The GCM–Reality Intercomparison Project for SPARC (GRIPS): scientific issues and initial results


It is advisable to refer to the publisher’s version if you intend to cite from the work. See Guidance on citing.
Published version at: http://dx.doi.org/10.1175/1520-0477(2000)081<0781:TGIPFS>2.3.CO;2
To link to this article DOI: http://dx.doi.org/10.1175/1520-0477(2000)081<0781:TGIPFS>2.3.CO;2
Publisher: American Meteorological Society

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur

CentAUR
Central Archive at the University of Reading
Reading’s research outputs online
The GCM-Reality Intercomparison Project for SPARC (GRIPS): Scientific Issues and Initial Results


ABSTRACT

To investigate the effects of the middle atmosphere on climate, the World Climate Research Programme is supporting the project “Stratospheric Processes and their Role in Climate” (SPARC). A central theme of SPARC, to examine model simulations of the coupled troposphere-middle atmosphere system, is being performed through the initiative called GRIPS (GCM-Reality Intercomparison Project for SPARC). In this paper, an overview of the objectives of GRIPS is given. Initial activities include an assessment of the performance of middle atmosphere climate models, and preliminary results from this evaluation are presented here. It is shown that although all 13 models evaluated represent most major features of the mean atmospheric state, there are deficiencies in the magnitude and location of the features, which cannot easily be traced to the formulation (resolution or the parameterizations included) of the models. Most models show a cold bias in all locations, apart from the tropical tropopause region where they can be either too warm or too cold. The strengths and locations of the major jets are often misrepresented in the models. Looking at three-dimensional fields reveals, for some models, more severe deficiencies in the magnitude and positioning of the dominant structures (such as the Aleutian high in the stratosphere), although undersampling might explain some of these differences from observations. All the models have shortcomings in their simulations of the present-day climate, which might limit the accuracy of predictions of the climate response to ozone change and other anomalous forcing.

*Universities Space Research Association, Seabrook, Maryland.
*Data Assimilation Office, NASA GSFC, Greenbelt, Maryland.
*Freie Universität Berlin, Berlin, Germany.
*Meteorological Research Institute, Tsukuba, Japan.
*NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey.
*University of Toronto, Toronto, Ontario, Canada.
*York University, Toronto, Ontario, Canada.
*NCAR, Boulder, Colorado.
*Department of Atmospheric Sciences, University of California, Los Angeles, Los Angeles, California.
*Science and Technology Corporation, Hampton, Virginia.
*Centre for Global Atmospheric Modelling, University of Reading, Reading, United Kingdom.
*Max-Planck-Institut für Meteorologie, Hamburg, Germany.
*NASA Goddard Institute for Space Studies, Palisades, New York.
*United Kingdom Meteorological Office, Bracknell, United Kingdom.
*CNRM, Météo-France, Toulouse, France.
*General Sciences Corporation, Beltsville, Maryland.
*NASA Langley Research Center, Hampton, Virginia.
*NIESDP, Tsukuba, Japan.
*PCMDI, LLNL, Livermore, California.
*European Centre for Medium-Range Weather Forecasts, Reading, United Kingdom.
*AES, Victoria, British Columbia, Canada.

Corresponding author address: Dr. Steven Pawson, Universities Space Research Association, Data Assimilation Office, NASA Goddard Space Flight Center, Code 910.3, Building 33, Greenbelt, MD 20771.

E-mail: pawson@dao.gsfc.nasa.gov

In final form 1 October 1999.

©2000 American Meteorological Society
1. Introduction

a. The need for GRIPS

There is an increasing recognition of the importance of the role played by the middle atmosphere in forcing other parts of the climate system. A major reason is the importance of stratospheric ozone as a climate gas (e.g., Ramanathan and Dickinson 1979) and the resulting implications of ozone depletion on the atmospheric circulation (e.g., Fels et al. 1980; Kiehl and Boville 1988; Mahlman et al. 1994; Santer et al. 1996; Ramaswamy et al. 1996; Forster and Shine 1997; Shindell et al. 1998). A further example of how anomalous radiative forcing in the middle atmosphere might impact the troposphere is the observational evidence of global-scale circulation anomalies following major volcanic eruptions (Robock and Mao 1992; Kodera 1994), which arise from the sulfate aerosols deposited in the tropical lower stratosphere. Even in the absence of anomalous radiative forcing, there is support for stratosphere–troposphere dynamical coupling mechanisms. There is a growing body of observational evidence that zonal-mean circulation anomalies can propagate from the upper stratosphere into the troposphere through the course of a winter (Kodera et al. 1990), which is consistent with the concept of “downward control” (Haynes et al. 1991). Other studies have isolated coupled structures in the wintertime stratosphere and troposphere (e.g., Baldwin et al. 1994; Perlwitz and Graf 1995; Thompson and Wallace 1998). There is also evidence that including the stratosphere in numerical weather prediction models can improve tropospheric forecasts (Mechoso et al. 1982). Such concepts were amongst the factors which motivated the World Climate Research Programme (WCRP) to support the project “Stratospheric Processes and Their Role in Climate” (SPARC), whose aims are summarized in the initial SPARC document (WCRP 1993) and Implementation Plan (WCRP 1998). Through a set of focused initiatives, often in conjunction with other international bodies, SPARC aims to investigate processes through which the middle atmosphere can exert an influence on the climate. In recognition of the central role of modeling studies in our understanding of climate, SPARC included an initiative to assess middle atmosphere climate models (MACMs): GRIPS (GCM-Reality Intercomparison Project for SPARC).

The large-scale thermal structure and trace gas distributions in the middle atmosphere are determined by the interaction of chemical, radiative, and dynamical processes. In the winter hemisphere, the dynamical forcing by waves that propagate out of the troposphere plays a first-order role in determining the residual circulation, which impacts trace gas transport and the thermal structure in the middle atmosphere. Even models with specified trace gas distributions have shortcomings in their representation of the present-day climate: two typical deficiencies are the cold-pole problem in the winter hemisphere (e.g., Fels et al. 1980) and the inability of most models to generate a quasi-biennial oscillation in the tropical lower stratosphere. [See Takahashi (1996) and Horinouchi and Yoden (1998) for exceptions to this statement, albeit with nonstandard climate models. Hamilton et al. (1999) showed the importance of including high-vertical resolution to simulate downward-propagating shear zones for QBO-like oscillations.] GRIPS will document how widespread and how severe such problems are in current MACMs. Some initial results of the intercomparison are presented in the current paper, which compares simulations by 13 models. Future studies will examine forcing mechanisms in more detail and will attempt to give more quantitative explanations of the simulated climate than is possible in the present study. Other issues for future comparison include the uniformity of model responses to predicted changes in greenhouse gases and chlorine/bromine loading of the atmosphere.

b. The structure of GRIPS

The primary purpose of GRIPS is to make a comprehensive evaluation of the performance of MACMs. This encompasses a whole range of issues, necessitating a multiphase structure to be adopted. In phase 1 an evaluation of MACMs is being performed; using the terminology of Gates (1992), this is a level-1 intercomparison, since the model integrations compared were not necessarily made with identical boundary conditions or specifications of radiatively active trace gases. Later phases, which will be discussed at the end of this paper, will include assessments of specific components of the MACMs and examinations of the response to several models to imposed forcing mechanisms. Additionally, collaborations between GRIPS and the Atmospheric Model Intercomparison Project (AMIP, Gates 1992; Gates et al. 1999), in which a special diagnostic subproject to examine the stratosphere has been initiated, will enable more rigorous evaluations to be made.

A list of tasks has been compiled for the phase-1 evaluations (Table 1). These address the basic perfor-
mance of the MACMs, including aspects of their climatology and variability. These tasks are tractable with the limited resources currently available for GRIPS. In order to perform the analysis, a set of diagnostic quantities on standard pressure levels was requested (see WCRP 1998); these include standard meteorological variables and the dynamical and physical forcing terms. In this paper, some results from task 1b are presented; these illustrate the quality of the present-day climate simulated by the MACMs. Results from the other tasks in Table 1, as well as more detailed examinations of the forcing mechanisms in the models, will be presented elsewhere as they are completed. The 13 participating MACMs are briefly described in section 2; the results of the comparison are in section 3; the paper ends with the conclusions and discussion (section 4).

2. The participating models

Participation in GRIPS is restricted to general circulation models with complete interactive representations of the hydrological cycle and radiative transfer, which can be used to simulate the atmospheric climate and its change. A further criterion must also be met: the upper boundary must be at the stratopause or higher. Thirteen groups have so far chosen to participate (Table 2). Most of these have an upper boundary well above the stratopause. More details of these models are given in the literature, the primary reference (generally describing a model version identical with, or close to, that used here) being given in Table 2.

As stressed in the introduction, this is a level-1 intercomparison. No constraints were placed on either the incoming solar radiation or on the sea surface temperature and sea-ice fields used as lower boundary conditions. All integrations submitted to GRIPS used climatological-average monthly values, repeating from year to year, so any interannual variability in the simulations is caused by internal dynamical processes. There were also no restrictions on the trace gas distributions, so that different ozone climatologies or CO₂ concentrations were used. Note that in AMIP, although concentrations of the well-mixed gases are mandatory, the ozone distributions are not. Future GRIPS exercises will examine the impacts of the ozone distribution. At this stage, we examine fields that may depend in some of their details on these “free” parameters, but which should not be fundamentally dependent on them. Note that 2 of the 13 models predict their ozone distributions internally: the ARPEGE-climate model uses simplified schemes carrying ozone as the only predicted middle atmospheric trace gas (Cariolle et al. 1990), while the LaRC-IMPACT model carries 24 chemical families (updated from Eckman et al. 1995).

Water vapor, the most important greenhouse gas in the troposphere, plays a lesser direct role in the radiative balance of the middle atmosphere and is treated differently in the different models; while some groups specify the middle atmospheric H₂O seen by the radiation scheme, other groups use the predicted fields. Of the MACMs in GRIPS, eight are based on spectral codes (Bourke 1972) while the remainder use one of Arakawa’s (1966) grid configurations (Table 3). Generally, the horizontal resolution used in the simulations is at the low end of that currently used in climate simulations (e.g., Gates 1992). This is at least partly a consequence of the additional computational burden of increasing the number of levels to accommodate the middle atmosphere.

Several groups have experimented with different horizontal resolutions in their models; there is evidence that the small-scale momentum and energy transport by gravity waves simulated at high horizontal resolution can result in a better simulation of the extratropical climate (e.g., Boville 1991, 1995; Hamilton et al. 1999). Clearly, these small-scale waves play an important role in forcing the middle atmospheric circulation, but it is not yet possible to perform long simulations at a resolution adequate to resolve the

<table>
<thead>
<tr>
<th>No.</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a</td>
<td>Documentation of the models, AMIP standards</td>
</tr>
<tr>
<td>1b</td>
<td>Basic model climatology</td>
</tr>
<tr>
<td>1c</td>
<td>Statistical connection between the troposphere and stratosphere</td>
</tr>
<tr>
<td>1d</td>
<td>Sudden warmings (synoptics, etc.)</td>
</tr>
<tr>
<td>1e</td>
<td>Traveling waves and tides</td>
</tr>
<tr>
<td>1f</td>
<td>Tropical oscillations and waves</td>
</tr>
<tr>
<td>1g</td>
<td>Troposphere–stratosphere exchange</td>
</tr>
<tr>
<td>1h</td>
<td>Spatial wavenumber spectra, including dissipation</td>
</tr>
<tr>
<td>1i</td>
<td>Southern polar vortex and its variability</td>
</tr>
</tbody>
</table>
entire gravity wave spectrum. Including parameterizations of orographically generated gravity waves leads to a reduction of systematic forecast model errors in the lower stratosphere (Palmer et al. 1986). Similarly, aspects of the modeled tropospheric climate can be improved (Boer et al. 1984; McFarlane 1987). Many of the GRIPS MACMs include some representation of orographic gravity waves (Table 3). Gravity waves generated by other phenomena also act as a momentum source for the middle atmosphere. These waves, which can have nonzero phase speeds, generally become more important for the local forcing at higher altitudes, but impact the mean meridional circulation and thereby the temperature at lower levels. They are parametrized in three of the GRIPS models discussed here: the GISS model (Rind et al. 1988a,b) uses a set of discrete waves with fixed phase speeds, excited by convection and shear instabilities, based on the parameterization of Lindzen (1981); the ARPEGE-climate model includes a similar representation of gravity waves forced by convection (Bossuet et al. 1998), but which are quasi-stationary in the frame of reference.

---

**Table 2.** List of the MACMs participating in GRIPS, along with the length of the integration, their home institution and nationality, and the primary reference for the model.

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Length (yr)</th>
<th>Group and location</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>UCLA</td>
<td>5</td>
<td>University of California, Los Angeles, CA</td>
<td>Kim et al. (1998)</td>
</tr>
<tr>
<td>CMAM</td>
<td>10</td>
<td>Canadian Middle Atmosphere Model (AES, University of Toronto and York University), Canada</td>
<td>Beagley et al. (1997)</td>
</tr>
<tr>
<td>GISS</td>
<td>10</td>
<td>NASA GISS, New York, NY</td>
<td>Rind et al. (1988)</td>
</tr>
<tr>
<td>MRI/lrf</td>
<td>10</td>
<td>MRI, Tsukuba, Japan (Long-Range Forecasting)</td>
<td>Chiba et al. (1996)</td>
</tr>
<tr>
<td>MRI/clim</td>
<td>10</td>
<td>MRI, Tsukuba, Japan Climate Group</td>
<td>Kitoh et al. (1995)</td>
</tr>
<tr>
<td>FUB</td>
<td>10</td>
<td>Freie Universität, Berlin, Germany</td>
<td>Langematz and Pawson (1997)</td>
</tr>
<tr>
<td>ARPEGE-climate</td>
<td>20</td>
<td>CNRM, Météo-France, Toulouse, France</td>
<td>Déqué et al. (1994)</td>
</tr>
<tr>
<td>MA/ECHAM4</td>
<td>10</td>
<td>MPI für Meteorologie, Hamburg, Germany</td>
<td>Manzini et al. (1997)</td>
</tr>
<tr>
<td>GSFC GEOS-2 (version 6.0)</td>
<td>5</td>
<td>NASA GSFC, Greenbelt, MD</td>
<td>DAO (1996)</td>
</tr>
<tr>
<td>UK-UM</td>
<td>5</td>
<td>United Kingdom Unified Model (UKMO, Bracknell and University of Reading, Reading), United Kingdom</td>
<td>Swinbank et al. (1998)</td>
</tr>
<tr>
<td>LaRC IMPACT</td>
<td>4</td>
<td>NASA LaRC, Hampton, VA</td>
<td>Fairlie et al. (1997)</td>
</tr>
</tbody>
</table>
moving with the convection; the MA/ECHAM4 model uses the more complex spectral scheme of Hines (1997), as described in Manzini et al. (1997). Some other models use an additional linearized drag (Rayleigh friction) in the mesosphere (see Table 3).

The vertical resolution of the MACMs varies considerably, with as few as 15 and as many as 70 levels in the atmosphere (Table 4). This presumably impacts the models’ ability to represent shallow structures, especially tropical waves. In the lower stratosphere, there is a factor of 3 range in the vertical resolution, which impacts the representation of the tropopause (and likely the water vapor transport). The choice of the vertical coordinate also differs; most models follow Simmons and Strüfing (1983) in adopting a hybrid vertical coordinate, with a smooth transition from terrain-following surfaces in the lower troposphere to pressure levels in the stratosphere; two models (GSFC GEOS-2 and LaRC-IMPACT) retain sigma coordinates at all altitudes.

### Table 3. Summary of the horizontal structure and resolution of the MACMs. The type of grid, the spatial resolution, and the use of a gravity wave drag formulation are documented. MRF refers to the use of mesospheric Rayleigh friction. Note that the MA/CCM2 has trapezoidal truncation, with a maximum 2D wavenumber of 42, but zonal truncation at maximum of wavenumber 21; other spectral models have triangular truncation (TN), with a maximum 2D wavenumber of N (equal to 21, 30, or 31 in these models).

<table>
<thead>
<tr>
<th>Model</th>
<th>Formulation</th>
<th>Horizontal resolution</th>
<th>Gravity wave drag</th>
</tr>
</thead>
<tbody>
<tr>
<td>MA/CCM2</td>
<td>Spectral</td>
<td>trap42 × 21</td>
<td>Orographic</td>
</tr>
<tr>
<td>UCLA</td>
<td>C-grid</td>
<td>5.0 × 4.0</td>
<td>Orographic</td>
</tr>
<tr>
<td>CMAM</td>
<td>Spectral</td>
<td>T32</td>
<td>Orographic</td>
</tr>
<tr>
<td>SKYHI</td>
<td>A-grid</td>
<td>3.6 × 3.0</td>
<td>None</td>
</tr>
<tr>
<td>GISS</td>
<td>B-grid</td>
<td>5 × 4</td>
<td>Orographic, convective, shear instabilities</td>
</tr>
<tr>
<td>MRI/lrf</td>
<td>Spectral</td>
<td>T21</td>
<td>MRF</td>
</tr>
<tr>
<td>MRI/clim</td>
<td>C-grid</td>
<td>5.0 × 4.0</td>
<td>Orographic</td>
</tr>
<tr>
<td>FUB</td>
<td>Spectral</td>
<td>T21</td>
<td>MRF</td>
</tr>
<tr>
<td>ARPEGE-climate</td>
<td>Spectral</td>
<td>T21</td>
<td>Orographic, convective</td>
</tr>
<tr>
<td>MA/ECHAM4</td>
<td>Spectral</td>
<td>T30</td>
<td>Spectral</td>
</tr>
<tr>
<td>GSFC GEOS-2</td>
<td>C-grid</td>
<td>2.5 × 2.0</td>
<td>Orographic</td>
</tr>
<tr>
<td>UKMO</td>
<td>B-grid</td>
<td>3.75 × 2.5</td>
<td>Orographic (p &gt; 20 hPa); MRF</td>
</tr>
<tr>
<td>LaRC</td>
<td>Spectral</td>
<td>T32</td>
<td>Orographic</td>
</tr>
</tbody>
</table>

### 3. Climatological structure

#### a. Reference datasets

A brief summary of some observational datasets is given before the model evaluation is described. An ideal climatological dataset would be accurate enough for all model evaluations, but there are several estimates of the “observational truth” and the differences between some of these are sufficient to impact the interpretation of the models.

European Centre for Medium-Range Weather Forecasts (ECMWF) reanalyses (ERA-15; Gibson et al. 1997) provide a good climatological dataset for the troposphere. They are used here to validate the thermal structure at pressures higher than 30 hPa. Pawson and Fiorino (1998) showed that they are in good agreement with radiosonde-based analyses near the tropical tropopause. Tropospheric geopotential heights from ERA-15 are also used; this is a robust field, with little difference between the observational datasets. Stratospheric geopotential heights (pressures between 50 and 1 hPa) are from an analysis of the TIROS Operational Vertical Sounder (TOVS) data described by Bailey et al. (1993). They use tropospheric heights from operational analyses and geopotential thicknesses derived from TOVS radiance data at the United Kingdom Meteorological Office (UKMO) to build up the fields.

Randel (1992) presented a climatology derived from the (former) National Meteorological Center
(NMC) analyses of geopotential heights. Winds are derived using a nonlinear balance equation. The zonal-mean winds are used to evaluate the MACMs between the surface and the stratopause. In the extratropics, they agree well with the winds from other datasets over a comparable period (1979–96). Temperatures in this dataset are accurate enough to serve as a reference in the stratosphere, but near the tropical tropopause they are several degrees higher than the ERA-15 values.

In collaboration with another SPARC initiative (Randel et al. 1997; WCRP 1998), GRIPS is examining the various climatological datasets available, with the objective of placing reasonable bounds on the observational truth; the results of that exercise will be valuable for this and other model evaluations.

b. Thermal structure

The model evaluation begins with the multiannual, global-mean temperature (Fig. 1). In common with the AMIP models (e.g., Boer et al. 1992) there is a cold bias at most levels in most MACMs. The exceptions (with a warm bias) are in the lower-middle stratosphere (100–30 hPa) in the MA/CCM2 and UCLA MACMs, and at pressures less than 10 hPa in the CMAM. Note also that in most of the troposphere of the SKYHI and LaRC-IMPACT models are colder (and further from the observations) than the others. In contrast, in the stratosphere the model biases are more uniform. This persistent cold bias in most of the models is indicative of a systematic underestimate of the radiative heating rates, such as too little absorption of solar radiation (by ozone or other gases in the near-infrared part of the spectrum) or too much longwave emission in the middle atmosphere.

Near the tropical tropopause, ERA-15 temperatures are lower than those in the Randel (1992) and the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalyses (Kalnay et al. 1996). ERA-15 temperatures agree best with radiosondes (Pawson and Fiorino 1998). The magnitude of the global-mean cold bias in the MACMs (Fig. 1) depends on the choice of reference dataset: near 100 hPa it is smaller when compared to the ERA-15 data. This emphasizes the necessity of using an accurate climatology. An accurate simulation of the temperatures near the tropical tropopause is essential for climate studies because of the impact on the stratospheric water vapor budget.

The latitudinal temperature structure at 100 hPa in January is also shown in Fig. 1, with the ERA-15 data as a reference. It is representative of all months. This reveals more about the nature of the model biases: the extratropical lower stratosphere in both hemispheres is too cold in almost all MACMs. The most pronounced bias is exhibited by the MRI/clim model, which is more than 20 K too cold at the North Pole and about 30 K too cold at the South Pole, with a slight warm bias in the Tropics. This is clearly an outlier,
most MACMs being within 5 K of the climatological estimate in the winter polar region, but exhibiting much more scatter in summer. In the Tropics, six MACMs (UCLA, MRI/clim, FUB, CMAM, MRI/lrf, and MA/CCM2, ranked according to their equatorial temperature) are warmer than the ERA-15 estimates.

A different aspect of the 100-hPa temperatures is illustrated by Fig. 2, which shows time series of modeled and observed monthly mean tropical and extratropical values for each year (for UCLA the multiyear average is shown). There is little interannual variability in these area means. The tropical values \( \langle T_{\text{trop}} \rangle \) are averaged over the 30°S–30°N latitude band, and the extratropical values \( \langle T_{\text{extr}} \rangle \) are for latitudes outside this band. Figure 2 illustrates the biases and the seasonal progressions of \( \langle T_{\text{trop}} \rangle \) and \( \langle T_{\text{extr}} \rangle \) in the MACMs. Yulaeva et al. (1994) showed how the annual cycle in \( \langle T_{\text{trop}} \rangle \) and \( \langle T_{\text{extr}} \rangle \) are out of phase, a feature clearly seen in the ERA-15 values. Most MACMs reproduce this feature of the observations, despite their systematic temperature biases.

There are different types of systematic errors. First, MACMs that are too cold at all latitudes but show a realistic annual cycle (SKYHI, MRI/lrf, ARPEGE-climate, MA/ECHAM4, GSFC GEOS-2, UK-UM, and LaRC IMPACT); the cold bias suggests that the primary problem is with the radiation. Second, models that are too warm in the Tropics but too cold at high latitudes, despite having a realistic annual cycle (MA/CCM2, CMAM, FUB). This structure is consistent with an approximately correct radiative forcing and a weak residual circulation linking the tropical and extratropical temperatures, with too little downwelling/upwelling at high/low latitudes. Clearly, the degree of asymmetry in the cold bias in the first group of MACMs is impacted by the same process. In some cases in each of those groups the annual cycle in the tropics (MA/CCM2 and GISS) or at high latitudes (UK-UM and LaRC IMPACT) is suppressed, suggesting a misrepresentation of the tropical–extratropical links in these models. Rind et al. (1999) found that such deficiencies in the GISS model were related to wave driving in the winter stratosphere and were insensitive to tropical convection. The third group contains the UCLA and MRI/clim MACMs, which evolved from a common model some years ago and now have very different global-mean biases. Both show an overly strong semiannual cycle in tropical temperatures, although the amplitude is suppressed in the MRI/clim model; the extratropical seasonal cycle is also unrealistic in these models.

---

**Fig. 1.** The top panel shows the vertical structure of the long-term, annual global-mean temperature (K) from observations (thick black line) and the 13 models (thin colored lines). Observations are a 17-yr mean, updated from Randel (1992). The lower panel shows the latitudinal structure of the multiyear monthly mean temperature (K) at 100 hPa in Jan. The observational estimate (thick black line) is specified from ERA-15 data (Gibson et al 1997). The models are represented by the same style and color as in the upper panel.
degrees of success by the models: their magnitude, location, and shape can differ from the observed climatology. In the tropical troposphere, observations show weak zonal-mean easterlies with very weak westerlies in the upper troposphere; in some models (MA/CCM2, CMAM, SKYHI, MRI/lrf, MRI/clim, FUB, GSFC GEOS-2, and UK-UM) these weak westerlies are replaced by weak easterlies.

The observed stratospheric winds clearly show the summertime easterlies and the wintertime polar night jet, which slopes equatorward with increasing altitude in the stratosphere. The models capture the stratospheric winds with varying degrees of success (Fig. 3). Typical problems occurring in more than one model are (i) a failure to close the summertime tropospheric westerlies in the lower stratosphere, (ii) a vertically aligned polar night jet in the winter stratosphere, and (iii) differing degrees of separation between the winter subtropical jet and the polar night jet.

All except two of the MACMs use a topographic gravity wave drag parameterization (Table 3), which is intended to provide adequate lower stratospheric forcing to close the subtropical jet (and warm the lower polar stratosphere). Despite this, Fig. 3 shows that many of the MACMs reproduce the closure of the subtropical jet quite poorly. There is clearly a need for more detailed studies of the lower stratospheric momentum budget.

In the southern lower stratosphere the persistence of the westerlies into January could also be due to gravity wave forcing. Alexander and Rosenlof (1996) showed that gravity waves with nonzero phase speeds can dissipate in this region and lead to a westward forcing of the mean flow. The GISS, MA/ECHAM4, and ARPEGE-climate models include parameteriza-
tions of nonzero phase speed waves and all reproduce reasonably accurate vertical wind gradients. Again, however, several models without a representation of these waves (UK-UM, GFDL SKYHI, and MRI/lrf) also represent this region quite well. Other effects could be responsible, including (i) the behavior of the upper levels, in the sense that weaker winds aloft (associated with stronger eddy forcing) in middle winter lead to a more rapid downward propagation of the seasonal cycle (Kodera et al. 1990) and an earlier transition to summertime easterlies; (ii) the amount of dynamical forcing in the lower stratosphere by medium-scale waves propagating from the troposphere; (iii) the impacts of the parametrization of subgrid-scale diffusion; and (iv) the effectiveness of the radiative heating rates in the models in warming this region as the seasonal cycle progresses (which depends on both the radiation scheme and the trace gas distributions used). All of these mechanisms will be investigated more fully in GRIPS.

In the higher stratosphere both the strength of the winds and the location differ considerably among the models. Generally, the polar night jet remains too far poleward near the stratopause, a problem that has been linked to the absence of nonorographic gravity waves (e.g., Rind et al. 1988a). Note, however, that the current ARPEGE-climate, GISS and MA/ECHAM4 models (which parametrize these waves) fit well within the range of simulated values, being no more realistic than some of the other MACMs.

d. Geopotential height fields

An overly strong zonal-mean polar night jet can arise from a vortex that is either too strong or too zon-
ally symmetric (or both). The multiyear mean 10-hPa geopotential heights (Fig. 4) illustrate several features: (i) the systematic differences in the mean values (related to thermal biases at lower levels), (ii) the different vortex strengths, and (iii) the various shapes and orientations of the modeled vortices. The observational basis (TOVS data) reveals a polar vortex displaced toward Europe with the Aleutian high dominating the circulation over the Pacific. Eleven of the models simulate a predominantly wavenumber-1 structure, the other two (UCLA and LaRC-IMPACT) have an excessive wavenumber-2 component. Even though 11 models have more realistic patterns, the phases are incorrect in some of them: this is clearly seen by the longitude of the Aleutian high, which is too far to the west in SKYHI, FUB, MA/ECHAM4, and UK-UM, but too far to the east in MRI/lrf. The strength of the wave pattern is also incorrect in many of the models.

Fig. 4. Polar stereographic projections, from 20° to 90°N, of the multiyear mean geopotential height (dam) at 10 hPa in Jan. The contour interval is 16 dam and values in the range 2872–2936 dam are shaded. The observational estimate is a 17-yr mean from UKMO TOVS analyses (Bailey et al. 1993).
One issue, which will be the subject of further study, is the effect of sampling. Table 2 shows that the model integrations range in length from 20 years (ARPEGE-climate) down to 3 years (UCLA). The sampling issue is most severe for the runs of five years or less. For instance, the bias in the LaRC-IMPACT model arises because two of the Januarys are dominated by a strong wavenumber-2 pattern while the other two show a wavenumber-1 structure; the mean state of a longer run would depend on how often these two states were populated. Other reasons for the climatological biases include the relationship of the middle atmospheric waves to their tropospheric sources, their propagation and dissipation through the atmosphere, and their relationship (if any) to such features as the model resolution, diffusion parameterizations, and the presence of gravity wave drag. For instance, Pawson et al. (1998) showed (in the FUB model) that topographic gravity wave drag can be used to correct biases in the time-mean, zonal-mean structure to the detriment of the planetary waves; while that result might be model specific, it illustrates the sensitivity of such MACMs to the representation of physical processes.

One of these issues, the structure of the tropospheric planetary waves, is briefly discussed here. The 500-hPa geopotential height fields are examined (Fig. 5). Again, the sampling issue must be considered, as well as the fact that the models were integrated with sea surface temperatures and sea-ice distributions that repeat from year to year, while the observed patterns arise in a fully coupled system. Even though most MACMs do capture the dominant feature of the observed climate, the precise locations and strengths of the peaks and troughs vary considerably from model to model, and sometimes lie well outside the observed range of variability, indicating that there are fundamental problems with some MACMs. This has significant implications for climate studies: in the troposphere the mean state and the variability are interlinked, so that substantial biases in the mean state can lead to important discrepancies in the strength and structure of such features as the North Atlantic oscillation and the Pacific–North American mode. It also has consequences for the middle atmospheric state: errors in the strength of the tropospheric waves affect the amount of eddy energy input into the middle atmosphere, so that one of the dominant dynamical driving mechanisms for the stratospheric circulation may be significantly misrepresented.

e. Interannual variability of monthly mean velocities

An important aspect of the middle atmospheric flow is the interannual variability of the polar night jet. In the Northern Hemisphere the climatological structure of the vortex (e.g., Fig. 4) represents an average of cold winters, in which the polar temperatures remain essentially undisturbed over long periods, and warmer winters when the cyclonic circulation is interrupted by dynamical disturbances known as stratospheric sudden warmings. A measure of the interannual variability is given by the monthly averaged zonal wind at 10 hPa and 60°N (Fig. 6). Note that examination of time series at a single point can be slightly misleading because of the systematic differences in the climatological structures already discussed, but this discussion does illustrate aspects which supplement the descriptions of the mean state. The build up of the westerly jet from the summer into the winter shows little year-to-year variation, but after this phase there are large differences, some years showing strong jets (more than 50 m s\(^{-1}\)) and other years showing weaker winds. Many of the models reproduce this type of variability, but in some the oscillations are much larger than the observations (e.g., SKYHI and FUB), while other MACMs show too little year-to-year variation (e.g., GISS and MRI/lrf).

The different characteristics of the southern hemisphere polar winter are illustrated by the equivalent time series at 60°S (Fig. 7). The peak winter Southern Hemisphere jet is stronger than that in the Northern Hemisphere (more than 80 m s\(^{-1}\) vs 50 m s\(^{-1}\)). The interannual variability in the Southern Hemisphere jet speed is significantly less than that in the Northern Hemisphere, and the magnitude of the variability peaks later in the winter than in the Northern Hemisphere. In the Southern Hemisphere, the midwinter westerlies in this location are much too weak in the UCLA, GISS, and LaRC-IMPACT models, much too strong in MA/CCM2, Canadian MAM, and FUB, and within about 10 m s\(^{-1}\) of the observations in the other models. In the GISS model the parameterized traveling gravity waves, which benefit the simulation of the summertime stratosphere, provide too much forcing in the Southern Hemisphere winter causing the polar night jet to be too weak.

The models with strong midwinter jets all show a late transition into the summer circulation, with overly strong mean winds and too much variability in the springtime. This is a significant climatological
anomaly, first because it implies some missing forcing, but second because this feature would have significant implications if ozone were to be predicted in the models: the delayed breakdown of the polar vortex would lead to a later recovery of the modeled ozone hole, because the transport of midlatitude air into the polar cap would be inhibited.

4. Conclusions and discussion

In this study, some initial results from the GRIPS initiative, performed under the auspices of the WCRP/SPARC program, have been presented. The objectives of the paper are to give an initial overview of what is planned in the intercomparison and to show how the models compare at a basic level. Specifically, some results from task 1b (Table 1) have been shown; more

Fig. 5. Polar stereographic projections, from 20° to 90°N, of the multiyear mean geopotential height (dam) at 500 hPa in Jan. The contour interval is 16 dam and values in the range 548–564 dam are shaded. The observational estimate is a 14-yr mean from ERA-15 data (Gibson et al. 1997).
complete assessments of these and the other tasks are nearing completion and should be the subjects of future publications.

There is a tendency for the models to show a global-mean cold bias at all levels, but this increases with altitude through the troposphere and remains of similar magnitude through the stratosphere. There are latitudinal variations, with scatter to both sides of reasonable climatological estimates in the neighborhood of the tropical tropopause, and with differing cold biases elsewhere.

Generally, all of the models can simulate qualitatively the dominant features of the zonal-mean zonal wind in the extratropics, but there are many important departures from the observed climatology. These involve differences in the strength, location, and shape of the predominant jets in the atmosphere, and the level of the reversal from summertime tropospheric westerlies to stratospheric easterlies.

The time-averaged, three-dimensional structures show more severe departures from the observational reference. The strength and shape of the northern stratospheric polar vortex are simulated with a wide variety of success in the different models—some show an overly zonal structure while others generate an unrealistically strong wavenumber-2. Similar discrepancies are evident in the 500-hPa geopotential height fields, where some models fail to reproduce the ridges and troughs seen in observations. Although some results are sensitive to the length of the sample, some of these biases are severe and throw doubt upon the adequacy of some models to simulate the current climate and make predictions of future climate change.

Level-1 intercomparisons, such as that presented here, are valuable in that they allow some evaluation of the status of climate modeling. However, more quantitative studies are required to place better bounds on our understanding. This requires model intercomparisons with stricter controls on the boundary conditions. To this end, GRIPS is interacting with AMIP, encouraging the participants to make integrations conforming to the AMIP standards.

Other more process-oriented studies are being conducted. Phase 2 of GRIPS aims to understand the limitations of some aspects of the physical parameterizations. The first of these concerns the radiation schemes. As an extension of the Intercomparison of Radiation Codes in Climate Models (e.g., Ellingson

Fig. 6. Time series of the monthly mean zonal-mean zonal wind at 60°N and 10 hPa for each of the models and the observations (Randel 1992). The thin lines show results for each year of the dataset, except for the UCLA model where only the multiyear mean was provided. The mean value from observations is continued through all panels. The time axis goes from Jul (month 7) through Jun (month 18).
et al. 1991) and the comparison by Shine et al. (1995), a group is evaluating the radiative heating rates, calculated off-line using a standard set of input data (temperature, trace gas distributions) and boundary conditions. This will allow evaluation of the radiation schemes in the middle atmosphere. Further, by repeating the calculations with different ozone distributions (as used in the models) it will be possible to determine whether it is the specified trace gases or the radiation parameterizations that have the largest impact on the temperature biases. Such a study complements other radiative issues of more direct relevance to the troposphere, such as cloud–radiation interactions. The second sensitivity test is directed at the impact of mesospheric gravity wave drag on the climate: models are being integrated with their normal parameterizations replaced by specified forcing at high levels, with a view toward assessing model responses to given forcings.

Other important aspects that are being considered by GRIPS are centered on the ozone–climate issue. Possibly in conjunction with other international bodies, GRIPS is looking toward evaluations of the transport characteristics of the models and of their performance when coupled with interactive ozone chemistry. These studies will form a prelude to coordinated research aimed at addressing controversial issues linking ozone loss and climate change (e.g., Shindell et al. 1998).

Phase 3 of GRIPS will involve experiments designed to investigate the impact of stratospheric perturbations on climate. Important issues are impacts of volcanic aerosols in the lower stratosphere, impacts of ozone loss, and atmospheric response to solar radiation anomalies. All of these issues have been considered by one or more modeling groups, but to date there has been no coordinated effort to run several different models under the same controlled conditions. Given the spread in the present-day climate simulated by the models, there are likely to be deviations in their responses to imposed perturbations. An understanding of how anomalous forcing mechanisms impact the circulation of the various MACMs will be an important input to future assessments of the possible impacts of middle atmospheric change and variability on climate. This should be an important contribution of GRIPS to the scientific community.

Fig. 7. Time series of the monthly mean zonal-mean zonal wind at 60°S and 10 hPa for each of the models and the observations (Randel 1992). The thin lines show results for each year of the dataset, except for the UCLA model where only the multiyear mean was provided. The mean value from observations is continued through all panels. The time axis goes from Jan (month 1) through Dec (month 12).
Acknowledgments. We thank Marie-Lise Chanin and Marvin Geller, the cochairs of the SPARC scientific steering committee, for their encouragement in setting up and proceeding with GRIPS. We thank Alan Robock for suggesting that we submit this status paper. The model intercomparison has been funded by the European Union Framework IV “EuroGRIPS” project (ENV4-ST95-0123). SP acknowledges the FUB, the USRA, and the DAO for their support at different times during this project. KK acknowledges the MRI for allowing him to participate as cochair of GRIPS. Three annual GRIPS workshops were partially financed by Environment Canada (1996), the DAO at NASA GSFC (1998), USRA (1998), EuroGRIPS (1997, 1998), and the WCRP (1996, 1997, 1998). We are especially grateful to Roger Newson at WCRP for providing generous support. Also, all participants have been able to draw on resources of their national funding agencies and their home institutes for the scientific development of the models and the integrations necessary for such an intercomparison.

References


