Spread of global 2-meter global temperature analyses: disentangling a forecast's systematic errors from mis-estimation of ensemble spread

Tom Hamill

NOAA Earth System Research Lab

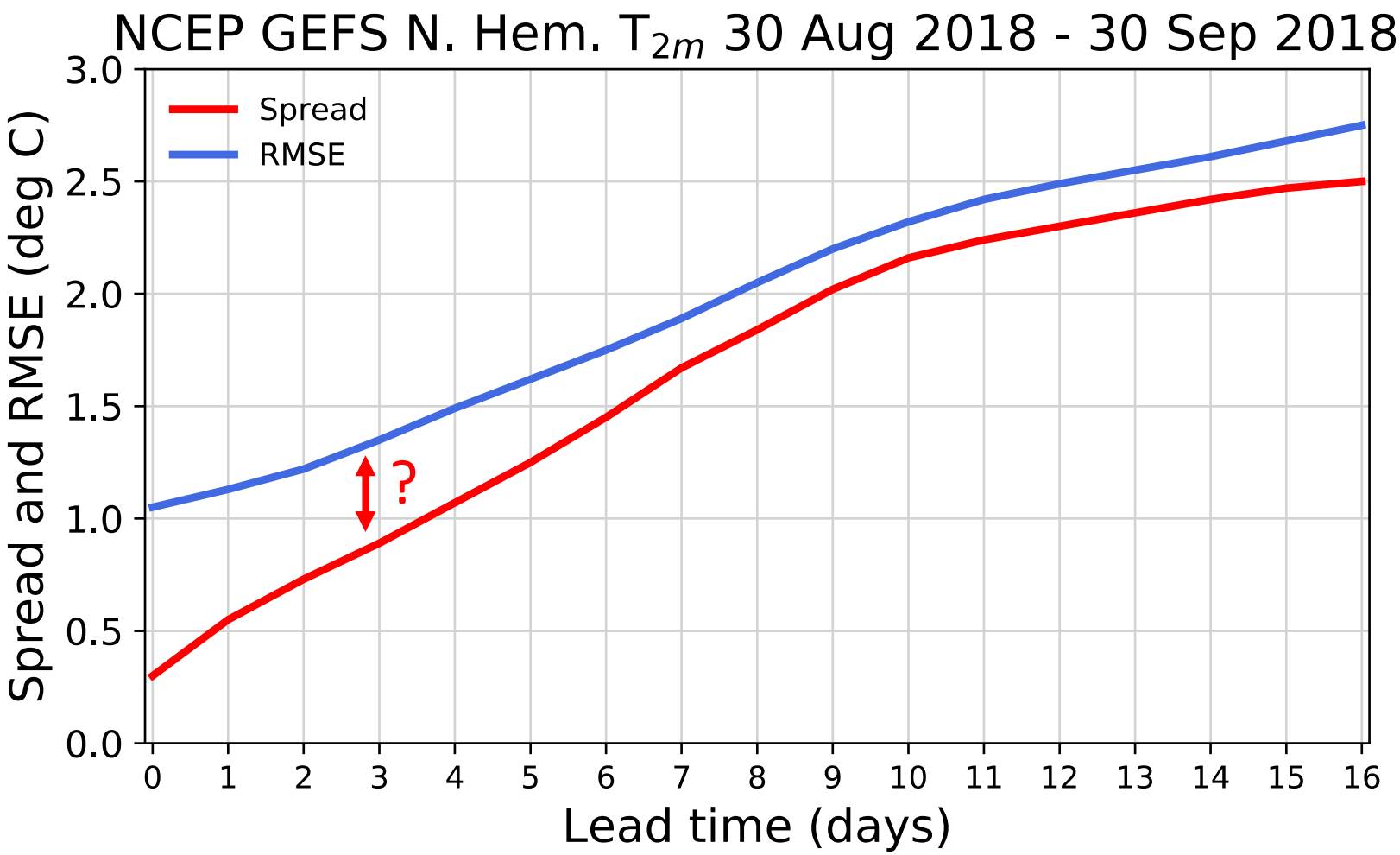
Physical Sciences Division

tom.hamill@noaa.gov

+1 (303) 497-3060

a presentation to the ECMWF workshop on predictability, dynamics and applications research using the TIGGE and S2S ensembles, Mar-Apr 2019

Research question

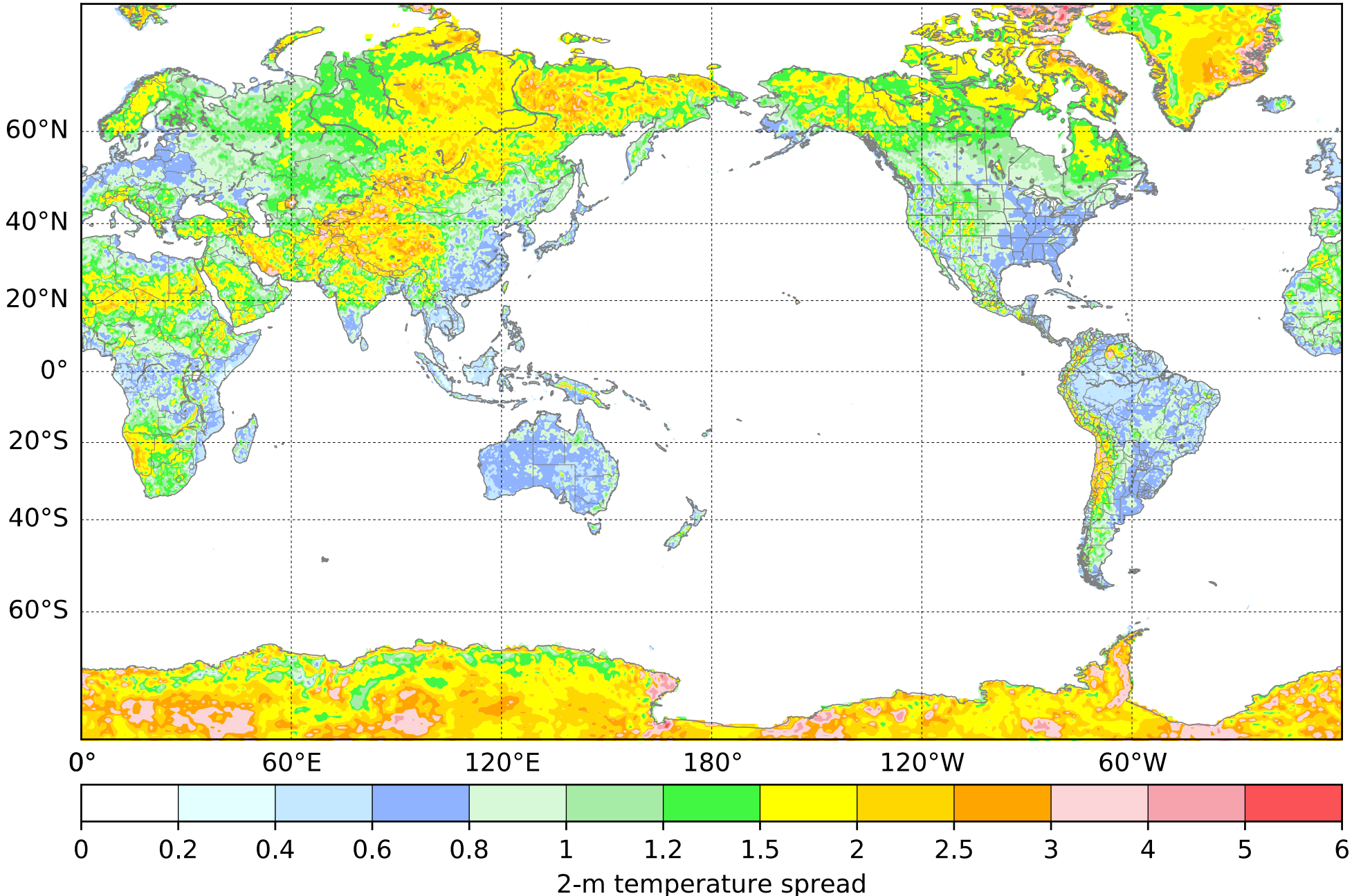


- 2-m temperatures are one of the most under-spread and important forecast variables, e.g. with NCEP system here. This affects the quality of medium-range and S2S forecasts.
- What can we learn about needed improvements in sub-seasonal ensemble prediction systems from an examination of differences between operational 2-m temperature analyses?
- Is it model bias?
- Is it poor ensemble initialization?

Data and methods used in this study.

- 2-m temperature analyses, every day, 00 and 12 UTC, from ECMWF, JMA, and UK Met Office on ½-degree grid, 2018.
  - Why no NCEP, CMC? Data missing in TIGGE.
- I will generate a multi-model ensemble mean which I then regard as an estimate of the true state (big assumption, perhaps wrong).
- Spread of analyses with respect to the instantaneous mean analysis value was calculated every day at every model grid point, then averaged over many days.
- Later, a proposed decomposition of this spread to disentangle systematic and random errors.

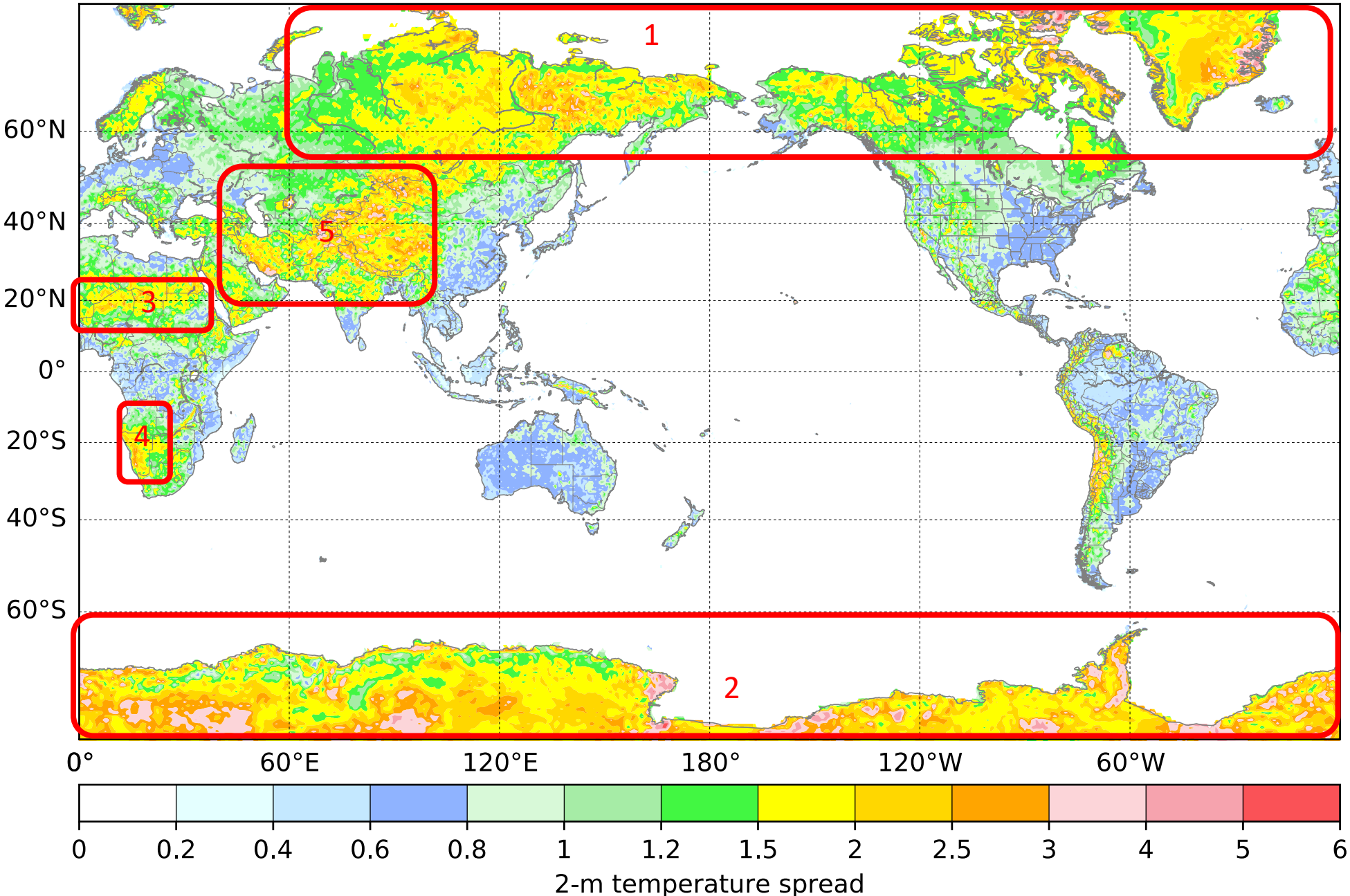
00 UTC ECMWF 2-m temp spread with respect to multi-model analysis daily mean



Much of the variability appears to be related to **mountainous areas**, where there may be differences between the various orographic data used in each prediction system.

But there are other areas.

00 UTC ECMWF 2-m temp spread with respect to multi-model analysis daily mean



Areas with overall large T_{2m} spread:

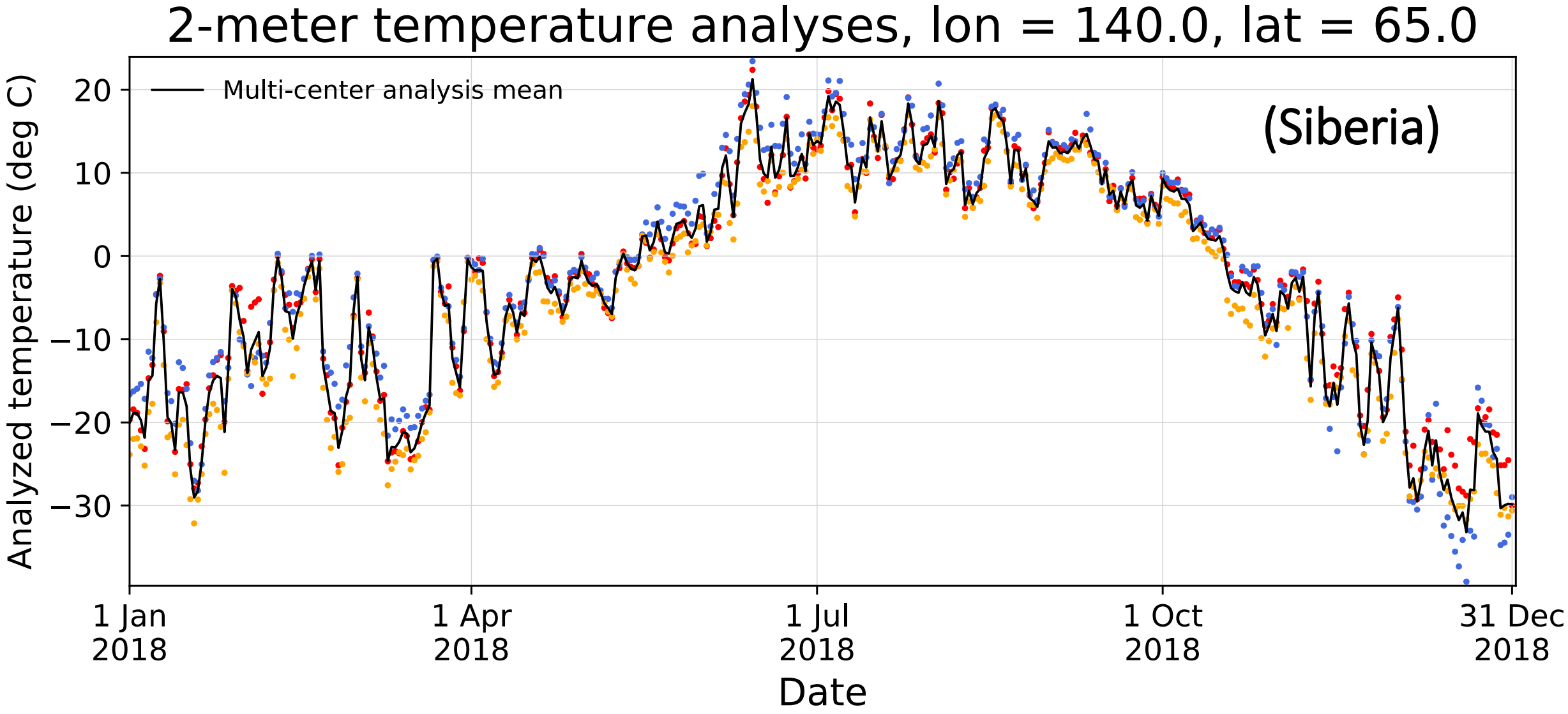
1. Tundra, Greenland.
2. Antarctic.
3. Sub-Saharan Africa.
4. Namibia
5. Middle East thru Kazakhstan, Tibet

How much spread is systematic error, how much reflects initial condition uncertainty?

Any dependence on winter vs. summer?

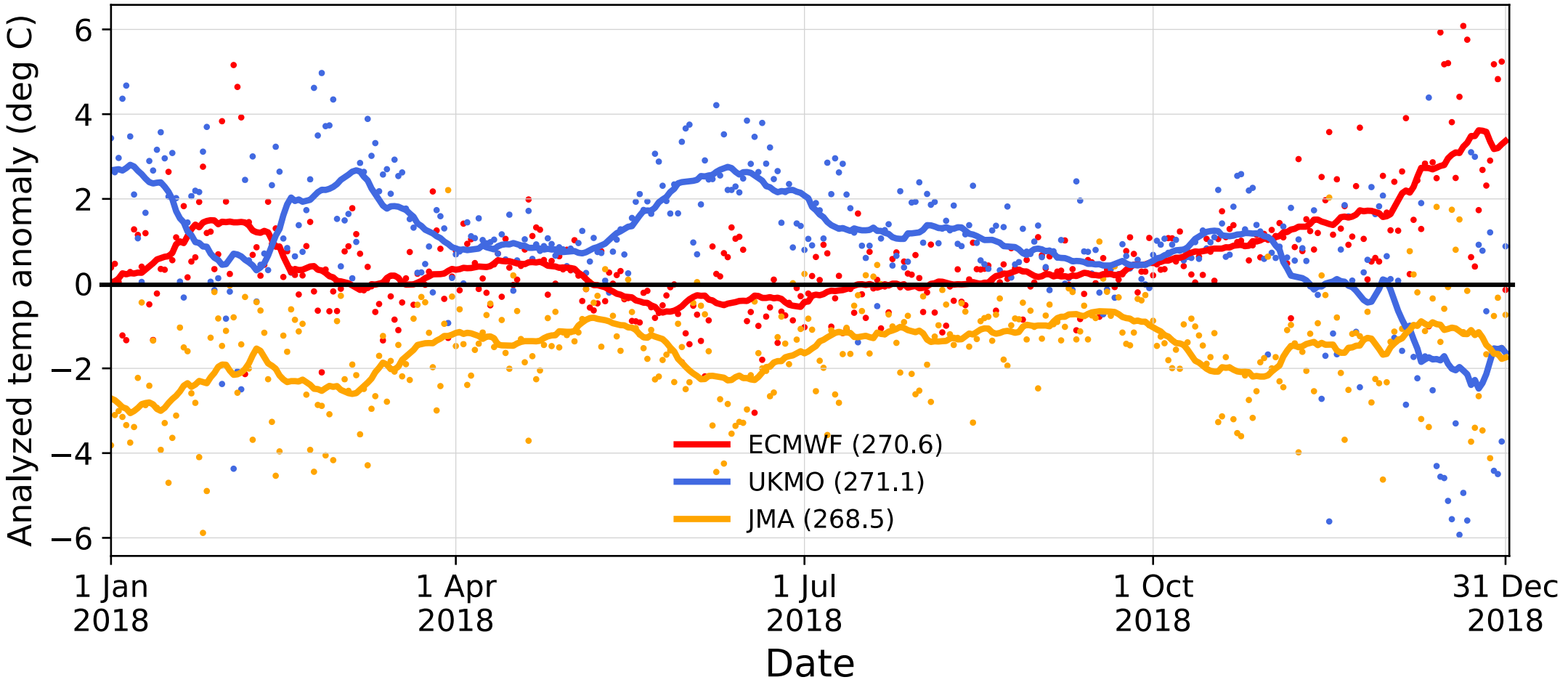
Is spread between global surface temperature analyses a reasonable estimate of initial condition uncertainty?

- No. 2-m temperature errors in analyses have **systematic** as well as **random** component.
  - Systematic → fix the underlying model deficiency (or vertical interpolation error).
  - Random error mis-estimated → fix the ensemble initialization.
- Can we then make some intelligent guesses at what is systematic and what is random, informed by differences between TIGGE analyses?



The procedure starts with subtracting the mean analysis in black.

2-m temp. analysis perturbations from mean analysis, lon = 140.0, lat = 65.0

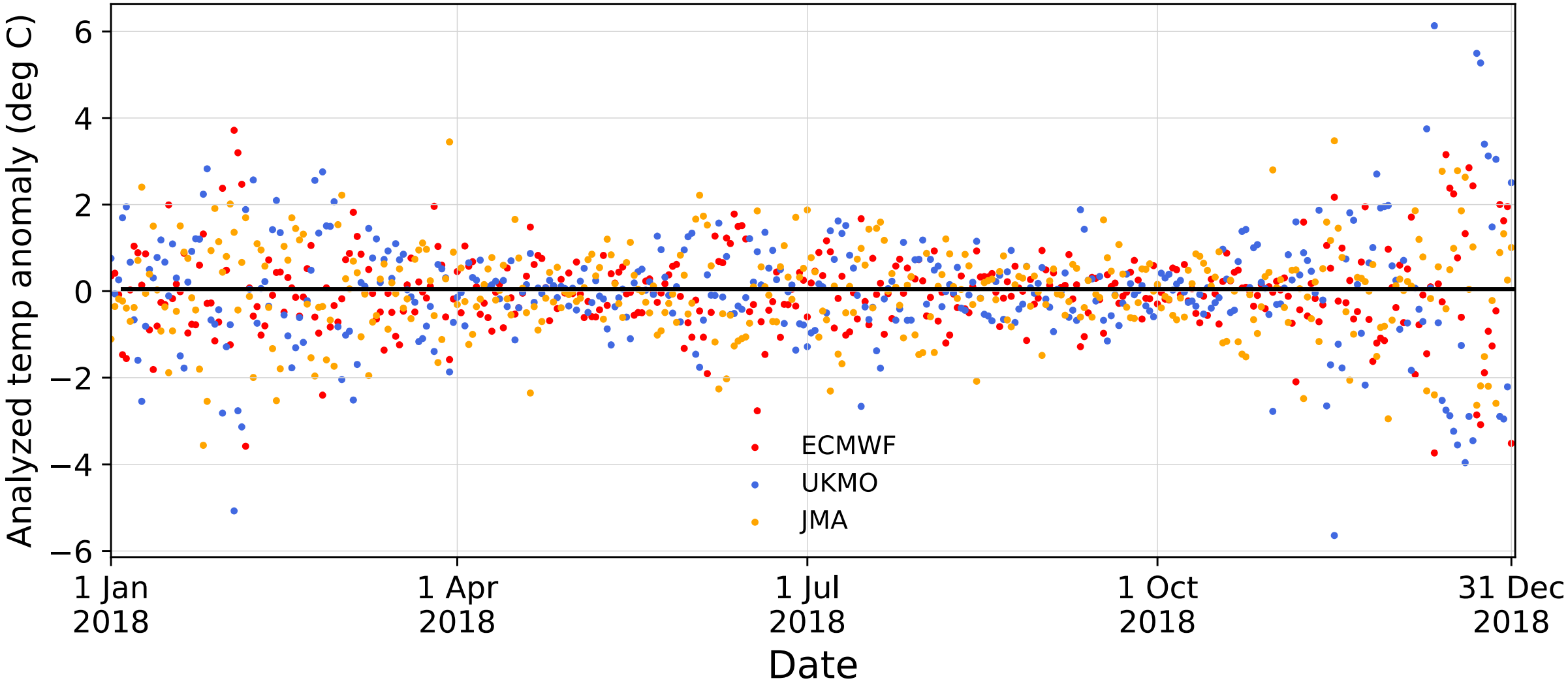


colored lines: biases may exist in the analyses,e.g., contributed from biased background estimates.

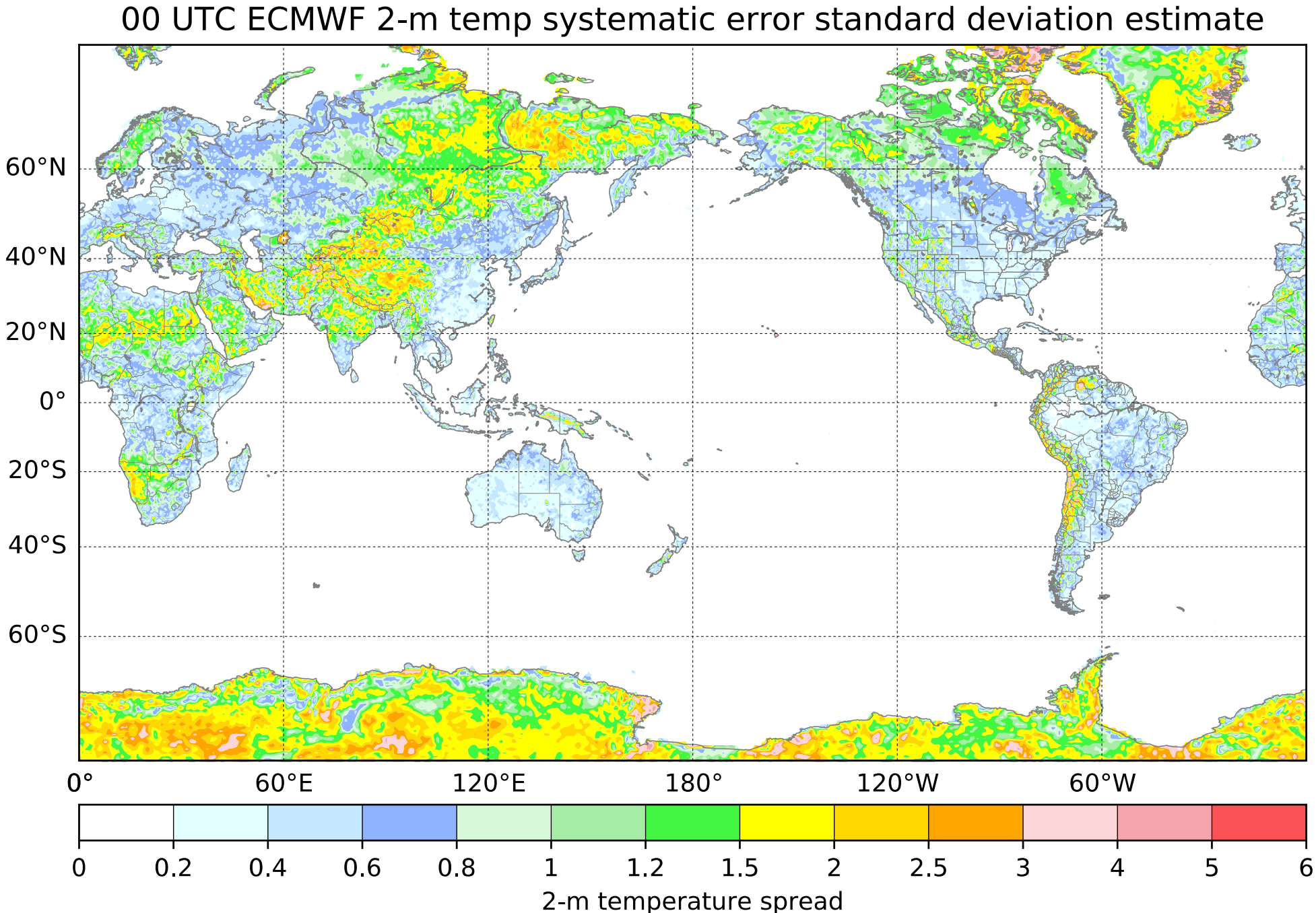
(1) Assume differences of mean from zero represent an estimate of the systematic error → fix the model.

(2) Differences of dots with respect to running mean an estimate of the random error → check the ensemble initialization.

2-m temp. initial-condition uncertainty, lon = 140.0, lat = 65.0

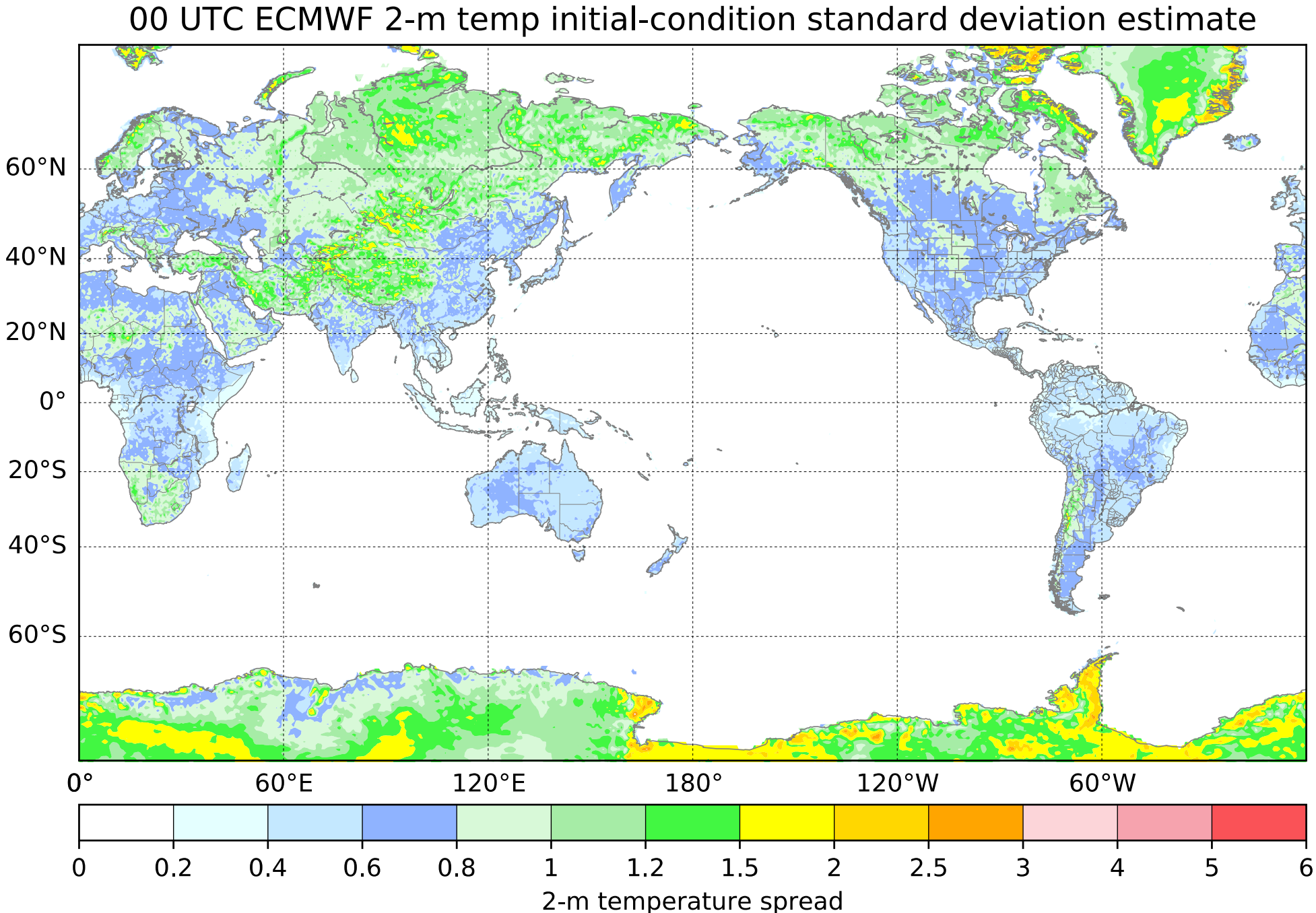


This presents each centre's differences with respect to each centre's running time mean, an estimate of initial condition uncertainty.



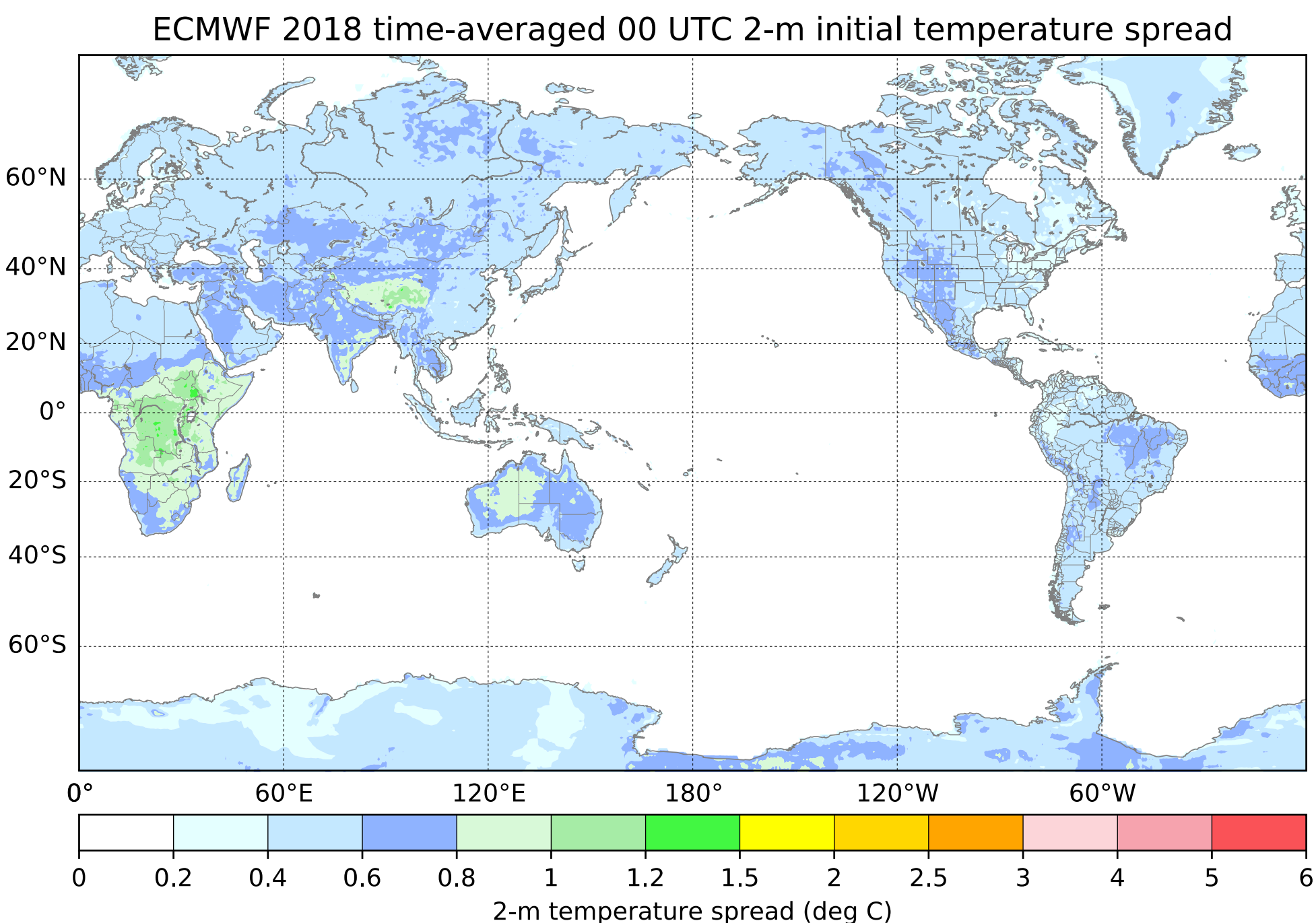
ECMWF 2018 T_{2m} systematic error estimate

Sub-Saharan Africa, Namibia, Arctic tundra, Greenland, Antarctica are fruitful areas to examine further for systematic error.

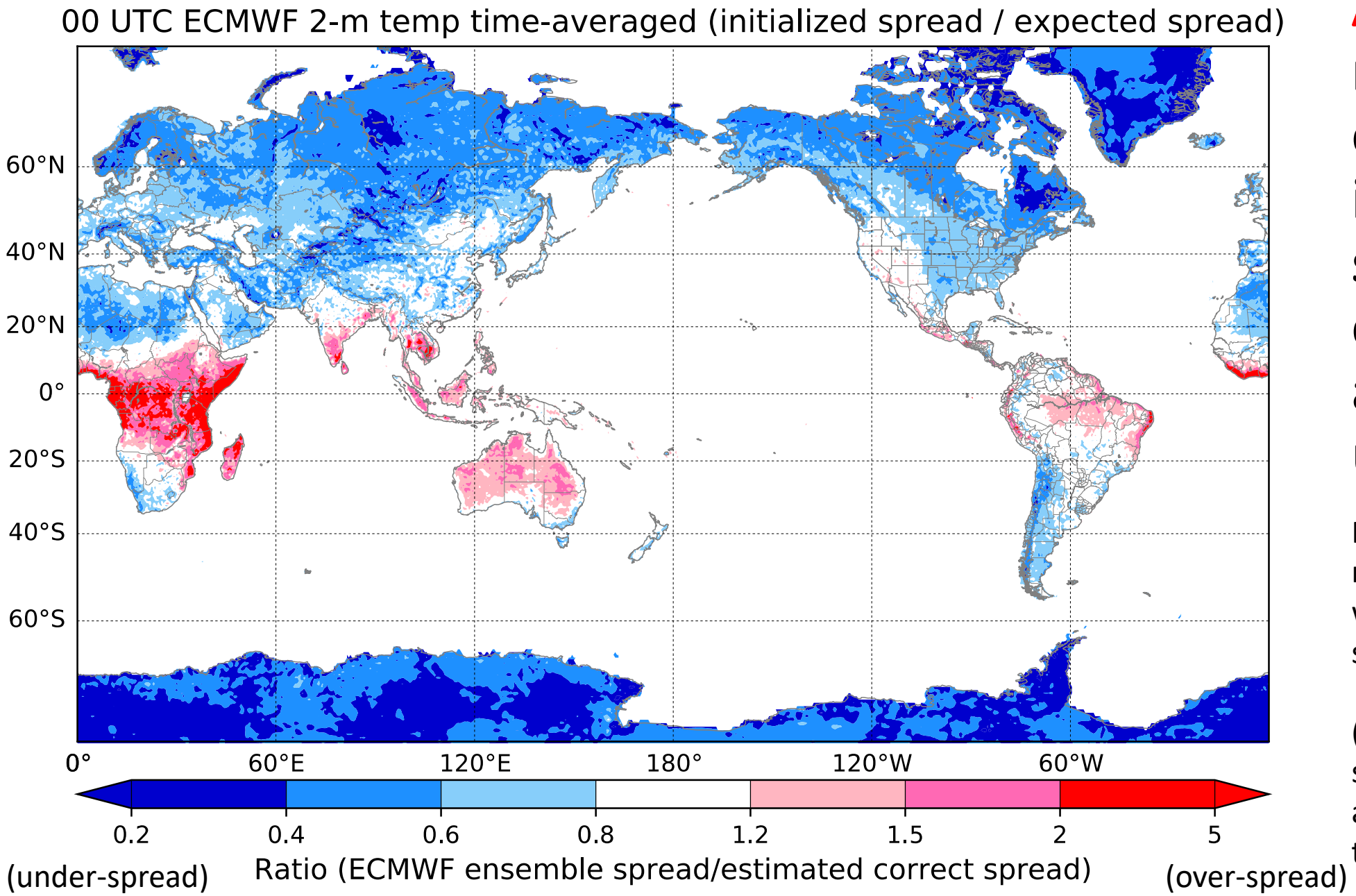


Initial condition uncertainty estimated from TIGGE residuals.

Let's compare this with the time average of ECMWF 2-m temperature perturbations.



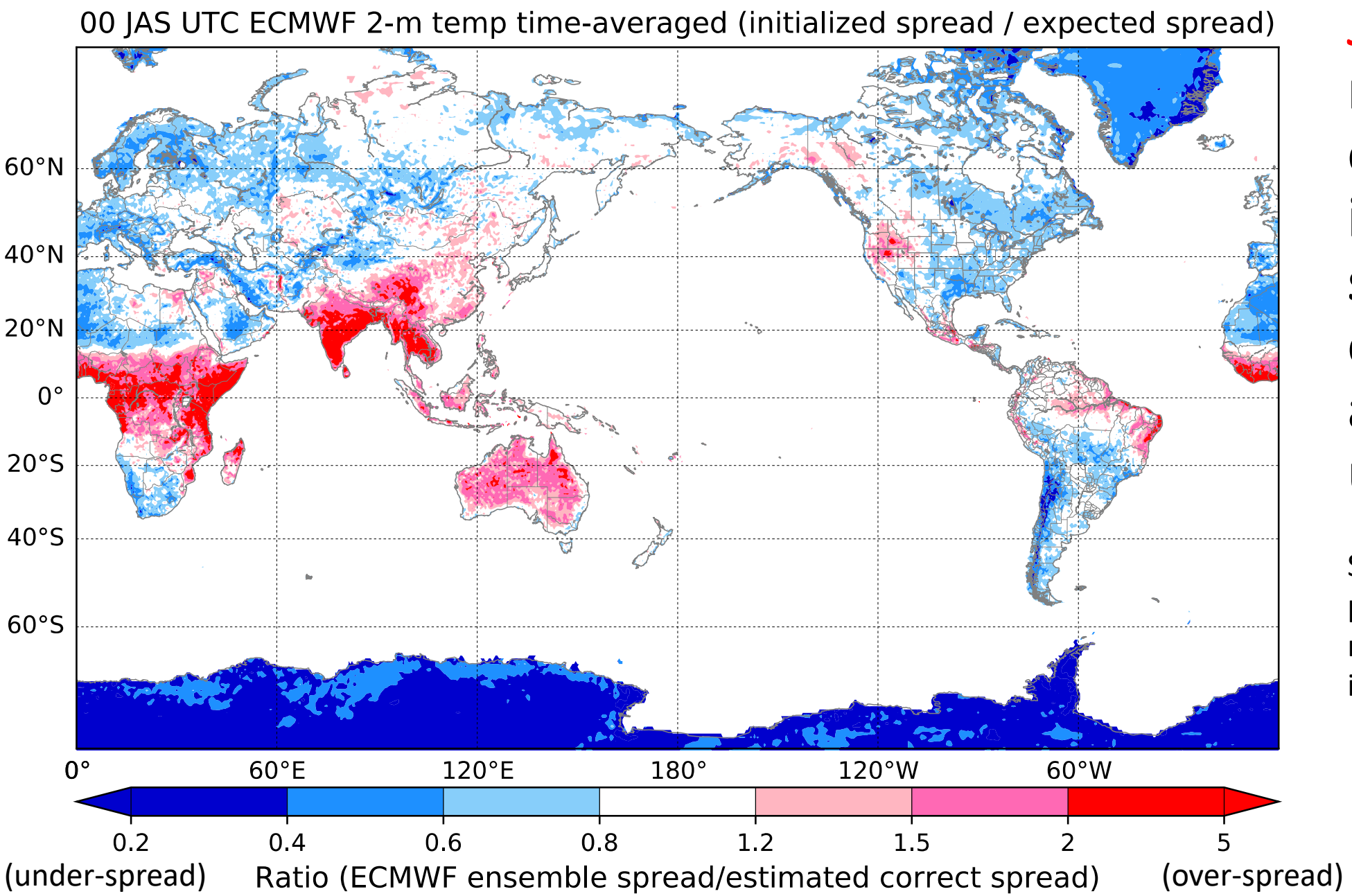
Time-averaged spread of ECMWF 2-m temperature initial perturbations



All of 2018: ratio of ensemble-initialized spread to estimated analysis uncertainty.

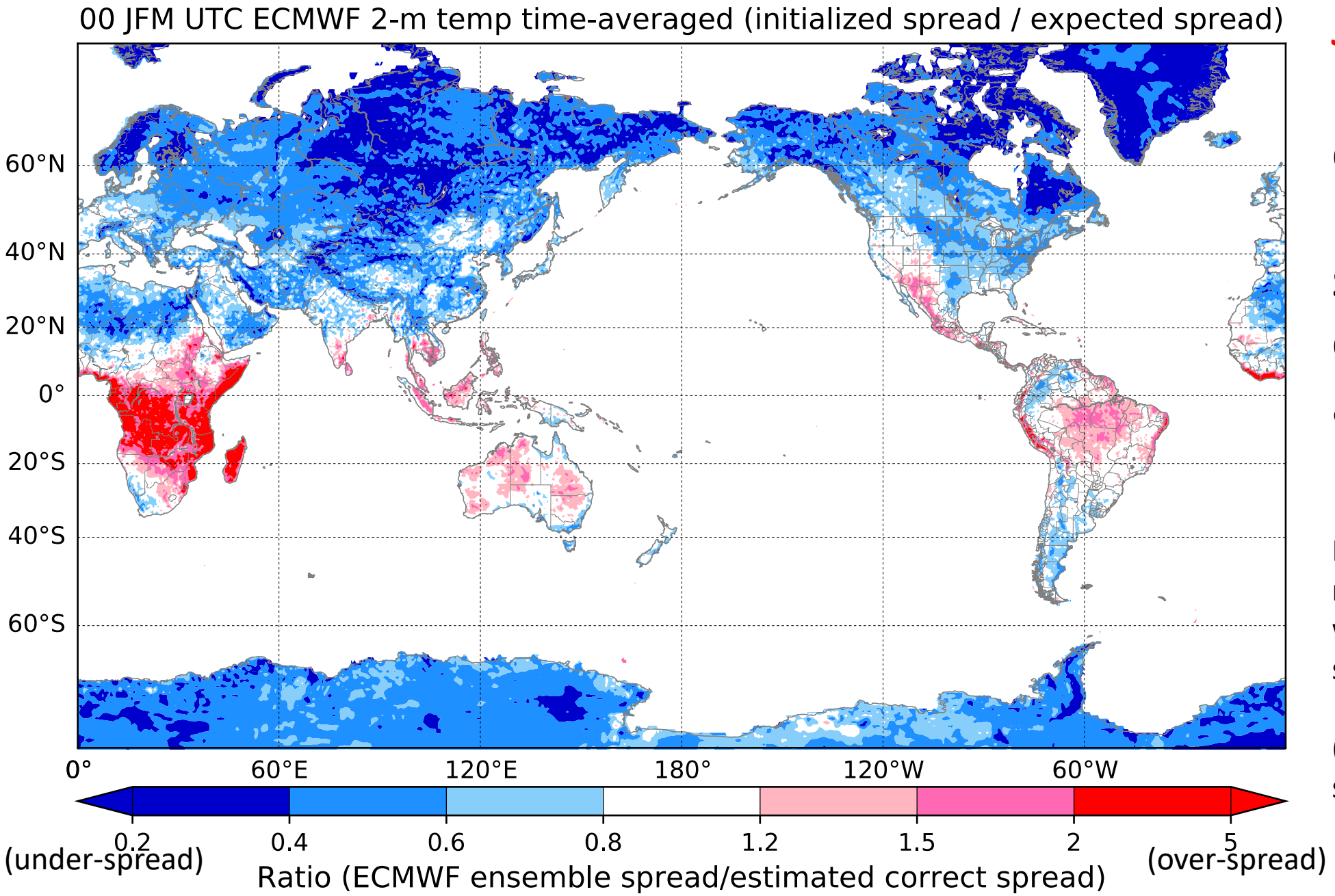
It appears snowy regions are areas where initial ensemble spread should be larger

(and probably land-surface temperature and moisture spread too)



Jul-Aug-Sep ratio of ensemble-initialized spread to estimated analysis uncertainty.

Smaller initialization problems through much of N. Hem. in summer.



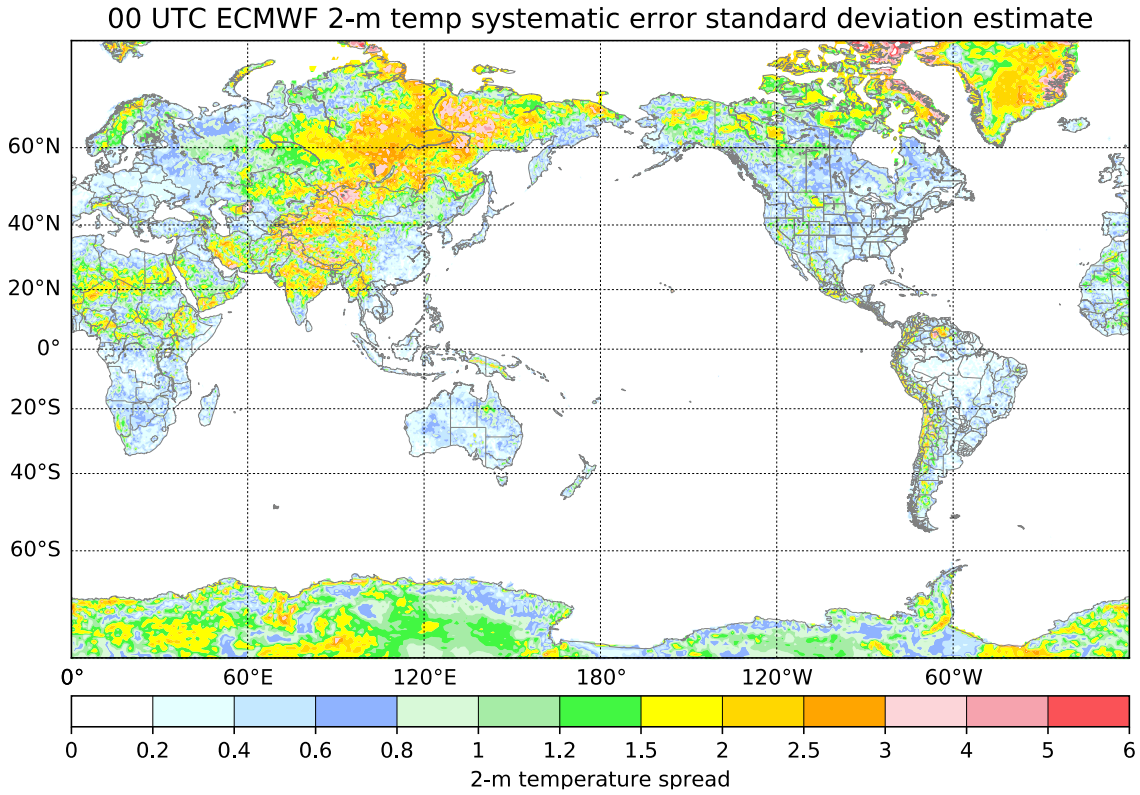
Jan-Feb-Mar ratio of ensemble-initialized spread to estimated analysis uncertainty.

It appears snowy regions are areas where initial ensemble spread should be larger

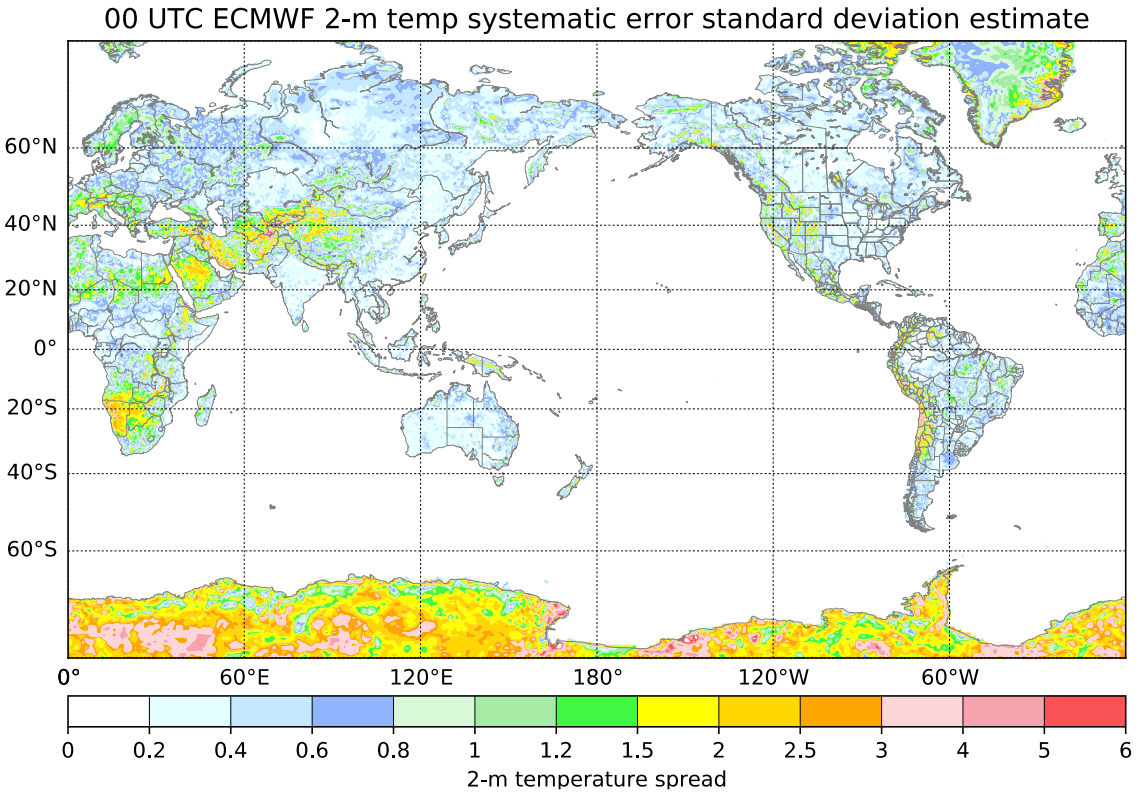
(& soil temp, moisture, snow cover)

Systematic errors of T_{2m}: seasonal dependence?

Jan – Feb – Mar 2018



Jul – Aug – Sep 2018



Notable potential systematic errors in ECMWF surface temperature initialization:

(1) Siberian wintertime;

(2) Namibia in S. Hem winter;

(3) Antarctica, especially in winter.

Discussion and conclusions

- The results presented here depend on the assumptions of:
  - A reduced error in ensemble-mean analysis relative to each individual centre.
  - A particular centre's time-averaged mean difference from the multi-model mean is representative of the systematic errors.
- ECMWF's initial surface temperature spread is too small in wintertime high latitudes.
- Possible systematic errors in wintertime Siberia, Greenland, Antarctica, Namibia.
- How to fix the initialization? Remember that spread of 2-m temperatures will not persist in forecast unless underlying land surface state also has realistic spread.
  - Soil temperature.
  - Soil moisture.
  - Snow cover.
  - → Strongly coupled land-atmosphere ensemble DA?
- Fixing the systematic errors in winter polar regions: other diagnostics might be necessary to isolate the specific causes such as cloud microphysical parameterization or boundary layer.