CREATING A GLOBALLY-COMPLETE ANALYSIS

NICK RAYNER, SCIENCE COORDINATOR

JOINT EUSTACE AND GLOBTEMPERATURE USER MEETING, LISBON, 27-29TH NOVEMBER 2017

nick.rayner@metoffice.gov.uk

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

MOTIVATION

An attempt to use temperature data directly in applications is prone to a number of difficulties:

- Results might be unduly biased to the places and times where there are the most observations (sampling bias).
- Results are affected by measurement bias in data sources.
- It is hard to write software and devise mathematical methods to analyse irregularly sampled data; many off-the-shelf tools assume that data is available on a regularly sampled grid.

Met Office
Hadley Centre

EUSTACE

MOTIVATION

To solve these problems, we would like to consider these irregular data points as samples from a continuous temperature field, and to use prior understanding of the statistical properties of this field, so that:

- We can correct for data source biases
- We can use the properties of the field to avoid spatial bias in the data values, as well as in derived statistics such as frequency spectra
- We can take regularly-spaced samples from this field which work well with commonly available software tools and mathematical methods

Met Office
Hadley Centre

EUSTACE

TWO METHODS

AMBITIOUS

- Novel statistical methods created specifically for EUSTACE
- State of the art
- Idea is to use CEMS computing facility to the full
- Challenging for implementation

ADVANCED STANDARD

- Extend current methods to proposed EUSTACE resolution and data volumes
- Low risk
- Designed to be extended to greater complexity
- Sharing techniques with the ambitious method for efficient processing.

Met Office
Hadley Centre

EUSTACE

MEANS, MAX AND MIN

- Initial idea was to produce mean temperature
- User feedback said maximum and minimum temperatures more useful over land than mean

$$T_{mean} = (T_{max} + T_{min})/2$$
$$DTR = T_{max} - T_{min}$$

$$T_{max} = T_{mean} + 0.5\ DTR$$
$$T_{min} = T_{mean} - 0.5\ DTR$$

Met Office Hadley Centre

EUSTACE

Temperature Observation Model

Linear error model

Daily mean air temperatures are decomposed into variability at different scales:

$$y^i = T(s^i, t^i) + \beta^i + \epsilon^i$$

Where β^i is a sum of observational biases affecting observation i and ϵ^i are non-bias related observational errors.

y^i = An air temperature observation index by i

$T(s^i, t^i)$ = Temperature at space/time location (s^i, t^i)

β^i = Additive bias associated with observation i

ϵ^i = Error associated with observation i

Met Office Hadley Centre

EUSTACE

Advanced Standard Air Temperature Model

Temperature Process Decomposition

- Temperature variability is decomposed into model sub-components with defined structure in space/time:

$$T(s,t) = T^{\text{clim}}(s,t) + T^{\text{large}}(s,t) + T^{\text{local}}(s,t)$$

$T(s,t)$ = Temperature at space/time location (s,t)

$T^{\text{clim}}(s,t)$ = Climatological temperature

$T^{\text{large}}(s,t)$ = Large spatial/temporal scale component

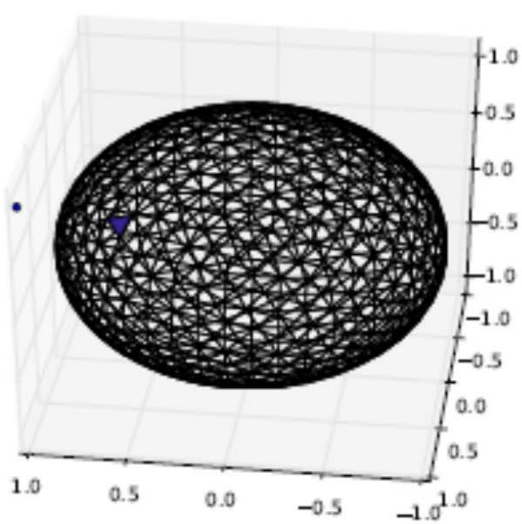
$T^{\text{local}}(s,t)$ = Daily, short spatial scale component

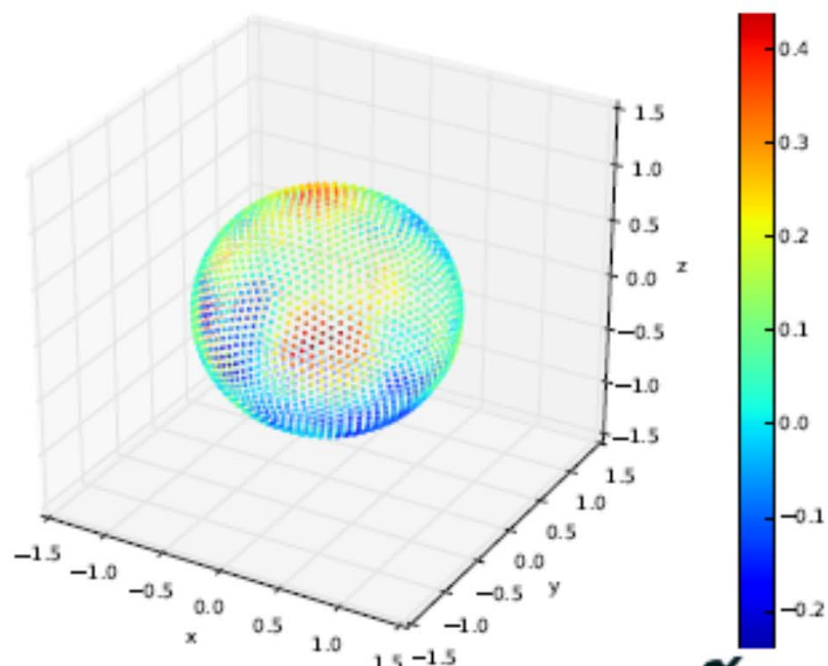
Met Office
Hadley Centre

EUSTACE

Spatial Estimation on a Sphere

- A nested triangulation of a sphere is used to define local (piecewise linear) basis functions;
- Also allows fast localisation of observations within the triangulation;
- Estimate basis function weights in a probabalistic manner given data;
- Observational bias terms estimated simultaneously.





Met Office
Hadley Centre

EUSTACE

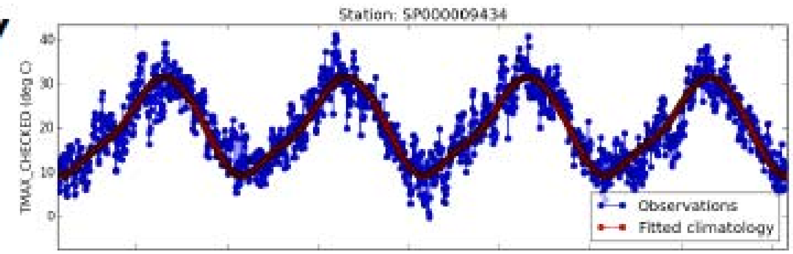
MODEL COMPONENT:

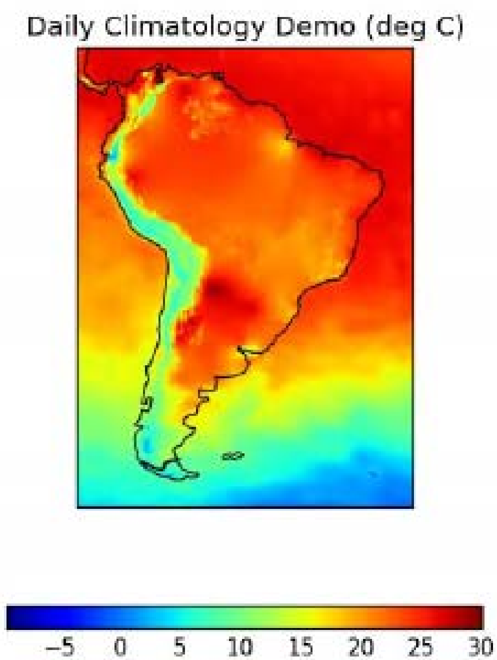
CLIMATOLOGY

Linear model in weights (the **c**'s):

$$\tau_m^{\text{climatology}}(\boldsymbol{s},t) = \sum^{K^{\text{harmonic}}} f_k^{\text{harmonic}}(\boldsymbol{s},t)c_k^{\text{harmonic}}$$

$$+ \sum_{k=1}^{K^{\text{covariate}}} f_k^{\text{covariate}}(\boldsymbol{s},t)c_k^{\text{covariate}}$$





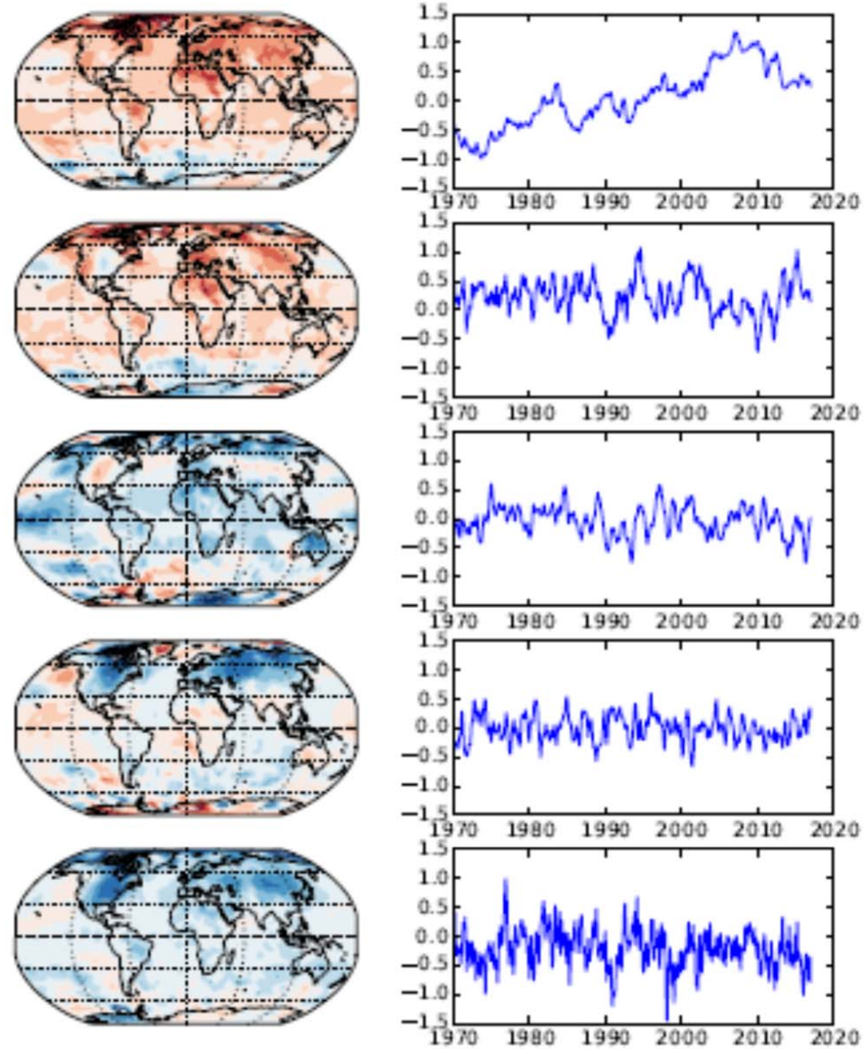
Climatology Component	Description
Seasonal Component	
Local Harmonic	Basis functions are temporal fourier series (2 to 3 sine and cosine harmonics) with local spatial influence. Basis function coefficients vary smoothly in space with GMRF prior.
Local Offset	Technically not seasonal, but considered to be the zeroth ordered temporal harmonic with same spatial basis function structure as the Local Harmonic model.
Covariate Component	
Grand mean	A constant offset for the whole globe. The climatological global average.
Harmonics of laititude	Spatial basis functions $\sin(2\lambda)$, $\cos(2\lambda)$, $\sin(4\lambda)$, $\cos(4\lambda)$.
Altitude	Currently learning a weighted function of altitude. Tests imply a spatially varying coefficient needed. Overlapping *altitude shaped* basis functions modulated by local windowing proposed.
Distance from water	For large water bodies. Needed for coastal effect. Also for lakes? Local variation needed? Not yet included.
Climatological surface type	Basis functions are indicators of climatological surface type. Not yet included.

Met Office
Hadley Centre

Model Component: Large-Scale

Large-scale model:

- A sum of *"factors"*;
- Each factor is a product of:
  - A spatial pattern;
  - A time series of pattern weights.



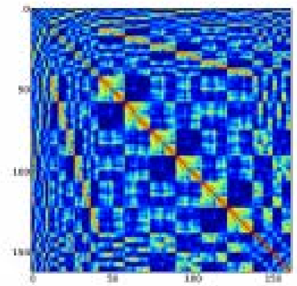
Met Office
Hadley Centre

EUSTACE

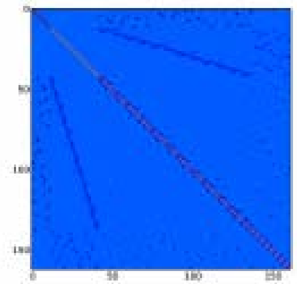
MODEL COMPONENT:

DAILY LOCAL ANALYSIS

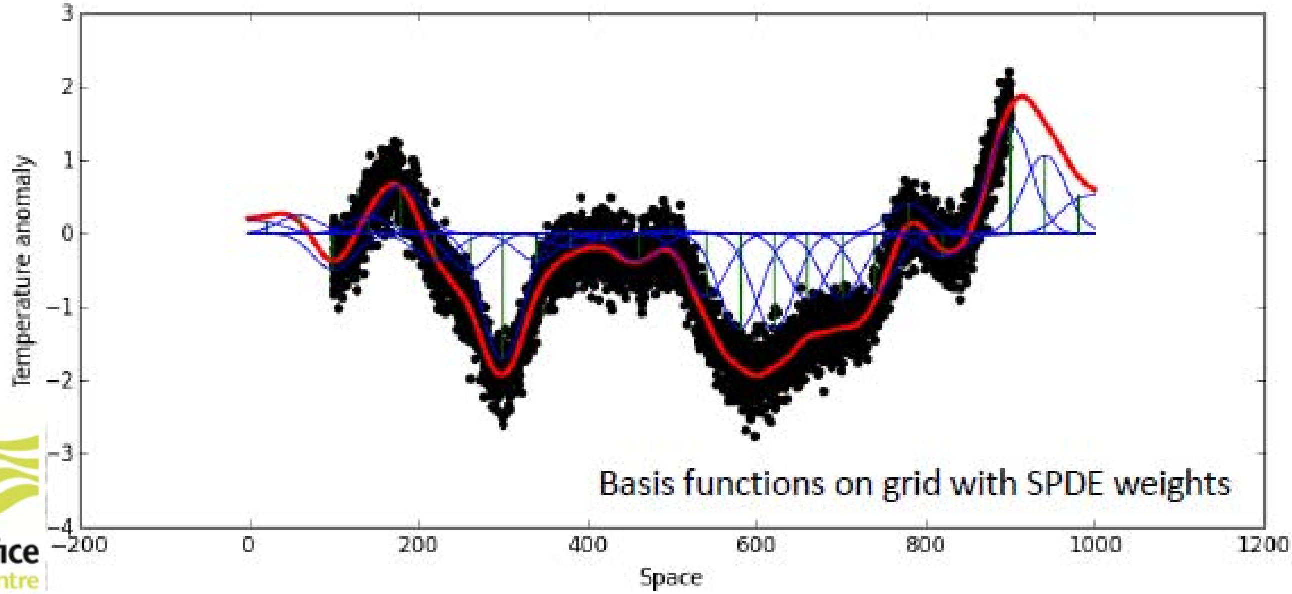
- Independent, daily analyses.
- Build as a local basis function model in space with smoothly varying coefficients.
- Lindgren et al. (2011) result allows efficient computation.
  - Directly compute sparse precision matrices corresponding to commonly used covariance functions.
  - Precision matrix parameters control smoothness.



Dense covariance matrix



Sparse precision matrix



Met Office Hadley Centre

EUSTACE

ANALYSIS PROTOTYPE

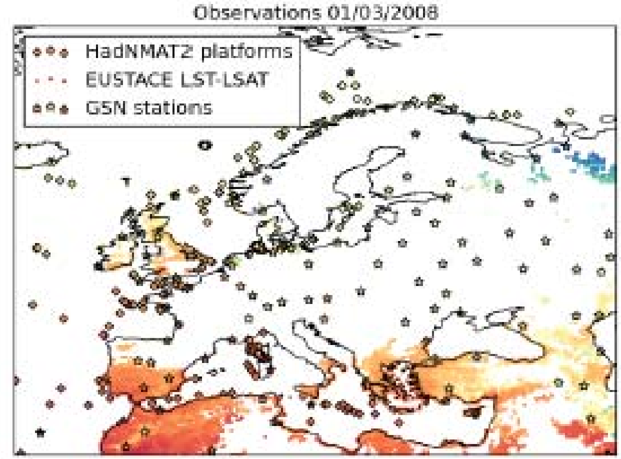
CLIMATOLOGY/LOCAL COMPONENT

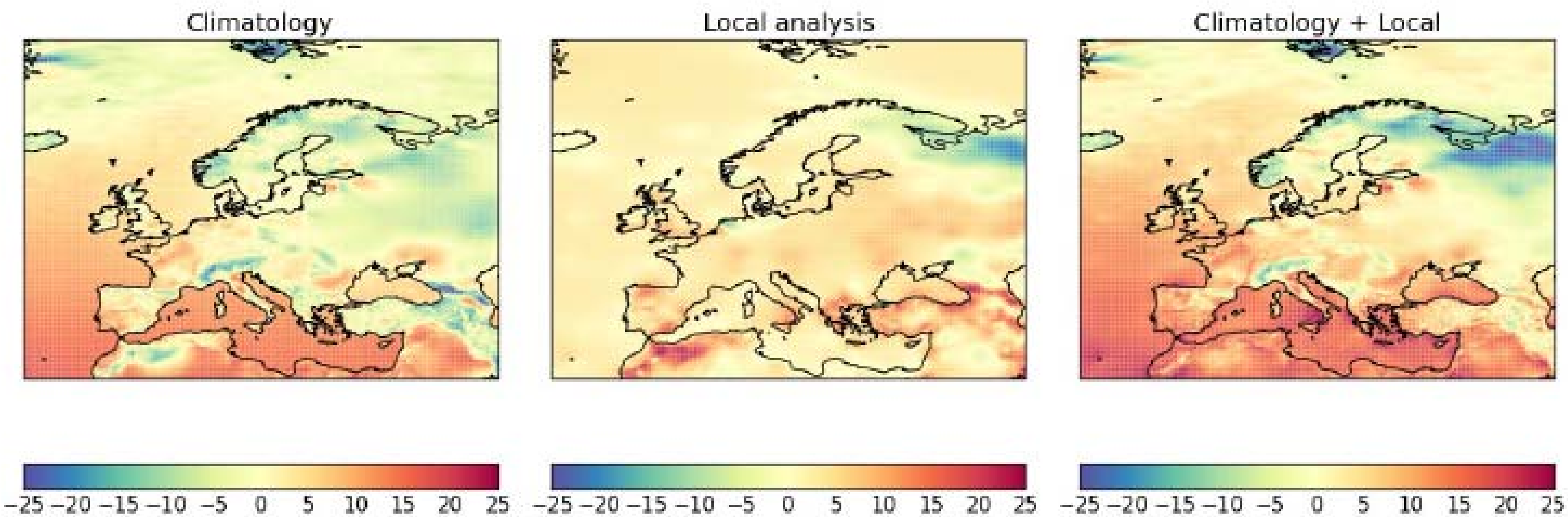
Applied to in situ NMAT/SAT & LSAT derived SAT:

- Fitted climatology – spatial SPDE's for coefficients of Fourier components in time.
- Fitted local – daily spatial SPDE.

Climatology used observations in 1961-1990 period.

Currently extending to include all model components.





Met Office
Hadley Centre

Ambitious model components

Climate and weather:

- Seasonal component
- Long term climate
- Medium variability (~months/years);
- Short component, weather
- Diurnal range modelled as $\exp(\text{climate}(\mathrm{s}, t))G_{\mathrm{s},t}(\text{Gaussian}(\mathrm{s}, t))$
- Error components

Model representation:

- Space: Subdivided icosahedron;
  ~ 4 levels for local computations, ~ 8 for full model
- Time: 2nd order B-spline basis functions at different scales
  Using this also on the daily scale avoids solar time aliasing effects

THE UNIVERSITY OF EDINBURGH

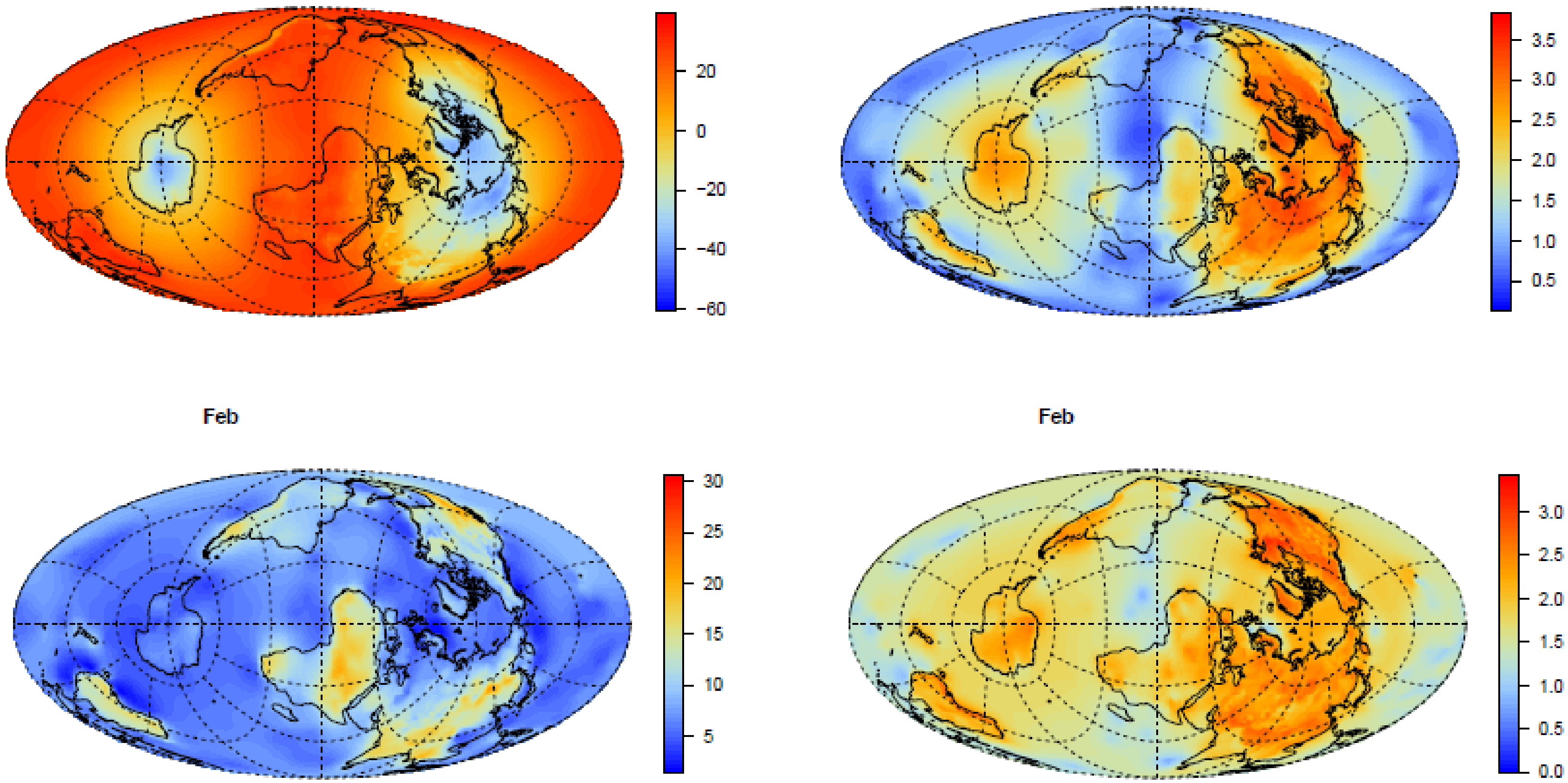
EUSTACE

AVERAGE

STANDARD DEVIATION

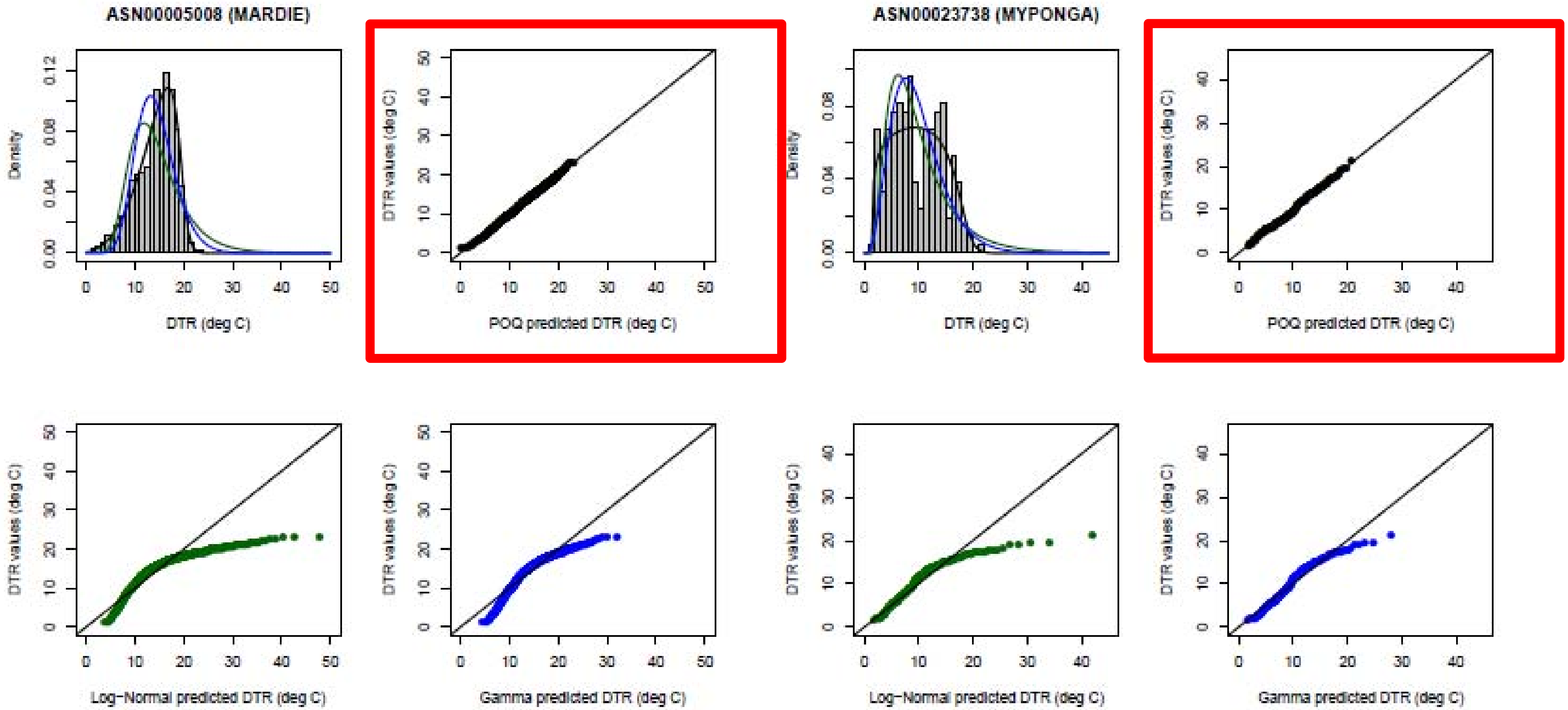
T_{mean}

DTR



EUSTACE

Diurnal range distributions



QUESTIONS

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