## 1. INTRODUCTION

The organisation of workshops is a part of the Centre's research activity. This publication contains the proceedings - sub-group recommendations and formal papers - of a workshop on **"Intercomparison of large-scale models used for extended range forecasts"** which was held at ECMWF, Shinfield Park, Reading, from 30 June to 2 July, 1982.

The increasing availability of large scale computer resources has encouraged many groups to design their own models and to carry out extensive experimentation, both extended range forecasts and climate simulation. This had lead, not unnaturally, to a wide diversity in model design and development and it was felt that it was an opportune time in the Centre's development to hold a workshop on the intercomparison of model performance in the extended range. There can be no doubt that many workshops, symposia and conferences will be held on this important subject during the next 5 to 10 years.

The purpose of this workshop was threefold:-

(i) To obtain an intercomparison of the performance of different models used for extended range forecast and general circulation simulation;

(ii) To assess, quantitatively, their systematic error;

(iii) To discuss the nature and the reasons for these systematic errors.

Since the quality of our forecasts, particularly in the medium range, is largely affected by systematic errors, the Centre and the meteorological community will benefit from this workshop. It was intended that this workshop should not only provide a forum for model intercomparison, but should also provide suggestions for model improvements, and to this end experts were invited from the following fields:

- General circulation modelling;
- Theoretical aspects of the general circulation;
- Observational studies.

The workshop was split into two sessions; in the first session invited scientists and scientists from the Centre presented papers on special topics (their contributions are included in these proceedings as Appendices). In the second session, the following subjects were discussed in working groups.

(i) Similarities and differences between model simulations and their systematic errors;

(ii) likely causes for the systematic errors;

(iii) suggestions for improvements of models.

The results of these discussions are summmarized in the following sections.

## 2. SIMILARITIES AND DIFFERENCES BETWEEN MODEL SIMULATIONS AND THEIR SYSTEMATIC ERRORS

### 2.1 Model performance

It is somewhat surprising that so little is known about how the various models behave and, consequently, how they compare. We will consider various properties in the order of increasing complexity, but which roughly corresponds to the inverse order of understanding and comparison. We examined only models which have been presented in this Workshop and our survey is therefore not complete. Other models such as those described in an early comparison study (GARP Report No.22) could not be considered during the short time available.

a) Overall thermal structure: Of the models examined, there is a systematic trend towards cool tropospheric temperatures. The GLAS model is a possible exception. The impact of these errors on the distribution of the static stability in the models is not well documented, but could be serious. There is a common model problem of cold polar stratospheric temperatures except for the NCAR model.

b) Zonally-averaged mean flow: The zonally averaged zonal wind usually shows tropospheric jets in both hemispheres which are somewhat too strong and extend too high and too far poleward. However, the NCAR model reproduces the winter jet quite accurately but is too weak in the summer hemisphere. The new GLAS model also has a weak summer jet. The only model to simulate the polar night jet properly is NCARs. The wintertime surface westerlies around 45° latitude are systematically too large in all models. The information on $[v]$, $[\omega]$, $[q]$, $[p_s]$ is not easily available in a form that can be compared. Observational uncertainties for some of these quantities are relatively large.

c) Geographical variation of the time-mean flow: A useful small set of variables to examine would be Z and T, $\psi$ and $\chi$ or $\underline{V}$ (e.g. at surface, 850, 500, 200 and 50 mb). At present most models seem to be reproducing mean tropospheric fields fairly well, aside from the temperature bias mentioned above. However, the size and position of the stationary eddies is not well documented in many models. The Icelandic low extends too far to the east changing the south-westerlies to westerlies. The Siberian/Alaska ridge appears to be too weak in the sea level pressure field and the westerly flow is too strong across the northern Rockies and over the eastern Atlantic and western Europe. The January simulation in the EC model is characterized by a Hadley-cell which is too weak and an Indonesian monsoon circulation which is excessively strong when compared with features at other longitudes. The July simulation is characterized by an excessively strong Walker circulation.

d) Geographic variation of the transient variability: Again a useful small set of variables would be $T'^2$ and $Z'^2$ at a few levels. To examine the high frequency amplitudes, bandpass filtering should be used. There is a tendency for models to be deficient in low frequency eddy activity. Low frequency eddy activity is important if the model is to be useful in simulating blocking for example. Models do a better job of reproducing the high frequency eddies and tend to underestimate the low frequency component of the variability.

e) Precipitation, cloudiness, boundary fluxes and diabatic forcing: Precipitation is the only one of these fields documented to any degree. Most models tend to generate too little precipitation over the tropical Atlantic. The GLAS model simulates the Atlantic ITCZ reasonably well.

f) Energetics: Not much is known about model characteristics here. The ECMWF model has baroclinic conversions which are too strong and momentum fluxes are overestimated.

## 2.2 Model evaluation (diagnostics and verification)

### a) How to verify?

GCM evaluation is a difficult question and no standard for doing it can be set at the present time. Generally speaking, one can see four different stages in the development of an evaluation strategy.

Stage 1: Verification of the zonally-averaged structures of the model simulation.

Stage 2: Verification of the full geographic distribution of time-mean fields and of some important (time) variance and covariance fields. For an example see Blackmon and Lau (1980).

Stage 3: Study of how different physical balance requirements are fulfilled in the model (as compared with the real atmosphere) not only for zonally or globally averaged cases but in the whole ($\lambda$, $\phi$, p) space. It is likely that with improved analysis techniques the "residual method" can provide estimates of the geographical distribution of the net diabatic forcing not only for time-mean conditions but also for the long-period transients. If the individual heating components are estimated simultaneously by using the parameterization scheme, significant new information can possibly be obtained on the role of diabatic processes in the maintenance of different modes of atmospheric circulation. Examples of new quantities, which in such a diagnosis could be considered as possibly useful measures of transient eddy forcing, are the extended E-P flux vector defined in the report of Hoskins and $X_q$ and $C_a$ as defined in the report of Holopainen.

Stage 4: Special regional verification could be done for characteristic climate regimes (e.g. desert climate). Special verification may also become appropriate in some continental regions in which good information exists, e.g. for the variables of the hydrological cycle.

These stages can be seen as consecutive phases such that e.g. Stage 3 implies also that 1-2 has been applied. At present, GCM evaluation is still largely in Stage 1 but is beginning to enter Stages 2 and 3.

b) Data for model verification

Stage 1: The statistics of Oort and Rasmusson (1971) and Newell et al (1972/1974) are basic material.

Stage 2: For the northern extratropics the atlas by Lau et al (1981) is already available. Oort (1982) will contain global patterns of time-means, variances, etc. for different pressure levels, as worked out from radiosonde data. The modelling group at Reading University is producing an annual digest of global seasonal and shorter period diagnostics based on ECMWF analyses, (see White, 1982). Newell et al (1972/74) is useful for verification in the tropical belt.

c) Uncertainties in the "observed" conditions

These are particularly large in the Southern Hemisphere but also in the data-sparse oceanic regions of the Northern Hemisphere (see Lau and Oort, 1981/1982). Some further measures of these uncertainties should be worked out,(e.g. through analysis intercomparison) and they should be considered when model simulations are compared with each other or with the reality. FGGE data are of special interest in this respect. Large uncertainties exist particularly in the diabatic forcing. A comprehensive Atlas would be very useful.

d) Lack of information

Cloudiness and soil moisture are examples of climate variables of great importance for which there is little information. The International Satellite Cloud Climatology Project (ISCCP) is probably going to provide important verification material for GCM cloudiness verification. Some indirect measures (e.g. daily temperature range) may have to be worked out for soil moisture verification.

2.3 Recommendations

Model Performance

a) We recommend that difference maps (model minus observation or model A minus model B) should be presented in future comparisons. Likewise, deviations from zonal means is a useful way to display the standing waves.

It was noted that in some respects different models exhibited different systematic errors and it is therefore recommended that the Centre should continue active participation in the comparison of the performance of its models with those of other groups. Attention should be paid to finding ways of estimating the significance of model differences and of illustrating them.

b) Since it is difficult, if not impossible, to specify a definitive set of model diagnostics or model comparison statistics, a variety of approaches is necessary. Statistics calculated from long runs are valuable. However, study of the systematic evolution of the forecast flow from ensembles of many forecasts, as have been shown in studies by Heckley and by Wallace, is a very useful approach which should be encouraged. The adjustments taking place particularly in the first few days of the forecast may reflect deficiencies in the models physics. Low- and high-frequency fluctuations should be examined separately. Geographical variations should be examined as well as integrated (either zonally or areally or vertically averaged) statistics.

In addition to variances other useful quantities, which could be used to look at the model's ability to simulate transient eddy activity, are the "observed" and "calculated" fields of extended E-P vectors by Hoskins and of the eddy forcing parameter $\psi_{\chi_q}$ and the baroclinic conversion parameter $C_a$ by

Holopainen. Also, attempts should be made to estimate the low-frequency diabatic forcing for "observed" and "calculated" conditions.

Uncertainties in the "observed" conditions

A comprehensive intercomparison of analyses made in different centres from essentially the same input data should be made in order to further improve our knowledge concerning the uncertainties in the "observed" conditions. Special importance must be given to the comparison of the different global (level III) FGGE data. A comparative study could also be made of the two level IIIb data sets produced by ECMWF and GFDL; this intercomparison would probably gain further significance if some operational (level IIIa) FGGE data sets were included in it.

## 3. LIKELY CAUSES FOR THE SYSTEMATIC ERRORS

### 3.1 Discussion

Clearly all of the processes represented in atmospheric models are candidates for being causes of the systematic errors that occur in these models and the groups' free-ranging discussions reflected this. The particular systematic errors of the ECMWF model as listed in Tiedtke's report were briefly considered in turn in the context of whether they were common to other models. This should provide clues to the likely sources of these errors, although a comprehensive picture did not emerge. Attention was then concentrated on individual components of atmospheric models. The points considered will be mentioned here only where specific recommendations arise as it was anticipated that present research at the Centre in all these areas would continue.

The basic model equations

The standard dynamical equations for a moist atmosphere involve approximations associated with, for example, the shape of the planet, the variation of gravity, the effects of moisture on specific heats and surface pressure, and differences between dynamic and static pressure. Some sensitivity to the moisture treatment has already been suggested by some preliminary experiments at the Centre.

Spatial discretization

It was considered that current model resolution and numerical schemes are capable of describing, with reasonable accuracy, large-scale adiabatic dynamical processes, with the possible exception of an adverse influence of the upper boundary on stratospheric dynamics. The latter point should be kept in mind and re-assessed at intervals, particularly when considering reduction in horizontal resolution at stratospheric levels. It was nevertheless felt that the representation of physical processes on the smallest resolved scales was worthy of further investigation, as discussed below.

Time scheme

Although the temporal discretization is not considered a prime source of systematic error, it is thought that the treatment of gravity waves by the semi-implicit scheme could possibly be detrimental to the tropical adjustment problem. Comparison with simulations using explicit schemes should occasionally be performed.

The interaction between the adiabatic model and physical parameterizations.

It was questioned whether parameterization involving input from, and feedback to, the smallest retained (and thus poorly represented) scales of the model is the best way to proceed, particularly bearing in mind the poorer climate simulations found in higher resolution models.

Mountains

The sensitivity of the simulated large-scale flow to the specifications of the orography has been demonstrated. Detailed synoptic analysis of the short-range errors in mountainous regions should be investigated for specific cases, in particular those available for study using the ALPEX data sets. Coupled with these case studies, the model behaviour in response to changes in horizontal resolution of the orography and/or model, and to changes in surface drag and vertical mixing in mountainous regions should be investigated.

Surface processes

Rowntree showed results from the British Met Office models indicating strong sensitivity of evaporation, precipitation, and surface wind patterns to the parameterization of evapotranspiration from land surfaces. This sensitivity implies that the values and/or functional forms of the various evapotranspiration parameters must be better determined. Randall reported some improvements in the precipitation and sea level pressure distributions simulated with the GLAS climate model resulting from a revised parameterization of evapotranspiration from land surfaces. He also described a tendency for the model's land surfaces to become too dry in extended simulations. This problem may be remedied by introducing a surface soil wetness as well as a deep soil wetness.

Radiation

Ramanathan reported experiments with the NCAR model indicating that if the middle and high latitude cirrus clouds are modelled unrealistically as black bodies, spurious polar cooling results. Randall showed improved results with the GLAS climate model when high level clouds are treated as transparent to both solar and terrestrial radiation. He also found some improvement in the simulated radiation budget at the model top when the cloudiness due to cumulus convection is neglected. The absence of a diurnal cycle in the ECMWF model is considered a serious defect and makes assessment of parameterizations difficult, particularly in the tropics. The performance of the parameterizations will need to be re-evaluated when the diurnal cycle is included in the model.

Boundary layer treatment

Sensitivity of features such as the depth and location of the Icelandic and Aleutian lows to the specification of turbulent vertical diffusion has been found and work in this area should continue. Randall reported that stratiform cloudiness is excessively widespread in the lowest layer of the GLAS model; a similar problem has occurred in some other GCMs as well. In order to overcome the problem some improvements are needed in the parameterization of those processes by which water can be transported upwards from the lowest model layer. The modelling of the interaction of the free atmosphere with the boundary layer is clearly a major difficulty. The value of high order closure schemes requires clarification as a potential candidate for solving this problem.

Convection

Tiedtke described experiments with the ECMWF model in which the Kuo and Arakawa-Schubert cumulus parameterizations were compared. The results showed significant differences, particularly over Indonesia and South America; however, the ECMWF parameterization of the effects of clouds on the radiation field made interpretation of the results difficult. The Arakawa-Schubert parameterization led to a drier troposphere giving much less cloudiness. The results suggest also that the parameterization of convection, though very important, cannot be treated in isolation. Green thought that the kinetic energy budget of convective processes should be clarified.

## 3.2 Recommendation

### (a) Some diagnostics and case studies

The interactive nature of the many processes represented in models is so crucial that studies of individual cases of particular synoptic systems or weather regimes provide a particularly useful tool for determining likely sources of error. Diagnoses of both atmospheric and model behaviour should be performed. The summer of 1976 in North-Western Europe is an example of an extreme regime worthy of further study which would provide information on the representation of some surface processes and of interactions between transients and time-mean flow. In this latter respect, diagnostics based on the anisotropy of the transients may prove useful.

Energetics calculations have suggested a significant, and perhaps spurious, generation of eddy available potential energy by diabatic processes in the ECMWF model. The source of this generation should be determined both geographically and synoptically by examining the correlation between diabatic heating and temperature. If the generation is found to be associated with particular synoptic systems, case studies of these will be required.

The behaviour of mature mid-latitude systems is crucial both to the weather of certain regions and to the momentum budget of the atmosphere in general. There is evidence that this behaviour is sensitive to the atmospheric structure on the equatorial side of the baroclinic region. Diabatic heating in the tropical regions, by influencing both the structure of the jet and the behaviour of extended mid-latitude troughs, is important in this context. Emphasis should be placed in this area through individual case studies.

(b) Approximations of overall equations

A re-assessment of the major approximations of the model equations should be made.

(c) Interaction between adiabatic and diabatic processes

The problem of interaction between adiabatic and diabatic processes should be re-assessed. At the first stage it is recommended that experiments be carried out in which both the input to, and output from, the parameterizations are spatially smoothed. This should first be tried in the horizontal but the important problem in the vertical should also be considered.

(d) Radiation

We recommend that the diurnal cycle is introduced into the ECMWF model as soon as possible and its performance evaluated. Also we strongly recommend the reassessment of the radiation scheme in the stratosphere in the light of results obtained at NCAR and the poor performance of the Centre's scheme at high levels.

(e) Cloud-radiation interaction

Since clouds are at the centre of the feedback problem with radiation, experiments should be conducted with zonally averaged and/or fixed clouds and by imposing on them different optical properties. The same should also be done with water vapour (using observed values) to compare the importance of both problems.

References

Blackmon, M.L. and N.-C. Lau, 1980: Regional characteristics of the Northern Hemisphere wintertime circulation: A comparison of thesimulation of aGFDL general circulation model with observations. J.Atmos.Sci., 37, 497-514.

Lau, N.-C. and A.H.Oort, 1981: A comparative study of the observed northern hemisphere circulation statistics based on GFDL and NMC analyses. Part I: the time-mean fields. Mon.Wea.Rev., 109, 1380-1403.

Lau, N.-C. and A.H.Oort, 1982: A comparative study of the observed northern hemispheric circulation statistics based on GFDL and NMC analyses. Part II: The transient eddy statistics. Submitted to Mon.Wea.Rev.

Lau, N.-C., G.H.White and R.L.Jenne, 1981: Circulation statistics for the extratropical Northern Hemisphere based on NMC analyses. NCAR Technical Note,Boulder, Colorado, pp.138.

Newell, R.E., J.W.Kidson, D.G.Vincent and G.J.Boer, 1972 and 1974: The General Circulation of the Tropical Atmosphere and Interactions with Extratropical Latitudes. Vols.1 and 2. The MIT Press, 258 and 371 pp.

Oort, A.H.and E.M.Rasmusson, 1971: Atmospheric Circulation Statistics. NOAA Prof.Paper 5, U.S.Department of Commerce, 323pp. NTIS order no. COM-72-50295.

Oort, A.H. 1982: Global Atmospheric Circulation Statistics, 1958-1973. NOOA Prof.Paper, in preparation.

White, G.H., 1982: The Global Circulation of the Atmosphere. Dec.1980-Nov.1981 based upon ECMWF analyses. University of Reading, pp.211.

WMO, 1979: Report of the FOC Study Conference on Climate Models: Performance, Intercomparison and Sensitivity Studies. Vol.I and II GARP Publications



Series No.22