Working group on S2S and TIGGE databases: technical aspects

Developer: Makes a simple, intuitive UI

Users

ECMWF

WG2, S2S and TIGGE databases: technical aspects

Participants

- Michael Denhard (DWD)
- Kenta Ochi (JMA)
- Daniele Mastrangelo (CNR-ISAC)
- Xing Hu (CMA)
- Feifei Yang (CMA)
- Manuel Fuentes (ECMWF)
- Frederic Vitart (ECMWF)
- Richard Mladek (ECMWF)

- Harry Hendon (BoM) – inputs provided in advance
- S2S & TIGGE user surveys (~400 users)

ECMWF

Q1 What where the biggest technical challenges to be able to participate in S2S/TIGGE projects at the beginning?
What could improve next time?

Challenges:

- conversion to required GRIB2 format is a challenge (especially from NetCDF)
- fixed reforecast production
- numerous S2S parameters

Ideas for future improvement:

- conversion tools could be shared among the centres (e.g. using NetCDF)
  - even within one organisation links to other projects and sharing of tools are important (KMA, BoM, CMA..)
  - production checking tools from ECMWF should be shared
  - new tools always welcomed (e.g. to convert to required accumulated sums since the forecast start)

Q2   What are the biggest challenges in the current production phase to ensure long-term minimal effort operations?
How could the data archive centres support you more in that role?

Challenges:

- <mark>keeping contribution to TIGGE/S2S database after model's upgrade</mark>
  - ➢ can even prevent to continuation (human/computing resources..)
    - ❖**official project support** would help
- any operational issue can impact products for archives (which has not highest priorities)
- fixed forecast computation is big job after each model upgrade (CPU, storage..)
- TIGGE high resolution is really needed?
  - ➢ who is using it? (limited domain with high res?)
    - ❖some stats could help to understand if it is used

ECMWF

Q3   How could we improve cooperation between data providers and archiving centres (how to communicate/automate tasks/implement checking tools/test/implement model upgrades etc)?

- <mark>share available tools and checking procedures from the archive centres</mark>
    - ➢ WMO GRIB2 compliance encoding check
    - ➢ value limits checks
    - ➢ input file checks (number of fields, reference field list creation..))
- notify each time a cycle has been ingested
    - ➢ automatic email alerts / some status files

ECMWF

Q4   What are good/bad features in the design of S2S/TIGGE databases (data format/structure/encoding/compression etc)?

- Advanced GRIB packing (second order/complex/JPEG 2000) might cause problems (higher CPU needed for data processing)
  - ➢ smaller file sizes for data dispatch can be still achieved using standard unix compression like gzip

ECMWF

Q5 What are good/bad features in the interfaces for getting S2S/TIGGE data (web portals/Web API/direct MARS access etc)?

Problematic features:

- GRIB to NetCDF conversion
    - Not available in ECMWF data portal interface (data discovery only!)
        - ❖ => via WebAPI works (should be better documented?)
    - CMA provides on-the-flight conversion (using GRIB-API)
- pre-computed data (e.g. eps means) before download (like IRI for S2S)
    - not planned at the moment (e.g. in Copernicus CDS (Climate Data Store))
    - can be partially done by MARS compute (for experienced users with full access ECMWF account)
- <mark>slow access and problems with downloading the data from ECMWF archives</mark>
    - mainly related to TIGGE
        - ❖ big archive with data gaps affecting efficient data retrievement
    - S2S data access should have already improved a lot thanks to staging most data on disc (form tapes)
        - ❖ additional "tuning" still might help

ECMWF

Q5   What are good/bad features in the interfaces for getting S2S/TIGGE data (web portals/Web API/direct MARS access etc)?

Problematic features:

- interpolation is different in CMA & ECMWF archives
  - ➤ In extreme cases can impact user results based on the identical data from 2 archives

Good features:

- Interpolation is available!
- See full surveys results!

ECMWF

Q6   What would be the most welcomed technical update of S2S/TIGGE databases from user point of view (data formats/data access/new products like time-series etc)?

- addition of pre-computed values (like eps means, weekly means, climatology computed from reforecasts..)

- faster download

ECMWF

Q8   Are you generally satisfied with the way the S2S/TIGGE databases have been created and supported until now? What would you like to change or improve the most?

- from surveys, number of related research articles etc the archives seem to be very useful for scientific community
    - ➢ Harry Hendon (BoM): *"Support seems to be outstanding."*

- ==ECMWF & CMA are keen on archives continuation==
    - ➢ But to keep them running in long term the official high level support is needed
    - ➢ TIGGE has been agreed until the end of 2019 only
        - ❖ Risk of losing project stakeholders (regular data providers) support

ECMWF

It would be great if one of main outcomes from this Workshop would be to ==highlight== the importance usefulness of the TIGGE and S2S databases, and recommend their continuation.

(particularly relevant for a TIGGE Phase III)

ECMWF