UNCERTAINTIES

EUSTACE has received funding from the European Union's Horizon 2020 Programme for Research and Innovation, under Grant Agreement no 640171

Uncertainty or Error?

- Error is the result of a measurement minus a true value (often unknown) of the measurand
- If the true value cannot be determined we therefore talk about uncertainty rather than error
- Another way of thinking about this is that the error is the "wrongness" of the measured value, whereas the uncertainty is the "doubt" given our knowledge of the measured value and the effects causing the errors

Common Uncertainty Format

- A common uncertainty format has been applied to the satellite data sets for all the surfaces
- EUSTACE produces a consistent set of multiyear satellite surface skin temperature records and validated uncertainties for all surfaces of the Earth.
- Uncertainties have been validated

<u>Uncertainties</u>

Uncertainties categorised by effects whose errors have distinct correlation properties:

- Random (level 1 NEdTs, Geolocation)
- Locally systematic (atmospheric effects, emissivity)
- (large-scale) Systematic

This three-component model can be applied to all processing levels and products

Propagation of uncertaintes:

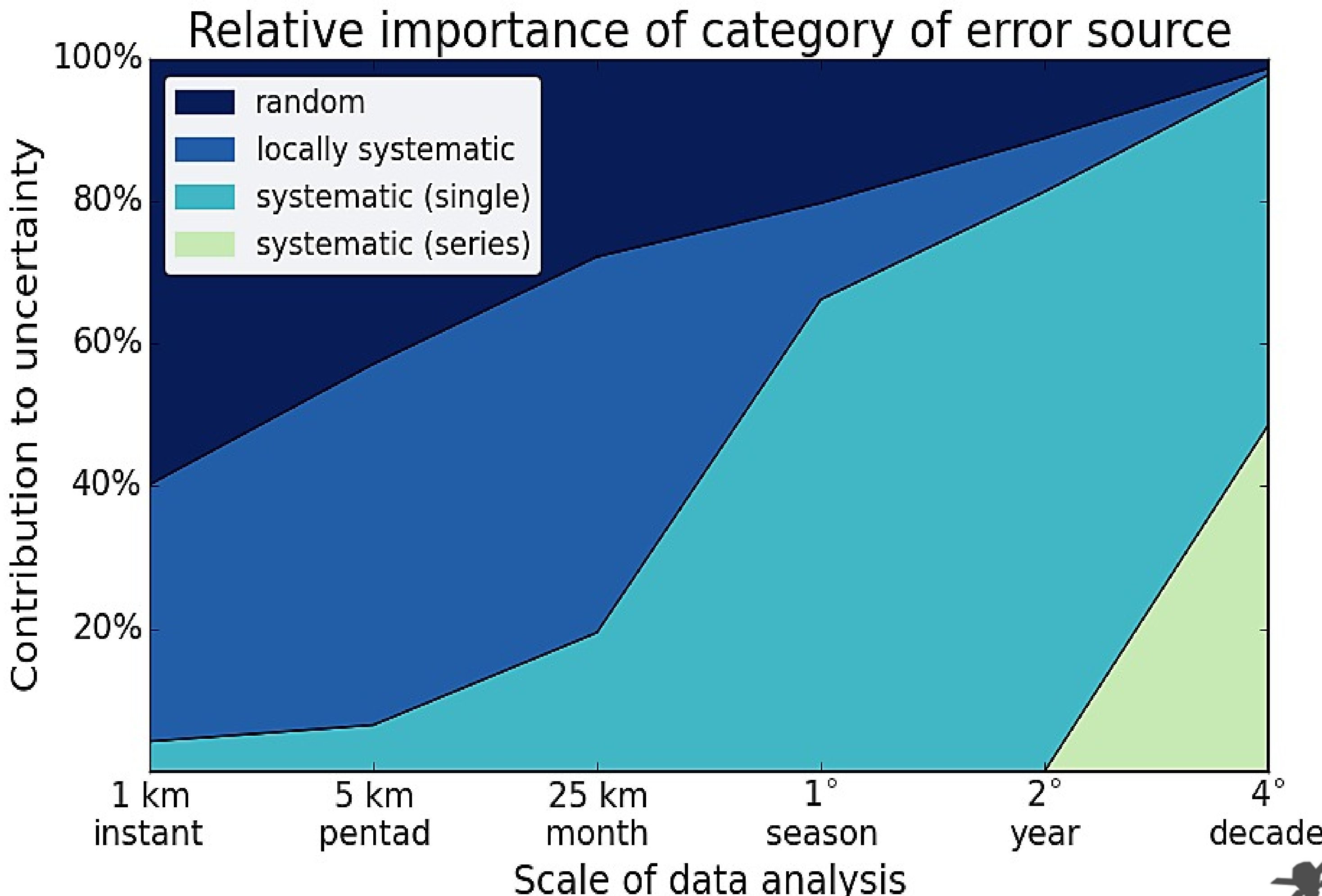
- L1 → L2 → L3 → L4 (Merged Product)

UNIVERSITY OF LEICESTER

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

EUSTACE

Importance of error sources in climate data on different analysis scales



EUSTACE

UNCERTAINTIES IN SATELLITE DATA

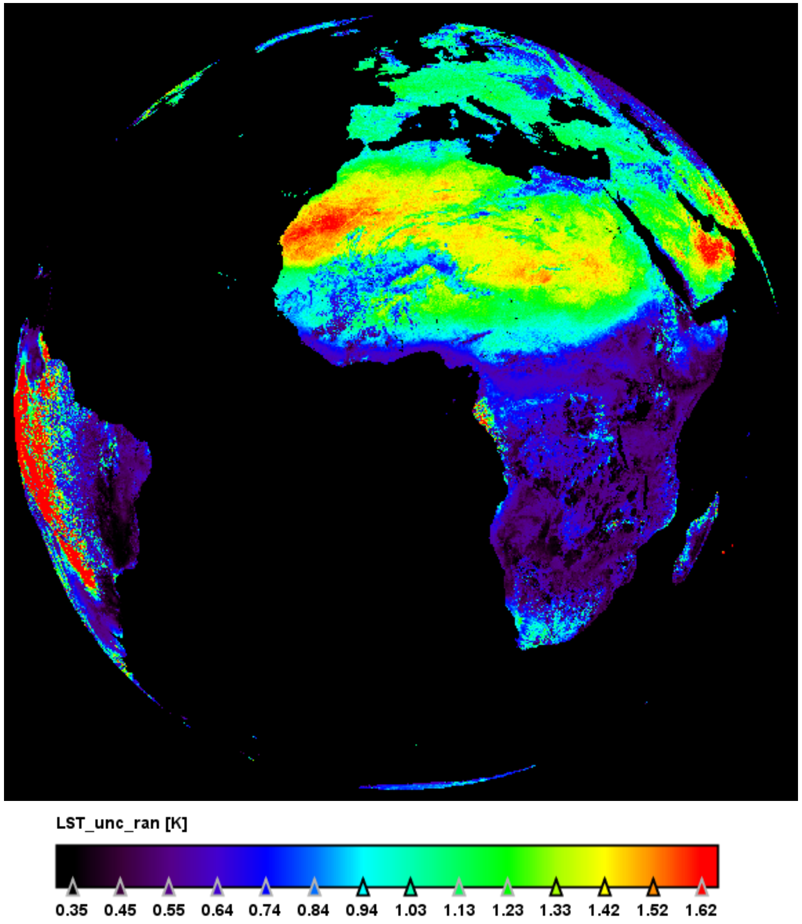
EUSTACE has received funding from the European Union's Horizon 2020 Programme for Research and Innovation, under Grant Agreement no 640171

Example: LST Uncertainty Components

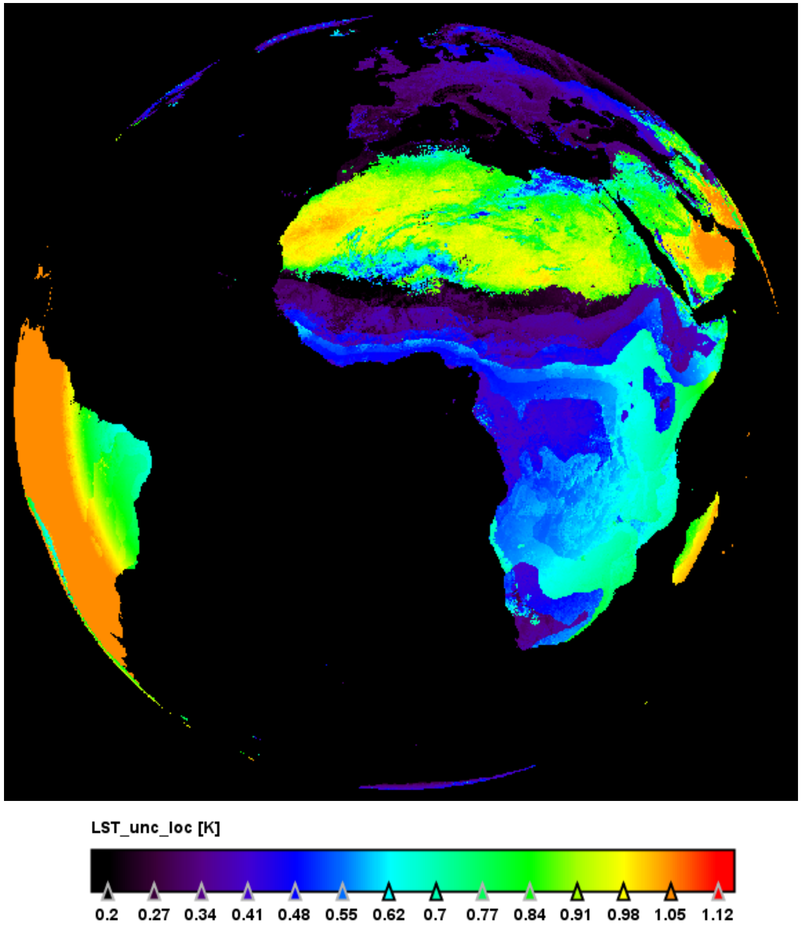
VARIABLE	METHOD	COMMENTS	
LST_UNC_RAN	**L2 Random 1 / Radiance noise Propagation**	$u_{ran,y}(x) = \sqrt{\sum_{c=1}^{n}\left(\frac{\partial R}{\partial y_c} u_{ran}(y_c)\right)^2}$	Random component of L1 channel uncertainties propagated through the retrieval
	L2 Random 2 / Emissivity noise Propagation	$u_{ran,\varepsilon}(x) = \sqrt{\sum_{c=1}^{n}\left(\frac{\partial R}{\partial \varepsilon_c} u_{ran}(\varepsilon_c)\right)^2}$	Estimate of the magnitude of pixel-to-pixel scale emissivity variability within areas based on land cover class
LST_UNC_LOC	**L2 Local 2 / Uncertainty from atmosphere/fit for regression-based retrieval**	$u_{loc,fit}(x) = \sqrt{Var(\hat{x} - x_{in})}$	Atmospheric fields correlated on timescales >1 day and length scales >100 km. For coefficient based retrieval methods the retrieval ambiguity is a contributor of residuals in the fit
	L2 Local 2 / Uncertainty from Emissivity	$u_{loc,\varepsilon}(x) = \sqrt{\sum_{c=1}^{n}\left(\frac{\partial R}{\partial \varepsilon_c} u_{loc}(\varepsilon_c)\right)^2}$	Across a particular land class area, there may be a mean difference between the assumed and true mean emissivity
LST_UNC_SYS	**L2 Systematic 1 / Reasoned estimate**	Assumed that known corrections have been applied by data producers and what remains is describable as an uncertainty in the bias of the satellite surface temperatures relative to other data sources of temperature (ie from validation)	

Example: UoL SEVIRI Uncertainties

Random



Locally correlated



UNIVERSITY OF LEICESTER

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

EUSTACE

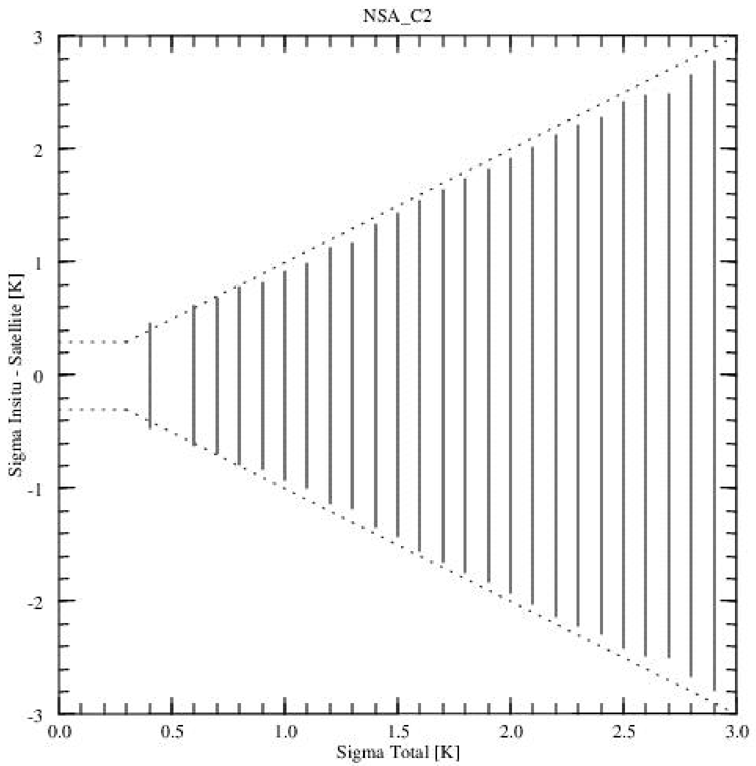
VALIDATION OF SATELLITE UNCERTAINTIES

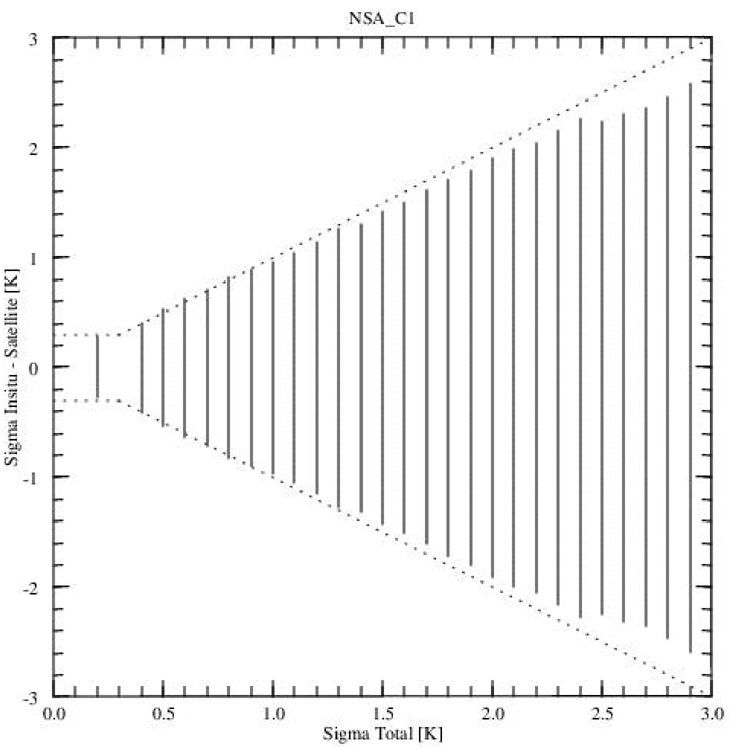
- Test the goodness-of-fit between the uncertainty from in situ validation ($\sigma_{sat\text{-}ground}$) and the total satellite product uncertainty for each associated matchup (σ_{total})

- $\sigma_{total} = \sqrt{\sigma_{sat}^2 + \sigma_{ground}^2 + \sigma_{space}^2 + \sigma_{time}^2 (+\sigma_{depth}^2)}$

- σ_{sat} is the total LST uncertainty for satellite pixel
- σ_{ground} is the uncertainty associated with the ground-based measurement
- σ_{space} is the uncertainty associated with matching a satellite and ground observation in a spatial context
- σ_{time} is the uncertainty associated with matching a satellite and ground observation in time
- σ_{depth} is the uncertainty due to the difference in depth of the measurements (SST only)

ICE TSKIN UNCERTAINTY VERIFICATION





AASTI IST uncertainty validation with respect to ARM in-situ data for 2008. Dashed lines show ideal uncertainty model accounting for uncertainties in the in situ data and geophysical uncertainties arising from spatial and temporal collocation. Solid black lines show one standard deviation of the retrieved minus in situ IST differences for each 0.1 K bin.

UNIVERSITY OF LEICESTER

National Centre for Earth Observation

NATURAL ENVIRONMENT RESEARCH COUNCIL

EUSTACE

Uncertainty propagation to SAT

- Propagation from Level-2 skin uncertainties to Level-3 gridded product:
  - Uncertainty from random effects reduces as 1/√n
  - Uncertainty from locally correlated effects do not reduce down
  - Sampling uncertainty is an additional random term introduced when the cell is not fully sampled in space and time

- Propagation from Skin to SAT is:

$$T_{type} = M(x, \chi)$$

- Uncertainty propagation from X_{L3} to T_{type}:

$$u_{comp,x}(T_{type}) = \left| \frac{\partial M}{\partial x_{L3}} u_{comp}(x_{L3}) \right|$$

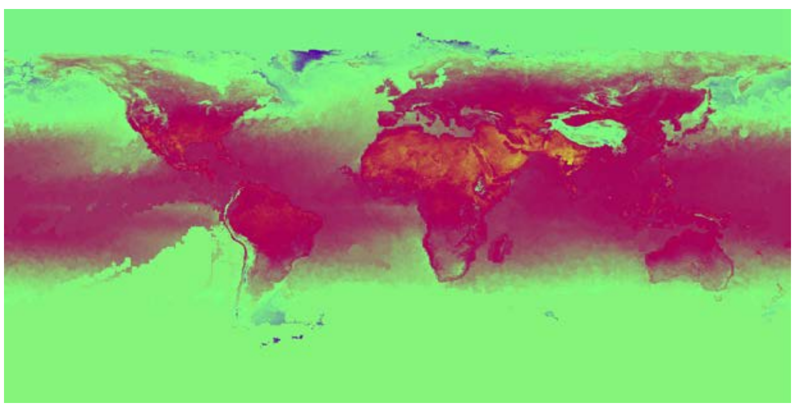
UNIVERSITY OF LEICESTER National Centre for Earth Observation NATURAL ENVIRONMENT RESEARCH COUNCIL EUSTACE

UNCERTAINTY INFORMATION IN OUTPUT PRODUCTS

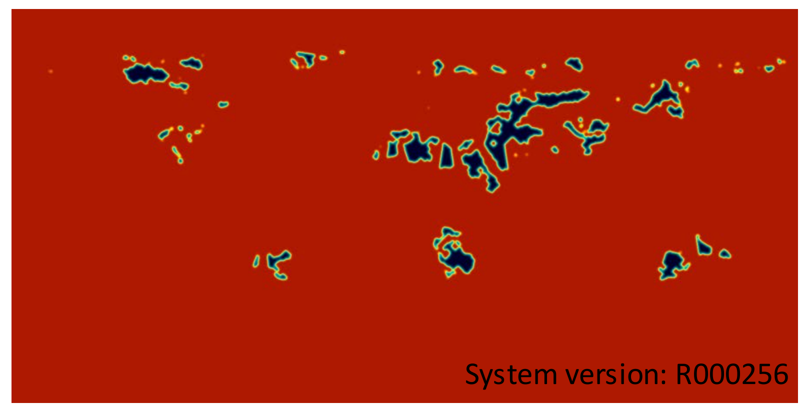
EUSTACE has received funding from the European Union's Horizon 2020 Programme for Research and Innovation, under Grant Agreement no 640171

CLIMATOLOGY FRACTION

- Alongside the complete analysis fields, there is a field called **climatology fraction** which indicates the extent to which the output is constrained by local observations (the weather) rather than long-range effects (the climate and long-distance interactions).
- The mockups do not include proper statistical inference for long-range effects and so there are large areas entirely dependent on climatology. It won't be like this in the final version.



Surface air temperature field in mockup



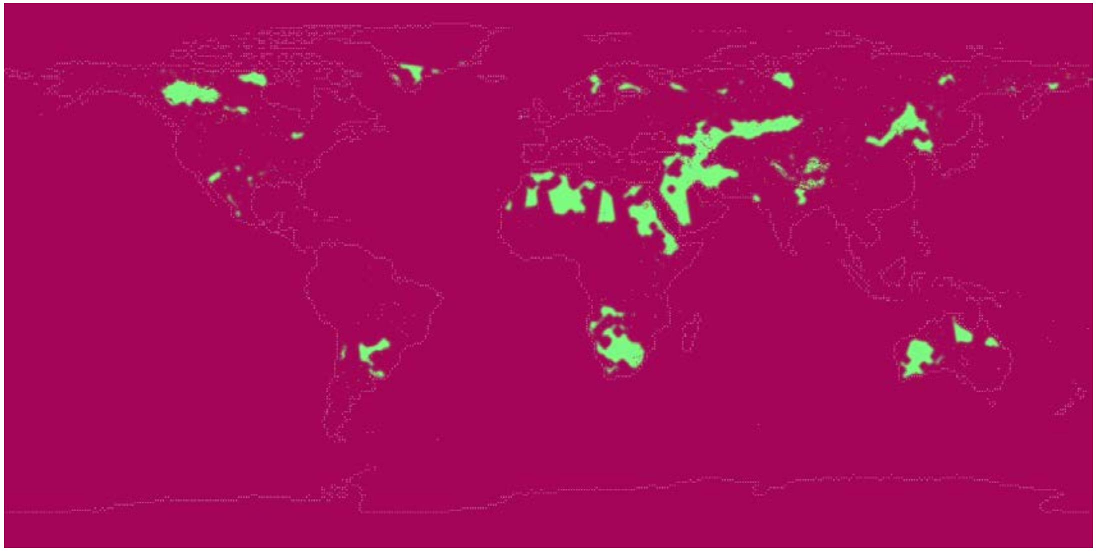
Climatology fraction in mockup

Met Office
Hadley Centre

EUSTACE

STANDARD DEVIATION

- Outputs include a standard deviation field for each variable:



- BUT: when the climatology fraction is high, if a user assumed that a measurement in a grid box could be anywhere within +- 1 standard deviation this would allow very unlikely temperature fields, as it does not account for temporal and spatial dependencies.

Met Office
Hadley Centre

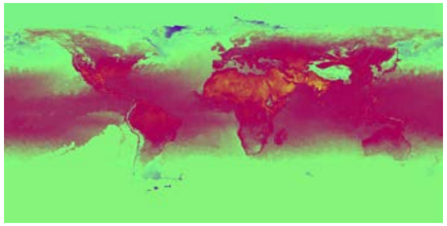
EUSTACE

EXPRESSING ENSEMBLES IN NETCDF OUTPUT

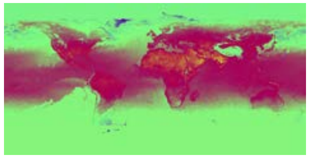
- A file contains one main variable indicating the ‘best’ output, represented as a global field. This has dimensions of time, latitude, and longitude.

- In the same file, there is a variable with an array of global fields. This is the ensemble. It has dimensions of ‘ensemble member’, time, latitude and longitude.

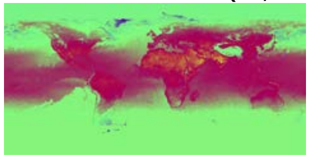
`NetCDF variable: tas`



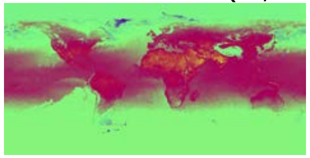
`tasensemble(0,...)`



`tasensemble(1,...)`



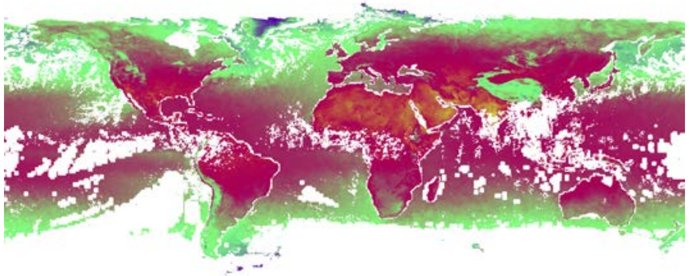
`tasensemble(2,...)`



Met Office
Hadley Centre

EUSTACE

SATELLITE-DERIVED PRODUCT



- The satellite-air temperature model for each surface type produces a set of uncertainty information (approximately 40 variables in total for each grid box)
- To perform statistical inference to produce global fields we take all of these individually.
- Who would use this set of uncertainty information and for what purpose?
- Is it more useful to provide a single combined field, with flags to indicate which surface model was used, and a summary of uncertainty?
- If we summarise uncertainty, what is useful?

Met Office
Hadley Centre

EUSTACE