



Title	The Aqua-Planet Experiment (APE): CONTROL SST Simulation
Author(s)	Blackburn, Michael; Williamson, David L.; Nakajima, Kensuke; Ohfuchi, Wataru; Takahashi, Yoshiyuki O.; Hayashi, Yoshi-Yuki; Nakamura, Hisashi; Ishiwatari, Masaki; Mcgregor, John L.; Borth, Hartmut; Wirth, Volkmar
Citation	Journal of the meteorological society of japan, 91A, 17-56 <a href="https://doi.org/10.2151/jmsj.2013-A02">https://doi.org/10.2151/jmsj.2013-A02</a>
Issue Date	2013-09
Doc URL	<a href="http://hdl.handle.net/2115/54604">http://hdl.handle.net/2115/54604</a>
Type	article
File Information	ishiwatari1.pdf



[Instructions for use](#)

## **The Aqua-Planet Experiment (APE): CONTROL SST Simulation**

**Michael BLACKBURN**

*National Centre for Atmospheric Science, University of Reading, Reading, UK*

**David L. WILLIAMSON**

*National Center for Atmospheric Research, Boulder, Colorado, USA*

**Kensuke NAKAJIMA**

*Faculty of Sciences, Kyushu University, Fukuoka, Japan*

**Wataru OHFUCHI**

*Japan Agency for Marine-Earth Science and Technology, Yokohama, Japan*

**Yoshiyuki O. TAKAHASHI, Yoshi-Yuki HAYASHI**

*Center for Planetary Science, Kobe, Japan  
Faculty of Science, Kobe University, Kobe, Japan*

**Hisashi NAKAMURA**

*Research Center for Advanced Science and Technology, The University of Tokyo, Tokyo, Japan*

**Masaki ISHIWATARI**

*Graduate School of Science, Hokkaido University, Sapporo, Japan*

**John L. McGREGOR**

*CSIRO Marine and Atmospheric Research, Aspendale, Australia*

**Hartmut BORTH**

*Theoretical Meteorology, University of Hamburg, Hamburg, Germany*

**Volkmar WIRTH**

*Institute for Atmospheric Physics, University of Mainz, Mainz, Germany*

**Helmut FRANK**

*Research and Development, Deutscher Wetterdienst, Offenbach, Germany*

**Peter BECHTOLD, Nils P. WEDI**

*European Centre for Medium-Range Weather Forecasts, Reading, Berkshire, UK*

**Hirofumi TOMITA**

*Advanced Institute for Computational Science, RIKEN, Kobe, Japan*

**Masaki SATOH**

*Atmosphere and Ocean Research Institute, The University of Tokyo, Kashiwa, Japan*

**Ming ZHAO, Isaac M. HELD**

*Geophysical Fluid Dynamics Laboratory, Princeton University, Princeton, New Jersey, USA*

**Max J. SUAREZ**

*Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA*

**Myong-In LEE**

*Ulsan National Institute of Science and Technology, Ulsan, Korea*

**Masahiro WATANABE, Masahide KIMOTO**

*Atmosphere and Ocean Research Institute, The University of Tokyo, Kashiwa, Japan*

**Yimin LIU**

*State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics,  
Institute of Atmospheric Physics, CAS, Beijing, China*

**Zaizhi WANG**

*National Climate Center, China Meteorological Administration, Beijing, China*

**Andrea MOLOD**

*Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA*

**Kavirajan RAJENDRAN**

*Center for Mathematical Modelling and Computer Simulation, National Aerospace Laboratories, Bangalore, India*

**Akio KITOH**

*Meteorological Research Institute, Tsukuba, Japan*

**and**

**Rachel STRATTON**

*Met Office, Exeter, UK*

*(Manuscript received 15 September 2011, in final form 17 August 2012)*

## Abstract

Climate simulations by 16 atmospheric general circulation models (AGCMs) are compared on an aqua-planet, a water-covered Earth with prescribed sea surface temperature varying only in latitude. The idealised configuration is designed to expose differences in the circulation simulated by different models. Basic features of the aqua-planet climate are characterised by comparison with Earth.

The models display a wide range of behaviour. The balanced component of the tropospheric mean flow, and mid-latitude eddy covariances subject to budget constraints, vary relatively little among the models. In contrast, differences in damping in the dynamical core strongly influence transient eddy amplitudes. Historical uncertainty in modelled lower stratospheric temperatures persists in APE.

Aspects of the circulation generated more directly by interactions between the resolved fluid dynamics and parameterized moist processes vary greatly. The tropical Hadley circulation forms either a single or double inter-tropical convergence zone (ITCZ) at the equator, with large variations in mean precipitation. The equatorial wave spectrum shows a wide range of precipitation intensity and propagation characteristics. Kelvin mode-like eastward propagation with remarkably constant phase speed dominates in most models. Westward propagation, less dispersive than the equatorial Rossby modes, dominates in a few models or occurs within an eastward propagating envelope in others. The mean structure of the ITCZ is related to precipitation variability, consistent with previous studies.

The aqua-planet global energy balance is unknown but the models produce a surprisingly large range of top of atmosphere global net flux, dominated by differences in shortwave reflection by clouds. A number of newly developed models, not optimised for Earth climate, contribute to this. Possible reasons for differences in the optimised models are discussed.

The aqua-planet configuration is intended as one component of an experimental hierarchy used to evaluate AGCMs. This comparison does suggest that the range of model behaviour could be better understood and reduced in conjunction with Earth climate simulations. Controlled experimentation is required to explore individual model behaviour and investigate convergence of the aqua-planet climate with increasing resolution.

**Keywords** comparison of atmospheric general circulation models (GCMs); idealized model configuration; global energy budget; tropical wave spectrum; precipitation

## 1. Introduction

The Aqua-Planet Experiment (APE) is a coordinated comparison of Atmospheric General Circulation Model (AGCM) simulations on a water-covered Earth. It compares idealized climates simulated by global AGCMs being developed and used for numerical weather prediction and climate research. The experiment involves complete AGCMs but on an idealized planet with simplified lower boundary interactions, namely an ocean surface everywhere with no land, orography or sea ice. The sea surface temperature (SST) is specified everywhere with a number of idealized analytic distributions designed by Neale and Hoskins (2000a).

APE was conceived by Neale and Hoskins as one component of a modelling hierarchy of increasing complexity, both of the models themselves and of the experimental configurations to which they are applied. The hierarchy has two distinct roles: an evaluation role in the development and testing of complex atmospheric models, and a conceptual role in linking complex models with theory and observation. This context of

APE is discussed in more detail by Blackburn and Hoskins (2013), so only a summary is given here.

In the conceptual context, theory and more idealised models provide constraints on the character of the global circulation expected in APE. The zonally averaged SST distributions in APE are broadly similar to Earth, so the equator-to-pole temperature difference is expected to give rise to a jet stream and storm track in mid-latitudes. In the tropics, the analytic model of Held and Hou (1980) predicts that a Hadley circulation with equatorial ascent must exist when the latitudinal profile of temperature in equilibrium with the underlying SSTs is steeper than quartic, in order to maintain a thermal wind balance consistent with angular momentum conservation. The tropical SST profiles in APE were designed to include this limiting case. The CONTROL SST case considered in this paper is quadratic, so the tropical circulation is expected to consist of a Hadley circulation in each hemisphere with equatorial ascent.

Previous modelling studies using aqua-planet configurations have confirmed that the tropical circulation and the location of the inter-tropical convergence zone (ITCZ) produced by AGCMs do depend on SST.

The most systematic study, by Hess et al. (1993), used a number of widely varying tropical SST profiles and two different convective parameterizations in a single model. Hess et al. and other studies found that the tropical circulation and ITCZ location also differ between models, with some models producing a double-ITCZ straddling the equator even when SST is peaked on the equator. Such a double-ITCZ was found in the first modern aqua-planet simulation, by Hayashi and Sumi (1986). The modelled tropical circulation can be affected not only by major modelling choices, such as the type of convective parameterization (e.g., Hess et al. 1993), but also by more detailed choices such as model time-step (Williamson and Olson 2003). ITCZ location can also be affected by disabling specific physical processes and interactions in models, such as the wind speed dependence of surface evaporation (Numaguti 1993; Chao and Chen 2004).

These studies, in particular that of Hess et al., provided motivation for Neale and Hoskins (2000a) to propose a benchmark test suite of AGCM aqua-planet experiments. In this evaluation context, APE is a bridge in the modelling hierarchy between experiments with models of reduced complexity that are used in the development of individual model components, and realistic simulations of Earth's climate using complete AGCMs, coordinated through the Atmospheric Model Intercomparison Project (AMIP, described in its original form by Gates 1992).

Neale and Hoskins (2000b) used a single AGCM to characterise the climate for each of the APE SST distributions. The aims of the APE coordinated comparison described here are to provide a benchmark of current model behaviours and to stimulate research to understand the cause of inter-model differences. For this reason APE is sponsored by the World Meteorological Organization (WMO) joint Commission on Atmospheric Science (CAS), World Climate Research Program (WCRP) Working Group on Numerical Experimentation (WGNE).

In this and a companion paper (Williamson et al. 2013) we provide an overview of a variety of aspects of the aqua-planet simulations. Detailed analysis of specific aspects is presented in other papers in this Special Issue or left for future individual studies. This paper considers simulations based on the CONTROL SST distribution (defined below). The simulated responses to varying the latitudinal profile of SST are discussed in the companion paper, using experiments based on the other zonally symmetric SST distributions defined by Neale and Hoskins. More comprehensive diagnostics are presented in Williamson et al. (2012),

a collection of model simulation statistics that will be referred to in the following as the APE ATLAS.

## 2. APE experimental strategy

### 2.1 Sea surface temperature

The APE design originated with Neale and Hoskins (2000a) who specified experiments based on eight different SST fields, of which five are zonally symmetric varying in latitude only and three impose tropical longitudinal perturbations on one of those zonally symmetric distributions. The philosophy of the design is described in Neale and Hoskins (2000a) and in Blackburn and Hoskins (2013) and is not repeated here. Only one of those 8 profiles is considered here which is zonally symmetric and labelled CONTROL. The specified zonally symmetric SST in °C is given by  $27[1 - \sin^2(3\phi/2)]$  for latitude  $\phi$  between  $\pm\pi/3$ , and 0 for  $|\phi| \geq \pi/3$ . Compared to present day Earth, this SST profile is more sharply peaked about the equator and has smaller equator-to-pole temperature difference because of the 0°C limit beyond 60° latitude.

### 2.2 Detailed specification

Certain model specifications were included in the APE design to obtain uniformity across the models. They are divided into Requirements and Recommendations. Complete details can be found in the APE ATLAS and on the APE website. [Available at <http://climate.ncas.ac.uk/ape/>.]

The insolation is perpetual equinoctial, symmetric about the equator, but including the diurnal cycle. The recommendation to achieve this is to modify the Earth orbit parameters, setting eccentricity and obliquity to zero, giving a circular equinoctial orbit. The distribution of insolation is then independent of the calendar. The solar constant is  $1365 \text{ W m}^{-2}$ . Ozone is taken from the annual mean climatology used in AMIP II (Wang et al. 1995; Liang and Wang 1996), symmetrized about the equator. CO<sub>2</sub> is to be set at 348 ppmv, as in AMIP II. Recommendations for the other well-mixed radiatively active gases also follow AMIP II: CH<sub>4</sub> at 1650 ppbv and N<sub>2</sub>O at 306 ppbv. Halocarbon concentrations should yield  $\sim 0.24 \text{ W m}^{-2}$  radiative forcing. The use of an “equivalent” CO<sub>2</sub> is discouraged. It is also recommended to exclude radiatively active aerosols. Any aerosol specification for cloud condensation should use an oceanic distribution which is fixed in time, zonally symmetric and symmetric about the equator. Recommended values are provided for geophysical constants and parameters.

It is recommended that the initial dry mass of the

Table 1. Requested and derived diagnostic data

DESCRIPTION		TIME AVERAGING	DOMAIN
GT	Global Time-Series	Daily	Global, Area Average
TR	Transients	6-hourly (year 3)	Global/tropical lat-lon
SH	Single-Level 2-D Means	3-yearly	Global lat-lon
ML	Multiple-Level 3-D Mean	3-yearly	Global lat-lon, P17 <sup>1</sup>
MF	Multiple-Level Fluxes	3-yearly	Zonal average, P17
PF <sup>2</sup>	Parameterization Forcing	3-yearly	Global lat-lon, model-level
TE <sup>3</sup>	Transformed Eulerian Means	3-yearly	Zonal average, P17
VB <sup>3</sup>	Vertically Integrated Budgets	3-yearly	Zonal average

<sup>1</sup>P17 denotes 17 WMO pressure levels, 10–1000 hPa.

<sup>2</sup>Optional.

<sup>3</sup>Derived.

atmosphere be equivalent to a global mean surface pressure of 101080 Pa. This is 101325 Pa minus 245 Pa, which corresponds to a global moisture content of 25.006 kg m<sup>-2</sup>. It is desirable that dry mass be conserved throughout the integration, but this, of course, depends on the characteristics of the model dynamical core.

### 2.3 Simulation period

The simulations are run for 3.5 years, with the analysis made over the last 3 years, omitting the first 6 months as spin-up. A model-simulated state, from either a previous aqua-planet integration or an earth-like simulation, should be used for the initial conditions. In independent tests of the spin-up in several models, aqua-planet climate equilibrated in a matter of weeks from these types of initial condition, so a 6 month spin-up is considered adequate. Nevertheless, the modelling groups were instructed to check that equilibration was achieved during the discarded initial 6-month period. Daily time-series of global averages were requested for a number of variables, which allows a gross check that equilibrium was reached.

The book-keeping calendar can be a 365 or 360 day year, with natural or 30-day months respectively. Insolation does not follow the calendar. Note that the 3.5 year integration length means that a realistic calendar can be used if integrations are started in March of a leap year.

### 2.4 Diagnostic data

The APE proposal requested the submission to a data archive of a variety of data from the final three years of the simulations, in the categories listed in Table 1. The diagnostic list is greatly expanded beyond the variables suggested by Neale and Hoskins (2000a). A complete list of variables in each category is included in the

ATLAS and is available from the APE website.

The required diagnostics comprise the top 5 categories in Table 1. Global energy and moisture balances are analysed using the time series of globally averaged single level variables (GT). These also allow simulations to be checked for residual drift. Temporal variability of selected dynamical variables and tropical cloud and precipitation are analysed using 6-hourly time-series of global/tropical distributions of selected single-level variables (TR). The time average state is obtained from the 3-D atmospheric variables (ML) and 2-D single level variables (SH). The balance of processes that maintains the zonal average atmospheric state is obtained from the dynamical variances and covariances (“fluxes”, MF) and the requested (optional) parameterization forcing (PF) diagnostics.

The APE proposal also included specification of derived diagnostics that can be computed from the required and requested diagnostics. These comprise individual variables in each of the preceding categories, plus two additional categories of time average data: zonal averages of Transformed Eulerian Means (TE) and vertically integrated budgets (VB).

The standard diagnostics were considered necessary for the adequate analysis and comparison of the experiments. The list is an amalgamation of requested diagnostics taken from the AMIP II standard model output, the WGNE standard diagnostics of mean climate and selected diagnostics that have proved useful in analyzing previous aqua-planet and dynamical core experiments. Many, but not all, of those diagnostics along with additional ones that have proven informative are included in the ATLAS. A selection of those is discussed here to illustrate the general characteristics of the aqua-planet climate and the variation among models.

Table 2. Participating models

	GROUP SYMBOL	LOCATION	MODEL	OPTIMISED TOA? <sup>1</sup>
1	AGU	Japan (consortium)	AFES	No
2	CGAM	Reading, UK	HadAM3	Yes
3	CSIRO	Aspendale, Australia	CCAM	No
4	DWD	Offenbach/Mainz, Germany	GME	Yes (NWP)
5	ECM-CY29	Reading, UK	IFS cy29r2	No
6	ECM-CY32	Reading, UK	IFS cy32r3	No
7	FRCGC	Yokohama, Japan	NICAM	No
8	GFDL	Princeton, USA	AM2.1	Yes
9	GSFC	Maryland, USA	NSIPP-1	No
10	K1JAPAN	Japan (collaboration)	CCSR/NIES 5.7	Yes
11	LASG	Beijing, China	SAMIL	No
12	MIT	Cambridge, USA	MIT-GCM	No
13	MRI	Tokyo, Japan	MRI/JMA98	No
14	NCAR	Boulder, USA	CCSM-CAM3	Yes
15	UKMO (N48)	Exeter, UK	pre-HadGAM1	Weakly <sup>2</sup>
16	UKMO (N96)	Exeter, UK	pre-HadGAM1	Weakly <sup>2</sup>

<sup>1</sup>Was the top of atmosphere radiative balance optimised for present day Earth climate?

<sup>2</sup>During the development phase of HadGEM1 the TOA fluxes of AMIP runs were monitored to check that they did not widely diverge from balance but were not actively tuned.

All standard diagnostic data received from the APE modelling groups are available from a Data Archive at the APE website.

### 2.5 Data processing

Diagnostic data were supplied on a regular latitude-longitude grid unique to each model. In most cases this is the model's native grid, in the case of spectral transform models the transform Gaussian grid. Data from models with non-regular grids were first interpolated to a regular grid by each modelling group. In addition, ECMWF interpolated from the model's transform grid to a uniform 2° latitude-longitude grid. Multi-level data were interpolated to 17 standard pressure levels by each modelling group, with the exception of parameterization forcing (PF) data, which were requested on model levels to retain the complete vertical structure.

For the diagnostics presented here and in the ATLAS, data for individual models are plotted on the submitted grid without further interpolation. Multi-model means and standard deviations are obtained by first linearly interpolating to a common 1° latitude-longitude grid before averaging the available models.

## 3. Participating Models

Table 2 lists the 16 models that participated in the

APE intercomparison. These include a number of established models used in production applications for both numerical weather prediction (NWP) and climate, including the Intergovernmental Panel on Climate Change (IPCC), and a number of novel models less tested and iterated in real-world applications. The APE intercomparison may therefore be expected to produce a wider spread in many circulation metrics than for previous intercomparisons such as AMIP.

Table 2 includes the commonly accepted model name, the group that contributed the data and the location of the group's home institution: more details including names of the contributing modellers are provided on the APE website. The group symbol is used to identify the models in figures, tables and discussions. The final column identifies whether each model was optimised to give top of atmosphere radiative balance for present day Earth climate. This will be used in assessing the global energy budget for the APE CONTROL experiment.

The participating models span a range of horizontal and vertical resolution reflecting their climate and NWP applications. In addition there is one global cloud-system resolving model, although this was only integrated for a 30 day period. The models include a diverse range of both dynamical core and parameterization choices. Table 3 provides some basic charac-

Table 3. Dynamical properties of participating models

GROUP SYMBOL	DYNAMICAL CORE	WATER VAPOR TRANSPORT	HORIZONTAL RESOLUTION	VERTICAL RESOLUTION
AGU	Eulerian spectral	Eulerian spectral	T39	L48
CGAM	lat-lon grid point	Eulerian grid	$3.75^\circ \times 2.5^\circ$	L30
CSIRO	C-C <sup>1</sup> semi-Lag <sup>2</sup>	semi-Lag	~210 km (C48)	L18
DWD	icosahedral grid	semi-Lag grid	~1°	L31
ECM-CY29	semi-Lag spectral	semi-Lag grid	T159	L60
ECM-CY32	semi-Lag spectral	semi-Lag grid	T159	L60
FRCGC	icosahedral Eulerian	Eulerian	~7 km	L54
GFDL	lat-lon finite volume	finite volume	$2.5^\circ \times 2^\circ$	L24
GSFC	lat-lon grid point	Eulerian centered	$3.75^\circ \times 3^\circ$	L34
K1JAPAN	Eulerian spectral	semi-Lag grid	T42	L20
LASG	Eulerian spectral	Eulerian grid	R42	L9
MIT	cubed sphere	Eulerian grid	~280 km	L40
MRI	Eulerian spectral	Eulerian spectral	T42	L30
NCAR	Eulerian spectral	semi-Lag grid	T42	L26
UKMO (N48)	semi-Lag lat-lon grid	semi-Lag	$3.75^\circ \times 2.5^\circ$	L38
UKMO (N96)	semi-Lag lat-lon grid	semi-Lag	$1.875^\circ \times 1.25^\circ$	L38

<sup>1</sup>C-C denotes conformal cubic

<sup>2</sup>semi-Lag denotes semi-Lagrangian

teristics of the models including the type of dynamical core, the method used for water vapor transport if it differs from that of the dry dynamics, the horizontal resolution and the number of model levels. Table 4 lists major choices in the model parameterizations of convection and boundary layer turbulence. These can only provide labels for the different model characteristics that are indicative of their definitions and serve as a reminder. The actual schemes must be determined from the additional information and references on model formulation presented in the ATLAS.

#### 4. Forcing of the global circulation

Before analysing circulation statistics from the aqua-planet simulations, the forcing of the aqua-planet circulation is compared with that of Earth. Although APE is configured as a perpetual equinox, it is compared with annual averages for Earth, to average over the observed large annual cycle of poleward energy transport and to avoid differences associated with seasonal uptake and release of heat by the ocean. Figure 1 shows the equinoctial insolation for APE, which varies simply with cosine of latitude, and the annual average for Earth as calculated by the NCAR model. The APE insolation has a larger latitudinal gradient than the annual mean at all latitudes, so tends to force a stronger circulation. By prescribing SST, the fraction of insolation reaching

the surface is lost, although an implied ocean poleward heat transport can be computed to balance the net surface energy flux. The SST profile for the APE CONTROL experiment, also shown in Fig. 1, is more peaked about the equator than the annual average for Earth, with unlimited moisture availability at all longitudes: this is also expected to force a stronger meridional circulation in the tropical atmosphere. In contrast, the APE SST is constant beyond 60° latitude, opposing the effect of increased insolation gradient on poleward energy transport there.

Figure 2 shows the poleward energy transport for the APE CONTROL SST experiment, calculated following Gleckler et al. (1995). The multi-model mean is shown as solid lines and the standard deviation of the models about the mean is shaded. The total transport by the atmosphere and an implicit underlying ocean is that required to balance the net top of atmosphere (TOA) energy flux; the oceanic transport is that required to balance the net surface energy flux and the atmospheric transport is the difference between the two. Globally averaged net fluxes are not relevant to the transport and have been removed: the global energy budget is discussed in Section 7. Observational estimates of the annual average poleward transport for Earth from Fasullo and Trenberth (2008) are included in Fig. 2 for comparison.

Table 4. Parameterizations of participating models

GROUP SYMBOL	PBL	SHALLOW CONVECTION	DEEP CONVECTION
AGU	Mellor-Yamada	None	Emanuel
CGAM	Smith	Gregory-Rowntree	Gregory-Rowntree
CSIRO	Holtstlag-Boville	None	McGregor
DWD	Louis	Tiedtke	Tiedtke
ECM-CY29	Louis-Beljaars	Tiedtke	Bechtold et al. 2004
ECM-CY32	Louis-Beljaars	Bechtold et al. 2008	Bechtold et al. 2008
FRCGC	Mellor-Yamada	None	None
GFDL	Lock	RAS <sup>1</sup>	RAS
GSFC	Louis	RAS	RAS
KIJAPAN	Mellor-Yamada	None	Pan-Randall
LASG	Local vert diffusion	None	Manabe
MIT	Mellor-Yamada	RAS	RAS
MRI	Mellor-Yamada	Randall-Pan	Randall-Pan
NCAR	Holtstlag-Boville	Hack	Zhang-McFarlane
UKMO (N48)	Lock/Richardson	Gregory 1990/Grant	Gregory 1999
UKMO (N96)	Lock/Richardson	Gregory 1990/Grant	Gregory 1999

<sup>1</sup>RAS denotes relaxed Arakawa-Schubert.

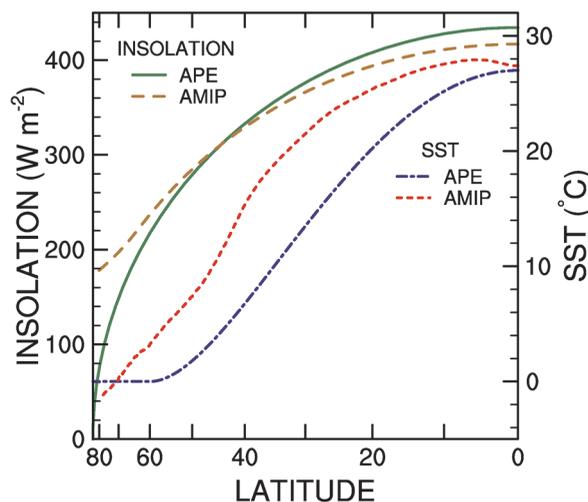


Fig. 1. Incident solar short wave flux at the top of atmosphere (left axis), zonally averaged, comparing the equinoctial insolation specified in APE (solid) and the annual average for Earth (dashed). Sea surface temperature (right axis), zonally averaged, specified for the APE CONTROL experiment (dash-dot) and the annual average for Earth (short-dash). Earth values are averaged over northern and southern latitudes and SST is set to  $-1.8^{\circ}\text{C}$  under sea-ice.

The total energy transport for the APE models is rather similar to Earth, reflecting similar TOA net flux distributions, but the peak transport is slightly stronger and shifted poleward in APE. However, the differences of approximately 0.5 PW are of the same order as uncertainties in the observational estimate (Fasullo and Trenberth 2008) and, except at high latitudes, are less than the standard deviation of the APE models, which is large, reaching around 1 PW in mid-latitudes. The atmospheric transport dominates, with APE having a similar magnitude but an equatorward shift relative to Earth in mid-latitudes. In particular the APE transport is larger by around 1 PW in the tropics, giving a sub-tropical “shoulder” to the distribution. The atmospheric transport is significantly more constrained among the APE models than the total transport, with the inter-model standard deviation reaching only 0.4 PW in the sub-tropics. This combines with the large spread in total energy transport to give a large standard deviation, around 1 PW, in the implied ocean transport required for energy balance. This implies that significantly different equilibrated climates would result if the APE models were coupled to either thermodynamic slab oceans or dynamic ocean models.

The opposing latitudinal shifts in total and atmospheric transports are associated with significantly different net surface energy flux and implied ocean transport. For APE the ocean transport is weak in the tropics and increases through mid-latitudes, whereas

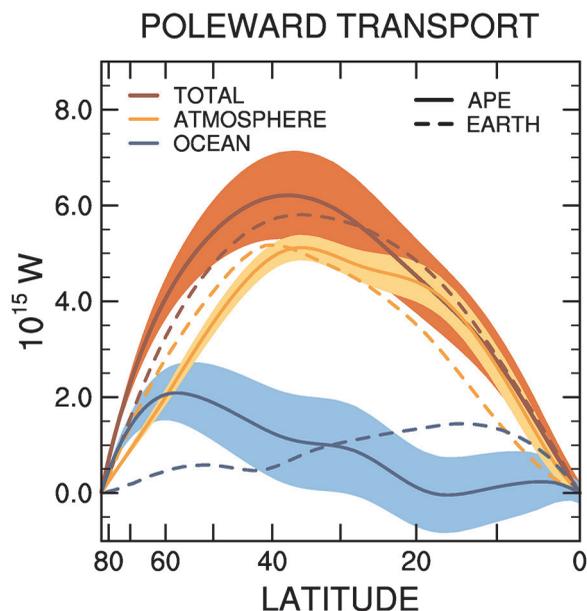


Fig. 2. Poleward energy transport for the APE CONTROL experiment: multi-model mean (solid lines) and standard deviation (shading). Total, ocean and atmosphere transports are inferred, respectively, from the top of atmosphere and surface net fluxes and their difference. Annual mean observational estimates for Earth from Fasullo and Trenberth (2008) (dashed). Values are zonal-time averages, averaged over northern and southern latitudes.

for Earth it peaks in the sub-tropics and decreases through mid-latitudes. Despite the large spread, all the APE models share this distribution of ocean transport: most models have weak poleward transport throughout the tropics but three outlying models have significant equatorward transport and reduce the average to near zero (see the ATLAS).

The differing latitudinal profiles of poleward transport for the APE CONTROL and Earth reflect different partitioning of the net TOA heat flux between heating of the atmosphere and ocean. At high latitudes the total poleward transport and its partition between atmosphere and ocean are insensitive to changes in the APE SST profile, as discussed by Williamson et al. (2013). The total poleward transport is  $\sim 0.75$  PW stronger in APE than for Earth at  $60^\circ$  latitude, forced by the stronger latitudinal gradient of insolation. The atmospheric transport is  $\sim 0.75$  PW weaker than Earth, presumably due to the absence of SST gradient poleward of  $60^\circ$  in APE. The warm isothermal SST also results in higher upward net heat flux at the polar

surface, requiring larger oceanic heat flux convergence for energy balance and thus a much larger poleward transport than Earth, by  $\sim 1.5$  PW at  $60^\circ$  latitude.

In the deep tropics, the TOA net downward flux is around  $75 \text{ W m}^{-2}$  for both APE and Earth. Of this, less than one third goes into the surface for the APE CONTROL, compared with two thirds for Earth (see Fasullo and Trenberth 2008, Fig.4 and the ATLAS). As the APE SST profile is flattened, going from the CONTROL to QOBS to FLAT cases defined by Neale and Hoskins (2000a), the TOA net flux changes little but the surface net flux increases substantially equatorward of  $15^\circ$  and decreases poleward to around  $40^\circ$ . The tropical atmospheric transport therefore decreases and the implied ocean transport increases with a relatively fixed total. The atmospheric dominance of tropical poleward energy transport in the APE CONTROL simulation is therefore ultimately determined by the more peaked SST profile than on Earth. This is discussed further in Williamson et al. (2013).

## 5. Zonal mean state

In this section we consider the time average, zonal averages of the multi-model mean and the variation among the models. The multi-model mean here includes all the models in Table 2 except FRCGC. Only a 30 day period was provided for the CONTROL experiment for this global cloud-system resolving model, due to its computational cost, so the model climate is unlikely to have reached equilibrium. The remaining models are equally weighted despite some models being outliers in various statistics. However, the mean includes a sufficient number of models that the impact of outliers is generally small.

The variability among models is summarized by plots of inter-model standard deviation or is described in words without showing the individual models. These discussions are intended to emphasize the great variation among the models in most of the statistics examined. The corresponding plots are available in the ATLAS but specific figure numbers are not provided here. In general all the individual models differ significantly from the multi-model mean.

The time average zonal averages are symmetric about the equator, which is to be expected given the hemispheric symmetry of the forcing. The two hemispheres are therefore averaged together to reduce the noise slightly and only one hemisphere is plotted, in the sense of the Northern Hemisphere. In addition, sine of latitude is used for the abscissa, to provide more detail of the tropical regions which exhibit more structure for many statistics.

### 5.1 Zonal wind and temperature

The multi-model mean zonal average state shown in Fig. 3 is very earth-like in structure, comparable to a symmetrized observational estimate, but with significantly stronger westerly jets. The jet core is  $62 \text{ m s}^{-1}$  and occurs at  $30^\circ$  latitude and 200 hPa. The maximum low-level westerly wind occurs only  $5^\circ$  poleward of the jet at  $35^\circ$ , so there is no separation between an upper tropospheric sub-tropical jet and the deep eddy-driven jet. This is due to the SST distribution, which is more peaked about the equator than on Earth, as discussed in Williamson et al. (2013), where the impact of changes in latitudinal SST profile on jet location and structure is considered. The zonally symmetric nature of the SST also contributes to the overall strength of the zonally averaged flow: the trade winds and low level polar easterly flow are also stronger than observed in the annual mean and in any individual season in ECMWF reanalyses (Källberg et al. 2005). The polar easterlies extend further equatorward towards the more equatorward mid-latitude westerlies.

The inter-model standard deviation generally increases with height in the troposphere, along with mean wind speed. However, it is a relative minimum close to the jet axis, with peak values both poleward and equatorward of the jet co-located with the maximum wind gradients. This is generally due to a subtle poleward or equatorward shift of the jet in the individual models (seen as a dipole structure in differences from the multi-model mean in the ATLAS) rather than to a widening or narrowing of the jet. The peak in standard-deviation is aligned vertically on the poleward flank of the jet but tilts poleward with height on the equatorward side, reminiscent of the structure of the leading mode of annular variability found in dynamical core simulations, which describes latitudinal displacements of the jet (see e.g., Sparrow et al. 2009). The standard deviation at the jet maximum is around  $3 \text{ m s}^{-1}$ . Individual model maximum values range from  $60 \text{ m s}^{-1}$  to  $67 \text{ m s}^{-1}$ , with the exception of MIT which is the lowest at  $54 \text{ m s}^{-1}$ . The peak standard deviation on the jet flanks reaches around  $5 \text{ m s}^{-1}$ , while near the surface it is  $1 \text{ m s}^{-1}$  or less. The standard deviation is somewhat higher on the equatorward flank of the jet, but the extent to which model variations in the Hadley circulation may contribute to this is unclear.

In contrast to these modest values of inter-model standard deviation in the troposphere, the spread of zonal mean zonal wind increases significantly above 150 hPa in the lower stratosphere. There is a localized peak above the equatorial tropopause, but the largest values occur above the poleward flank of the jet. This

is due to upward and poleward extension of the tropospheric jet into the stratosphere in several models, forming a second zonal wind maximum near  $50^\circ$  at the model top, while a larger number of models close off the tropospheric jet and form a well separated second maximum further poleward (evident in individual model zonal winds in the ATLAS). This large variation in stratospheric flow may be associated with the top boundary conditions of the models. For example, the NCAR CAM3 applies a  $\nabla^2$  horizontal diffusion at the top three levels of the model as a simple sponge to absorb vertically propagating planetary wave energy and also to control the strength of the stratospheric jets. The diffusion coefficient is chosen arbitrarily to yield reasonable polar night jets in earth-like simulations (Collins et al. 2004). This diffusion provides a strong control on the jet and thermal structure near the top of the model.

The multi-model mean zonal average temperature, shown in Fig. 3, is also superficially very earth-like. The temperature difference between the tropics and  $60^\circ$  latitude is similar to the annual average on Earth, but the tropospheric temperature gradient is more limited in latitudinal extent and located further equatorward, associated with the latitudinal SST profile. This is consistent with the stronger and more equatorward westerly jet in the aqua-planet than in reality. Beyond  $60^\circ$  latitude the temperature gradient is weak in the aqua-planet by design, particularly in the lower troposphere. Here the polar atmosphere is less statically stable than reality, so temperature gradients are closer to observed in the upper troposphere.

The inter-model standard deviation of temperature generally increases with height, with a strong increase immediately below the tropopause at all latitudes. At each level there is a localized maximum in standard deviation, co-located with the maximum temperature gradient. In the upper troposphere this is close to the centre of the jet at  $30^\circ$  latitude and is co-located with the minimum standard deviation of zonal wind between the two peaks either side of the jet, consistent with thermal wind balance. The localized maximum extends downward following the maximum temperature gradient on the equatorward flank of the jet. Throughout the mid-troposphere the largest standard deviation is therefore in the sub-tropics. This may also reflect the relative lack of thermal contact between descending air above the trade wind inversion and the underlying SST: it is notable that the standard deviation increases strongly between 925 hPa and 850 hPa in the sub-tropics. In contrast, equatorial deep convection communicates the SST throughout the troposphere. In mid-latitudes deep

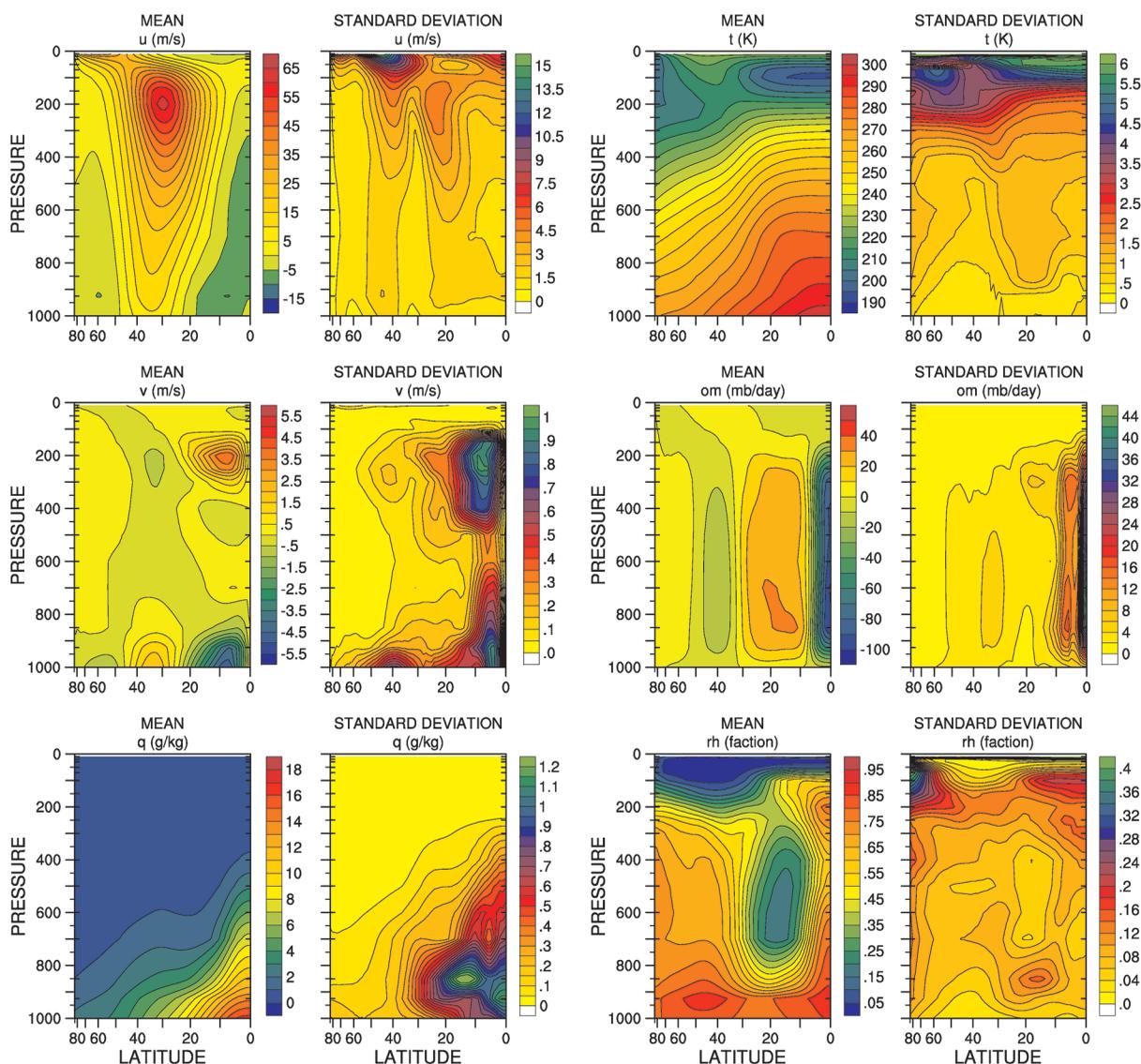


Fig. 3. Zonal-time average multi-model mean and standard deviation for zonal wind ( $u$ ),  $\text{m s}^{-1}$ ; temperature ( $t$ ), K; meridional wind ( $v$ ),  $\text{m s}^{-1}$ ; pressure vertical velocity  $\omega$  ( $om$ ),  $\text{mb day}^{-1}$ ; specific humidity ( $q$ ),  $\text{g kg}^{-1}$  and relative humidity ( $rh$ ).

convection also occurs in equatorward moving air in baroclinic waves, communicating the SST higher into the troposphere. Larger values of standard deviation occur over the polar cap, due to a variety of convective behaviour among the models there.

The inter-model standard deviation of temperature is significantly larger in the stratosphere than the troposphere, with a maximum value of 6 K in the region of strong vertical temperature gradient above the tropical tropopause. This is influenced by one outlier model

but a number of models differ from the multi-model mean by almost 10 K (individual models are shown in the ATLAS). Historically AGCMs have exhibited large cold biases at the polar tropopause, as discussed by e.g. Boer et al. (1992) and Gates et al. (1999). There is evidence that use of a semi-Lagrangian dynamical core reduces the bias, although it remains a stubborn problem in some models (Williamson and Olson 1994; Chen and Bates 1996; Williamson and Olson 1998). Of the APE models, two of the three that use semi-La-

grangian dynamical cores (ECMWF and UKMO) are significantly warmer than the others in the sub-polar lower stratosphere, on the poleward side of the tropospheric jet, confining the jet on its poleward upper flank. The CGAM model (HadAM3 with 30 levels) is also warmer than the multi-model mean in the polar lower stratosphere, consistent with the impact of increased vertical resolution in that model seen by Pope et al. (2001). The remaining semi-Lagrangian model, CSIRO, exhibits the opposite sense of temperature anomalies in the lower stratosphere, possibly due either to its lower vertical resolution or its treatment of the vertical trajectory calculation.

The majority of the features of the APE inter-model variation discussed above are also seen in AMIP simulations of Earth climate performed as part of phase 3 of the Coupled Model Intercomparison Project (CMIP3, Meehl et al. 2007). These are plotted for the annual mean climatology in the ATLAS. The standard-deviation of zonal wind scales approximately with the multi-model mean wind, which is lower in the AMIP annual mean than in APE by a factor of two. The zonal wind variation in AMIP is proportionally less than this only on the equatorward flank of the mid-latitudes jets, while it is higher in the tropical upper troposphere. Temperature variation among the AMIP models is generally similar to APE, but the peak variation above the tropical tropopause is somewhat smaller in AMIP, presumably reflecting the use of established and optimized AGCMs, compared to a mix of established and recently developed models in APE. The AMIP temperature variation is larger than APE at low levels in the northern (but not southern) mid-latitudes, reflecting the interactive land surface in AMIP compared to prescribed SST in APE.

### 5.2 Meridional circulation

The Eulerian mean meridional circulation in Fig. 3 shows a three-cell circulation in each hemisphere similar to that for Earth, with a strong Hadley cell in the tropics, a thermally indirect circulation that straddles the jet and surface westerlies near 35° latitude, further equatorward than for Earth, and a weak thermally direct polar circulation. The Hadley circulation is much stronger than for Earth at equinox, with meridional wind reaching 5 m s<sup>-1</sup> in the boundary layer and at 200 hPa for the multi-model mean. This is mainly due to the SST profile, as discussed earlier in Section 4, but zonal symmetry of the SST and the proximity of the extra-tropical jet may also contribute. Strong upward motion is centered on the equator and downward motion extends from around 10° to just beyond

30° latitude. The multi-model mean meridional wind shows a slight hint of secondary cells in the vertical in the tropics between the main equatorward surface flow and the poleward upper tropospheric flow. There is also a barely discernable mid-tropospheric minimum in the equatorial ascent. In fact all models except MIT show equatorward flow to some degree below the upper level poleward flow and many models also show a poleward flow above the low level equatorward flow.

The standard deviations of mean meridional wind and vertical motion are superficially similar to their respective mean distributions, so variation among the models is primarily due to differences in the strength of the meridional overturning. However there are also significant differences in both vertical and latitudinal structure.

The deep structure of the standard deviation of mean meridional wind compared to its mean reflects variations in the vertical structure of the Hadley circulation among the models. The tropical upper tropospheric poleward flow peaks at 200 hPa in all but two models and the top of the equatorial upward motion varies from 200 hPa to 150 hPa. However, the strength and depth of both the upper level outflow and the boundary layer inflow vary significantly. Variations in the strength of the secondary cells also contribute to the standard deviation of meridional wind near 700 hPa and 400 hPa and to the peak in variation of upper tropospheric vertical motion near 15° latitude. Here, some models show a distinct mid-level minimum descent, with a second peak in the upper troposphere.

In the lower troposphere, the variation in low-level equatorward flow maximises equatorward of the multi-model mean, pointing to significant differences in the latitudinal structure of low level convergence in the ITCZ among the models. The variation in vertical velocity is largest in the equatorial region and is due to large variations in both the strength of the upward motion and its latitudinal structure. All models have ascent at the equator and the majority form a single maximum there, but the width of the ascent varies and several models form a double maximum straddling the equator. This differing structure is responsible for the very large variation at the equator itself and for the double peak in latitude, with minimum variation near 5°.

In mid-latitudes, the maximum variations in meridional wind and vertical motion are shifted in latitude relative to their means. This is due to joint variations in the strength and latitude of the Ferrel cell: in a subset of models the jet is further poleward and the Ferrel cell is stronger and extends further poleward.

The different sub-grid scale parameterizations employed in the models are likely to play a major role in driving the large variations in the strength and spatial structure of the tropical meridional and vertical motion. However, sensitivity studies in the NCAR model by Williamson and Olson (2003) and Williamson (2008b) show that details of the implementation of the parameterizations and the resolution at which they are applied may also lead to significant differences. In contrast, dynamical processes play a more significant role in variation of the mid-latitude Eulerian mean circulation (see Section 6.1 below).

### 5.3 Water vapour

The multi-model mean specific humidity in Fig. 3 is determined predominantly by the distribution of temperature, but the narrow equatorial peak in mid-tropospheric humidity reflects the strong sub-tropical descent centered on 20° latitude in the APE CONTROL. Here the relative humidity is below 20% for the multi-model mean. On the equator, high relative humidity extends well above the boundary layer to 600 hPa. There is a separate maximum at the Hadley cell outflow level and a minimum below this at 400 hPa. It is unclear whether this minimum is associated with the secondary cells discussed above, since the secondary inflow (equatorward meridional wind) is located below the minimum relative humidity. The height of the boundary layer is most strongly confined in the sub-tropics, by descent of dry air above. Further poleward, high relative humidity extends higher, due to boundary layer ventilation by baroclinic waves in mid-latitudes and by convection over the polar cap. Very dry air in the lower stratosphere, with relative humidity below 5% in the multi-model mean, extends above the tropopause temperature minimum at most latitudes. The exception to this is the polar cap, where colder air extends up to 10 hPa and the mean relative humidity exceeds 10%.

In the troposphere the largest humidity variation among the models occurs at the sub-tropical boundary layer inversion and is seen in both absolute and relative humidity variation at 850 hPa. The boundary layer top is not well resolved in the interpolated isobaric data but, in a number of models 850 hPa is clearly above the boundary layer in relatively dry air, while in other models the relative humidity gradient occurs mainly between 850 hPa and 700 hPa (individual models are shown in the ATLAS). An additional level at 775 hPa would have enabled better resolution of these vertical structures. Relative humidity also varies widely among the models in the region of strong vertical gradient at

the tropopause. In the tropics the maximum variation is located at the cold point, 100 hPa. This is distinct from the region of maximum temperature variation higher up, and is instead associated with variations in the vertical extent of high tropospheric humidity. In the majority of models the relative humidity maximum is coincident with the poleward meridional flow maximum at or below 200 hPa, but in three models (ECM-CY32, GFDL, K1JAPAN) it is located at 100 hPa, above the layer of poleward flow. This structure is worthy of further investigation. It could be associated either with differences in the vertical profiles of convective heating and detrainment of moisture or with large scale advection in these models. There is a secondary maximum of humidity variation at 300 hPa, also associated with these surprisingly large variations in vertical humidity structure in the equatorial upper troposphere.

Globally, the largest variation in relative humidity among the models occurs in the polar lower stratosphere at 70 hPa, where the mean relative humidity is only 11% but its standard deviation exceeds 40%. These statistics are significantly biased by three outlier models. MIT has large negative humidity there<sup>1</sup> and omitting it increases the multi-model mean to 21% and reduces the standard deviation to 24%. Two further models with greater than 50% relative humidity at 70 hPa contribute the majority of this residual standard deviation.

The regions of large humidity variation described above are associated with major differences in tropopause relative humidity structure in individual models (shown in the ATLAS). At one extreme, the LASG hygropause is virtually flat between 150–200 hPa at all latitudes, due to this model's very low vertical resolution (9 levels). In contrast, the ECMWF and MIT hygropause is above 100 hPa in the tropics but below 200 hPa in the extratropics. At the poles, AGU is relatively dry in the upper troposphere but much moister above this at 100 hPa.

The distribution of humidity variation among the APE models is generally similar to that for the AMIP annual average (shown in the ATLAS), taking account of the stronger Hadley circulation in APE and weaker tropical features in AMIP due to averaging over the annual cycle. Only two regions show significant differences. First, in the equatorial upper tropo-

<sup>1</sup>LASG is the only other model with negative specific humidity (at the tropical tropopause), although its relative humidity remains positive everywhere, suggesting interpolation problems from the 9 model levels.

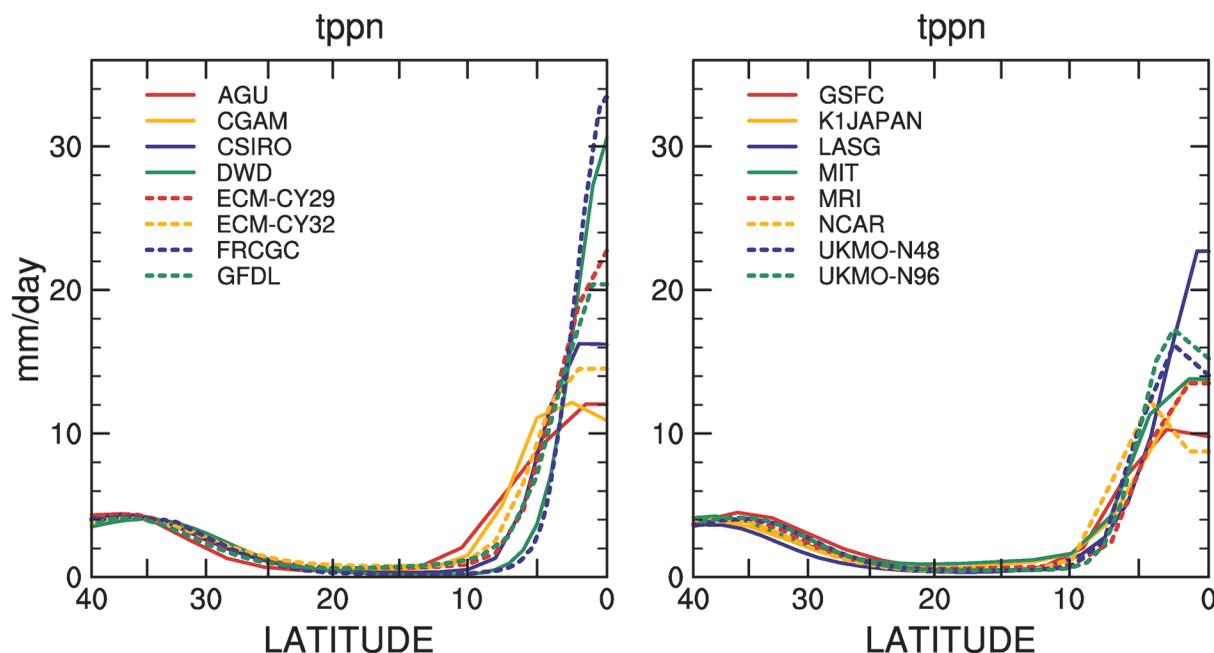


Fig. 4. Zonal-time average total precipitation (tppn) for individual models,  $\text{mm day}^{-1}$ . The 16 models are split between two panels for clarity.

sphere the region of large standard deviation extends further below the Hadley cell outflow in APE, despite the multi-model mean vertical profile being similar to AMIP. Second, the AMIP variation is much smaller in the polar lower stratosphere, presumably because the AMIP ensemble does not include such large outlier models, although circulation differences and the annual averaging may also contribute.

#### 5.4 Precipitation

The single versus double equatorial ITCZ structure in vertical velocity is reflected in the zonal average total precipitation shown in Fig. 4 for the individual models. The preferential formation of a single or double equatorial precipitation maximum has long been a feature of AGCM simulations of Earth climate and idealised aqua-planets, and it remains something of a modern modelling mystery. Due to feedbacks between low level winds and equatorial SSTs, an unrealistic double-ITCZ is a major systematic bias in many coupled ocean-atmosphere simulations (see e.g., Lin (2007) and the discussion in Blackburn and Hoskins (2013)). In Fig. 4 the total precipitation is plotted equatorward of  $40^\circ$  latitude and with the models split over two panels to better expose the tropical behaviours. The mid-latitude maximum occurring at around  $35^\circ$  latitude is associated with baroclinic waves. The curves in the

unplotted region poleward of  $40^\circ$  show little variation between them as can be verified in the ATLAS. It has been verified that averaging the two hemispheres does not misrepresent the ITCZ structures. All models with a double-ITCZ in the hemispheric average do contain a true double-ITCZ in the 3-year mean. Using daily data for 1 year, the ITCZ structures appear stable down to monthly timescales but relative north-south intensities do vary significantly on this timescale.

There is a large variation in the maximum ITCZ precipitation, from 10 to  $34 \text{ mm d}^{-1}$ . The CGAM, GSFC, NCAR and UKMO(N48,N96) models have clear double structures, where double is defined as the presence of a minimum equatorward of the maximum. The DWD, FRCGC and ECM-CY29 models have a single maximum at the grid point on the equator. This might be influenced by interpolation, since these three models submitted data interpolated to uniform latitude-longitude grids. For the ECMWF model, an equatorial maximum represents a maximum at the first Gaussian latitude on each side of the equator. The AGU, GFDL, LASG, K1JAPAN, MIT and MRI models have no equatorial grid point, but form a single structure across the two points spanning the equator: thus the curves are flat approaching the equator and represent a single structure. The CSIRO and ECM-CY32 models form a broader single structure across the equator and

two adjacent points (one in each hemisphere), i.e. the equatorial value is not significantly different from the adjacent points. Again, this might be influenced by interpolation.

Figure 5 shows the tropical-average total precipitation for the individual models, obtained by zonally and temporally averaging between 20°S and 20°N, and the contributions from convective and large-scale processes. The tropical-average total precipitation is around 4.5 mm d<sup>-1</sup>, which corresponds approximately to a radiative convective equilibrium averaged over the entire tropics<sup>2</sup>. The range among the models, 1.2 mm d<sup>-1</sup> or 25–30% of the mean, is modest but significantly larger than the range for the global average shown below in Section 7. This is due to differences among the models in several processes: radiative cooling rates, the latitudinal extent of the Hadley circulation descent region that is in exact mass balance with the ITCZ ascent and transient eddy poleward moisture flux in the sub-tropics. Convective precipitation dominates the total in all models, but the variation in large-scale precipitation fraction is large, ranging from negligible in the UK Met Office family of models to almost 50% in K1JAPAN. FRCGC, with a horizontal resolution of 7 km, does not include a convective parameterization.

The modest variation in tropical-average total precipitation also contrasts with the large variation in precipitation intensity at the equator, indicating compensating effects of ITCZ intensity and width. This is evident in the latitudinal profiles in Fig. 4, where precipitation curves for individual models cross close to 4° latitude, giving minimum spread there. Therefore, the large variation in ITCZ structure and intensity among the models occurs despite a relatively strong budget constraint on tropical-average precipitation.

## 6. Maintenance of zonal mean state

This section considers the processes that maintain the zonal mean state, focusing on the transient eddy dynamical fluxes and components of the parameterization forcing. The primary interest is not the detailed budgets, which must balance closely for the 3 year climate averages in each model, but rather the spatial structures of the dominant processes and their variation among the models. The relevant diagnostics are available for the majority of models, apart from the parameterization forcing from individual processes. Details are given below.

<sup>2</sup>Latent heating from 4 mm d<sup>-1</sup> balances a radiative flux divergence from the atmospheric column of around 120 W m<sup>-2</sup>, equivalent to a cooling of 1 K d<sup>-1</sup> throughout the depth.

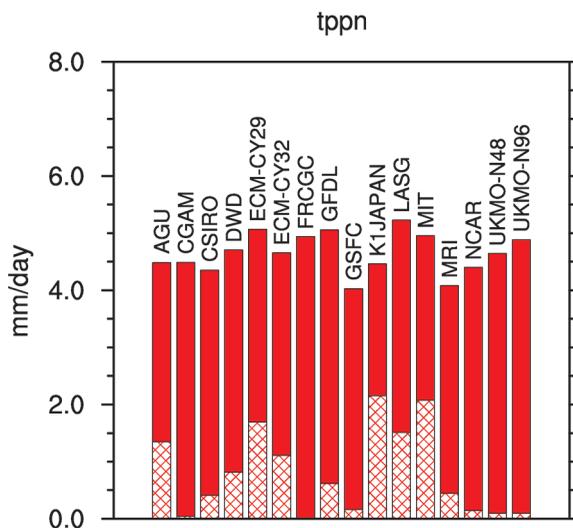


Fig. 5. Tropical average from 20°S to 20°N, time average total precipitation (tppn), divided into convective precipitation (cppn, solid) and large-scale precipitation (dppn, cross-hatched), mm day<sup>-1</sup>.

### 6.1 Transient eddies

The zonally averaged covariance of the quantities  $\alpha$  and  $\beta$  is partitioned into stationary zonal mean (SM), transient zonal mean (TM), stationary eddy (SE) and transient eddy (TE) components, respectively, in

$$\overline{[\alpha\beta]} = \overline{[\alpha][\beta]} + \overline{[\alpha']'[\beta]'} + \overline{[\alpha^* \beta^*]} + \overline{[\alpha'^* \beta'^*]}$$

where the overbar ( $\overline{\quad}$ ) denotes the time average and prime ( $'$ ) the deviation from it, square brackets [ $\quad$ ] denote the zonal average and star ( $*$ ) the deviation from it.

SM quantifies the zonal-time average state and the variances are simply the squares of variables in Fig. 3. TM quantifies time variations of the zonal average state and is small for many of the variables discussed below. For Earth both stationary eddy (SE) and transient eddy (TE) fluxes contribute to the mean state budgets. For the APE CONTROL, forced by zonally symmetric SST, SE contributions are small, so the zonally symmetric overturning and transient eddies dominate the transports and budgets. The Hadley circulation has already been shown to be stronger than for Earth, consistent with the atmosphere contributing a greater fraction of total poleward energy transport in the tropics for APE than for Earth (Fig. 2). In the extratropics the APE and Earth atmospheric transports are more comparable, so the APE TE fluxes are expected to be of similar magnitude to the total eddy (SE + TE) fluxes for Earth.

The TE fluxes have not been time filtered, so synoptic timescale eddies and lower frequencies contribute to the diagnostics in this section. Wavenumber-frequency analysis of 250 hPa meridional wind at 30° latitude, shown in the ATLAS, confirms that a wide range of timescales is present in the storm-tracks in all the models. The 3 year samples do contain a zonal wavenumber 5 structure that has non-zero SE transport, but the noise of a single 3 year sample is too large to allow systematic comparison of the models. This feature is discussed separately in Section 9.

Figure 6 shows the multi-model mean and the inter-model standard deviation, to summarize the variability among the models, for selected TE covariance quantities. The variances  $[(u')^2]$ ,  $[(v')^2]$ ,  $[(\omega')^2]$ , show the location and magnitude of transient eddy activity. The covariances  $[u'v']$ ,  $[v'T']$  and  $[v'q']$  show the TE poleward fluxes of westerly momentum, temperature and moisture. The SM, TM and TE components of these and additional quantities, including vertical fluxes, are shown in the ATLAS. Table 5 lists the mid-latitude maximum value of each quantity for the eleven (out of sixteen) models that submitted covariance data.

The multi-model mean TE covariances in Fig. 6 are mostly dominated by transient activity in the mid-latitude storm-tracks and resemble those for Earth (Källberg et al. 2005). The kinetic energy (KE) components  $[(u')^2]$  and  $[(v')^2]$  peak in the upper-troposphere while  $[(\omega')^2]$  peaks in mid-troposphere directly below  $[(u')^2]$ . The magnitudes of all three quantities vary greatly among the models, with standard deviations generally 20–25% of the multi-model mean. The maximum values, shown in Table 5, vary by roughly a factor of two, with standard deviations also in the range 20–25%. The GSFC values are generally the smallest, indicating least transient activity, while K1JAPAN is among the largest, with particularly large transient KE. Values do not vary systematically with horizontal resolution, implying that factors such as numerical damping and parameterization characteristics help to determine eddy amplitudes. The ratio  $[(v')^2]/[(u')^2]$  in Table 5 is greater than unity in all the models except DWD, but varies widely between 1.10 and 1.52. This ratio appears (indirectly) in the E-vector,  $\mathbf{E} = (v'^2 - u'^2, -u'v')$ , defined by Hoskins et al. (1983), so its large range implies that the horizontal anisotropy and zonal group propagation characteristics of transient eddies vary widely among the models. This could be associated either with high frequency transients alone or with differences in the

residual of opposing high and low frequency contributions to the zonal E-vector component, which is a feature of observations found by Hoskins et al. (1983).

The poleward flux patterns for the multi-model mean in Fig. 6 are again similar to those for Earth. Now the corresponding standard deviations (SDs) are significantly different from the means. For  $[u'v']$  the mean and SD are both largest in the upper troposphere, but maximum SD coincides with the maximum gradient between the poleward and equatorward flux maxima. For  $[v'T']$  the SD maximises on both sides of the upper tropospheric peak in the mean. This quadrature relationship between mean and SD reflects variations in both the magnitude and location of these particular TE flux structures among the models. For momentum flux, the standard deviation is again around 20% of the multi-model mean in Fig. 6, and also in Table 5 for the poleward and equatorward flux maxima. The poleward flux maximum at 30° latitude is stronger than the equatorward flux maximum at 50° in all models, but the ratio varies widely, between 1.2 and 2.2 in all but the MIT model, for which it is 4.0. This indicates large differences in meridional propagation of TE wave activity in the storm-tracks, in addition to the zonal variations discussed above.

The variation among the models is considerably smaller for the lower tropospheric temperature and moisture fluxes, with the SD of the maximum values in Table 5 being only 10% of the mean for  $[v'T']$  and 13% for  $[v'q']$ . This points to stronger budget constraints on these quantities than for momentum. There is scope for further analysis of the fluxes and vertically integrated budget constraints.

A number of the TE covariances additionally show tropical structure. This is generally of smaller amplitude than in mid-latitudes, with the notable exception of vertical velocity variance, for which the multi-model mean in Fig. 6 is similar in mid-latitudes and at the equator. Moreover, the standard deviation is significantly larger than the mean at the equator, whereas it is only 20–25% of the mean in mid-latitudes. The large differences in mean ITCZ intensity and structure among the models, discussed previously, are therefore accompanied by large differences in the transient zonal asymmetries.

The horizontal wind variances reveal that the differences in vertical velocity variance are associated with winds in the equatorial upper troposphere, at the same level as the mean Hadley cell outflow, 200 hPa. The variability is dominated by east-west winds, since  $[(u')^2]$  exceeds  $[(v')^2]$  by a factor of around four (note the different colour scales in Fig. 6). This

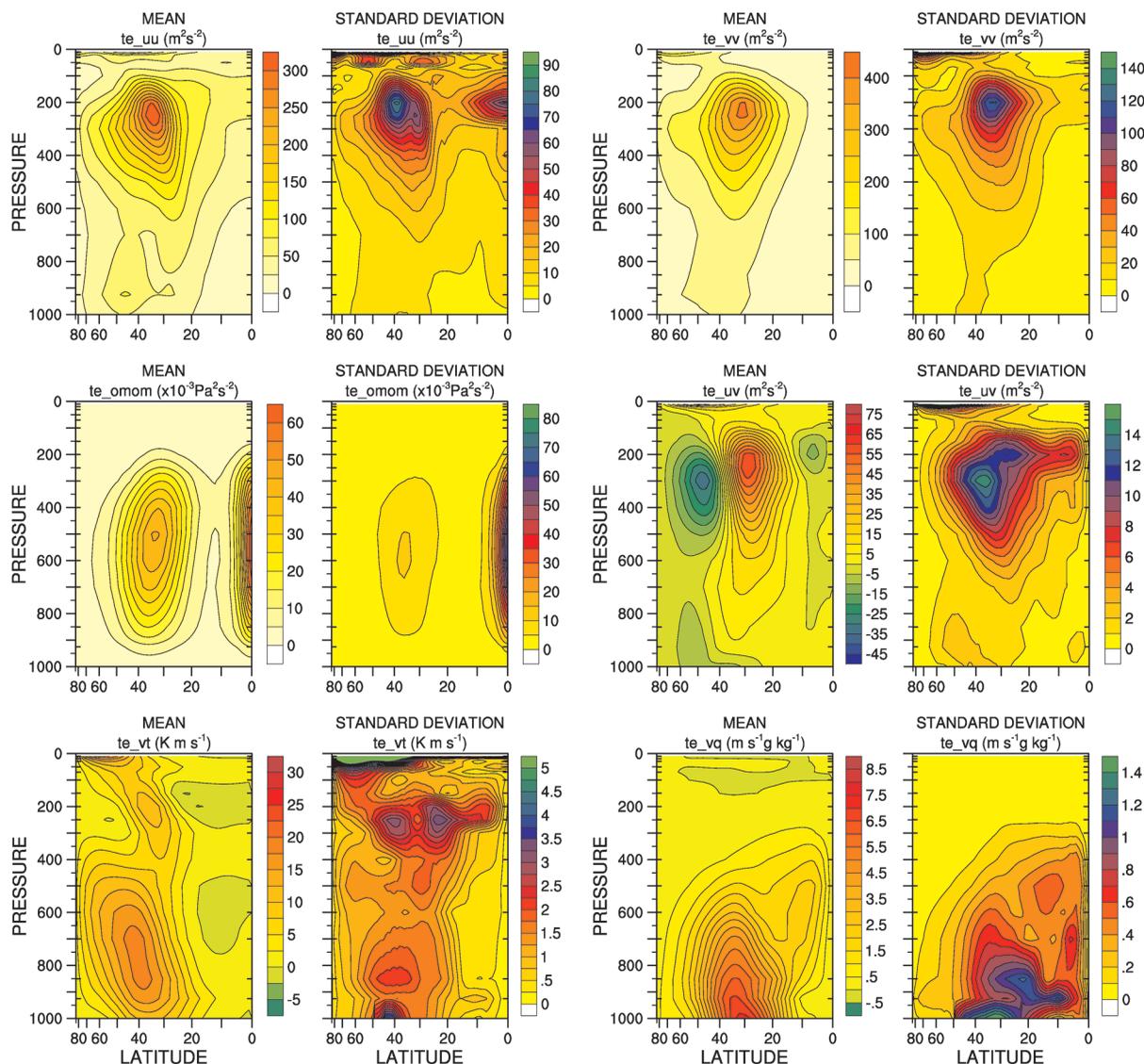


Fig. 6. Multi-model mean and standard deviation of transient eddy fluxes,  $[\overline{(u'*)^2}]$  (te\_uu),  $\text{m}^2 \text{s}^{-2}$ ;  $[\overline{(v'*)^2}]$  (te\_vv),  $\text{m}^2 \text{s}^{-2}$ ;  $[\overline{(\omega'*)^2}]$  (te\_omom),  $\times 10^{-3} \text{Pa}^2 \text{s}^{-2}$ ;  $[\overline{u' * v' *}]$  (te\_uv),  $\text{m}^2 \text{s}^{-2}$ ;  $[\overline{v' * T' *}]$  (te\_vt),  $\text{K m s}^{-1}$  and  $[\overline{v' * q' *}]$  (te\_vq),  $\text{m s}^{-1} \text{g kg}^{-1}$ .

suggests significant differences in the equatorial wave spectra, principally in low zonal wavenumbers and/or wave types for which zonal flow dominates. Although LASG has significantly higher tropical variance than other models, it does not dominate the inter-model standard deviations.

The covariances  $[\overline{u' * v' *}]$  and  $[\overline{v' * T' *}]$  also show significant variation about the multi-model mean in the tropical upper troposphere. The momentum flux is downgradient, with equatorward flux of westerly

momentum towards the peak easterly wind at the equator. The mean temperature flux is small but also equatorward. In addition there is a poleward moisture flux maximum near  $10^\circ$  latitude throughout the mid-troposphere in the multi-model mean, with weak equatorward flux near the surface. The variation among the models is more confined to the lower troposphere and peaks near  $5^\circ$  latitude, presumably associated with the differences in zonal-time average ITCZ structure. These fluxes may be evidence of equatorial

Table 5. Maximum Transient Eddy statistics

model	te_uu	te_vv	te_omom	te_vt <sup>1</sup>	te_uv	te_vq	
units	m <sup>2</sup> s <sup>-2</sup>	m <sup>2</sup> s <sup>-2</sup>	× 10 <sup>-3</sup> Pa <sup>2</sup> s <sup>-2</sup>	K m s <sup>-1</sup>	m <sup>2</sup> s <sup>-2</sup>	m s <sup>-1</sup> g kg <sup>-1</sup>	
AGU	330	451	48	19	78	-39	8.0
CSIRO	309	438	41	20	59	-33	8.8
DWD	416	357	54	19	50	-31	7.3
ECM-CY29	404	444	56		65	-42	7.6
ECM-CY32	417	459	57	19	65	-44	7.3
GFDL	274	330	36	16	57	-41	6.7
GSFC	281	302	24	15	37	-31	5.6
KIJAPAN	453	690	51	19	73	-37	8.1
LASG	289	382	48	21	76	-34	6.9
MIT	264	370	53	21	61	-15	8.8
NCAR	370	544	40	20	63	-44	8.5

<sup>1</sup>Transient eddy  $\overline{[v'^*T'^*]}$  was not available for ECM-CY29.

wave activity in the presence of mean state gradients, not part of the classical theory by Matsuno and Gill which requires a resting basic state. The equatorial wave spectra are discussed in more detail in section 10.1. The transient mean covariances (TM, shown in the ATLAS) reveal significant temporal variation of the zonally averaged tropical moisture and Hadley circulation, and that these contribute to the total moisture fluxes.

### 6.2 Parameterized thermodynamic forcing

Parameterization forcing diagnostics were not submitted for all the models and, of those that were, not all included the complete set of requested variables. For this reason and to avoid loss of vertical structures by interpolation to a standard grid, a multi-model mean has not been computed. Instead, selected variables are shown for individual models on the model vertical grid. Only heating terms that are well defined are presented here. Forcing by moist processes is partitioned differently between models and so individual parameterizations cannot be sensibly compared. More complete forcing diagnostics are included in the ATLAS.

Figure 7 shows zonal and time averaged total parameterized heating in the simulations. The largest inter-model variation is in the tropics. Heating variations are much smaller in the mid-latitude storm-track region, except for AGU with about 25% stronger heating and GSFC with double the heating rate.

Variations in equatorial heating are associated with the convective and cloud parameterizations (shown in the ATLAS) and for individual models are consistent with the variations in ITCZ vertical motion and precip-

itation seen earlier. Most of the models have minimum heating near 600–750 hPa, perhaps associated with evaporation below the freezing level or with outflow from convection that reaches mid-troposphere (representing cumulus congestus). Each model's heating minimum is below its mid-level minimum in vertical motion (cf. Fig. 3) and closer to the level of secondary poleward mean flow in the low/mid troposphere, to the extent that these dynamical structures are resolved by the standard isobaric level data.

In the sub-tropical subsidence region, large variations in the rate of cooling at the boundary layer inversion are dominated by longwave radiative cooling, shown in Fig. 8, partly offset by qualitatively similar but weaker variations in shortwave heating (shown in the ATLAS). This large variation in radiative cooling is associated with tropical low level clouds, which have been identified as important for estimates of climate sensitivity. Medeiros and Stevens (2011) argue that in aqua-planet simulations these are shallow cumulus clouds, not stratocumulus, and that they occur in similar dynamical environments to those in Earth simulations and in observations. Therefore, the shallow cumulus regime in the APE simulations can be compared with observations, just as well as the earth-like configurations can be.

Above the sub-tropical inversion, the parameterization cooling in Fig. 7 is in better agreement between the models. However, its latitudinal structure is non-trivially determined, since longwave cooling (Fig. 8) peaks further equatorward while shortwave heating minimises further poleward. The total parameterized cooling does not extend as far poleward as the mean

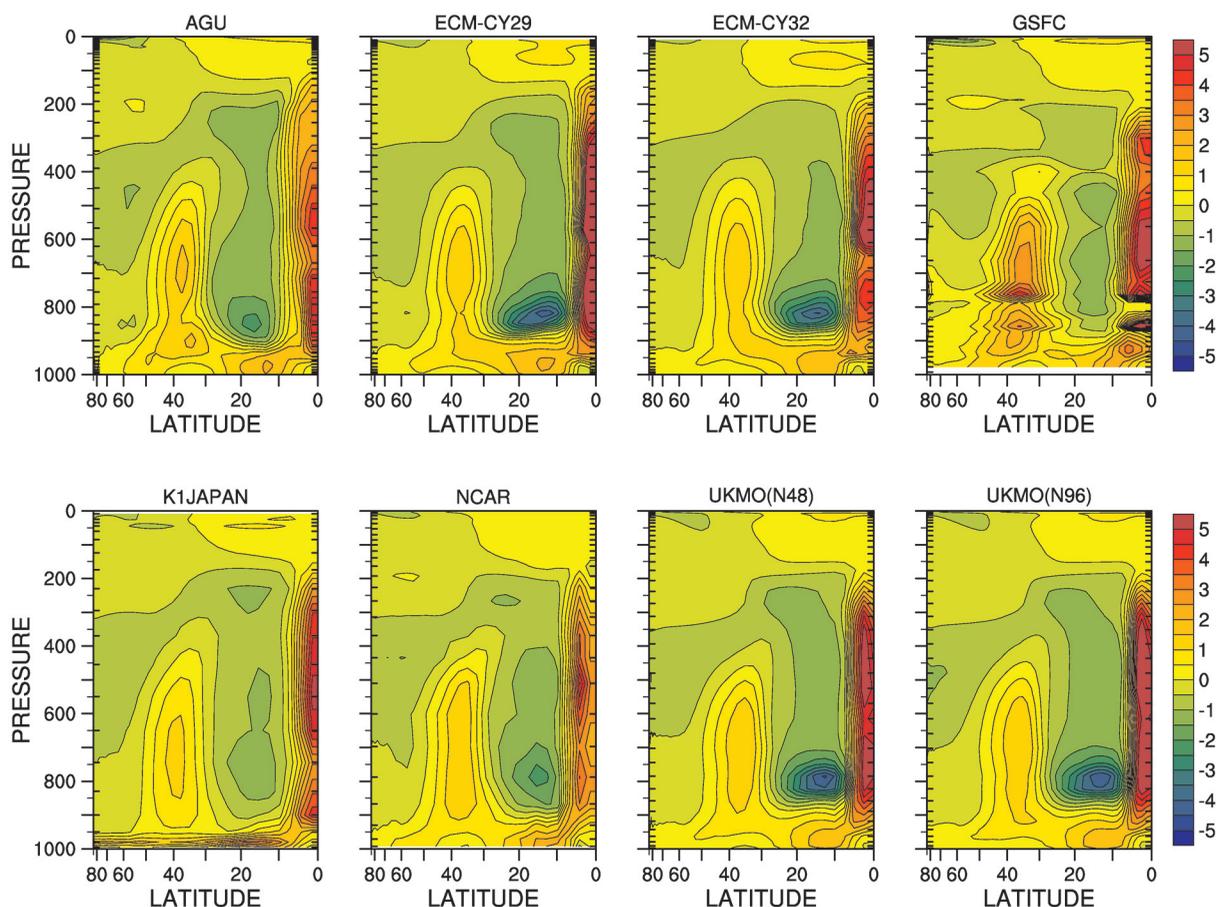


Fig. 7. Total parameterized temperature tendency,  $\text{K day}^{-1}$ , for individual models. Inner tick marks on each panel show the model's vertical grid.

descent or, to a lesser extent, the minimum in relative humidity in Fig. 3. This implies that cooling due to divergence of the transient eddy poleward temperature flux (Fig. 6) makes an increasing contribution to thermodynamic balance with latitude in the sub-tropics.

Two models include heating structures that differ significantly from the other models. In the boundary layer, parameterization heating is stronger and shallower in K1JAPAN than in the other models shown in Fig. 7. Turbulence is the only contributing process in K1JAPAN, while convection deepens the boundary layer heating in the other models (individual process contributions are shown in the ATLAS). GSFC exhibits a large amplitude grid-scale structure in the lower tropospheric heating profile in Fig. 7. This is due to the convection and cloud parameterizations, partly offset by similar structures in longwave radiation and turbulent mixing. Sub-grid turbulence is also active above the equatorial boundary layer in AGU, opposing grid-

scale heating structures due to convection and cloud.

All models except LASG show longwave heating at the tropical tropopause in Fig. 8. LASG has very coarse vertical resolution in this region and probably cannot resolve the processes involved. The heating is due to absorption of upwelling radiation from the warmer troposphere, so is stronger in the sub-tropics, where clear skies and low tropospheric humidity lead to a larger upwelling longwave flux than from colder deep convective cloud tops at the equator. The heating varies significantly among the other models, despite both ozone and carbon dioxide concentrations being prescribed in the APE protocol. Thuburn and Craig (2002) found that, although longwave heating at the tropopause cold point is dominated by ozone absorption,  $\text{CO}_2$  dominates the sensitivity of heating with respect to temperature in the tropical substratosphere, between the convection top and cold point. Whichever gas dominates the tropopause heating variation in the

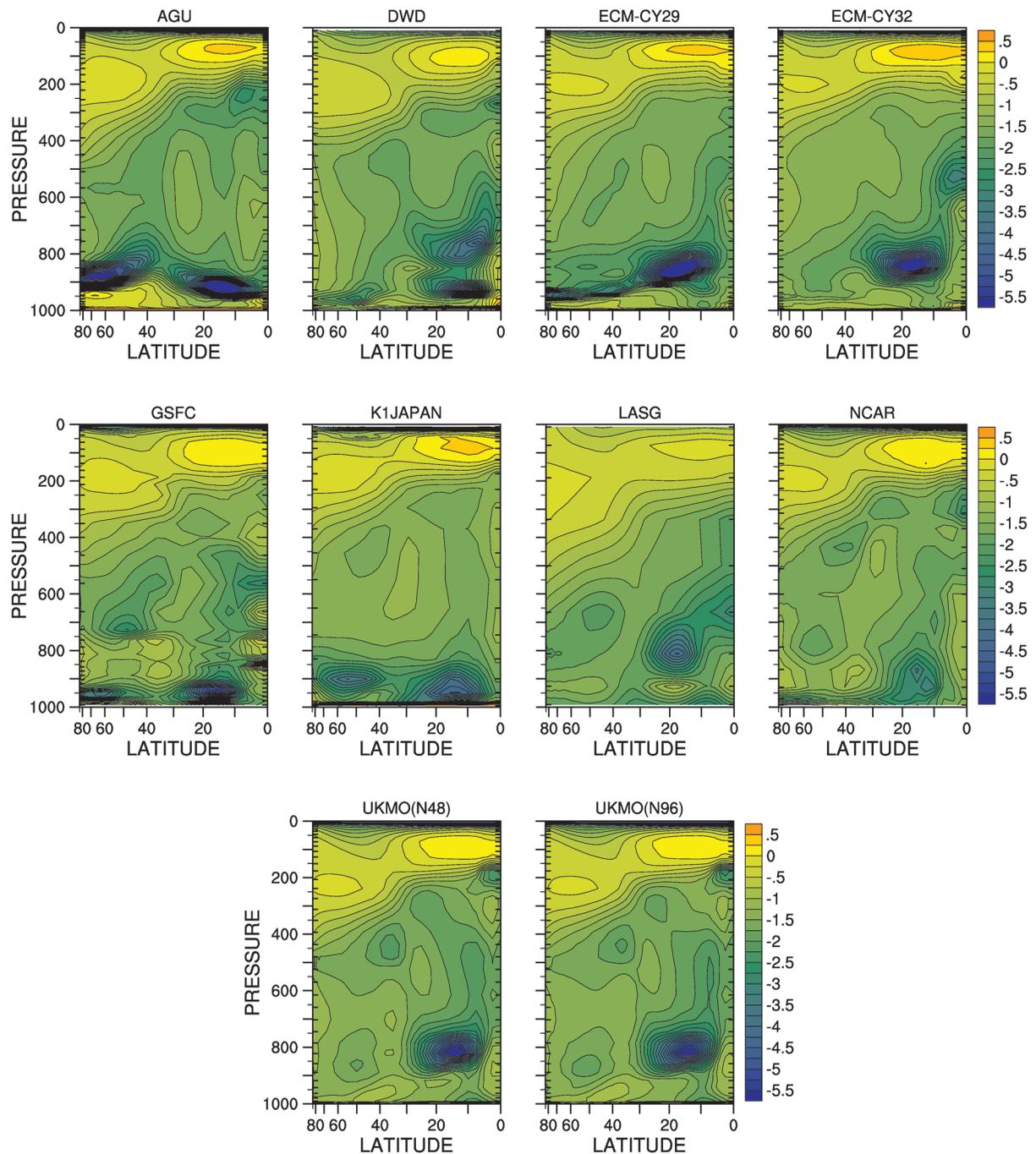


Fig. 8. Longwave radiation temperature tendency,  $\text{K day}^{-1}$ , for individual models.

APE models, Fig. 9 shows that, to first order, the longwave heating variations damp differences in temperature between the models. Shortwave variations (shown in the ATLAS) are small, so the longwave variations appear in the total parameterized heating (Fig. 7). This

implies that the variations in tropopause temperature structure are dynamically maintained, by differing numerical computations of adiabatic cooling associated with ascent. There is no obvious association of heating with vertical resolution in Fig. 9, but this inter-

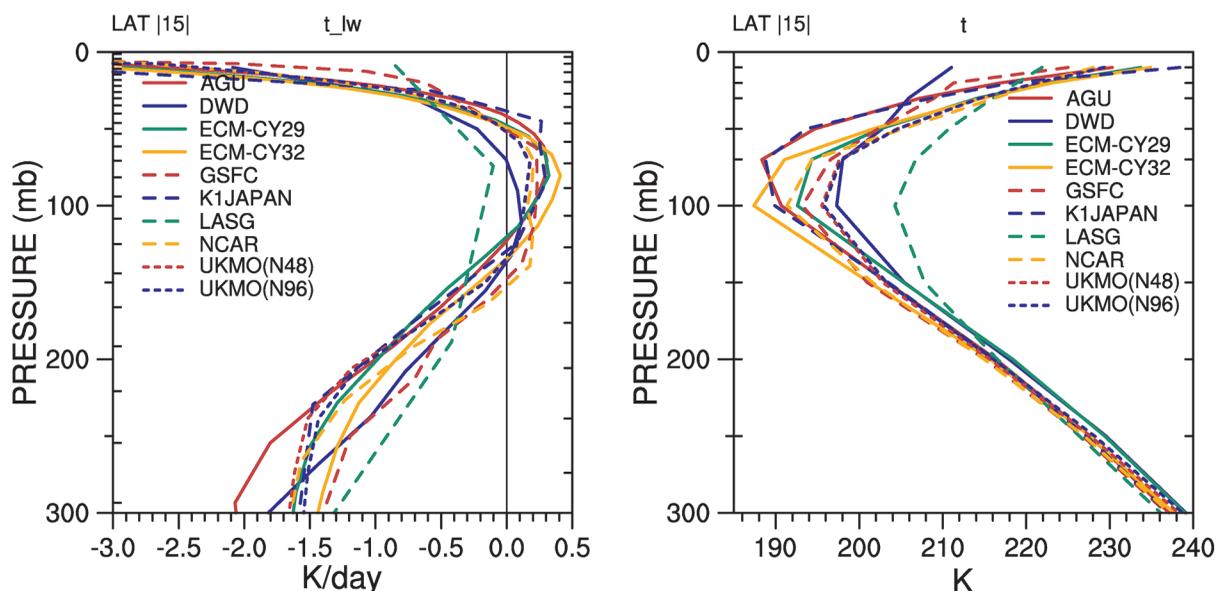


Fig. 9. Vertical profiles above 300 hPa of longwave radiation temperature tendency (left),  $\text{K day}^{-1}$ , and temperature (right), K, at each model's grid point closest to  $15^\circ$  latitude.

pretation is consistent with the finding by Williamson et al. (1998) that tropical tropopause temperature and structure are highly sensitive to dynamical formulation at vertical resolutions typical of those in the APE models.

## 7. Global average budgets

### 7.1 Water

Figure 10 compares the global average water budget displayed as bar plots for each model, calculated from the daily average, global average time series. The multi-model mean precipitation is almost  $3 \text{ mm d}^{-1}$ , very close to observational estimates for Earth. There is a 20% variation in total precipitation among the models, smaller than the 25–30% range of tropical averages seen in Fig. 5. A much larger variation is evident in the partition between convective and large scale stable precipitation. Neglecting FRCGC, which has no convective parameterization, convective precipitation varies by a factor greater than two and large scale varies by a factor of three. The temporal variability of global average precipitation is shown by the whiskers in Fig. 10, which range from plus one to minus one standard deviation of the time series of daily average values. The standard deviation is around 4–5% in most models, with the exception of DWD which shows very little day-to-day variation in all global averages, most likely due to a loss of precision in the output data.

Most models conserve water over the three year integration period, with evaporation minus precipitation being zero or nearly so. Exceptions are DWD, ECMWF and LASG, whose water budget residuals are approximately 1.3%, 2% and 4% respectively of global precipitation or evaporation. The main contributor to the ECMWF residuals is believed to be interpolation to a regular latitude-longitude grid before global integration, since climate simulations of the ECMWF model produce a residual of about  $-0.4\%$  on the model grid, much closer to those of the remaining APE models. A loss of numerical precision in the post-processing is believed to be the main contributor to the DWD residual. Advection and adjustment of negative vapour account for the LASG residual, the largest at  $\sim 4\%$  of global precipitation. The 30 day period for FRCGC is too short to rule out a change in the water vapor content of the atmosphere dominating the average. GFDL, GSFC, MIT and NCAR are the closest to zero. We note that NCAR has a fixer to ensure water is conserved. The remaining models show a minimal residual value.

### 7.2 Energy

Prescribing both insolation and SST in APE leads to an unknown global imbalance in TOA and surface net energy fluxes, even for models optimised for simulation of Earth climate. Atmospheric energy storage is negligible in multi-annual averages, so the imbalance

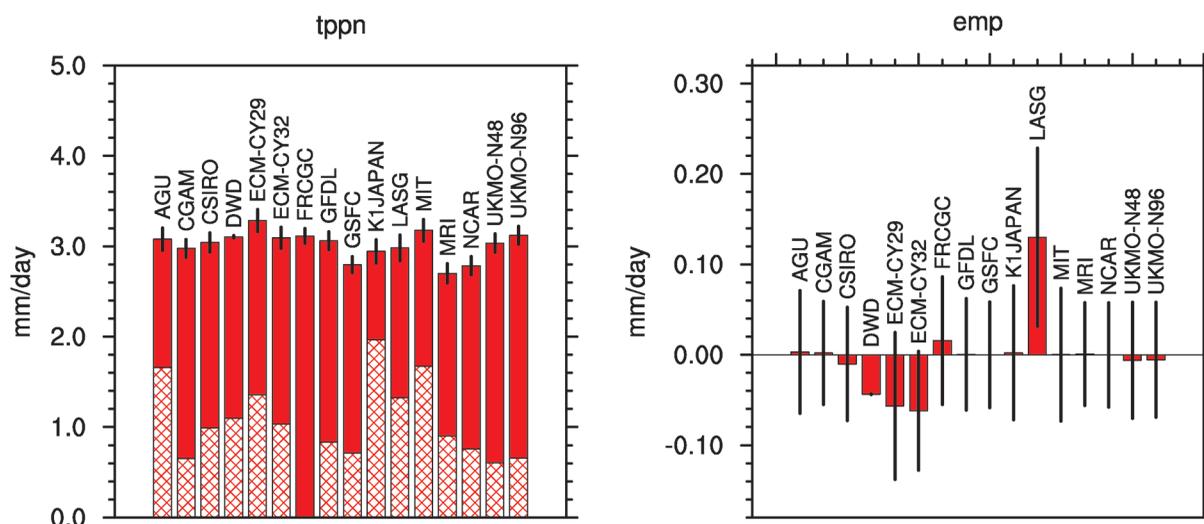


Fig. 10. Global-time average total precipitation (tppn) and evaporation minus precipitation (emp). The total precipitation is divided into convective precipitation (cppn, solid) and large-scale precipitation (dppn, cross-hatched),  $\text{mm day}^{-1}$ .

would heat or cool an underlying ocean. On Earth the present-day global imbalance is estimated to be  $0.9 \text{ W m}^{-2}$  downwards, causing global warming (Hansen et al. 2005; Trenberth et al. 2009), but it may be much larger for each of the aqua-planet SST distributions.

The TOA and surface residual energy fluxes are shown for each model in Fig. 11, together with the cloud fraction and albedo. For the multi-model mean, the TOA residual flux is around  $13 \text{ W m}^{-2}$  downwards, going into the ocean. The range across the models is surprisingly large, greater than  $30 \text{ W m}^{-2}$ , even omitting FRCGC which is not in equilibrium, and despite the simplified aqua-planet configuration which might be expected to constrain differences between the models. Several factors may contribute to the large range.

First, experimental and numerical artifacts are relatively minor contributors. Most of the models accurately implemented the prescribed global insolation, to within  $0.1 \text{ W m}^{-2}$ , but a small number did not, either by prescribing a different value of the solar constant or by using perpetual vernal equinox for Earth's elliptical orbit<sup>3</sup>. The largest discrepancy is less than  $4 \text{ W m}^{-2}$  in global insolation, contributing only around  $3$

$\text{W m}^{-2}$  to net downward shortwave flux after allowing for global albedo. Lack of formal energy conservation by the model atmospheres leads to a range of atmospheric energy gain/loss of  $5 \text{ W m}^{-2}$ , calculated as the difference between the TOA and surface residual fluxes. This contributes little to the range of residual flux, which differs by only  $1.5 \text{ W m}^{-2}$  at the TOA and surface, although it contributes more to differences between individual models.

Second, only a subset of the models listed in Table 2 has been optimised to give global energy balance for present day Earth climate. The range of TOA residual flux is only moderately reduced for the optimised models but, omitting DWD which was optimised for short NWP timescales, the range is approximately halved to around  $15 \text{ W m}^{-2}$ . However, the contribution of this effect is uncertain, because the global imbalance for Earth climate in the non-optimised models is unknown.

Third, optimisation of TOA flux for Earth climate may compensate differing systematic biases over land and ocean. Removing land in the aqua-planet then reveals the uncompensated bias over ocean, which will differ between models. To quantify this effect, it would be necessary to compute fluxes averaged separately over land and ocean in AMIP simulations for each of the models, and compare with the global residual fluxes in APE.

Fourth, the APE CONTROL may be considered an extreme SST sensitivity experiment, with earth-like

<sup>3</sup>We strongly recommend that future APE simulations use a circular orbit with zero obliquity. The Earth-Sun distance at vernal equinox is less than the annual average, increasing solar irradiance relative to the prescribed "solar constant". It is also difficult to identify the precise point in the calendar at which insolation is symmetric about the equator.

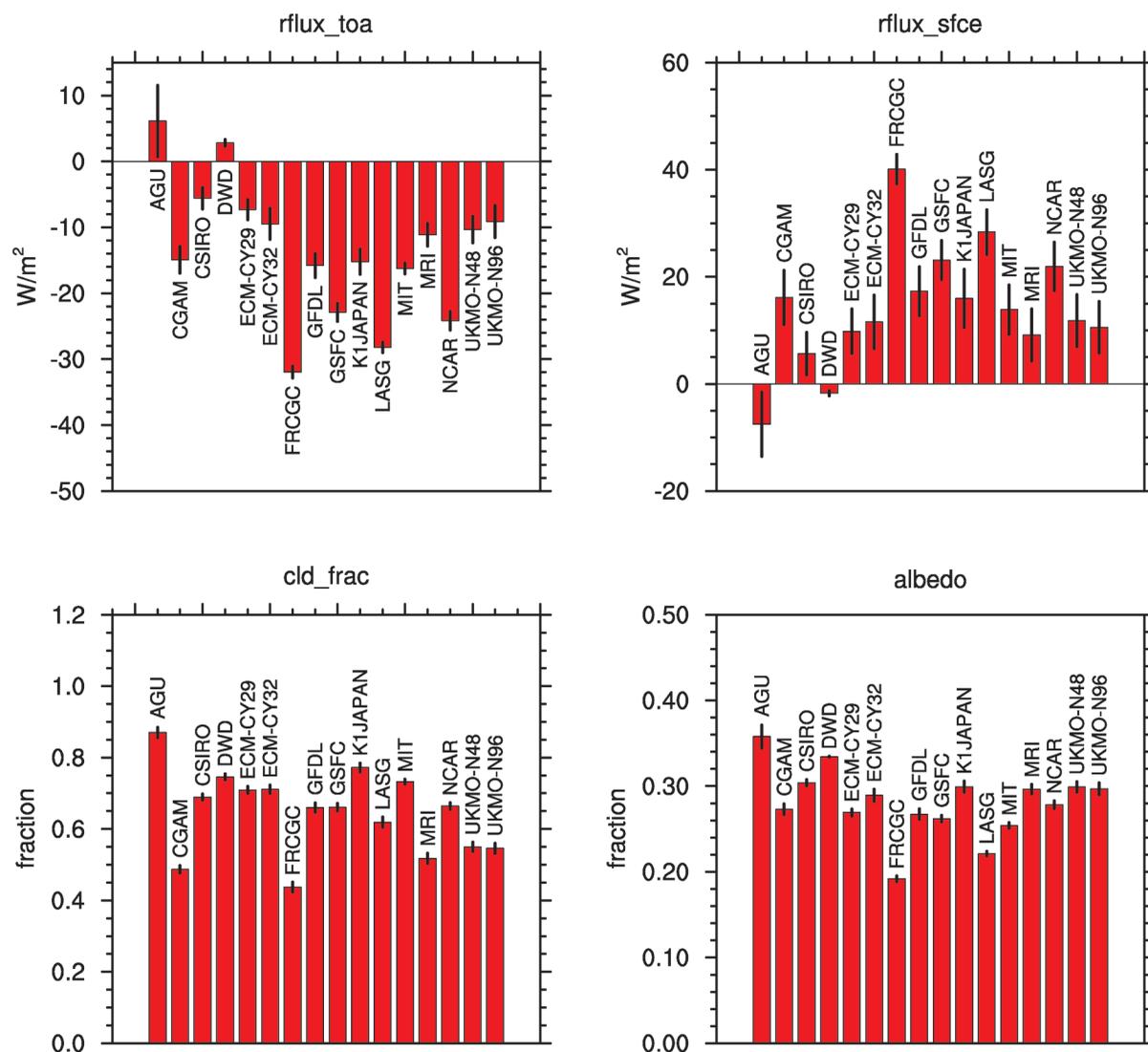


Fig. 11. Global-time average residual TOA radiative flux (rflux\_toa, positive upwards) and residual surface flux (rflux\_sfce, positive downwards)  $\text{W m}^{-2}$ , cloud fraction (cld\_frac) and albedo.

SSTs and land surface temperatures replaced by the specified aqua-planet values. This is analogous to the Cess type experiment (Cess et al. 1989, 1990). If the APE CONTROL SST is eliciting a range of sensitivities in the global energy balance, we would expect the global imbalance to vary significantly for the gross variation of SST from the PEAKED through CONTROL and QOBS to FLAT aqua-planet prescriptions. This is certainly not the case for the multi-model mean (shown in the ATLAS), for which TOA residual flux varies by only around  $2 \text{ W m}^{-2}$ . For individual models, over half approximately follow this mean vari-

ation, though with varying magnitudes, while a smaller number show a reversed or less monotonic trend. The maximum variation with SST is  $15 \text{ W m}^{-2}$  (for AGU) but it is more typically around  $5 \text{ W m}^{-2}$ , including the subset of models whose TOA balance has been explicitly optimised for Earth climate. The inter-model range of residual flux is therefore rather insensitive to SST. It is slightly larger for the PEAKED and FLAT cases, whose tropical climates are least similar to Earth, as discussed in Williamson et al. (2013).

The relative insensitivity of TOA global imbalance to aqua-planet SST suggests that the large range among

the models is mainly due to a lack of optimisation for Earth climate in the majority of the models and/or removal of differing biases in the TOA energy balance over land in the optimised models. A combination of APE and AMIP simulations by individual models would be required to quantify these effects and it is hoped that the inclusion of both AMIP and APE experiments in the protocol for the fifth phase of the Coupled Model Intercomparison Project (CMIP5, Taylor et al. 2009) will facilitate this.

Global cloud fraction and albedo<sup>4</sup>, also shown in Fig. 11, reveal that the variations in global energy balance among the models are dominated by short wave reflection by clouds. Variations in albedo correlate well with the TOA residual flux, and correspond to a range of  $46 \text{ W m}^{-2}$  in reflected (and net) shortwave flux. This is more than double the  $20 \text{ W m}^{-2}$  range in outgoing longwave radiation (OLR, shown in the ATLAS), which is uncorrelated with the residual flux. However, variations in OLR tend to oppose those in the shortwave, leading to the smaller range in residual flux noted earlier, around  $30 \text{ W m}^{-2}$ , compared to that in net shortwave.

Global cloud fraction varies by a factor of almost two, from a particularly high value in AGU (0.87) to less than 0.55 in a cluster of four models (MRI and the 3 UK Met Office models, CGAM and UKMO, again neglecting FRCGC). The latter group all have close to average albedo, so there is evidently a compensation between cloud fraction and cloud optical properties in this group relative to the remaining models.

Figure 12 shows the latitudinal variation of zonal average cloud fraction and albedo from the individual models. Cloud fraction varies by more than a factor of two in both the sub-tropical and polar regions and is one of the most variable of all the metrics studied. In the sub-tropics, most of the models do cluster more closely, with 3 outliers having cloud fraction below 0.4 (FRCGC plus two from the group with low global cloud fraction), plus two in which the equatorial maximum extends further into the sub-tropics. There is less general agreement at the poles. As was the case for the global averages, albedo is in somewhat better agreement than cloud fraction, because of compensating cloud optical properties in a number of outlying models. The albedos in many models cluster within 0.1 of each other over most latitudes, but the range increases to more than 0.3 at the equator. This exceeds the fractional spread in equatorial OLR, so

shortwave variations dominate the energy balance at TOA and also at the surface there, driving variations in poleward energy transport in the tropics, as discussed in Section 4.

Cloud radiative forcing (CRF) diagnostics were not collected for the APE models, so it is not possible to formally attribute the differences in the TOA radiative balance among the APE models to cloud versus clear sky effects. However, previous studies have shown that the radiative response of clouds, and in particular sub-tropical boundary layer clouds, is a major source of uncertainty in estimates of climate sensitivity in models (Bony and Dufresne 2005; Webb et al. 2006). Low-level clouds are expected to have a large radiative impact because their shortwave CRF dominates, whereas shortwave and longwave CRF more strongly oppose each other for deep tropical cloud. Variations in sub-tropical cloud fraction and optical properties among the APE models might therefore be expected to have the largest impact on the global energy balance. However this is not the case, since the TOA residual flux (shown in the Atlas) varies more in the ITCZ than in the sub-tropics. This large equatorial variation is due to the combination of the large range of albedo and the dominance of shortwave over longwave effects of deep cloud in the equatorial ITCZ noted above.

## 8. Kinetic energy spectra

Figure 13 shows the kinetic energy (KE) spectra of the models at 250 hPa, split over four panels for clarity. The figure includes the total kinetic energy (solid lines) and the divergent component (dashed lines), as a function of 2-dimensional total spherical wavenumber  $n$ . The zonal average flow, i.e., zonal wavenumber  $m = 0$ , is included in the KE for all  $n$ . Energy at a given total wavenumber  $n$  is associated with the spherical harmonics of vorticity and divergence at that wavenumber.

The KE spectra provide some indication of the damping from the dynamical core. The damping is due to explicit horizontal diffusion terms added to the dynamical equations, is inherent in the numerical approximations themselves, or is a combination of the two. In contrast to Earth simulations, which include large amplitude planetary waves forced by land-sea contrast and orography, the idealised APE configuration is dominated by a zonal average flow and transient eddies. It is hoped that this simpler circulation will help to isolate differences between models in the behaviour of the internal energy cascade process.

The KE spectra in Fig. 13 are in good agreement among the models for wavenumbers up to  $n \sim 15\text{--}20$

<sup>4</sup>Global albedo is defined here as the ratio of the upward and downward global average shortwave fluxes.

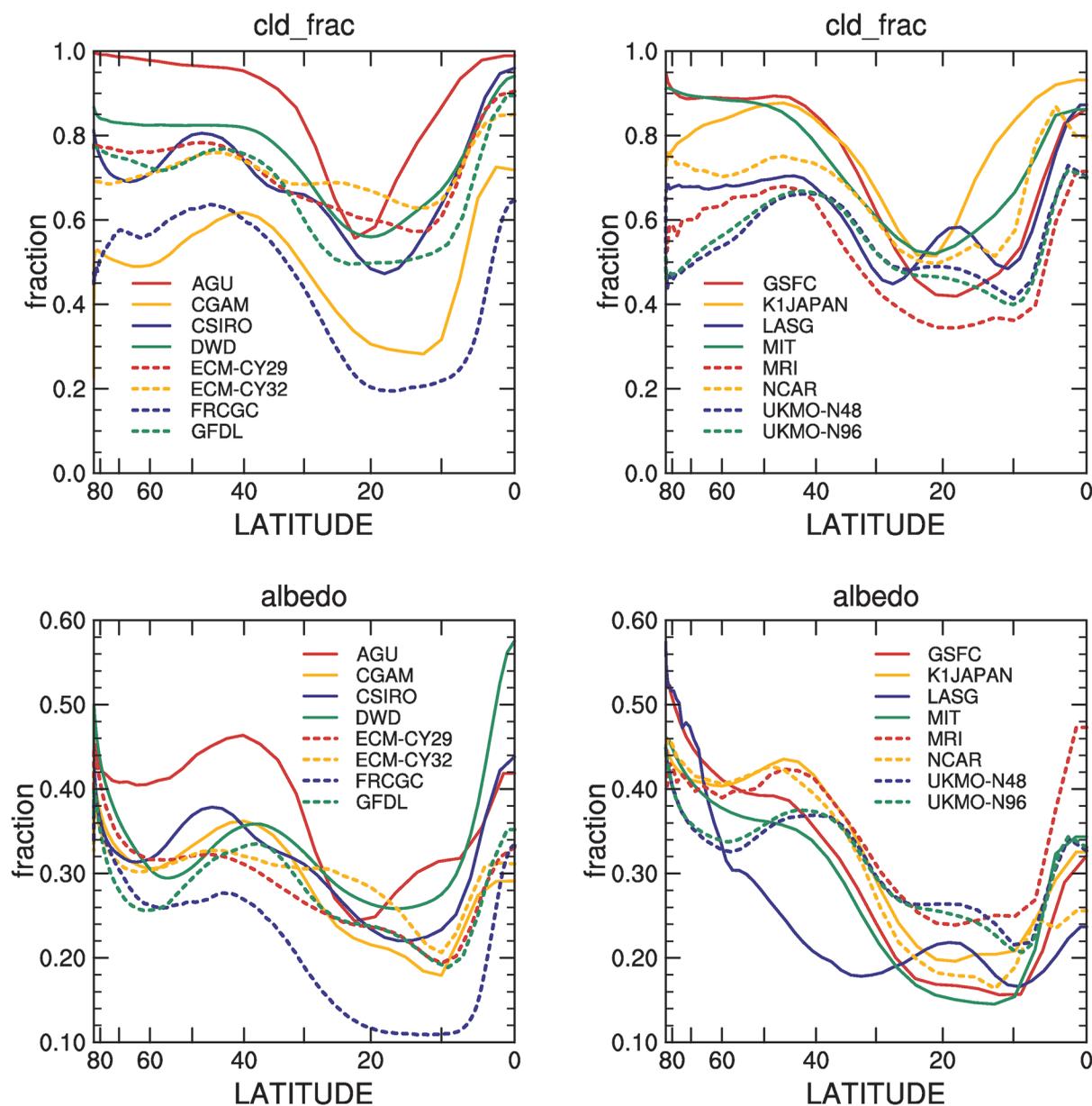


Fig. 12. Zonal-time average cloud fraction (*cld\_frac*) and albedo.

(apart from some exceptions discussed below). At higher wavenumbers the effects of truncation and damping lead to a wide range of energies and spectral slopes. There is an approximate correspondence between the spectral slope in Fig. 13 and the transient eddy KE maxima in Table 5, discussed in Section 6.1, with shallow spectral slope corresponding to high transient eddy KE and vice versa. Resolution is an additional factor, with higher resolution contributing energy at additional wavenumbers, more so for spectra

with shallow slope. Thus, the overall level of transient eddy KE is strongly controlled by the explicit and implicit damping in the dynamical core.

The Eulerian spectral transform models (AGU, K1JAPAN, LASG, MRI and NCAR) all maintain a slope of  $-3$  to the truncation limit for the total kinetic energy. Generally, this is by design and obtained by adding a  $\nabla^4$  diffusion term with the coefficient chosen to yield a slope of  $-3$ , to be consistent with observational evidence of Nastrom and Gage (1985) at larger

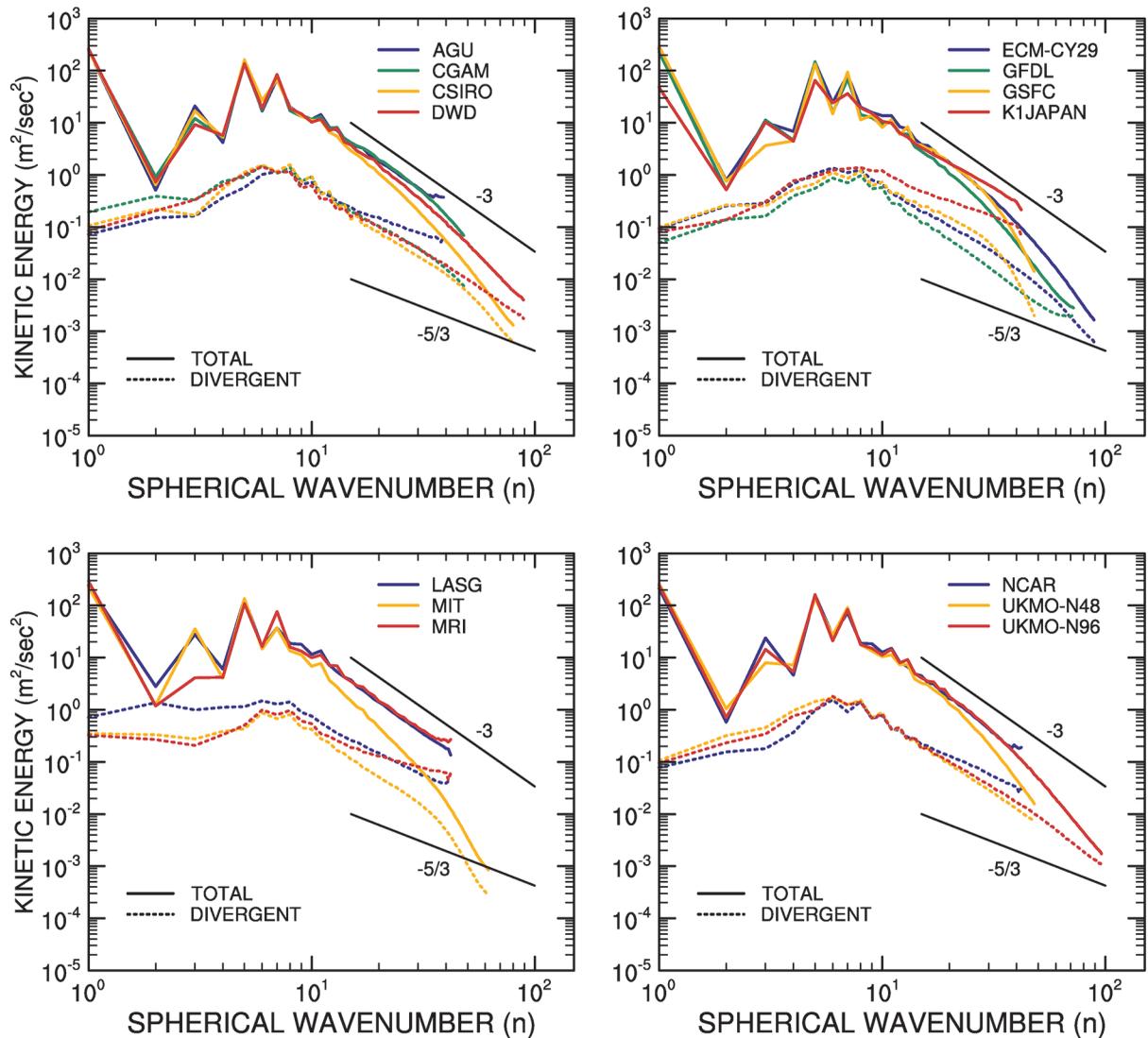


Fig. 13. Kinetic energy spectra with divergent component at 250 hPa ( $\text{m}^2 \text{s}^{-2}$ ).

length scales ( $n \leq 80$ ). In these cases with a  $-3$  kinetic energy slope, the divergent component has a slope of  $-5/3$ . It has been argued that a  $\nabla^6$  or  $\nabla^8$  diffusion term which yields a steeper slope is desirable to reduce spectral ringing (MacVean, 1983) but  $\nabla^4$  has remained the most popular choice for Eulerian spectral transform models. The spectra from ECM-CY29, which is based on a semi-Lagrangian spectral transform core, drop off faster than  $-3$  for the kinetic energy and faster than  $-5/3$  for the divergent component. This behaviour is strongly influenced by the truncation error of the (non-linear) semi-Lagrangian advection operator and explicit or implicit diffusion in the

dynamical core. The grid point models also all show faster drop-offs, some models faster than others, due to damping inherent in the grid point methods. The extra steep segment above wavenumber 40 of MIT might be due to the postprocessing interpolations to the latitude-longitude grid. Interpolation might also affect the spectra of CSIRO and DWD, but it has little impact on the plotted component of the spectra for ECM-CY29, i.e. for wavenumbers  $n < 90$  that are resolved by the regular  $2^\circ$  interpolation grid.

The model KE spectra are in much better agreement below wavenumber  $n \sim 10$ . This is mainly due to similarity of the zonal average flow among the

models, which contributes most of the energy at low wavenumbers in the absence of significant stationary wave amplitudes. Separate KE spectra for the zonal average state (zonal wavenumber  $m = 0$ ) and eddy ( $m > 0$ ) components in the NCAR model confirm this (not shown).

The low wavenumber behaviour is consistent with a high resolution aqua-planet simulation reported by Hamilton et al. (2008), hereafter HTO, (their Fig. 12). Compared to Earth simulations and observational analyses, also shown by HTO, the shape of the spectrum differs in APE, with lower energy in wavenumbers 2–4 and higher energy in wavenumbers 1 and 5–7. Lower energy is consistent with weak stationary wave amplitudes, as concluded by HTO, but higher energy and the differing shape is associated with the latitudinal structure of the zonal average flow. In the APE CONTROL the westerly jet is stronger and at lower latitude than Earth, implying a stronger super-rotation and therefore higher energy in  $n = 1$  (HTO, Fig. 12), and possibly higher energy at other wavenumbers. There is a large range of energy, around  $5\text{--}30\text{ m}^2\text{ s}^{-2}$ , for a single wavenumber,  $n = 3$ , among the APE models in Fig. 13. Comparison with zonal-time average zonal wind for individual models in the ATLAS shows that a more poleward jet and weaker westerlies throughout the tropics are associated with higher KE in  $n = 3$ , and vice versa. Variations of the zonal average state among the APE models therefore project particularly strongly onto the  $m = 0$ ,  $n = 3$  vorticity harmonic.

Compared to Earth simulations, there is an exaggerated difference between the energy in odd and even wavenumbers below  $n = 10$  in APE, also seen in dry dynamical core simulations shown by Takahashi et al. (2006) and HTO. This is due to weak stationary waves in the idealised configurations and strong north-south symmetry of the zonal average state about the equator, which increases odd  $n$  and reduces even  $n$  vorticity harmonics for zonal wavenumber  $m = 0$  (confirmed in the NCAR model, not shown).

The divergent spectra in Fig. 13 show less agreement among the models at low wavenumber. Variations appear to be due to both the zonally symmetric overturning circulation and low zonal wavenumber tropical variability (discussed later in Section 10). The Hadley cell outflow is generally above 250 hPa, so differences in both vertical structure and strength of meridional flow affect divergent KE at 250 hPa. LASG, which has the most active tropical variability, exhibits the largest divergent KE at low wavenumbers.

## 9. Mid-latitude low frequency variability

A very noticeable wavenumber-five pattern often appears in mid-latitudes in long time averages of APE simulations. It is most noticeable in maps of the meridional velocity where wavenumber-five can be seen as the dominant pattern. Examples are included in the ATLAS. Watanabe (2005), Cash et al. (2007) and Watanabe (2007) have discussed this mode and its relationship to the annular mode. Watanabe (2005) describes it as a slowly propagating, wavenumber-five disturbance forced by high-frequency eddies. More recently, Zappa et al. (2011) provide evidence that the feature is a weakly unstable baroclinic wave with very low phase speed, belonging to a baroclinic spectrum that obeys a well defined dispersion relation. In the ECHAM5 model studied by Zappa et al (not included in APE), the inverse energy cascade and phase locking with tropical convection each provide only a supplementary positive feedback.

Unfortunately the design of the APE experiment and the data collected do not allow a thorough examination of the phenomenon. As is seen in Fig. 14, the 3-year averages of 250 hPa meridional velocity at  $30^\circ$  latitude show amplitudes of order  $1\text{ m s}^{-1}$  in zonal wavenumber-five. This gives an apparent amplitude of the mode at a single timescale, averaging over different phases due to slow phase propagation, the speed of which may differ between models. The true amplitude, following the phase propagation, cannot be robustly deduced from the data collected in APE, but is expected to decrease with increasing averaging period. The one year of 6-hourly data of a few two-dimensional fields is too short and does not allow examination of vertical aspects. However, the 3-year averages do provide an estimate of the mode's linear and covariance structure in the APE models. The APE boundary conditions mean that there is nothing to set the phase of the wave in long-period averages and, indeed, apparently random phase variation is seen among the APE models and also for many 3-year samples from the NCAR model (not shown).

Nevertheless, most of the APE models indicate a significant mid-latitude wavenumber-five structure in the 3-year average which explains most of the “stationary” eddy variance on this timescale. The top row of Fig. 14 shows the amplitude of wavenumber-five of 3-year average meridional velocity at  $30^\circ\text{N}$  and  $30^\circ\text{S}$  at 250 hPa. The left panel shows all the APE models except the very high resolution FRCGC which was only integrated for 30 days. The models are identified numerically in Table 2. Clearly, there is no consis-

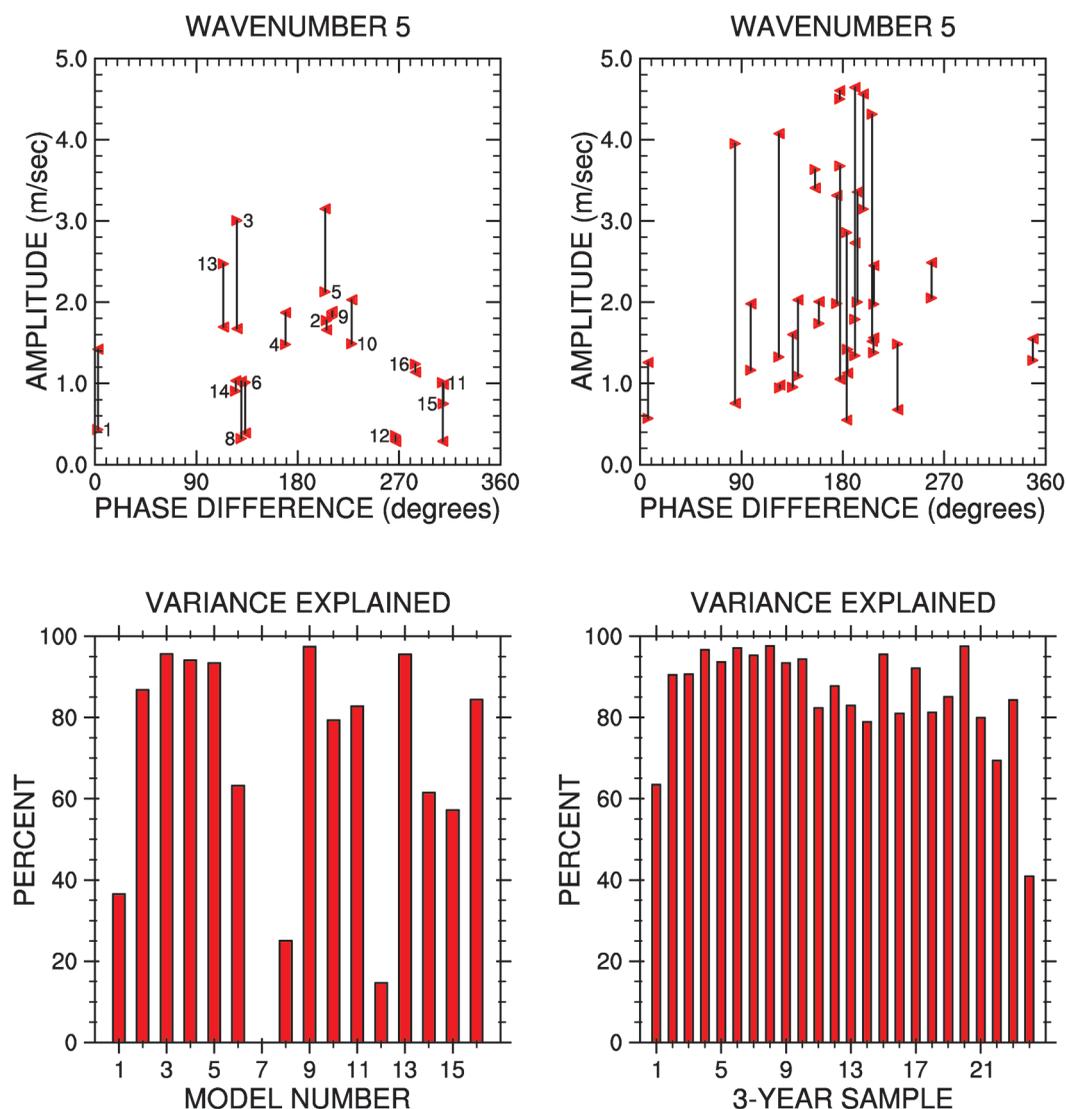


Fig. 14. Top row: Amplitude of wavenumber-five of 3-year average meridional velocity at 30°N (right arrow head) and 30°S (left arrow head) at 250 hPa, versus harmonic phase difference between the hemispheres, for (left) 15 APE models omitting FRCGC, numbered as in Table 2 and (right) 24 sequential 3-year averages from the NCAR model. Bottom row: percent variance explained by wavenumber-five at 30°N.

tency in the apparent amplitude of the mode, although this may be due to differences in both phase propagation and true amplitude among the models. The right panel shows the amplitude for 24 successive 3-year averages from a 72 year integration of the NCAR model. This provides an indication of the variability in amplitude for 3-year averages. The APE models are consistent with this variability, with models 1, 8 and 12 slightly low in amplitude. The bottom row shows the percent variance explained by wavenumber-five at 30°N in the

APE models (left) and in the 24 NCAR samples (right). The cases with relatively small amplitude show less variance explained. In the cases with more than 80% variance explained, maps of the meridional velocity shown in the ATLAS are clearly dominated by wavenumber-five. In those around 60% (models 6, 14, 15), a strong wavenumber-five is modulated by other waves. In those below 40%, wavenumber-five is not visible.

As found by Zappa et al. (2011) for the ECHAM5 model, a number of the APE models also show weak

wavenumber-five modulation of the 3-year average tropical circulation, including low level winds, evaporation and moisture content in the sub-tropics and, less clearly, modulation of ITCZ precipitation (not shown). To see whether this produces a tendency for phase locking of wavenumber-five between northern and southern mid-latitudes, the top panels of Fig. 14 show the amplitude of 3-year average wavenumber-five at 30°N and 30°S as a function of its harmonic phase difference between the hemispheres. For the NCAR model the phase differences are clustered around 180° (meridional velocity out of phase, geopotential in-phase), suggesting significant phase interaction between the hemispheres. There is little evidence of phase interaction across all the APE models, but NCAR does show higher apparent amplitude of wavenumber-five than many of the other models. A lack of symmetry about 180° also suggests that the APE sample size is insufficient.

Selected “stationary” eddy covariances for the 3-year averages are shown in the ATLAS for individual models, for all zonal wavenumbers and for wavenumber-five separately. For wavenumber-five the models generally have very weak poleward temperature flux in the lower troposphere between ~30–40° latitude, consistent with weak baroclinic instability. In the upper troposphere there is very weak poleward eddy momentum flux near 30–40° latitude, with equatorward flux positioned further poleward in a few models. These structures are broadly similar to those of the dominant transient eddy fluxes in Fig. 6. The wavenumber-five “stationary” eddy fluxes therefore appear consistent with the broader spectrum of baroclinic waves in the model storm-tracks, consistent with Zappa et al. (2011), but analysis of longer integrations would be necessary for a more complete, quantitative analysis.

## 10. Tropical variability

The overall magnitude of tropical variability in the models is compactly summarised by the standard deviation of six hour average precipitation on the model/data grid about the zonal-time average, shown in Fig. 15. This varies by a factor of four among the models, excluding FRCGC whose standard deviation on its fine-scale grid is almost double the maximum among the remaining coarser scale models. The range of standard deviation, and the value in many individual models, approximately scales with the zonal average precipitation at the equator in the ITCZ, which varies by a factor of around 3.5 in Fig. 4. Since there is a compensation between ITCZ intensity and width, due

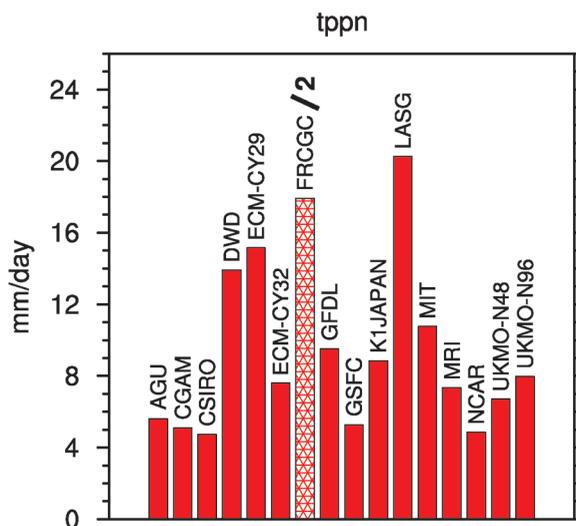


Fig. 15. Standard deviation of precipitation about the zonal-temporal average from 20°S to 20°N (tppn), mm day<sup>-1</sup>. The FRCGC value has been divided by 2 to show the variation among the models more clearly.

to the tropical average energy budget discussed in Section 5.4, models with a narrow single ITCZ tend to have stronger precipitation variability, while models with a broader or double ITCZ tend to have weaker variability. This relationship has been seen in a number of previous studies (e.g., Slingo et al. 1996; Wang and Schlesinger 1999; Lee et al. 2003, 2008).

Differing horizontal resolution among the models is likely to be a contributing factor to the intensity of precipitation on the model grid, as Williamson (2008a) found in the NCAR model. The smallest horizontal scales of latent heating are not directly relevant to large scale forcing, but they are part of the forcing spectrum whose interaction with the dynamics determines the nature of the tropical variability, in particular its organisation in large scale equatorial waves. The following sections address these aspects of the simulations.

### 10.1 Tropical waves

The models show a particularly large variation in their equatorial wave propagation characteristics. This is most readily visualised in Hovmöller plots of equatorial precipitation, averaged between 5°S and 5°N for sample 30 day periods for the individual models, in Fig. 16. It would be an understatement to say that the models display a wide variety of propagation characteristics! Eastward propagating features dominate in many of the models, with a phase speed that is remark-

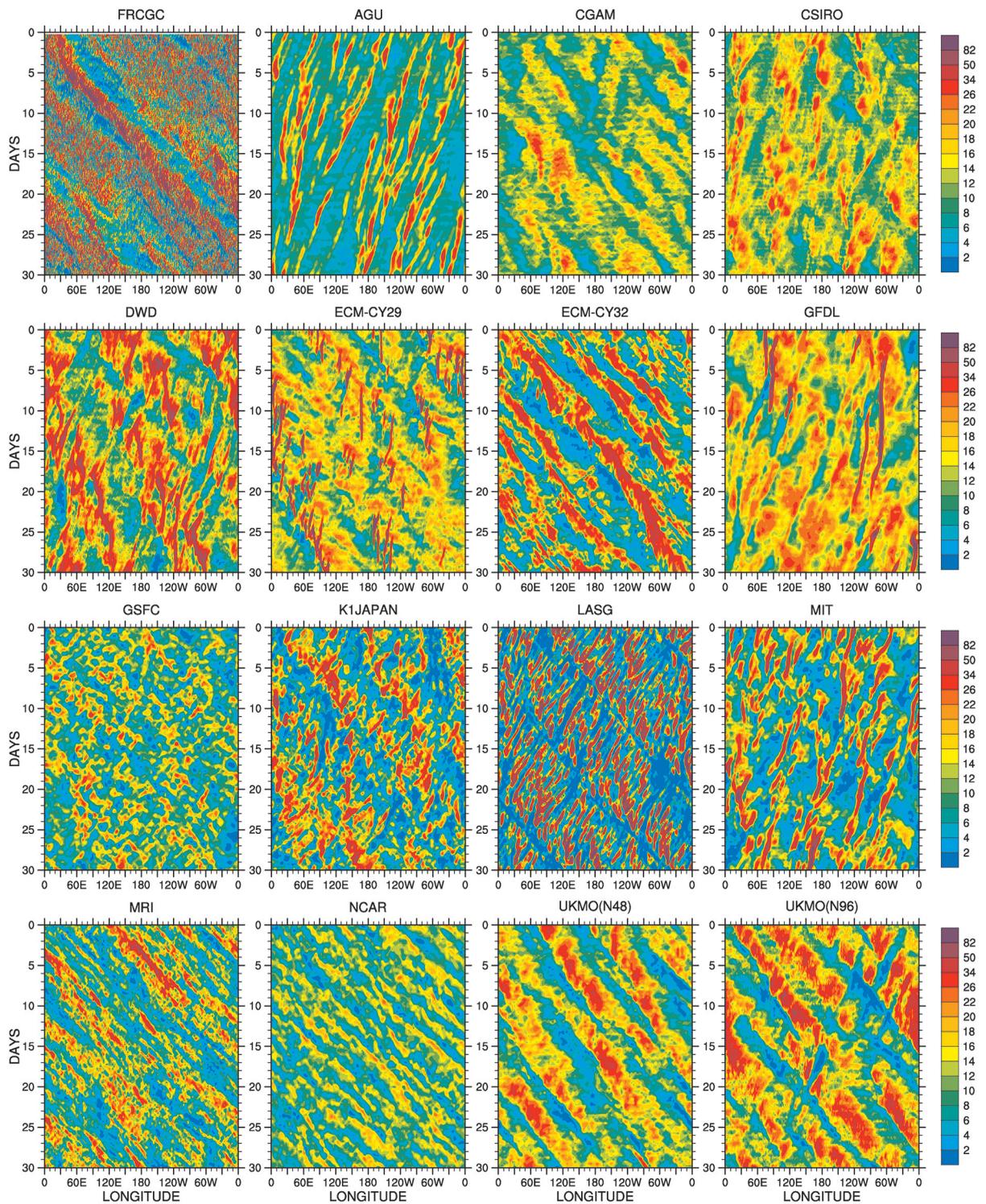


Fig. 16. Hovmöller plots of equatorial precipitation averaged from  $-5^{\circ}$  to  $+5^{\circ}$  latitude, for an arbitrary 30 day period from each model,  $\text{mm day}^{-1}$ .

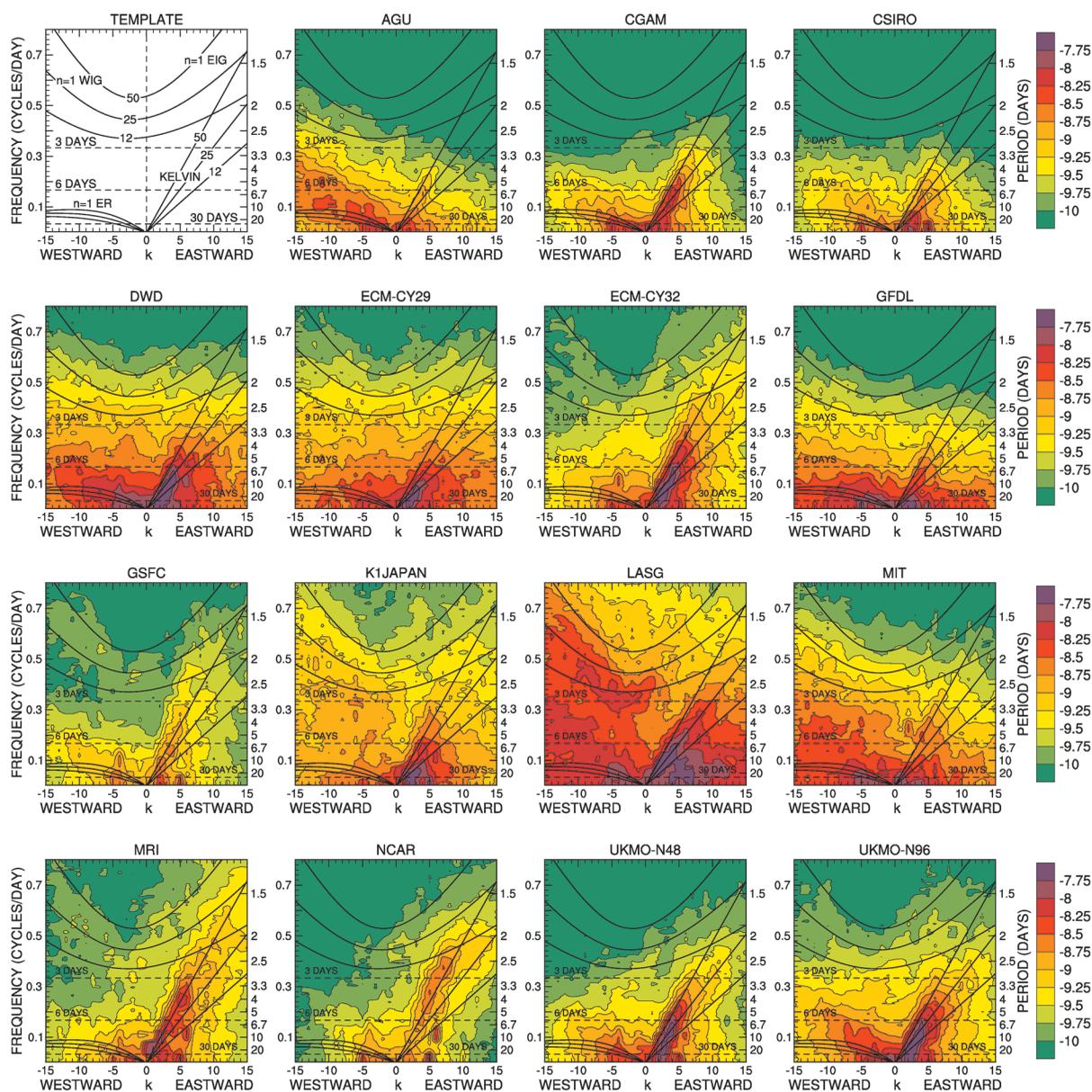


Fig. 17. Wavenumber-frequency diagrams of log of power of symmetric equatorial precipitation averaged from  $-10^\circ$  to  $+10^\circ$  latitude. Dispersion curves for the symmetric theoretical equatorial modes are included for reference. See text for details.

ably constant among the models, around  $15 \text{ m s}^{-1}$ , and independent of zonal scale. Westward propagation dominates in only two models, AGU and MIT. In a number of models, westward propagation of smaller scale features occurs within an eastward propagating envelope. Intense small scale quasi-stationary features occur in two further models, ECM-CY29 and GFDL. The intensity of the propagating features also varies

widely. A diurnal cycle is visible in several models.

This widely varying behaviour may be quantified using wavenumber-frequency diagrams, following the spectral analysis methodology of Wheeler and Kiladis (1999). Figure 17 shows the log of the power of the symmetric component of the unnormalized spectra of the 6-hour averaged precipitation averaged from  $10^\circ\text{S}$  to  $10^\circ\text{N}$ . The full power is plotted without removing a

background spectrum in order to allow a comparison of the overall power of the waves. Normalization by a smooth background field, such as done by Wheeler and Kiladis, isolates spectral peaks that are often associated with specific normal modes or waves. However, since the normalization is done individually for each model, the signal associated with the overall power is contained in the background spectrum and would be lost. In addition, in the aqua-planet the wave characteristics emerge from the total without needing a normalization. Figure 17 includes the conventional dispersion curves (labelled in the template panel) for odd meridional mode-numbered equatorial waves for equivalent depths of 12, 25, and 50m (see Wheeler and Kiladis 1999). The curves in the upper half of the plots show westward and eastward propagating inertio-gravity waves. Those at lower frequency show eastward propagating Kelvin modes and westward propagating equatorial Rossby modes. The period decreases with increasing equivalent depth in each set. The ordering of the models in Fig. 17 is the same as for the Hovmöller diagrams, for direct comparison.

There are spectral peaks associated with the equatorial wave modes in all the models, but the fraction of total power in these modes varies greatly. In a few models the modes are embedded in a background spectrum whose power increases with period, while in others most of the power projects onto the Kelvin and equatorial Rossby modes. A stronger background spectrum is characterised by more intermittent large scale eastward propagation in the Hovmöller diagrams, so occurs in those models which contain short lived, smaller scale westward propagation within an eastward propagating envelope.

Eastward propagating Kelvin modes are present in all models, and dominate the spectrum in almost half. Kelvin mode power generally peaks at low frequencies but extends further to high frequencies in some models, while 2–5 day periods dominate in the NCAR model. These differences appear in the Hovmöller diagrams as variations in the zonal wavelength of eastward propagating features whose phase speed varies remarkably little between the models. The small phase speed variations that do exist, e.g. faster in GSFC and NCAR, are visible in both figures. The agreement on eastward propagating phase speed is in marked contrast to aqua-planet simulations by Lee et al. (2008) in which AGCMs were coupled to a mixed layer ocean. Lee et al.'s finding of significant differences in phase speed may be due to widely varying equatorial SSTs among their coupled models, leading to differences in static stability, precipitation rates and heating profiles.

Westward propagation dominates in a smaller number of models, either as a westward bias in the background spectrum in LASG and MIT, or as a stronger projection onto the equatorial Rossby modes in models such as AGU. However, westward low frequency power is generally less dispersive than the Rossby modes, with higher zonal wavenumbers “lifted” to higher frequency. This is likely to be a doppler shifting due to the presence of easterly zonal flow throughout the equatorial troposphere in the APE CONTROL (Gui-Ying Yang, 2011, personal communication).

A few models contain some clearly isolated waves, such as NCAR and to a lesser extent MRI with a 10 day eastward propagating wavenumber 6. The low frequency wavenumber 5 signal in mid-latitudes, discussed in Section 9, can also be seen here in the tropics in some models such as NCAR.

The anti-symmetric precipitation spectra (shown in the ATLAS) contain westward propagating non-dispersive modes similar to their symmetric counterparts, though with slightly shorter periods in a few models. Almost half the models have eastward propagating power consistent with the mixed-Rossby-gravity (MRG) mode. However, in some of these (ECM-CY32, CGAM and UKMO-N48) there is a clear overlapping Kelvin mode signature, suggesting that ITCZ precipitation not entirely symmetric about the equator can excite and propagate with Kelvin modes. In fact calculations over 3-years of 6-hourly precipitation data from the NCAR model, which has a double ITCZ, show that the time-longitude correlation between precipitation at the northern and southern hemisphere latitudes of maximum zonal average precipitation is near zero. Spectra of outgoing longwave radiation (also shown in the ATLAS) are very similar to those of precipitation, but lack the background precipitation spectrum seen in a few models. Apparently the more stochastic precipitation in these models does not prevent their generation of equatorial waves.

Tropical wave spectra are available for Earth climate simulations in CMIP3 for several of the APE models, documented by Lin et al. (2006). These are GFDL, K1JAPAN (named MIROC-medres in CMIP3), MRI and NCAR (run at higher resolution). Of these models, only K1JAPAN has a similar spectrum to observations. The others have much weaker power, particularly at higher frequencies, and weaker projection onto the equatorial wave modes. However, in CMIP3 each model retains the character of its APE spectrum compared to the other models. It is notable that, in APE, K1JAPAN shows perhaps the clearest hierarchical organisation of equatorial convection in the

Hovmöller diagrams (Fig. 16).

It is difficult to isolate the causes of specific behaviour in individual models or groups of models, since there are multiple differences in formulation between each model. Nor has a clear dependence been found on the background flow (mean state). However, comparison of ECM-CY29 and ECM-CY32 indicates the effect the parameterizations can have on the equatorial wave spectrum. These two models have the same dynamical core and differ in their parameterization suites, including the closure for convective triggering and entrainment (Bechtold et al. 2008). The changes in ECM-CY32 led to significant improvements in the level and realism of forecast tropical variability, documented by Bechtold et al., implying that the increased Kelvin mode projection in the APE spectrum in ECM-CY32 is more “realistic” than ECM-CY29. Williamson (2008a) studied the dependence of the tropical wave spectrum on horizontal resolution and timestep in the NCAR model and found that the low resolution (T42) version used in APE is far from convergence. In fact the tropical wave characteristics do not converge for zonal wavenumbers less than 16 until T170 truncation in the NCAR model. Compared to T42, the converged spectrum contains relatively more power in lower frequency Kelvin modes, rather similar to ECM-CY32 but with a stronger background spectrum. However, it is not yet known whether a unique converged spectrum will emerge in multiple models for the APE CONTROL.

This analysis has been restricted to parameterization aspects of equatorial variability. Dynamical variables would also need to be analysed to formally show whether the equatorial wave characteristics seen in precipitation and OLR in the APE models are indeed convectively coupled with the dynamical structures predicted by equatorial wave theory.

### 10.2 Frequency distributions of precipitation

The frequency distribution of precipitation provides information about the extremes simulated by the models. In this section we show the frequency distribution of 6-hour averaged precipitation in the equatorial region, calculated between 10°S and 10°N to span the ITCZ in all the models in Fig. 4. Each column in Fig. 18 shows a duet of plots. The top row gives the fraction of time the precipitation is in each 1 mm d<sup>-1</sup> bin ranging from 0 to 120 mm d<sup>-1</sup>. The bottom row gives the fraction of time the precipitation is in each 10 mm d<sup>-1</sup> bin ranging from 0 to 600 mm d<sup>-1</sup>. In each plot the left-most bin is the fraction of time the precipitation is exactly 0, and the right-most bin in the top row is the fraction of time the precipitation exceeds 120 mm d<sup>-1</sup>.

Williamson (2008a) has shown that the distributions depend on the grid sizes. To eliminate this variation the fraction is calculated after the model grid data are conservatively mapped to a 5° latitude-longitude grid. Plots for the frequency distributions calculated on the original model grids are included in the ATLAS.

The smallest rates are difficult to discern in Fig. 18, so Fig. 19 is included to provide details of zero and light precipitation. For this figure the fractions are calculated on the original model/data grid, not from the data averaged to the 5° grid. For each model this shows the fraction of time the 6-hour averaged precipitation at grid points between 10°S to 10°N is zero and the fraction of time it is positive and less than 0.01, less than 0.1 and less than 1.0 mm d<sup>-1</sup>. Note that the smallest category,  $0 < p \leq 0.01$  mm d<sup>-1</sup>, is 1% of the smallest bin represented in the first row of Fig. 18.

The characteristics of trace and no precipitation are very different between models. Three models, AGU, GFDL and NCAR, rain almost all the time, with zero precipitation only 0.005, 0.001 and 0.008% of the time, respectively. A few models have zero precipitation for a significant fraction of the time: CSIRO, LASG, MRI, UKMO-N48 and UKMO-N96 do not rain 27.5, 37.4, 13.6, 10.8 and 9.1% of the time, respectively. The remaining models range from 0.1 to 4.0%. For the  $0 < p \leq 0.01$  range most models rain a few percent of the time. For the extreme cases, FRCGC rains 68% and AGU 0.03% of the time. There is much less model spread for  $0 < p \leq 1.0$ . Most models are in this range around 20% to 30% of the time. The exceptions are AGU at 3%, CSIRO at 9% and FRCGC at 80%. In this range, however, FRCGC is dominated by the small  $0 < p \leq 0.01$  component.

Figure 18 shows that, away from the very small values, the fraction of occurrence decreases monotonically with increasing rate, except for AGU where it peaks at around 7 mm d<sup>-1</sup> and CSIRO which has a slight peak at around 14 mm day<sup>-1</sup>. ECM-CY29 and GFDL show plateaus between 6 and 12 mm d<sup>-1</sup> and between 10 and 20 mm d<sup>-1</sup> respectively.

The models show a wide range of behaviour at large precipitation rates. ECM-CY29 has values exceeding 400 mm d<sup>-1</sup>, LASG exceeds 250 mm d<sup>-1</sup> and FRCGC, GFDL, K1JAPAN and MIT reach around 200 mm d<sup>-1</sup>. The remaining models have maximum values from 40 mm d<sup>-1</sup> to 100 mm d<sup>-1</sup>. On the original grids the higher resolution models have much higher maximum values, with rates greater than 1200 mm d<sup>-1</sup> in FRCGC and almost reaching 1200 mm d<sup>-1</sup> in ECM-CY29 (shown in the ATLAS).

As might be expected, the behaviour at large precip-

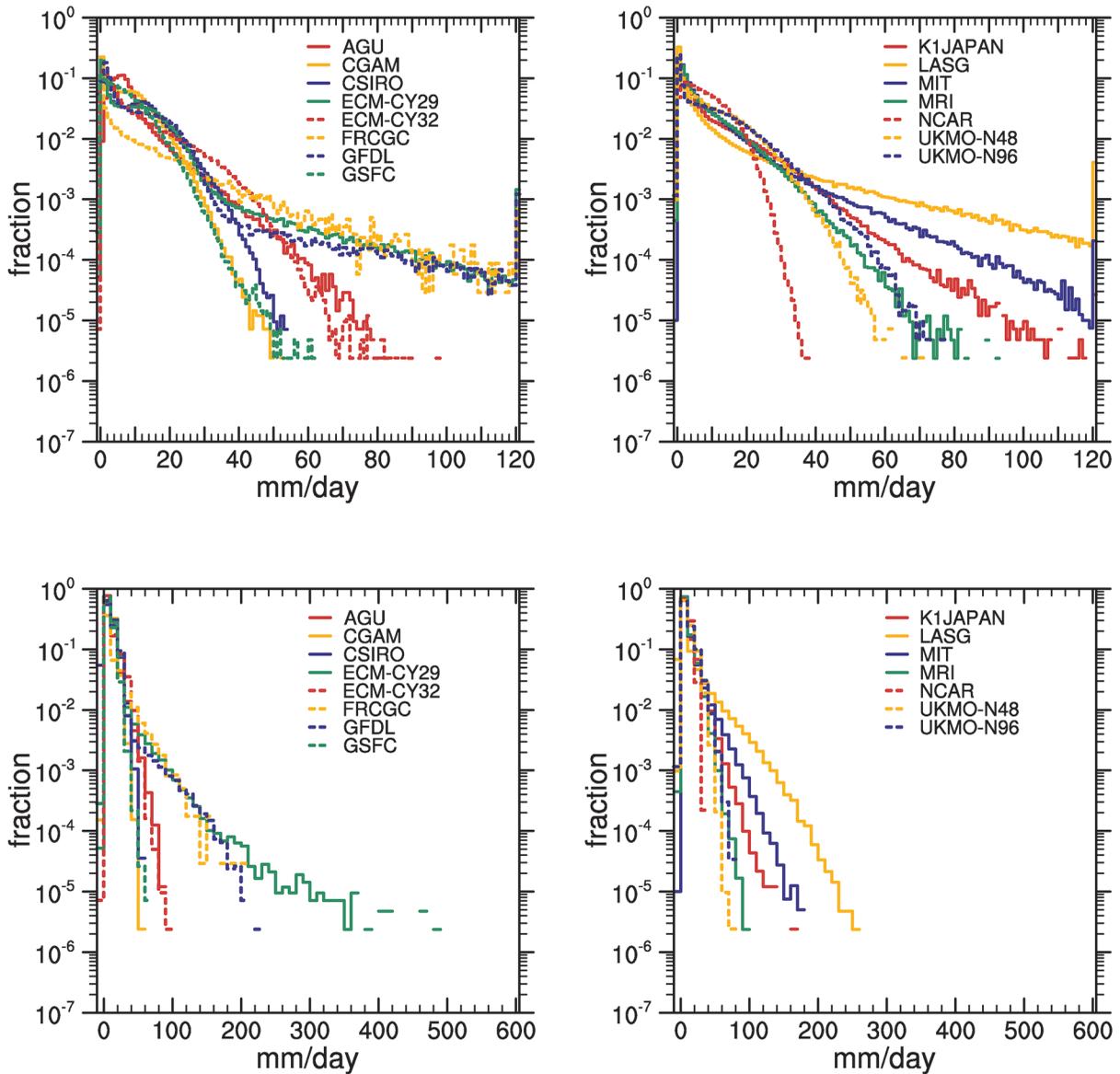


Fig. 18. Fraction of time precipitation is in (top row)  $1 \text{ mm day}^{-1}$  bins ranging from 0 to  $120 \text{ mm day}^{-1}$ , and (bottom row)  $10 \text{ mm day}^{-1}$  bins ranging from 0 to  $600 \text{ mm day}^{-1}$ . Grid values between  $10^{\circ}\text{S}$  and  $10^{\circ}\text{N}$  have been conservatively averaged to a  $5^{\circ}$  latitude-longitude grid. A gap in the curve indicates the fraction is zero for that bin.

itation rate is consistent with and clearly dominates variations in the standard deviation of tropical precipitation in Fig. 15. The models with a long tail in precipitation frequency and largest maximum values on the  $5^{\circ}$  grid are those with a significant background spectrum in Fig. 17 and intense quasi-stationary or westward propagating features in the Hovmöller diagrams. The models with a more limited frequency distribution have tropical wave spectra that are restricted to lower

frequencies and project strongly onto the Kelvin and equatorial Rossby modes. Recalling the correlation of standard deviation of precipitation with ITCZ width noted earlier, a pattern emerges that a narrow single ITCZ generally occurs in models that are able to generate more intense equatorial precipitation features, whereas a broader or double ITCZ generally occurs in models with more limited precipitation rates associated with dominant lower frequency variability.

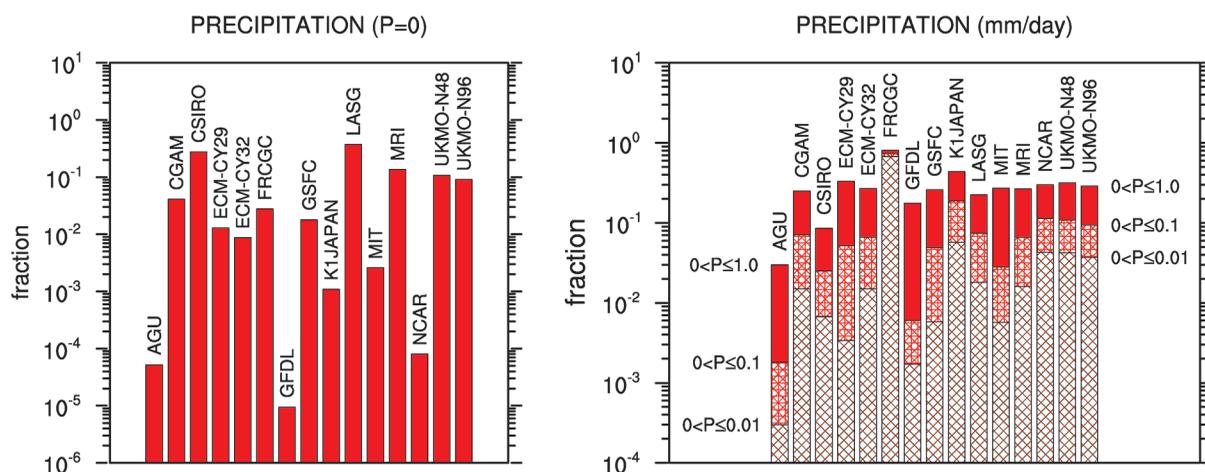


Fig. 19. Fraction of time precipitation,  $p$ , on the model grid between  $10^{\circ}\text{S}$  and  $10^{\circ}\text{N}$  is (left) equal to zero, and (right) in the ranges  $0 < p \leq 0.01$ ,  $0 < p \leq 0.1$  and  $0 < p \leq 1.0 \text{ mm day}^{-1}$ .

## 11. Discussion

The preceding sections describe a range of circulation statistics from comparison of the APE models for the aqua-planet CONTROL SST experiment. The conclusions of the individual sections are not repeated here and the reader is referred back to each section for more detailed discussion. Here, aspects of the aqua-planet simulations worthy of particular note are discussed and more general conclusions are drawn from the model comparison.

A multi-model mean and standard deviation have been computed, as a means of showing variation between the models in a concise form. The multi-model mean is not intended to be a reference solution for the CONTROL aqua-planet, whose circulation cannot be precisely known. The aqua-planet climate is very earth-like in structure, but the SST profile, its zonal symmetry, the absence of orography and perpetual equinoctial insolation combine to create a stronger circulation, with both a stronger westerly jet and tropical meridional circulation.

In comparing atmospheric model simulations on an aqua-planet, it is not clear a priori whether the range of model behaviour will be larger or smaller than for an equivalent comparison of simulations of Earth climate. On the one hand, removing land, orography and zonal SST variations results in a simpler circulation, which might lead to an expectation of reduced model spread. Conversely, the lack of zonally asymmetric forcing (additional to the diurnal cycle of insolation) implies that all longitudinal structure and variability must

be internally generated by the model dynamics and parameterizations and by interactions between them. This might lead to an expectation of increased model spread and was an important motivation in the design of the idealised aqua-planet configuration by Neale and Hoskins (2000a).

Indeed, the comparison does reveal a wide range of model behaviour in most of the atmospheric circulation statistics considered. However, certain parameters are more constrained, with relatively small differences between the models. Generally, it is aspects of the circulation determined mainly by the resolved dynamics that display less variation, whereas aspects that involve interaction between dynamical and parameterized moist processes appear to be most variable between the models. To illustrate this for the most basic statistic, namely the time average zonal average tropospheric state, the balanced component (temperature and zonal wind) varies relatively little between the models, while the tropical divergent circulation varies significantly, associated with the strength, width and vertical structure of the equatorial ITCZ. The trade wind inversion and longwave cooling in the sub-tropics associated with boundary layer cloud also vary significantly between the models. Furthermore, even for the dynamical circulation, quantities influenced by budget constraints have smaller variation. For example, the transient eddy covariances (poleward fluxes of heat and momentum) show significantly less variation between the models than the transient eddy variances (including kinetic energy), which are particularly sensitive to model resolution and numerical damping.

In contrast to the tropospheric mean state, the lower stratosphere varies greatly between the models, due, at least in part, to the proximity and differing numerical treatments of the upper boundary. The historical uncertainty in both polar and tropical lower stratospheric temperatures persists in APE, with a larger range of tropical tropopause temperature and vertical structure than in a recent AMIP comparison (part of CMIP3). This may reflect the inclusion in APE of a number of development models, compared to the more established models in AMIP and other intercomparisons.

All aspects of tropical precipitation vary greatly between the models. In the time average, the ITCZ appears as either a single peak on the equator or a double peak with a relative minimum at the equator in individual models, in common with simulations of Earth climate (Lin 2007, and references therein). This behaviour is sensitive to the latitudinal profile of SST, with evidence of a transition from a single equatorial ITCZ to a widely separated double ITCZ as the tropical SST profile flattens, as discussed by Neale and Hoskins (2000b) and Williamson et al. (2013). In single model sensitivity studies (using the NCAR model), Williamson and Olson (2003) and Williamson (2008a) have shown that the ITCZ structure is sensitive to the convective parameterization, to details of its implementation and to its interaction with the model dynamics.

The spatio-temporal variability of tropical precipitation displays varying degrees of projection onto the theoretical equatorial wave modes. In most models with a strong projection, convective features propagate predominantly eastward, but in a few models westward propagation dominates. Eastward propagation generally projects onto the non-dispersive Kelvin modes and has remarkably constant phase speed among the models. Westward propagation is less dispersive than the theoretical equatorial Rossby modes. Several models display the hierarchical convective organisation seen in Earth observations (Nakazawa 1988), with smaller scale westward propagation appearing within an envelope of eastward propagation. This is particularly true of the higher resolution models that have grid spacings of order 100 km.

A relationship exists between the tropical mean state and variability among the APE models, consistent with the results of previous studies. A narrow single ITCZ tends to occur in models with more intense equatorial precipitation features, and a wider or double ITCZ in models dominated by lower frequency tropical waves and weaker precipitation variability. The reason for this relationship is unclear, beyond a scaling of precip-

itation variability by the equatorial zonal average. Nor does the relationship appear predictive of the transition to a widely separated double ITCZ when the tropical SST profile is flattened in the QOBS and FLAT SST experiments, discussed by Williamson et al. (2013). Two of the models with a narrow single ITCZ in the CONTROL experiment are among those most strongly split in QOBS.

The global energy balance is surprisingly unconstrained between the APE models, because it is one parameter that is strongly constrained in simulations of Earth climate. The global average TOA net flux is of course unknown for the aqua-planet, but it has a range of  $30 \text{ W m}^{-2}$  in the APE models, dominated by variations in shortwave reflection by clouds. In most of the models the TOA net flux is relatively insensitive to the aqua-planet SST profile, so the main contributing factors to this large range appear to be a lack of optimisation for present-day Earth climate in a majority of the models and, more speculatively, the presence of differing biases in TOA net flux over land in the models that are optimised. To resolve this uncertainty, it would be necessary to quantify the global energy balance both in aqua-planet and AMIP simulations of Earth climate for individual models. This illustrates the intention of Neale and Hoskins (2000a) that APE be considered one component of a modelling hierarchy. The detailed aqua-planet circulation is ultimately unknown, though its overall character is predicted by theory and idealised models. Aqua-planet simulations are therefore most usefully compared and assessed in relation to simulation of the known Earth climate.

The large differences seen in global average TOA net energy flux among the APE models also extend to its latitudinal distribution, forcing large differences, with a standard deviation of around 1 PW in mid-latitudes, in the total poleward energy transport of the atmosphere-ocean system required for balance. This implies that significantly different equilibrated climates would result if the APE models were coupled to either thermodynamic slab oceans or dynamic ocean models.

It is commonly thought that much of the difference between atmospheric GCM simulations is due to the sub-grid scale parameterizations. The results of APE are not sufficiently generic to allow us to comment on systematic impacts of particular modelling choices such as parameterization scheme or physics-dynamics coupling across the APE models. This requires experiments specifically targeted to such questions in a more constrained modelling environment. However, part of the difference in APE might be primarily due to the truncation error of the numerical approximations to the

resolved fluid flow component of the models, especially as many of the model resolutions are rather coarse, in the area of T42 spectral truncation or 2.5° grids. These do not fully capture baroclinic wave development and differ significantly from the same models at higher resolution (Jablonowski and Williamson 2006). This will influence the eddy variances and covariances in particular. However, even in this case, parameterization forcing may play a role.

Resolution will also influence phenomena that arise fundamentally from interaction between the model dynamics and physical parameterizations, particularly if the response of either model component to changes in the other is non-linear. In the APE simulations this most clearly applies to the convectively coupled equatorial waves. Williamson (2008a) has shown that in the NCAR CAM the tropical wave characteristics depend on resolution and for zonal wavenumbers less than 16 they do not converge until T170 truncation, while T42 is common in the APE experiments here. This lack of convergence at the resolutions that APE models have been run contributes to the difference among the models.

It is important, therefore, that convergence of the APE mean state and variability with increasing resolution is studied in a number of models. It is conceivable that individual models could converge towards a unique climate but, equally, they might converge to significantly different climates, because of their wide range of parameterizations and, possibly to a lesser extent, their dynamical approximations.

A future possibility might be the application of global cloud resolving models (CRMs) to the aqua-planet configuration, as computational capacity continues to increase. CRMs do retain parameterizations, of radiation, turbulence and cloud microphysics, but the more empirically based closure assumptions of convective parameterizations in large scale models are not required. If global CRMs were to converge to a single climate, more confidence would be gained that this was indeed the aqua-planet climate. However, confidence would also be required in the CRMs themselves, for example by accurate simulation of Earth's global climate or at least the particular phenomena relevant to an aqua-planet. Progress has already been made in this regard by coarser resolution cloud-system resolving models (CSRMs, one of which is included in APE). In a number of local case studies organised within the GEWEX Cloud Systems Studies (GCSS), CSRMs have proved successful in simulating observed convection more accurately than have parameterizations extracted from large scale models.

## Acknowledgments

We wish to thank Richard Neale and the anonymous reviewers whose comments helped to improve the manuscript. Richard Neale (NCAR), Peter Gleckler (Program for Climate Model Diagnosis and Intercomparison, PCMDI) and staff at PCMDI contributed to the APE diagnostic specification, data protocol and initial collection and quality control of the data. Williamson was partially supported by the Office of Science (BER), U.S. Department of Energy, Cooperative Agreement No. DE-FC02-97ER62402. He thanks J. Olson for carrying out the NCAR APE simulations. The National Center for Atmospheric Research is sponsored by the National Science Foundation. Bechtold and Wedi thank J.J. Morcrette and M. Hortal for their help in setting up the ECMWF/IFS simulations. Stratton thanks Chris Dearden for his help in reformatting and transferring the output from the Met Office simulations to PCMDI. Contributions from the Met Office are Crown Copyright. McGregor thanks Martin Dix for his assistance. The numerical calculation of AFES was carried out on the Earth Simulator under support of JAMSTEC.

## References

- Bechtold, P., M. Köhler, T. Jung, F. Doblas-Reyes, M. Leutbecher, M. J. Rodwell, F. Vitart, and G. Balsamo, 2008: Advances in simulating atmospheric variability with the ECMWF model: From synoptic to decadal time-scales. *Quart. J. Roy. Meteor. Soc.*, **134**, 1337–1351.
- Blackburn, M., and B. J. Hoskins, 2013: Context and aims of the Aqua-Planet Experiment. *J. Meteor. Soc. Japan*, **91A**, 1–15, doi:10.2151/jmsj.2013-A01.
- Boer, G. J., K. Arpe, M. Blackburn, M. Déqué, W. L. Gates, T. L. Hart, H. Le Treut, E. Roeckner, D. A. Sheinin, I. Simmonds, R. N. B. Smith, T. Tokioka, R. T. Wetherald, and D. L. Williamson, 1992: Some results from an intercomparison of the climates simulated by 14 atmospheric general circulation models. *J. Geophys. Res.*, **97**, 12771–12786.
- Bony, S., and J.-L. Dufresne, 2005: Marine boundary layer clouds at the heart of tropical cloud feedback uncertainties in climate models. *Geophys. Res. Lett.*, **32**, L20806, doi:10.1029/2005GL023851.
- Cash, B. A., P. Kushner, and G. Vallis, 2007: Comment on “On the presence of annular variability in an aqua-planet model” by Masahiro Watanabe. *Geophys. Res. Lett.*, **34**, L03707, doi:10.1029/2006GL027274.
- Cess, R. D., G. L. Potter, J. P. Blanchet, G. J. Boer, S. J. Ghan, J. T. Kiehl, H. Le Treut, Z.-X. Li, X.-Z. Liang, J. F. B. Mitchell, J.-J. Morcrette, D. A. Randall, M. R. Riches, E. Roeckner, U. Schlese, A. Slingo, K. E. Taylor, W. M. Washington, R. T. Wetherald, and I. Yagai, 1989:

- Interpretation of cloud-climate feedback as produced by 14 atmospheric general circulation models. *Science*, **245**, 513–515, doi:10.1126/science.245.4917.513.
- Cess, R. D., G. L. Potter, J. P. Blanchet, G. J. Boer, A. D. Del Genio, M. Déqué, V. Dymnikov, V. Galin, W. L. Gates, S. J. Ghan, J. T. Kiehl, A. A. Lacis, H. Le Treut, Z.-H. Li, X.-Z. Liang, B. J. McAvaney, V. P. Meleshko, J. F. B. Mitchell, J.-J. Morcrette, D. A. Randall, L. Rikus, E. Roeckner, J. F. Royer, U. Schlese, D. A. Sheinin, A. Slingo, A. P. Sokolov, K. E. Taylor, W. M. Washington, R. T. Wetherald, I. Yagai, and M.-H. Zhang, 1990: Intercomparison and interpretation of climate feedback processes in 19 atmospheric general circulation models. *J. Geophys. Res.*, **95**, 16601–16637, doi:10.1029/JD095iD10p16601.
- Chao, W. C., and B. Chen, 2004: Single and double ITCZ in an aqua-planet model with constant sea surface temperature and solar angle. *Climate Dyn.*, **22**, 447–459.
- Chen, M., and J. R. Bates, 1996: A comparison of climate simulations from a semi-Lagrangian and an Eulerian GCM. *J. Climate*, **9**, 1126–1149.
- Collins, W. D., P. J. Rasch, B. A. Boville, J. J. Hack, J. R. McCaa, D. L. Williamson, J. T. Kiehl, B. P. Briegleb, C. M. Bitz, S.-J. Lin, M. Zhang, and Y. Dai, 2004: Description of the NCAR Community Atmosphere Model (CAM3.0). Tech. Rep. NCAR/TN-464+STR, National Center for Atmospheric Research.
- Fasullo, J. T., and K. E. Trenberth, 2008: The annual cycle of the energy budget. part II: Meridional structures and poleward transports. *J. Climate*, **21**, 2313–2325.
- Gates, W. L., 1992: AMIP: The Atmospheric Model Intercomparison Project. *Bull. Amer. Meteor. Soc.*, **73**, 1962–1979.
- Gates, W. L., J. S. Boyle, C. Covey, C. G. Dease, C. M. Doutriaux, R. S. Drach, M. Fiorino, P. J. Gleckler, J. J. Hnilo, S. M. Marlais, T. J. Phillips, G. L. Potter, B. D. Santer, K. R. Sperber, K. E. Taylor, and D. N. Williams, 1999: An overview of the results of the Atmospheric Model Intercomparison Project (AMIP I). *Bull. Amer. Meteor. Soc.*, **80**, 29–55.
- Gleckler, P. J., D. A. Randall, G. Boer, R. Colman, M. Dix, V. Galin, M. Helfand, J. Kiehl, A. Kitoh, W. Lau, X.-Y. Liang, V. Lykossov, B. McAvaney, K. Miyakoda, S. Planton, and W. Stern, 1995: Cloud-radiative effects on implied oceanic energy transports as simulated by atmospheric general circulation models. *Geophys. Res. Lett.*, **22**, 791–794.
- Hamilton, K., Y. O. Takahashi, and W. Ohfuchi, 2008: Mesoscale spectrum of atmospheric motions investigated in a very fine resolution global general circulation model. *J. Geophys. Res.*, **113**, D18 110, doi:10.1029/2008JD009785.
- Hansen, J., L. Nazarenko, R. Ruedy, M. Sato, J. Willis, A. Del Genio, D. Koch, A. Lacis, K. Lo, S. Menon, T. Novakov, J. Perlwitz, G. Russel, G. A. Schmidt, and N. Tausnev, 2005: Earth's energy imbalance: Confirmation and implications. *Science*, **308**, 1431–1435, doi:10.1126/science.1110252.
- Hayashi, Y.-Y., and A. Sumi, 1986: The 30–40 day oscillations simulated in an "aqua planet" model. *J. Meteor. Soc. Japan*, **64**, 451–467.
- Held, I. M., and A. Y. Hou, 1980: Nonlinear axially symmetric circulations in a nearly inviscid atmosphere. *J. Atmos. Sci.*, **37**, 515–533.
- Hess, P. G., D. S. Battisti, and P. J. Rasch, 1993: Maintenance of the intertropical convergence zones and the large-scale tropical circulation on a water-covered earth. *J. Atmos. Sci.*, **50**, 691–713.
- Hoskins, B. J., I. N. James, and G. H. White, 1983: The shape, propagation and mean-flow interaction of large-scale weather systems. *J. Atmos. Sci.*, **40**, 1595–1612.
- Jablonski, C., and D. L. Williamson, 2006: A baroclinic instability test case for atmospheric model dynamical cores. *Quart. J. Roy. Meteor. Soc.*, **132**, 2943–2975.
- Källberg, P., P. Berrisford, B. J. Hoskins, A. J. Simmons, S. Uppala, S. Lamy-Thépaut, and R. Hine, 2005: ERA-40 Atlas. ERA-40 Report Series No. 19, European Centre for Medium-range Weather Forecasts, Reading, England, 191 pp. [Available at [http://www.ecmwf.int/research/era/ERA-40\\_Atlas](http://www.ecmwf.int/research/era/ERA-40_Atlas).]
- Lee, M.-I., I.-S. Kang, and B. E. Mapes, 2003: Impacts of cumulus convection parameterization on aqua-planet AGCM simulations of tropical intraseasonal variability. *J. Meteor. Soc. Japan*, **81**, 963–992.
- Lee, M.-I., M. J. Suarez, I.-S. Kang, I. M. Held, and D. Kim, 2008: A moist benchmark calculation for Atmospheric General Circulation Models. *J. Climate*, **21**, 4934–4954.
- Liang, X.-Z., and W.-C. Wang, 1996: Atmospheric ozone climatology for use in General Circulation Models. Atmospheric Sciences Research Center, State University of New York at Albany. [Available at <http://www-pcmdi.llnl.gov/projects/amip/AMIP2EXPDSN/OZONE/OZONE2/03wangdoc.html>.]
- Lin, J.-L., 2007: The double-ITCZ problem in IPCC AR4 coupled GCMs: Ocean-atmosphere feedback analysis. *J. Climate*, **20**, 4497–4525.
- Lin, J.-L., G. N. Kiladis, B. E. Mapes, K. M. Weickmann, K. R. Sperber, W. Lin, M. C. Wheeler, S. D. Schubert, A. Del Genio, L. J. Donner, S. Emori, J.-F. Guérémy, F. Hourdin, P. J. Rasch, E. Roeckner, and J. F. Scinocca, 2006: Tropical intraseasonal variability in 14 IPCC AR4 climate models. Part I: Convective signals. *J. Climate*, **19**, 2665–2690.
- MacVean, M. K., 1983: The effects of horizontal diffusion on baroclinic development in a spectral model. *Quart. J. Roy. Meteor. Soc.*, **109**, 771–783.
- Medeiros, B., and B. Stevens, 2011: Revealing differences in GCM representations of low clouds. *Climate Dyn.*, **36**, 385–399.
- Meehl, G. A., C. Covey, T. Delworth, M. Latif, B. McAvaney, J. F. B. Mitchell, R. J. Stouffer, and K. E. Taylor, 2007:

- The WCRP CMIP3 multi-model dataset: A new era in climate change research. *Bull. Amer. Meteor. Soc.*, **88**, 1383–1394.
- Nakazawa, T., 1988: Tropical super clusters within intraseasonal variations over the western Pacific. *J. Meteor. Soc. Japan*, **6**, 823–839.
- Nastrom, G. D., and K. S. Gage, 1985: A climatology of atmospheric wavenumber spectra of wind and temperature observed by commercial aircraft. *J. Atmos. Sci.*, **42**, 950–960.
- Neale, R. B., and B. J. Hoskins, 2000a: A standard test for AGCMs including their physical parameterizations. I: The proposal. *Atmos. Sci. Lett.*, **1**, 101–107, doi:10.1006/asle.2000.0022.
- Neale, R. B., and B. J. Hoskins, 2000b: A standard test for AGCMs including their physical parameterizations. II: Results for the Met Office model. *Atmos. Sci. Lett.*, **1**, 108–114, doi:10.1006/asle.2000.0024.
- Numaguti, A., 1993: Dynamics and energy balance of the Hadley circulation and the tropical precipitation zones: Significance of the distribution of evaporation. *J. Atmos. Sci.*, **50**, 1874–1887.
- Pope, V. D., J. A. Pamment, D. R. Jackson, and A. Slingo, 2001: The representation of water vapor and its dependence on vertical resolution in the Hadley Centre climate model. *J. Climate*, **14**, 3065–3085.
- Slingo, J. M., K. R. Sperber, J. S. Boyle, J.-P. Ceron, M. Dix, B. Dugas, W. Ebisuzaki, J. Fyfe, D. Gregory, J.-F. Gueremy, J. Hack, A. Harzallah, P. Inness, A. Kitoh, W. K.-M. Lau, B. McAvaney, R. Madden, A. Matthews, T. N. Palmer, C.-K. Parkas, D. Randall, and N. Renno, 1996: Intraseasonal oscillations in 15 atmospheric general circulation models: results from an AMIP diagnostic subproject. *Climate Dyn.*, **12**, 325–357.
- Sparrow, S. N., M. Blackburn, and J. D. Haigh, 2009: Annular variability and eddy-zonal flow interactions in a simplified atmospheric GCM: Part I - Characterization of high and low frequency behaviour. *J. Atmos. Sci.*, **66**, 3075–3094.
- Takahashi, Y. O., K. Hamilton, and W. Ohfuchi, 2006: Explicit global simulation of the mesoscale spectrum of atmospheric motions. *Geophys. Res. Lett.*, **33**, L12812, doi:10.1029/2006GL026429.
- Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2009: A summary of the CMIP5 experiment design. Tech. Rep., Programme for Climate Model Diagnosis and Intercomparison, 33 pp. [Available at [http://cmip-pcmdi.llnl.gov/cmip5/docs/Taylor\\_CMIP5\\_design.pdf](http://cmip-pcmdi.llnl.gov/cmip5/docs/Taylor_CMIP5_design.pdf).]
- Thuburn, J., and G. C. Craig, 2002: On the temperature structure of the tropical stratosphere. *J. Geophys. Res.*, **107**, doi:10.1029/2001JD000448.
- Trenberth, K. E., J. T. Fasullo, and J. Kiehl, 2009: Earth's global energy budget. *Bull. Amer. Meteor. Soc.*, **90**, 311–323.
- Wang, W., and M. E. Schlesinger, 1999: The dependence on convection parameterization of the tropical intraseasonal oscillation simulated by the UIUC 11-layer atmospheric GCM. *J. Climate*, **12**, 1423–1457.
- Wang, W.-C., X.-Z. Liang, M. P. Dudek, D. Pollard, and S. L. Thompson, 1995: Atmospheric ozone as a climate gas. *Atmos. Res.*, **37**, 247–256.
- Watanabe, M., 2005: On the presence of annular variability in an aquaplanet model. *Geophys. Res. Lett.*, **32**, L05701, doi:10.1029/2004GL021869.
- Watanabe, M., 2007: Reply to comment by B. A. Cash et al. on “On the presence of annular variability in an aquaplanet model”. *Geophys. Res. Lett.*, **34**, L03708, doi:10.1029/2006GL028669.
- Webb, M. J., C. A. Senior, D. M. H. Sexton, W. J. Ingram, K. D. Williams, M. A. Ringer, B. J. McAvaney, R. Colman, B. J. Soden, R. Gudgel, T. Knutson, S. Emori, T. Ogura, Y. Tsushima, N. Andronova, B. Li, I. Musat, S. Bony, and K. E. Taylor, 2006: On the contribution of local feedback mechanisms to the range of climate sensitivity in two GCM ensembles. *Climate Dyn.*, **27**, 17–38.
- Wheeler, M. C., and G. N. Kiladis, 1999: Convectively coupled equatorial waves: Analysis of clouds and temperature in the wavenumber-frequency domain. *J. Atmos. Sci.*, **56**, 374–399.
- Williamson, D. L., 2008a: Convergence of aqua-planet simulations with increasing resolution in the Community Atmospheric Model, Version 3. *Tellus*, **60A**, 848–862.
- Williamson, D. L., 2008b: Equivalent finite volume and spectral transform horizontal resolutions established from aqua-planet simulations. *Tellus*, **60A**, 839–847.
- Williamson, D. L., and J. G. Olson, 1994: Climate simulations with a semi-Lagrangian version of the NCAR Community Climate Model. *Mon. Wea. Rev.*, **122**, 1594–1610.
- Williamson, D. L., and J. G. Olson, 1998: A comparison of semi-Lagrangian and Eulerian polar climate simulations. *Mon. Wea. Rev.*, **126**, 991–1000.
- Williamson, D. L., and J. G. Olson, 2003: Dependence of aqua-planet simulations on time step. *Quart. J. Roy. Meteor. Soc.*, **129**, 2049–2064.
- Williamson, D. L., J. G. Olson, and B. A. Boville, 1998: A comparison of semi-Lagrangian and Eulerian tropical climate simulations. *Mon. Wea. Rev.*, **126**, 1001–1012.
- Williamson, D. L., M. Blackburn, B. J. Hoskins, K. Nakajima, W. Ohfuchi, Y. O. Takahashi, Y.-Y. Hayashi, H. Nakamura, M. Ishiwatari, J. L. McGregor, H. Borth, V. Wirth, H. Frank, P. Bechtold, N. P. Wedi, H. Tomita, M. Satoh, M. Zhao, I. M. Held, M. J. Suarez, M.-I. Lee, M. Watanabe, M. Kimoto, Y. Liu, Z. Wang, A. Molod, K. Rajendran, A. Kitoh, and R. Stratton, 2012: The APE Atlas. NCAR Technical Note NCAR/TN-484+STR, National Center for Atmospheric Research, Boulder, Colorado, xxii+508 pp, doi:10.5065/D6FF3QBR. [Available at <http://nldr.library.ucar.edu/repository/collections/TECH-NOTE-000-000-000-865/>.]

- Williamson, D. L., M. Blackburn, K. Nakajima, W. Ohfuchi, Y. O. Takahashi, Y.-Y. Hayashi, H. Nakamura, M. Ishiwatari, J. L. McGregor, H. Borth, V. Wirth, H. Frank, P. Bechtold, N. P. Wedi, H. Tomita, M. Satoh, M. Zhao, I. M. Held, M. J. Suarez, M.-I. Lee, M. Watanabe, M. Kimoto, Y. Liu, Z. Wang, A. Molod, K. Rajendran, A. Kitoh, and R. Stratton, 2013: The Aqua-Planet Experiment (APE): Response to changed meridional SST profile. *J. Meteor. Soc. Japan*, **91A**, 57–89, doi:10.2151/jmsj.2013-A03.
- Zappa, G., V. Lucarini, and A. Navarra, 2011: Baroclinic stationary waves in aquaplanet models. *J. Atmos. Sci.*, **68**, 1023–1040.