



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

Jonathan Eliashiv¹ · Aneesh C. Subramanian^{1,2,3} · Arthur J. Miller¹

Received: 22 February 2019 / Accepted: 26 October 2019 / Published online: 3 December 2019
© The Author(s) 2019

Abstract

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 differences 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 difference. 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.

Keywords CESM · Data Assimilation · Tropical Variability · Madden–Julian Oscillation · DART · Ensemble Kalman Filter

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 does not assure the improvement of the surface fluxes, which are typically diagnostically assessed. A coupled data assimilation framework (Penny and Hamill 2017)

✉ Aneesh C. Subramanian
aneeshcs@colorado.edu

Jonathan Eliashiv
jeliashi@ucsd.edu

¹ Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA, USA

² Department of Atmospheric and Oceanic Sciences, University of Colorado Boulder, Boulder, CO, USA

³ Department of Physics, University of Oxford, Oxford, UK

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-vector of the entire coupled system is 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 (Table 1), 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, which could be associated with the propagation of the coupled error covariance or other aspects of 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 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 (NCEP-DOE) 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

Table 1 Resolution and assimilation techniques of each reanalysis product

	Horizontal resolution	DA method
CESM-DART	1°	30 member 6 h EnKF
ERA 20C	125 km atm	24 h 4DVar
NCAR R1	T62	3DVar
CERA 20C	125 km atm, 110 km ocn	10 member ensemble with 24 h 4DVar in atmosphere and 3DVar in ocean
JRA 55	TL319	6 h 4DVar

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), NASA's Modern-Era Retrospective Analysis for Research and Applications (MERRA) (Rienecker et al. 2011) 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 h for atmospheric fields and 24 h for oceanic components. The components are coupled together every 24 h. 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 Adjustment Kalman Filters allow a model to prescribe the error using ensemble variance which can then be compared to observational uncertainty. This will allow the data assimilation to correct the mean model state as well as the ensemble variance. This not only improves the accuracy of the model state, but additionally adjusts the precision of it by adjusting the ensemble variance towards 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-h 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 $\approx 2.5^\circ$, ERA-20C has a horizontal resolution of ≈ 125 km, JRA55 has ≈ 55 km resolution, CERA20C has ≈ 125 km resolution, and CESM-DART has a horizontal resolution of $\approx 1^\circ$. The R1 uses 3Dvar, while the ERA20C, JRA5, and CERA20C all use incremental 4DVar, in contrast to the EAKF of CESM-DART. These different characteristics are summarized in Table 1.

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.

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 five reanalyses datasets (Table 1). We summarize the details of our analysis procedures in the following subsections.

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 OAFflux 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. OAFflux uses SST, humidity, wind fields at 2 m, and air temperature at 2 m 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 OAFflux product (Yu et al. 2008), which is treated as truth in this study. The OAFflux product is a combination of data-assimilated reanalysis products for the time period between 1958–2018, which includes the common period of 1970–1979 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 OAFflux 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 OAFflux. For all products, the latitude weighted spatial correlation and spatial bias against GPCP and OAFflux are calculated.

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 200 hPa, which serves as smoother measure of

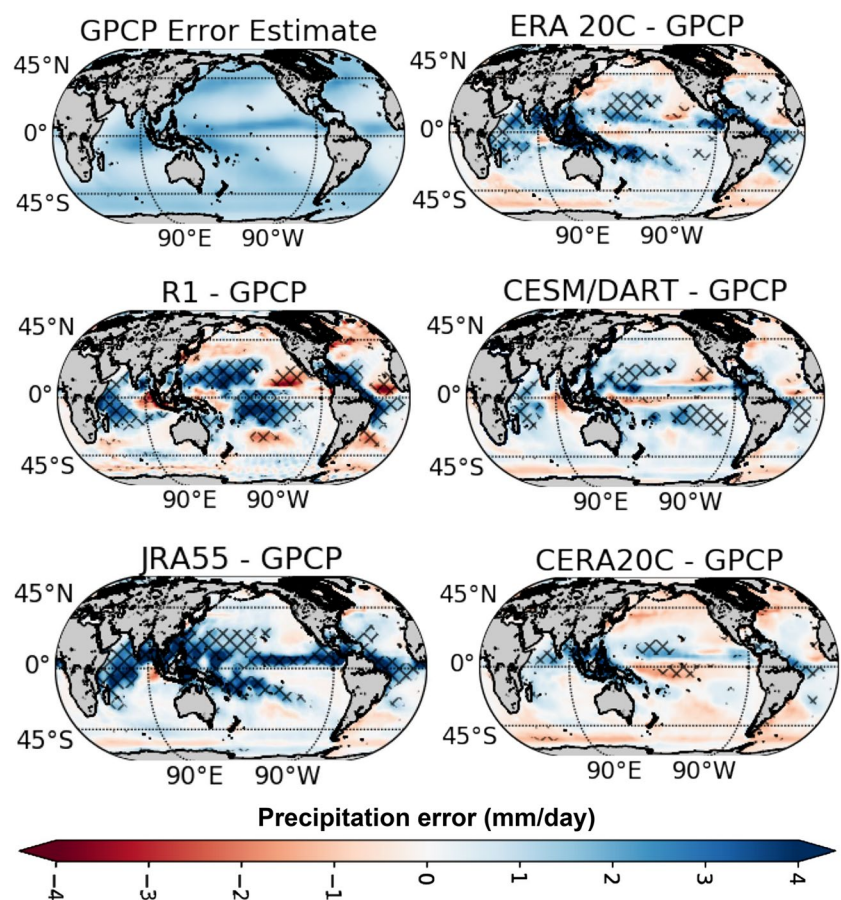
convective activity, is calculated for each reanalysis. Anomalies of the velocity potential and zonal wind anomalies at pressure levels 200 hPa and 850 hPa are then band-pass filtered between 20 and 100 days, and symmetrically averaged between 15° N and 15° S. The combined EOF is then calculated for 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 following the procedure advocated by Wheeler and Kiladis (1999) to isolate the spectral peaks that arise relative to a red “background spectrum”. Zonal wind fields between 15° N and 15° S are taken from each reanalysis product at 850 hPa and 200 hPa levels. All fields are first detrended in time and then high-pass filtered with a 1-year cutoff. The Fast Fourier Transform (FFT) in time and longitude is computed for every latitude of the resulting fields to produce the raw spectrum. The “background spectrum” is then computed from this raw spectrum by heavily smoothing the spectral coefficients with a 1-2-1 filter applied roughly ten times in both frequency and wavenumber for each reanalysis product. The raw spectrum is then decomposed into its equatorially symmetric and asymmetric components and the spectral coefficients are lightly smoothed by applying a 1-2-1 filter only once and only in wavenumber. The resulting symmetric component spectrum is then divided by the corresponding background spectrum for each reanalysis field. The resulting ratio spectrum for each data assimilated product can then be compared with each other to search for differences in amplitude and propagation rates for various equatorial waves relative to the individual product’s red background.

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.

Fig. 1 Annual mean (1970–1979) 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



4 Results

4.1 Precipitation

Figure 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 1 also reveals that JRA55 and R1 have the largest areas of

significant bias. Table 2 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 Fig. 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.

4.2 Surface heat fluxes

Figure 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 structures of the differences vary widely among the models, with the largest errors associated with JRA55. Among all the products, there is an extension of positive bias within the equatorial cold tongue, and increased biases within the 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 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

Fig. 2 Monthly-mean standard deviation of precipitation, calculated relative to the annual mean climatology (1970–1979), 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

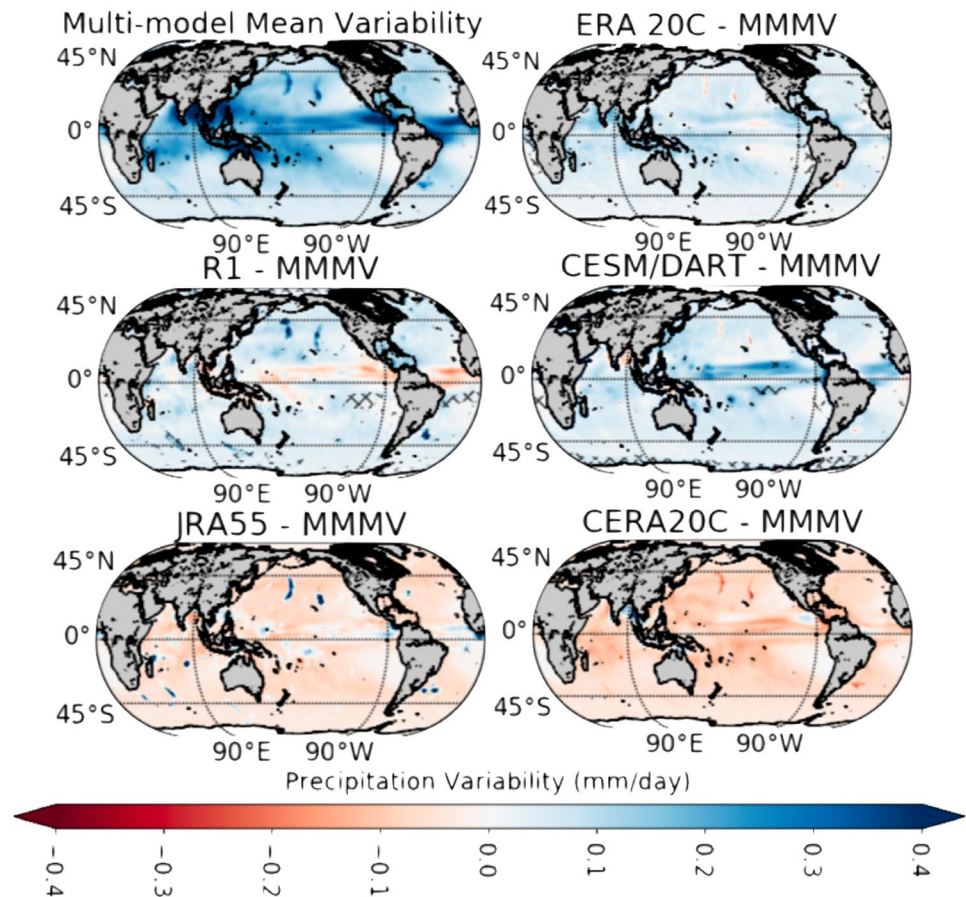


Table 2 shows an overall very high pattern correlation in performance for mean latent heat flux across all models, but with much lower pattern correlations for mean sensible heat flux, especially for CESM-DART. Table 3 additionally quantifies some of the differences seen in Fig. 3 by showing the spatial 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 4 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.

4.3 Equatorial waves

The frequency–wavenumber content of the atmospheric equatorial waves is displayed in Fig. 5, which was produced using the symmetric lower tropospheric zonal wind response normalized against the red “background spectrum” for all the different reanalyses, which we term the “ratio spectra”. The top plot displays the five-product-average ratio spectrum, with the lower plots showing the spectral deviations from this average. (The physical meaning of these deviations is such that a value of 0.05 would mean the spectral peak of that product rises 5% higher in power above its background red spectrum than does the average of the products.) The ratio 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 0.15–0.3 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

Fig. 3 Annual mean (1970–1979) latent (left) and sensible (right) surface heat flux differences for the five reanalyses relative to OAFflux observations (1970–1979). Differences are scaled by a normalization factor of 60 W/m² for latent heat and 20 W/m² for sensible heat

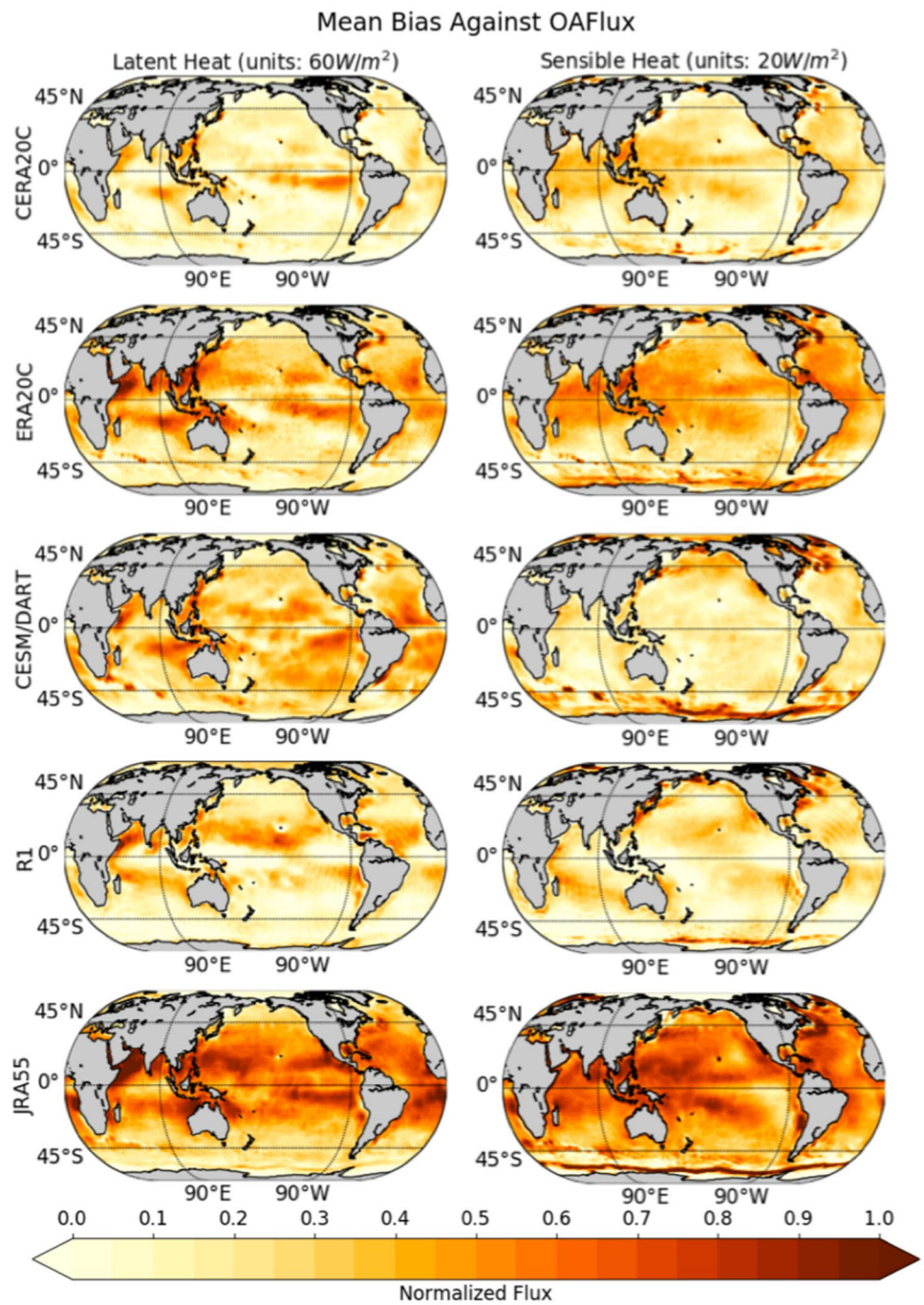
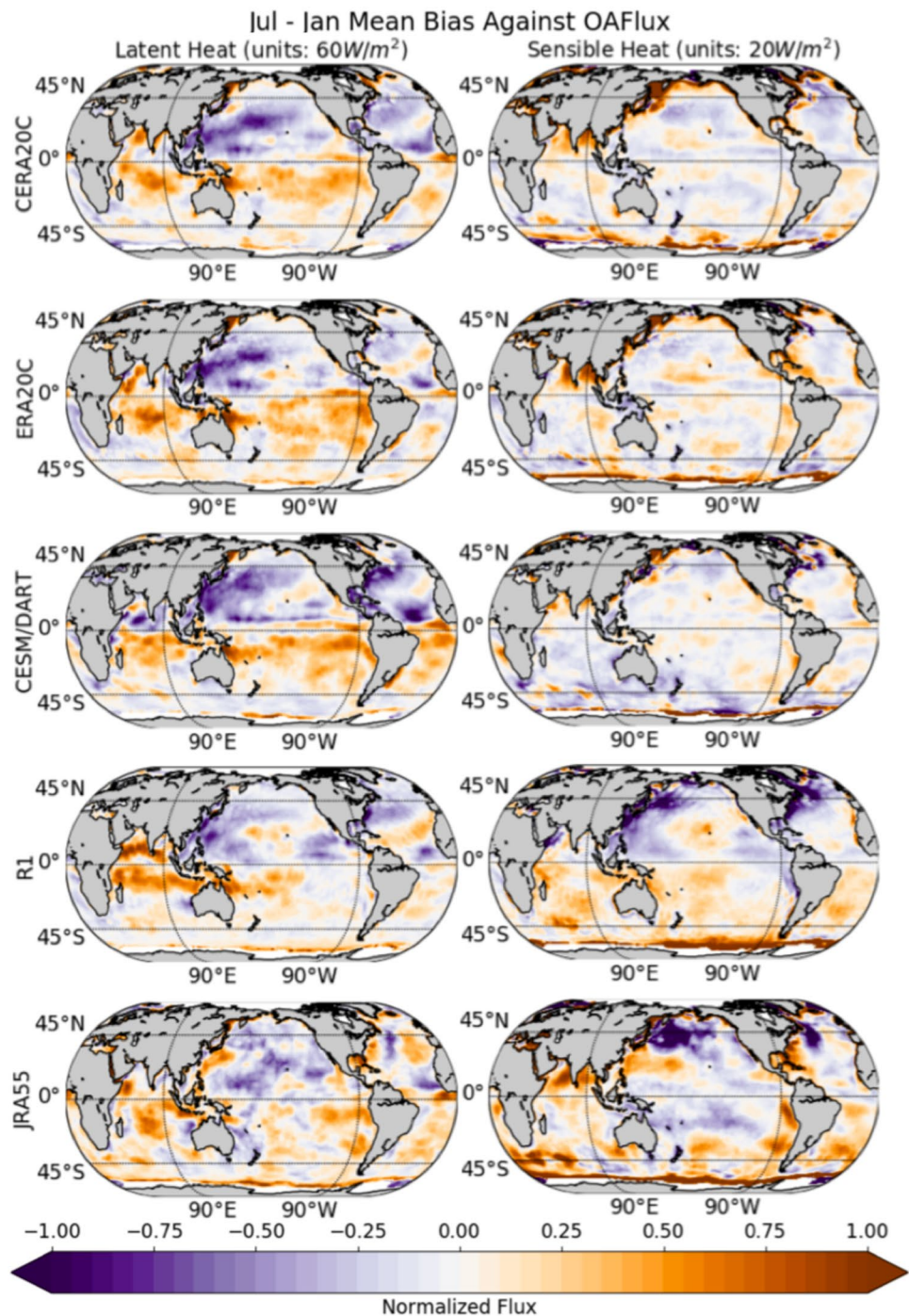


Table 3 Spatially averaged mean and rms differences between reanalysis products and observations (OAFflux) of annual mean latent heat flux and sensible heat flux for Pacific Ocean and Indian Ocean

	Basin	NCAR R1	CESM-DART	ERA20C	CERA 20C	JRA55
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

Fig. 4 Differences between July and January reanalysis biases relative to OAFflux, calculated as in Fig. 3



ratio 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, CEM-DART and JRA55 producing slower waves). CERA20C and ERA20C 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 ratio spectrum,

while R1, CEM-DART and JRA55 all have much weaker IG energy.

Figure 6 shows the effect of using DART on CEM for upper and lower tropospheric symmetrical zonal wind responses. These ratio spectra clearly reveal that data assimilation in CEM 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,

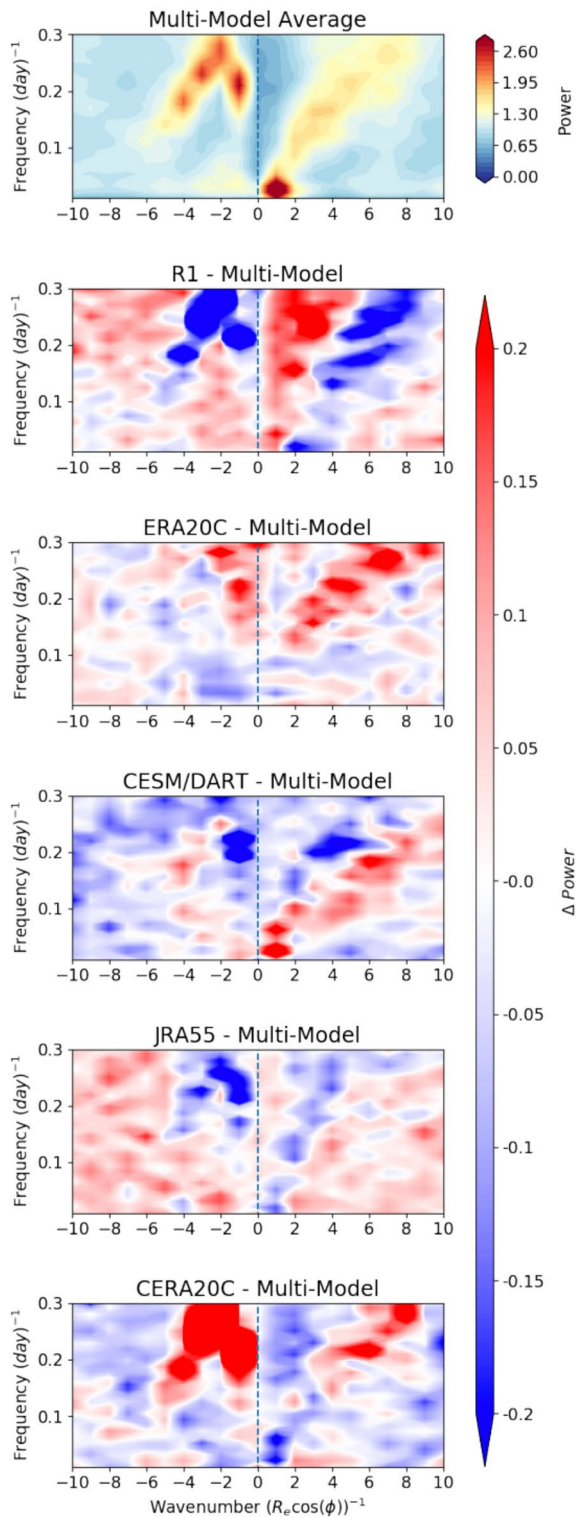


Fig. 5 Frequency–wavenumber spectra of equatorially symmetric lower tropospheric (850 hPa) zonal winds, computed as a ratio to the corresponding red “background spectrum”, and plotted as the five-product mean ratio spectrum and as differences from that mean ratio spectrum for each reanalysis (1970–1979). The power is the spectral power of the ratio of two spectra and hence are unitless

which differ highly between the reanalysis products in that band, as seen in Fig. 5.

4.4 MJO

The structure of the leading combined EOFs found in the formulation of the VPM index is shown in Fig. 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 Fig. 8. MJO events tend to all occur at the same times across all reanalyses. There exist

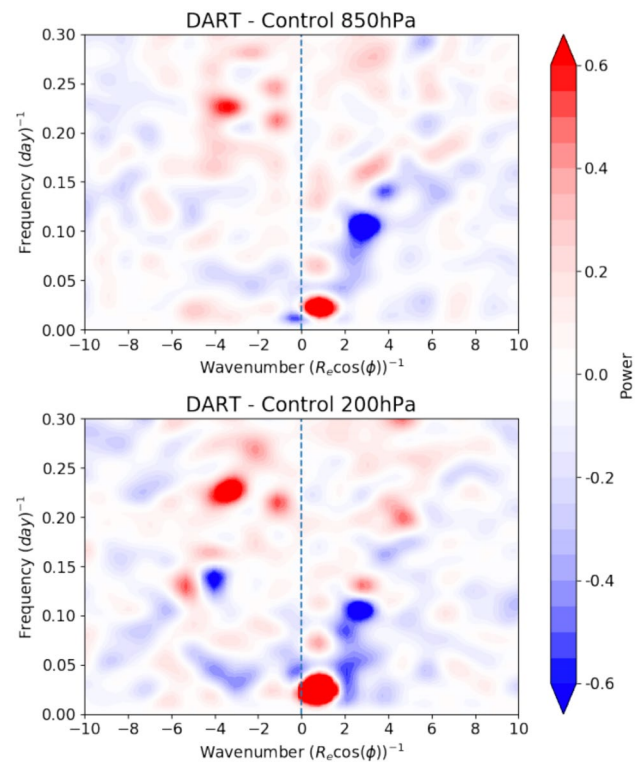


Fig. 6 Frequency–wavenumber spectra of equatorially symmetric lower tropospheric (850 hPa) zonal winds and upper tropospheric (200 hPa) zonal winds, computed as a ratio to the corresponding red “background spectrum”, and plotted as the difference between the ratio spectrum for CESM-DART and the ratio spectrum for Free-CESM. The power represents the spectral power of the ratio of two spectra and hence are unitless

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 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

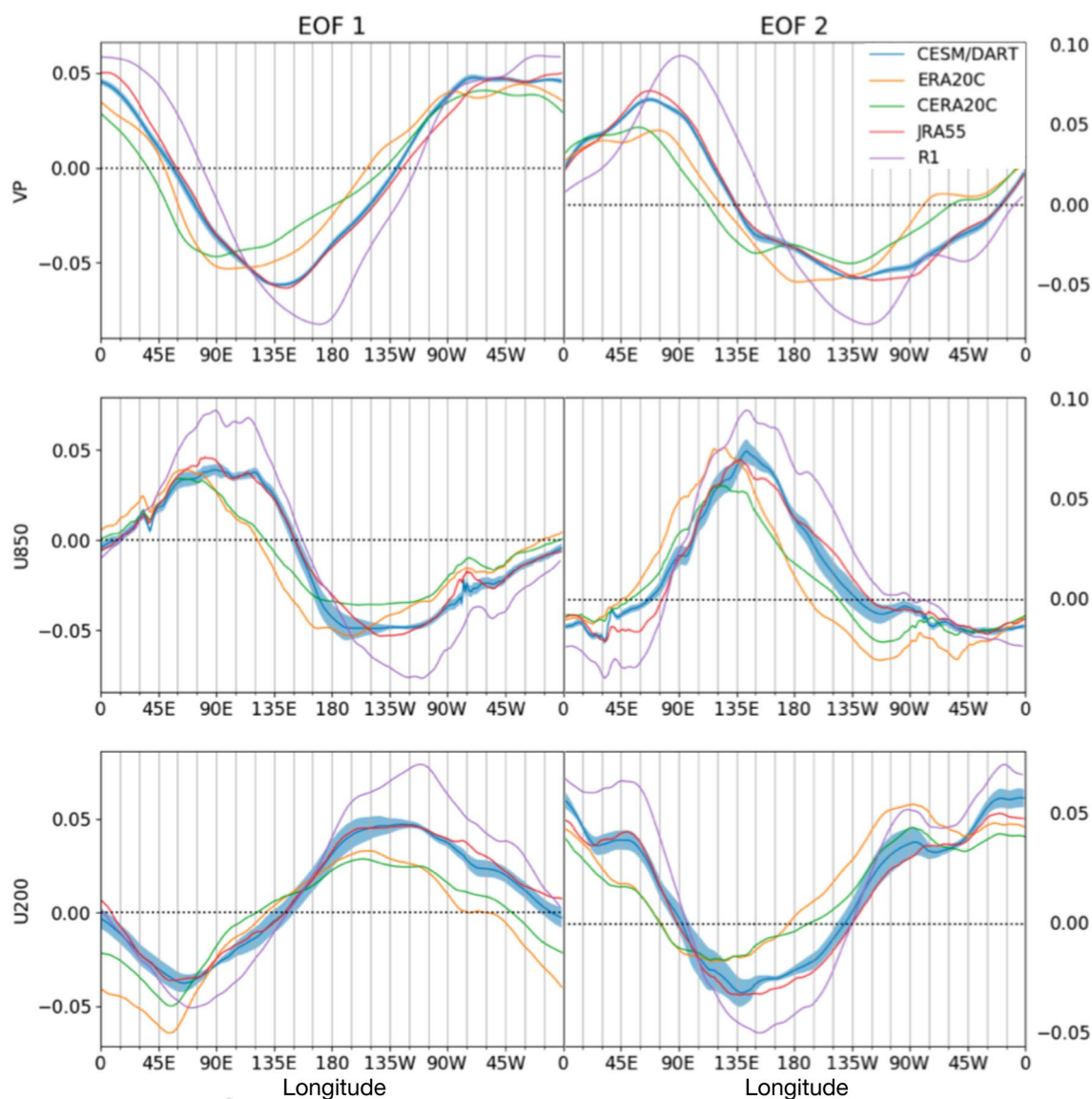


Fig. 7 Combined EOF of lower tropospheric zonal winds, upper tropospheric zonal winds, and upper tropospheric velocity potential for each reanalysis product, computed for the MJO period band in the tropics, 1970–1979, following the procedure of Ventrice et al. (2013).

The blue line shows the ensemble mean of the CESM-DART EOF structure with the shading indicating one ensemble standard deviation from the mean

errors reflects somewhat similar behavior found in Fig. 7, especially with R1 being shifted too far eastward.

These composite map differences are quantified in Table 4, which shows the spatial mean and rms errors of each composited MJO field (from Fig. 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.

5 Discussion and conclusion

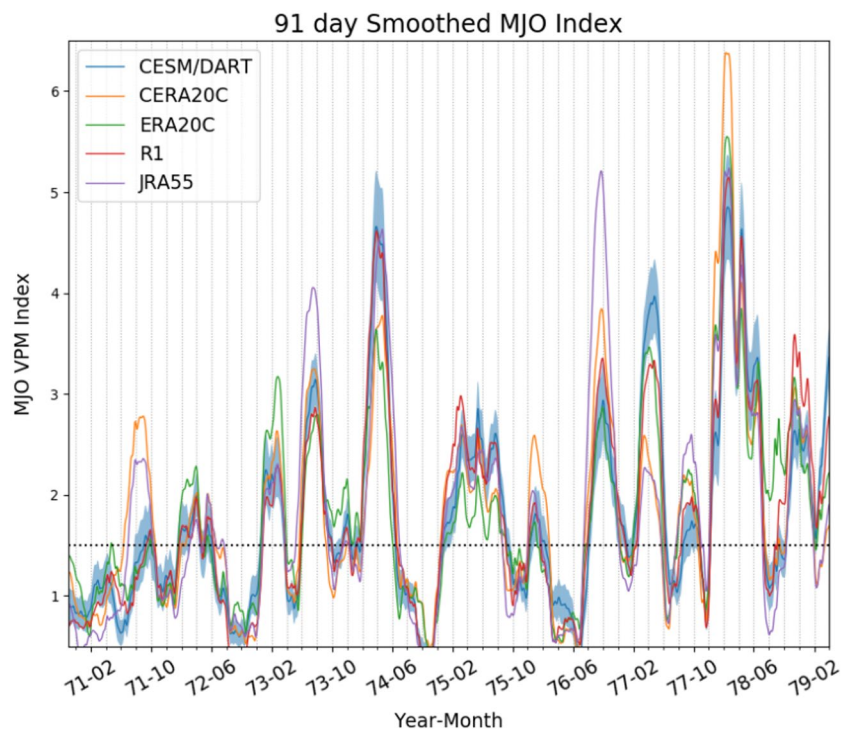
In this study, we analyzed a new prototype coupled ocean–atmosphere Ensemble Kalman Filter reanalysis product of the 1970s (Karspeck et al. 2018) by comparing its tropical climate variability to other reanalysis products, available observations, and a free-running version of the model. The results reveal 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 differences in the fields of 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 is broadly consistent the study of Kim et al. (2014) who showed that R1 and ERA-Interim were deficient in MJO energy compared to the more recent reanalysis products CFSR, MERRA and NCEP-DOE. Our findings suggest that the flux coupling at the ocean–atmosphere interface may be important in organizing the convective activity inherent in MJO. In addition, high-frequency free Kelvin waves in CESM-DART are reduced in amplitude compared to the Free-CESM run and the other products, which may be due to the assimilation procedure itself or potentially due to oceanic coupling. These Kelvin waves are often overly energetic in free-running climate models due to weak convective coupling (e.g., Hung et al. 2013). 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

Fig. 8 As in Fig. 7, but for the corresponding MJO VPM Index PC calculated for each reanalysis product



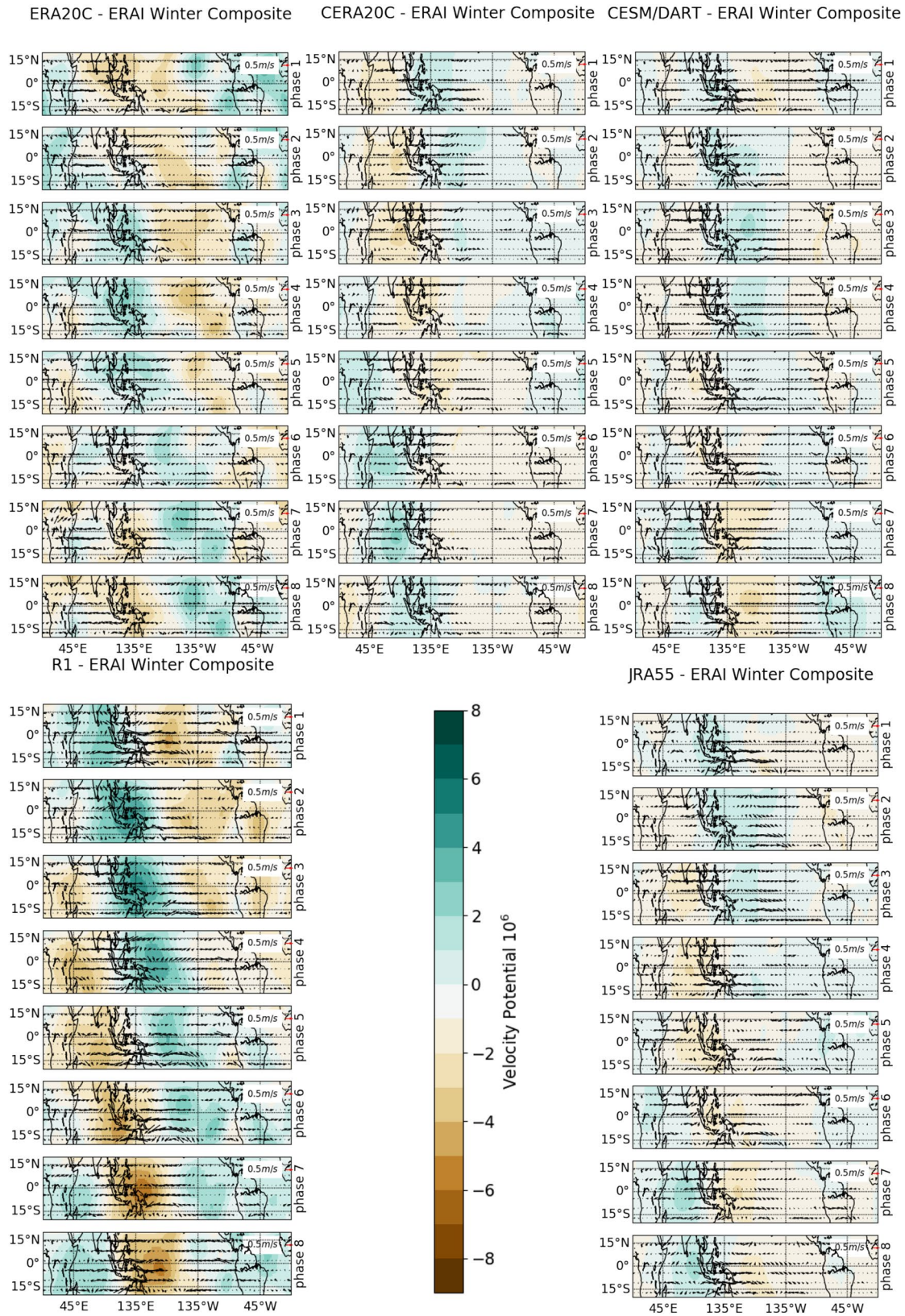


Fig. 9 Differences in phases through a composite MJO cycle in winter months shown as differences for each reanalysis product (1970–1979) relative to the corresponding composite computed from

the ERA-I (1991–2010). The contours indicate velocity potential at 200 hPa and the vectors indicate 850 hPa winds

Table 4 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
	JRA55	-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

precipitation field. The typical local bias is comparable to both CERA20C 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. The other coupled product, CERA20C, in contrast, had low tropical precipitation variability, but lacked the assimilation of tropospheric variables that could aid in organizing convective activity.

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 3) 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 related work (Eliashiv 2019) will help guide the development and improvement of ensemble climate forecast strategies for diagnostics and predictions.

Acknowledgements This study forms a portion of the Ph.D. dissertation of J.E. 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 NCAR's 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. We thank the two anonymous referees who provided extensive and insightful comments that significantly improved the clarity and interpretation of results in our manuscript.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use,

distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Anderson JL (2003) A local least squares framework for ensemble filtering. *Mon Weather Rev* 131(4):634–642. [https://doi.org/10.1175/1520-0493\(2003\)131<0634:ALLSFF>2.0.CO;2](https://doi.org/10.1175/1520-0493(2003)131<0634:ALLSFF>2.0.CO;2)
- Anderson J, Hoar T, Raeder K, Liu H, Collins N, Torn R, Avellano A (2009) The data assimilation research testbed: a community facility. *Bull Am Meteorol Soc* 90(9):1283–1296. <https://doi.org/10.1175/2009BAMS2618.1>
- Arnold NP, Branson M, Kuang Z, Randall DA, Tziperman E (2015) MJO intensification with warming in the superparameterized CESM. *J Clim* 28(7):2706–2724
- Balmaseda MA, Trenberth KE, Källén E (2013) Distinctive climate signals in reanalysis of global ocean heat content. *Geophys Res Lett* 40(9):1754–1759
- Carton JA, Giese BS (2008) A reanalysis of ocean climate using simple ocean data assimilation (SODA). *Mon Weather Rev* 136(8):2999–3017
- Chelliah M, Bell GD (2004) Tropical multidecadal and interannual climate variability in the NCEP-NCAR reanalysis. *J Clim* 17(9):1777–1803
- Dee DP, Uppala S, Simmons A, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda M, Balsamo G, Bauer P et al. (2011) The era-interim reanalysis: configuration and performance of the data assimilation system. *Q J R Meteorol Soc* 137(656):553–597
- Eliashiv J (2019) Exploring the usage of data assimilation on the community earth system model, PhD dissertation. PhD thesis, Scripps Institution of Oceanography, University of California, San Diego
- Frolov S, Bishop CH, Holt T, Cummings J, Kuhl D (2016) Facilitating strongly coupled ocean-atmosphere data assimilation with an interface solver. *Mon Weather Rev* 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). *J Clim* 30(14):5419–5454
- 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. *J Clim* 24(19):4973–4991
- Gill A (1982) *Atmosphere–ocean dynamics*. Academic Press, London
- Hodges KI, Lee RW, Bengtsson L (2011) A comparison of extratropical cyclones in recent reanalyses ERA-interim, NASA MERRA, NCEP CFSR, and JRA-25. *J Clim* 24(18):4888–4906
- Hung M-P, Lin J-L, Wang W, Kim D, Shinoda T, Weaver SJ (2013) MJO and convectively coupled equatorial waves simulated by CMIP5 climate models. *J Clim* 26(17):6185–6214
- Houtekamer PL, Mitchell HL (1998) Data assimilation using an ensemble Kalman filter technique. *Mon Weather Rev* 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. *Bull Am Meteorol Soc* 94(9):1339–1360
- Jones PW (1999) First-and second-order conservative remapping schemes for grids in spherical coordinates. *Mon Weather Rev* 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. *Bull Am Meteorol Soc* 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). *Bull Am Meteorol Soc* 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. *Q J R Meteorol Soc* 144(717):2404–2430
- Kim D, Lee MI, Kim D, Schubert SD, Waliser DE, Tian B (2014) Representation of tropical subseasonal variability of precipitation in global reanalyses. *Clim Dyn* 43(1–2):517–534
- Lalouaux P, De Boissesson E, Dahlgren P (2017) CERA-20C: an earth system approach to climate reanalysis. *ECMWF Newsl* 150:25–30
- Onogi K, Tsutsui J, Koide H, Sakamoto M, Kobayashi S, Hattushika H, Matsumoto T, Yamazaki N, Kamahori H, Takahashi K et al. (2007) The JRA-25 reanalysis. *J Meteorol Soc Jpn Ser II* 85(3):369–432
- Penny SG, Hamill TM (2017) Coupled data assimilation for integrated earth system analysis and prediction. *Bull Am Meteorol Soc* 97(7):ES169–ES172
- 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. *J Clim* 25(18):6304–6317
- Rienecker MM, Suarez MJ, Gelaro R, Todling R, Bacmeister J, Liu E, Bosilovich MG, Schubert SD, Takacs L, Kim GK, Bloom S et al. (2011) MERRA: NASA’s modern-era retrospective analysis for research and applications. *J Clim* 24(14):3624–3648
- Rodwell M, Lang S, Ingleby N, Bormann N, Hólm E, Rabier F, Richardson D, Yamaguchi M (2016) Reliability in ensemble data assimilation. *Q J R Meteorol Soc* 142(694):443–454
- 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. *Bull Am Meteorol Soc* 91(8):1015–1058
- Subramanian AC, Jochum M, Miller AJ, Murtugudde R, Neale RB, Waliser DE (2011) The Madden–Julian oscillation in CCSM4. *J Clim* 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. *Q J R Meteorol Soc* 131(612):2961–3012
- Ventrice MJ, Wheeler MC, Hendon HH, Schreck CJ III, Thorncroft CD, Kiladis GN (2013) A modified multivariate Madden–Julian oscillation index using velocity potential. *Mon Weather Rev* 141(12):4197–4210
- Wheeler M, Kiladis GN (1999) Convectively coupled equatorial waves: analysis of clouds and temperature in the wavenumber–frequency domain. *J Atmos Sci* 56(3):374–399
- Yu L, Jin X, Weller R (2008) Multidecade global flux datasets from the objectively analyzed air-sea fluxes (oafux) project: Latent and sensible heat fluxes, ocean evaporation, and related surface meteorological variables. OAFux project technical report. OA-2008-01

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.