Product design for satstace and fullstace system outputs

Expected files, variables, and attributes in the output of EUSTACE satstace and fullstace systems

Joel R. Mitchelson

*Revisions*

|  |  |  |
| --- | --- | --- |
| **Date** | **Author** | **Comment** |
| 13/12/2016 | Joel Mitchelson | Created. |
| 14/12/2016 | Joel Mitchelson | Add separate tasmaxminclimatologyfraction. |

Contents

[1 Overview 2](#_Toc469499695)

[1.1 Scope 2](#_Toc469499696)

[1.2 Dependencies 2](#_Toc469499697)

[1.3 Objectives 2](#_Toc469499698)

[1.4 Key technical requirements 2](#_Toc469499699)

[2 Design of satstace outputs 3](#_Toc469499700)

[2.1 Data provided 3](#_Toc469499701)

[2.2 Evidence from user interactions (WP4) 3](#_Toc469499702)

[2.3 Container format for satstace files 4](#_Toc469499703)

[2.4 File and directory names for satstace 4](#_Toc469499704)

[2.5 Global field variables in satstace surface air temperature files 4](#_Toc469499705)

[2.6 Global field variables in satstace ancillary files 5](#_Toc469499706)

[3 Design of fullstace outputs 6](#_Toc469499707)

[3.1 Data provided 6](#_Toc469499708)

[3.2 Evidence from user interactions (WP4) 6](#_Toc469499709)

[3.3 Container format for fullstace files 6](#_Toc469499710)

[3.4 File and directory names for fullstace 6](#_Toc469499711)

[3.5 Variable names for fullstace files 6](#_Toc469499712)

[3.5.1 Global file 6](#_Toc469499713)

[3.5.2 Ensemble file 8](#_Toc469499714)

[4 References 8](#_Toc469499715)

# Overview

## Scope

The design described by this document applies to the following planned EUSTACE products which are outputs from work package 2:

* Surface air temperature estimates (with estimates of uncertainty) for all surfaces of Earth, derived from satellite surface skin temperature retrievals
* Globally-complete daily analyses of surface air temperature (with estimates of uncertainty) for the whole of Earth since 1850, based on combined information from satellite and in situ data sources

## Dependencies

This document is not stand-alone but is **intended to be read alongside implementation details provided in EUSTACE Milestone 12** ‘Recommendations on data formats and management’ , and subsequent revisions.

## Objectives

Whereas gives NetCDF file structure in general terms, here we describe the specific lists of files, variables, and attributes which appear in the outputs. These must meet the technical requirements and draw on evidence from users .

In addition we provide guidelines on how some of the desirable (but not essential) features in section 5 can be met.

## Key technical requirements

The following technical requirements are of particular importance to the part of the design addressed by this document:

|  |  |  |  |
| --- | --- | --- | --- |
| PRDREQ005 | Outputs include daily maximum, minimum, and mean surface air temperature | D4.1 | “Temperature variable: Mean, Minimum, Maximum” |
| PRDREQ006 | Uncertainty information must be provided | DA, MS30, D4.1 | DA: “report new, consistent, validated estimates of uncertainty...”  MS30: “All stakeholders recognize that information on uncertainties is important...”  D4.1: “Within box variability” |
| PRDREQ012 | Output is for all locations over all surfaces of Earth | DA | “...for all locations over all surfaces of Earth (i.e. land, ocean, ice and lakes),...” |
| PRDREQ015 | The sampling period for daily records at each location is from midnight to midnight measured in local solar time for the location. | GA1 |  |

In addition the following desirable features from may bring additional requirements regarding uncertainty information:

|  |  |  |
| --- | --- | --- |
| OPT003 | Products are as small as possible to facilitate quick download | MS30 |
| OPT007 | Format facilitates extraction of sub-periods | MS30 |
| OPT008 | Format facilitates extraction of sub-regions | MS30 |

# Design of satstace outputs

## Data provided

The satstace system provides daily surface air temperature estimates with uncertainty information for the globe over the four categories of land, sea, ice, and lakes. Values expected 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
* Average (temporal mean) daily temperature over lakes

For each temperature measure a total per-grid-box uncertainty figure is available. In addition the system produces a breakdown of total uncertainty into a random component (assumed independent per grid-box), one or more systematic components with associated spatio-temporal patterns, and one or more locally correlated components with associated spatio-temporal scales.

## Evidence from user interactions (WP4)

The following key points of feed back are taken from draft feedback on mockups directly affect the design choices for satstace 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 will be provided for each surface, with masked values where no data is present. For ease of implementation separate surface types will be in separate files. This also meets user needs for fast download times (OPT003). Temperature estimates together with total uncertainties will be provided together in one file as requested.

In addition there is provision in the design of a separate ancillary file to store uncertainty components. This is not required to meet the user needs nor the core requirements. However it may be needed to implement additional desirable features OPT007, OPT008, if there is a change in grid resolution. This is because producing a total uncertainty figure at different resolution requires knowledge of correlation structure.

## Container format for satstace files

All files follow the recommendations set out in with respect to global fields (section 1 ‘Global field data product format’).

## File and directory names for satstace

The following values of predicates are used as described in section 1.4 to define file and directory names:

|  |  |
| --- | --- |
| **Predicate** | **Value** |
| *project* | eustace |
| *product* | satellitederived |
| *institute* | mohc |
| *framework* | eustace |
| *collection* | land | sea | ice | lakes |
| *frequency* | day |
| *Table* | 0 |
| *realization* | 0 |
| *product\_version* | *Version taken from EUMOPPS* |
| *version* | *Date of launch of processing* |
| *variable* | tas | ancillary |
| *YYYYmmdd* | *Date at longitude 0E in the contained data* |

## Global field variables in satstace surface air temperature files

**Variable: tas (int16)**

standard\_name: air\_temperature

long\_name: Average daily surface air temperature

units: K

cell\_methods: time: mean

scale\_factor: 0.002

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: 0.002

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: 0.002

add\_offset: 273.15  
\_FillValue: -32768

**Variable: tasuncertainty (int16)**

standard\_name: air\_temperature standard\_error

long\_name: Total uncertainty in average daily surface air temperature

units: K

scale\_factor: 0.002

add\_offset: 0  
\_FillValue: -32768

**Variable: tasminuncertainty (int16)**

standard\_name: air\_temperature standard\_error

long\_name: Total uncertainty in minimum daily surface air temperature

units: K

scale\_factor: 0.002

add\_offset: 0  
\_FillValue: -32768

**Variable: tasmaxuncertainty (int16)**

standard\_name: air\_temperature standard\_error

long\_name: Total uncertainty in maximum daily surface air temperature

units: K  
scale\_factor: 0.002

add\_offset: 0  
\_FillValue: -32768

## Global field variables in satstace ancillary files

Global fields describing individual uncertainty components can be specified in ancillary files according to the guidelines in section 1.8.2.

# Design of fullstace outputs

## Data provided

The fullstace system provides ‘best estimate’ fields of maximum, minimum, and mid-range daily surface air temperature estimates for the globe (using combined inputs from satellites and in-situ data), together with a value for total uncertainty and a ‘climatology fraction’ which is a dimensionless quantity in range 0-1 indicating the extent to which a temperature value relies on climatology versus being constrained by available observations local in space and time. Further, there is an ensemble of temperature fields to represent the uncertainty in each variable.

## Evidence from user interactions (WP4)

The following key points of feed back are taken from draft feedback on mockups directly affect the design choices for fullstace file and variable names:

* If both variables tasmax and tasmin were in the same netcdf file, it would simplify the processing for the end user (less files to download/manage)

## Container format for fullstace files

All files follow the recommendations set out in with respect to global fields (section 1 ‘Global field data product format’).

## File and directory names for fullstace

The following values of predicates are used as described in to define file and directory names:

|  |  |
| --- | --- |
| **Predicate** | **Value** |
| *Project* | eustace |
| *product* | combined |
| *institute* | mohc |
| *framework* | eustace |
| *collection* | global | ensemble |
| *frequency* | day |
| *Table* | 0 |
| *realization* | 0 |
| *product\_version* | *Version taken from EUMOPPS* |
| *version* | *Date of launch of processing* |
| *variable* | tas |
| *YYYYmmdd* | *Date at longitude 0E in the contained data* |

## Variable names for fullstace files

### Global file

**Variable: tas (int16)**

standard\_name: air\_temperature

long\_name: Mean daily surface air temperature

units: K

cell\_methods: time: mid\_range  
scale\_factor: 0.002

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: 0.002

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: 0.002

add\_offset: 273.15  
\_FillValue: -32768

**Variable: tasclimatologyfraction (int16)**

long\_name: Climatology fraction for mean daily surface air temperature

scale\_factor: 1.5625e-05

add\_offset: 0.5

\_FillValue: -32768

**Variable: tasminmaxclimatologyfraction (int16)**

long\_name: Climatology fraction for minimum and maximum daily surface air temperature

scale\_factor: 1.5625e-05

add\_offset: 0.5

\_FillValue: -32768

**Variable: tasuncertainty (int16)**

long\_name: Total uncertainty in mean daily surface air temperature

units: K

scale\_factor: 0.002

add\_offset: 0  
\_FillValue: -32768

**Variable: tasminuncertainty (int16)**

long\_name: Total uncertainty in minimum daily surface air temperature

units: K

scale\_factor: 0.002

add\_offset: 0  
\_FillValue: -32768

**Variable: tasmaxuncertainty (int16)**

long\_name: Total uncertainty in maximum daily surface air temperature

units: K

scale\_factor: 0.002

add\_offset: 0  
\_FillValue: -32768

### Ensemble file

**Variable: tasensemble (int16)**

standard\_name: air\_temperature

long\_name: Mean daily surface air temperature ensemble

units: K

cell\_methods: time: mid\_range

scale\_factor: 0.002

add\_offset: 273.15  
\_FillValue: -32768

**Variable: tasminensemble (int16)**

standard\_name: air\_temperature

long\_name: Minimum daily surface air temperature ensemble

units: K

cell\_methods: time: minimum

scale\_factor: 0.002

add\_offset: 273.15  
\_FillValue: -32768

**Variable: tasmaxensemble (int16)**

standard\_name: air\_temperature

long\_name: Maximum daily surface air temperature ensemble

units: K

cell\_methods: time: maximum

scale\_factor: 0.002

add\_offset: 273.15  
\_FillValue: -32768

# References

|  |  |
| --- | --- |
| [1] | European Commission, *Grant agreement number: 640171 — EUSTACE — H2020-EO-2014, Annex 1 - Description of Action,* 2014. |
| [2] | J. R. Mitchelson, C. Morice, F. Lindgren, A. Stephens and A. Waterfall, “Recommendations on data formats and management,” EUSTACE Project Milestone 12, 2015. |
| [3] | J. R. Mitchelson, “EUSTACE System Technical Requirements,” Met Office Hadley Centre, Exeter, UK, 2016. |
| [4] | J. Bessembinder, “Report on user requirements based on first user workshop and literature review,” EUSTACE Project Milestone 30, 2015. |
| [5] | J. Bessembinder, “Report on user interaction workshop 1,” EUSTACE Project Milestone 36, 2016. |
| [6] | J. Bessembinder, “First concept report on user feedback on mockups of data files and satstace,” 2016. [Online]. Available: http://proj.badc.rl.ac.uk/eustace/attachment/wiki/ProjMstones/EUSTACE%20user%20feedback%20on%20mockups%20data%20files.docx. |