

Demonstrating scales of predictability at tropical and middle latitudes

Richard J. Keane^{1,2}, D. J. Parker^{2,3}, E. Dunn-Sigouin³, E. W. Kolstad³, J. H. Marsham²
¹Met Office, Exeter, UK; ²Leeds University, UK; ³NORCE, Bergen, Norway

INTRODUCTION

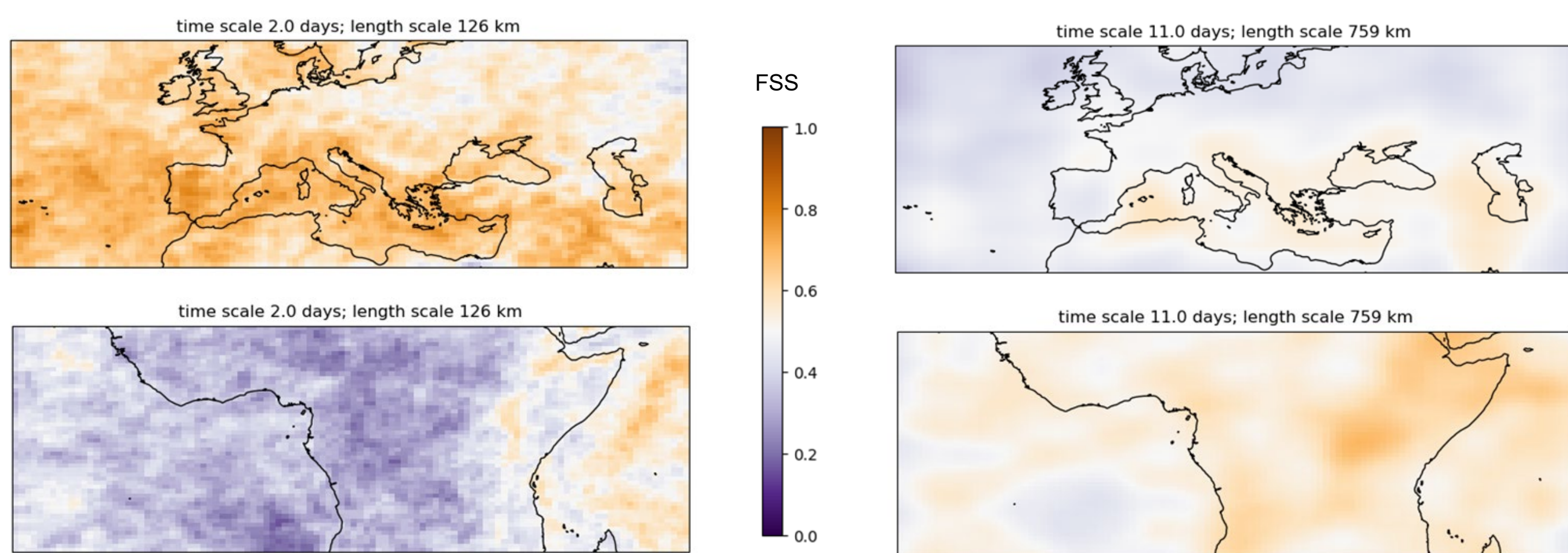
- Weather forecasts generally perform worse in the tropics than at mid-latitudes.
- However, seasonal forecasts generally perform better in the tropics.
- Here we quantify this systematically for a range of forecasting systems.
- As well as being an important scientific exercise, this provides crucial information for developing forecast technologies in tropical regions.

METHOD

- The widely-used fractions skill score (FSS) is applied to verify the forecasts.
- This provides a way of evaluating forecasts at different spatial scales.
- Threshold-based binary precipitation fields are averaged over different numbers of grid points (and ensemble members for GloSea).
- These are compared between observations (IMERG data regridded to the model grid) and model forecasts.
- Scores range from 0 (no skill) to 1 (perfect forecast **at the relevant scale**).
- Percentile thresholds are applied, based on climatology of the verification period, separately for each grid point.
- Five modelling systems are evaluated:
 - Met Office seasonal hindcasts (GloSea6, 2001—2016).
 - Met Office (2020) and ECMWF (2022) operational weather forecasts.
 - Two sets of hindcasts run at ECMWF for 2022 using machine-learning-based models FuXi and GraphCast.
- Verification is carried out over tropical Africa and over Europe.

RESULTS: GloSea

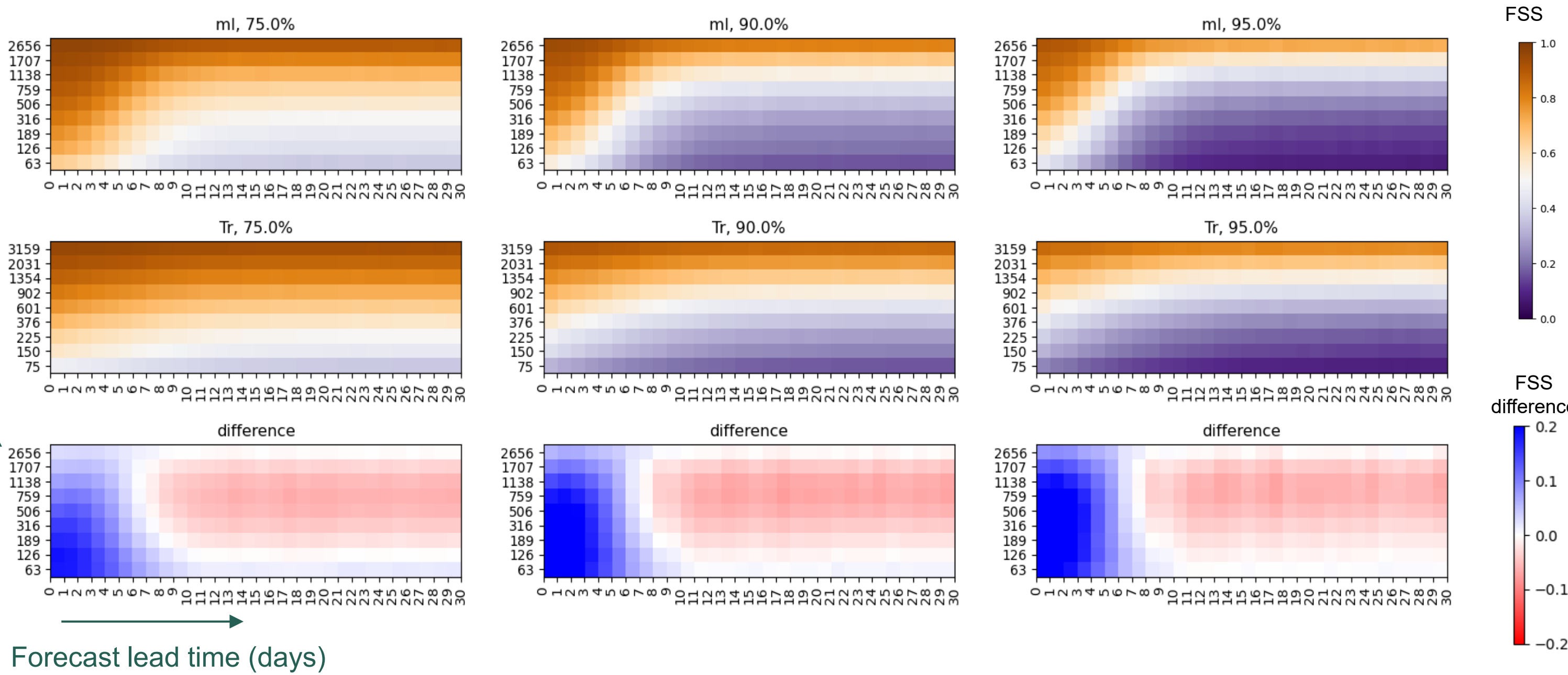
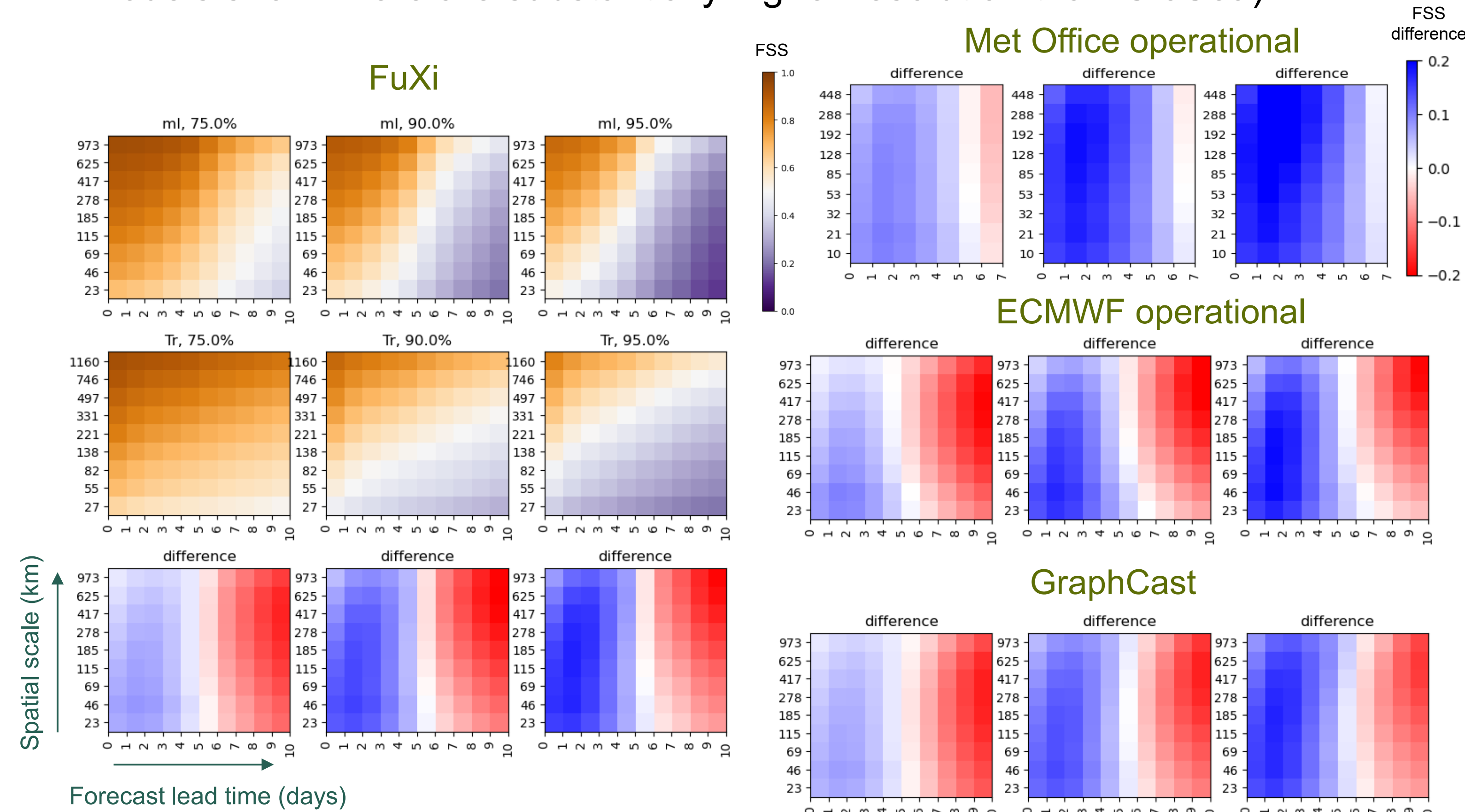
- Spatial FSS maps for GloSea show that it performs better on shorter/smaller scales at mid-latitudes and on longer/larger scales in the tropics.



RESULTS: Other models

- Only the difference panels are shown for three of the models.
- All the models show similar behaviour to GloSea, suggesting that it is persistent across model scale and model construction method, and holds for both physical and machine-learning-based models.
- The 'crossover' time scale is slightly shorter than for GloSea, at 4 to 7 days.
- Could this have a systematic dependence on grid spacing? (The four models shown here are substantially higher resolution than GloSea).

- This is investigated systematically by plotting FSS values as a function of forecast lead time and spatial scale.
- Spatial scale is based on multiplying FSS number of grid points by the average grid spacing (slightly different for the two domains).
- Difference plots show mid-latitude minus tropics.
 - Tropical FSSs, as a function of spatial scale, are interpolated from the tropical scales to the mid-latitude scales for a direct comparison.
 - Blue means higher skill in mid-latitudes.
 - Red means higher skill in tropics.



CONCLUSIONS

- The results confirm that rainfall forecasts are worse in the tropics than at mid-latitudes on shorter scales but better on longer scales.
- This could be due to inherent predictability differences.
- Worse performance at shorter scales could also be due to less advanced observational infrastructure.
- The results hold for a range of models, including machine-learning-based models.
- This suggests that it is a fundamental property of the atmosphere rather than an artifact of model construction.

- Forecasts are overall better on larger spatial scales and shorter lead times (as would be expected).
- The reduction in skill with lead time is slower in the tropics.
- Performance is better in the tropics at longer lead times and larger spatial scales.
- Performance is worse in the tropics at shorter lead times and smaller spatial scales.
- Performance varies more strongly with temporal scale than spatial scale.
- There is a 'crossover' in relative performance at 6 to 9 days (dependent on the spatial scale).

IMPLICATIONS

- Forecasting methods that work well at mid-latitudes may not be the best solution for forecasting at tropical latitudes.
- There may be fundamental constraints on predictability, varying with latitude, that cannot be overcome using machine-learning models.