UC San Diego
UC San Diego Electronic Theses and Dissertations

Title
Exploring the Usage of Data Assimilation on the Community Earth System Model

Permalink
https://escholarship.org/uc/item/775611rt

Author
Eliashiv, Jonathan

Publication Date
2019

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA SAN DIEGO

Exploring the Usage of Data Assimilation on the Community Earth System Model

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Oceanography

by

Jonathan D. Eliashiv

Committee in charge:

Arthur J. Miller, Chair
Ian Eisenman
Myrl Hendershott
Petr Krysl
Joel Norris
Aneesh Subramanian

2019

Copyright
Jonathan D. Eliashiv, 2019
All rights reserved.

The dissertation of Jonathan D. Eliashiv is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California San Diego

2019

EPIGRAPH

The fluttering of a butterfly's wing in Rio de Janeiro, amplified by atmospheric currents, could cause a tornado in Texas two weeks later.

—Edward Lorenz

TABLE OF CONTENTS

Signature Page iii

Epigraph iv

Table of Contents v

List of Figures vii

List of Tables ix

Acknowledgements x

Vita xii

Abstract of the Dissertation xiii

Chapter 1 Introduction 1
- 1.1 General Circulation Models 1
  - 1.1.1 Community Earth System Model 2
- 1.2 Representation of Climate Modes in GCMS 3
  - 1.2.1 Madden Julian Oscillation 3
- 1.3 Data Assimilation 4
  - 1.3.1 EAKF 6
- 1.4 Summary of Results 6

Chapter 2 Tropical climate variability in the Community Earth System Model - Data Assimilation Research Testbed 11
- 2.1 Introduction 12
- 2.2 Reanalysis products and CESM-DART 15
- 2.3 Methodology 17
  - 2.3.1 Surface Fluxes 17
  - 2.3.2 MJO metrics and equatorial waves 19
- 2.4 Results 20
  - 2.4.1 Precipitation 20
  - 2.4.2 Surface heat fluxes 21
  - 2.4.3 Equatorial waves 22
  - 2.4.4 MJO 23
- 2.5 Discussion and Conclusion 24
- 2.6 Acknowledgements 26

Chapter 3 A Reliability Budget analysis of CESM-DART 40
3.1 Introduction 41
3.2 Methodology 43
3.2.1 Reliability Budget 43
3.2.2 CESM-DART 47
3.2.3 MJO index 48
3.3 Results 48
3.3.1 Reliability Budget 48
3.3.2 Expanded Reliability Budget 49
3.3.3 Single dimension exploration 50
3.3.4 Ocean temperature exploration 52
3.3.5 Intraseasonal exploration 53
3.4 Summary and Conclusion 55
3.5 Appendix 58
3.6 Acknowledgements 61

Chapter 4 Conclusions and Comments 75
4.1 Advantages of DART in CESM 75
4.2 Isolated sources of error in CESM 76
4.3 Future Outlook 78

Bibliography 79

LIST OF FIGURES

Figure 1.1: A diagram representing the Madden Julian Oscillation through its cycle. 8
Figure 1.2: Composite of boreal winter MJO events in ERA-I. MJO events and phases were determined using Ventrice et al (2013). 9
Figure 1.3: Schematic of assimilated measurements into CESM using an EAKF provided by DART. (Karspeck et al, 2018) 10

Figure 2.1: Annual mean precipitation differenes between reanalyses and GPCP. . . 31
Figure 2.2: Monthly-mean standard deviation of precipitation, calculated relative to the annual mean climatology (1970-79), shown as a five-product mean and differences from that mean for each reanalysis product. 32
Figure 2.3: Annual mean (1970-79) latent (left) and sensible (right) surface heat flux differences for the five reanalyses relative to OAFlux observations (1970-79). Differences are scaled by a normalization factor of 60 W/m^2 for latent heat and 20 W/m^2 for sensible heat. 33
Figure 2.4: Differences between July and January reanalysis biases relative to OAFlux, calculated as in Figure 2.3. 34
Figure 2.5: Frequency-wavenumber spectra of equatorially symmetric lower tropospheric (850hPa) zonal winds, plotted as the five-product mean and as differences from that mean spectrum for each reanalysis (1970-79). . . 35
Figure 2.6: Frequency-wavenumber spectra of equatorially symmetric lower tropospheric (850hPa) zonal winds and upper tropospheric (200hPa) zonal winds, plotted as the difference between the spectrum for CESM-DART and the spectrum for Free-CESM. 36
Figure 2.7: Combined EOF of lower tropospheric zonal winds, upper tropospheric zonal winds, and upper tropospheric velocity potential for each reanalysis product, computed for the MJO period band in the tropics, 1970-79, following the procedure of Ventrice et al. (2013). 37
Figure 2.8: As in Figure 7, but for the corresponding MJO VPM Index PC calculated for each reanalysis product. 38
Figure 2.9: Differences in phases through a composite MJO cycle in winter months shown as differences for each reanalysis product (1970-79) relative to the corresponding composite computed from the ERA-I (1991-2010). . 39

Figure 3.1: Observation Density of of atmospheric (temperature, zonal wind, meridional wind) and oceanic (temperature, salinity) measurements that are assimilated into CESM using DART. 63
Figure 3.2: Reliability budget terms for atmospheric zonal velocity, vertically averaged. From top to bottom, Departure (total error variance), Bias of the ensemble mean, Ensemble Variance, Observational Uncertainty, Residual. Observation Density is also plotted for reference. 64

Figure 3.3: As in Figure 3.2, but including Expanded reliability budget terms for Model Error Variance, Observational Error Variance, and linearly estimated Bias. 65
Figure 3.4: Reliability budget for atmospheric zonal velocity, plotted as a function of latitude, longitude, time, and height, and otherwise averaged in the other dimensions. Shaded regions represent the observation density at each grid point. 66
Figure 3.5: Reliability budget for atmospheric zonal velocity plotted as (top) a seasonal cycle and (bottom) anomalies around the seasonal cycle, after subtracting out the annual mean, and otherwise averaged over latitude, longitude, and height. 67
Figure 3.6: Sensitivity of the terms in the reliability budget to latitude, longitude, height, time and observation density. 68
Figure 3.7: As in figure 3.2 but for ocean temperature, vertically averaged 69
Figure 3.8: Reliability budget terms for each phase of the MJO and for non-MJO conditions, averaged for the Western Pacific and Indian basins for lower tropospheric (850±50 hPa) and upper tropospheric (200±50 hPa) zonal winds. 70
Figure 3.9: Reliability budget of zonal wind observations in the western equatorial Pacific region. Top panel shows the budget without an MJO events, while subsequent panels show the budget for individual phases of MJO. Histogram shows the observation density for associated calculations. . 71
Figure 3.A1: Error in the estimates of e_o, e_m, A, and B as a function of (a) e_o, (b) e_m, (c) linear bias, A and (d) background bias, B. In each subplot, values for the other parameters are set to $\sqrt{\overline{x^2}}/10$. These plots are made using 1000 randomly selected values of Truth. 72
Figure 3.A2: Error in estimating e_o (red) and e_m (black) as a function of both e_o and e_m. Values for the other parameters are set to $\sqrt{\overline{x^2}}/10$. 73
Figure 3.A3: Error in estimating e_o, e_m, A, and B with increasing data points (observations). For each plotted point, simulations are created with random values of the above parameters 100 times, and estimates from each simulation are averaged, as described in the text. 74

LIST OF TABLES

Table 2.1: Pattern correlation (latitude weighted) between reanalysis products and observations (GPCP and OAFlux) of annual mean precipitation, latent heat flux, and sensible heat flux. 28
Table 2.2: Spatially averaged mean and rms differences between reanalysis products and observations (OAFlux) of annual mean latent heat flux and sensible heat flux for Pacific Ocean and Indian Ocean. 29
Table 2.3: Spatially averaged mean and rms differences between reanalysis products and observations (ERA-I) across all eight MJO composite phases of U850, V850, and VP200 for summer and winter composite MJO. 30

Table 3.1: Squared correlation between terms in the reliability budget (from Figure 3.4) and $log(n+1)$ in the variable specified by each row. 62
Table 3.2: Reliability budget terms for ocean temperatures averaged globally for the indicated range of ocean temperature. 62
Table 3.3: Reliability budget terms for zonal winds and ocean temperatures during MJO and non-MJO conditions averaged over western equatorial Pacific Ocean and Indian Ocean. 62

ACKNOWLEDGEMENTS

I would like to thank Art Miller for providing me with the support to stay at Scripps beyond my initial Masters studies. His enthusiasm and support drove me to tackle research problems far more complicated that I initially gave them credit for. Even as the complexities of life provided their own turbulent mixing, his care provided a beacon and aid for which I am extremely grateful.

I would also like to thank Asst. Professor Subramanian, for whose advice and ceaseless support proved to make this work possible to begin with. Aneesh's patience, ceaseless efforts and careful critiques allowed me to have all the support I could want for. Additionally, his advice has further helped me in all areas of my life and provided me with necessary support. He extends himself to everyone he knows and always helps however he can. For this and other reasons, Aneesh is by far the most compassionate academic I have ever had the privilege of meeting.

I would like to thank Osinachi Ajoku. His friendship and perspective have helped me through the years since first being accepted into Scripps Institution of Oceanography.

The subsequent chapters have only been possible because of the efforts of Dr. Alicia Karspeck, who provided the model runs that are used tirelessly in this analysis. Her talks and advice allowed me to maintain perspective through the process of this PhD candidacy.

This research was supported by a grant from NOAA Climate Variability and Prediction Program (NA14OAR4310276) and the NSF Earth System Modeling Program (OCE1419306). High-performance computing support from Yellowstone (ark:/85065/d7wd3xhc) was provided by NCARs Computational and Information Systems Laboratory (CISL), which is sponsored by the National Science Foundation. The data were provided by the Data Support Section of CISL at NCAR.

Chapter 2 has been submitted for publication in its entirety:

- **Eliashiv, J**, Subramanian A, Miller AJ (2019). Tropical climate variability in the

Community Earth System Model - Data Assimilation Research Testbed. *Climate Dynamics*, submitted, *sub judice*.

The dissertation author was the primary researcher and author of this paper.

Chapter 3 has been submitted for publication in its entirety:

- **Eliashiv, J**, Subramanian A, Miller AJ (2019). A Reliability Budget analysis of CESM-DART, *Journal of Advanced in Modeling Earth Systems*, submitted, *sub judice*.

The dissertation author was the primary researcher and author of this paper.

VITA

2013	B. S. in Physics w/ specialization in Computational Physics, and B. S. in Earth Sciences University of California San Diego
2014-2018	Graduate Teaching Assistant University of California San Diego
2019	Ph. D. in Oceanography Scripps Institution of Oceanography University of California San Diego

PUBLICATIONS

Eliashiv, J., A. C. Subramanian and A. J. Miller, 2019a: Tropical climate variability in the Community Earth System Model - Data Assimilation Research Testbed, *Climate Dynamics, sub judice*.

Eliashiv, J., A. C. Subramanian and A. J. Miller, 2019b: A Reliability Budget analysis of CESM-DART, *Journal of Advances in Modeling Earth Systems, sub judice*.

ABSTRACT OF THE DISSERTATION

Exploring the Usage of Data Assimilation on the Community Earth System Model

by

Jonathan D. Eliashiv

Doctor of Philosophy in Oceanography

University of California San Diego, 2019

Arthur J. Miller, Chair

A new prototype coupled ocean-atmosphere Ensemble Kalman Filter reanalysis product, the Community Earth System Model using the Data Assimilation Research Testbed (CESM-DART), is studied by comparing its tropical climate variability to other reanalysis products, available observations, and a free-running version of the model. The results reveal that CESM-DART produces fields that are comparable in overall performance with those of four other uncoupled and coupled reanalyses. The clearest signature of improvements in CESM-DART is in the analysis of the Madden-Julian Oscillation (MJO) and other tropical atmospheric waves. MJO energy is enhanced over the free-running CESM as well as compared to the other products, suggesting the importance of the surface flux coupling at the

ocean-atmosphere interface in organizing convective activity. In addition, high-frequency Kelvin waves in CESM-DART are reduced in amplitude compared to the free-running CESM run and the other products, again supportive of the oceanic coupling playing a role in this improvement. CESM-DART also exhibits a relatively low bias in the mean tropical precipitation field and mean sensible heat flux field. Conclusive evidence of the importance of coupling on data assimilation performance will require additional detailed direct comparisons with identically formulated, uncoupled data assimilation runs.

A reliability budget is used to diagnose potential sources of error (departure from observations) in the CESM-DART product. In areas with sufficient observations, the mean bias in zonal wind was generally very low compared to the spread due to ensemble variance, which did not exhibit patterns associated with Northern Hemisphere jet streams but did have regional enhancement over the Maritime Continent. However, the Residual term was often the largest contributor to the budget, which is problematic, suggesting improper observational error statistics and inadequately represented ensemble variance statistics. The departure and residual exhibit significant seasonal variability, with a strong peak in boreal winter months, indicating the model's deficiencies during the energetic Northern Hemisphere winter. Ocean temperature contained large error in areas with eddy production indicating inadequate ensemble variance due to poor model resolution. Periods when the the Madden-Julian Oscillation (MJO) was active exhibited lower error, especially in the western equatorial Pacific during MJO phases with reduced convection. In contrast, during MJO phases with enhanced convection in that region, the ensemble variance is increased yet the error is comparable to non-MJO conditions, suggesting a controlling effect of the convection parameterization. Further studies evaluating the impact of the coupled assimilation procedure on the reliability budget will be illuminating.

Chapter 1

Introduction

The physical equations that define the earth system are few yet complex. Each is chaotic and complex in their own right. These equations are further forced by additional systems such as biogeochemical forcing, and external forcing agents such as anthropogenic agents. However, it is still just a small handful of equations, and therefore they are able to be isolated and studied. The study of those equations and the subsequent systems they generate have been tackled in force, and through a large series of discoveries, the progression of the earth system can be modeled to a coarse degree of accuracy.

1.1 General Circulation Models

The Navier-Stokes equations, alluded to above, have been numerically modeled and refined for individual fluid systems over the years, including the coupled Earth system. Coupled Earth system models are especially important for their usefulness such as a full earth representation (Edwards, 2001), disaster projection (Shen et al, 2006), and climate change (Stocker, 2014). Full earth representation is an important avenue of global climate models because they allow areas with low data representation to be shown and modeled in turn. This is important for areas such as developing countries that do not have the resources to deploy

and share data of their own domains, as well as regions like the Southern Ocean or deep oceans that are too difficult to gather much data from, yet provide powerful mechanisms that affect the rest of the coupled climate system (Gould and Turton, 2006). Disaster projection and climate change have been explored extensively using global climate models due to the unfortunate lack of data from the future, making the models the only peer-reviewed source of material for such research.

Numerical modeling of fluid systems have always proven challenging (Sengupta, 2004). A key issue when modeling Earth's fluid systems is choosing which resolution to model. The higher the resolution, the better the model will be able to show features like eddies (Cuxart et al, 2000) and singular Inter Tropical Convergence Zone, ITCZ, bands (Lin, 2007). However, as the resolution of the model increases, so does the computational demand, therefore the resolution of the model must be limited to a reasonable amount in order to limit the need to build additional power plants and computer clusters. There have been many decades of improvement from simplified frameworks such as single column models (Randall et al, 1996) to modern day global coupled climate models, often referred to as General Circulation Models (GCMs). These modern global climate models split the earth system into separate ocean, atmosphere, ice, land, and other component models, which are then coupled together. Each component is forced by each other as well as external forcing components such as solar radiation and oceanic vents.

1.1.1 Community Earth System Model

The National Center for Atmospheric Research (NCAR) released multiple versions of the Community Earth Systems Model, CESM. CESM is a GCM widely adopted and used in various studies, including the Intergovernmental Panel on Climate Change Assessment Reports (Stocker, 2014). Created in collaboration between NCAR and a plethora of experts across the globe, CESM splits the planet into ocean, atmosphere, atmospheric chemistry,

land, sea ice, biogeochemistry, land ice, and external forcing model components (Hurrell et al, 2013). As an example of what each of these components process, the atmosphere model solves horizontal momentum, thermodynamic energy, mass continuity, and water vapor flux in its own finite volume grid. CESM is an incredibly useful tool in climate research, as it allows users to independently pose and answer scientific questions using the model. The model is designed to allow for alterations to model physics parameters and additions of different forcings and dampings.

1.2 Representation of Climate Modes in GCMS

CESM, as well as other GCMs, are far from perfect, and should not be taken as gospel. Though they have many areas that perform well, there are still mechanisms that are yet to be ideally modeled. These are normally processes around coupled climate modes such as representation of the ITCZ, and Madden Julian Oscillations, MJO. The ITCZ is well understood but is often represented in global climate models with two distinct bands (Li and Xie, 2014), rather than a singular band. MJO events are the dominant source of intraseasonal tropical climate variability, however their simulation is often overdamped while running and poorly predicted (Arnold et al, 2015).

1.2.1 Madden Julian Oscillation

MJO events are responsible for over half of the tropical climate variability in the intraseasonal time scale (Hendon and Salby, 1994). Figure 1.1 shows a diagram of the structure of a possible mechanism for MJO. In the diagram, it can be seen that MJO has a convective cloud structure with enhanced evaporation to the west, and increased shortwave flux to the east. This in turn creates an SST gradient with positive SST to the east. The SST gradient builds a surface pressure system with a mean westerly wind component, moving the

convective MJO structure to the increased SST region. This region, which has undergone increased evaporation has more water content that is then convected to the upper troposphere. In the upper troposphere the air diverges. To the east of the divergence the air then converges and sinks, causing a reduction in clouds and a warming of the surface. This moves the shortwave and SST gradient eastward which reinforces the surface westerly wind and moves the entire structure eastward.

The MJO event loses power as it hits landmasses, including the maritime continent. These sudden perturbations in the SST dipole structure can result in baroclinic Matusuno-Gill responses, which send much of the MJO power westward, dampening the MJO response (Lau et al, 2012). The events typically range from 30 to 90 days in time period. These events are identified using combined empirical orthogonal functions, EOFs, and separated into phases associated with the events longitudinal location, as seen in Figure 1.2. The figure shows velocity potential at 200 hPa because it is associated with precipitation.

1.3 Data Assimilation

Data assimilation is the class of methodologies which use observations and observation uncertainty to obtain state estimates and correct models. Using data assimilation assumes that the model used to make system forecast state vectors are accurate to a good degree of certainty, though not necessarily perfect. These techniques are used to correct the model forecast state vector to an analysis state vector. The analysis state vector is a superposition of the model state vector with the observation state vector in model space. In order to assimilate data, error covariance of both the forecast state vector and observation must be estimated. After estimation, the analysis state can be described below:

$$\psi^a = \psi^f + \frac{\mathbf{C}^f}{\mathbf{C}^f + \mathbf{H}\varepsilon\mathbf{H}^T}\left(\psi^f - \mathbf{Hp}\right)$$

where ψ^f and ψ^a is the state vector for the forecast and analysis respectively, $\mathbf{C}^f$ is the error covariance matrix of the forecast, $\mathbf{H}$ is the transform from observation to model space, ε is observation error matrix, and $\mathbf{p}$ is the observation vector.

Observation error is estimated with representation error, which is normally constant for each type of measurement. The forecast state vector is estimated based on the constraints of the model physics, which can be derived from the model resolution and model physics.

Variational methods use the given covariance matrix to adjust the initial state vector until the error between observations and the forecast are minimized. In contrast, sequential assimilation schemes, such as Kalman filters, calculate the probability densities of model state vectors given observations using the given error covariances. These probability densities are then used to weight the contributions of the forecast state vector and the observation projected into model space for the analysis state vector.

When using data assimilation methods, GCMs are able to perform with high precision over the the regions with large observation densities, and relatively good precision over all areas. As will be later discussed, multiple different GCMs, each with their own data assimilation methods, have all been able to replicate MJO responses at correct times with relatively accurate magnitude. This shows that the use of data assimilation significantly improves recreating an accurate representation of the climate system, even without necessarily high observation density.

1.3.1 EAKF

This dissertation explores the use of Ensemble Adjustment Kalman Filters, EAKF, on CESM. This uses a large ensemble system of CESM. By using ensembles, the error covariance matrix of the forecast state vector can be set to the ensemble variance. Following the definitions used above, the analysis step is derived as follows:

$$\psi^a = \mathbf{C}^a \left(\psi^f / \mathbf{C}^f + \mathbf{H}^T \varepsilon^{-1} \mathbf{p} \right)$$
$$\mathbf{C}^a = \left((\mathbf{C}^f)^{-1} + \mathbf{H}^T \varepsilon^{-1} \mathbf{H} \right)^{-1}$$

where $\mathbf{C}^a$ and $\mathbf{C}^f$ are the ensemble variance of the analysis and forecast state vector respectively. The Data Assimilation Research Testbed (Anderson et al, 2009) uses the EAKF after a 6-hour forecast before the component models are coupled together. The data used to assimilate include atmosphere wind and temperature data, and ocean temperature and salinity data. This type of data assimilation is a weakly-coupled system because the variables being directly assimilated are not directly coupling the different component models together. Examples of strongly coupled systems include ECMWF ERA20C which uses surface temperature and pressure data to relax the forecast towards. Figure 1.3 shows a diagram of the procedure of DART being used on the CESM system.

1.4 Summary of Results

Chapter 2 explores the tropical climate variability of CESM with DART as compared against other reanalysis systems that cover the same time period. Precipitation is found to have relatively low bias in CESM-DART as compared to other models, despite not being directly assimilated into the model. The MJO responses across the different models are found

to accurately represent MJO events, and more so, are found to have the most improvement in behavior in CESM-DART. The response of spurious equatorial Kelvin waves are additionally found to be reduced in CESM-DART compared to other models. This suggests an important effect from the surface fluxes during assimilation in appropriately dampening these dynamical waves.

Chapter 3 uses a reliability budget to diagnose error growth and propagation within CESM before the EAKF is applied to the system. In this system mean square error, departure, is separated into square mean error, bias, ensemble variance, observational uncertainty, and a residual term to close the budget. Although the bias term was low, the residual term was often the largest contributor to the budget, indicating improper observational error statistics and inadequately represented Ensemble Variance statistics. During periods when the Madden-Julian Oscillation (MJO) was active, the error was lower, especially in the western equatorial Pacific during MJO phases with reduced convection, suggesting a controlling effect of the convection parameterization.

As a note to the readers, the subsequent chapters were written to stand alone and interpreted independently. There will therefore undoubtedly be repeated information, and possible differences in notation.

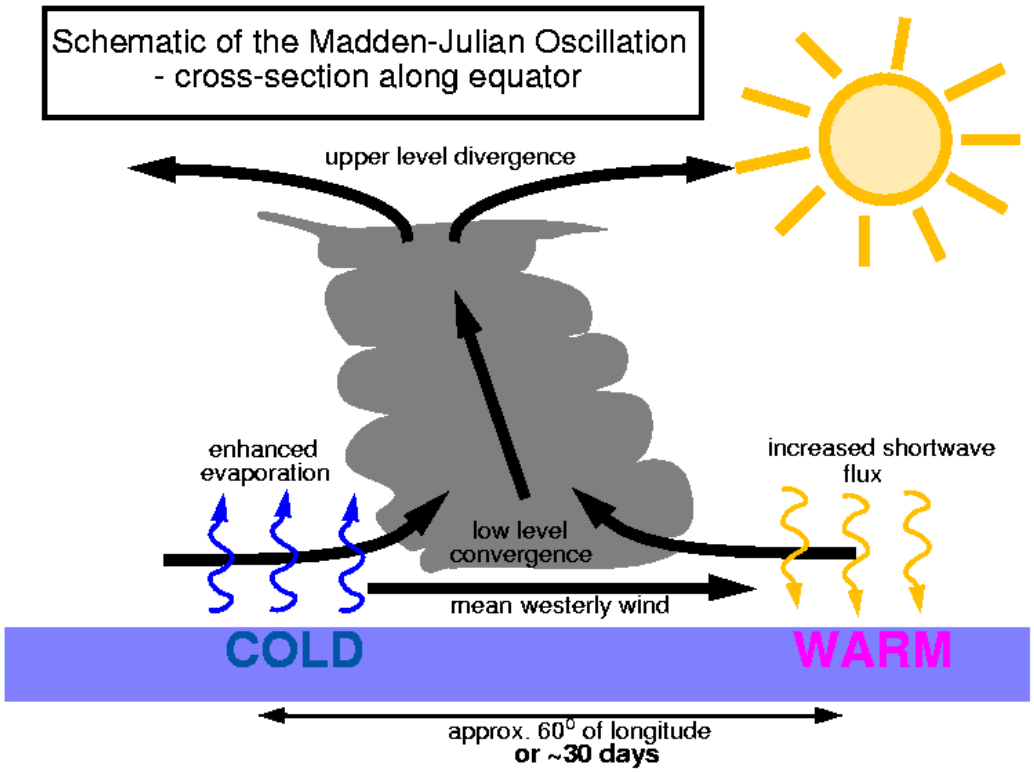

Figure 1.1: A diagram representing the Madden Julian Oscillation through its cycle (Source: http://www.met.reading.ac.uk/ pete/mjo.html).

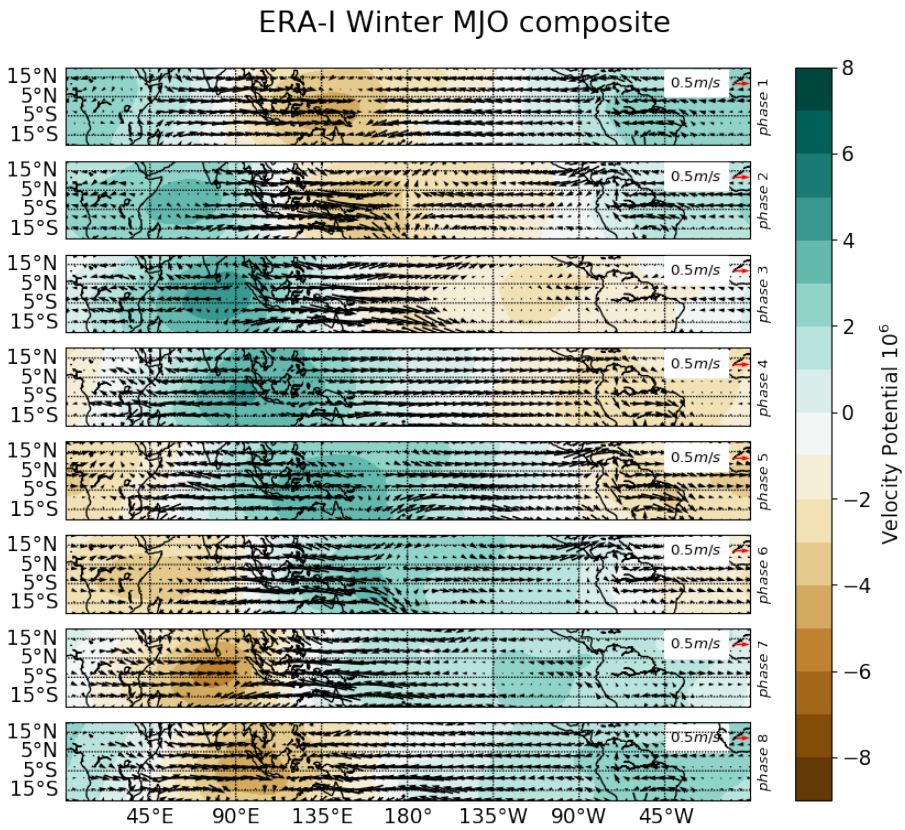

Figure 1.2: Composite of boreal winter MJO events in ERA-I from 1990-2010. Arrows show horizontal winds at 850 hPa while shaded contours of velocity potential at 200 hPa for the different MJO phases. MJO events and phases were determined using a VPM index derived in Ventrice et al (2013).

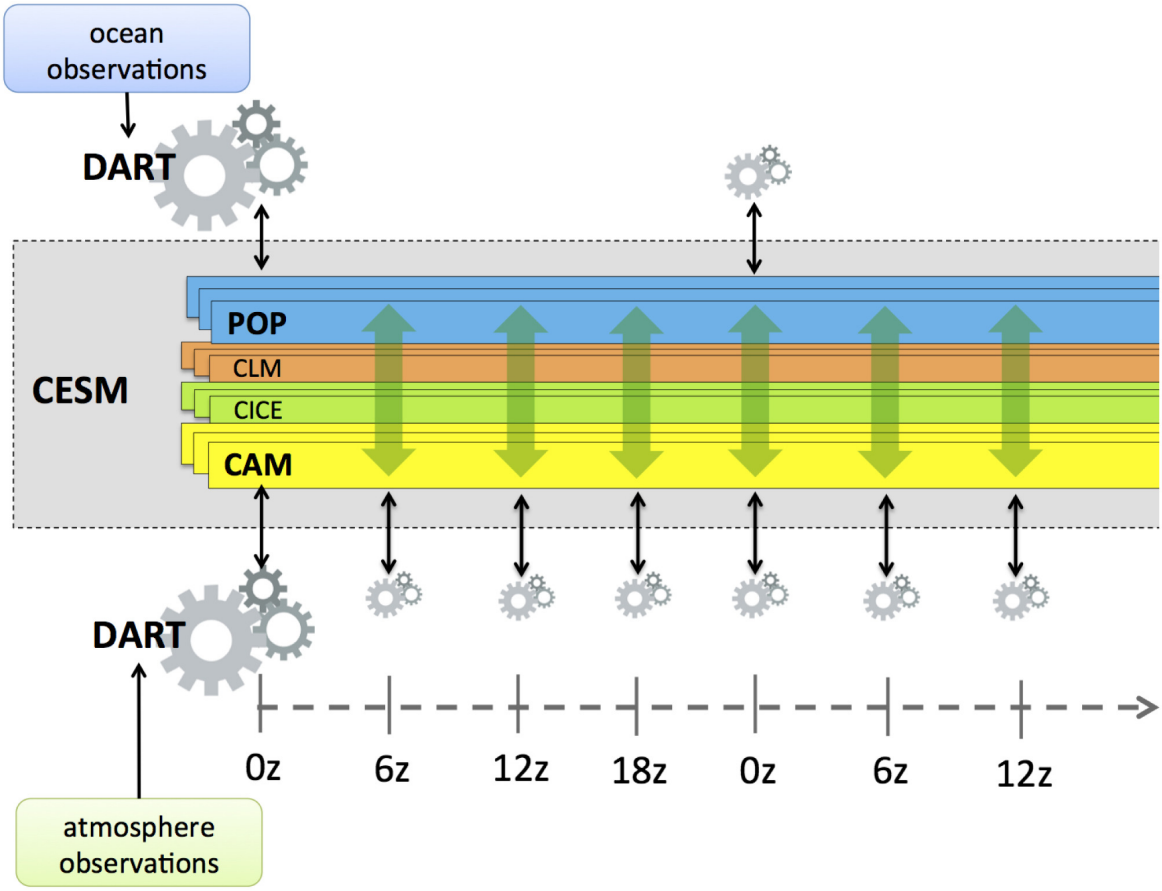

Figure 1.3: Schematic of assimilated measurements into CESM using an EAKF provided by DART. (Karspeck et al, 2018)

Chapter 2

Tropical climate variability in the Community Earth System Model - Data Assimilation Research Testbed

A new prototype coupled ocean-atmosphere Ensemble Kalman Filter reanalysis product, the Community Earth System Model using the Data Assimilation Research Testbed (CESM-DART), is studied by comparing its tropical climate variability to other reanalysis products, available observations, and a free-running version of the model. The results reveal that CESM-DART produces fields that are comparable in overall performance with those of four other uncoupled and coupled reanalyses. The clearest signature of improvements in CESM-DART is in the analysis of the Madden-Julian Oscillation (MJO) and other tropical atmospheric waves. MJO energy is enhanced over the free-running CESM as well as compared to the other products, suggesting the importance of the surface flux coupling at the ocean-atmosphere interface in organizing convective activity. In addition, high-frequency Kelvin waves in CESM-DART are reduced in amplitude compared to the free-running CESM run and the other products, again supportive of the oceanic coupling playing a role

in this improvement. CESM-DART also exhibits a relatively low bias in the mean tropical precipitation field and mean sensible heat flux field. Conclusive evidence of the importance of coupling on data assimilation performance will require additional detailed direct comparisons with identically formulated, uncoupled data assimilation runs.

2.1 Introduction

Atmospheric and oceanic data assimilation schemes are now routinely combined with climate observations to produce reanalysis products that are used to quantify and diagnose climate variability processes (e.g. Chelliah and Bell, 2004; Carton and Giese, 2008; Balmaseda et al, 2013). Until recently, these assimilation frameworks have treated the atmosphere and ocean independently instead of as a coupled system. Over the past decade, research and operational forecasting centers around the world have recognized the potential benefits of coupled ocean-atmosphere assimilation and rapid progress has been made in efforts to develop coupled data assimilation systems (e.g. Penny and Hamill, 2017). Recently, a weakly-coupled data assimilation system (Karspeck et al, 2018) was developed at the National Center for Atmospheric Research (NCAR) using the Community Earth System Model (CESM) in the Data Assimilation Research Testbed (DART), which is an Ensemble Adjustment Kalman Filter-based (EAKF) scheme (Anderson et al, 2009). Invoking such a procedure in an Earth System Model could provide new insights into the role of ocean-atmosphere coupling on the climate processes that have occurred over the historical record.

Surface heat fluxes, momentum flux, and fresh-water flux are the primary means for communication between the ocean and atmosphere systems. The estimates of these fluxes from coupled climate models, however, are frequently cited as containing biases and random errors that result in inaccurate representations of climate processes and internal modes of

variability (e.g. Arnold et al, 2015). The process of data assimilation can adjust the model state variables closer to measured data, but assimilation is not assured of improving the surface surfaces, which are typically diagnostically assessed. A coupled data assimilation framework (Penny and Hamill, 2017) may improve the estimation of these surface fluxes since they transit information between the ocean and atmosphere during the assimilation process itself. This may be especially important for properly representing climate modes of variability that rely on or are affected by the surface fluxes, such as the Madden-Julian Oscillation (MJO) (e.g. Raeder et al, 2012).

The CESM-DART uses a weakly-coupled data assimilation scheme, where state vectors of individual models run the ensemble Kalman filter that adjusts their individual state vectors, as opposed to a strongly-coupled system where the state-vectors of the entire combined system are adjusted with each run. The advantage is that by allowing the models to be assimilated separately, the individual systems may be more likely to retain their independent features while still improving surface flux estimates. This is in contrast to a strongly-coupled system where mechanisms in one system may potentially be improperly adjusted by covariances of data from the other system (Frolov et al, 2016).

Previous studies have shown substantial differences among various reanalysis products, such as the work of Hodges et al (2011) who highlighted the differences in mid-latitude storms tracks, and Kim et al (2014), who examined differences in tropical precipitation, intraseasonal variability, and MJO representation. Differences in tropical circulation, rainfall, MJO characteristics are expected because these fields are largely defined by the deep convection in the tropics, which is not directly constrained by observations and is highly controlled by widely varying model physics schemes (e.g. Subramanian et al, 2011). Hence, the representation of the tropical basic state and variability will depend on the different assimilation systems, observation datasets, and model physics parameterizations used in the reanalyses systems.

Here we consider the CESM in a data assimilation framework using DART over a prototype 10-year time period. Our aim is to evaluate the impact of the coupled data assimilation on the estimation of fields relevant to tropical climate and its variability in both the atmosphere and ocean. We first approach the problem by comparing key surface flux fields among five distinct coupled and uncoupled analysis products, CESM-DART, ECMWF 20th-Century Reanalysis (ERA20C), ECMWF Coupled 20th-Century Reanalysis (CERA20C), NCEP/NCAR Reanalysis 1 (R1), and Japanese 55-Year Reanalysis 55 (JRA55) to determine if the CESM-DART fluxes exhibit significantly different estimates over the traditional analysis products. We also evaluate if the CESM-DART analysis results in improved estimates of the surface fluxes due to the propagation of the coupled error covariance for the data assimilation procedure. This can be assessed by comparing the fields of CESM-DART and the other coupled and uncoupled assimilation products with the in-situ observations-based Objectively Analyzed Air-Sea Flux (OAFlux) fields. We then assess the ability of the CESM-DART product to represent a key tropical climate mode, the MJO, by comparing its differences among the five analysis products. We also attempt to assess if some of these differences are due to the surface flux field differences or if they are more strongly controlled by model physics differences.

In the next section, we explain the reanalysis products and the details of CESM-DART. We then present our methodology in comparing the products to available observations and among themselves. We then describe our results concerning the precipitation field, the surface heat fluxes, equatorial wave activity and MJO. We lastly provide a discussion and some concluding remarks.

2.2 Reanalysis products and CESM-DART

Data assimilation is a method in which observations are combined with model dynamics to derive a better estimate of the state at a given time. The first global, multi-decadal reanalysis using a consistent data assimilation method and model was produced jointly at the National Center for Environmental Prediction (NCEP) and NCAR (Kalnay et al, 1996). Many reanalyses products have been produced since to evaluate and compare with model outputs as well as to help improve our understanding of the weather and climate system. These include NCEP- Department of Energy reanalysis (Kanamitsu et al, 2002), the ERA20C (ECMWF) reanalysis (Uppala et al, 2005), the CERA20C (Laloyaux et al, 2017), and the JRA55 (Onogi et al, 2007). The large observational network including satellite records starting in the late 1970s as well as improvements in our data assimilation techniques have vastly improved the reanalyses products. This has led to the production of the most recent reanalyses products: the NCEP Climate Forecast System Reanalysis (CFSR) (Saha et al, 2010) and the European Reanalysis Interim Project (ERA-I) (Dee et al, 2011). Comparisons among the reanalysis products requires understanding the impacts of the various types of data assimilation techniques employed and the effects of the oceanic boundary conditions, whether specified or via coupled modeling, on the atmosphere.

In this study, we focus on CESM1.1.1, which is a coupled model which combines the community atmospheric model version 4 (Gent et al, 2011), with the Community Land Model (CLM4), the Community Ice Code (CICE4), CESM River Transport Model (RTM), and the LANL Parallel Ocean Program (POP2) for a coupled earth climate model (Hurrell et al, 2013). The DART system is an open-source community developed software used to implement the ensemble adjustment Kalman filter (EAKF) data assimilation technique into different models (Anderson et al, 2009). The application of DART to the CESM (Karspeck et al, 2018) provides a novel depiction of climate data for the period that it was run (1970-

1979). CESM-DART is run with a 30-member ensemble, with data assimilated every 6 hours for atmospheric fields and 24 hours for oceanic components. The components are coupled together every 24 hours. ACARS data is used to adjust CAM state vectors and BUFR data is used for POP state vectors. Localization and inflation methods are explained by Karspeck et al (2018).

Ensemble Kalman filters allow a model to prescribe the error using model variance which can then be compared to observational uncertainty which will allow the data assimilation to not only correct the mean model state. This not only improves the accuracy of the model run, but adjusts the precision of it as well by prescribing the ensemble variance to be related to the observational uncertainty. This technique becomes increasingly beneficial as the number of ensembles is increased (Houtekamer and Mitchell, 1998). In contrast to the 3dvar and 4dvar DA schemes, the Ensemble Kalman filter uses the covariance of the different ensemble members to correct its analysis against observation data. The EAKF scheme is well described by Anderson (2003) as a sequence of scalar computations that involves assimilating single observations to adjust the ensemble, and subsequently updating each variable in the model state vector with linear regression. This approach is implemented in parallel by the DART framework and is coupled to several modeling systems for assimilation applications (Anderson et al, 2009).

We have selected several of the available reanalysis products to compare with CESM-DART. These include the uncoupled assimilation products, R1, ERA20C, JRA55, as well as the coupled product, CERA20C. The CESM-DART reanalysis is limited to the time period January 3rd, 1970 0:00 UTC and December 31st, 1979 18:00 UTC, and all the reanalysis products have atmospheric fields analyzed at 6-hour intervals. Any unavailable data was bilinearly interpolated over to make a continuous dataset. For all data sets, zonal and meridional wind variables were gathered with daily sampling at 200 hPa and 850 hPa, while surface heat fluxes and precipitation were gathered at monthly intervals. R1 has a

horizontal resolution of 2.5, ERA-20C has a horizontal resolution of 125 km, JRA55 has 55km resolution, CERA20C has 125km resolution, and CESM-DART has a horizontal resolution of 1. The R1 uses 3Dvar, while the ERA20C, JRA5, and CERA20C all use incremental 4D-Var, in contrast to the EAKF of CESM- DART.

Given these different reanalysis products, we now objectively compare and contrast the surface fluxes and tropical climate behavior between the CESM-DART reanalysis and four other reanalyses for the 10-year time period of 1970-1979, based on the availability of data from CESM-DART. Note that because these data assimilated products span the 1970s, the segment that we are comparing does not include remotely sensed satellite data. Therefore since that the volume of data that was assimilated is much lower than for products that include later periods, the results of our analysis may be different for later decades.

2.3 Methodology

The combination of looking at surface heat fluxes, tropical waves, and MJO events, will give a comprehensive analysis of the tropical climate variability for the physical system produced by the three reanalyses datasets. We summarize the details of our analysis procedures in the following subsections.

2.3.1 Surface Fluxes

The three surface fluxes considered in this study are precipitation flux, latent heat flux, and sensible heat flux. For comparison to model estimates of mean surface fluxes we use observation datasets from OAFlux and the Global Precipitation Climatology Project (GPCP). These in-situ observation-based products are used to compare against the three reanalysis products for a direct evaluation of these fields. OAFlux uses SST, humidity, wind fields at 2m, and air temperature at 2m to make best guess analysis of latent heat flux and sensible

heat flux. All fields and products are detrended to remove any climate change signature.

Precipitation fields for the reanalyses are first compared against the GPCP product to compare annual means over the oceans. The temporal mean for each reanalysis product covers the common 1970-1979 period, while the temporal mean of the GPCP product covers the entire available dataset, 1979-2015. Although these are non-overlapping time intervals, the basic state structure of tropical precipitation should be relatively stable. GPCP products are subtracted from the temporal mean of each reanalysis product to show the difference maps. Mismatched grids use bilinear interpolation (Jones, 1999) to regrid the lower resolution GPCP grid to the higher resolution grids.

Since the GPCP product does not span the same period as the reanalysis products, the variability of the products can only be directly compared against each other. The variability is calculated based on the monthly mean precipitation time series. The resulting spatial products are compared directly using the same bilinear interpolation scheme discussed above.

For latent heat flux and sensible heat flux, all reanalyses products are compared against the OAFlux product (Yu et al, 2008), which is treated as truth in this study. The OAFlux product is a combination of data-assimilated reanalysis products for the time period between 1958-2018, which includes the common period of 1970-79 used in this study. Although the direct observations were rather limited in the common period, this product was well validated using the large numbers of observations from recent decades. This lends weight to our regarding the OAFlux fields as observations in this study (Gelaro et al, 2017).

For these flux analyses, the reanalysis products are converted into monthly time series and have their temporal variability calculated from those datasets. When compared, the higher resolution datasets are bilinearly interpolated to the coarser scheme for a direct comparison. Since all the ensemble members of CESM-DART are available, the ensemble mean of their temporal means is used to compare with GPCP and OAFlux. For all products, the latitude weighted spatial correlation and spatial bias against GPCP and OAFlux are

calculated.

2.3.2 MJO metrics and equatorial waves

A variety of higher-frequency tropical atmospheric equatorial waves are associated with important energy transport, such as MJO, Kelvin waves, and inertia-gravity waves. Since all the products are available at subdaily sampling rates, they are suitable for comparing their intraseasonal behavior.

We use MJO events as one measure of intraseasonal coupled climate variability. To isolate MJO events, the velocity potential MJO (VPM) index is computed using the methods described in (Ventrice et al, 2013). Velocity potential at pressure level 200hPa, which serves as smoother measure of convective aqctivity, is calculated for each reanalysis. Anomalies of the velocity potential and zonal wind anomalies at pressure levels 200hPa and 850hPa are then band-pass filtered between 20 and 100 days, and symmetrically averaged between 15°N and 15°S. The combined EOF is then calculated for the resulting fields, which provides the structures of the coherent meridional variability. The sum of squares of the two leading principal components provide the VPM index, which is used as the time series of the occurrence of MJO events.

This method is applied to each of the five reanalyses (including each ensemble member of the CESM-DART reanalysis) as well as a 20-year record of ERA-I between 1991 and 2010. ERA-I is used as a proxy for true MJO structures since it uses data assimilation schemes over a time period with many observations and is therefore heavily validated (Dee et al, 2011).

Frequency-wavenumber spectra are also computed for key daily sampled fields for each data assimilated product. Zonal wind fields between 15N and 15S are taken from each reanalysis product at 850hPa and 200hPa levels. The symmetric power spectrum for the wind fields are smoothed and divided by the average zonal wind power spectrum for each

reanalysis field, which is the procedure advocated by Wheeler and Kiladis (1999). The resulting spectra for each data assimilated product can then compared with each other to search for differences in amplitude and propagation rates for various equatorial waves.

In order to isolate the impact of assimilation on the freely propagating waves within CESM, we also ran a 10-year non-assimilated CESM run with the same parameter set as CESM-DART and computed the same diagnostic. This Free-CESM spectrum can then be compared directly to the CESM-DART spectrum.

2.4 Results

2.4.1 Precipitation

Figure 2.1 shows the difference maps of the reanalyses mean precipitation relative to GPCP observations. Hashed regions show areas in which the difference of mean precipitation structures exceed the errors prescribed by GPCP estimates, which are also plotted. These high-error areas are largely confined to the climatological high-precipitation tropical and subtropical regions across all reanalyses, which typically tend to rain too intensely in the reanalyses. Figure 2.1 also reveals that JRA55 and R1 have the largest areas of significant bias. Table 2.1 shows CERA20C has the highest correlation with GPCP precipitation structure, while R1 has the lowest.

The variability of precipitation also is of interest, which was computed as monthly-mean standard deviations around the long-term mean. This is shown in Figure 2.2 as a five-product mean by averaging the precipitation variability from all five products. The figure additionally shows how all products compare against the averaged variability. CESM-DART shows the highest amount of additional variability, while CERA20C shows the lowest amount of variability.

2.4.2 Surface heat fluxes

Figure 2.3 shows the difference maps of the temporal mean structure of latent heat flux and sensible heat flux, relative to the OAFlux. The magnitude and and structures of the differences vary widely among the models, with the largest errors associated with JRA55. responses between OAFlux and the different reanalyses. Among all products, there is an extension of positive bias within the equatorial cold tongue, and increased biases within tropics for both the Atlantic and Indian basins. For sensible heat flux there is broad tropical bias across the basins with a reduction in the eastern boundary currents.

Table 2.1 shows an overall very high pattern correlation in performance for mean latent heat flux across all models, but with much lower pattern correlations for mean sensible heat flux, especially for CESM-DART. Table 2.2 additionally quantifies some of the differences seen in Figure 2.3 by showing the spatial mean error and the rms error averaged over two regions, the Pacific and Indian Oceans. CERA20C tends to outperform other reanalysis products while CESM-DART tends to perform within the range of the other products.

The seasonal cycle of the monthly climatology of surface heat fluxes was also computed for the five products and compared to that of the OAFlux. Figure 2.3 shows the differences in July bias and January bias, which gives a perspective of the magnitude of the error in the seasonal cycle. The largest errors in the latent heat flux seasonal cycle magnitude tend to occur in the northwestern Pacific and the low-latitude South Pacific and Indian Oceans for all the products but JRA55, which has much lower amplitude errors. The largest errors in the sensible heat flux seasonal cycle magnitude occur in R1 and JRA55, and are smallest in CESM-DART.

2.4.3 Equatorial waves

The frequency-wavenumber content of the atmospheric equatorial waves is displayed in Figure 2.5, which was produced using the symmetric lower tropospheric zonal wind response normalized against the background spectrum for all the different reanalyses. The top plot displays the five- product average spectrum, with the lower plots showing the spectral deviations from this average. The spectra highlight three types of energetic waves. The wavenumber 1-to-2 band, with frequencies lower than 0.05, represents the MJO response. The response along the linear slope for wavenumbers 2 to 8 represents the nondispersive equatorial Kelvin waves (Gill, 1982). Finally the response for wavenumbers -6 to -1 and frequencies 0.15 to 0.3 represent inertia- gravity waves (Gill, 1982).

CESM-DART and R1 both exhibit enhanced MJO activity compared to the other products, while CERA20C has the lowest amplitude MJO peak relative to the mean spectrum. In the Kelvin wave band, several models tend to exhibit coherent spectral differences that increase or decrease the slope of the dispersion relation, indicating slight differences in preferred propagation speeds for each model (R1 producing faster waves, CESM-DART and JRA55 producing slower waves). CERA20C and ERA2C both exhibit enhanced Kelvin wave energy compared to the mean spectrum. CERA20C and ERA20C both exhibit far more energy in the IG band than the mean spectrum, while R1, CESM-DART and JRA55 all have much weaker IG energy.

Figure 2.6 shows the effect of using DART on CESM for upper and lower tropospheric symmetrical zonal wind responses. These spectra clearly reveal that data assimilation in CESM increases the magnitude of the MJO response and dampens the long and slow Kelvin waves with wavenumbers around 3. The assimilation process does not have a strong effect on Kelvin waves with smaller scales, which differ highly between the reanalysis products in that band, as seen in Figure 2.5.

2.4.4 MJO

The structure of the leading combined EOFs found in the formulation of the VPM index is shown in Figure 2.7. The basic zonal wavenumber-1 structure and regional phasing occurs for all the products. However, a surprisingly large disparity between the reanalysis products is evident. CESM-DART is most similar to JRA55, while R1 generally differs the most from the other products. The ensemble spread within CESM-DART generally encapsulates the JRA55 structures, but not so for the other products. The tropospheric wind convergence and divergence centers are shifted eastward for R1 and westward for ERA20C and CERA20C compared to CESM-DART and JRA55 in the upper tropospheric divergence centers.

The time series of the VPM index for each reanalysis product of presented in Figure 2.8. MJO events tend to all occur at the same times across all reanalyses. There exist significant variations in the amplitudes, however, with individual events sometimes varying by 30-50% among certain products. CESM-DART is again most similar to JRA55, although a few events fall outside the envelop of the ensemble spread. These results show the sensitivity to the assimilation scheme when representing even a dominant and well-known intraseasonal climate mode.

The composite MJO structure and phasing is frequently shown as spatial maps of velocity and convective activity during eight phases as it propagates around the globe. These are constructed by averaging the band-passed MJO activity when the MJO index (here, the VPM) exceeds a certain value (here, VPM¿1.5). Figure 2.9 shows the differences between the composites of MJO for the five different reanalyses and the composite MJO for ERA-I (taken here to represent its true structure). CESM-DART, JRA55, and CERA20C exhibit the closest response to the observed in upper tropospheric velocity potential while R1 shows the largest disparity. The differences in lower troposphere wind vectors are less pronounced among the fields, although R1 again tends to exhibit the largest errors. The phase of the errors

reflects somewhat similar behavior found in Figure 2.7, especially with R1 being shifted too far eastward.

These composite map differences are quantified in Table 2.3, which shows the spatial mean and rms errors of each composited MJO field (from Figure 2.9) averaged from 15N to 15S circumglobally. The spatial mean bias of most of the fields, compared with ERA-I, is minimal and near zero mainly because the fields have wavelike regions of positive and negative error. For both seasons, CERA20C and JRA55 both exhibit the smallest rms errors in lower tropospheric winds. For summer MJO events, CESM-DART has the smallest velocity potential rms error, while R1 has the largest. For winter events, CERA20C and JRA55 both outperform the other three products for all the fields in terms of rms error.

2.5 Discussion and Conclusion

In this study, we analyzed a new prototype coupled ocean-atmosphere Ensemble Kalman Filter reanalysis product of the 1970s (Karspeck et al, 2018) by comparing its tropical climate variability to other reanalysis products, available observations, and a free-running version of the model. The results reveal that CESM-DART produces fields that are comparable to those of other reanalyses, but not overwhelmingly different or highly superior or inferior. However, several things stand out that we highlight here.

The clearest signature of the improvements in the fields CESM-DART is in the analysis of MJO and other tropical atmospheric waves. MJO energy is enhanced over the free-running CESM (Fig. 6) as well as compared to the other products (Fig. 5). This suggests that the flux coupling at the ocean-atmosphere interface may be important to organizing the convective activity inherent in MJO. In addition, the spurious high-frequency free Kelvin waves in CESM-DART are reduced in amplitude compared to the Free-CESM run and the other products, again supportive of the oceanic coupling playing a role in this improvement.

Conclusive evidence, however, of the importance of the coupling could only be achieved by comparison with a separate uncoupled CESM-DART run, which is not available but should be explored in future work.

An increase in equatorial atmospheric tropical wave fidelity could increase ENSO forecast accuracy, since they constitute one of the primary mechanisms for instigating ENSO development. This should also be explored in further studies that include seasonal to interannual forecasts using states initialized from the coupled product.

Another positive feature of the CESM-DART reanalysis is the relatively low bias it exhibits in the tropical precipitation field. The typical local bias is comparable to both CESM20C and ERA20C, and is much lower than that of JRA55 and R1 (Fig. 1). Tropical precipitation variability is also found to be largest in CESM-DART (Fig. 2), suggestive of an important feedback between the coupled SST field and tropical convection fields.

An additional highlight of the CESM-DART product is that it produced the lowest bias in mean sensible heat flux in the tropical regions (Fig. 3). Globally averaged, the mean sensible heat flux bias (Table 2.2) was also comparable to the lowest biased products, CERA20C and R1. However, the lowest overall bias in latent heat fluxes occurred in the CERA20C product. We note that these comparisons rely on treating OAFlux as observations, when OAFlux actually uses both the ERA20C and the R1 reanalysis output in the creation of the product. One would expect that this would render a more favorable comparison of OAFlux with those two models. But each of the two coupled products yield better results in their own way, supporting the idea that coupling can enhance the skill of state estimates.

Radiative fluxes were not studied here, but they are known to vary widely among the reanalysis products, particularly in the tropical Pacific. Future studies should address the potentially important effects of these differences on the estimation of the state vectors.

When analyzing these results, there are important issues to note about the data used and the assimilation frameworks. Surface flux data is not directly assimilated into any of

the analysis products, but is estimated in different ways. CESM/DART assimilates available wind and temperature measurements from the ACARS record for the atmosphere, and salinity and temperature data from BUFR records for the ocean, with surface fluxes determined as a consequence of the coupled dynamics. While CERA20C also produces surface fluxes from coupled dynamics, it assimilates only sea surface pressure and marine winds in the atmosphere, plus SST and salinity data in the ocean. ERA20C, on the other hand, assimilates only sea surface pressure and marine winds, while using specified SST to produce the surface fluxes. JRA55 and R1 assimilate available atmospheric data while using specified SST to produce the surface fluxes. The often large disparity in surface heat flux estimates found here, and their impact on the entire climate system, is a consequence of these different approaches and requires further study to implicate mechanisms.

In conclusion, the CESM-DART framework holds great promise for improving reanalysis products by including ocean-atmosphere coupling. In addition, the application of new analysis techniques, experimental forecasts, and complementary control runs, can provide insight on the coupling mechanisms involved in producing the assimilated CESM-DART fields. For example, additional analyses with CESM-DART are currently in progress to analyze the ensemble members of the state estimates in a reliability budget (Rodwell et al, 2016) to determine where and when the biases arise the system and how they are linked to climate modes of variability. The results of the current study and our future work will help guide the development and improvement of ensemble climate forecast strategies for diagnostics and predictions.

2.6 Acknowledgements

This research was only possible because of the immense effort by Alicia Karspeck and the NCAR team, who generously provided us with the CESM-DART datasets. This

research was supported by a grant from NOAA Climate Variability and Prediction Program (NA14OAR4310276) and the NSF Earth System Modeling Program (OCE1419306). High-performance computing support from Yellowstone (ark:/85065/d7wd3xhc) was provided by NCARs Computational and Information Systems Laboratory (CISL), which is sponsored by the National Science Foundation. The data were provided by the Data Support Section of CISL at NCAR. This chapter has been published in: **Eliashiv, J**, Subramanian A, Miller AJ (2019). Tropical climate variability in the Community Earth System Model - Data Assimilation Research Testbed. *Climate Dynamics*, submitted, *sub judice*. The dissertation author was the primary researcher and author of this paper.

Table 2.1: Pattern correlation (latitude weighted) between reanalysis products and observations (GPCP and OAFlux) of annual mean precipitation, latent heat flux, and sensible heat flux.

	Precip	LH	SH
CESM-DART	0.84	0.96	0.46
ERA 20C	0.80	0.95	0.69
NCAR R1	0.62	0.95	0.67
CERA 20C	0.91	0.97	0.72
JRA 55	0.88	0.97	0.72

Table 2.2: Spatially averaged mean and rms differences between reanalysis products and observations (OAFlux) of annual mean latent heat flux and sensible heat flux for Pacific Ocean and Indian Ocean.

		NCAR R1	CESM-DART	ERA20C	CERA 20C	JRA55
	Basin					
Precip (mm/day)	Pac	0.56±1.0	0.03±0.86	0.62±1.1	0.81±1.7	3.7±2.8
	Ind	0.46±0.93	0.29±0.85	0.65±1.2	0.53±1.4	3.3±2.7
Latent Heat (W/m^2)	Pac	19.±8.4	7.5± 8.5	18.±10.	13.±12	35.±9.4
	Ind	19.±10.	6.8±12.	25.±13.	16.±12.	32.±16.
Sensible Heat (W/m^2)	Pac	2.9±1.5	4.8±4.3	7.8±2.1	3.5±2.5	12.±3.8
	Ind	2.9±1.7	6.3±4.2	8.4±2.8	4.6±2.7	13.±2.4

Table 2.3: Spatially averaged mean and rms differences between reanalysis products and observations (ERA-I) across all eight MJO composite phases of U850, V850, and VP200 for summer and winter composite MJO.

	Reanalysis	U850 (m/s)	V850 (m/s)	VP200 $(10^6\ m^2/s)$
Summer	ERA 20C	-0.01 ± 0.37	-0.01 ± 0.19	-0.09 ± 1.3
	CESM-DART	0.00 ± 0.37	-0.01 ± 0.19	-0.00 ± 0.70
	NCAR R1	0.00 ± 0.51	-0.01 ± 0.23	-0.01 ± 1.9
	CERA 20C	-0.01 ± 0.05	-0.00 ± 0.02	-0.03 ± 1.2
	JRA 55	-0.01 ± 0.07	-0.00 ± 0.03	-0.03 ± 1.8
Winter	ERA 20C	-0.00 ± 0.37	0.00 ± 0.19	0.00 ± 1.4
	CESM-DART	-0.00 ± -0.40	-0.01 ± 0.18	0.00 ± 0.72
	NCAR R1	0.00 ± 0.51	0.00 ± 0.21	0.00 ± 2.0
	CERA 20C	0.01 ± 0.03	-0.00 ± 0.02	-0.02 ± 0.48
	JRA 55	0.01 ± -0.04	-0.00 ± 0.02	-0.01 ± 0.43

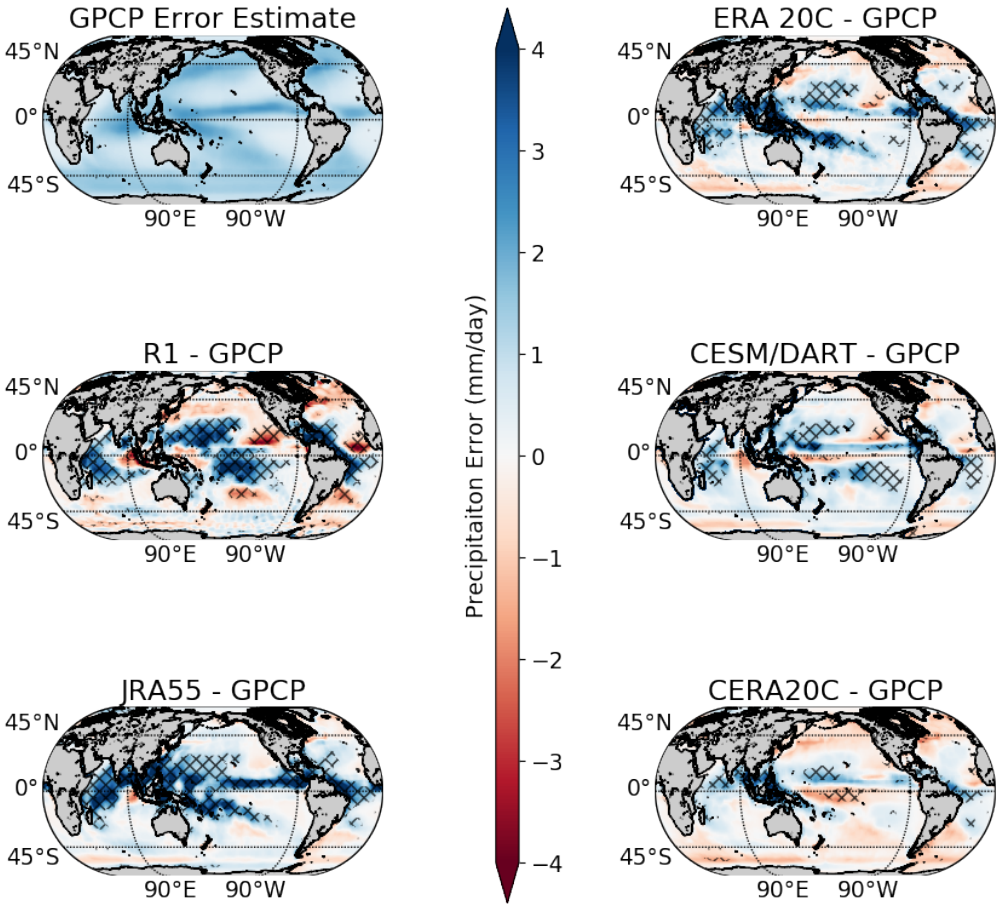

Figure 2.1: Annual mean (1970-79) precipitation differences for the five reanalyses relative to GPCP observations (1979-2015). All datasets are detrended with respect to 1975. The error estimates from GPCP as shown in the top left. Hashed regions in the other panels indicate where the mean difference is greater than GPCP error.

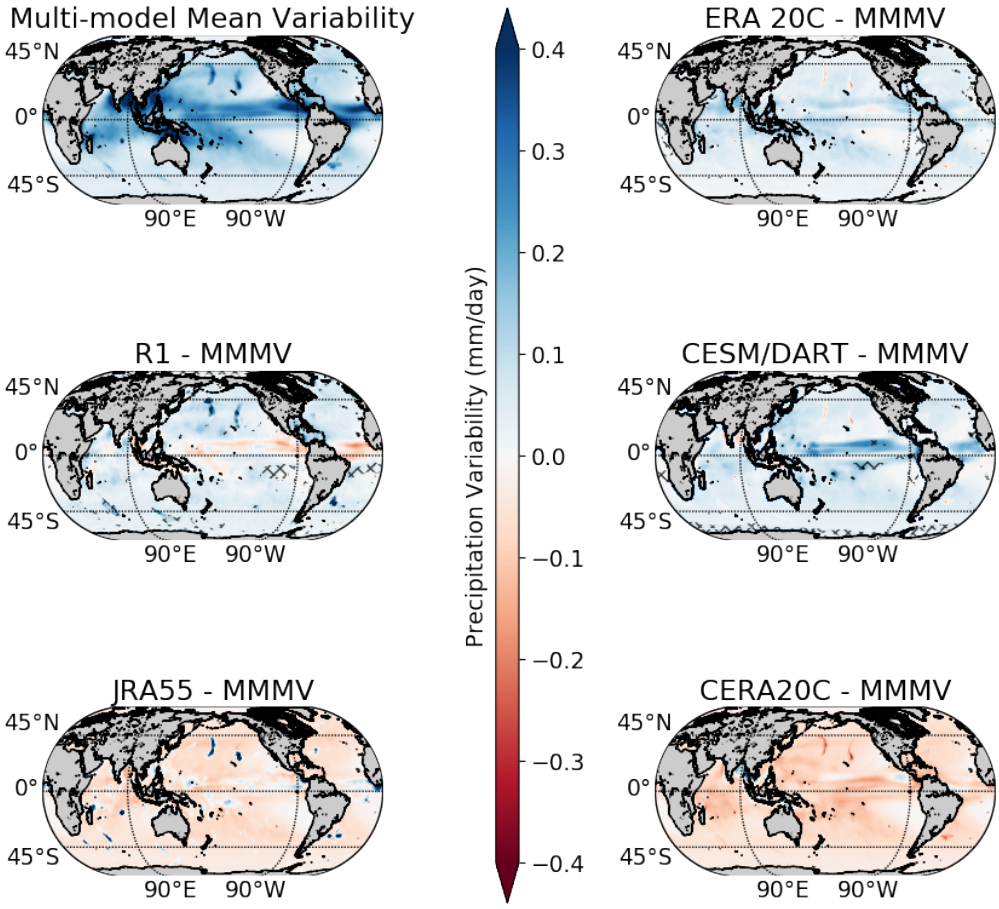

Figure 2.2: Monthly-mean standard deviation of precipitation, calculated relative to the annual mean climatology (1970-79), shown as a five-product mean and differences from that mean for each reanalysis product. Hashed regions in the other panels indicate where the mean difference is greater than GPCP error.

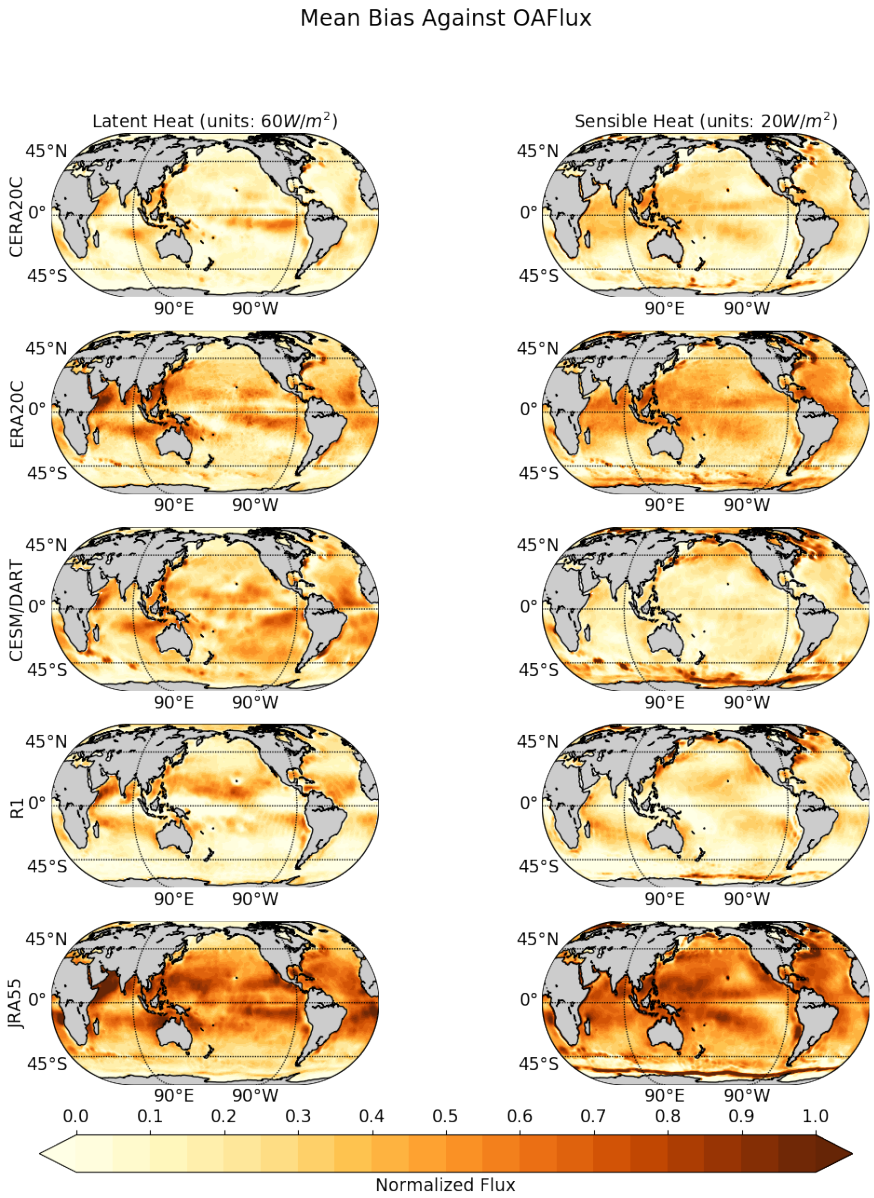

Figure 2.3: Annual mean (1970-79) latent (left) and sensible (right) surface heat flux differences for the five reanalyses relative to OAFlux observations (1970-79). Differences are scaled by a normalization factor of 60 W/m^2 for latent heat and 20 W/m^2 for sensible heat.

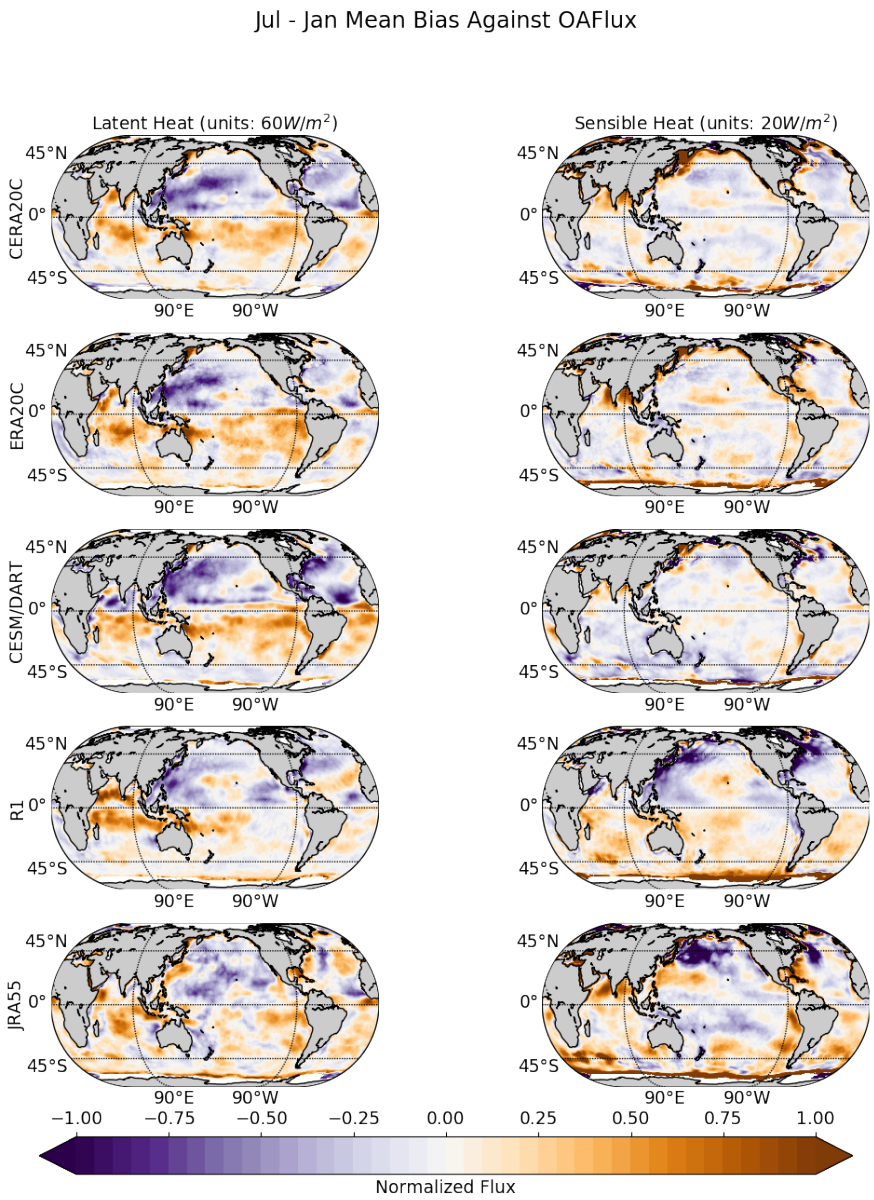

Figure 2.4: Differences between July and January reanalysis biases relative to OAFlux, calculated as in Figure 2.3.

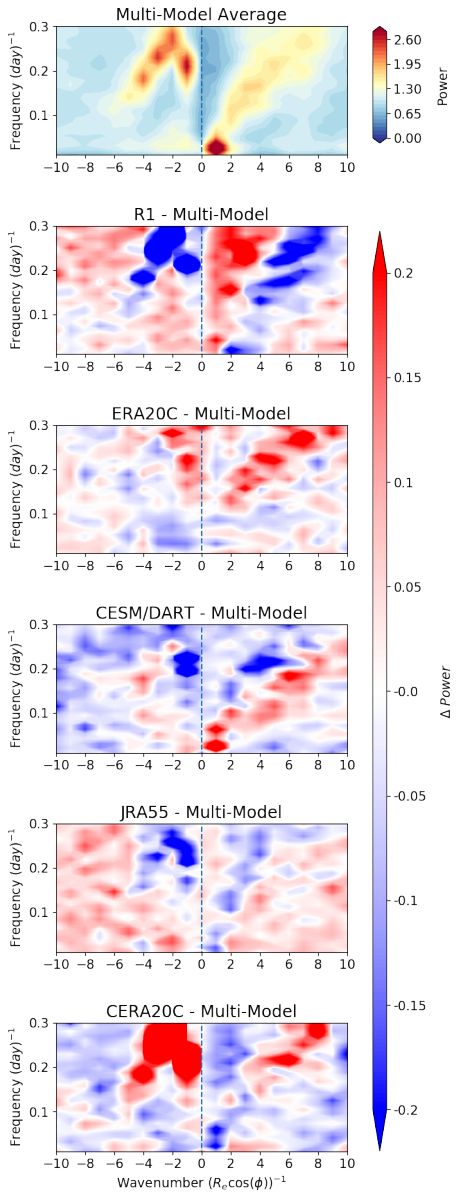

Figure 2.5: Frequency-wavenumber spectra of equatorially symmetric lower tropospheric (850hPa) zonal winds, plotted as the five-product mean and as differences from that mean spectrum for each reanalysis (1970-79).

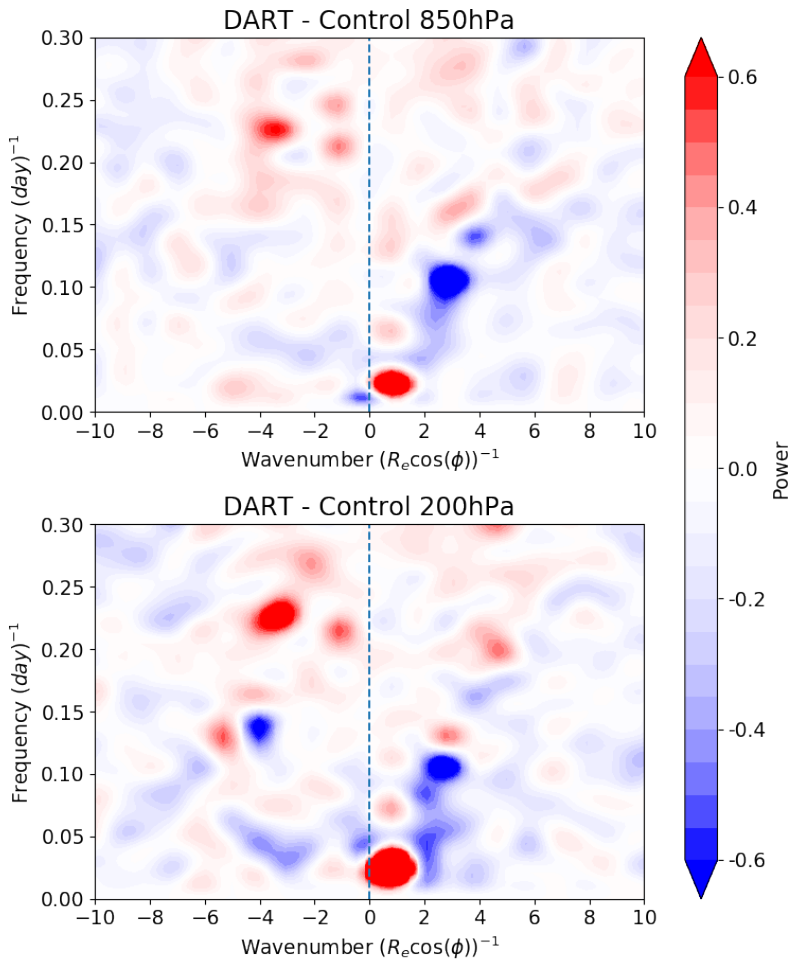

Figure 2.6: Frequency-wavenumber spectra of equatorially symmetric lower tropospheric (850hPa) zonal winds and upper tropospheric (200hPa) zonal winds, plotted as the difference between the spectrum for CESM-DART and the spectrum for Free-CESM.

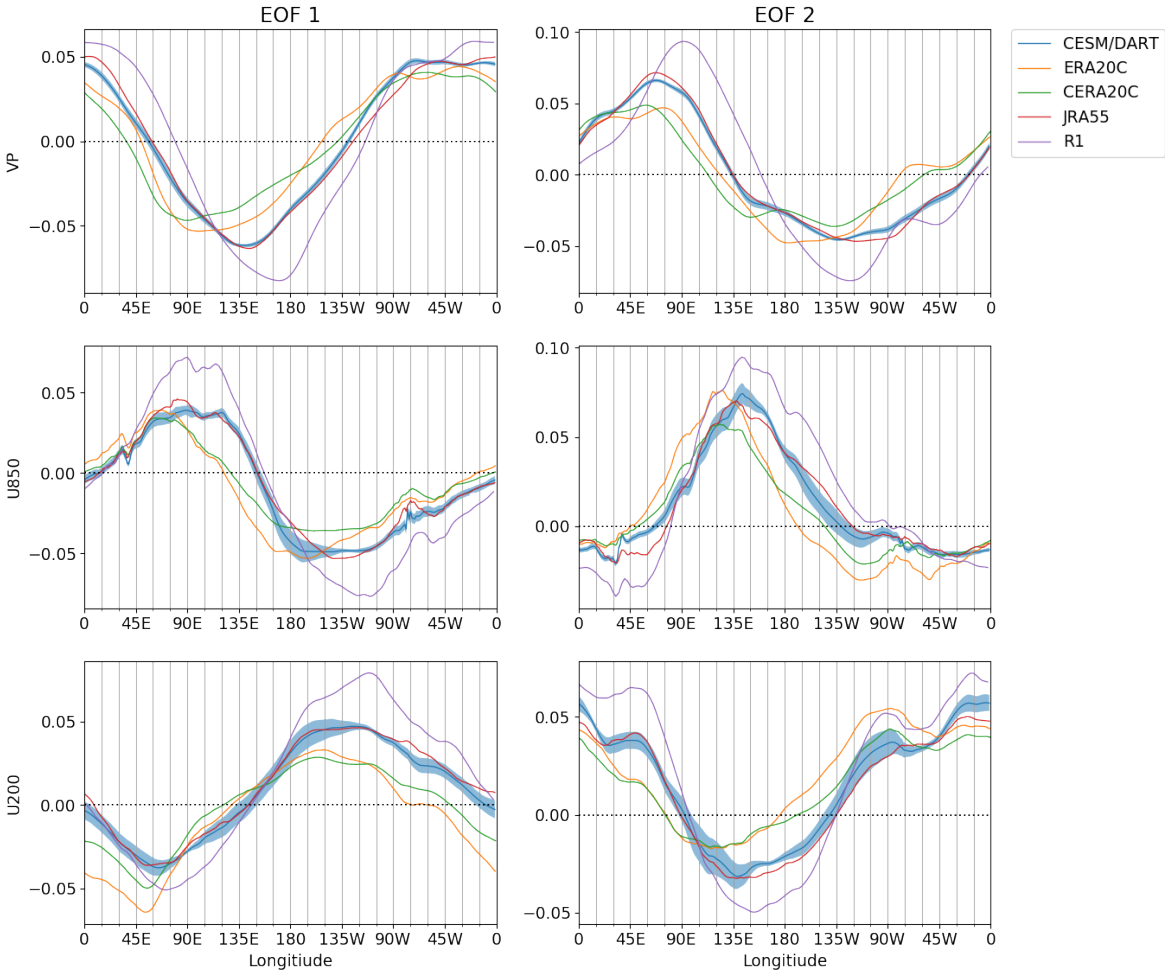

Figure 2.7: Combined EOF of lower tropospheric zonal winds, upper tropospheric zonal winds, and upper tropospheric velocity potential for each reanalysis product, computed for the MJO period band in the tropics, 1970-79, following the procedure of Ventrice et al. (2013). The blue line shows the ensemble mean of the CESM-DART EOF structure with the shading indicating one ensemble standard deviation from the mean.

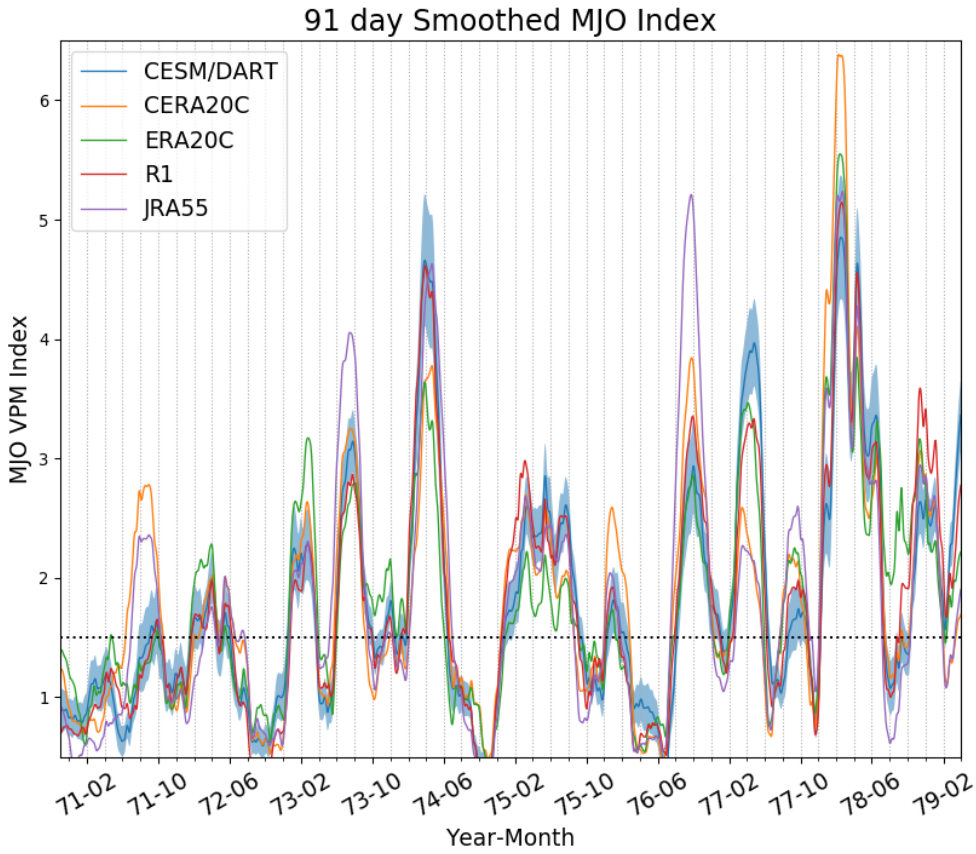

Figure 2.8: As in Figure 7, but for the corresponding MJO VPM Index PC calculated for each reanalysis product.

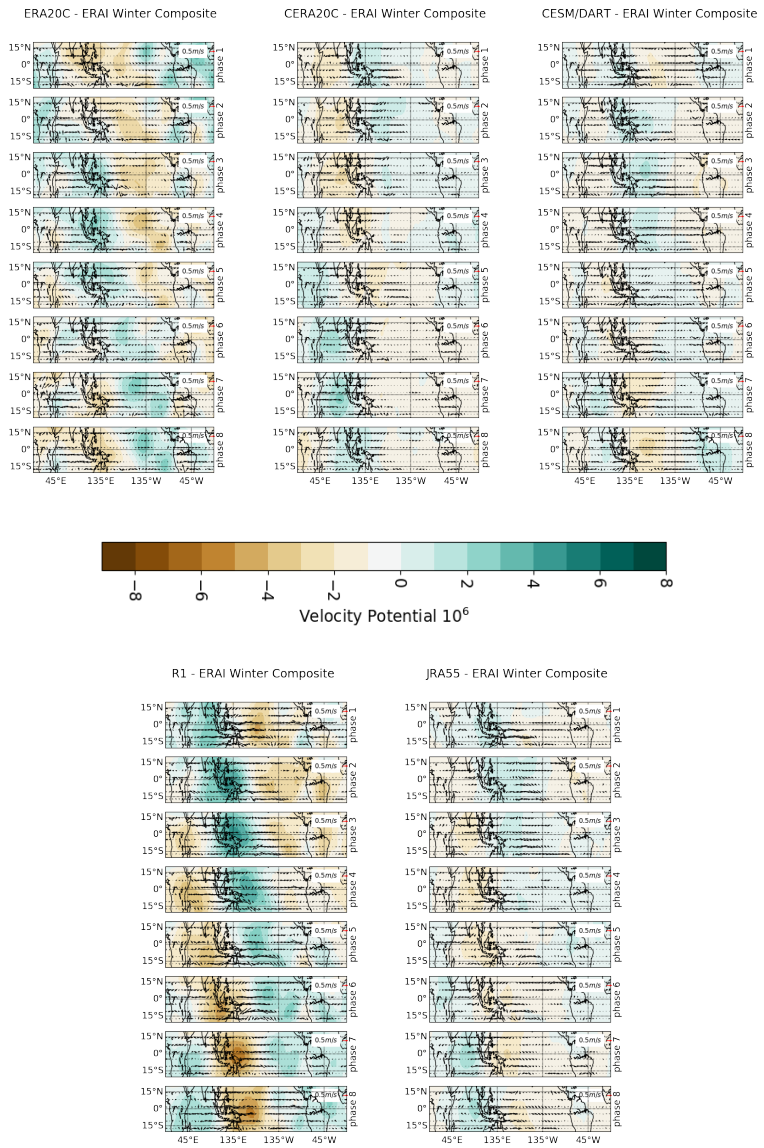

Figure 2.9: Differences in phases through a composite MJO cycle in winter months shown as differences for each reanalysis product (1970-79) relative to the corresponding composite computed from the ERA-I (1991-2010). The contours indicate velocity potential at 200 hPa and the vectors indicate 850 hPa winds.

Chapter 3

A Reliability Budget analysis of CESM-DART

A Reliability Budget is used to diagnose potential sources of error (Departure from observations) in a new prototype coupled ocean-atmosphere Ensemble Kalman Filter reanalysis product, the Community Earth System Model using the Data Assimilation Research Testbed (CESM-DART). In areas with sufficient observations, the mean Bias in zonal wind was generally very low compared to the spread due to Ensemble Variance, which did not exhibit patterns associated with Northern Hemisphere jet streams but did have regional enhancement over the Maritime Continent. However, the Residual term was often the largest contributor to the budget, which is problematic, suggesting improper observational error statistics and inadequately represented Ensemble Variance statistics. The Departure and Residual exhibit significant seasonal variability, with a strong peak in boreal winter months, indicating the model's deficiencies during the energetic Northern Hemisphere winter. Ocean temperature contained large error in areas with eddy production indicating inadequate Ensemble Variance due to poor model resolution. Periods when the the Madden-Julian Oscillation (MJO) was active exhibited lower error, especially in the western equatorial Pacific during MJO phases

with reduced convection. In contrast, during MJO phases with enhanced convection in that region, the Ensemble Variance is increased yet the error is comparable to non-MJO conditions, suggesting a controlling effect of the precipitation parameterization. Further studies evaluating the impact of the coupled assimilation procedure on the reliability Budget will be illuminating.

3.1 Introduction

Global climate models are well known to contain errors that lead to biases in the mean state and seasonal cycle, as well as to inaccurate representation of variability (e.g. Bretherton et al, 2012; Neale et al, 2013). A great deal of research has been executed to attempt to identify the source of these biases and to alleviate them by developing better parameterizations or increasing resolution (e.g. Zhang and McFarlane, 1995; Phillips et al, 2004; Jochum, 2009; Neale et al, 2008; Gent et al, 2010; Hwang and Frierson, 2013; Subramanian and Zhang, 2014). But the problem remains and is still being attacked so that new approaches to find and fix these errors are actively being sought because of their critical importance to improving weather and climate forecasting.

A promising technique for potentially unravelling some of the sources of model error is the reliability budget of Rodwell et al (2016). In this strategy, the differences that arise between a model's ensemble members and the target observations of forecasts are broken down into the parts due to mean bias, ensemble spread, observational error, and residuals. This can be applied for any selected field in the model state vector, for any time period, and for any region. The results can potentially shed light on what processes are most likely to be responsible for error growth, where and when they arise, and how they relate to changing background states associated with climate modes of variability (Rodwell et al, 2018).

Here we use the reliability budget to study errors in a new prototype coupled reanalysis

product, the Community Earth System Model (CESM) in the Data Assimilation Research Testbed (DART) (Karspeck et al, 2018). CESM-DART invokes a weakly-coupled data assimilation scheme using an Ensemble Adjustment Kalman Filter (EAKF) (Anderson et al, 2009). Applying the reliability budget to this ensemble-based coupled reanalysis could provide new insights into the climate processes that conspire to generate model errors in CESM. This is because the surface fluxes, which are the primary means of communication between the atmosphere and ocean, are computed as part of the assimilation scheme itself (Penny and Hamill, 2017). And the estimates of these fluxes from coupled climate models are often cited as containing the biases and random errors that result in inaccurate representations of climate processes and internal modes of variability (e.g. Arnold et al, 2015).

Our reliability budget analysis takes a careful look at how the CESM-DART performs overall during the sequential updating that occurs every six hours over the available time period of 1971-82. What regions contain the most mean bias development? Where does ensemble spread due to instability dominate the model-data mismatch? How does the importance of different terms in the reliability Budget change with the seasons or from year to year? How does the model perform in data-rich versus data-sparse regions? We then focus on model performance in representing the Madden-Julian Oscillation (MJO), which is an important and predictable mode of climate variability known to be often poorly represented in climate models (e.g. Waliser et al, 2009; Kim et al, 2009; Subramanian et al, 2011). We break up the budget into times when MJO is active compared to when it is not active. Does the presence of MJO increase or reduce error? Which phases of MJO generate the most error? These and other issues are addressed in this study, as we try to relate this statistical breakdown of errors to focal points in the dynamics that require attention for future model improvements.

In the next section, we present the basics of the reliability budget (including an expanded version of the budget that is supported by a simple model diagnosis in the Appendix),

explain the CESM-DART reanalysis in more details, and describe how MJO conditions are isolated in the model. In section 3, the results are presented for applying the budget to atmospheric and oceanic variables, as well as for MJO vs. non-MJO conditions. Section 4 contains a summary discussion and conclusion.

3.2 Methodology

3.2.1 Reliability Budget

The framework for our analysis is the reliability budget of Rodwell et al (2016) that can partition the mean-squared differences of the ensemble mean error (i.e., relative to observed data values) into Bias, Ensemble Variance and Observational Uncertainty contributions, along with a Residual that indicates a breakdown of the assumptions used in deriving the budget. This is applied to an ensemble of forecasts, which in this case are the ensemble members used to compute the ensemble-mean reanalysis fields of CESM-DART, at each update time. Averaging the budget terms over time and/or space gives an aggregate perspective of the model performance and the source of mismatch between forecast and observed.

Under the assumptions of perfect reliability and a perfect knowledge of the observation-error characteristics (Rodwell et al, 2016), the reliability budget can be written as:

$$\overbrace{\overline{(m-o)^2}}^{\text{Departure}} = \overbrace{\overline{m-o}^2}^{\text{Bias}} + \overbrace{\overline{e_m^2}}^{\text{Ensemble Variance}} + \overbrace{\overline{e_o^2}}^{\text{Observation Uncertainty}} + \underbrace{(Var(o) - \overline{e_o^2}) + (Var(m) - \overline{e_m^2}) - 2Cov(m,o)}_{\text{Residual}} \tag{3.1}$$

where o is an observation (including error), m is the ensemble mean model state variable, $\overline{e_m^2}$

is the Ensemble Variance of the model system, and $\overline{e_o^2}$ is the variance of the Observational Uncertainty. Variance and covariance for arbitrary variables, x and y, are defined as

$$Var(x) = \overline{x^2} - \overline{x}^2$$

$$Cov(x,y) = \overline{xy} - \overline{x}\,\overline{y}$$

so that one can see that equation (1) simply corresponds to the expansion of the variance of the model-mean and observation differences (that is, the Departure from observations) as follows:

$$\begin{aligned} Var(m-o) &= Var(m) + Var(o) - 2Cov(m,o) \\ \overline{(m-o)^2} &= \overline{m-o}^2 + Var(m) + Var(o) - 2Cov(m,o) \end{aligned} \tag{3.2}$$

If the reliability budget does not balance so that the Residual is non-zero, then we have either incorrectly assigned observational error statistics or we have inadequately represented forecast variance statistics, or both. We shall see below that the Residual is indeed frequently non-zero, and significant, in our analysis of CESM-DART.

As a further exploration of the 'structural bias' in the model (Williamson et al, 2015), we expand the model ensemble mean state variable, m, into its true value in nature, T, plus two errror terms, one representing the mean error, termed the structural bias, $f(x)$, plus the random error around that mean, e_m. The observation data can additionally be broken down into T plus Observational Uncertainty e_o, which is generally prescribed in the development of the data assimilation scheme. Using these expansions, derivations for the structural bias can be written as:

$$m \equiv T - e_m + f(x) \tag{3.3}$$

$$o \equiv T - e_o \tag{3.4}$$

$$\begin{aligned}
Var(m) &= Var(T) + Var(e_m) + Var(f(x) - 2Cov(t, e_m) \\
&\quad + 2Cov(T, f(x)) - 2Cov(e_m, f(x)) \\
&= Var(T) + Var(e_m) + Var(f(x)) + 2Cov(T, f(x)) \\
Var(o) &= Var(T) + Var(e_o) - 2Cov(T, e_o) \\
&= Var(T) + Var(e_o) \\
Cov(m, o) &= Cov(T + f(x), T) \\
&= Var(T) + Cov(f(x), T) \\
\delta &\equiv m - o \\
Var(\delta) &= Var(e_m) + Var(e_o) + Var(f(x)) \\
Var(e_o) &= Var(\delta) - Var(m) + Cov(m, o) + Cov(T, f(x))
\end{aligned}$$

If the structural bias is assumed to be linear or approximately linear, the strength of the Bias and Ensemble Variance can be isolated and estimated:

$$f(x) \approx Ax + B$$

$$Cov(o,x) = Cov(T,x)$$

$$Cov(m,x) \approx Cov(T,x) + AVar(x)$$

$$A \approx \frac{Cov(\delta,x)}{Var(x)}$$

$$Var(e_o) \approx Var(\delta) - Var(m) + Cov(m,o) + ACov(o,x) \tag{3.5}$$

$$Var(e_m) \approx Var(\delta) - Var(e_o) - A^2 Var(x) \tag{3.6}$$

Both (5) and (6) can then be directly calculated from the observations and the model reanalysis products. The results can then used to breakdown their influence on the Residual, i.e., determine possible reasons for the unbalanced budget.

We will apply the relations from equations 3.5 and 3.6 to CESM-DART in the next section. The purpose of using the above technique is to test whether the true Ensemble Variance is of the same order as the variance produced by the model output, as well as whether the Observational Uncertainty is specified correctly. However, the approximations invoked may not hold true, so we apply the technique to a simplified theoretical model (Lorenz, 1963) where we can specify the errors as known values in both model and observations and thereby test under what conditions the linearity approximation for structural bias is valid. This analysis is presented in the Appendix. The results reveal that one of the limitations found for this approximation to the budget terms is that model error variance, e_m^2, and the Observational Uncertainty variance, e_o^2, need to be of the same order of magnitude. This analysis does show that using this approach allows for a quantification of linear bias, A, with less than 8% error when using at least 1000 data points, and less than 20% error in the quantification of background bias, B, when using at least 40,000 points, even when sampled

randomly. This quantification may not be as effective when using the CESM-DART product of the 1970's because there is a limitation due to the low Observational Density. However for systems using modern sensors and data, the structural bias can be likely be identified effectively using this linearized estimate.

3.2.2 CESM-DART

The DART system is an open-source community developed software used to implement the Ensemble Adjustment Kalman filter (EAKF) data assimilation technique into different models (Anderson et al, 2009). The application of DART to the CESM (Karspeck et al, 2018) provides a novel depiction of climate data for the 12-year period that it was run and made available to us (1971-1982). CESM-DART uses CESM1.1.1, which is a coupled model which combines the community atmospheric model version 4 (Gent et al, 2011), with the Community Land Model (CLM4), the Community Ice Code (CICE4), CESM River Transport Model (RTM), and the LANL Parallel Ocean Program (POP2) for a coupled earth climate model (Hurrell et al, 2013). CESM-DART is run with a 30-member ensemble, with data assimilated every 6 hours for atmospheric fields and 24 hours for oceanic components. The components are coupled together every 24 hours. ACARS data is used to adjust CAM state vectors and BUFR data is used for POP state vectors. Aircraft, radiosonde, and a few satellite measurements of wind and temperature are used for the atmosphere model. Bottle, DBT, STD, CTD, MBT, and XBT measuruments of temperature and salinity are used for the ocean model. These reanalysis fields data are compared to the individual assimilated observations before the next Kalman Filter update is applied at each sequential update time and for each observation.

3.2.3 MJO index

We use MJO events as one measure of intraseasonal coupled climate variability. To isolate MJO events, the velocity potential MJO (VPM) index is computed using the methods described by Ventrice et al (2013). Velocity potential is calculated at pressure level 200hPa, which serves as smoother measure of convective activity. Anomalies of the velocity potential and zonal wind anomalies at pressure levels 200hPa and 850hPa are then band-pass filtered between 20 and 100 days, and symmetrically averaged between 15°N and 15°S. The combined EOF (Wheeler and Hendon, 2004) is then calculated for the resulting fields, which provides the structures of the coherent meridional variability. The sum of squares of the two leading principal components provide the VPM index, which when exceeding 1.5 is used to indicate the occurrence of MJO events.

3.3 Results

3.3.1 Reliability Budget

Since the reliability budget is applied in observation space, we first examine where the analysis can be effectively used during the available CESM-DART reanalysis period, 1971-82. Figure 3.1 shows the observation density of all the different variable types assimilated over 1971-82 in CESM-DART. Each variable uses multiple products from ACARS and BUFR for atmosphere and ocean measurements, respectively. Observational density is poor in the Southern Hemisphere, but reasonably high in many areas of the Northern Hemisphere and in the tropical oceanic regions.

The reliability budget was initially applied to each of the different model state variables, averaged vertically and plotted globally. As an example of these results, Figure 3.2 shows the terms of the budget for zonal winds. The top plot shows the rms error (Departure

from observations) for the entire 12-year reanalysis. Notice that this can only be plotted where there was data to be assimilated (plotted at the bottom as observational density), so some remote areas are blank. The two largest terms that contribute to the error are Ensemble Variance and the Residual. Surprisingly, the Ensemble Variance does not exhibit patterns associated with Northern Hemisphere jet streams, but it does have regional enhancement over the Maritime Continent region where deep convection frequently occurs. The Bias is low in most areas, although significant Bias can be found in some regions. The Uncertainty in the input observations (which is specified before assimilation) is generally very small.

In the initial representation of the reliability budget (Figure 3.2), large amounts of Departure can be found in two different types of regions. The first and most dramatic are locations on the southern edges of where data is available in the mid-to-high-latitude Southern Hemisphere. There are energetic flows in those regions, but also very limited observations to constrain the atmospheric activity. This will be further explored below. Another location of higher Departure is in the Northern Pacific and Atlantic basin. This region is not associated with an increase in Ensemble Variance or Observational Uncertainty, nor is the increased Departure visible in the Bias. This indicates that those regions display a higher variability in the model or observation noise that can be explained by either the Ensemble Variance or the Observational Uncertainty. For those areas, we will use the expanded budget analysis derived from equations 3.5 and 3.6 in order to isolate their response.

3.3.2 Expanded Reliability Budget

Figure 3.3 shows the same reliability budget but including the terms $Var(e_m)$, $Var(e_o)$, and $\sum_i A_i^2 Var(x_i)$ from equations 3.6 and 3.5. In Figure 3.3, e_o is derived from data under the assumption of linear bias. Though the estimates for e_o and e_m are not likely to be quantitatively accurate, as found in the Appendix for a simplified model, the spatial structures of the variability of e_o and e_m are likely to be meaningful. However, the plots show that e_m is

not spatially similar to the Ensemble Variance, suggesting that the model does not produce enough variability within the ensemble. The structure of e_o resembles the Residual because the Observational Uncertainty and Ensemble Variance cannot span the entire variability of the climate system. The variability of $\sum_i A_i^2 Var(x_i)$ shows that there is small amount of structural bias distributed through the system, but it is found to be negligible in comparison to the Departure. It is therefore nonlinear in not readily explored in this linearized framework.

This initial foray into exploring the reliability budget motivates us to pursue a more detailed examination of the budget by considering different regions, timescales, variables, and processes, in order to attempt to isolate dominant mechanisms controlling the error development. In particular, we focus later on the MJO as a target climate process, since it gives insight into the growth of errors associated with atmospheric convectively coupled equatorial waves on intraseasonal timescales, which are key to extended-range predictions.

3.3.3 Single dimension exploration

Figure 3.4 shows equation 3.1 for zonal wind as a function of longitude, latitude, height, and time. In the figure, the histogram in the background shows the observation density for the measurements in each dimension. The main feature that is evident is a reduction of the entire budget with higher observation density when plotted versus latitude or longtidue.

The correlation between observation density and each term of the reliability budget is shown in Table 3.1. In the table r_x^2 represents the R^2 value calculated against $log(n+1)$, where n is observation density in the x domain. We find that Departure and Observational Uncertainty have the strongest dependence on observation density when considered as a function of latitude. There is also some weak longitudinal relationship between Departure and observational density, but essentially no vertical or temporal relationship.

Figure 3.4 shows two interesting features in the time domain. There is a general trend towards lower Departures in the latter half of the record as the Observational Density

increases. There is also a clear seasonal cycle in the Departure and Residual, likely indicating a seasonal cycle in the ability of the model to produce observed levels of variability in the Northern Hemisphere where the data is concentrated. This is further explored in Figure 3.5, which shows the seasonal cycle climatology (after subtracting the annual mean) and the anomalies of the budget. The Departure and Residual both have the most seasonal variability, with a strong peak in winter months, and weaker signals in other months, indicating the model's deficiencies during the energetic Northern Hemisphere winter. The anomalies of the budget, however, do not seem to correlate with observational density in time or with any particular climate mode of variability during this 12-year period. Note that he combination of seasonal cycle and interannual variability account for less than 15 % of the total budget.

To quantify the affect of observation density, first order approximations of its effect are derived as follows:

$$
\begin{aligned}
n &\equiv n(\mathbf{x}) = n(t, \theta, \phi, z) \\
R(n) &= R\left(n(\mathbf{x})\right) \\
\frac{\partial R}{\partial n} &= \sum_i \frac{\frac{\partial R}{\partial x_i}}{\frac{dn}{dx_i}} \\
\overline{\frac{\partial R}{\partial n}} &\equiv \int \frac{dn}{d\mathbf{x}} \frac{\partial R}{\partial n} \mathbf{dx} \Big/ \int \frac{dn}{d\mathbf{x}} \mathbf{dx}
\end{aligned}
\tag{3.7}
$$

where n is observation density, t is time, θ is longitude, ϕ is latitude, z is height, and R is the respective reliability budget term.

Using this derivation, Figure 3.6 shows the relative contributions of observation density, time, and spatial dimensions. It can easily be deduced that observation density is by far the dominant producer of error for most aspects of the reliability budget. It also confirms that the overwhelming contribution of the strong meridional dependence on the budget was

from observation density, and not meridional biases. Surprisingly, figure 3.6 also shows that Departure has a very strong dependence on height overwhelmingly over other variables. Though the Bias terms are dominated mostly by observation density, their contributions to the budget are 2 orders of magnitude lower than any other contribution, and are therefore negligible in this context.

Both Ensemble Variance and Observational Uncertainty rise with height, but surprisingly, Ensemble Variance increases with observation density as well. These effects may be caused by sudden influx of observations in areas that were not well covered, which would lead to a sudden model correction. In order to explore the mechanism for this rise, the model would have to be examined with data input in more regular intervals in both time and space for comparison.

3.3.4 Ocean temperature exploration

Figure 3.7 shows the reliability budget for ocean temperatures taken around the globe and throughout the water column, although these are concentrated in the upper ocean. There is a large production of error (Departure) in areas with eddy production in western boundary currents, as well as in equatorial regions where there are tropical instability waves. These errors are reflected in the Ensemble Variance, and to a lesser degree in Bias, but have very small amounts of Observational Uncertainty. In fact, the Departure and Ensemble Variance are highly correlated to each other, showing that areas with unstable currents are being regularly corrected to observations, as would be expected.

The areas with unstable current features have approximately the same order of magnitude of Residual and Ensemble Variance. This shows that Ensemble Variance is actually underestimated through the dynamical processes of the model, indicative of the coarse resolution being incapable of producing energetic eddies. Additionally, the assimilated measurements are given heavy weight even in the unstable regions, and thereby prevent the

growth of error. In more quiescent regions, such as in the subtropical gyres, the Departures are small.

Table 3.2 looks at the effect of ocean water temperature on the reliability budget. Areas with warmer waters have larger amounts of errors than areas with colder waters. This is found in Departure, Bias, and Ensemble Variance. The Ensemble Variance surpasses the Departure for the warmest waters. This shows that the model produces exaggerated variability in especially warm waters, but weaker variability otherwise.

3.3.5 Intraseasonal exploration

In a related study using CESM-DART (Eliashiv et al, 2019, submitted), the data assimilation in CESM significantly improved the performance of MJO events. Focusing on MJO events, which have well known mechanisms involved, may give us added insight into the processes leading to the error growth in CESM-DART. Using the VPM index, time periods when MJO are active (VPN > 1.5) are separated from those when it is not. We then averaged the reliability budget terms over the tropical (30S-30N) Indian (40E-100E) and western Pacific Ocean (130E-170E) regions, where MJO is most convectively active. Ocean temperature is included for this analysis since MJO events convectively couple the oceanic and atmospheric systems within the tropical Indian and Pacific basins.

Table 3.3 shows that there is, surprisingly, a slight reduction in error (Departure) for both atmospheric zonal winds and ocean temperature during MJO events, even though the atmosphere can be more convectively vigorous at these times. To further explore this reduction in Departure during MJO, the VPM index was compared to the anomalies of the global zonal wind reliability budget. They were found to be uncorrelated, indicating that this reduction in error in the Indian and Pacific basins is not simply a coincidental effect that occurs circumglobally during times of MJO.

The composite MJO structure and phasing is frequently shown as spatial maps of

velocity and convective activity during eight phases as it propagates around the globe. We next broke up the reliability budget into each of the eight phases of MJO, compared to times when MJO is not active. Figure 3.8 shows how the different MJO phases influence the reliability budget for the Indian and Pacific basins, specifically within 850 $\pm$ 50 hPa and 200 $\pm$ 50 hPa. These heights are chosen as they are heights at which divergence and convergence specific to the MJO structure occur. There are two key characteristics to notice for this figure. The first is that upper tropospheric errors are reduced for these two basins for all phases, with or without MJO events, in contrast to the full column results shown in Table 3.3 as well as in Figure 3.4 and Figure 3.6, which show an increase in error with height. The second feature is that the Departure exhibits its lowest values for MJO phases 2 to 3 in the lower troposphere of the western Pacific. These are the times of least active convection locally there, as the MJO propagates from the Indian Ocean towards the western Pacific Ocean.

As a final exploration for the mechanism that causes the reduction in error, Figure 3.9 shows how zonal velocity affects the occurrence of error in the western Pacific basin. Plotted as histograms are the available zonal wind observations throughout the troposphere in the equatorial western Pacific region for non-MJO conditions as well for each MJO phase. The zonal wind distribution peaks at negative values due to the easterly trade winds averaged over this area for all MJO phases as well as for non-MJO conditions. Also plotted are the terms in the reliability budget for the zonal winds that land in the histogram bins, averaged over all available observations in the vertical in that region, with the full budget plotted for non-MJO conditions and the differences from the non-MJO budget plotted for each phase. The top figure shows that without MJO events, higher Departures occur when the zonal velocity magnitude increases, with westerlies producing more error than easterlies. It additionally shows that there is no strong dependence of the error production on wind speed except for exceedingly strong winds, for which there are not many measurements to support the accuracy of the budget.

An important feature during the MJO phases is that the Departure is reduced for easterly winds during phases 1, 2, 3, 7, and 8. These are times when the MJO convective activity in the region is reduced. This indicates that the model representation of organized MJO propagation is sufficiently good that it is able to reduce errors in the reanalysis during MJO events. Surprisingly, the Ensemble Variance is not concomitantly reduced, indicating that it is not simply the non-linearity in the local model dynamics that is controlling the reduction in these Departures. Note, in contrast, that the Ensemble Variance does increase significantly in phase 5, when MJO convective activity is maximum in this region, yet the Departure is not changed from non-MJO (normal) conditions. This suggests that the model precipitation parameterization may be responsible for producing the ensemble spread through feedbacks with the dynamics, thereby limiting the overall skill of the reanalysis in this region for both MJO and non-MJO conditions.

3.4 Summary and Conclusion

We used a reliability budget (Rodwell et al, 2016) analysis to diagnose potential sources of error (termed Departure) in the ensemble mean from an Ensemble Kalman Filter-based CESM-DART climate reanalysis of the time period 1971-82 (Karspeck et al, 2018). We focused on zonal winds as a typical example of our results for other atmospheric variables. In areas where there were sufficient numbers of observations, the mean error Bias was generally very low compared to the spread due to Ensemble Variance, indicating the important controls exerted by the unstable dynamics in the atmospheric flows. The Ensemble Variance did not exhibit patterns associated with Northern Hemisphere jet streams, but did have regional enhancement over the Maritime Continent. However, the largest Ensemble Variance was in the mid-to-high-latitude Southern Hemisphere, where there are energetic flows with limited observations to control them. The Residual terms were often the largest

contributors to the budget, which is problematic. This suggests that either the statistics of Observational Uncertainty, which was very small and specified in the application of EaKF, was improperly assigned, or the model inadequately represented Ensemble Variance statistics, or both. Additional experimental designs are needed to better assess these shortcomings in the model, especially for recent time periods where more data is available.

We also attempted to develop a strategy to explore the relative importance of model error and Observational Uncertainty in contributing to the Residual. We used an 'identical twin' framework with a simple, but non-linear, model (Lorenz, 1963) to prescribe known errors for both model and observations, such that the observations contain only random error and the model contains both random error and structural bias (assumed to be linear). However, we found that the estimate of structural bias will only be sufficiently accurate when the observational and model errors are of similar magnitude, which is generally not applicable in more complicated models. Applying the strategy qualitatively to CESM-DART suggested that the model does not produce enough variability within the ensemble in many regions around the globe.

Note that the Bias in our reliability budget is distinct from the mean bias that develops in a free-running climate model in statistical equilibrium after non-linear interactions have acted to balance the flows. In our analysis, the Bias corresponds to the recurring systematic drift of the ensemble of 'forecasts' away from the 'observed' reanalysis initial state and towards the targeted observation, in a tangent-linear sense. There may be no relation between the spatial patterns seem in our Bias maps and the climatologist mean bias maps from a long-term current-climate simulation using CESM. But patterns in the reliability budget Bias (and the other terms) may reveal the tendencies that suggest why the model consistently behaves incorrectly in some region, for a particular variable, during a certain season, or while a particular climate mode of variability is active.

Towards this goal, we next examined the temporal and spatial variability of the

reliability budget. The Departure and Residual exhibit significant seasonal variability, with a strong peak in winter months, and weaker signals in other months, indicating the model's deficiencies during the energetic Northern Hemisphere winter. The year-to-year variability in the budget exhibited no obvious relationship to any climate mode. A sensitivity analysis revealed that, over this time period with limited data, the observation density predominantly controlled the quality of the reliability budget, which explained some of features seen in the latitudinal and longitudinal structures of the budget, since far more data was available over land areas of the Northern Hemisphere.

We also studied the ocean temperature field and, not surprisingly, found a large production of error (Departure) in areas with eddy production in western boundary currents, as well as in equatorial regions where there are tropical instability waves. These areas with unstable current features have approximately the same magnitude of Residual and Ensemble Variance. This is consistent with the coarse resolution of the model being incapable of dynamically producing energetic eddies thereby underestimating Ensemble Variance.

We then studied intraseasonal variability by identifying periods when MJO was active and compared them to non-active time periods to provide insight into the growth of errors associated with atmospheric convectively coupled equatorial waves. We found a slight reduction in error (Departure) for both atmospheric zonal winds and ocean temperature during MJO events, even though this includes times when the atmosphere is more convectively vigorous than normal. After breaking down the budget by the phase of MJO, we noted that the Departures exhibit their lowest values in the lower troposphere of the tropical western Pacific for MJO phases 2 to 3, which are the times of least active convection there. This suggests that weaker convective activity in the atmosphere allows a locally more skillful model representation of the flows. We also found that in the tropical western Pacific during the more convectively quiescent phases of MJO (1, 2, 3, 7, 8) the departure is reduced for moderate-to-strong easterly wind conditions. In contrast, during phase 5, when convection is

enhanced by MJO in that region, the Ensemble Variance is increased over non-MJO times, but the skill is left unchanged. Taken together, this suggests that the model precipitation physics produces enhanced ensemble spread through feedbacks with the dynamics, thereby limiting the overall skill of the reanalysis in this region for both MJO and non-MJO conditions.

These results provide a novel perspective of the quantitative performance of CESM-DART in providing a reliable estimate of the atmospheric and oceanic state during 1971-82. But further studies evaluating the impact of the coupled ensemble-based assimilation procedure on the reliability budget are needed. For instance, since the study was executed in observations space, there was no assessment of the reliability of the surface fluxes because they were not included in the data assimilation procedure (cf. Eliashiv et al, 2019, submitted). Additionally, evaluations using ensemble-based reanalysis products with the high-density observations that are available in more recent decades may also provide clearer perspectives on how to pinpoint the processes in the model that need the most improvement in their parameterization.

3.5 Appendix

We use an 'identical twin' framework with a simple, but non-linear, model to explore the relative importance of model error to Observational Uncertainty. We apply the Lorenz (1963) system, which is derived from the dominant modes of the solution to Rayleigh-Benard problem which describes convective tropospheric behavior as follows:

$$\phi(x,z,t) \propto x_1(t)\sin(\pi x/k)\sin(\pi z/l)$$

$$\theta(x,z,t) \propto \sqrt{2}x_2(t)\cos(\pi x/k)\sin(\pi z/l) - x_3(t)\sin(2\pi z/l)$$

$$\frac{d}{dt}x_1 = \sigma(x_2 - x_1)$$

$$\frac{d}{dt}x_2 = x_1(\rho - x_3) - x_2$$

$$\frac{d}{dt}x_3 = x_1 x_2 - \beta x_3$$

Where ϕ is the stream function, θ is temperature anomalies, $\sigma = 10$ is the Prandtl number, $\rho = 28$ is the rescaled Rayleigh number, and $\beta = 8/3$ is derived from the wavenumbers k and l. Using this system, data is derived from 50 high resolution cycles (loops around the entire attractor) made using a forward Runga-Kutta fourth-order time step. The state variable x_1 is then ascribed as Truth, and used as input to create both an erroneous model state and and uncertain observation.

Errors are then prescribed and added to both model and observations following equations 3.3 and 3.4, so that the observations contain only random error, e_o, and the model contains both random error, e_m, and structural bias, At (linear growth with time) + B (constant background bias). Unless otherwise specified the initial default values of e_o, e_m, A, and B are set to be $\sqrt{var(x_1)}/10$, where $\sqrt{var(x_1)}$ is the rms of the anomalies produced by the model. The random error is generated as a uniform distribution.

Using equations 3.5 and 3.6, we can estimate values of e_o, e_m, A, and B, and examine whether or not they are close to their prescribed values. Figure 3.A1 shows the error in the estimation of these values while varying the magnitude of the specified errors. For example, in the upper left plot, if the observational error, e_o, is very low, the estimate of A and B are both very good (lower than 0.1), but the estimates of e_o and e_m are large and unusable. But if the observational error is near 0.1, then all the variables are estimated well. If the

observational error, e_o, is comparable to the variability of the system ($\tilde{1}.0$), the estimate of A is still useful (lower than 0.1), but the estimates of B, e_o and e_m are large and inaccurate. Likewise, in the upper right panel, the same situation occurs for e_m, but with the roles of e_m and e_o reversed.

The upper panels of Figure 3.A1 clearly show that if e_o or e_m increase to be as large as the natural variance in the system, the error in estimating the other variables increases to levels that render the estimates invalid. Interesting, however, when A or B grow in magnitude, the other value estimates are hardly affected, while the bias estimates themselves grow more and more accurate.

Another major point found by inspecting Figure 3.A1 is that when e_o and e_m when are os similar magniture, the estimation becomes the most accurate for all four parameters. to isolate this effect, Figure 3.A2 shows the structure of the error when varying both e_o and e_m. The plot shows a minimum in error when e_o and e_m are equal in magnitude, and increasing error when their magnitudes begin to differ. This shows that in order for the approximations in equations 3.5 and 3.6 to hold true, e_m and e_o must be approximately the same magnitude.

These plots were made using 1000 randomly selected values of Truth. We repeated this experiment using regularly spaced sampling time intervals, and with e_m derived using the known e_o, all which showed no major effects on the experiment result. We also explored how the number of observations affect our approximate estimate of the structural bias. Figure 3.A3 shows the effect of sampling density on error estimation. This plot was made using randomly generated values for e_o, e_m, A, and B (within 20% of the canonical value of 0.1). For each plotted point, the experiment was run 100 times, and the average error in estimation was recorded. The figure shows that even with $O(10^6)$ observations, e_o and e_m will have estimation errors that range from from 10% to over 100% of the value itself. Exceptions to this are of course found in Figure 3.A2, when e_o and e_m are nearly equal, as discussed above. Figure 3.A3 additionally shows that structural biases, A, can be found with less than

8% estimation error even with only O(10^3) randomly sampled observations, or more. Finally, background biases, B, can be estimated with less than 10% error when using at least O(10^5) points.

3.6 Acknowledgements

This research was only possible because of the immense effort by Alicia Karspeck and the NCAR team, who generously provided us with the CESM-DART datasets. This research was supported by a grant from NOAA Climate Variability and Prediction Program (NA14OAR4310276) and the NSF Earth System Modeling Program (OCE1419306). High-performance computing support from Yellowstone (ark:/85065/d7wd3xhc) was provided by NCARs Computational and Information Systems Laboratory (CISL), which is sponsored by the National Science Foundation. The data were provided by the Data Support Section of CISL at NCAR. This chapter has been published in: **Eliashiv, J**, Subramanian A, Miller AJ (2019). A Reliability Budget analysis of CESM-DART, *Journal of Advanced in Modeling Earth Systems*, submitted, *sub judice*. The dissertation author was the primary researcher and author of this paper.

Table 3.1: Squared correlation between terms in the reliability budget (from Figure 3.4) and $log(n+1)$ in the variable specified by each row.

	Departure	Bias	Ens. Var.	Obs. Unc.	Residual
r^2_{lat}	0.56	0.30	0.17	0.69	0.45
r^2_{lon}	0.19	0.10	0.36	0.14	0.09
r^2_{height}	0.05	0.01	0.02	0.07	0.05
r^2_{time}	0.03	0.00	0.15	0.05	0.06

Table 3.2: Reliability budget terms for ocean temperatures averaged globally for the indicated range of ocean temperature.

	$\theta < 10K$	$10K \leq \theta < 20K$	$20K \leq \theta$
Departure	1.69	2.18	2.70
Bias	0.08	0.17	0.19
Ens. Var	0.09	1.50	4.00
Obs.Unc	0.11	0.10	0.09
Residual	0.64	0.04	-1.59

Table 3.3: Reliability budget terms for zonal winds and ocean temperatures during MJO and non-MJO conditions averaged over western equatorial Pacific Ocean and Indian Ocean.

		Departure	Bias	Ens.Var.	Unc.	Residual
Pacific (no MJO)	U (m^2/s^2)	17.	.037	.26	2.3	13.
	θ (K^2)	2.0	.013	.003	2.3	-.25
Pacific (MJO)	U (m^2/s^2)	16.	.047	.26	2.2	13.
	θ (K^2)	1.9	.011	.003	2.2	-.29
Indian (no MJO)	U (m^2/s^2)	14.	.023	.25	1.9	12.
	θ (K^2)	1.6	.057	.004	6.0	-4.5
Indian (MJO)	U (m^2/s^2)	14.	.032	.25	1.8	12.
	θ (K^2)	1.4	.017	.004	4.3	-2.9

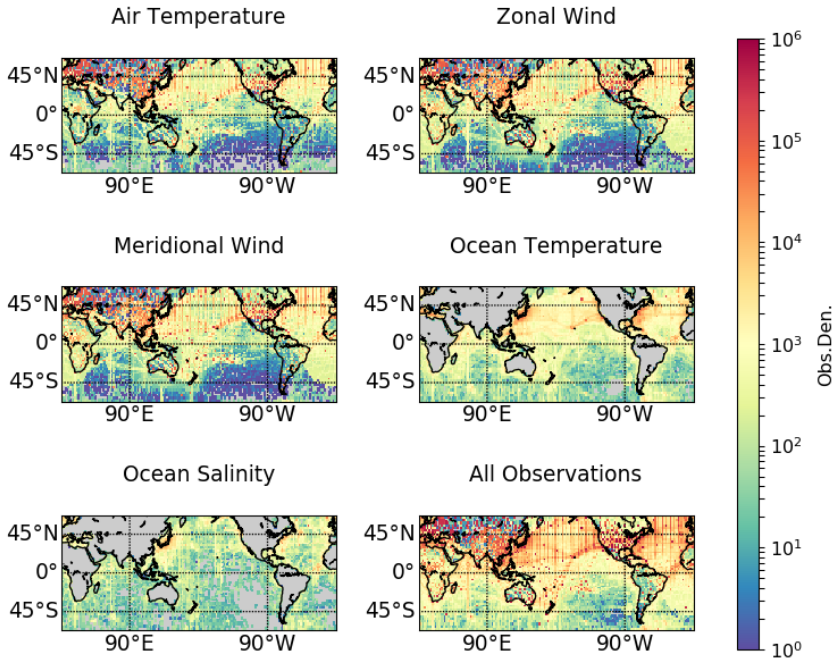

Figure 3.1: Observation Density of of atmospheric (temperature, zonal wind, meridional wind) and oceanic (temperature, salinity) measurements that are assimilated into CESM using DART.

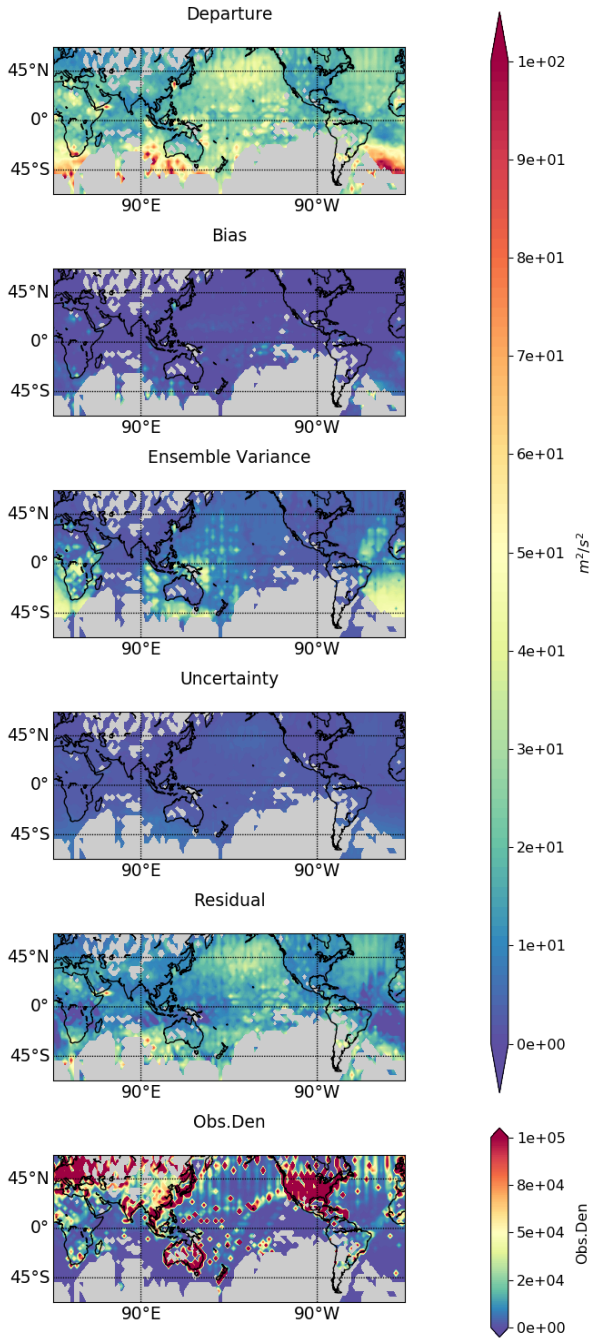

Figure 3.2: Reliability budget terms for atmospheric zonal velocity, vertically averaged. From top to bottom, Departure (total error variance), Bias of the ensemble mean, Ensemble Variance, Observational Uncertainty, Residual. Observation Density is also plotted for reference.

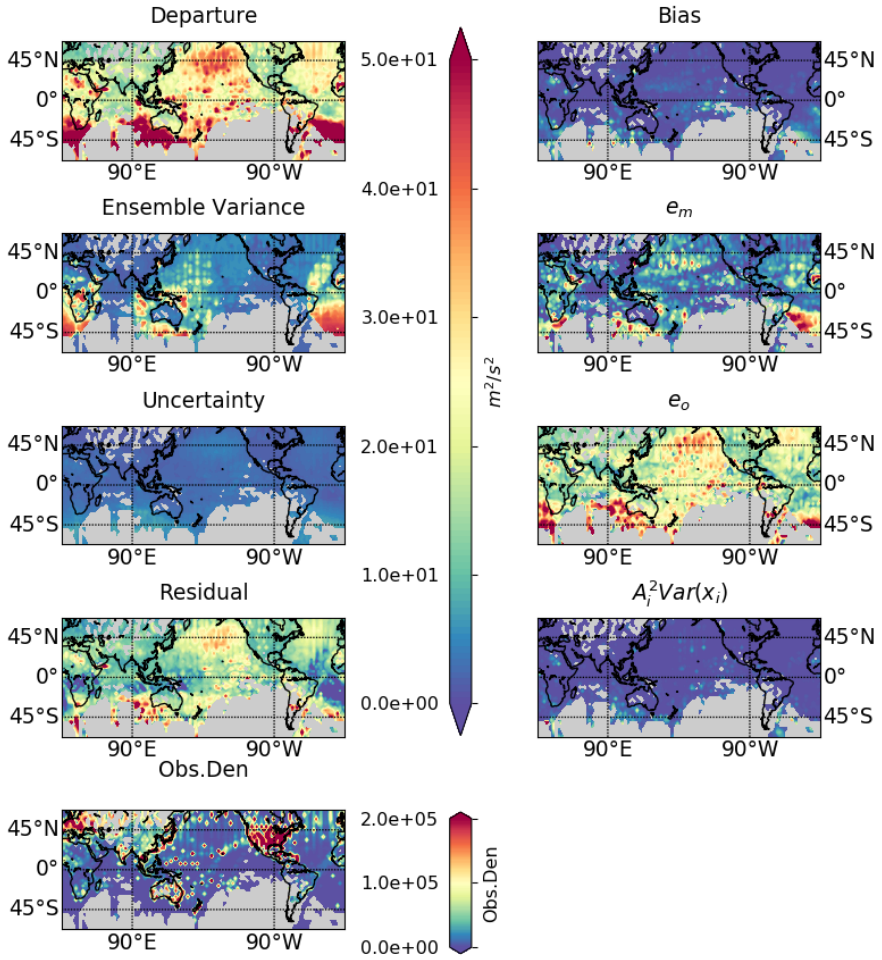

Figure 3.3: As in Figure 3.2, but including Expanded reliability budget terms for Model Error Variance, Observational Error Variance, and linearly estimated Bias.

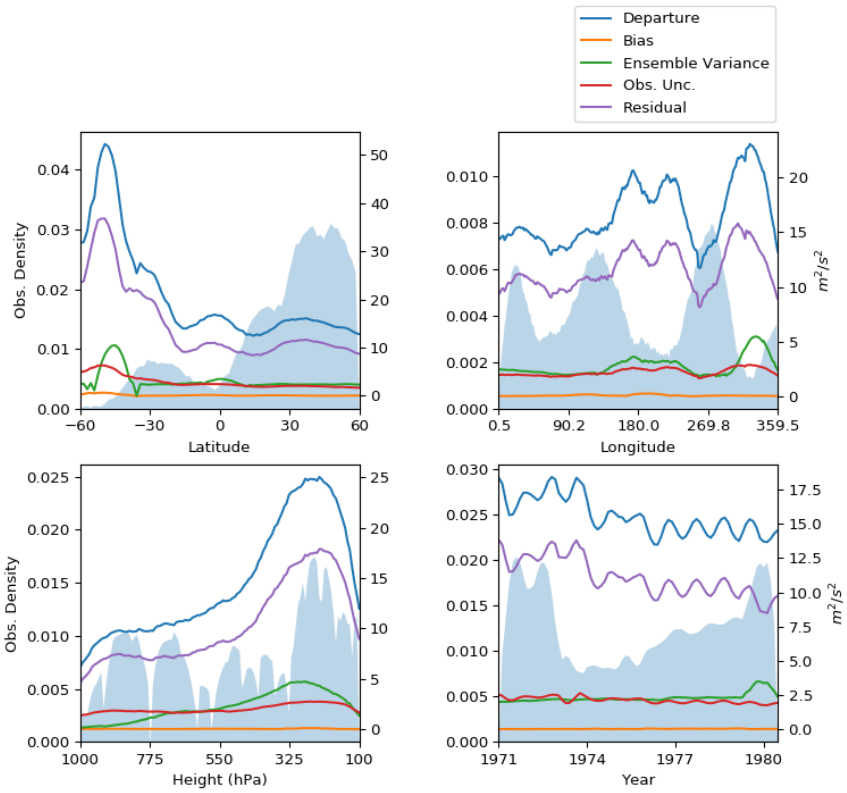

Figure 3.4: Reliability budget for atmospheric zonal velocity, plotted as a function of latitude, longitude, time, and height, and otherwise averaged in the other dimensions. Shaded regions represent the observation density at each grid point.

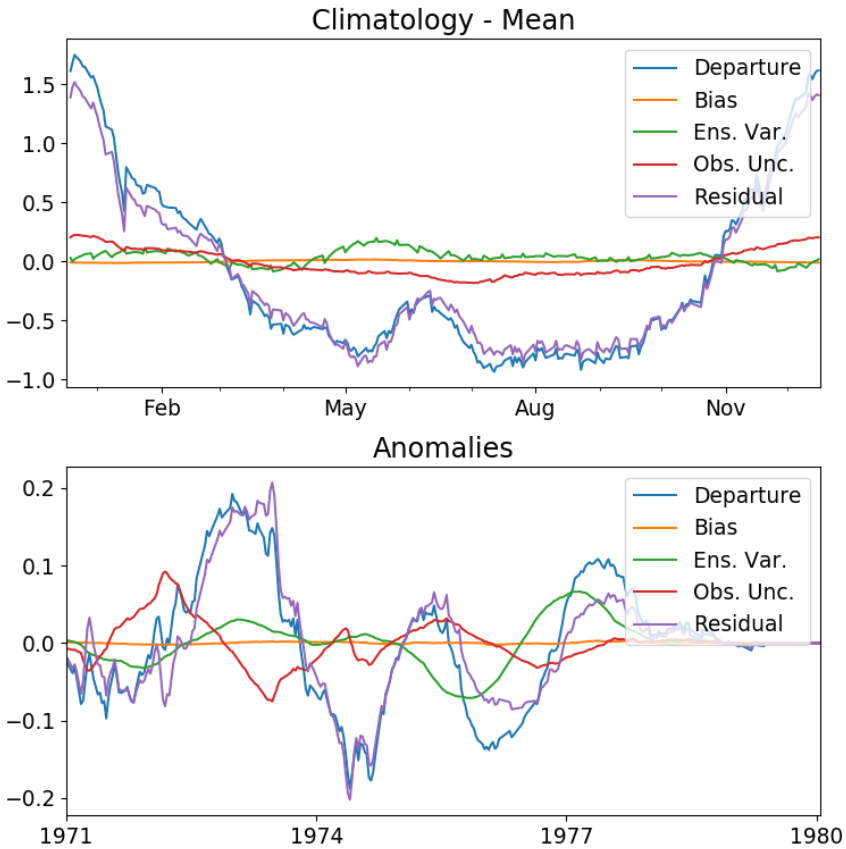

Figure 3.5: Reliability budget for atmospheric zonal velocity plotted as (top) a seasonal cycle and (bottom) anomalies around the seasonal cycle, after subtracting out the annual mean, and otherwise averaged over latitude, longitude, and height.

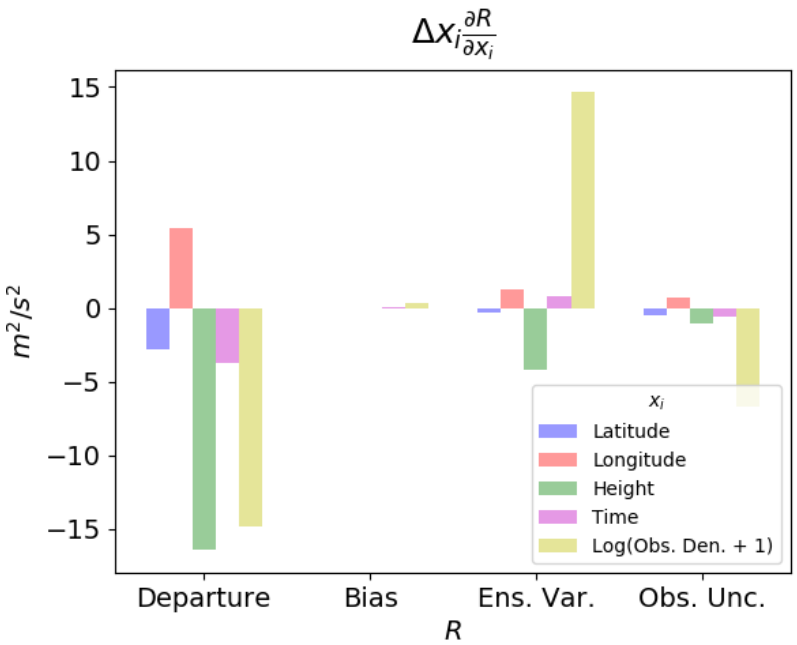

Figure 3.6: Sensitivity of the terms in the reliability budget to latitude, longitude, height, time and observation density.

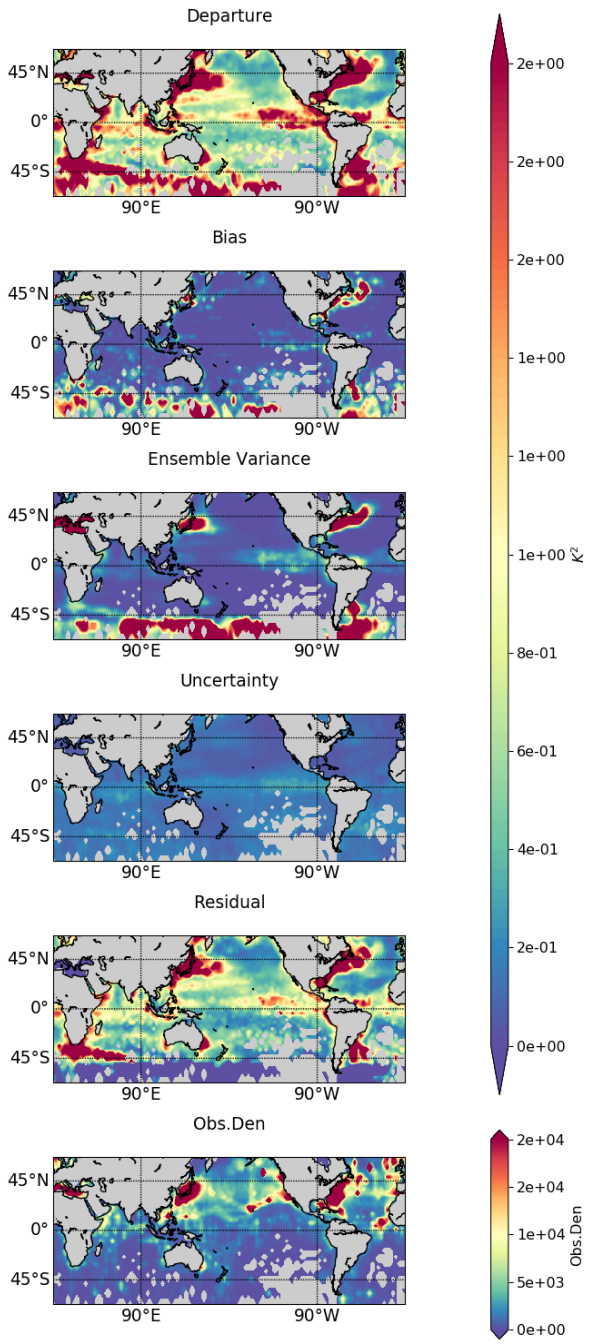

Figure 3.7: As in figure 3.2 but for ocean temperature, vertically averaged

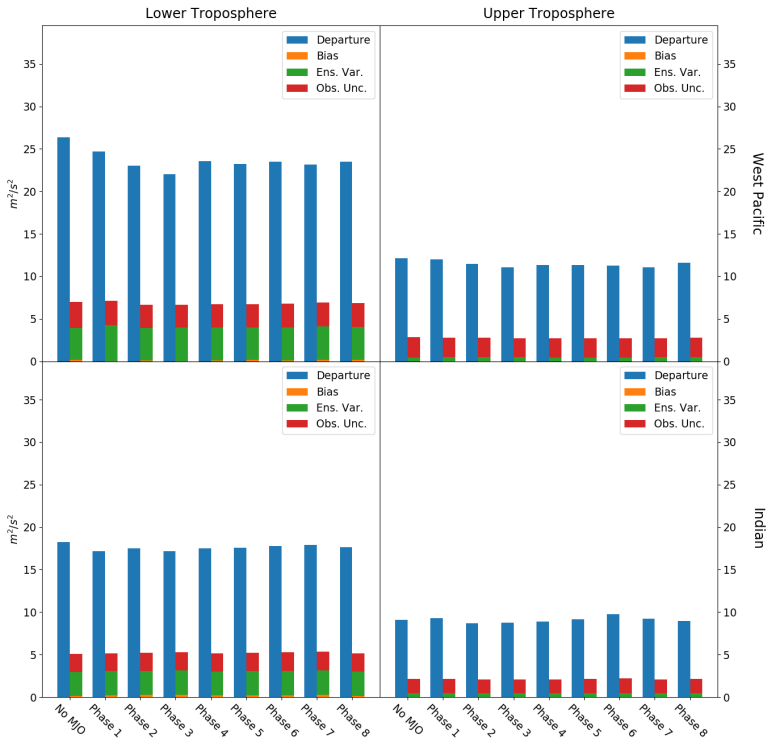

Figure 3.8: Reliability budget terms for each phase of the MJO and for non-MJO conditions, averaged for the Western Pacific and Indian basins for lower tropospheric (850±50 hPa) and upper tropospheric (200±50 hPa) zonal winds.

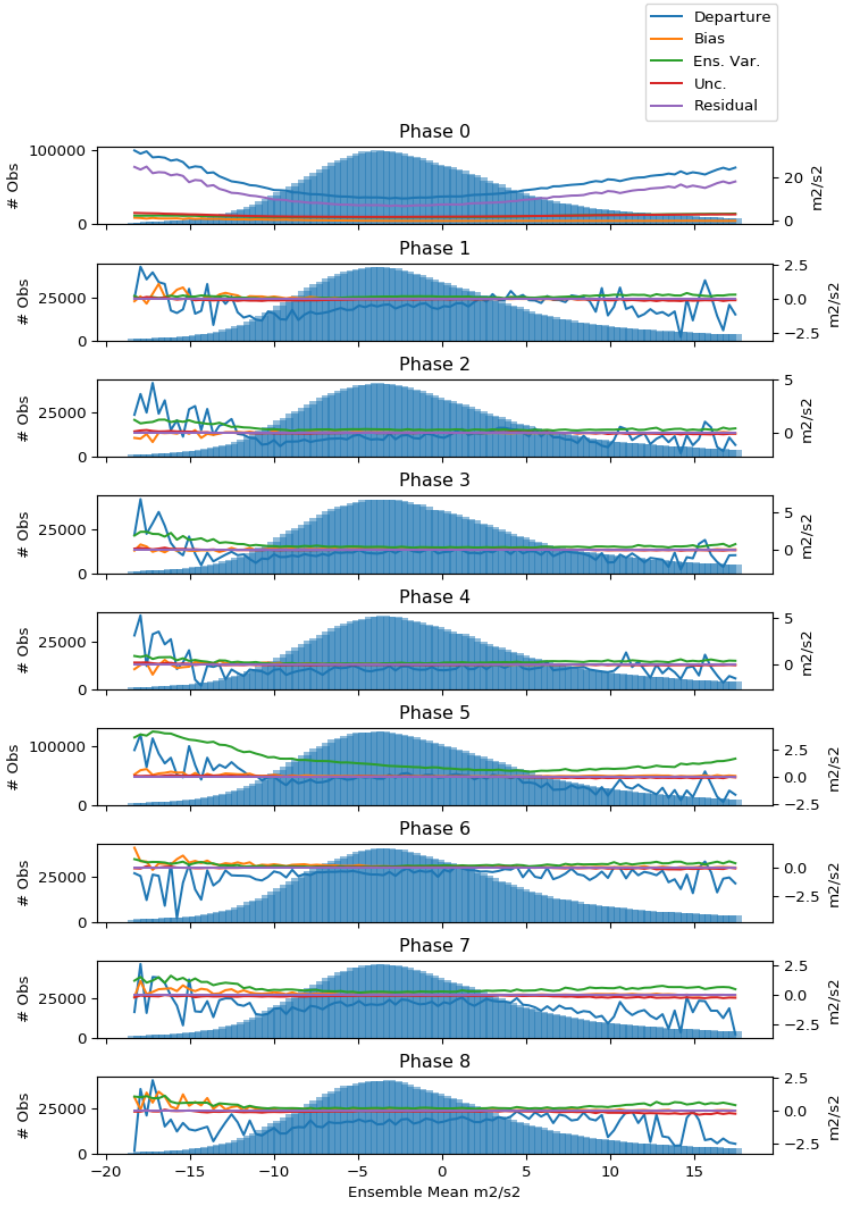

Figure 3.9: Histograms of zonal winds (blue shading), for all available observations in the vertical in the western equatorial Pacific region, for (top panel) non-MJO conditions and (lower panels) each phase of the MJO. Also plotted as colored lines are the corresponding reliability budget terms for the zonal winds that occur in each bin of the histogram, to show the impact of wind strength on the budget. Top panel is the full budget for non-MJO time conditions, and remaining panels are plotted as differences from the top panel for each phase of MJO.

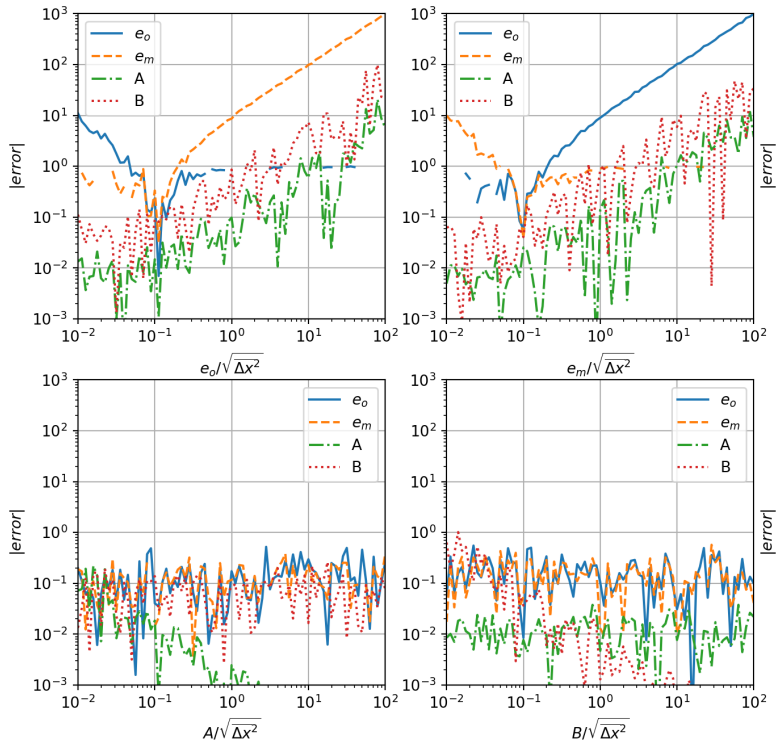

Figure 3.A1: Error in the estimates of e_o, e_m, A, and B as a function of (a) e_o, (b) e_m, (c) linear bias, A and (d) background bias, B. In each subplot, values for the other parameters are set to $\sqrt{\overline{x^2}}/10$. These plots are made using 1000 randomly selected values of Truth.

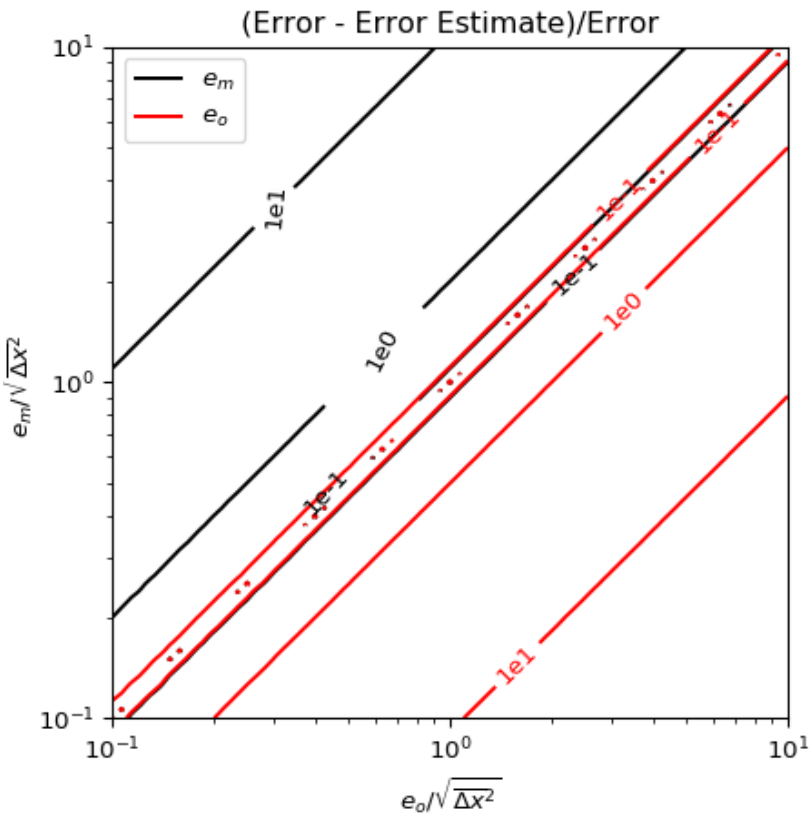

Figure 3.A2: Error in estimating e_o (red) and e_m (black) as a function of both e_o and e_m. Values for the other parameters are set to $\sqrt{\overline{x^2}}/10$.

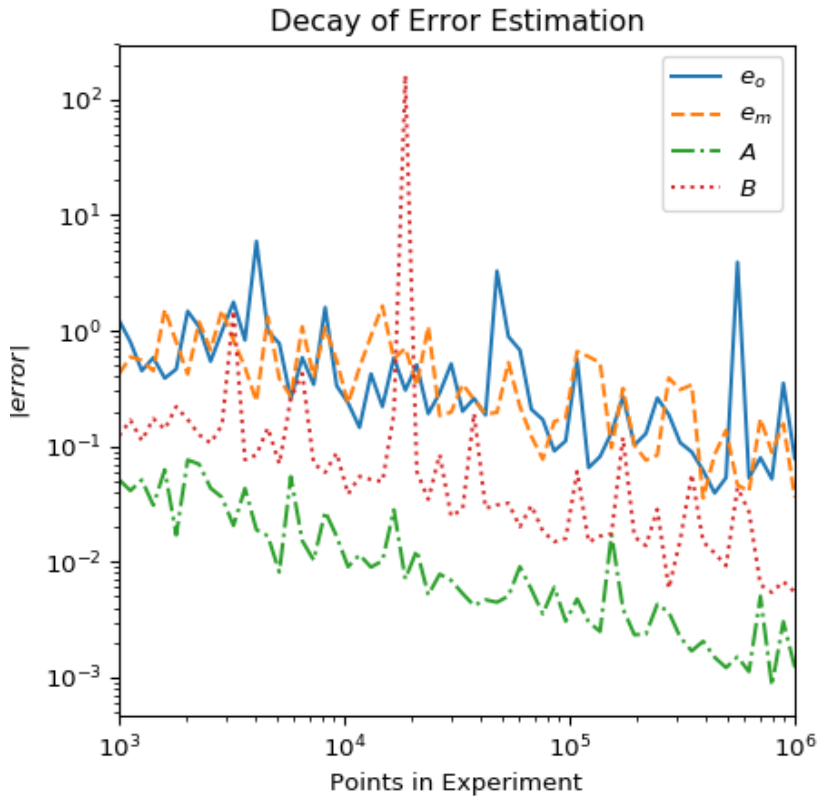

Figure 3.A3: Error in estimating e_o, e_m, A, and B with increasing data points (observations). For each plotted point, simulations are created with random values of the above parameters 100 times, and estimates from each simulation are averaged, as described in the text.

Chapter 4

Conclusions and Comments

4.1 Advantages of DART in CESM

CESM-DART shows a great deal of improvement for CESM, especially in the analysis of MJO and other tropical atmospheric waves. MJO events do not dampen too quickly, as they do in free running CESM analyses. MJO energy is appropriately increased, as compared to free running CESM, to adequately replicate the necessary MJO wave energy. This improvement in convective behavior is possibly due to the flux coupling, where measurements from both ocean and atmosphere datasets are assimilated into CESM. Additionally, high frequency Kelvin waves are reduced in CESM-DART, which supports ocean coupling posing an improvement in GCM reanalysis.

Another positive feature of the CESM-DART reanalysis is the relatively low bias it exhibits in the tropical precipitation field. The typical local bias is comparable to both CESM20C and ERA20C, and is much lower than that of JRA55 and R1. CESM-DART also displays the largest amount of tropical precipitation variability, suggesting an important feedback between the coupled SST field and tropical convection fields.

Additionally CESM-DART produced low amounts of bias in sensible heat flux for

tropical regions, similar to CERA20C and R1. In contrast, latent heat flux was often found to be worse in CESM-DART than other models, however we note that these comparisons rely on treating OAFlux as observations, when OAFlux actually uses both the ERA20C and the R1 reanalysis output in the creation of the product. One would expect that this would render a more favorable comparison of OAFlux with those two models. But each of the two coupled products yield better results in their own way, supporting the idea that coupling can enhance the skill of state estimates.

When analyzing these results, there are important issues to note about the data used and the assimilation frameworks. Surface flux data is not directly assimilated into any of the analysis products, but is estimated in different ways. CESM-DART assimilates available wind and temperature measurements from the ACARS record for the atmosphere, and salinity and temperature data from BUFR records for the ocean, with surface fluxes determined as a consequence of the coupled dynamics. While CERA20C also produces surface fluxes from coupled dynamics, it assimilates only sea surface pressure and marine winds in the atmosphere, plus SST and salinity data in the ocean. ERA20C, on the other hand, assimilates only sea surface pressure and marine winds, while using specified SST to produce the surface fluxes. JRA55 and R1 assimilate available atmospheric data while using specified SST to produce the surface fluxes. The often large disparity in surface heat flux estimates found here, and their impact on the entire climate system, is a consequence of these different approaches and requires further study to implicate mechanisms.

4.2 Isolated sources of error in CESM

We focused on zonal winds as a typical example of our results for other atmospheric variables. In areas where there were sufficient numbers of observations, the mean error Bias was generally very low compared to the spread due to Ensemble Variance, indicating

the important controls exerted by the unstable dynamics in the atmospheric flows. The largest ensemble variance was in the mid-to-high-latitude Southern Hemisphere, where there are energetic flows with limited observations to control them. The Residual terms were often the largest contributors to the budget, which suggests that either the statistics of observational uncertainty was improperly assigned, or the model inadequately represented Ensemble Variance statistics, or both.

The departure exhibits significant seasonal variability, with a strong peak in winter months, and weaker signals in other months. This seasonal variability is not seen in the other terms, apart from the residual, indicating the model's deficiencies during the energetic Northern Hemisphere winter. The year-to-year variability in the budget exhibited no obvious relationship to any climate mode. A sensitivity analysis revealed that, over this time period with limited data, the observation density predominantly controlled the quality of the reliability budget, which explained some of features seen in the latitudinal and longitudinal structures of the budget, since far more data was available over land areas of the Northern Hemisphere.

Ocean temperature fields found a large production of departure in areas with eddy production in western boundary currents, as well as in equatorial regions where there are tropical instability waves. These areas with unstable current features have approximately the same magnitude of residual and ensemble variance. This is consistent with the coarse resolution of the model being incapable of dynamically producing energetic eddies thereby underestimating ensemble variance.

We then studied intraseasonal variability by identifying periods when MJO was active and compared them to non-active time periods to provide insight into the growth of errors associated with atmospheric convectively coupled equatorial waves. We found a slight reduction in Departure for both atmospheric zonal winds and ocean temperature during MJO events. After breaking down the budget by the phase of MJO, we noted that

the Departures exhibit their lowest values in the lower troposphere of the tropical western Pacific for MJO phases 2 to 3, which are the times of least active convection there. This suggests that weaker convective activity in the atmosphere allows a locally more skillful model representation of the flows. We also found that in the tropical western Pacific during the more convectively quiescent phases of MJO the departure is reduced for moderate-to-strong easterly wind conditions. In contrast, during phase 5, when convection is enhanced by MJO in that region, the Ensemble Variance is increased over non-MJO times, but the skill is left unchanged. Taken together, this suggests that the model precipitation physics produces enhanced ensemble spread through feedbacks with the dynamics, thereby limiting the overall skill of the reanalysis in this region for both MJO and non-MJO conditions.

4.3 Future Outlook

To further explore the impacts of Data Assimilation on DART, CESM can be run from restarts of the CESM-DART output to explore the sensitivity of CESM. This can also be used to explore the predictability of a start of MJO, despite being unable to sustain an event for the duration of its run. However, the sensitivity runs can be used to determine where MJO begins to deviate most from nature within CESM.

The observed increase in equatorial atmospheric tropical wave fidelity could increase ENSO forecast accuracy, since they constitute one of the primary mechanisms for instigating ENSO development. Interannual forecasts using states initialized from the coupled product could determine whether the increased response would improve prediction

Radiative fluxes were not studied here, but they are known to vary widely among the reanalysis products, particularly in the tropical Pacific. Future studies should address the potentially important effects of these differences on the estimation of the state vectors.

Bibliography

Anderson J, Hoar T, Raeder K, Liu H, Collins N, Torn R, Avellano A (2009) The data assimilation research testbed: A community facility. Bulletin of the American Meteorological Society 90(9):1283–1296, DOI 10.1175/2009BAMS2618.1

Anderson JL (2003) A local least squares framework for ensemble filtering. Monthly Weather Review 131(4):634–642, DOI 10.1175/1520-0493(2003)131¡0634:ALLSFF¿2.0.CO;2

Arnold NP, Branson M, Kuang Z, Randall DA, Tziperman E (2015) Mjo intensification with warming in the superparameterized cesm. Journal of Climate 28(7):2706–2724

Balmaseda MA, Trenberth KE, Källén E (2013) Distinctive climate signals in reanalysis of global ocean heat content. Geophysical Research Letters 40(9):1754–1759

Bretherton C, Balaji V, Delworth T, Dickinson R, Edmonds J, Famiglietti J, Smarr L (2012) A national strategy for advancing climate modeling

Carton JA, Giese BS (2008) A reanalysis of ocean climate using simple ocean data assimilation (soda). Monthly Weather Review 136(8):2999–3017

Chelliah M, Bell GD (2004) Tropical multidecadal and interannual climate variability in the ncep–ncar reanalysis. Journal of Climate 17(9):1777–1803

Cushman-Roisin B, Beckers JM (2011) Introduction to geophysical fluid dynamics: physical and numerical aspects, vol 101. Academic press

Cuxart J, Bougeault P, Redelsperger JL (2000) A turbulence scheme allowing for mesoscale and large-eddy simulations. Quarterly Journal of the Royal Meteorological Society 126(562):1–30

Dee DP, Uppala S, Simmons A, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda M, Balsamo G, Bauer dP, et al (2011) The era-interim reanalysis: Configuration and performance of the data assimilation system. Quarterly Journal of the royal meteorological society 137(656):553–597

Edwards PN (2001) Representing the global atmosphere: Computer models, data, and knowledge about climate change. Changing the atmosphere: Expert knowledge and environmental governance 31:33

Eliashiv J, Subramanian A, Miller AJ (2019, submitted) Tropical climate variability in the community earth system model - data assimilation research testbed. Climate Dyn

Frolov S, Bishop CH, Holt T, Cummings J, Kuhl D (2016) Facilitating strongly coupled ocean–atmosphere data assimilation with an interface solver. Monthly Weather Review 144(1):3–20

Gelaro R, McCarty W, Suárez MJ, Todling R, Molod A, Takacs L, Randles CA, Darmenov A, Bosilovich MG, Reichle R, et al (2017) The modern-era retrospective analysis for research and applications, version 2 (merra-2). Journal of Climate 30(14):5419–5454

Gent PR, Yeager SG, Neale RB, Levis S, Bailey DA (2010) Improvements in a half degree atmosphere/land version of the ccsm. Climate Dynamics 34(6):819–833

Gent PR, Danabasoglu G, Donner LJ, Holland MM, Hunke EC, Jayne SR, Lawrence DM, Neale RB, Rasch PJ, Vertenstein M, et al (2011) The community climate system model version 4. Journal of Climate 24(19):4973–4991

Gill AE (1982) Atmosphere-Ocean dynamics (International Geophysics Series), academic press, p 662

Gould WJ, Turton J (2006) Argo–sounding the oceans. Weather 61(1):17–21

Hendon HH, Salby ML (1994) The life cycle of the madden julian oscillation. Journal of the Atmospheric Sciences 51(15):2225–2237

Hodges KI, Lee RW, Bengtsson L (2011) A comparison of extratropical cyclones in recent reanalyses era-interim, nasa merra, ncep cfsr, and jra-25. Journal of Climate 24(18):4888–4906

Houtekamer PL, Mitchell HL (1998) Data assimilation using an ensemble kalman filter technique. Monthly Weather Review 126(3):796–811

Hurrell JW, Holland MM, Gent PR, Ghan S, Kay JE, Kushner PJ, Lamarque JF, Large WG, Lawrence D, Lindsay K, et al (2013) The community earth system model: a framework for collaborative research. Bulletin of the American Meteorological Society 94(9):1339–1360

Hwang YT, Frierson DM (2013) Link between the double-intertropical convergence zone problem and cloud biases over the southern ocean. Proceedings of the National Academy of Sciences 110(13):4935–4940

Jochum M (2009) Impact of latitudinal variations in vertical diffusivity on climate simulations. Journal of Geophysical Research: Oceans 114(C1)

Jones PW (1999) First-and second-order conservative remapping schemes for grids in spherical coordinates. Monthly Weather Review 127(9):2204–2210

Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woollen J, et al (1996) The ncep/ncar 40-year reanalysis project. Bulletin of the American meteorological Society 77(3):437–472

Kanamitsu M, Ebisuzaki W, Woollen J, Yang SK, Hnilo J, Fiorino M, Potter G (2002) Ncep–doe amip-ii reanalysis (r-2). Bulletin of the American Meteorological Society 83(11):1631–1644

Karspeck AR, Danabasoglu G, Anderson J, Karol S, Collins N, Vertenstein M, Raeder K, Hoar T, Neale R, Edwards J, et al (2018) A global coupled ensemble data assimilation system using the community earth system model and the data assimilation research testbed. Quarterly Journal of the Royal Meteorological Society 144(717):2404–2430

Kim D, Sperber K, Stern W, Waliser D, Kang IS, Maloney E, Wang W, Weickmann K, Benedict J, Khairoutdinov M, et al (2009) Application of mjo simulation diagnostics to climate models. Journal of Climate 22(23):6413–6436

Kim D, Lee MI, Kim D, Schubert SD, Waliser DE, Tian B (2014) Representation of tropical subseasonal variability of precipitation in global reanalyses. Climate dynamics 43(1-2):517–534

Laloyaux P, De Boisseson E, Dahlgren P (2017) Cera-20c: An earth system approach to climate reanalysis. ECMWF newsletter 150:25–30

Lau WK, Waliser DE, Majda AJ, Stechmann SN (2012) Multiscale theories for the mjo. In: Intraseasonal variability in the atmosphere-ocean climate system, Springer, pp 549–568

Leutbecher M, Lock SJ, Ollinaho P, Lang ST, Balsamo G, Bechtold P, Bonavita M, Christensen HM, Diamantakis M, Dutra E, et al (2017) Stochastic representations of model uncertainties at ecmwf: State of the art and future vision. Quarterly Journal of the Royal Meteorological Society 143(707):2315–2339

Li G, Xie SP (2014) Tropical biases in cmip5 multimodel ensemble: The excessive equatorial pacific cold tongue and double itcz problems. Journal of Climate 27(4):1765–1780

Lin JL (2007) The double-itcz problem in ipcc ar4 coupled gcms: Ocean–atmosphere feedback analysis. Journal of Climate 20(18):4497–4525

Lorenz EN (1963) The mechanics of vacillation. Journal of the atmospheric sciences 20(5):448–465

Neale RB, Richter JH, Jochum M (2008) The impact of convection on enso: From a delayed oscillator to a series of events. Journal of climate 21(22):5904–5924

Neale RB, Richter J, Park S, Lauritzen PH, Vavrus SJ, Rasch PJ, Zhang M (2013) The mean climate of the community atmosphere model (cam4) in forced sst and fully coupled experiments. Journal of Climate 26(14):5150–5168

Onogi K, Tsutsui J, Koide H, Sakamoto M, Kobayashi S, Hatsushika H, Matsumoto T, Yamazaki N, Kamahori H, Takahashi K, et al (2007) The jra-25 reanalysis. Journal of the Meteorological Society of Japan Ser II 85(3):369–432

Penny SG, Hamill TM (2017) Coupled data assimilation for integrated earth system analysis and prediction. Bulletin of the American Meteorological Society 97(7):ES169–ES172

Phillips TJ, Potter GL, Williamson DL, Cederwall RT, Boyle JS, Fiorino M, Hnilo JJ, Olson JG, Xie S, Yio JJ (2004) Evaluating parameterizations in general circulation models: Climate simulation meets weather prediction. Bulletin of the American Meteorological Society 85(12):1903–1916

Raeder K, Anderson JL, Collins N, Hoar TJ, Kay JE, Lauritzen PH, Pincus R (2012) Dart/cam: An ensemble data assimilation system for cesm atmospheric models. Journal of Climate 25(18):6304–6317

Randall DA, Xu KM, Somerville RJ, Iacobellis S (1996) Single-column models and cloud ensemble models as links between observations and climate models. Journal of Climate 9(8):1683–1697

Rodwell M, Lang S, Ingleby N, Bormann N, Hólm E, Rabier F, Richardson D, Yamaguchi M (2016) Reliability in ensemble data assimilation. Quarterly Journal of the Royal Meteorological Society 142(694):443–454

Rodwell MJ, Richardson DS, Parsons DB, Wernli H (2018) Flow-dependent reliability: A path to more skillful ensemble forecasts. Bulletin of the American Meteorological Society 99(5):1015–1026

Saha S, Moorthi S, Pan HL, Wu X, Wang J, Nadiga S, Tripp P, Kistler R, Woollen J, Behringer D, et al (2010) The ncep climate forecast system reanalysis. Bulletin of the American Meteorological Society 91(8):1015–1058

Sengupta TK (2004) Fundamentals of computational fluid dynamics, vol 364. Universities Press Hyderabad (India)

Shen BW, Atlas R, Reale O, Lin SJ, Chern JD, Chang J, Henze C, Li JL (2006) Hurricane forecasts with a global mesoscale-resolving model: Preliminary results with hurricane katrina (2005). Geophysical Research Letters 33(13)

Stocker T (2014) Climate change 2013: the physical science basis: Working Group I contribution to the Fifth assessment report of the Intergovernmental Panel on Climate Change. Cambridge University Press

Subramanian AC, Zhang GJ (2014) Diagnosing mjo hindcast biases in ncar cam3 using nudging during the dynamo field campaign. Journal of Geophysical Research: Atmospheres 119(12):7231–7253

Subramanian AC, Jochum M, Miller AJ, Murtugudde R, Neale RB, Waliser DE (2011) The madden–julian oscillation in ccsm4. Journal of Climate 24(24):6261–6282

Uppala SM, Kållberg P, Simmons A, Andrae U, Bechtold VDC, Fiorino M, Gibson J, Haseler J, Hernandez A, Kelly G, et al (2005) The era-40 re-analysis. Quarterly Journal of the royal meteorological society 131(612):2961–3012

Van Ulden AP, Van Oldenborgh GJ (2006) Large-scale atmospheric circulation biases and changes in global climate model simulations and their importance for climate change in central europe. Atmospheric Chemistry and Physics 6(4):863–881

Ventrice MJ, Wheeler MC, Hendon HH, Schreck III CJ, Thorncroft CD, Kiladis GN (2013) A modified multivariate madden–julian oscillation index using velocity potential. Monthly Weather Review 141(12):4197–4210

Waliser D, Sperber K, Hendon H, Kim D, Maloney E, Wheeler M, Weickmann K, Zhang C, Donner L, Gottschalck J, et al (2009) Mjo simulation diagnostics. Journal of Climate 22(11):3006–3030

Wang J, Chen J, Du J, Zhang Y, Xia Y, Deng G (2018) Sensitivity of ensemble forecast verification to model bias. Monthly Weather Review 146(3):781–796

Watson PA, Berner J, Corti S, Davini P, von Hardenberg J, Sanchez C, Weisheimer A, Palmer TN (2017) The impact of stochastic physics on tropical rainfall variability in global climate models on daily to weekly time scales. Journal of Geophysical Research: Atmospheres 122(11):5738–5762

Wheeler M, Kiladis GN (1999) Convectively coupled equatorial waves: Analysis of clouds and temperature in the wavenumber–frequency domain. Journal of the Atmospheric Sciences 56(3):374–399

Wheeler MC, Hendon HH (2004) An all-season real-time multivariate mjo index: Development of an index for monitoring and prediction. Monthly Weather Review 132(8):1917–1932

Williamson D, Blaker AT, Hampton C, Salter J (2015) Identifying and removing structural biases in climate models with history matching. Climate dynamics 45(5-6):1299–1324

Yamaguchi M, Lang ST, Leutbecher M, Rodwell MJ, Radnoti G, Bormann N (2016) Observation-based evaluation of ensemble reliability. Quarterly Journal of the Royal Meteorological Society 142(694):506–514

Yu L, Jin X, Weller R (2008) Multidecade global flux datasets from the objectively analyzed air-sea fluxes (oaflux) project: Latent and sensible heat fluxes, ocean evaporation, and related surface meteorological variables. oaflux project technical report. oa-2008-01, 64pp. Tech. rep., OAFlux

Zhang GJ, McFarlane NA (1995) Role of convective scale momentum transport in climate simulation. Journal of Geophysical Research: Atmospheres 100(D1):1417–1426