CopERnIcus climate change Service Evolution

D3.5 - Documentation of the final version of land surface initialisation systems

Due date of deliverable	30 June 2026
Submission date	29 June 2026
File Name	CERISE-D3-5-V1.3
Work Package /Task	WP3
Organisation Responsible of Deliverable	MF
Author name(s)	J-C. Calvet, C. Ardilouze, J. Day, L.G.G. De Goncalves, D. Fairbairn, K. Froehlich, N. Noll, M. Lange, O. Rojas-Munoz, G. Narvaez-Campo, V. Romanova, T. Stockdale, A. Vasconcelos
Revision number	1.3
Status	Revised version
Dissemination Level	Public

Funded by the European Union

The CERISE project (grant agreement No 101082139) is funded by the European Union.
Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the Commission. Neither the European Union nor the granting authority can be held responsible for them.

1 Executive Summary

The initialisation data sets for seasonal forecasts from CERISE Phase 2 were produced by the institutions participating in Work Package 3 (WP3): the Deutscher Wetterdienst (DWD), Meteo-France (MF), the Euro-Mediterranean Centre on Climate Change (CMCC) and the European Centre for Medium-Range Weather Forecasts (ECMWF). These data sets were delivered to the ECMWF and are stored in the MARS archive. These datasets include near-real-time prototypes demonstrating that the various partners can compute usable land initial conditions (LICs) within operational time constraints and make them available in a timely manner for seasonal forecast production within the Copernicus framework. For intercomparison and further assessment, a set of monthly mean land surface data from each initialisation system is also provided.

Table of Contents

Contents

1 Executive Summary 2
2 Introduction 5
2.1 Background 5
2.2 Scope of this deliverable 5
2.2.1 Objectives of this deliverables 6
2.2.2 Work performed in this deliverable 6
2.2.3 Deviations and counter measures 6
2.2.4 Reference Documents 6
2.2.5 CERISE Project Partners: 6
3 Methods and experiment setup for Phase 2 7
3.1 DWD 7
3.2 MF 8
3.3 CMCC 8
3.4 ECMWF 9
4 Methods and experiment setup for NRT LICs in 2021 10
4.1 DWD 10
4.2 MF 11
4.3 CMCC 11
5 Results for LICs' reanalyses 11
5.1 Global maps of mean LST and SWE 12
5.2 Global monthly mean time series of LST 13
5.3 Global monthly mean time series of SWE 14
5.4 Verification of the consistency across demonstrators 15
6 Results for NRT LICs (2021) 16
6.1 DWD 16
6.1.1 LAI 17
6.1.2 SM 17
6.1.3 SWE 19
6.2 MF 20
6.2.1 LAI 20
6.2.2 SM 20
6.2.3 SWE 22
6.3 CMCC 22
6.3.1 LAI 22
6.3.2 SM 24
6.3.3 SWE 25

7 Conclusion 26
8 References 27

2 Introduction

The different initialisation data sets used for the seasonal forecast in CERISE Phase 2 are produced using improved Earth climate models, which include or refine land surface analysis systems. Different institutions considered different control variables depending on the current development. These are selected based on their impact on the current state and should mainly control soil moisture, leaf area index and snow depth. In order to be useful for current predictions, timely observations are needed to capture rapid changes, such as a sudden dry spell affecting soil moisture or early snowmelt. Without these observations, errors in the LICs can accumulate and affect temperature and precipitation forecasts further down the line. In addition to reanalyses, LICs were also created using Near Real-Time (NRT) observations to assess whether LICs could be generated and delivered within the operational time constraints of the Copernicus seasonal forecasting schedule.

2.1 Background

The scope of CERISE is to improve the quality of the Copernicus Climate Change Service (C3S) reanalysis and seasonal forecasting capabilities, with a particular focus on land-atmosphere interactions. Over the project's four-year timescale and beyond, it will support the evolution of C3S by improving the integrity and coherence of the C3S Earth system Essential Climate Variables (ECVs) through enhanced climate reanalysis and seasonal prediction systems and products.

CERISE will develop new and innovative ensemble-based coupled land-atmosphere data assimilation approaches and land surface initialisation techniques to pave the way for the next generations of the C3S reanalysis and seasonal prediction systems.

These developments will be combined with innovative work on observation operator developments integrating Artificial Intelligence (AI) to ensure optimal data fusion fully integrated in coupled assimilation systems. They will drastically enhance the exploitation of past, current, and future Earth system observations over land surfaces, including from the Copernicus Sentinels and from the European Space Agency (ESA) Earth Explorer missions, moving towards an all-sky and all-surface approach. For example, land observations can simultaneously improve the representation and prediction of land and atmosphere and provide additional benefits through the coupling feedback mechanisms. Using an ensemble-based approach will improve uncertainty estimates over land and lowest atmospheric levels.

Improving coupled land-atmosphere assimilation methods, land surface evolution and the exploitation of satellite data will enhance the representation of long-term trends and regional extremes in the C3S reanalysis and seasonal prediction systems by incorporating CERISE inputs.

In addition, CERISE provides proof of concept to demonstrate the feasibility of integrating the developed approaches into the core C3S. This is achieved by delivering reanalysis prototype datasets in a pre-operational environment and seasonal prediction demonstrator datasets in a relevant environment.

The aim of CERISE is to improve the quality and consistency of C3S reanalysis systems and components of the seasonal prediction multi-system. This directly addresses evolving user needs for improved, consistent C3S Earth system products.

2.2 Scope of this deliverable

This section outlines the scope of the deliverable, including its main objectives and the work performed. It also covers any deviations from the planned activities. It also provides references to relevant documents and an overview of the project partners involved.

2.2.1 Objectives of this deliverable

This report describes the Phase 2 dataset of initial conditions for the seasonal forecast, provided by the DWD, MF, CMCC and ECMWF models. The methods and experimental protocol developed up to the delivery stage in month 34 (October 2025), after Phase 2, are presented alongside the results. The results include NRT LICs for the year 2021.

2.2.2 Work performed in this deliverable

In Work Package 3 (WP3) of the CERISE project, each partner developed Phase 1 and Phase 2 demonstrators that analyse surface data and produce initial conditions as close as possible to the observational datasets. This data has been transferred to the MARS archive (see D3.2 for more details) and is currently available to the project consortium for analysis, investigation, and initialisation of the seasonal forecast. This includes a set of monthly mean land surface data from each initialisation dataset for DWD, MF, CMCC and ECMWF. All demonstrators include land surface data assimilation in order to accurately initialise seasonal predictions from 1993 to 2023 (CMCC's predictions start in 2002). Participating institutions use coupled or advanced Earth system models, incorporating the evolution of the atmospheric (and oceanic) background state over time, as well as land and vegetation models. The aim is to create balanced initial conditions for seasonal numerical predictions. Land surface analysis schemes were adapted from the participating institutions' current operational analyses and implemented in the development of the systems. Here, we present a set of initialisation datasets completed by all participating institutions during the WP3 CERISE Phase 2 period.

2.2.3 Deviations and counter measures

No deviations have been encountered.

2.2.4 Reference Documents

[1] Project 101082139- CERISE-HORIZON-CL4-2021-SPACE-01 Grant Agreement

[2] CERISE D3.2 Land surface initial conditions for 1993-2022 [data description report, “SENSITIVE”, available on request]

[3] CERISE D4.1 ERA6-Land-Pv2 Land prototype 1940-2019 [data description report, “SENSITIVE”, available on request]

[4] CERISE D7.4 Time varying Lake cover, Land cover and LAI and extension of CONFESS vegetation data back to 1925 datasets

2.2.5 CERISE Project Partners:

ECMWF	European Centre for Medium-Range Weather Forecasts
Met Norway	Norwegian Meteorological Institute
SMHI	Swedish Meteorological and Hydrological Institute
MF	Météo-France
DWD	Deutscher Wetterdienst
CMCC	Euro-Mediterranean Center on Climate Change
BSC	Barcelona Supercomputing Centre
DMI	Danish Meteorological Institute
Estellus	Estellus
IPMA	Portuguese Institute for Sea and Atmosphere
NILU	Norwegian Institute for Air Research
MetO	Met Office

3 Methods and experiment setup for Phase 2

3.1 DWD

DWD's data assimilation configuration of ICOsahedral Nonhydrostatic model for eXtended Predictions and Projections (ICON XPP, Müller et al., 2025) was enhanced by incorporating a two-dimensional variational data assimilation (DA) scheme for snow depth (2Dvar), and an independent analysis for LAI, into the prediction system. Vegetation, soil, and land surface state are simulated with the Jena Scheme for Biosphere–Atmosphere Coupling in Hamburg (JS-Bach, Scheck et al., 2022) instead of TERRA (Schrodin and Heise, 2001), DWD's land surface component for the operational NWP system. JS-Bach has 5 soil layers, configured with top and bottom layer boundaries at layer depth [0-6.5 / 6.5-31.9 / 31.9-123.2 / 123.2-413.4 / 413.4-983.4] in cm, corresponding to layer thickness of [6.5, 25.4, 91.3, 290.2, 570.0] cm. The forecast system runs on 25 ensemble members, each updated with separate snow analysis, using its own background state. A variational univariate method is used for LAI analysis, applied to one selected member, with resulting increments distributed to the whole ensemble, keeping the spread from the different model runs. The Basic Cycling Environment (BACY) was used to test the land surface initialisation components integrated into the seasonal prediction system. BACY is well established at DWD for conducting pre-operational tests, with modifications made to the NWP suite. It controls the entire data assimilation cycle, including analyses of the atmosphere, ocean surface temperature and various land components, as well as ICON model forecast runs for the deterministic and ensemble systems. As part of the CERISE project, the new land DA components are implemented in addition to the reference configuration used for Phase 0. The reference experiment is initialised with ERA5 in the ICON atmosphere using Nudging of 6-hourly wind and temperature fields above a sea level height of 1.5 km at each time step, with given relaxation parameters. The ocean is constrained to the EN4 monthly mean vertical profiles of temperature and salinity on the first day of each month using the Local Singular Evolutive Interpolated Kalman Filter (LSEIK) filter (Nerger et al., 2006). The Phase 2 system now includes two land initialisation methods. First, the 2D-Var snow analysis, tested in Phase 1 for two cycling frequencies: one month and five days, with the latter leading to an appropriate adaptation of the model state to the observations. Second, a variational DA scheme for LAI is implemented as an intermediate step on the way to an integrated soil moisture analysis. The 2D-Var analysis package is designed to assimilate observations from synoptic stations for snow depth and screen-level variables temperature, dew point temperature, and relative humidity. It is an iterative procedure that minimises the cost functional using the conjugate gradient descent algorithm. Technically it is implemented with abstract function pointers to enhance the flexibility of using parameter-specific procedures that differ for the screen-level variables T2m, Rh2m, and snow. Here, snow depth observations from SYNOP reports are retrieved from ECMWF's Meteorological Archival and Retrieval System (MARS) archive with "era5" stream specified for the reanalysis experiment. LAI is assimilated sequentially in a separate univariate variational scheme every 10 days, in accordance with each update of the observation dataset THEIA GEOV2/AVHRR. Ancillary physiographic data and soil and vegetation parameters are generated using the EXTPAR software package, which is maintained by the Center for Climate Systems Modelling (C2SM). They can be accessed at https://c2sm.ethz.ch/news/archive/2025/05/zonda-v10-is-here-making-icon-simulations-more-accessible.html. It is used within the COSMO consortium and the ICON community to create grid-specific files of aggregated and processed datasets based on high-resolution raw data. The basic datasets used in EXTPAR are described in Asensio et al. (2023). Apart from the 2Dvar assimilation of the SYNOP stations, the initialisation run in Phase 2 also includes LAI assimilation of THEIA GEOV2/AVHRR satellite data every 10 days. This is the first time that LAI assimilation has been implemented in the DWD system.

3.2 MF

The joint assimilation of surface soil moisture (SSM) and leaf area index (LAI) satellite-based products is possible within the SURFEX modelling platform using a Land Data Assimilation System (LDAS), the LDAS-Monde tool. MF used the LDAS-Monde tool within the SURFEX version 9 modelling platform in offline mode, which was forced by ERA5. THEIA AVHRR GEOV2 LAI (1993–2018) (Verger et al., 2025) and Copernicus Land Monitoring Service (CLMS) GEOV2-AVHRR LAI (2019–2022) were assimilated every 10 days using a simplified extended Kalman filter (SEKF) to update leaf biomass and surface soil moisture. The system follows a two-step sequence in which the ISBA model forecast is corrected via an analysis stage involving the propagation of observational information to control variables using finite-difference Jacobians. This flow-dependent structure enables consistent updates of LAI and soil moisture, which indirectly impact other land surface states over 24-hour assimilation windows. In Phase 2, analysed soil moisture derived from LAI assimilation is used. Conversely, an LAI climatology is used instead of analysed LAI. Vegetation growth in ISBA was represented using the ISBA-A-gs configuration (Calvet et al., 1998; Gibelin et al., 2006; Calvet et al., 2008). This version of ISBA simulates the net CO_2 assimilation rate (A) and stomatal conductance (gs) of vegetation at leaf and canopy levels. This enables the simulation of LAI, respiration, and energy and water fluxes. ISBA-A-gs can represent feedback between LAI and root-zone soil moisture. Increasing LAI values tend to increase plant transpiration and reduce root-zone soil moisture through root water extraction. Conversely, decreasing root-zone soil moisture reduces photosynthesis, stomatal conductance (gs) and LAI. To more accurately represent root-zone soil moisture, the ISBA diffusion multilayer soil representation (Decharme et al., 2019) was employed. Soil moisture and temperature were calculated for 14 layers down to 12 m and 8–10 layers down to 1 and 2 m, depending on the characteristics of the vegetation. LDAS-Monde has been employed in numerous studies to assimilate and validate various satellite products across diverse regions and scales (e.g. Albergel et al., 2017; Mucia et al., 2020). In this study, the variables analysed were leaf biomass and soil moisture at several depths (up to 1 m). Each analysed variable had 12 different values, corresponding to 12 land surface patch classes. The patch fraction for each model grid cell was calculated using the ECOCLIMAP-II land cover database (Faroux et al., 2013). The assimilation process was performed every 24 hours, with the analysed variables being used as the initial conditions for the subsequent 24-hour period. In Phase 2, in addition to soil moisture and soil temperature variables initialized from LDAS-Monde (Phase 1), the snow water equivalent and snow density variables are directly initialized from the ERA5 reanalysis.

3.3 CMCC

The CMCC CESM2/CLM5-BGC configuration has been enhanced through the integration of a daily Ensemble Adjustment Kalman Filter (EAKF) within an offline, 30-member land reanalysis system driven by expanded ECMWF EDA forcings. The system assimilates ESA-CCI surface soil moisture (SM) and snow cover fraction (SCF) daily products and GLASS LAI (every 8 days) after strict quality control (QC) and continuous adaptive inflation adjustments. The asynchronous schedule (SM and SCF daily, and LAI every eighth day) enables updates to SM, snow water equivalent (SWE), snow cover and vegetation states to be made in a physically consistent way for seasonal forecast initialisation. Updates to SM, SCF and LAI were propagated through model couplings to evapotranspiration, energy partitioning and carbon pools. Validation uses independent datasets: ISMN in situ soil moisture, FLUXNET towers and FLUXCOM (fluxes/productivity), GLEAM ET and satellite snow/albedo data were used to verify cryospheric seasonality and melt timing. Core diagnostics include innovation statistics, bias, root mean square error (RMSE) and unbiased root mean square error (ubRMSE), anomaly correlation, and ensemble spread–skill, with all results benchmarked against a no-data assimilation (DA) control under identical forcing. An offline CESM2/CLM5-BGC land reanalysis was performed for the period 2002–2022 using a 30-member ensemble forced by ECMWF EDA members. This yielded flow-dependent variations in precipitation, radiation, temperature, humidity and winds. The Ensemble Adjustment Kalman Filter (EAKF) performs deterministic daily updates via CMCC/SPREADS (Cardinali et al., 2025). For each cycle, we archive the ensemble means and standard deviations (SDs) before and after assimilation. Observation operators map ESA-CCI SSM to the top layer of H_2O-SOI, ESA-CCI SCF to the CLM snow cover state (with covariances adjusting SWE) and GLASS LAI to the prognostic TLAI. Strict QC is applied to all observations during preprocessing, and assimilation runs asynchronously (SM+SCF daily and LAI every eighth day), so increments from each stream adjust the coupled water–energy–carbon states via cross-covariances. Covariance localisation and adaptive multiplicative inflation maintain filter reliability and spread–skill consistency. The outputs follow CMCC CLM5 conventions and expose daily states/fluxes and their ensemble uncertainty for diagnostics. The novelties versus Phase 0 are as follows: (i) joint assimilation of SM+SCF+LAI (dynamic LAI is rarely assimilated operationally); (ii) a dynamic vegetation constraint — LAI DA corrects phenology and plant pools rather than prescribing seasonality; (iii) end-to-end uncertainty quantification, enabling coupled water–carbon–energy analyses. Routed runoff and discharge are generated using the HYDROS routing scheme for basin-scale evaluation, which is consistent with DA-corrected land states. The Phase 1 setup (2002–2022) consisted of an offline CESM2/CLM5-BGC 30-member ensemble, which was forced by an expanded ECMWF EDA of 30 and cycled daily using a deterministic EAKF (SPREADS). The assimilated observations are ESA-CCI SSM (daily), ESA-CCI SCF (daily) and GLASS LAI (8-day, aggregated to the model grid), and the soil moisture is comparable to the model climatology. Assimilation is asynchronous (SM and SCF daily, and LAI according to its cycle), with operators mapping each product to the corresponding CLM5 state, thereby adjusting the covariances of SWE, canopy carbon and related variables. Spin-up follows a GSWP3 equilibrium-to-transient protocol to year-2000 conditions. The DA run and a parallel no-DA control share identical forcing. The outputs follow CLM5 history conventions on a ~0.5° global grid and provide daily water, energy and carbon states and fluxes (e.g. H_2OSOI/TSOI, H_2OSNO, TLAI, NEP/NEE), with ensemble means and standard deviations. Routed runoff and discharge are produced using the HYDROS scheme for hydrological evaluation. In Phase 2, the CMCC used the offline CESM2/CLM5-BGC land configuration within the SPREADS Ensemble Adjustment Kalman Filter (EAKF) framework to create a global land reanalysis comprising 30 members for the period 2002–2022. The system was driven by an expanded ECMWF EDA-based atmospheric forcing ensemble and cycled at a daily frequency. The assimilated observation streams were ESA-CCI surface soil moisture and ESA-CCI snow cover fraction (both assimilated daily) and GLASS leaf area index (assimilated every eight days after aggregation to the CLM5 grid). The multivariate EAKF updates adjusted the states of soil moisture, snow, and vegetation carbon and nitrogen through the CLM5 state vector and background covariances. Adaptive prior inflation and horizontal covariance localisation were used to maintain ensemble reliability. Phase 2 outputs include daily land states and fluxes, as well as ensemble means and standard deviations before and after assimilation and routed runoff/discharge, which was generated using the HYDROS routing scheme.

3.4 ECMWF

ECMWF has developed an offline (standalone) LDAS that reproduces the main features of the coupled LDAS used operationally at ECMWF. An SEKF (de Rosnay *et al.*, 2013) assimilates scatterometer-derived surface SM, IMS (Interactive Multisensor Snow and Ice Mapping System) snow cover, and pseudo screen-level (T2m/RH2m) observations over 12-hour assimilation windows, in a manner similar to the coupled integrated forecasting system (IFS). Furthermore, lake temperature and ice cover are also assimilated using a lake nudging approach. The ecLand model is used to propagate in time and space the information from the observations, accounting for physiographic information (soil texture, orography and LC), meteorological conditions and land surface processes such as soil evaporation and vegetation transpiration (Boussetta *et al.*, 2021). While the offline ecLand DA system is similar to that of the IFS, its implementation in "offline" mode means it is constrained by the atmospheric reanalysis ERA5. It is also much faster to run than the IFS. The computational efficiency of ecLand is very useful for running long re-analyses, which is important for evaluating the impact of the new time-varying land cover and vegetation. Seasonal hindcasts were initialised with three different surface analyses: An “open-loop” land analysis (EC-land forced with ERA5, but without land-DA) with monthly-varying land cover and LAI, one forced with the offline LDAS with climatological land cover and LAI and a second with the offline LDAS and time-varying land cover and LAI. Comparison of these three hindcasts allows us to understand the relative impacts of the time-varying land cover, LAI and land DA on the seasonal forecasts. For Phase 1, seasonal hindcasts were initialised with four different surface analyses: ERA5, An “open-loop” land analysis (EC-land forced with ERA5, but without land-DA), one forced with the Offline-LDAS (described above) and a second LDAS where the soil-moisture assimilation was switched off, but the snow DA was switched on. Comparison of the last two allows us to understand the impact coming from the different land-surface components. For Phase 2, monthly varying leaf area index (LAI) and annually-varying land cover were integrated in ecLand, together with the LDAS, as it was showed that time-varying LAI and land cover can be used in the LDAS (D4.1). The most recent version of the offline LDAS, includes the new lake temperature/ice analysis and it was used for Phase 2. More information on the time-varying land cover and LAI can be found in the CERISE D7.4 report (https://cerise-project.eu/sites/default/files/2025-05/CERISE-D7-4-V1.0.pdf). Phase 2 ECMWF simulations cover the period from January 1998 to December 2019.

4 Methods and experiment setup for NRT LICs in 2021

This section describes the data assimilation systems and model configurations used by DWD, MF and CMCC for the NRT demonstrators. It provides an overview of the modelling frameworks, assimilation approaches, and operational workflows adopted to generate LICs. A common framework has been established to enable partners to describe their systems in a standardised way. All demonstrators incorporate land surface data assimilation to provide accurate initialisation for seasonal predictions. ECMWF was not involved in this exercise.

4.1 DWD

The DWD seasonal forecast system used for the 2021 NRT experiment is based on the Phase 2 configuration, with differences in observations that require real time availability. To demonstrate the NRT capabilities, the system uses Nudging against ERA5T for the atmosphere and ARGO data (Wong et al., 2020) as observation input for the ensemble-based LSEIKF Ocean DA system. Over land, the 2DVAR for snow depth assimilation was used, along with the operational stream of SYNOP data from the MARS archive. The LAI assimilation scheme uses the CLMS NRT product detailed below. The assimilation methods differ according to the demands of the respective land and ocean variables. Land and ocean data assimilation cycles run at different times, but within a coupled system. The current operational workflow is designed to run on the third day of each month, as all necessary input data for the previous month should be available from ERA5T, including days for interpolation. The restart files for the first day of the current month are used as initial conditions for the forecasts. This procedure, as tested within CERISE, represents a prototype for the upcoming seasonal forecast cycle. The NRT products are specified as the CLMS GLOBE OLCI RT1 Leaf Area Index (300 m resolution, ~10 d latency), ARGO ocean temperature and salinity, ERA5T thermodynamic variables and wind, SYNOP snow depth data from the MARS archive (stream “oper” instead of “era5”). The restart files for the 25 ensemble members are used to initiate the forecasts. The ICON XPP (Icosahedral Non-hydrostatic Model for Extended Predictions and Projections) system is currently in use at DWD and will be used in the future to run seasonal forecasts. The ICON XPP seasonal prediction system differs from reanalyses for numerical weather prediction (NWP) in terms of spatial resolution, the number of assimilated variables and the land surface model. It uses the JS-Bach land surface model (Schneck et al., 2022) instead of the TERRA land surface model (Schrodin and Heise, 2001), which is used in the operational NWP system. During the 2021 demonstration year, the LICs from the 30-year Phase 2 assimilation run, with the same land assimilation scheme components, are used as the baseline reanalysis for comparison with the NRT experiment.

4.2 MF

The SEKF LAI assimilation scheme, which was used for the Phase 2 reanalysis, is also used for the NRT demonstrator. The only difference is that a NRT version of the CLMS LAI is assimilated. CLMS GLOBE OLCI RT1 LAI observations are used, where RT1 refers to the NRT processing stream of the CLMS LAI product. The system assimilates NRT products retrieved from the CLMS GLOBE OLCI RT1 LAI (with a resolution of 300 m and a latency of ~10 days). Each daily cycle updates the land surface state using the most recent RT1 LAI observations available. Due to the latency of the LAI product, the assimilation cycle is performed up to the most recent date for which observations are available. For the 2021 demonstrator, the workflow was configured to ensure operational feasibility under realistic latency constraints. When LAI observations were unavailable for a given day, the model proceeded in forecast mode until new observations were ingested in the next assimilation window. The land surface states produced by LDAS-Monde at the end of each assimilation cycle are used as LICs for the seasonal forecast system. Updated LAI and SM states ensure physically consistent land initialisation, thereby improving the representation of land–atmosphere coupling in subsequent forecasts.

4.3 CMCC

The CMCC seasonal forecasting system uses the CERISE results to test the initial land surface conditions derived from a 30-member land reanalysis. This setup employs the SPREADS Ensemble Adjustment Kalman Filter (EAKF) data assimilation system. The operational workflow is designed to run on a daily basis. However, the primary bottleneck is the availability of the 2-day latency ERA5 Data Assimilation Ensemble) (EDA) atmospheric forcing. Due to this latency and the input requirements of the Community Land Model (CLM), it is not possible to run the system up to the current date. Consequently, on the first day of the month, the analysis runs up to midnight on the third day prior. Example: For a forecast on 1 February, the analysis runs up to 28 January. The initial conditions (ICs) used are the restart state at 00:00 UTC on 29 January. The system assimilates NRT products retrieved daily from the EODATA repository, which is the Copernicus data access platform for Earth observation products. The selected NRT variables are snow cover fraction from CLMS (300 m resolution, latency of less than 24 hours), surface soil moisture from SMOS (ESA, latency of less than 24 hours) with ASCAT (EUMETSAT, latency of less than 6 hours) as a backup, and LAI from CLMS (300 m resolution, updated every 10 days). The assimilation cycle extracts 30-member land states. For the seasonal forecast runs, the initial conditions are selected from these 30 ensemble members. This is different from the previous operational strategy, which relied on three independent CLM runs.

5 Results for LICs’ reanalyses

In this section, intercomparisons of simulated land surface temperature (LST), and snow water equivalent (SWE), are presented for Phase 2 simulations:

- global maps of mean LST and SWE,
- global monthly mean time series of LST and SWE,
- a basic verification of the consistency across demonstrators, including leaf area index (LAI).

5.1 Global maps of mean LST and SWE

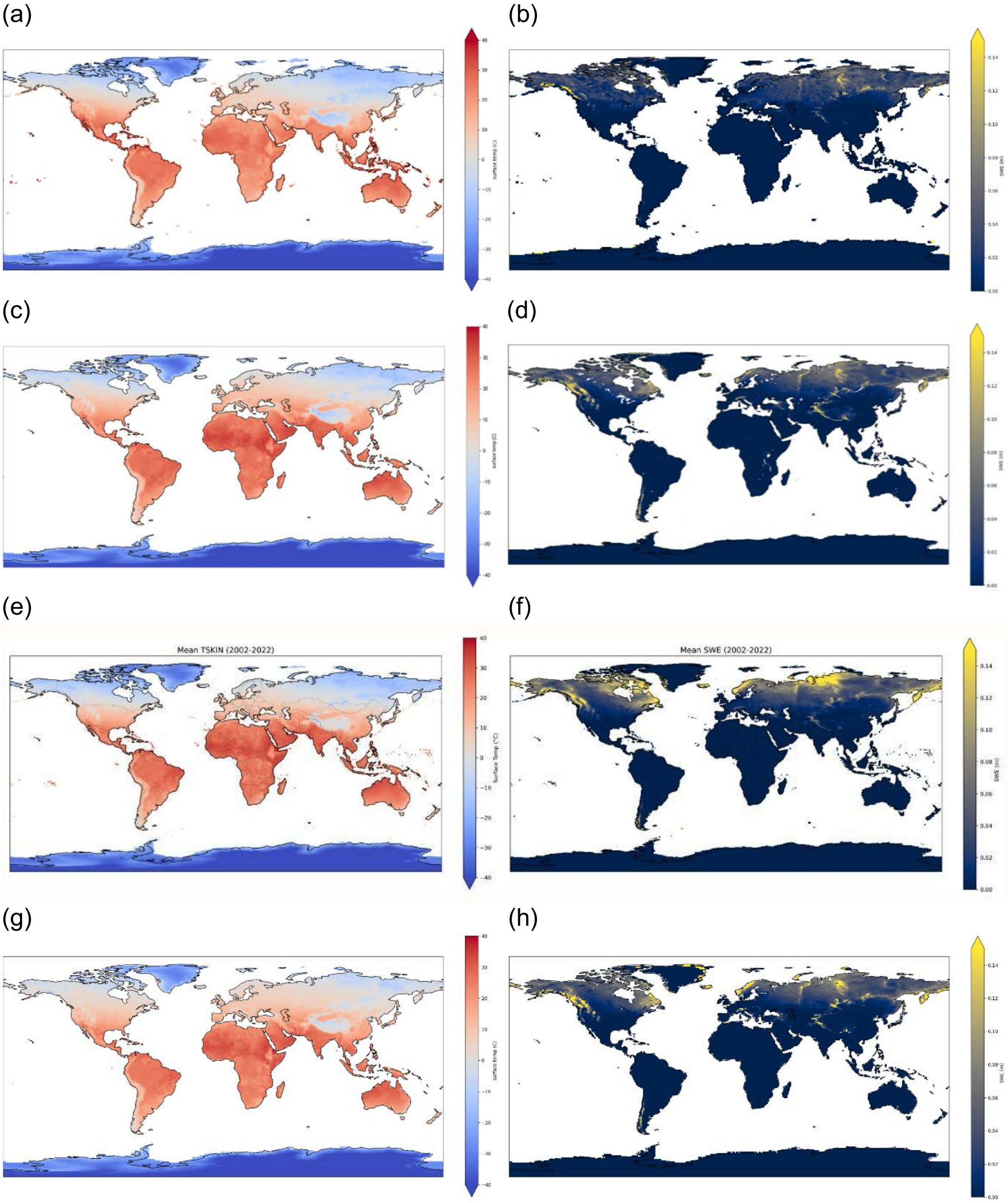
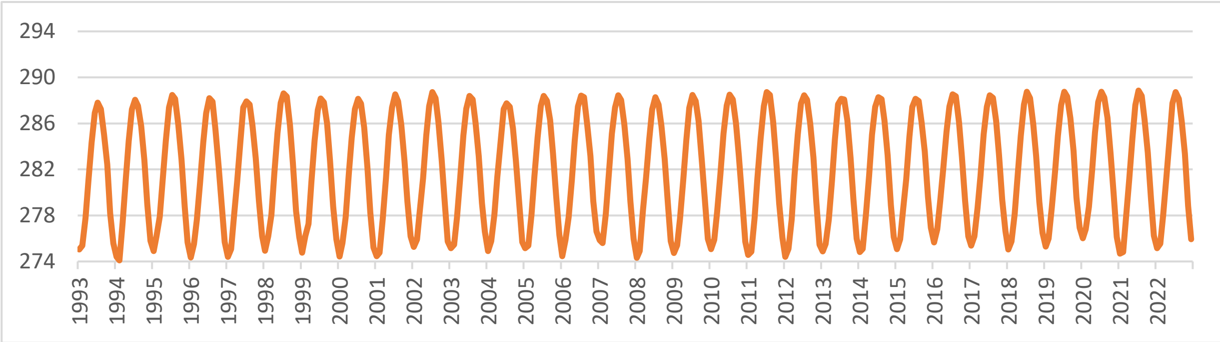


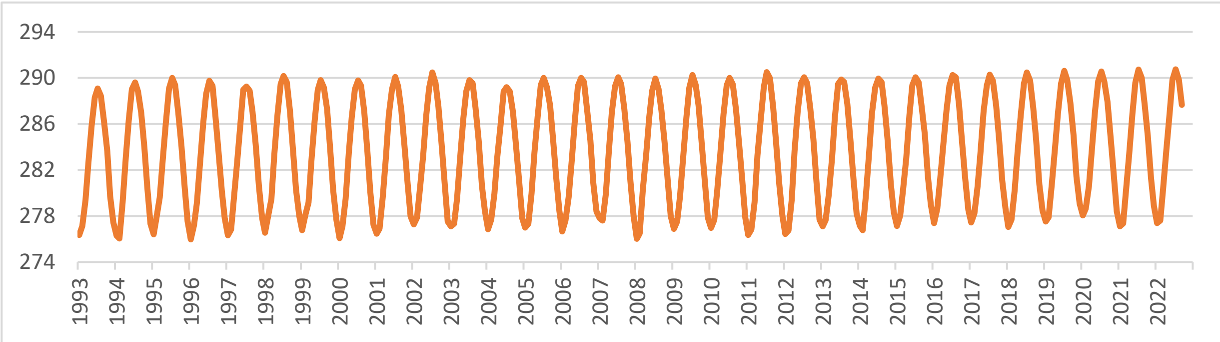
Figure 1 – Mean (a,c,e,g) LST and (b,d,f,h) SWE, for Phase 2 simulations of (a,b) DWD, (c,d) MF, (e,f) CMCC, and (g,h) ECMWF. DWD and MF simulations are averaged over the time period 1993-2022, CMCC over 2002-2022, and ECMWF over 1998-2019.

5.2 Global monthly mean time series of LST

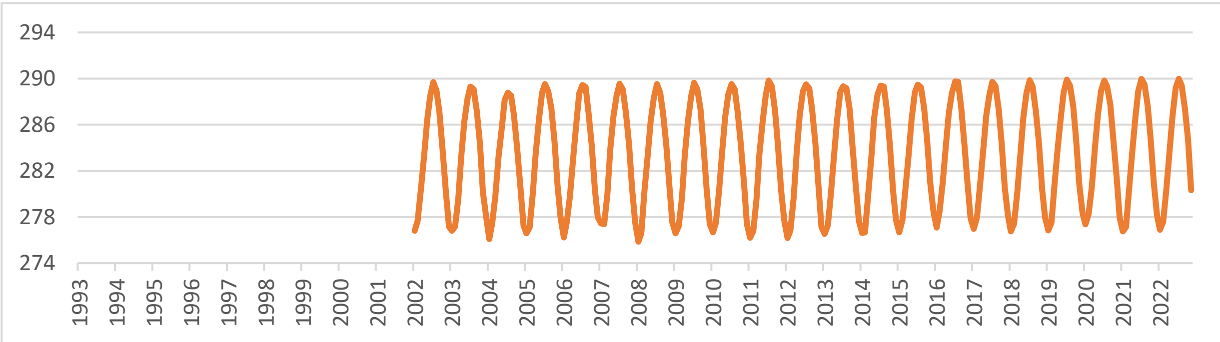
(a)



(b)



(c)



(d)

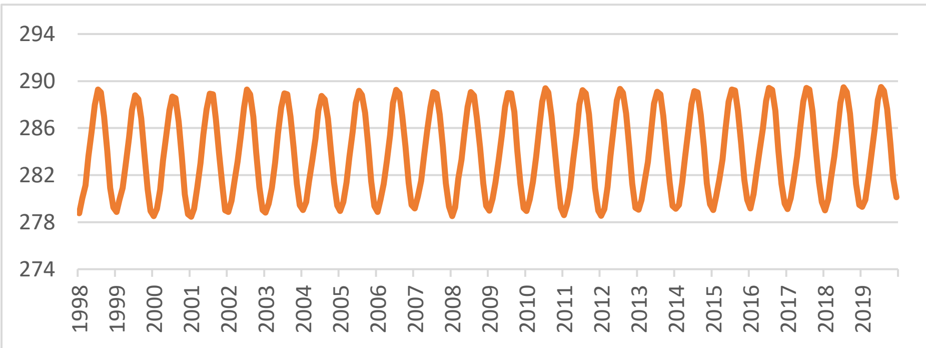
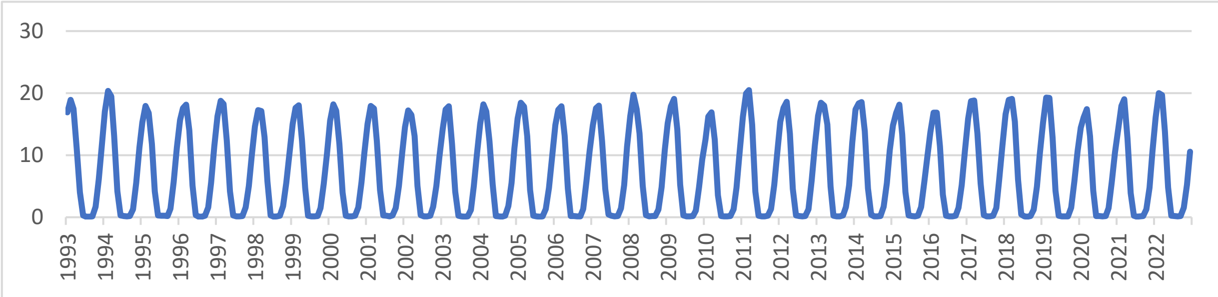


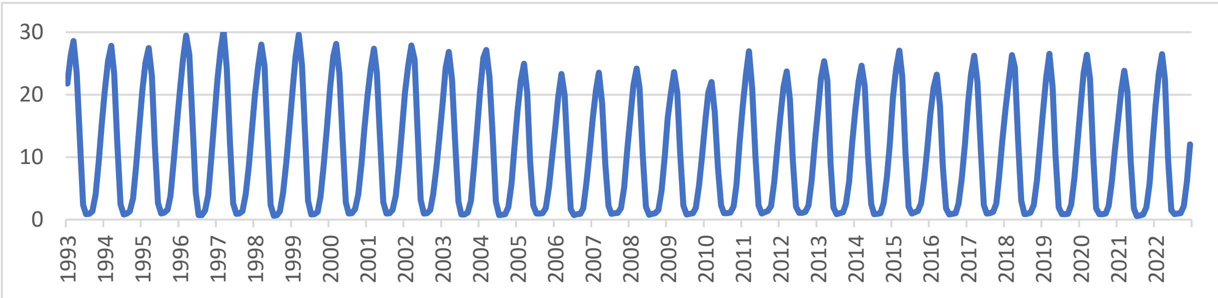
Figure 2 – Time series of the LST global monthly mean for (a) DWD, (b) MF, (c) CMCC, (d) ECMWF. LST is given in units of K.

5.3 Global monthly mean time series of SWE

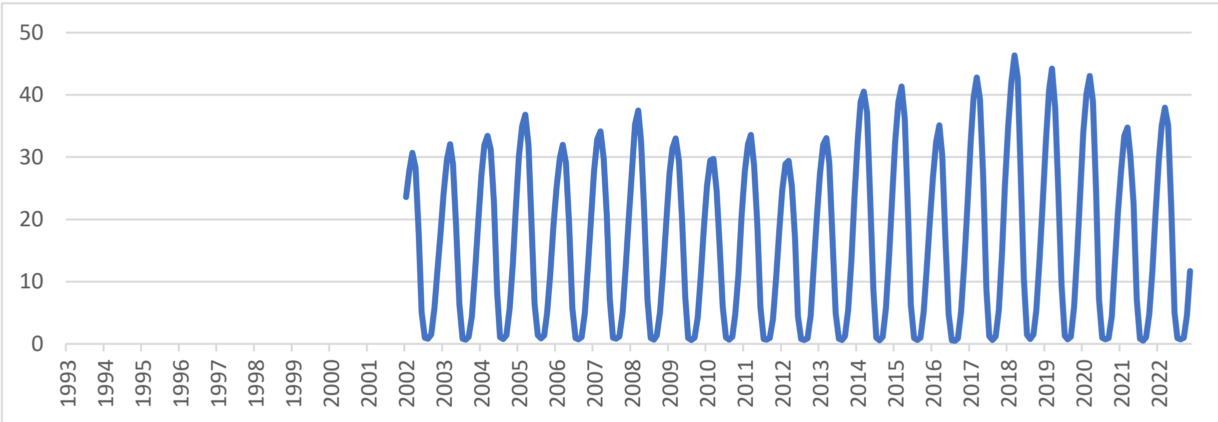
(a)



(b)



(c)



(d)

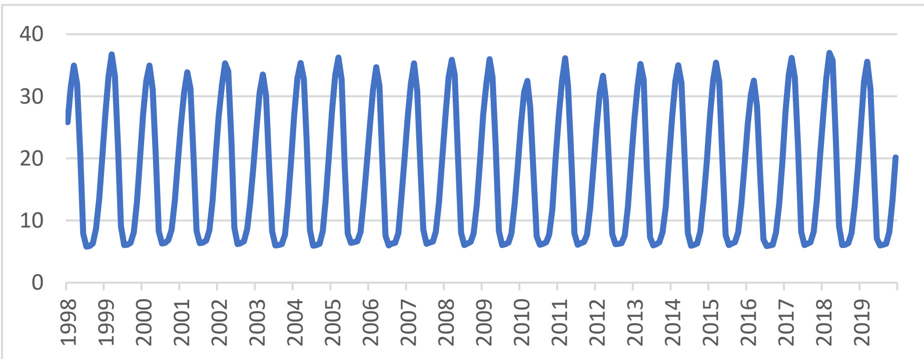


Figure 3 – Time series of the SWE global monthly mean for (a) DWD, (b) MF, (c) CMCC, (d) ECMWF. SWE is given in units of kg m^{-2}.

5.4 Verification of the consistency across demonstrators

Table 1 summarises the global mean values of the simulated snow water equivalent (SWE), leaf area index (LAI), and land surface temperature (LST) for all models from January 2002 to December 2019.

Table 1 – Global mean values of simulated SWE, LAI, and LST from Phase 2 initial conditions produced by DWD, MF, CMCC, and ECMWF. The time period from 2002 to 2019 is considered. LAI was not simulated by ECMWF.

	SWE (kg m^{-2})	LAI (m^2 m^{-2})	LST (K)
DWD	7	2.1	282.0
MF	10	1.6	283.7
CMCC	17	1.4	283.4
ECMWF	18	-	284.0

All SWE simulation values exclude continental ice sheets and glaciers. However, small permanent ice masses may be present in the obtained snow mass values. For instance, the ECMWF's global SWE values are consistently greater than 5.8 kg m^{-2} (Fig. 3) in July-August, whereas the minimum SWE values produced by the other models are smaller than 1 kg m^{-2}. The SWE values simulated by the MF are larger than those simulated by the DWD (see Fig. 4), but the global mean values correlate well (R^2= 0.93). The global mean LST values range from 282.0 for DWD to 284.0 K for ECMWF. Note that, on average, the minimum LST values in January and the maximum LST values in July are 4 K and 0.8 K larger, respectively, than those of DWD, as shown in Fig. 5. The ECMWF minimum LST values in January are 2 K higher and the maximum LST values in July are 0.9 K lower than those of MF. The ECMWF minimum LST values in January are 2.3 K higher and the maximum LST values in July are 0.3 K lower than those of CMCC. This discrepancy may be due to the fact that the ECMWF LST values correspond to soil temperature at a depth of 0.07 m, whereas the other models use a shallower soil temperature.

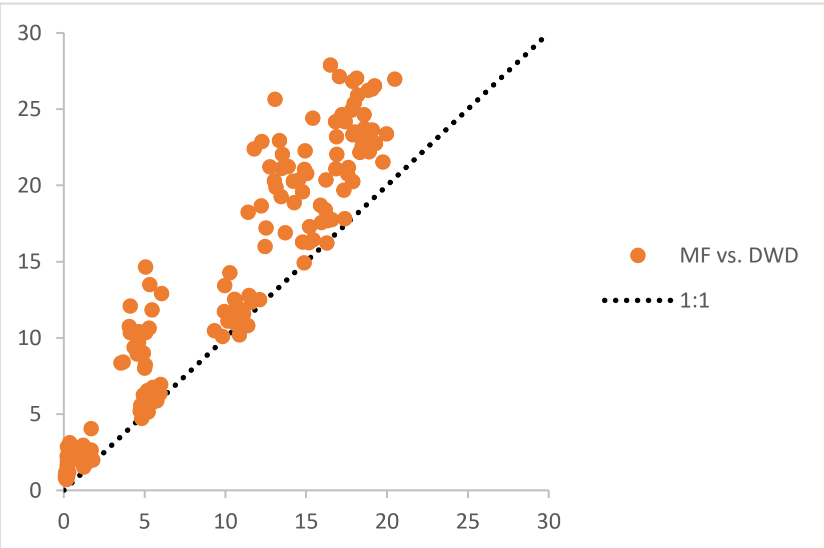


Figure 4 – Global monthly mean Snow Water Equivalent (in kg m^{-2}) from Phase 2 initial conditions produced by MF (y-axis) vs. DWD (x-axis). The time period from 2002 to 2019 is considered.

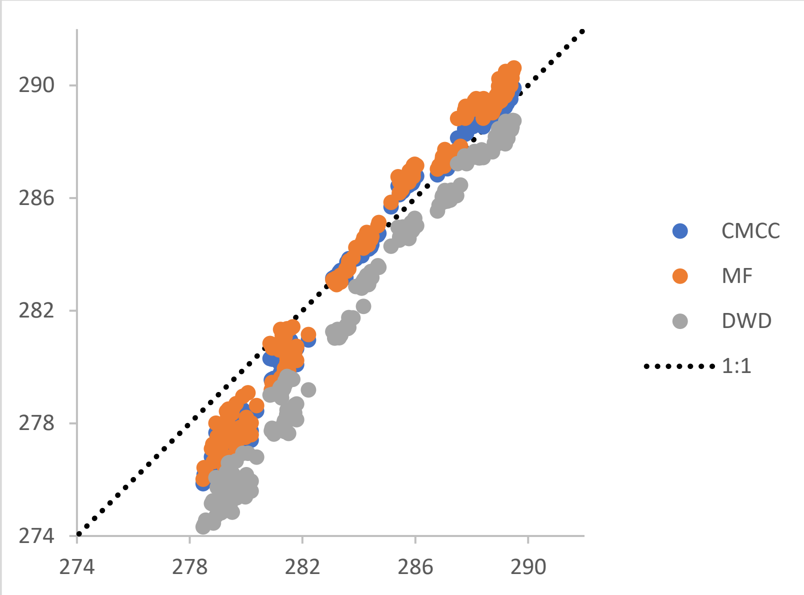


Figure 5 – Global monthly mean Land Surface Temperature (in K) from Phase 2 initial conditions produced by DWD, MF, CMCC (y-axis) vs. ECMWF (x-axis). The time period from 2002 to 2019 is considered.

In DWD Phase 2 simulations, the prognostic LAI variable of the land model differs substantially from the observed LAI, resulting in significant changes. These large increments, in turn, have a significant impact on LST, producing values that are, on average, 1.5 °C lower than in Phase 1. Work is ongoing to improve LAI data assimilation by using a diagnostic LAI variable closer to the observed LAI as input for the data assimilation routine. This variable incorporates a vegetation fraction correction which explicitly excludes desert regions. The resulting increments are smaller than those obtained when using the prognostic LAI as input for the data assimilation routine. Although these increments will still be added to the prognostic variable, the smaller ones will have a weaker impact on surface temperature, soil moisture and air temperature.

6 Results for NRT LICs (2021)

The following operational metrics have been defined in order to assess the feasibility of the NRT demonstrator:

- Latency compliance (the percentage of days on which NRT observations were ingested within 24 hours of acquisition; the target availability for all variables is >90%) and computing time (documentation of the CPU hours required per month to generate the LICs).

- Scientific quality is evaluated by comparing NRT LICs with retrospective land reanalysis (the baseline). The following metrics are calculated for the reference year of 2021, comparing the NRT LICs with the reanalysis LICs for soil moisture, snow cover fraction, and leaf area index. Root Mean Squared Error (RMSE), correlation, mean bias and, when applicable, ensemble spread (a comparison of ensemble spread versus RMSE is used to assess filter consistency).

6.1 DWD

This subsection evaluates the performance of the NRT LAI initial conditions LICs generated by the DWD system, comparing them with the retrospective reanalysis baseline. The analysis focuses on spatial and temporal consistency, using global maps and time series to evaluate variations in magnitude, variability and seasonal dynamics.

6.1.1 LAI

There are two main differences between the LAI products used for reanalysis and the NRT experiment. Firstly, the THEIA GEOV2-AVHRR product (THEIA Advanced Very High Resolution Radiometer; Verger et al., 2025) is provided on the 1st, 11th and 21st of each month, whereas the Copernicus NRT RT1 LAI product (i.e. CLMS GLOBE) provides data on the 5th, 16th and 25th. Consequently, the assimilation times and the data used differ by about 4–5 days. Secondly, the products originate from different data streams, and while GEOV2-AVHRR is a consolidated product, the NRT CLMS RT1 LAI is based on fewer observations than its RT6 consolidated counterpart. In Figure 6, the annual mean differences indicate large areas in the northern part of the Northern Hemisphere and over Asia where the NRT LAI analysis shows lower LAI values than using the consolidated product. In contrast, positive differences dominate the equatorial region and northern Africa. Negative differences dominate in southern South America. These regions of significant bias correspond to increased values in the RMSE metric and generally reflect areas of the globe with higher vegetation dynamics. The low correlations observed over the Sahara, central Australia and southern Argentina reflect the low variability present in these dry regions. Satellite retrievals of vegetation parameters in these regions are difficult and statistical evaluation is not very meaningful. In contrast, for large parts of the globe with vegetated areas, such as Europe, Eurasia, North America, and southern Asia, the correlation is close to one. The good overall agreement in vegetation dynamics between the two products is based on a similar data processing algorithm. Figure 6 shows the time series of the global mean LAI differences and the global mean LAI values, indicating that the NRT experiment has slightly lower LAI values overall, at around 0.05 m²/m², except during the northern hemisphere's growing season (March–June). The differences between the NRT experiment and the fully processed analysis are clearly visible here. However, since the model generally produces larger LAI values than the satellite data, negative increments occur at cycle time, resulting in the subsequent regrowth of vegetation in the model. The excessive impact of the assimilation triggers the pronounced LAI difference cycling shown in Figure 6 and RMSE is large (about 40 %). In such conditions, it is difficult to determine how closely the NRT LAI anomaly matches the anomaly in the reanalysis compared to the long-term climatology. This behaviour is the subject of an ongoing investigation.

6.1.2 SM

Figure 7 shows the differences in soil moisture between the reanalysis and the NRT experiment. These differences are caused by variations in evaporation from vegetation and snowmelt, resulting from the use of different observation datasets. Investigating the top soil layer (0–6.5 cm depth) reveals drier conditions for NRT in mid-latitude western Eurasia, with a difference absolute maximum value of around 3 mm soil water equivalent, corresponding to a relative difference of less than 5 % with respect to layer thickness. Other regions show a slight positive bias of around 1 mm of water equivalent or less, with more pronounced patterns over Europe and northern North America. These patterns generally correspond to regions with higher RMSE values and lower correlations. The time series reveals only minor differences in the global mean surface soil moisture between the two experiments, with a zero bias in the global mean during the winter. During the snowmelt and northern hemisphere vegetation growth season (March to May), slightly lower values of NRT evolve to around 0.25 mm, and these values persist throughout the summer. This dynamic period is also reflected in the highest RMSE metric values. These differences result from the smaller accumulation of snow depth during winter due to the absence of zero snow depth observations in the snow analysis for reanalysis.

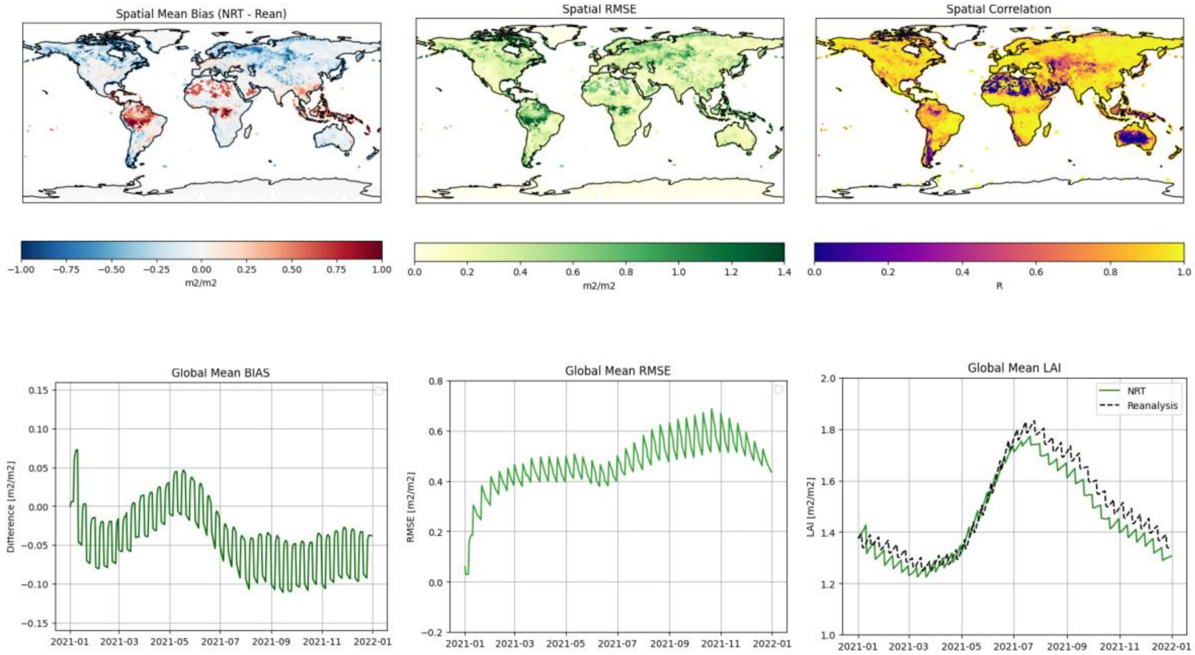


Figure 6 – Global maps and time series comparison of LAI (m^2m^{-2}) between the NRT experiment and the reanalysis LICs for the year 2021, for DWD. Top panels show spatial distributions of bias, RMSE, and correlation, while bottom panels present the global mean bias, RMSE, and LAI time series.

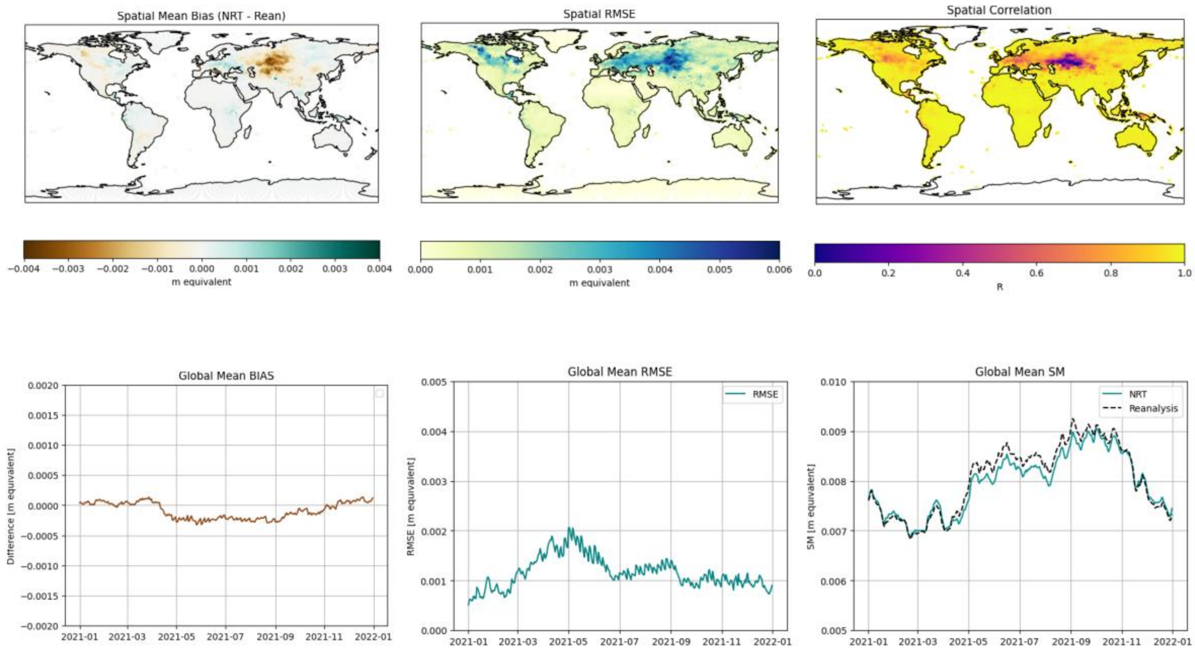


Figure 7 – Global maps and time series comparison of SM (m) between the NRT experiment and the reanalysis LICs for the year 2021, for DWD. Top panels display bias, RMSE, and correlation, while bottom panels illustrate the global mean bias, RMSE, and SM temporal evolution. The units for SM describe total column soil water for the top layer (see text).

6.1.3 SWE

The two-dimensional variational (2D-Var) scheme for snow depth analysis, developed within CERISE, uses SYNOP snow depth measurements from the MARS archive. For the Phase 2 reanalysis, data are extracted from the ERA5 stream, while the 'oper' stream is used for NRT. Zero snow depth observations were excluded from the retrieval for the reference, whereas they are used for NRT, which removes some of the high bias in SWE, showing up during the snow melt phase over selected regions in March and April. Regional differences are primarily observed in the mid-latitude regions of Europe and extend to the northern parts of Eurasia and North America. This reflects the global regions with the greatest snow accumulation dynamics (Figure 8). Negative biases are more prevalent in areas with greater observation coverage, particularly over central Europe and Canada. They are also more pronounced in regions where the change between snow-covered and non-snow-covered surfaces occurs more frequently, i.e. towards the southern edge of the snow-affected area. In these areas, zero snow depth measurements have the greatest impact. The annual cycle is well represented by the global mean values and the largest root mean squared errors during the snow season. The small overall biases between the two products indicate the feasibility of replacing the consolidated products with the new ones for NRT seasonal forecasting. The smaller SWE values in the NRT system are due to the inclusion of previously excluded valid observations of snow-free conditions.

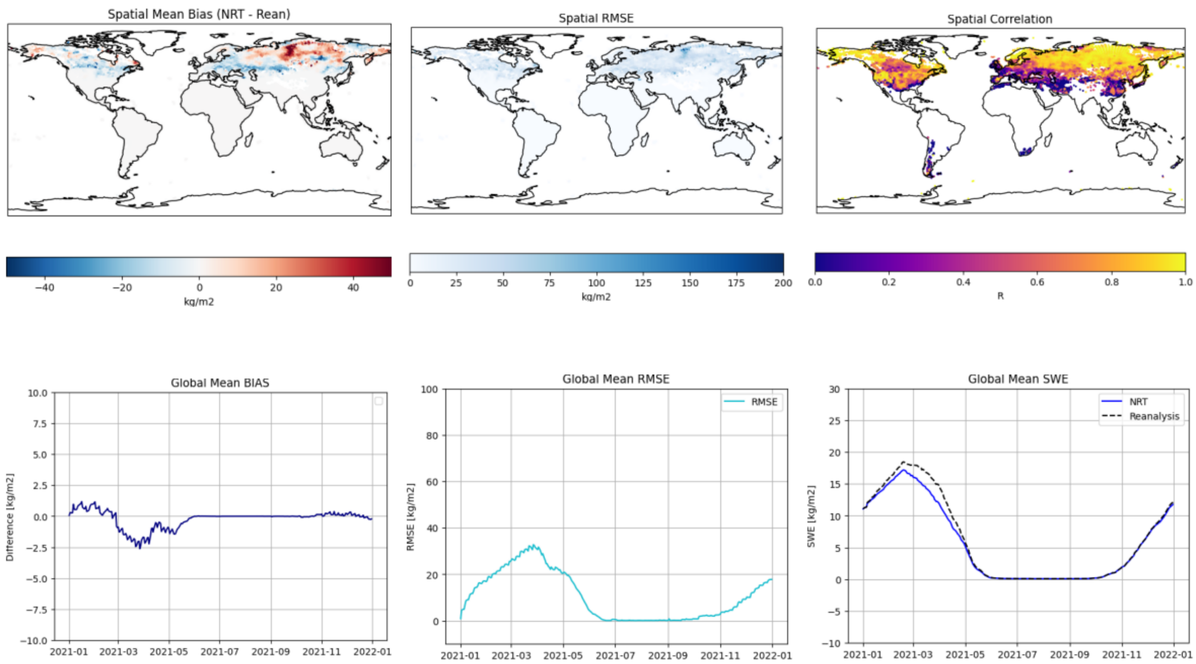


Figure 8 – Global maps and time series comparison of SWE (kg m^{-2}) between the NRT experiment and the reanalysis LICs for the year 2021, for DWD. Top panels present spatial distributions of bias, RMSE, and correlation, and bottom panels show the global mean bias, RMSE, and SWE time series.

6.2 MF

This subsection evaluates the performance of the NRT LAI initial conditions LICs generated by the Météo-France system, comparing them with the retrospective reanalysis baseline. The analysis focuses on spatial and temporal consistency, using global maps and time series to evaluate variations in magnitude, variability and seasonal dynamics.

6.2.1 LAI

In Figure 9, spatial differences in LAI between the reanalysis (RT6 LAI 300 m from CLMS) and the NRT IC (RT1 LAI 300 m from CLMS) experiment are mainly observed in regions with strong vegetation dynamics, such as North America, Europe, and parts of Asia and South America. These areas correspond largely to regions dominated by croplands or deciduous vegetation, where seasonal changes in leaf development are more pronounced. RMSE is generally low (5 % or less) but increases in productive regions during the growing season, reflecting stronger vegetation variability. Despite these localised differences, correlation remains high across most land areas, indicating a consistent temporal evolution of vegetation between the two experiments. Given that both experiments rely on similar observational sources, such agreement is expected. Therefore, the small differences observed should not be interpreted as a systematic improvement or degradation, but rather as the result of using near-real-time observations, which are less temporally consolidated than delayed-mode products. The temporal analysis shows a consistent seasonal cycle in both experiments, with very similar global mean LAI evolution. Small fluctuations in the NRT time series are visible, particularly during periods of rapid vegetation growth. These variations reflect the direct response of the deterministic SEKF to incoming observations and the absence of ensemble-based smoothing, which can lead to slightly higher short-term variability. Overall, the results indicate that the NRT configuration provides LAI initial conditions that are consistent with the reanalysis in terms of large-scale spatial patterns and seasonal dynamics, supporting the use of RT1 observations as a suitable alternative to RT6 in an operational context. While the strong agreement between the NRT (RT1) and reanalysis (RT6) configurations is expected due to the use of similar observational products, this does not imply that the assimilation system has a negligible impact. Sensitivity experiments using NRT products with shorter latency (e.g. RT0, not shown here) reveal significantly larger differences in the analysed LAI fields. This indicates that the system effectively responds to observational inputs, but that the magnitude of the impact depends on the consistency and robustness of the assimilated product. In the case of RT1, the relatively high level of consolidation leads to behaviour closer to the reanalysis, whereas less mature products introduce larger deviations. This highlights the role of the deterministic SEKF framework in balancing observational information and model constraints, ensuring both responsiveness to observations and stability of the land surface state.

6.2.2 SM

There are only small differences in the spatial distribution of soil moisture between the reanalysis and the NRT IC experiment (Figure 10). Bias patterns are mainly observed in regions with strong hydrological variability, such as parts of North America, Eurasia, and some semi-arid areas. RMSE values remain generally low, although slightly higher values appear in regions influenced by snow processes or strong evapotranspiration variability. Correlations are very high across most regions, indicating that both experiments capture similar temporal variations in SM. The global mean bias remains close to zero throughout the year, suggesting a very consistent SM state between the two configurations, with no significant differences or systematic improvement. The RMSE gradually decreases during the year, while the global mean SM follows the expected seasonal cycle, with wetter conditions in winter and early spring and drier conditions during summer. Overall, the differences between the NRT and reanalysis configurations are minimal, indicating that using NRT initial conditions does not substantially impact the large-scale SM dynamics.

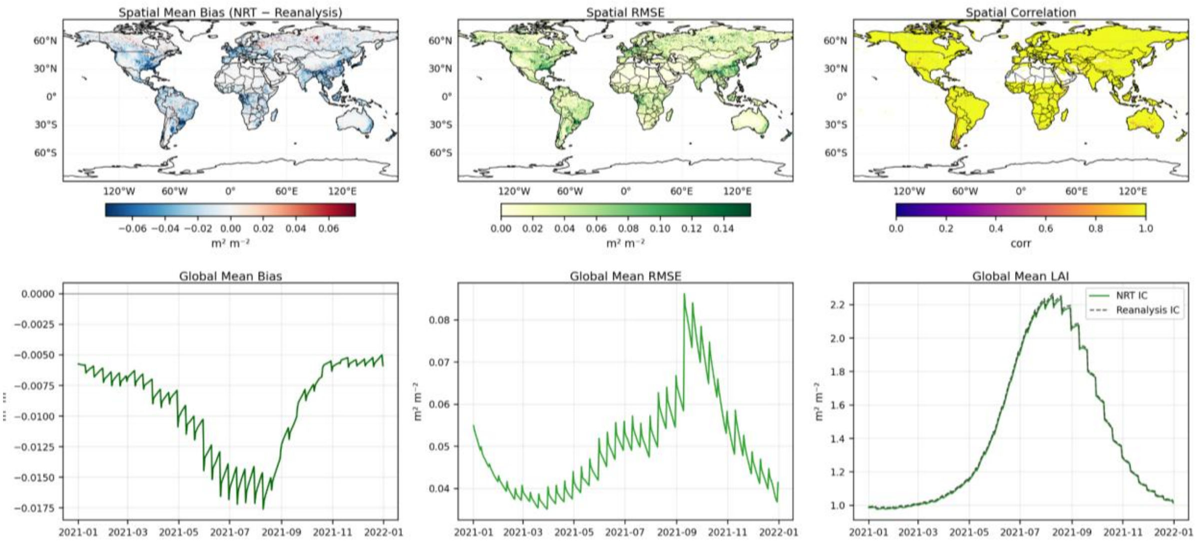


Figure 9 – Global maps and time series comparison of LAI (m^2m^{-2}) between the NRT experiment and the reanalysis LICs for the year 2021, for Météo-France. Top panels show spatial distributions of bias, RMSE, and correlation, while bottom panels present the global mean bias, RMSE, and LAI time series.

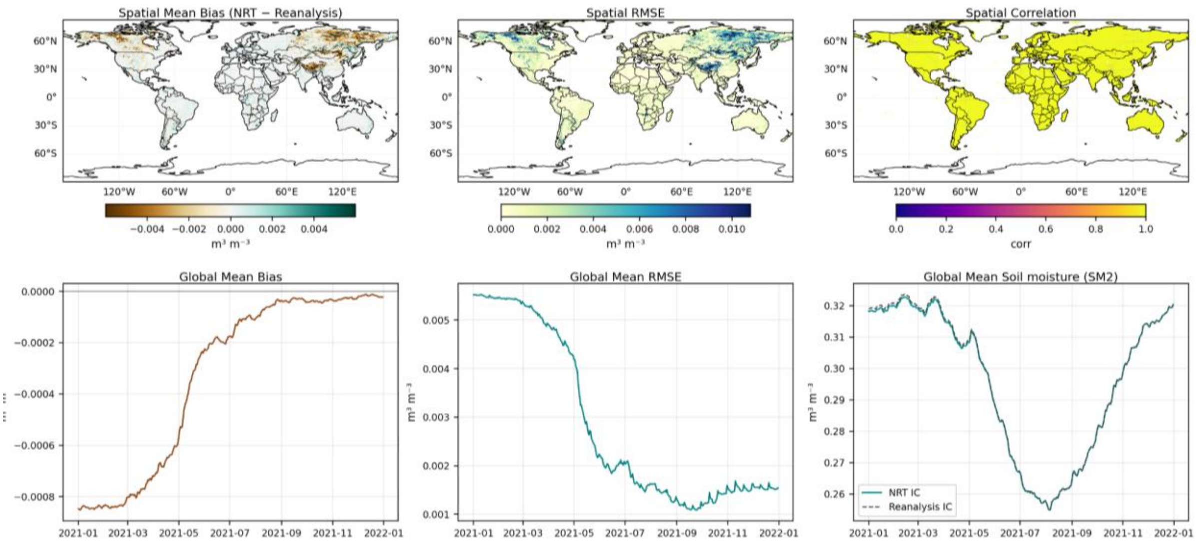


Figure 10 – Global maps and time series comparison of SM (m^3m^{-3}) between the NRT experiment and the reanalysis LICs for the year 2021, for Météo-France. Top panels display bias, RMSE, and correlation, while bottom panels illustrate the global mean bias, RMSE, and SM temporal evolution.

6.2.3 SWE

Spatial differences in SWE are concentrated in the mid- and high-latitude regions of the Northern Hemisphere, where seasonal snow accumulation dominates (Figure 11).

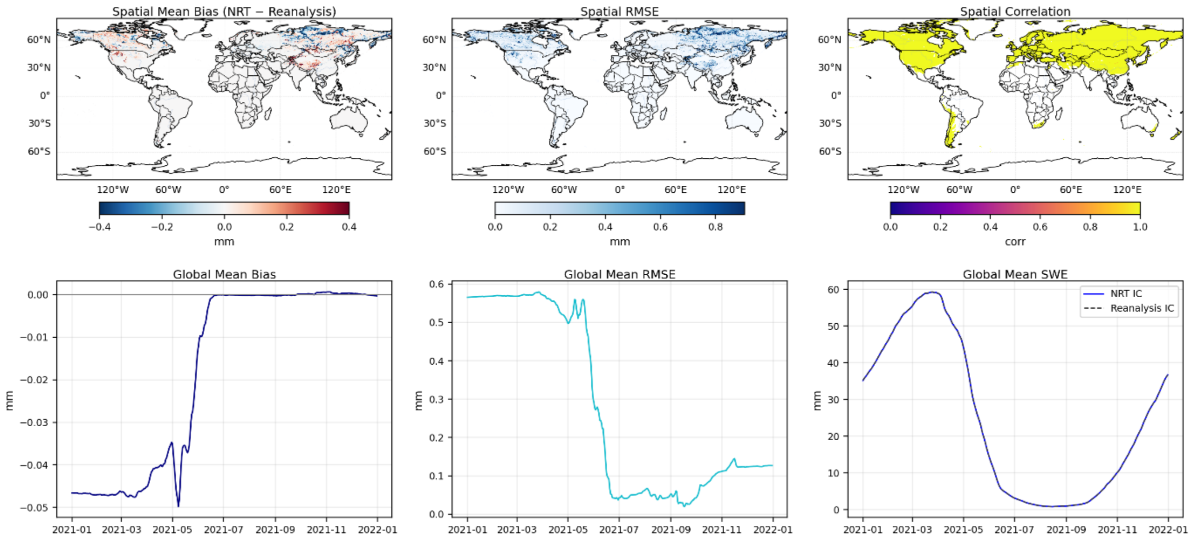


Figure 11 – Global maps and time series comparison of SWE (kg m^{-2}) between the NRT experiment and the reanalysis LICs for the year 2021, for Météo-France. Top panels present spatial distributions of bias, RMSE, and correlation, and bottom panels show the global mean bias, RMSE, and SWE time series.

Bias and RMSE differences are primarily observed across northern Eurasia and North America, reflecting differences in snow accumulation and melt processes between the reanalysis and the NRT IC experiment. Outside these snow-dominated regions, differences remain negligible. The temporal evolution reflects the typical seasonal snow cycle. RMSE is higher during winter, when snow accumulation is widespread, and decreases rapidly during spring as the snowpack melts. The global mean SWE shows a clear accumulation phase during winter, followed by a rapid decline during spring and summer. Both the reanalysis and the NRT experiments exhibit almost identical seasonal behaviour. This indicates that the NRT IC configuration preserves the large-scale snow dynamics represented in the reanalysis. This is consistent with the behaviour observed for LAI and SM. Please note that the SWE results in Figure 11 correspond to ISBA simulations, whereas Phase 2 uses ERA5 SWE inputs. This explains the different SWE peak values shown in Figure 3b.

6.3 CMCC

This section demonstrates the performance of the land ICs produced by the CMCC land data assimilation system. The analysis examines the impact of a 30-member Ensemble Adjustment Kalman Filter (EAKF) and a dynamic bias-handling approach utilising adaptive inflation on the spatial and temporal consistency of the assimilation of LAI, SM and snow cover fraction.

6.3.1 LAI

Figure 12 shows that differences in LAI can be identified between the reanalysis and the NRT IC in subtropical regions, which are likely associated with intensive agricultural activities in South America, Africa, North America and Asia. In these regions, as well as in Australia, lower correlations between the experiments are observed ($r < 0.6$ in transitional zones), highlighting discrepancies related to phenological phases or agricultural management practices. These practices include planting and harvesting cycles, which may influence the assimilation of remote sensing data. The bias fluctuates around zero, showing a slight negative trend in the middle of the year, which indicates underestimation during the peak of vegetative growth. The RMSE reaches its highest values during the Northern Hemisphere summer (JJA), which coincides with maximum variability in LAI in agricultural and forested areas. Nevertheless, the correlation remains consistently high, confirming strong temporal agreement despite localised spatial variability. Time series analysis reveals a pronounced NRT peak during the growing season, when vegetation dynamics are more active and responsive to atmospheric forcing. These differences likely reflect the influence of assimilated satellite observations on the timing and magnitude of leaf development, particularly in regions dominated by croplands or deciduous forests, where phenological transitions occur rapidly. Discrepancies during the onset of greening or senescence directly affect model LICs since LAI regulates surface fluxes by controlling photosynthesis, transpiration, canopy interception and surface roughness. Variations in LAI also influence the partitioning of net radiation into sensible and latent heat fluxes, as well as soil moisture depletion rates. The faster growth and decay rates observed in the NRT initial conditions compared to the reanalysis are primarily driven by differences in the temporal compositing algorithms of the satellite products combined with the physical sensitivity of the model. The retrospective reanalysis uses GLASS LAI, a product designed to be consistent over time. It uses a symmetric observation window and applies strong temporal filtering by aggregating satellite acquisitions from before and after the target date. While this extensive smoothing creates a stable time series, it inherently dampens rapid phenological extremes. By contrast, the NRT demonstrator assimilates the Copernicus CLMS RT1 LAI. To meet operational near-real-time constraints, the RT1 product uses an asymmetric window that relies almost entirely on observations leading up to the target date. As it lacks future data on which to base the smoothing, this approach results in weaker temporal regularisation. Consequently, RT1 is more responsive to rapid changes in vegetation (e.g. sudden greening or harvesting) and exhibits greater short-term variability. As the NRT system processes this responsive, less-smoothed preliminary data, rapid phenological transitions are preserved rather than suppressed. This effect is amplified by the CLM5 model's high physiological sensitivity to LAI: the model regulates photosynthesis, transpiration and surface energy fluxes based on this input. Consequently, CLM5 translates these strong observational gradients directly into the other prognostic states, thereby impacting their faster growth and decay rates.

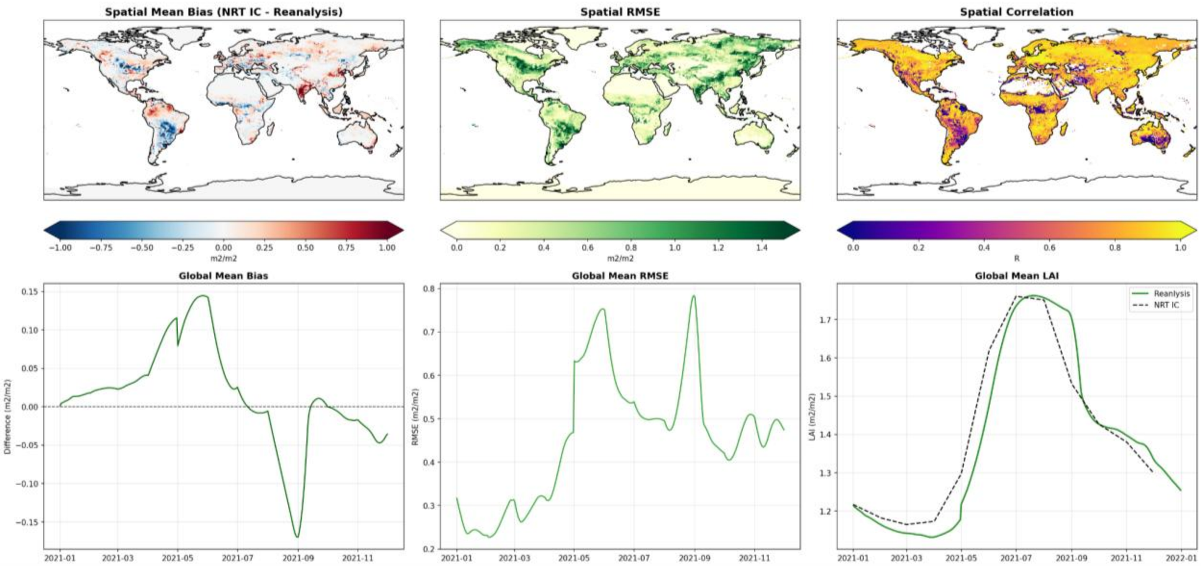


Figure 12 – Global maps and time series comparison of LAI (m^2m^{-2}) between the NRT experiment and the reanalysis LICs for the year 2021, for CMCC. Top panels show spatial distributions of bias, RMSE, and correlation, while bottom panels present the global mean bias, RMSE, and LAI time series.

6.3.2 SM

Analysis of soil moisture reveals more pronounced spatial differences between the reanalysis and the NRT ICs in regions influenced by snow, particularly in the high latitudes of North America and Eurasia (see Figure 13). In these regions, higher bias and RMSE are associated with seasonal freeze–thaw processes affecting soil hydrology and surface energy exchanges. By contrast, arid and semi-arid regions such as the Sahara, parts of Australia and the Middle East exhibit lower correlations, suggesting discrepancies in the representation of short-term variability. The reanalysis captures episodic responses to sparse precipitation events, whereas the NRT LICs exhibit comparatively smoother dynamics in their forecasts. In terms of time, the global mean bias remains close to zero throughout the year, with slightly positive values in the first half of the year and negative values in late summer and early autumn. The RMSE increases during late spring and early summer, coinciding with snowmelt and increased evapotranspiration, and then decreases. The global mean SM follows a consistent seasonal cycle, peaking in winter and early spring, then declining towards summer. Differences between the reanalysis and the NRT LIC forecast are more evident during the transition between seasons, when rapid cycles of wetting and drying amplify contrasts in short-term SM variability. The high variability and spike in RMSE observed during the transitional seasons (particularly in May and June) can be interpreted as physical and statistical responses to the processing and assimilation of NRT data. The retrospective reanalysis uses the ESA CCI soil moisture product. As this is a historical dataset designed for long-term stability, it eliminates high-frequency noise. However, the NRT demonstrator relies on preliminary SMOS and ASCAT retrievals. These real-time streams bypass extensive retrospective filtering, preserving much more day-to-day variance. The timing of the peak in RMSE in May and June is also significant. This occurs during the most active period for vegetation greening and snowmelt in the Northern Hemisphere. This suggests that the system simultaneously assimilates SM, LAI and Snow Cover Fraction, and that the EAKF filter updates the state and short-term variability from highly responsive NRT LAI and snow observations. This statistical effect is then compounded by the physical coupling inside the CLM5 model. Rapid, unsmoothed updates to vegetation alter transpiration and canopy interception immediately, while sharp changes in snow cover alter the flow of meltwater.

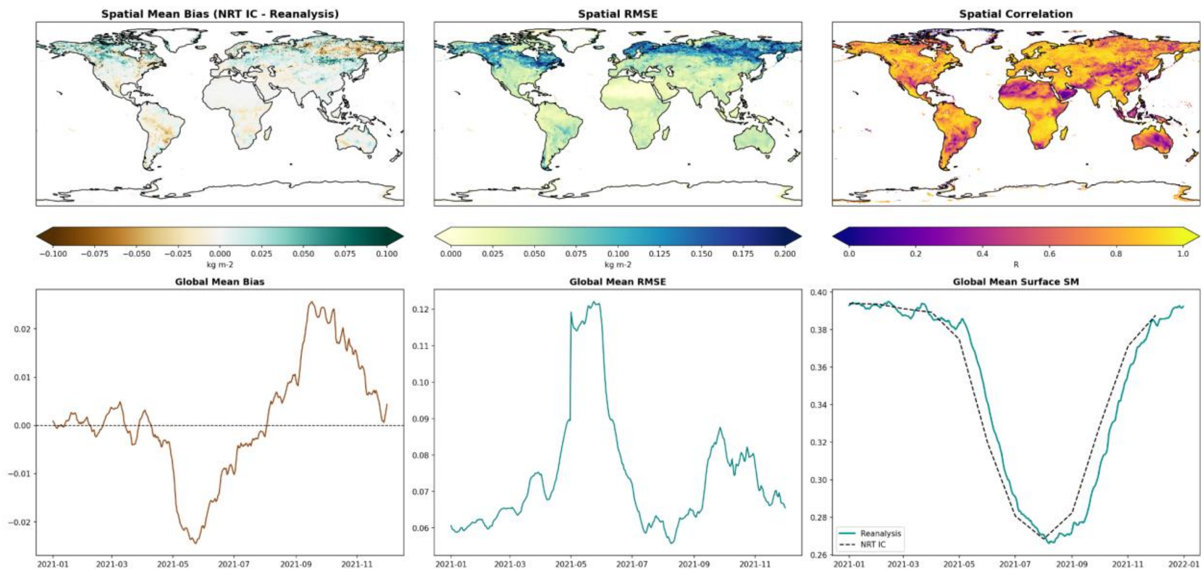


Figure 13 – Global maps and time series comparison of SM (m^3m^{-3}) between the NRT experiment and the reanalysis LICs for the year 2021, for CMCC. Top panels display bias, RMSE, and correlation, while bottom panels illustrate the global mean bias, RMSE, and SM temporal evolution.

Together, these factors cause abrupt, short-term fluctuations in SM levels during this time. Nevertheless, the magnitude of the May–June peak warrants closer inspection. As this NRT demonstration was run without a dedicated multi-year spin-up phase, the system may be experiencing residual imbalances from the initial state. These imbalances could be amplifying the differences during such a highly variable seasonal transition.

6.3.3 SWE

Figure 14 shows that the analysis of SWE highlights spatial differences concentrated in the mid- and high-latitude regions of the Northern Hemisphere, where seasonal snow dynamics dominate.

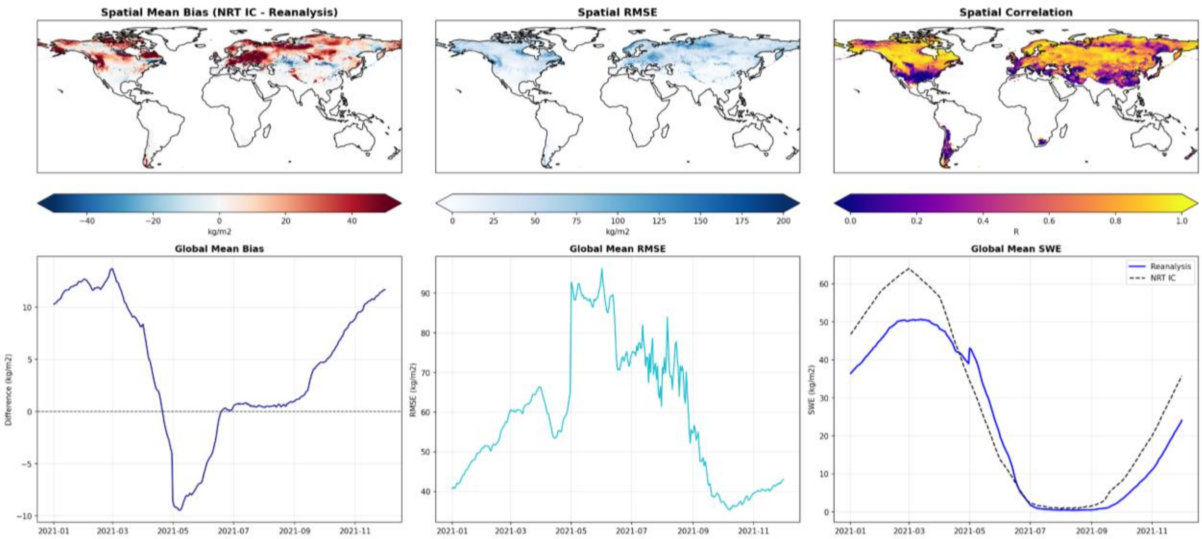


Figure 14 – Global maps and time series comparison of SWE (kg m^{-2}) between the NRT experiment and the reanalysis LICs for the year 2021, for CMCC. Top panels present spatial distributions of bias, RMSE, and correlation, and bottom panels show the global mean bias, RMSE, and SWE time series.

The bias is predominantly positive throughout the year, indicating that the reanalysis generally underestimates SWE volumes compared to the NRT LIC forecast, particularly during the period of maximum accumulation. However, the sign of the bias changes during the melt phase. The RMSE increases progressively from winter to spring, reaching higher values in late spring and early boreal summer. This coincides with the snowmelt period when small differences in melt rates can produce substantial discrepancies in water equivalent content. The temporal evolution of the global indicators reflects the annual snow cycle. The mean bias becomes more negative during spring and early summer, suggesting that the NRT either anticipates or intensifies the melting process more quickly than the reanalysis does. The RMSE follows a similar pattern, peaking during the transition from accumulation to ablation — a period characterised by strong spatial variability and sensitivity to atmospheric forcing. The time series of the global mean SWE shows gradual growth from autumn to a spring maximum, followed by a pronounced decline in summer. The reanalysis exhibits a lower amplitude than the NRT LIC forecast. The differences observed during the freezing and melting phases directly impact LICs, as snow storage influences the surface energy balance, post-melt soil moisture and sensible and latent heat fluxes. The higher short-term variability and accelerated phase shifts observed in the NRT SWE time series are primarily driven by the observational streams and their multivariate assimilation. The retrospective reanalysis relies on the ESA-CCI Snow Cover Fraction over Ground (SCFG) product, a dataset which applies symmetric temporal smoothing in order to generate a consistent long-term record. The NRT demonstrator utilises the preliminary Copernicus CLMS SCF product. Operating under operational constraints with less than 24-hour latency, this data stream lacks retrospective harmonisation and is therefore significantly more sensitive to high-frequency, short-term fluctuations. This highly responsive SCF data is assimilated simultaneously with LAI and soil moisture through the multivariate EAKF. This means that its undamped temporal variability is statistically propagated into the SWE state vector during the analysis step. Furthermore, this variability immediately affects the CLM5 surface energy balance. Constant modifications to the surface albedo and net radiation partitioning caused by high-frequency fluctuations in both vegetation density and snow cover drive the observed short-term changes in snowmelt rates, resulting in a highly responsive, oscillating SWE time series.

7 Conclusion

Creating balanced initialisation datasets and delivering them to the C3S community is an essential step in producing seasonal forecasts. This process is also a critical first step in developing prototype reanalyses and seasonal forecast demonstrators, which could eventually be approved as systems for predicting climate variability on seasonal timescales. In this study, DWD, MF, CMCC, and ECMWF have developed reanalysis prototypes. However, despite observations being assimilated, significant differences remain in land surface variables between the various analyses (e.g. global average SWE and LAI). This calls into question the idea that we can currently use land surface data to meaningfully drive seasonal predictions. However, progress has been made in that, while land surface variable values still vary widely, the range is smaller than it would be without assimilation. The NRT demonstrators implemented by DWD, MF and CMCC show that LICs can be produced within operational time constraints. Results for the shared simulation year of 2021 demonstrate that NRT observations from sources such as CLMS, SMOS and ASCAT can be successfully assimilated with acceptable latency. However, quality assessments show that the closeness of the NRT analyses to corresponding baseline reanalyses varies between systems. The MF NRT analysis introduces little bias into the forecast relative to the hindcast baseline, showing small differences and a high degree of correlation for all variables. By contrast, the DWD and CMCC analyses reveal significant discrepancies and large areas of low correlation. This suggests that NRT and reanalysis data from these two systems cannot be used in a reforecast/forecast system at this stage. To reach the highest levels of operational readiness for these systems, continued development beyond the duration of the CERISE project is essential, building on the significant progress achieved within it.

8 References

Albergel, C., Munier, S., Leroux, D. J., Dewaele, H., Fairbairn, D., Barbu, A. L., Gelati, E., Dorigo, W., Faroux, S., Meurey, C., Le Moigne, P., Decharme, B., Mahfouf, J.-F., and Calvet, J.-C.: Sequential assimilation of satellite-derived vegetation and soil moisture products using SURFEX_v8.0: LDAS-Monde assessment over the Euro-Mediterranean area, Geosci. Model Dev., 10, 3889–3912, https://doi.org/10.5194/gmd-10-3889-2017, 2017.

Asensio, H., Messmer, M., Lüthi, D., Osterried, K., Jucker, J., Canton, J., Sommer, P., Helmert, J., and Jähn, M.: EXTPAR User and Implementation Guide (v5.14). Center for Climate Systems Modeling (C2SM). Available at https://www.cosmo-model.org/content/support/software/ethz/EXTPAR_user_and_implementation_manual.pdf, 2023.

Boussetta, S., Balsamo, G., Arduini, G., Dutra, E., McNorton, J., Choulga, M., Agustí-Panareda, A., Beljaars, A., Wedi, N., Munõz-Sabater, J. and De Rosnay, P., 2021. ECLand: The ECMWF land surface modelling system. *Atmosphere*, 12(6), p.723. https://doi.org/10.3390/atmos12060723

Calvet, J.-C., Noilhan, J., Roujean, J.-L., Bessemoulin, P., Cabelguenne, M., Olioso, A., and Wigneron, J.-P.: An interactive vegetation SVAT model tested against data from six contrasting sites, Agr. Forest Meteorol., 92, 73–95, https://doi.org/10.1016/S0168-1923(98)00091-4, 1998.

Calvet, J.-C., Gibelin, A.-L., Roujean, J.-L., Martin, E., Le Moigne, P., Douville, H., and Noilhan, J.: Past and future scenarios of the effect of carbon dioxide on plant growth and transpiration for three vegetation types of southwestern France, Atmos. Chem. Phys., 8, 397–406, https://doi.org/10.5194/acp-8-397-2008, 2008.

Cardinali, C., Conti, G., Guatura, M., Saarinen, S., Gonçalves De Gonçalves, L. G., Anderson, J., and Raeder, K.: SPREADS: From Research to Operational Open-Source Data Assimilation System, EGUsphere [preprint], https://doi.org/10.5194/egusphere-2025-4294, 2025.

Decharme, B., Delire, C., Minvielle, M., Colin, J., Vergnes, J., Alias, A., Saint-Martin, D., Séférian, R., Sénési, S., and Voldoire, A.: Recent Changes in the ISBA-CTRIP Land Surface System for Use in the CNRM-CM6 Climate Model and in Global Off-Line Hydrological Applications, J. Adv. Model. Earth Syst., 11, 1207–1252, https://doi.org/10.1029/2018MS001545, 2019.

de Rosnay, P., Drusch, M., Vasiljevic, D., Balsamo, G., Albergel, C. and Isaksen, L.: A simplified Extended Kalman Filter for the global operational soil moisture analysis at ECMWF. Q.J.R. Meteorol. Soc., 139: 1199-1213. https://doi.org/10.1002/qj.2023, 2013.

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

Gibelin, A.-L., Calvet, J.-C., Roujean, J.-L., Jarlan, L., and Los, S. O.: Ability of the land surface model ISBA-A-gs to simulate leaf area index at the global scale: Comparison with satellites products, J. Geophys. Res., 111, D18102, https://doi.org/10.1029/2005JD006691, 2006.

Mucia, A., Bonan, B., Zheng, Y., Albergel, C., and Calvet, J.-C.: From monitoring to forecasting land surface conditions using a land data assimilation system: Application over the contiguous United States, Remote Sens., 12, 12, https://doi.org/10.3390/rs12122020, 2020.

Müller, W. A. et al.: ICON: Towards vertically integrated model configurations for numerical weather prediction, climate predictions, and projections, Bulletin of the American Meteorological Society 106 (6) E1017-E1031, https://doi.org/10.1175/BAMS-D-24-0042.1, 2025.

Nerger, L., Danilov, S., Hiller, W., Schröter, J.: Using sea-level data to constrain a finite-element primitive-equation ocean model with a local SEIK filter, Ocean Dynamics, 56(5/6), 634-649, https://doi.org/10.1007/s10236-006-0083-0, 2006.

Schneck, R., Gayler, V., Nabel, J. E. M. S., Raddatz, T., Reick, C. H., and Schnur, R.: Assessment of JSBACHv4.30 as a land component of ICON-ESM-V1 in comparison to its predecessor JSBACHv3.2 of MPI-ESM1.2, Geosci. Model Dev., 15, 8581–8611, https://doi.org/10.5194/gmd-15-8581-2022, 2022.

Schrodin, R., and Heise, E.: The multi-layer version of the DWD soil model TERRA_LM, Technical report, DWD, Offenbach am Main, https://dwdbib.dwd.de/retrosammlung/content/titleinfo/87001, 2001.

Verger, A., Weiss, M., Baret, F.: GEOV2-AVHRR: Continuous and consistent time series of global leaf area index and fraction absorbed PAR from 1981 to 2022, Remote Sensing of Environment, 331, 115029, https://doi.org/10.1016/j.rse.2025.115029, 2025.

Wong, A. P. S., et al.: Argo Data 1999–2019: Two Million Temperature-Salinity Profiles and Subsurface Velocity Observations From a Global Array of Profiling Floats, Frontiers in Marine Science, 7(700), https://doi.org/10.3389/fmars.2020.00700, 2020.

Document History

Version	Author(s)	Date	Changes
1.0	JC Calvet	30/04/2026	Initial version
1.1	All co-authors	04/06/2026	Submitted version
1.2	JC Calvet	24/06/2026	Draft revised version
1.3	JC Calvet	25/06/2026	Revised version

Internal Review History

Internal Reviewers	Date	Comments
Jeff Knight (UK Met Office), Patricia de Rosnay (ECMWF)	June 2026	Minor revisions