Improving Predictions of Surface Air Temperature Anomalies over Japan by the Selective Ensemble Mean Technique

J. V. RATNAM, a TAKESHI DOI, a YUSHI MORIOKA, a PASCAL OETTLI, a MASAMI NONAKA, a AND SWADHIN K. BEHERA a

a Application Laboratory, Japan Agency for Marine-Earth Science and Technology, Yokohama, Japan

(Manuscript received 21 June 2020, in final form 20 November 2020)

ABSTRACT: The selective ensemble mean (SEM) technique is applied to the late spring and summer months (May–August) surface air temperature anomaly predictions of the Scale Interaction Experiment–Frontier Research Center for Global Change, version 2 (SINTEX-F2), coupled general circulation model over Japan. Using the Köppen–Geiger climatic classification we chose four regions over Japan for applying the SEM technique. The SINTEX-F2 ensemble members for the SEM are chosen based on the anomaly correlation coefficients (ACC) of the SINTEX-F2 predicted and observed surface air temperature anomalies. The SEM technique is applied to generate the forecasts of the surface air temperature anomalies for the period 1983–2018 using the selected members. Analysis shows the ACC skill score of the SEM prediction to be higher compared to the ACC skill score of predictions obtained by averaging all the 24 members of the SINTEX-F2 (ENSMMEAN). The SEM predicted surface air temperature anomalies also have higher hit rate and lower false alarm rate compared to the ENSMEAN predicted anomalies over a range of temperature anomalies. The results indicate the SEM technique to be a simple and easy to apply method to improve the SINTEX-F2 predictions of surface air temperature anomalies over Japan. The better performance of the SEM in generating the surface air temperature anomalies can be partly attributed to realistic prediction of 850-hPa geopotential height anomalies over Japan.

KEYWORDS: Climate prediction; Forecasting techniques; Seasonal forecasting

1. Introduction

Japan experiences a hot and humid climate in the summer season. The combination of high temperatures and humidity in the season increases the wet bulb temperatures, enhancing the likelihood of heatstroke and related illnesses (Akihiko et al. 2015). Predicting surface air temperatures over Japan well in advance would benefit planners and society as a whole. Numerical models are often used to predict temperatures at seasonal time scales. Usually an ensemble of model predictions with varying initial and boundary conditions is generated to quantify the uncertainty in the predictions. The mean of the ensemble of the model predictions is taken as a more reliable prediction compared to the prediction by individual members of the ensemble (Palmer and Anderson 1994). However, there is a likelihood of some members of the ensemble performing better than others in predicting the surface air temperatures over a region. Identifying and generating a mean of those members of the ensemble may enhance the skill of the regional predictions of surface air temperatures (SAT). This technique of selectively averaging the members of a seasonal forecasting system to improve predictions is called the selective ensemble mean (SEM; Qi et al. 2014; Nishimura and Yamaguchi 2015). This technique is similar to the method often adopted by researchers in choosing a particular model from a large number of Coupled Model Intercomparison Project (CMIP) models for analysis (Sabeerali et al. 2013).

Qi et al. (2014) and Nishimura and Yamaguchi (2015) applied the SEM technique to improve the predictions of the tracks of tropical cyclones of an ensemble prediction system. The ensemble members that have smaller positional errors at short lead times of 6 or 12 h were considered for the SEM in their studies. Their studies found that the SEM technique improves the track predictions compared to the ensemble mean of all the members of the ensemble prediction system. Scher and Messeri (2019) applied the SEM technique to improve the predictions of severe European windstorms. Their study found that the SEM technique could improve the prediction of extreme windstorm events up to a lead time of 36–48 h. We apply such a technique of SEM to late spring and summer (May–August) seasonal predictions of the Scale Interaction Experiment–Frontier Research Center for Global Change, version 2 (SINTEX-F2) (Doi et al. 2016, 2017), over Japan in this study. The anomaly correlation coefficient (ACC), which is widely used to verify model predictions (Miyakoda et al. 1986; Jolliffe and Stephenson 2012), is the value of correlation between the predicted anomalies and the verifying or observed anomalies. Positive (negative) values of ACC indicate the predicted anomalies to have same sign (opposite sign) as the verifying anomalies. The predictions with positive ACC values are considered skillful. Thus, ACC is a good measure for identifying the members for the SEM.

The Köppen–Geiger climatic classification (Peel et al. 2007; Beck et al. 2018) classifies the climate of Japan into three climatic zones. The northernmost island of Japan, Hokkaido, falls under the category of a cold, without dry season, and warm summer (Dfb) climate. The region of Japan between 38° and 41°N latitudes (hereafter north) falls under the category of a cold, without dry season, hot summer (Cfa) climate. The region of Japan to the west of 137°E (hereafter west) has a temperate, without dry season, hot summer (Cfa) climate.
The region between 137°–141°E and 34°–38°N (hereafter central) has a combination of Cfa and Dfa climates. Based on the above climatic classification types, we choose the four regions of Hokkaido, north, central, and west (Fig. 1) of Japan in this study for applying the technique of SEM to the SAT anomaly predictions of SINTEX-F2.

In the following sections we analyze if the SEM technique has advantages over the traditional method of averaging all the members of the ensemble when applied to SAT anomaly predictions over Japan.

2. Model and methodology

In this study, SAT predictions of the 24 members of the SINTEX-F2 forecasting system initialized in April are analyzed. The period of analysis is the late spring and summer months, May–August, which correspond to 1–4-month lead predictions from April initial conditions, covering the years 1983–2018. The area of interest are the four regions (viz. Hokkaido, north, central, and west; Fig. 1) of Japan. The SINTEX-F2 seasonal forecasting system is a coupled ocean–atmosphere global model (Masson et al. 2012; Sasaki et al. 2013) and is used for generating experimental seasonal predictions every month at Japan Agency for Marine-Earth Science and Technology (JAMSTEC). The atmospheric component of the SINTEX-F2 coupled system is the fifth generation ECHAM model (Roeckner et al. 2003) and the oceanic component is the Nucleus for European Modeling of the Ocean (NEMO) (Madec 2008). The atmospheric component has a horizontal resolution of about 100 km and 31 vertical levels. The oceanic component also has 31 vertical levels but has a horizontal resolution of 0.5° × 0.5° (~50 × 50 km²). The atmospheric and oceanic components of SINTEX-F2 forecasting system are coupled using the Ocean Atmosphere Sea Ice Soil coupler (OASIS3; Valcke et al. 2004) and fluxes are exchanged between the models every 2 h. The SINTEX-F2 predictions are initialized by updating the oceanic state by three-dimensional assimilation and also by nudging technique. The 24 members of the SINTEX-F2 differ in the way the oceanic conditions are initialized by applying various nudging strengths; 1) two kinds of observational SST datasets: the low-resolution weekly data and the high-resolution daily; 2) three large negative feedback values for the nudging (Luo et al. 2005); 3) two different modeling ways for the ocean vertical mixing (Sasaki et al. 2012); and 4) with/without monthly three-dimensional variational ocean data assimilation (3DVAR) correction (Doi et al. 2017). Therefore, the 24-ensemble prediction system considers to some extent uncertainties of both initial conditions and model physics. The details of the model spin up and retrospective prediction experiments are given in Doi et al. (2016, 2017). The SINTEX-F2 model predictions are available for the years 1983–2018 and we used those predictions in this study.

As mentioned in the previous section, the members for the selective ensemble mean are identified based on the anomaly correlation coefficient of the predictions with respect to a validation data. The detailed procedure followed for the selection of SINTEX-F2 members for the SEM is as given below:

1) The monthly anomalies of the predicted surface air temperature are generated for each of the 24 SINTEX-F2 ensemble members by removing the monthly climatology of that member. The anomalies are detrended by removing the linear trend. The period 1983–2018 is used for the calculation of the monthly climatology.

2) The Global Historical Climatological Network, version 3 and the Climate Anomaly Monitoring System (GHCN-CAMS; Fan and van den Dool 2008), a product of Climate Prediction Center, National Centers for Environmental Prediction, United States, surface air temperature is used as the validation dataset. The gridded data are at 0.5° × 0.5° horizontal resolution and is available on real time. The monthly verifying SAT anomalies are derived by removing the monthly climatology. The anomalies are detrended by removing the linear trend. The period for the calculations of climatology is 1983–2018.

3) Leave-one-out or the jackknife cross validation technique is used in the calculation of ACC for identification of the SINTEX-F2 members for the SEM. In this technique the data are repeatedly divided into calibration (35 years) and independent validation period (1 year). The ACC between the GHCN-CAMS and SINTEX-F2 surface air temperature anomalies is calculated for each training period.

4) Bootstrapping (Efron and Tibshirani 1994) with replacement is used in the calculation of the ACC. The bootstrapping technique is useful to estimate the confidence intervals (Mason 2008) of the calculated ACC. A bootstrap sample of size 500 is used to generate 500 ACC values for each training period in step 3. In the resampling, the forecast and observations pairs are kept together, so that confidence levels for the ACC are obtained. The 500 ACC values are sorted in descending order of magnitude. We choose those members of SINTEX-F2 as members suitable for the SEM ensemble that have at least 90% of the 500 ACC values, leaving the top 5% and bottom 5% values, greater than 0.0 in the training period.
5) Initial analysis showed the SINTEX-F2 members to vary among training periods in step 4 with some SINTEX-F2 members being common in most of the training periods. So, we selected those SINTEX-F2 members that were common in at least 80% of the training periods as the members of the SEM. This step is a slight modification of the traditional application of leave-one-out approach in which the weights obtained in the training period are applied to get the value.

<table>
<thead>
<tr>
<th></th>
<th>May</th>
<th>June</th>
<th>July</th>
<th>August</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hokkaido</td>
<td>2, 3</td>
<td>13, 15, 21</td>
<td>10, 21</td>
<td>16</td>
</tr>
<tr>
<td>North</td>
<td>19</td>
<td>13, 15</td>
<td>2, 10, 12, 15</td>
<td>16, 18</td>
</tr>
<tr>
<td>Central</td>
<td>19</td>
<td>1, 13, 23</td>
<td>2, 8, 12, 15</td>
<td>12, 13</td>
</tr>
<tr>
<td>West</td>
<td>6, 12, 19, 20</td>
<td>1, 8, 13, 15, 17, 19, 23</td>
<td>2, 8, 15</td>
<td>1, 6, 12, 13, 21</td>
</tr>
</tbody>
</table>

Table 1. Ensemble members of SINTEX-F2 model used for the SEM.

Figure 2. (a) Observed (black), ENSMEAN (orange), and SEM (blue) surface air temperature anomalies (°C) over Hokkaido in May of 1983–2018. (b)–(d) As in (a), but for the north, central, and west regions. (e)–(h) As in (a)–(d), but for the month of June.
in the validation year. The traditional application of leave-one-out approach assumes that the relations between the variables do not change much with time, which is not strictly true in our case. This step is useful for applying the SEM technique to predictions beyond the study period 1983–2018.

6) After identifying the SINTEX-F2 members for SEM in step 5, the SEM predicted surface temperature anomalies are generated for the period 1983–2018.

7) Steps 3–6 are applied to generate SEM forecasts over the Hokkaido, north, central, and west regions of Japan and for the months May–August.

3. Results

a. Selective ensemble mean of SINTEX-F2 members

Table 1 shows the members of the SINTEX-F2 used in the SEM forecasts of SAT anomalies over the four regions and over the months May–August. It is interesting to note that the SINTEX-F2 members of SEM vary among all the regions and over all the months May–August (Table 1). The variation of the SINTEX-F2 members in the SEM among the regions indicates that the SEM technique is more suitable for application to regional forecasts.
FIG. 4. (a) Hit rate vs false alarm rate plot of ENSMEAN (dashed lines) and SEM (solid lines) in May of positive temperature anomaly ranges over the four regions of Hokkaido, north, central, and west. The numbers correspond to the ranges of temperature anomalies as indicated at the top of each panel. (b)–(d) As in (a), but for the months June, July, and August, respectively. (e)–(h) As in (a)–(d), but for various ranges of negative temperature anomalies.
The surface air temperature predicted anomalies obtained from the ensemble mean of the 24 members of SINTEX-F2 (ENSMEAN) and the mean of the selected members (SEM) (Table 1) for all the four months and the four regions are calculated for the years 1983–2018 and presented in Figs. 2a–h and 3a–h. In the month of May, positive SAT anomalies greater than 1°C were observed in 1990, 1991, 1994, 1995, 2004, and 2016 over Hokkaido (Fig. 2a). Both the ENSMEAN and SEM predictions failed to capture SAT anomalies with magnitude greater than 1°C in all the six years i.e., the number of hits were 0 and number of misses were 6 leading to a hit rate [hit rate = hits/(hits + misses); HR] of 0 for both ENSMEAN and SEM. The false alarm, i.e., the number of years in which the predicted temperature anomalies were greater than 1°C but the observed anomalies were less than 1°C, is 0 for ENSMEAN and 1 in SEM (Fig. 2a). SEM predicted the temperature anomalies to be greater than 1°C in 2018 (Fig. 2a). The number of correct rejections, i.e., the number of years in which the observed temperature anomalies as well as predicted were less than 1°C, is 30 for ENSMEAN and 29 for SEM. Thus, the false alarm rate [false alarm rate = false alarm[(false alarm + correct rejections)]; FAR hereafter] is 0.0 for ENSMEAN and 0.03 for SEM. The calculations of HR and FAR are based on the definitions of Mason and Graham (1999). Plots of HR versus FAR are shown in Figs. 4a–d and 4e–h, respectively, for a range of positive (>1°C, >0.5°C, >0.2°C, >0.1°C, and >0.0°C) and negative (<−1°C, <−0.5°C, <−0.4°C, <−0.3°C, <−0.2°C, <−0.1°C, <0.0°C) temperature anomalies for the months May–August over the four regions (Figs. 2 and 3). From Fig. 4, it is evident that SEM has higher hit rate and lower false alarm rate compared to ENSMEAN over most ranges of temperature anomalies in all the months May–August and over all the four regions. However, SEM predicted SAT anomalies similar to ENSMEAN predicted values have lower hit rate for temperature anomalies over 1°C and below 1°C in all the regions and in all the months (Fig. 4). In the month of May, the positive temperature anomalies predicted by ENSMEAN over central and north regions have higher FAR compared to HR (Fig. 4a). Similarly, in the month of August, the ENSMEAN predicted negative temperature anomalies over Hokkaido and north regions have higher FAR compared to HR over most of the temperature ranges (Fig. 4h). These results indicate that the ENSMEAN predicted positive temperature anomalies over the central and north in May, and negative surface temperature anomalies over Hokkaido and north regions in August, are not very useful.

The ACC and RMSE values of ENSMEAN and SEM predicted SAT anomalies with respect to GHCN-CAMS surface temperature anomalies over the four regions and over the months May–August are given in Table 2. As seen from Table 2, the ACC skill score of SEM predicted SAT anomalies is higher compared to that of ENSMEAN predicted values over all the regions and in all the months May–August (Table 2). However, the RMSE of the SEM predicted SAT anomalies are higher than that of ENSMEAN predicted SAT anomalies in some of the months and over some regions (Table 2). In May, the ACC of ENSMEAN over north and central is low with values of 0.09 and −0.05, respectively (Table 2). This may partly be attributed to the high FAR compared to HR over these regions (Fig. 4a). The ACC in May over north and central regions is much improved to 0.33 and 0.46 in SEM along with the increase in HR and reduction of FAR values over these regions (Figs. 4a,e). The ACC of the ENSMEAN over Hokkaido, north, and central regions in June are low, 0.09, 0.08, and 0.08 (Table 2). The HR is almost equal to FAR over these regions in predicting the positive SAT anomalies in ENSMEAN (Fig. 4b) though the HR is higher than FAR in forecasting the negative temperature anomalies (Fig. 4f). The ACC increased to 0.45, 0.38, and 0.35 over those regions in SEM predictions for June over Hokkaido as well as over north and central regions (Table 2). The HR is also higher in SEM predictions compared to that of ENSMEAN in June over these regions (Figs. 4b,f). The ENSMEAN has higher ACC over west in predicted SAT anomalies compared to the other three regions in May and June (Table 2) though the ACC is less than that of SEM predicted values. In July, the ACC of the ENSMEAN prediction over Hokkaido, north, and central is higher than that in May and June (Table 2). Also, the HR of ENSMEAN predictions is higher over these regions in July compared to that in May and June (Figs. 4c,g). However, the SEM predicted SAT anomalies have higher ACC values (Table 2) and HR (Figs. 4c,g) compared to that of ENSMEAN in July. In August, the ACC values are higher for the SEM predicted SAT anomalies compared to the ENSMEAN predicted values (Table 2). The HR is higher and FAR is lesser in the SEM predictions compared

<table>
<thead>
<tr>
<th>May</th>
<th>June</th>
<th>July</th>
<th>August</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anomaly correlation coefficients: ENSMEAN (SEM)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hokkaido</td>
<td>0.17 (0.34)</td>
<td>0.09 (0.45)</td>
<td>0.22 (0.36)</td>
</tr>
<tr>
<td>North</td>
<td>0.09 (0.33)</td>
<td>0.08 (0.38)</td>
<td>0.17 (0.50)</td>
</tr>
<tr>
<td>Central</td>
<td>−0.05 (0.46)</td>
<td>0.08 (0.35)</td>
<td>0.19 (0.63)</td>
</tr>
<tr>
<td>West</td>
<td>0.33 (0.48)</td>
<td>0.21 (0.52)</td>
<td>0.18 (0.49)</td>
</tr>
<tr>
<td>RMSE: ENSMEAN (SEM)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hokkaido</td>
<td>0.98 (1.00)</td>
<td>1.11 (0.98)</td>
<td>1.14 (1.15)</td>
</tr>
<tr>
<td>North</td>
<td>0.83 (0.95)</td>
<td>0.92 (0.95)</td>
<td>1.27 (1.11)</td>
</tr>
<tr>
<td>Central</td>
<td>0.74 (0.76)</td>
<td>0.78 (0.85)</td>
<td>1.12 (0.89)</td>
</tr>
<tr>
<td>West</td>
<td>0.50 (0.50)</td>
<td>0.63 (0.54)</td>
<td>0.86 (0.77)</td>
</tr>
</tbody>
</table>

TABLE 2. Anomaly correlation coefficient and RMSE of ENSMEAN (SEM) for all the regions and for the months May–August.
to that of ENSMEAN prediction of SAT anomalies over all of the four regions (Figs. 4d,h).

The above results indicate that the SEM technique is useful for improving the sign and amplitude of the SAT anomaly predictions of SINTEX-F2 over different regions of Japan in the late spring and summer months (May–August).

b. Causes of differences between ENSMEAN and SEM predictions

In this section, we try to understand the causes of better performance of SEM compared to ENSMEAN in predicting SAT anomalies over the four regions of Japan. The sign and location of 850-hPa geopotential height anomalies can to some extent explain the sign of the SAT anomalies. Positive (negative) 850-hPa geopotential height anomalies located over a region indicate open sky (cloudy) conditions and hence can lead to increase (decrease) of incoming solar radiation thereby increase (decrease) the SAT anomalies over the region. Also, the circulation anomalies associated with the 850-hPa height anomalies can advect warm or cold air from the surrounding seas and contribute to the variation of the SAT anomalies. Thus, to understand the causes for the better performance of SEM compared to ENSMEAN, we correlated the 850-hPa geopotential height detrended anomalies of NCEP–NCAR reanalysis (Kalnay et al. 1996), ENSMEAN, and SEM with the GHCN-CAMS estimated SAT detrended anomalies over the four regions Hokkaido, north, central, and west regions of Japan for the months May–August. Correlating the observed SAT values with the predicted 850-hPa geopotential anomalies would bring out the biases in the prediction of 850-hPa geopotential height anomalies in ENSMEAN and SEM and thereby partly explain the performances of ENSMEAN and SEM in predicting the SAT anomalies.

FIG. 5. (a)–(c) Spatial distribution of ACC values of correlation between the GHCN-CAMS estimated temperature anomalies in May over Hokkaido with the 850-hPa geopotential height (m) anomalies of (a) NCEP–NCAR reanalysis, (b) ENSMEAN, and (c) SEM. (d)–(f) As in (a)–(c), but for June; (g)–(i) as in (a)–(c), but for July; (j)–(l) as in (a)–(c), but for August. The hashed region is significant at 90% using Student’s two-tailed t test.
The spatial distribution of ACC of SAT anomalies over Hokkaido with the NCEP 850-hPa geopotential height anomalies shows a region of positive values over Japan and over the western North Pacific (Fig. 5a) in the month of May. However, the spatial distribution of the ACC values of ENSMEAN predictions shows negative values over Japan and over the surrounding regions (Fig. 5b). The SEM prediction of 850-hPa geopotential height anomalies is positively correlated with the observed SAT values over parts of Japan and western North Pacific (Fig. 5c). In the months of June, July, and August, which are similar to May, the correlation of SAT anomalies over Hokkaido with the NCEP 850-hPa anomalies over Japan are positive. The ACC values are also positive over the surrounding regions to the east, west, and south of Japan (Figs. 5d,g,j). The spatial distribution of correlation coefficients of SAT anomalies over Hokkaido and the 850-hPa anomalies of ENSMEAN predictions though positive over all regions Japan (Figs. 5e,h,k) are negative over Hokkaido in the months of June and August (Figs. 5e,k). The spatial distribution of ACC values of SEM predictions shows positive values over Japan and the surrounding regions similar to the spatial distribution of the counterparts with NCEP though with some biases in the months of June, July, and August (Figs. 5f,i,l). The spatial distribution of ACC values shows the SEM predictions to better predict the sign of the 850-hPa geopotential height anomalies compared to the ENSMEAN predictions thereby partly explain the better performance of SEM in predicting the SAT anomalies over Hokkaido compared to the SAT predictions by ENSMEAN.

The SAT anomalies over north are positively correlated with the NCEP estimated 850-hPa geopotential height anomalies over Japan in all the months May–August (Figs. 6a,d,g,j) indicating the sign of the 850-hPa geopotential height anomalies over Japan in all the months May–August (Figs. 6a,d,g,j) indicating the sign of the 850-hPa geopotential height anomalies to explain the sign of the SAT anomalies over the region. The ENSMEAN predicted 850-hPa geopotential height anomalies are negatively correlated with the GHCN-CAMS estimated...
SAT anomalies over north in the months of May and July (Figs. 6b,h). However, the ACC values are positive over Japan in the months of June and August though small in magnitude (Figs. 6e,k). The spatial distribution of ACC values of the SEM predictions over north region is positive over Japan in all the months May–August (Figs. 6c,f,i,l) similar to the distribution of the distribution of observed ACC values (Figs. 6a,d,g,j).

Over the central region, the correlation of GHCN-CAMS estimated SAT anomalies and the NCEP estimated 850-hPa geopotential height anomalies is positive over Japan (Figs. 7a,d,g,j). The ENSMEAN predictions show negative ACC values in May and July (Figs. 7b,h) and positive ACC values over Japan in June and August (Figs. 7e,k) over Japan. On the other hand, the SEM predictions have positive ACC values over Japan in all the months (Figs. 7c,f,i,l).

Similar to the other three regions, the correlation of GHCN-CAMS estimated SAT anomalies over west are positively correlated with the 850-hPa NCEP geopotential height anomalies over Japan (Figs. 8a,d,g,j). The ACC values of the ENSMEAN over Japan are negative in May and July (Figs. 8b,h) and positive in June and August (Figs. 8c,k). The ACC values over Japan are positive in all the months in the SEM forecasts (Figs. 8c,f,i,l).

The above analysis indicates that the ENSMEAN forecasts have large biases in the prediction of the 850-hPa geopotential height anomalies. The SEM predictions, compared to ENSMEAN predictions, have less biases in the prediction of 850-hPa geopotential height anomalies thereby are better at predicting the SAT anomalies over the four regions in the months of May–August.

4. Conclusions

In this study, we investigated if the technique of SEM has any advantages over the ENSMEAN technique in the prediction of SAT anomalies over four regions of Japan: Hokkaido, north, central, and west (Fig. 1) in the months May–August. The predictions of the SINTEX-F2 seasonal forecasting model...
with 24 ensemble members for the period 1983–2018 are used in the study. The SINTEX-F2 members for the SEM are selected based on the ACC values computed between the GHCN-CAMS SAT estimates and the SINTEX-F2 predicted anomalies. After the selection of the members, the SEM predicted SAT anomalies are generated for the period 1983–2018. The SINTEX-F2 members used for the SEM are shown in Table 1. The understanding of physical processes responsible for the better performance of certain members over a particular region in a particular month is challenging, and we would be taking up such an exercise in future.

A detailed calculation of the HR and FAR for different ranges of positive and negative temperature anomalies showed the SEM technique to better predict the SAT anomalies compared to the ENSMEAN predicted SAT anomalies. The HR (FAR) of SEM forecast is higher (lower) compared to ENSMEAN over most temperature ranges (Fig. 4). The ACC values of the SEM are also higher compared to that of ENSMEAN in all the months of May–August and for all four study regions (Table 2). Analysis of the spatial distribution of the correlation between the GHCN-CAMS estimated surface temperature anomalies and the predicted 850-hPa geopotential height anomalies shows the ENSMEAN to have larger biases in forecasting the 850-hPa geopotential height anomalies compared to SEM, which could have led to the smaller ACC scores in ENSMEAN.

Even though there are many statistical methods to improve the mean surface air temperature forecasts, there are only a few techniques to improve the amplitude and sign of anomalies of seasonal forecasts (Feudale and Tompkins 2011; Ratnam et al. 2019a,b). Our study shows the SEM is a promising technique for improving the prediction of SAT anomalies over Japan.

The results of the study encourage us to further investigate the technique of SEM to improve regional predictions of temperature anomalies over other parts of the globe. The results are also useful for improving the predictions using...
dynamical downscaling. The SEM technique may be used as one of the techniques to generate optimum boundary conditions for the dynamical downscaling models to improve the regional predictions. We plan to carry out such studies in future. One of the caveats of the SEM technique is that it is highly dependent on the choice of the region and the members need to be selected for changes in the area of interest.

Acknowledgments. The research was partly supported by Japan Agency for Marine-Earth Science and Technology (JAMSTEC) under Project-B, an initiative to develop AI techniques for climate predictions. The authors thank the three anonymous reviewers for their insightful comments, which improved the manuscript substantially. The NOAA_OI_SST_V2, NCEP reanalysis, and GHCCN-CAMS data were provided by NOAA/OAR/ESRL PSD, Boulder, Colorado, from their website at https://www.esrl.noaa.gov/psd/.

Data availability statement. Due to confidentiality agreements, supporting SINTEX-F2 model data can only be made available to bona fide researchers subject to a nondisclosure agreement. Details of the data and how to request access are available from the authors.

REFERENCES


