



HAL
open science

Overview of experiment design and comparison of models participating in phase 1 of the SPARC Quasi-Biennial Oscillation initiative (QBOi)

Neal Butchart, James Anstey, Kevin Hamilton, Scott Osprey, Charles Mclandress, Andrew Bushell, Yoshio Kawatani, Young-Ha Kim, Francois Lott, John Scinocca, et al.

► To cite this version:

Neal Butchart, James Anstey, Kevin Hamilton, Scott Osprey, Charles Mclandress, et al.. Overview of experiment design and comparison of models participating in phase 1 of the SPARC Quasi-Biennial Oscillation initiative (QBOi). *Geoscientific Model Development Discussions*, 2018, 11 (3), pp.1009-1032. 10.5194/gmd-11-1009-2018 . hal-03049573

HAL Id: hal-03049573

<https://hal.science/hal-03049573>

Submitted on 7 Jun 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



Overview of experiment design and comparison of models participating in phase 1 of the SPARC Quasi-Biennial Oscillation initiative (QBOi)

Neal Butchart¹, James A. Anstey², Kevin Hamilton³, Scott Osprey⁴, Charles McLandress^{5,2}, Andrew C. Bushell⁶, Yoshio Kawatani⁷, Young-Ha Kim⁸, Francois Lott⁹, John Scinocca², Timothy N. Stockdale¹⁰, Martin Andrews¹, Omar Bellprat¹¹, Peter Braesicke¹², Chiara Cagnazzo¹³, Chih-Chieh Chen¹⁴, Hye-Yeong Chun¹⁵, Mikhail Dobrynin¹⁶, Rolando R. Garcia¹⁴, Javier Garcia-Serrano¹¹, Lesley J. Gray⁴, Laura Holt¹⁷, Tobias Kerzenmacher¹², Hiroaki Naoe¹⁸, Holger Pohlmann¹⁹, Jadwiga H. Richter¹⁴, Adam A. Scaife^{1,20}, Verena Schenzinger²¹, Federico Serva^{13,22}, Stefan Versick¹², Shingo Watanabe⁷, Kohei Yoshida¹⁸, and Seiji Yukimoto¹⁸

¹Met Office Hadley Centre (MOHC), Exeter, UK

²Canadian Centre for Climate Modelling and Analysis (CCCma), Victoria, Canada

³International Pacific Research Center (IPRC), Honolulu, USA

⁴National Centre for Atmospheric Science (NCAS), University of Oxford, Oxford, UK

⁵University of Toronto, Toronto, Canada

⁶Met Office, Exeter, UK

⁷Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Yokohama, Japan

⁸Ewha Womans University, Seoul, South Korea

⁹Laboratoire de Météorologie Dynamique (LMD), Paris, France

¹⁰European Centre for Medium-Range Weather Forecasts (ECMWF), Reading, UK

¹¹Barcelona Supercomputing Center (BSC), Barcelona, Spain

¹²Karlsruher Institut für Technologie (KIT), Karlsruhe, Germany

¹³Istituto di Scienze Dell' Atmosfera e del Clima (ISAC-CNR), Rome, Italy

¹⁴National Center for Atmospheric Research (NCAR), Boulder, USA

¹⁵Yonsei University, Seoul, South Korea

¹⁶Universität Hamburg, Hamburg, Germany

¹⁷NorthWest Research Associates (NWRA), Boulder, USA

¹⁸Meteorological Research Institute (MRI), Tsukuba, Japan

¹⁹Max-Planck-Institut für Meteorologie (MPI), Hamburg, Germany

²⁰University of Exeter, Exeter, UK

²¹Universität Wien, Vienna, Austria

²²Università degli Studi di Napoli "Parthenope", Naples, Italy

Correspondence: Neal Butchart (neal.butchart@metoffice.gov.uk)

Received: 6 August 2017 – Discussion started: 26 October 2017

Revised: 22 January 2018 – Accepted: 29 January 2018 – Published: 16 March 2018

Abstract. The Stratosphere–troposphere Processes And their Role in Climate (SPARC) Quasi-Biennial Oscillation initiative (QBOi) aims to improve the fidelity of tropical stratospheric variability in general circulation and Earth system models by conducting coordinated numerical experiments and analysis. In the equatorial stratosphere, the QBO is the most conspicuous mode of variability. Five coordinated experiments have therefore been designed to (i) evaluate and compare the verisimilitude of modelled QBOs under present-day conditions, (ii) identify robustness (or alternatively the spread and uncertainty) in the simulated QBO response to commonly imposed changes in model climate forcings (e.g. a doubling of CO₂ amounts), and (iii) examine model dependence of QBO predictability. This paper documents these experiments and the recommended output diagnostics. The rationale behind the experimental design and choice of diagnostics is presented. To facilitate scientific interpretation of the results in other planned QBOi studies, consistent descriptions of the models performing each experiment set are given, with those aspects particularly relevant for simulating the QBO tabulated for easy comparison.

1 Introduction

Over the last decade or so, there has been a move toward global climate, Earth system, and weather forecasting models having properly resolved stratospheres and elevated upper boundaries. In some cases (e.g. Marsh et al., 2013) these boundaries are above 100 km and thus nominally located in space (as defined by the Fédération Aéronautique Internationale). Despite this, tropical stratospheric variability, in particular the Quasi-Biennial Oscillation (QBO), has generally been rather poorly represented (Butchart et al., 2011) in models used in recent international assessments of stratospheric ozone depletion (WMO, 2011, 2015). Likewise only a handful of the models central to the last international assessment of climate change (IPCC, 2013) simulated tropical variability approaching a realistic QBO (see Fig. 1). Even with the latest generation of models the representation of the QBO remains problematic in many cases (Schenzinger et al., 2017). For instance, several of the state-of-the-art chemistry–climate models participating in the concurrent Chemistry–Climate Model Initiative (CCMI) prescribe a QBO in order to “improve” the accuracy of their simulations (Morgenstern et al., 2017). Consequently the World Climate Research Programme (WCRP) Stratosphere–troposphere Processes And their Role in Climate (SPARC) core project has promoted a new QBO initiative (QBOi) to improve the simulation of tropical stratospheric variability in general circulation models and Earth system models (GCMs and ESMs). While QBOi is focused on modelling studies, it is also closely aligned with other SPARC activities including the SPARC Reanalysis Intercomparison Project (S-RIP; Fujiwara et al.,

2017) providing supporting analysis of observations and re-analyses and the SPARC gravity waves activity (Alexander and Sato, 2015) that is studying an important driver of the QBO.

Unlike the Coupled Model Intercomparison Project Phase 6 (CMIP6; Eyring et al., 2016), and to a lesser extent CCMI, the design of experiments for QBOi is not governed by the huge and rather diverse requirements from policy makers and scientists that have presented such a massive cultural and organizational challenge to the modelling community (Eyring et al., 2016). Instead, QBOi has adopted a less onerous approach for experimental set-up using stand-alone experiments (Sect. 3) specifically focused on improving the representation of the QBO in GCMs and addressing scientific questions related to advancing the understanding of the QBO (Sect. 2). This is an essential prerequisite to improving the representation in models of important QBO influences (Baldwin et al., 2001), such as the modulation of the transport of aerosols and chemical constituents into and within the stratosphere (e.g. Strahan et al., 2015) or the dynamical teleconnections to the extra-tropics (Anstey and Shepherd, 2014), and their subsequent surface climate and weather impacts. These aspects are expected to be included more prominently in the next phase of QBOi. The purpose of this paper is to describe the experiments to be used in phase 1 of QBOi and provide supporting documentation for other publications analysing and interpreting the output from the experiments. To help promote the widest possible participation in the experiments and thereby maximize the size of the multi-model ensembles, the design of the experiments has involved input from the community throughout (Anstey et al., 2015; Hamilton et al., 2015). The scientific rationale for the experiments also evolved through community discussion (Anstey et al., 2015) and is presented in the next section.

An important part of the multi-model analysis and interpretation of the experiments is the availability of a consistent set of relevant diagnostics from each model. For this QBOi follows best practices and, where possible, variable and file naming conventions of CMIP5 and CCMI (see the Supplement). In particular the recommended output diagnostics are adapted from those requested by the Dynamics and Variability Model Intercomparison Project (DynVarMIP; Gerber and Manzini, 2016a). These will allow the zonal-mean zonal momentum budgets to be examined in detail in the transformed Eulerian mean (TEM) framework (e.g. Andrews et al., 1987, pp. 127–130) including contributions from parameterized (sub-grid-scale) gravity waves. Other requested diagnostics are aimed at characterizing the sources, propagation, and filtering (i.e. breaking) of both resolved and unresolved waves in the participating models, particularly in the equatorial region. Precise specification of the requested diagnostics can be found in Sect. 4. To facilitate the comparison of these diagnostics among participating models, salient model features that are important for capturing QBO-like behaviour are described and tabulated in Sect. 5, with some emphasis in

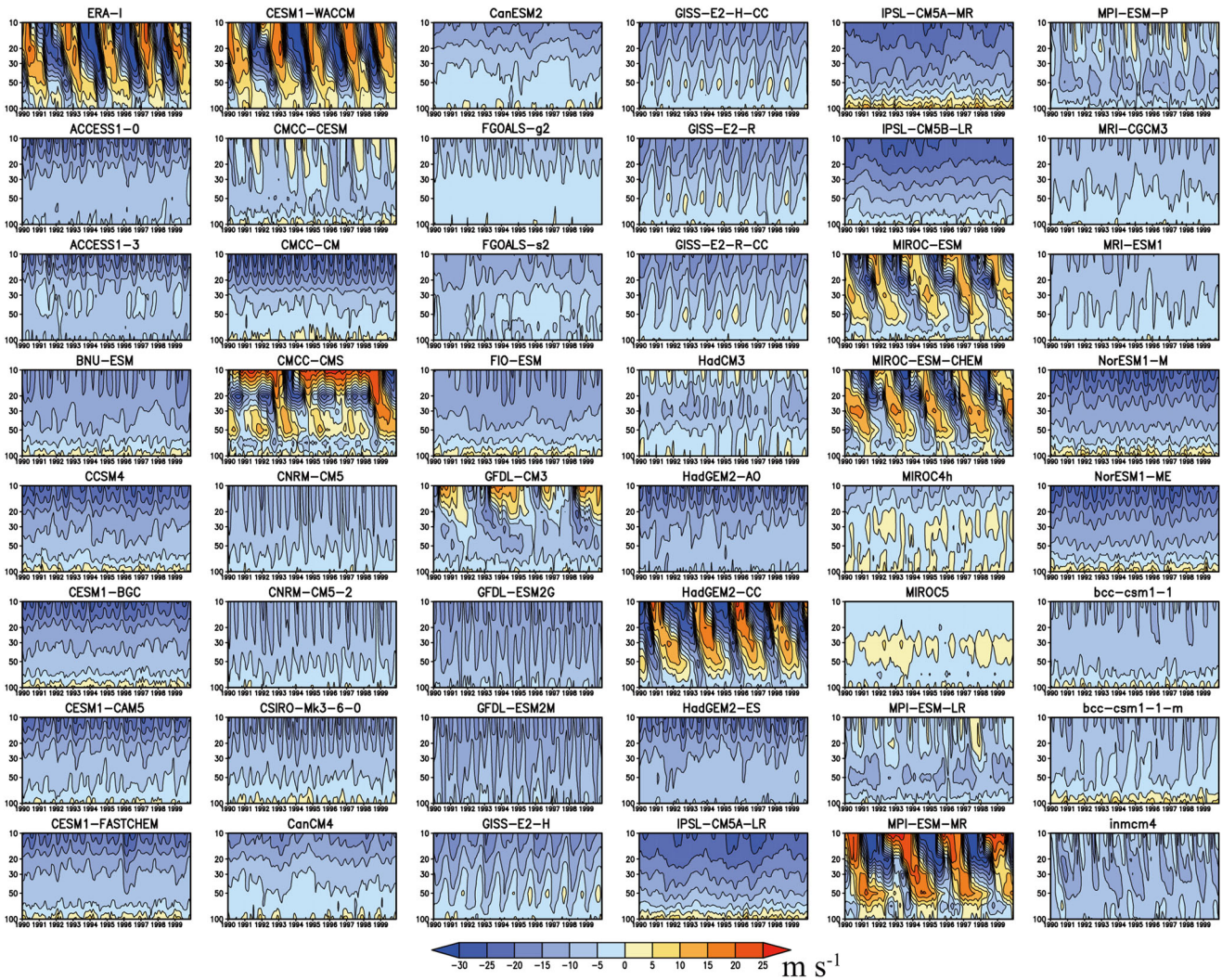


Figure 1. A 10-year (1990–1999) time series of monthly and zonal-mean zonal wind at the Equator from 100 to 10 hPa for 47 models that uploaded data to the Coupled Model Intercomparison Project Phase 5 (CMIP5) data repository. Only five models (CMCC-CMS, HadGEM2-CC, MIROC-ESM, MIROC-ESM-CHEM, and MPI-ESM-MR) spontaneously produce the iconic QBO behaviour of alternating descending layers of eastward and westward winds as indicated in the upper left-hand panel for the ERA-Interim reanalysis (Dee et al., 2011). Equatorial stratospheric winds in the CMIP5 version of CESM1(WACCM) are strongly relaxed (“nudged”) toward observations, which is why it shows a close resemblance to ERA-Interim in this figure. (Note that the version of WACCM participating in QBOi, described in Sect. 5, is a different version of this model.)

particular on the non-orographic gravity wave drag (GWD) parameterizations used by almost all of the QBOi models. Closing remarks including future plans follow in Sect. 6.

2 Scientific rationale

A crucial test of our understanding and ability to model the QBO occurred around the beginning of 2016 when the QBO cycle was unexpectedly disrupted for the first time since its discovery in the late 1950s (Dunkerton, 2016; Newman et al., 2016; Osprey et al., 2016; Coy et al., 2017). The well-established QBO paradigm, originating from the 1960s,

of alternate eastward and westward momentum deposition from vertically propagating equatorial waves (Baldwin et al., 2001) could not account for this disruption (Osprey et al., 2016). Despite the fact that the QBO is normally highly predictable (Pohlmann et al., 2013; Scaife et al., 2014) the disruption was completely missed by seasonal forecasts, and this failure illustrates the difficulty models have in capturing the complex phenomenology of the QBO and its full range of variability. Similar disruptions have only very rarely been seen in multi-decadal simulations and from just a few models with QBO-like oscillations (e.g. Osprey et al., 2016). It is possible that the models may be over-tuned to ensure that

they capture the mean behaviour of selected metrics (e.g. mean period and amplitude) of the present-day QBO. Furthermore, the disruption itself raises the possibility that the real QBO is less robust than previously thought, although it has since returned to its usual cycling as predicted.

With the advent of non-orographic GWD parameterizations and/or the use of increased vertical resolution in the stratosphere, a growing number of global models have been able to reproduce QBO-like variability in the equatorial stratosphere (e.g. Takahashi, 1996; Scaife et al., 2000; Hamilton et al., 2001; Giorgetta et al., 2002; Shibata and Deushi, 2005; Anstey et al., 2010; Kawatani et al., 2010; Orr et al., 2010; Lott and Guez, 2013; Richter et al., 2014; Rind et al., 2014; McCormack et al., 2015; Molod et al., 2015). However, common deficiencies exist in all current simulations, notably with QBO winds often being unrealistically weak in the lowermost stratosphere and having unrealistically small cycle-to-cycle variability (e.g. Schenzinger et al., 2017). The simulated QBOs can also be quite “fragile”, which means sensitive to many different aspects of model formulation depending on the model. For example, the QBO in the Canadian Middle Atmosphere Model (AGCM3-CMAM) is sensitive to the balance of resolved and parameterized wave forcing (Anstey et al., 2016), while in different versions of the Met Office Unified Model (MetUM) the QBO is sensitive to the specification of stratospheric ozone (Butchart et al., 2003; Bushell et al., 2010) and/or the parameterized gravity waves (Bushell et al., 2010; Kim et al., 2013). Sensitivity to vertical resolution has been reported by numerous studies, for example by Giorgetta et al. (2006) for the middle atmosphere version of the ECHAM5 (MAECHAM5) model and by Geller et al. (2016) for the NASA Goddard Institute for Space Studies (GISS) climate model. In addition Yao and Jablonowski (2015) identified a sensitivity to the choice of dynamical core. Other key questions concerning simulation of the QBO regard its possible synchronization with other modes of variability, such as the annual cycle (e.g. Rajendran et al., 2016) and El Niño–Southern Oscillation (e.g. Christiansen et al., 2016), the QBO’s predictability (e.g. Pohlmann et al., 2013; Scaife et al., 2014), and finally the robustness of the QBO response to climate change (e.g. Kawatani and Hamilton, 2013; Schirber et al., 2015).

Phase 1 of QBOi focuses on reducing these uncertainties in simulated QBOs by conducting coordinated experiments that will allow for more rigorous intercomparison of models than is otherwise possible from individual studies. The aim is to address the ability of GCMs to capture the QBO in the present climate, to predict its behaviour under climate-change forcings, and to predict its evolution when initialized with observations (i.e. hindcasts).

3 Experiments

Anstey et al. (2015) and Hamilton et al. (2015) briefly describe a set of five QBO experiments which are designed to be simple and accessible to a wide range of modelling groups. The motivation and specific goals for each of these experiments is presented below with the technical specifications given in Appendix A. The aim is for modelling groups to perform all five experiments, and even if this is not possible, it is important that the same model version is used for the subset of experiments that are conducted; i.e. there should be no tuning of free parameters between experiments. Use of the same model version for the different experiments is crucial for learning the most from this study. The model version used should be that which the group considered gave the “best” representation of the QBO under present-day conditions (e.g. in Experiment 1 or similar preparatory simulations). Of course there are situations when two different versions of a model might be used to perform the experiment set, such as when high- and low-resolution versions or alternative non-orographic GWD parameterizations are available. In these situations the results would then be treated for the purpose of the QBOi multi-model analysis as if they were obtained from two separate models (although interpretation of results will need to be aware of and test for sensitivity to the possible dominance of the results by one particular family of models). All experiments are for AGCMs apart from an option to perform Experiment 5 with a coupled ocean, which is denoted as Experiment 5a (see below).

3.1 Experiment list and goals

3.1.1 Present-day climate

The first two experiments are designed with the goal of identifying and distinguishing the properties of and mechanisms underlying the variety of model simulations of the QBO in present-day conditions:

- Experiment 1 (“AMIP”) involves specified observed inter-annually varying sea surface temperatures (SSTs), sea ice, and external forcings for 1 January 1979 to 28 February 2009 (one- to three-member ensemble); and
- Experiment 2 (present-day time slice) is identical to Experiment 1 except employing repeated annual cycle SSTs, sea ice, and external forcings (100 years or ensemble of 3×30 years).

The main differences between these two experiments are expected to arise from the differences between their specified SSTs. Figure 2 compares the variability in the tropics (5° N– 5° S) of the prescribed SSTs for Experiments 1 and 2. Averaged over all longitudes the differences are relatively small (Fig. 2a), although regionally there are large differences, for

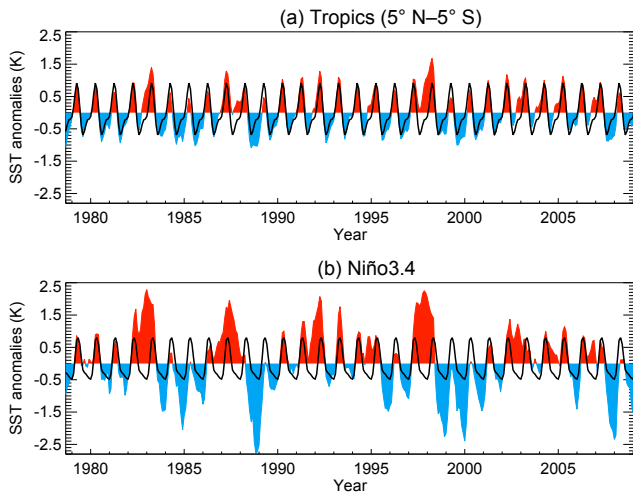


Figure 2. Comparison of monthly mean sea surface temperature (SST) anomalies (red and blue) from the 30-year mean (1979–2008) of the CMIP5 AMIP SSTs used in Experiment 1 with the mean annual cycle (black curve) for the same period. **(a)** Average for all longitudes between 5° N and 5° S. **(b)** Average for the Niño 3.4 region (120°–170° W, 5° N–5° S).

instance due to the effects of the El Niño–Southern Oscillation (Fig. 2b).

These experiments will allow for an evaluation of the accuracy of modelled QBOs under present-day climate conditions by employing the diagnostics and metrics discussed in Sect. 4. The impact of inter-annually varying forcing (e.g. Fig. 2) on the model QBO will be assessed through a comparison of the two experiments. Experiment 2 also provides the control for the climate projection in Experiments 3 and 4.

3.1.2 Climate projections

Two further experiments are designed to subject the modelled QBOs (i.e. the QBO simulated by the present-day experiments) to an external forcing similar to that typically applied for climate projections:

- Experiment 3 ($2\times\text{CO}_2$ time slice) is identical to Experiment 2, but with a change in CO_2 concentration and specified SSTs appropriate for a $2\times\text{CO}_2$ world (100 years or ensemble of 3×30 years); and
- Experiment 4 ($4\times\text{CO}_2$ time slice) is identical to Experiment 2, but with a change in CO_2 concentration and specified SSTs appropriate for a $4\times\text{CO}_2$ world (100 years or ensemble of 3×30 years).

These experiments will allow the response (i.e. $2\times\text{CO}_2$ – $1\times\text{CO}_2$ and $4\times\text{CO}_2$ – $1\times\text{CO}_2$) of the QBO, its forcing mechanisms, and its impact and influence to be evaluated using the same diagnostics and metrics used in the analysis of Experiments 1 and 2. Key questions that will be addressed are the following.

- What is the spread and uncertainty of the forced model response?
- Do different models cluster in any particular way?
- Can a connection or correlation be made between QBOs that exhibit similar values of metrics and diagnostics under present-day climate forcing and the behaviour of the QBO in these same models under future climate forcing?

The motivation is to investigate what aspects of modelled QBOs determine the spread, or uncertainty, of the QBO response to CO_2 forcing. These aspects are considered high priority by QBOi in order to reduce uncertainty in future projections. These experiments also will provide context for the uncertainty in climate change projections of QBO behaviour among the state-of-the-art GCMs being used in CMIP6. Furthermore, the possibility was noted in Sect. 2 that some models may be over-tuned to ensure that they capture the behaviour of the present-day QBO. If so, then a large multi-model spread in the forced response may indicate that such tuning constitutes, in effect, an “overfitting” of models to present-day conditions.

3.1.3 QBO hindcasts

The goal of the final experiment is to evaluate and compare the predictive skill of modelled QBOs in a retrospective hindcast context, quantify this predictive capability in multiple models, and study the model processes driving the evolution of the QBO.

- Experiment 5 (hindcasts) involves a set of initialized QBO hindcasts of 9–12 months using the observed SSTs and forcings specified as in Experiment 1. Specified start dates are 1 May and 1 November for the years 1993–2007 (i.e. 15 years, 30 start dates) with initial atmospheric conditions obtained from reanalyses (at least three-member ensemble).

Because of the prescribed SSTs these are not true prediction experiments; nonetheless they provide an important test of how well models can predict the evolution of the QBO from specified initial conditions that reasonably sample the full range of QBO phases, despite some clustering of the 1 May initial profiles (Fig. 3). Key questions that will be addressed are the following.

- How does prediction skill vary among models, and to what extent and for how long are models able to predict the QBO evolution correctly at different vertical levels and different phases of the QBO?
- How does the forecast skill relate to the behaviour of the QBO in Experiment 1? Are realistic QBO simulations in a multi-decadal simulation well correlated with skillful long-term deterministic predictions?

- Do the models that cluster and/or do well in the prediction experiments cluster in the CO₂ forcing experiments?

One aim is to investigate which aspects of modelled QBOs determine the quality of QBO prediction and therefore where development needs to be focused for model improvement. The hindcast framework can also be helpful for directly assessing model changes, possibly driving improvements in free-running models. A further motivation for these experiments is to investigate the possibility of using the hindcast results to narrow the range of plausible models for climate change experiments.

It is recognized that some groups may already have completed for the period 1993–2007 operational seasonal hindcasts using a coupled ocean–atmosphere model, and therefore for the QBOi multi-model analysis an acceptable alternative (or addition) to Experiment 5 is

- Experiment 5a (hindcasts), which involves a set of initialized QBO hindcasts of 9–12 months identical to Experiment 5 apart from replacing the specified SSTs with a coupled ocean model appropriately initialized (at least three-member ensemble).

Full comparison with the other models providing Experiment 5 output will nonetheless depend on most of the diagnostics discussed in Sect. 5 being available from those groups providing Experiment 5a output.

3.1.4 Process studies

A secondary purpose of Experiment 5 is to investigate and evaluate differences in wave dissipation and momentum deposition to understand the processes driving the QBO in each model and separate the contributions from resolved and unresolved waves (e.g. Scaife et al., 2000; Shibata and Deushi, 2005). Due to the initialization of the hindcasts, each model will have essentially the same initial basic state, and its evolution immediately after the start of the forecast will allow the properties of wave dissipation and momentum deposition to be compared and contrasted between different models given a near-identical basic state. Specifying the same observed SST in all models (rather than allowing each model to predict its own SST evolution) facilitates the comparison as it eliminates any differences resulting from the evolving ocean. Short periods of additional high-frequency diagnostics are requested to maximize the benefits of the multi-model comparison.

4 Diagnostics

The diagnostics requested by QBOi draw on those requested by other major multi-model intercomparison projects, in particular DynVarMIP (Gerber and Manzini, 2016a), though

they have been specifically tailored through community discussion for the analysis of the QBO in Experiments 1–5. The requested diagnostics are described in this section; additional technical information on how they should be formatted and uploaded to the shared QBOi repository is available in the Supplement.

4.1 Spatial and temporal resolution

For ease of comparison among models most output variables are requested on a standard set of 30 pressure levels: 1000, 925, 850, 700, 600, 500, 400, 300, 250, 200, 175, 150, 120, 100, 85, 70, 60, 50, 40, 30, 20, 15, 10, 7, 5, 3, 2, 1.5, 1.0, and 0.4 hPa. These are adapted from the extended levels set requested by DynVarMIP for CMIP6 (e.g. Gerber and Manzini, 2016a) to obtain a vertical resolution in the upper tropical troposphere and lower stratosphere (i.e. between 200 and 40 hPa) of 1.0 to 1.5 km. There are two exceptions, however.

- Data to be used for calculating equatorial wave spectra (6-hourly instantaneous fields) should be provided at a vertical resolution equivalent to the model resolution (i.e. with the same number of levels in the specified altitude range) to ensure accurate calculation of QBO wave forcing (e.g. Kim and Chun, 2015a); see below for further details.
- To reduce data volume, daily mean three-dimensional (3-D) variables are requested for only the eight pressure levels used by CMIP5: 1000, 850, 700, 500, 250, 100, 50, and 10 hPa. These data will be used mainly to examine the QBO influence on other regions of the atmosphere (e.g. on the North Atlantic Oscillation, NAO) and higher vertical resolution is not considered necessary.

Horizontal resolution should be the same as the model but if data volume is an issue then a reduced grid is acceptable, provided the reduction method is documented.

To examine the daily mean and monthly mean QBO zonal mean momentum budget, terms making up the TEM zonal momentum equation (e.g. Andrews et al., 1987, pp. 127–130) are requested following the recipe given by Gerber and Manzini (2016a, Appendix A3), but also see their corrigendum (Gerber and Manzini, 2016b). In particular note the importance of calculating the individual terms from 6-hourly or higher-frequency data (e.g. every time step) and the need for sufficient vertical resolution (e.g. the standard pressure levels listed above) for accurate estimates of the vertical derivatives. Furthermore to examine the wavenumber–frequency spectra of the equatorial waves (e.g. Horinouchi et al., 2003; Lott et al., 2014) instantaneous values of 3-D winds and temperature are requested every 6 h on model levels or on pressure levels at roughly equivalent vertical resolution to the model levels but, to reduce data volumes, only for levels between

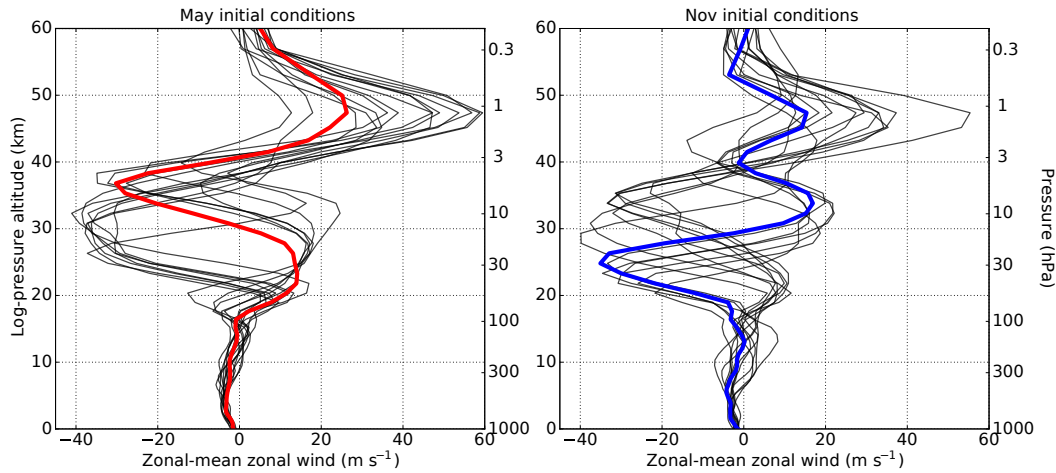


Figure 3. Zonal-mean and daily mean zonal wind (m s^{-1}) profiles at the Equator for 1 May and 1 November for the 15 years 1993–2007 from the ERA-Interim reanalyses (Dee et al., 2011). The two profiles shown in coloured lines (May 1993 and November 2005, taken as representative of eastward and westward QBO phases in the lower stratosphere, respectively) are those used in offline comparison of the gravity wave drag parameterizations presented in Sect. 5.1.

Table 1. Climate and variability. Monthly and daily means, with 2-D indicating a longitude–latitude–time (XYT) field and 3-D indicating a longitude–latitude–pressure–time (XYPT) field. XY is typically the model’s horizontal output grid and P is the standard 30-level set of diagnostic pressure levels described in Sect. 4.1: 1000, 925, 850, 700, 600, 500, 400, 300, 250, 200, 175, 150, 120, 100, 85, 70, 60, 50, 40, 30, 20, 15, 10, 7, 5, 3, 2, 1.5, 1.0, and 0.4 hPa.

Name	Long name (units)	Dimension
psl	sea level pressure (Pa)	2-D
prc	convective precipitation flux ($\text{kg s}^{-1} \text{m}^{-2}$)	2-D
pr	total precipitation flux ($\text{kg s}^{-1} \text{m}^{-2}$)	2-D
tas	near-surface air temperature (K)	2-D
uas	eastward near-surface wind (m s^{-1})	2-D
vas	northward near-surface wind (m s^{-1})	2-D
ta	air temperature (K)	3-D*
ua	eastward wind (m s^{-1})	3-D*
zg	geopotential height (m)	3-D*

* For daily 3-D variables P is reduced to eight pressure levels: 1000, 850, 700, 500, 250, 100, 50, 10 hPa.

100 and 0.4 hPa and for latitudes between 15°N and 15°S . For ease of analysis, pressure levels at model-level resolution are preferred over actual model levels.

An additional benefit of requesting these 6-hourly data is that they can provide a check on the requested TEM budget terms (Table 2), albeit only for tropical latitudes. Calculating the budget terms from the 3-D 6-hourly wind and temperature fields (Table 4) in a self-consistent way across all models removes the possibility that some of the inter-model differences in the requested TEM diagnostics (Table 2) are due to differences in how the calculation of these terms was performed by different modelling groups.

4.2 Output period

Monthly mean output is requested for the full duration of all experiments and all ensemble members. Likewise for Experiment 5 daily mean output is requested for the full duration of each ensemble member. On the other hand, for Experiments 1–4 daily mean output is only requested for the first 30 years and/or the first ensemble member.

High-frequency (6-hourly) diagnostics for calculating equatorial wave spectra are requested for the following periods and ensemble members for each experiment:

- Experiment 1, 1997–2002 (note that this period encompasses positive, negative, and neutral El Niño–Southern Oscillation (ENSO) phases) of first ensemble member;
- Experiments 2–4, years 1–4 of first ensemble member; and
- Experiment 5, first 3 months of all ensemble members.

4.3 Requested output variables

Similarly to DynVarMIP (Gerber and Manzini, 2016a), the requested variables are separated into three categories: standard variables (Table 1) for diagnosing the climate and variability in the models, dynamical variables (Table 2) for analysing momentum transport and budgets, and thermodynamic quantities (Table 3). In addition a fourth category of variables (Table 4) will enable the equatorial wave spectra (e.g. Horinouchi et al., 2003; Lott et al., 2014) to be compared among the models.

Table 2. Dynamics. (a) Monthly mean and daily mean fields and contributions to zonal-mean zonal momentum equation (YPT). (b) Monthly mean tendencies and fluxes from parameterized gravity waves (XYPT). (c) Daily mean sources for orographic and non-orographic gravity waves (XYT). P is the standard 30-level set of diagnostic pressure levels described in Sect. 4.1 (also Table 1 caption).

(a) Monthly mean and daily mean zonal mean fields – YPT		
Name	Long name (units)	Dimension
ua	eastward wind (m s^{-1})	2-D
ta	air temperature (K)	2-D
zg	geopotential height (m)	2-D
vstar	residual northward wind (m s^{-1})	2-D
wstar	residual upward wind (m s^{-1})	2-D
fy	northward EP flux (N m^{-1})	2-D
fz	upward EP flux (N m^{-1})	2-D
utenddivf	u tendency by EP flux divergence (m s^{-2})	2-D
utend	u tendency (m s^{-2})	2-D
utendogw	u tendency by orographic gravity waves (m s^{-2})	2-D
utendnogw	u tendency by non-orographic gravity waves (m s^{-2})	2-D
psistar	residual stream function (kg s^{-1})	2-D
(b) Monthly mean gravity wave tendencies and fluxes – XYPT		
utendogw	u tendency by orographic gravity waves (m s^{-2})	3-D
utendnogw	u tendency by non-orographic gravity waves (m s^{-2})	3-D
vtendogw	v tendency by orographic gravity waves (m s^{-2})	3-D
vtendnogw	v tendency by non-orographic gravity waves (m s^{-2})	3-D
taunoge	eastward wind stress of non-orographic gravity waves (Pa)	3-D
taunogs	southward wind stress of non-orographic gravity waves (Pa)	3-D
taunogw	westward wind stress of non-orographic gravity waves (Pa)	3-D
taunogn	northward wind stress of non-orographic gravity waves (Pa)	3-D
(c) Daily mean gravity wave sources – XYT		
taougu	surface eastward wind stress by orographic gravity waves (Pa)	2-D
taougw	surface northward wind stress by orographic gravity waves (Pa)	2-D
taunoge*	launch eastward wind stress of non-orographic gravity waves (Pa)	3-D
taunogs*	launch southward wind stress of non-orographic gravity waves (Pa)	3-D
taunogw*	launch westward wind stress of non-orographic gravity waves (Pa)	3-D
taunogn*	launch northward wind stress of non-orographic gravity waves (Pa)	3-D

* Only if non-isotropic and/or non-stationary at launch level (e.g. coupled to convection or fronts).

Table 3. Thermodynamics. Monthly mean and daily mean zonal mean fields (YPT). P is the standard 30-level set of diagnostic pressure levels described in Sect. 4.1 (also Table 1 caption).

Monthly mean and daily mean zonal mean fields – YPT		
Name	Long name (units)	Dimension
hus	specific humidity (kg kg^{-1})	2-D
zmtnt	diabatic heating rate (K s^{-1})	2-D
tntlw	longwave heating rate (K s^{-1})	2-D
tntsw	shortwave heating rate (K s^{-1})	2-D
o3*	mole fraction of ozone in air (mole mole^{-1})	2-D

* Only if model has prognostic ozone.

5 Participating models

All the experiments for phase 1 of QBOi have been designed for atmosphere-only GCMs. From the experiment descriptions in Sect. 3 it is also clear that for an AGCM to participate in these experiments it must be configured with a number of essential characteristics (e.g. land–ocean contrast, annual cycle, and a radiation scheme that can accommodate changes in CO_2 amounts). Apart from this QBOi does not impose any restrictions on the representation in participating models of any physical process or, indeed, chemical process for those models with interactive ozone. Of course, participating models are expected to properly resolve the stratosphere with an average vertical resolution of the order of 2 km or less between 100 and 1 hPa and an upper boundary somewhere above that (cf. high- and low-top results in Osprey et

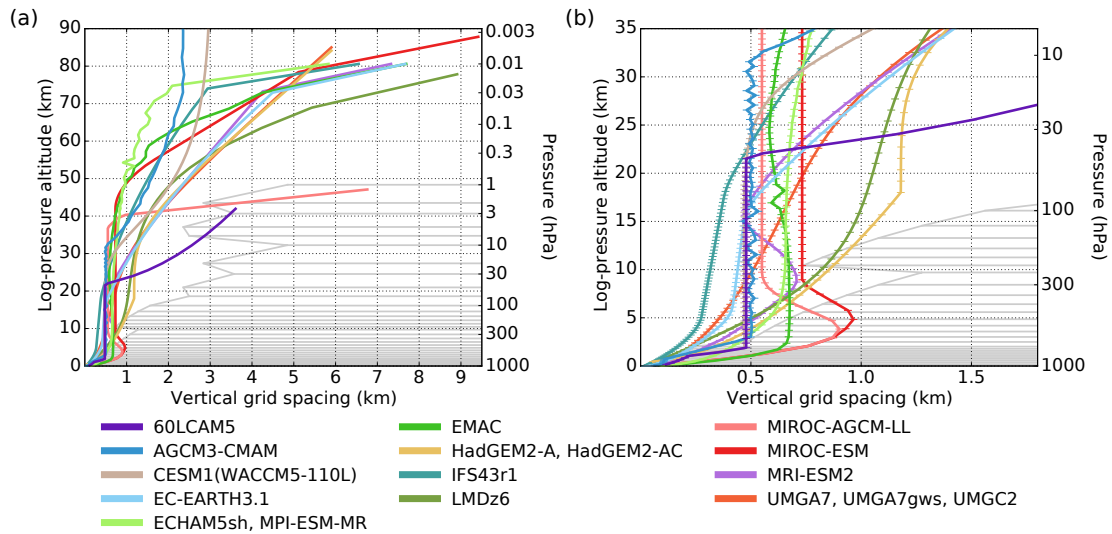


Figure 4. (a) Vertical profiles of vertical grid spacing, Δz (km), for models participating in QBOi. Log-pressure altitude on the model levels is calculated by assuming a surface pressure of 1013.25 hPa and a fixed scale height of 7 km. The grey horizontal lines denote the set of 30 QBOi diagnostic pressure levels (see Sect. 4.1), while the grey vertical profile (left end of the grey horizontal lines) indicates the Δz of the diagnostics. (b) Same as in (a), but zoomed in to the altitude range most relevant for the QBO. Thin horizontal lines in (b) indicate the locations of the model levels.

al., 2013). However, it is not strictly necessary for a model to display QBO-like variability in the equatorial stratosphere as additional insight can be gained by comparing models with and without this property. Models with QBO-like variability but without a properly resolved stratosphere (e.g. with upper boundary below 1 hPa) are also considered since, again, this potentially provides guidance on the level of stratospheric detail that is required in order to reproduce a QBO.

There are 17 models or model versions participating in phase 1 of QBOi (i.e. data from 17 models have been uploaded or are planned for upload to the shared QBOi repository; see the Supplement for details of this repository). These models are listed in Table 5 along with the institutes and investigators using the models and their contact information. The model names given refer to the names used in the repository, while the information given in Tables 6–8 refers specifically to the configuration and parameter settings used by each model when producing the uploaded data. More comprehensive descriptions of the individual models can be found in the references given in the last column of Table 5.

It should be noted that common model development history can lead to a lack of full independence among models. For example, 60LCAM5 and CESM1(WACCM5-110L) have developed from the NCAR Community Atmosphere Model (CAM); HadGEM2-A, HadGEM2-AC, UMGA7, UMGA7gws, and UMGC2 have developed out the Met Office Unified Model (UM); EC-EARTH3.1 and IFS have their origins in the ECMWF Integrated Forecasting System (IFS); MIROC-AGCM and MIROC-ESM belong to the family of MIROC models; and ECHAM5sh, EMAC, and MPI-ESM-

Table 4. Equatorial wave spectra. The 6-hourly instantaneous 3-D (XYPT) equatorial field (15° N to 15° S) output for selected sub-periods of each experiment (see Sect. 4.2). Here P is *not* the standard set of pressure levels used in Tables 1–3. Rather, as described in Sect. 4.1, it indicates a set of pressure levels with equivalent vertical resolution to the model levels, covering the altitude range 100 to 0.4 hPa. Alternatively the data can be provided on actual model levels, although in this case the data required for conversion between model and pressure levels must also be provided.

The 6-hourly equatorial fields – XYPT		
Name	Long name (units)	Dimension
ta	air temperature (K)	3-D
ua	eastward wind (m s^{-1})	3-D
va	northward wind (m s^{-1})	3-D
wa	vertical wind (m s^{-1})	3-D

MR all originate from the MPI ECHAM line of model development. Tables 6–8 indicate that some model components are shared by different models. The extent to which shared development history affects model independence can be difficult to assess and varies among models (e.g. Knutti et al., 2013). Apart from describing those aspects of model formulation that are expected to be relevant to the QBO (Tables 6–8), detailed consideration of model independence is outside the scope of this paper. However, note that out of the 17 QBOi models, there are two pairs of models that are identical in all respects but one: HadGEM2-A and UMGA7 used fixed sources for their non-orographic gravity wave pa-

Table 5. Participating models and contact information.

Model names	Expts.	Institutes	Investigators	Email address	References
60LCAM5	1–4	NCAR	J. Chen J. Richter	cchen@ucar.edu jrichter@ucar.edu	Richter et al. (2014)
AGCM3-CMAM	1–3, 5	CCCma U. Toronto	J. Anstey J. Scinocca C. McLandress	james.anstey@canada.ca john.scinocca@canada.ca charles@atmos.physics.utoronto.ca	Scinocca et al. (2008) Anstey et al. (2016)
CESM1- (WACCM5-110L)	1–4	NCAR	R. Garcia J. Richter	rgarcia@ucar.edu jrichter@ucar.edu	Garcia and Richter (2017)
EC-EARTH3.1	5	BSC	J. Garcia-Serrano	javier.garcia@bsc.es	Christiansen et al. (2016)
ECHAM5sh	1–4	ISAC-CNR	F. Serva C. Cagnazzo	federico.serva@artov.isac.cnr.it c.cagnazzo@isac.cnr.it	Serva et al. (2018) Manzini et al. (2012)
EMAC	1–4	KIT	P. Braesicke T. Kerzenmacher S. Versick	peter.braesicke@kit.edu tobias.kerzenmacher@kit.edu stefan.versick@kit.edu	Jöckel et al. (2005) Jöckel et al. (2010)
HadGEM2-A	1	Ewha W. U. Yonsei U.	Y.-H. Kim H.-Y. Chun	young-ha.kim@ewha.ac.kr chunhy@yonsei.ac.kr	The HadGEM2 Development Team (2011)
HadGEM2-AC	1	Ewha W. U. Yonsei U.	Y.-H. Kim H.-Y. Chun	young-ha.kim@ewha.ac.kr chunhy@yonsei.ac.kr	The HadGEM2 Development Team (2011) Kim and Chun (2015b)
IFS43r1	1–5	ECMWF	T. Stockdale	tim.stockdale@ecmwf.int	ECMWF (2016); Orr et al. (2010)
LMDz6	1–4	IPSL-LMD	F. Lott	flott@lmd.ens.fr	Lott et al. (2005, 2012)
MIROC-AGCM-LL	1–5	MIROC	Y. Kawatani	yoskawatani@jamstec.go.jp	Kawatani et al. (2011)
MIROC-ESM	1–5	MIROC	S. Watanabe	wnabe@jamstec.go.jp	Watanabe et al. (2011)
MPI-ESM-MR	5a	MPI U. Hamburg	H. Pohlmann M. Dobrynin	holger.pohlmann@mpimet.mpg.de mikhail.dobrynin@uni-hamburg.de	Pohlmann et al. (2013) Dobrynin et al. (2018)
MRI-ESM2	1–5	MRI-JMA	K. Yoshida H. Naoe S. Yukimoto	kyoshida@mri-jma.go.jp hnaoe@mri-jma.go.jp yukimoto@mri-jma.go.jp	Adachi et al. (2013) Yukimoto et al. (2012)
UMGA7	1–4	Met Office MOHC U. Oxford	A. Bushell N. Butchart S. Osprey	andrew.bushell@metoffice.gov.uk neal.butchart@metoffice.gov.uk scott.osprey@physics.ox.ac.uk	Walters et al. (2017b)
UMGA7gws	1–4	Met Office MOHC U. Oxford	A. Bushell N. Butchart S. Osprey	andrew.bushell@metoffice.gov.uk neal.butchart@metoffice.gov.uk scott.osprey@physics.ox.ac.uk	Bushell et al. (2015) Walters et al. (2017b)
UMGC2	5a	MOHC	A. Scaife M. Andrews	adam.scaife@metoffice.gov.uk martin.andrews@metoffice.gov.uk	Dunstone et al. (2016)

parameterizations, while their counterparts HadGEM2-AC and UMGA7gws, respectively, use parameterized gravity wave sources; this distinction is described in more detail below.

Properties of the models (Tables 6–8) that are of particular relevance for simulating a QBO are the following.

- *Vertical domain and resolution.* A high upper boundary is potentially important depending on how much influence the semi-annual oscillation has on the timing of the start of each new descending QBO cycle. Likewise vertical resolution is important both for accurately simulating vertically propagating equatorial waves and for representing the wave dissipation and descending sharp shear zones that are a characteristic feature of the QBO. Figure 4 (see also columns 3 and 4 of Table 6) shows the different vertical resolutions used by the QBOi participating models (although note that a small number of models share common vertical grids), along with the vertical resolution of the set of 30 diagnostic pressure levels described in Sect. 4.1.
- *Horizontal resolution.* This is likely to have a significant impact on the development and evolution of wave sources in the tropical troposphere, which are important for forcing the QBO. Horizontal resolution may also affect the propagation and breaking of large-scale Rossby waves propagating from the extra-tropics, which are now known to affect the QBO (e.g. Osprey et al., 2016). Figure 5 (see also column 2 of Table 6) shows the horizontal resolution of each model and how the differences

Table 6. Model domain and resolution.

Model name	Horizontal resolution	No. of levels	Upper boundary	Dynamical time step
60LCAM5	100 km	60	2.5 hPa (41 km)	30 min
AGCM3-CMAM	T47	113	0.00074 hPa (98 km)	7.5 min
CESM1(WACCM5-110L)	$1.25^\circ \times 0.94^\circ$	110	6.1×10^{-6} hPa (132 km)	30 min
EC-EARTH3.1	T255	91	0.01 hPa (80 km)	45 min
ECHAM5sh	T63	95	0.01 hPa (80 km)	7.5 min
EMAC	T42	90	0.01 hPa (80 km)	12 min
HadGEM2-A	$1.875^\circ \times 1.25^\circ$	60	0.006 hPa (84 km)	20 min
HadGEM2-AC	$1.875^\circ \times 1.25^\circ$	60	0.006 hPa (84 km)	20 min
IFS43r1	T255	137	0.01 hPa (80 km)	30 min
LMDz6	$2.5^\circ \times 1.25^\circ$	79	0.015 hPa (77 km)	3 min
MIROC-AGCM-LL	T106	72	1.2 hPa (47 km)	5 min
MIROC-ESM	T42	80	0.0036 hPa (87 km)	30 min
MPI-ESM-MR	T63	95	0.01 hPa (80 km)	7.5 min
MRI-ESM2	T159	80	0.01 hPa (80 km)	30 min
UMGA7	$1.875^\circ \times 1.25^\circ$	85	0.0053 hPa (85 km)	20 min
UMGA7gws	$1.875^\circ \times 1.25^\circ$	85	0.0053 hPa (85 km)	20 min
UMGC2	$0.833^\circ \times 0.556^\circ$	85	0.0053 hPa (85 km)	15 min

For spectral models the horizontal resolution is given in terms of triangular truncation of spectral coefficients, from which a grid spacing can be estimated as described in the Fig. 5 caption. For example, T63 $\sim 2.8^\circ \times 2.8^\circ$, T159 $\sim 1.125^\circ \times 1.125^\circ$, and T255 $\sim 0.7^\circ \times 0.7^\circ$, corresponding roughly to grid lengths of 310, 130, and 80 km, respectively. Upper boundary altitude is given in terms of pressure and log-pressure altitude as described in the Fig. 4 caption.

Table 7. Dynamical core.

Model name	Advection	Horizontal diffusion	Vertical diffusion	Sponge layer
60LCAM5	SE	fourth order	no	no
AGCM3-CMAM	EUL/ST	Leith (1971)	explicit	non-zonal above 80 km
CESM1(WACCM5-110L)	FV	implicit ¹	implicit	no
EC-EARTH3.1	SL/SI	fourth order ²	Richardson no. dependent ³	above 10 hPa ⁴
ECHAM5sh	EUL/ST	fourth order	eddy diffusion ⁵	uppermost level
EMAC	EUL/ST	fourth order	eddy diffusion ⁵	uppermost level
HadGEM2-A	SL/SI	no ⁶	no	no
HadGEM2-AC	SL/SI	no ⁶	no	no
IFS43r1	SL/SI	fourth order	Richardson no. dependent ⁷	above 10 hPa ⁴
LMDz6	EUL	fourth order	no	no
MIROC-AGCM-LL	EUL/ST	fourth order	eddy diffusion ⁸	five uppermost levels ⁹
MIROC-ESM	EUL/ST	sixth order	eddy diffusion ⁸	two uppermost levels
MPI-ESM-MR	EUL/ST	fourth order	eddy diffusion ⁵	uppermost level
MRI-ESM2	SL/SI	fourth order	explicit	no
UMGA7	SL/SI	no ⁶	no	above ~ 40 km ¹⁰
UMGA7gws	SL/SI	no ⁶	no	above ~ 40 km ¹⁰
UMGC2	SL/SI	no ⁶	no	above ~ 40 km ¹⁰

EUL: Eulerian, EUL/ST: Eulerian/spectral transform, FV: finite volume, SE: spectral element, SL/SI: semi-Lagrangian/semi-implicit.

¹ CESM1(WACCM5-110L) momentum equations also include fourth-order divergence damping (Lauritzen et al., 2011). ² Second order in uppermost levels.

³ ECMWF (2010). ⁴ Increasing with height with a more powerful mesospheric sponge being added above 1 hPa. ⁵ The vertical turbulent fluxes are related to the gradient of the respective variable (Roeckner et al., 2003, chap. 5). ⁶ All the UM family include some implicit damping from the SL/SI advection though this is much reduced in UMGA7, UMGA7gws, and UMG2 (Walters et al., 2017a). ⁷ A short-tail formulation was used in the lower stratosphere to reduce vertical diffusion compared to that given in ECMWF (2016) for the standard IFS configuration. ⁸ Using the level-2 closure scheme of Mellor and Yamada (1982). ⁹ Rayleigh friction was applied at five levels above 5 hPa and the strength of the fourth-order horizontal diffusion was successively doubled in this layer (Kawatani et al., 2011). ¹⁰ Weak Rayleigh damping applied only to vertical velocity.

Table 8. Non-orographic gravity waves, convection, and ozone chemistry.

Model name	Non-orographic gravity waves	Non-orographic GW source	Convection	Ozone chemistry
60LCAM5	Lindzen (1981) ^b	Richter et al. (2010)	Zhang and McFarlane (1995)	no
AGCM3-CMAM	Scinocca (2003) ^a [WM]	fixed	Zhang and McFarlane (1995)	no
CESM1(WACCM5-110L)	Lindzen (1981) ^b	Richter et al. (2010)	Zhang and McFarlane (1995)	no
EC-EARTH3.1	Scinocca (2003) ^a [WM]	fixed	Davini et al. (2017)	no
ECHAM5sh	Hines (1997a, b) ^a [H]	fixed	Tiedtke (1989), Nordeng (1994)	no
EMAC	Hines (1997a, b) ^a [H]	fixed	Tiedtke (1989)	no
HadGEM2-A	Warner and McIntyre (1999) ^a [WM]	fixed	Gregory et al. (1990)	no
HadGEM2-AC	Warner and McIntyre (1999) ^a [WM]	fixed	Gregory et al. (1990)	no
	Warner and McIntyre (1999) ^a [WM]	Choi and Chun (2011)		
IFS43r1	Scinocca (2003) ^a [WM]	fixed	Bechtold et al. (2008)	yes
LMDz6	Lott et al. (2012) ^b [L]	de la Cámara and Lott (2015), Lott and Guez (2013)	Emanuel (1991), Hourdin et al. (2013)	no
MIROC-AGCM-LL	none	—	Emori et al. (2001)	no
MIROC-ESM	Hines (1997a, b) ^a [H]	fixed	Emori et al. (2001)	no
MPI-ESM-MR	Hines (1997a, b) ^a [H]	fixed	Tiedtke (1989), Nordeng (1994)	no
MRI-ESM2	Hines (1997a, b) ^a [H]	fixed	Yoshimura et al. (2015)	yes
UMGA7	Warner and McIntyre (1999) ^a [WM]	fixed	Gregory et al. (1990)	no
UMGA7gws	Warner and McIntyre (1999) ^a [WM]	Bushell et al. (2015)	Gregory et al. (1990)	no
UMGC2	Warner and McIntyre (1999) ^a [WM]	fixed	Gregory et al. (1990)	no

Schemes marked ^a are non-orographic GWD parameterizations based on a wave-spectrum approach, while in schemes marked ^b the wave spectrum is treated as a collection of monochromatic waves. For the models using the Warner and McIntyre (1999) scheme (HadGEM2-A, HadGEM2-AC, UMGA7, UMGA7gws, and UMGC2), the use of the scheme to generate a QBO is described in Scaife et al. (2002). For IFS43r1, the use of the Scinocca (2003) scheme to generate a QBO is described in Orr et al. (2010). For MPI-ESM-MR, the use of the Hines (1997a, b) scheme is described in Schmidt et al. (2013). The abbreviation in square brackets for each scheme (second column; [WM], [H], or [L]) denotes the type of dissipation used in the scheme as labelled in Figure 7. “Fixed” in column 3 refers to sources of parameterized gravity waves that are not linked to any other model physical variable (see footnote 1, Sect. 5). Note, however, that “fixed” includes sources that vary in time and/or space in a prescribed way and stochastically (e.g. as is done in the ECHAM5sh model).

in horizontal resolution compare to the differences in stratospheric vertical resolution.

- *Time step.* The increasing use of inherently stable advection schemes such as semi-implicit semi-Lagrangian methods allows for longer time steps than are possible, for example, with a more traditional Eulerian advection. While this can lead to significant savings in computing requirements, particularly at higher spatial resolution, an adverse effect is the filtering or damping of high-frequency equatorial waves (e.g. Shutts and Vosper, 2011) that can potentially make a significant contribution to the QBO momentum budget. See column 5 of Table 6 for the different dynamical time steps used by the participating models.
- *Dynamical core.* In a set of idealized experiments Yao and Jablonowski (2015) demonstrated that spontaneous generation of QBO-like behaviour in general circulation models was sensitive to the dynamical core chosen. This involved both the choice of numerical advection scheme (including any associated numerical diffusion) and the dissipation mechanisms. As well as impacting wave generation, propagation, and dissipation mechanisms, the choices can impact (Yao and Jablonowski, 2015) the simulation of the Brewer–Dobson circulation (Butchart, 2014), in particular its tropical upwelling

component which opposes the descending QBO cycles in the standard paradigm (Baldwin et al., 2001). The different advection schemes used by the QBOi models are given in column 2 of Table 7. In the equatorial stratosphere most of the unresolved mechanical (e.g. GWD; see below) and thermal (e.g. radiative heating and cooling) dissipation in these models is from complex physical parameterizations though many of the models also include some explicit diffusion and a “sponge layer” to prevent spurious reflections from the upper boundary. A condensed summary of the diffusion and sponge layer information for the QBOi models is also given in Table 7. For more details see the references given in column 6 of Table 5.

- *Parameterized sub-grid-scale waves (non-orographic gravity waves).* A very significant development in models that has led to increased success in simulating QBO-like variability has been the introduction of non-orographic GWD parameterizations. Early schemes focused on parameterizing the (vertical) propagation and dissipation of sub-grid-scale waves from spatially and temporally fixed sources, while more recent developments have also included parameterized sources (e.g. Beres et al., 2005; Choi and Chun, 2011; Lott and Guez, 2013; Schirber et al., 2014; Bushell et al., 2015). Broadly speaking there have been two approaches to pa-

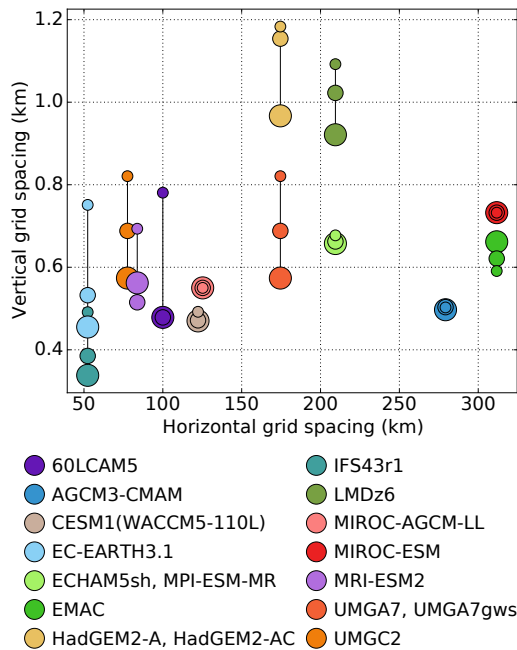


Figure 5. Vertical resolution, Δz , vs. horizontal resolution of models participating in QBOi. Since Δz can vary with altitude, as shown in Fig. 4, here Δz is shown for the three layers 10–15, 15–20, and 20–25 km spanning the tropical upper troposphere and lower stratosphere, with the size of the markers scaled by the average density in each layer. See column 3 of Table 6 for the total number of vertical levels in each model. The horizontal grid spacing is estimated by calculating the average of the zonal and meridional grid spacings, $(\Delta\lambda + \Delta\phi)/2$, and converting this to a value in kilometres at the Equator. For spectral models with triangular truncation we assume $\Delta\lambda = \Delta\phi = \frac{2}{3}180^\circ/(T + 1)$ as an estimate of the transform grid resolution, where T is the truncation wavenumber as given in column 2 of Table 6.

parameterizing the propagation and dissipation. The first, followed by Hines (1997a, b) and Warner and McIntyre (1996), aims to represent a broad spectrum of unresolved gravity waves generated by a variety of sources, while the alternative method is to represent the wave spectrum by a finite number or collection of monochromatic waves such as described by Lindzen (1981) or Alexander and Dunkerton (1999). All models or model versions participating in QBOi, with the exception of MIROC-AGCM-LL, include at least one parameterization of non-orographic GWD, with the superscripts ^a or ^b in the second column of Table 8 indicating, respectively, whether the spectrum or collection of monochromatic waves method is used. A comparison of how the different schemes attenuate parameterized eastward and westward momentum fluxes of non-orographic gravity waves propagating upward through typical wind profiles with opposite phases of the QBO is shown in Fig. 7 and described in detail in Sect. 5.1.

Five of the 17 models (60LCAM5, CESM1(WACCM5-110L) HadGEM2-AC, LMDz6, and UMGA7gws) have extended their non-orographic GWD parameterizations to include parameterized gravity wave sources¹. References giving details of these extended parameterizations are listed in column 3 of Table 8. In most cases this has simply involved replacing an ersatz “fixed” source with one that is more physically based, although for the LMDz6 model the previously used Hines scheme was replaced with a new GWD parameterization (Lott et al., 2012; Lott and Guez, 2013). There are two pairs of models that are identical except for their gravity wave source being fixed or parameterized: UMGA7–UMGA7gws and HadGEM2-A–HadGEM2-AC. Hence it will be possible to assess the impact that these model developments have on the simulation of the QBO and how it responds to changes in climate forcings, at least for a small subset of the participating models.

- *Convection.* An important source of equatorial waves in the models is convection and its associated diabatic heating. Gravity wave source parameterizations also typically couple the generation of parameterized GWD to parameters obtained from the convection schemes such as the precipitation (e.g. Lott and Guez, 2013). The different convection schemes used by the participating models are listed in column 4 of Table 8 for easy comparison.
- *Ozone.* Although differences in ozone climatologies can potentially impact simulated QBOs (e.g. Bushell et al., 2010), precise specifications for the ozone forcing were not included in the experiment descriptions (Sect. 3; Appendix A) to allow for the inclusion of models with prognostic ozone and also to keep the experiment specifications as simple as possible. Therefore for those models without ozone chemistry (see column 5 of Table 8) there are some variations among the ozone climatologies that have been prescribed. Figure 6 illustrates these variations in the tropics for the ozone climatologies used in the QBOi experiments, except for MPI-ESM-MR and UMGC2 which provided existing hindcast results for Experiment 5a. Any sensitivity of the simulated QBOs to these variations in the ozone climatology is, however, not considered critical for phase 1 of the QBOi analysis, as each model was tuned to give its “best” QBO with its prescribed ozone. On the other hand, it is important for the analysis that for a particular model the *same* ozone

¹A “source parameterization” denotes a gravity wave source that is coupled with other physical fields in the model, such as precipitation or deep convective heating, and therefore varies temporally and spatially. In contrast, “fixed” gravity wave sources are not coupled to other physical fields. Fixed sources are often constant in time, although this category could also include sources that have a prescribed temporal variation (e.g. seasonal cycle) or are stochastic.

was used across all the experiments performed with that model (see Sect. 3). Sensitivity of the QBO to ozone is expected to be considered in the next phases of QBOi (see Sect. 6).

5.1 Offline comparison of non-orographic gravity wave drag schemes used in the participating models

As noted above, non-orographic GWD parameterizations have been important for the generation of a QBO in many climate models. Only MIROC_AGCM-LL does not use parameterized GWD for QBOi (see Table 8). The non-orographic GWD parameterization schemes used in all the other QBOi models, except for 60LCAM5 and CESM1(WACCM5-110L), are compared by performing offline calculations for prescribed equatorial wind and temperature profiles (see Appendix B for details and why it was not possible to include results for 60LCAM5 and CESM1(WACCM5-110L), which use the same parameterization based on Lindzen, 1981). The 1 May 1993 and 1 November 2005 start dates for Experiment 5 (Fig. 3) are used since they have oppositely phased QBOs. Three experiments are performed. The first two use a prescribed amount of momentum flux (MF) at a launch height² of 100 hPa, namely 1 and 10 mPa. Inter-model differences in GWD in these two experiments arise solely from differences in the phase speed spectrum at the launch height and the non-linear dissipation mechanism inherent in the schemes (e.g. Hines' Doppler spreading or Warner and McIntyre's imposed saturated spectrum). The purpose of the 10 mPa experiment is to see how linearly the MF (and GWD) scales with the MF at 100 hPa in comparison with the 1 mPa experiment. The third experiment uses the models' own launch heights and amplitudes; hence this experiment most closely matches the set-up used in the QBOi simulations. For all three experiments the GWD is computed at each longitude and the results are zonally averaged.

Vertical profiles of zonal mean GWD for the 1 mPa experiment are shown in Fig. 7b and e. Results for the 10 mPa experiment (not shown) are quite similar to the 1 mPa results but are larger by a factor of 10, confirming that to a good first approximation the GWD at these heights scales linearly with the MF at 100 hPa. This is perhaps not too surprising given that critical level absorption by the background winds, as opposed to non-linear dissipation resulting from the exponential growth with height of the gravity wave amplitudes, is the primary cause of the momentum flux deposition in these highly sheared wind profiles. The results of the third experiment are shown in Fig. 7c and f. Compared to the 1 mPa results, these show much more inter-model spread. Since the source specifications used in this experiment are

²For the models with parameterized gravity wave sources, this "launch height" is instead a reference height at which the offline scheme are tuned to have the specified properties; see Appendix B for further details.

the ones that produce each model's best QBO, the larger inter-model spread in the third experiment is a reflection of model-dependent biases in, for instance, the mean winds and temperatures and resolved waves that must be overcome by tuning the gravity wave sources. The GWD profiles between 20 and 40 km are approximately Gaussian in form and can be simplified by fitting the zonal mean GWD to a function of the form $A \exp[-((z - B)/C)^2]$. The three fit parameters are shown in the insets in Fig. 7b–c and e–f. The increase in inter-model spread of the maximum GWD (fit parameter A) in the experiment using the models' launch amplitudes and heights is more readily seen. As observed (not simulated) precipitation is used in the offline calculations for two of the models using parameterized gravity wave sources (LMDz6 and UMGA7gws), the results in Fig. 7c and f may not accurately reflect what the models themselves would produce. Hence the parameterized-source and fixed-source results are not entirely comparable. A case in point is the rather large difference in the peak GWD in the UMGA7 (fixed source) and UMGA7gws (parameterized source) results; for this reason the UMGA7gws results have been scaled to fit on the plot. Also note in the 1 mPa experiment that the GWD peaks are wider in the vertical and weaker for the models that use Hines than for the others. This is consistent with the vertical smoothing of the momentum fluxes that is conventionally applied in the Hines scheme before the GWD is computed. The differences in the 1 mPa Hines results are a consequence of the different amount of smoothing used by the different models; if the smoothing is removed from the offline calculation, the 1 mPa Hines results for the different models are identical.

In summary, the offline comparison shows that most of the inter-model differences in the parameterized GWD in the equatorial stratosphere arise from the differences in their launch height and launch amplitude, not from differences in the wave dissipation mechanism and the shape of the assumed launch spectrum.

6 Closing remarks and future plans

The QBO is arguably the most conspicuous and regular mode of variability observed anywhere in the atmosphere that is not directly related to either the annual or diurnal cycles. At a fundamental level and for current conditions, it can be considered to be purely an atmospheric dynamical mode of variability, despite possible external influences from variability in the oceans, the solar cycle, or changes in atmospheric composition. Therefore the primary goals of phase 1 of QBOi are achievable using atmosphere-only global models that are computationally relatively inexpensive to run. To date (January 2018) output from 17 models and/or model versions (Table 5) has been uploaded or is planned for uploading to the shared database.

The goals of phase 1 of QBOi are to

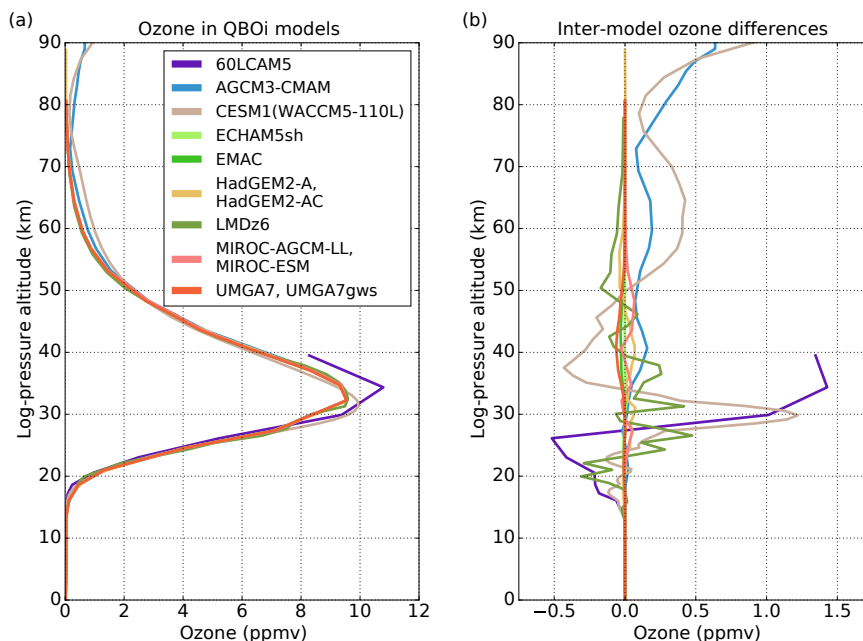


Figure 6. (a) Vertical profiles in the tropics of the climatological ozone concentration prescribed in QBOi experiments for models that do not include ozone chemistry (as indicated in Table 8) and excluding MPI-ESM-MR and UMGC2, which uploaded existing hindcast data for Experiment 5a. Each vertical profile is an average over the 5° S– 5° N latitude band, zonal mean, and annual mean. (b) As (a), but showing differences from a reference profile so that inter-model variations are more clearly visible. The reference profile is the 1988–2007 climatology of the SPARC ozone referred to in Appendix A (item A1). The period 1988–2007 is the same as that recommended for Experiment 2 for the SST and sea ice climatologies (Appendix A, item A2).

- compare, for present-day conditions, the accuracy of the morphology of the simulated QBOs across models and relate this to differences between models in the representation of the forcing mechanisms (e.g. terms contributing to the zonal-mean zonal momentum equation) and other model properties such as the resolution and sources of waves;
- compare how the morphology of the simulated QBOs and QBO forcing mechanisms respond to climate change (i.e. a doubling and quadrupling of CO_2 amounts) and identify which aspects of these responses are robust; and
- compare QBO predictive skill between models and its dependence on the QBO’s initialized phase, the underlying state of the atmosphere, and/or properties of the individual models (e.g. why was there an absence of skill in predicting the disruption of the QBO in 2016?).

Phase 1 of QBOi therefore addresses the challenges associated with modelling, predicting the evolution of, and projecting long-term changes in the QBO. Results from planned studies are expected to provide information on requirements for future model development leading to more accurate representations of the QBO and its variability in the individual models and across the multi-model ensemble. Benefits, how-

ever, are likely to extend well beyond this and range from potential enhancements in skill in seasonal to decadal predictions resulting from concomitant improvements in QBO extra-tropical dynamical teleconnections, to better capabilities for assessing the consequences of geoengineering proposals involving the injection of aerosol into the equatorial stratosphere where its redistribution away from the tropics is likely to be significantly influenced by the QBO.

Beyond phase 1, QBOi is expected to focus more on QBO extra-tropical dynamical teleconnections and couplings to other aspects of the climate system. In this respect QBOi again differs from multi-model activities like CMIP and CCMI that are largely policy-driven and hence place considerable emphasis on continually updating projections using the latest generation of models. Instead the developing consensus in the QBOi community, which has emerged primarily from the September 2016 QBO workshop (see Anstey et al., 2017, for a workshop summary), is to build on the experiments described in this paper, though of course results from phase 1 studies are expected to feed through into improving the representation of the QBO in the next generation of models. Some new coordinated studies that have been proposed for future endorsement by QBOi include

- increasing the ensemble size of Experiment 1 (“AMIP”) to examine the robustness across models of possible

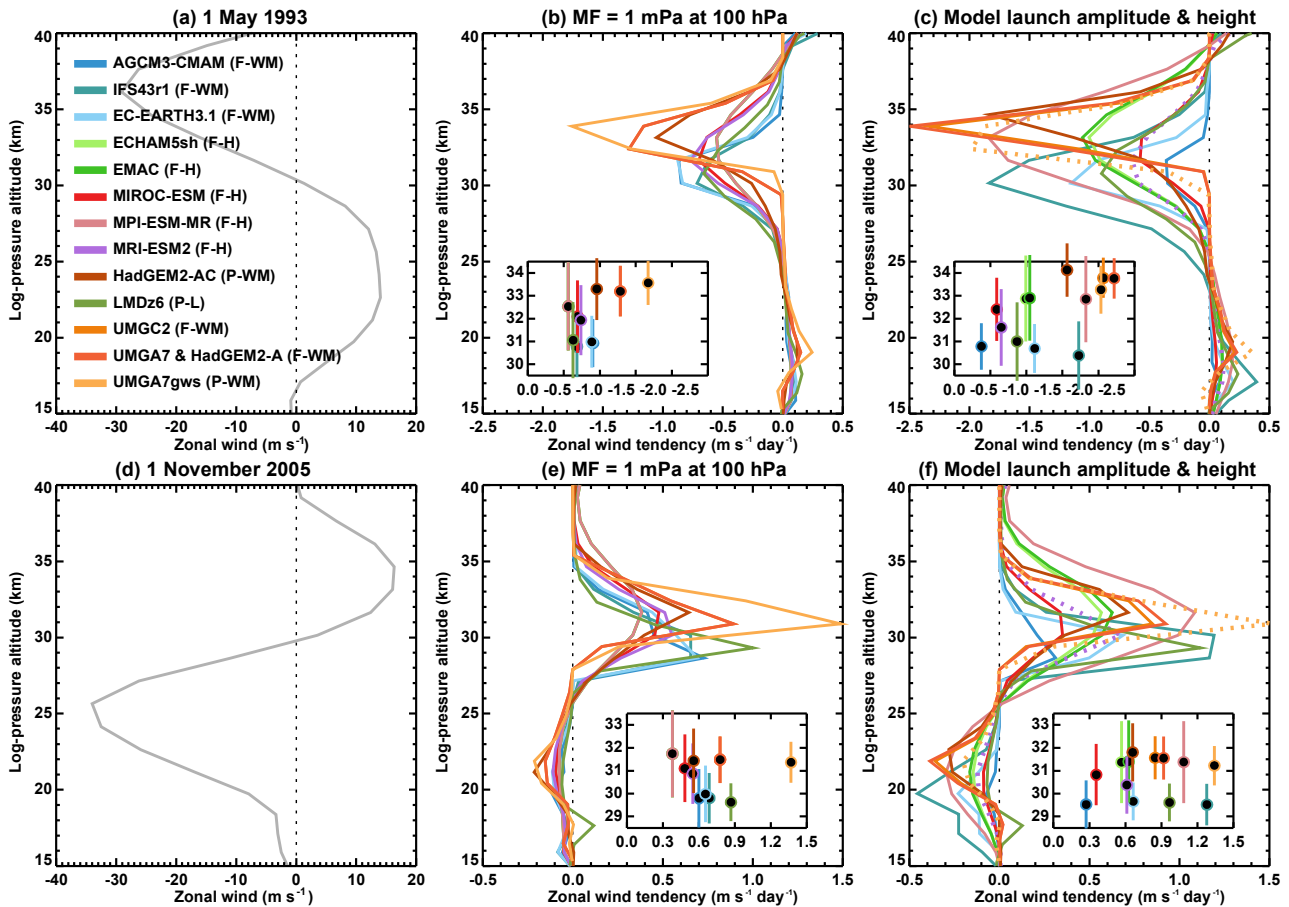


Figure 7. Vertical profiles of zonal mean non-orographic GWD computed using the parameterization schemes used by the different models. The offline calculations are performed using ERA-Interim equatorial zonal and meridional winds and temperatures for 1 May 1993 (a, b, c) and 1 November 2005 (d, e, f). The middle panels show results for the case in which the momentum flux is set to 1 mPa at 100 hPa (≈ 16 km). The right panels show results for the case in which the models' own launch amplitudes and launch heights are used. Note that the results in the right-hand panel for MRI-ESM2 and UMGA7gws have been multiplied by 0.1 and 0.6, respectively, and the GWD profiles plotted using dotted lines (see Appendix B). The labels in parentheses to the right of the model names denote the type of GWD scheme: “F” or “P” for fixed or parameterized sources; “H” for Hines, “WM” for Warner–McIntyre, or “L” for Lott et al. (2012) for the type of dissipation used. Note that “WM” here includes both the Warner and McIntyre (1999) and Scinocca (2003) schemes (Table 8), which are both implementations of the Warner and McIntyre (1996) framework for gravity wave parameterization. The insets show the parameters of Gaussian fits, $A \exp[-((z - B)/C)^2]$, to the zonal mean GWD profiles. The peaks of the Gaussians (A , m s^{-1} per day, horizontal axes) and their heights (B , km, vertical axes) are denoted by the filled circles. The e -folding widths of the Gaussians (C , km) are given by the vertical bars. See text and Appendix B for more details.

synchronization between ENSO events and the QBO (e.g. Christiansen et al., 2016) and other ENSO influences on the QBO (e.g. Barton and McCormack, 2017);

- extending Experiment 2 (present-day time slice) to increase the sample size to examine QBO teleconnection robustness in an idealized framework in which there is no other externally forced variability, apart from the annual and diurnal cycles;
- repeating Experiment 2 (present-day time slice) with idealized perpetual El Niño–La Niña SST anomalies to

examine the interaction of ENSO and QBO teleconnections;

- empirically separating the effects of stratospheric and tropospheric climate change on the QBO by modifying Experiments 3 and 4 (future time slice) such that the increases in CO_2 amount (\sim forcings stratospheric climate change only) and SSTs (\sim forcing tropospheric climate change only) are applied separately;
- extending Experiment 5–5a (retrospective hindcasts) to examine the 2016 QBO disruption and its predictability; and

- examining the impact of ozone on the QBO either through prescribed ozone perturbations or through ozone feedbacks for models that can rerun with and without ozone chemistry.

The above list is by no means exhaustive and other possible extensions of the research plans for QBOi include more idealized studies comparing simulations using only “dynamical cores” (e.g. Yao and Jablonowski, 2015) or perhaps simulations in which the QBO is artificially removed (e.g. by turning off the non-orographic GWD parameterization in the tropics). However, in line with current QBOi practices, details of any new coordinated studies will again be formulated through community discussion at forthcoming QBOi workshops and will depend on the outcomes of the phase 1 studies.

Code and data availability. For information on the code availability for the individual models considered in this paper see the appropriate references given in Table 5. Details of the QBOi data repository and how to access it are provided in the Supplement.

Appendix A: Experiments – technical specifications

A1 Experiment 1 – “AMIP”

Experiment 1 is based on the CMIP5 experiment 3.3 alternatively referred to as the “Atmospheric Model Inter-comparison Project (AMIP)” experiment (Taylor et al., 2012). It is a one- to three-member ensemble of 30-year simulations using observed SSTs and sea ice amounts from 1 January 1979 to 28 February 2009. These can be obtained from http://www-pcmdi.llnl.gov/projects/amip/AMIP2EXPDSN/BCS/amipbc_dwnld.php.

The corresponding external forcings for the CMIP5 AMIP experiment (e.g. radiative trace gas concentrations, aerosol distributions, solar irradiance, and appropriate forcings from explosive volcanoes) can be found here: <http://cmip-pcmdi.llnl.gov/cmip5/forcing.html#amip> apart from ozone which, for high-top models, can be obtained from Osprey et al. (2011): <https://doi.org/10.5281/zenodo.1035142>.³

Initial conditions are not prescribed and it is left to individual groups to use whatever is appropriate for their model and to include any spin-up if this is considered necessary.

A2 Experiment 2 – $1\times\text{CO}_2$

Experiment 2 is similar to Experiment 1 but with a repeated annual cycle for the SSTs and sea ice amounts plus all the other forcings (i.e. there is no inter-annual variability or any secular changes in the forcings). It can either be a one- to three-member ensemble of 30-year simulations or *preferably* a single 100-year (or longer) simulation. The long single integration has the additional potential of providing information on very low-frequency variations.

Ideally the external annual cycle forcings should be 30-year climatologies based on Experiment 1, although as these are generally not readily available, a suitable alternative is to apply annually repeating forcings based on the 2002 CMIP5 forcings. The year 2002 is well removed from any explosive volcanic eruptions and the ENSO and Pacific Decadal Oscillation (PDO) are both in their neutral phases and hence conditions in this year can be considered as a useful proxy for the multi-year mean for most quantities. However, 2002 ozone amounts are likely to be strongly perturbed because of the Southern Hemisphere sudden stratospheric warming (e.g. Shepherd et al., 2005) and for ozone a 2-D climatological field representative of the 1990s is preferable. For SSTs and sea ice amounts CMIP5 1988–2007 climatologies are available from http://www-pcmdi.llnl.gov/projects/amip/AMIP2EXPDSN/BCS/amipbc_dwnld.php.

As Experiment 2 is the control for Experiments 3 and 4 ($2\times\text{CO}_2$ and $4\times\text{CO}_2$, respectively) the average CO_2 amount for 2002 should be used as the baseline $1\times\text{CO}_2$ amount.

³These ozone data are identical to those which previously could be obtained from https://groups.physics.ox.ac.uk/climate/osprey/QBOi_O3/.

Although the use of different length climatologies for different forcings is not ideal and does not provide direct comparison to the 30-year period of Experiment 1, the observed dependence of the QBO on a changing climate through this period appears to be negligible. Thus for QBOi the benefits of the simpler experimental set-up is considered to far outweigh any possible disadvantages. Nonetheless it important to emphasize that the same idealized set of climatologies and forcings are to be used throughout Experiments 2–4, apart from the changes to the CO_2 amounts and SSTs described below.

As with Experiment 1, atmospheric initial conditions are not prescribed.

A3 Experiments 3 – $2\times\text{CO}_2$, and 4 – $4\times\text{CO}_2$

Experiments 3 and 4 are the same as Experiment 2 but for $2\times\text{CO}_2$ and $4\times\text{CO}_2$ climates, respectively. Again these can either be a one- to three-member ensemble of 30-year simulations or *preferably* a single 100-year simulation after allowing for a suitable spin-up to the new climate (without a coupled ocean this is expected to be fairly rapid though for the $4\times\text{CO}_2$ experiment this can be of the order of 5 years). Compared to the amount specified for Experiment 2 the CO_2 concentration should either be doubled (Experiment 3) or quadrupled (Experiment 4) with a corresponding idealized adjustment made to the SSTs of a spatially uniform perturbation of +2 K for $2\times\text{CO}_2$ and +4 K for $4\times\text{CO}_2$. Sea ice amounts should be kept the same as in Experiment 2.

All other forcings in these two experiments should be *exactly* the same as in Experiment 2 including the amounts of all radiatively active greenhouse gases other than CO_2 . If ozone is prescribed (i.e. if the model does not have interactive chemistry) then this too should be exactly the same as in Experiment 2. Alternatively if the model does have interactive chemistry then the source gases and/or emissions should be kept exactly the same as in Experiment 2. This idealized set-up for Experiments 3 and 4 is appropriate as these are sensitivity experiments and not attempts to predict specific periods in the future.

As with Experiment 2 atmospheric initial conditions are not prescribed, but note the need to allow for spin-up to the new climates.

A4 Experiment 5 – QBO hindcasts

These are atmosphere-only experiments initialized from re-analysis data and providing multiple ensembles of short integrations from a relatively large set of start dates sampling different phases of the QBO. The prescribed start dates (i.e. atmospheric initial conditions) are 1 May and 1 November for the years 1993–2007 (i.e. 15 years with a total 30 start dates). The duration of each hindcast should be at least 6 months but preferably 9–12 months.

As with Experiment 1 the boundary conditions and external forcings should be the same as those specified for the CMIP5 AMIP experiment (Taylor et al., 2012). CMIP5 inter-annually varying sea ice and SSTs can be obtained from http://www-pcmdi.llnl.gov/projects/amip/AMIP2EXPDSN/BCS/amipbc_dwnld.php, while the CMIP5 external forcings for radiative trace gas concentrations, aerosols, solar, explosive volcanoes, etc., can be obtained from <http://cmip-pcmdi.llnl.gov/cmip5/forcing.html#amip>. Ozone forcing datasets appropriate for use in high-top models are available from Osprey et al. (2011): <https://doi.org/10.5281/zenodo.1035142>. Initial data for the hindcasts should be taken from the ERA-Interim reanalysis (Dee et al., 2011), which can be downloaded from <http://apps.ecmwf.int/datasets>.

Registration is required; if downloading many start dates from this site, it may be easier to use the “batch access” method described on the site, although interactive download of each date is also possible. Data are available on either standard pressure levels or original model levels and in either grib or netCDF formats. The ensemble is expected to be generated by perturbing the initial conditions by a small anomaly, which requires no more than changing the bit pattern of the simulation. For some models this is possible through stochastic physics; however, each group should use an ensemble generation method that is most appropriate to their model and that is most readily available to them.

A5 Experiment 5a – QBO forecasts

This experiment is as Experiment 5, but using a coupled ocean–atmosphere model and predicting the SST instead of specifying observed values. External forcings should also be fixed at the initial start time so as not to use future information. This is then a true forecast experiment for the QBO and can be compared with the results of Experiment 5. Some groups may have already performed these hindcasts as part of their operational seasonal forecasts but note that for QBOi purposes it is important that the majority of the diagnostics discussed in Sect. 4 are available for a full comparison to Experiment 5 results.

Appendix B: Offline non-orographic gravity wave drag calculations

This appendix provides details about the offline GWD calculations shown in Fig. 7. The background equatorial winds and temperatures are from a single day (daily mean) of ERA-Interim data on a 1° longitude grid and on pressure levels at the ECMWF model-level resolution.

For models that use “fixed” gravity wave sources (e.g. AGCM3-CMAM), the calculations are straightforward and simply involve computing the GWD above the launch height. Since these models all use a horizontally isotropic gravity wave source, the MF in a single azimuth is set to either 1 or 10 mPa for the first two experiments. All fixed-source calculations are done using offline versions of the Scinocca (2003), Hines (1997a, b), and Warner and McIntyre (1999) non-orographic GWD schemes using each model’s parameter settings. Results for the third offline experiment, in which the models’ own source amplitudes (i.e. momentum flux for Scinocca, root mean square (RMS) winds for Hines) and launch heights are used, are validated by comparing to results from QBOi Experiment 5 for models that provided daily mean GWD. With the exception of one model, the agreement is reasonably good, which is all that can be expected given that the resolution of the models differs from that used in the offline calculations. For MRI-ESM2 the offline results for the third experiment are 10 times larger than the Experiment 5 results and have been scaled in Fig. 7c and f. The reason for this large discrepancy is unknown. For models that tie their non-orographic gravity wave sources to parameterized processes in the troposphere (referred to in the Fig. 7 caption as parameterized sources), the calculations are more involved.

For the models that were able to perform the offline calculations for parameterized-source schemes (LMDz6, UMGA7gws, and HadGEM2-AC) the procedure was as follows. For LMDz6, daily precipitation observations were used to generate an ensemble of monochromatic waves. The background winds and temperatures are held fixed in time using either the 1 May or 1 November data. A similar procedure is used for the other two models, except that the launch momentum fluxes in HadGEM2-AC are obtained by sampling from the Experiment 1 result for the month since the source parameterization in HadGEM2-AC requires convective heating profiles not provided by observations. As momentum flux is not prescribed for these models, tuning the gravity wave parameters is required to achieve the desired MF at 100 hPa for the first two experiments such that $(|MF_{\text{east}}| + |MF_{\text{west}}|)/2 = 1$ or 10 mPa at 100 hPa. Due to time constraints, the NCAR group, which also ties its GWD scheme to convection in the 60LCAM5 and CESM1(WACCM5-110L) models, was unable to participate in this comparison.

The Supplement related to this article is available online at <https://doi.org/10.5194/gmd-11-1009-2018-supplement>.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The design of the experiments described here grew out of community discussions at the first QBOi workshop in March 2015 in Victoria, Canada. Funding for the workshop from the UK Natural Environment Research Council (NE/M005828/1), the World Climate Research Programme (WCRP), Stratosphere–troposphere Processes And their Role in Climate (SPARC) activity, and the Canadian Centre for Climate Modelling and Analysis is gratefully acknowledged. We further acknowledge the scientific guidance of the WCRP for helping motivate this work, coordinated under the framework of the SPARC QBO initiative (QBOi) led by James Anstey, Neal Butchart, Kevin Hamilton, and Scott Osprey. The Centre for Environmental Data Analysis (CEDA) have very kindly offered to host the QBOi data archive. Neal Butchart and Adam Scaife were supported by the Joint UK BEIS/Defra Met Office Hadley Centre Climate Programme (GA01101). Scott Osprey and Lesley Gray were supported by NERC projects NE/M005828/1 and NE/P006779/1. Shingo Watanabe and Yoshio Kawatani used the Earth simulator for QBOi simulations and were supported by the SOUSEI programme, MEXT Japan, and the Japan Science and Technology Agency (JST) as part of the Belmont Forum. Yoshio Kawatani was supported by Grant-in-Aid for Scientific Research B (26287117), joint international research (15KK0178) from the Japan Society for the Promotion of Science, and the Environment Research and Technology Development Fund (2-1503) of the Ministry of the Environment, Japan. Francois Lott and Scott Osprey were supported by the ANR/JPI-Climate/Belmont Forum project GOTHAM (ANR-15-JCLI-0004-01). Federico Serva was supported by the European Commission under grant StratoClim-603557-FP7-ENV.2013.6.1-2, with computing resources for the ECHAM5sh simulations provided by an ECMWF special project. Young-Ha Kim was supported by the Basic Science Research Program through the National Research Foundation of Korea funded by the Ministry of Science, ICT & Future Planning (NRF-2015R1C1A1A02036449). Holger Pohlmann was supported by the German Federal Ministry for Education and Research (BMBF) project MiKlip (FKZ 01LP1519A) and thanks Elisa Manzini for providing additional information on the MPI model. BSC contribution is supported by the Spanish MINECO-funded DANAE project (CGL2015-68342-R) and Red Española de Supercomputación (RES project AECT-2017-3-0015).

Edited by: Paul Ullrich

Reviewed by: three anonymous referees

References

- Adachi, Y., Yukimoto, S., Deushi, M., Obata, A., Nakano, H., Tanaka, T. Y., Hosaka, M., Sakami, T., Yoshimura, H., Hirabara, M., Shindo, E., Tsujino, H., Mizuta, R., Yabu, S., Koshiro, T., Ose, T., and Kitoh, A.: Basic performance of a new earth system model of the Meteorological Research Institute (MRI-ESM1), *Pap. Meteorol. Geophys.*, 64, 1–19, 2013.
- Alexander, M. J. and Dunkerton, T. J.: A spectral parameterization of mean-flow forcing due to breaking gravity waves, *J. Atmos. Sci.*, 56, 4167–4182, [https://doi.org/10.1175/1520-0469\(1999\)056<4167:ASPOMF>2.0.CO;2](https://doi.org/10.1175/1520-0469(1999)056<4167:ASPOMF>2.0.CO;2), 1999.
- Alexander, M. J. and Sato, K.: Gravity wave dynamics and climate: An update from the SPARC gravity wave activity, *SPARC Newsletter*, 44, 9–13, 2015.
- Andrews, D. G., Holton, J. R., and Leovy, C. B.: *Middle Atmospheric Dynamics*, Academic Press, San Diego, California, 489 pp., 1987.
- Anstey, J. A. and Shepherd, T. G.: High-latitude influence of the quasi-biennial oscillation, *Q. J. Roy. Meteor. Soc.*, 140, 1–21, <https://doi.org/10.1002/qj.2132>, 2014.
- Anstey, J. A., Scinocca, J. F., and Keller, M.: Simulating the QBO in an atmospheric general circulation model: Sensitivity to resolved and parameterized forcing, *J. Atmos. Sci.*, 73, 1649–1665, <https://doi.org/10.1175/JAS-D-15-0099.1>, 2016.
- Anstey, J. A., Shepherd, T. G., and Scinocca, J. F.: Influence of the quasi-biennial oscillation on the extratropical winter stratosphere in an atmospheric general circulation model and in reanalysis data, *J. Atmos. Sci.*, 67, 1402–1419, <https://doi.org/10.1175/2009JAS3292.1>, 2010.
- Anstey, J., Hamilton, K., Osprey, S., Butchart, N., and Gray, L.: Report on the 1st QBO modelling and reanalyses workshop, 16–18 March 2015, Victoria, BC, Canada, *SPARC Newsletter*, 45, 19–25, 2015.
- Anstey, J., Osprey, S., Butchart, N., Hamilton, K., Gray, L., and Baldwin, M.: Report on the SPARC QBO Workshop: The QBO and its Global Influence – Past, Present and Future 26–30 September 2016, Oxford, UK, *SPARC Newsletter*, 48, 33–41, 2017.
- Baldwin, M. P., Gray, L. J., Dunkerton, T. K., Hamilton, K., Haynes, P. H., Randel, W. J., Holton, J. R., Alexander, M. J., Hirota, I., Horinouchi, T., Jones, D. B. A., Kinniersley, J. S., Marquardt, C., Sato, K., and Takahashi, M.: The quasi-biennial oscillation, *Rev. Geophys.*, 39, 179–229, <https://doi.org/10.1029/1999RG000073>, 2001.
- Barton, C. A. and McCormack, J. P.: Origin of the 2016 QBO disruption and its relationship to extreme El Niño events, *Geophys. Res. Lett.*, 44, 11105–11157, <https://doi.org/10.1002/2017GL075576>, 2017.
- Bechtold, P., Köhler, M., Jung, T., Doblas-Reyes, F., Leutbecher, M., Rodwell, M. J., Vitart, F., and Balsamo, G.: Advances in simulating atmospheric variability with the ECMWF model: From synoptic to decadal time-scales. *Q. J. Roy. Meteor. Soc.*, 134, 1337–1351, <https://doi.org/10.1002/qj.289>, 2008.
- Beres, J. H., Garcia, R. R., Boville, B. A., and Sassi, F.: Implementation of a gravity wave source spectrum parameterization dependent on the properties of convection in the Whole Atmosphere Community Climate Model (WACCM), *J. Geophys. Res.*, 110, D10108, <https://doi.org/10.1029/2004JD005504>, 2005.

- Bretherton, C. S., McCaa, J. R., and Grenier, H.: A new parameterization for shallow cumulus convection and its application to marine subtropical cloud-topped boundary layers. Part I: description and 1D results, *Mon. Weather Rev.*, 132, 864–882, [https://doi.org/10.1175/1520-0493\(2004\)132<0864:ANPFSC>2.0.CO;2](https://doi.org/10.1175/1520-0493(2004)132<0864:ANPFSC>2.0.CO;2), 2004.
- Bushell, A. C., Jackson, D. R., Butchart, N., Hardiman, S. C., Hinton, T. J., Osprey, S. M., and Gray, L. J.: Sensitivity of GCM tropical middle atmosphere variability and climate to ozone and parameterized gravity wave changes, *J. Geophys. Res.*, 115, D15101, <https://doi.org/10.1029/2009JD013340>, 2010.
- Bushell, A. C., Butchart, N., Derbyshire, S. H., Jackson, D. R., Shutts, G. J., Vosper, S. B., and Webster, S.: Parameterized gravity wave momentum fluxes from sources related to convection and large-scale precipitation processes in a global atmosphere model, *J. Atmos. Sci.*, 72, 4349–4371, <https://doi.org/10.1175/JAS-D-15-0022.1>, 2015.
- Butchart, N.: The Brewer-Dobson circulation, *Rev. Geophys.*, 52, 157–184, <https://doi.org/10.1002/2013RG000448>, 2014.
- Butchart, N., Scaife, A. A., Austin, J., Hare, S. H. E., and Knight, J. R.: Quasi-biennial oscillation in ozone in a coupled chemistry-climate model, *J. Geophys. Res.*, 108, 4486, <https://doi.org/10.1029/2002JD003004>, D15, 2003.
- Butchart, N., Charlton-Perez, A. J., Cionni, I., Hardiman, S. C., Haynes, P. H., Krüger, K., Kushner, P. J., Newman, P. A., Osprey, S. M., Perlwitz, J., Sigmond, M., Wang, L., Akiyoshi, H., Austin, J., Bekki, S., Baumgaertner, A., Braesicke, P., Brühl, C., Chipperfield, M., Dameris, M., Dhomse, S., Eyring, V., Garcia, R., Garny, H., Jöckel, P., Lamarque, J.-F., Marchand, M., Michou, M., Morgenstern, O., Nakamura, T., Pawson, S., Plummer, D., Pyle, J., Rozanov, E., Scinocca, J., Shepherd, T. G., Shibata, K., Smale, D., Teyssèdre, H., Tian, W., Waugh, D., and Yamashita, Y.: Multimodel climate and variability of the stratosphere, *J. Geophys. Res.*, 116, D05102, <https://doi.org/10.1029/2010JD014995>, 2011.
- Choi, H.-J. and Chun, H.-Y.: Momentum flux spectrum of convective gravity waves. Part I: An update of a parameterization using mesoscale simulations, *J. Atmos. Sci.*, 68, 739–759, <https://doi.org/10.1175/2010JAS3552.1>, 2011.
- Christiansen, B., Yang, S., and Madsen, M. S.: Do strong warm ENSO events control the phase of the stratospheric QBO?, *Geophys. Res. Lett.*, 43, 10489–10495, <https://doi.org/10.1002/2016GL070751>, 2016.
- Coy, L., Newman, P. A., Pawson, S., and Lait, L. R.: Dynamics of the disrupted 2015/16 quasi-biennial oscillation, *J. Climate*, 30, 5661–5674, <https://doi.org/10.1175/JCLI-D-16-0663.1>, 2017.
- Davini, P., von Hardenberg, J., Corti, S., Christensen, H. M., Juricke, S., Subramanian, A., Watson, P. A. G., Weisheimer, A., and Palmer, T. N.: Climate SPHINX: evaluating the impact of resolution and stochastic physics parameterisations in the EC-Earth global climate model, *Geosci. Model Dev.*, 10, 1383–1402, <https://doi.org/10.5194/gmd-10-1383-2017>, 2017.
- Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, I., Biblot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Greer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J. N., and Vitart, F.: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system, *Q. J. Roy. Meteor. Soc.*, 137, 553–597, 2011.
- de la Cámara, A. and Lott, F.: A parameterization of gravity waves emitted by fronts and jets, *Geophys. Res. Lett.*, 42, 2071–2078, <https://doi.org/10.1002/2015GL063298>, 2015.
- Dobrynin, M., Domeisen, D.I.V., Müller, W. A., Bell, L., Brune, S., Bunzel, F., Fröhlich, K., Pohlmann, H., Baehr, J.: Improved seasonal prediction of boreal winter through ensemble subsampling, *Geophys. Res. Lett.*, accepted, 2018.
- Dunkerton, T. J.: The quasi-biennial oscillation of 2015–2016: Hiccup or death spiral?, *Geophys. Res. Lett.*, 43, 10547–10552, <https://doi.org/10.1002/2016GL070921>, 2016.
- Dunstone, N., Smith, D., Scaife, A., Hermanson, L., Eade, R., Robinson, N., Andrews, M., and Knight, J.: Skilful predictions of the winter North Atlantic Oscillation one year ahead, *Nat. Geosci.*, 9, 809–814, <https://doi.org/10.1038/ngeo2824>, 2016.
- ECMWF: IFS documentation CY36R1, published by ECMWF, available at: <https://www.ecmwf.int/en/forecasts/documentation-and-support/changes-ecmwf-model/ifs-documentation>, 2010.
- ECMWF: IFS documentation CY43R1, published by ECMWF, available at: <https://www.ecmwf.int/en/forecasts/documentation-and-support/changes-ecmwf-model/ifs-documentation>, 2016.
- Emanuel, K. A.: A scheme for representing cumulus convection in large-scale models, *J. Atmos. Sci.*, 48, 2313–2335, [https://doi.org/10.1175/1520-0469\(1991\)048<2313:ASFRCC>2.0.CO;2](https://doi.org/10.1175/1520-0469(1991)048<2313:ASFRCC>2.0.CO;2), 1991.
- Emori, S., Nozawa, T., Numaguti, A., and Uno, I.: Importance of cumulus parameterization for precipitation simulation over East Asia in June, *J. Meteorol. Soc. Jpn.*, 79, 939–947, 2001.
- Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, *Geosci. Model Dev.*, 9, 1937–1958, <https://doi.org/10.5194/gmd-9-1937-2016>, 2016.
- Fujiwara, M., Wright, J. S., Manney, G. L., Gray, L. J., Anstey, J., Birner, T., Davis, S., Gerber, E. P., Harvey, V. L., Hegglin, M. I., Homeyer, C. R., Knox, J. A., Krüger, K., Lambert, A., Long, C. S., Martineau, P., Molod, A., Monge-Sanz, B. M., Santee, M. L., Tegtmeier, S., Chabrillat, S., Tan, D. G. H., Jackson, D. R., Polavarapu, S., Compo, G. P., Dragani, R., Ebisuzaki, W., Harada, Y., Kobayashi, C., McCarty, W., Onogi, K., Pawson, S., Simmons, A., Wargan, K., Whitaker, J. S., and Zou, C.-Z.: Introduction to the SPARC Reanalysis Intercomparison Project (S-RIP) and overview of the reanalysis systems, *Atmos. Chem. Phys.*, 17, 1417–1452, <https://doi.org/10.5194/acp-17-1417-2017>, 2017.
- Garcia, R. R. and Richter, J. H.: The momentum budget of the quasi-biennial oscillation in the high vertical resolution CESM1(WACCM), *J. Atmos. Sci.*, submitted, 2017.
- Geller, M. A., Zhou, T., Shindell, D., Ruedy, R., Aleinov, I., Nazarenko, L., Tausnev, N. L., Kelley, M., Sun, S., Cheng, Y., Field, R. D., and Faluvegi, G.: Modeling the QBO-improvements resulting from higher-model vertical resolution, *J. Adv. Model. Earth Sy.*, 8, 1092–1105, <https://doi.org/10.1002/2016MS000699>, 2016.

- Gerber, E. P. and Manzini, E.: The Dynamics and Variability Model Intercomparison Project (DynVarMIP) for CMIP6: assessing the stratosphere–troposphere system, *Geosci. Model Dev.*, 9, 3413–3425, <https://doi.org/10.5194/gmd-9-3413-2016>, 2016a.
- Gerber, E. P. and Manzini, E.: Corrigendum to “The Dynamics and Variability Model Intercomparison Project (DynVarMIP) for CMIP6: assessing the stratosphere–troposphere system” published in *Geosci. Model Dev.*, 9, 3413–3425, <https://doi.org/10.5194/gmd-9-3413-2016-corrigenum>, 2016b.
- Giorgetta, M. A., Manzini, E., and Roeckner, E.: Forcing of the quasi-biennial oscillation from a broad spectrum of atmospheric waves, *Geophys. Res. Lett.*, 29, 86-1–86-4, <https://doi.org/10.1029/2002GL014756>, 2002.
- Giorgetta, M. A., Manzini, E., and Roeckner, E., Esch, M., and Bengtson, L.: Climatology and forcing of the quasi-biennial oscillation in the MAECHAM5 Model, *J. Climate*, 19, 3882–3901, <https://doi.org/10.1175/JCLI3830.1>, 2006.
- Gregory, D. and Rowntree, P. R.: A massflux convection scheme with representation of cloud ensemble characteristics and stability dependent closure, *Mon. Weather Rev.*, 118, 1483–1506, [https://doi.org/10.1175/1520-0493\(1990\)118<1483:AMFCSW>2.0.CO;2](https://doi.org/10.1175/1520-0493(1990)118<1483:AMFCSW>2.0.CO;2), 1990.
- The HadGEM2 Development Team: Martin, G. M., Bellouin, N., Collins, W. J., Culverwell, I. D., Halloran, P. R., Hardiman, S. C., Hinton, T. J., Jones, C. D., McDonald, R. E., McLaren, A. J., O’Connor, F. M., Roberts, M. J., Rodriguez, J. M., Woodward, S., Best, M. J., Brooks, M. E., Brown, A. R., Butchart, N., Dearden, C., Derbyshire, S. H., Dharssi, I., Doutriaux-Boucher, M., Edwards, J. M., Falloon, P. D., Gedney, N., Gray, L. J., Hewitt, H. T., Hobson, M., Huddleston, M. R., Hughes, J., Ineson, S., Ingram, W. J., James, P. M., Johns, T. C., Johnson, C. E., Jones, A., Jones, C. P., Joshi, M. M., Keen, A. B., Liddicoat, S., Lock, A. P., Maidens, A. V., Manners, J. C., Milton, S. F., Rae, J. G. L., Ridley, J. K., Sellar, A., Senior, C. A., Totterdell, I. J., Verhoef, A., Vidale, P. L., and Wiltshire, A.: The HadGEM2 family of Met Office Unified Model climate configurations, *Geosci. Model Dev.*, 4, 723–757, <https://doi.org/10.5194/gmd-4-723-2011>.
- Hamilton, K., Wilson, R. J., and Hemler, R. S.: Spontaneous stratospheric QBO-like oscillations simulated by the GFDL SKYHI general circulation model, *J. Atmos. Sci.*, 58, 3271–3292, 2001.
- Hamilton, K., Osprey, S., and Butchart, N.: Modeling the stratosphere’s “heartbeat,” *Eos*, 96, p. 8, <https://doi.org/10.1029/2015EO032301>, 2015.
- Hines, C. O.: Doppler spreading parametrization of gravity-wave momentum deposition in the middle atmosphere. 1. Basic formulation, *J. Atmos. Sol.-Terr. Phys.*, 59, 371–386, 1997a.
- Hines, C. O.: Doppler spreading parametrization of gravity-wave momentum deposition in the middle atmosphere. 2. Broad and quasi-monochromatic spectra, and implementation, *J. Atmos. Sol.-Terr. Phys.*, 59, 387–400, 1997b.
- Horinouchi, T., Pawson, S., Shibata, K., Manzini, E., Giorgetta, M. A., Sassi, F., R. Wilson, R. J., Hamilton, K., DeGrandpe, J., and Scaife, A. A.: Tropical cumulus convection and upward propagating waves in middle-atmospheric GCMs, *J. Atmos. Sci.*, 60, 2765–2782, 2003.
- Hourdin, F., Grandpeix, J.-Y., Rio, C., Bony, S., Jam, A., Cheruy, F., Rochetin, N., Fairhead, L., Idelkadi, A., Musat, I., Dufresne, J.-L., Lahellec, A., Lefebvre, M.-P., and Roehrig, R.: LMDZ5B: the atmospheric component of the IPSL climate model with revisited parameterizations for clouds and convection, *Clim. Dynam.*, 40, 2193–2222, <https://doi.org/10.1007/s00382-012-1343-y>, 2013.
- IPCC: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, UK and New York, NY, USA, 2013.
- Jöckel, P., Sander, R., Kerkweg, A., Tost, H., and Lelieveld, J.: Technical Note: The Modular Earth Submodel System (MESSy) – a new approach towards Earth System Modeling, *Atmos. Chem. Phys.*, 5, 433–444, <https://doi.org/10.5194/acp-5-433-2005>, 2005.
- Jöckel, P., Kerkweg, A., Pozzer, A., Sander, R., Tost, H., Riede, H., Baumgaertner, A., Gromov, S., and Kern, B.: Development cycle 2 of the Modular Earth Submodel System (MESSy2), *Geosci. Model Dev.*, 3, 717–752, <https://doi.org/10.5194/gmd-3-717-2010>, 2010.
- Kawatani, Y. and Hamilton, K.: Weakened stratospheric quasibiennial oscillation driven by increased tropical mean upwelling, *Nature*, 497, 478–481, <https://doi.org/10.1038/nature12140>, 2013.
- Kawatani, Y., Hamilton, K., and Watanabe, S.: The quasi-biennial oscillation in a double CO₂ climate, *J. Atmos. Sci.*, 68, 265–283, <https://doi.org/10.1175/2010JAS3623.1>, 2011.
- Kawatani, Y., Sato, K., Dunkerton, T. J., Watanabe, S., Miyahara, S., and Takahashi, M.: The roles of equatorial trapped waves and internal inertia-gravity waves in driving the quasi-biennial oscillation. Part I: Zonal mean wave forcing, *J. Atmos. Sci.*, 67, 963–980, <https://doi.org/10.1175/2009JAS3222.1>, 2010.
- Kim, Y.-H. and Chun, H.-Y.: Momentum forcing of the quasi-biennial oscillation by equatorial waves in recent reanalyses, *Atmos. Chem. Phys.*, 15, 6577–6587, <https://doi.org/10.5194/acp-15-6577-2015>, 2015a.
- Kim, Y.-H. and Chun, H.-Y.: Contributions of equatorial wave modes and parameterized gravity waves to the tropical QBO in HadGEM2, *J. Geophys. Res.-Atmos.*, 120, 1065–1090, <https://doi.org/10.1002/2014JD022174>, 2015b.
- Kim, Y.-H., Bushell, A. C., Jackson, D. R., and Chun, H.-Y.: Impacts of introducing a convective gravity-wave parameterization upon the QBO in the Met Office Unified Model, *Geophys. Res. Lett.*, 40, 1873–1877, <https://doi.org/10.1002/grl.50353>, 2013.
- Knutti, R., Masson, D., and Gettelman, A.: Climate model genealogy: Generation CMIP5 and how we got there, *Geophys. Res. Lett.*, 40, 1194–1199, <https://doi.org/10.1002/grl.50256>, 2013.
- Lauritzen, P. H., Mirin, A., Truesdale, J., Raeder, K., Anderson, J., Bacmeister, J., and Neale, R. B.: Implementation of new diffusion/filtering operators in the CAM-FV dynamical core, *Int. J. High. Perform. C.*, 26, 63–73, <https://doi.org/10.1177/1094342011410088>, 2011.
- Leith, C. E.: Atmospheric Predictability and Two-Dimensional Turbulence, *J. Atmos. Sci.*, 28, 145–161, [https://doi.org/10.1175/1520-0469\(1971\)028<0145:APATDT>2.0.CO;2](https://doi.org/10.1175/1520-0469(1971)028<0145:APATDT>2.0.CO;2), 1971.
- Lindzen, R. S.: Turbulence and stress owing to gravity wave and tidal breakdown, *J. Geophys. Res.*, 86, 9707–9714, 1981.
- Lott, F., and Guez, L.: A stochastic parameterization of the gravity waves due to convection and its impact on the equatorial stratosphere, *J. Geophys. Res.-Atmos.*, 118, 8897–8909, <https://doi.org/10.1002/jgrd.50705>, 2013.

- Lott, F., Fairhead, L., Hourdin, F., and Levan, P.: The stratospheric version of LMDz: dynamical climatologies, arctic oscillation, and impact on the surface climate, *Clim. Dynam.*, 25, 851–868, <https://doi.org/10.1007/s00382-005-0064-x>, 2005.
- Lott, F., Guez, L., and Maury, P.: A stochastic parameterization of non-orographic gravity waves, Formalism and impact on the equatorial stratosphere, *Geophys. Res. Lett.*, 39, L06807, <https://doi.org/10.1029/2012GL051001>, 2012.
- Lott, F., Denvil, S., Butchart, N., Cagnazzo, C., Giorgetta, M., Hardiman, S., Manzini, E., Krishmer, T. T., Duvel, J.-P., Maury, P., Scinocca, J., Watanabe, S., and Yukimoto, S.: Kelvin and Rossby gravity wave packets in the lower stratosphere of some high-top CMIP5 models, *J. Geophys. Res.*, 119, 2156–2173, <https://doi.org/10.1002/2013JD020797>, 2014.
- Manzini, E., Cagnazzo, C., Fogli, P. G., Bellucci, A., and Müller, W. A.: Stratosphere-troposphere coupling at inter-decadal time scales: Implications for the North Atlantic Ocean, *Geophys. Res. Lett.*, 39, L05801, <https://doi.org/10.1029/2011GL050771>, 2012.
- Marsh, D. R., Mills, M. J., Kinnison, D. E., Lamarque, J.-F., Calvo, N., and Polvani, L. M.: Climate change from 1850 to 2005 simulated in CESM1(WACCM), *J. Climate*, 26, 7372–7391, <https://doi.org/10.1175/JCLI-D-12-00558.1>, 2013.
- McCormack, J. P., Eckermann, S. D., and Hogan, T. F.: Generation of a quasi-biennial oscillation in an NWP model using a stochastic gravity wave drag parameterization, *Mon. Weather Rev.*, 143, 2121–2147, <https://doi.org/10.1175/MWR-D-14-00208.1>, 2015.
- Mellor, G. L. and Yamada, T.: Development of a turbulence closure model for geostrophic fluid problems, *Rev. Geophys.*, 20, 851–875, <https://doi.org/10.1029/RG020i004p00851>, 1982.
- Molod, A., Takacs, L., Suarez, M., and Bacmeister, J.: Development of the GEOS-5 atmospheric general circulation model: evolution from MERRA to MERRA2, *Geosci. Model Dev.*, 8, 1339–1356, <https://doi.org/10.5194/gmd-8-1339-2015>, 2015.
- Morgenstern, O., Hegglin, M. I., Rozanov, E., O'Connor, F. M., Abraham, N. L., Akiyoshi, H., Archibald, A. T., Bekki, S., Butchart, N., Chipperfield, M. P., Deushi, M., Dhomse, S. S., Garcia, R. R., Hardiman, S. C., Horowitz, L. W., Jöckel, P., Josse, B., Kinnison, D., Lin, M., Mancini, E., Manyin, M. E., Marchand, M., Maréchal, V., Michou, M., Oman, L. D., Pitari, G., Plummer, D. A., Revell, L. E., Saint-Martin, D., Schofield, R., Stenke, A., Stone, K., Sudo, K., Tanaka, T. Y., Tilmes, S., Yamashita, Y., Yoshida, K., and Zeng, G.: Review of the global models used within phase 1 of the Chemistry–Climate Model Initiative (CCMI), *Geosci. Model Dev.*, 10, 639–671, <https://doi.org/10.5194/gmd-10-639-2017>, 2017.
- Newman, P. A., Coy, L., Pawson, S., and Lait, L. R.: The anomalous change in the QBO in 2015–2016, *Geophys. Res. Lett.*, 43, 8791–8797, <https://doi.org/10.1002/2016GL070373>, 2016.
- Nordeng, T. E.: Extended versions of the convective parametrization scheme at ECMWF and their impact on the mean and transient activity of the model in the tropics, *Tech. Rep.*, 206, ECWMF, 1994.
- Orr, A., Bechtold, P., Scinocca, J., Ern, M., and Janiskova, M.: Improved middle atmosphere climate and forecasts in the ECMWF model through a nonorographic gravity wave drag parameterization, *J. Climate*, 23, 5905–5926, <https://doi.org/10.1175/2010JCLI3490.1>, 2010.
- Osprey, S., Bell, C., Cionni, I., Eyring, V., and Lamarque, J.-F.: AC&C/SPARC Ozone Database for CMIP5 – Hitop Version [Data set], Zenodo, <https://doi.org/10.5281/zenodo.1035142>, 2011.
- Osprey, S. M., Gray, L. J., Hardiman, S. C., Butchart, N., and Hinton, T. J.: Stratospheric variability in twentieth-century CMIP5 simulations of the Met Office climate model: high top versus low top, *J. Climate*, 26, 1595–1606, <https://doi.org/10.1175/JCLI-D-12-00147.1>, 2013.
- Osprey, S. M., Butchart, N., Knight, J. R., Scaife, A. A., Hamilton, K., Anstey, J. A., Schenzinger, V., and Zhang, C.: An unexpected disruption of the atmospheric quasi-biennial oscillation, *Science*, 353, 1424–1427, [10.1126/science.aah4156](https://doi.org/10.1126/science.aah4156), 2016.
- Pohlmann, H., Müller, W. A., Kulkarni, K., Kameswarrao, M., Matei, D., Vamborg, F. S. E., Kadow, C., Illing, S., and Marotzke, J.: Improved forecast skill in the tropics in the new MiKlip decadal climate predictions, *Geophys. Res. Lett.*, 40, 5798–5802, <https://doi.org/10.1002/2013GL058051>, 2013.
- Rajendran, K., Moroz, I. M., Read, P. L., and Osprey, S. M.: Synchronisation of the equatorial QBO by the annual cycle in tropical upwelling in a warming climate, *Q. J. Roy. Meteor. Soc.*, 142, 1111–1120, <https://doi.org/10.1002/qj.2714>, 2016.
- Richter, J. H., Sassi, F., and Garcia, R. R.: Towards a physically based gravity wave source parameterization in a general circulation model, *J. Atmos. Sci.*, 67, 136–156, <https://doi.org/10.1175/2009JAS3112.1>, 2010.
- Richter, J. H., Solomon, A., and Bacmeister, J. T.: On the simulation of the quasi-biennial oscillation in the Community Atmosphere Model, version 5, *J. Geophys. Res.-Atmos.*, 119, 3045–3062, <https://doi.org/10.1002/2013JD021122>, 2014.
- Rind, D., Jonas, J., Balachandran, N. K., Schmidt, G. A., and Lean, J.: The QBO in two GISS global climate models: 1. Generation of the QBO, *J. Geophys. Res.-Atmos.*, 119, 8798–8824, <https://doi.org/10.1002/2014JD021678>, 2014.
- Roeckner, E., Bäuml, G., Bonaventura, L., Brokopf, R., Esch, M., Giorgetta, M., Hagemann, S., Kirchner, I., Kornbluh, L., Manzini, E., Rhodin, A., Schlese, U., Schulzweida, U., and Tompkins, A.: The atmospheric general circulation model ECHAM5 – Part I: Model description, *Tech. Rep.* 349, Max-Planck-Institut für Meteorologie, Hamburg, Germany, 2003.
- Scaife, A. A., Butchart, N., Warner, C. D., Stainforth, D., Norton W., and Austin, J.: Realistic quasi-biennial oscillations in a simulation of the global climate, *Geophys. Res. Lett.*, 27, 3481–3484, <https://doi.org/10.1029/2000GL011625>, 2000.
- Scaife, A. A., Butchart, N., Warner, C. D., and Swinbank, R.: Impact of a spectral gravity wave parameterization on the stratosphere in the Met Office Unified Model, *J. Atmos. Sci.*, 59, 1473–1489, [https://doi.org/10.1175/1520-0469\(2002\)059<1473:IOASGW>2.0.CO;2](https://doi.org/10.1175/1520-0469(2002)059<1473:IOASGW>2.0.CO;2), 2002.
- Scaife, A. A., Athanassiadou, M., Andrews, M., Arribas, A., Baldwin, M., Dunstone, N., Knight, J., MacLachlan, C., Manzini, E., Müller, W. A., Pohlmann, H., Smith, D., Stockdale, T., and Williams, A.: Predictability of the quasi-biennial oscillation and its northern winter teleconnection on seasonal to decadal timescales, *Geophys. Res. Lett.*, 41, 1752–1758, <https://doi.org/10.1002/2013GL059160>, 2014.
- Schenzinger, V., Osprey, S., Gray, L., and Butchart, N.: Defining metrics of the Quasi-Biennial Oscillation in global

- climate models, *Geosci. Model Dev.*, 10, 2157–2168, <https://doi.org/10.5194/gmd-10-2157-2017>, 2017.
- Schirber, S., Manzini, E., and Alexander, M. J.: A convection-based gravity wave parameterization in a general circulation model: Implementation and improvements on the QBO, *J. Adv. Model. Earth Sy.*, 6, 264–279, <https://doi.org/10.1002/2013MS000286>, 2014.
- Schirber, S., Manzini, E., Krismer, T., and Giorgetta, M.: The quasi-biennial oscillation in a warmer climate: sensitivity to different gravity wave parameterizations, *Clim. Dynam.*, 45, 825–836, <https://doi.org/10.1007/s00382-014-2314-2>, 2015.
- Schmidt, H., Rast, S., Bunzel, F., Esch, M., Giorgetta, M., Kinne, S., Krismer, T., Stenchikov, G., Timmreck, C., Tomassini, L., and Walz, M.: Response of the middle atmosphere to anthropogenic and natural forcings in the CMIP5 simulations with the Max Planck Institute Earth system model, *J. Adv. Model. Earth Sy.*, 5, 98–116, <https://doi.org/10.1002/jame.20014>, 2013.
- Scinocca, J. F.: An accurate spectral nonorographic gravity wave drag parameterization for general circulation models, *J. Atmos. Sci.*, 60, 667–682, [https://doi.org/10.1175/1520-0469\(2003\)060<0667:AASNGW>2.0.CO;2](https://doi.org/10.1175/1520-0469(2003)060<0667:AASNGW>2.0.CO;2), 2003.
- Scinocca, J. F., McFarlane, N. A., Lazare, M., Li, J., and Plummer, D.: Technical Note: The CCCma third generation AGCM and its extension into the middle atmosphere, *Atmos. Chem. Phys.*, 8, 7055–7074, <https://doi.org/10.5194/acp-8-7055-2008>, 2008.
- Serva, F., Cagnazzo, C., Riccio, A., and Manzini, E.: Impact of a stochastic nonorographic gravity wave parameterization on the stratospheric dynamics of a General Circulation Model, *J. Adv. Model. Earth Sy.*, submitted, 2018.
- Shepherd, T. G., Plumb, R. A., and Wofsy, S. C.: Preface to *J. Atmos. Sci.* special issue on the Antarctic stratospheric sudden warming and split ozone hole of 2002, *J. Atmos. Sci.*, 62, 565–566, <https://doi.org/10.1175/JAS-9999.1>, 2005.
- Shibata, K. and Deushi, M.: Partitioning between resolved wave forcing and unresolved gravity wave forcing to the quasi-biennial oscillation as revealed with a coupled chemistry-climate model, *Geophys. Res. Lett.*, 32, L12820, <https://doi.org/10.1029/2005GL022885>, 2005.
- Shutts, G. J. and Vosper, S. B.: Stratospheric gravity waves revealed in NWP model forecasts, *Q. J. Roy. Meteor. Soc.*, 137, 303–317, <https://doi.org/10.1002/qj.763>, 2011.
- Strahan, S. E., Oman, L. D., Douglass, A. R., and Coy, L.: Modulation of Antarctic vortex composition by the quasi-biennial oscillation, *Geophys. Res. Lett.*, 42, 4216–4223, <https://doi.org/10.1002/2015GL063759>, 2015.
- Takahashi, M.: Simulation of the stratospheric quasi-biennial oscillation using a general circulation model, *Geophys. Res. Lett.*, 23, 661–664, 1996.
- Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An overview of CMIP5 and the experiment design, *B. Am. Meteorol. Soc.*, 93, 485–498, <https://doi.org/10.1175/BAMS-D-11-00094.1>, 2012.
- Tiedtke, M.: A comprehensive mass flux scheme for cumulus parameterization in large-scale models, *Mon. Weather Rev.*, 117, 1779–1800, [https://doi.org/10.1175/1520-0493\(1989\)117<1779:ACMFSF>2.0.CO;2](https://doi.org/10.1175/1520-0493(1989)117<1779:ACMFSF>2.0.CO;2), 1989.
- Walters, D., Boutle, I., Brooks, M., Melvin, T., Stratton, R., Vosper, S., Wells, H., Williams, K., Wood, N., Allen, T., Bushell, A., Copsey, D., Earnshaw, P., Edwards, J., Gross, M., Hardiman, S., Harris, C., Heming, J., Klingaman, N., Levine, R., Manners, J., Martin, G., Milton, S., Mittermaier, M., Morcrette, C., Riddick, T., Roberts, M., Sanchez, C., Selwood, P., Stirling, A., Smith, C., Suri, D., Tennant, W., Vidale, P. L., Wilkinson, J., Willett, M., Woolnough, S., and Xavier, P.: The Met Office Unified Model Global Atmosphere 6.0/6.1 and JULES Global Land 6.0/6.1 configurations, *Geosci. Model Dev.*, 10, 1487–1520, <https://doi.org/10.5194/gmd-10-1487-2017>, 2017a.
- Walters, D., Baran, A., Boutle, I., Brooks, M., Earnshaw, P., Edwards, J., Furtado, K., Hill, P., Lock, A., Manners, J., Morcrette, C., Mulcahy, J., Sanchez, C., Smith, C., Stratton, R., Tennant, W., Tomassini, L., Van Weverberg, K., Vosper, S., Willett, M., Browse, J., Bushell, A., Dalvi, M., Essery, R., Gedney, N., Hardiman, S., Johnson, B., Johnson, C., Jones, A., Mann, G., Milton, S., Rumbold, H., Sellar, A., Ujiie, M., Whitall, M., Williams, K., and Zerroukat, M.: The Met Office Unified Model Global Atmosphere 7.0/7.1 and JULES Global Land 7.0 configurations, *Geosci. Model Dev. Discuss.*, <https://doi.org/10.5194/gmd-2017-291>, in review, 2017b.
- Warner, C. D. and McIntyre, M. E.: On the propagation and dissipation of gravity wave spectra through a realistic middle atmosphere, *J. Atmos. Sci.*, 53, 3213–3235, 1996.
- Warner, C. D. and McIntyre, M. E.: Toward an ultra-simple spectral gravity wave parameterization for general circulation models, *Earth Planets Space*, 51, 475–484, 1999.
- Watanabe, S., Hajima, T., Sudo, K., Nagashima, T., Takemura, T., Okajima, H., Nozawa, T., Kawase, H., Abe, M., Yokohata, T., Ise, T., Sato, H., Kato, E., Takata, K., Emori, S., and Kawamiya, M.: MIROC-ESM 2010: model description and basic results of CMIP5-20c3m experiments, *Geosci. Model Dev.*, 4, 845–872, <https://doi.org/10.5194/gmd-4-845-2011>, 2011.
- WMO: Scientific Assessment of Ozone Depletion 2010. Global Ozone Research and Monitoring Project Report, 52. World Meteorological Organisation, Geneva, Switzerland, 2011.
- WMO: Scientific Assessment of Ozone Depletion 2014. Global Ozone Research and Monitoring Project Report, 55. World Meteorological Organisation, Geneva, Switzerland, 2015.
- Yao, W. and Jablonowski, C.: Idealized quasi-biennial oscillations in an ensemble of dry GCM dynamical cores, *J. Atmos. Sci.*, 72, 2201–2226, <https://doi.org/10.1175/JAS-D-14-0236.1>, 2015.
- Yoshimura, H., Mizuta, R., and Murakami, H.: A spectral cumulus parameterization scheme interpolating between two convective updrafts with semi-Lagrangian calculation of transport by compensatory subsidence, *Mon. Weather Rev.*, 14, 597–621, <https://doi.org/10.1175/MWR-D-14-00068.1>, 2015.
- Yukimoto, S., Adachi, Y., Hosaka, M., Sakami, T., Yoshimura, H., Hirabara, M., Tanaka, T. Y., Shindo, E., Tsujino, H., Deushi, M., Mizuta, R., Yabu, S., Obata, A., Nakano, H., Koshiro, T., Ose, T., and Kitoh, A.: A new global climate model of the Meteorological Research Institute: MRI-CGCM3 -Model description and basic performance-, *J. Meteorol. Soc. Jpn.*, 90A, 23–64, <https://doi.org/10.2151/jmsj.2012-A02>, 2012.
- Zhang, G. and McFarlane, N. A.: Sensitivity of climate simulations to the parameterization of cumulus convection in the Canadian Climate Centre general circulation model, *Atmos.-Ocean*, 33, 407–446, <https://doi.org/10.1080/07055900.1995.9649539>, 1995.