

Assessment of the Met Office Global Coupled model version 4 (GC4) configurations

The Forecasting Research Technical Report No: 661

<https://doi.org/10.62998/uzui3766>

June 2024

Prince Xavier¹, Martin Willett¹, Tim Graham¹, Paul Earnshaw¹, Dan Copsey¹, Charline Marzin¹, Alistair Sellar¹, Duncan Ackerley¹, Oscar Alves², Ed Blockley¹, Alejandro Bodas-Salcedo¹, Andrew Bushell¹, Neal Butchart¹, Daley Calvert¹, Jeong-A Cho³, Dan Copsey¹, Catherine de Burgh-Day², John Edwards¹, Paul Earnshaw¹, Kalli Furtado¹, Paul Field¹, Catherine Guiavarch¹, Sam Hardy⁶, Chris Harris¹, Karen Heywood⁹, Julian Heming¹, Harry Hendon², Helene Hewitt¹, Pat Hyder¹, Yu-Kyung Hyun³, Seung-Hwon Hyun³, Sarah Ineson¹, Richard Jones¹, Jiyeong Kim³, Ki-Young Kim³, Nick Klingaman⁴, Richard Levine¹, Sang-Min Lee³, Kyriaki Lekakou⁹, Adrian Lock¹, Gill Martin¹, Pierre Mathiot¹, Alex Megann⁷, Andrew Meijers⁸, Ja-Yeon Moon³, Olaf Morgenstern⁵, Rachel North¹, George Nurser⁷, Yeon-Hee Park³, Leighton Regayre^{1, 9, 10}, Malcolm Roberts¹, José Rodriguez¹, Jeff Ridley¹, Rick Rawlins¹, Bablu Sinha⁷, Beoncheol Shin³, Adrian Semple¹, Dave Storkey¹, David Stephens¹, Taehyoun Shim³, Lorenzo Tomassini¹, Yoko Tsushima¹, Helen Titley¹, Warren Tennant¹, Vidya Varma⁵, Michael Vellinga¹, Graham Weedon¹, Keith Williams¹, Young-Min Yang³, Maggie Zhao², Xiaobing Zhou², Hongyan Zhu²

¹ Met Office, Exeter, UK

² The Bureau of Meteorology, Melbourne, Australia

³ National Institute of Meteorological Sciences, Korea Meteorological Administration, Jeju, Korea

⁴ National Centre for Atmospheric Science (NCAS), University of Reading, Reading, United Kingdom

⁵ NIWA - The National Institute of Water and Atmospheric Research, New Zealand

⁹ Centre for Environmental Modelling and Computation, University of Leeds, Leeds, United Kingdom

¹⁰ School of Earth and Environment, University of Leeds, Leeds, United Kingdom

⁷ National Oceanography Centre, Southampton, United Kingdom

⁸ Joint Marine Modelling Programme (JMMP)

Disclaimer

- This document is published by the Met Office on behalf of the Secretary of State for Business, Energy and Industrial Strategy, HM Government, UK. Its content is covered by © Crown Copyright 2024.
- This document is published specifically for the readership and use of Customer Name and may not be used or relied upon by any third party, without the Met Office's express written permission.
- The Met Office aims to ensure that the content of this document is accurate and consistent with its best current scientific understanding. However, the science which underlies meteorological forecasts and climate projections is constantly evolving. Therefore, any element of the content of this document which involves a forecast or a prediction should be regarded as our best possible guidance, but should not be relied upon as if it were a statement of fact. To the fullest extent permitted by applicable law, the Met Office excludes all warranties or representations (express or implied) in respect of the content of this document.
- Use of the content of this document is entirely at the reader's own risk. The Met Office makes no warranty, representation or guarantee that the content of this document is error free or fit for your intended use.
- Before taking action based on the content of this document, the reader should evaluate it thoroughly in the context of his/her specific requirements and intended applications.
- To the fullest extent permitted by applicable law, the Met Office, its employees, contractors or subcontractors, hereby disclaim any and all liability for loss, injury or damage (direct, indirect, consequential, incidental or special) arising out of or in connection with the use of the content of this document including without limitation any and all liability:
 - relating to the accuracy, completeness, reliability, availability, suitability, quality, ownership, non-infringement, operation, merchantability and fitness for purpose of the content of this document;
 - relating to its work procuring, compiling, interpreting, editing, reporting and publishing the content of this document; and
 - resulting from reliance upon, operation of, use of or actions or decisions made on the basis of, any facts, opinions, ideas, instructions, methods, or procedures set out in this document.
- This does not affect the Met Office's liability for death or personal injury arising from the Met Office's negligence, nor the Met Office's liability for fraud or fraudulent misrepresentation, nor any other liability which cannot be excluded or limited under applicable law.
- If any of these provisions or part provisions are, for any reason, held to be unenforceable, illegal or invalid, that unenforceability, illegality or invalidity will not affect any other provisions or part provisions which will continue in full force and effect.

Contents

1	Executive summary	2
2	Introduction	3
2.1	Model experiments	4
3	Global Characteristics	6
3.1	Ocean and sea ice	6
3.2	Global land and precipitation	8
3.3	Hydrological cycle	10
3.4	NWP trials	10
3.5	Aerosol	12
4	Large-scale variability	15
4.1	El Niño Southern Oscillation (ENSO)	15
4.2	Madden-Julian Oscillation (MJO)	15
4.3	Mid-latitude storm tracks	16
5	Spatial Variability	19
5.1	Near real time N1280 simulations	19
5.2	Spatial power spectra	22
6	Regional aspects	23
6.1	South Asian Summer Monsoon (SASM)	23
6.2	East Asian monsoon	24
6.3	Maritime Continent	25
7	Summary	27
7.1	Improvements in GC4	27
7.2	Neutral/slightly improved in GC4, yet a significant issue	28
7.3	Degradations in GC4	29
7.4	Towards GC5	29

Chapter 1

Executive summary

The Global Coupled model version 4 (GC4) is an upgraded configuration of the MetUM system, building upon its predecessor, GC3.0/3.1. It incorporates various improvements and changes in the atmospheric and land components (Global Atmosphere 8 and Global Land 9 - GA8GL9) while keeping the ocean component (Global Ocean 6 - GO6) unchanged, except for minor bug fixes. The GC4 model introduces several enhancements, such as the drag package for land surface and hydrology, improvements in radiation and large-scale precipitation parametrisations, advancements in the boundary layer and convection representation (including the prognostic-based convective entrainment rate - ProgEnt), and updates in aerosol properties. Additionally, the inclusion of a multi-grid solver in the dynamics module aims to improve model stability and reduce computational costs.

Key improvements in GC4 include better representation of the diurnal cycle of convection over land, reduced Southern Ocean warm bias, increased rainfall over India during the JJA season, improved distribution of precipitation, enhanced representation of low-medium clouds over Northern Europe, and positive impacts of atmosphere-ocean coupling on NWP scores. However, challenges and areas for further improvement persist, including excessive global precipitation, warm biases over coastal regions of East Asia, wet biases over East Asia, weak cloud forcing over certain regions, hydrological cycle discrepancies, biases in gross primary productivity, persistent Southern Ocean biases, enhanced warming and weakened trade winds in the equatorial east Pacific, excessive surface warming in the North Atlantic, weakening of monsoon low-pressure systems and tropical cyclones, drying over Africa, and excessive thick cloud biases in mid-latitudes. The next version of GC (GC5) will attempt to address some of these biases in the next development and assessment cycle with inputs from relevant evaluation groups and partners.

Chapter 2

Introduction

Global Coupled model version 4 (GC4) is the MetUM configuration built on its predecessor GC3.0/3.1 ([Williams et al., 2018](#)). The atmospheric and land components of GC4 (Global Atmosphere 8 and Global Land 9 - GA8GL9) are built on GA7GL7, which was released in January 2016. GA8GL9 incorporates most of the science changes added to GL8, GA7.1, GA7.2, and GA7.2.1 ([Willett et al., 2024](#)) with the exceptions on system-specific tuning or where a newer science option has made a change in previous branch configurations redundant. A full list of the tickets included in GA8GL9 are listed in the GA8GL9 documentation report (Willett et al, in preparation). A brief description of a few important changes included in GA8GL9 is below:

- Land surface and hydrology: A drag package has been developed that removes the need to use an aggregate surface tile in the numerical weather prediction (NWP) configuration. This package has been shown to improve large-scale circulation and reduce differences between the representation of the land surface in global and regional configurations. The changes include an improved representation of orography with reduced smoothing of the mean orography, new land coverage ancillaries from the European Space Agency - Climate Change Initiative (CCI), improvements to the thermal and drag properties of the land surface, and a new distributed form drag. Improvements in sea drag and the representation of snow grains have been carried forward from GL8.
- Radiation: The [Liu et al. \(2008\)](#) spectral dispersion, which was included in GA7.1, improves the simulation of the aerosol indirect effect and consequently reduces the bias in the effective radius of the cloud droplets.
- Large-scale precipitation: A new riming parametrisation has been introduced that increases the liquid water content at mid- and high-latitudes and consequently reduces shortwave flux biases in the Southern Ocean.
- Boundary layer: Several changes have been made to the boundary layer (BL) to improve the diagnosis and representation of shear-driven BLs. The numerical stability of stable boundary layers is improved due to these changes.

- **Convection:** The “prognostic based convective entrainment rate” (ProgEnt) adds memory of sub-grid activity into the convection parameterisation ([Willett and Whittall, 2017](#)). This improves the diurnal cycle of convection over land, produces realistically more intense instantaneous precipitation rates, and more realistic spatial structures, and reduces the overprediction of light rain. Additional changes include the convective snow and convective rain melting and freezing, respectively, over a physical depth rather than a single model level. The computational stability of the Gregory-Kershaw convective momentum transport is improved and applied to deep convection. The coupling between convection and dynamics is improved through the time-smoothing of convective temperature and humidity increments. This largely eliminates the dynamical effects of any convective intermittency and reduces the intermittency itself. An improvement to the convective ascent termination condition which prevents convection from reaching physically unrealistic depths and consequently reduces stratospheric humidity biases is also implemented.
- **Aerosols changes** include black carbon and aerosol absorption properties, updated Dimethyl sulfide (DMS) concentrations (primarily emitted from the ocean, is a dominant natural precursor of non-sea-salt sulfate in the Pre-Industrial and pristine present-day (PD) atmospheres), and volcanic emissions are all carried forward from GA7.1. A new representation of marine organic aerosol replaces the simple scaling of the DMS emission in GA7.1.
- **Dynamics:** The multigrid solver is included in GA8. This reduces the cost of the pressure solver and can potentially improve the stability of the model.

The ocean component Global Ocean 6 (GO6) remains unchanged in GC4 compared to GC3.1 except for minor bug fixes which affect the conservation of all tracers. These bug fixes, however, have a neutral impact overall, and hence it may be reasonable to assume that all changes in GC4 are driven by changes to the atmospheric model.

2.1 Model experiments

The GC4 assessment process adopts a seamless strategy and a number of assessment experiments are conducted as NWP case studies, NWP trials (using both the GA8GL9 and GC4 configurations), AMIP simulations and coupled climate simulations with both CMIP5 and CMIP6 forcings. GC4 was initially developed using the CMIP5-based climate forcing datasets. The definition of a given GC constrains only the science settings of the model and is independent of the forcing.

The updated CMIP6 forcing includes changes to Greenhouse gases, Aerosol emissions and Ozone. As in GC3, concentrations of CO₂, CH₄, and N₂O are prescribed, while all CFCs are represented by a radiatively equivalent concentration of CFC-12 and all HFCs are represented by a radiatively equivalent concentration of HFC-134a. See [Sellar et al. \(2020\)](#) for details. Emissions of aerosol precursor gases

SO₂, DMS, monoterpene, as well as primary black carbon and organic carbon aerosols, are specified. Only anthropogenic and biomass burning emissions are updated, as CMIP6 does not provide natural emissions. Natural emissions of SO₂, DMS, and monoterpene are the same as in GC3.1, though DMS and monoterpene could be replaced with newer datasets in a future update.

CMIP6 ozone is applied as a 3D monthly field, rather than the zonal mean used in GC3.1. Note that unlike the CMIP6 production runs there is no redistribution of ozone below the tropopause ([Hardiman et al., 2019](#)). This decision is made on practical grounds, as the redistribution scheme was not designed to be used very widely and is not considered suitable for inclusion in a standard GC release. Up to 2014, the impact of the redistribution is small but non-negligible so the lack of ozone redistribution should be borne in mind when comparing against CMIP6 historical runs.

CMIP6 experiments include changing land use, but the CMIP6 ancillaries for GC4 are derived from IGBP land cover data and as a result, are very similar to IGBP in the present day. Using these in GC4 would be inconsistent with the move from IGBP to CCI land cover data in GC4. Additionally, the standard GC configuration uses the albedo-obs scheme which would suppress the albedo effect of land use change. It was therefore decided not to include any land use change in GC4 climate experiments, and leave land cover as prescribed by the CCI LC data together with the albedo-obs scheme switched on. This setup, therefore, differs from the CMIP6 GC3.1 experiments but is consistent with the historical runs performed with GC2.

GC4 assessment runs (with CMIP6 forcing) compared to the standard runs with CMIP5 forcing has the following changes: i) Using Easy Aerosol to model the effect of volcanoes on long-wave (LW) and short-wave (SW) radiation. ii) Using 2014 timeslice ancillary files for Ozone and all aerosol emissions. iii) Using constant 2014 values of CO₂, methane, N₂O, CFC-12 and HFC-134a.

GC4 will be delivered and be used with the CMIP6 forcing and hence the GC4 assessment workshop focused on a comparison between GC4 with CMIP6 forcing as a new configuration assessed against GC3.1 with CMIP5 forcing. However, the purpose of this report is to provide an evaluation of the modifications made to the physical model from GC3.1 to GC4, independent of the changes in forcing data from CMIP5 to CMIP6. The assessment is based on the results obtained from the second 50 years of the 100-year simulation of GC4 compared to GC3.1. These simulations were conducted using the N216 ORCA025 configuration with CMIP5 forcings. The results of two sets of standard NWP trials (Winter 2018-19 and Summer 2019) at N320 ORCA025 resolutions compared to the PS43 trial with GC3.2 (that is, the coupled version of the standard PS43 science) are also included in this report.

Chapter 3

Global Characteristics

The climate mean state derived from the years 51-100 of GC4 at N216O025 resolution with CMIP5 forcing is shown against a similar climatology of GC3.1 (Fig. 3.1). A comparison of the SSTs from these two runs with the ESA CCI SST data ([Merchant et al., 2019](#)) is also shown. GC3.1 generally tends to have cooler tropical and subtropical regions, especially in the west Pacific and in the west Atlantic sectors, where the warm anomalies reach up to 200 m depth. The southern ocean, on the other hand, is warmer by over 1°C in GC3.1 despite the fact that this bias was significantly reduced in the GC3.0 / 3.1 versions compared to GC2.0 ([Williams et al., 2018](#)). GC4 reduces the SST biases of the Southern Ocean further by about 50% overall, but with slightly warmer tropics and subtropical oceans by 0.2-0.3 degrees. Since the ocean model has not changed compared to GC3.1 it could be concluded that the improvements seen here are due to the atmospheric model changes through the changes in surface fluxes. These changes also contribute to a slight reduction in the annual mean tropical Sea Surface Salinity (SSS) over the Maritime Continent and the Atlantic (figure not shown). Some of the improvements over the Southern Ocean is attributed to the improvements in shortwave radiation at the surface due to a parametrisation for riming that increases supercooled liquid water cloud. GC4 demonstrates an increase in SST in the equatorial East Pacific, which may be related to the reduced equatorial Pacific wind stress. Furthermore, it shows better wind patterns in the Southern Ocean, which can be attributed to the reduced warm bias in this region, achieved through surface shortwave improvements.

3.1 Ocean and sea ice

With the ocean model being essentially the same except for some minor bug fixes between GC4 and GC3.1, the sea ice in GC3.1 and GC4 are very similar. There is little change in the northern hemisphere (NH) ice extent and area. The NH ice volume and thickness are increased. In some regions, this improves the thickness, but overall the sea ice volume increase represents a degradation. The southern hemisphere (SH) ice extent is increased at all times of the year, more so in winter than summer which

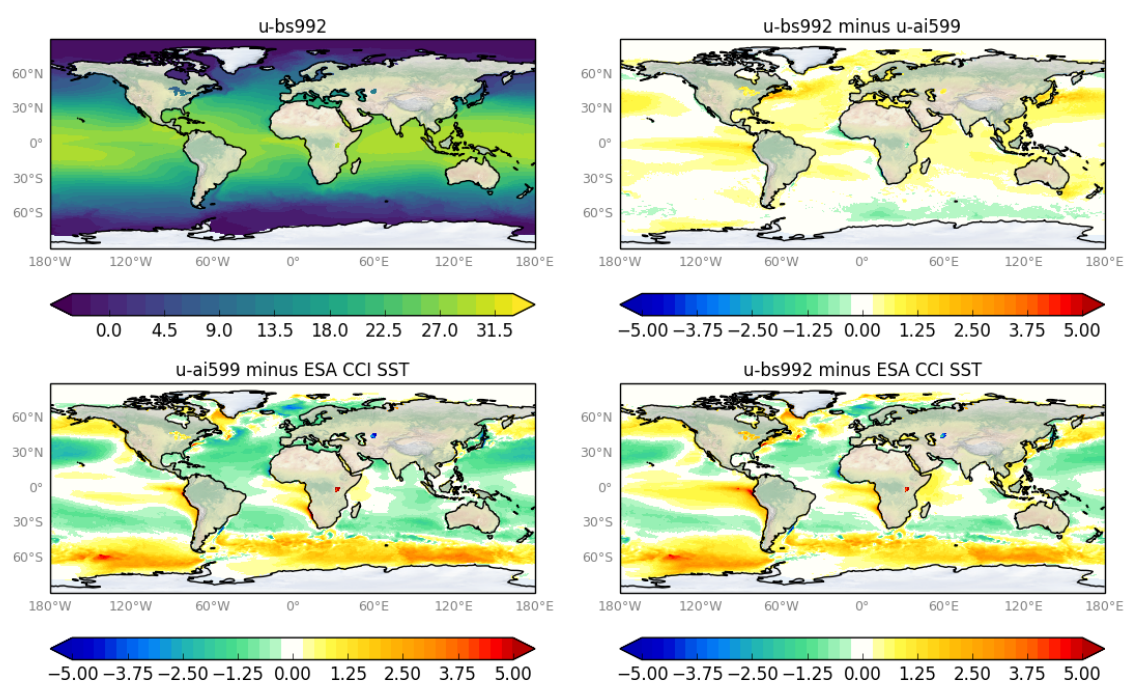


Figure 3.1: Annual mean SST (top left) from GC4, difference between GC3.1 SSTs and the ESA CCI SST (bottom left), difference between GC4 SSTs and the ESA CCI SST (bottom right) and GC4 minus GC3.1 (top right). Units are in $^{\circ}\text{C}$.

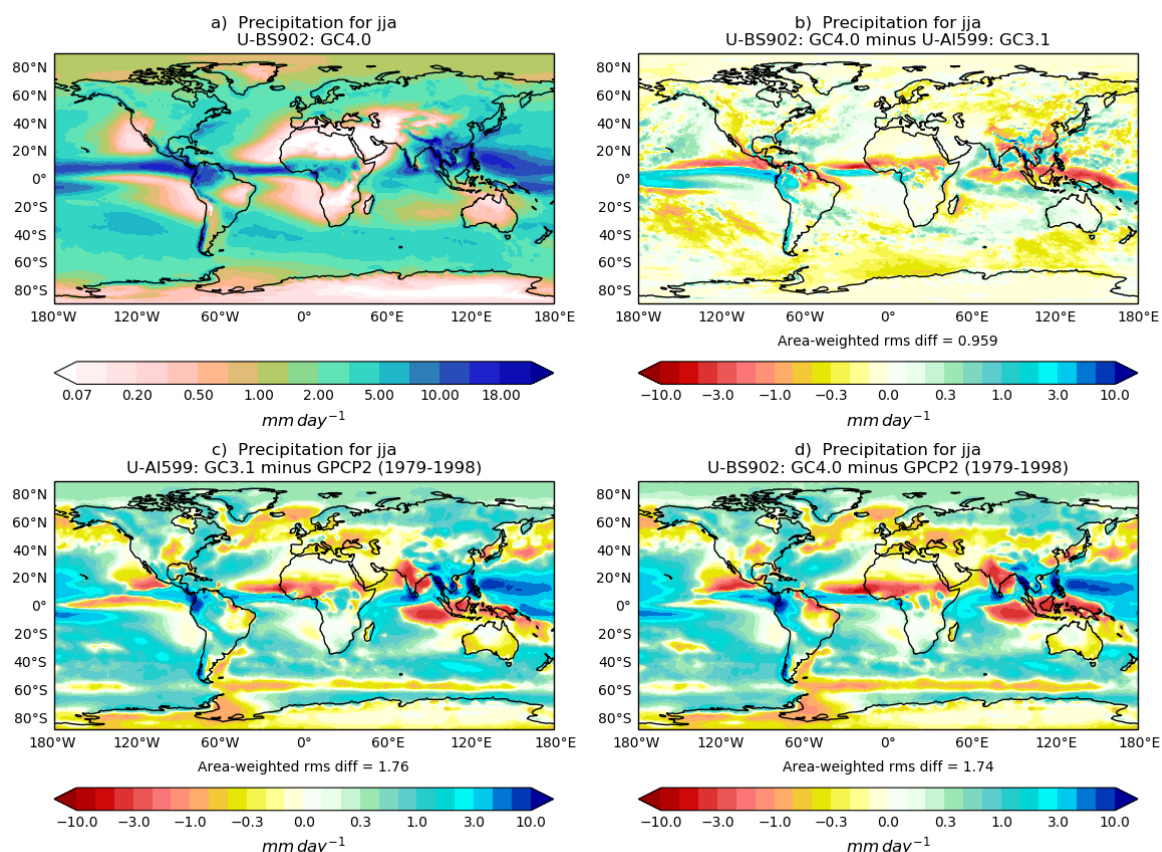


Figure 3.2: DJF mean air temperature (top left) from GC4, difference between GC3.1 SSTs and the ERA Interim reanalysis (bottom left), difference between GC4 SSTs and the ERA Interim reanalysis (bottom right) and GC4 minus GC3.1 (top right). Units are in K.

is an improvement. The SH ice volume is also slightly increased which is seen as an improvement.

Some of the ocean biases may be traced to the biases in wind stress at the surface, which is the primary source of momentum transfer from the atmosphere across the air-sea interface. GC4 shows a significant reduction in the equatorial Pacific wind stress compared to GC3.1 (and the global wind stress climatology observed from the Southampton Oceanography Centre (SOC) (Josey et al., 2002)). This degradation could result in a reduction in Ekman upwelling and could further contribute to warmer equatorial SSTs.

3.2 Global land and precipitation

The analysis of precipitation changes reveals a mixed pattern. Global precipitation on land is a better match to observations due to the improvement in cloud cover (using net downward SW radiation observations). Annually there is too much rain in northwestern South America and in June-August (JJA)

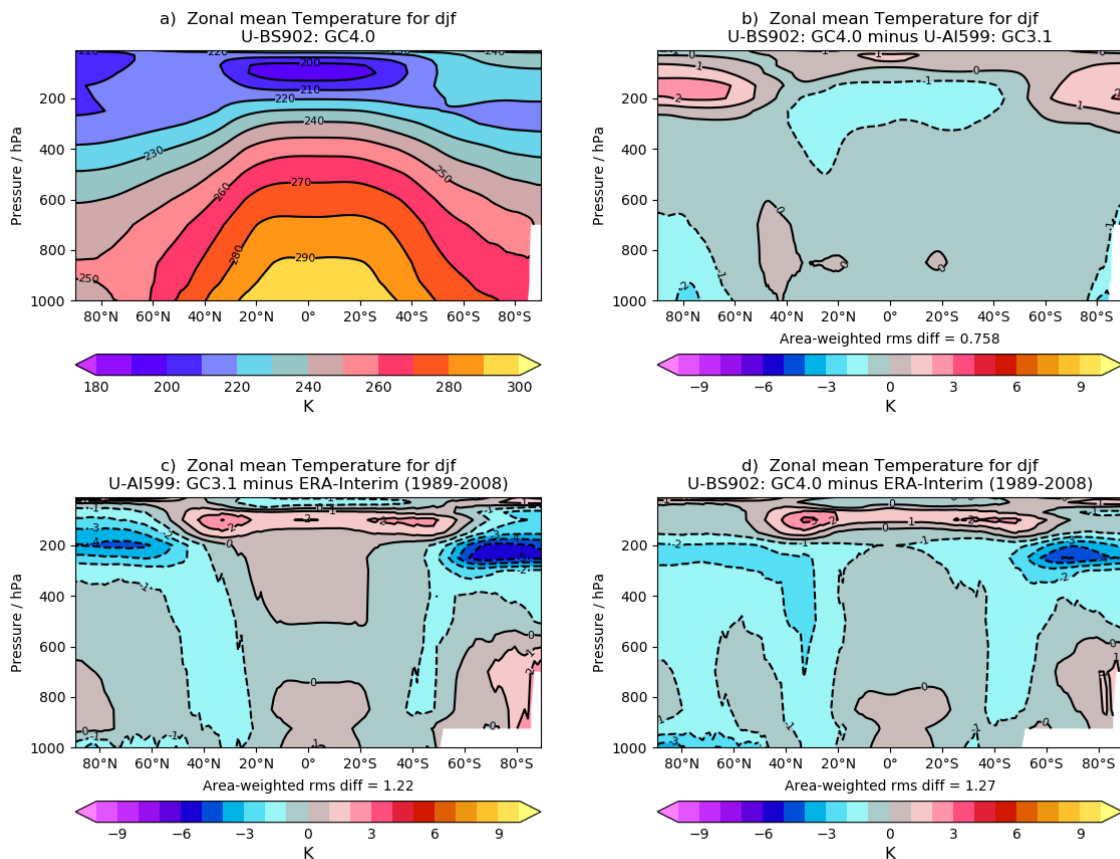


Figure 3.3: DJF mean air temperature (top left) from GC4, difference between GC3.1 SSTs and the ERA Interim reanalysis (bottom left), difference between GC4 SSTs and the ERA Interim reanalysis (bottom right) and GC4 minus GC3.1 (top right). Units are in K.

(Fig.3.2) and too little in east-central India, Sahel and NE South America. The lack of precipitation in NE South America is new (not found in GC3.1) and may be linked to lack of cloud and perhaps to the biases offshore.

There were only minor changes in latent heat. The soil moisture biases during JJA still exhibit spatial patterns that match the IGBP land cover. The modelled layer 1 soils in forests are too wet, while the bare soil, grassland, savanna, desert, and shrubland layer 1 soils are too dry. The gross primary productivity (GPP) biases compared to observations throughout the year follow the biases in precipitation: GPP is too high in NW South America and too low in NE South America and India. Land GPP is too high in cropland areas in the southeast United States, Europe, and southeast China. Global land GPP is still too low (85-90 v obs of 120 GT C yr^{-1}) – despite the use of $can_rad_mod = 6$, $diff_frac = 0.4$ in the GL9 Drag Package #435).

GC4 vs GC3.1 N216ORCA025

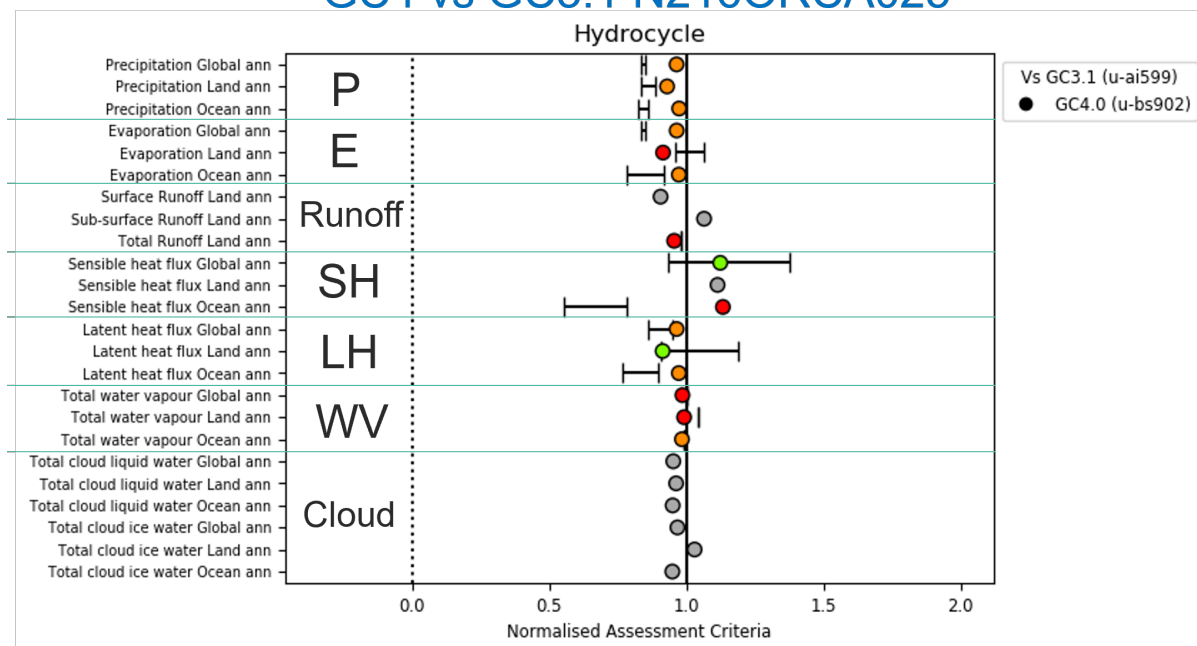


Figure 3.4: Annual mean hydrological parameters of GC4 (dots) compared with GC3.1 (displayed as the x=1 vertical line and the range of observed values (whiskers)). Green colours indicate improvements in GC4 compared to GC3.1 and in the observed range, amber for improved variables in GC4 compared to GC3.1 but still outside the observed range and red for variables that are deteriorated in GC4.

3.3 Hydrological cycle

The precipitation and evaporation rates experience a decline from approximately 3.15 to 3.05 mm/day, primarily attributed to the tuning of convective clouds. Notably, the reduction is more pronounced over land than over the ocean, indicating improved precipitation but worsened evaporation patterns. The changes in precipitation and evaporation have varying effects on the P-E (precipitation minus evaporation) values, leading to shifts in different directions. However, the overall difference remains within the range of $\pm 5 \times 10^{-4} \text{ mm day}^{-1}$. There is a decrease in cloud ice and cloud water, except over land where an increase in ice concentration is observed, predominantly in tropical regions. The total runoff experiences a decrease, primarily due to a higher proportion of sub-surface runoff. In contrast, there is an increase in sensible heat flux.

3.4 NWP trials

Coupled NWP trials for winter 2018-19 and summer 2019 as N320O025 configuration was compared against the PS43 trial with GC3.2 i.e. coupled version of standard PS43 science. This trial uses a weakly-coupled ocean-atmosphere data Assimilation via a forecast model. Full set of assessment plots are available at https://www-nwp.metoffice.com/~hadtd/ges/GC4_Win18_N320/imt/page.html. Fig.3.5

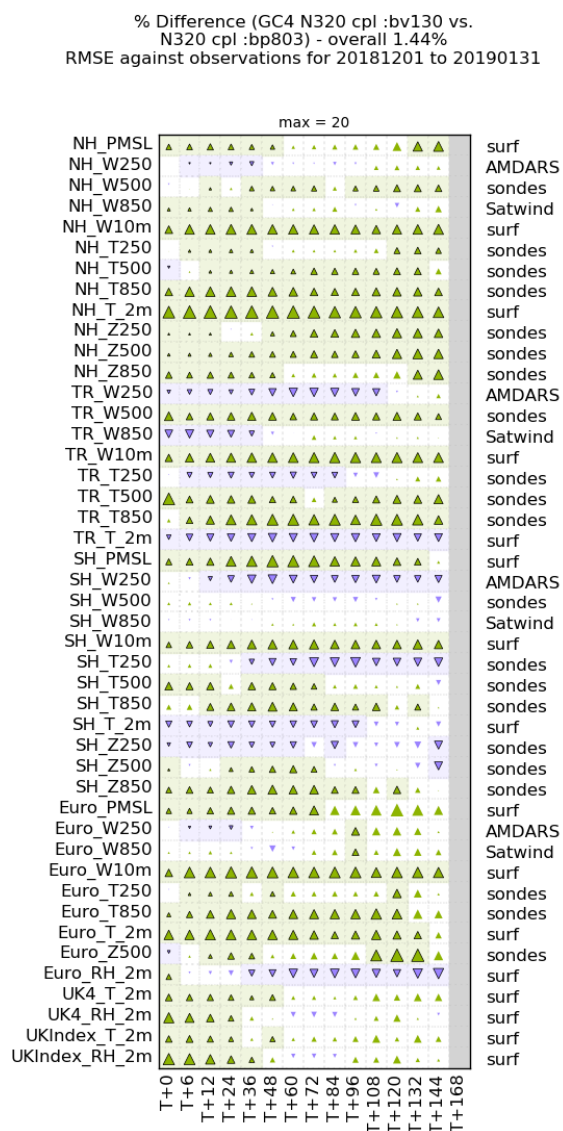


Figure 3.5: Scorecards of deterministic forecast RMSE for various parameters vs lead-time for the period 1 Dec 2018 to 31 Jan 2019 at N320O025 resolution. Green (purple) upwards (downwards) pointing triangles indicate improved (deteriorated) RMSE in GC4 compared to the GC3.2, and their size is proportional to the magnitude of change.

shows the RMSE scorecards for GC4 compared to the GC3.2 counterpart. Major improvements include the NH 2m temperature and near-surface wind speeds in both hemispheres when compared against observations. There are also minor degradations in the tropical 2m temperature and upper-level tropical winds. The impact of coupling alone is mostly positive when assessed against the atmosphere-only trials.

3.5 Aerosol

Model aerosol optical depth (AOD) agrees slightly better with observed values from the Aerosol Robotic Network (AERONET; [Holben et al. 1998](#)) for GC4 than GC3.1 (Pearson correlation coefficient increases from 0.64 to 0.66 for AOD at 870 nm across all sites). This modest improvement in simulated AOD is dominated by closer agreement with observations in marine environments with relatively high aerosol concentrations. Surface level sulfate aerosol concentrations are effectively unchanged from GC3.1 to GC4. However, organic carbon (OC) aerosol concentrations, which were biased low in GC3.1 compared to observed concentrations from the International Chemistry Experiment in the Arctic Lower Troposphere (ICEALOT [Russel et al. 2009](#); [Reddington et al. 2017](#)), are much higher in GC4 (now biased high), particularly in marine regions (e.g. Fig. 3.6). Revised representation of primary marine organic aerosol in GC4 ([Mulcahy et al., 2020](#)) likely causes these higher OC concentrations and improved AOD in marine regions.

Improvements in AOD and OC aerosol concentrations do not translate into improved aerosol size distribution comparisons. Higher overall aerosol concentrations in GC4 exacerbate high biases in GC3.1 accumulation mode aerosol in some regions (e.g. Fig. 3.7 a,b), in part because of excessive dust concentrations. Yet, multiple regions where aerosol concentrations were biased low in GC3.1 (e.g. rural mainland China; [Hu et al. 2012](#); [Reddington et al. 2017](#)) remain biased low in GC4 (Fig. 3.7 c,d). Additionally, observed nucleation mode aerosol concentrations are not always simulated in GC3.1 nor GC4. These results suggest a need to include a) missing aerosol species such as nitrate aerosol ([Jones et al., 2021](#)) to improve aerosol concentrations in relatively polluted regions, b) an improved representation aerosol scavenging (particularly of dust [Jones et al. 2022](#)) to prevent excessive long-distance transport into the remote Arctic and c) an improved representation of new particle formation in the troposphere to improve simulation of aerosol size distributions, in future GC versions.

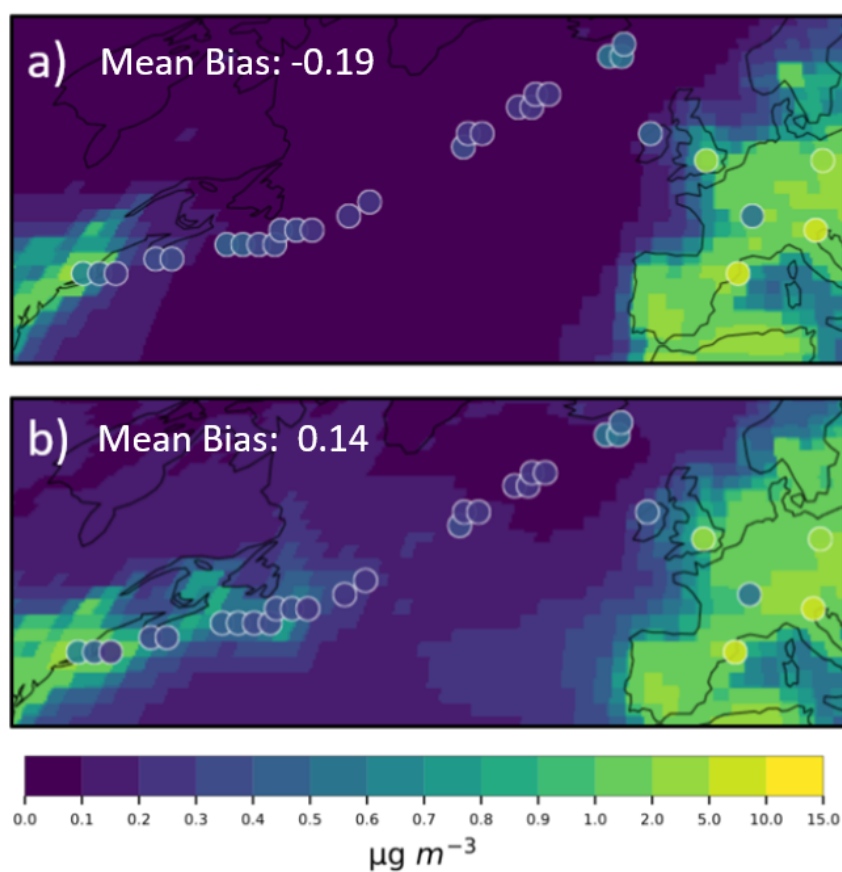


Figure 3.6: Organic aerosol concentrations in a) GC3.1 and b) GC4, compared to observations from ICEALOT [Russel et al. \(2009\)](#); [Reddington et al. \(2017\)](#). Background shading shows model March values between 2006 and 2010. Observed values for March 2008 are shown as shaded dots. Bias values are the means of observation-model comparisons over marine regions.

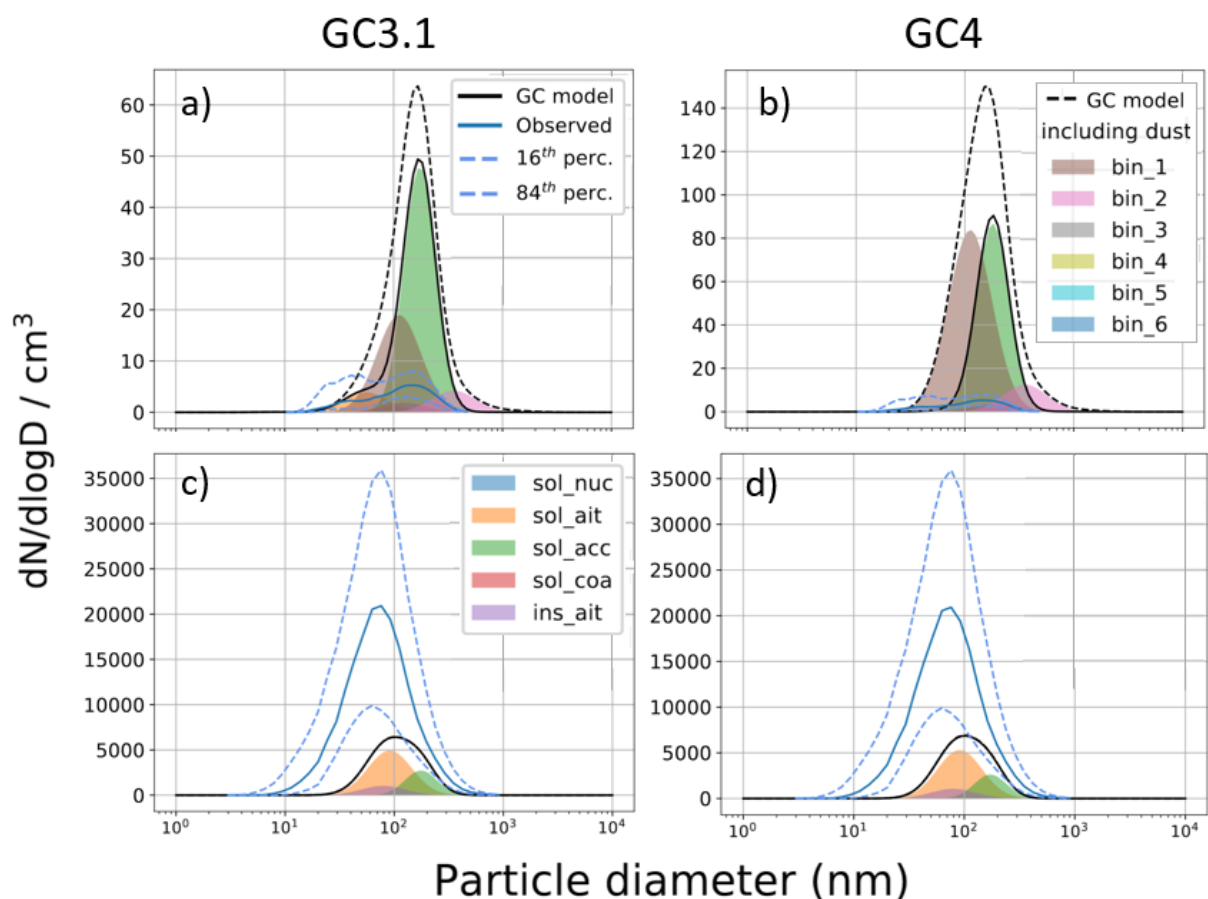


Figure 3.7: Aerosol number size distributions at a,b) the polar research station at ALERT for May 2011 (Leaitch et al., 2013; Reddington et al., 2017) and c,d) at the PRIDE rural mainland China station for July 2006 (Hu et al., 2012; Reddington et al., 2017). Model concentrations with and without dust contributions are shown alongside concentrations for aerosol size modes and six dust size bins. Observed values and percentiles that represent variability in the high time frequency observation data are shown. Note y-axes values are distinct for each sub-figure.

Chapter 4

Large-scale variability

4.1 El Niño Southern Oscillation (ENSO)

Fig.4.1 shows the annual mean SST difference between GC3.1 and HadISST observations and the difference between GC4 and GC3.1. Compared to GC3.1, GC4 exhibits several notable differences. Firstly, it demonstrates enhanced warming in the East Pacific and a reduction in trade winds. This is accompanied by weaker easterly wind stress over the Pacific region (not shown). Analyzing the east Pacific seasonal cycle of sea surface temperatures (SST), both GC4 and GC3.1 demonstrate a tendency to warm too quickly after reaching the September minimum. This increased amplitude of the ENSO cycle in GC4 can be attributed to physics changes.

In terms of spatial patterns, GC4 shows no westward extension, which is similar to GC3.1 and earlier versions of HadGEM3. However, a notable issue arises in the size of La Niña events, as they appear too large, with El Niño and La Niña events showing excessive symmetry. On a positive note, the phase locking of the El Niño-Southern Oscillation (ENSO) cycle is well reproduced, with the maximum amplitude occurring during November-December-January (NDJ) and the minimum amplitude during March-April-May (MAM). Rainfall patterns generally align with observations, although a weak response is noted over southern Africa. An unusual response is observed in India, where there is a wet pattern instead of the expected dry response to El Niño.

4.2 Madden-Julian Oscillation (MJO)

To evaluate the MJO simulation skill in GC3.1 and GC4, ERA5 reanalysis data were analyzed from November through the following April (NDJFMA) covering the years 1979 to 2019. The MJO was defined as 20-70 day anomalies derived by applying the Lanczos bandpass filter (Duchon, [year]) to the daily datasets. A composite life cycle of OLR associated with the MJO, ranging from phase 3 to phase

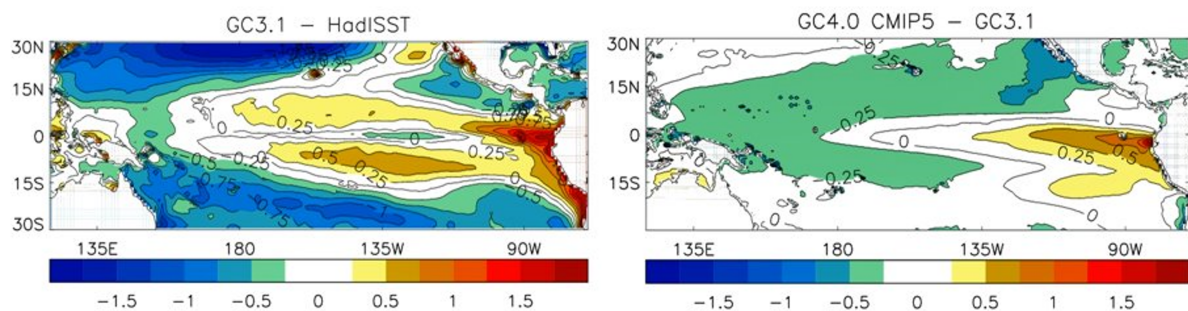


Figure 4.1: Annual mean SST differences (K) between GC3.1 and HadISST observations (left) and the difference between GC4 and GC3.1 (right) from the N216 O025 simulations.

6, was compared between GC3.1, GC4, and ERA5, as shown in Figure 4.2a-c.

The models successfully capture the eastward evolution of the MJO from the Indian Ocean to the western Pacific, with the exception of an overestimated convective anomaly over the Maritime Continent (MC) in phases 3 and 4. To assess the agreement between the model and ERA5, the pattern correlation coefficient (PCC) and the amplitude of OLR over the Indo-Pacific region (40°E - 140°W , 15°S - 15°N) were calculated for MJO phases 1 to 8 and are displayed in Figure 4.3. The OLR amplitude was computed as the standard deviation of composite OLR anomalies in the model divided by that of ERA5 (Taylor, 2001; Wang et al., 2020).

Both GC3.1 and GC4 demonstrate good agreement with the OLR pattern at all phases, with correlations greater than 0.6. More realistic patterns are simulated in phases 1-3 and 5-6, with pattern correlations exceeding 0.75. A value less than 1 indicates an underestimate, while a value greater than 1 indicates an overestimation of the OLR amplitude in the model. The OLR amplitude in both versions of the model shows comparable results with ERA5 at all phases. GC4 exhibits a consistent and smaller gap between the phases, whereas GC3.1 shows a slightly larger difference. The highest pattern correlation coefficient of 0.88 is achieved by GC3.1 at phase 2.

4.3 Mid-latitude storm tracks

Cyclone tracking uses the TRACK algorithm (Hodges, 1995) and tracks 850 hPa vorticity maxima/minima and produces track and genesis densities. In Fig.4.4, the statistics of cyclogenesis in various model configurations are shown, with GC3.1 and GC4 represented by dark red colours. In the northern hemisphere, GC4 exhibits some minor improvements over GC3.1 in the North Atlantic during the months of March-May. These improvements are attributed to better temperature distribution and possibly a more accurate simulation of 850 hPa winds around 60°S . Similarly in the Southern Hemisphere (60°S - 80°S) during the months of September to November, GC4 demonstrates an improved median number of cy-

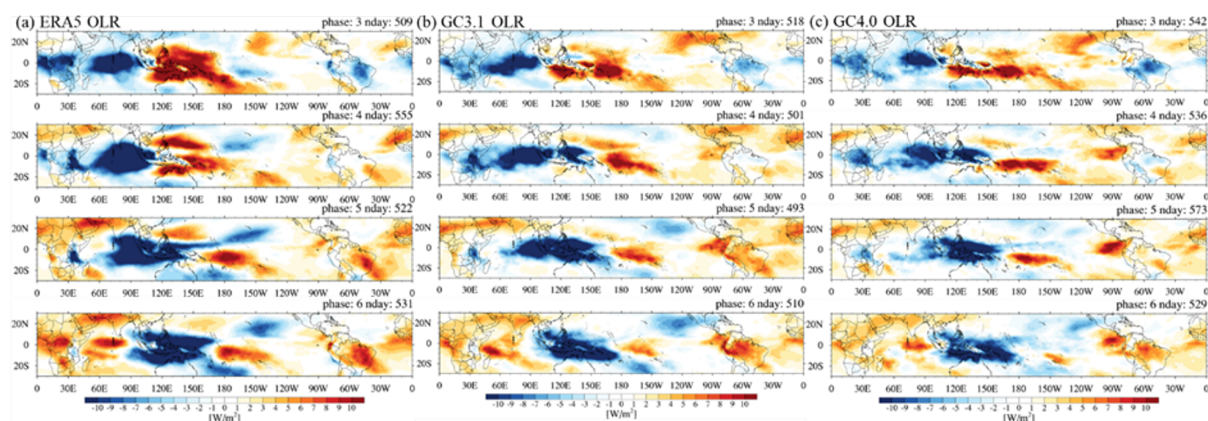


Figure 4.2: Composite of 20-70-day OLR based on a function of the MJO phase 3 to phase 6 during boreal winter (November-April). (a) ERA5 reanalysis, (b) GC3.1, and (c) GC4. The numbers of each panel represent the number of days used for composite for each phase.

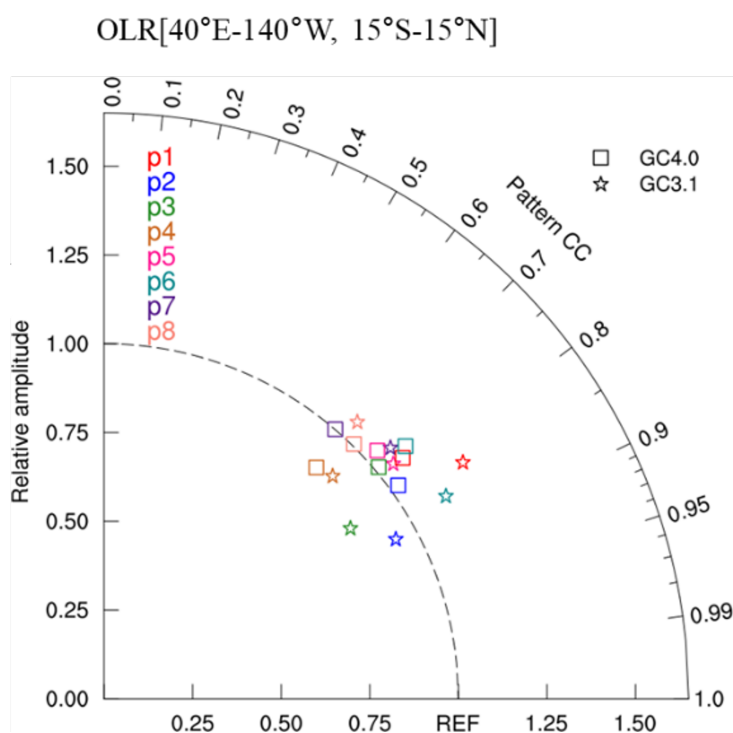


Figure 4.3: Taylor diagram of OLR [40°E-140°W, 15°S-15°N] pattern CC and OLR amplitude for MJO phase 1 to 8. The distance between each model and the reference point “REF” indicates the root-mean-square error.

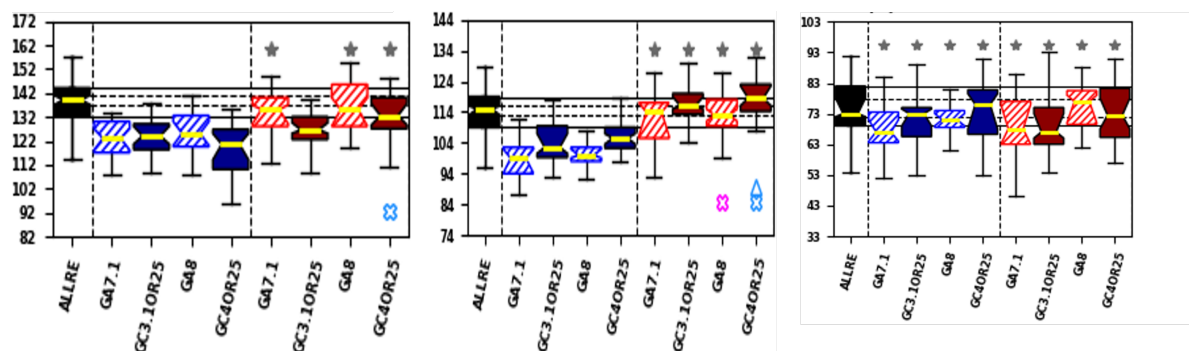


Figure 4.4: Seasonal cyclogenesis statistics from the ERA5 and various model configurations for (left) northern Atlantic sector in March-May season, (middle) Asia Pacific sector in DJF and (right) southern hemisphere (60°S-80°S) in September-November season. GC3.1 and GC4 and show in dark red colours. * sign indicates the median value is indistinguishable from reanalyses, X sign indicates the median value is significantly different at N216 relative to N96 and Δ indicates that the median value is significantly different between GC4 and the AMIP run at N216.

clogenesis compared to other models. Note that no major changes is seen in most other sectors (e.g. the Asia Pacific).

For the Winter 2018-19 Trials in the Northern Hemisphere, TRACK can be employed to match cyclones in Numerical Weather Prediction (NWP) simulations. This involves comparing cyclones tracked in own analyses with those tracked in the 6-day forecasts. If the tracks overlap within a certain distance threshold for a significant proportion of time, they are considered to be matching. By examining these matched cyclones, differences in cyclone intensity (maximum vorticity) and separation distance can be estimated. Mean-matched cyclone displacement and intensity errors can be calculated for each forecast lead time, and statistics for the entire simulation period can be derived. Importantly, the science changes in GC4 do not have a detrimental impact on cyclones in NWP. Overall, although small improvements are observed, there are no notable degradations in the performance of cyclone tracking between GC3.1 and GC4 models.

The report does not include the results of the mid-latitude blocking diagnostics due to the minimal differences observed between the GC3.1 and GC4 configurations.

Chapter 5

Spatial Variability

5.1 Near real time N1280 simulations

N1280 GA8GL9 (the atmosphere and land components of GC4) was initialised from operational analyses and run in near real time (NRT) between 2020 and GC4's operational implementation in 2022. This allowed direct comparisons between operational GA7.2.1GL8.1 and the NRT GA8GL9 forecasts. Figure 5.1 compares precipitation forecasts from GA7.2.1GL8.1 and GA8GL9 for region around the Philippines with the size of the region being approximately 3000km east-west by 2500km north-south. The instantaneous precipitation from GA7.2.1GL8.1 (figure 5.1a) tends to be very noisy within a larger active region: the instantaneous precipitation from GA8GL9 (figure 5.1b) is far more spatially coherent and produces spatially continuous structures. Over longer time-averaging periods GA7.2.1GL8.1 produces precipitation structures that are diffuse, still quite noisy and have relatively low peak precipitation rates (figure 5.1c). In contrast GA8GL9 produces more smaller and well-defined coherent structures with higher peak precipitation rates (figure 5.1d). Subjective comparisons with GPM suggest the spatial structure and precipitation rates in GA8GL9 are much more realistic than those in GA7.2.1GL8.1. The changes in precipitation structure are limited to regions or situations where the convection scheme is active such as the tropics and mid-latitudes in the summer season; little or no difference is seen in frontal precipitation. The improvement is primarily due to prognostic based entrainment rate that was added to the convection scheme in GA8. It should be noted that the improvement in instantaneous precipitation is less obvious at very low resolutions although some benefit is still seen when time-averages are considered.

The convection scheme is the primary source of cloud and cloud condensate in the tropics, and therefore the improvements to convective structures also affects these fields. Figure 5.2 shows the total column frozen water (TCFW), outgoing longwave radiation (OLR) and outgoing shortwave radiation (OSW) for GA7.2.1GL8.1 and GA8GL9. Again, the GA7.2.1GL8.1 forecasts are noisy and diffuse whereas the GA8GL9 forecasts are more spatially coherent and identify higher peak values.

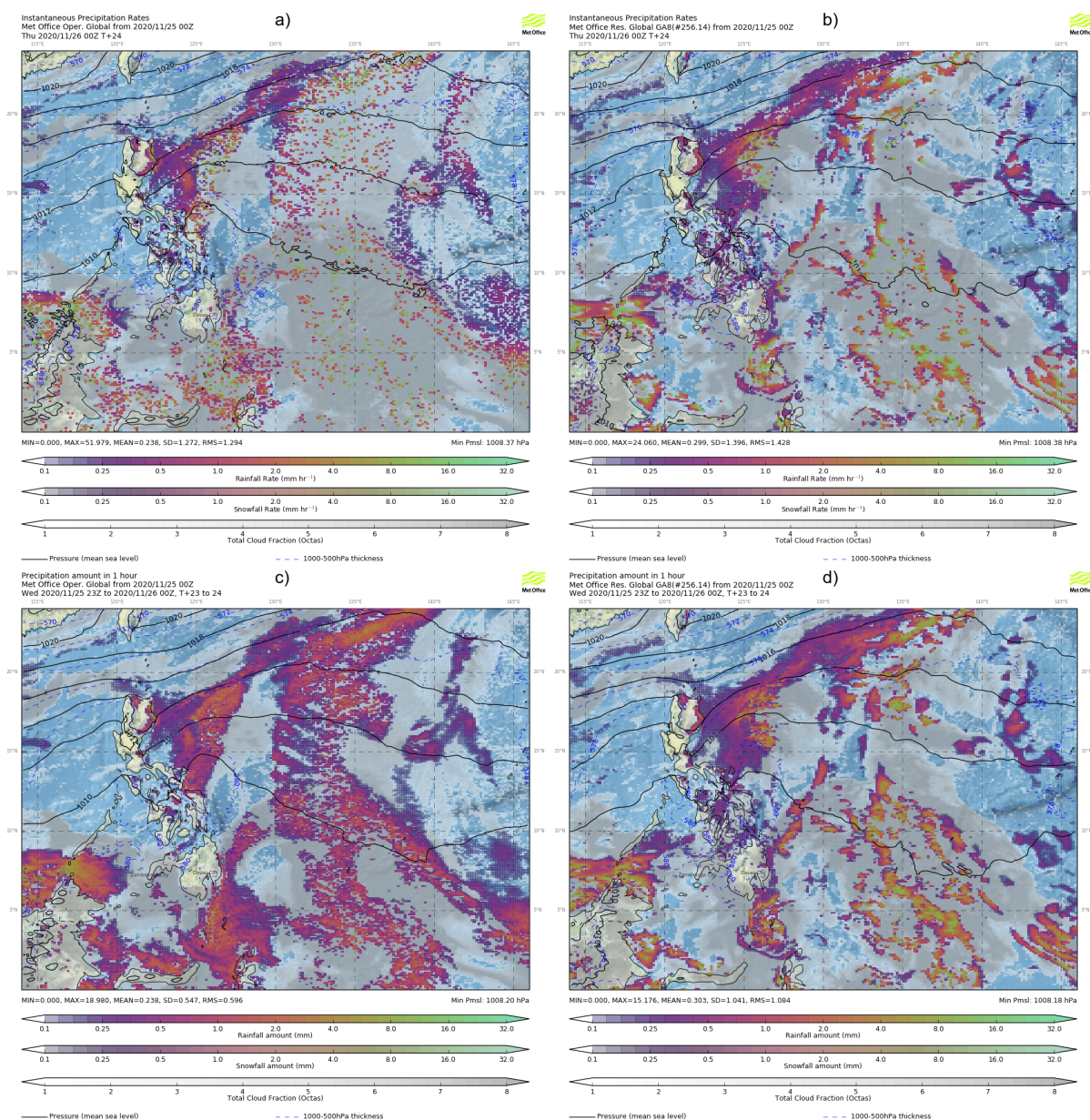


Figure 5.1: forecasts of instantaneous (top row) and 1-hourly averaged (bottom row) precipitation rates in colour and total cloud fraction in grey from N1280 simulations using GA7.2.1GL8.2 (a,c) and GA8GL9 (b,d). Instantaneous rate is value is valid at T+24, and 1-hourly is averaged between T+23 and T+24. Instantaneous precipitation rate is the precipitation rate over a single time step which in this case is 4 minutes.

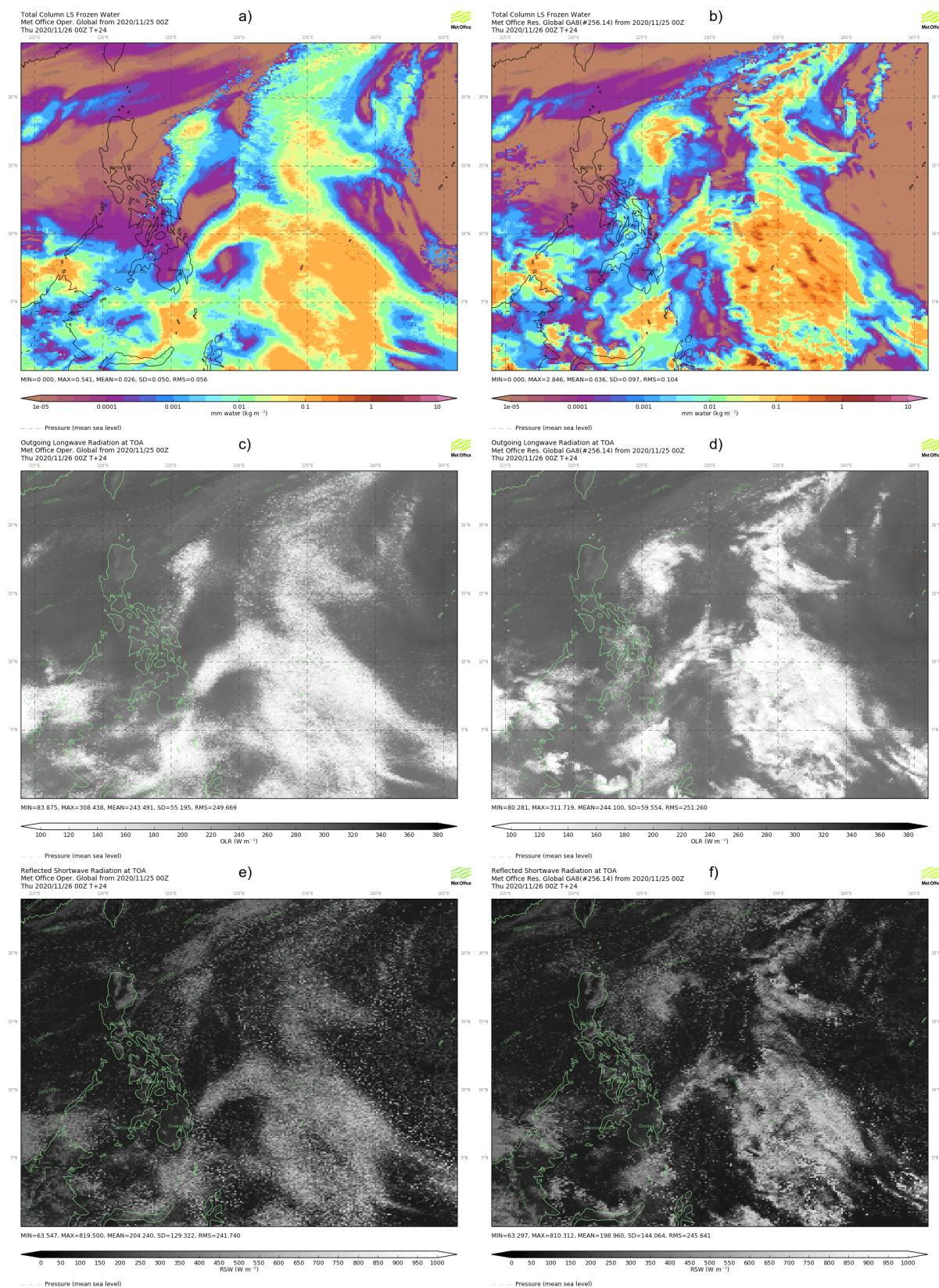


Figure 5.2: T+24 forecasts of total column frozen water (top row), outgoing longwave radiation (middle row) and outgoing shortwave radiation (bottom row) from N1280 simulations using GA7.2.1GL8.2 (a,c,e) and GA8GL9 (b,d,f) from the same forecast as the precipitation in figure 5.1.

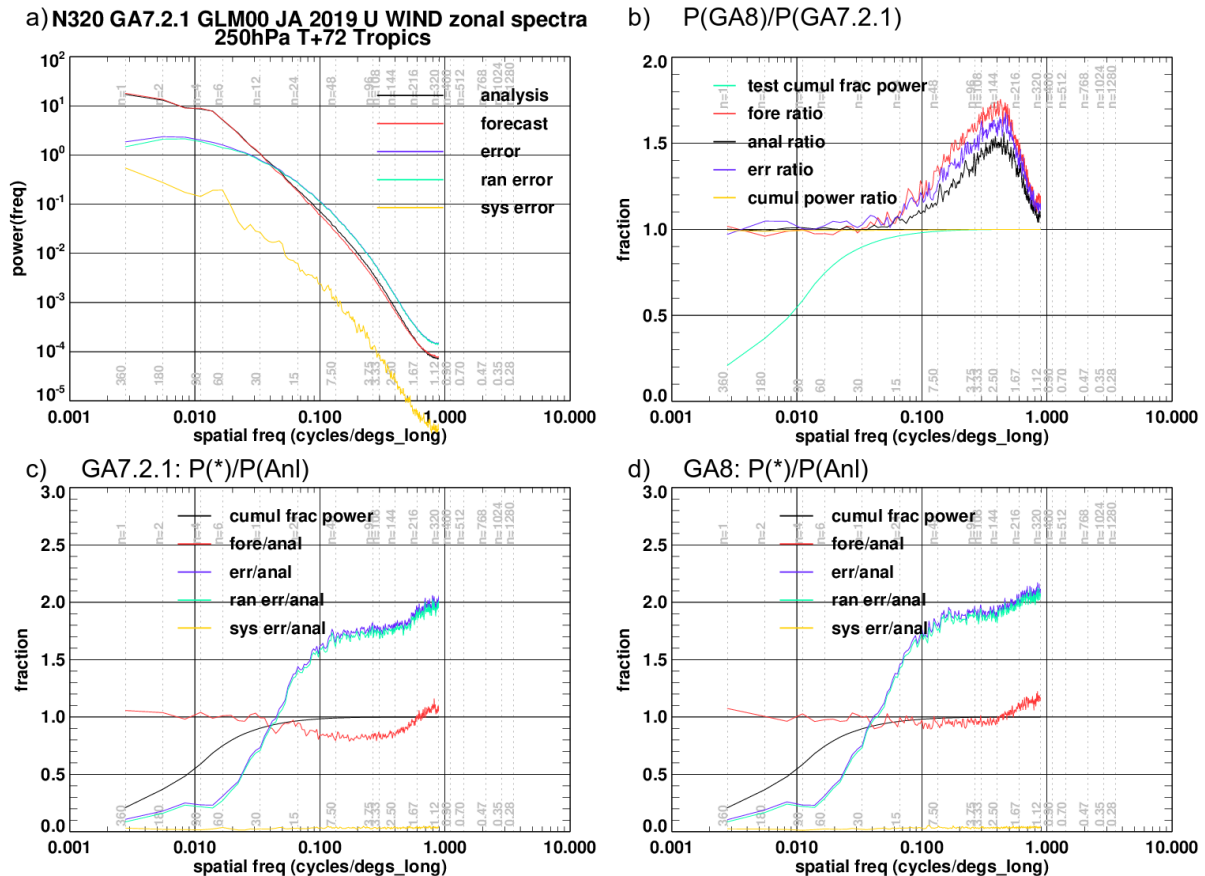


Figure 5.3: Power spectra for U-wind at 250hPa in the tropics (20S-20N) from summer N320 DA trials of GA7.2.1 and GA8 showing (a) analysis, T+72 forecast and error spectra from GA7.2.1, (b) the ratio of GA8 and GA7.2.1 analysis, T+72 forecast and error power spectra, (c) the T+72 forecast and error spectra for GA7.2.1 normalised by the power in its own analysis, and (d) the forecast and error spectra for GA8 normalised by the power in its own analysis.

5.2 Spatial power spectra

Figure 5.3 shows the zonal power spectra of U-wind at 250hPa in the tropics from N320 Data Assimilation (DA) trials of GA7.2.1GL8.1 and GA8GL9. GA7.2.1GL8.1 loses power relative to its own analysis between wavenumbers 20 and 200 (Fig 5.3c) whereas GA8GL9 loses very little (Fig 5.3d). GA8GL9 analyses also have greater power than GA7.2.1GL8.1 analyses because the reduced losses in the forecast feed through to the background forecast for the DA (Fig 5.3b). The power in the GA8GL9's error also increases, but this is to be expected because the power in its forecast and analyses has increased. Increases in power are also seen to varying degrees at other pressure levels and in other prognostic fields. This increased spatial variability in GA8GL9 is caused by the greater spatial coherence in the diabatic heating that in turn is primarily due to the addition of the prognostic entrainment.

Chapter 6

Regional aspects

6.1 South Asian Summer Monsoon (SASM)

The results depicted in Figure 3.2 demonstrate that despite a decrease in rainfall in the east-central Indian region, the GC4 runs exhibit an improvement in the mean rainfall during the June-July-August (JJA) season for the entire Indian region. The wet bias in central equatorial Indian Ocean rainfall is diminished in GC4, while there is an increase in rainfall over the south and southwest Indian Ocean (Figure 3.2). In GC4, this rise in rainfall can be attributed to a warm bias on the western side of the Indian Ocean, which reinforces the east-west sea surface temperature (SST) bias resembling a positive Indian Ocean Dipole (Saji et al., 1999). Consequently, this bias promotes compensatory increases in Indian rainfall. Additionally, the low-level monsoon winds are slightly weaker in the Indian region in GC4 compared to GC3.1, particularly over the Indian sector.

GC4 simulations demonstrate a reduced occurrence of monsoon low-pressure systems (LPS) and a decreased contribution of rainfall from such systems, as indicated in Figure 6.1. Previous studies have shown that these systems are better represented at higher resolutions, such as N216, compared to N96 (not shown here). This conclusion remains valid for the GC4 runs as well.

Both GC4 and GC3.1 reasonably reproduce the spatial distribution of tracks with the highest density over the northern Bay of Bengal and the northeastern part of the Indian peninsula, in accordance with observations. However, they underestimate the number of tracks extending towards the northwest. Consequently, there is a lack of rainfall contribution from these systems over the regions of Rajasthan, Gujarat, and Maharashtra, which significantly contributes to the overall mean bias in the June-September (JJAS) season.

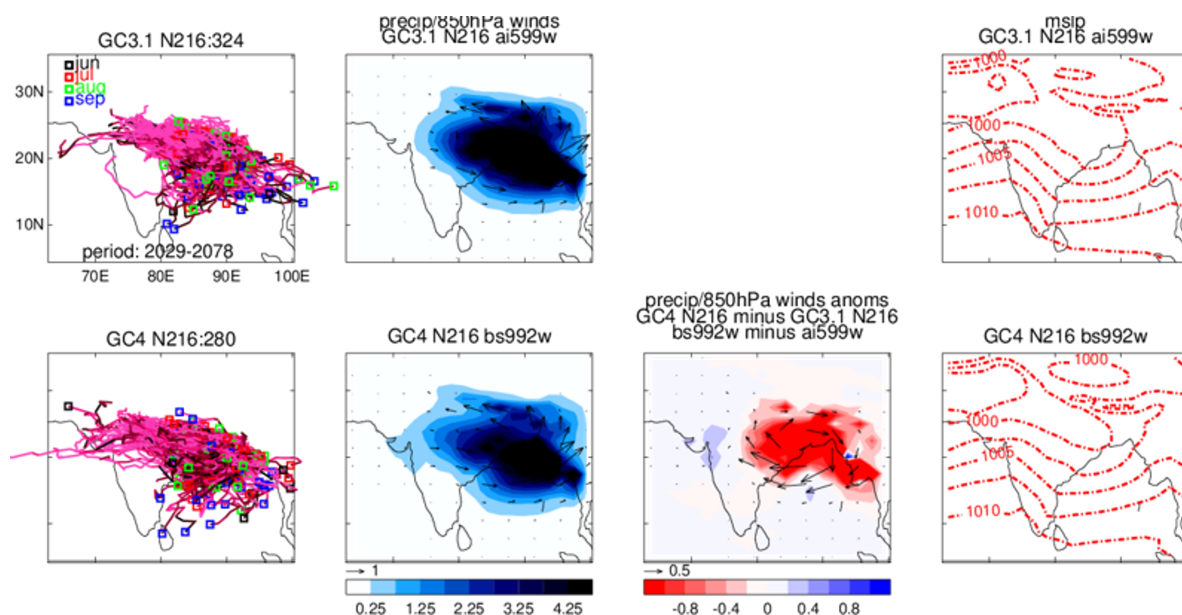


Figure 6.1: Tracks of LPS (left columns), the rainfall and 850 hPa winds associated with them (second column), the difference in rainfall contributions between GC4 and GC3.1 (third column). The mean sea level pressure values associated with the LPS are also shown in the rightmost columns.

6.2 East Asian monsoon

The East Asian monsoon, being a subsystem of the Asia-Pacific monsoon, is influenced by complex teleconnections from the Tropics and Arctic. [Zhou et al. \(2009\)](#) demonstrated lower predictability of the East Asian (EA) monsoon in climate models compared to other monsoon subsystems.

In this section, the performance of GC3.1 and GC4 was assessed using the East Asian summer monsoon, utilizing the EA monsoon metrics developed by [Boo et al., 2011](#)). Figure 6.2 (left) illustrates the normalized assessment criteria of GC3.1 (represented by the red circle) and GC4 (represented by the green circle) in comparison to observational data. The central reference line corresponds to the observations, while the grey bars indicate the range of observation data. The light grey bars represent the 75th to 25th percentile range. The metrics employed in this assessment include ERA5, ERA-interim, NCEP/NCAR 1, and NCEP/DOE 2 reanalysis, as well as CMAP, GPCP, and APHRODITE precipitation data. The proximity of the circle to the central line indicates the proximity to the observation.

The performance of surface temperature (SAT), mean sea level pressure (MSLP), 500hPa geopotential height (GPH), precipitation, and wind (850 and 200hPa GPH) were compared, revealing that the overall performance of GC4 is closer to the observations (Fig. 6.2). The 500hPa GPH falls within the observational range, and the 200hPa upper jets and wind shear, which are crucial for water vapor convergence in the EA monsoon, exhibited significant improvements close to the observations.

Although there was a slight improvement in mean sea level pressure, both GC3.1 and GC4 demon-

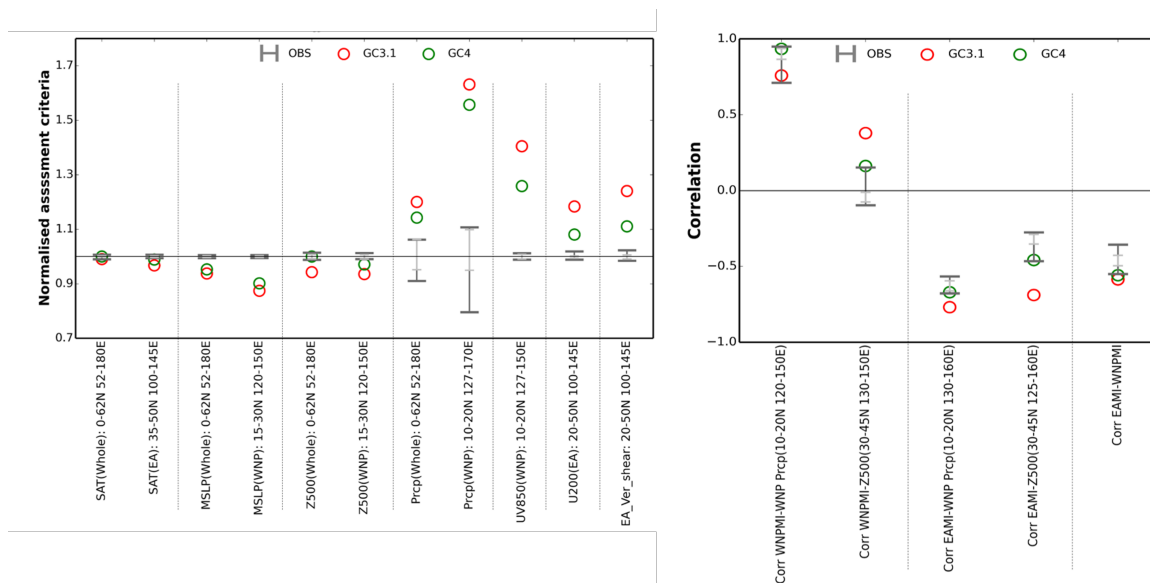


Figure 6.2: (Left) Normalized GC3.1 (red circle) and GC4 (green circle) model biases relative to observations (1.0 line) with the observation ranges (grey bars). Biases are normalized by the standard deviation of the observed interannual variation of each variable. (Right) Correlation of WNPMI and EAMI monsoon indices between precipitation and Z500, and the correlation between the two indices for GC3.1 (red circle) and GC4 (green circle). Grey bars are observation ranges.

strated good performance in surface temperature. The enhancements in vertical wind shear within the EA region and the 850hPa wind vector were instrumental in improving the WNPMI monsoon index.

While the improvement in precipitation was not substantial, the improved geopotential heights were deemed influential in the intensity of the Northwestern Pacific High, along with upper and low-level wind simulations. These factors are essential for the lower-level convergence of humidity in the East Asian summer monsoon. Figure 2 displays the correlations between two EA monsoon indices (WNPMI and EAMI) with precipitation and 500hPa geopotential height (Z500). Notably, all correlations exhibited improvements in GC4, closely resembling the observations.

6.3 Maritime Continent

Figure 6.3 shows the biases in cold surge composites for GC4 and the difference between GC4 and GC3.1. A detailed description of the impact of cold surges on extreme precipitation and how they interact with the other large-scale systems such as the MJO is given in [Lim et al. \(2017\)](#) and [Xavier et al. \(2020\)](#). Cold surges generally are associated with increased rainfall over the western side of the Malay peninsula, north of Borneo, the Java sea and the southeastern Philippines. In GC3.1 and most previous model versions, the precipitation associated with cold surges broadly follows the biases in mean DJF precipitation with drier biases over south-eastern parts of the Malay peninsula and a distinct lack of precipitation over northern Borneo and north of the island. There is a general improvement in the pre-

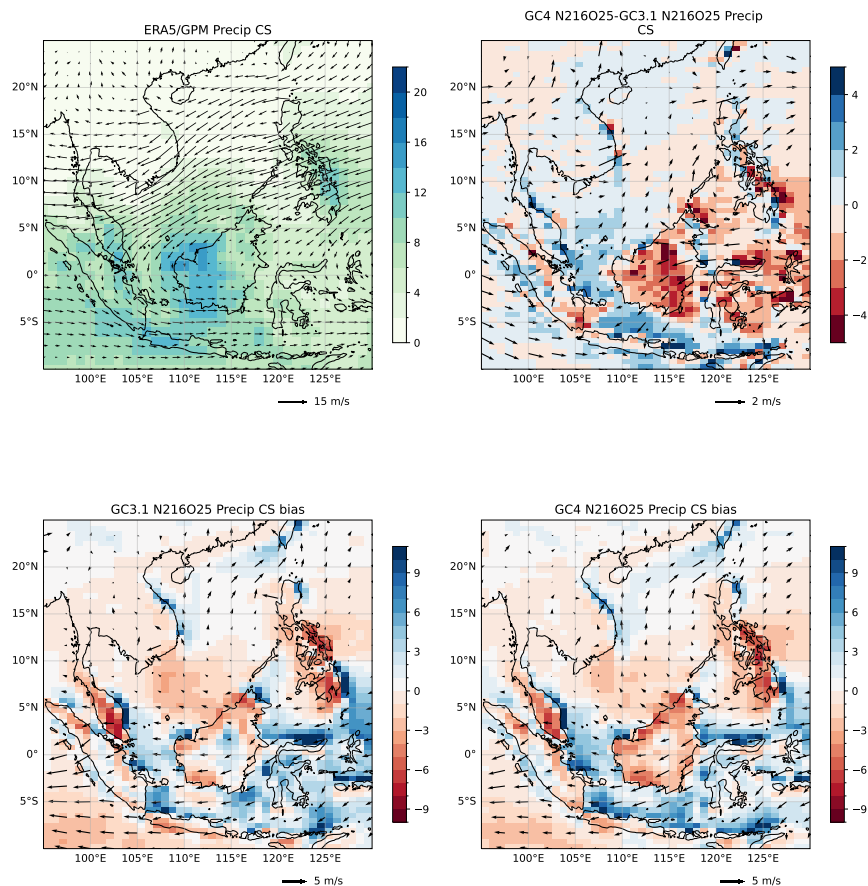


Figure 6.3: Precipitation (mm day^{-1}) and 850 hPa winds (m s^{-1}) associated with cold surges in the ERA5 reanalysis and GPM precipitation (top left), GC3.1 bias (bottom left), GC4 bias (bottom right) and GC4 minus GC3.1 (top right).

precipitation related to the cold surges over peninsular Malaysia with precipitation extending into the land rather than accumulating at the coastal points. Similar improvements are seen over the Java sea which is related to the cross-equatorial cold surges (Xavier et al., 2020). Most regions of Borneo however have drier bias in GC4, which may be related to the weaker northeasterly cold surge winds.

Chapter 7

Summary

7.1 Improvements in GC4

- The prognostic-based convective entrainment rate (ProgEnt) included in the GC4 model adds memory of subgrid activity into the parameterisation, which improves the representation of the diurnal cycle of convection over land (e.g. Africa), more intense instantaneous precipitation rates, more realistic spatial structures of precipitation, reduced overprediction of light rain. It also targets to reduce the convective intermittency and hence a better organization of convective features, and improved convection-dynamics relationship.
- Significant reduction of the Southern Ocean warm bias ($\sim 0.5^{\circ}\text{C}$). Other improvements to a number of inter-related Southern Ocean and atmospheric biases are primarily due to the cloud improvements as there are no ocean changes at GC4. Cloud changes include modifying parameters controlling the convective cloud and altering the convective parcel's condensate profile to give an improved model performance.
- Increased rainfall over India GC4 during the JJA season which is an improvement over GC3.1 and many previous model versions which had a long-standing dry bias over the region [Levine and Turner \(2012\)](#). This improvement in land precipitation is also linked to a reduction in the wet equatorial Indian Ocean rainfall. These improvements also suggest improved distribution of precipitation and improved representation of rainfall categories.
- It is noted that there is a slight reduction of the global hydrological cycle in GC4 compared to GC3.1, primarily attributed to the tuning of convective clouds.
- Other improvements include better representation of low-medium clouds over Northern Europe which are related to [Liu et al. \(2008\)](#) spectral dispersion. This also improves the subtropical marine low clouds.
- Improvements to the primary marine organic aerosol parametrisation ([Mulcahy et al., 2020](#)) in-

crease model skill at simulating marine aerosol concentrations (improved AOD and OC concentration).

- Even though it is not the focus of this report to compare the role of atmosphere-ocean coupling, it is worth noting that significant positive impacts of coupling on NWP scores were observed during the development stages of the NWP case studies and trials.
- In the JJAS 2019 N320 DA trials, tropical 250 hPa zonal wind spectra show that GC4 does not lose power relative to its own analysis at spatial scales less than ~ 1000 km and hence retains more variability at these scales. Similar behaviour is seen in summer mid-latitudes but the effects appear to be smaller at lower atmospheric levels.
- Comparisons of atmospheric components (GA8 with GA7.2.1) revealed that GA8 produce more spatially coherent and realistic precipitation structures, particularly in regions where the convection scheme was active, attributed mainly to the addition of a prognostic-based entrainment rate in the convection scheme.

7.2 Neutral/slightly improved in GC4, yet a significant issue

- As noted above, even though there are some aspects of the global hydrological cycle and energetics that show slight improvements, it remains an important model problem as GC4 still produces too much global precipitation with respect to the observations.
- The cold bias in the 2m temperature annual mean is improved in GC4, it now shows warm bias over coastal regions of East Asia. It might be related to the strong short-wave radiation bias in this region.
- Even though there is a general improvement in precipitation over the South Asian monsoon regions, GC4 shows a slight wet bias over East Asia compared to a dry bias in GC3.
- Despite the improvements in shortwave and longwave cloud forcing in GC4 over equatorial Africa, the Atlantic and equatorial East Pacific, they remain too weak, particularly over the Maritime Continent, Indian Ocean and East Asian sector.
- Gross primary productivity (GPP) biases v observations year-round follow biases in precipitation: too high GPP in NW South America, too low NE South America and India which is quite similar to that in GC3.1. GPP on land is too high in cropland areas of Southeast USA, Europe and Southeast China. Global land GPP is still too low (85-90 v ~ 120 GT C yr⁻¹ in observations) despite the use of `can_rad_mod = 6`, `diff_frac = 0.4` in the GL9 Drag Package. This can potentially impact the moisture fluxes in the physical model and the carbon cycle in the Earth System Model (ESM) ([Collins et al., 2011](#)).

- Southern Ocean biases have improved slightly or remain neutral at GC4 due to cloud improvements as there are no ocean changes in GC4 compared to GC3.1. However, substantial biases in the Antarctic Circumpolar Current transport and ice-shelf temperatures west of the Antarctic Peninsula remain, which may be indicative of wider Southern Ocean problems. Longer runs are needed to confirm this but these biases are expected to be improved with numerics tuning at GC5. However, recent work suggests the remaining problem might be linked to the representation of topography so there may be potential for more substantial future improvements in the GC5 development cycle.
- Aerosol concentrations in some relatively polluted environments remain biased low in GC4, suggesting a need to identify and include missing aerosol sources in future model versions.

7.3 Degradations in GC4

- Sea surface temperatures in all seasons in GC4 show enhanced warming in the equatorial east Pacific (which was already 0.5-1.5 degrees warmer in GC3.1 compared to observations). This warming is accompanied by weaker trade winds which may contribute via Bjerknes feedback ([Bjerknes, 1969](#)) to further warming of the eastern Pacific. GC4 also shows a related weakening of the Walker Circulation ([Walker, 1924](#)) and hence may introduce some regional and remote teleconnection biases, particularly at the seasonal time scales.
- The atmospheric model changes in GC4 and the related feedback with the ocean introduces excessive surface warming of about 1°C in the North Atlantic.
- GC4 has a general reduction in the numbers and strength of monsoon low-pressure systems and tropical cyclones in both the climate and NWP simulations.
- Annual mean precipitation over Africa shows a general drying in GC4 compared to GC3.1 and the low-level wind convergence appears to be further south, particularly in the context of West Africa.
- Excessive low-top thick cloud bias in mid-latitudes gets worse in GC4. A similar increase is noted in the excessive high-top thick clouds in both the tropics and mid-latitudes.
- Aerosol concentrations are generally too high in remote marine regions in part due to increased dust transport.

7.4 Towards GC5

Seasonal hindcasts using GC4 were not included in this assessment cycle due to time and resource constraints. The Bureau of Meteorology (BoM) will be assessing the impact of GC4 seasonal hindcast test cases. Currently, BoM is working on initializing the GC4 model with GC3 ocean initial conditions

from the Met Office seasonal group, as the ocean model remains unchanged in GC4. The results from these hindcasts are expected in the development cycle of GC5.

There are stratosphere issues that may require further understanding and investigation, such as the Southern Hemisphere polar vortex cold biases and their impact on ozone chemistry, and the equatorial Semi-Annual Oscillation (SAO) and its descent from the mesosphere down to the stratopause (1hPa), which may be linked to the absence of resolved wave forcing propagating into this region. Additionally, there are issues with aerosol size distributions due to missing sources and processes. These issues are crucial to address in order to improve the accuracy and reliability of the Earth System Models used for climate prediction and analysis.

Bibliography

- Bjerknes, J., 1969: Atmospheric teleconnections from the equatorial pacific. *Monthly weather review*, **97 (3)**, 163–172.
- Boo, K.-O., G. Martin, A. Sellar, C. Senior, and Y.-H. Byun, 2011: Evaluating the east asian monsoon simulation in climate models. *Journal of Geophysical Research: Atmospheres*, **116 (D1)**.
- Collins, W., and Coauthors, 2011: Development and evaluation of an earth-system model–hadgem2. *Geoscientific Model Development*, **4 (4)**, 1051–1075.
- Hardiman, S. C., and Coauthors, 2019: The impact of prescribed ozone in climate projections run with hadgem3-gc3. 1. *Journal of Advances in Modeling Earth Systems*, **11 (11)**, 3443–3453.
- Hodges, K., 1995: Feature tracking on the unit sphere. *Monthly Weather Review*, **123 (12)**, 3458–3465.
- Holben, B. N., and Coauthors, 1998: Aeronet - a federated instrument network and data archive for aerosol characterization. *Remote Sens. Environ.*, **(66)**, 1–16, doi:10.1016/50034-4257(98)00031-5.
- Hu, W. W., and Coauthors, 2012: The characteristics and origins of carbonaceous aerosol at a rural site of PRD in summer of 2006. *Atmos. Chem. Phys.*, **(12)**, 1811–1822, doi:10.5194/acp-12-1811-2012.
- Jones, A. C., A. Hill, J. Hemmings, P. Lemaître, A. Querel, C. L. Ryder, and S. Woodward, 2022: Below-cloud scavenging of aerosol by rain: a review of numerical modelling approaches and sensitivity simulations with mineral dust in the Met Office’s Unified Model. *Atmos. Chem. Phys.*, **(22)**, 11 381–11 407, doi:https://doi.org/10.5194/acp-22-11381-2022.
- Jones, A. C., and Coauthors, 2021: Exploring the sensitivity of atmospheric nitrate concentrations to nitric acid uptake rate using the Met Office’s Unified Model. *Atmos. Chem. Phys.*, **(21)**, 15 901–15 927, doi:https://doi.org/10.5194/acp-21-15901-2021.
- Josey, S. A., E. C. Kent, and P. K. Taylor, 2002: Wind stress forcing of the ocean in the soc climatology: Comparisons with the ncep–ncar, ecmwf, uwm/coads, and hellerman and rosenstein datasets. *Journal of physical oceanography*, **32 (7)**, 1993–2019.
- Leaitch, W. R., and Coauthors, 2013: Dimethyl sulfide control of the clean summertime Arctic aerosol and cloud. *Elem. Sci. Anth.*, doi:10.12952/journal.elementa.000017.

- Levine, R. C., and A. G. Turner, 2012: Dependence of indian monsoon rainfall on moisture fluxes across the arabian sea and the impact of coupled model sea surface temperature biases. *Climate Dynamics*, **38 (11)**, 2167–2190.
- Lim, S. Y., C. Marzin, P. Xavier, C.-P. Chang, and B. Timbal, 2017: Impacts of boreal winter monsoon cold surges and the interaction with MJO on Southeast Asia rainfall. *Journal of Climate*, **30 (11)**, 4267–4281.
- Liu, Y., P. H. Daum, H. Guo, and Y. Peng, 2008: Dispersion bias, dispersion effect, and the aerosol–cloud conundrum. *Environmental Research Letters*, **3 (4)**, 045 021.
- Merchant, C. J., and Coauthors, 2019: Satellite-based time-series of sea-surface temperature since 1981 for climate applications. *Scientific data*, **6 (1)**, 1–18.
- Mulcahy, J. P., and Coauthors, 2020: Description and evaluation of aerosol in UKESM1 and HadGEM3-GC3.1 CMIP6 historical simulations. *Geosci. Mod. Dev.*, **13**, 6383–6423, doi:10.5194/gmd-13-6383-2020.
- Reddington, C. L., and Coauthors, 2017: The Global Aerosol Synthesis and Science Project (GASSP): Measurements and modeling to reduce uncertainty. *Bull. Amer. Met. Soc.*, **98**, 1857–1877, doi:10.1175/BAMS-D-15-00317.1.
- Russel, L. M., L. N. Hawkins, A. A. Frossard, P. K. Quinn, and T. S. Bates, 2009: Carbohydrate-like composition of submicron atmospheric particles and their production from ocean bubble bursting. *Proc. Nat. Acad. Sci.*, **107**, 6652–6657, doi:10.1073/pnas.0908905107.
- Saji, N., B. N. Goswami, P. Vinayachandran, and T. Yamagata, 1999: A dipole mode in the tropical Indian Ocean. *Nature*, **401 (6751)**, 360–363.
- Sellar, A. A., and Coauthors, 2020: Implementation of uk earth system models for cmip6. *Journal of Advances in Modeling Earth Systems*, **12 (4)**, e2019MS001 946.
- Taylor, K. E., 2001: Summarizing multiple aspects of model performance in a single diagram. *Journal of geophysical research: atmospheres*, **106 (D7)**, 7183–7192.
- Walker, G. T., 1924: Correlation in seasonal variations of weather. IV : A further study of world weather. **24**, 275–332.
- Wang, J., H. Kim, D. Kim, S. A. Henderson, C. Stan, and E. D. Maloney, 2020: Mjo teleconnections over the pna region in climate models. part ii: Impacts of the mjo and basic state. *Journal of Climate*, **33 (12)**, 5081–5101.
- Willett, M., and M. Whittall, 2017: A simple prognostic based convective entrainment rate for the unified model: Description and tests. *Met Office internal) Forecasting Research Technical Reports*, **617**.

Willett, M. R., M. E. Brooks, J. M. Edwards, A. P. Lock, A. J. Malcolm, E. H. Muller, , and W. J. Tennant, 2024: The Met Office Unified Model GA7.2GL8.1 and GA7.2.1GL8.1.1 configurations: Developments from GA7GL7. Tech. Rep. 654, Met Office, U.K. doi:<https://doi.org/10.62998/yrta1217>.

Williams, K., and Coauthors, 2018: The Met Office global coupled model 3.0 and 3.1 (GC3. 0 and GC3. 1) configurations. *Journal of Advances in Modeling Earth Systems*, **10** (2), 357–380.

Xavier, P., and Coauthors, 2020: Seasonal Dependence of Cold Surges and their Interaction with the Madden–Julian Oscillation over Southeast Asia. *Journal of Climate*, **33** (6), 2467–2482.

Zhou, T., and Coauthors, 2009: The clivar c20c project: which components of the asian–australian monsoon circulation variations are forced and reproducible? *Climate dynamics*, **33**, 1051–1068.

Met Office
FitzRoy Road
Exeter
Devon
EX1 3PB
United Kingdom