Horizon 2020

H2020-EO-2014 New ideas for Earth-relevant Space Applications

EUSTACE

(Grant Agreement 640171)

EU Surface Temperature for All Corners of Earth

Deliverable D2.2

A consistent set of air temperature fields and uncertainties across all surfaces

Deliverable Title	*A consistent set of air temperature fields and uncertainties across all surfaces*
Brief Description	*As an input to the statistical infilling and as part of a sustainable system, air temperature datasets will be created based on the best available inputs, that are produced in a consistent fashion and which have ancillary information such as surface type.*
WP number	2
Lead Beneficiary	*Met Office*
Contributors	*Met Office*
Creation Date Version Number Version Date	23 June 2017 3 12 March 2019
Deliverable Due Date Actual Delivery Date	Month 30

Nature of the Deliverable		
		R - Report
		DEM – Demonstrator, Pilot, Prototype
		DEC – Dissemination, Exploitation or Communication
	X	*O - Other*
Dissemination Level/ Audience	X	*PU - Public*
		CO - Confidential, only for members of the consortium, including the Commission services

Version	Date	Modified by	Comments
1	27 June 2017	John Kennedy	Putting report into template
1.1	29 June 2017	Joel Mitchelson	Writing additional text
1.2	2 July 2017	John Kennedy	Finalising for internal review
1.3	26 July 2017	John Kennedy	Revised following internal review
1.4	1 Aug 2017	John Kennedy	Finalised for submission
2	24 Apr 2018	John Kennedy	Updating for v2 of data
3	12 Mar 2019	John Kennedy	Updating for final product

Table of Contents

1. Executive Summary

- A set of consistent air temperature fields with uncertainty estimates across land, ice and ocean areas has been created. This is one of the EUSTACE products which will be publicly released (EUSTACE Air temperature estimates from satellite) and is sometimes also referred to as SATSTACE.

- The data set meets external user requirements specified in the final product design and gathered via Work Package 4 and internal user requirements captured through regular project discussions.

- The relationships between skin and air temperature derived in Work Package 1 have been brought into the EUSTACE processing system and a consistent set of outputs has been produced.

- A reprocessing of the data set brought in a recently reprocessed version of the land surface temperature data, the ATSR2 data were added for the ocean and the product design was refined to allow better, more consistent communication of uncertainty in the ancillary files.

- This version documents final changes to the files prior to release relating to the exclusion of some air temperature estimates over land (those coming from regression model 3).

- This product will be released as EUSTACE Air temperature estimates from satellite, v1.0

2. Project Objectives

With this deliverable, the project has contributed to the achievement of the following objectives (DOA, Section B1.1):

No.	Objective	Yes	No
1	Intensively develop the hitherto immature use of Earth Observation estimates of Earth's surface **skin** temperature to enable new Climate Data Records of the surface **air** temperature Essential Climate Variable (ECV) to be created, for all locations over all surfaces of Earth (i.e. land, ocean, ice and lakes), for every day since 1850. EUSTACE will achieve this by: combining information estimated from multiple satellites with surface air temperature measurements made *in situ* and **creating complete analyses** of surface air temperature, through the application of novel statistical in-filling methods.	X	
2	Integrate these new daily surface air temperature Climate Data Records into a range of applications in Earth System Science and Climate Services and research, amongst others. EUSTACE will achieve this via the active and continuous engagement of trail-blazer users, and the provision of products through already-existing user community data portals and service mechanisms, in standard formats.	X	

3	Undertake and report detailed research into the relationships between surface skin temperature estimated from Earth Observation satellite measurements and surface air temperature observed *in situ* by conventional measurements, over all surfaces of the Earth, including the polar regions. This is likely to provide information useful for refining coupling in Earth system models.	X	
4	Create a sustainable, automated system at an appropriate level of maturity for the potential production of the products beyond the lifetime of the project. To enable this, EUSTACE will also identify Earth Observation and conventional data streams that could be used to update the surface air temperature Climate Data Records in the future, including those from Sentinel missions.	X	
5	Extensively validate the new surface air temperature Climate Data Records against independent, surface-based reference data, sourced by the project for this purpose.		X
6	Develop and report new, consistent, validated estimates of uncertainty both in already-existing Earth Observation surface skin temperature estimates and in the new surface air temperature Climate Data Records, at all locations and times across the Earth's surface.	X	
7	Develop links with related activities within Europe and beyond to help to ensure the execution of a joined-up work programme, the Copernicus Services and to enable the provision of requirements for the future surface skin temperature and surface air temperature observing system.		X
8	Other – not directly linked to one of the above objectives		X

3. Detailed Report

3.1 Summary

Here we deliver a consistent set of air temperature fields and uncertainties across all surfaces. We use the relationships that were derived in Work Package 1 (WP1, Hoyer et al. 2018, D1.5)

between skin temperature, as measured by satellites, and near-surface air temperature and apply these to satellite skin temperature retrievals to estimate air temperature across three surfaces: land, ice and oceans.

The presentation of the data is informed by the requirements for later creation of globally-complete fields and by user requirements gathered in WP4.

The data are provided in a set of NetCDF files and the format is described in Appendix A. In short, for each surface type there is one file per day for each surface type containing a field, or fields, of air temperature estimates with corresponding estimates of the total uncertainty. Data are formatted in a consistent way. A more-complete breakdown of uncertainty information is provided in an ancillary file. Temperature and uncertainty information is provided split according to correlation structure as used throughout the EUSTACE project and a consistent nomenclature has been used for all surfaces

Processing of the entire data set takes around one week on the CEMS computing platform. The current limiting factors in processing are the availability of IDL licenses needed to process the land and ocean components and the reliability of the system (Rose) used to manage job submission. The code is kept under version control in the EUSTACE code repository.

3.2 Differences from the product foreseen in the EUSTACE Description of Action

The Description of Action (DoA) suggests that ancillary information would include such things as surface type, an implication being that multiple surface types could be combined in a single file per day of data. In fact the user feedback informed us that users would prefer separate files for the different surface types, so that they may efficiently download the surfaces of interest to them.

The project plan shows that development of a complete analysis system happens concurrently with development of the consistent air temperature product. We have taken this a step further and made use of code components intended for the end-to-end system as the basis for producing this data set. This helps to prove out the system architecture and contributes to code maturity both for this deliverable and later system deliverables. Repeated use suggests typical compute times of a week for the calculation of all satellite-derived consistent air temperature fields from gridded satellite surface-temperature products; which could facilitate straightforward update if new input data becomes available.

Lake air temperatures are not included in the data files because no air temperature fields were produced over lakes that are equivalent to those produced over land, oceans and ice.

3.3 Difficulties encountered and solutions

Earlier comparison of the derived air temperatures over different surfaces, at pre-determined checkpoints in the EUSTACE project, highlighted persistent differences between the satellite-derived air temperature estimates over ice and those derived over land when relationships derived between LST and air temperature were used in ice-covered regions. It was decided that, where areas were entirely ice covered, the air temperatures derived from ice surface temperature should be used, as those relationships were derived specifically for that surface type. The land surface temperature (LST) to land surface air temperature (LSAT) relationships

were derived for a wider range of surface types and very few *in situ* data were available to constrain the LST-LSAT relationships over ice-covered areas, particularly Antarctica and Greenland.

Overlaps between air temperatures over ice and over oceans were few due to the lack at high latitudes of the ship data that are required to build the air-sea temperature relationships. Differences in the overlap between land and ocean were generally consistent with the estimated uncertainties and persistent differences in most areas were consistent with expected land-sea temperature differences. For example, the sea was warmer than the land during winter over Europe.

Comparisons of the LSAT derived from LST and in situ air temperatures showed large offsets of several degrees between the estimated air temperatures over tropical areas and in eastern Asia. Investigations showed that of the three regression models used to estimate LSAT from LST, one gave results that showed particularly large differences from the in situ data. Consequently, it was decided to exclude results from this model. This reduces the offset between the satellite and in situ estimates of air temperature over land, but at the cost of reducing the observational coverage.

3.4 Methods and Results

Details of the derivation of relationships between skin temperatures and air temperatures are given in D1.5 (Hoyer et al. (2018) "Report on the relationship between satellite surface skin temperature and surface air temperature observations"). The processing code developed in WP1 was incorporated into the EUSTACE system and correct functioning of the code was tested using sample outputs provided by WP1. In the case of the ice-surface temperature relationships, the processing code was successfully recoded in Python from the original Matlab scripts.

The relationships were used with satellite-derived skin temperatures to estimate air temperature fields for each surface for each day. A combined uncertainty was calculated for the daily fields by adding all contributing terms in quadrature and the temperatures, uncertainties and ancillary information were written out to NetCDF files in a consistent format (see Appendix A for details).

For the oceans, there were separate retrievals from ATSR-2 and AATSR, these were separately converted to air temperatures and averaged where they overlapped. The uncertainties were combined assuming that systematic and locally-correlated error components were fully correlated within a daily, 0.25° grid cell; uncorrelated errors were treated appropriately.

Some example fields for the three surface types are shown in Figure 1. For land, the variables Tmax and Tmin are available from 2002 to 2016; for ice, Tmax, Tmin and Tavg are available from 2001-2009; for the oceans only Tavg is available from 1995-2012.

At milestones in the development of EUSTACE, fields calculated for different surface types were compared by calculating daily temperature differences where comparable variables were available in the same 0.25° grid cells. A total uncertainty in the difference was also calculated by taking the root of the sum of the squared uncertainty components. Differences were

normalised by the total estimated uncertainty, which naturally accounts for the large variation in expected differences due to known uncertainties (less than 0.1°C to greater than 5°C). Differences were aggregated at daily, monthly and annual time scales producing maps (e.g. Figure 2) and histograms (e.g. Figure 3).
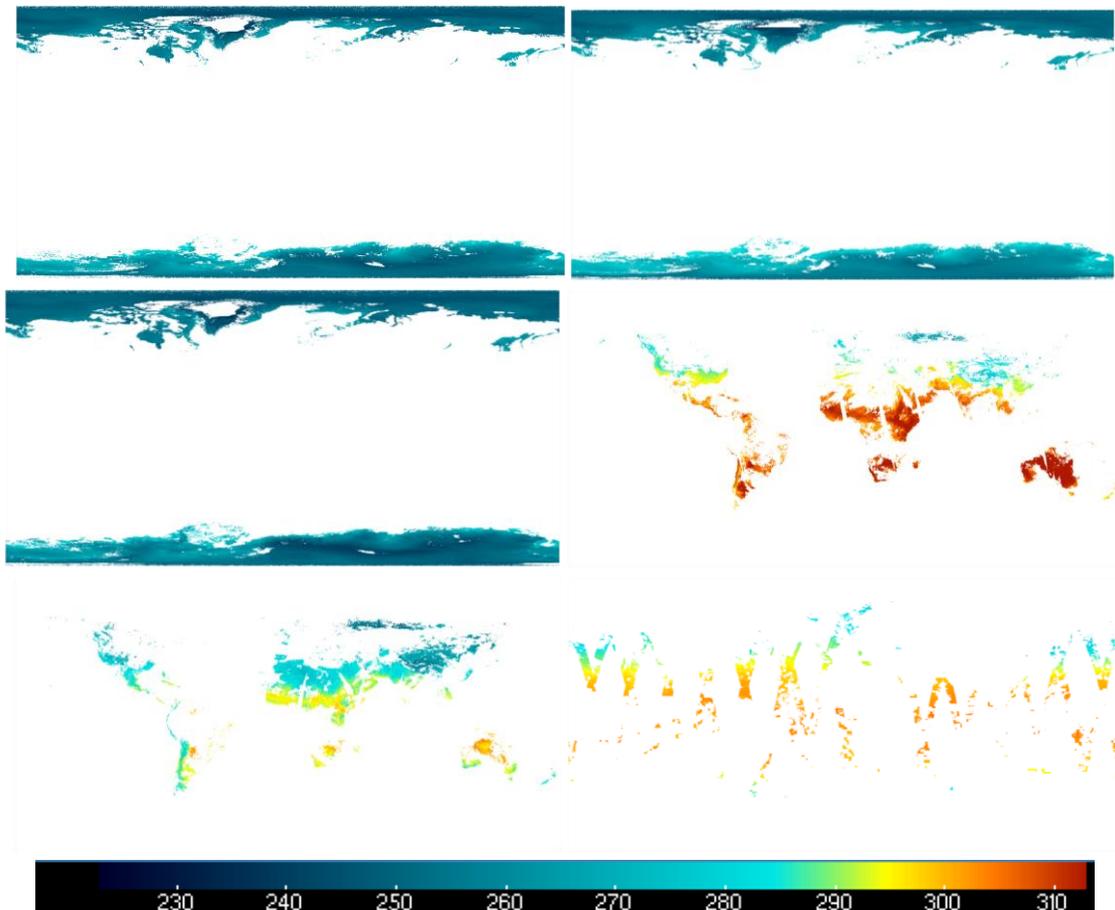

Figure 1: Example estimated air temperature fields (K) for 4th February 2003. (from top left): Tavg Ice surface air temperature, Tmax ice surface air temperature, Tmin ice surface air temperature, Tmax land surface air temperature, Tmin land surface air temperature and Tavg marine surface air temperature. All images use the same temperature scale.

Figure 2 shows the mean difference between air temperatures derived from LST and IST for April 2008 and for the whole of 2006. The overlapping areas are very limited in any particular month or year because air temperatures derived from LST are screened out in ice-covered areas. Nonetheless in areas where the two do overlap, the distribution of normalised differences is typically narrower and more sharply peaked with larger numbers of outliers than would be expected if the uncertainties were correct and the distributions were Gaussian. This suggests that the uncertainties are somewhat overestimated and that the distributions do not exactly conform to theory.
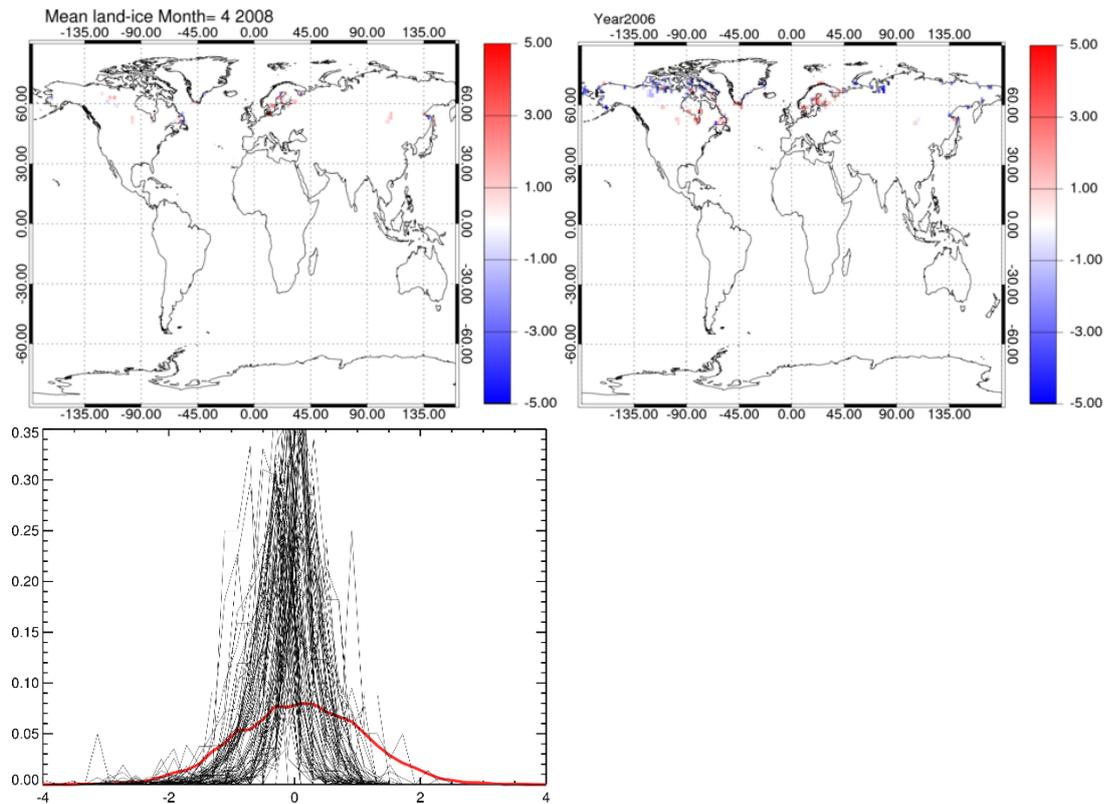
Figure 2: Mean difference (K) between air temperature derived from Land Surface Temperature and air temperature derived from Ice Surface Temperature for April 2008 (top left) and for the whole of 2006 (top right). Histogram of differences (bottom left) between co-located air temperature over land and ice normalised by the estimated combined uncertainty (x-axis) for all individual years in the comparison (black lines). The red line is a reference distribution with mean zero and a standard deviation of one.

Figure 3 shows a histogram of differences between air temperature estimates based on LST and SST, at locations where both land and ocean are present, normalised by the estimated uncertainty (red line). If the uncertainties are correctly specified, the width of the red histogram should be similar to the grey reference histogram. In this case, the red histogram is narrower than the reference histogram suggesting that the combined uncertainties are overestimated (similar to the land-ice comparison). This is generally the case for land-ocean comparisons. There is a slight offset between the average air temperature over land and the air temperature over the oceans. This arises from actual temperature differences over these two surfaces, which are never exactly co-located.
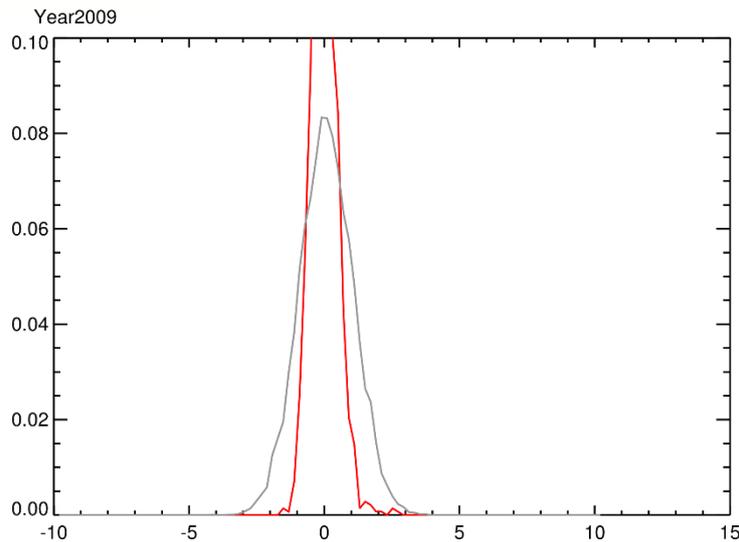
Year2009

Figure 3: Histogram of differences between co-located air temperature over land and ocean normalised by the estimated combined uncertainty for September 2009. The red line shows the normalised differences and the grey line is a reference distribution with mean zero and a standard deviation of one. If the uncertainties are perfectly specified, the red and grey lines should overlap. As the red histogram sits "inside" the grey histogram, it is likely that the uncertainties are over-estimated.

3.5 Known issues

In the current version of the data set, there are twenty days of missing land surface air temperatures. They are missing because there are no vegetation data for these dates and these are needed to estimate the land surface air temperature from the land surface skin temperature.

Ocean and land processing are carried out by running IDL sub-processes controlled by the EUSTACE end-to-end system. This runs effectively as part of the system used here, but the use of IDL currently limits the speed at which the whole of the consistent air temperature data set can be run as there are a limited number of IDL licenses available on CEMS.

References

Hoyer, J., E. Good, P. Nielsen-Englyst, K. S. Madsen, R. I. Woolway, J. Kennedy (2018) Report on the relationship between satellite surface skin temperature and surface air temperature observations for oceans, land, sea ice and lakes, EUSTACE Deliverable 1.5

Appendix A

A.1 Data provided

The EUSTACE Air temperature estimates from satellite product provides daily surface air temperature estimates with uncertainty information for the globe over the three categories of land, sea, and ice. Values provided are:

- Maximum and minimum daily temperatures over land
- Average (temporal mean) daily temperature over sea
- Maximum, minimum, and average (temporal mean) daily temperatures over ice

For each temperature measure a total per-grid-box uncertainty figure is available. In addition the ancillary files contain a breakdown of total uncertainty into components arising from random effects (assumed independent per grid-box), one or more components arising from systematic effects with widespread spatial and/or temporal correlation (possibly with associated spatio-temporal correlation patterns), and one or more components arising from locally-correlated effects, with associated spatio-temporal scales.

A.2 Evidence from user interactions

The following key points of feedback after provision of product mock-ups informed the design choices for file and variable names:

- Probably best not to make assumptions on how to estimate the average air temperature for grid cells with more than one surface type.
- Good to have the air temperature estimates and the related (total) uncertainty in the same file (as separate variables).
- Conflicting evidence on whether all surface types should be in one file or separate ones.
- More detailed uncertainty information (other than total uncertainty) may be helpful to some users but not all, and therefore should be in separate files so as to minimise download time for those who don't need it.

Based on this information, a separate field of values has been provided for each surface, with masked values where no data are present. For ease of implementation, separate surface types will be in separate files. This also meets user needs for fast download times. Temperature estimates together with total uncertainties are provided together in one file as requested.

In addition, we provide a separate ancillary file to store uncertainty components. This enables production of a total uncertainty estimate at different resolutions, because this requires knowledge of correlation structure in the uncertainty components.

A.3 Global field variables in surface air temperature files

The variables described below have been extracted from the NetCDF files and show the range of information available in the product described in this document.

Variable: tas (int16)
standard_name: air_temperature

```
long_name: Average daily surface air temperature
units: K
cell_methods: time: mean
scale_factor: 0.005
add_offset: 273.15
_FillValue: -32768
```

Variable: tasmin (int16)
```
standard_name: air_temperature
long_name: Minimum daily surface air temperature
units: K
cell_methods: time: minimum
scale_factor: 0.005
add_offset: 273.15
_FillValue: -32768
```

Variable: tasmax (int16)
```
standard_name: air_temperature
long_name: Maximum daily surface air temperature
units: K
cell_methods: time: maximum
scale_factor: 0.005
add_offset: 273.15
_FillValue: -32768
```

Variable: tasuncertainty (int16)
```
long_name: Total uncertainty in average daily surface air
temperature
units: K
scale_factor: 0.001
add_offset: 0
_FillValue: -32768
```

Variable: tasminuncertainty (int16)
```
long_name: Total uncertainty in minimum daily surface air
temperature
units: K
scale_factor: 0.001
add_offset: 0
_FillValue: -32768
```

Variable: tasmaxuncertainty (int16)
```
long_name: Total uncertainty in maximum daily surface air
temperature
units: K
scale_factor: 0.001
```

```
add_offset: 0
_FillValue: -32768
```

A.4 Global field variables in ancillary files

Global fields describing individual uncertainty components are specified in ancillary files these correspond to the uncertainty components described in Deliverable 1.5. Each locally-correlated error component has a time_scale and a length_scale which describe the characteristic distances and times over which they act. In some cases, the length scale is "unknown" because the underlying spatial correlations are not well understood.

For ice, the following fields are available for Tavg (tas) with equivalents for Tmin (tasmin) and Tmax (tasmax):

```
tas_unc_no_cloud: Total uncertainty excluding cloud on average daily
surface air temperature

tas_unc_rand: Random uncertainty on average daily surface air
temperature

tas_unc_corr_local: Locally correlated uncertainty on average daily
surface air temperature

tas_unc_sys: Systematic uncertainty on average daily surface air
temperature

tas_unc_cloud: Cloud component of uncertainty on average daily surface
air temperature.
```

For land, we have the following fields for Tmin with equivalents for Tmax:

```
tasmin_unc_rand: Random uncertainty on minimum daily surface air
temperature

tasmin_unc_corr_atm: Locally correlated atmospheric uncertainty on
minimum daily surface air temperature

tasmin_unc_corr_sfc: Locally correlated surface uncertainty on
minimum daily surface air temperature

tasmin_unc_sys: Systematic uncertainty on minimum daily surface air
temperature

tasmin_model_number: Model number used for estimating Tmin from
satellite data
```

For the oceans the following fields are available for Tavg:

```
tas_unc_rand: Random uncertainty on average daily surface air
temperature

tas_unc_corr_sat: Locally correlated uncertainty (from satellite
retrieval) on average daily surface air temperature
```

tas_unc_sys: Systematic uncertainty on average daily surface air temperature

tas_unc_corr_mod: Locally correlated uncertainty (from surface-air model) on average daily surface air temperature

tas_unc_sys_mod: Systematic uncertainty (from surface-air model) on average daily surface air temperature

tas_unc_parameter_0: Systematic uncertainty mean offset on average daily surface air temperature

tas_unc_parameter_1: Systematic uncertainty first fourier component on average daily surface air temperature

tas_unc_parameter_2: Systematic uncertainty second fourier component on average daily surface air temperature

tas_unc_parameter_3: Systematic uncertainty third fourier component on average daily surface air temperature

tas_unc_parameter_4: Systematic uncertainty fourth fourier component on average daily surface air temperature