Recent Challenges in the APCC Multi-Model Ensemble Seasonal Prediction: Hindcast Period Issue

Young-Mi Min[1], Chang-Mook Im[1], V. N. Kryjov[2], and Daeun Jeong[1]

[1]APEC Climate Center, Busan, Republic of Korea

[2]Busan National University, Busan, Republic of Korea

Corresponding author: Young-Mi Min (ymmin@apcc21.org)

Key Points:

- APCC, which combines all information from different ensemble prediction systems, recently faced challenges on the hindcast period issue
- Proposed solution leads to increase total number of models contributing to MME prediction, particularly recently developed newer models
- It shows improved skills in both temperature and precipitation prediction over most of the globe and seasons

Abstract

Seasonal forecasts are commonly issued in the form of anomalies as departures from average in a specified multiyear reference period (climatology). The model climatology is estimated as average of retrospective forecasts over the hindcast period. However, different operational centers providing seasonal ensemble predictions use different hindcast periods based on their own model climatology. In addition, the hindcast period of recently developed/upgraded newer models tends to shift to the recent years. In this paper, we discuss recent challenges faced by the APCC multi-model ensemble (MME) operations, especially changes in the hindcast period for individual models. Based on the results of various sensitivity experiments for the MME prediction, we proposed to change the hindcast period that is the most appropriate solution for the APCC operations. It makes the newly developed models join the MME and increase the total number of participating models, which facilitates the skill improvement of the MME prediction.

Plain Language Summary

In seasonal forecasting, it is well known that the MME, which combines different single-model prediction from various operational and research centers, is more effective way to improve the forecast skill. Since 2005, the APCC has been providing the MME seasonal forecasts and the models participating in the APCC MME operations have been continuously changing. Particularly, as the hindcast period of newly developed models is shifted to the latest, they cannot participate in the operational MME forecasts because of discrepancy in climatologies. However, over time, as the number of new models expected to have skillful forecasts gradually increases, APCC faced the challenge of continuously reducing the number of participating models or changing the hindcast period to the more recent years. Considering various aspects such as the number of participating models, skills, and climatology period, we have selected the most appropriate way for the APCC operations. Through this, the MME prediction skill has been improved over most of the globe and seasons due to increasing of the number of participating models, particularly inclusion of newer models.

1 Introduction

Seasonal forecasts are commonly expressed in terms of anomalies as departures from a climatological mean and/or probabilities of an event occurring with respect to a climatological distribution (usually, tercile-based categorical forecasts). This allows users to see if the predicted seasonal mean variables are anomalously positive or negative in respect to climatological means, and/or what probability of the events (e.g., above, near, or below-normal category) is expected. So, climatology is used as a benchmark or reference against which expected conditions to be likely experienced. It also provides a way for removing systematic biases in forecasts from dynamical prediction systems by subtracting model climatology because they are not perfect representations of the real world (Stockdale, 1997; Kumar et al., 2012). The model climatology is estimated using retrospective forecasts (hindcasts) over a specified long-term reference period.

World Meteorological Organization (WMO) recommends climatology (normals) to be estimated as 30-year averages computed for the most-recent 30-year period finishing in a year ending with 0 (WMO, 2007), i.e., 1991-2020 at present. National meteorological and hydrological services (NMHs) estimate forecasts as departures from these 30-year normals in their locations. However, different operational and research centers use different hindcast periods for estimation of the model climatology. Furthermore, hindcast periods of recently developed and improved climate models, particularly beginning of the hindcast period, tends to be shifted to recent years. The Asia-Pacific Economic Cooperation (APEC) Climate Center (APCC) is one of major operational centers providing well-validated multi-model ensemble (MME) seasonal forecasts. Since its establishment in 2005, APCC has been collecting dynamical ensemble forecasts through multi-institutional cooperation and coordinating the MME prediction. At present, 15 leading operational and research institutes from 11 countries are involved in the APCC operational MME prediction. The MME operational centers, e.g., APCC (Min et al., 2014, 2017), the WMO Lead Center (WMOLC; Kim et al., 2021), the North American MME (NMME; Becker et al., 2014; Kirtman et al., 2014), and the Copernicus Climate Change Service (C3S; Manazanas et al., 2019) use a common hindcast period of all participating models, which results in a relatively short period compared to that of single-model prediction systems.

As the hindcast period for recently developed newer models has gradually shifted to the later years, the full range of hindcast periods for the dynamical models routinely running in operational centers is widening, from early-1980s to late-2010s nowadays. However, the common hindcast period is rather shortening due to shift of the newer models' hindcast period beginning to the early 1990s. This raised new issue at APCC that combines all information from different climate prediction systems, particularly in 2019. This is because some of contributing models in the APCC MME prediction were changed to their upgraded versions in 2020, and their hindcast periods were shifted to more recent years. That is, with implementation of new models, if the common hindcast period, 1983-2010, were kept, the number of participating models in MME would have been reduced and would be gradually reducing in future because recently developed models that are expected to have better skill do not match with this common period. It might lead to deterioration of the MME prediction skill. Therefore, the time has come for APCC to consider the issue on the hindcast period that could affect the number of participating models in MME and eventually the MME skill. In this study, we discuss challenges faced by MME operations caused by upgrading of the participating model set. In particular, we focus on the issue of a decrease in the number of participating models in the MME prediction with a shift to the later years of the hindcast periods of recently developed models. We suggest most appropriate solution for APCC operations based on several sensitivity experiments on the different hindcast periods and different number of participating models in MME.

2 Method and Data

With the most recent joining of SYS8 from METFR, APCC currently collects the ensemble predictions from 15 state-of-the-art climate models, and the models are being continuously improved with the great efforts of their own operational and research centers. The APCC MME model suits of 2019 and 2020 are listed in Table 1. In 2019, the real-time operational MME prediction suit comprised eight models from APCC, BOM, CWB, JMA, MSC/ECCC, NASA, NCEP, and PNU that matched with the common hindcast period of 1983-2010. However, APCC was not able to involve recently upgraded models from CMCC, KMA, and UKMO in the real-time operations due to the mismatch of the hindcast period. Furthermore, several models were scheduled to be changed to their upgraded versions in 2020. To test the sensitivity in terms of predictability as the participating models in MME change due to their improvements, we performed several experiments with varying reference periods and contributing models in MME. Here, the MME forecast is a simple average of individual models with equal weights.

We focus on 1-month lead 3-month mean (seasonal) MME forecasts of 2m temperature and precipitation over globe and sub-regions: Northern Extratropics (NE; 20°N-90°N), Southern Extratropics (SE; 20°S-90°S), Tropics (TR; 20°N-20°S), East Asia (EAs; 75°E-150°E, 15°N-60°N), South Asia (SAs; 60°E-140°E, 10°S-35°N), North America (NAm; 190°E-310°E, 10°N-75°N), South America (SAm; 270E-330E, 60°S-10°N), Australia (Aus; 110E-180E, 50°S-0°N), and Northen Eurasia (NE; 25°E-190°E, 40°N-80°N). For skill assessment, we use the National Center for Environmental Prediction (NCEP)-Department of Energy (DOE) Reanalysis 2 data (Kanamitsu et al., 2002) for temperature and the Climate Anomaly System and Outgoing longwave radiation Prediction Index data (CAMS-OPI, Janowiak and Xie, 1999) for precipitation. All model forecasts and observations are interpolated onto a 2.5 x 2.5 common grid. We use the anomaly pattern correlation coefficient (ACC) and temporal correlation coefficient (TCC) to assess the prediction skill. We use the ACC-based skill score (Murphey, 1988) for assessment of the prediction skill improvement and deterioration of the MME forecast with another model set compared to the reference model set. The Student's t-test and the Mann-Kendall test is used to assess the statistical significance of the difference between means and trends.

3 Results

A history of operational exploiting of the seasonal prediction models exceeds two decades. It has been marked by numerous upgrades of models, resolution, ensemble size, lead time, etc. Operational long-range forecasting centers perform essential efforts to improve their climate prediction system. Particularly, they tend to extend the period of hindcasts over which climatology is estimated and to move it to more recent years. As shown in Fig. 1(a), the number of models providing their ensemble forecasts to APCC and contributing to the real-time APCC MME predictions vary from year to year, depending on the operational environmental at the time of the given forecast issuing. Here, the proportion of models that do not participate in the real-time MME prediction has been gradually increasing and it was expected to increase up to nearly 50% in 2020 (red line in Fig. 1(a)). It is noteworthy that about 80% of these models do not participate in real-time MME prediction because of inconsistency with the common hindcast period, 1983-2010 (black line in Fig. 1(a)). That is, model developers continue to improve their models and gradually move their hindcast periods to more recent years, however, with the current 1983-2010 MME hindcast remaining unchanged, the number of the models participating in the APCC real-time operations matching the current MME hindcast period will gradually decreases.

Table 1. Evolved models in the real-time APCC MME prediction in 2019 (v2019) and in 2020 (v2020)

	v2019		v2020	
Institute	Model Name	Hindcast Period	Model Name	Hindcast Period
APCC	**SCoPS**	1982-3013	SCoPS	1982-3013
BCC	CSM_1.1m	1991-2015	CSM_1.1m	1991-2015
BOM	**POAMA**	1983-2011	ACCESS-S	1990-2012
CMCC	SPSv2	1993-2016	SPSv3	1993-2016
CWB	**GFST119**	1982-2011	GFST119	1982-2011
HMC	SL-AV	1985-2010	SL-AV	1985-2010
JMA	**MRI-CPS2**	1979-2014	MRI-CPS2	1979-2014
KMA	GloSea5GC2	1991-2010	GloSea5GC2	1991-2016
MGO*	MGOAM-2	1982-2013	MGOAM-2	1982-2013
MSC/ECCC	**CanSIP**	1981-2010	CanSIPSv2	1981-2010
NASA	**GEOS-S2S-2**	1981-2016	GEOS-S2S-2	1981-2016
NCEP	**CFSv2**	1982-2010	CFSv2	1982-2010
PNU	**CGCMv1.0**	1980-2018	CGCMv1.0	1980-2018
UKMO	GloSea5	1993-2016	GloSea5	1991-2016

Bold in 2019 indicates the models contributing to the real-time APCC MME operation.

* Not participating model due to the inconsistency of the hindcast experiment (i.e., the AMIP-type)

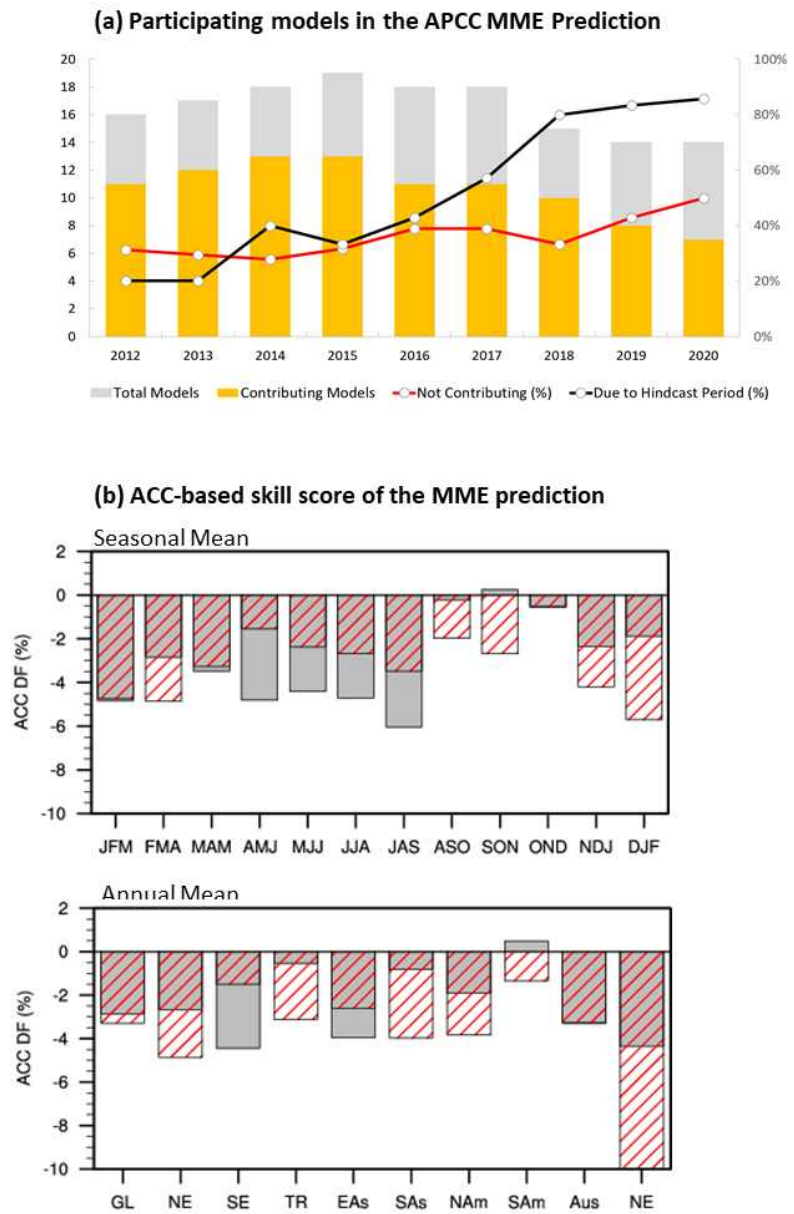

Figure 1. (a) Changes in the number of models providing their seasonal forecasts to APCC (grey bar) and participating in the real-time operational APCC MME prediction (yellow bar) for the period 2012-2019 and expected changes in 2020. Lines indicate the proportion of models that do not participate in the MME operation (red, %) and the proportion of the not-participating models due to the common hindcast period mismatch (black, %). (b) ACC-based skill score of the MME predictions comprising forecasts from seven models of v2020 in respect to that comprising eight models of v2019 for the period of 1983-2010 of seasonal mean temperature and precipitation forecasts over globe and annual mean for several sub-regions.

The more important issue is the MME skill that is affected by the mean skill of individual models and models' diversity (Yoo and Kang 2005; Alessandri et al., 2018). If the number of contributing models in the MME predictions continues to reduce, particularly by excluding of recently developed and improved newer models, it may result in the MME skill decreases. Fig. 1(b) shows the ACC-based skill score of the MME predictions comprising forecasts from seven models of v2020 in respect to that comprising eight models of v2019 (Table 1) selected under condition that the MME system matches the common 28-year hindcast period (1983-2010). The skill score is mainly negative that indicates deterioration in the MME skill caused by the decrease in the number of participating models even by one. The hindcast skill of both global temperature and precipitation using v2020 decreases across almost all seasons. It is also true for sub-regions in terms of annual means, with exception of temperature in South America. Thus, if we keep the 28-year long climatology period of 1983-2010, we will not be able to incorporate recently developed/upgraded models to the MME prediction system and as a result the MME prediction skill may continue to decrease along with decrease in the number of the models participating in real-time MME operations.

APCC considered several solutions to solve this hindcast issue and to take advantage of a large set of models participating in the MME prediction. The first solution would be the use of forecast anomalies with respect to climatologies estimated over the models' own hindcast periods, which varies among the groups producing the model forecast, like IRI ENSO forecast. That is, all the models can participate in the MME prediction by using forecast anomalies with respect to different base periods and, consequently, to the different climatologies. However, discrepancies may arise if the climatologies are significantly different. We have assessed significance of the difference between climatologies estimated over two periods, 1983-2010 and 1993-2016, which cover the common hindcast period and the most recent hindcast period of the 14 models in v2019, based on the Student's *t*-test (Fig. S1). The results show that the differences between the two climatologies of seasonal mean temperature in observations are statistically significant in many regions. The most significant differences are evident in the high latitudes of Northern and Southern Hemispheres throughout all seasons. These are the regions of the mostly pronounced global warming accelerating in the recent years. It is also evident in the South Indian Ocean in MAM and JJA and in the Western Pacific in SON and DJF. Furthermore, for the model with the longest hindcast period spanning from the early 1980s to the most recent years the differences between climatologies from the periods 1983-2010 and 1993-2016 are statistically significant (not shown). Thus, the first solution may cause another issue in forecast anomalies due to the significant differences in the climatologies because of the different reference (hindcast) periods of individual models, and eventually in the MME prediction that combine the forecast anomalies of individual models (Wallace and Arribas, 2012). Furthermore, this solution does not suit the users of our seasonal forecasts. The users formulate their local forecasts in terms of anomalies in respect to their local normals estimated over the 30-year periods appointed/defined by WMO. As a rule, they make their local corrections to the MME forecasts accounting for the difference between the local normals estimated over, e.g., 1991-2020 and local MME climatology estimated over, e.g., 1983-2010. However, this solution does not provide distinctness of the period of MME climatology that makes impossible any corrections. Particularly, another issue arises as to which a reference period should be applied to observation to assess the MME forecasts combined with models using different climatologies.

In this situation, we suggested an alternative solution that is to change the current hindcast period to unified 1991-2010, for which almost all of the models are included. Although, the 20-year climatology is shorter, it is comparable with the climatologies of other MME groups (e.g., WMOLC (1993-2009; 17 years), C3S (1993-2016; 24 years). In case of models of CMCC and HMC whose data start from 1993, it was treated as missing values of 1991-1992 to allow more models to participate in MME and to extend the common hindcast period. To estimate the forecast skill according to the changes in the number of participating models, we examine the skill of MME hindcast (1991-2010) in three different model combinations within model suites of 2020 (v2020). Those are the seven models matched with the 1983-2010 hindcast period (+7M; APCC, CWB, JMA, NASA, NCEP, PNU, MSC/ECCC); additional six models matched with the 1991-2010 (+6M; BCCM, HMC, KMA, UKMO, BOM, CMCC); the whole model set comprising all 13 models(13M). The diagrams shown in Fig. 2 demonstrate that the skills of the MMEs based on 7M (MME_7M) and +6M (MME_+6M) are comparable with each other, showing ACC=0.36 (0.44) for annual mean temperature (precipitation) for both MMEs. Meanwhile, the MME comprising all 13 models (MME_13M) definitely outperforms both MME_7M and MME_+6M for both temperature and precipitation and for all 12 seasons throughout the year. The skill improvement of MME_13M forecasts as compared with MME_7M for both variables appears in most oceans and lands, only except for precipitation in the Arctic region (Fig. 3). This improvement is mainly due to increase in the number of models and corresponding increase of diversity of the contributing models and the relatively high skills of newly contributing models (Yoo and Kang 2005; Alessandri et al., 2018; Fig. S2).

Based on the results from the hindcast experiments, we have changed the common base period to 1991-2010 for APCC MME operations since 2020, which is covered by almost all the models (Oper). Finally, we assessed the MME skill of the real-time forecasts for 2020JFM-2022DJF because it is important for operational centers to assess whether the skill improvement exists in real-time forecast as well as hindcast, although the 3-year periods is too short for the collection of a sufficient number of real-time forecasts to obtain some well-grounded conclusions. For comparison, we produced the real-time MME forecasts for the period of 2020-22 with the models (Exp) that could participate if the hindcast period has not been changed. By changing the hindcast period to 1991-2010, the number of participating models in the real-time MME operations in 2020-22 increased by 100%, and recently the difference between Oper and Exp has gradually widened (Fig. 4a). As a result, substantial improvement in temperature over the globe is observed in recent years and the skill increases by approximately 2.8% for temperature (Fig. 4b). However, precipitation shows little change in a global scale. Based on the results from hindcast and real-time forecasts, it turned out that the change of common hindcast period for MME prediction in 2020 was an appropriate action for APCC operations.

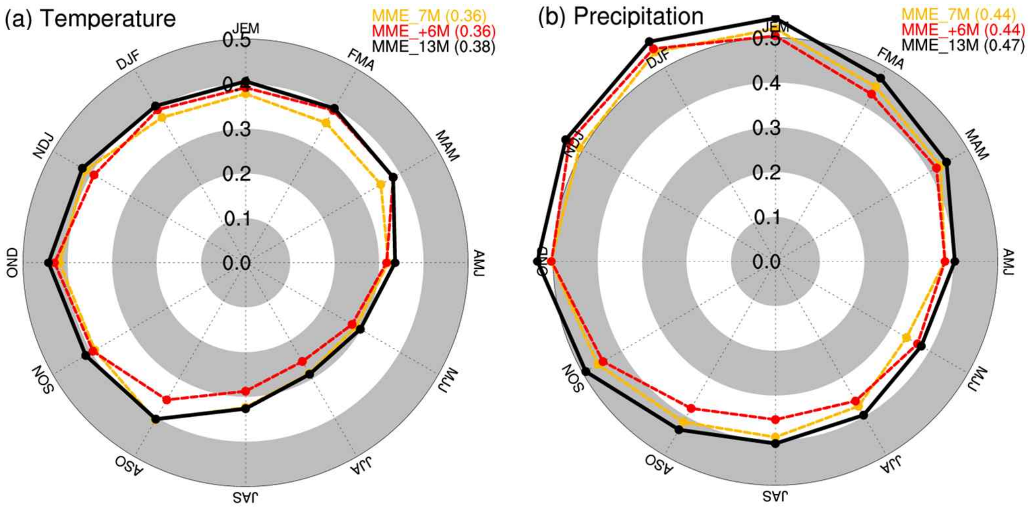

Figure 2. ACC for MME hindcast (1991-2010) with 7 models, additionally contributing 6 models, and with 13 models of (a) seasonal mean temperature and (b) precipitation over globe. Annual mean ACC for each MME is shown in parentheses.

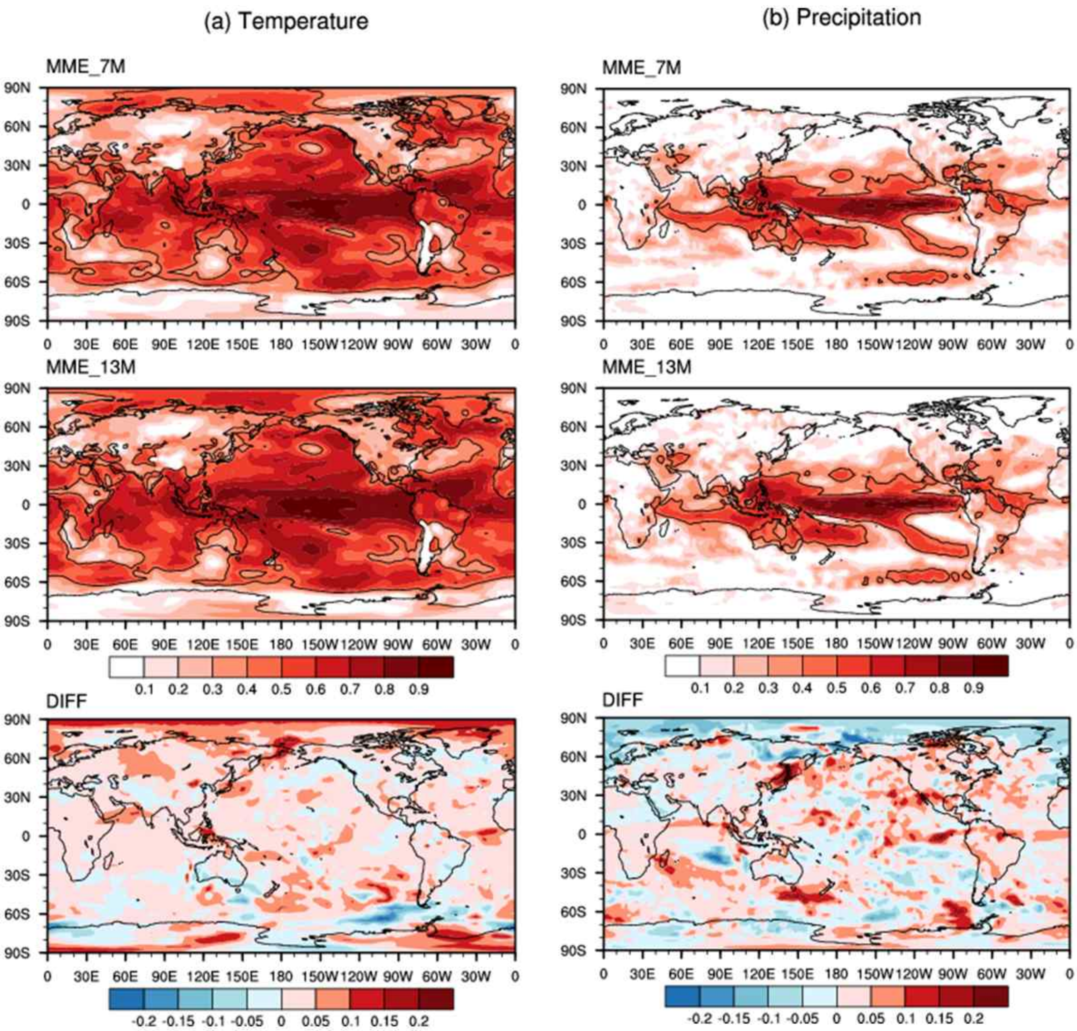

Figure 3. Spatial distribution of 12-season averaged (annual mean) temporal correlation coefficients (TCCs) for MME hindcast (1991-2010) with 7 models and with 13 models of (a) seasonal mean temperature and (b) precipitation. The contour lines enclose the areas where the TCCs are statistically significant at the 5% level using the two-tailed Student's t-test. The skill differences (DIFF) shows the difference two MMEs (MME_13M minus MME_7M).

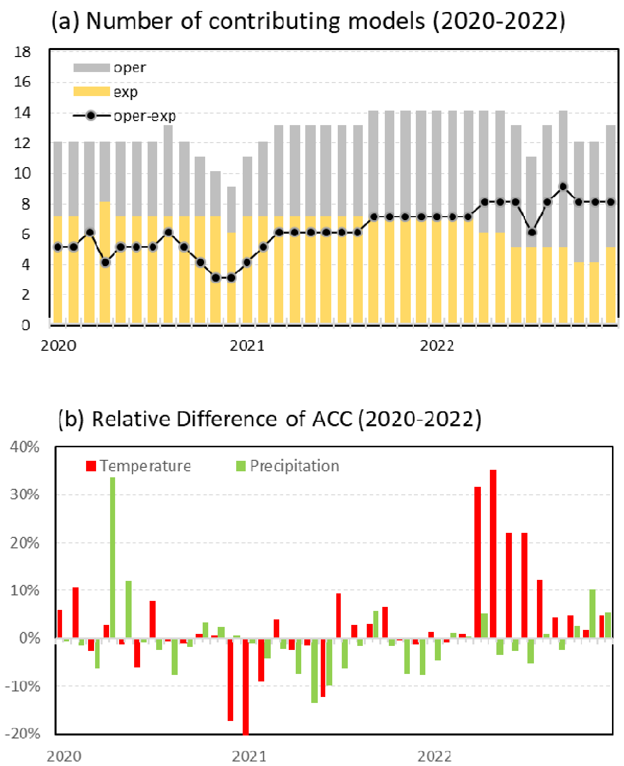

Figure 4. (a) Number of participating models in real-time MME operations for 2020-22 and ratio of its change from Exp to Oper. (b) Relative difference of ACC from Exp to Oper for global temperature and precipitation.

5 Conclusions

Construction of the MME is a compromise between the number of participating models and length of a common hindcast period. Increase in the number of participating models with sufficient model diversity provided decreases random and model formulation errors. On the other hand, increase in the length of the common hindcast period decreases errors in climatology but increase random and model formulation errors in MME forecasts because of decrease in the number of participating models. In this situation, permanent development and improvement of the newly developed/improved models, being a sign of progress, is gradually destroying existing compromise and requires achievement of a new compromise with participation of the newer models. In the early of 2020s, APCC faced new challenges while maintaining a common hindcast period exceeding 20 years that maintained for many years, resulting that the proportion of the models that could not participate in the operational MME prediction due to their hindcast periods started in the late 1980s-early 1990s achieved about 50% by 2020. Based on the results of several experiments, we proposed a solution to change the common hindcast period to 1991-2010, which is the most appropriate way for APCC operation, reflecting the recently developed models. That is, by changing the reference period for MME prediction, APCC provides opportunities for participation in operational MME prediction system for the newly developed/upgraded models, which results in a twofold increase in the number of participating models and improvement of the forecast skill. Certainly, the payment for the increase in the number of the participating models is shortening of the common hindcast period down to 20 years, which is less than 30 years recommended by WMO for climatology estimation but which is comparable with other MME prediction producing centers, e.g., WMOLC and C3S.

Although not the scope of this study, the most important issue in recent years is that since late 2021, the NMHSs worldwide have used the WMO recommended 1991-2020 normals (https://www.wmo.int/edistrib_exped/grp_prs/_en/08791-2019-CLW-CLPA-DMA-CLIN8110_en.pdf). Meanwhile, there are still some limitations to match with the WMO recommended period of normal and nowadays, no climate center providing MME seasonal forecasts to the NHMSs uses climatology matching with the WMO references. The NMHSs needs to adjust the model predicted departures from the model climatologies to departures from the WMO normals. So, adjustment of the period of the MME climatology to the period of the WMO normals would be one of the forthcoming research tasks.

Acknowledgments

This research was supported by APEC Climate Center. The authors acknowledge the APCC MME Producing Centres (PCs) for making their hindcast/forecast data available for analysis and the APCC for collecting and archiving them and for organizing APCC MME prediction. We also thanks to the APCC MME PCs to discuss this issue together at the 3rd APCC MME Providers' meeting in 2019.

Data Availability Statement

The single-model and MME predictions used in the study are available at the platform-based climate data service, CLimate Information toolkit (CLIKs; https://cliks.apcc21.org). The NCEP-DOE reanalysis 2 was obtained from the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their website (https://psl.noaa.gov/data/gridded/index.html). The CAMS OPI monthly estimates were available at https://www.cpc.ncep.noaa.gov/products/global_precip/html/wpage.cams_opi.html.

References

Alessandri, A., Catalano, F., Lee, J. Y., Wang, B., Lee, D. Y., Yoo, J. H., & Weisheimer, A. (2018). Grand European and Asian-Pacific multi-model seasonal forecasts: maximization of skill and of potential economical value to end-users. *Climate Dynamics*, *50*, 2719-2738. https://doi.org/10.1007/s00382-017-3766-y

Becker, E., den Dool, H. V., & Zhang, Q. (2014). Predictability and forecast skill in NMME. *Journal of Climate*, *27*(15), 5891-5906. https://doi.org/10.1175/JCLI-D-13-00597.1

Janowiak, J. E,, & Xie, P. (1999). CAMS_OPI: a global satellite-raingauge merged product for real-time precipitation monitoring applications. *Journal of Climatology*, 12:3335-3342. https://doi.org/10.1175/1520-0442(1999)012<3335:COAGSR>2.0.CO;2

Kanamitsu, M., Ebisuzaki, W., Woollen, J., Yang, S.-K., Hnilo, J. J., Fiorino, M., & Potter, G. L. (2002). NCEP-DOE AMIP-II Reanalysis (R-2). *Bulletin of the American Meteorological Society*, 83:1631-1643. https://doi.org/10.1175/Bams-83-11-1631

Kim, G., Ahn, J. B., Kryjov, V. N., Lee, W. S., Kim, D. J., and Kumar, A (2021). Assessment of multi-model ensemble methods for WMO LC-LRFMME. *International Journal of Climatology, 41*, E2462-E248. https://doi: 10.1002/joc.6858.

Kirtman, B. P., Min, D., Infanti, J. M., et al. (2014). The North American Multimodel Ensemble: Phase-1 Seasonal-to-Interannual Prediction; Phase-2 toward Developing Intraseasonal Prediction. *Bulletin of the American Meteorological Society*, *95*(4), 585-601. https://doi.org/10.1175/BAMS-D-12-00050.1

Kryjov, V. N., & Min, Y.-M. (2016). Predictability of the wintertime Arctic Oscillation based on autumn circulation. *International Journal of Climatology*, *36*, 4181–4186, https://doi.org/10.1002/joc.4616

Kumar, A., Chen, M., Zhang, L., Yang, W., Xue, Y., Wen, C., Marx, L., & Huang, B. (2012). An analysis of the nonstationarity in the bias of sea surface temperature forecasts for the NCEP Climate Forecast System (CFS) version 2. *Monthly Weather Review*, *140*, 3003-3016. https://doi.org/10/1175/MWR-D-11-00335.1

Lim, J., Dunstone, N. J., Scaife, A. A., and Smith D. M., (2019). Skillful seasonal prediction of Korean winter temperature, *Atmospheric Science Letter*, *20*, https://doi.org/10.1002/asl.881

Manzanas, R., Gutierrez, J. M., Bhend, K., Hemri, S., Doblas-Reyes, F. J., Torralba, V., Penabad, E., & Brookshaw, A. (2019). Bias adjustment and ensemble recalibration methods for seasonal forecasting: a comprehensive intercomparison using the C3S dataset. *Climate Dynamics*, 53:1287-1305. https://doi.org/10.1007/s00382-019-04640-4

Min, Y. M., Kryjov, V. N., & Oh, S. M. (2014). Assessment of APCC multimodel ensemble prediction in seasonal climate forecasting: Retrospective (1983-2003) and real-time forecasts (2008-2013). *Journal of Geophysical Research: Atmospheres*, 199, 12,132-12,150. https://doi.org/10.1002/2014JD022230

Min, Y. M., Kryjov, V. N., Oh, S. M., & Lee, H. J. (2017) Skill of real-time operational forecasts with the APCC multi-model ensemble prediction system during the period 2008-2015. *Climate Dynamics*, 49(11–12), 4141–4156. https://doi.org/10.1007/s00382-017-3576-2

Murphy, A. H. (1988). Skill scores based on the mean squared error and their relationships to the correlation coefficient. *Mon.Weather Rev*., *116*, 2417–24. https://doi.org/10.1175/1520-0493(1988)116<2417:SSBOTM>2.0.CO;2

Stockdale, T. N. (1997). Coupled ocean-atmosphere forecasts in the presence of climate drift. *Monthly Weather Review*, 235, 809-818. https://doi.org/10.1175/1520-0493(1997)125<0809:COAFIT>2.0.CO;2

Wallace, E., Arribas, A. (2012). Forecasting with reference to a specific climatology. *Monthly Weather Review*, 140, 3795-3802. https://doi.org/10.1175/MWR-D-12-00159.1

World Meteorological Organization (2007). *The role of climatological normals in a changing climate*. WCDMP-No. 61, WMO/TD-No. 1377, Geneva, World Meteorological Organization

Yoo, J, H., & Kang, I. S. (2005). Theoretical examination of a multi-model composite for seasonal prediction. *Geophysical Research Letter*, *32*, L18707. https://doi.org/10.1029/2005GL023513

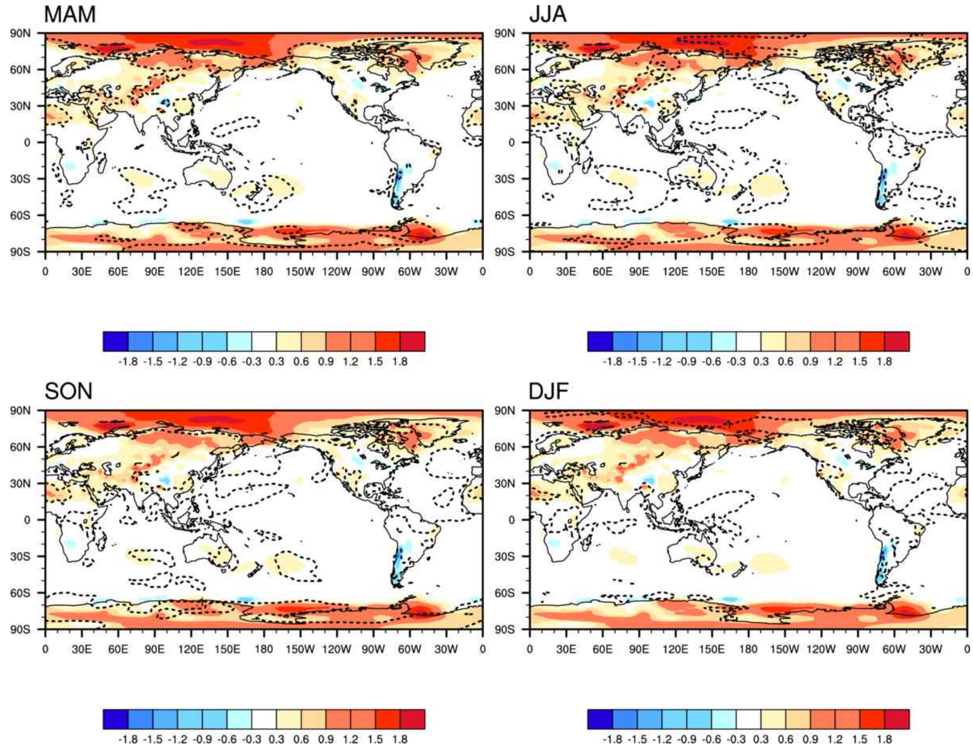

Figure S1. Differences between two climatologies over the period of 1979-2014 and 1993-2016 (dashed line) and trends of seasonal mean temperature for the whole 37-year period (shading). Differences and trends are only displayed at a 10% significance level.

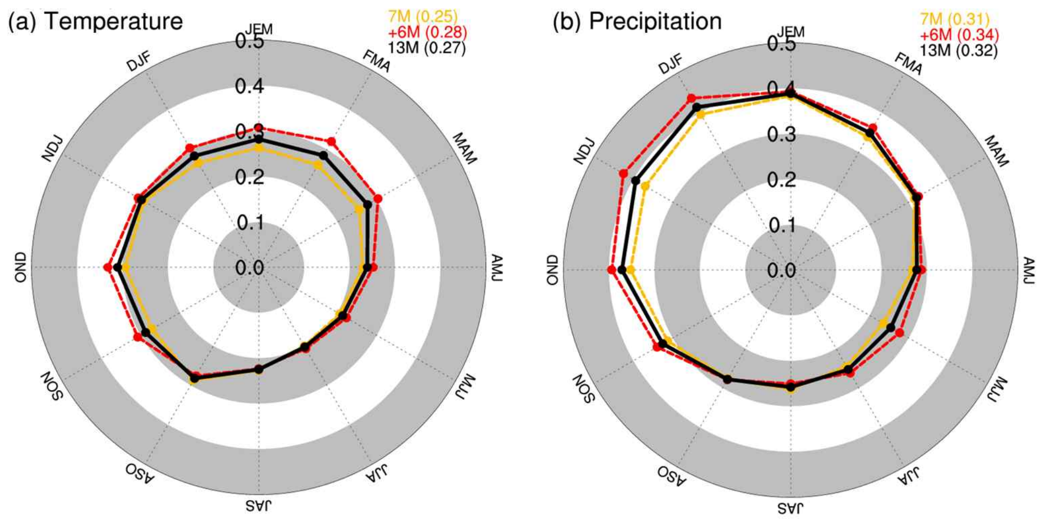

Figure S2. The averaged ACC for hindcast (1991-2010) with 7 models, additionally contributing 6 models, and all 13 models of (a) seasonal mean temperature and (b) precipitation over globe. Annual mean ACC for the average of individual models' skill is shown in paretheses.