

## ACTIVITIES OF THE CAS/JSC WORKING GROUP ON NUMERICAL EXPERIMENTATION (WGNE)

### 1. Role of WGNE in support of WCRP and CAS

The Working Group on Numerical Experimentation (WGNE), as a joint working group of the Joint Scientific Committee (JSC) of the WCRP and Commission for Atmospheric Sciences (CAS) of the WMO, has the basic responsibility of fostering the development of atmospheric models for use in weather prediction and climate studies on all space and time-scales. In the WCRP, WGNE is at the core of the global modelling effort and co-ordination between WGNE, the Working Group on Coupled Modelling (WGCM) and the Working Group on Seasonal to Inter annual Prediction (WGSIP) is maintained primarily through ex officio meeting attendances. WGNE also works in close conjunction with the WCRP Global Energy and Water Cycle Experiment (GEWEX) particularly in the development of atmospheric model parametrizations, with WGNE sessions held jointly with the GEWEX Modelling and Prediction Panel (GMPP) (but not in 2006). The WGNE Chair is a member of WCRP Modelling Panel (WMP), with WGNE represented on the WCRP Observations and Assimilation Panel (WOAP) also.

WGNE also has specific THORPEX sessions at its meetings. The close relationship that exists between WGNE and operational (NWP) centres underpins many of the activities of WGNE, and it is the work of these centres that provides much of the impetus for the development and refinement of the physics and dynamics of atmospheric models. As usual, WGNE sessions include reviews of progress at operational centres in all aspects of NWP including data assimilation, numerical methods, physical parametrizations, ensemble predictions, seasonal prediction, verification of precipitation and tropical cyclone track forecasts. WGNE also follows progress in various relevant national initiatives such as the Frontier Research Center for Global Change in Japan. The need for good metrics for climate-type models is under discussion. WGNE will discuss this further also in the context of the new 'unified' prediction systems.

The following paragraphs briefly review the main activities of WGNE in support of WCRP/CAS objectives, emphasizing items arising at its twenty second session which was kindly hosted by the National Centre for Atmospheric Research, Boulder, Colorado USA, 24 -27 October 2006. On this occasion the 24 October was a joint session with the WMP. Emphasis has been given to activities where international coordination is paramount and facilitated by the working group's existence, or where new scientific initiatives are involved.

### 2. Studies and comparisons of atmospheric model simulations

Model inter-comparison exercises are a key element in meeting a basic WGNE objective of identifying errors in atmospheric models, appreciating their causes and reducing or eliminating these errors.

#### PCMDI, CMIP and a Workshop on Model systematic errors

WGNE congratulated PCMDI for continuing to maintain and enhance a valuable infrastructure for processing model outputs at PCMDI and establishing efficient data formats etc. for such exchanges of model simulations. The recent outstanding achievements in the context of the IPCC/AR4 are of particular note. PCMDI has offered to receive high resolution NWP AMIP-type runs to complement their ongoing CMIP activities. PCMDI is the local host for a pan-WCRP/CAS workshop on Model systematic errors in February 2007. This is being organized by PCMDI and WGNE with input from WGCM and GMPP, and the programme is structured by time-scales to emphasize the 'seamlessness' of many model errors.

#### "Transpose" AMIP

The goal of the WGNE-Transpose AMIP is to obtain the benefits for climate model development and evaluation that have been invaluable for weather prediction model development, by applying climate models to weather forecasting. The method allows direct comparison of parametrized variables such

as clouds and precipitation with synoptic observations, satellite and field programmes. In general, development of a complete analysis system is not needed with initial conditions obtained from NWP (re-)analyses. The method allows direct comparison of parametrized variables such as clouds and precipitation with observations including field programmes (such as ARM), early in the forecast while the model state is still near that of the real atmosphere. This WGNE initiative was prototyped jointly by PCMDI and NCAR and known as CAPT. The intention is to encourage climate modelling groups to implement this forecast strategy into their development process. The formal proposal for Transpose AMIP has been sent to climate modelling groups.

### Aqua-Planet Experiments (APE)

WGNE recognizes the value of applying atmospheric models to simplified surface conditions for examining the behaviour of physical parametrizations and the interactions of parametrizations with the dynamical cores. In particular, "aqua-planet" experiments with a basic sea surface temperature distribution offer a useful vehicle in this regard. The details of the experiment and schedule are available at <http://www.met.reading.ac.uk/~mike/APE>.

The experiment is designed to provide a benchmark of current model behaviour and to stimulate research to understand differences arising from: (1) different subgrid-scale parametrization suites, (2) different dynamical cores, and (3) different methods of coupling model dynamics and parametrizations. Using the APE database, analysis of the APE experiments is continuing for another year. Following the workshop held in April 2005, a second workshop is planned to discuss the more complete analyses in late 2007 at the University of Tokyo.

The basic experiments are deliberately done at "climate model" resolutions but a few groups are examining convergence with resolution and the results and interpretation of resolution studies will be an important outcome of this work.

### **3. Regional Climate Modelling**

Following the WMO/WCRP sponsored RCM Workshop in Lund, Sweden in 2004, the Transferability Working Group (TWG) was created. The aims of this group are to assess the global applicability of RCMs in regions remote from their home domain of development. Particular emphasis is being placed on the simulation of regional scale water and energy cycles in a wide variety of climatic regimes and the Inter Continental Scale Experiment Transferability Study (ICTS) focussing on GEWEX Continental Scale Experiment sites is in progress. A second RCM workshop is planned for 2008, in Trieste (ICTP). WGNE also discussed results from SGMIP (Stretched Grid Model Inter-comparison Project). It will continue to monitor the developments in this area in its future sessions.

### **4. Climate Model Metrics**

WGNE has been involved in developing standard climate model diagnostics and metrics for some years. The goal of such metrics is to objectively measure model quality or skill and suitable metrics depend on the intended applications. The application for climate models includes the prediction of future climates for which no verification data will be available within the lifetime of the model. WGNE discussed the issue of climate model metrics at some length with many questions and issues resulting. A sub group with a member from each of PCMDI, WGCM, WGNE, GMPP and the JWGV (Joint Working Group on Verification) will define the climate model metrics and standard verification data sets with the intention of asking WCRP to encourage usage of these metrics for climate models. It was decided to ensure some emphasis on climate model metrics at the February 2007 model systematic errors workshop.

The need for good metrics for climate-type models is under discussion. WGNE will discuss this further also in the context of the new 'unified' prediction systems.

### **5. Physical parametrizations in models**

WGNE's close working relationship with GMPP (the GEWEX modelling and prediction panel),

provides the focus for the development, refinement and evaluation of atmospheric model parametrizations, notably those of cloud and radiation, land surface processes and soil moisture, and the atmospheric boundary layer. WGNE reiterated the value of the interaction with GMPP for parametrization work, particularly with GCSS. A joint WGNE/GCSS model intercomparison study of a Pacific cross section (GPCI) to evaluate physical parametrizations along the atmospheric cross section following the trade winds is in progress, with excellent support from both NWP and climate modeling groups. The need for an expert group on parametrization to advise both WCRP and WWRP (and their Working Groups) was discussed, and further consideration will be given to this in consultation with the GMPP.

## **6. Numerical weather prediction**

### Reanalysis projects and data assimilation

The ERA-40 reanalysis at ECMWF is complete and an "interim reanalysis" has begun. This is running from 1989 onwards. It contains improvements that greatly alleviate deficiencies identified in ERA-40. A comprehensive atlas of the atmospheric general circulation as depicted by ERA-40 has been produced in collaboration with the Meteorology Department of the University of Reading. The Japanese 25-year Reanalysis Project (JRA-25, 1979-2004) has been completed. WGNE reiterated its strong support for the reanalysis work, the desirability of maintaining a core of experts without excessive duplication of effort and ensuring efficient phasing of these efforts.

### Earth System assimilation

The new developments in the assimilation of parameters pertinent to the Earth System but not routinely analysed by current data assimilation systems are being monitored by WGNE. These include analyses of greenhouse gases, aerosols and reactive gases. Earth system science such as the GEMS (Global and regional Earth-system Monitoring using Satellite and in-situ data) project will increasingly demand cross-project liaison within WCRP and CAS.

### Model developments

WGNE noted the substantial improvements in the resolution of global and deep convection permitting forecast models in progress or planned in the next few years. There exists a dichotomy of opinion regarding the use and interpretation of grid-lengths of several kms for forecasting. These resolutions will become affordable for GCM use in the coming years, and the prospect of climate simulations with grids of order one kilometre is an issue of international activity and debate, and WGNE will continue to monitor such developments.

Recent results showing the need for model resolutions of 100 kms or better to properly define the statistics of extra-tropical storm tracks were noted. This contrasts with typical climate model resolutions substantially poorer than this, a matter of serious concern to the group.

WGNE noted that plans for unified (coupled) forecast systems that will provide forecasts from days out to seasons, typically by progressively degrading the resolution with forecast range, will provide new opportunities for ensemble techniques, including initial perturbations, stochastic parametrizations and metrics, and bring even closer collaboration between the NWP and climate communities.

WGNE will contribute to the WCRP Task Force on Seasonal Prediction (TFSP) meeting in Barcelona, June, 2007 including a report on the Systematic Errors workshop.

### Performance of the main global operational forecasting models

WGNE routinely reviews the skill of daily forecasts from a number of the main operational centres in terms of verification scores (such as anomaly correlation and root mean square error) for various fields at different lead times. For most centres, a distinct increase in skill continues.

### Model Verification

With global models attaining much higher resolutions, and mesoscale models being routinely run at most operational centres, consideration is being given to additional skill scores to the conventional ones that are more appropriate for such resolutions. Furthermore there is an increasing requirement to provide measures of model performance for predicting weather elements and severe weather events. The joint WGNE/WWRP working group on verification (JWGC) is now considering this important subject.

There are a number of WGNE projects involved with the validation of forecasts. New developments were discussed including the development of methods to verify high resolution spatial forecasts; verification methods for rare events; incorporation of scaling methods into verification processes; approaches to account for observational uncertainty in verification measures and analyses; development of methods that are customer dependent and appropriate for studies of forecast value; and verification of probability distribution functions.

#### Inter-comparison of typhoon track forecasts

The inter-comparison of forecasts of typhoon tracks has been an ongoing project that has been conducted by the Japan Meteorological Agency on behalf of WGNE for a number of years. This now includes all ocean basins, and data from operational forecasts is now available from eight Centres. The overall gradually improving performance of these models in predicting cyclone tracks over the past few years has been maintained. In future statistics will be gathered to assess the skill in intensity forecasts and forecasts of cyclone genesis. Many results related to typhoon track forecast including a multi-model ensemble are presented on the web site ([http://nwp-verif.kishou.go.jp/wgne\\_tc/index.html](http://nwp-verif.kishou.go.jp/wgne_tc/index.html) (user id and password are required)).

#### Verification and inter-comparison of precipitation forecasts

This WGNE initiative is being conducted at the DWD, NCEP, BMRC, CMA, JMA, CMC, the Met Office and Meteo-France. Quantitative global precipitation forecasts from the above are being verified against surface stations in these relatively data rich areas (some Centres also include their limited area model forecasts in the verification). A series of scores such as bias, Heike skill score, equitable threat score are used. It was noted that there is clear evidence from several Centres that the skill of precipitation forecasts in mid-latitudes was increasing.

#### Model-derived estimates of ocean-atmosphere fluxes (SURFA)

SURFA will evaluate and inter-compare global surface flux products (over ocean and land) from the operational products of a number of the main NWP centres and this will provide a good opportunity for estimating and determining the quality of model surface fluxes, of considerable relevance to atmospheric and coupled modelling communities and oceanographers. Following a joint session at WGNE-22 with the WCRP Working Group on Surface Fluxes (WGSF) it has been agreed to revitalize SURFA, and an agreed set of NWP fields etc. will be routinely archived at the National Climate Data Centre from a number of NWP Centres (after a preliminary pilot study currently in progress).

## **7. THORPEX**

At the 22nd WGNE meeting there was a session, which reviewed the status and plans of THORPEX and the wide-ranging opportunities for collaboration and synergy with WCRP and other bodies. The plans for T-PARC were of particular note, and this 'campaign' promises to make a major contribution to our understanding of meteorology in the Pacific basin.

The use of ensemble methods now forms a cornerstone of forecasting on all time-scales, and WGNE hoped that the rapidly progressing TIGGE project will help accelerate the effective use of ensemble forecasting information.

## **8. A Year of Tropical Convection**

WGNE discussed the proposal for 'A Year of Tropical Convection' (YOTC) which as currently envisaged, is aiming to assemble a dataset that will enable focussed research on many aspects of tropical convection, which in turn should lead to significant/important advances in our NWP abilities on all time-scales currently labelled under 'seamless' prediction. The discussions strongly supported the idea but felt that it was less clear how the aims of the YOTC would be achieved. Some concern was also expressed that the proposed time-scales were somewhat too tight.

As this YOTC dataset will be a judicious combination of many existing datasets in a variety of forms and repositories, questions were asked as to whether this is an opportunity to harness the powers of the new WMO Information System (WIS), and what was the YOTC relationship to other planned 'global' activities such as IPY and a possible Monsoon' focus. It was suggested that WWRP and WCRP should consider these questions and the efficacy of having a working group and/or a workshop in 2007.

Recognizing that convection is central to many problems in WCRP modelling research on almost all space and time scales, WGNE/GMPP were already jointly considering a high resolution modelling experiment specifically directed towards aiding and accelerating parametrization development. This could be part of a coordinated effort to benefit the entire WCRP community.

## **9. The Third International Workshop on Systematic Errors in Climate and NWP Models, San Francisco, February 12-16, 2007.**

### **9.1 Introduction and summary**

This was the Third JSC/CAS-sponsored WGNE workshop on model systematic errors, the previous ones being in Toronto in 1988 and Melbourne in 2000. On this occasion PCMDI provided substantial logistic and financial help for this well-attended meeting (~170 people). The workshop was structured to study model errors across multiple time-scales, from NWP to climate integrations. Errors in both atmospheric and coupled ocean-atmosphere models were high on the agenda. The workshop was structured with a limited number of presentations, a large number of posters with plenty of time to view and discuss them, a number of breakout groups to discuss various issues, and a plenary session to review and discuss the meeting as a whole and to identify and address the salient themes emerging from the workshop.

Systematic errors in climate and weather prediction models are evident on a wide range of space and time scales. The root causes of these errors are often difficult to address, because the many complex processes and phenomena of the climate system interact, both in the real world and in model simulations. A key motivation for this workshop was to bring together a variety of diagnostic approaches, with the expectation that awareness and understanding of the causes of systematic errors would be increased, and lead to a more coherent strategy for future advances.

### **9.2 Main Issues**

There were several main issues which emerged from the presentations and the discussions of the breakout groups:

**The importance of metrics.** In the NWP community, standard metrics that gauge the skill of forecasts have been routine for years. There is now increased interest in developing performance metrics for climate models. Establishment of a set of standard metrics could encourage all modelling groups to provide at least a minimal standardized summary of model strengths and weaknesses, which would facilitate monitoring and documenting of changes in model performance. A hierarchy of metrics could be designed to help assess the simulation of a variety of processes and phenomena on a range of time and space scales. Although work on optimizing the utility of metrics is in its early stages, it is widely believed that the metrics of most value will almost certainly be application dependent. Community-based efforts are underway to explore and establish a set of standard metrics relevant for climate models. The IPCC-AR4 archive typically includes results from more than thirty models and it is evident (both a priori and on looking at the model output) that not all models are created equal! The climate modelling community have traditionally been reluctant to “rank” model performance, but maybe the workshop has encouraged them to be more open about where they stand. Metrics may also be able to guide the interpretation of the model results - some models may be given more weight when making predictions of future climate change. This is a difficult and sometimes controversial area yet it is essential to perform weighting. The issue of appropriate metrics (typically based on simulation of past and present climate) is an area of ongoing work. Metrics that assess phenomena are important for intercomparison but weighting climate predictions really needs to be based on a more systematic assessment of model physics/dynamics.

**The importance of short range forecasts from NWP analyses.** Increasingly, our confidence in climate simulations (decades and longer) is dependent on how well they perform on much shorter time scales. What is wrong in a 100 year climate run is often going perceptibly wrong after 5 days of integration, because many errors are in the treatment of fast processes (boundary layer, convection, radiation, clouds). Short integrations from realistic initial conditions allow both detailed comparison with the latest observational data and diagnosis of the processes and tendencies in the model. This is

a much simpler and cheaper experimental framework than that of a fully coupled ocean-atmosphere general circulation model (OAGCM) run for decades to centuries. A number of climate modelling groups are beginning to use such techniques (as championed by WGNE for several years with the Transpose AMIP project)

Experience thusfar with several climate models in this mode has shown that the growth of systematic errors can be so prominent that residual problems of 'spin-up' due to incompatibilities between analysis and forecast model are not critical for many of these studies when using the highest quality NWP analyses. Undoubtedly there is a limit to this especially when looking at surface or near-surface issues and there are opportunities for research in this area. The ability to simulate the observed climate record over long periods still remains a crucial model test.

**The difficulties of accurate simulation of the diurnal cycle.** It clearly stretches the capability of current models to realistically capture the coupled and local physical processes that constitute the diurnal cycle. It is generally poorly simulated in GCMs, but models with very high horizontal resolution (e.g., less than 4km) do a little better, and cloud resolving models (CRMs) also do well with sufficient resolution. Poor simulation of the diurnal cycle impacts weather forecasts, but it is also important for climate via the Earth's radiation balance or the terrestrial carbon budget. It remains unclear what impact this deficiency might have on climate change projections. Climate models have yet to be run at convection-resolving resolutions, however, careful experimentation at high enough resolutions to capture cloud systems (gridlengths ~10kms) may benefit parametrization development in ways that could lead to better simulation of the diurnal cycle at typical climate resolutions. It was noted that the impact of explicitly resolving deep convection (gridlengths ~1km) in a climate simulation remains to be seen and is a clear challenge in the coming decade.

Results showing that the diurnal cycle had a strong impact on the momentum budget in the equatorial ocean suggest that the diurnal cycle of forcing might be important in ocean data assimilation systems. Much improved complete physics packages are needed to better handle these highly coupled situations involving a range of time and space scales.

**The value of running suitably initialized coupled models in forecast mode over seasonal time-scales.** This is analogous to atmosphere-only runs from NWP, but allows examination of somewhat slower processes, particularly those associated with ENSO and the seasonal cycle in the tropics. Relatively short coupled runs are also natural tools for comparing modelled and observed cloud/SST interactions. An analysis of ENSO in the AR4 models shows that ENSO amplitude has a big scatter - many models are overly strong, quite a few models are overly weak, very few models look anything like reality. The experience of seasonal forecasters is that simple initialization with wind and SST data was capable of giving very good ENSO forecasts, and that by selecting a relatively limited set of initial dates a model's ability to handle a range of EL Niño / La Nina / neutral conditions could be assessed. WGSIP will try to provide some "recommended" procedures and dates through the auspices of CLIVAR.

An outstanding challenge in modelling the MJO and monsoons was that active-break transitions are not forecast, and typically not represented in GCMs. The broader implications of this is that this limits medium-range and seasonal predictability, as well as ENSO forecasting, and that the simulation of extreme events is compromised. To date, many of the root causes behind errors in simulating the monsoon have not been identified.

**The need for much higher resolution.** The highest resolution simulations in AR4 are around T85, but the sentiment of the workshop suggested that the argument supporting much higher resolutions is now overwhelming, with several presentations demonstrating positive impacts of much higher resolution both from a dynamical and physical viewpoint. Recent experiments with high-resolution (60-90km) coupled models show that, in the tropics, the full potential of high resolution emerges if coupled models are used. Moreover, it seems crucial that high-resolution is used in both the ocean and the atmosphere.

### 9.3 Further workshop conclusions

There are a number of persistent model errors for which there is limited understanding of the underlying processes, and for which there are no clear solutions. Model errors affecting intermediate time-scales (e.g., monsoons and the Madden Julian Oscillation) are often subtle, and the processes

responsible for them need not be local. On longer time-scales, the El Niño Southern Oscillation (ENSO) is a dominant mode of climate variability, and there continue to be simulation errors in its structure, frequency and amplitude. Other coupled atmosphere-ocean modes of variability that require improvement include the Pacific Decadal Oscillation and the Atlantic Multi-decadal Oscillation. Increasing the use of climate models for seasonal time-scale experimentation is a practical recommendation from this workshop to tackle some modelling deficiencies associated with these modes of variability.

Although there is a growing appreciation of how the climate may change, century-long simulations are still very uncertain. How the global cloud fields respond to small changes in the Earth's energy budget is a key issue, with systematic errors over the sub-tropical oceans, for example, being responsible for substantial uncertainty.

The development of Earth System Models brings new challenges not least via their need for greatly increased resources, both computational and human. It was suggested that this posed a genuine threat to the necessary studies required for minimizing existing major systematic errors evident in less complex models, and without which, reliable/stable ESM simulations will pose a major challenge for some time.

The extent to which systematic errors limit the veracity of climate model projections is a key issue. Some systematic errors are clearly sensitive to horizontal resolution, while other errors seem not to be, and are presumed to be attributable to deficiencies in the parametrized formulations of unresolved processes. Nevertheless, recent experimentation suggests that current climate model resolutions are significantly too coarse to properly resolve important atmospheric and oceanic phenomena. The exploration of systematic errors should be conducted at much higher resolution than is typical for current global climate models and hopefully high enough to be operating in a numerically convergent regime for the realistic representation of the most important climate phenomena.

Progress will also be aided by emerging observational technologies for crucial physical processes in the climate system (e.g., clouds, aerosols, precipitation, surface energy exchanges), which will help to constrain the formulation of these processes in climate models. New types of data now becoming available such as from the CloudSat/Calipso satellite, will be a great resource for looking at model errors in simulating cloud and rain.

Increased computing resources will undoubtedly accelerate progress in reducing systematic errors in climate models. In this workshop, there were striking examples of how increased atmosphere and ocean horizontal resolution (substantially higher than typical for climate) can improve the simulation of some key climate processes. However, the meeting consensus was that, while continuing enhancements in computing resources were needed, having the right scientific manpower to work with interfacing the increasingly abundant observational data with the models was just as important. Since progress is at best incremental, there was concern expressed that model development was unattractive to young scientists and that it was difficult to attract and keep young talent in a publication-driven environment.

Martin Miller (Chair of WGNE)

