Evaluation of Mean State in NCEP Climate Forecast System (Version 2) Simulation Using a Stochastic Multicloud Model Calibrated With DYNAMO RADAR Data

Kumar Roy1,2, Parthasarathi Mukhopadhyay1, R. P. M. Krishna1, B. Khouider3,4, and B. B. Goswami1

1Indian Institute of Tropical Meteorology, Ministry of Earth Sciences, Pune, India, 2Department of Meteorology and Oceanography, College of Science and Technology, Andhra University, Vishakhapatnam, India, 3Department of Mathematics and Statistics, University of Victoria, Victoria, BC, Canada, 4Irreversible Climate Change Research Center (IRCC), Yonsei University, Seoul, South Korea

Abstract Stochastic parameterizations are continuously providing promising simulations of unresolved atmospheric processes for global climate models (GCMs). One of the stochastic multi-cloud model (SMCM) features is to mimic the life cycle of the three most common cloud types (congestus, deep, and stratiform) in tropical convective systems. To better represent organized convection in the Climate Forecast System version 2 (CFSv2), the SMCM parameterization is adopted in CFSv2 (SMCM-CTRL) in lieu of the pre-existing revised simplified Arakawa–Schubert (RSAS) cumulus scheme and has shown essential improvements in different large-scale features of tropical convection. But the sensitivity of the SMCM parameterization from the observations is yet to be ascertained. Radar data during the Dynamics of the Madden-Julian Oscillation (DYNAMO) field campaign is used to tune the SMCM in the present manuscript. The DYNAMO radar observations have been used to calibrate the SMCM using a Bayesian inference procedure to generate key time scale parameters for the transition probabilities of the underlying Markov chains of the SMCM as implemented in CFSv2 (hereafter SMCM-DYNAMO). SMCM-DYNAMO improves many aspects of the mean state climate compared to RSAS, and SMCM-CTRL. Significant improvement is noted in the rainfall probability distribution function over the global tropics.

The global distribution of different types of clouds, particularly low-level clouds, is also improved. The convective and large-scale rainfall simulations are investigated in detail.

1. Introduction

Although there has been considerable progress made in the representation of clouds in the global climate models in recent years, a large spread in cloud feedback remains one of the primary sources of uncertainties for climate sensitivity estimation (Cess et al., 1990; Houghton et al., 2001; Stephens, 2005). One of the possible reasons for this issue may be the inadequate parameterization of certain aspects of cloud processes (Bony & Dufresne, 2005; Roy et al., 2020; Senior & Mitchell, 1993; Zhang, 2005; Zhao et al., 2016). In fact, low-level clouds have been identified in the Intergovernmental Panel on Climate Change's (IPCC’s) report on model evaluation (Flato et al., 2013; Randall et al., 2007) as a primary source of uncertainty for the sensitivity of climate models. Notably, climate models face difficulties in simulating tropical marine stratocumulus, trade cumulus, and the transition to deep convection, which may cause the significant uncertainties in making projections of future climate scenarios (Klein et al., 2017). The difficulty in accurately modeling clouds in global climate models (GCMs) arises not only due to the limited spatial resolution of the climate models but also due to the inadequate representation of the highly interactive nature of these clouds and associated complex physical processes (Jakob, 2001; Moncrieff & Klinker, 1997; Siebesma et al., 2004; Jakob (2010); Stevens and Bony (2013); Jakob (2014) and Bony et al. (2015). Cumulus parameterization schemes are designed so that they can represent convective effects from large-scale dynamics, but they cannot simulate sub-grid scale variability (Lin & Neelin, 2002). Furthermore, it may also help to argue in favor of sub-grid scale resolving approaches for the benefit of improved convection parameterizations for future generations of GCMs as workhorses of climate science, because emerging work suggests that convection-permitting simulations provide significant added-value in simulating convection-circulation interactions (e.g., Klocke...
et al., 2017; Peters et al., 2019; Satoh et al., 2019; Stevens et al., 2019, 2020). The goal should be to capture as adequately as possible the features detailed in those studies by improving convection parameterizations and of which the stochastic multi-cloud model (SMCM) approach is an excellent example. It was suggested by Palmer (2001) that the sub-grid variability should belong to the parameterization scheme executed by the dynamic-stochastic systems coupled with the resolved system over the diverse range of scales rather than using deterministic strategies. Neglecting small scale variability can lead to the accumulation of more errors in the climatology of large-scale variables.

For this reason, some researchers have opted for the incorporation of stochastic elements in those schemes. Buizza et al. (1999) showed more accuracy in the probabilistic prediction of precipitation when the parameterized tendencies were perturbed stochastically. Skillful precipitation variance is achieved by Lin and Neelin (2003) when adding a stochastic term to the convective available potential energy (CAPE) and the vertical structure of the heating of a traditional convective scheme. Previous studies by Khoury and Majda (2006a, 2006b) suggest that different types of clouds (congestus, deep convective, and stratiform) play an essential role in convectively coupled waves. Khoury et al. (2010) developed the SMCM to simulate the variability of the small unresolved features of organized tropical convection in GCMs. This model was introduced as a “stochasticization” of the multicloud parameterization, which focuses on clouds’ lifetime in the tropical atmosphere. It was first introduced by Khoury and Majda (2006b). The deterministic and stochastic flavors of the multi cloud model (MCM) have been successfully implemented in an idealized aquaplanet-GCM, and it shows better agreement with the simulated features of tropical convection and wave characteristics (Ajayamohan et al., 2016; Deng et al., 2015, 2016; Khoury et al., 2010). The SMCM was first shown to be able to reproduce observed characteristics of tropical convection by Peters et al. (2013). Recent studies (Ma et al., 2019; Peters et al., 2017) used the SMCM-onescales informed by observations deduced in Peters et al. (2013) and showed that the fidelity of the SMCM, which helps an existing cumulus scheme in a GCM to trigger better and modulate convection, which further improves the simulation of tropical intraseasonal variability.

Goswami et al. (2017c; see also Goswami et al., 2017a, 2017b) implemented the SMCM in the National Center for Environmental Prediction (NCEP) Climate Forecasting System second version 2 (CFSv2) climate model. It is seen that with the incorporation of SMCM in lieu of the pre-existing cumulus scheme, the tropical modes of atmospheric variability are simulated reasonably well compared to observations without the deterioration of the mean climatology as it is often the case with the traditional cumulus scheme (Kim et al., 2011; Klingaman & Demott, 2020). However, the SMCM largely depends on its large number of tuning parameters. Further, the improvements in GCM simulations based on the SMCM are subject to a better estimation of those parameters. Recently, De La Chevrotière et al. (2014, 2016) developed a Bayesian inference method to estimate key time scale parameters of the transition probabilities that define the time evolution of the emulated distributions from observations to reduce uncertainties in the SMCM. Cardoso-Bihlo et al. (2019) recently introduced and reassessed a refinement of the SMCM by introducing new dynamical predictors such as convective inhibition and vertical subsidence and the previously used mid-level moisture and CAPE. They then extended the Bayesian inference algorithm of De La Chevrotière et al. (2014) to estimate the transition times between the various cloud types and applied Dynamics of the Madden-Julian Oscillation (DYNAMO) data. They considered three regimes associated with an MJO event, namely the suppressed phase, the initiation phase, and the mature phase.

In this work, we use the DYNAMO transition time scales of Cardoso-Bihlo et al. (2019) in CFSv2-SMCM (hereafter SMCM-DYNAMO). After replacing the tuning parameters in SMCM-DYNAMO, we have run the model for 25 years and compared the results with the pre-existing revised, simplified Arakawa–Schubert (RSAS as in Ganai et al., 2015) cumulus scheme and with the SMCM implemented in CFSv2 (SMCM-CTRL as in Goswami et al., 2017b).

This is the first time the observations tune the SMCM in the CFSv2 model, and the model shows promising improvement compared with RSAS which is based on deterministic closure and SMCM-CTRL, for which the time-scale transition parameters are inferred from large eddy simulation data of tropical Atlantic convection by De La Chevrotiere et al. (2016).
The study is organized as follows. A brief overview of the SMCM parameterization, the data, and the methodology are presented in Section 2. The main simulation results are reported in Section 3, while a concluding summary is given in Section 4.

2. Model Description, Data, and Methodology

The NCEP CFSv2 has been adopted under the Monsoon Mission program of the Ministry of Earth Sciences, Government of India for improving dynamical monsoon prediction at various space and time scales (Rao et al., 2019). The model considered for the present study is coupled NCEP CFSv2 with an atmospheric spectral resolution of T126 (~100 km) and 64 hybrid vertical levels. Additionally, the Geophysical Fluid Dynamics Laboratory Modular Ocean Model, version 4p0d (Griffies et al., 2004), is utilized as the oceanic part in CFSv2 with a zonal resolution of 0.25–0.5° and 40 vertical layers. More details about CFSv2 model components can be found in Saha et al. (2014). The deep convective parameterization scheme is the RSAS in CFSv2 based on Han and Pan (2011) study. The deep convective parameterization scheme is replaced by the stochastic multicloud model (SMCM) component, and it is implemented in NCEP CFSv2 (Goswami et al., 2017a, 2017b, 2017c). Readers may see the details of SMCM in Khouider et al. (2010). The SMCM component largely depends on its large number of tuning parameters. One set of such key parameters is represented by the transition times between the various cloud types. Earlier, the transition times were inferred from large eddy simulation data of tropical Atlantic convection by De La Chevrotiere et al. (2016) and used in CFSv2 (Goswami et al., 2017b). In the present study, the transition times are calculated from...
Figure 2. Rainfall probability distribution function (PDF) averaged over 10°S–10°N for TRMM-3B42, simplified Arakawa–Schubert (RSAS), SMCM-CTRL, and SMCM-DYNAMO respectively. The ranges of rain rate are (mm day$^{-1}$) along X axis and rainfall PDF (in fraction) along Y axis.

Figure 3. (a–c) Annual mean climatological convective rainfall (mm day$^{-1}$) over the 50°S–50°N band from simplified Arakawa–Schubert (RSAS), SMCM-CTRL, and SMCM-DYNAMO respectively, (d) difference in convective rainfall (mm day$^{-1}$) between SMCM-CTRL and RSAS, (e) difference in convective rainfall (mm day$^{-1}$) between SMCM-DYNAMO and RSAS, (f) difference in convective rainfall (mm day$^{-1}$) between SMCM-DYNAMO and SMCM-CTRL.
the initiation phase of the MJO event of DYNAMO observational data-sets (Cardoso-Bihlo et al., 2019) by applying the Bayesian inference technique by De La Chevrotière et al. (2014, 2016).

Thus, the newly generated transition times, based on DYNAMO observational data-sets are used in CFSv2 (hereafter SMCM-DYNAMO) by replacing the ones used by Goswami et al. (2017a, 2017b, 2017c).

In the present study, a new free run, with SMCM-DYNAMO in CFSv2 at T126 spectral resolution is carried out for 25 years, and the output is stored at a 24 h interval. The initial conditions (both atmospheric and oceanic) are taken from NCEP CFS Reanalysis (Saha et al., 2010). Except for the use of different transition times in the SMCM for the parameterization of deep-convection, the other components of CFSv2 are the same for SMCM-CTRL and SMCM-DYNAMO simulations. The model experiment has been executed in the Ministry of Earth Sciences high-performance computing system facility “Aditya” at Indian Institute of Tropical Meteorology, Pune, India.

Various observational and reanalysis data sets are used for the validation of the model simulations. The Tropical Rainfall Measuring Mission (TRMM) 3B42 version 7 (V7) (Huffman et al., 2007) daily rainfall data-sets is used at a horizontal resolution of 0.25° × 0.25° for the year 1999–2012. National Oceanic and Atmospheric Administration (NOAA) interpolated outgoing long-wave radiation (OLR) data from National Center for Atmospheric Research (NCAR) (Liebmann & Smith, 1996) is used from 1985 to 2005 with a spatial resolution of 1° X 1°. Daily data of Clouds and the Earth’s Radiant Energy System (CERES) Energy Balance and Filled (EBAF) product for the period for 2004–2018 is used as an observational counterpart for shortwave cloud forcing (SWCF) and Longwave cloud forcing (LWCF) in the present study. CALIPSO-GOC-
CP data of low-level \( (z < 3.36 \text{ km}) \), mid-level \( (3.36 \text{ km} < z < 6.72 \text{ km}) \), high-level \( (z > 6.72 \text{ km}) \) clouds (Chepfer et al., 2010) for the period of 2007–2018 have been utilized in the present study.

3. Results

To check the impact of SMCM calibrated with DYNAMO radar data, we have investigated the simulation of mean state climate over 50°S–50°N band in comparison with the pre-existing RSAS and SMCM_CTRL simulations.

3.1. Precipitation and OLR Climatologies

Figure 1 displays the annual mean precipitation and associated biases over the 50°S–50°N band. The left panels (Figures 1a–1d) show the annual mean precipitation from TRMM, RSAS, SMCM-CTRL, and SMCM-DYNAMO. Although the annual mean precipitation is overestimated in SMCM-DYNAMO, more than RSAS and SMCM-CTRL, overall, the simulated climatologies look similar, and the double ITCZ problem remains unresolved. Moreover, the double ITCZ bias is even exaggerated in SMCM-DYNAMO simulation. This is perhaps an indication that the double ITCZ problem has more to do with atmosphere-ocean feed-

Figure 5. Figure 5 Scatter diagram of the convective precipitation versus convective precipitation efficiency over the Global tropics. The ranges of convective precipitation are (unit: mm day\(^{-1}\)) along X axis and convective precipitation efficiency (unit: x 104 day\(^{-1}\)) along Y axis.

Figure 6. (a–c) Annual Joint distribution (in %) of OLR-stratiform rainfall over 50°S–50°N band, for simplified Arakawa–Schubert (RSAS), SMCM-CTRL, and SMCM-DYNAMO respectively.
back and the associated surface fluxes than it has to do with the convective parameterization (Lin, 2007; Xie & Philander, 1994). Nonetheless, SMCM-DYNAMO shows a few improvements regionally, for example, the dry biases over the Indian summer monsoon (ISM) region, northern Australia, central Pacific, African landmass and Amazonia have substantially reduced in SMCM-DYNAMO as compared to RSAS and SMCM-CTRL simulations. The right panels (Figures 1e–1g) depict the corresponding biases. Some exaggeration in precipitation is noticeable in the equatorial Atlantic and the north and south of the central Pacific in SMCM-DYNAMO simulation as compared to other two simulations. Also, the lower RMSE and higher spatial correlations (CC) are evident in SMCM-DYNAMO as compared to SMCM-CTRL. Although, the RMSE is less in RSAS as compared to SMCM-DYNAMO. This may be due to the overly estimated precipitation by the SMCM-DYNAMO simulation.

To get a better insight of model performance, it is worthy of computing and analyzing the rainfall probability distribution function (PDF) over the global tropics (Figure 2). It is found that SMCM-DYNAMO improves the simulation of the lighter rain-rate categories. Moderate rain-rate categories also are better simulated in SMCM simulations as compared to RSAS. These distributions of rainfall beyond 50 mm day$^{-1}$ are also better simulated by SMCM-DYNAMO as compared to RSAS. Interestingly, Peters et al. (2017) also had reported very similar improvements in the rainfall PDF. The overestimation of the heavier rain-rate categories (>50 mm day$^{-1}$) in SMCM-CTRL is reduced in SMCM-DYNAMO as compared to observations. The RSAS simulated PDF appears to significantly underestimate the heavy rain rate frequencies and overestimates the light rain events. This is consistent with the general idea that climate models using deterministic parameterizations tend to produce too much light rain as it has been reported by several studies.
One of the major concerns in climate modeling is that they tend to generate excessive amounts of convective rain at the expense of so little stratiform (i.e., large scale) rain (Dai, 2006). It has been documented that climate models generate around 80%-90% convective versus approximately 20% stratiform rain (Dai, 2006). The global (50°S-50°N) annual mean convective and stratiform (large-scale) rainfall, for the RSAS, SMCM_CTRL, and SMCM-DYNAMO, are shown in Figures 3 and 4 respectively. Figures 3a–3c show the convective precipitation from RSAS, SMCM-CTRL, and SMCM-DYNAMO, respectively. The right panels (Figures 3d–3f) show the difference between model convective rainfall from SMCM-CTRL (Figure 3d) and SMCM-DYNAMO (Figure 3e) with respect to RSAS and between SMCM-DYNAMO and SMCM-CTRL (Figure 3f), respectively. It is visibly evident from (Figures 3a–3c) that the dominantly huge convective rainfall (Figure 3a) production from RSAS (Ganai et al., 2015) is reduced significantly in both SMCM-CTRL and SMCM-DYNAMO simulations, with SMCM-DYNAMO produces more convective rainfall than SMCM-CTRL (comparing Figures 3b and 3c). But, the magnitude of the annual mean convective rainfall is substantially less in SMCM-DYNAMO as compared to RSAS. Consistently, from (Figures 4a–4c), both the SMCM simulations produce a significant amount of stratiform rain as compared to RSAS. The reduction of convective rainfall substitutes the enhancement of stratiform rainfall in both SMCM simulations which may be due to less intense deep-convective mass-fluxes in the SMCM versions as shown in Peters et al. (2017). The right panels (Figures 4d–4f) show the difference between model stratiform rainfall from SMCM-CTRL (Figure 4d) and SMCM-DYNAMO (Figure 4e) with respect to RSAS and between SMCM-DYNAMO and SMCM-CTRL (Figure 4f), respectively. This figure further demonstrates more production of stratiform rainfall in both SMCM simulations as compared to RSAS. Figure 4f
shows blue shades all over 50°S–50°N which depicts that SMCM-CTRL produce more stratiform rainfall as compared to SMCM-DYNAMO.

Figures 3 and 4 depict that SMCM versions produce less (more) convective (stratiform) rainfall than RSAS. These results are further supported by scatter diagram (Figure 5) of the convective precipitation efficiency versus convective precipitation over the Global tropics. The convective precipitation efficiency (Li et al., 2012, unit, day⁻¹) is computed by dividing the total daily convective precipitation amount by the corresponding total daily cloud water path (LWP). This diagram shows RSAS has too much frequency of generating convective rain in all thresholds of convective precipitation where both the SMCM versions show lower convective precipitation efficiency as compared to RSAS. This may be due to less intense deep convection in SMCM versions as shown earlier in Peters et al. (2017). Thus, there is still room for improvement in the SMCM versions of CFSv2. Now, to check the contribution of large-scale rainfall in heavier categories in model simulations, OLR-Stratiform rainfall Joint PDF is computed and analyzed for RSAS, SMCM-CTRL, and SMCM-DYNAMO in Figure 6. This figure shows RSAS has less stratiform rainfall (Figure 6a) for lower OLR values as compared to SMCM versions (Figures 6b and 6c) which is consistent with more convective rainfall in RSAS as compared to SMCM versions in Figure 3. The high values of stratiform rainfall in SMCM versions are coming from low values of OLR with broad spectrum as compared to RSAS with narrow spectrum of OLR-Stratiform rainfall distribution.

Further, the spatial distribution of OLR is shown in Figure 7. The left panels (Figures 7a–7d) establish the annual mean distribution of OLR over the 50°S–50°N from NOAA observations, RSAS, SMCM-CTRL, and SMCM-DYNAMO, repetitively while the right panels (Figures 7e–7g) show the respective model biases to NOAA observations. The annual mean value of OLR is 243.43 Wm⁻² from NOAA observations. Although all the model simulations overestimate the value, the mean value of OLR from SMCM-CTRL appears to be closer to the observed one (244.83 Wm⁻² vs. 249.43 Wm⁻² in RSAS). The mean OLR value of SMCM-
DYNAMO (245.25 Wm$^{-2}$) is closer to SMCM-CTRL (244.83 Wm$^{-2}$). The right panels (Figures 7e–7g) show the corresponding bias plots. Overall bias in SMCM-CTRL is improved than RSAS and SMCM-DYNAMO as evident by lower RMSE value. Although, CC value is more in RSAS than two SMCM versions, SMCM-DYNAMO shows higher CC value than SMCM-CTRL.

3.2. Mean Cloud Cover

The high-level, mid-level, low-level clouds are estimated for the three model simulations and the CALIPSO-GOCCP observations in Figures 8–10. Figures 8a–8d show the high-level clouds from CALIPSO-GOCCP observation and the model simulations by RSAS, SMCM-CTRL and SMCM-DYNAMO. As indicated on the top-right corners of each of the left panels (Figures 8a–8d), the high-level cloud global mean value has not improved in SMCM-DYNAMO compared to RSAS and SMCM-CTRL when assessed against the CALIPSO-GOCCP observations. There is also a significant underestimation in terms of the overall global pattern in the SMCM-DYNAMO simulation, as indicated by the higher RMSE error and lower CC values on the right panels (Figures 8e–8g). Moreover, the negative bias over Asian Summer Monsoon (ASM) region and Amazonia has gotten significantly worse in SMCM-DYNAMO than SMCM-CTRL and RSAS. But, the positive bias over the central Pacific is improved in SMCM-DYNAO simulation. Figures 9a–9d show mid-level cloud from CALIPSO-GOCCP observation and the model simulations. All the models qualitatively simulate the overall spatial pattern of mid-level cloud over 50°S–50°N band but high annual mean value of mid-level cloud in all the simulations are evident as compared to CALIPSO-GOCCP observations. The right panels (Figures 9e–9g) show the corresponding biases. All the models overestimate mid-level cloud over 50°S–50°N, except a marginal improvement in the SMCM-DYNAMO simulations over central India and Amazonia landmass. Lower RMSE value also is
evident in SMCM-DYNAMO as compared to other two simulations. But SMCM-CTRL shows higher CC value than SMCM-DYNAMO.

Figure 10 depicts the annual mean simulation of low-level clouds and corresponding biases over 50°S–50°N. Left panels (Figures 10a–10d) show spatial distribution of annual mean low-level cloud amount. Annual mean low-level cloud amount in RSAS is less than SMCM versions as compared to CALIPSO-GOCCP observations. Both SMCM versions qualitatively show the same spatial pattern and the annual mean values of both the simulations are also close to the observation. Although, the annual mean value of low-level cloud amount in SMCM-CTRL is slightly more than SMCM-DYNAMO. The right panels (Figures 10e–10g) show the corresponding biases. RSAS shows negative bias over oceanic regions with respect to CALIPSO-GOCCP observation. The negative bias in RSAS is reduced in SMCM versions with respect to CALIPSO-GOCCP observations. The positive bias over the south and south-east of central Pacific in SMCM-CTRL is marginally improved in SMCM-DYNAMO. Improvement in low-level cloud biases in SMCM versions than RSAS is also evident from lower RMSE and higher CC values. However, RMSE (CC) is slightly higher (lesser) in SMCM-DYNAMO than SMCM-CTRL.

Shortwave cloud forcing (SWCF) is one of the most important metrics to examine for improvement in low-level clouds’ representation. The observation and simulated SWCF are plotted in Figures 11a–11d. All the models qualitatively capture the main features and patterns of the observed SWCF, as indicated by the Clouds and the Earth’s Radiant Energy System (CERES) Energy Balance and Filled (EBAF) product. However, none of the models could capture the spatial pattern of SWCF over north and south of the central Pacific as in the observations. Over the Indian landmass all the models underestimate SWCF. The annual mean value of SWCF is 1.9 W m⁻² less in SMCM-DYNAMO, 9.9 m⁻² less in SMCM-CTRL and 4.6 W m⁻² more in RSAS than the observations. Figure 12 depicts annual mean longwave cloud forcing (LWCF) over 50°S–50°N from CERES-EBAF observation, RSAS, SMCM-CTRL, and SMCM-DYNAMO. The spatial pattern appears to be similar in all the models. LWCF is underestimated over the central pacific, Amazonia and the equatorial Indian Ocean in all the models compared to observations. The annual mean value of SMCM-DYNAMO is 5.65 W m⁻² less than the observations, whereas the mean value of RSAS is 3.13 W m⁻² less than the observation. The annual mean value of LWCF in SMCM-CTRL is close to the observation.

4. Summary and Conclusions

In the present study, we have used the new transition time parameters computed from DYNAMO observations in CFSv2-SMCM (SMCM-DYNAMO). We have run the SMCM-DYNAMO model for 25 years and compared the results with the pre-existing revised simplified Arakawa–Schubert (RSAS, Ganai et al., 2015) cumulus scheme and default SMCM (SMCM-CTRL, Goswami et al., 2017b). Among all the simulations of CFSv2, the SMCM versions simulate many aspects of mean climate state reasonably better compared to RSAS performance. Given that CFSv2 is one of the better state-of-the-art climate models, this is a satisfactory result. The use of observation-based transition time parameters has made the SMCM simulation more realistic. In the SMCM framework, these transition time parameters are crucial for capturing convective organization (Goswami et al., 2017b). Therefore, by improving these parameters we expect to see improvements in the distribution of convection and precipitation more than in the mean state. Nonetheless, the SMCM-DYNAMO simulation is found to have a fairly good mean state, at least as good as the RSAS and CTRL-SMCM model if not better in some aspects. As expected, we found significant improvements in
the rainfall PDF over the global tropics and improvement in the distribution of different types of clouds, particularly low-level clouds. These improvements indicate that observation guided well trained SMCM parameters are a key to utilize the full potential of the SMCM framework. This study is the step forward in the direction of using machine learning to tune model parameters. It is true that the SMCM-CTRL seems to perform better than the SMCM_DYNAMO in some metrics but this is in no means constitutes a negative result. One has to keep in mind that the SMCM-CTRL runs has been heavily tuned (see Goswami et al., 2017a, JAMES) while the SMCM-DYNAMO is a first attempt dry/run so the fact the results comes very close to the SMCM-CTRL and remain overly better than RSAS is a huge success.

This study shows many gap areas with scopes of further improvement. Thorough training of the SMCM models over the whole five-month period of DYNAMO is warranted. There is also the possibility of expanding the training data to other regions of the globe to include, for instance, convection regimes over tropical landmasses such as India.

This study is the first attempt to calibrate the CFS-implementation of the SMCM with transition timescales informed by observations and investigate the impact on the mean climate state. Further, while this is the first attempt to do so in CFS, an observation-informed version of the SMCM was already tested earlier in Peters et al. (2017). This study considered the observational data-sets for a concise period to calculate the transition time-scales for the SMCM, which would confidently say that SMCM should use more realistic transition time-scales extracted from actual observations. This study concludes the requirement of more observational data-sets in diverse convective environments to calibrate SMCM and improve GCMs for proper simulation of different aspects of clouds and convection.

Data Availability Statement

The CALIPSO-GOCCP datasets were obtained from https://climserv.ipsl.polytechnique.fr/cfmip-obs/CALIPSO_goccp.html. CERES-EBAF plus data-sets were obtained from CERES data products (https://ceres.larc.nasa.gov/data/). All model data-sets are available at https://data.mendeley.com/datasets/jjxk2sbmmt/.

References

Ajayamohan, R. S., Khoudier, B., Majda, A. J., & Deng, Q. (2016). Role of stratiform heating on the organization of convection over the monsoon trough. Climate Dynamics, 47(12), 3641–3660. https://doi.org/10.1007/s00382-016-3033-7

Bony, S., & Dufresne, J. L. (2005). Marine boundary layer clouds at the heart of tropical cloud feedback uncertainties in climate models. Geophysical Research Letters, 32, L20806. https://doi.org/10.1029/2005GL023851

Bony, S., Stevens, B., Frierson, D. M., Jakob, C., Kageyama, M., Pincus, R., et al. (2015). Clouds, circulation and climate sensitivity. Nature Geoscience, 8, 261–268. https://doi.org/10.1038/ngeo2398

Buizza, R., Miller, M., & Palmer, T. N. (1999). Stochastic representation of model uncertainties in the ECMWF Ensemble Prediction System. Quarterly Journal of the Royal Meteorological Society, 125, 2887–2908.

Cardoso-Bihlo, E., Khoudier, B., Schumacher, C., & De La Chevrotière, M. (2019). Using radar data to calibrate a stochastic parametrization of organized convection. Journal of Advances in Modeling Earth Systems, 11, 1655–1684. https://doi.org/10.1029/2018MS001537

Cess, R., Potter, G. L., Blanchet, J. P., Boer, G. J., Del Genio, A. D., Deque, M., et al. (1990). Intercomparison and interpretation of cloud-climate feedback processes in nineteen atmospheric general circulation models. Journal of Geophysical Research, 95, 16601–16615. https://doi.org/10.1029/JD095iD10p16601

Chepfer, H., Bony, S., Winker, D. M., Cesana, G., Dufresne, J. L., Minnis, P., et al. (2010). The GCM Oriented CALIPSO Cloud Product (CALIPSO-GOCCP). Journal of Geophysical Research, 115, D00H16. https://doi.org/10.1029/2009JD012251

Dai, A. (2006). Precipitation characteristics in eighteen coupled climate models. Journal of Climate, 19, 4605–4630. https://doi.org/10.1175/jcli3884.1

Acknowledgments

Indian Institute of Tropical Meteorology, Pune, is fully funded by the Ministry of Earth Sciences (MoES), Government of India. Authors (Kumar Roy, Parthasarathi Mukhopadhyay, and R.P.M. Krishna) thank the Director, IITM, Pune for the study’s motivation and encouragement. Kumar Roy would like to thank Dr. Malay Ganai for useful discussions. The authors would like to thank the anonymous reviewers and the editor for their valuable suggestions which helped in improving the manuscript. The authors would like to thank GSFC/DAAC, NASA for providing TRMM3B42 (http://mirador.gsfc.nasa.gov/cgi-bin/mirador/presentNavigation.pl?tree=project&project=TRMM&dataGroup=Gridded) data sets and ESRL for NOAA (https://www.esrl.noaa.gov/psd/data/gridded/) data sets. All the model runs are being carried out in MoES, High Power Computing facility “Aaditya” at IITM, Pune. The authors would like to sincerely thank two anonymous reviewers and the editor for their valuable comments that greatly helped to improve the manuscript.

Figure 12. Same as Figure 10 but for longwave cloud forcing (LWCF, Units: Wm−2).
De La Chevrotière, M., Khouider, B., & Majda, A. J. (2014). Calibration of the stochastic multicloud model using Bayesian inference. SIAM Journal on Scientific Computing, 36(3), B538–B560. https://doi.org/10.1137/13094267x

De La Chevrotière, M., Khouider, B., & Majda, A. J. (2016). Stochasticity of convection in Giga-LES data. Climate Dynamics, 47, 1845–1861. https://doi.org/10.1007/s00382-015-2936-z

Deng, Q., Khouider, B., & Majda, A. (2015). The MJO in a coarse-resolution GCM with a stochastic multicloud parameterization. Journal of the Atmospheric Sciences, 72, 55774. https://doi.org/10.1175/JAS-D-14-0120.1

Deng, Q., Khouider, B., Majda, A., & Ajayamohan, R. (2016). Effect of stratiform heating on the planetary-scale organization of tropical convection. Journal of the Atmospheric Sciences, 73, 371–392. https://doi.org/10.1175/JAS-D-15-0178.1

Flato, M., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S. C., Collins, W., et al. (2013). Evaluation of Climate Models. Climate Change 2013: The Physical Science Basis. In Stocker, T. et al. (Eds.). Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change. United Kingdom and New York: Cambridge University Press.

Ganai, M., Mukhopadhyay, P., Krishna, R. P. M., & Mahakur, M. (2015). The impact of revised simplified Arakawa-Schubert convection parameterization scheme in CFSv2 on the simulation of the Indian summer monsoon. Climate Dynamics, 45(3–4), 881–902. https://doi.org/10.1007/s00382-014-2320-4

Goswami, B. B., Khouider, B., Krishna, R. P. M., Mukhopadhyay, P., & Majda, A. J. (2017a). Implementation and calibration of a stochastic convective parameterization in the NCEP Climate Forecast System. Journal of Advances in Modeling Earth Systems, 9, 1721–1739. https://doi.org/10.1002/2017ms001014

Goswami, B. B., Khouider, B., Krishna, R. P. M., Mukhopadhyay, P., & Majda, A. J. (2017b). Improved tropical modes of variability in the NCEP climate forecast system (version 2) via a stochastic multicloud model. Journal of the Atmospheric Sciences, 74, 3339–3366. https://doi.org/10.1175/jas-d-17-0113.1

Goswami, B. B., Khouider, B., Krishna, R. P. M., Mukhopadhyay, P., & Majda, A. J. (2017c). Improving synoptic and intraseasonal variability in CFSv2 via stochastic representation of organized convection. Geophysical Research Letters, 44, 1104–1113. https://doi.org/10.1002/2016GL071542

Griffies, S. M., Harrison, M. J., Pacanowski, R. C., & Ronati, A. (2004). A technical guide to MOM4 GFDL ocean group (Technical Report 5 (p. 337). GFDL.

Han, J., & Pan, H.-L. (2011). Revision of convection and vertical diffusion schemes in the NCEP Global Forecast System. Weather and Forecasting, 26(4), 520–533. https://doi.org/10.1175/waf-d-09-05038.1

Houghton, J. T., Ding, Y., Griggs, D. J., Noguer, M., van der Linden, P. J., Dai, X., et al. (2001). Climate Change 2001. In The scientific basis (p. 49). New York: Cambridge University Press.

Huffman, G. J., Bolvin, D. T., Nelkin, E. J., Wolff, D. B., Adler, R. F., Gu, G., et al. (2007). The TRMM multisatellite precipitation analysis (TMPA): Quasiglobal, multiyear, combined-sensor precipitation estimates at fine scales. Journal of Hydrometeorology, 8(1), 38–55. https://doi.org/10.1175/jhm560.1

Jakob, C. (2001). The representation of cloud cover in atmospheric general circulation models (PhD thesis). Ludwig-Maximilians-UniversitätMünchen.

Jakob, C. (2010). Accelerating progress in global atmospheric model development through improved parameterizations: Challenges, opportunities, and strategies. Bulletin of the American Meteorological Society, 91, 869–876. https://doi.org/10.1175/2009bams2898.1

Jakob, C. (2014). Going back to basics. Nature Climate Change, 4, 1042–1045. https://doi.org/10.1038/nclimate2445

Khouider, B., Biello, J., & Majda, A. J. (2010). A stochastic multicloud model for tropical convection. Communications in Mathematical Sciences, 8(1), 187–216.

Khouider, B., & Majda, A. J. (2006a). Model multi-cloud parameterizations for convectively coupled tropical waves: Detailed nonlinear wave evolution. Dynamics of Atmospheres and Oceans, 42, 59–80. https://doi.org/10.1016/j.dynatmoce.2005.12.001

Khouider, B., & Majda, A. J. (2006b). A simple multicloud parameterization for convectively coupled tropical waves. Part I: Linear analysis. Journal of the Atmospheric Sciences, 63, 308–323. https://doi.org/10.1175/jas3677.1

Kim, D., Sobel, A. H., Frierson, D. M. W., Maloney, E. D., & Kang, I.-S. (2011). A systematic relationship between intraseasonal variability and mean state bias in AGCM simulations. Journal of Climate, 24, 5506–5520. https://doi.org/10.1175/2011jcli4177.1

Klein, S. A., Hall, A., Norris, J. R., & Pincus, R. (2017). Low-cloud feedbacks from cloud-controlling factors: A review. Surveys in Geophysics, 38, 1307–1329. https://doi.org/10.1007/s10712-017-9433-3

Klingaman, N. P., & Demott, C. A. (2020). Mean-state biases and interannual variability affect perceived sensitivities of the Madden-Julian oscillation to air-sea coupling. Journal of Advances in Modeling Earth Systems, 12.e2019MS0017999. https://doi.org/10.1029/2019MS0017999

Klocke, D., Brueck, M., Hohenegger, C., & Stevens, B. (2017). Rediscovery of the doldrums in stormresolving simulations over the tropical Atlantic. Nature Geoscience, 10, 891–896. https://doi.org/10.1038/ngeo3140

Li, F., Rosa, D., Collins, W. D., & Wehner, M. F. (2012). Superparameterization: A better way to simulate regional extreme precipitation? Journal of Advances in Modeling Earth Systems, 4, M04002. https://doi.org/10.1029/2011MS000106

Liebmann, B., & Smith, C. A. (1996). Description of a complete (interpolated) outgoing longwave radiation dataset. Bulletin of the American Meteorological Society, 77, 1275–1277. https://doi.org/10.1175/1520-0477-77.111

Lin, I. (2007). The double-ITCZ problem in IPCC AR4 Coupled GCMs: Ocean–atmosphere feedback analysis. Journal of Climate, 20, 4497–4525. https://doi.org/10.1175/jcli4272.1

Lin, J., & Neelin, J. D. (2002). Considerations for stochastic convective parameterization. Journal of the Atmospheric Sciences, 59(5), 959–975. https://doi.org/10.1175/1520-0469(2002)059<0959:csfcp>2.0.co2

Lin, J., & Neelin, J. D. (2003). Toward stochastic deep convective parameterization in general circulation models. Geophysical Research Letters, 30(4), 1162. https://doi.org/10.1029/2002GL016203

Ma, L., Peters, K., Wang, B., & Li, J. (2019). Revisiting the impact of Stochastic Multicloud Model on the MJO using low-resolution ECHAM6.3 atmosphere model. Journal of the Meteorological Society Japan Series II, 97, 977–993. https://doi.org/10.2151/jmsj.2019-053

Moncrieff, M. W., & Klinker, E. (1997). Organized convective systems in the tropical western Pacific as a process in general circulation models: A Toga-COARE case study. Quarterly Journal of Royal Meteorological Society, 123, 805–827. https://doi.org/10.1002/qj.49712354002

Palmer, T. N. (2001). A nonlinear dynamical perspective on model error: A proposal for non-local stochastic-dynamic parametrization in weather and climate prediction models. Quarterly Journal of the Royal Meteorological Society, 127, 279–304. https://doi.org/10.1002/qj.49712375702

Peters, K., Crueger, T., Jakob, C., & Möbis, B. (2017). Improved MJO-simulation in ECHAM6.3 by coupling a stochastic multicloud model to the convection scheme. Journal of Advances in Modeling Earth Systems, 9, 193–219. https://doi.org/10.1002/2016MS000809
Peters, K., Hohenegger, C., & Klocke, D. (2019). Different representation of mesoscale convective systems in convection-permitting and convection-parameterizing NWP models and its implications for large-scale forecast evolution. *Atmosphere, 10*, 503. https://doi.org/10.3390/atmos10090503

Peters, K., Jakob, C., Davies, L., Khroulanguage, B., & Majda, A. J. (2013). Stochastic behavior of tropical convection in observations and a Multicloud Model. *Journal of the Atmospheric Sciences, 70*, 3558–3575. https://doi.org/10.1175/JAS-D-13-031.1

Randall, D. A., Wood, R. A., Bony, S., Colman, R., Fichefet, T., Fyfe, J., et al. (2007). Climate models and their evaluation. *Climate Change 2007: The physical science basis*. In In Solomon, S., et al. (Eds.), *Contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change*. United Kingdom and New York, NY: Cambridge University Press.

Rao, S. A., Goswami, B. N., Sahai, A. K., Rajagopal, E. N., Mukhopadhyay, P., Rajeevan, M., et al. (2019). Monsoon mission: A targeted activity to improve monsoon prediction across scales. *Bulletin of the American Meteorological Society, 100*(12), 2509–2532.

Roy, K., Mukhopadhyay, P., Krishna, R. M., Ganai, M., Mahakur, M., Rao, T. N., et al. (2020). Assessment of climate models in relation to the low-level clouds over the southern Indian Ocean. *Quarterly Journal of the Royal Meteorological Society, 146*, 1–3325. https://doi.org/10.1002/qj.3847

Saha, S., Moorthi, S., Pan, H.-L., Wu, X., Wang, J., Nadiga, S., et al. (2010). The NCEP climate forecast system reanalysis. *Bulletin of the American Meteorological Society, 91*, 1015–1058. https://doi.org/10.1175/2010Bams3001.1

Saha, S., Moorthi, S., Wu, S., Wang, X. J., Nadiga, S., Tripp, P., et al. (2014). The NCEP climate forecast system version 2. *Journal of Climate, 27*(6), 2185–2208. https://doi.org/10.1175/JCLI-D-12-00823.1

Satoh, M., Stevens, B., Judi, F., Khairoutdinov, M., Lin, S.-J., Putman, W. M., & Du’ben, P. (2019). Global cloud-resolving models. *Current Climate Change Reports, 5*, 172–184. https://doi.org/10.1007/s40641-019-00131-0

Senior, C. A., & Mitchell, J. F. B. (1993). CO and climate: The impact of cloud parametrizations. *Journal of Climate, 6*, 393–418. https://doi.org/10.1175/1520-0442(1993)006<0393:CDACTI>2.0.CO;2

Siebesma, A. P., Jakob, C., Lenderink, G., Neggers, R. A. J., Teixeira, J., van Meijgaard, E., et al. (2004). Cloud representation in general circulation models over the northern Pacific Ocean: A EUROCS intercomparison study. *Quarterly Journal of the Royal Meteorological Society, 130*, 3245–3267. https://doi.org/10.10125/qj.03.146

Stephens, G. L. (2005). Cloud feedbacks in the climate system: A critical review. *Journal of Climate, 18*, 237–273. https://doi.org/10.1175/1jcli4324.1

Stevens, B., Acquistapace, C., Hansen, A., Heinze, R., Klingler, C., Klocke, D., et al. (2020). The added value of large-eddy and storm-resolving models for simulating clouds and precipitation. *Journal of Meteorological Society of Japan Ser II, 98*, 395–435. https://doi.org/10.2151/jmsj.2020-021

Stevens, B., & Bony, S. (2013). What are climate models missing? *Science, 340*, 1053–1054. https://doi.org/10.1126/science.1237554

Stevens, B., Satoh, M., Auger, L., Biercamp, J., Bretherton, C. S., Chen, X., et al. (2019). DYAMOND: The dynamics of the atmospheric general circulation modeled on non-hydrostatic domains. *Progress in Earth and Planetary Science, 6*, 1–17. https://doi.org/10.1186/s40645-019-0304-z

Xie, S.-P., & Philander, S. G. H. (1994). A coupled ocean-atmosphere model of relevance to the ITCCZ in the eastern Pacific. *Tellus, 46A*, 340–350. https://doi.org/10.1034/j.1600-0870.1994.01-1-00001.x

Zhang, C. (2005). Madden–Julian oscillation. *Reviews of Geophysics, 43*, RG2003. https://doi.org/10.1029/2004RG000158

Zhao, M., Golaz, J.-C., Held, I. M., Ramaswamy, V., Lin, S. J., Ming, Y., et al. (2016). Uncertainty in model climate sensitivity traced to representations of cumulus precipitation microphysics. *Journal of Climate, 29*, 543–560. https://doi.org/10.1175/JCLI-D-15-0191.1