THE THORPEX INTERACTIVE GRAND GLOBAL ENSEMBLE (TIGGE)

Philippe Bougeault[1] , Zoltan Toth[2], Craig Bishop[3] , Barbara Brown[4], David Burridge[5], De Hui Chen[6], Beth Ebert[7], Manuel Fuentes[1], Thomas M. Hamill[8], Ken Mylne[9], Jean Nicolau[10], Tiziana Paccagnella[11], Young-Youn Park[12], David Parsons[5], Baudouin Raoult[1], Doug Schuster[4], Pedro Silva Dias[13], Richard Swinbank[9], Yoshiaki Takeuchi[14], Warren Tennant[9], Laurence Wilson[15] and Steve Worley[4]

(1) European Centre for Medium-range Weather Forecasts, Reading, UK

(2) National Centers for Environmental Prediction, Camp Springs, MD

(3) Navy Research Laboratory, Monterey, CA

(4) National Center for Atmospheric Research, Boulder, CO

(5) World Meteorological Organization, Geneva, Switzerland

(6) Chinese Meteorological Administration, Beijing, China

(7) Bureau of Meteorology, Melbourne, Australia

(8) NOAA Earth System Research Laboratory, Boulder, CO

(9) Met Office, Exeter, UK

(10) Météo-France, Toulouse, France

(11) Agenzia Regionale Prevenzione e Ambiante dell'Emilia-Romagna, Bologna, Italy

(12) Korean Meteorological Administration, Seoul, Korea

(13) Centra de Previsao de Tempo e Estudos Climatico, Brazil

(14) Japan Meteorological Agency, Tokyo, Japan

(15) Meteorological Service of Canada, Montreal

Contact: *P.Bougeault@ecmwf.int*

(from 1st April 2009, Philippe.Bougeault@meteo.fr)

Capsule

The leading NWP centres have agreed to create a database of their operational ensemble forecasts and open access to the research community, in order to accelerate the development of probabilistic forecasting

Abstract

The objectives of TIGGE (GEO task WE-06-03) are (1) to facilitate closer cooperation between the academic and operational meteorological communities by expanding the availability of operational products for research, and (2) to facilitate exploring the concept and benefits of multi-model probabilistic weather forecasts, with a particular focus on high-impact weather prediction. Ten operational weather forecasting centres producing daily global ensemble forecasts to 1-2 weeks ahead have agreed to deliver in near-real-time a selection of forecast data to the TIGGE data archives at the China Meteorological Agency, the European Centre for Medium-range Weather Forecasts and the National Center for Atmospheric Research. The volume of data accumulated daily is 245 GB (1.6 million global fields). This is offered to the scientific community as a new resource for research and education. The TIGGE data policy is to make each forecast accessible via the internet 48h after it was initially issued by each originating centre. Quicker access can also be granted for field experiments or projects of particular interest to the World Weather Research Programme and THORPEX. Examples of initial research results based on TIGGE data are discussed and they point to several directions for additional research. The extension of the TIGGE concept to exchange of data from Limited-Area ensemble systems is currently underway under the auspices of a TIGGE-LAM group. Prototype forecast products based on TIGGE data are also being prepared. They are paving the way towards the "Global Interactive Forecast System" (GIFS), one of the long-term goals of THORPEX.

Keywords *– THORPEX, WWRP, WMO, TIGGE, GIFS, Multi-model forecasts, ensemble prediction*

1. Objectives and Concept

The THORPEX Interactive Grand Global Ensemble (TIGGE) is a key component of THORPEX (TIP, 2005) which in turn is a major component of the World Weather Research Programme (WWRP) under the World Meteorological Organization (WMO). A key goal of THORPEX is to accelerate improvements in the accuracy of 1-day to 2-week high-impact weather forecasts for the benefit of humanity. TIGGE was initiated in 2005 at a workshop hosted by the European Centre for Medium-Range Weather Forecasts (ECMWF). A full report of this event was prepared by Richardson et al. (2005).

The following objectives of TIGGE were agreed to at the workshop:

(i) enhance collaboration on ensemble prediction, internationally and between operational centres and universities;

(ii) develop new methods to combine ensembles from different sources and to correct for systematic errors (biases, spread over-/under-estimation);

(iii) achieve a deeper understanding of the contribution of observation, initial and model uncertainties to forecast error;

(iv) explore the feasibility and the benefit of interactive ensemble systems responding dynamically to changing uncertainty;

(v) enable evolution towards an operational system, the "Global Interactive Forecast System (GIFS)".

To meet these objectives, it was agreed that ensemble forecasts generated by a number of NWP centres (hereafter "Data Providers") would be accumulated in real time in databases operated by three TIGGE "Archive Centres" (see Table 1) and made accessible to the scientific community for research and education with only a slight (2-day) time delay. The highest priority data accumulated in the TIGGE archive are the ensemble forecasts generated

routinely (operationally) at major forecast centres around the world. This core data stored in the TIGGE archive is accumulating at a daily rate of approximately 245 GB from ten providers around the world (see Table 1). Additional special datasets may be added in the future for specific research and application areas. Ensemble forecasts from a number of limited-area systems are being considered for addition to the archive.

The TIGGE project has been developed under the leadership of the THORPEX GIFS-TIGGE Working Group, to which most of the authors belong. The WMO Working Group on Numerical Experimentation (WGNE)/WWRP Joint Working Group on Forecast Verification Research advises the project on verification methodology. In addition the WMO Expert Team on Ensemble Prediction Systems advises the project on a number of issues, for instance metadata formulation. TIGGE has strong links with the North American Ensemble Forecasting System (NAEFS, see Toth et al, 2005), which synthesizes ensemble products from NCEP and the Meteorological service of Canada. Although NAEFS uses data from only two centres and produces real-time, operational products, TIGGE and NAEFS share many technical aspects, and NAEFS plans to implement results from TIGGE. It is believed that TIGGE and the NAEFS will ultimately evolve into a single operational system. TIGGE is also registered as Task WE-06-03 of the Group on Earth Observations (GEO, 2007). It has general relevance to GEO societal benefit areas that will benefit from access to advanced multi-model global weather forecasts and the derived products, especially in areas related to risk management, disaster mitigation, energy, agriculture and health.

2. Building the TIGGE databases

The implementation of TIGGE has been quite challenging. Data must be collected from ten different centres and redistributed to a potentially large number of users very rapidly, using only readily available communication technologies such as the Internet. The content of the

database must be as homogeneous and have as few gaps as possible. The Archive Centres must operate user-friendly interfaces, enabling researchers to obtain subsets of ensemble data, especially over geographic regions of their choice. This post-processing of archived data, done at the Archive Centres, typically includes grid conversions, format conversions, and the extraction of sub-areas, parameters and levels. Archive Centres must also provide links to associated regional and user-specific observational data sets.

a. Content and format of the archive

As a starting point, all partners have agreed on a common way of referencing data within the TIGGE dataset. Fields are described using the following attributes: *analysis date, analysis time, forecast time step, origin centre, ensemble member number, level* and *parameter*. In this context "parameter" refers to the physical quantity represented by the field: temperature, pressure etc. Furthermore, all partners have agreed to provide data in the same units and with the same period of accumulation (when applicable). This led to the definition of the TIGGE core dataset to which all Data Providers must adhere (Table 2).

When the first data transfers were being set up between the partners, it became clear that most Data Providers could not contribute to the full agreed list of products, mainly because these products were not produced by their models. It was decided that waiting for all partners to upgrade their systems to produce the missing fields was an unnecessary delay in the building of the archive. As all Data Providers were producing the most important fields (the usual surface parameters and upper air data on pressure levels), a staged approach was adopted. Data Providers would join the project by sending currently available parameters, and would add more parameters during the course of the project. The actual data accumulation started between October 2006 and January 2008, depending on parameter and Data Provider. The TIGGE database now contains most requested data from all Data

Providers, and holds more than 180 TBytes of data (1.1 billion fields, see Table 3). Forecast data has now been archived for more than two years for several parameters.

To guarantee the best precision, original model grids and resolutions are preserved whenever possible. Data Providers supply data on a horizontal grid of their choice, as close as possible (identical if possible) to the computational grid of their model. These data are stored in the database without any modification. On the other hand, users generally want data interpolated on common regular grids of their choice. The archive centres offer this interpolation service. Before delivery, data may be interpolated to a single point or to a regular, limited-area or global, lat-lon grid specified by the user. To respect the unique features of each model, data providers are encouraged to supply and regularly update the interpolation software used by the archive centres. Alternatively, the archive centres can use other available interpolation software.

As a common archive data format, it was decided to use GRIB edition 2 – it is the only WMO standard that supports ensemble data without the need for local extensions. Moreover, the NAEFS community is committed to using it. Data Providers are requested to provide data to Archive Centres directly in the archive format. The Archive Centres defined the list of GRIB2 codes, tables and templates to be used for each of the fields. They also provided guidelines on how all TIGGE fields should be coded in GRIB2, as well as examples of properly encoded model outputs. For some TIGGE parameters proper rules did not yet exist. Requests for clarifications and proposals for new parameters were submitted to the WMO Expert Team on Data Representation systems and Codes; as a result, a substantial number of amendments were made to the Guide to the WMO Table Driven Code Form. (see *http://www.wmo.ch/pages/prog/www/WDM/Guides/BUFRCREXGuide-English.html*).

Because of the large data volumes involved, an archive of the full forecast model output fields was not possible; consequently, the TIGGE archive does not include all the fields necessary to provide lateral boundary conditions to run limited-area models. The possibility to do this in special circumstances will be discussed in Section 5 below.

b. Data transfers

After extensive testing, it was shown that IDD/LDM (Internet Data Distribution system, Local Data Manager), an Internet based distribution system developed by UNIDATA, suits TIGGE requirements. In particular it can support the parallel transfers needed to exchange large volumes of data. LDM is a broadcasting system, based on a subscription mechanism: a "downstream" LDM can subscribe to "products" from an "upstream" LDM. When a product is inserted in the upstream LDM, it is automatically sent to all the downstream LDMs that have subscribed to this product. Unfortunately, such a broadcasting system does not guarantee that products will be received by all downstream LDM, in particular if some are temporarily not running. To overcome this problem, a protocol has been defined on top of LDM to exchange fields by specifying a file name convention and a series of messages to request retransmission of missing fields. A complete description of the protocol is available on the TIGGE web site (see *http://tigge.ecmwf.int*).

Although LDM is the preferred solution for the exchange of data between the TIGGE partners, it was not always possible for Data Providers to install an LDM server at their site. Some decided to use either FTP or HTTP to transfer the data to one of the Archive Centres which would in turn relay it to the two others. Figure 1a shows the various transfer protocols used between the Data Providers and Archive Centres.

The available network bandwidth between Europe, the USA and China is sufficient to meet the needs of TIGGE. Nevertheless this would become a limiting factor if TIGGE partners decided to engage in real-time exchange for operational products.

c. Operational aspects

The Archive Centres are in charge of the technical coordination of the project. For day-to-day operations, tools have been created to monitor the data transfers. Each Archive Centre maintains a web page showing volumes, date of data and date of reception from each Data Provider. This information is used to cross-validate the data arrivals at the three Archives Centres. If an Archive Centre does not receive the expected data from a Data Provider, or if the data are incomplete or corrupted, it first checks with other Archive Centres and determine whether the failure is an isolated case. In that case recovery is initiated between Archive Centres. If not, the Data Provider must re-initiate the data delivery, even if this means rerunning a forecast cycle. In any case, incidents are investigated and documented. The Archive Centres have agreed to define and collect common metrics that can be used to create combined TIGGE-wide reports. This information will be used as a basis for the further evolutions of the system.

Participation in TIGGE does not interfere with the operational activities of Data Providers; they are able to upgrade models, introduce higher resolutions, and make all customary changes as needed. On the other hand, Data Providers must take into account their participation in TIGGE when planning changes to their forecasting systems, and must inform Archive Centres accordingly. Metadata information files are maintained to allow TIGGE users to keep track of changes in any of the contributing systems.

d. Quality control

Quality control procedures are implemented to guarantee the homogeneity of the dataset. Many tests have been designed to ensure that the TIGGE database does not contain any badly encoded data and that researchers can use the archive with confidence. When a Data Provider starts sending new data to the Archive Centres, all new fields are marked as being in "test" mode. The new fields are checked against the agreed list of TIGGE fields, and then checked for proper encoding and units. Once all the fields have been validated, they are tagged as being in "production" mode and are sent routinely. Although all the fields have been thoroughly validated during the test phase, changes may be introduced by the Data Providers due to evolutions in their operational environments and this may have an unforeseen effect on the TIGGE data exchange. This is why all the fields which are received in production mode are also validated, and any field that fails this validation are quarantined for further investigation. This guarantees that no unexpected change in data will compromise the homogeneity of the archive.

Fields may be properly encoded, and pass checks by validation tools, but their values may still be incorrect. This is usually the case if the units are wrong, or the fields are instantaneous instead of accumulated. To spot these problems, a series of plots are produced every day, comparing the data from all providers.

In order to ensure that the TIGGE archive is as complete as possible, a web page has been set up to show the status of availability of each cycle from each Data Provider. This web page shows the whole history of the dataset: the addition of new fields is indicated, as well as any incidents. It is used by the providers to check the completeness of their contributions to the database (see *http://tigge.ecmwf.int/tigge/d/tigge_history*)

3. Access to TIGGE Data for Research and Education

Access to TIGGE data is provided for Research & Education through a simple electronic registration process, with valid e-mail address and acknowledgment of conditions of supply. Under the simple registration process, access is given with a delay (48 hours) after initial time of each forecast. Real-time access is granted (subject to bandwidth limitations) in some cases, e.g. for field experiments and projects of special interest to THORPEX. Registration for this real-time access is handled via the THORPEX International Project Office.

Data access is operated via the three TIGGE data portals operated by NCAR, ECMWF and CMA (see the portals URLs in Table 1). The current functionalities of the data portal are (i) registration; (ii) Search, discover, and download files; (iii) Select data by initialization date/time, data provider, file type and forecast time; (iv) Interpolate data on a regular, limited-area or global, lat-lon grid specified by the user; (v) Check volume and download data.

All three Archive Centres are currently able to distribute data in GRIB2 format. NETCDF format is also available from NCAR and should soon become available from the other Centres. Plans to expand the services available include, inter alia, the possibility of setting up standing data requests (for example order specific data to be sent routinely every day to interested users).

At the moment the three data portals have a total of about 230 registered users, of which a third are active. Figure 1b shows the country of origin of the registered users.

4. Early results from research based on TIGGE

A list of research papers based on TIGGE data is continuously updated at *http://tigge.ecmwf.int/references.html*. Only a few of them are being reviewed here.

a. Performance of individual systems

Park et al. (2008) have investigated the performance of various single- and multi-model ensemble systems available from TIGGE up to December 2007 (thus, their results reflect the performance of the various systems only up to this time). This study focussed on 500hPa geopotential height and 850hPa temperature and was the first extensive comparison of the global operational ensemble prediction systems. Each system was verified primarily against its own analysis, but the sensitivity to the choice of the verification analysis was also investigated. This highlighted large differences in the forecast quality of the various contributed systems, both for the deterministic forecasts based on the control runs or on the ensemble mean, and, even more, for the probabilistic forecasts. Differences in the accuracy of probabilistic forecasts were shown to be due both to model error characteristics and to the quality of the spread-error relationship. Ideally, the spread of an ensemble should be equal to the RMSE of the ensemble mean throughout the forecast range, for all forecast parameters. This turns out to be a very challenging goal to attain. The best calibrated ensemble systems have now reached this optimal calibration for upper air parameters such as the geopotential height at 500 hPa or the temperature at 850hPa. For other parameters (e.g. surface temperatures and precipitation) this has not been yet reached, and some systems are still quite far from it for all parameters. This, on top of model error differences, was shown to result in differences of up to three days in forecast skill between the various systems. Another result worth mentioning is that in the tropics all systems (in 2007) were substantially underestimating the spread compared to the RMSE of the ensemble mean. This finding

formed a strong incentive for several data providers to address more vigorously the issue of improving the quality of ensemble forecasts in the tropics.

The choice of the verification analysis was shown to have relatively small impact for upper-air parameters in the mid-latitudes as long as one of the best analyses was used. On the other hand in the tropics or generally for the near-surface parameters, despite considerable work at NWP centres, there are still large differences between analyses from various systems, and therefore the forecasts from most systems verify significantly better when scored against their own analysis than when scored against the analysis of a different system. As a consequence it may for some time remain unclear how to optimally combine and verify forecasts from different systems.

To complement the above results, a more recent assessment of the spread-error relation in TIGGE systems is shown in Figs. 2 and 3, based on forecasts from December 2008. Fig. 2 shows how the spread in sea-level pressure develops with forecast range as a function of the latitude. It can be readily compared to Fig. 3, where the RMS errors of the ensemble means are shown with the same units and colour code. Note that in order to obtain a fairer comparison, sub-ensembles of 10 members have been used for each system, resulting in some degradation of the results for the largest ensembles. The spread in this recent period is still often smaller than the RMSE of the ensemble mean. This is especially true in the Southern Hemisphere, and in the tropics. For some systems, this situation is actually expected as they do not use initial perturbations in these regions (e.g. the JMA system in the Southern Hemisphere). Even in the Northern Hemisphere there are large differences from system to system, showing that beyond the size of the ensemble, the methods used to represent initial and model uncertainty are important.

b. Skill of multi-model systems

Park et al. (2008) have also compared the performances of various single and multi-model systems, with and without bias correction. They assessed several methods to compute the bias correction and showed that this is a sensitive issue. One particular result is reproduced here in Fig. 4 (cf Figure 17 of Park et al.). It compares the performance of the single ECMWF ensemble, with and without bias correction, and two bias-corrected multi-model ensembles (ECMWF + UKMO and ECMWF + UKMO + JMA + CMA). Both the root-mean-squared-error of the ensemble means (RMS) and the Ranked Probability Skill Score (RPSS) are shown. The RPSS computation was based on ten climatologically equally likely categories. The results cover 86 cases from June to August 2007. It can be seen that the performance for the geopotential height at 500 hPa over the Northern Hemisphere benefits very little from either the bias-correction or the addition of the extra members. On the other hand, for temperature at 850hPa over the tropics, bias correction has a large positive impact on the quality of ECMWF-only ensemble. The addition of extra members from other systems also has a positive impact, although the authors note that some saturation effect can be seen when many systems are used. Qualitatively similar results were found with other combinations of models and other periods: multi-model forecasts only gave small benefits for forecasts of NH 500 hPa geopotential height, but generally better results for tropical 850 hPa temperatures. A possible weakness of Park et al. results lies in their choice of ECMWF analysis as the verification for all the above systems. As discussed above, the choice of the optimal verification analysis in the tropic is both a difficult and a sensitive one, and additional results are needed, especially using direct observations, before drawing final conclusions about the relative merits of the various systems.

Matsueda and Tanaka (2008) compared the ECMWF ensemble with three different multi-model ensembles, each of which combined forecasts from MSC, ECMWF, JMA, NCEP and UKMO. The first ensemble was designed to be the same size as the ECMWF ensemble (51 members); it was created by taking 10 members from each model (except 11 from ECMWF). The second ensemble was 154 members, all the members then available from the 12 UTC forecasts from each centre. The size of the third ensemble was 279, comprising all members available each day. Fig. 5 shows the gain in Ranked Probability Score of 500 hPa geopotential height, relative to the ECMWF ensemble. Verification was again performed against the ECMWF analyses. These results show relatively small benefit from the multi-model ensemble, and in some cases a negative impact in the early stages of the forecasts (perhaps due to the use of the ECMWF analysis for verification). The 51-member ensemble shows that a small benefit results from using multiple models, but keeping the same sized ensemble - even though the ECMWF model is the most skilful. The 154-member ensemble shows most benefit, while the 279-member ensemble suffers from some of its component members being taken from older forecasts. Overall the study of Matsueda and Tanaka seems to confirm the results of Park et al.

Johnson and Swinbank (2008) investigated the benefit of a 3-model ensemble, using ECMWF, NCEP and UKMO ensembles. Fig. 6 shows Brier Skill Scores for mean sea level pressure and surface air (2m) temperature, verifying the skill of categorical probabilistic forecasts, with category boundaries set as the climatological quantiles defined using ERA-40 data. Each forecast was bias-corrected and forecasts were verified against a multi-model analysis (the mean of the three analyses). Three variations of multi-model ensemble were assessed: first, each ensemble was weighted equally; second, each ensemble was weighted to take account of its estimated RMS error, and third, both weights and variance of each ensemble were adjusted. Fig 6a shows that the skill in forecasting sea level pressure greater

than the climatological mean is very similar for both the ECMWF and multi-model ensembles. Fig 6b compares scores for forecasts of 2m temperature, relative to the mean; in this case all three multi-model ensembles give a significant improvement over any single ensemble. The largest benefit of multi-model ensembles is shown for forecasts of 2m temperature greater than the 90th percentile (Fig 6c). The results show relatively small impacts from varying the ensemble weighting, consistent with earlier results (e.g. Peña and van den Dool, 2008).

These statistical studies of the benefits of multi-model ensembles were complemented by a series of case studies by Titley et al (2008). They compared the forecasts of several high-impact weather events, based on diagnostics from three different ensemble prediction systems (UKMO, ECMWF and NCEP). Having access to different ensemble forecasts was shown to add value at both short and medium range. There is value in the multi-model ensemble approach both in cases where there is agreement between models (increasing confidence in the forecast) and where there are significant differences (giving a better representation of uncertainties). Different case studies had a different "best" model. There were several cases where a significant signal of the high-impact weather was forecast well into week two of the forecast, justifying running the ensemble forecast models out to 15 days.

In summary, results from TIGGE have so far shown only limited benefit of multi-model ensembles over the best available model for forecasts of 500 hPa height and sea-level pressure. The improvement for 2-m level and, to a lesser extent, 850 hPa level temperatures appears larger, especially in the tropics. One possible explanation is that large-scale, mid-tropospheric dynamical fields are generally consistently predicted by current NWP models. There is less consistency among models for near surface variables, as these forecasts are

more dependent on details of physical parameterizations and are thus affected by different model biases. These results are also consistent with the notion that benefits from multi-model combination are more significant when ensembles with comparable skill are combined whilst the benefits are less clear when poorer performing ensembles are added to a better performing system. The verification statistics do seem to be sensitive to the verification data and climate reference data. Although we have only shown examples of one type of score from each study, all studies showed clearer benefits of multi-model ensembles for probabilistic scores than for deterministic scores.

More work is needed to confirm the above conclusions on longer time series and by direct comparison to observations. There is also an urgent need to explore the forecast skill for other parameters, such as 10m winds, rainfall and clouds. Above all it is necessary to explore the impact of multi-model systems on severe-weather forecasts. It is likely that the benefits of multi-model systems vary depending on weather parameter, on lead-time, and on the user. They may also vary rapidly in time due to variations in the quality of component systems. It is important to fully document these aspects as the cost of maintaining operational multi-model systems are likely to be significant, and must not exceed the benefits.

c. Other applications of TIGGE

Beyond the derivation of probabilistic weather forecasts, ensembles have a wide variety of applications. They can be used in decision support systems to explore the sensitivity of user-relevant consequences of weather conditions. For example Pappenberger et al (2008) applied both single-model and multi-model ensembles to the prediction of a particular flood event in Romania in October 2007. Results reveal that, in this case, warnings could have been issued as early as 8 days before the event. A comparison of 5-day forecasts, shown in Fig 7, illustrates the positive impact of the multi-model approach at this lead time. The subsequent

forecasts provided increasing insight into the range of possible flood conditions. This case study illustrates the potential value of the TIGGE archive and the multi-model ensembles approach to raise preparedness and reduce the negative socio-economic impact of floods. He et al (2009) is another application of TIGGE ensemble forecasts to flood forecasting.

Finally the TIGGE database is opening the possibility of more upstream studies on how various systems (including multi-model) treat some features of the atmosphere. As an example Froude (2008, private communication) investigated the representation of extra-tropical cyclones in medium-range forecasts present in the database. A notable result is that models generally underestimate the speed of propagation - although in different proportions (see Figure 8). Champion (2008) compared the different methods used for defining initial perturbations. He found that these result in large differences in initial amplitude of the perturbations and subsequent growth rates. Significant differences were found even between systems using similar methods, which points to the different behaviour of the data assimilation systems. In particular he found that singular-vector based methods create perturbations with a westward tilt with height at initial time, experiencing a rapid baroclinic growth. On the other hand perturbations based on the Ensemble Transform Kalman Filter method have no tilt with height initially and progress to having an eastward tilt with height, which is consistent with decay.

Those few examples are just meant to show how TIGGE-based research will help understand the behaviour of the various current approaches to ensemble forecasting.

5. Ensemble forecasting on limited areas: TIGGE-LAM

Recently an expert group (TIGGE-LAM panel) was formed to coordinate the contribution of Limited Area Ensemble systems to TIGGE and, in a longer perspective, to the Global

Interactive Forecasting System (see below). So far the group has been focussing on three topics.

Firstly, there is a plan to create a database of limited-area ensemble products, similar to the global TIGGE database. Discussions are on-going to agree the content and format of these database and their access conditions for research and education usage. In a first phase, the three TIGGE Archive Centres have agreed in principle to host a sub-set of high priority TIGGE-LAM data, on the basis of a geographical share of responsibility. In the future this should evolve towards a decentralized database system based on Regional Centres where the full data-sets will be available.

Secondly, there is a large interest in the meteorological community to make the various global and regional systems "interoperable". This includes among other aspects the development of tools to enable use of initial and lateral boundary conditions from the various global ensembles to drive the various limited-area ensembles. The concept of GIFS, and in particular of Adaptive Forecasting Systems, would benefit of an improved level of interoperability allowing the desirable flexibility in coupling the different LAMs and the different Global Models. In Europe a specific Interoperability programme, lead by the UK Met Office, has started under the auspices of the SRNWP/EUMETNET programme (the Short Range Numerical Weather Prediction programme of the Network of European Met Services) to facilitate the interoperability among the different numerical forecasting systems. Due to the many common interests and purposes, TIGGE LAM and the Interoperability programme will cooperate actively.

Thirdly, there is interest in the possibility to relocate existing LAM EPS systems already implemented and tested on specific regions, in other areas not covered by analogous forecasting systems. This assumes that the systems are first undergoing the necessary

scientific testing in the different target regions with due regards to the capability to predict local severe weather events. Quite recently, a TIGGE LAM workshop was held in Bologna, Italy, where the above mentioned issues were discussed together with several other aspects related to the LAM EPS contribution to TIGGE (*http://www.smr.arpa.emr.it/tiggelam/workshop*). One major outcome of this workshop is the decision to concentrate most of the activity on Testbeds, to be organized in relation with specific Forecast Demonstration Projects (FDPs). Such projects are organized under WMO auspices to facilitate the intercomparison of various ensemble systems and to investigate specific forecasting issues. These activities will be planned and carried out in close cooperation with the WWRP WG-MWFR (Working Group on Mesoscale Weather Forecasting Research), with the JWGFVR (WWRP/WGNE Joint Working Group on Forecast Verification Research), with the local contact people and especially with the THORPEX Regional Committee representatives, who are in the right position to stress the relevant regional issues and to set priorities. For more information on TIGGE LAM see *http://www.smr.arpa.emr.it/tiggelam/.*

6. Towards the Global Interactive Forecasting System

The Global Interactive Forecasting System (GIFS) is central to the THORPEX vision of accelerating the improvement of one-day to two-week forecasts, focusing on high-impact weather (see TIP, 2005). The objective of the GIFS is the production of internationally coordinated advance warnings and forecasts for high-impact weather, to mitigate loss of life and property and contribute to the welfare of all WMO nations, with a particular emphasis on least developed and developing countries. Ensemble predictions will play a critical role in assessing and mitigating weather- and climate-related risks by quantifying forecast

uncertainty. GIFS will be based on forecast products and services contributed voluntarily by NWP centres and other providers around the globe.

As its name indicates, the GIFS-TIGGE working group is in charge of developing concepts for the GIFS and fostering discussions with other THORPEX and WMO groups. The following issues have been identified:

- *Science*: TIGGE results have not yet demonstrated unambiguously the benefits of multi-model systems; more research is needed. In addition multi-model systems may become less attractive in the future if the skill of one or two individual systems exceed that of the rest.
- *Resource*: Much hardware and manpower will be needed to develop reliable exchange mechanisms for real-time production. Resources for multi-model systems will only be made available by the operational centres if research results provide enough evidence that real benefits can be expected.
- *Operational continuity*: It will be a challenge to manage operational changes occurring at different times for the various component systems, to guarantee a smooth progress of the multi-model system skill and to supply proper information on system upgrades to the users.
- *Data policy*: Several TIGGE providers will want to protect their commercial revenues from probabilistic forecasts. Negotiations will be needed to agree on a scheme that satisfies all partners.

As a way forward, the GIFS-TIGGE working group decided to develop pilot products clearly related to severe weather. In relation with the T-PARC experiment of THORPEX an exercise of real-time exchange of tropical cyclone tracks predicted by the various TIGGE systems has been defined and monitored by the Bureau of Meteorology (Australia). A special format easy

to read for academic partners (CXML, the Cyclone XML format, see Ebert et al., 2008) was defined, and the TIGGE data providers were requested to provide tropical cyclone tracks on ftp sites in real time for the duration of the T-PARC experiment. These data will also be distributed by the TIGGE Archive Centres in addition to the usual TIGGE data.

Figure 9 shows an example of multi-model tropical cyclone tracks and strike probability charts generated from track data distributed in CXML format. This example takes data from only two ensembles, (ECMWF and UKMO), but the technique can easily be extended to more. In this case there was a large overlap between the spreads of the two individual ensembles, but the ECMWF EPS showed a larger probability of a more southerly track while the UKMO EPS gave a higher probability to a more northerly track. Research continues into the optimal combination of ensembles in this way, for example whether the contributions from individual ensembles should be weighted according to ensemble size or past performance.

7. Conclusions

The TIGGE project has attracted a high level of interest from both operational centres and the research community. TIGGE has already reached two key targets: firstly, it has led to the agreement of a data format to be used by all partners to exchange forecast, facilitating comparison and combination of forecasts from different systems; secondly, an increased level of communication between the communities developing and using the ensemble forecasts. This will certainly promote the use of probabilistic forecasts.

We are convinced that the TIGGE databases will constitute a key resource to reach the objective of THORPEX: the acceleration of the progress of the forecast skill for severe weather events from one day to two weeks ahead. This will be reached by a robust combination of research on the scientific basis of ensemble prediction, of experimentation

with new products and of developments of new protocols and policies for data exchange across WMO Member States and across the science and application communities.

Acknowledgments

We are indebted to Gordon Brent, Piers Buchanan, Roberto Buizza, O. Champion, Steve Chiswell, Luca Cinquini, Lizzie Froude, Renate Hagedorn, Christine Johnson, Honglian Lang, Mio Matsueda, Florian Pappenberger, Yves Pelletier, Louis Poulin, David Richardson, Kelvyn Robertson, Kiyo Sato, Peiliang Shi, Dave Stepaniak, Simon Thompson, Hao Tian, Helen Titley, Joerg Urban, Hannah Wilcox, Nathan Wilhelmi, Xin Yang, Shintaro Yokoi, Tom Yoksas, and many other colleagues from the data providers and archive centres for contributing to the TIGGE project and/or providing material to this paper.

REFERENCE*S*

Champion, A, 2008: Properties of ensemble forecasts from different weather centres. Master Thesis dissertation, available from University of Reading.

Ebert, B., Z. Toth, M. Charles, and G. Ross, 2008: An XML format for cyclone analyses and forecasts. Available at *http://www.bom.gov.au/bmrc/projects/THORPEX/CXML/CXMLspecification.doc*

GEO, 2007: Group on Earth Observations Work Plan, 2007-2009. Available at http://www.earthobservations.org/documents/wp0709_v6.pdf

He, Y., Wetterhall, F. Cloke, H.L. Pappenberger, F. Wilson, M. Freer K. and McGregor, G. (2009), Tracking the uncertainty in flood alerts driven by grand ensemble weather predictions, Meteorological Applications, 15 - in press

Johnson, C. and R. Swinbank, 2008: 'Medium-range multi-model ensemble combination and calibration', *Q. J. R. Meteor. Soc.* (submitted). Also issued as Forecasting Research Technical Report no. 517, Met Office.

Matsueda, M. and H.L Tanaka, 2008: 'Can MCGE outperform the ECMWF ensemble?' SOLA, **4**, 77-80. doi:10.2151/sola.2008-020.

Pappenberger, F., J. Bartholmes, J. Thielen, H. L. Cloke, R. Buizza, and A. de Roo, 2008: 'New dimensions in early flood warning across the globe using grand-ensemble weather predictions', *Geophys. Res. Lett.*, **35**, L10404, doi:10.1029/2008GL033837.

Park, Y.-Y., R. Buizza, and M. Leutbecher, 2008: 'TIGGE: preliminary results on comparing and combining ensembles', *Q. J. R. Meteor. Soc.*, 134, 2029-2050.

Peña, M., and H. van den Dool, 2008: Consolidation of Multi Model Forecasts by Ridge Regression: Application to Pacific Sea Surface Temperature. J. Climate, in press, available at: *http://ams.allenpress.com/perlserv/?request=res-loc&uri=urn%3Aap%3Apdf%3Adoi%3A10.1175%2F2008JCLI2226.1*

Richardson, D., R., Buizza and R. Hagedorn, 2005: Final report of the 1st Workshop on the THORPEX Interactive Grand Global Ensemble (TIGGE). WMO TD No. 1273, WWRP-THORPEX No. 5

TIP, 2005: THORPEX International Research Implementation Plan, World Meteorological Organization (WMO), WMO/TD-No. 1258 WWRP/THORPEX No. 4, available at: *http://www.wmo.int/pages/prog/arep/wwrp/new/documents/CD_ROM_implementation_plan_v1.pdf*

Titley, H., N. Savage, R. Swinbank and S. Thompson, 2008: Comparison between Met Office and ECMWF medium-range ensemble forecast systems, Meteorology R&D Technical Report no. 512, Met Office.

Toth, Z., J. Desmarais, G. Brunet, Houtekamer, Y. Zhu, R. Wobus, R. Hogue, R. Verret, L. Wilson, B. Cui, G. Pellerin, B. Gordon, E. O'Lenic, D. Unger, 2005: The North American Ensemble Forecast System (NAEFS). Proceedings of the 1st THORPEX International Science Symposium, December 2004, Montreal, Canada. Available at *http://wwwt.emc.ncep.noaa.gov/gmb/ens/papers/ toth_naefs_thorpex_montreal.pdf*

Table 1: TIGGE portals and data providers

TIGGE Archive Centres and Data Portals

- China Meteorological Administration (CMA) *http://wisportal.cma.gov.cn/tigge/*
- National Centre for Atmospheric Research (USA) *http://tigge.ucar.edu*
- European Centre for Medium-range Weather Forecasts (ECMWF) *http://tigge-portal.ecmwf.int*

Centres supplying daily forecasts to the TIGGE archive

- European Centre for Medium-range Weather Forecasts (ECMWF)
- US National Centers for Environmental Prediction (NCEP)
- Meteorological Service of Canada (MSC)
- Australian Bureau of Meteorology (BoM)
- China Meteorological Administration (CMA)
- Brazilian Centra de Previsao de Tempo e Estudos Climatico (CPTEC)
- Japan Meteorological Administration (JMA)
- Korea Meteorological Administration (KMA)
- Météo-France
- UK Met Office

Table 2: Agreed list of parameters and units to be delivered to the TIGGE database. Note that temperature, u-velocity, v-velocity and specific humidity are provided on the following isobaric surfaces: 1000, 925, 850, 700, 500, 300, 250 and 200 hPa. The geopotential height is provided on the same surfaces plus 50 hPa. All parameters have to be provided six-hourly, included the initial time of the forecast. All fluxes are accumulated since the beginning of the forecast.

Parameter	**Unit**
Surface level parameters	
Mean sea-level pressure	Pa
Surface Pressure	Pa
10-m u velocity	m s^{-1}
10-m v velocity	m s^{-1}
Surface temperature	K
Surface dew point temperature	K
Surface max temperature	K
Surface min temperature	K
Skin temperature	K
Soil moisture	kg m^{-3}
Soil temperature	K
Total precipitation (liquid + frozen)	kg m^{-2}
Snow fall water equivalent	kg m^{-2}
Snow depth water equivalent	kg m^{-2}
Total cloud cover	0–100%

Total column water	kg m^{-2}
Time-integrated surface latent heat flux	W m^{-2} s
Time-integrated surface sensible heat flux	W m^{-2} s
Time-integrated surface net solar radiation	W m^{-2} s
Time-integrated surface net thermal radiation	W m^{-2} s
Time-integrated outgoing long-wave radiation	W m^{-2} s
Sunshine duration	s
Convective available potential energy	J kg^{-1}
Convective inhibition	J kg^{-1}
Orography (Geopotential height at the surface)	m
Land-sea mask	0-1
Parameters on isobaric surfaces	
Temperature on 8 isobaric surfaces	K
Geopotential height on 9 isobaric surfaces	m
U-velocity on 8 isobaric surfaces	m s^{-1}
V-velocity on 8 isobaric surfaces	m s^{-1}
Specific humidity on 8 isobaric surfaces	kg kg^{-1}
Parameters on potential temperature surfaces	
Potential vorticity on θ = 320 K surface	K m^2 kg^{-1} s^{-1}
Parameters of potential vorticity surfaces	
Potential temperature on 2PVU surface	K
U-velocity 2PVU surface	m s^{-1}
V-velocity 2PVU surface	m s^{-1}

Table 3: Parameter availability and configuration of ensemble for each data provider.

	BOM	CMA	CMC	CPTEC	ECMWF	JMA	KMA	MF	NCEP	UKMO
Standard Fields (Out of 73 requested)	**55**	**60**	**56**	**55**	**70**	**61**	**46**	**62**	**69**	**70**
Ensemble Members	**33**	**15**	**21**	**15**	**51**	**51**	**17**	**11**	**21**	**24**
Forecast Length (Day)	**10**	**10**	**16**	**15**	**15**	**9**	**10**	**3**	**16**	**15**
Forecast cycles per Day	**2**	**2**	**2**	**2**	**2**	**1**	**2**	**1**	**4**	**2**

Figure legends

Figure 1: (a) Protocols for exchange of data between Data Providers and Archive Centres. (b) Number of registered TIGGE users (by country).

Figure 2: Spread of the mean sea-level pressure for the various TIGGE ensembles as function of the forecast range and the latitude. For a fair comparison only the first ten members of each ensemble have been used. The period covered is December 2008.

Figure 3: Root- mean-squared error of the ensemble mean (sea-level pressure) for the various systems (the computation was not possible for the Australian system as no verification analysis is available for this period).

Figure 4: RMSE and RPSS for four different ensemble systems: ECMWF alone and non bias-corrected (dashed line), bias-corrected ECMWF (solid black), ECMWF + UKMO bias-corrected (solid grey), ECMWF + UKMO + JMA + CMA bias corrected (dotted). Results on 86 cases from June, July and August 2007.

Figure 5: Improvements in RPS of 51-, 154- and 279-member multi-model ensembles relative to the ECMWF ensemble for 500hPa heights over 20°-90°N for three-month periods from December 2006 to November 2007. From Matsueda and Tanaka (2008).

Figure 6: Brier skill scores for a) mean sea level pressure greater than the climatological mean, b) 2m temperature greater than the climatological mean, c) 2m temperature greater than 90th percentile. In addition to the individual systems (ECMWF, UKMO and NCEP), three almost equivalent variants of the multi-model system are shown (multi, weighted and adjusted). The data are globally averaged over 120 days ending 29 April 2008. From Johnson and Swinbank (2008).

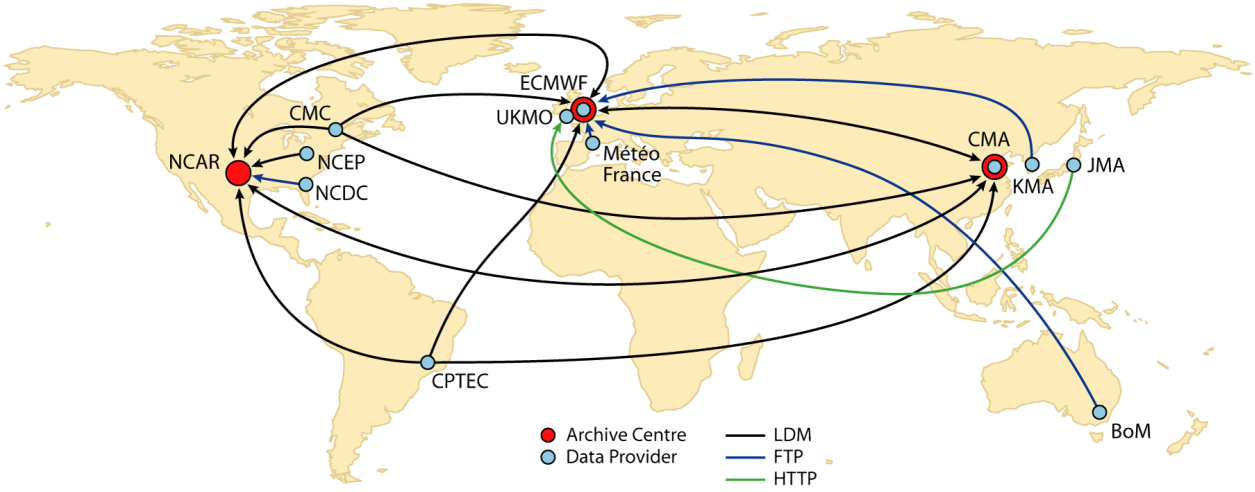
Figure 7: Flow discharge "observed" and predicted by several TIGGE systems (called here Systems I to VII) and one multi-model system (Grand Ensemble) for a point on the river Jiu (in Romania) where flooding was observed. The 5th and 95th percentile of river discharge predictions are shown for the different forecasts with a 5-day lead time. The dashed horizontal lines show four classic flood warning thresholds. "Observed" discharges refer in fact to simulations forced by observed rainfall. From Pappenberger et al. (2008).

Figure 8: Performance of the individual TIGGE systems for several characteristics of extra-tropical cyclones (courtesy of Lizzie Froude). All results are for the northern hemisphere extra-tropics and cover the 6-month period from 1st February to 31st July 2008. The verification is the ECMWF analysis. Upper left: mean vorticity bias of the ensemble (excluding the control). A positive (negative) bias in vorticity corresponds to an over (under) prediction of cyclone intensity. Upper right: mean propagation speed bias of the ensemble (excluding the control). The negative bias in propagation speed corresponds to the forecast cyclones propagating too slowly. Lower Left: Ensemble mean error in cyclone position. Lower Right: Ensemble mean error in cyclone intensity. The ensemble mean error is calculated by taking the mean of the ensemble member storm tracks and then calculating the error compared to the corresponding analysis storm track. For cyclone position the error is calculated as the mean separation distance between the ensemble mean and analysis tracks and for intensity it is the mean absolute intensity difference.

Figure 9: Multi-model ensemble forecast tracks (a) and strike probabilities (b) for Hurricane Ike initiated at 12UTC on 4th Sep 2008, combining outputs from the ECMWF and UK Met Office ensembles. These charts were generated at the Met Office using track data distributed using the CXML format.

Figure 1: (a) Protocols for exchange of data between Data Providers and Archive Centres. (b) Number of registered TIGGE users (by country).

(a)



(b)

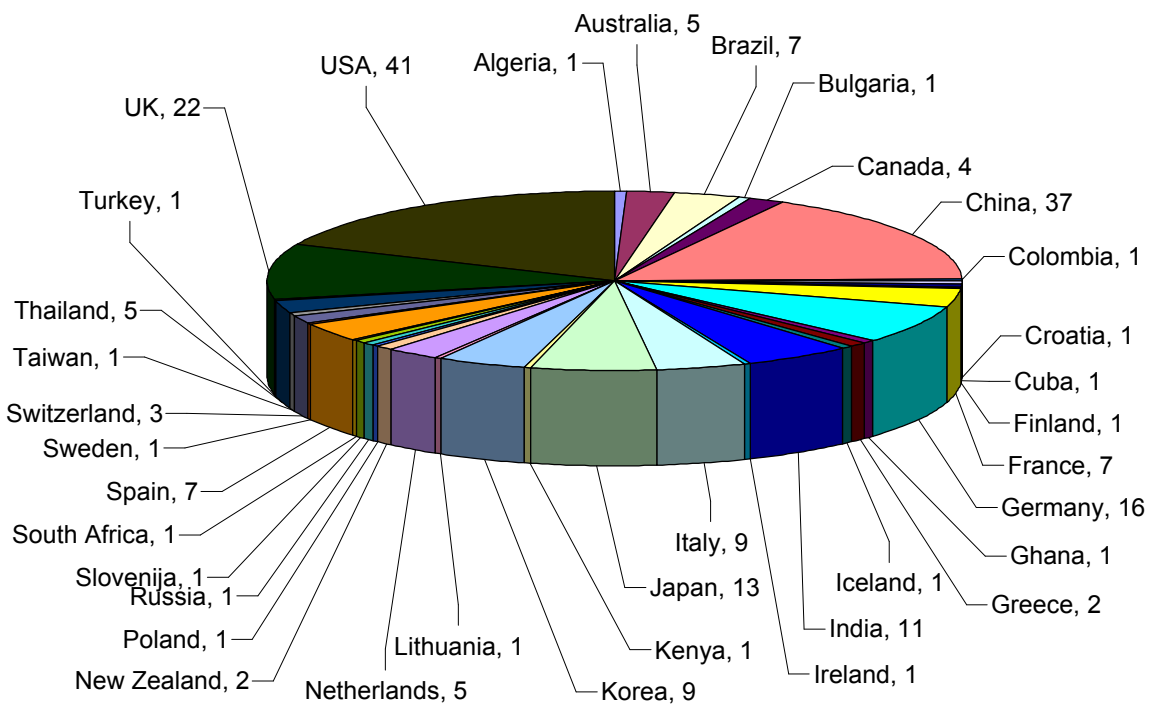


Figure 2: Spread of the mean sea-level pressure for the various TIGGE ensembles as function of the forecast range and the latitude. For a fair comparison only the first ten members of each ensemble have been used. The period covered is December 2008.

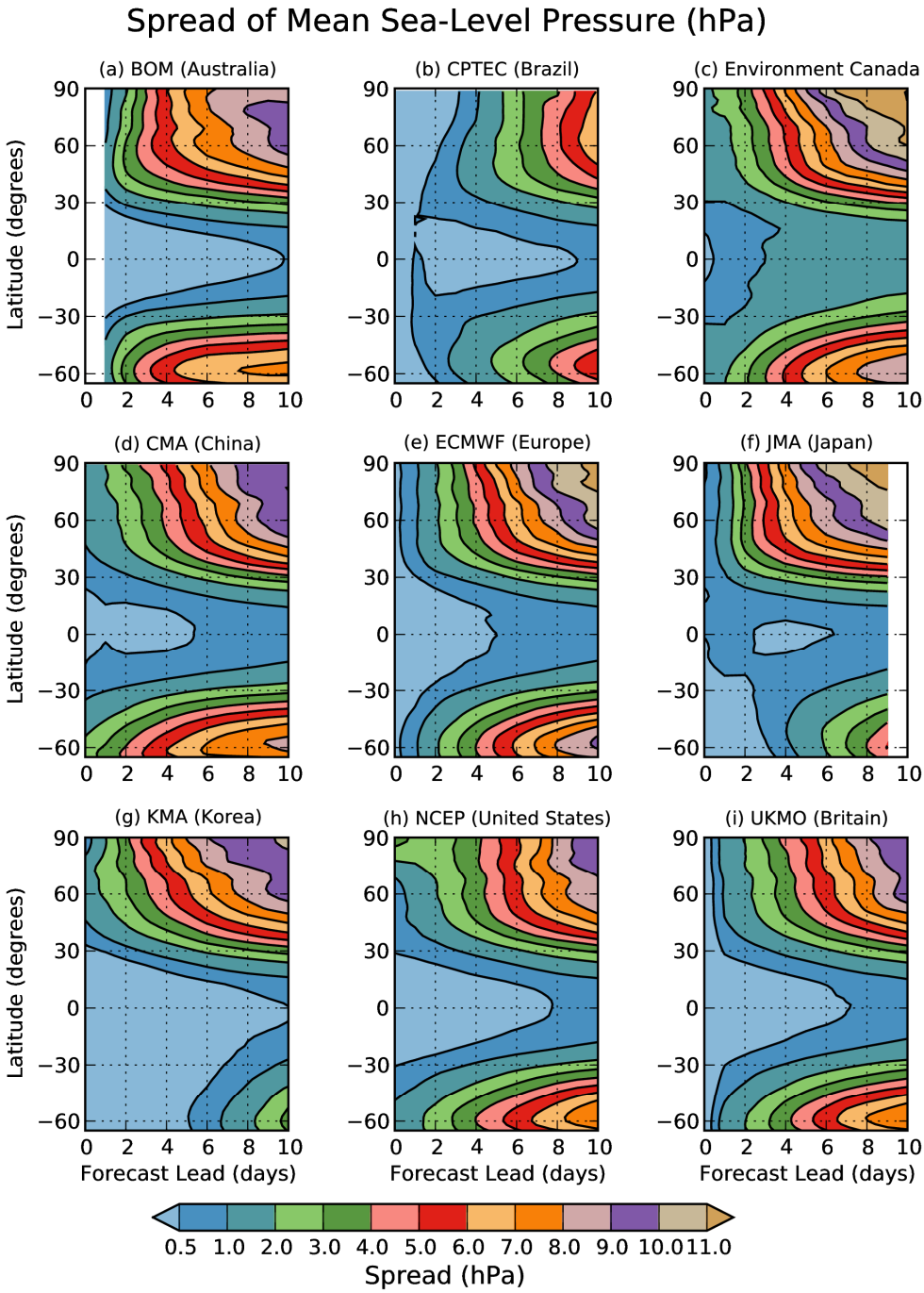


Figure 3: Root- mean-squared error of the ensemble mean (sea-level pressure) for the various systems (the computation was not possible for the Australian system as no verification analysis is available for this period).

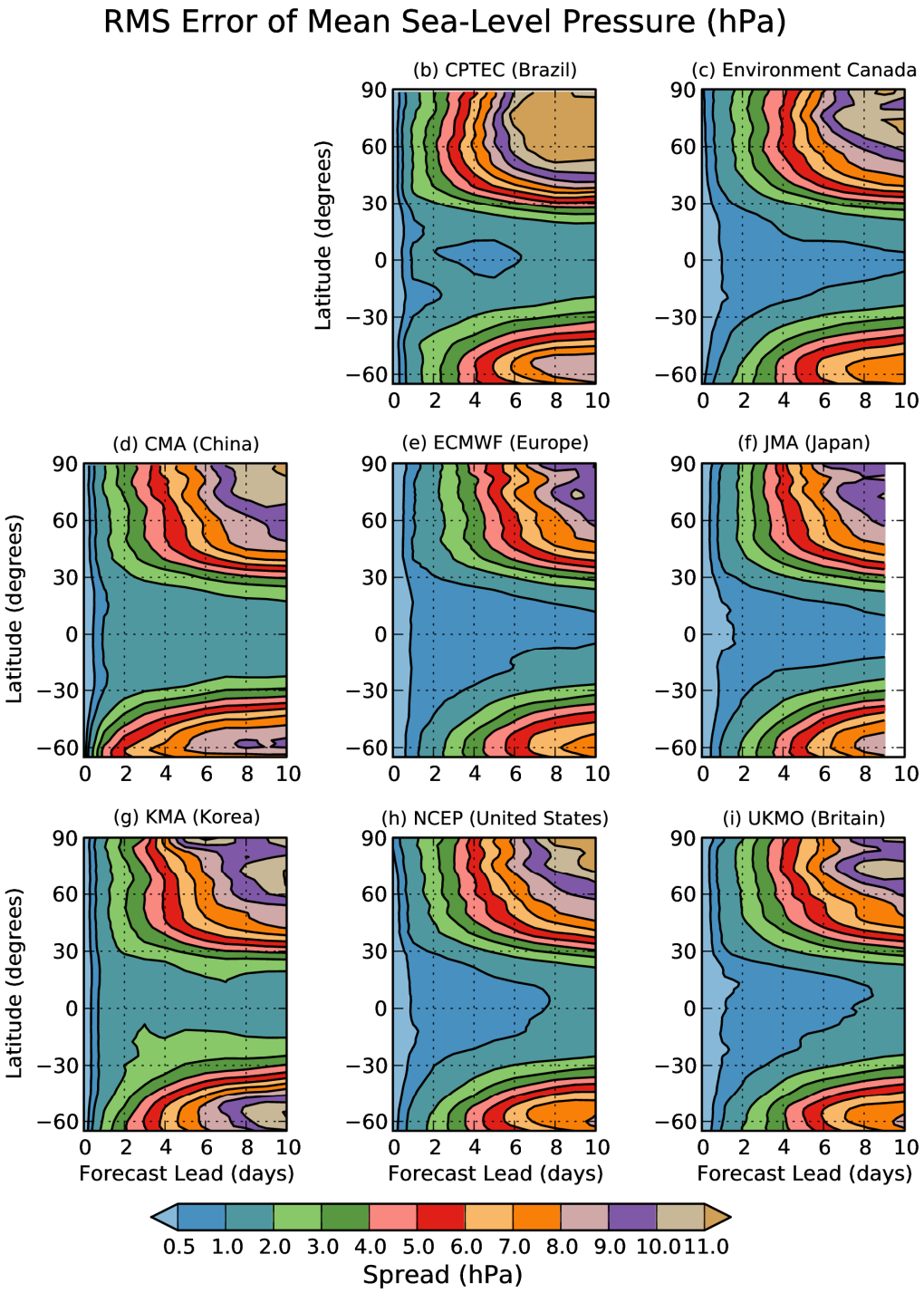


Figure 4: RMSE and RPSS for four different ensemble systems: ECMWF alone and non bias-corrected (dashed line), bias-corrected ECMWF (solid black), ECMWF + UKMO bias-corrected (solid grey), ECMWF + UKMO + JMA + CMA bias corrected (dotted). Results on 86 cases from June, July and August 2007.

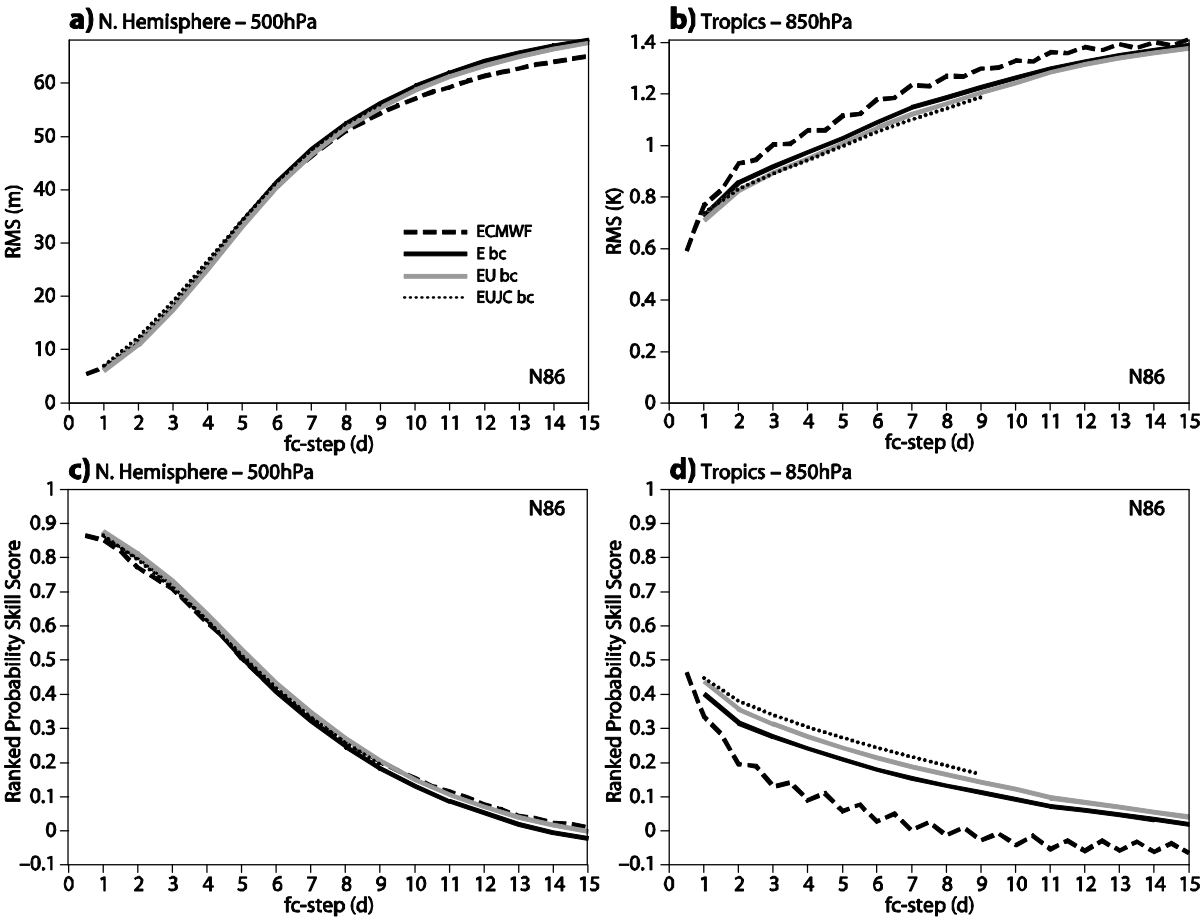


Figure 5: Improvements in RPS of 51-, 154- and 279-member multi-model ensembles relative to the ECMWF ensemble for 500hPa heights over 20°-90°N for three-month periods from December 2006 to November 2007. From Matsueda and Tanaka (2008).

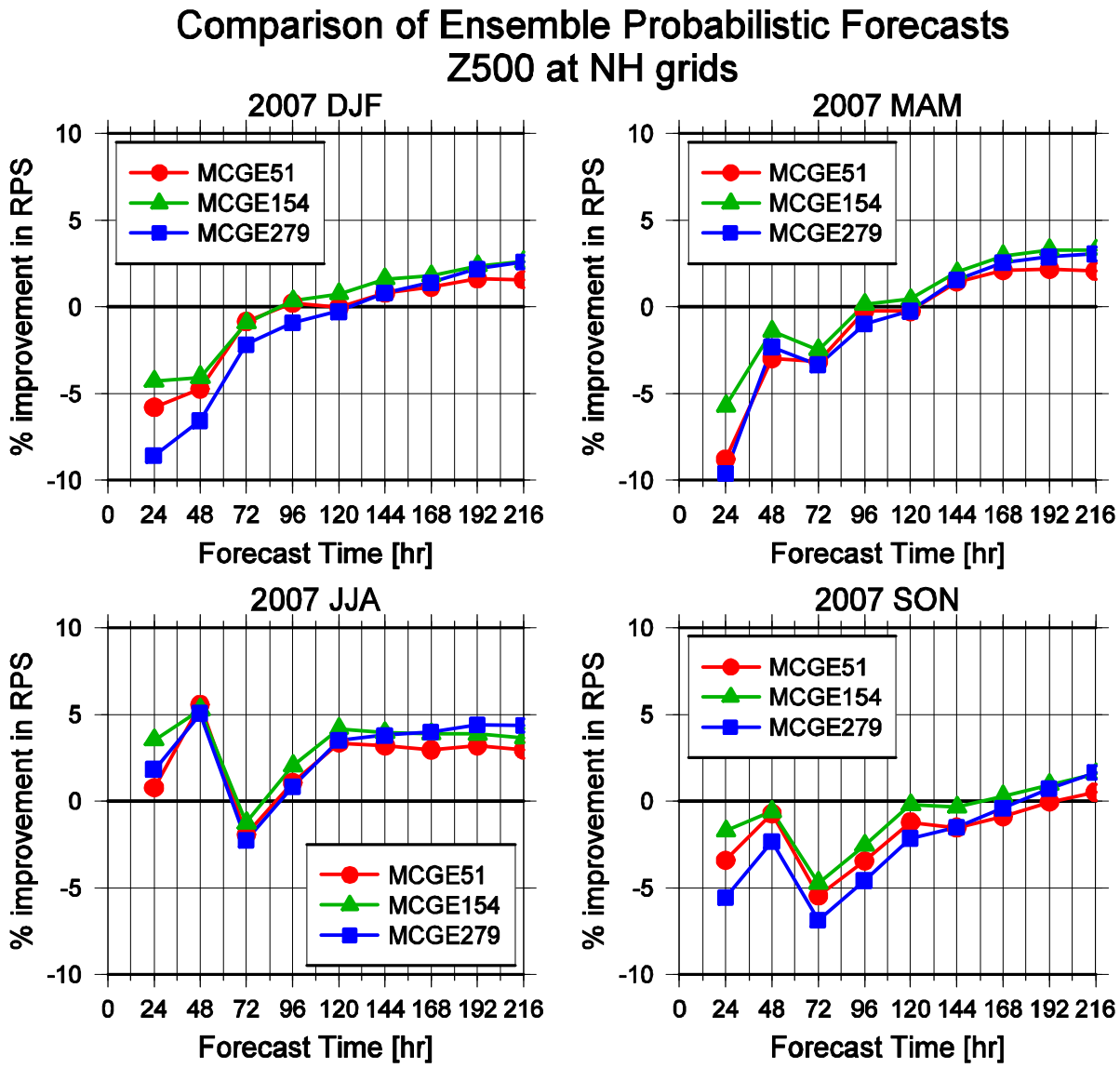


Figure 6: Brier skill scores for a) mean sea level pressure greater than the climatological mean, b) 2m temperature greater than the climatological mean, c) 2m temperature greater than 90th percentile. In addition to the individual systems (ECMWF, UKMO and NCEP), three almost equivalent variants of the multi-model system are shown (multi, weighted and adjusted). The data are globally averaged over 120 days ending 29 April 2008. From Johnson and Swinbank (2008).

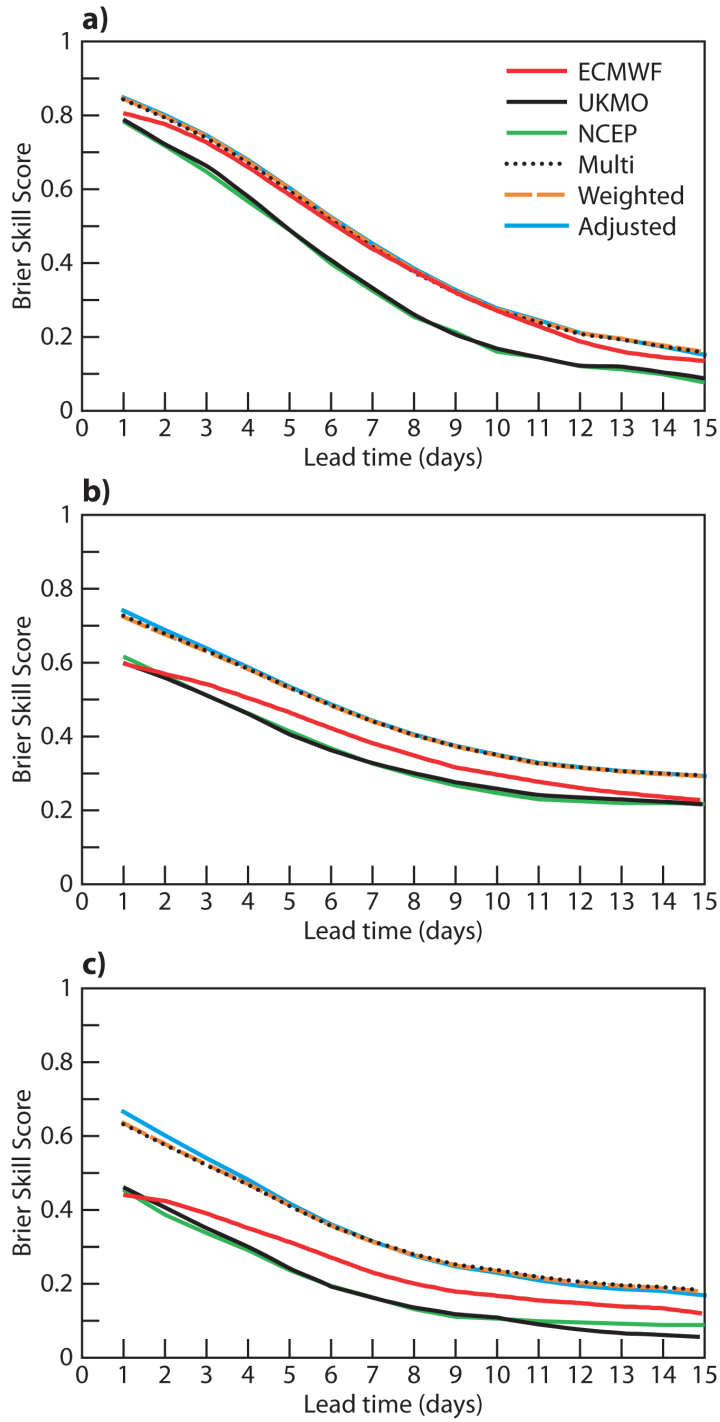


Figure 7: Flow discharge "observed" and predicted by several TIGGE systems (called here Systems I to VII) and one multi-model system (Grand Ensemble) for a point on the river Jiu (in Romania) where flooding was observed. The 5th and 95th percentile of river discharge predictions are shown for the different forecasts with a 5-day lead time. The dashed horizontal lines show four classic flood warning thresholds. "Observed" discharges refer in fact to simulations forced by observed rainfall. From Pappenberger et al. (2008).

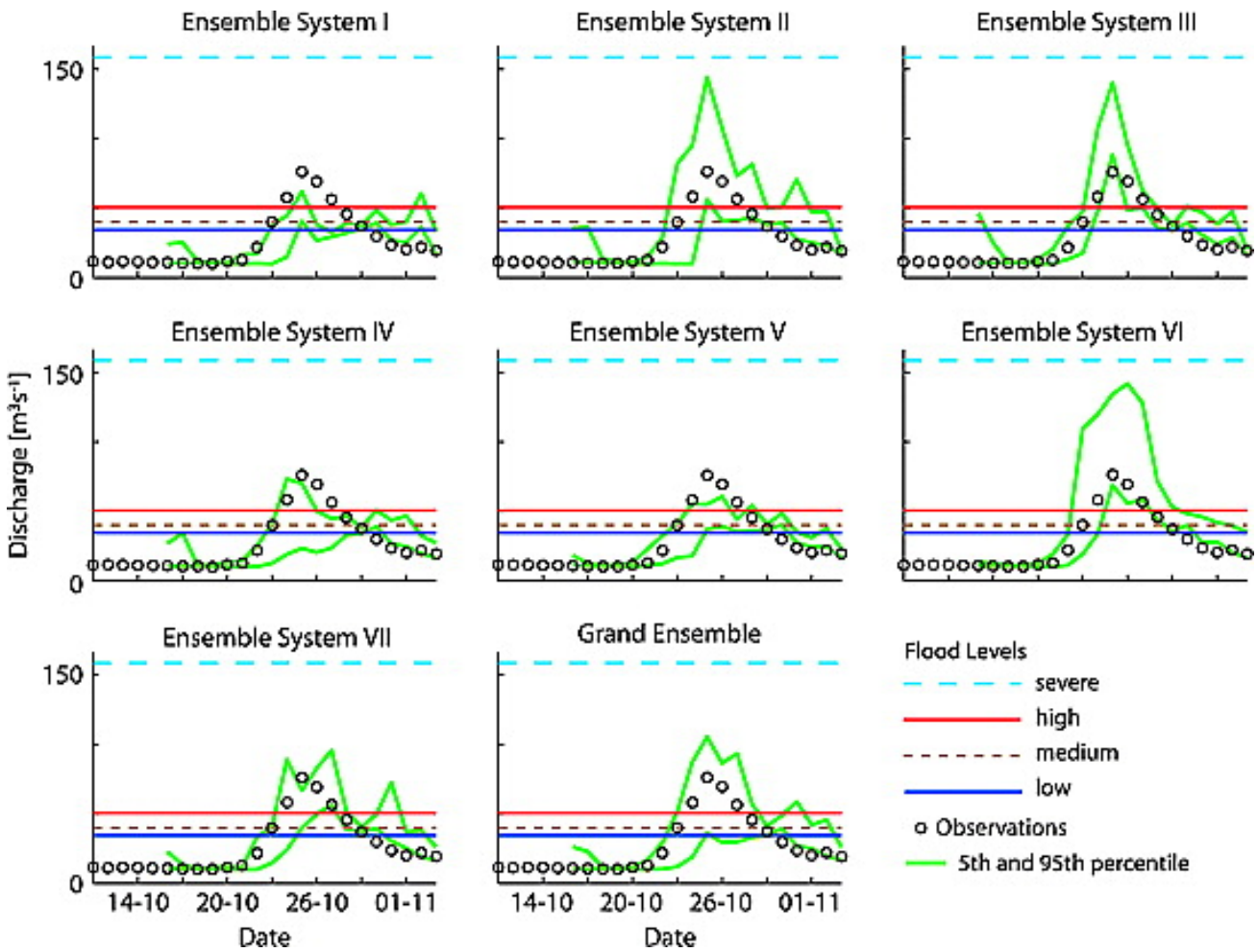


Figure 8: Performance of the individual TIGGE systems for several characteristics of extra-tropical cyclones (courtesy of Lizzie Froude). All results are for the northern hemisphere extra-tropics and cover the 6-month period from 1st February to 31st July 2008. The verification is the ECMWF analysis. Upper left: mean vorticity bias of the ensemble (excluding the control). A positive (negative) bias in vorticity corresponds to an over (under) prediction of cyclone intensity. Upper right: mean propagation speed bias of the ensemble (excluding the control). The negative bias in propagation speed corresponds to the forecast cyclones propagating too slowly. Lower Left: Ensemble mean error in cyclone position. Lower Right: Ensemble mean error in cyclone intensity. The ensemble mean error is calculated by taking the mean of the ensemble member storm tracks and then calculating the error compared to the corresponding analysis storm track. For cyclone position the error is calculated as the mean separation distance between the ensemble mean and analysis tracks and for intensity it is the mean absolute intensity difference.

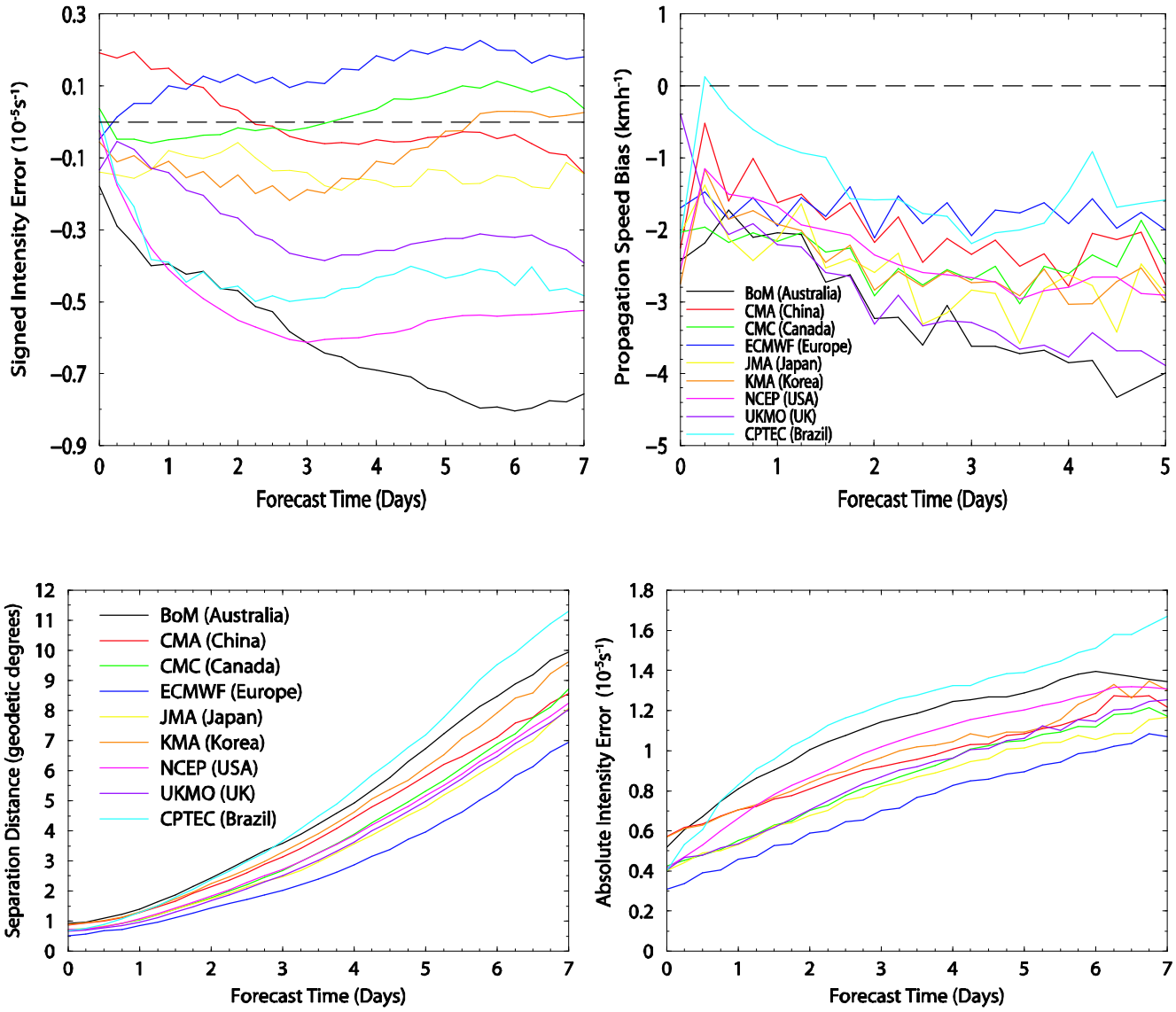
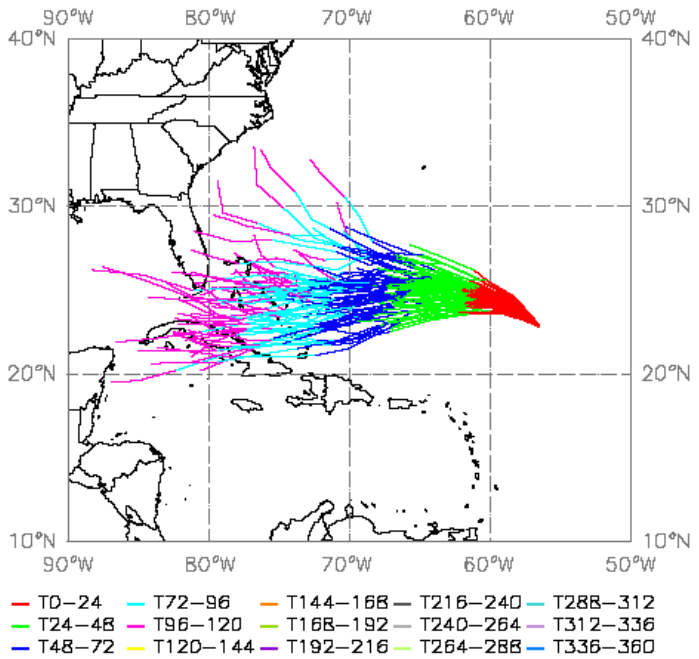
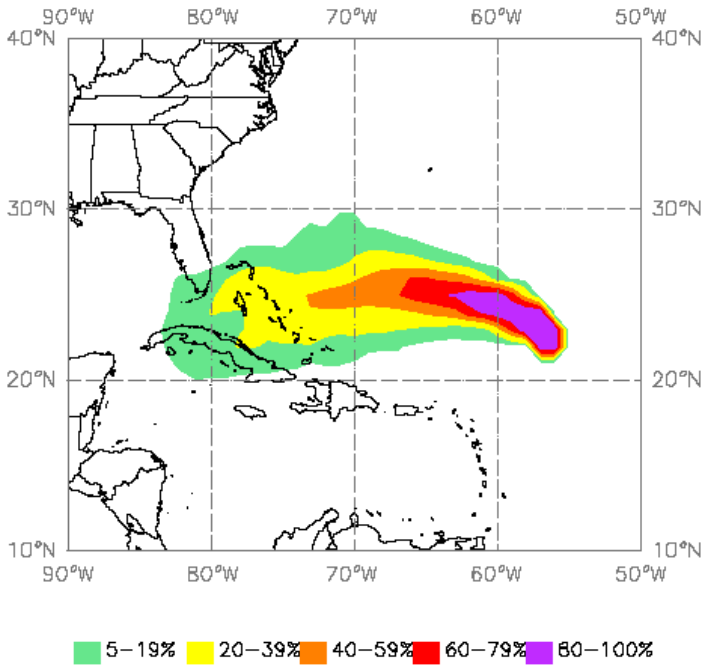



Figure 9: Multi-model ensemble forecast tracks (top) and strike probabilities (bottom) for Hurricane Ike initiated at 12UTC on 4th Sep 2008, combining outputs from the ECMWF and UK Met Office ensembles. These charts were generated at the Met Office using track data distributed using the CXML format.