NOTES AND CORRESPONDENCE

Interdecadal Trend of Prediction Skill in an Ensemble AMIP-Type Experiment

TOSIYUKI NAKAEGAWA AND MASAO KANAMITSU

Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California

THOMAS M. SMITH

NOAA/National Climate Data Center and Cooperative Institute for Climate Studies, University of Maryland, College Park, College Park, Maryland

8 September 2003 and 12 February 2004

ABSTRACT

This study addresses the interdecadal trend in potential skill score as estimated from the 500-hPa height temporal correlation coefficient (TCC), based on a 50-yr 10-member ensemble GCM integration with observed SST. The skill scores are based on the perfect model assumption, in which one of the members of the ensemble is assumed to be true. A distinct decadal positive trend in the TCC in boreal winter (December–January–February) was found. This trend is shown to be consistent with the positive trend in the interdecadal time-scale temporal variance of SST. The geographical pattern of the differences of the TCC between each decade and the 50-yr period resembles the Matsuno–Gill pattern, suggesting that the increase in the TCC is due to the Rossby wave excitation induced by the anomalous diabatic heating caused by the anomalous SST. Similar interdecadal trends in the variance of the Southern Oscillation index and Pacific–North American pattern were found in both the observation and the simulation. The interdecadal trend in the variance of 500-hPa geopotential height over the continental United States, however, existed only in the simulation.

1. Introduction

Since dynamical seasonal forecasting is primarily a boundary value problem, its skill is expected to depend on the nature of the external forcing. Among others, the sea surface temperature anomaly (SSTA) in the tropical Pacific imposes a significant influence on the forecast skill. Skill is known to be high during events of large interannual variability in SST, such as El Niño and La Niña episodes (Brankovic and Palmer 2000). In fact, the operational 5-month lead forecasts had high skill over the United States during the very strong 1997/98 El Niño episode (Barnston et al. 1999). In contrast to the importance of the interannual variability of SST, forecast skill is also heavily influenced by the trend in the monthly mean SST itself. Canonical correlation analysis by Shabbar and Barnston (1996) and Hwang et al. (2001) demonstrated that the forecast skills over Canada and the Far East are associated with a long-term trend in global SST. A significant increase in forecast skill is expected if the interdecadal trend of precipitation in the Tropics (due to the tropical SSTA), together with associated teleconnection patterns into the midlatitude, is predicted (Higgins et al. 2000). It should be noted, however, that the constant trend in mean SST increases the skill but does not contribute to the trend in forecast skill since the former contributes to the skill to the same extent for any period. The trend in mean

Corresponding author address: Dr. Tosiyuki Nakaegawa, University of California, San Diego, 9500 Gilman Drive, MC 0224, La Jolla, CA 92037-0224.
E-mail: tnakaegawa@ucsd.edu

© 2004 American Meteorological Society
SST can contribute to the trend in forecast skill only when the former varies with time. For example, the positive (negative) trend in forecast skill will occur when the trend in mean SST increases (decreases) with time, if nontrend variability remains the same.

In this paper, we present the relationship between the forecast skill in boreal winter and the temporal variance of SST using 50-yr climate simulations forced with observed SST.

### 2. Experiments, data, and analysis

**a. Experiments**

The model used in this study is the Scripps Institution of Oceanography (SIO) Experimental Climate Prediction Center (ECPC) version of the National Centers for Environmental Prediction (NCEP) Seasonal Forecast Model (SFM; Kanamitsu et al. 2002). The model horizontal resolution is T62 with 28 layers in the vertical. The model physics are very similar to those of the NCEP SFM except for some minor improvements in the land surface parameterizations. The integration period is 54 yr, from 1948 to 2001. The number of ensembles is seven and the initial conditions are taken somewhat arbitrarily from 1 January of different years from an independent Atmospheric Model Intercomparison Project (AMIP) run performed separately. The observed SSTs and sea ice concentrations used in this integration consist of two datasets: Hadley Centre Sea Ice and SST (HadISST, 1948–81; Rayner et al. 1996) and Optimum Interpolation sea ice concentration and SST (OISST, 1982–2001; Reynolds and Smith 1994). It is known that artificial discontinuities exist in sea ice concentration between the two analyses (Simmons and Gibson 2000). The dataset utilized in this paper is the one generated for the European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA40), in which the discontinuity was smoothed out (Fiorino 2004).

**b. Data**

We used two additional SST datasets for comparison. One set is the Extended Reconstructed SST (ERSST; Smith and Reynolds 2003), and the other is the HadISST. The equatorial Southern Oscillation index (EQSOI) is obtained from NCEP’s Climate Prediction Center (Bell and Halpert 1998), and the Southern Oscillation index (SOI) is obtained from the University of East Anglia’s Climate Research Unit (CRU; Ropelewski and Jones 1987). The observed geopotential height is taken from NCEP–National Center for Atmospheric Research (NCAR) reanalysis (Kalnay et al. 1996). Precipitation comes from the New et al. (1999, 2000) dataset, compiled by CRU. The datasets were aggregated onto the T62 Gaussian model grid. Pacific decadal oscillation (PDO; Mantua et al. 1997) and Atlantic multidecadal oscillation (AMO; Enfield et al. 2001) indices are used in the analysis. Table 1 presents the spatial resolutions, periods, and references of these datasets.

**Table 1. Observation data used for analysis.**

<table>
<thead>
<tr>
<th>Data</th>
<th>Resolution</th>
<th>Period</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>HadISST SST</td>
<td>360 × 180</td>
<td>1855–Present</td>
<td>Rayner et al. (1996)</td>
</tr>
<tr>
<td>EQSOI</td>
<td>point data</td>
<td>1949–Present</td>
<td>Bell and Halpert (1998)</td>
</tr>
<tr>
<td>NCEP–NCAR R-1</td>
<td>144 × 73</td>
<td>1949–Present</td>
<td>Kalnay et al. (1996)</td>
</tr>
<tr>
<td>PDO</td>
<td>point data</td>
<td>1900–Present</td>
<td>Mantua et al. (1997)</td>
</tr>
<tr>
<td>AMO</td>
<td>point data</td>
<td>1948–94</td>
<td>Enfield et al. (2001)</td>
</tr>
</tbody>
</table>

**c. Analysis**

The temporal correlation coefficient (TCC) of the simulation is computed according to Sardeshmukh et al. (2000) with a “perfect model” assumption and is expressed as

\[
\rho = \frac{s^2}{[(s^2 + 1)(s^2 + 1/n)]^{1/2}}, \tag{1}
\]

where \(s = \langle (x - \bar{x})(x' - \bar{x}') \rangle^{1/2} \) and is the signal-to-noise ratio, \(n\) is the ensemble number, \(x\) is the state vector in an ensemble member, \(\bar{x}\) is the ensemble mean of \(x\), and \(x'\) is the anomaly state vector. The value \(\rho\) is the same as the mean of the correlation coefficients between the ensemble mean and any realization of the ensemble members. It is the maximum value that the dynamical seasonal forecast can reach and is a measure of the potential prediction skill. We show the results for December–January–February (DJF) in the following section.

### 3. Results

Figure 1 shows the geographical distributions of 500–hPa geopotential height TCC for each decade from the 1950s to the 1990s. The values are the difference between the TCC of the 50-yr dataset (1951–2000; Fig. 1f) and that of each 10-yr dataset. The area of negative differences clearly decreased over time. The negative difference larger than 0.1 in the 1950s covered the whole globe except the equatorial zone and both Northern Hemisphere (NH) and Southern Hemisphere (SH) high latitudes. Particularly, the larger negative differences were found over the Pacific subtropical zone and North
America. The large negative differences in the 1960s covered the NH mid- and high latitudes and the South Pacific. Overall, the differences became smaller than those of the 1950s, especially over the Tropics and the SH. The large negative differences in the 1970s were limited over North Africa, the Middle East, the northern west Pacific, the United States, and the southern east Pacific. The large negative differences in the 1980s were confined over the 20°N zone, 30°S zone, and Antarctic coastal zone, and those in the 1990s were confined over the South America midlatitude and the Sahel. Large positive differences for the period from the 1970s to the 1990s existed but were confined over very small regions. Thus the skill of the seasonal forecast improves from the 1950s to the 1990s over most of the global domain. This trend is consistent with the results of Sugi et al. (1997) who showed greater skill in the period 1976–88 compared to the period 1955–88.

Since the TCC in this paper does not refer to the actual observations, the trend in observed atmospheric parameters is not responsible for the trend in forecast skill. There is a possibility of a change of the trend in atmospheric parameters due to the change of trend in SST used as a surface forcing. This will be investigated in a future study.

Figure 2 shows the geographical distribution of the ratio of the temporal variance of DJF SST of each decade over that of the 1990s. The variances clearly in-
increased with time, especially in the tropical and Antarctic oceans. The variances in the 1950s were small in almost all oceans except in the eastern North Pacific, the eastern equatorial Indian Ocean, and a portion of the South Pacific. The pattern of variances of the 1960s was similar to that of the 1950s, but the magnitude increased over most areas except the eastern North Atlantic. In the 1970s, the western Pacific and Antarctic Ocean had small SST variances, but the subtropical eastern Pacific had large variances. The large differences disappeared after the 1980s.

Sampling of SST increased over the analysis period, with the most complete sampling after November 1981 when Advanced Very High Resolution Radiometer (AVHRR) retrievals from satellites became available. Before then, only in situ sampling was available. The weak variance in the Southern Ocean before the 1980s is a result of sparse in situ sampling. Sampling problems are less severe in the Tropics and NH for this analysis period, but there is a slight increase in in situ sampling from 1950 to 1981, which may influence some of these results.

The time-dependent characteristics of simulated 500-hPa geopotential height TCC are found to be consistent with those of the SST variance. An increase in the temporal variance of ensemble mean 500-hPa geopotential height is found as well (figure not shown). Therefore, the increasing trend of 500-hPa geopotential height TCC corresponds well to the increasing trend in the temporal variance of the SST. The larger variance in SST forcing increased the signal-to-noise ratio of the 500-hPa geopotential height forecast.

4. Discussion

a. Physical processes

The difference of 500-hPa geopotential height TCCs between the 1950s and 1950–2000 shows patterns symmetric to the equator over the eastern and western Pacific. Figure 3 shows the streamlines of the warm minus cold event composites of the anomalous 500-hPa wind fields for DJF. Ten events are selected based on the SST index in the Niño-3.4 area (5°N–5°S, 170°–120°W). The symmetric dipolelike patterns are presented in Fig. 3. The positive (negative) tropical SSTA leads to decreasing (increasing) precipitation over the western Pacific and to increasing (decreasing) precipitation over the eastern equatorial Pacific, which excites the Rossby wave and forms a dipolelike forcing pattern (Matsuno 1966; Gill 1980) in the extratropics as well as the Tropics as shown in Fig. 3. The same mechanism seems to
apply to the formation of dipole patterns of high TCC in the extratropics (Fig. 1f). The excitation of the Matsuno–Gill pattern is weak during the 1950s when the interannual variability of SST is small, thus forming symmetric difference patterns in Fig. 1a. Although the magnitudes are smaller, similar patterns are observed in later years.

The latitudinal patterns over the subtropical regions are due to the trend in the zonal-mean component of the response to the SSTA that enhances/reduces the zonal-mean Hadley circulation. This mechanism is responsible for the predictability of the boreal winter extratropical height field (Schubert et al. 2002). The weak zonal-mean Hadley circulation during the 1950s associated with the small interannual variability of SST also contributes to the latitudinal difference patterns in Fig. 1.

b. Other interdecadal variabilities related to SST

Thus far, we have suggested that the interdecadal variability of the 500-hPa geopotential height TCC is controlled by the SST. Xue et al. (2003) examined the seasonal standard deviation of Niño-3 SST using 30-yr base periods and ERSST and found its upward trend since 1950. This agrees with our finding of an increased trend in the variance of each decade over the tropical eastern Pacific (Fig. 2). Similar trends were found in the standard deviation and variance of SOI (Kestin et al. 1998; Torrence and Compo 1998), the index of central equatorial Pacific rainfall (Kestin et al. 1998), Indian monsoon variance (Torrence and Webster 1998), the tropical Pacific sea level (Smith 2000), and the Darwin, Australia, sea level pressure (Xue et al. 2003).

We computed the interannual variability of the temporal variance of the SST for each decade using ERSST and HadISST. The same trends as in Fig. 2 were found in both datasets. The variance of ERSST was generally larger than that of HadISST. These characteristics are the same as those of Xue et al. (2003), who computed the variances for a 30-yr base period.

c. Interdecadal variability of the variance of parameters in observations and simulations

If the interdecadal variability of the SST variance exists, then we might find corresponding interdecadal variability in the atmospheric parameters. We examined this in both the observation and our simulation and compared the time series of the variances.

1) EQSOI

Figure 4 shows the time variation of EQSOI temporal variance in an 11-yr time window for DJF, computed by shifting the starting year by 1 (similar to a running mean, but running temporal variance instead). The simulated EQSOI is computed based on the NCEP CPC method, which is the same as that of Ropelewski and Jones (1987). The Mann–Kendall nonparametric test for the trend is used. The small values of the two-sided significance level ($P_{\text{mk}}$) indicate a statistically significant trend. The simulated trend of EQSOI well reproduced the observed EQSOI (TCC was 0.93), and the variability for the entire period. The variance of the observed EQSOI was small in the 1950s and 1960s, and it had an increasing trend ($P_{\text{mk}} = 0.1\%$). The local minimum
2886 VOLUME 17 JOURNAL OF CLIMATE

Fig. 4. (a) EQSOI time series for DJF. (b) Moving window variance of overlapping 11-yr windows of EQSOI series. The horizontal axis indicates the year at the center of each segment. Solid lines denote the observed data and dashed lines show the ensemble mean of the experiment.

Fig. 5. Rotated EOF (REOF) of 500-hPa geopotential height for DJF: (a) the second REOF loading factor of R-1 for 1949–98, (b) the first REOF loading factor of the ensemble mean for 1951–2000, (c) the score for (a), and (d) the score for (b). The horizontal axis indicates the year at the center of each segment. Solid lines denote the score, while dashed lines show the moving window variance of overlapping 11-yr windows of the score series. The percent of variance explained by each mode is indicated.

occurs around 1990. The variance of the simulated EQSOI has a similar interdecadal variability ($P_{mk} = 1.2\%$) and was very similar to the observation. The increasing rate of variance in the simulation is almost the same as that of the observation except around 1970. These results are consistent with those of the interdecadal variability of the SST variance (see Fig. 2) and with previous studies (Torrence and Compo 1998; Kestin et al. 1998). We also analyzed SOI and found a significant trend in the observed SOI variance, but not in the simulation. The insignificant trend in the simulated SOI variance was consistent with small temporal correlation differences in Fig. 1 over Tahiti (17°S, 149°E) and Darwin, Australia (12°S, 131°E), thus the use of a point value such as SOI may yield misleading results.

2) PNA PATTERN

The most significant rotated EOF mode of the observed 500-hPa geopotential height is the North Atlantic Oscillation (NAO) pattern, and the second mode is the Pacific–North American (PNA) pattern (Kanamitsu et al. 2002). In contrast, the PNA-like pattern is the first mode of the ensemble mean simulated 500-hPa geopotential height. The NAO pattern is weak in the ensemble mean. This is thought to occur because the NAO is not likely to be forced by the SST variability (Visbeck et al. 2001). Figure 5 shows the second mode of the observed (NCEP–NCAR reanalysis) 500-hPa geopotential height and the first mode of the ensemble mean of the simulations for DJF. The two patterns are similar, but the amplitude in the simulation is smaller than that of the observation. The reason for smaller first-mode
amplitudes is due to the ensemble mean procedure applied to the simulation. The first mode in the simulation made a large contribution to total variance. The TCC between the two was 0.22. The time variation of the 11-yr temporal variance from the observation had no statistically significant trend ($P_{mk} = 25.3\%$). The variance of the ensemble mean had an increasing trend ($P_{mk} = 0.1\%$) reaching maximum values in the mid-1980s and 1990s. We analyzed the trends of the scores of the PNA pattern of each member and found that four out of seven members had a significant level of trend (larger than 10%) and that the mean of all the significant levels was 19.2%. Therefore, the simulation did not show a robust trend in agreement with the observation. The statistically significant trend in the score obtained from the ensemble mean is due to the trend in the variance of SST, since the ensemble mean represents the low-frequency variations forced by the lower boundary condition. This is in accord with the trend found in the interdecadal variability of the SST variance as well as in the trend in the EQSOI variance.

3) Precipitation over tropical South America

Precipitation is a good measure for estimating the major components of diabatic heating, but its dataset over the eastern Pacific is not available. We selected a rectangular area of tropical South America ($10^\circ S$–$10^\circ N$, $60^\circ W$–$80^\circ W$), where precipitation data is available, to examine the interdecadal change in the variance of observed precipitation using the CRU dataset for DJF (Fig. 6). The observed precipitation was well reproduced in the simulation (TCC was 0.63). The trend in the 11-yr temporal variance in the observation clearly increased ($P_{mk} = 0.4\%$), and the variance became large around 1980. A statistically significant trend in the variance in the simulation is found ($P_{mk} = 9.6\%$), though the trends between the observation and the simulation behaved differently. The anomalous diabatic heating due to anomalous precipitation excites the Rossby wave, thus forming the Matsuno–Gill pattern in Fig. 3. This result supports the mechanism that is responsible for the equatorial symmetric difference pattern in Fig. 1.

d. Decadal variation in the oceans

Decadal variations in the oceans, such as the PDO and AMO, are well known. We computed the 11-yr temporal variance of PDO and AMO indices. The trend in the variance of the PDO index was not statistically significant ($P_{mk} = 46.9\%$). The trend in the variance of the AMO index was significant ($P_{mk} = 0.64\%$), but it was negative. Therefore, these decadal variations in the oceans are not responsible for the interdecadal trend in the forecast skill.

5. Summary

The forecast skill of an ensemble climate simulation using observed SSTs has had a distinct positive trend since the 1950s. This corresponds well with the increase in the interdecadal variance of the tropical Pacific SST. The geographical pattern of the difference of the TCC between individual decades and the 1950–2000 period resembles the Matsuno–Gill pattern, suggesting that the difference is due to the Rossby wave propagation induced by the anomalous diabatic heating from the anomalous tropical precipitation.

Similar interdecadal trends in the variance of the EQSOI and precipitation over tropical South America were found in both the observations and the simulation. The interdecadal trend in the variance of the PNA index was found only in the ensemble mean of the simulation.

We extended our analysis from a 500-hPa geopotential height to other variables and found that the temporal
correlations of almost all variables in the simulation, for all seasons, had increasing trends in the interdecadal time scale. There is a possibility that the model atmosphere is more sensitive to the tropical SST than the real atmosphere. The difference of TCCs between each decadal period and the 1950–2000 period was significantly reduced when the decadal period was extended. For example, for the period of 1951–70, the negative difference of less than 0.1 was confined only over areas such as North America. In this longer time window, the prediction skill is more strongly controlled by the interdecadal trend in the mean rather than the temporal variance (Shabbar and Barnston 1996; Hwang et al. 2001). We believe that a similar interdecadal trend of the variances might also be found in other atmospheric models. Unfortunately, it requires significant computer resources to verify predictability for earlier years, since an additional 50–70-yr integration with several ensemble members is needed. These integrations are left for a future project.

The positive trend in the variance of SST may be real, as indicated by Xue et al. (2003). Some of the signals are also seen in sea level observations (Smith 2000) and atmospheric parameters (Torrence and Webster 1998; Torrence and Compo 1998; Kestin et al. 1998). However, there remains a possibility that the increase is an artifact of the change in the observational coverage. In Fig. 7, we present the interdecadal change of the decadal temporal variance of tropical SST for DJF over 10°S–10°N, 80°W–180° from HadISST and ERSST.

Acknowledgments. This project is supported by GEWEX Grant GC99-016 and by Cooperative Agreement NOAA-NA17J1231. HadISST analyses were supplied by the British Atmospheric Data Center, SOI data by the University of New Anglia, and SST used for the simulation by Dr. Fiorino. Thanks to Ms. De Haan who helped with AMIP-type simulations. Ms. Diane Boomer helped to improve the text. The authors would like to express their sincere thanks to Mr. Sugi and Dr. John Roads for arranging this cooperative study.

REFERENCES
Kanamitsu, M., W. Ebisuzaki, J. Wollen, S.-K. Yang, J. Hnilo, M.