $90^{\circ}$ , in which aggregated means for each region are considered. In this evaluation, the operational NWP analysis was taken as reference to compute the scores, which for the winter season only considers the months of January and February 2019.

**Table 3: Weather forecasts simulations configuration.**

Simulation	Setup Details
CTR0	Control simulation, with default configuration.
CCI0	Simulation with the revised land cover and LAI climatology.

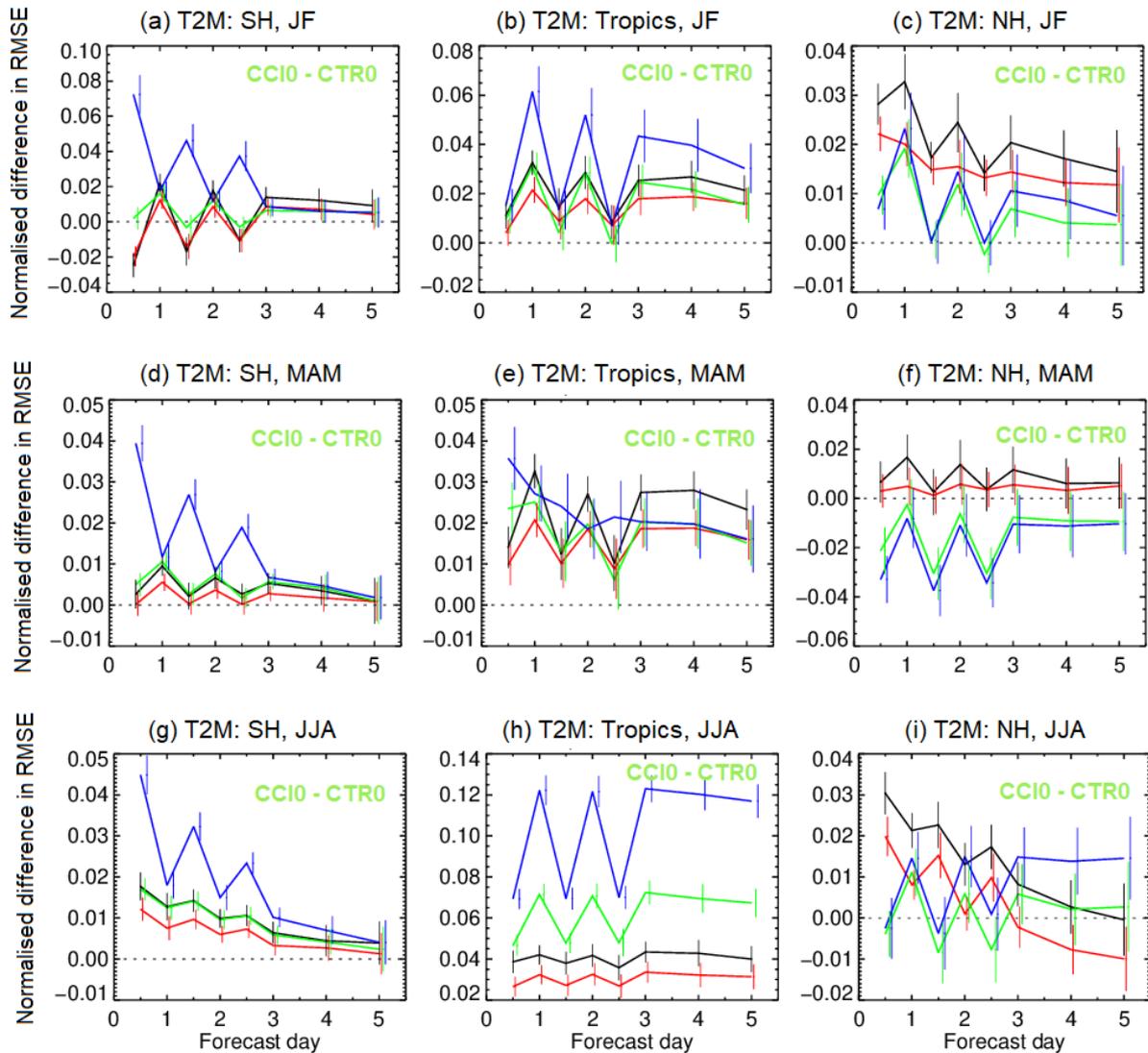
## 5.2 Results

### 5.2.1 Meteorological impact

The starting point for this evaluation consists in assessing the regional scores found between the CCI0 and CTR0 for the 2-metre air temperature (T2M), as shown by the corresponding aggregated means (green lines) in [Figure 15](#). It should be noted that, as previously mentioned, different land cover iterations were performed before reaching the current configuration. Therefore, regional scores from previous experiments are also shown in the results, but not discussed in detail. The black and red lines in [Figure 15](#) to [Figure 17](#) are from an initial version of the land cover mapping and cross-walking table with the red line a test of changing roughness lengths, while the blue line was from an experiment with active clumping for low vegetation (Nogueira et al., 2021b).

The first noticeable feature in all plots concerns the variability of the forecast error, where beyond 72-hour ahead forecasts the variation between CCI0 and CTR0 deviations tends to be smoother than within shorter forecasting ranges. It is worth noting that, although the forecast variability between 24, 48, and 72-hour ahead forecasts does not show a significant change in the differences between the two simulations (the same behaviour is observed between the 12, 36, and 60-hour head forecasts), the deviations between the two simulations at the 12, 36, and 60-hour ahead forecasts are shown to be smaller (tending to decrease as the forecast horizon increases) than at the 24, 48, and 72-hour ahead forecasts. Generally, DRMSE regional scores show that CCI0 has better scores (i.e., lower errors) than CTR0 during the spring period, especially in the NH region ([Figure 15f](#)), where CCI0 reduces the errors (between 0 and -0.03 approximately), at all lead times. Regarding CCI0 lower scores (i.e., higher errors) in comparison to CTR0, these are generally found to be more significant in the Tropics, particularly during the summer period ([Figure 15h](#)), where CCI0 values are predominantly positive (between 0.04 and 0.07, approximately). T2M forecasts from CCI0 also show errors during the winter period in all regions (particularly in the Tropics, [Figure 15b](#)),

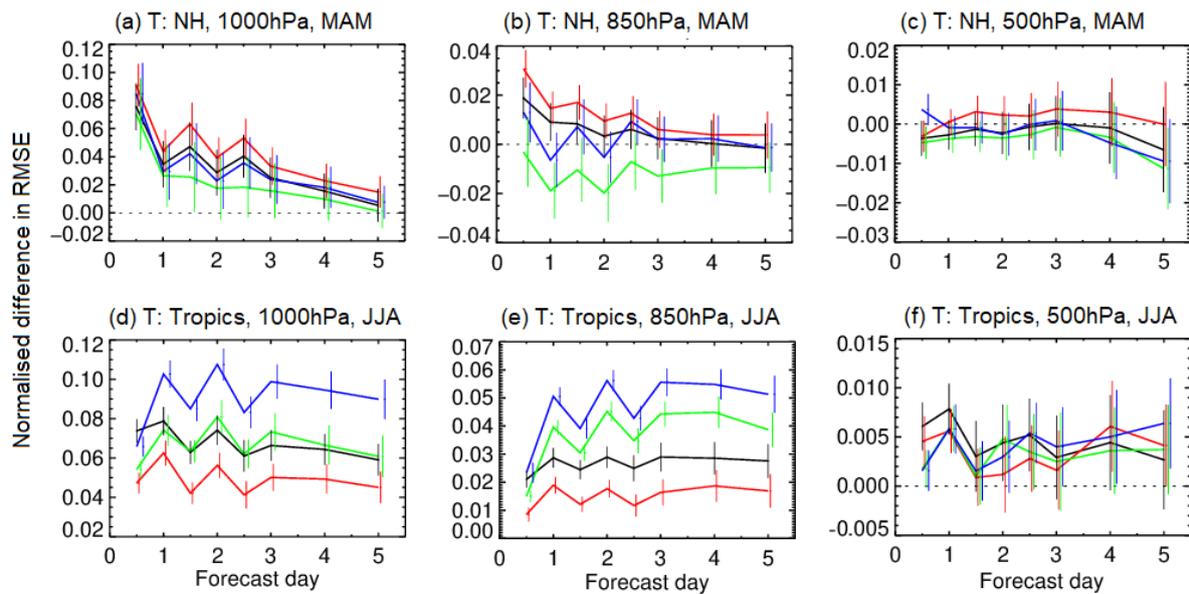
where DRMSE are generally positive (between 0 and 0.03, approximately). Although this analysis is informative, the use of the operational NWP analysis as reference brings some limitations, in particular over regions with a small coverage of observations. In those areas the NWP analysis are mostly dominated by the model background, which will tend to benefit the CTR0 simulation. In the following section, a more detailed evaluation with the direct use of in-situ GHCN observations will provide a better insight into the model performance.



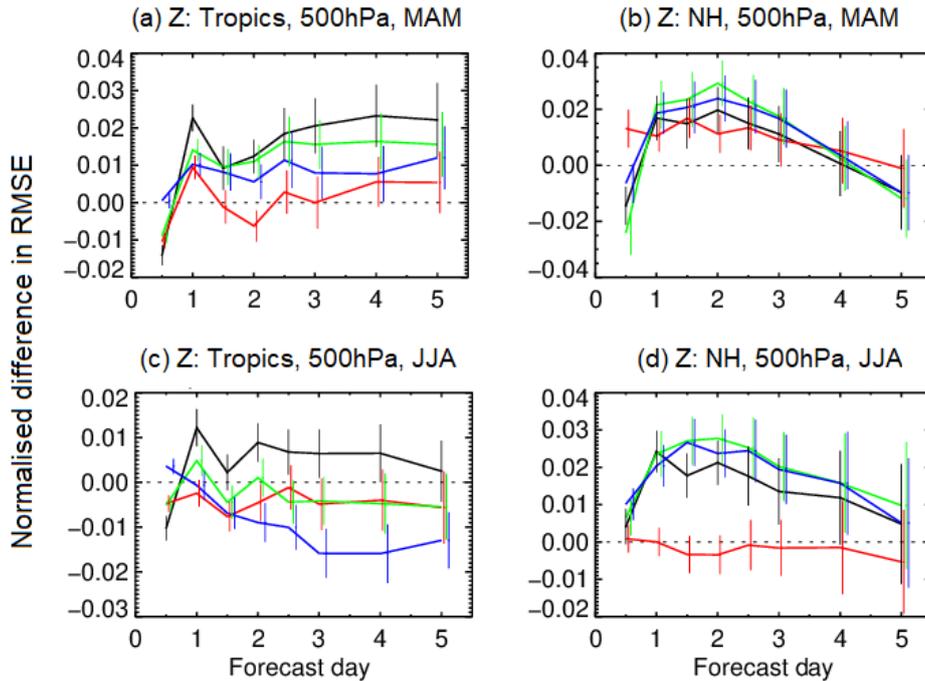
**Figure 15: Aggregated means for the 2-metre temperature (T2M) normalised RMSE differences between CCI0 and CTR0 forecasts (depicted by the green lines) for winter (JF), spring (MAM), and summer (JJA) seasons in 2019 with lead times up to 5 days. Regions considered: Southern Hemisphere (SH), between  $-90^{\circ}$  and  $-20^{\circ}$ ; Tropics, between  $-20^{\circ}$  and  $20^{\circ}$ ; and Northern Hemisphere (NH), between  $20^{\circ}$  and  $90^{\circ}$ . Confidence range of 95% with AR(1) inflation and Sidak correction for 16 independent tests. Additional experimental forecasts from previous versions are also depicted (black and red with a previous version of the land-cover mapping, and blue with low vegetation clumping).**

Following these results, and to further assess the CCI0 performance in producing skilful near-surface forecasts, additional analyses towards other meteorological fields are performed focusing on the spring and summer periods. Air temperature (T) at different pressure levels (1000, 850, and 500 hPa) are also evaluated for the NH and the tropical regions during spring and summer periods (**Figure 16**). In the NH, during spring, high deviations towards air temperature at 1000 hPa (**Figure 16a**) occur between CCI0 and CTR0, where CCI0 errors are found to be larger than in CTR0 for shorter forecasting ranges, with values between 0.07 and

0.03. Again, these errors near the surface can be partially attributed to the use of the NWP analysis. At 850 hPa, CCI0 has an opposed behaviour, with improvements towards CTR0 (**Figure 16b**), where values are found between 0 and -0.02, which is also observed at 500 hPa although with less significance (**Figure 16c**). In the Tropics, during the summer, larger air temperature errors generally occur in CCI0 (i.e., DRMSE values are predominantly positive). At 1000 and 850 hPa (**Figure 16d** and **Figure 16e**, respectively), significantly larger deviations are found, with a maximum (about 0.08) and minimum (about 0.04) difference respectively found between the two simulations, while at 500 hPa (**Figure 16f**) the deviations closer to zero. When looking at the DRMSE for the geopotential height (Z) at 500 hPa (**Figure 17**), it is possible to observe a similar behaviour in the NH during spring and summer (**Figure 17b** and **Figure 17d**, respectively). In particular, shorter forecasting ranges (i.e., at 12-hour ahead) show lower CCI0 errors in spring (about -0.02) than in summer (about 0.01), while a maximum error (about 0.03) is found at the 48-hour ahead forecasts in both periods. Regarding the tropical region, Z differences are generally lower between the two simulations, although CCI0 shows higher errors than CTR0 during spring (except at 12-hour ahead), while lower CCI0 errors tend to be produced during summer (**Figure 17a** and **Figure 17c**, respectively). The impact on the NH geopotential is likely due to changes in momentum transfer induced by the reduction of high vegetation and requires further attention and possible some changes in roughness lengths.

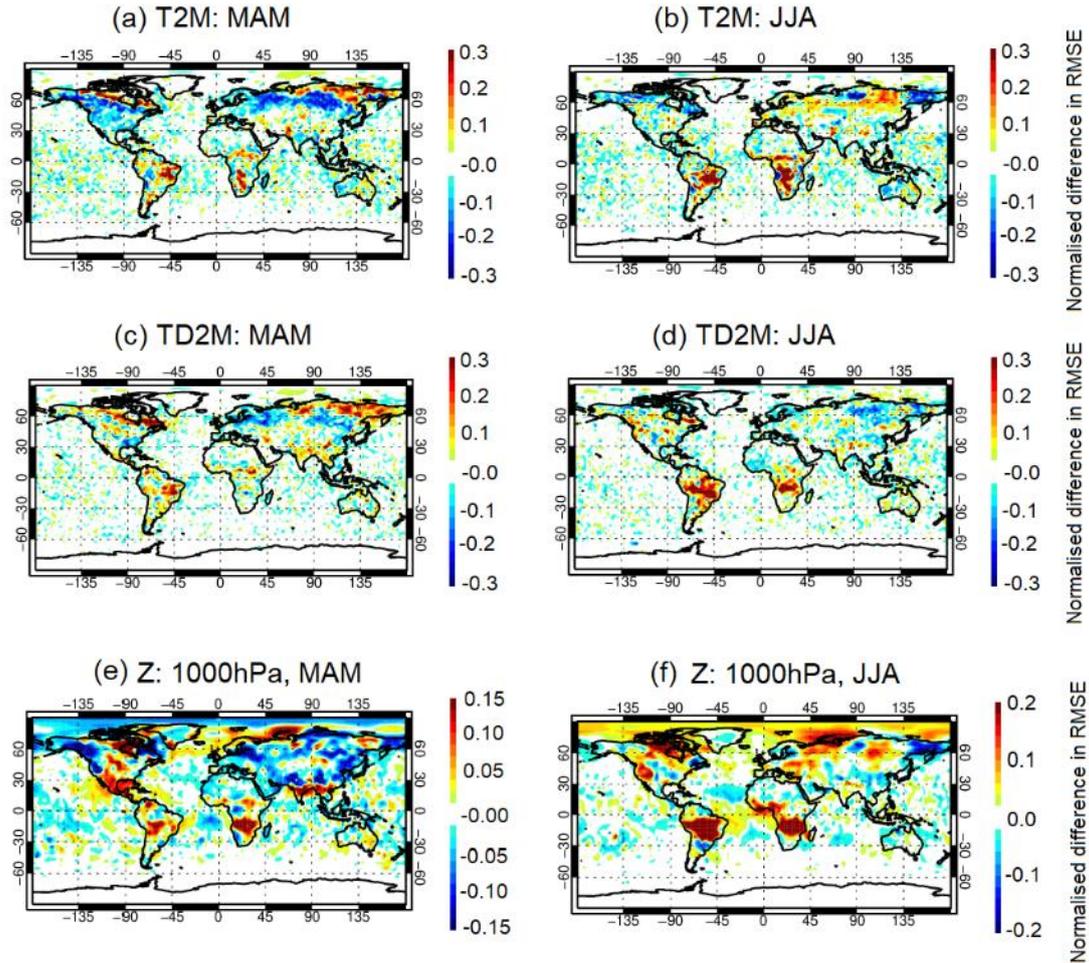


**Figure 16: As Figure 15 but for temperature (T) during spring in the Northern Hemisphere (NH) (a-c) and summer in the Tropics (d-f) at different pressure levels: 1000hPa (a, d), 850 hPa (b, e), and 500 hPa (c, f).**



**Figure 17:** As Figure 15 but for the geopotential height (Z) at 500hPa during spring (a, b) and summer (c, d) for the Tropics (a, c) and Northern Hemisphere (NH) (b, d).

In addition to the previous aggregated scores for the NH and Tropics, the following global difference maps of DRMSE between CCI0 and CTR0 for the 2-days lead time are discussed in more detail. **Figure 18** shows the different regions where improved (blue contours) and deteriorated (red contours) performances arise in CCI0 in several surface fields. For the T2M forecast, CCI0 improvements occur in the NH, particularly during spring over Eastern EU, RUS, and USA (**Figure 18a**), while lower scores are found at higher latitudes, namely in Northern USA and Canada, and North-eastern RUS. During the same period, in the Tropics, T2M deviations are higher, showing more errors for CCI0 over the SA, CAF, and SAF regions. A similar behaviour is observed in the dewpoint temperature (TD2M) forecast (**Figure 18c**). During the summer period, the tropical regions are characterized by an increase of the error in both T2M and TD2M (**Figure 18b** and **Figure 18d**, respectively). Additionally, the impact of CCI0 on the geopotential height at 1000hPa (a proxy for surface pressure) is shown in **Figure 18e** and **Figure 18f** for the spring and summer periods, respectively. Although there is a reduction of the errors in CCI0 over most of the NH in spring, there is also a marked deterioration in Eastern Brazil and South Africa, being further amplified in summer. These results identify a consistent seasonal and geospatial signal with a positive impact of CCI0 in the NH in spring and negative in the tropical region, particularly during summer. The negative impact in the Eastern Brazil and south-central Africa regions is likely associated with the transition of high vegetation cover to low vegetation in those regions and the associated impact on surface roughness for momentum and heat. Further investigations are needed to identify the root causes and to mitigate its impact, as well as data assimilation experiments (which are computationally expensive) to exclude possible signals arising from model changes in regions with a reduced constrain by the data assimilation system.



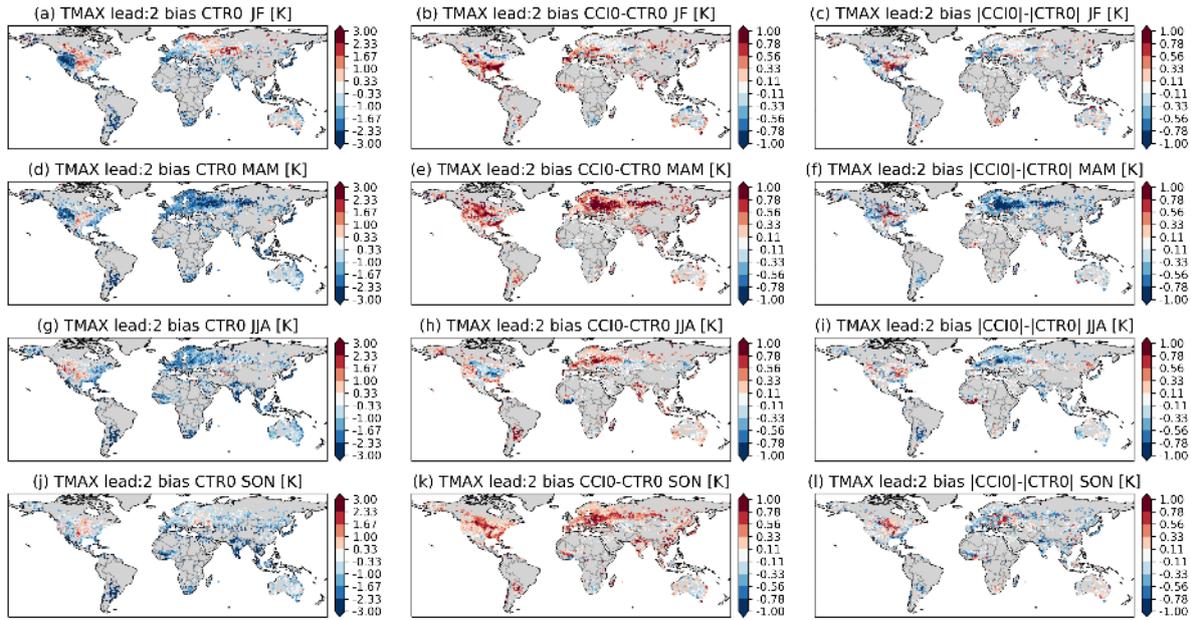
**Figure 18: Normalised Root Mean Square Error (RMSE) differences between CCI0 and CTR0 forecasts for 2-days lead time forecasts of 2-metre air temperature (T2M a, b), 2-metre dewpoint temperature (TD2M, c, d) and geopotential height (Z) at 1000hPa (e, f), for spring (MAM, left panels) and summer (JJA, right panels) periods in 2019. No statistical significance testing was applied.**

## 5.2.2 Impact on near-surface temperature

In this section, GHCN ground measurements are used to further evaluate the impact that the revised land cover and LAI has over near-surface temperature forecasts. Considering the adopted data screening procedure (described in subsection 5.1.1) between the forecast and GHCN daily values, seasonal mean values for the daily maximum and minimum temperatures (TMAX and TMIN, respectively) are used. It should be noted that although up to 5 days lead-time forecasts have been analysed in the previous section, only the 2 days lead-time forecasts are hereafter used, since TMAX and TMIN forecast errors at different lead times do not change significantly (not shown).

The global bias maps of TMAX and TMIN (**Figure 19** and **Figure A3**, respectively) show that the biases in CTR0 are mostly negative for TMAX and positive for TMIN, indicating an underestimation of the diurnal cycle range. The differences between CCI0 and CTR0 show that the impact of the revised land cover and LAI is stronger in TMAX. These differences (middle panels), show mostly positive changes, i.e., CCI0 is warmer, with values over 1.00 K in the USA and Eastern EU regions, particularly during the spring period (**Figure 19e**). This behaviour is further noted in the results obtained towards the magnitude change in the bias (right panels), where CCI0 improvements (i.e., higher scores) are depicted with blue, while

CCI0 higher errors are represented by the red colours. It is worth noting the presence of a warm/cold bias effect in the USA region throughout the entire experimental period, where in winter (**Figure 19a**) the forecasts produced by CTR0 depict an overestimation of TMAX towards GHCN in Central USA with values between 0.33 and 2.33 K, while a significant underestimation is obtained in the western region with values reaching - 3.00 K. In spring (**Figure 19d**), CTR0 shows a high bias underestimation, occurring in Eastern EU and RUS, with values going from (close to) null to -3.00 K. Other notable (and frequent) features concern CTR0 underestimation towards GHCN in lower sampling regions, such as in CAF, SA, and AUS. When evaluating the performance of CCI0 against CTR0, it is possible to observe areas where the revised land cover and LAI shows improvements. In particular, when compared with GHCN, lower bias deviations are attained in CCI0 than with CTR0 in: most of Eastern EU and RUS, especially during spring (**Figure 19f**), while slightly lower bias deviations are obtained in Southern EU during the remaining seasons; Western USA, particularly during winter and spring (**Figure 19c** and **Figure 19f**, respectively); and in SA throughout the entire year. Additionally, for the case of AUS, there is a mixed bias variation during the entire experimental period, where in summer there seems to be very small improvements near coastal areas and (to some extent) higher errors in the central region of the continent (**Figure 19i**). Improvements are also found with CCI0 in ICH, particularly for the spring and autumn periods (**Figure 19f** and **Figure 19i**, respectively), which cover the typical monsoon rainfall months of the region. Improvements with CCI0 are also obtained in IND for spring, summer, and autumn (**Figure 19f**, **Figure 19i**, and **Figure 19l**, respectively), which cover the three typical monsoon periods occurring in the region. On the other hand, periods of lower performances (i.e., higher errors) with CCI0 are also noted, namely in: Central USA throughout the entire year, having less significance during summer; SAF during winter (**Figure 19c**); CAF during summer (**Figure 19i**); and over a small area in Eastern EU during autumn (**Figure 19i**). Regarding TMIN differences (**Figure A3**), as previously mentioned, there is a general bias overestimation in CTR0 forecasts towards GHCN throughout the entire year (left panels), with higher values being reached in the autumn over the Eastern USA, Eastern EU, SA, SAF, ICH, and AUS regions (**Figure A3j**). The previous noted warm/cold bias effect in the USA region is still visible, particularly in winter and spring (**Figure A3a** and **Figure A3b**, respectively). In terms of CCI0 improvements for TMIN, the magnitude changes in the bias forecasts (right panels) show that there are lower deviations between CCI0 and CTR0 than for TMAX. Few noticeable features are observed for TMIN forecasts, such as the lower bias error in CCI0 towards GHCN than in CTR0 over Eastern EU and RUS during spring (**Figure A3f**), and Western USA during summer (**Figure A3i**), with values reaching about -1.00 K. In addition to these results, CCI0 depicts a slightly general increase of the deviations in comparison to CTR0 through the entire period of study, particularly in RUS, Eastern USA, SAF, ICH, and Western AUS.

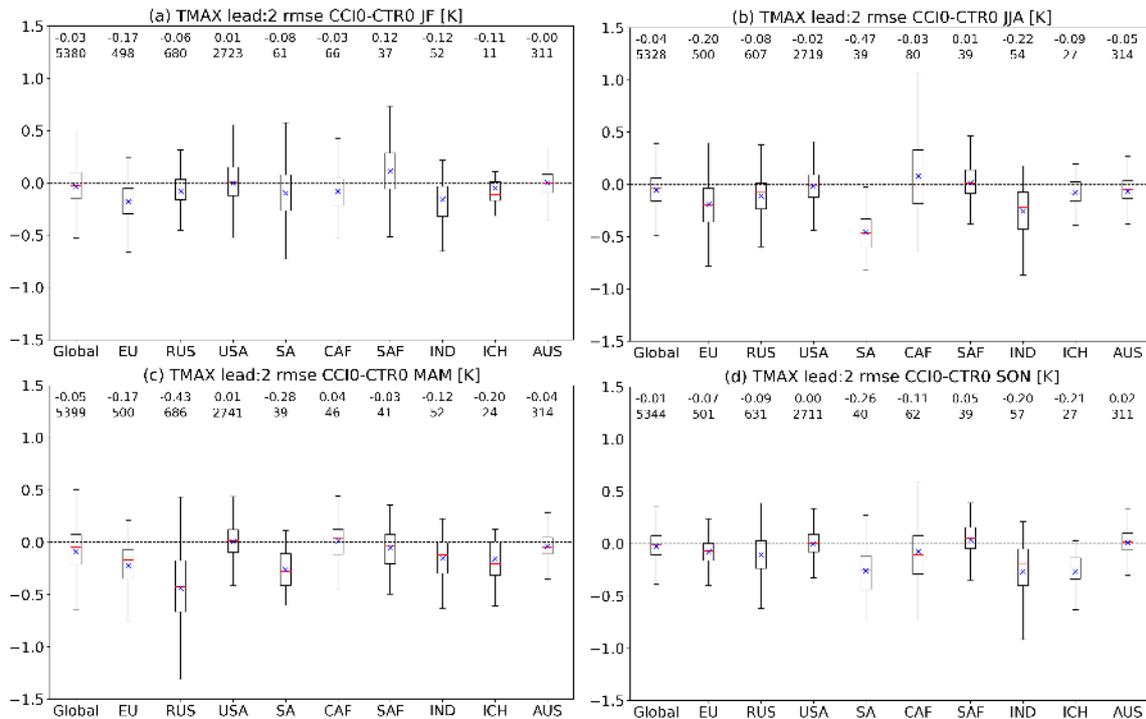


**Figure 19: Daily maximum temperature (TMAX) for 2 days lead-time forecasts mean bias differences: CTR0-CHCN (left panels), CCI0-CTR0 (centre panels) and absolute bias changes between CCI0 and CTR0:  $\text{abs}(\text{CCI0-GHCN})-\text{abs}(\text{CTR0-GHCN})$  (left panels) for winter (JF a, b, c), spring (MAM d, e, f), summer (JJA g, h, i), and autumn (SON j, k, l) of 2019.**

The previous results are further explored through a series of boxplots showing the differences of the RMSE between CCI0 and CTR0, i.e.,  $\text{RMSE}(\text{CCI0,GHCN})-\text{RMSE}(\text{CTR0,GHCN})$  for TMAX and TMIN (Figure 20 and Figure A4., respectively). Within each set of boxplots, a statistical summary is provided for all data (Global), and separately for the nine different regions, where each individual domain is delimited in Figure 14. This includes the respective mean and median of the distributions obtained for each region (depicted by a red line and blue cross, respectively), while on top of each boxplot is the corresponding median and sampling values. It should be noted that due to the spatial heterogeneity of the station's location, the distributions of the metrics over the different regions does not represent a homogeneous spatial sampling.

As expected, for TMAX (Figure 20), there are generally lower errors being produced with CCI0 than CTR0, as shown by the overall negative values for the Global outputs throughout the experimental period, with the highest median score (i.e., the lowest value of about -0.05 K) being reached during spring (Figure 20c), where CCI0 has shown to provide best performances. The CCI0 performance in producing TMAX forecasts for the remaining seasons is characterized by a reduction of the changes from summer (about -0.04 K), winter (about -0.03 K), to autumn (about -0.01 K), as shown by Figure 20b, Figure 20a, Figure 20d (respectively). In terms of behaviour for the entire experimental period over the different regions, RMSE differences show that CCI0 has lower errors than CTR0 in: EU during most of the year, especially during summer (about -0.20 K); RUS during spring (about -0.43 K); SA during summer (about -0.47 K) and autumn (about -0.26 K). Lower scores (i.e., higher errors) are found with CCI0 in SAF during winter, summer, and autumn (about 0.12, 0.01, and 0.05 K, respectively), and CAF during spring (about 0.04 K). Moreover, the effects of the previously observed warm/cold bias in the USA region are also shown here through an apparent neutral impact, with the distribution being centred around zero in all seasons. Similarly, the errors found in AUS also show an overall neutral impact. In comparison to TMAX, the impact of CCI0 on TMIN tends to be smaller, and mostly neutral. Overall, for TMIN, the lowest errors between both simulations are found with CCI0, except in autumn (Figure A4:d), where there is a general deterioration (about 0.06 K). For the different regions, during winter (Figure A4:a),

there are lower errors in CCI0 than CTR0 for the ICH region (about -0.07 K), with the same values being reached in RUS during spring. As opposed to these, CCI0 higher errors for TMIN forecasts are generally found during the autumn (Figure A4:d), particularly in USA and SAF (about 0.10 and 0.06 K, respectively), as well as in summer (Figure A4:b) in SAF and SA (about 0.10 and 0.06 K, respectively).



**Figure 20: Boxplots for the daily maximum temperature (TMAX) Root Mean Square Error (RMSE) differences (in K) between CCI0 and CTR0 forecasts for (a) winter (JF), (c) spring (MAM), (b) summer (JJA), and (d) autumn (SON) of 2019. The location of mean and median values is depicted inside each boxplot by a blue cross and a red line (respectively), while at top a reference is made to the median and sampling values used for each region: all data (Global), Europe (EU), Russia (RU), United States of America (USA), South America (SA), Central Africa (CAF), South Africa (SAF), India (IND), Indochina (ICH), and Australia (AUS). The forecast lead time of 2 days is considered.**

A more detailed analysis is also performed towards the new land cover corrections to infer possible sources of error linked to the previous warm/cold bias effect (more pronounced for TMAX than TMIN) observed over the USA region, which occurs throughout the entire year. Scatter density plots of temperature differences between CCI0 and CTR0 bias are produced for both TMAX (Figure 21, left panels) and TMIN (Figure A5, left panels) as function of the CTR bias. This allows us to better understand the positive/negative impacts of CCI0 considering the pre-existing biases in CTR0. Based on this analysis, four general conditions (referred here as Q1, Q2, Q3, and Q4) are established to identify the corresponding stations that show CCI0 higher and lower bias errors in comparison to CTR0, as shown by the following relations:

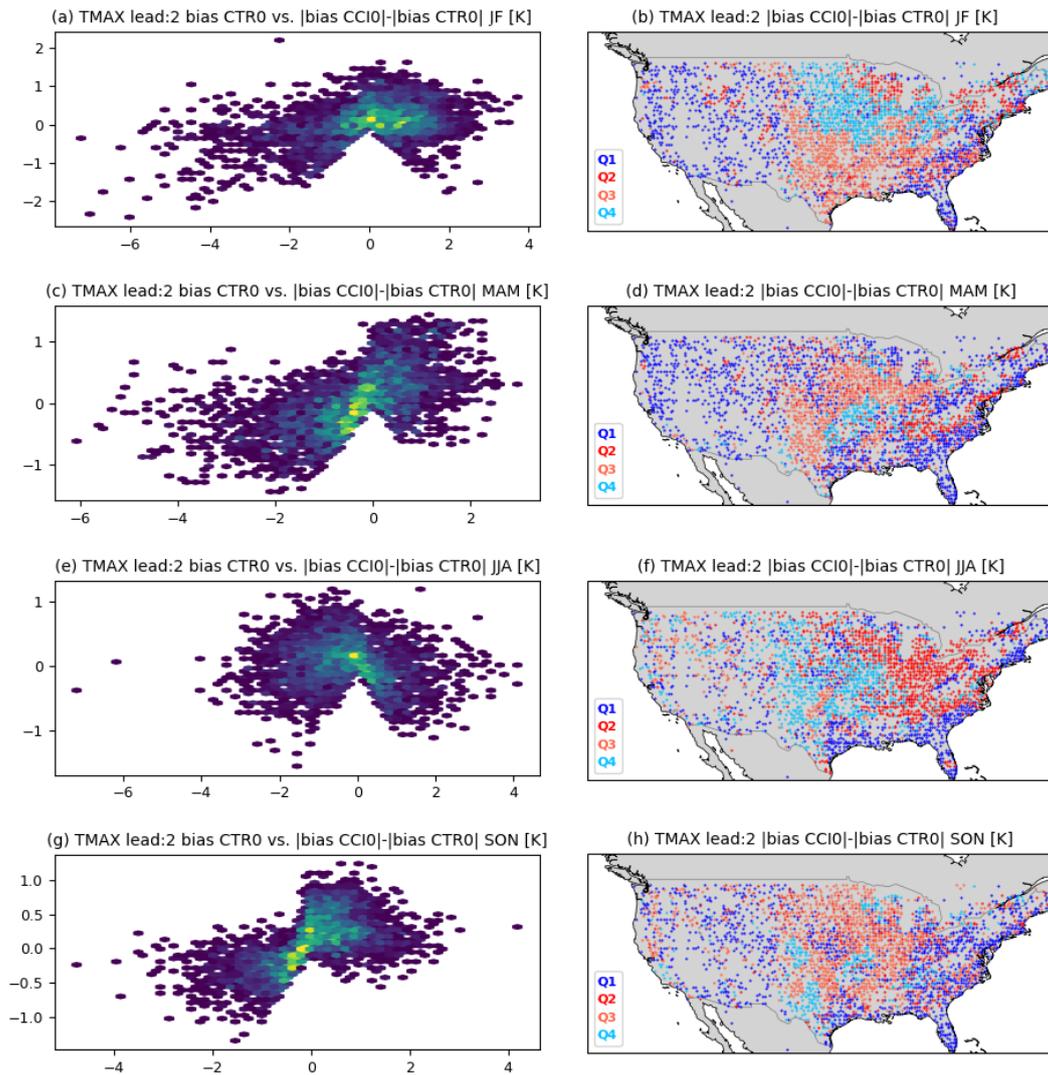
- Q1:  $\text{bias CTR0} < 0.1$  &  $|\text{bias CCI0}| - |\text{bias CTR0}| < 0.1$ ;
- Q2:  $\text{bias CTR0} < 0.1$  &  $|\text{bias CCI0}| - |\text{bias CTR0}| > 0.1$ ;
- Q3:  $\text{bias CTR0} > 0.1$  &  $|\text{bias CCI0}| - |\text{bias CTR0}| > 0.1$ ;
- Q4:  $\text{bias CTR0} > 0.1$  &  $|\text{bias CCI0}| - |\text{bias CTR0}| < 0.1$ ;

where the difference between the absolute values of CCI0 bias and the absolute values of CTR0 bias represents the magnitude change in the temperature bias. Both Q1 and Q4 report stations where CCI0 performs better than CTR0 (i.e., lower errors), while the opposed behaviour is shown in Q2 and Q3 conditions. The corresponding stations are then depicted in regional maps of the USA region according to each performance condition for TMAX (**Figure 21**, right panels) and TMIN (**Figure A5**., right panels). Additionally, since there is a higher impact of the new model corrections over TMAX forecasts than for TMIN, a summary regarding the number of stations found within each condition is presented for TMAX in **Table 4**.

Considering the previous adopted conditions, for TMAX forecasts it is possible to notice that in the USA region CCI0 generally has better performances than CTR0 during spring. In particular, for this period, a total of 1821 stations (**Figure 21d**), i.e., the sum of stations found within conditions Q1 and Q4 (**Table 4**), show that improvements are mostly located in the eastern and western regions, corresponding to more than half of the available stations during this season. At the same time, despite a lower number of stations showing higher errors in CCI0 (i.e. a total of 1349, considering the sum of stations found within conditions Q2 and Q3), the revised land cover bias seems to increase more in the central region of the country. This behaviour is similarly observed in the remaining seasons, being more significant during the autumn period (**Figure 21h**). Regarding the obtained results for TMIN (**Figure A5**), there is no clear signal between the pre-existing biases in CTR0 and the magnitude change in the temperature bias between CCI0 and CTR0. There is an overall improvement with CCI0 for the USA region, particularly during the summer period (**Figure A5:f**), while CCI0 higher errors are found in the autumn period (**Figure A5:h**). Despite achieving an overall improvement towards the representation of near-surface temperature with CCI0, these results show that there are still locations with systematic errors that the revised land cover is not able to resolve. The relation that such errors can have with the different land cover components was also explored through additional analysis (not shown). In particular, a comparison analysis towards the use of different ranges of vegetation coverage parameters (e.g. total LAI, total vegetation, types of vegetation, etc.) and the impact of these in the resulting temperature forecasts between CCI0 and CTR0, revealed that there is no clear relation between the warm/cold bias effect and the vegetation parameters selected for testing.

**Table 4: Summary for the number of samples (#) found within each performance condition (Q1, Q2, Q3, and Q4) for the IFS land cover model defined between control and revised simulations of TMAX, including the respective percentage in reference to the total sampling (i.e., total number of stations found in the USA region) for each season in 2019.**

		<b>Q1</b>	<b>Q2</b>	<b>Q3</b>	<b>Q4</b>
<b>JF</b>	Samples (#)	982	529	941	696
	Total (%)	31.19	16.80	29.89	22.11
<b>MAM</b>	Samples (#)	1520	545	804	301
	Total (%)	47.95	17.19	25.36	9.50
<b>JJA</b>	Samples (#)	909	1094	303	838
	Total (%)	28.91	34.80	9.64	26.65
<b>SON</b>	Samples (#)	1307	389	1116	324
	Total (%)	41.68	12.40	35.59	10.33



**Figure 21: (left panels) scatter plots for the daily maximum temperature (TMAX) bias magnitude change between CCI0-CTR0 (vertical axis) as function of the CTR0 bias (horizontal axis); (right panels) regional maps (USA) for TMAX bias differences between CCI0 and CTR0. Results are displayed for winter (JF a, b), spring (MAM c, d), summer (JJA e, f), and autumn (SON g, h) of 2019, with the forecast lead of 2 days. All units are in K.**

### 5.3 Discussion

The meteorological impact of CCI0 in comparison with CTR0 in terms of NWP forecast scores was assessed using the ECMWF IVER tool, while a more detailed analysis focusing on daily maximum and minimum 2-metre air temperature was performed with in situ observations from GHCN as reference. The T2M results were mostly coherent, showing a clear improvement in CCI0 in Eurasia during spring with higher errors found in the Tropics, particularly during summer. Specifically, temperature improvements (i.e., with lower errors), including TD2M, air temperature up to 500 hPa, are essentially obtained in the NH, namely in USA, EU, and RUS. Despite these improvements, systematic errors in temperature forecasts persist throughout the entire experimental period of 2019 (at very high latitudes), having a significant increase during the summer over tropical regions. The main areas that experience a negative impact in CCI0 are in: SA (particularly North-eastern Brazil) during spring, while most of the country

is affected negatively during the summer; and central-south Africa. Moreover, the impact on 1000 hPa geopotential height is also consistent with the temperature errors.

Considering the daily maximum and minimum temperature evaluation with GHCN, results show a general improvement in CCI0 in comparison to CTR0, particularly for TMAX. When compared with GHCN daily values, TMAX and TMIN forecasts from CTR0 are characterized by an overall under and overestimation, respectively, indicating an under-estimation of the diurnal cycle range. A closer analysis for TMAX over different regions shows how CCI0 has lower errors than CTR0 in EU and RUS during most of the experimental period, particularly in spring, and in Western USA, during winter and spring. For the case of TMIN, CCI0 improvements are found in EU, mostly during spring, and Western USA, during summer. During the same period, CCI0 higher errors are also observable in some regions, having a mixed behaviour throughout the entire experimental period. Additionally, it is worth noting the presence of a warm/cold bias anomaly occurring in the USA region for both TMAX and TMIN. The general warming of CCI0 results in a mixed impact, with regions with a prior cold bias improving, while regions with a prior warm bias deteriorating. Despite significant advances in improving the IFS land cover and LAI realism there are still systematic errors present in the model's forecasts, which can be related to the tuning of the model and respective parameter changes. Although positive impacts were found, there are still some regions that require further attention but also more detailed consideration for regions with poor observation network.

## 6 Conclusion

This deliverable was produced in the framework of the CoCO2 project aiming at the development of a new observation-based operational anthropogenic CO<sub>2</sub> emissions Monitoring and Verification Support capacity. Special attention was given to the improvement of the ECMWF's Integrated Forecasting System (IFS) and its land surface component, ECLand, via the revision of the land cover and LAI representation using new satellite information. This revision was evaluated in terms of the meteorological impact in NWP and in surface-atmosphere exchanges of carbon fluxes, including SURFEX simulations and using FLUXCOM as reference. In addition to the evaluation of the model developments, the work carried out in this report sets the foundations for an integrated NWP and biogenic fluxes assessment of the system, which will be used throughout the project and beyond since the technical infrastructure has been developed under a version control system and is available on request by any project partner interested.

The surface offline simulations evaluation identified the added values of the revised land cover and LAI in term of GPP when used in conjunction with a model configuration using the Farquhar photosynthesis model (CLIM\_FvCB). Results also suggest that time-varying LAI conditions are relevant to the GPP estimates during large-scale extreme events. Limitations in the evaluation of NEE and TER arise from uncertainties in the FLUXCOM reference data used, and in the ECLand and SURFEX simulations, suggesting that further developments are required and that flux adjustments (Agustí-Panareda et al., 2016) are paramount to mitigate biased in global CO<sub>2</sub> analysis.

The online simulations of weather forecasts for the year of 2019 focused on the evaluation of the meteorological impact of the revised land cover and LAI. The results identified a clear improvement of 2-metre air temperature in Eurasia during spring while during summer the results were mostly negative in the Tropics, particularly over North-Eastern Brazil and central-south Africa. Similar results were found when considering 1000 hPa geopotential height. The geopotential height at 500 hPa indicated an error growth in CCI0 between forecast lead times of 1 to 3 days. A more detailed analysis over the USA with a high density of GHCN stations indicated that the pre-existing bias in CTR0 explains the impact of CCI0. There is a general increase of temperature in CCI0, mostly induced by the reduction of surface roughness (due

to the reduction of high vegetation) that is positive in regions where CTR0 had a negative bias, while negative in regions with already a pre-existing warm bias.

Ongoing developments in ECLand for the next operational IFS cycle include a revision of the handling of roughness lengths disaggregation for the tiles and post-processing of 2-metre air temperature. Preliminary results (not included in the report) indicate a positive impact on LST, however these require further testing with the revised land cover and LAI. Similarly, ongoing tests (performed during the preparation of this report) with revised model parameters for TER show the potential to improve the global NEE budgets. Therefore, the work in progress will continue to provide improvements, with an analysis of the latest version of the system, being proposed to be carried out in a second version of this report, which is expected to take place in about one year time (i.e., mid-2023).

## 7 Acknowledgements

The Farquhar model was implemented in ECLand/CAMS by Vladislav Bastrikov, Fabienne Maignan and Philippe Peylin (LSCE) and we also thank their support in the land cover maps processing.

## 8 References

- Agustí-Panareda, A., Massart, S., Chevallier, F., Balsamo, G., Boussetta, S., Dutra, E., & Beljaars, A. (2016). A biogenic CO<sub>2</sub> flux adjustment scheme for the mitigation of large-scale biases in global atmospheric CO<sub>2</sub> analyses and forecasts. *Atmospheric Chemistry and Physics*, *16*(16), 10399–10418. <https://doi.org/10.5194/acp-16-10399-2016>
- Albergel, C., Calvet, J. C., Gibelin, A. L., Lafont, S., Roujean, J. L., Berne, C., et al. (2010). Observed and modelled ecosystem respiration and gross primary production of a grassland in southwestern France. *Biogeosciences*, *7*(5), 1657–1668. <https://doi.org/10.5194/bg-7-1657-2010>
- Albergel, Clément, Munier, S., Jennifer Leroux, D., Dewaele, H., Fairbairn, D., Lavinia Barbu, A., et al. (2017). Sequential assimilation of satellite-derived vegetation and soil moisture products using SURFEX-v8.0: LDAS-Monde assessment over the Euro-Mediterranean area. *Geoscientific Model Development*, *10*(10), 3889–3912. <https://doi.org/10.5194/gmd-10-3889-2017>
- Barriopedro, D., Fischer, E. M., Luterbacher, J., Trigo, R. M., & García-Herrera, R. (2011). The hot summer of 2010: Redrawing the temperature record map of Europe. *Science*, *332*(6026), 220–224. <https://doi.org/10.1126/science.1201224>
- Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., & Wood, E. F. (2018). Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Scientific Data*, *5*(1), 180214. <https://doi.org/10.1038/sdata.2018.214>
- Boussetta, S., Balsamo, G., Beljaars, A., Kral, T., & Jarlan, L. (2013). Impact of a satellite-derived leaf area index monthly climatology in a global numerical weather prediction model. *International Journal of Remote Sensing*, *34*(9–10), 3520–3542. <https://doi.org/10.1080/01431161.2012.716543>
- Boussetta, S., Balsamo, G., Beljaars, A., Panareda, A. A., Calvet, J. C., Jacobs, C., et al. (2013). Natural land carbon dioxide exchanges in the ECMWF integrated forecasting system: Implementation and offline validation. *Journal of Geophysical Research Atmospheres*, *118*(12), 5923–5946. <https://doi.org/10.1002/jgrd.50488>

- Boussetta, S., Balsamo, G., Arduini, G., Dutra, E., McNorton, J., Choulga, M., et al. (2021). Ecland: The ecmwf land surface modelling system. *Atmosphere*, 12(6). <https://doi.org/10.3390/atmos12060723>
- Collatz, G., Ribas-Carbo, M., & Berry, J. (1992). Coupled Photosynthesis-Stomatal Conductance Model for Leaves of C4 Plants. *Functional Plant Biology*, 19(5), 519. <https://doi.org/10.1071/PP9920519>
- Collier, N., Hoffman, F. M., Lawrence, D. M., Keppel-Aleks, G., Koven, C. D., Riley, W. J., et al. (2018). The International Land Model Benchmarking (ILAMB) System: Design, Theory, and Implementation. *Journal of Advances in Modeling Earth Systems*, 10(11), 2731–2754. <https://doi.org/10.1029/2018MS001354>
- Defourny, P., Boettcher, M., Bontemps, S., Kirches, G., Krueger, O., Lamarche, C., et al. (2014). Algorithm theoretical basis document for land cover climate change initiative. *Eur Space Agency*, 191.
- Delbeke, J., Runge-Metzger, A., Slingenberg, Y., & Werksman, J. (2019). The paris agreement. *Towards a Climate-Neutral Europe: Curbing the Trend*, 24–45. <https://doi.org/10.4324/9789276082569-2>
- Delire, C., Séférian, R., Decharme, B., Alkama, R., Calvet, J. C., Carrer, D., et al. (2020). The Global Land Carbon Cycle Simulated With ISBA-CTRIP: Improvements Over the Last Decade. *Journal of Advances in Modeling Earth Systems*, 12(9), 1–31. <https://doi.org/10.1029/2019MS001886>
- Durre, I., Menne, M. J., Gleason, B. E., Houston, T. G., & Vose, R. S. (2010). Comprehensive automated quality assurance of daily surface observations. *Journal of Applied Meteorology and Climatology*, 49(8), 1615–1633. <https://doi.org/10.1175/2010JAMC2375.1>
- ECMWF. (2021). IFS Documentation CY47R3 - Part IV Physical processes. *IFS Documentation*.
- EU Commission. (2021). *Tracking EU Citizens' Interest in EC Priorities Using Online Search Data: The European Green Deal*. <https://doi.org/10.2760/18216>
- Farquhar, G. D., von Caemmerer, S., & Berry, J. A. (1980). A biochemical model of photosynthetic CO<sub>2</sub> assimilation in leaves of C<sub>3</sub> species. *Planta*, 149(1), 78–90. <https://doi.org/10.1007/BF00386231>
- Flach, M., Sippel, S., Gans, F., Bastos, A., Brenning, A., Reichstein, M., & Mahecha, M. D. (2018). Contrasting biosphere responses to hydrometeorological extremes: revisiting the 2010 western Russian heatwave. *Biogeosciences*, 15(20), 6067–6085. <https://doi.org/10.5194/bg-15-6067-2018>
- Freitas, S. C., Trigo, I. F., Bioucas-dias, J. M., & Göttsche, F. (2010). Quantifying the Uncertainty of Land Surface Temperature Retrievals From SEVIRI / Meteosat, 48(1), 523–534.
- Friedlingstein, P., Jones, M. W., O'Sullivan, M., Andrew, R. M., Bakker, D. C. E., Hauck, J., et al. (2022). Global Carbon Budget 2021. *Earth System Science Data*, 14(4), 1917–2005. <https://doi.org/10.5194/essd-14-1917-2022>
- Geer, A. (2015). Significance of changes in medium-range forecast scores. *ECMWF Technical Memoranda*, 766.
- Gibelin, Anne Laure, Calvet, J. C., & Viovy, N. (2008). Modelling energy and CO<sub>2</sub> fluxes with an interactive vegetation land surface model-Evaluation at high and middle latitudes. *Agricultural and Forest Meteorology*, 148(10), 1611–1628.

<https://doi.org/10.1016/j.agrformet.2008.05.013>

- Guo, Z., Dirmeyer, P. A., & Delsole, T. (2011). Land surface impacts on subseasonal and seasonal predictability. *Geophysical Research Letters*, *38*(24), L24812. <https://doi.org/10.1029/2011GL049945>
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. <https://doi.org/10.1002/qj.3803>
- Johannsen, F., Ermida, S., Martins, J. P. A., Trigo, I. F., Nogueira, M., & Dutra, E. (2019). Cold bias of ERA5 summertime daily maximum land surface temperature over Iberian Peninsula. *Remote Sensing*, *11*(21), 2570. <https://doi.org/10.3390/rs11212570>
- Jung, M., Koirala, S., Weber, U., Ichii, K., Gans, F., Camps-Valls, G., et al. (2019). The FLUXCOM ensemble of global land-atmosphere energy fluxes. *Scientific Data*, *6*(1), 74. <https://doi.org/10.1038/s41597-019-0076-8>
- Jung, M., Schwalm, C., Migliavacca, M., Walther, S., Camps-Valls, G., Koirala, S., et al. (2020). Scaling carbon fluxes from eddy covariance sites to globe: synthesis and evaluation of the FLUXCOM approach. *Biogeosciences*, *17*(5), 1343–1365. <https://doi.org/10.5194/bg-17-1343-2020>
- Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., et al. (2005). A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system. *Global Biogeochemical Cycles*, *19*(1), 1–33. <https://doi.org/10.1029/2003GB002199>
- Kumar, S., Mocko, D., Vuyovich, C., & Peters-Lidard, C. (2020). Impact of surface albedo assimilation on snow estimation. *Remote Sensing*, *12*(4), 1–27. <https://doi.org/10.3390/rs12040645>
- Loveland, T. R., Reed, B. C., Ohlen, D. O., Brown, J. F., Zhu, Z., Yang, L., & Merchant, J. W. (2000). Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *International Journal of Remote Sensing*, *21*(6–7), 1303–1330. <https://doi.org/10.1080/014311600210191>
- Masson, V., Le Moigne, P., Martin, E., Faroux, S., Alias, A., Alkama, R., et al. (2013). The SURFEXv7.2 land and ocean surface platform for coupled or offline simulation of earth surface variables and fluxes. *Geoscientific Model Development*, *6*(4), 929–960. <https://doi.org/10.5194/gmd-6-929-2013>
- Menne, M. J., Durre, I., Vose, R. S., Gleason, B. E., & Houston, T. G. (2000). An overview of the global historical climatology network-daily database. *ECMWF Technical Memorandum, No.295*(7), 897–910. <https://doi.org/10.1175/JTECH-D-11-00103.1>
- Mutti, P. R., Dubreuil, V., Bezerra, B. G., Arvor, D., de Oliveira, C. P., & Santos E Silva, C. M. (2020). Assessment of gridded cru ts data for long-term climatic water balance monitoring over the são francisco watershed, brazil. *Atmosphere*, *11*(11), 1–25. <https://doi.org/10.3390/atmos11111207>
- Nogueira, M., Albergel, C., Boussetta, S., Johannsen, F., Trigo, I. F., Ermida, S. L., et al. (2020). Role of vegetation in representing land surface temperature in the CHTESSEL (CY45R1) and SURFEX-ISBA (v8.1) land surface models: A case study over Iberia. *Geoscientific Model Development*, *13*(9), 3975–3993. <https://doi.org/10.5194/gmd-13-3975-2020>
- Nogueira, M., Boussetta, S., Balsamo, G., Albergel, C., Trigo, I. F., Johannsen, F., et al. (2021a). Upgrading Land-Cover and Vegetation Seasonality in the ECMWF Coupled

System: Verification With FLUXNET Sites, METEOSAT Satellite Land Surface Temperatures, and ERA5 Atmospheric Reanalysis. *Journal of Geophysical Research: Atmospheres*, 126(15). <https://doi.org/10.1029/2020JD034163>

Nogueira, M., Boussetta, S., Balsamo, G., Albergel, C., Trigo, I. F., Johannsen, F., et al. (2021b). Upgrading Land-Cover and Vegetation Seasonality in the ECMWF Coupled System: Verification With FLUXNET Sites, METEOSAT Satellite Land Surface Temperatures, and ERA5 Atmospheric Reanalysis. *Journal of Geophysical Research: Atmospheres*, 126(15), e2020JD034163. <https://doi.org/10.1029/2020JD034163>

Orth, R., Dutra, E., & Pappenberger, F. (2016). Improving weather predictability by including land surface model parameter uncertainty. *Monthly Weather Review*, 144(4), 1551–1569. <https://doi.org/10.1175/MWR-D-15-0283.1>

Poulter, B., MacBean, N., Hartley, A., Khlystova, I., Arino, O., Betts, R., et al. (2015). Plant functional type classification for earth system models: Results from the European Space Agency’s Land Cover Climate Change Initiative. *Geoscientific Model Development*, 8(7), 2315–2328. <https://doi.org/10.5194/gmd-8-2315-2015>

Quaife, T., Quegan, S., Disney, M., Lewis, P., Lomas, M., & Woodward, F. I. (2008). Impact of land cover uncertainties on estimates of biospheric carbon fluxes. *Global Biogeochemical Cycles*, 22(4), n/a-n/a. <https://doi.org/10.1029/2007GB003097>

Still, C. J., Berry, J. A., Collatz, G. J., & DeFries, R. S. (2003). Global distribution of C<sub>3</sub> and C<sub>4</sub> vegetation: Carbon cycle implications. *Global Biogeochemical Cycles*, 17(1), 6-1-6–14. <https://doi.org/10.1029/2001GB001807>

Trigo, I. F., Dacamara, C. C., Viterbo, P., Roujean, J. L., Olesen, F., Barroso, C., et al. (2011). The satellite application facility for land surface analysis. *International Journal of Remote Sensing*, 32(10), 2725–2744. <https://doi.org/10.1080/01431161003743199>

Verger, A., Baret, F., & Weiss, M. (2014). Near Real-Time Vegetation Monitoring at Global Scale. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(8). <https://doi.org/10.1109/JSTARS.2014.2328632>

Wang, X. C., Wang, C. K., & Yu, G. R. (2008). Spatio-temporal patterns of forest carbon dioxide exchange based on global eddy covariance measurements. *Science in China, Series D: Earth Sciences*, 51(8), 1129–1143. <https://doi.org/10.1007/s11430-008-0087-3>

Yin, X., & Struik, P. C. (2009). C<sub>3</sub> and C<sub>4</sub> photosynthesis models: An overview from the perspective of crop modelling. *NJAS: Wageningen Journal of Life Sciences*, 57(1), 27–38. <https://doi.org/10.1016/j.njas.2009.07.001>

Yuan, W., Luo, Y., Li, X., Liu, S., Yu, G., Zhou, T., et al. (2011). Redefinition and global estimation of basal ecosystem respiration rate. *Global Biogeochemical Cycles*, 25(4), 1–14. <https://doi.org/10.1029/2011GB004150>

## Document History

Version	Author(s)	Date	Changes
0.1	Emanuel Dutra & Francisco Lopes (FC.ID)	08/11/2021	First draft with table of contents and associated work to each partner

0.2	""	06/06/2022	First version with all chapters prepared and ready for partners review
1.0	All contributing authors	20/06/2022	Final version with contributions from all authors. For internal review
2.0	All contributing authors	05/07/2022	Final version after review

## Internal Review History

Internal Reviewers	Date	Comments
Peter Wouter (Wageningen University)	24/06/2022	Minor editorial changes and further discussion on the TER and NEE budgets.
Ruben Urraca and Nadine Gobron (JRC)	30/06/2022	Minor comments and editorial comments.

## Estimated Effort Contribution per Partner

Partner	Effort
FC.ID	16
Meteo-France	6
ECMWF	2
MPG	1
<b>Total</b>	<b>25</b>

## 9 Appendix I: Supplementary Figures

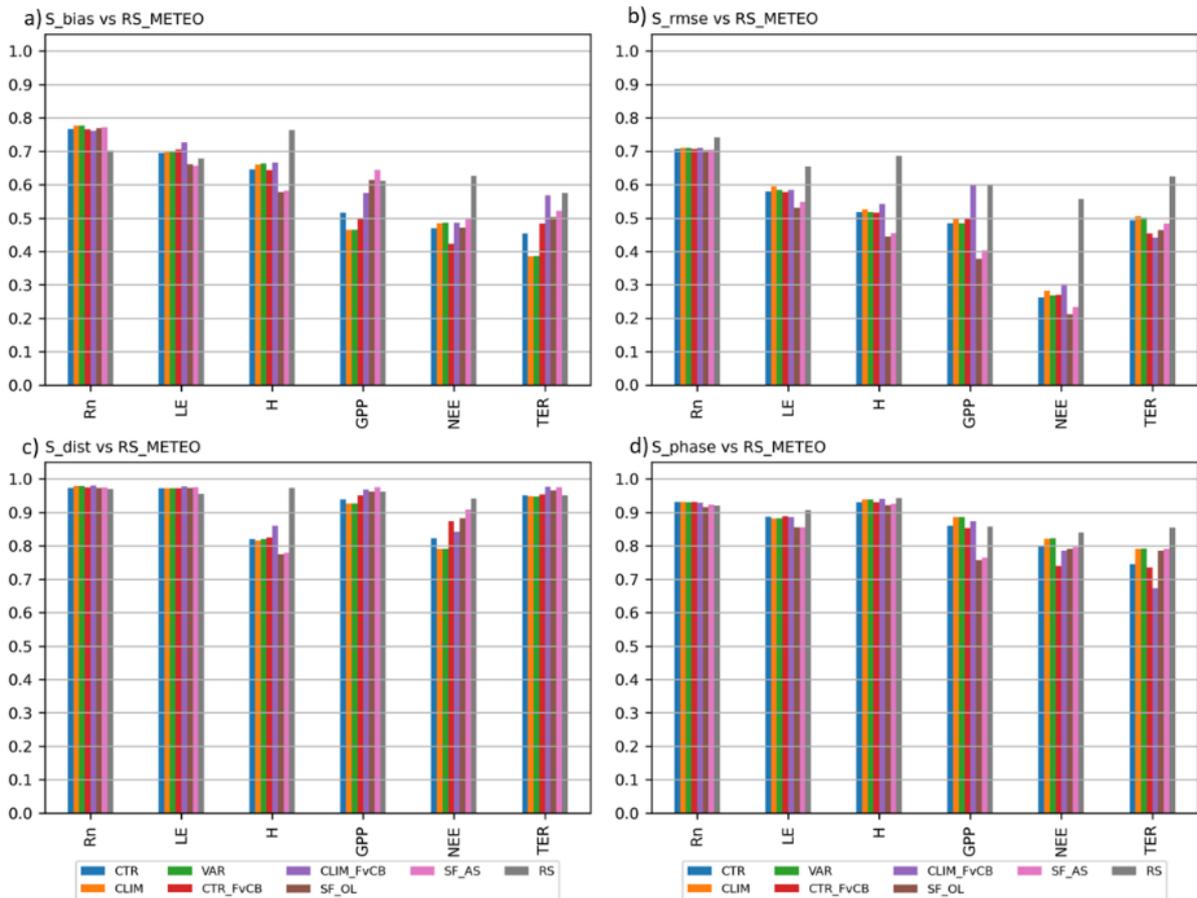
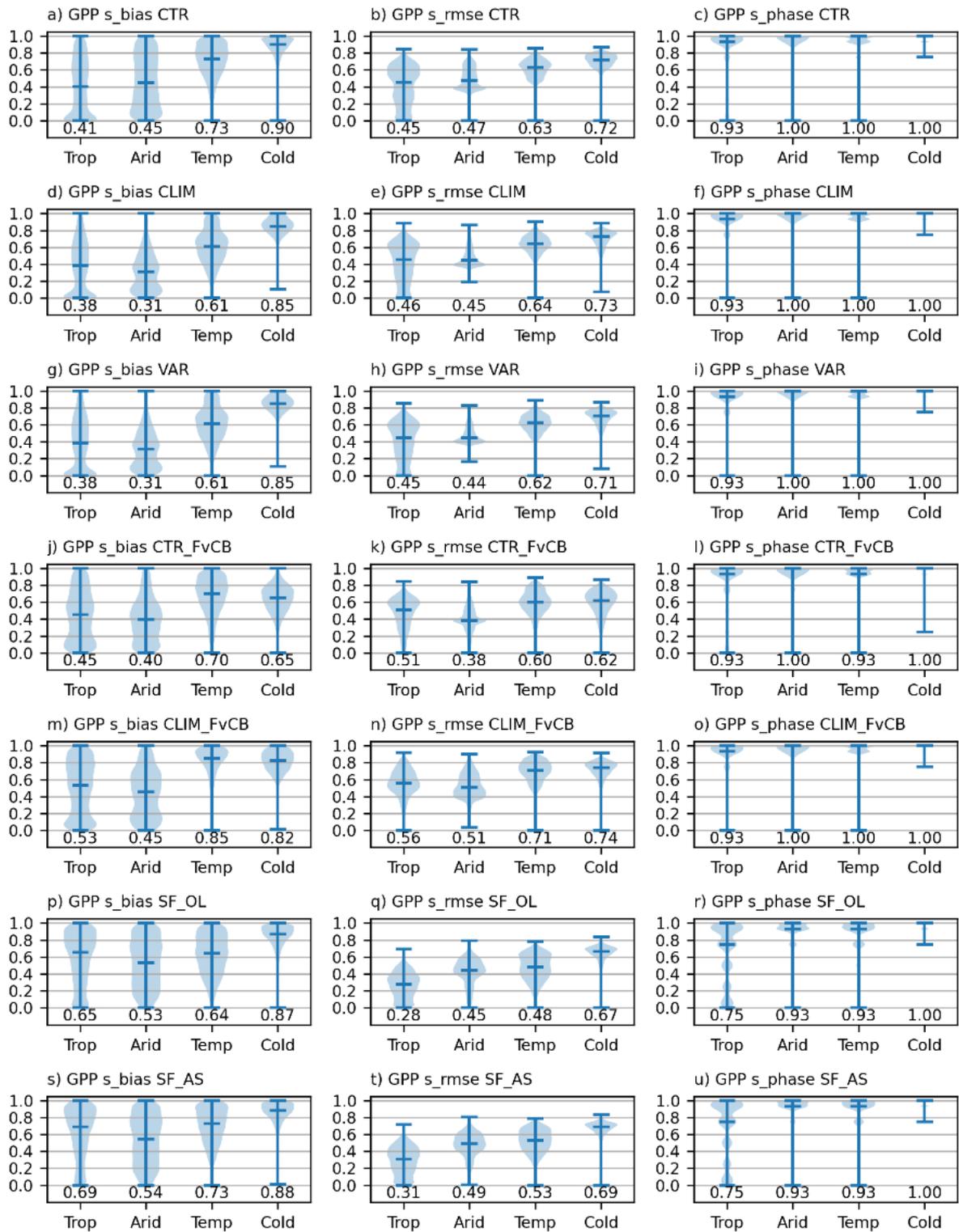


Figure A1: Scores for each of the considered metrics (a) S\_bias, (b) S\_rmse, (c) S\_dist, and (d) S\_phase for the 6 considered variables (horizontal axis) shown for each simulation (bars) considering FLUXCOM R\_METEO as reference and including the metrics of FLUXCOM RS using FLUXCOM RS\_METEO as reference (grey bars).



**Figure A2: Distribution of S\_bias (left panel), S\_rmse (centre panels) and S\_phase (right panels) for each simulation CTR (a, b, c), CLIM (d, e, f), VAR (g, h, i), CTR\_FvCB (j, k, l), CLM\_FvCB (m, n, o), SF\_OL (p, q, r), and SF\_AS (s, t, u) for the 4 climate regions Tropical, Arid, Temperate, and Cold (horizontal axis). Each violin distribution displays the extremes and median as a horizontal line, with the value of the median shown at the base of the plot.**

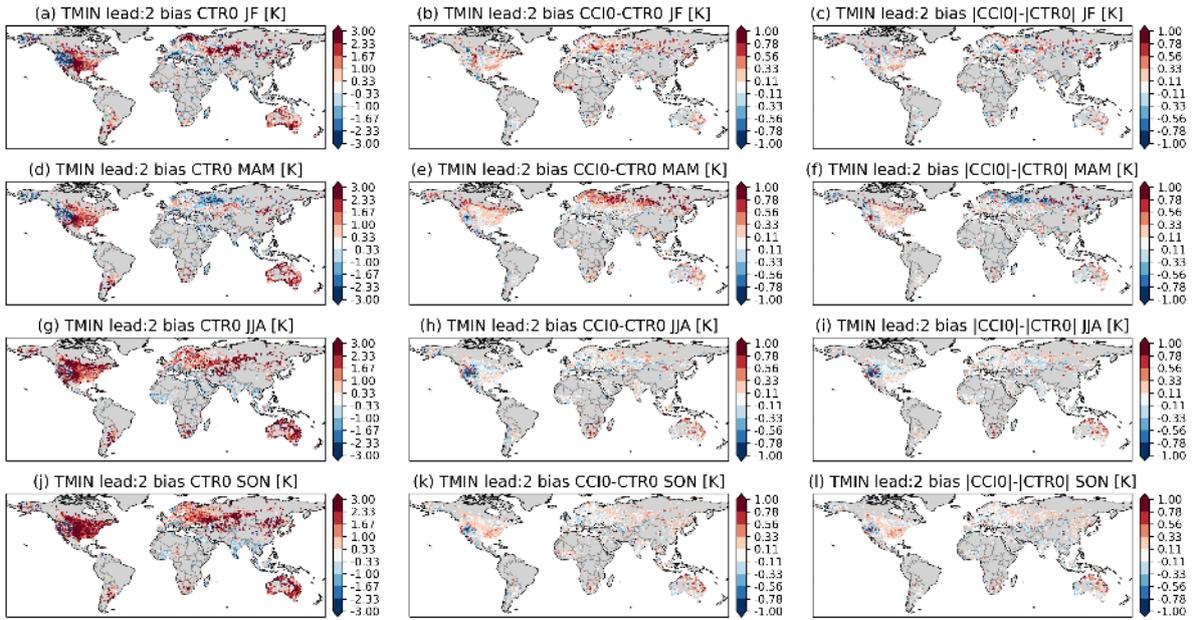


Figure A3: Same as Figure 19 but for the daily minimum temperature (TMIN) forecasts.

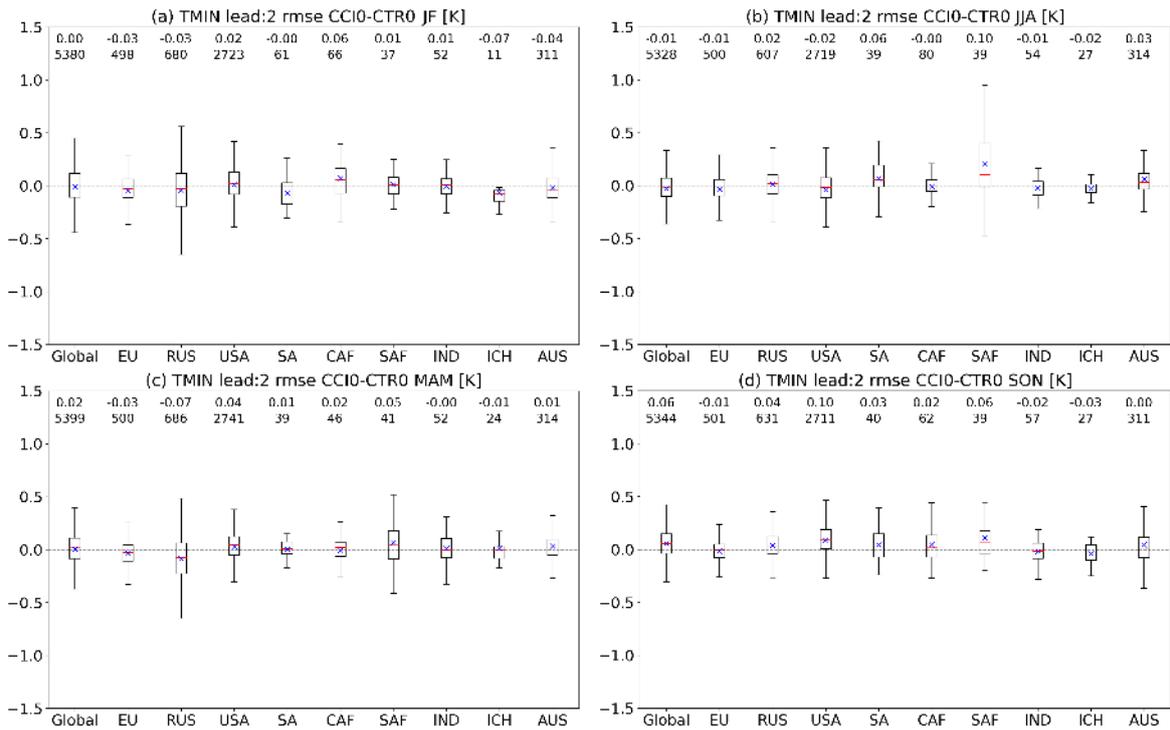
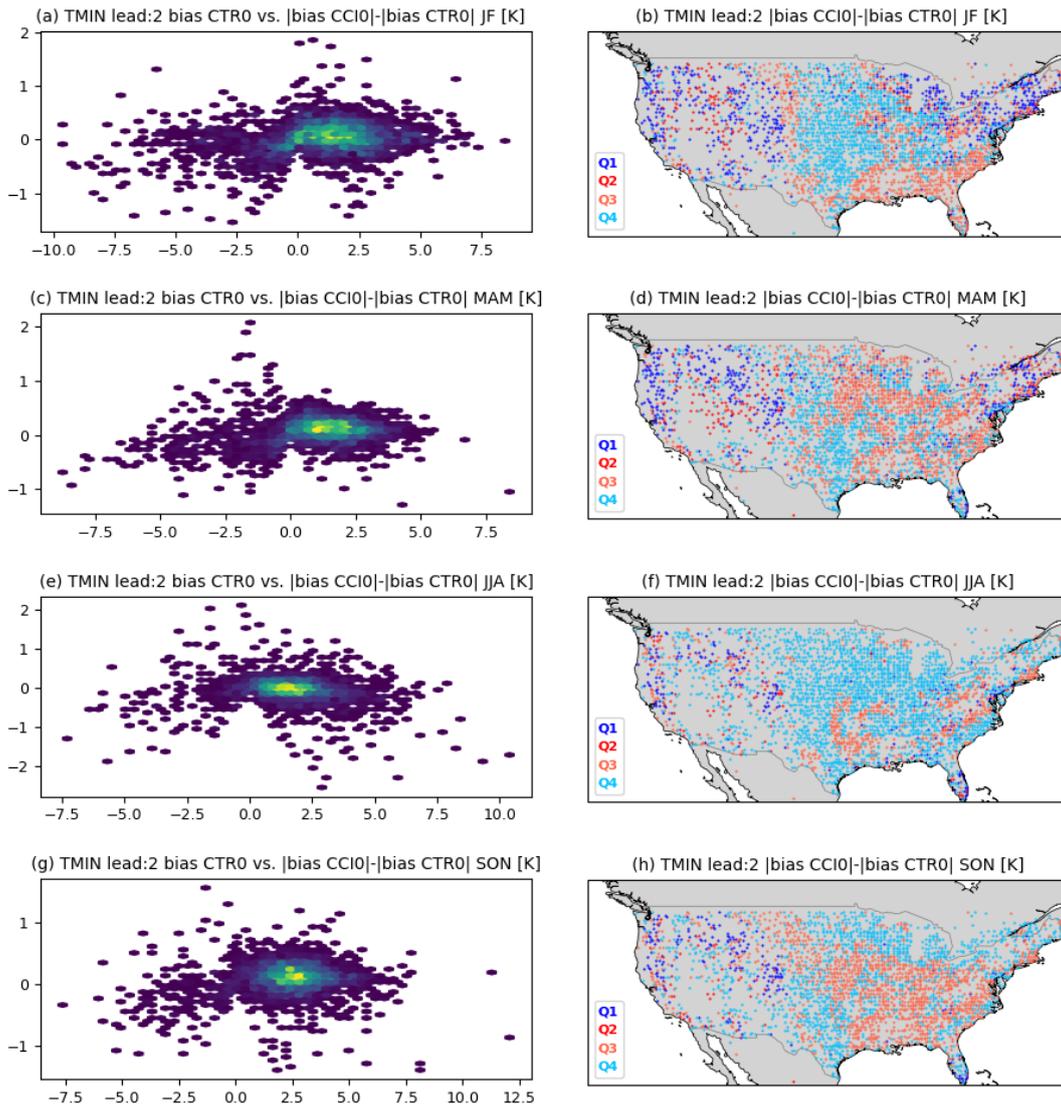


Figure A4: Same as Figure 20 but for the daily minimum temperature (TMIN) forecasts.



**Figure A5: Same as Figure 21 but for daily minimum temperature (TMIN) forecasts.**

This publication reflects the views only of the author, and the Commission cannot be held responsible for any use which may be made of the information contained therein.