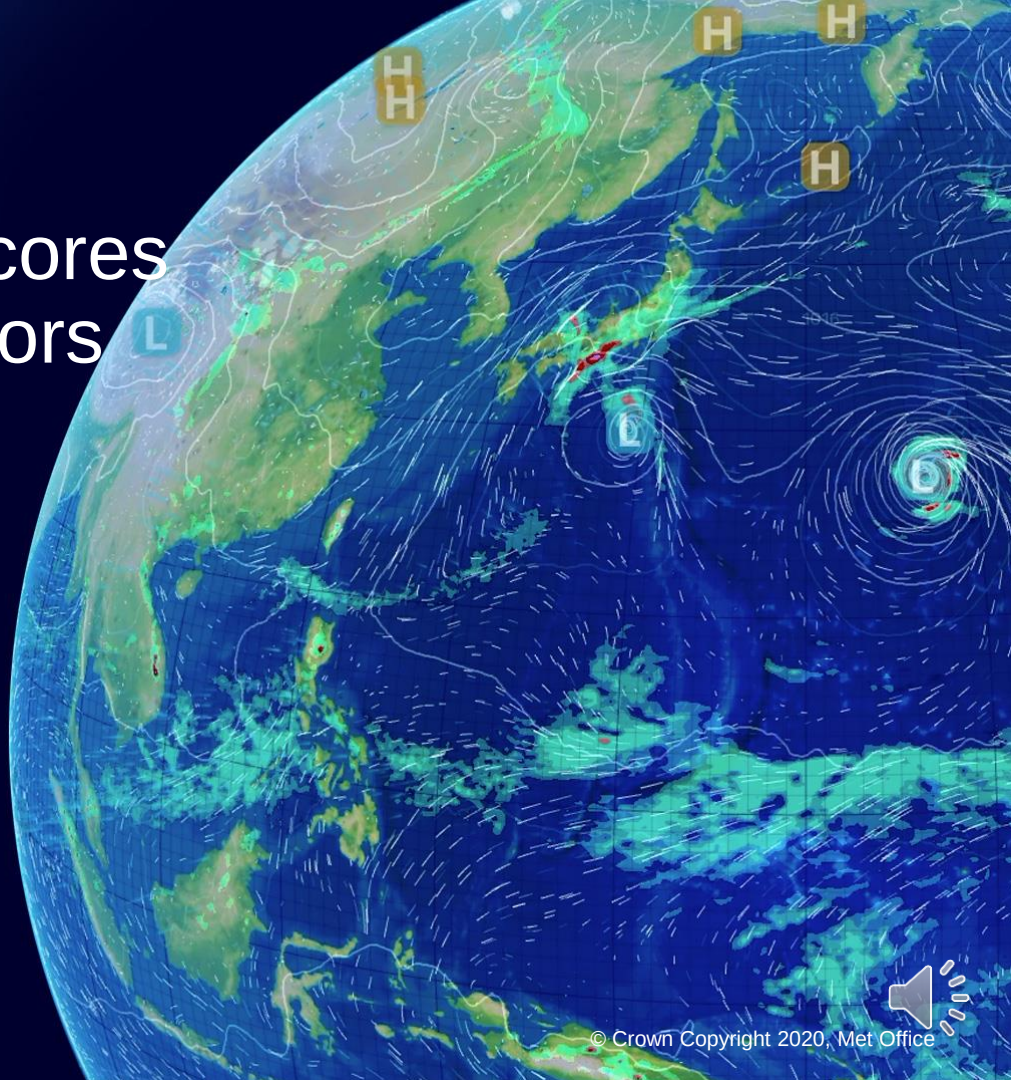
Using diagnostics from calculating verification scores to identify systematic errors

Rachel North, Met Office, UK

Marion Mittermaier, Met Office, UK

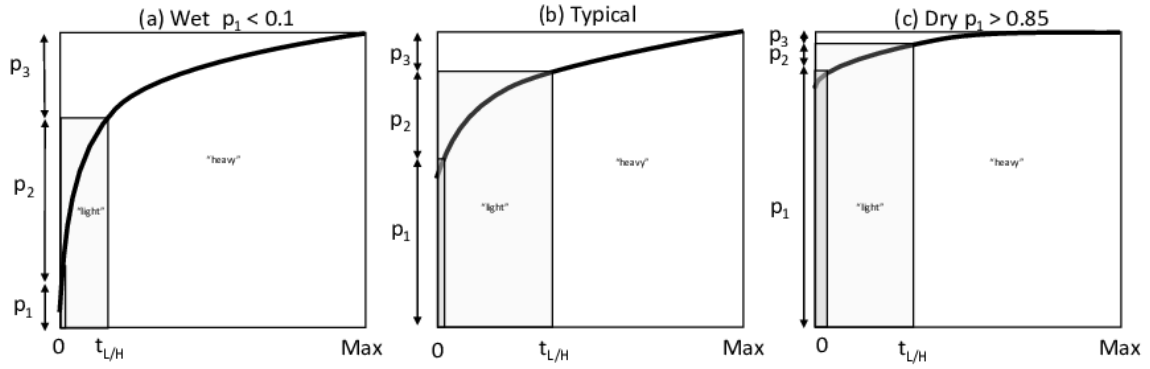


Overview

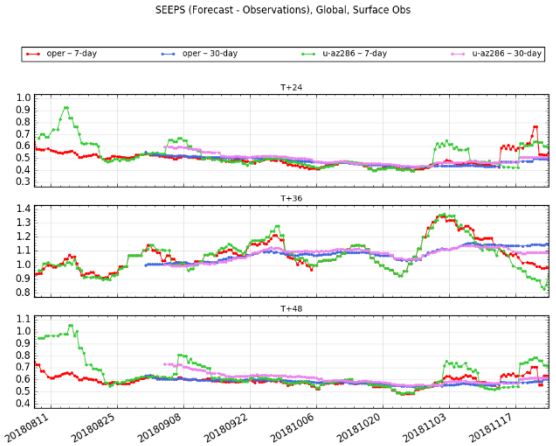
- One slide recap of SEEPS
- Creation of a satellite-based climatology for use calculating SEEPS
- Diagnostics from SEEPS calculation and their application in evaluation of an NWP model upgrade candidate
  - Timescales
- Questions

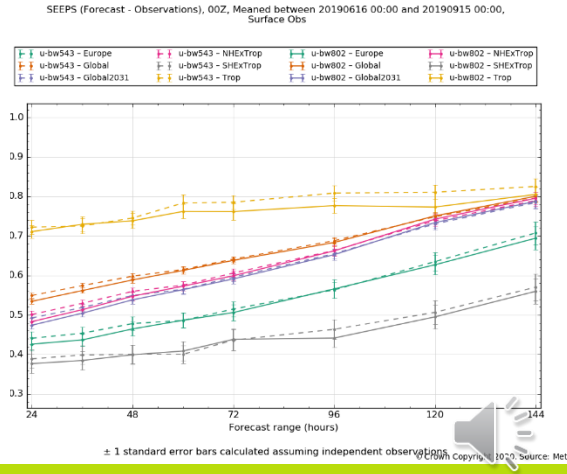
Stable Equitable Error in Probability Space (SEEPS)

- Recap of SEEPS
  - A verification metric (detail in Rodwell et al., 2010) which can be used as part of ongoing monitoring activities, which was developed for QPF
  - Uses 3 categories; dry, light and heavy
  - A climatology derived at station locations accounts for local variations in QPF and the score rewards forecasts which predict a full range of possible outcomes using locally appropriate thresholds
  - Can then be aggregated spatially due to having accounted for local variations in climate



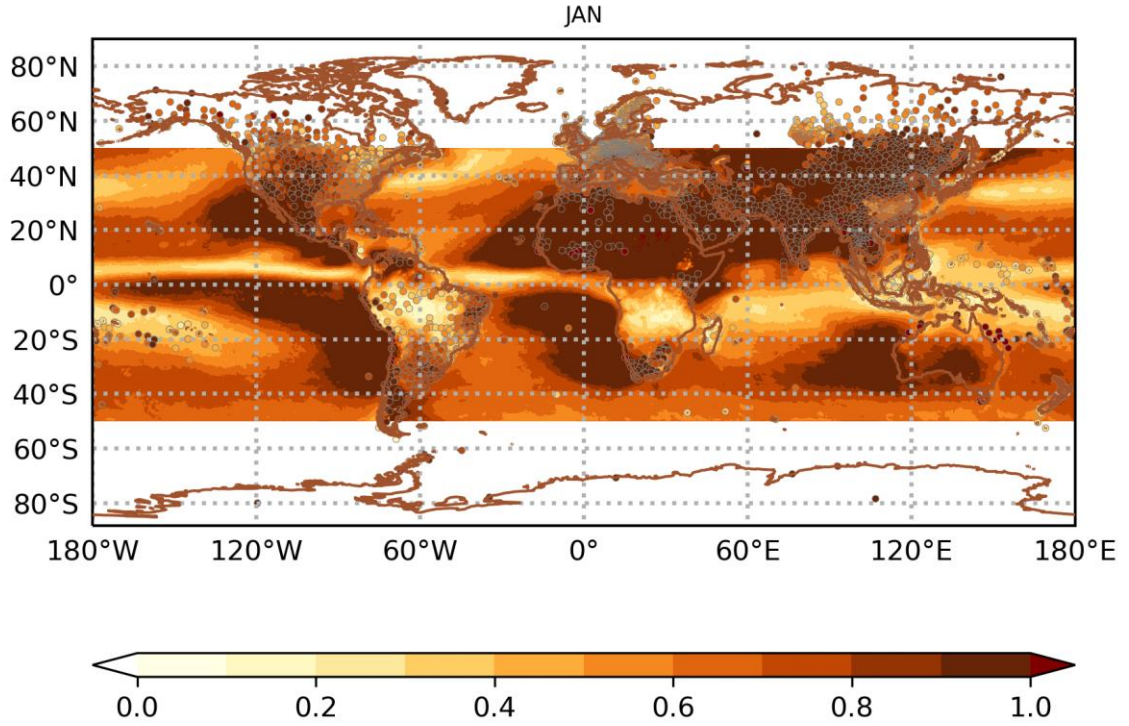
Observed precipitation





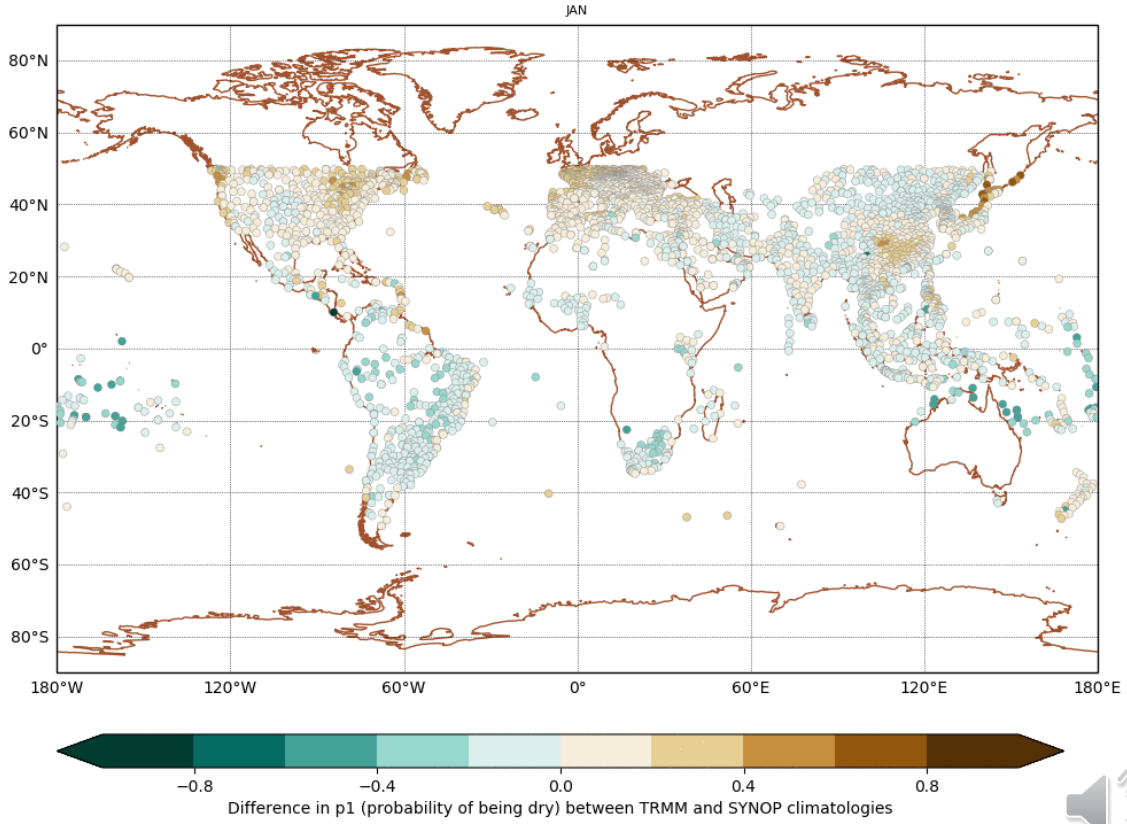
Satellite-derived climatology for use with SEEPS

- TRMM 3B42 v7 research product
- Jan 1998 to Jun 2015
- For daily accumulations, ending at 00UTC
- Calculated probability of being dry, light precipitation and heavy precipitation at each satellite grid point (~25km footprint)
- **Probability of being dry**
- Comparison with gauge climatology (1980-2009)



Satellite-derived climatology for use with SEEPS

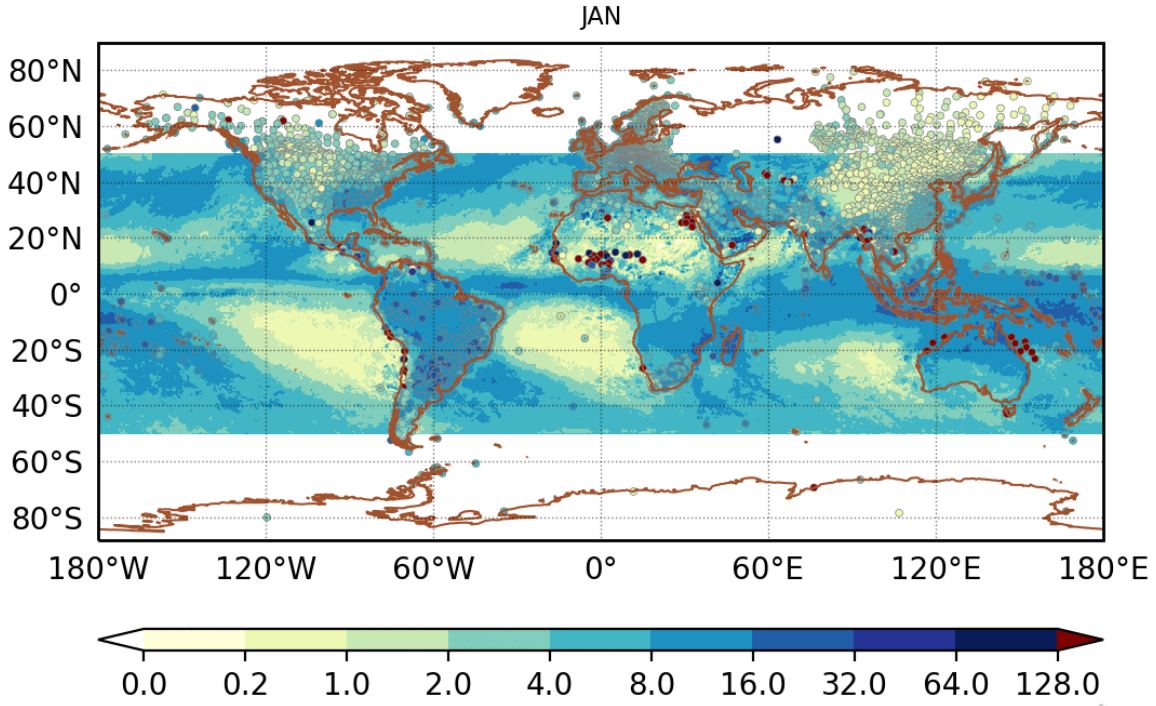
- Differences at gauge locations
- Interpolation (nearest grid point from TRMM)
- Green => TRMM wetter
- Brown => TRMM drier



Met Office

Satellite-derived climatology for use with SEEPS

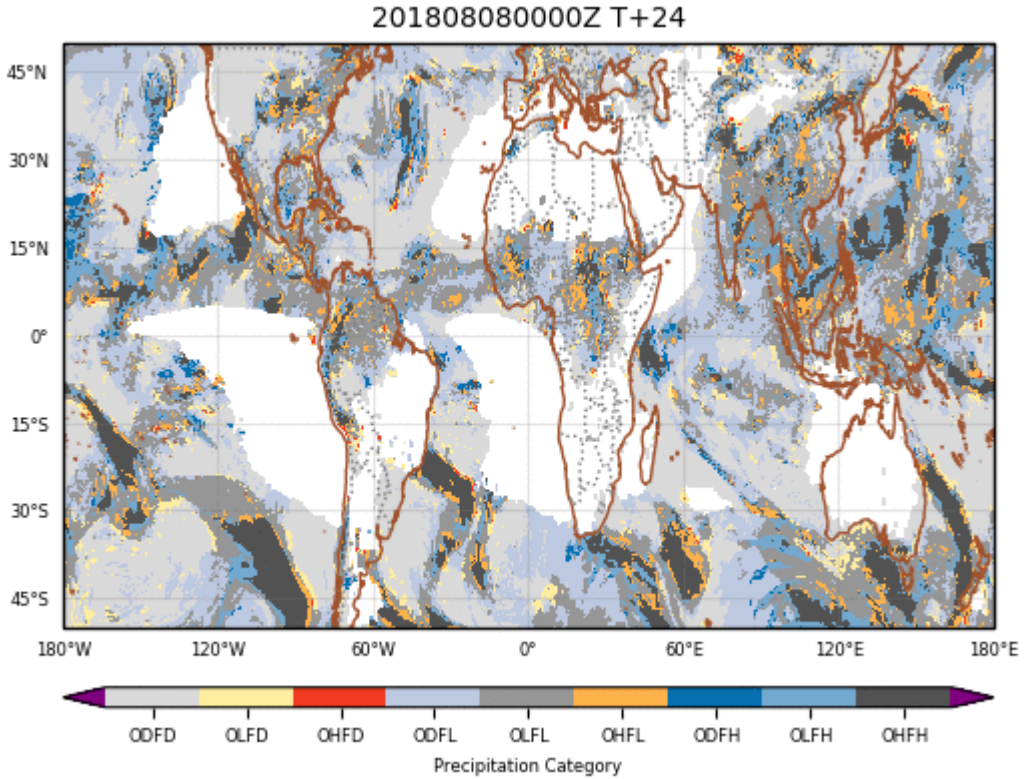
- From daily accumulations ending at 00UTC
- TRMM 3B42 from Jan 1998 – Jun 2015
- **Light-heavy threshold**
- Comparison with gauge climatology (1980-2009)



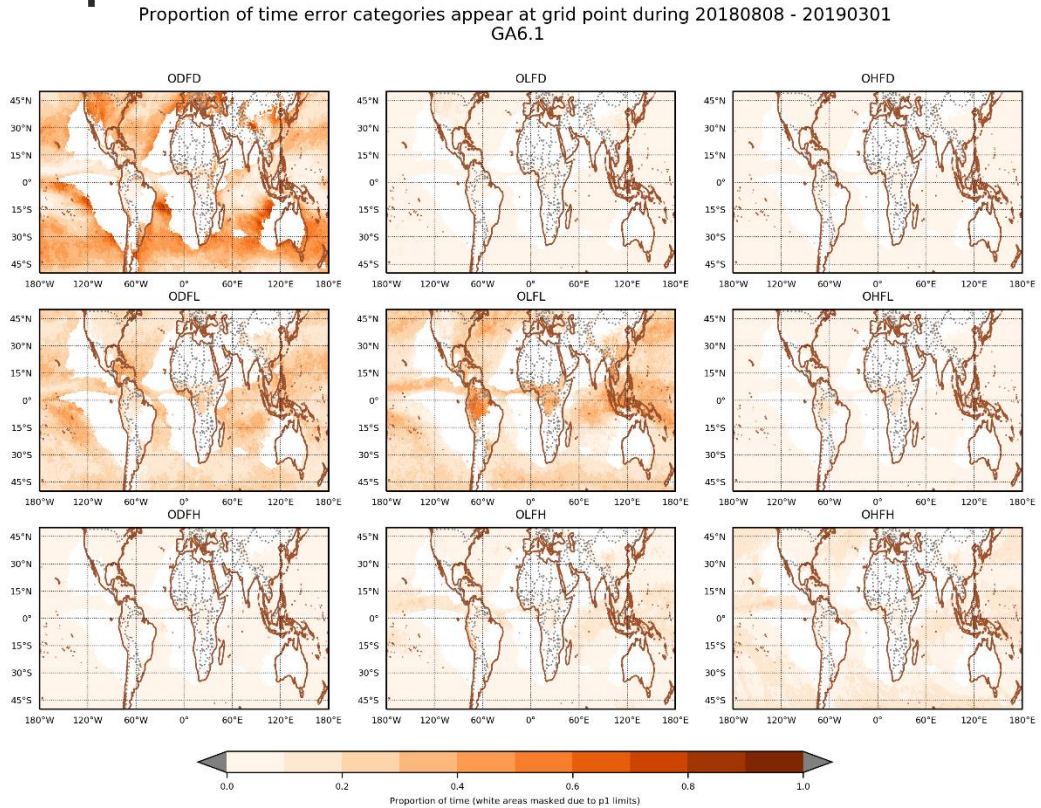
Diagnostics produced as part of the SEEPS calculation

GA6.1 Day 1 error categories

- Daily errors
- White regions masked due to p_1 limits
- Precipitation error categories determined for both observed and forecast accumulations
- Grey shades => correct categories!

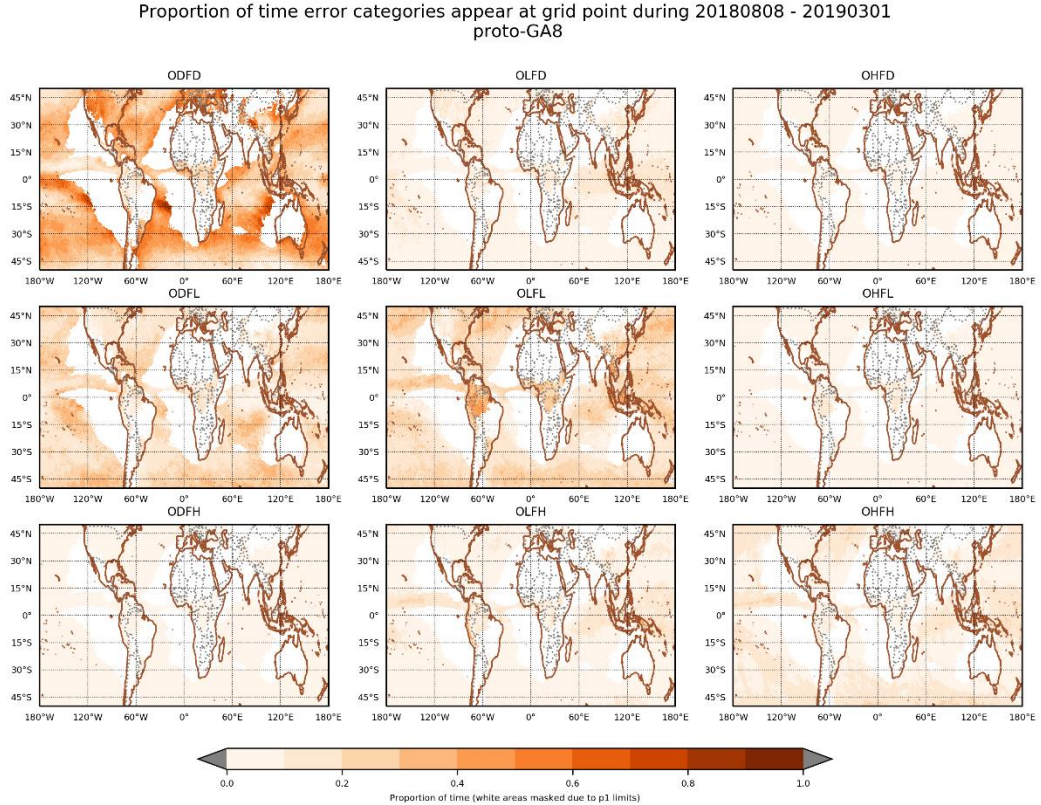


Proportion of Errors over time - operational



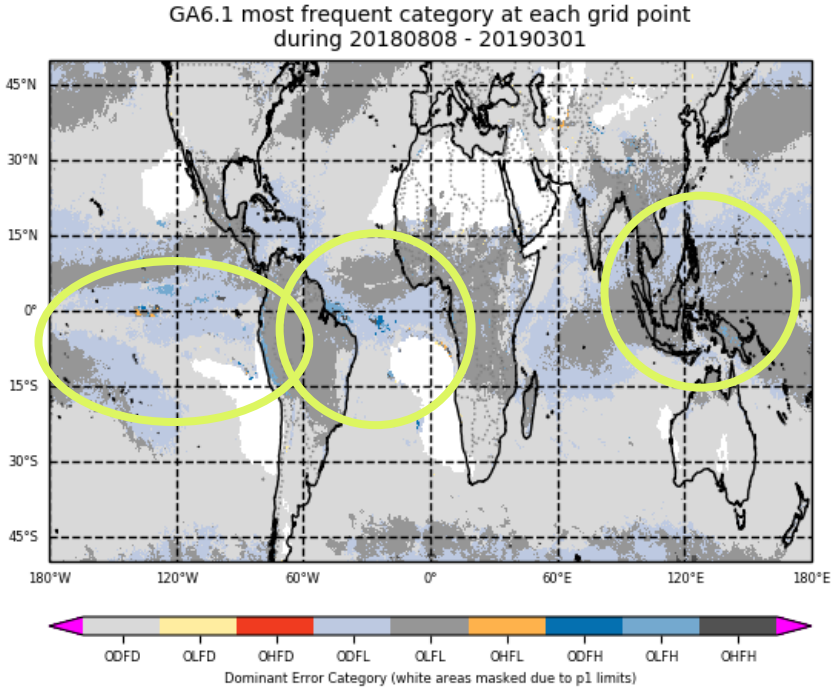
- Observed dry, light, heavy
- Forecast dry, light, heavy
- Used with error matrix (Rodwell et al. 2010) to calculate SEEPS
- Useful to map these – use of satellite-derived climatology => values over sea
- First example, proportion of time each category appears at each grid point (over full analysis period

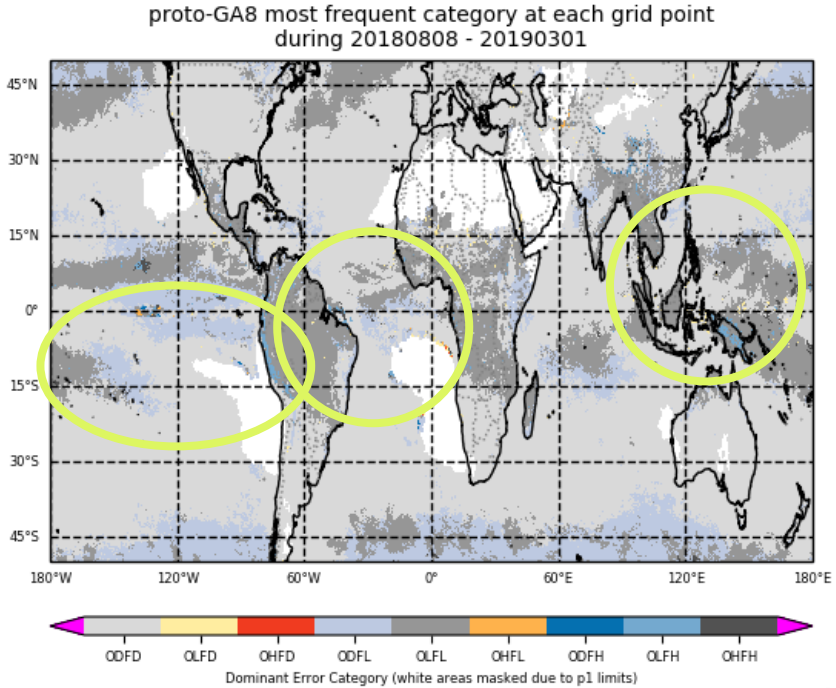
Proportion of Errors over time – early GA8 candidate



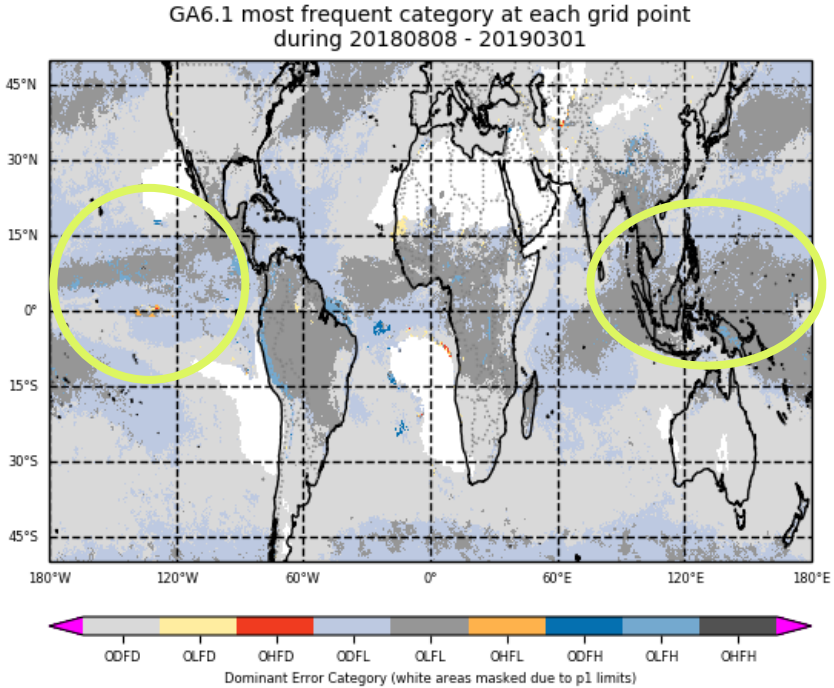
- Proportion of time each error category is seen at each grid point
- Across whole period of study (Aug 2018 – Feb 2019)
- Early GA8 candidate, day 1 accumulations

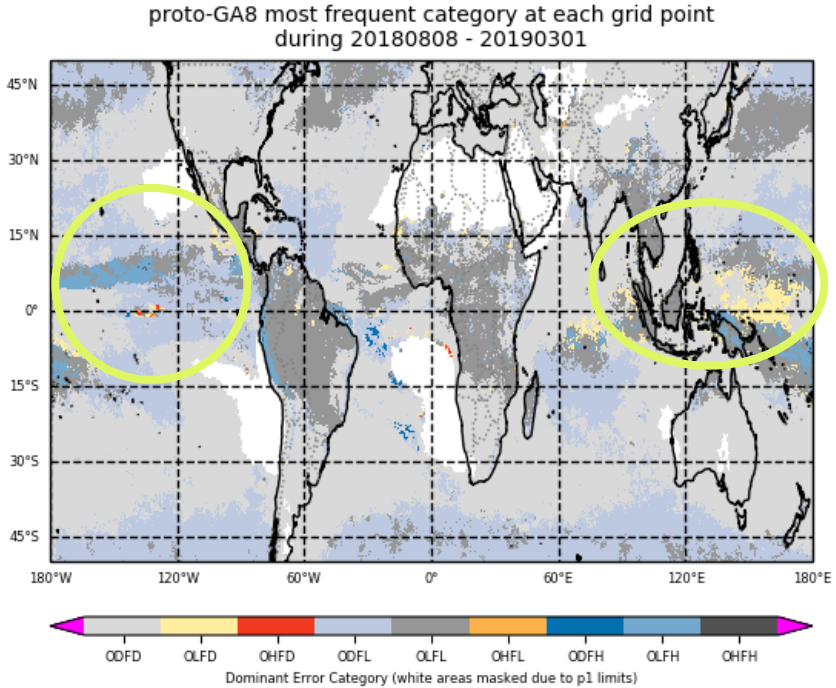
Most frequent error category, day 1 (full 7 months)



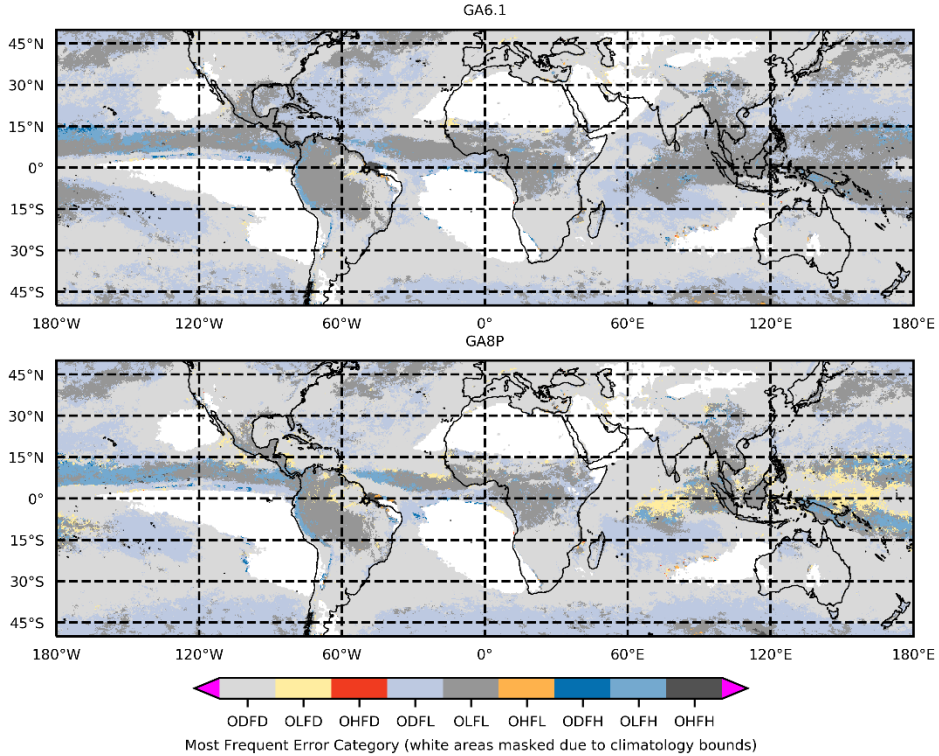


Most frequent error category, day 6 (full 7 months)

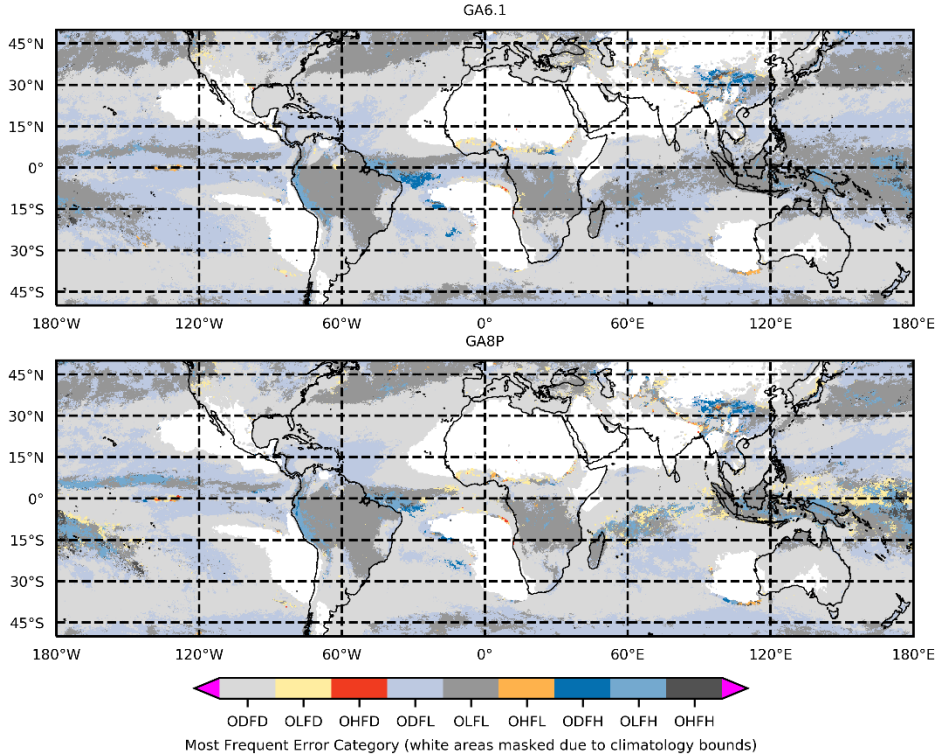




Most frequent error category – seasonal (SON)



Most frequent error category – seasonal (DJF)



Conclusions

- Useful diagnostic evaluation can be done using the error categories produced as part of the calculation of SEEPS
- In addition to the skill perspective given by the score itself (which we haven’t shown here)
- Analysis over different timescales can give different aspects of performance (e.g. as shown here over full trial period, on daily variations, or monthly/seasonal variations)
  - Can highlight systematic model errors

Next steps…

- Re-run climatology creation with GPM IMERG data released this year
- Create sub-daily accumulation climatology to accompany this capability
- Further investigation into behaviour of observation dataset compared to gauges
- Set up category monitoring capability from the operational model at the UK Met Office
- Investigate use with case studies for more process-focused evaluation

Questions?

Paper submitted to MWR:

Using SEEPS with a TRMM-derived climatology to assess global NWP precipitation forecast skill

North, R., Mittermaier, M. and Milton, S.