Workshop on Predictability, dynamics and applications research using the TIGGE and S2S ensembles

2-5 April 2019 – ECMWF

The workshop attracted about 120 participants representing around 60 organisations from Europe as well as from Australia, Brazil, Canada, China, India, Japan, Korea, Nigeria, Singapore, Taiwan and USA. Due to the conference room capacity, the number of participants had to be limited to 120. This large attendance demonstrates the wide interest of the TIGGE and S2S databases in the research and application communities.

The programme included 44 oral presentations. 45 posters were displayed in the ECMWF Weather Room during the 3 days of the workshop. Working group sessions took place on Thursday 4 April afternoon and Friday 5 April morning. Presentations and recordings as well as the recommendations from the 6 working groups are now available on the workshop web page.

The workshop provided an opportunity to review the main scientific advances in predictability, dynamical process studies and applications of ensemble forecasts across the medium and S2S forecast ranges. The oral presentations were organized around 5 themes:

1. Predictability and Dynamics
2. Database technical development
3. Prediction and verification
4. Multi-model ensemble approaches
5. Ensemble applications research

The first session on Predictability and Dynamics showed how the TIGGE and S2S databases provide important insights into the predictability and impacts of stratospheric sudden warmings, the Madden Julian Oscillation and European weather regimes. The databases are also used to evaluate the impact of QBO on MJO and sudden stratospheric warmings In the second session, the technical developments of the TIGGE and S2S databases at ECMWF and CMA were presented. Both report from CMA and ECMWF showed a strong and increasing usage of the databases as well as a large number of scientific publications based on the TIGGE and S2S databases. Several new models may contribute to the TIGGE and S2S databases, including the global ICON_EPS from DWD. The S2S database is also archived at IRI, which provides a maproom and online analysis tool for the S2S database particularly useful for users from developing countries. This session closed with the evaluation of sub-seasonal forecasts from two new models: the Navy Earth System Capability (EPSC) and the new FV3 based GEFS model.

In the third session, several oral presentations discussed the verification of ensemble forecasts. They included a discussion on the use of the Receiver Operating Characteristics (ROC) for ensemble forecast verification and the presentation of a general framework for verifying S2S probabilistic forecasts. A methodology for producing seamless transition from daily to time-

averaged forecasts which could be applied to the TIGGE and S2S models was also presented. Several presentations discussed the prediction of weather regimes for Europe and North America using the S2S database, tropical cyclones using TIGGE, sea-ice cover, stratospheric sudden warmings, precipitation. For instance, it was shown that the prediction skill horizon of sea ice cover exceeds 30 days in some S2S models and that the verification of tropical cyclone in TIGGE demonstrated value in multi-model ensemble forecasts, which led to operational multi-model tropical cyclone forecasts.

Session 4 included several talks demonstrating the benefit of multi-model ensembles for medium-range and extended-range forecasts, as well as frameworks for multi-model ensemble calibration and post-processing. Finally, Session 5 was dedicated to applications studies using the TIGGE and S2S databases. This session showed several areas where medium-range and S2S predictions could benefit society: energy sector, drought monitoring and prediction, telecommunications, flood alerts, agriculture...

The oral sessions were followed by working group sessions. The participants were divided into 6 working groups around the 5 following themes: database technical development, user-oriented variables, Processes and Forecasts (2 groups), Verification/calibration and potential value of multi-model ensembles. The working group reports can be found in the workshop main web page.

The recommendations from the 6 working groups are listed below:

Database technical Development:

1. Provide official project support for keeping contributions to S2S/TIGGE databases after model's upgrades
2. Provide user statistics to better understand the need for high resolution archiving in TIGGE which is very expensive.
3. Share available tools and checking procedures from the archiving centres as well as conversion tools to GRIB2 (e.g. from netcdf) as well as from GRIB to netcdf.
4. Data archiving centres should inform the operational centres by automatic emails each time a cycle has been successfully ingested.
5. **Add pre-computed values such as ensemble means, weekly means, climatology computed from re-forecasts, anomalies.**

User-oriented variables to facilitate communication and developments

As part of the meeting, there was an aim to promote the development of 'user-oriented variables' (UOVs) as a means of improving communication between the Forecasting and User communities. We define user-oriented variables as variables derived purely from forecast output, which represent the metrological aspects that a given user might be sensitive to. A simple example might be V^3 for wind power, but without regard to current turbine efficiencies, *etc*. Key questions are "How reliable and skilful are the ensemble forecasts within the TIGGE/S2S databases when predicting user-oriented variables? " and "How can forecasters learn from any calibration applied by the users?"

1. Identify and speak with important user sectors, learn from previous experience, and develop/adopt UOVs
2. Define actionable thresholds (predominantly these will be fairly extreme)
3. Averaging/accumulation period should generally increase with lead-time
4. Use of proper scores (with associated attributes of reliability, refinement, and sharpness) and user-oriented scores.
5. Score against analysis ("potential skill") and observations (closer to "actual skill", which might be lower due to lack of representativity *etc*.)
6. Need to develop flexible software framework (which could be augmented with additional UOVs)
7. Need for more model data (*e.g.* higher spatio-temporal resolution in S2S?)

Processes and Forecasts:

1. Add more variables:

    - Review TIGGE and S2S user surveys for suggestions.
    - Compare list of variables archived in re-analysis with S2S and TIGGE list of variables.
    - Stratospheric community needs more vertical levels and variables to make better use and evaluation of S2S database. Survey with wish list exists through SNAP: (https://www.dropbox.com/s/cc07fdplb4xoizn/snap_request.docx?dl=0)
    - PV more levels: TH@2PVU, and/or 310, 320, 330, 340, 350K PV, vertical average (500-150hPa)
    - Surface variables at 6-hourly frequency from S2S (diurnal cycle for increasing lead time)
    - Surface fluxes relating to coupling: ocean, land, sea ice.
    - "meridional eddy fluxes" such as $v'\theta'$, $u'v'$ are useful, but transient fluxes depend on frequency of output (or online calculation). Need to complete set for 3-D wave activity flux.
    - $\omega'\theta'$ would be useful (baroclinic energy conversion)
    - Diabatic heating (comprised of physical parameterisation tendencies).
    - Column integral quantities including TCWV and IVT (vector).

2. Expand provision of pre-computed indices (in addition to the existing MJO RMMs, Tropical cyclone tracks, North Atlantic weather regimes..).
3. Communicate via S2S website the fact that the stratospheric community already shares fields internally: heat fluxes, zonal means (SNAP group). Provide survey on existing pre-computed diagnostics. Provide links from the S2S website to data sites and diagnostics.
4. Make better use of existing Wikis/Webpage about the S2S scientific projects to facilitate sharing data/tools
5. Give more attention to the following areas:
    - Land-atmosphere interaction
    - Air-sea interaction also in relation to MJO. Impact of resolution of ocean eddies?
    - How biases affect Rossby Wave propagation and therefore deteriorate teleconnections?
    - How well are planetary stationary waves represented in S2S models?

- Better understanding of tropical-extratropical interactions (both directions)
- Better understanding of MJO-stratosphere-QBO link to establish teleconnections.
- Better diagnostics to capture sources / and teleconnections? (e.g. MJO envelope rather than global index)
- Large-scale patterns, how they relate to surface weather (multi-scale extremes, TCs) in different world regions? How do errors for these patterns grow?
- Conditional dependence on interannual indices (e.g. ENSO, QBO…).

6. Output more diagnostics during special forecasting periods (e.g. YOPP, YOTC…).
7. Review definition of patterns for each mid-latitude storm track region and define shared diagnostics including transition rates between regimes
8. Interact more with people working on error growth on shorter ranges (e.g. German Waves2Weather consortium) in order to better understand the evolution of model biases across time scales.
9. Encourage research on the optimal filtering for assessing skill in a window moving from medium-range to extended-range.
10. Use basic scores to compare TIGGE and S2S ensembles across week 1 (and week 2 where possible) and the role of calibration.
11. Characterise systematic error in large-scale spatial structures as function of lead time in TIGGE and S2S systems across common time range (to 7 days) (E.g., teleconnection patterns conditional on MJO phase)
12. Provide documentation on methodologies to compute climatologies for specific applications and variables.
13. Provide more information on how to best combine different ensemble data for research applications (problem of different initial times)
14. Move less time-critical tasks (reforecast, seasonal forecast) on cheaper hardware/cloud, to free resources to increase resolution / members.
15. More research is needed to link errors in large-scale patterns to errors in specific processes and better understand non-linear interactions between different processes
16. Need better representation of land surface uncertainty both in IC and model
17. Provision of initial condition (step 0h) for S2S would be helpful.
18. Compare rate of increase (in spread and error) in TIGGE and S2S ensembles
19. Add high res forecasts from each centre to TIGGE database (on same grid)
20. Diagnose if S2S ensemble spread saturates at the right lead time and is flow dependant.
21. Use PDEF/WGNE coarse-graining experiments to inform stochastic parametrization
22. Investigate the importance of vertical resolution at upper levels
23. Encourage the harmonization of real-time forecast schedule
24. Communicate existing datasets testing aspects of frequency/size/resolution/complexity (survey contact person)
25. Encourage the real-time provision of charts from the TIGGE and S2S Forecast websites.
26. Encourage the production of long-reforecasts every 5 years since the current S2S re-forecast datasets are too short for statistics conditional on ENSO/QBO

Verification/calibration:

1. Observations are crucial: provide better coverage/access to the observed data (e.g. station data not always shared).
2. Encourage convergence between forecast and re-forecast methodologies.
3. Explore optimal balance of number of members/number of years of re-forecasts for different applications.
4. Need for WMO guideline on sub-seasonal verification (including necessary calibration steps).
5. Define best practices for both calibration and verification
6. Encourage communication with the users about product definition, target, limitation of the forecast, systematic errors, meaning of the verification metrics, risk,… but start with defining the question to be answered
7. Inform about the quality of the observation dataset: estimate of uncertainties and methodology to account for them
8. Encourage the share of software
9. Promote seamless verification with (spatial and temporal) scale adaptation as a function of the lead time and "seamless calibration" with blending/post-processing possibly at different steps of the product generation
10. (Re)Explore machine learning for calibration and model diagnostics (with care)

Potential value of multi-model ensembles:

1. Recommend that the NMHS share their ensemble forecasts and make them available for research through publicly available data bases like TIGGE and S2S.
2. Recommend that the NMHS make the beta runs and/or hindcasts associated with their decision to transition to their current model, publicly available.
3. Use multi-model forecasting of parameters describing shape, position and intensity of common high impact events like cyclones to avoid the smearing of fields.
4. Explore the following promising research areas:
    - Better weighting even though training data sets inevitably inadequate.
    - Synthesizing multi-model ensembles using clustering and advanced visualization techniques
    - Deriving multi-variate pdfs from multi-model ensembles where the members themselves are not equally likely.
    - Post-processing for important spatio-temporal features.
    - Recognize that the inter-model variance is a tool for predicting uncertainty and understanding and characterize model differences.
    - Explore how the much larger ensemble size of multi-model ensembles could be used to explore higher order moments, pre-emptive forecasting.
    - Blending models of differing resolutions through time; e.g. convection resolving to medium range to S2S to decadal – climate model.