WORLD METEOROLOGICAL ORGANIZATION

WORLD WEATHER RESEARCH PROGRAMME

THORPEX
A World Weather Research Programme

FIRST WORKSHOP ON THE THORPEX INTERACTIVE GRAND GLOBAL ENSEMBLE (TIGGE)

FINAL REPORT

WMO/TD-No.1273

WWRP/THORPEX No. 5

Reading, United Kingdom, 1-3 March 2005

This report was edited by *David Richardson* (Met Office), *Roberto Buizza* (ECMWF) and *Renate Hagedorn* (ECMWF) and compiled with the contribution of some participants, invited speakers and members of the organizing committee of the 1st TIGGE Workshop held at ECMWF between the 1st and the 3rd of March 2005: *H Boettger* (ECMWF), *P Bougeault* (ECMWF), *E Ebert* (BMRC, Australia), *M Ehrendorfer* (Univ. Innsbruck, Austria), *T Hamill* (NOAA/CDC), *F Lalaurette* (ECMWF), *K Mylne* (Met Office), *T Paccagnella* (ARPA/SMR, Italy), *M Roulston* (Penn State University, US), *M Shapiro* (NOAA), *Y Takeuchi* (JMA, Japan), *J Thielen* (JRC, Italy), *Z Toth* (NCEP, USA), *L Treinish* (IBM, US), and *L Wilson* (MSC, Canada).

Executive Summary

TIGGE, the THORPEX Interactive Grand Global Ensemble, is a key component of THORPEX: a World Weather Research Programme to accelerate the improvements in the accuracy of 1-day to 2-week high-impact weather forecasts for the benefit of humanity.

The key objectives of TIGGE are:

- An enhanced collaboration on development of ensemble prediction, internationally and between operational centres and universities
- New methods of combining ensembles from different sources and of correcting for systematic errors (biases, spread over-/under-estimation)
- A deeper understanding of the contribution of observation, initial and model uncertainties to forecast error
- A deeper understanding of the feasibility of interactive ensemble system responding dynamically to changing uncertainty (including use for adaptive observing, variable ensemble size, on-demand regional ensembles) and exploiting new technology for grid computing and high-speed data transfer.
- Test concepts of a TIGGE Prediction Centre to produce ensemble-based predictions of high-impact weather, wherever it occurs, on all predictable time ranges
- The development of a prototype future Global Interactive Forecasting System

This report summarizes the key recommendations of the ***1st TIGGE Workshop***, held at ECMWF from 1 to 3 March 2005. This workshop aimed to address the strategy required to achieve these goals, with a particular focus on establishing the user requirements for TIGGE and proposing an infrastructure design to meet these requirements.

It has been concluded that to achieve TIGGE's key objectives, it is necessary to:

- Determine the user requirements for TIGGE data, including the types, volumes, format of data, access methods and timeliness
- Design the TIGGE infrastructure to meet these requirements
- Determine resource requirements and secure necessary funding if required
- Establish commitments from data contributors, TIGGE archive centres and prospective users
- Implement proposed infrastructure, collecting, archiving, and providing access to TIGGE data
- Develop and maintain close links with TIGGE users, including other THORPEX sub-programmes, field campaigns and Demonstration Projects, as well as other partners
- Have the flexibility to respond to evolving user needs as scientific understanding increases during the project

It is also necessary that the following data be shared between TIGGE users:

- Ensemble forecasts generated routinely (often operationally) at different centres around the world. This is the core data of the TIGGE archive. The total daily data volume is expected to be around 200GB, based on a preliminary list of required parameters developed at the workshop.
- Observational data and existing datasets including re-analyses and re-forecasts
- Additional special datasets generated during the TIGGE project for specific research and applications.

With respect to data policy, it is suggested that:

- TIGGE data should be available to all users for research purposes
- Consideration needs to be given to the issue of real-time access to data, in particular for demonstration projects and field experiments
- The process of obtaining approval for data access should be transparent, stream-lined and reasonably fast

- The user interface for access to the central archives should be user-friendly and should make it as easy as possible for researchers in disciplines not used to dealing with exceptionally high data volumes to obtain subsets of ensemble data.
- Open-source sharing of post-processing software (calibration, combination, decision-making) should be promoted in order to maximise benefit for both researchers and end-users.

A TIGGE infrastructure needs to be established rapidly to meet specific requirements. In particular it is proposed to use TIGGE's capability during the International Polar Year 2007-08 and during the Beijing Olympics in 2008, and to link TIGGE with the North American Ensemble Forecast System (NAEFS), which is planned to become fully operational by March 2006. The database needs to be available to allow research to begin in advance of the above applications, for the TIGGE research projects and for users in other THORPEX sub-programmes

Given the above requirements for data policy and access, it is suggested that the TIGGE infrastructure is developed in two phases:

- *Phase-1*, during which data are collected in near-real time (via internet ftp) at a small number of central TIGGE data archives. This can be implemented now at little cost and could handle the estimated 200 GB per day data volumes with current network and storage capabilities
- *Phase-2*, during which data archives are distributed over a number of repositories, instead of all being held centrally, but efficient and transparent access to users is maintained. This is a more flexible solution with the potential to eliminate routine transfers of large data volumes. But this will require substantial software development over a number of years, in coordination with the WMO Information System, and will require additional funding

To meet the time constraints of user access to TIGGE, it is proposed that Phase-1 is started as soon as possible and that planning for Phase-2 begins in parallel. In particular, the following outline timetable is proposed for the Phase-1 development and implementation.

Firstly, the THORPEX GIFS-TIGGE Working Group that is intended to coordinate the TIGGE activity should be established as soon as possible.

Secondly, in 2005, the GIFS-TIGGE WG should:

- Complete detailed planning (including costs and resource) and initiate Phase-1
- Secure commitment from institutes willing to act as central data archives for TIGGE's Phase-1 and from initial contributors to TIGGE database
- Agree data exchange protocols and data formats
- Finalise list of datasets and parameters to store in the TIGGE archive or provide alternative access to, based on provisional lists given in this report
- Help and encourage contributing centres to begin development to provide data in agreed formats
- Address data access policy, including consideration of requirements for real-time access, particularly for field campaigns and demonstration projects
- Consider the strategy required to initiate work on Phase-2 of TIGGE, including funding and resources
- Work within THORPEX and with external partners to develop detailed planning for the first real-time demonstrations using the TIGGE infrastructure (including evaluation of potential use during IPY and Beijing Olympics)

It is recommended that a follow-on workshop be held as soon as possible for contributors and Phase-1 archive centres to achieve the relevant above targets.

Thirdly, in 2006:

- The Phase-1 TIGGE data archives will begin collecting available ensemble contributions in near-real time

- Links will be available to other already established data archives
- Research access will be available to TIGGE data (through the existing access infrastructure of the archive centres)
- Work will begin on Phase 2 software developments (once required funding is secured)

The above timetable should enable TIGGE to contribute to proposed real-time THORPEX support for:

- International Polar Year field campaigns in 2007-08
- Beijing 2008 Olympics WWRP Research and Development project

The workshop considered that achieving Phase-1 within the above time constraints is feasible, but challenging, and depends on immediate action and commitments from participation centres.

It is suggested that this report is treated as a detailed proposal for the execution of the TIGGE project, and as such it is discussed by the WMO THORPEX Executive Board and by the WMO International Core Steering Committee. It is also suggested that this report is made available to the other THORPEX Working Groups, which are expected to oversee much of the research undertaken using the TIGGE data, and that the considerations outlined in this report are considered during the planning of the research projects promoted/managed by these sub-programmes.

Table of Contents

1 Introduction: THORPEX and TIGGE

Two major research challenges of the 21st century are to reduce and mitigate natural disasters and to realise the societal and economic benefits of improved weather forecasts. THORPEX (THe Observing system Research and Predictability Experiment) is a response to these challenges. THORPEX will provide the research underpinning the WMO Multi-Hazard Prevention Strategy to halve the number of deaths due to meteorological, hydrological and climate-related natural disasters over the next 15 years. THORPEX will contribute to the development of a future truly integrated Global Interactive Forecast System (GIFS), that will form the basis for the weather component of the WMO-lead Global Early Warning and Response System for multi-hazard prevention. The GIFS will provide appropriate risk management tools for decision makers of all nations.

The THORPEX Interactive Grand Global Ensemble (TIGGE) will be developed as a resource to support the substantial research effort needed to evaluate the potential of a Global Interactive Forecast System. TIGGE will provide a framework for international collaboration in the development of ensemble prediction systems (EPSs) and will facilitate the design of the main prediction tools in the THORPEX Forecast Demonstration Projects.

1.1 THORPEX research

The core research objectives of THORPEX are described within four sub-programmes: Predictability and Dynamical Processes; Observing Systems; Data Assimilation and Observing Strategies; and Societal and Economic Applications. Developments in all these areas will be needed to achieve the goal of developing truly interactive forecast systems in which observations, their use in data assimilation, and the design of the numerical modelling system are adapted to the requirements of the users of the forecasts.

The scientific research needed, and the strategy to achieve this, are given in the following documents, available from http://www.wmo.int/thorpex/publications.html:

- M. A. Shapiro and A. J. Thorpe, 2004: THORPEX International Science Plan. Version 3. 2 November 2004, WMO/TD-No. 1246, WWRP/THORPEX No. 2
- D. P. Rogers et al., 2005: THORPEX International Research Implementation Plan. Version 1. 14 February 2005. WMO/TD-No. 1258, WWRP/THORPEX No. 4

1.2 TIGGE

The THORPEX Interactive Grand Global Ensemble, TIGGE, is introduced in the THORPEX Science Plan:

> *"THORPEX will develop, demonstrate and evaluate a multi-model, multi-analysis and multi-national ensemble prediction system, referred to as the THORPEX Interactive Grand Global Ensemble (TIGGE). The TIGGE will integrate user requirements for forecast information, developments in observing systems, targeting, adaptive data assimilation, and model improvements into a multi-model/multi-analysis ensemble prediction system. TIGGE data bases, containing individual model EPS analyses and forecasts will be required to facilitate research on the design of best configuration of multi-model/multi-analysis ensemble forecast systems. TIGGE prototype forecast systems, resulting from this research, would be used to produce experimental real-time forecasts. The skill of these experimental forecasts will be evaluated relative to existing techniques. These forecasts should be tailored according to the severity of the predicted weather hazards. An operational forecast system of this design would allow international resources to be brought to bear on the most critical environmental forecast problems of the day."*

The TIGGE concept is expanded in the THORPEX Implementation Plan; the main points are summarised below. More information about TIGGE can be found in Chapter 8 of the THORPEX Implementation Plan.

1.2.1 Basic structure

The initial basic components of TIGGE will be global ensembles run to around 14 days, including those run currently at a number of operational centres (see list of current systems at Appendix B). These will be collected in near real time and stored in a common format in a number of central data servers for access by researchers in operational centres and the academic communities. This will facilitate research on combination and inter-comparison of different systems; it will become straightforward, for example, to compare the value of multi-model ensembles with those based on perturbations of a single model. Easy access to long series of data is necessary for applications such as bias correction and the optimal combination of ensembles from different sources; these will be accumulated, albeit with ever improving data assimilation and modelling techniques, as the TIGGE archive grows during the project.

1.2.2 Interactive forecast system

The concept of interactivity will be tested in the TIGGE framework. As a response to the day to day changes in uncertainty, extra observations could be called on in ‘sensitive areas’, ensemble size and resolution adjusted, and regional ensembles run as and when needed; all these adapting in real time to meet user needs. Research on the scientific aspects of interactivity that can be carried out through TIGGE include:

- Are there different ‘optimal’ configurations for different situations (e.g. relative benefits of resolution and ensemble size); how would the requirements vary from day to day.
- Links between global and regional ensembles, including the provision of appropriate boundary conditions, how global ensembles can be used for ‘on-demand’ requests for regional ensembles, whether global ensembles can guide the real-time composition of regional ensembles (‘representative members’).
- Observation targeting - some methods of sensitive area prediction such as the ETKF will benefit directly from the large ensemble size available in TIGGE; other methods will not rely directly on TIGGE. Other aspects of targeting are covered elsewhere in the THORPEX plans.

1.2.3 THORPEX field demonstration projects and real-time use

TIGGE data could be used in real time during the THORPEX field demonstration projects (FDPs). The timeframe for these FDPs will allow for 2-3 years of near-real-time TIGGE development before real-time availability is required. By the time of THORPEX FDPs (in ~2008) enough data and experience will have been gained to use TIGGE in real time as the main forecast tool, providing both global and regional quantifications of weather uncertainties to use in decision-making.

1.3 Expected outcomes of TIGGE

The overall expected outcomes of the TIGGE project are stated in the THORPEX Implementation Plan

- Enhanced collaboration on development of ensemble prediction, internationally and between operational centres and universities
- New methods of combining ensembles from different sources and of correcting for systematic errors (biases)

- Understanding of the feasibility of interactive ensemble system responding dynamically to changing uncertainty (including use for adaptive observing, variable ensemble size, on-demand regional ensembles) and exploiting new technology for grid computing and high-speed data transfer.
- Test (through FDPs) concept of TIGGE Prediction Centre to produce ensemble-based predictions of high-impact weather, wherever it occurs, on all predictable time ranges.
- A prototype future Global Interactive Forecasting System.

1.4 Links between TIGGE and other THORPEX sub-programmes

TIGGE will fulfil several roles within THORPEX:

- It will provide a framework to address scientific questions of predictability and ensemble design
- It will provide a resource for research being carried out under the four research sub-programmes
- It will provide a facility to evaluate the concept of a Global Interactive Forecast System (GIFS)
- It will contribute key prediction tools to the THORPEX field demonstration projects

Close collaboration will be maintained throughout the project between all THORPEX sub-programmes. It is anticipated that much of the research carried out using the TIGGE data will be done under the auspices of other sub-programmes.

The TIGGE data facility will need to be designed to meet the differing requirements of each of these activities. Outlines of the research issues and resources for each activity are given in the THORPEX Science and Implementation Plans. These provide initial guidelines for the development of the TIGGE user requirements. It is expected that the THORPEX Working Groups that will be established for each sub-programme will further refine the requirements in each area.

Since significant work is needed to develop the TIGGE infrastructure, and this is a necessary pre-requisite for much of the THORPEX research, work on TIGGE needs to begin as soon as possible. While ideally detailed specifications would be obtained from each prospective user group of TIGGE, it is recognised that some of these will take time to establish. The TIGGE Workshop has drawn on experts in each THORPEX research area to obtain the best current estimate of user requirements. It is anyway likely that user requirements will evolve during the lifetime of THORPEX as we better understand the scientific issues. The TIGGE design will reflect this expectation.

2 The 1st TIGGE Workshop

The 1st TIGGE workshop was held at ECMWF from 1 to 3 March 2005. The workshop was initiated by the WMO THORPEX programme, and was co-sponsored by ECMWF, The UK Met Office and WMO. Seventy delegates from international organizations, national and regional meteorological and hydrological services, universities and private companies attended it (Appendix A includes a full list of participants).

2.1 Workshop objectives

The workshop was held to discuss the scientific aims and user requirements in more detail with representatives of the expected TIGGE user community, and to make sure that the TIGGE infrastructure accurately meets the users' needs.

One of the objectives of the workshop was to produce a report outlining user requirements and the infrastructure design, to be presented to the WMO THORPEX Executive Board and International Core Steering Committee (ICSC). This report could be treated as a detailed proposal for the execution of the TIGGE project. The report could also be made available to the other THORPEX Working Groups, which are expected to oversee much of the research undertaken using the TIGGE data. The considerations of this workshop may be useful in planning of these research sub-programmes.

2.2 Organisation of the workshop

During the first half of the workshop, 14 invited speakers gave overview talks on some key aspects of TIGGE, such as the design of global and regional ensemble prediction systems, the simulation of initial and model uncertainties, the combination and calibration of forecasts produced by different systems, the verification of weather forecasts, high-impact weather and societal and economic applications, infrastructure and architecture of forecasting systems (Appendix A includes the workshop programme).

During the second half of the workshop, delegates joined four working groups designed to discuss the following issues:

- WG1: Design of TIGGE experiments and model systems
- WG2: Post-processing and verification
- WG3: Applications in TIGGE
- WG4: Infrastructure of TIGGE

Plenary sessions were held to provide feedback between the working groups and to draw some conclusions on three key areas: key scientific questions, ensemble design experiments and societal and economic applications. Conclusions and recommendations on these three areas are summarised, respectively, in sections 3, 4 and 5.

3 TIGGE's key scientific questions

The overall purpose, goals and expected outcomes of TIGGE, as given in the THORPEX Science and Implementation Plans are summarised in Section 1. The workshop considered the strategy required to achieve these goals. As outlined in Section 1, TIGGE is expected to be a resource for a wide range of users. Research on fundamental issues of predictability and ensemble design, information on analysis uncertainty, potential of interactive forecast systems and applications to users will all make use of TIGGE data. Sections 5 and 6 review the requirements in detail from the perspectives of ensemble design and end-user applications. Consideration is focused particularly on issues that need access to multi-centre ensemble data. However, the infrastructure and access facilities provided as a necessary part of the TIGGE infrastructure will have more general application to a broad range of THORPEX research activities that do not specifically require multi-ensemble data. This may be particularly valuable for users outside the main operational centres, such as academics, and those working on end-user applications.

During the workshop, the following questions were discussed:

- Is there value in combining ensemble forecasts from multiple NWP centres?
- Is model-related uncertainty better captured in multi-centre ensembles?
- Can an 'optimal' approach to the simulation of observation./initial uncertainties be identified?
- What is the best way to combine ensembles from multiple centres?
- Will less skilful sub-components add value to an overall product?
- Does a multi-centre EPS possess value beyond an ensemble of high-resolution multi-centre operational control forecasts?
- To what extent are multi-centre perturbations appropriate?
- Can we learn about analysis uncertainty from TIGGE and hence improve DA procedures and existing EPSs?
- Do the benefits of combining ensembles from multiple centres justify the additional operational costs?
- Can a single system be improved to a higher level by learning from a multi-centre EPS?

During the workshop, it was pointed out that these questions are relevant for a variety of situations, such as tropical storm track forecasts, or other high-impact weather situations, and that the answers may be different for different applications. It was also pointed out that the interfacing of LAMs with the global TIGGE data set was considered especially important for high-impact weather.

The workshop noted that answers to the questions posed above would depend on the time and space scales and lead time considered, and also on the type of high-impact events studied. Metrics for the evaluation of the results should include those related to practical applications. Particular weather phenomena suggested for further studies, including those with LAMs, included tropical storms, heavy precipitation events, etc., possibly in selected "test-beds" over specified areas and time periods, to be agreed upon by the collaborating investigators. It is the expectation of the workshop that the scientific investigations and experiments carried out with the help of the TIGGE data archive will also contribute to advancing our knowledge of atmospheric predictability.

4 Potential TIGGE ensemble investigations

TIGGE offers a unique opportunity to investigate ensemble forecasting related issues that can be addressed only through the use of multi-centre ensemble prediction data sets.

4.1 Investigations with routine TIGGE data

In order to address the key scientific questions listed in section 3, the workshop proposed the following experiments for the consideration of the THORPEX research community.

The first group of experiments are based on the use of the routine ensemble data sets that are proposed to be part of the TIGGE archive, and should address the following two key issues:

- *Is forecast uncertainty better captured with a multi-centre approach?*
    - How much analysis/forecast error is explained by single vs multi-centre ensembles?
    - To what extent do perturbations describe analysis error? (Check perturbations against known properties of analysis error).
    - Does the multi-centre approach provide a better description of uncertainty associated with moist/precipitation processes?
- *Explore and explain cases when sub-components of the TIGGE ensemble generated by different models bifurcate.*

4.2 Global new-data experiments

The second group of experiments considered by the workshop involved the coordinated generation of additional ensemble forecast data as part of a multi-centre inter-comparison.

4.2.1 Multi-Model / Multi-Analysis Ensemble (MUMMA)

A set of cross-model cross-analyses experiments was proposed, in which each centre's model would be run from the analyses of each of the other centres as well as from its own analysis. If N centres participate, this would yield NxN forecasts. The properties of the N forecasts from each initial condition may provide further insight into characteristics of model error. Consideration of the N forecasts using the same model may provide information about the impacts of using the different analyses available at the various centres. Integrations can be carried out in near-real time (preferred for operational feedback) or retrospectively for limited period of time (30-90 days?), possibly at low resolution (e.g. ~T150), to limit cost requirements. Difficulties in actually interfacing analyses and models were recognized and discussed. This experiment was referred to as MUMMA .

In particular, the workshop considered MUMMA suitable for testing the following science hypotheses related to multi-centre ensemble forecasting:

- The control initial fields (operational analyses) from the different NWP centres provide a good description and sampling of uncertainty in the initial conditions (when the same model is used)
- The use of different models provides a good description of model related forecast uncertainty (when the same initial conditions are used)
- The complete MUMMA ensemble will better describe forecast uncertainty in general as compared to an ensemble generated at a single centre.
- The combination of analyses from multiple centres provides an improved estimate of the initial state, and better quantification of the analysis uncertainty

The immediate requirement to enable MUMMA is that analyses from contributing centres are made available within TIGGE on ~20 pressure levels at the resolution available at contributing centres.

It was pointed out that one of the key difficulties in realizing this project would be the treatment of surface field: one approach that could be followed is to integrate each model using upper-air fields from analyses generated by the other centres, but using surface data from its own analysis. The fact that this project may face this and other technical problems, and require considerable resources should be considered in the planning phase. It was also mentioned that a similar initiative might be being undertaken or considered by WGNE, and has been suggested that before proceeding with any MUMMA project, the potential link with WGNE is investigated.

4.2.2 Controlled initial perturbation experiments

An experiment using a single model (possibly developed and/or installed at the Earth Simulator) to integrate all available perturbation fields was proposed. This would allow:

- the comparison of the multi-centre TIGGE ensemble with a same-size single-model system in which only the initial conditions are different between members
- The comparison of the effectiveness of different perturbation techniques that are used operationally at different centres, eliminating the effect of different initial conditions and NWP models as used in operations (Buizza et al 2005, Bourke et al. 2004)
- evaluation of the relative importance of ensemble size and resolution by for example running the system with 350 members at T100 and with 50 members at T200 (i.e. comparing systems with similar computational cost).

4.3 Links between global and regional EPSs

In view of the emphasis of THORPEX on high-impact weather, the interfacing of global ensemble prediction systems with LAM ensembles was considered to be of particular importance. In that context, a key question that TIGGE should address is the following:

- Will the use of TIGGE data as boundary conditions in LAM ensembles (for existing systems, see table in Appendix B) lead to enhanced guidance for high impact weather events?

Regional models require lateral boundary conditions from a global model at high temporal resolution and either on all model levels or on pressure levels, depending on the specifications of the regional model. Transmission and archiving of this data (full model fields for all ensemble members) is not currently feasible for TIGGE. However, the following proposals are feasible and should be considered:

- Retain the required global fields at the originating centre for a limited time (e.g. 2 days), during which LAM groups can access the data if they wish (for example if the situation is particularly interesting). Archiving of these data would be the responsibility of the LAM group; the data would not be archived by the global centre.
- Retain for a longer period (and possibly transmit and archive in TIGGE) a limited and pre-defined set of "frames" that provide the necessary boundary conditions for given regions. This is significantly less data than the full fields but does not allow flexible choice of LAM region.
- For particular high-impact cases, chosen in real time, archive the full model fields in TIGGE. This would allow many LAM groups to carry out case study work on significant events.
- For a pre-selected period (e.g. 1 month), archive required global fields at originating centres and transfer (by tape) to TIGGE archive, to allow intensive study of use of TIGGE global models to drive regional ensembles.

The first proposal in the above list is a potential real-time/operational configuration. Such a system provides part of the interactive component of TIGGE since it allows for specific on-demand action whenever severe weather seems imminent. It is suggested that in the planning phase these proposals are discussed with groups and consortia involved with limited area activities (e.g. HIRLAM, COSMO-LEPS).

5 Societal and economic applications of TIGGE research

There is a wide range of users in both the public and private sectors who could potentially act as partners in translating the information contained in TIGGE products into socio-economic benefits. Organizations in the following broad areas were identified by the workshop as being possible partners in TIGGE applications research:

- *Civil Protection and Aid Agencies*: e.g. preparation for and response to floods, storms, tropical and mid-latitude cyclones, heat-waves, wild-land fires, snowfall, landslides, atmospheric release of hazardous materials, and diseases with a weather-dependent component (including airborne disease).
- *Resource and infrastructure management:* e.g. water management, energy management, agriculture, forestry, fishing, health services, transportation, construction, retail, finance, tourism and sporting events.

It is anticipated that these sectors could exploit medium-range forecast information in various ways, uncluding:

- Preparation for deployment of resources: e.g. emergency disaster aid, forest fire-fighters, and equipment.
- Planning and scheduling of possible actions: e.g., evacuation planning, elective procedures in hospitals, making additional staff and equipment available at short notice, power-station maintenance.
- Monitoring of potential events: e.g. placement of mobile Doppler radar units.
- Scheduling of personnel shifts: e.g. city workers for ploughing and gritting.
- Increasing frequency of local, high-resolution modelling and/or monitoring: e.g. increase frequency of local flood forecasting models.

5.1 Obstacles to Application of Ensemble Forecasts

To allow and encourage these potential users to participate effectively in TIGGE-oriented research, the workshop identified several issues that need to be addressed:

- A lack of awareness of ensemble-based medium-range forecasts among potential end-users.
- A tradition of end-users dealing with forecasts presented deterministically and a lack of experience in exploiting probabilistic forecast skill.
- The need for end-users to formulate appropriate responses to relatively low probability, albeit highly significant, warnings issued at lead times of several days.
- A lack of verification of ensemble forecasts against quantities of interest and relevance to end-users.
- The non-quantitative formulation of some ensemble-based products.
- Restricted access to ensemble data products.
- The high data volumes associated with ensemble forecasts, and data interfaces that are unsuitable for end-users.
- The availability of ensemble-based forecasts in forms suitable for end-users and specific weather-risk problems. This is both a problem of visualization and dissemination and also applies to deterministic forecasts.
- An incomplete characterization of the global climatology of severe weather in conjunction with the model climatologies of severe weather.
- Periodic changes and updates to numerical models that possibly diminish the robustness of post-processing and verification methods.

5.2 Strategies for Promotion of Socio-economic Applications

The workshop made the following recommendations and suggestions for how the issues outlined above might be addressed:

- Awareness and visibility of ensemble-based forecasts should be raised. Two means of furthering this aim could be the creation and dissemination of brochures and establishing a web site aimed at potential end-users.
- A workshop format should be developed to stimulate discussion between meteorologists and potential end-users in how information contained in ensemble forecasts might be translated into socio-economic benefits. Such workshops should be held in multiple locations and targeted on specific applications or sectors to reach communities of potential end-users.
- Collaborative projects between meteorologists, academics and end-users in the public and private sectors should be promoted. Projects such as the Direct and Inverse Modelling in Environmental End-to-End Prediction (DIME) project at the London School of Economics and the European Flood Alert System (EFAS) developed and tested during the European Commission's Framework 6 programme may provide useful models for this type of joint research.
- Post-doctoral researchers should be placed within end-user organizations to facilitate a two-way exchange of information and experience concerning the integration of ensemble-based forecasts into decision-making. Such efforts should yield an understanding of the decision-making process, quantitative assessment of the impact of forecast information for specific weather-sensitive decisions, and a definition of the metrics for measuring forecast utility.
- An online catalogue of databases of observations relevant to potential end-users should be established. Examples include hydrological observations and energy demand.
- An effort to characterize the global climatology of severe weather in both observations and the models should be made to help end-users interpret low probability forecasts of severe events and allow tools such as the Extreme Forecast Index to be extended to the whole globe.
- The TIGGE project should be used as an opportunity to study and possibly quantify the impact of numerical model changes on the value of post-processing techniques.
- The TIGGE project should develop an early prototype deployment of some subset of the above ideas. This will serve to demonstrate the concepts and potential value to the intended beneficiaries as well as provide a test-bed for the technical and scientific requirements.

Implications for the design of the TIGGE facility are included in Section 8 on access and data policy.

5.3 Data Requirements for Socio-Economic Applications

The draft list of variables given in section 9.7 was considered by the workshop to be acceptable for socio-economic applications research, including driving hydrological models. In addition, it would be advantageous to have a higher temporal frequency than 6 hours in the first 72 hours of the forecasts. The actual temporal frequency should be consistent with the spatial resolution of the model.

If ensemble forecast information is to be translated into socio-economic benefits it is of paramount importance that all barriers to the use of TIGGE data be minimised, whether these barriers are administrative, technical or caused by a lack of understanding as to what the data are and how they should, and should not, be used. The success of socio-economic applications will depend upon data being available, accessible and usable.

6 TIGGE's links with other projects

TIGGE could benefit from strong links with existing international projects, in particular with close interactions with the North American Ensemble Forecast System, hydrological projects, monthly/seasonal projects and ad-hoc demonstration projects. Some ideas on why these links should be promoted are highlighted hereafter.

6.1 The North American Ensemble Forecast System (NAEFS)

The North American Ensemble Forecast System (NAEFS) is a multinational project being developed by the Meteorological Service of Canada (MSC), the National Meteorological Service of Mexico (NMSM), and the US National Weather Service (NWS), with the UK Met Office joining in 2006. The project involves the exchange and sharing of global ensemble forecasts among the participating countries. The NAEFS' Initial Operating Capability was established in September 2004.

NAEFS plans to become fully operational by March 2006, and further refinements are expected to be implemented in the following years. Each set of ensemble forecasts (currently from MSC and NWS) will be statistically post-processed to reduce systematic errors. The bias-corrected constituent ensembles will be combined into a joint ensemble that will be used to generate forecast products for intermediate and end users. All software products used in the processing of ensemble forecast data, as well as in the generation of products are shared among the participating organizations, and used consistently at each centre. As a result, forecasts issued by the participating national meteorological services will be seamless both in time and space.

TIGGE will co-ordinate with NAEFS, and will benefit from the research, development and experience of multi-model ensembles accumulated within that project. Not surprisingly, the TIGGE database and the associated research plans share many of the design features of the NAEFS project, since both are based on the use of combined multi-centre ensemble forecasts. The new techniques and tools that are being developed for NAEFS will also be shared with the THORPEX research and user community. It is expected that TIGGE-related THORPEX research will greatly expand the capabilities of multi-centre ensemble forecasting and that the enhanced techniques will find their operational use in the NAEFS and its successors.

The workshop considered it important to minimise any duplication of effort between TIGGE research activities and the operational NAEFS. Wherever possible, infrastructure should be integrated, software shared and research and development activities co-ordinated between the two projects. This needs to be addressed at the outset of TIGGE to avoid costly parallel developments. Formal contact between TIGGE/THORPEX and NAEFS should be established as soon as possible to discuss the issues highlighted above.

6.2 Links with hydrological projects (EFAS and HEPEX)

TIGGE could foster the development of closer links between meteorological and hydrological institutions, thus helping the realization of integrated meteo- and hydro-logical forecasting systems. To investigate these opportunities, TIGGE should explore the possibility to establish links with two existing meteo- and hydro-logical initiatives:

- EFAS, the European Flood Alert System
- HEPEX, the Hydrological Ensemble Prediction Experiment

These two initiatives aim to bring the international hydrological and meteorological communities together to develop reliable forecasting tools that can be used with confidence by the emergency management and water resources sectors to make decisions that have important consequences for public health and safety, and the economy. The TIGGE data-set should be designed taking into consideration the requirements of these and similar initiatives.

6.3 Links to monthly/seasonal data and COPES

Co-ordination between THORPEX and the WCRP/COPES initiative, to address the observational and modelling requirements for the prediction of weather and climate for two weeks and beyond, is one of the core objectives stated in the THORPEX Science Plan.

The workshop noted that a number of centres are exploring the concept of “unified” or “seamless” forecast production spanning a range of timescales from days to seasons. End users would benefit from a more consistent and standardised product suite. Potential benefits for modelling centres are both scientific and technical, with expected resource savings.

There is significant potential for mutual benefits in infrastructure, data access and applications development from collaboration and coordination between activities on different time-scales. The links being established between THORPEX and COPES should include consideration of these issues at an early stage.

6.4 THORPEX Campaigns and Demonstration Projects

THORPEX will conduct regional and global campaigns to demonstrate and evaluate new observing technologies and interactive forecast systems. THORPEX will explore the opportunities to carry out these campaigns in conjunction with major international programmes such as the International Polar Year (IPY) and the African Monsoon Multi-disciplinary Analysis (AMMA).

THORPEX Demonstration Projects are specifically intended to benchmark the societal and economic benefits of the THORPEX programme.
Possible demonstrations include

- The Beijing Olympics in 2008, in collaboration with the WWRP Research and Development Project (B08RDP) on use of short-range ensemble forecasts
- Demonstration of the usefulness of ensemble forecasts for risk-reduction strategies in African countries most vulnerable to weather-related diseases
- Improvements of forecasts of sand and dust storms (in collaboration with the WWRP Sand and Dust Storm Experiment)

It is important that assessment of the benefits to end users (through coupling to impacts models or specific user applications) is included in at least some of these demonstrations.

The Campaigns and Demonstrations will provide guidance to relevant bodies including the WMO Commission for Basic Systems (CBS), the World Weather Watch (WWW) and operational forecast centres on improvements to forecast systems, and optimisation of global and regional observing-systems.

TIGGE will be a major resource for many of these field campaign and demonstration activities. The TIGGE framework may, for example, be able to provide:

- Straightforward, consistent and real-time access to required forecast information to all involved in a given activity
- Advance warning of potential significant events based on multi-centre ensemble predictions
- Specific indications of where and when additional observations would be particularly useful (by generating and providing access to sensitive area prediction data through the TIGGE infrastructure).
- Predictions of when it may be especially important to be able to run regional/mesoscale ensembles (this might give some flexibility if a centre is not able to commit to continuous running of a regional system).

- Provision of boundary conditions for high resolution regional ensembles

The appropriateness and feasibility of these options will need to be assessed for each activity.

7 Calibration, combination and verification of TIGGE's forecasts

Once a set of ensemble forecasts has been generated it then becomes necessary to verify the forecasts. TIGGE will archive forecasts and observations to enable individual users to perform a wide range of diagnostics and verification studies. If the verification indicates that there are systematic errors then some adjustment (calibration) is necessary to optimize the forecasts. Because TIGGE will involve multiple ensemble forecast systems, techniques will be needed to combine them into synthesized forecast products. To assist TIGGE researchers in their investigations and data users in their choice of optimal ensemble configuration, it is recommended that a library of open-source verification, calibration, and combination routines be made available along with the data.

7.1 Observations and analyses

Observations and analyses of meteorological variables must form an important part of the TIGGE dataset. When verifying against analyses a weighted blend of analyses from several operational centers may be a fairer verification field than single-model analyses to avoid incestuous false skill, particularly in the early forecast range and in data-sparse regions (this is a debatable issue that should be investigated under TIGGE).

In order to produce statistically significant verification results, particularly for severe weather and other rare events, it is necessary to have a large number of samples spanning both time and space. However, to explore the dependence of forecast performance on region, season, etc., it is also desirable to have sample sets that are relatively homogeneous. Care must be taken to balance these two considerations.

Verification techniques are now beginning to account for uncertainty in the verification data. Therefore observation and analysis error statistics should also be part of the TIGGE dataset. Research should continue on investigating the best ways to make use of information on observational uncertainty. Use of re-sampling techniques in verification can help to quantify the impact of small data samples for rare events, and also of inhomogeneities in forecast datasets caused by changes in operational ensembles.

7.2 Verification

To provide information on forecast quality to help address many of the scientific questions listed in Section 3, TIGGE should host (or be instrumental in the establishment of) a website that contains a limited set of verification statistics for individual and multi-model ensembles. It is recommended that the WMO standard for ensemble forecast verification (section III of the WMO Manual of GDPS, Attachment II.7, Table F) be used as a baseline. A verification website following this standard is already hosted for WMO by the Japan Meteorological Agency and could form a prototype for a TIGGE website. The recommended scores and diagrams for probability forecasts are:

- Reliability diagrams
- ROC diagrams
- Brier skill score
- (Continuous) Ranked Probability skill score
- Relative economic value

Ensemble spread should be assessed using:

- Rank histogram
- Spread vs. skill

The WMO standard defines "reliability tables" (contingency tables) as the mechanism for exchange of results, from which all of the probabilistic scores and diagrams can be calculated. In addition to these probability-based metrics the verification should also include assessment of the ensemble mean in a similar fashion to traditional deterministic verification:

- RMSE
- Anomaly correlation
- Spatial variance

The WMO standard variables to be verified include:

- PMSL +/- 1, +/- 2 S.Ds w.r.t. climatology
- Z_{500} +/- 1, +/- 2 S.Ds w.r.t. climatology
- $|V|_{850}$ with thresholds of 10, 15, 25 m/s
- T_{850} anomalies with 2, 4, 8 C w.r.t. climatology
- Precipitation at 1, 5, 10, 25 mm/24h verified against observations from the GCOS network

To provide further verification for weather near the surface, which is where the greatest impact is felt, we additionally recommend the verification of:

- T_{2m} +/- 1, +/- 2 S.Ds w.r.t. climatology
- $|V|_{10m}$ with thresholds of 5, 10, 15 m/s

The climatology referred to by the WMO standard is a "centre specified climatology". Since the TIGGE dataset will include analyses of these fields from multiple operational centres it would be preferable to use a weighted blending of operational analyses as the reference state used for verification and for definitions of climatology. Alternatively, a re-analysis dataset (e.g. ERA-40) could be used for climatology. For station verification, the long-term station climatology should be used

Verification software, compatible with the TIGGE archive dataset, should be prepared for the variables and scores in the above list. This will provide the basis for consistent evaluation of TIGGE and needs to be done as soon as possible, and made available to all participants in TIGGE. This may eventually be a more consistent and effective way of carrying out the verification both for WMO and for TIGGE research projects, and could be automated to update monthly on the TIGGE dataset.

Statistics will be accumulated monthly for forecast periods of 24h, 48h, etc. Monthly results will be saved in a format that enables aggregation to longer time periods. The spatial domains over which verification should be done are specified following the WMO standard:

- Northern Hemispheric extratropics (90°N – 20°N)
- Tropics (20°N – 20°S)
- Southern Hemispheric extratropics (20°S – 90°S)

The reporting of confidence intervals with the verification statistics is highly recommended to help assess whether estimated ensemble skill is statistically significant, and whether differences in ensemble performance are significant.

The verification products specified above represent only a basic set. TIGGE also encourages the development and use of additional verification methods including feature-based and user-oriented verification methods.

Verification of extreme (rare) events remains a difficult problem, but one which is extremely important for meeting THORPEX goals. These fall well outside of the ±1, ±2 standard deviations w.r.t. long-term climatology specified above. Because of the small sample size, and also because of difficulties in making accurate observations in extreme events, there are large uncertainties associated with the verification results. Rather than relying on uncertain statistics, research groups should examine the TIGGE ensembles for the severe weather events that are of particular interest to them. Further research, possibly involving the statistical community, is required to develop methods of estimating ensemble skill for rare events.

7.3 Calibration

Here, calibration refers to the statistical adjustment of ensemble forecasts based upon a comparison of prior forecasts and observations. Calibration can apply corrections for biases in both the mean forecasts from individual models and the spread in ensembles. Within TIGGE research is particularly required in calibration methods for multi-model ensembles, which is closely linked with work on combination described in section 7.4 below.

There is a tension between the desire of operational forecast centers to calibrate using short training periods and the potential for greater improvement by using longer training periods (*Hamill et al.* 2004). TIGGE will include NOAA/CDC's reforecast dataset (*Hamill et al.* 2005) which will enable comparisons between probability forecasts made using low resolution models with large training datasets and those made using the latest EPS systems (single and multi-model) but with shorter training sets.

7.4 Combination

Combination refers to the generation of probabilistic forecast products based on multiple ensemble forecast systems with variable quality. TIGGE will investigate methods for optimal combination, building on related work done by the seasonal forecasting community. Some techniques such as Bayesian model averaging comprise both calibration and combination. As with verification and calibration, there is a tension between the need for large datasets to ensure statistical significance and avoid over-fitting based on small samples, and the need for systems to adjust to changes in operational ensembles.

8 User requirements: access and TIGGE's data policy

The workshop made the following recommendations for data policy and access to TIGGE data:

- The data policy should ensure that ensemble data are available to any user willing to publish the results of research involving the data. This policy should include real-time data required for demonstrations. The process of obtaining approval for data access should be transparent, streamlined and reasonably fast
- To meet user needs, data access must go beyond simple bulk access (e.g. ftp of GriB files).
- Ensemble variables most likely to have the most direct socio-economic benefits should be re-archived in such a way that time series of forecasts for limited regions or single locations can be efficiently extracted. Self-training systems capable of making such a selection adaptively should be considered.
- Regional subsets of ensemble data should be stored at regional data centres. These regional data centres would improve the service offered to local researchers by offering a selection of product tailored for potential applications in that region and making them quicker to access in developing countries.
- The user interface for making data queries to the central and regional re-archives should be designed to be user-friendly and should make it as easy as possible for researchers in disciplines not used to dealing with exceptionally high data volumes to obtain subsets of ensemble data. This may require the support of multiple application views layered above the "meteorological access"
- Ready access to information explaining the data sets should be available. This information should be designed to be usable by non-meteorologists. Information concerning the model (centre, resolution, version etc.) used to produce data should be included as meta-data as well as the impact of those meteorological aspects of the model on end-user applications. Background information and guidelines will be required for users to make appropriate use (and understand limitations) of the data.
- Open-source sharing of post-processing software (calibration, combination, verification and decision-making) should be promoted in order to maximise benefit for both researchers and end-users.

9 Data requirements for TIGGE

The success of TIGGE depends greatly on the data sets made available to the user community. In order to achieve the goals of the project, not only the forecasts produced during the project, but also re-forecasts and observational data sets have to be provided. This section summarises the data requirements identified at the workshop, based on the scientific research and application requirements presented in other sections of this report.

9.1 Forecasts

The operationally generated ensemble forecasts will form the core of the TIGGE database. This data set will be used to address basic scientific questions related to the performance of single vs. multi-models. A preliminary list of parameters to be archived in the common TIGGE database is given in section 9.7. Those tables, specifying the proposed content of the archive separately for surface and upper air levels, should be circulated in the TIGGE community for further consideration and eventual approval by the GIFS-TIGGE working group.

In order to perform extended studies on initial/evolving perturbations and the “Multi-Analyses-Multi-Model” experiments described in section 4.2.1, the full fields of the operational control run and perturbed analyses are also needed (on model or pressure levels). The immediate and non-trivial requirement to enable MUMMA is that analyses from contributing centres are made available within TIGGE on ~20 pressure levels at the resolution available at contributing centres.

9.2 Re-Forecasts

A number of forecast data sets for past periods exist and should be offered to the TIGGE community. This includes e.g. the NOAA-CIRES Climate Diagnostic Center’s ensemble re-forecast data set, which will be part of TIGGE. This data set consists of 15-member, 15-day ensemble forecasts from 1979 - current, created from an unchanging T62 version of NCEP’s medium-range forecast model, see Hamill et al. (2004, 2005) for details. Furthermore, both historical data (including verification) from operational ensemble forecast systems (ECMWF, NCEP, etc.), and re-analysis datasets (including forecasts, e.g. ERA-40, NCEP/NCAR) could be made available for TIGGE research. It is proposed to link to all these and any other data sets (NCEP re-analyses etc.) from the TIGGE website, whereas - at least for the time being - the actual access and download of the data should be realized via the original data access mechanisms of the respective data supplier. However, if during the project specific access standards are developed, it is desirable to apply such standards to the existing data sets as well.

9.3 Initial / boundary conditions for LAMS

In order to run limited area models with initial and boundary conditions from the TIGGE data set - in general - model level data have to be provided. Since it is not feasible to transfer model level data from all participating EPS’s to the central data repositories, other strategies for supplying LAMs with the necessary forcing data have to be developed. A possible solution would be to provide the full model level data only at the site of the producing centre and only for a limited period (up to 2 days after production). The LAM groups then have to download the forcing data (global model levels or frames) directly from the individual EPS centres they are interested in. Additionally, it was mentioned at the workshop that not all LAMs need model level data, but some LAMs can be run with a suitable number of pressure level fields. In order to address such needs, the archive should contain more pressure levels than needed for diagnostic and verification purposes only.

Since one of the main goals of TIGGE is to address severe weather and LAMs are an integral part in the process of improving severe weather predictions, it might be useful to circulate a specific questionnaire amongst the LAM community for an exact definition of the specific LAM requirements.

9.4 Observations

Various observational data sets are necessary for a thorough assessment of the TIGGE forecasts. In order to account for uncertainties in observations, observational error statistics should be made available where possible. For an efficient verification process it is desirable to have easy access to the following observational data sets:

- Surface network (GTS)
- Radiosonde network
- Gauge precipitation

Furthermore, the TIGGE database / website should provide the capability to link to regional and user-specific observational data sets.

9.5 Verification output data

In order to avoid duplication of verification efforts amongst the TIGGE user community, the TIGGE database could include intermediate verification output data, e.g. contingency tables, which could be used by individual users to calculate individual scores. WMO has already initiated a similar effort for verification of global ensemble prediction systems, and JMA is hosting the corresponding website (access is restricted to the participating centres). It should be investigated whether the facilities at JMA can be used and extended to incorporate a TIGGE verification output database.

9.6 User utility functions

In order to allow and encourage more user-relevant verification, tables of individual user utility functions should be made available. However, this information should be well documented to avoid inappropriate usage.

9.7 Proposed content of TIGGE archive

The following tables contain the proposed list of variables to be archived in the common TIGGE database. The proposed content encompasses basically all variables identified for ensemble exchange between MSC and NCEP, i.e. in NAEFS (see section 6.1). However, it does not mirror the exact content and definitions of NAEFS, in particular to avoid existing inconsistencies in the definitions between NCEP and MSC. It is strongly recommended to agree on common archive definitions regarding the content, units and aggregation method of the archived fields. Furthermore, the proposed archive goes over and above the agreed exchange in NAEFS to accommodate user requirements from the application and LAM communities. It might also be necessary to adapt the output frequency of selected parameters to specific end-user needs, e.g. for the hydrological community.

9.7.1 Surface data

It is proposed to archive 19 (+2) surface fields.

Parameter	Abbrev	Level	Unit	Output frequency	Accum
Mean sea level pressure	MSLP	MSL	Pa	6h	inst
Surface Presure	SP	surface	Pa	6h	inst
10m U-velocity	10U	10m	m s**-1	6h	inst
10m V-velocity	10V	10m	m s**-1	6h	inst
2m temperature	2T	2m	K	6h	inst
2m dew point temperature	2D	2m	K	6h	inst
2m max temperature	MX2T	2m	K	6h	det_lo
2m min temperature	MN2T	2m	K	6h	det_lo
Total precipitation (liquid+frozen)	TP	surface	m	6h	acc_st
Snow fall	SF	surface	m of water equivalent	6h	acc_st
Snow depth	SD	surface	m of water equivalent	6h	inst
Total cloud cover	TCC	surface	0-1	6h	inst
Total column water	TCW	surface	kg m**-2	6h	inst
Surface latent heat flux	SLHF	surface	W m**-2 s	6h	acc_st
Surface sensible heat flux	SSHF	surface	W m**-2 s	6h	acc_st
Surface solar radiation	SSR	surface	W m**-2 s	6h	acc_st
Surface thermal radiation	STR	surface	W m**-2 s	6h	acc_st
Sunshine duration	SUND	surface	s	6h	acc_st
Convective available potential energy	CAPE	surface	J kg**-1	6h	inst
Orography (Geopotential at the surface)	GH	surface	m**2s**-2	t_0 (and poss. t_res-change)	inst
Land-sea mask	LSM	surface	0-1	t_0 (and poss. t_res-change)	inst

inst: instantaneous output
det_lo: determined over period from last output time to current output time
acc_st: accumulated over period from start of forecast to current output time (or alternatively accumulated from last output time to current output time; to be decided)

9.7.2 Upper air data

It is proposed to archive 5 parameters on 9 pressure levels, i.e. 45 fields.

Parameter	Abbrev	Level	Unit	Output frequ.	Comments
Temperature	T	L9	K	6h	inst
Geopotential	G	L9	m**2 s**-2	6h	inst
U-velocity	U	L9	m s**-1	6h	inst
V-velocity	V	L9	m s**-1	6h	inst
Specific Humidity	Q	L9	kg kg**-1	6h	inst

L9: at levels 1000, 925, 850, 700, 600, 500, 300, 250, 200 hPa

10 TIGGE's implementation plan and infrastructure requirements

This Section describes some details of the infrastructure requirements to set up and manage the TIGGE data sets. These are based on the data and access requirements summarised in the previous two sections.

Datasets needed for TIGGE projects are of three main types:

- Existing (historic) data
- Data to be collected routinely in near-real time from the participating centres
- Special datasets for specific sub-projects collected either in near-real time for selected periods or archived in delayed mode, possible by exchange on tape.

10.1 Historic data

This consists of data that is already available from various centres, for example the CDC reforecast data, ERA-40, NCEP reanalysis, and historical EPS forecasts and verification results. Observation datasets, including satellite data, are also available.

It is recommended that the TIGGE archive contain data catalogues and documentation for as many of these existing datasets as possible, but that the data themselves would not be archived as part of TIGGE. Users are expected to make requests to the source organization for access to these datasets.

To facilitate access to historical data, access methods, and formats (GRIB, NetCDF etc) should be documented. Also, existing or future common data access methods and tools should be applied wherever possible. These data would be free for registered research users.

10.2 Data exchange, storage and access for operational EPSs

This is the routine (often operational) ensemble data generated every day by a number of NWP centres. Two possible general structures were defined for the storage and access to TIGGE data, a "Phase-1" structure, which could be implemented now at little cost, and a "Phase-2" structure, which would be possible to implement within several years. These are briefly described in this section:

10.2.1 TIGGE Phase-1

Data would be exchanged in "near-real time", which means within a few hours to a few days, via FTP over the Internet. The data would be collected and archived at one or more central data repositories (ECMWF, NCAR, and CMA have so far expressed interest in retaining the entire dataset) Regional subsets of data may also be archived at centres which have expressed an interest in taking on the task of providing a data service for their region (e.g. SAWS)

Each contributing centre would be responsible for providing a backup archive of the data, which it contributes to the archive.

In addition, complementary datasets may be exchanged in delayed mode for use in THORPEX-TIGGE demonstration projects, for example boundary conditions over local areas for limited area models. It is not considered feasible to routinely archive sufficient data from global ensembles in the central archive to provide boundary conditions to LAMs, due to the enormous data volumes.

10.2.2 TIGGE Phase-2

Within several years, and following significant software development, it will become possible to set up distributed data repositories, in coordination with the WMO/WIS initiative. These must provide efficient and transparent data delivery to users, and may eventually reduce or eliminate to need to transfer large quantities of data on a daily basis to centralized archive centres. It is expected that Phase 1 will continue during the development and demonstration period for phase 2, for at least 2 years. Funding will be required to support software development.

10.3 Data volumes and archive structure

By the time the TIGGE archive starts, ECMWF (which runs the largest and highest-resolution ensemble of all centres) will have increased the resolution of its ensemble model to T399. Assuming 75 forecast fields, 40 timesteps per run for two runs per day, Gaussian grid, 50+1 ensemble members, and including full archive of the analysis on both model and pressure levels, this leads to an archive volume of 130 Gb per day. With respect to this, addition of ensembles from up to 9 other centres will increase the data volume by 50-80 Gb per day, giving a total daily volume of about 200 Gb. This is judged to be feasible within current network constraints. Sending out more than 2 or 3 copies of the ECMWF ensemble would be problematic, however. It is expected that input volumes from other centres will increase rapidly over the next few years as upgrades to their ensembles are implemented.

Data exchange should be standardized in terms of parameters (parameter definitions may vary between models), physical units, format and file naming conventions.

A requirement for data processing and archiving at higher temporal resolution (3-hourly or hourly) for the first 3 days was identified for some applications (e.g. driving hydrological models). Inclusion of high temporal resolution data for the whole forecast period could further increase the data volume, while implementation of a variable temporal resolution would complicate the archive and access structures.

The requirement for provision of boundary conditions for LAM integrations can be handled separately from the central archive by accessing full model resolution data while they are online (within a few hours of the model run). Archive service for these data could be made available at a later stage. Data exchange of full model data on model levels is not possible in near-real time; demonstration datasets could be distributed by tape.

A large subset of users may want data in the form of temporal series of a single parameter or a small number of parameters rather than as a set of full global fields for each day. This leads to a requirement for the rearchiving of a subset of the parameters as timeseries, which is feasible and may benefit from technical advances.

Post-processing of archive data will be required before delivery to users. Routines are needed for grid conversion, format conversion, and for the extraction of sub-areas, parameters and levels, for example. At ECMWF, access through the MARS system would be available to registered research users, and the more popular data sets can be made available through public servers.

Applications and tools for data processing tailored to the needs of specific users can be prepared and shared among users. In this way, the effort required for development of data processing tools can be shared among participants. These tools will need to be catalogued and documented. The option for storing user-generated data should be considered.

It was noted that the NOAA Operational Model Archive Distribution System (NOMADS) data access software project, developed at NOAA, may also have addressed some of these issues and collaboration with that project may be useful.

Applications are potentially CPU-intensive and they may require additional resources.

11 TIGGE's timetable and milestones

The following outline timetable and links are proposed. It is suggested that this schedule is discussed and approved by the THORPEX GIFS-TIGGE Working Group that needs to be established as soon as possible.

11.1 TIGGE's time table

In 2005, the GIFS-TIGGE WG should:

- Complete detailed planning (including costs and resource) and initiate Phase-1
- Secure commitment from institutes willing to act as central data archives for TIGGE's Phase-1 and from initial contributors to TIGGE database
- Agree data exchange protocols and data formats
- Finalise list of datasets and parameters to store in the TIGGE archive or provide alternative access to, based on provisional lists given in this report
- Help and encourage contributing centres to begin development to provide data in agreed formats
- Address data access policy with the aim to provide as far as possible real-time access for research use
- Consider the strategy required to initiate work on Phase-2 of TIGGE, including funding and resources
- Work within THORPEX and with external partners to develop detailed planning for the first real-time demonstrations using the TIGGE infrastructure (including IPY and Beijing Olympics)

It is recommended that a follow-on workshop be held as soon as possible for contributors and Phase-1 archive centres to achieve the relevant above targets.

In 2006:

- the Phase-1 TIGGE data archives will begin collecting available ensemble contributions in near-real time
- links will be available to other already established data archives
- research access will be available to TIGGE data (through the existing access infrastructure of the archive centres)
- work will begin on Phase 2 software developments (once required funding is secured)

In 2007 and 2008 the TIGGE infrastructure should be sufficiently established to provide real-time access to TIGGE forecast data (and possibly access through the TIGGE infrastructure to observation targeting information including sensitive area predictions) to contribute to proposed THORPEX support for:

- International Polar Year field campaigns
- Beijing 2008 Olympics WWRP Research and Development project

11.2 Link with NAEFS

Operational NAEFS products will be available in early 2006 and so will be available for real-time use during IPY and future campaigns. The above timetable is designed to complement this by providing access to additional ensemble data for use and evaluation during these THORPEX demonstrations.

Plans are underway for the NAEFS to include ensemble forecasts from the Met Office in 2006; the Fleet Numerical Meteorology and Oceanography Centre (FNMOC) of the US Navy also intends to participate. NAEFS could be considered as a forerunner to a potential operational application of TIGGE. As more NWP centres join, this could evolve into an important part of the future GIFS, shaped by new scientific results, and benefiting from a new level of coordination among major NWP centres that will support it.

To achieve this TIGGE and NAEFS developments need to be co-ordinated as soon as possible. Formal contact between TIGGE/THORPEX and NAEFS should be established as soon as possible to minimise costly parallel developments and to maximise potential mutual benefits.

12 TIGGE’s funding

TIGGE work will incur specific costs: telecommunications, data servers, additional disk space in the organizations serving as global or regional data repositories, and labour costs.

It is expected that a significant fraction of these costs will be supported by the partner institutions as “in kind contributions”. However a remaining fraction may not be covered in this way. In Phase-1, disk space and manpower will be needed at the global repositories. The development of the needed software to access data from the distributed TIGGE system in Phase 2 will probably require some specific work by specialized analysts, and could become the common property of the partner institutes. The associated costs can be estimated only after detailed plans are developed for Phase 2. At that point it may make sense to establish mechanisms to share costs associated with this development. The use of the WMO trust fund established by THORPEX ICSC or of similar instruments was noted as a possibility. The partner institutions may also develop specific proposals to funding agencies (e.g. the EU FP6 programme). After a more detailed evaluation of the costs, a decision will be needed as to the amount of required contributions in order to start the project on a safe basis.

13 References

Buizza, R., Houtekamer, P. L., Toth, Z., Pellerin, G., Wei, M., & Zhu, Y., 2005:A comparison of the ECMWF, MSC and NCEP Global Ensemble Prediction Systems.Mon. Wea. Rev., in press.

Bourke, W., Buizza, R., & Naughton, M., 2004: Performance of the ECMWF and the BoM Ensemble Systems in the Southern Hemisphere. Mon. Wea. Rev. 132, 2338-2357.

Hamill, T.M., J.S. Whitaker, and S.L. Mullen, 2005: Reforecasts, an important new data set for improving weather predictions. Bull. Amer. Meteor. Soc., submitted.

Hamill, T.M., J.S. Whitaker, and X. Wei, 2004: Ensemble re-forecasting: improving medium-range forecast skill using retrospective forecasts. MWR, 132, 1434-1447

Jordan C. Alpert and J. Wang, A Client Application for Real Time NOMADS at NCEP To Disseminate NOAA's Information Data Base, 17.5, 21st IIPS, AMS 2005. Available at: http://ams.confex.com/ams/pdfpapers/83062.pdf

Rogers, D. P. et al., 2005: THORPEX International Research Implementation Plan. Version 1. 14 February 2005. WMO/TD-No. 1258, WWRP/THORPEX No. 4. Available from http://www.wmo.int/thorpex/publications.html

Shapiro, M. A. and A. J. Thorpe, 2004: THORPEX International Science Plan. Version 3. 2 November 2004, WMO/TD-No. 1246, WWRP/THORPEX No. 2. Available from http://www.wmo.int/thorpex/publications.html

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 first THORPEX International Science Symposium, Montreal, Canada, December 6-10, 2004. [Available at: http://wwwt.emc.ncep.noaa.gov/gmb/ens/papers/toth_naefs_thorpex_montreal.pdf.]

ANNEX I

List of Participants

Andersson, Erik	ECMWF	e.andersson@ecmwf.int
Barnes, Andrew	IBM	andrew.barnes@uk.ibm.com
Berner, Judith	ECMWF	j.berner@ecmwf.int
Boettger, Horst	ECMWF	horst.boettger@ecmwf.int
Bohra, Ashwini	NCMRWF	akbohra@ncmrwf.gov.in
Bougeault, Philippe	ECMWF	philippe.bougeault@ecmwf.int
Broecker, Jochen	LSE	j.broecker@lse.ac.uk
Brown, Barbara	NCAR	bgb@ucar.edu
Buizza, Roberto	ECMWF	roberto.buizza@ecmwf.int
Burridge, David	WMO	d.m.burridge@btinternet.com
Casati, Barbara	MSC	barbara.casati@ec.gc.ca
Chen, Dehui	CMA	chendh@cma.gov.cn
Chessa, Piero	SAR Sardinia	chessa@sar.sardegna.it
Clarke, Liam	LSE	l.clarke@lse.ac.uk
Coelho, Caio	U Reading	c.a.d.s.coelho@reading.ac.uk
da Silva Dias, Pedro	U Sao Paolo	pldsdias@master.iag.usp.br
Dell'Acqua, Matteo	ECMWF	m.dellacqua@ecmwf.int
Ebert, Beth	BMRC	e.ebert@bom.gov.au
Ehrendorfer, Martin	U Innsbruck Austria	martin.ehrendorfer@uibk.ac.at
Enomoto, Takeshi	Earth simulator	eno@jamstec.go.jp
Estrada, Bartolome	INM	orfila@inm.es
Franz, Kristie	U California Irvine	franzk@uci.edu
Frogner, Inge-Lise	NMINorway	inger.lise.frogner@met.no
Hagedorn, Renate	ECMWF	renate.hagedorn@ecmwf.int
Hamill, Tom	CDC	tom.hamill@noaa.gov
Hofstadler, Fredi	ECMWF	alfred.hofstadler@ecmwf.int
Iversen, Trond	U Oslo	trond.iversen@geo.uio.no
Lalaurette, François	ECMWF	francois.lalaurette@ecmwf.int
Lawrence, Andrew	ECMWF	a.lawrence@ecmwf.int
Lee, Woo-Jin	KMA	wjlee@kma.go.kr
Leutbecher, Martin	ECMWF	m.leutbecher@ecmwf.int
Marsigli , Chiara	ARPA	cmarsigli@smr.arpa.emr.it
Montani, Andrea	ARPA	amontani@smr.arpa.emr.it
Morse, Andy	U Liverpool	a.p.morse@liverpool.ac.uk
Motz, Martin	IBM	martin_motz@uk.ibm.com
Mylne, Ken	Met Office	ken.mylne@metoffice.gov.uk
Ndarana, Thando	SAWS	thando@weathersa.co.za
Paccagnella, Tiziana	ARPA	t.paccagnella@smr.arpa.emr.it
Pailleux, Jean	Meteo-France	jean.pailleux@meteo.fr
Palmer, Tim	ECMWF	tim.palmer@ecmwf.int
Parsons, Dave	NCAR	parsons@ucar.edu
Persson, Anders	SMHI	anders.persson@smhi.se
Raoult, Baudouin	ECMWF	b.raoult@ecmwf.int
Richardson, David	Met Office	david.s.richardson@metoffice.gov.uk
Roulston, Mark	Penn State	roulston@met.psu.edu
Shapiro, Mel	NOAA	mshapiro@ucar.edu
Silva Dias, Maria	CPTEC	assuncao@cptec.inpe.br
Stephenson, David	U Reading	d.b.stephenson@reading.ac.uk
Swinbank, Richard	Met Office	richard.swinbank@metoffice.gov.uk
Takeuchi, Yoshiaki	JMA	ytakeuchi@met.kishou.go.jp
Talagrand, Olivier	LMD	talagrand@lmd.ens.fr

Tennant, Warren	SAWS	tennant@weathersa.co.za
Thielen, Jutta	JRC-ISPRA	jutta.thielen@jrc.it
Thorpe, Alan	U Reading	a.j.thorpe@reading.ac.uk
Tolasz, Radim	CHMI	tolasz@chmi.cz
Toth, Zoltan	NCEP	zoltan.toth@noaa.gov
Treinish, Lloyd	IBM	lloydt@us.ibm.com
Vitart, Frederik	ECMWF	f.vitart@ecmwf.int
Wilson, Laurie	MSC	lawrence.wilson@ec.gc.ca
Yoden, Shigeo	U Kyoto	yoden@kugi.kyoto-u.ac.jp
Zsoter, Ervin	Hungary Met Service	zsoter.e@met.hu

ANNEX II

Workshop Programme

Tuesday 1 March 2005

9.00- 9.15: Registration
9.15- 9.30: *D Marbouty*: Welcome and opening
9.30-10.15: *D Richardson*: Introduction to THORPEX / TIGGE
10.15-11.00 *R Buizza*: Overview of global operational ensembles
11.15-12.00 *M Ehrendorfer*: Initial perturbations in ensemble prediction
12.00-12.45 *T Palmer*: Model imperfections
14.15-15.00 *T Hamill*: Calibration/combination
15.00-15.45 *E Ebert*: Validation
16.00-16.45 *T Paccagnella*: Limited area ensemble systems
16:45-17:30 *K Mylne*: Applications: forecaster perspective, training

Wednesday 2 March 2005

9.15-9.45 *J Thielen*: Applications: weather driven natural hazards
9.45 - 10.15 *M Roulston*: Applications: economics, energy, risk management
10.45.11.15 *Z Toth*: NAEFS: scientific and technical issues
11.15-11.45 *F Hofstadler*: Infrastructure and tools
11.45-12.15 *R Hagedorn*: User experience and requirements
12.15-12.45 *L Treinish*: architecture for a User-centric Global Interactive Forecast System testbed

14.15-17.00 Working groups

Thursday 3 March 2005

9.00-10.30 Plenary: preliminary reports from WGs and new brief
11.00-13.00 Working groups: revision of recommendations
14.30-17.00 Plenary: revised reports from WGs, and final discussion

Friday 4 March 2005

9.00-17.00 Organising Committee, speakers, WG chairs only

ANNEX III

Available EPSs and LEPSs

This table summarizes some key characteristics of the global and limited-area ensemble prediction systems run operationally and in research mode.

Status	SYSTEM	Runs/day	Domain	Resolution	Lead Time(hr)	Members / run	Num models	Pert method
GLOBAL OPERATIONAL	BMRC	2	globe	TL119	240	33	1	SVs
	CMA	1	globe	T106	240	33	1	SVs+BVs
	CPTEC	2	globe	T126	360	15	1	EOF
	ECMWF	2	globe	TL255	240	51	1	SVs
	FNMOC	1	globe	T119	240	17	1	BVs
	JMA	1	globe	T106	216	25	1	BVs
	KMA	1	globe	T106	192	17	1	BVs
	MSC	1	globe	TL149	240	17	2	EnKF
	NCEP	4	globe	T126-T62	384	11	1	BVs
GLOBAL RESEARCH	NCMRWF	?	globe	T80	384	9	1	BVs
	SAWS	?	globe	T62	336	17	1	BVs
	UK Met Office	2	globe	90km	336	20	1	ETKF
LAM OPERATIONAL	NCEP-SREF	2	North America	32 km	63	15	2 (ETA, RSM)	NCEP global EPS
	SRNWP-PEPS	4	Europe	7-22 km	30	18	18 European Models	EU global models
	COSMO-LEPS	1	Europe	10 km	120	10	1(Lokal Modell ,COSMO)	Selected ECMWF global EPS members
	PEACE Meteo France	1	Globe	TL358, ST2.4	60	11	1 (ARPEGE)	Targeted SVs
	NOR LAMEPS	1	Europe	20 km	72	41	2 (HIRLAM-TEPs)	TEPS, based on targeted SVs
	UK Met Office	2	Euro-Atlantic	24 km	72	16	1 (UK Unified Model)	UKMO-EPS, using ETKF
	INM-SREPS	4	Euro-Atlantic	0.25 deg	72	64(lagged system)	4 (Hirlam, HRM, MM5, UM)	4 (ECMWF.UKMO AVN-NCEP, GME-DWD)
LAM RESEARCH	BMRC	1	65S-17N; 65E-185	0.5 deg	72	16	1 (LAPS)	BMRC global EPS, using DA cycles with pert-obs
	CMC-EPS	1	North America	28 km	48	20	1 (CMC reg mod)	CMC global EPS, using targeted SVs
	DMI-HIRLAM (1)	1	Euro-Atalntic	0.6 deg	72	51	DMI-HIRLAM	ECMWF-EPS
	DMI-HIRLAM (2)	1	Euro-Atlantic	0.2 deg	72	5	DMI-HIRLAM	ECMWF Control, using pert moist physics
	HMS LAMEPS	1	Europe	12 km	54	11	ALADIN	Meteo-France PEACE using targeted SVs
	SWISS LEPS		Europe	10 km	72	variable	1 (Lokal Modell COSMO)	Selected ECMWF global EPS members
	SAR_MME	1	Euro-Atlantic	0.25 deg	72	6	3 (MM5, BOLAM, RAMS)	ICs/BCs from AVN and IFS deterministic fcs

	PIED_SE	1	Europe	7 km	48	3	4 (LM-DWD, aLmo MeteoSwiss, LAMI, IFS)	L:AM and IFS
	UWME (Univerity of Washington)	2	Pacific Nortwest	Outer 36 km Nested 12 km		17	MM5	8 global models
	MAP D-phase		Europe	?	120	Under development	Under development	Under development

ANNEX IV

Summary of Expressions of Interest

An informal survey of the centres represented at the meeting was conducted. The WG discussions indicated various possible levels of involvement for any centre: Provider of ensemble forecasts, user of ensembles forecasts, global repository and distribution centre for TIGGE products, regional repository and distribution centre for TIGGE products. The Table below summarizes the first expressions of interest received during the meeting. It is does not constitute a commitment of these Centres, nor does it constitute a full survey of the potential participants in TIGGE. It is included here to indicate the already substantial level of interest in participation in the TIGGE project.

Centre	Provider	User	Global Repository	Regional Repository
BMRC	x	x		
CPTEC	x	x		in Phase 2
UKMO	x	x (verification)		
Meteo-France	x	x		
MS Canada	x	x		
KMA	x	x		in Phase 2 (subject to funding)
CMA	x	x	x (to be confirmed)	
ECMWF	x + MARS software	x	x	
SMHI		x		
NCAR		x	x	
ARPA	x	x(verification)		
NCMRWF	x (in future)	x(agriculture)		
Earth Simulator Center	x	x		
NCEP	x	x		
South Africa				x (to be confirmed)
JMA	x	x		with Univ. Tsukuba?
CDC	x (reforecasts)			
JRC		x		

In addition, a strong interest to use TIGGE data was indicated by Univ Innsbruck, LMD (France), and the HEPEX project.

ANNEX V

List of Acronyms

BMRC	Bureau of Meteorology Research Centre
CAS	WMO Commission for Atmospheric Sciences
CDC	Climate Diagnostic Center
CMA	China Meteorological Administration
CMC	Canadian Meteorological Centre
COSMO	COnsortium of Small-scale MOdelling
CPTEC	Centro de Previsao de Tempo e Estudos Climaticos
DIME	Direct and Inverse Modelling in Environmental end-to-end prediction
DMI	Danish Meteorological Institute
ECMWF	European Centre for Medium-Range Weather Forecasts
EFAS	European Flood Alert System
EPS	Ensemble Prediction System
ETKF	Ensemble Transform Kalman Filter
FNMOC	Fleet Numerical Meteorology and Oceanography Center
GEPS	Global EPS
GIFS	Global Interactive Forecast System
GTS	Global Telecommunication Service
HEPEX	Hydrological Ensemble Prediction Experiment
HMS	
ICSC	International Core Steering Committee
INM	
JMA	Japan Meteorological Agency
JRC	Joint Research Center
KMA	Korea Meteorological Administration
LAM	Limited Area Model
LEPS	LAM Ensemble Prediction System
MARS	Meteorological Archive System
MSC	Meteorological Service of Canada
MUMMA	MUlti-Model Multi-Analysis
NAEFS	North American Ensemble Forecasting System
NCAR	National Center for Atmospheric Research
NCEP	National Centers for Environmental Prediction
NCMRWF	National Centre for Medium-Range Weather Forecasting
NNSM	National Meteorological Service of Mexico
NOAA	National Oceanic and Atmospheric Administration
NOR	Norwegian?
NWS	US National Weather Services
PEACE	
PIED-SE	
SAR	SArdinian Regional meteolorigcal service
SAWS	South-African Weather Service
SRNWP	Short-Range Numerical Weather Prediction
THORPEX	The Observing system Research and Predictability EXperiment
TIGGE	THORPEX Interactive Grand Global Ensemble
UWME	University of Washington Mesoscale Ensemble??
WGNE	Working Group on Numerical Experimentation (joint WCRP, CAS)
WMO	World Meteorological Organization
WCRP	WMO World Climate Research Programme
WWRP	WMO World Weather Research Programme