

## TIGGE, an International Data Archive and Access System

Douglas C Schuster, David Stepaniak, Steven Worley, Nathan Wilhelmi, Luca Cinquini and Hannah Wilcox – National Center for Atmospheric Research, Boulder, Colorado  
email: schuster@ucar.edu.

### 1 Introduction

TIGGE, the THORPEX Interactive Grand Global Ensemble, is a key component of the World Weather Research Programme intended to accelerate improvements in 1-day to 2-week weather forecasts. Centralized archives of ensemble model forecast data, from many international centers, are being used to enable extensive data sharing and research during Phase I of the project. The designated TIGGE archive centers include the Chinese Meteorological Administration (CMA), The European Center for Medium-Range Weather Forecasts (ECMWF), and The National Center for Atmospheric Research (NCAR). Scientific data requirements and archive planning solidified in late 2005, and archive collection was initiated in October 2006 with receipt of partial sets of parameters from multiple data providers. This paper will review the status and plans for TIGGE at NCAR, ECMWF, and CMA.

### 2 Data collection at archive centers

The Unidata Internet Data Distribution (IDD) (<http://www.unidata.ucar.edu/software/idd/>) system is the primary mode used to transport ensemble model data from the data providers to the Archive Centers. In some cases, this process can involve multiple steps, with data being transferred to an initial collection point before being fed into the IDD. ECMWF acts as one initial collection point, using FTP and HTTP to collect model output from the Japanese Meteorological Agency (JMA), MeteoFrance (MF), and Australia Bureau of Meteorology (BOM) global models. Additionally, ECMWF runs the UK Met Office (UKMO) model and its own model in house, and stores the output locally. ECMWF ends up transmitting 175 GB in combined volume across the IDD to CMA and NCAR daily. Special circumstances also exist between National Centers for Environmental Prediction (NCEP) and NCAR where NCAR collects native NCEP model data through the IDD, reformats the data into TIGGE compliant format, and transmits the reformatted NCEP data across the IDD to CMA and ECMWF (10 GB/day). Finally, output from the Meteorological Service of Canada (CMC) and CMA, totaling 36 GB/day, is transmitted directly to NCAR, ECMWF, and CMA using the IDD (Table 1). When the Korea Meteorological Administration (KMA), and Centro de Previsao Tempo e Estudos Climaticos (CPTEC - Brazil) join the TIGGE data effort, the combined flow is expected to be 230 GB/day and be well within the capacity of the IDD.

Data Provider	Conforming Parameters	Fcts/Day	GB/Day	Fields/Day	Ens. Members
ECMWF	70/73	2	139	428,712	51
UKMO	62/73	2	21	175,680	24
JMA	61/73	1	7	113,192	51
NCEP	59/73	4	10	316,596	21
CMA	60/73	2	28	72,510	15
CMC	56/73	2	8	163,674	21
BOM	55/73	2	8	147,972	33
MF	62/73	1	15	7,558	11

Table 1. Summary of TIGGE Forecast Data received at the Archive Centers (NCAR, ECMWF, and CMA) on Nov. 1, 2007. Number of conforming parameters of the requested set of 73, number of forecast runs per day, volume in GB delivered per day, number of 2D grids per day, and number of ensemble members per forecast for each of the contributing operational weather forecast centers.

### 3 Data archiving

NCAR uses file-based methods commonly employed by the research community to build its TIGGE data archive. This contrasts with the operationally driven GRIB2 message level data archive built at ECMWF. Over 1.4 million GRIB2 messages are collected at NCAR and combined to build 2600 forecast files daily. Forecast files are organized by initialization date/time, data provider, parameter level type (single level, pressure level, potential vorticity level, and potential temperature level), and forecast time-step. All ensemble members are included in each forecast file (deterministic model data are also included in ECMWF forecast files). Content metadata are extracted from each file and injected into a local database once the files are created. File location and content are stored in the database to drive data discovery and extraction through the NCAR TIGGE portal, and to provide an inventory of data archived in long-term storage (Figure 1). Files remain online on disk for a 2-3 week period until being aged off to the NCAR Mass Storage System (MSS).

### 4. Data access

TIGGE data are delivered in near real-time and by default are available to the public for non-commercial research, with a 48-hour delay after forecast initialization time. Registration is required for all users and will be handled electronically by simple agreement to the TIGGE data policies.

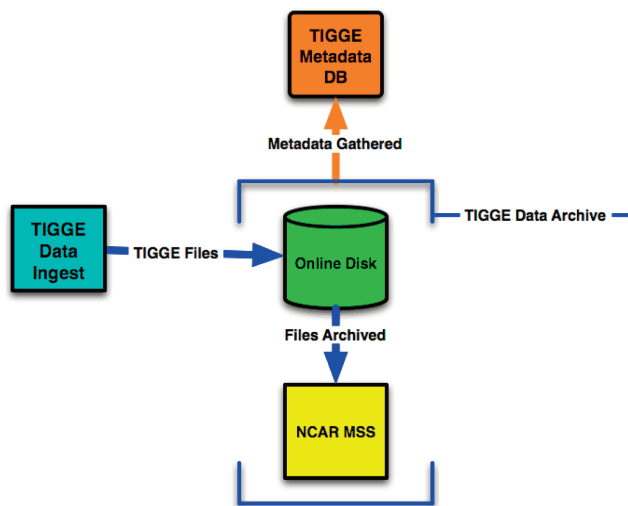


Fig. 1 Summary of TIGGE data ingest, metadata collection, and data archival at NCAR.

At NCAR users can discover data through the TIGGE web portal (<http://tigge.ucar.edu>), and choose to download files directly from the online file archive, or create customized files based on selected parameters from the most recent 2 – 3 weeks of available data. The online file archive is available for download in real-time, while customized files require delayed mode processing. Files can either be downloaded directly through a web browser, or through an automated script provided to the user. Offline data is available by request and through a delayed mode process. Access to both online and offline data has proven to be popular with over 1 TB of TIGGE data downloaded from NCAR servers in 2007.

At ECMWF, users can discover and download data through a web interface <http://tigge-portal.ecmwf.int> linked to the Meteorological Archival and Retrieval System (MARS) <http://www.ecmwf.int/services/archive/>. Parameter, spatial, and temporal sub-setting on the complete TIGGE archive are available through the ECMWF TIGGE portal. CMA plans on activating its TIGGE portal in 2008.

Planned 2008 additions to the NCAR TIGGE portal include user specified spatial sub-setting, and uniform interpolation across multiple center output. Automated access to the complete TIGGE archive will be added to the portal as resources allow. Web services that will give users a common interface to request data at all archive centers are in the initial planning and development stages. This functionality is to be implemented in phase II of the TIGGE project.

In cases where access to the TIGGE archive through the Internet is not practical, the data may be copied to writeable media. Media options will be commensurate with the data request size.

Finally, the entire TIGGE archive is also available on the NCAR MSS. Access to the MSS is restricted to users with NCAR computer accounts as provided by NCAR's Computational and Information Systems Laboratory (CISL). To apply for an account, please contact Database Services at [dbs@ucar.edu](mailto:dbs@ucar.edu) or visit the web page at <http://www.cisl.ucar.edu/docs/access/accounts.html>.

## 5. Data format and software

TIGGE parameter fields, atmospheric levels, and physical units are consistent across all data from the providers as initially defined by the TIGGE working groups and documented in the workshop report (WMO, 2005). After discussions a finalized list of fields and levels was agreed upon and is now available at <http://tigge.ecmwf.int/>. Data are stored in GRIB2 format with parameter encoding also consistent across all data providers. GRIB2 TIGGE parameter encoding definitions can also be found at <http://tigge.ecmwf.int/>.

In contrast to parameter fields and format, each data provider may submit their model output in a resolution they choose (Table 2). This allows for research opportunities to study the impact of model resolution.

Data Provider	Model Resolution
ECMWF	N200 (Reduced Gaussian) days 1- 10 N128 (Reduced Gaussian) days 11 - 15
UKMO	1.25 x 2/3 Deg Lat/Lon
JMA	1.25 x 1.25 Deg Lat/Lon
NCEP	1.00 x 1.00 Deg Lat/Lon
CMA	0.56 x 0.56 Deg Lat/Lon
CMC	1.00 x 1.00 Deg Lat/Lon
BOM	1.50 x 1.50 Deg Lat/Lon
MF	1.50 x 1.50 Deg Lat/Lon

Table 2. Model Grid Resolutions received from the providers at the TIGGE Archive Centers on Nov. 1, 2007

### 5.1 Software

TIGGE data analysis tools include the NCAR Command Language (NCL), ECMWF's GRIB-API, Unidata's Gempack software, and NOAA's GRIB2 libraries. Additionally, NCAR and ECMWF are collaborating to develop a GRIB2 to NetCDF conversion tool to open up additional software options in analyzing TIGGE data. Links to the supporting sites of analysis software can be found through the NCAR TIGGE web portal (<http://tigge.ucar.edu>).

## 6. Summary

The TIGGE data archive provides a valuable resource for research of probabilistic forecast techniques. Archive Centers at NCAR, ECMWF, and CMA, are currently collecting and archiving 221 GB/day of global ensemble model data from ECMWF, UKMO, NCEP, JMA, CMC, CMA, MeteoFrance, and BOM. Additional data providers including KMA, and CPTEC plan on initiating TIGGE data contributions in the near future (KMA has initiated its contribution as of Dec 2007). Access to the TIGGE archive at NCAR is offered in several ways:

- Forecast files organized by model initialization date/ time, data provider, parameter level type, and forecast time-step from the most recent two to three weeks of available forecast data (available for immediate download).
- Customized forecast files built according to user specified parameter subsets from the most recent two to three weeks of available forecast data (requires delayed mode processing).
- Data may be provided on writeable media, upon request.
- All TIGGE data are available to NCAR users from the MSS.
- Spatially subset and interpolated grids across all data providers will be available at NCAR by mid- 2008 (currently available through the ECWMF TIGGE portal).

The primary focus is to provide a high quality archive with an emphasis on data delivery. The NCAR TIGGE archive has been designed to promote efficient access for a large variety of users and requests. It is actively supported by data specialists who provide: consultation on the archive products; locally developed software; and individual requests that require nonroutine data processing and possible data delivery on media. Basic data services and access methods for the archive are now available with additional services coming in the near future.

## References

WMO, 2005: First Workshop on the THORPEX Interactive Grand Global Ensemble (TIGGE). WMO, TD-No.1237, [http://www.wmo.int/thorpex/pdf/tigge\\_first\\_workshop\\_report.pdf](http://www.wmo.int/thorpex/pdf/tigge_first_workshop_report.pdf).