CMUG Extension Phase 2 Deliverable
Reference: D5.3.v1: Task 4 annual report
Due date: Sep 2024
Submission date: Nov 2024
Version: 1.2.0

Climate Modelling User Group (CMUG)

Extension Phase 2 Deliverable D5.3.v1:

WP4: Annual report on progress achieved in integrating CCI ECVs into ESMValTool and the development of associated tools and diagnostics

Centres providing input: DLR, Met Office, University of Reading

Version	Date	Status
1.0.0	14 August 2024	first version
1.1.0	13 September 2024	input from U Reading
1.2.0	12 November 2024	version submitted to Met Office

Table of Contents

WP4: Annual report on progress achieved in integrating CCI ECVs into ESMValTool and the development of associated tools and diagnostics

Summary

This report summarizes the progress made within WP4 between September 2023 and August 2024. WP4 consists of two main tasks, of which the first is to enhance the ESMValTool with additional diagnostics for the evaluation of global climate models with ESA CCI and CCI+ data. This includes in particular the implementation of new CCI datasets such as SNOW and PERMAFROST as well as updating existing datasets where needed (WP4.1). The second main task is to explore possibilities to take advantage of the uncertainty information provided with the CCI datasets for model evaluation starting with LST.

Section 1 gives and overview of the CCI datasets that have been updated, extended or newly implemented into ESMValTool along with some example plots (WP4.1). Implementation of uncertainty estimates into ESMValTool for LST is described in Section 2 including details on the mathematical background and method as well as all steps taken to implement this approach into ESMValTool and recommendations for future work (WP4.2).

1 Implementation of CCIs SNOW and PERMAFROST into ESMValTool and update of existing datasets (WP4.1)

This section provides an overview of the CCI datasets newly implemented into ESMValTool (Section 1.1), already implemented datasets that have been extended (Section 1.2) and datasets implemented that have been updated to their recent version (Section 1.3). Table 1 summarizes all CCI datasets covered by this report.

For each of the datasets implemented into ESMValTool, scripts have been written that allow for automatic downloading and reformatting of the CCI data (so-called “CMORization”) for use with ESMValTool. Configuration files for each of the datasets defining e.g. the variables, filenames and time periods to be processed allow for an easy and user-friendly adaptation of the downloading and reformatting scripts to new versions of the CCI datasets once available.

Table 1 CCI datasets newly implemented, updated or extended in ESMValTool.

Dataset	Version	Variables
AEROSOL	SU (A)ATSR v4.3	od550aer, od870aer, od550lt1aer, abs550aer, od550aerStderr, od870aerSdterr (daily and monthly)
LAND SURFACE TEMPERATURE	v3.0	tsDay, tsLCDay, tsLCNight, tsLocalAtmErrDay, tsLocalAtmErrNight, tsLocalSfcErrDay, tsLocalSfcErrNight, tsNight, tsTotalDay, tsTotalNight,

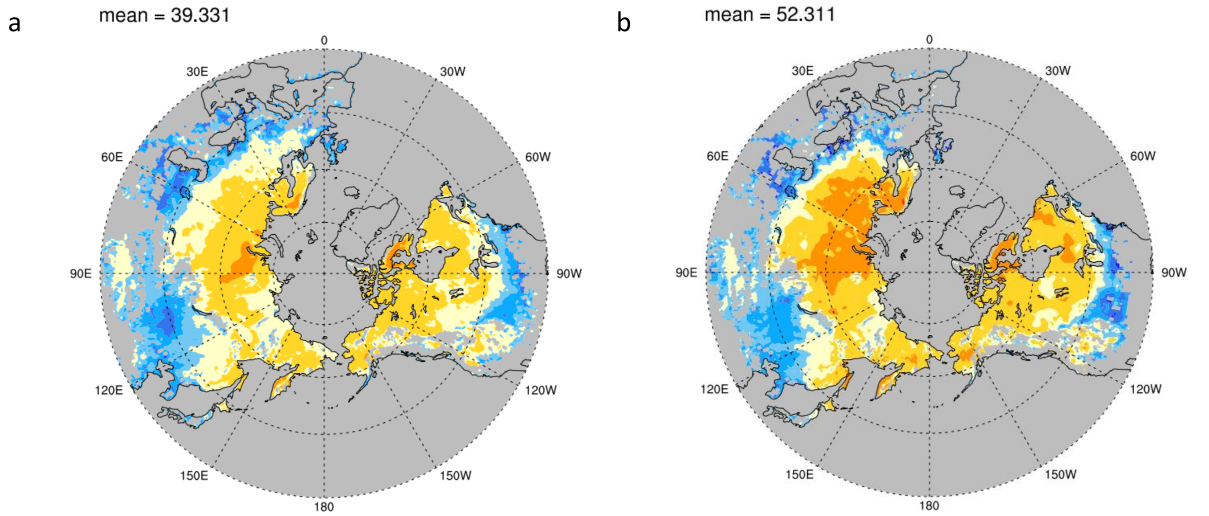
		tsUnCorErrDay, tsUnCorErrNight, tsVarDay, tsVarNight (monthly)
PERMAFROST	v3.0	alt, gtd, pfr (yearly)
SEA SURFACE TEMPERATURE	v3.0 L4 analysis	tos, tosStderr (daily)
		tos (monthly)
SNOW	v2.0 (AVHRR) MERGED	snc, snw (daily)
SOIL MOISTURE	v8.1	sm, smStderr (daily)
		sm (monthly)
WATER VAPOUR	v3.1 (CM SAF COMBI V001)	prw (daily and monthly)

1.1 Datasets newly implemented

The following two subsections provide a brief summary of the CCI datasets SNOW and PERMAFROST newly implemented into ESMValTool.

1.1.1 SNOW

Version 2.0 of the AVHRR MERGED dataset has been implemented into ESMValTool. Variables implemented include daily values of the surface snow area fraction in % (snc) and surface snow amount in kg m^{-2} (snw). As an example, Figure 1 shows the 5-year average (2000-2004) seasonal mean surface snow amount (snow water equivalent) for December-January-February (DJF), March-April-May (MAM), and September-October-November (SON) in the Northern Hemisphere (30°-90°N) in a polar stereographic projection. No data are available for June-July-August (JJA).



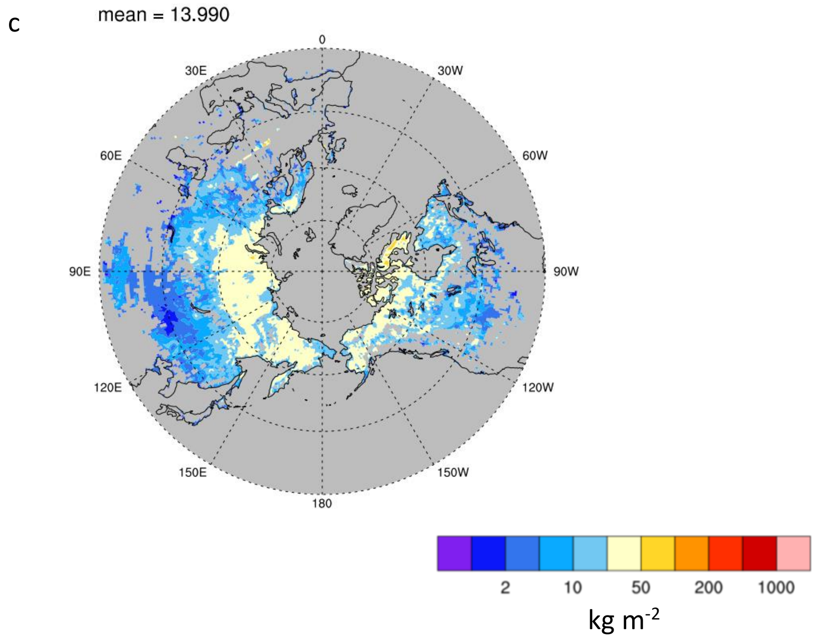


Figure 1 5-year average (2000-2004) seasonal mean surface snow amount (snow water equivalent, kg m^{-2}) in the Northern Hemisphere (30°-90°N) from ESA CCI SNOW for (a) December-January-February (DJF), (b) March-April-May (MAM), (c) September-October-November (SON). There are no data available in June-July-August (JJA).

1.1.2 PERMAFROST

PERMAFROST data v3.0 have been implemented into ESMValTool. This includes annual mean values of the permafrost extent (pfr), active layer thickness (alt) and permafrost ground temperature (gtd). For comparison with global climate model results, the method to derive the permafrost extent from CMIP6 model data by *Burke et al.* [2020] has been implemented into ESMValTool. This method assumes permafrost if

(1) the soil temperature in the deepest level is < 0°C,

(2) for at least 24 consecutive months, and

(3) the ice-covered part of grid cell is excluded.

Figure 2 shows as an example a comparison of the 5-year average permafrost extent from ESA CCI PERMAFROST over the Northern Hemisphere and results from the CMIP6 historical simulation of the MPI-ESM model over ice free regions.

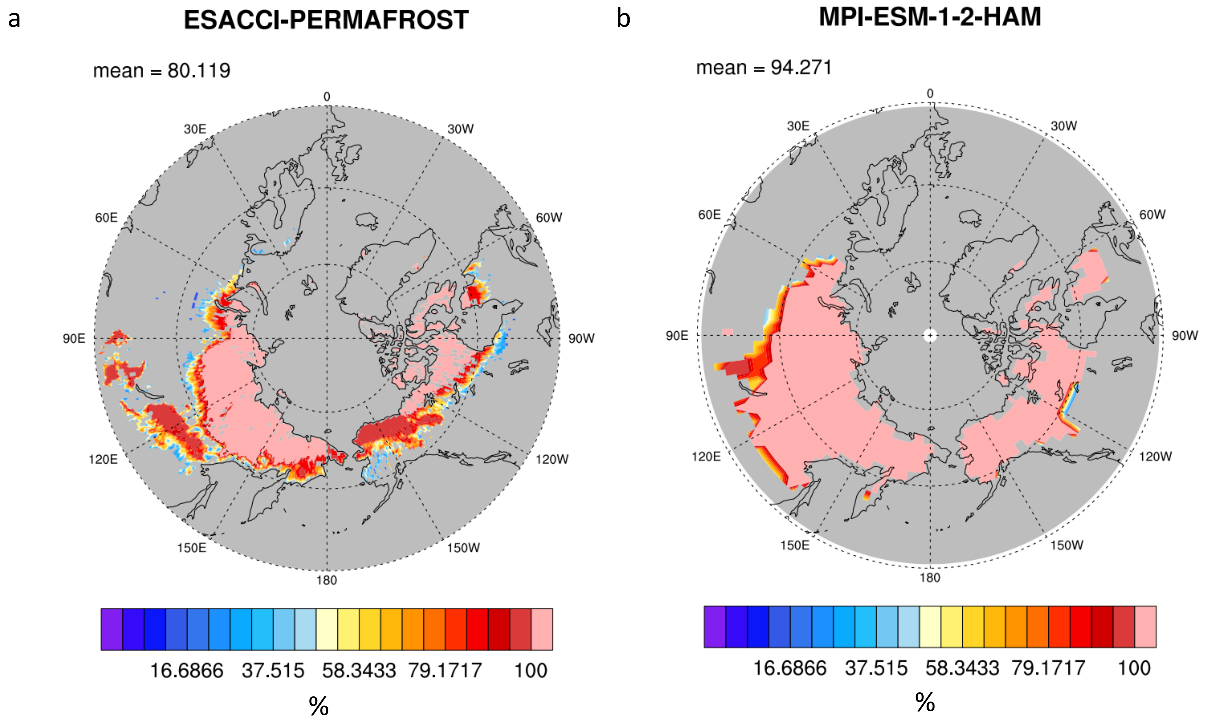


Figure 2 5-year average (2000-2004) permafrost extent (%) over ice-free regions in the Northern Hemisphere from (a) ESA CCI PERMAFROST compared with (b) the CMIP6 model MPI-ESM-1-2-HAM (historical simulation).

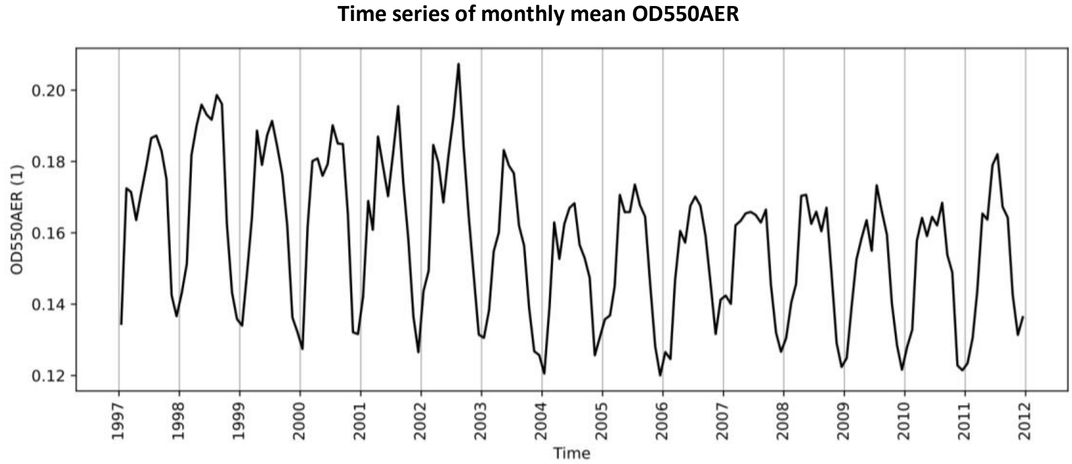
1.2 Datasets extended and updated

The following subsections briefly describe the three CCI datasets AEROSOL, SST and WATER VAPOUR that have been updated and for which now also daily values have been implemented in addition to the monthly means already implemented into ESMValTool.

1.2.1 AEROSOL

The AEROSOL data in ESMValTool has been updated to SU_(A)ATSR v4.3. In addition to the monthly mean values, also daily values have been newly implemented. Variables available include AOD at 550 and 870 nm (od550aer, od870aer), fine mode AOD at 550 nm (od550lt1aer) and absorbing AOD at 550 nm (abs550aer) as well .as the uncertainty estimates of AOD at 550 and 870 nm. Figure 3 shows time series from 1997 through 2011 of the global average AOD at 550 nm calculated from (a) monthly means and (b) daily values of ESA CCI AEROSOL.

a



b

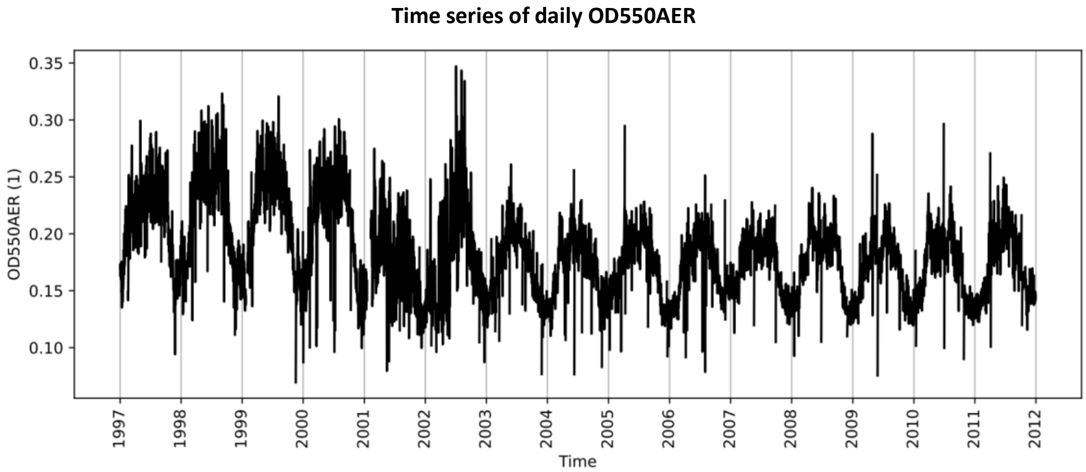


Figure 3 Time series of (a) monthly mean and (b) daily aerosol optical depth (AOD) at 550 nm averaged over the whole globe from SU_(A)ATSR v4.3.

1.2.2 SEA SURFACE TEMPERATURE

SST data have been updated to v3.0 (L4 analysis) and now also include daily values in addition to monthly means. As an example, Figure 4a shows a time series of the monthly mean anomalies in global average sea surface temperature from 1980 through 2021. As a reference period, the 30-year time period 1980-2009 has been used. Figure 4b shows the daily global mean sea surface temperature for each year in the time period 1980 through 2021 from ESA CCI-SEA SURFACE TEMPERATURE.

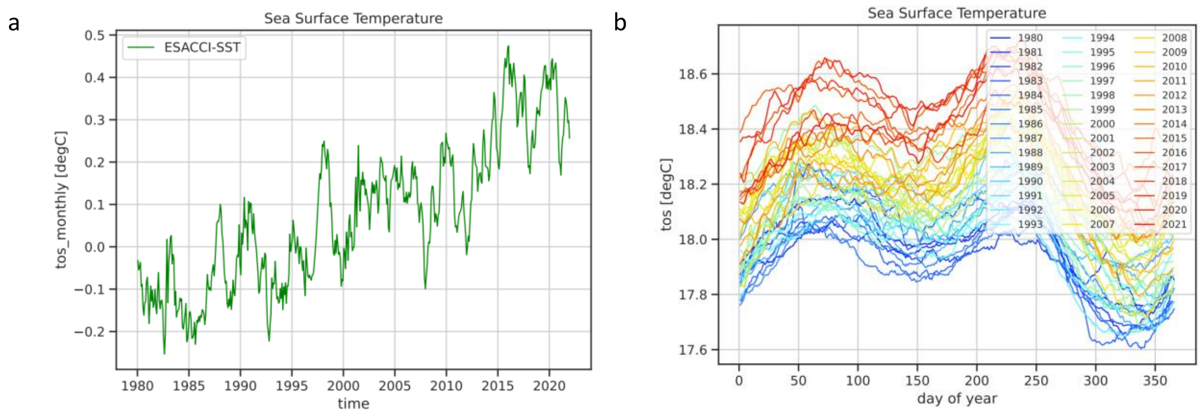


Figure 4 (a) Time series from 1980 to 2021 of global average monthly anomalies in sea surface temperature (reference period 1980-2009) from ESA CCI-SEA SURFACE TEMPERATURE. (b) Daily global mean sea surface temperature for each year in the time period 1980 to 2021 from ESA CCI SEA SURFACE TEMPERATURE.

1.2.3 WATER VAPOUR

ESA CCI WATER VAPOUR data already implemented in ESMValTool have been updated to version 3.1 (CM SAF COMBI V001). In addition to the monthly means of the water vapor path (prw), also daily values have been implemented. As an example, Figure 5 show the geographical distribution of the multi-year (2003-2018) annual average water vapor path from ESA CCI WATER VAPOUR.

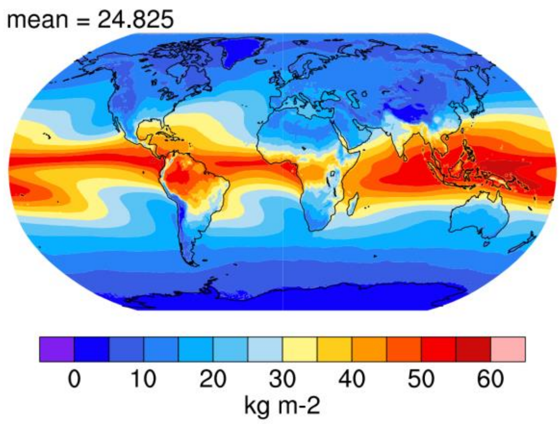


Figure 5 Multi-year (2003-2017) annual mean of the water vapor path (kg m^{-2}) from ESA CCI WATER VAPOUR.

1.3 Datasets updated

The following subsections describe the CCI datasets AEROSOL, LAND SURFACE TEMPERATURE and SOIL MOISTURE that were already implemented in ESMValTool but have now been updated to the recent dataset versions.

1.3.1 LAND SURFACE TEMPERATURE

LST data in ESMValTool have been updated to version 3.0. In addition to the daytime (tsDay) and night-time (tsNight) values of the land surface temperature, also the daytime and night-time uncertainty estimates have been implemented including the uncertainties from large-scale systematic errors (tsLSSysErr), from locally correlated errors on atmospheric (tsLocalAtmErr) and surface (tsLocalSrfErr) scales, from uncorrelated errors (tsUnCorErr), and the total uncertainty of the LST value in the original gridbox (tsTotal) and the auxiliary variable land cover class as defined by LST_cci (tsLC) (see also Table 3). These are used to explore possibilities to propagate the uncertainty estimates to the spatial and temporal scales used by climate models, which is needed in order to be able to take advantage of the uncertainty information provided with the CCI datasets for model evaluation. This is described in detail in Section 2. As an example, Figure 6 shows the multi-year (2003-2018) seasonal means of daytime (a, b) and night-time (c, d) land surface temperature from ESA CCI LAND SURFACE TEMPERATURE averaged over the months December, January, February (DJF) and June, July, August (JJA).

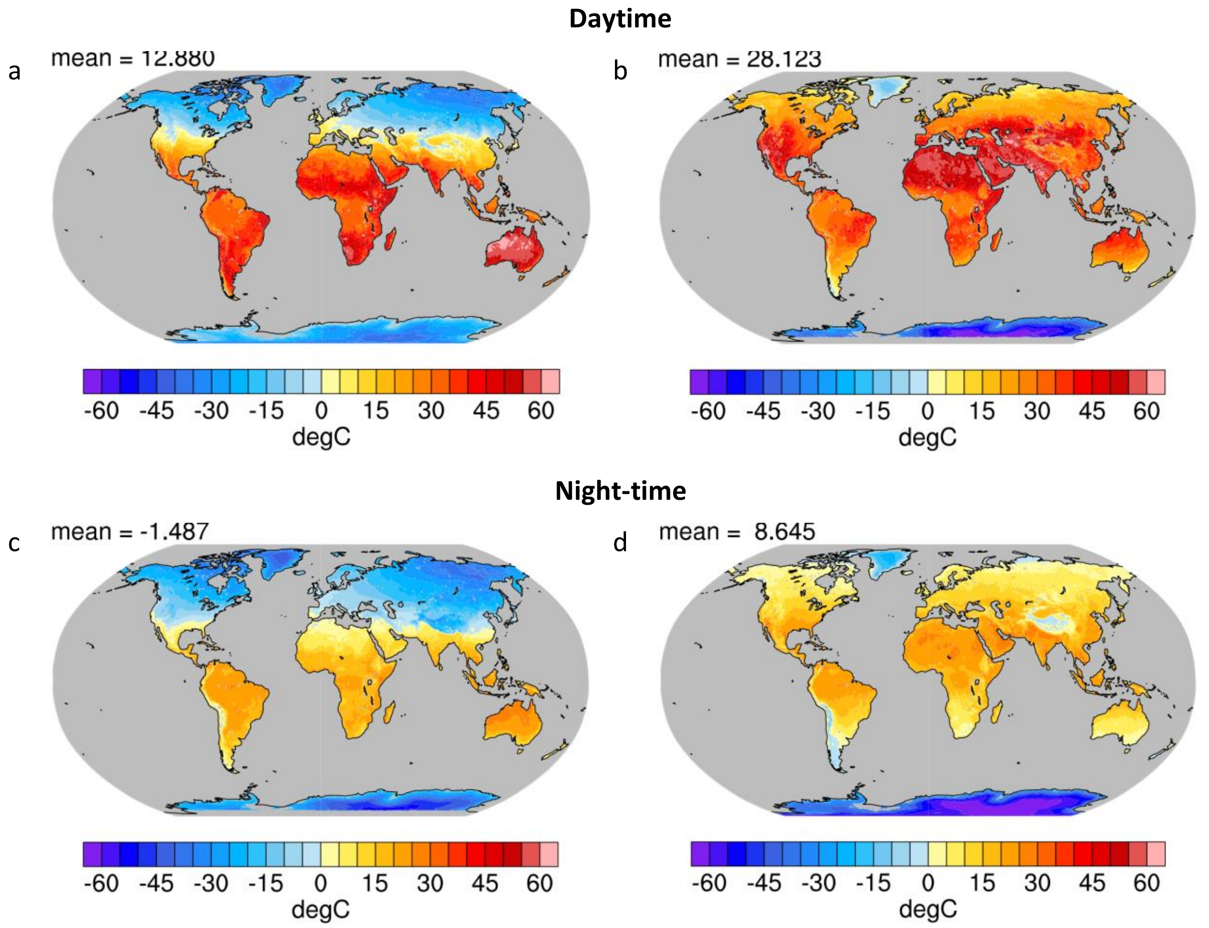


Figure 6 Multi-year (2003-2018) seasonal average land surface temperature (°C) from ESA CCI LAND SURFACE TEMPERATURE regridded to 0.5°×0.5°. (a) Daytime average for December-January-February, (b) daytime average for June-July-August, (c) night-time average for December-January-February, and (d) night-time average for June-July-August.

1.3.2 SOIL MOISTURE

The SOIL-MOISTURE dataset implemented in ESMValTool has been updated to v8.1. As an example, Figure 7 shows, the densities of the spatio-temporal distributions of monthly mean soil moisture from ESA CCI SOILMOISTURE in the period 1979-2022 for six IPCC AR6 regions. The regions shown include northern Europe, West & Central Europe, Mediterranean Europe, Central Africa, eastern North America and South Asia.

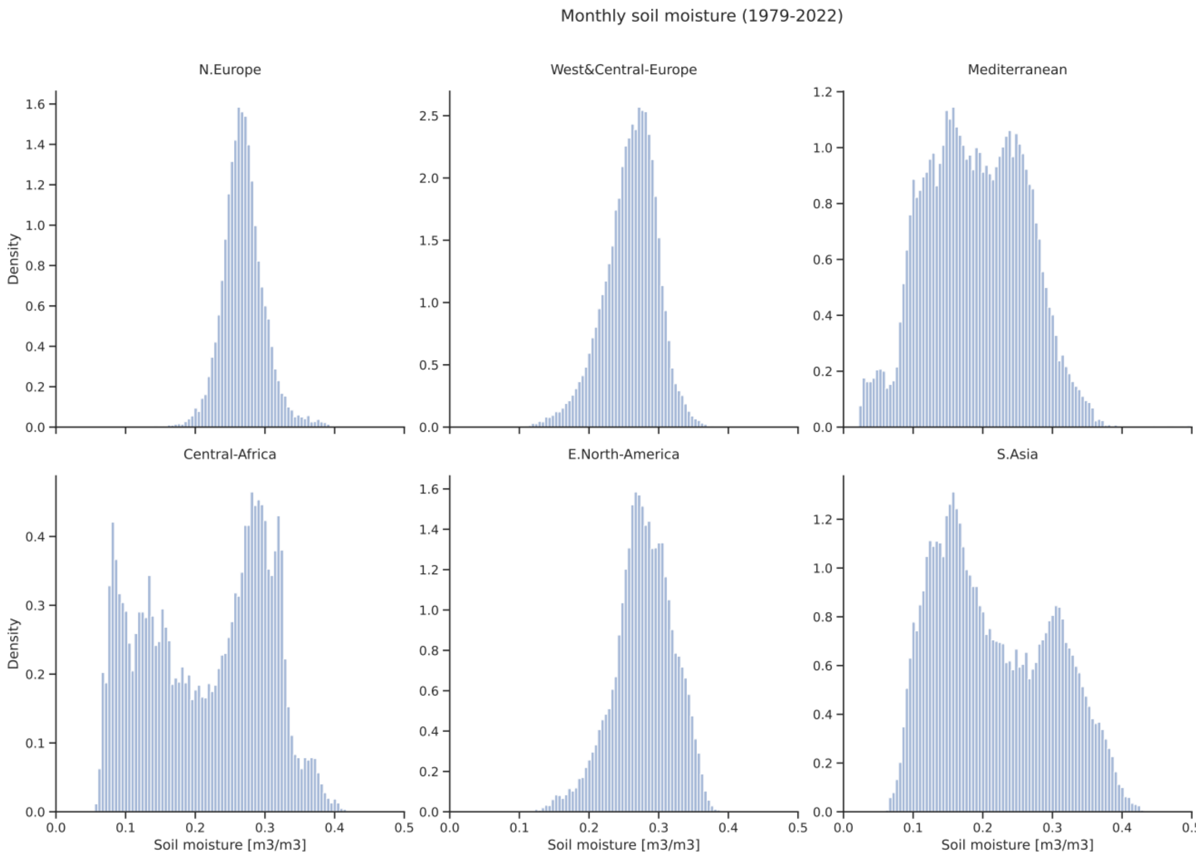


Figure 7 Spatiotemporal distribution of monthly mean soil moisture from ESA CCI SOILMOISTURE in the period 1979-2022 for six IPCC AR6 regions (from left to right) northern Europe, West & Central Europe, Mediterranean Europe, Central Africa, eastern North America and South Asia. Each month in each grid cell in the corresponding regions is considered with equal weight.

2 Implementation of uncertainty estimates into ESMValTool (WP4.2)

2.1 Introduction

All measurements have an uncertainty associated with them, which can be attributed to a variety of different error sources. For example, from the process of taking a measurement, instrument noise, satellite drift, and calibration methods. Furthermore, satellite retrievals that are used to calculate a desired geophysical quantity, e.g. temperature, have uncertainties associated with the physical constraints they operate within including the assumptions made and the use of auxiliary information, which also has an associated uncertainty. Uncertainties from these difference sources of error combine to form an uncertainty budget that can be propagated through the measurement and retrieval process, to the higher-level satellite products that are used in climate and other sciences.

The existence of uncertainty is not a bad thing. When it is expressed in accurate and accessible forms, it aids the use of measurements providing insight on signal stability, measurement processes, and evaluation of models.

For the climate model community to use the uncertainty information provided by contemporary satellite products, the format and description, and tools to propagate the uncertainty to model length scales and variables need to be available. This is important because different sources of uncertainty propagate through the processing chain in different ways. These differences need to be clearly explained and the results in an accessible format for them to be useful in model evaluation.

The expression of uncertainty in observations is now being presented and explained in the literature, bridging the metrology and climate communities [*Mittaz et al.*, 2019]. Some physical quantities, such as sea surface temperature (SST) have well established methodologies and expressions of uncertainties [*Bulgin et al.*, 2016a; *Bulgin et al.*, 2016b; *Merchant et al.*, 2019]. Furthermore, land surface temperature (LST) has well-documented processing chains for uncertainty information [*Ghent et al.*, 2019]. With this information available there is a desire to exploit it in climate model evaluation, and the tools to do this need to be developed; this work details an example of such development.

The aim of this work is to demonstrate how ESMValTool can be used to propagate uncertainty information given in satellite observation products to evaluate ESM outputs, focusing on LST. Each step is shown both in the process of accessing this information, and the algorithms used to achieve this.

2.2 Background / Method

Observation and model outputs are regularly averaged. However, the techniques for correctly propagating their uncertainties when calculating these averages are not well known or used within the user community. This averaging of multiple uncertainties from different sources is known as *uncertainty propagation*. The general form of uncertainty propagation is given in the Guide to the Expression of Uncertainty in Measurement [*GUM*, 2008]. Components in the uncertainty budget can be correlated with neighbouring measurements (in both space and time), or not, and can have different length scales over which the correlation persists. Correlation length scales change the form of the calculation when propagating the uncertainty values.

Satellite observation data pass through several stages in a processing chain before they are suitable for evaluation against model outputs. These stages give rise to different *levels* of satellite product, see Table 2. For climate model evaluation, data at level 3 is one of the most useful because these data are the geophysical quantities, such as LST, on a regular space and time grid. Level 3 products have associated uncertainty values, calculated from the uncertainty budget propagated through the processing chain from the raw observations. This work looks at propagating the level 3 uncertainties to give coarser spatial averages of the values.

Table 2 The different levels of satellite products, adapted from *Mittaz et al.* [2019].

Level	Indicative content	Example application
Level 0	Raw telemetry such as digital counts and instrument data	Input into Level 1 data.
Level 1	Calibrated radiances with location, view geometry and time information.	The input into geophysical retrieval algorithms.
Level 2	Geophysical variable estimates presented on the satellite swath/image grid.	Evaluation of the observation, process studies.
Level 3	Geophysical data on a regular space and time grid.	Model analysis.
Level 4	Complete in space and time, can be a synthesis of multiple products with gaps filled, often with some additional modelling to gap fill or retrieve the quantity.	Model analysis and prescribed fields for simulations.

The Climate Change Initiative (CCI) LST data (LST_cci) include level 3 LST products based on data from several satellite-based sensors, both infrared and microwave, on several platforms, both low earth orbit and geostationary [*Ghent et al.*, 2021]. LST_cci provides LST datasets both from individual sensors and using combinations of data from multiple sensors. All data products are presented in the same file format and with the same metadata standards for all sensors. This allows the user to easily switch between sensors and means that only one product is needed for this demonstration; the others should work with little modification of the workflow. The spatial resolution of LST_cci version 3 (v3) data is 0.01° latitude and longitude, with daily and monthly average products available. Here, the monthly averaged product is used. The low earth orbit platforms have a daytime and night-time overpass version of the data. These will be kept separate for the purposes of this assessment. The focus in this demonstration is using the LST from the Moderate Resolution Imaging Spectroradiometer (MODIS) on Aqua, an infrared (IR) sensor-based temperature product.

Using ESMValTool requires several steps:

Firstly, the observation data need to be prepared in the correct format - the Climate Model Output Rewriter (CMOR) format. Data were converted to CMOR format using the CMORiser functionality within ESMValTool. The LST_cci data was CMORised to ensure the correct format, make the uncertainty component variables available, and to make some basic checks on the file structure. This process is detailed in Section 2.3.1.

Secondly, a recipe and diagnostic were used to select and preprocess data, and then perform a scientific analysis. The recipe selects observation and model data and performs any required preprocessing of these data. ESMValTool has several built-in data preprocessors, including selecting regions, applying land/sea masks, and averaging in time and space. Ideally, propagating uncertainties across selected regions would be included in this list. The work presented here shows the first steps to achieving this. There are several steps required to make such preprocessors for uncertainty propagation and these are discussed below.

Once the required data has been prepared by the recipe, it is passed to one or more diagnostics. These perform analysis and calculations with the data, returning statistics and plots to evaluate climate models. There is a large built-in set of diagnostics already within ESMValTool, covering a diverse scientific focus including the ocean, atmosphere, and land domains. This work develops a diagnostic for propagating uncertainties across a region and compares the observed LST and its uncertainty to CMIP6 model LST results.

2.3 Data and CMORisers

As noted above, ESMValTool requires input data to be in CMOR format. This defines a standard for the dimension definitions and data presentation in the files. It is preferable for ESMValTool to have individual files for each input variable. The CMORiser and the script written in python perform checks on the data presentation and creates files that meet the required specification for ESMValTool to use.

2.3.1 LST_cci CMORiser

The CMORiser for the LST_cci V3 monthly data does several things that check the data formats and create suitable files for ESMValTool to read. It does not provide any data validation nor fill any missing data.

Table 3 The variables used from the LST_cci data. The names of both the daytime and night-time overpass versions of the variables are given, along with the var_name that gets assigned and a short description.

Day Variable Name	Night Variable Name	var_name	Description
tsDay	tsNight	ts	LST
tsLSSysErrDay	tsLSSysErrNight	lst_unc_sys	Uncertainty from large-scale systematic errors
tsLocalAtmErrDay	tsLocalAtmErrNight	lst_unc_loc_atm	Uncertainty from locally correlated errors on atmospheric scales
tsLocalSrfErrDay	tsLocalSrfErrNight	lst_unc_loc_sfc	Uncertainty from locally correlated errors on surface scales
tsUnCorErrDay	tsUnCorErrNight	lst_unc_ran	Uncertainty from uncorrelated errors
tsLCDay	tsLCNight	lcc	Land cover class as defined by LST_cci
tsTotalDay	tsTotalNight	lst_uncertainty	Total uncertainty of the LST value in the original gridbox

2.3.2 Create a Single File for each Variable, Month and Day/Night

Each variable in Table 3 is treated separately by the CMORiser and returned in its own file. This is done for each month individually, and the daytime and night-time satellite overpass versions of the data are also considered separately.

The following things are performed to create each file:

Use a common time coordinate definition. This gives the time for the data as hours since 1/1/1970 to midnight on the first of the month.

Surface temperature is used as the LST's variable name. This makes the LST variable name consistent with CMIP6's surface temperature variable.

Add day/night to variable names. Although daytime and night-time files are treated separately, adding day and night to the variable names means the overpass time can be easily identified when working with the CMORised data.

Land Cover Class made a float. The land cover class is given in the original LST_cci files as an integer however this appears to not be properly read by ESMValTool. It is not known if this is an issue with the LST_cci data files, or ESMValTool. Saving the CMORised files with this variable as a float solves this problem when the data is loaded by ESMValTool. It does mean that a simple equality test cannot be done in the diagnostic, but this is detailed below. It also increases the file size of the output.

Longitude dimension check. In version 1 of the LST_cci data, the latitude coordinate was named longitude. This check and correction have been left in place to allow older data to be used if necessary.

2.4 Mathematical Background

When creating a new data product, in this case coarsening the spatial resolution of the input data, the propagation of uncertainty information must be handled correctly. The Law of Propagation of Uncertainty [*GUM*, 2008] gives the full functional form for propagating uncertainty, but under some circumstances (either fully correlated uncertainties or completely independent uncertainties) the maths simplifies. The full details of the end-to-end uncertainty budget, including propagation from the initial measurement through to level 3 and level 4 satellite data products can be found in general in *Mittaz et al.* [2019], and for the LST_cci data in *Bulgin et al.* [2023]. The method for propagating the uncertainty when averaging the data over a larger area, as defined by the ESMValTool recipe in this work is detailed here. The necessary details of how these are algorithmically implemented are given in Section 2.5.

The following notation and definitions are used throughout.

- $\boldsymbol{z}$ is the vector of LST values in the region, one entry for each gridbox. LST could generalised to another physical quantity. Although this work propagates uncertainties over a two-dimensional region, this spatial structure is not required in the calculations, therefore $\boldsymbol{z}$ can be considered as a one-dimensional vector.
- $\langle \boldsymbol{z} \rangle$ is the average of all the values of the vector $\boldsymbol{z}$.
- $u(\langle \boldsymbol{z} \rangle)$ is the uncertainty of the averaged vector $\boldsymbol{z}$**.** A subscript will denote which uncertainty component is being referred to.
- n is the number of gridboxes of data in the region in which the uncertainties are being propagated across.

The descriptions of uncertainty components and methodology is taken from *Bulgin et al.* [2023] and *Ghent and Bulgin* [2023]. With the Level 3 LST value, which is used as the input to this recipe, there are four components of uncertainty given, as well as a total gridbox uncertainty value. This total uncertainty is the sum of the individual components, in quadrature. The four components are the random, systematic, surface and atmospheric uncertainties. These are provided independently of one another in LST product because they have different correlation length scales, and therefore the way in which they are propagated into new, derived products (such as the coarser grid LST defined here) differs. The random component is fully independent, uncorrelated with neighbouring observations. The systematic component is fully dependent, correlated across all observations, and the surface and atmospheric components are locally correlated with defined correlation length scales. The definitions and mathematical method for propagating the uncertainties and averaging the LST are given below.

2.4.1 LST

The average LST value is used to represent the whole area. The arithmetic mean of all suitable gridboxes is used. In this case there are *n* gridboxes of data in the region,

$$LST_{region} = \frac{1}{n}\sum_{i=1}^{n} LST_i$$

2.4.2 Random Uncertainty

Random uncertainty arises from effects such as instrument noise and is characterised by it affecting two separate measurements differently. They are uncorrelated on all temporal and spatial length scales. Uncorrelated uncertainties scale by a factor of $\frac{1}{\sqrt{n}}$ when propagated. This gives:

$$u(\langle \boldsymbol{z} \rangle)_{random} = \frac{1}{\sqrt{n}}\sum_{n} u(\boldsymbol{z})_{random}{}^{2}$$

When calculating the average LST over a given region, the perfect calculation of the average would require LST values for all gridboxes within the area over which the average is calculated. However, it is frequently the case that some LST values are missing as LST cannot be measured using infrared wavelengths when there is cloud in the field of view of the satellite. Consequently, there is a sampling uncertainty introduced in the calculation of the average LST, that arises due to these missing data. This sampling uncertainty needs to be quantified and then added to the uncertainty budget for the new, coarsened LST product. The effect of the sampling methods has been modelled for SST by *Bulgin et al.* [2016b]. The method used by the LST_cci product is estimates the sampling uncertainty using:

$$u(\langle \boldsymbol{z} \rangle)_{sampling} = \frac{n_{fill}\sigma_{\boldsymbol{z}}^2}{n-1}$$

where, n_{fill} is the number of unavailable gridboxes, and $\sigma_{\boldsymbol{z}}^2$ is the variance of the LST values in $\boldsymbol{z}$.

Sampling uncertainty is uncorrelated between observations and is therefore grouped with the propagated random uncertainty from the input data to give the total value of the random uncertainty component for the output product. These are combined in quadrature:

$$u(\langle \boldsymbol{z} \rangle)_{random} = \sqrt{u(\langle \boldsymbol{z} \rangle)_{random}{}^2 + u(\langle \boldsymbol{z} \rangle)_{sampling}{}^2}$$

2.4.3 Systematic Uncertainty

The systematic uncertainty component arises from effects that are common to all measurements. Such effects can include calibration uncertainties that are common to the instrument and common geolocation offsets. The LST_cci product currently provides a single value across every gridbox, but it is best practice to propagate this assuming different values are possible, which is the arithmetic mean of the values:

$$u(\langle \boldsymbol{z} \rangle)_{systematic} = \frac{1}{n}\sum_{n} u(\boldsymbol{z})_{systematic}$$

2.4.4 Locally Correlated Atmospheric Uncertainty

Two components are given for uncertainties that are locally systematic, an atmospheric and surface component. They need to be propagated separately on account of the assumed correlation length scales. LST_cci data have a correlation length scale of 5 km is for the atmosphere. This means that for grid resolutions <= 0.05°, the associated uncertainty should be propagated as fully correlated. Beyond this space scale this uncertainty component is assumed to be uncorrelated. Presently the data coarsening is done in two stages: first propagating to 0.05° (5 km resolution) and then coarsening further to the target resolution. This approach has been chosen on account of the methods used to propagate the surface component (see section 2.4.5) and the fact that for products at a resolution >= 0.05°, biome information is not retained in the product. Propagating in a two-stage process is not mathematically ideal but is the pragmatic solution given the available LST data products. A full description of the two-step method is given in section 2.4.5; the atmosphere case is mathematically the same as just one surface biome.

The first step is to propagate the 0.01° data to a 0.05° grid. This uses the 5 km correlation length and each new 0.05° gridbox is calculated assuming full correlation:

$$u(\langle z\rangle)_{local_{atmospheric}} = \frac{1}{n}\sqrt{\sum_n u(z)_{local_{atmospheric}}{}^2}$$

where *n* is the number of 0.01° gridboxes used to make the new 0.05°gridbox. There are in general *m* 0.05°gridboxes created.

The second step is to propagate the *m* 0.05° gridboxes' values to the whole region. This is done assuming that each 0.05° gridbox is uncorrelated to each other using the equation:

$$u(\langle z\rangle)_{local_{atmospheric}} = \frac{1}{\sqrt{m}}\sqrt{\sum_m u(z)_{local_{atmospheric}}{}^2}$$

2.4.5 Locally Correlated Surface Uncertainty

Similarly to the atmospheric component, a two-step process is used to propagate the locally correlated surface component. The spatial correlation length scale is 5 km, meaning that gridboxes within this distance are assumed to have an influence on that gridbox's LST measurement. The first step is to create a 0.05° spatial resolution grid (5 km is approximately 0.05°). This is a re-gridding step, not an attempt to look at every possible 5 km radius around every original gridbox. Only original 0.01° gridboxes that are within each new gridbox are used to create the new value. For each new gridbox, the number of unique land cover classes is found – this information is a variable contained in the LST_cci product. For each land cover class present, the locally correlated surface uncertainty values are found, and these are propagated assuming that they are fully systematic. This gives a single uncertainty value for each land cover class found within the 0.05° gridbox domain. This is then repeated for each new 0.05° gridbox. This step assumes there is a correlation between the LST measurements of every gridbox of the same land cover class within 5 km. The final uncertainty for this component is then calculated by taking each of the land-cover specific uncertainties and propagating these assuming that these uncertainties (between biomes) are independent.

The second step is to take these uncertainties in the 0.05° product and then propagate them assuming that they are uncorrelated into subsequent coarser products.

This method treats each 0.05° grid cell independently and neglects correlations between 0.05° gridboxes. For example, there could be two adjacent grassland 0.01° gridboxes, but they end up in different 0.05° gridboxes. The uncertainties between the two are likely to be correlated, but in using this two-step approach, they are treated as uncorrelated. This is a limitation of the method and can only be overcome by a more complicated approach that fully defines an off-diagonal correlation matrix and propagates directly from 0.01° to the target resolution in a single step.

2.4.5.1 Worked Example

A 5×5 set of gridboxes have been identified to contribute to a gridbox at the new re-gridded resolution. They have a land cover class associated to them, Figure 8; in this case there are three: Tree, Grass, and Shrub. They have locally correlated surface uncertainty values given in Figure 8. This gives us three mean values to use (given to one decimal place):

$$\langle Tree\rangle = \sqrt{\frac{0.1^2+0.6^2+0.2^2+0.9^2+0.3^2+0.7^2+0.2^2+0.5^2+0.4^2+0.8^2}{10}} = 0.5$$

$$\langle Grass \rangle = \sqrt{\frac{0.7^2 + 0.1^2 + 0.7^2 + 0.3^2 + 0.5^2 + 0.1^2 + 0.6^2 + 0.6^2 + 0.1^2}{9}} = 0.5$$

$$\langle Shrub \rangle = \sqrt{\frac{0.1^2 + 0.8^2 + 0.6^2 + 0.5^2 + 0.7^2 + 0.9^2}{6}} = 0.7$$

These give us a value for the complete set as:

$$u(\langle \boldsymbol{z} \rangle)_{0.05^\circ\,gridbox} = \frac{1}{\sqrt{3}} \sqrt{\frac{0.5^2 + 0.5^2 + 0.7^2}{3}} = 0.3$$

This gives 0.3 K as the uncertainty from locally correlated surface effects for whole 5×5 grid.

Tree	Tree	Tree	Shrub	Grass
Tree	Grass	Grass	Grass	Tree
Shrub	Shrub	Tree	Tree	Tree
Grass	Shrub	Grass	Grass	Shrub
Tree	Grass	Grass	Shrub	Tree

0.1	0.6	0.2	0.1	0.7
0.9	0.1	0.7	0.3	0.3
0.8	0.6	0.7	0.2	0.5
0.5	0.5	0.1	0.6	0.7
0.4	0.6	0.1	0.9	0.8

Figure 8 An example of creating the 0.05°×0.05° target grid from 0.01°×0.01° original data. (top) The land cover classes, coloured blue (tree), green (grass), and purple (shrub). These are indicative classes, and the full range and exact descriptions are given in the user guide to the LST_cci data [*Ghent et al.*, 2021]. (bottom) Uncertainty values coloured by their land cover class (as defined in the top figure). These land cover classes and uncertainty values are used in the worked example in the text.

2.4.6 Total Uncertainty

The total uncertainty across the region is the sum in quadrature of all the components. For LST_cci data, there are four components so

$$u(\langle \boldsymbol{z} \rangle)_{total} = \sqrt{u(\langle \boldsymbol{z} \rangle)_{random}{}^2 + u(\langle \boldsymbol{z} \rangle)_{systematic}{}^2 + u(\langle \boldsymbol{z} \rangle)_{local_{surface}}{}^2 + u(\langle \boldsymbol{z} \rangle)_{local_{atmosphere}}{}^2}$$

2.5 ESMValTool Recipe and Diagnostic

This demonstration of uncertainty propagation in ESMValTool consists of two parts: i) the recipe, and ii) the diagnostic. This demonstration uses the CMORised LST_cci data and the recipe cuts out a region for all variables. No other preprocessing is done to the LST_cci data. It is an aspiration in the future to include the uncertainty propagation functions detailed in the diagnostic as ESMValTool preprocessor functions.

The diagnostic contains five functions used to propagate each of the four uncertainty components. The two non-trial ones, random uncertainty including sampling uncertainty, and locally correlated surface effects uncertainty, are detailed here.

2.5.1 Random Uncertainty

Section 2.4.2 detailed the equations need to propagate the random uncertainty. The first step is to obtain $u(\langle \boldsymbol{z} \rangle)_{random}$ as propagated from the input product. The sampling uncertainty, $u(\langle \boldsymbol{z} \rangle)_{sampling}$, needs n_{fill} to be calculated. This is taken as the number of masked gridboxes in the LST variable. If the LST data is presented with no gridboxes masked, then there is no sampling uncertainty to calculate. The built in Iris function for calculating variance is used to obtain $\sigma_{\boldsymbol{z}}^2$. A separate function is used to calculate the sum in quadrature to arrive at the final values for $u(\langle \boldsymbol{z} \rangle)_{random}$ for the output product. Sum in quadrature is given its own function and is also used for combining all four components and propagating the given total uncertainty values.

2.5.2 Locally Correlated Surface Uncertainty

This is a multistage calculation involving two different variables, the uncertainty values themselves, and the land cover classes of the gridboxes. The first step is to define a target 0.05°×0.05° grid. This is done by taking 5×5 gridbox areas from the original 0.01°×0.01° grid. It is possible that if the target region does not exactly fit, new gridboxes in the eastern-most column, and southern most row of the domain will have less than 25 values. In this case, the sampling uncertainty calculation will represent the uncertainty introduced from not fully sampling the grid domain.

The second stage is to associate each uncertainty value to its corresponding land cover class. A list is created for each unique land cover class in the area, and the uncertainty values looped over and appended to the appropriate list. At this stage no information as what the land cover class is kept; it is not needed for the calculation. The number of land cover classes in the area is not important for the calculation and could be any number between 1 and 25 depending on the region. It is possible that different numbers of land cover classes within the area has implications for the propagated uncertainty value, but this is not explored here.

For each list of uncertainty values, the mean value is calculated. This is equivalent to the set of average values shown in the worked example in Section 2.4.5.1. The final stage for the 5×5 gridbox area is to propagate the biome specific uncertainties, assuming that these are uncorrelated. This gives the propagated locally correlated due to surface effects uncertainty value for the 0.05°×0.05° gridbox and is repeated for all gridboxes in this target grid.

2.6 Example

This section shows a set of two example plots generated by the ESMValTool diagnostic described in Section 2.5.

2.6.1 Individual Uncertainty Components

A region in France is selected (2.60°-3.00° E, 46.05°-47.45° N) to propagate the uncertainties over. No other preprocessing of the observation data is carried out. The diagnostic propagates the four uncertainty components separately and combines them to give a total uncertainty in the LST over the region. The mean LST is calculated and used as the LST of the region. The daytime and night-time overpass times from the LST_cci data are calculated and plotted separately.

Figure 9 shows the daytime values calculated by the diagnostic. The top panel shows the mean LST and gives +/- 1 times the uncertainty as a shaded region. The seasonal and annual variations can be seen. The shading shows the variation in the total uncertainty through the whole timeseries 2003-2014 (to correspond to when CMIP6 data are available).

The bottom panel shows the four individual components propagated across the region, along with their combined total uncertainty. The random and locally correlated (surface) components show the largest contribution. The systematic uncertainty is given in this data as a constant for the whole globe and timeseries. The locally correlated (surface) uncertainty is shown to have a seasonal cycle. The method used to propagate this considered the land cover classifications; changes in this classification over time, or changes in the uncertainty value due to seasonal effects such as leaf area index or soil moisture could be contributing to this.

The corresponding plots for the night-time data are given in Figure 10. The land surface class types are shown in Figure 11 to illustrate how the locally correlated surface uncertainty calculation is performed. In this case there are sixteen land cover types within the region and not every type is present in each new 0.05°×0.05° grid box.

The day and night-time overpass data have large random uncertainty 2-4 K compared to less than 1 K for the other three components.

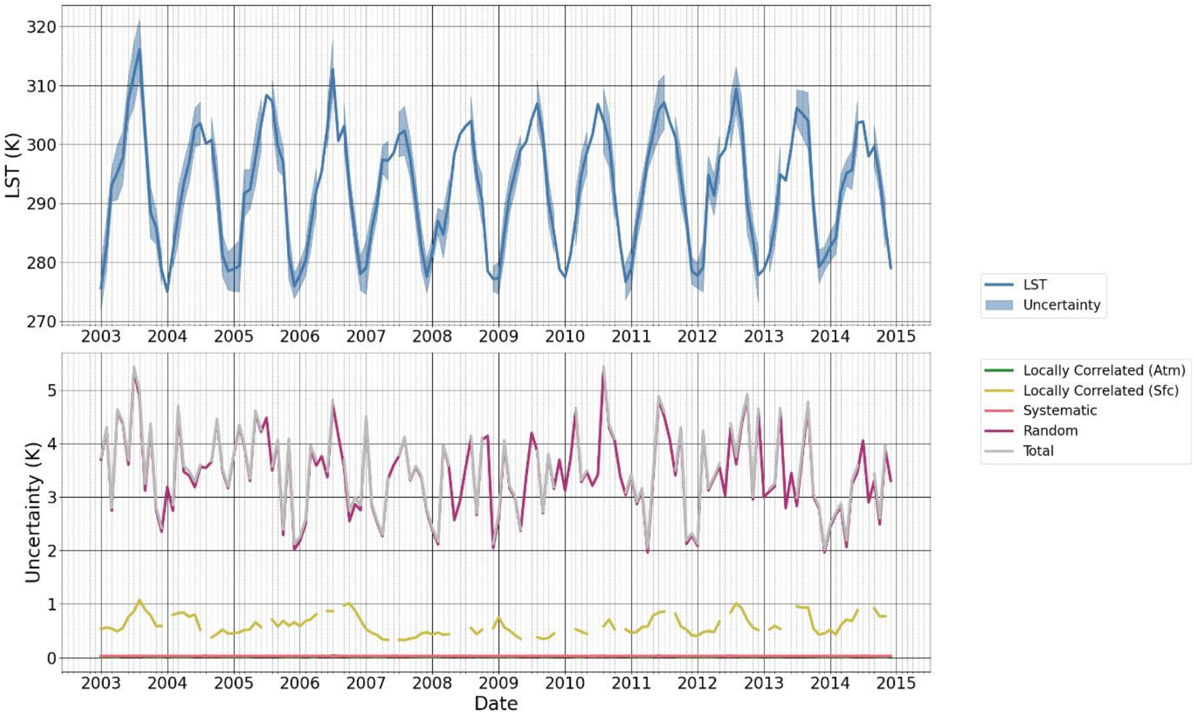


Figure 9 Timeseries plots of the daytime Land Surface Temperature (LST) and the propagated uncertainty components for the region of France defined in the text. (top) The mean LST shown with +/-1 times the uncertainty shown as a shaded region. (bottom) The four uncertainty component values propagated across the region. The total combined uncertainty is given as the grey line.

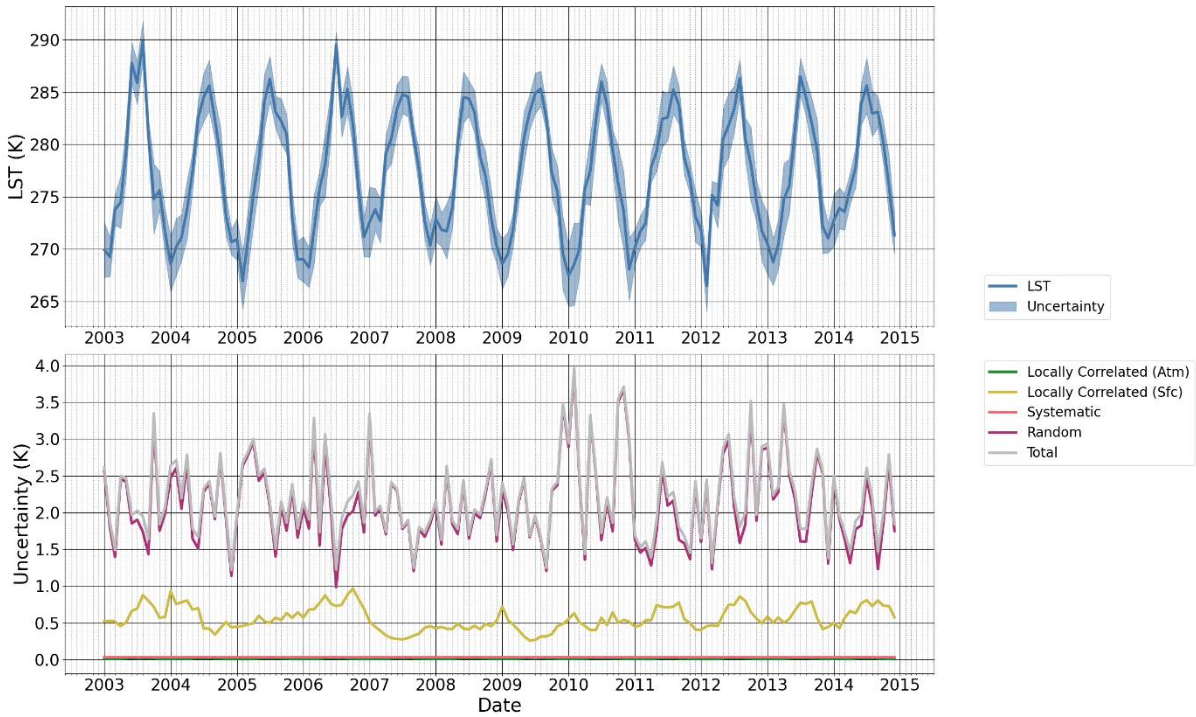


Figure 10 The same plots as in Figure 9 but for the night-time overpass data.

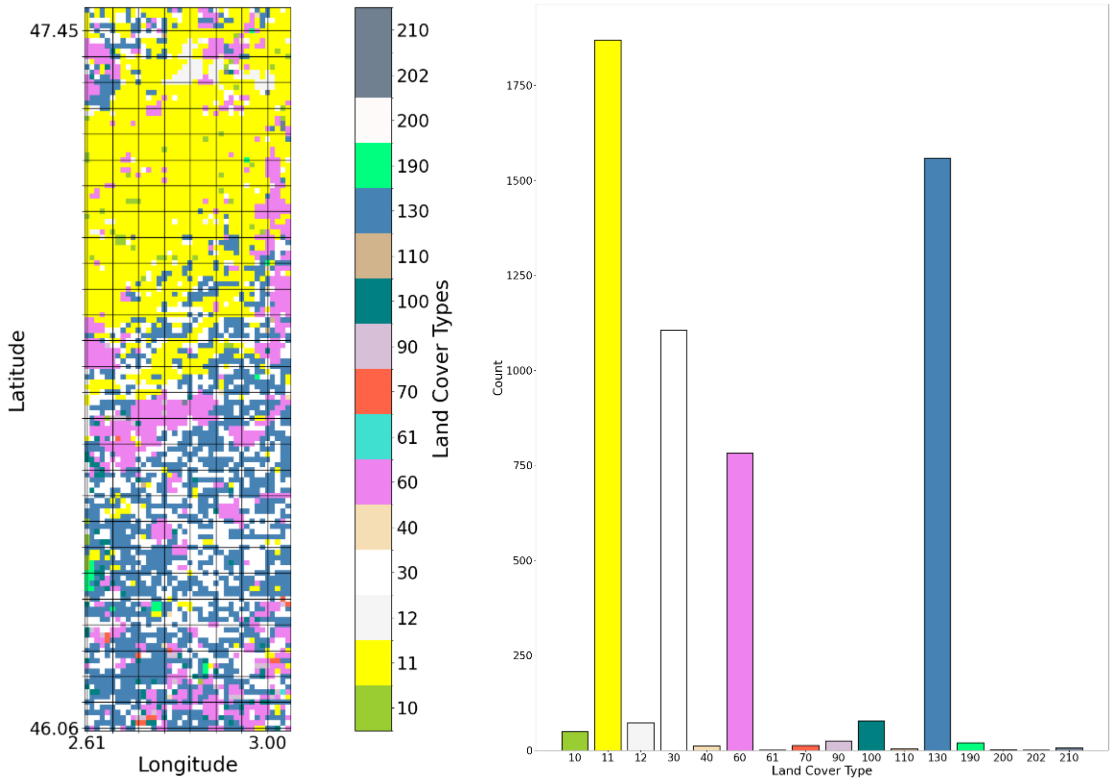


Figure 11 Illustration of land cover classes used to perform the propagation of the surface locally correlated uncertainty component, shown for January 2003. (left) A map of each gridbox coloured by land cover class with new 0.05°×0.05° gridboxes shown. (right) A chart of the counts of each individual land cover classification. The full mapping of classification numbers to their descriptions can be found in *Ghent et al.* [2021]; but the dominant types here are 11 (cropland rainfed herbaceous cover), 30 (mosaic cropland), 60 (tree broadleaved deciduous closed to open) and 130 (grassland).

2.6.2 CMIP6 Comparison

This example shows a comparison of the LST_cci with its calculated total uncertainty to the LST from a CMIP6 model. This is shown in Figure 12.

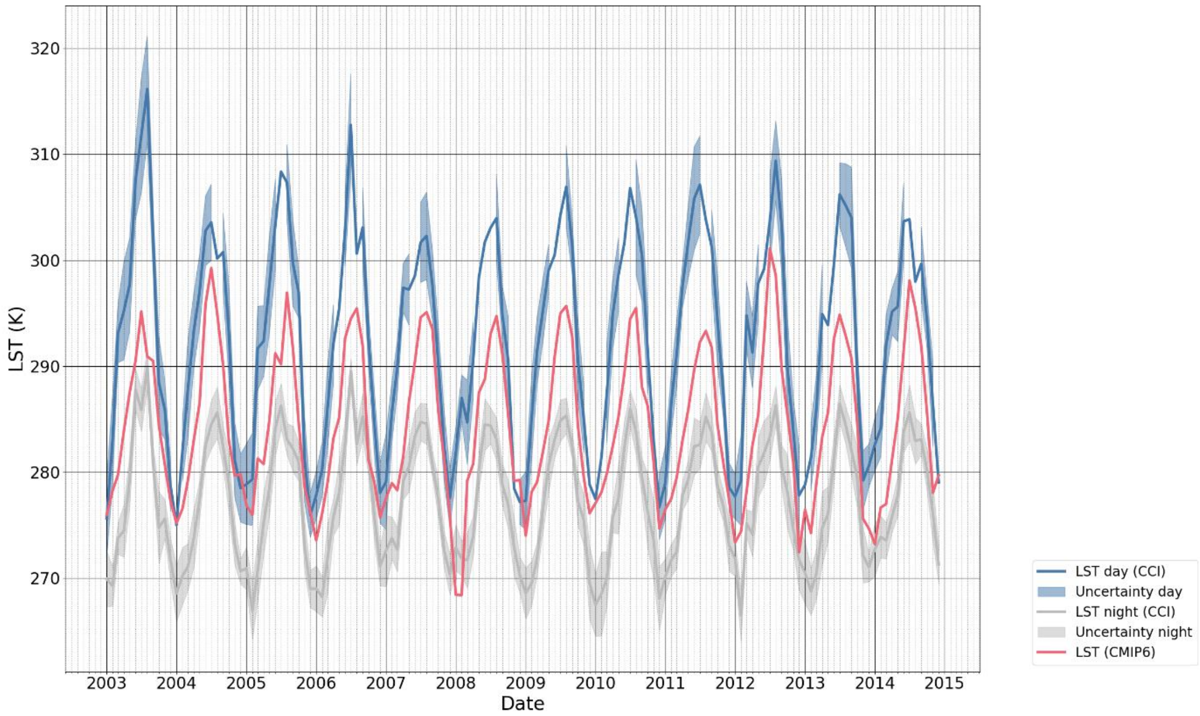


Figure 12 The LST_cci and calculated total uncertainty, and model comparison. The daytime and night-time overpasses from the LST_cci are given in blue and grey respectively, with shaded regions showing +/- 1 total uncertainty as in Figure 9. The red line is the land surface temperature from the model discussed in the text.

The CMIP6 model is a single ensemble member from United Kingdom Earth System Model (UKESM) [*Sellar et al.*, 2019]. Figure 12 shows limited overlap between the model and LST_cci values and the uncertainty range. This highlights the model is an average of all times of the day, and the satellite observations are for fixed times of the day. This is an artefact of LST being sensitive to the time-of-day of the observation and this would not be expected for observations that are slower varying.

2.7 Recommendations

The diagnostic demonstrated provides an example to other essential climate variables (ECVs) to create their own ESMValTool diagnostics that propagate uncertainties. The European Space Agency (ESA) CCI is developing 21 data records of the current 55 specified by the Global Climate Observing System (GCOS). The similarity in approaches and file formats should allow any of them to adapt the approach given here to their own variable and propagate the uncertainty information spatially. To do this, and to improve this demonstration the following recommendations are made:

Write Preprocessor functions for the uncertainty propagation. Section 2.4 showed that some types of uncertainty do not propagate in a trivial manner, e.g. the locally correlated surface effects and random uncertainty. These uncertainty variables require two input data values respectively. This is not a functionality currently available in ESMValTool.

Correlated uncertainty from local surface effects. The approach used in Section 2.4.5 is one method presented in *Bulgin et al.* [2023]. This two-step process is a pragmatic choice based on the auxiliary biome data available in the LST products. The method used could underestimate the true uncertainty if the correlation length scale is underestimated or by the way that the 0.05° grid cells are ‘artificially’ imposed for the initial propagation step but is achievable with the current LST_cci data. The LST_cci team recommend that they themselves consider the full metrological approach as part of the CCI project.

The algorithm presented in Section 2.5.2 could be optimised to give better performance when working on larger areas. This will become important if the region selected in the recipe spans multiple CMIP6 gridboxes.

Number of Land Cover Classifications. There are 42 different land cover types specified in the LST_cci data. Some of these are very similar and would be given the same plant functional type (PFT) in a climate model, for example there are 6 needle leaf tree classifications, 3 for deciduous and 3 for evergreen. When comparing the observations to a model, it is a question as to whether similar land cover types should be grouped together into similar PFTs like the model or left separate.

Non-adjacent land cover types. The first step to propagate the locally correlated due to surface effects uncertainty component is to create a new 0.05°×0.05° grid. Within each gridbox every original resolution gridbox of the same land cover type is assumed to be correlated to each other, even if they are not adjacent. Similarly, adjacent gridboxes at the original resolution can end up in different 0.05°×0.05° gridboxes after the regridding. This can prevent some correlations from being considered by the propagation calculation leading to a possible underestimate of the surface uncertainty component.

Review of methods and data. The method and data used is under review. The final version of this report will include any updates and further recommendations that arise from the review.

References

Bulgin, C. E., O. Embury, G. Corlett, and C. J. Merchant (2016a), Independent uncertainty estimates for coefficient based sea surface temperature retrieval from the Along-Track Scanning Radiometer instruments, *Remote Sens Environ*, *178*, 213-222, doi:10.1016/j.rse.2016.02.022.

Bulgin, C. E., O. Embury, and C. J. Merchant (2016b), Sampling uncertainty in gridded sea surface temperature products and Advanced Very High Resolution Radiometer (AVHRR) Global Area Coverage (GAC) data, *Remote Sens Environ*, *177*, 287-294, doi:10.1016/j.rse.2016.02.021.

Bulgin, C. E., S. Ermida, and C. Jimenez (2023), End-To-End ECV Uncertainty Budget, CCI Land Surface Temperature, *10 Feb 2023 Rep.*

Burke, E. J., Y. Zhang, and G. Krinner (2020), Evaluating permafrost physics in the Coupled Model Intercomparison Project 6 (CMIP6) models and their sensitivity to climate change, *Cryosphere*, *14*(9), 3155-3174, doi:10.5194/tc-14-3155-2020.

Ghent, D., and C. E. Bulgin (2023), Re-gridding and Sub-setting ATBD, CCI Land Surface Temperature *Rep.*

Ghent, D., S. Ermida, C. Jimenez, and E. Dodd (2021), CCI Land Surface Temperature, Product User Guide *Rep.*

Ghent, D., K. Veal, T. Trent, E. Dodd, H. Sembhi, and J. Remedios (2019), A New Approach to Defining Uncertainties for MODIS Land Surface Temperature, *Remote Sens-Basel*, *11*(9), doi:10.3390/rs11091021.

GUM, I. (2008), Guide to the Expression of Uncertainty in Measurement,(1995), with Supplement 1, Evaluation of measurement data, JCGM 101: 2008, *Organization for Standardization, Geneva, Switzerland*, *785*.

Merchant, C. J., et al. (2019), Satellite-based time-series of sea-surface temperature since 1981 for climate applications, *Scientific Data*, *6*(1), 223, doi:10.1038/s41597-019-0236-x.

Mittaz, J., C. J. Merchant, and E. R. Woolliams (2019), Applying principles of metrology to historical Earth observations from satellites, *Metrologia*, *56*(3), doi:10.1088/1681-7575/ab1705.

Sellar, A. A., et al. (2019), UKESM1: Description and Evaluation of the UK Earth System Model, *J Adv Model Earth Sy*, *11*(12), 4513-4558, doi:10.1029/2019ms001739.