

NOTES AND CORRESPONDENCE

Improvement of the Multimodel Superensemble Technique for Seasonal Forecasts

W. T. YUN

Department of Meteorology, The Florida State University, Tallahassee, Florida, and Korea Meteorological Administration, Seoul, South Korea

L. STEFANOVA AND T. N. KRISHNAMURTI

Department of Meteorology, The Florida State University, Tallahassee, Florida

5 June 2002 and 23 May 2003

ABSTRACT

The superensemble technique has previously been demonstrated to provide an improved seasonal forecast compared to the bias-removed ensemble of equally weighted models. This paper offers a further improvement to the superensemble method by modifying the regression coefficients used in the weighting of the models for the construction of the superensemble. The improvement is achieved by use of singular value decomposition of the covariance matrix, and selecting only the largest singular value, corresponding to maximal explained variance, for the calculation of the regression coefficients. The results shown here are based on calculations done with 10 yr worth of monthly forecasts from the Atmospheric Model Intercomparison Project (AMIP) dataset, using cross validation.

1. Introduction

A considerable amount of research is necessary to address the seasonal forecast issue. Can one provide useful guidance on whether during the coming season a given region will be wetter or drier, or warmer or colder than usual? Except for some dynamical modeling and empirical statistical studies on the seasonal guidance of ENSO scenarios, such forecasts have, in general, rather low skills. Our experience shows that multimodel superensemble skills, while better than those of a multimodel bias-removed ensemble, are still generally only slightly higher than those of climatology, except for some isolated seasonal case studies. That being the state of real-time seasonal climate forecasting, it is necessary to gradually improve not only the model's data assimilation, physics, resolution, surface parameterizations, and ocean-atmospheric coupling, but also the statistical postprocessing methods. This will at best occur slowly and a continual thrust is necessary. This paper is one such effort to improve the forecast skills of atmospheric general circulation models (AGCMs). We will show that the use of singular value decomposition (SVD) for multimodel superensemble seasonal forecasts provides an

incremental improvement in forecast skills. Sometimes it is confusing to describe ensemble forecasting schemes because similar terms are used in the literature referring to different concepts. In this paper, we use the term multimodel superensemble in the sense of Krishnamurti et al. (1999).

Various ensemble methods have been used to reduce climate noise in model prediction, such as a lagged ensemble forecasting method introduced by Hoffman and Kalnay (1983), breeding techniques by Toth and Kalnay (1993), or a singular vector method by Buizza and Palmer (1995). Ensemble techniques are routinely used at operational weather forecasting centers (Molteni et al. 1996; Buizza et al. 1998; Toth and Kalnay 1997; Houtekamer et al. 1996; Stephenson and Doblak-Reyes 2000) and are also applied in longer-range climate studies (Brankovic and Palmer 1997; Zwiers 1996; Pavan and Doblak-Reyes 2000; Doblak-Reyes et al. 2000; Kharin and Zwiers 2002; Peng et al. 2002). A recent improvement in ensemble forecasting was the development of the multimodel superensemble approach. It has been shown that the multimodel superensemble forecasts are superior in skill not only to the individual member models, but also, more importantly, to the unbiased ensemble of member models (Krishnamurti et al. 1999, 2000, 2001). In this technique, the different model forecasts are statistically combined during a training

Corresponding author address: Dr. W. T. Yun, Department of Meteorology, The Florida State University, Tallahassee, FL 32306.
E-mail: wtyun@io.met.fsu.edu

phase, with the skill of each ensemble member factored into the superensemble forecast. The forecast resulting from the projection of these solutions into a forecast phase has smaller root-mean-square (rms) errors than most conventional models and conventional ensemble techniques. Various studies have extensively discussed multiple linear regression approaches for seasonal predictions. Pavan and Doblas-Reyes (2000) combined seasonal forecasts from four different AGCMs and found minimal skill improvement. Kharin and Zwiers (2002) assessed different ways of constructing multimodel forecasts and found a disagreement with the results of Krishnamurti et al., in that their regression-improved multimodel forecast (i.e., the superensemble) performed worse than the multimodel ensemble. This discrepancy is due to the fact that in their calculations the seasonal mean is removed only after the regression coefficients are calculated, while in our case the seasonal mean is removed prior to the calculation of regression coefficients.

In this study, we introduce the SVD technique for improvement of the long-term prediction skill of the multimodel superensemble. The SVD multimodel superensemble method has been applied to a set of six long-term forecast models. The results are compared to observations and to the conventional superensemble forecast as defined in Krishnamurti et al. (2000).

2. Datasets

We use datasets from several model simulations of the January 1979–December 1988 period that were produced for the Atmospheric Model Intercomparison Project (AMIP; Gates 1992). Philips (1996) describes the AMIP models. AMIP consists of 31 different global models. We arbitrarily chose six different models from AMIP for the construction of the superensemble following Krishnamurti et al. (2000). These datasets include all basic meteorological variables in the form of monthly averages. The observed analysis fields used in the study are based on the European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis. All meteorological fields are interpolated to a common resolution of 2.5° latitude by 2.5° longitude. It should be noted that the AMIP models are atmospheric general circulation models that utilize prescribed SSTs and snow cover, and the forecasts made with them are in actuality hindcasts. All calculations are done using cross validation, with each year being successively withheld from the training dataset, and the remaining nine years used for calculation of the model and observed statistics (i.e., the monthly means and regression coefficients). These means and regression coefficients are used for calculating the forecast for the verification year.

3. Construction of multimodel superensemble forecast models

This section introduces some basic concepts in multimodel superensemble forecasting. The conventional

multimodel superensemble forecast (Krishnamurti et al. 2000) constructed with bias-corrected data is given by

$$S_t = \bar{O} + \sum_{i=1}^n x_i (F_{i,t} - \bar{F}_i), \quad (1)$$

where $F_{i,t}$ is the i th model forecast for time t , \bar{F}_i is the appropriate monthly mean of the i th forecast over the training period, \bar{O} is the observed monthly mean over the training period, x_i are regression coefficients obtained by a minimization procedure during the training period, and n is the number of forecast models involved. The systematic errors of forecast models in Eq. (1) are removed because the anomalies term ($F_{i,t} - \bar{F}_i$) in the equation accounts for each model's own seasonal climatology. At each grid point for each model of the multimodel superensemble the respective weights are generated using pointwise multiple regression technique based on the training period.

a. Multimodel superensemble using standard linear regression

For obtaining the weights, the covariance matrix is built with the seasonal cycle-removed anomaly (F'),

$$C_{i,j} = \sum_{t=0}^{\text{Train}} F'_{i,t} F'_{j,t}, \quad (2)$$

where Train denotes the training period, and i and j —the i th and j th forecast models, respectively.

The goal of regression is to express a set of data as a linear function of input data. For this, we construct a set of linear algebraic equations,

$$\mathbf{C} \cdot \mathbf{x} = \tilde{\mathbf{o}}', \quad (3)$$

where $\tilde{\mathbf{o}}'_j = \sum_{t=0}^{\text{Train}} \mathbf{O}'_t F'_{j,t}$ is a ($n \times 1$) vector containing the covariances of the observations with the individual models for which we want to find a linear regression formula, and \mathbf{o}' is seasonal mean-removed observation anomaly, \mathbf{C} is the ($n \times n$) covariance matrix, and \mathbf{x} is an ($n \times 1$) vector of regression coefficients (the unknowns). In the conventional superensemble approach, the regression coefficients are obtained using Gauss–Jordan elimination with pivoting. The covariance matrix \mathbf{C} and $\tilde{\mathbf{o}}'$ are rearranged into a diagonal matrix \mathbf{C}' and $\tilde{\mathbf{o}}''$, and the solution vector is obtained as

$$\mathbf{x}^T = \left(\frac{\tilde{\mathbf{o}}''_1}{\mathbf{C}'_{11}}, \dots, \frac{\tilde{\mathbf{o}}''_n}{\mathbf{C}'_{nn}} \right), \quad (4)$$

where the superscript T denotes the transpose.

The Gauss–Jordan elimination method for obtaining the regression coefficients between different model forecasts is not numerically robust. Problems arise if a zero pivot element is encountered on the diagonal, because the solution procedure involves division by the diagonal elements. Note that if there are fewer equations than unknowns, the regression equation defines an under-

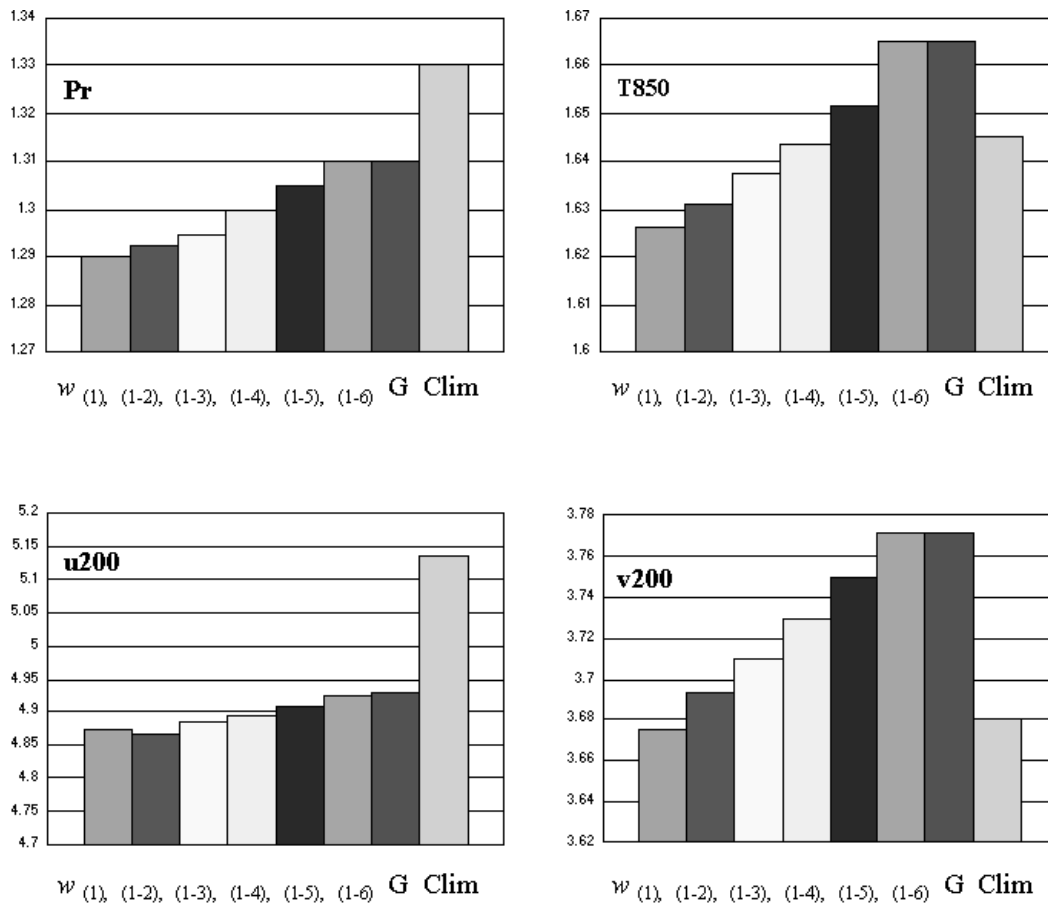


FIG. 1. Global mean rms during the forecast period from Jan to Dec 1988. The rms is computed with zeroing of the small singular values. The parenthesized numbers indicate the used singular values w_j . "Clim" denotes the rms of climatology, "G" denotes the rms of Gauss–Jordan method. Units are mm day^{-1} for precipitation, $^{\circ}\text{C}$ for temperature, and m s^{-1} for zonal and meridional wind.

determined system such that there are more regression coefficients than the number of $\{\tilde{\mathbf{o}}'_j\}$. In such a situation, there is no unique solution and the covariance matrix is said to be singular. In general, use of the Gauss–Jordan elimination method for solving the regression problem is not recommendable since singularity problems like the above are occasionally encountered. In practice, when a singularity is detected, the superensemble forecast is replaced by an ensemble forecast.

b. Multimodel superensemble using SVD

SVD is applied to the computation of the regression coefficients for a set of different model forecasts. The SVD of the covariance matrix \mathbf{C} is its decomposition into a product of three different matrices. The covariance matrix \mathbf{C} can be rewritten as a sum of outer products of columns of a matrix \mathbf{U} and rows of a transposed matrix \mathbf{V}^T , represented as

$$\mathbf{C}_{i,j} = (\mathbf{U}\mathbf{W}\mathbf{V}^T)_{i,j} = \sum_{k=1}^n w_k \mathbf{U}_{ik} \mathbf{V}_{jk}. \quad (5)$$

Here \mathbf{U} and \mathbf{V} are $(n \times n)$ matrices that obey the orthogonality relations and \mathbf{W} is an $(n \times n)$ diagonal matrix, which contains rank k real positive singular values (w_k) arranged in decreasing magnitude. Because the covariance matrix \mathbf{C} is a square symmetric matrix, $\mathbf{C}^T = \mathbf{V}\mathbf{W}\mathbf{U}^T = \mathbf{U}\mathbf{W}\mathbf{V}^T = \mathbf{C}$. This proves that the left and right singular vector \mathbf{U} and \mathbf{V} are equal. Therefore, the method used can also be called principal component analysis (PCA). The decomposition can be used to obtain the regression coefficients:

$$\mathbf{x} = \mathbf{V} \cdot \left[\text{diag} \left(\frac{1}{w_j} \right) \right] \cdot (\mathbf{U}^T \cdot \tilde{\mathbf{o}}'). \quad (6)$$

The pointwise regression model using the SVD method removes the singular matrix problem that cannot be entirely solved with the Gauss–Jordan elimination method. Moreover, solving Eq. (6) with zeroing of the small singular values gives better regression coefficients than the SVD solution where the small w_j values are left as nonzero. If the small w_j values are retained as nonzero, it usually makes the residual $|\mathbf{C} \cdot \mathbf{x} - \tilde{\mathbf{o}}|$ larger (Press

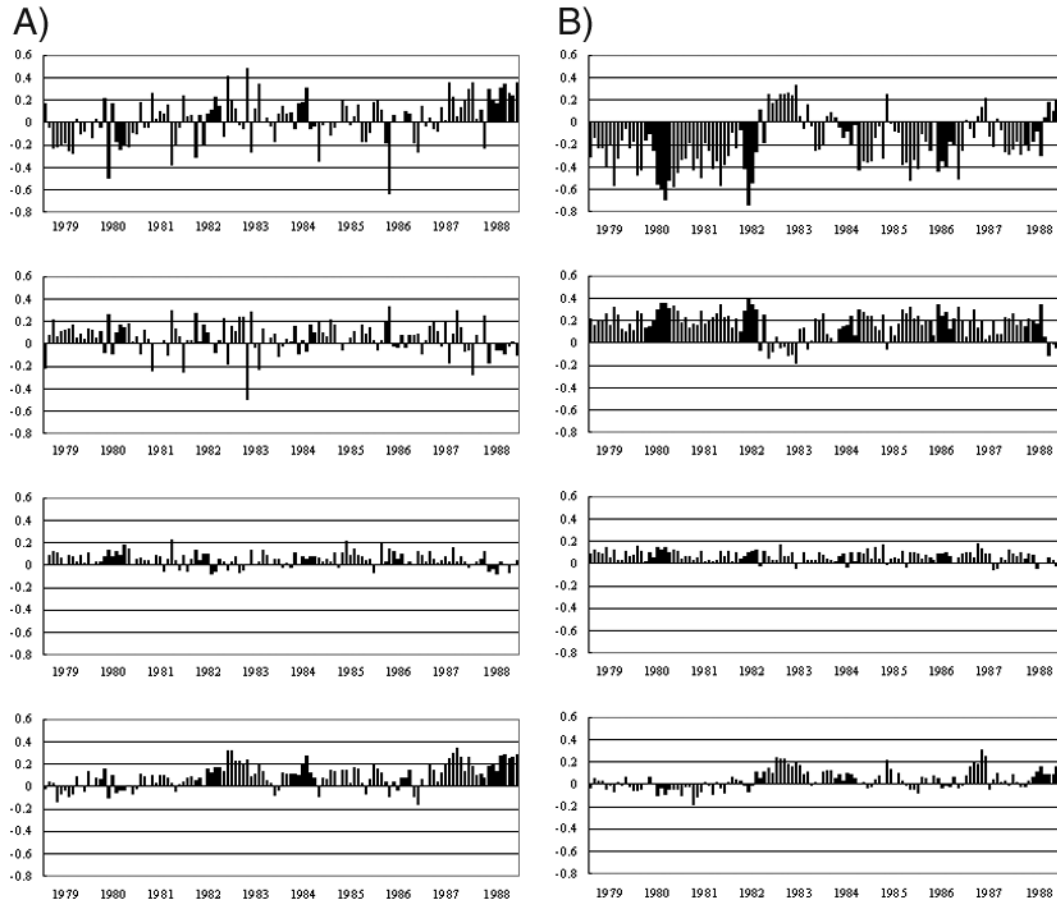


FIG. 2. Cross-validated skill scores S (a) for the temperature at 850 hPa and (b) for precipitation in the Tropics (30°S – 30°N). Skill scores S are computed with $1 - (\text{rms}_{\text{BCE}}^2/\text{rms}_{\text{CLI}}^2)$, $1 - (\text{rms}_{\text{SVD}}^2/\text{rms}_{\text{BCE}}^2)$, $1 - (\text{rms}_{\text{SVD}}^2/\text{rms}_{\text{G-J}}^2)$, $1 - (\text{rms}_{\text{SVD}}^2/\text{rms}_{\text{CLI}}^2)$ from top to bottom, where SVD and G–J indicate the forecasts based on SVD-0 and Gauss–Jordan method, CLI indicates climatology forecast, BCE indicates bias-corrected ensemble mean forecast.

et al. 1992). This means that if we have a situation where most of the w_j singular values of a matrix \mathbf{C} are small, then \mathbf{C} will be better approximated by only a few large w_j singular values in the sum of Eq. (5).

Since the regression coefficients x_i in Eq. (1) are obtained by a minimization of mean-squared error procedure during the training period, the rms error is a natural choice for assessment of the accuracy of our forecast:

$$\text{rms}_t = \left[\overline{(F_t^* - O_t^*)^2} \right]^{1/2}, \quad (7)$$

where the prime denotes the departure from the monthly mean, the asterisk denotes departure from the area mean, and the double overbar denotes an area average. The skill score describing the performance of a forecast method A versus a forecast method B can be calculated as

$$S = 1 - \frac{\text{rms}_A^2}{\text{rms}_B^2}, \quad (8)$$

where A , B can be SVD, Gauss–Jordan (G–J), bias-corrected ensemble mean (BCE), or climatology (CLI). If S is positive (negative) forecast method A is

more (less) skillful than forecast method B . We also assess the anomaly correlation, defined as $AC_t = \overline{F_t^* O_t^*} / (\overline{F_t^{*2}} * \overline{O_t^{*2}})^{1/2}$ as a complementary measure of forecast quality.

4. Multimodel superensemble forecast

This section describes and compares the performance of the aforementioned superensemble techniques. The methods in this study are demonstrated by applying them to six different model forecasts of AMIP models. All calculations are done with cross validation.

Figure 1 shows the global mean rms of precipitation, 850-mb temperature, and 200-mb zonal and meridional wind forecasts during the forecast period. The rms is computed for the climatological forecast (zero anomaly), the conventional superensemble by the Gauss–Jordan method, and the new SVD method with zeroing of a different number of small singular values. When all singular values are considered for the computation of the member models' regression coefficients, the rms obtained by the SVD method is virtually indistinguishable

from that by the Gauss–Jordan method. This is to be expected, since both represent the solution to the same set of equations, with slight differences only in the rare instances of a singular solution matrix. The rms of the SVD superensemble progressively decreases with fewer singular values considered, and is found to be smallest if only the largest one or two singular values are used. The comparison with the climatological forecast can be used as a benchmark for the usefulness of the forecast. It is particularly important to note that the SVD method with zeroing of all but the largest singular value (this method will be further referred to as SVD-0) can yield a better-than-climatology forecast even when the conventional superensemble method performs poorer than climatology.

The cross-validated skill scores S for the Tropics (30°S – 30°N) are shown in Fig. 2 for the 850-hPa temperature (a) and for the precipitation (b) forecast. In the case of 850-hPa temperatures, the bias-corrected ensemble generally performs better than the climatological forecast (average improvement of 2%) while in the case of precipitation, the bias-corrected ensemble performs worse than climatology by an average of 18%. In both precipitation and temperature cases, the SVD-0 method offers a great improvement over the bias-corrected ensemble (average of 17% for precipitation, and 5% for temperature). SVD-0 is an improvement over G–J as well (average of 6% for precipitation, 5% for temperature). The percentage improvement over G–J is less than that over the bias-corrected ensemble, which implicitly shows that G–J is better than the bias-corrected ensemble. Finally the average improvement of SVD-0 over climatology is 4% for precipitation and 9% for temperature. In the Tropics, the superensemble forecast based on the SVD-0 method shows the highest skill score. In most of verification period the performance of SVD-0 superensemble forecast is better than that of Gauss–Jordan method-based superensemble forecast. The SVD-0 superensemble forecast outperforms both the bias-corrected ensemble mean forecast and the climatological forecast. In the extratropical (25° – 50°N) Northern Hemisphere, where atmospheric predictability is relatively low, the SVD-0 superensemble forecast is better than the superensemble forecast based on the Gauss–Jordan method. Nevertheless the SVD-0 superensemble forecast barely outperforms the climatology forecast (the figures are not shown here).

A comparison of precipitation anomaly correlations in the Tropics (30°S – 30°N) is shown in Fig. 3. It can be seen that, in terms of anomaly correlation, SVD-0 is generally an improvement over the bias-corrected ensemble mean, except when the bias-corrected ensemble mean's correlation is already relatively high (Fig. 3a). The average anomaly correlation of SVD-0 is about 6% higher than the bias-corrected ensemble forecast anomaly correlation. For most cases, SVD-0 is an improvement over G–J regardless of the magnitude of the G–J

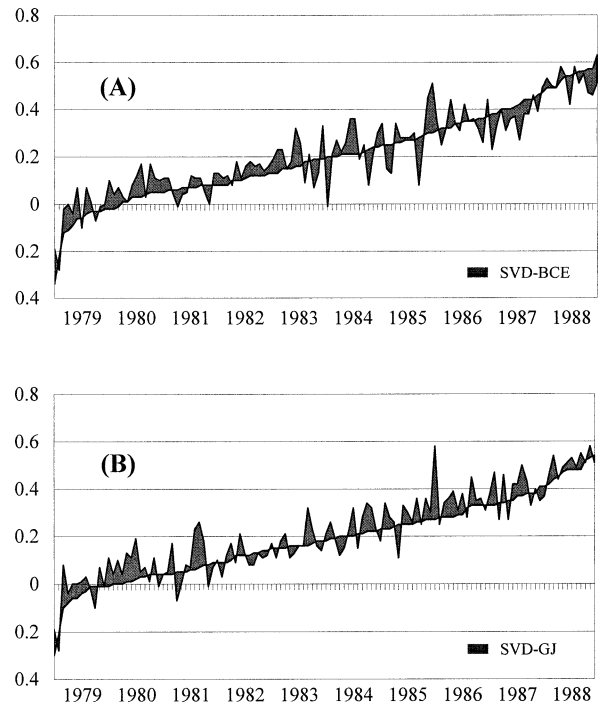


FIG. 3. Cross-validated monthly precipitation anomaly correlation in the Tropics (30°S – 30°N) of SVD-0 plotted around the anomaly correlation of (a) the bias-corrected ensemble mean (values on abscissa arranged in order of increasing anomaly correlation of bias-corrected ensemble mean forecast), and (b) G–J (values on abscissa arranged in order of increasing anomaly correlation of G–J forecast). The shaded area indicates the differences.

anomaly correlation (Fig. 3b), with an average increase of anomaly correlation by about 17%.

Figure 4 illustrates the difference in forecasts made by the conventional superensemble and the superensemble constructed using SVD-0. The plot shows the forecast period (January–December 1988) averaged precipitation anomaly over the Asian monsoon region (20°S – 30°N , 50° – 150°E). The observed anomaly is shown in Fig. 4a. The forecast anomaly by Gauss–Jordan method is in Fig. 4b, and in Fig. 4c is the forecast anomaly by SVD-0. Both forecasts underestimate the strength of the anomalies, but do capture the general structure of the anomaly field, with below-normal precipitation in the southern part of the region and generally above-normal precipitation between the equator and 20°N . The difference between the conventional superensemble forecast and the observed anomaly is shown in Fig. 4d. Apparently, the forecast is too wet south of 5°S , too dry between 5°S and 10°N , and from 10° to 30°N it is too wet in the western part of the domain and too dry in the eastern part. If now we compare this picture to the difference between the SVD-0 and Gauss–Jordan methods (Fig. 4e), we see the improvement of forecast brought about by use of the new scheme. Although the magnitudes of the differences are relatively small compared to the discrepancies with the observations, the

PR Forecast

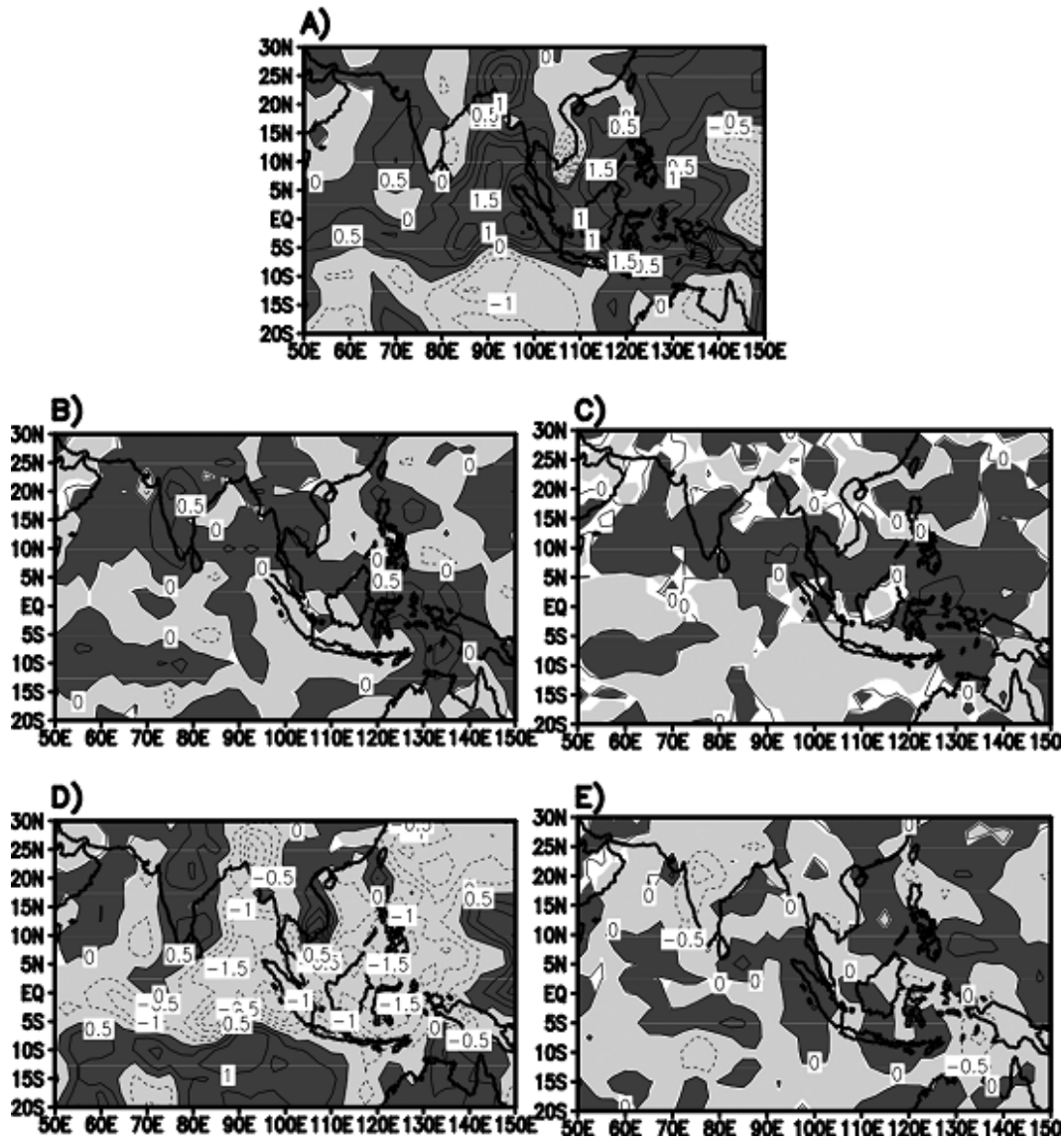


FIG. 4. Forecast period (Jan–Dec 1988) averaged precipitation for Asian monsoon region (20°S – 30°N , 50° – 150°E). (a) Observed anomaly; (b) anomaly forecast by Gauss–Jordan method; (c) anomaly forecast by SVD-0; (d) difference between (b) and (a); (e) difference between (c) and (b). Units are mm day^{-1} . Negative values have light gray shading and dashed contours, positive values have dark gray shading and solid contours; contour interval is 0.5 mm day^{-1} .

changes are of the appropriate sign. Overall, the areas that are forecast too wet by the conventional superensemble are now somewhat drier, and those forecasts that are too dry are now somewhat wetter. Even though these corrections are small in amplitude, they do lead to a better forecast and a reduction in the rms errors.

5. Summary

Our study has focused on improving the prediction skill through the construction of a multiple regression model using an SVD-based approach for the generation

of multimodel superensemble forecasts. The regression model is constructed using covariance matrices where the bias and the annual cycle were removed. For obtaining the optimal regression coefficients, the squared uncertainties of estimated parameter x in Eq. (6) are minimized by setting the small singular values to zero, based on the premise that since smaller squared uncertainties of estimated parameter x explain the relative variance better, that would enhance the multimodel superensemble forecast.

We have shown that the results of the proposed technique are better than those of the conventional super-

ensemble method, and that the multimodel superensemble forecast based on the SVD method with zeroing of the small singular values (SVD-0) yields the best result. A postprocessing algorithm based on multiple regression of multimodel solutions toward observed fields during a training period is one of the best solutions for long-term prediction. Due to the removal of biases of the different models, the forecast multimodel ensemble and superensemble errors are already quite small. Our study shows that SVD-0 further reduces the forecast errors below those of the conventional superensemble technique, whose errors are already below those of the bias-removed ensemble. The forecast produced by the proposed method generally outperforms the climatological forecast, particularly in the Tropics, even when the bias-corrected ensemble fails to do so.

In the context of seasonal climate forecasts, it remains to be seen whether the proposed SVD-0 method can improve upon the previously used standard linear regression method. That needs to be addressed using coupled atmosphere–ocean models. In previous studies we had noted that the multimodel bias-removed ensemble demonstrated a skill slightly below that of climatology, whereas the superensemble provided a skill slightly above that of climatology. Whether the SVD-0 method can increase the skill for coupled models to the extent it is able to do so for the uncoupled atmospheric models in the hindcast setting of the present studies will be a challenging problem for the future.

Acknowledgments. The research reported here was supported by NSF Grant ATM-0108741, NOAA Grant NA06690512, and FSURF Grant 1338-831-45.

REFERENCES

- Brankovic, C., and T. N. Palmer, 1997: Atmospheric seasonal predictability and estimates of ensemble size. *Mon. Wea. Rev.*, **125**, 859–874.
- Buizza, R., and T. N. Palmer, 1995: The singular-vector structure of the atmospheric global circulation. *J. Atmos. Sci.*, **52**, 1434–1456.
- , T. Petroliagis, T. Palmer, J. Barkmeijer, M. Hamrud, A. Hollingsworth, A. Simmons, and N. Wedi, 1998: Impact of model resolution and ensemble size on the performance of an ensemble prediction system. *Quart. J. Roy. Meteor. Soc.*, **124**, 1935–1960.
- Doblas-Reyes, F. J., M. Déqué, and J.-P. Piedelievre, 2000: Multimodel spread and probabilistic forecasts in PROVOST. *Quart. J. Roy. Meteor. Soc.*, **126**, 2069–2087.
- Gates, W. L., 1992: AMIP: The Atmospheric Model Intercomparison Project. *Bull. Amer. Meteor. Soc.*, **73**, 1962–1970.
- Hoffman, R. N., and E. Kalnay, 1983: Lagged average forecasting, an alternative to Monte Carlo forecasting. *Tellus*, **35A**, 100–118.
- Houtekamer, P. L., L. Lefaiivre, J. Derome, H. Ritchie, and H. L. Mitchell, 1996: A system simulation approach to ensemble prediction. *Mon. Wea. Rev.*, **124**, 1225–1242.
- Kharin, V. V., and F. W. Zwiers, 2002: Climate predictions with multimodel ensembles. *J. Climate*, **15**, 793–799.
- Krishnamurti, T. N., C. M. Kishtawal, T. E. LaRow, D. R. Bachiochi, Z. Zhang, C. E. Williford, S. Gadgil, and S. Surendran, 1999: Improved weather and seasonal climate forecasts from multimodel superensemble. *Science*, **285**, 1548–1550.
- , —, Z. Zhang, T. E. LaRow, D. R. Bachiochi, C. E. Williford, S. Gadgil, and S. Surendran, 2000: Multi-model ensemble forecasts for weather and seasonal climate. *J. Climate*, **13**, 4196–4216.
- , and Coauthors, 2001: Real-time multianalysis–multimodel superensemble forecasts of precipitation using TRMM and SSM/I products. *Mon. Wea. Rev.*, **129**, 2861–2883.
- Molteni, F., R. Buizza, T. N. Palmer, and T. Petroliagis, 1996: The ECMWF ensemble prediction system: Methodology and validation. *Quart. J. Roy. Meteor. Soc.*, **122**, 73–119.
- Pavan, V., and J. Doblas-Reyes, 2000: Multimodel seasonal hindcasts over the Euro-Atlantic: Skill scores and dynamic features. *Climatic Dyn.*, **16**, 611–625.
- Peng, P., A. Kumar, H. Van den Dool, and A. G. Barnston, 2002: An analysis of multimodel ensemble predictions for seasonal climate anomalies. *J. Geophys. Res.*, **107**, 4710, doi:10.1029/2002JD002712.
- Phillips, T. J., 1996: Documentation of the AMIP models on the World Wide Web. *Bull. Amer. Meteor. Soc.*, **77**, 1191–1196.
- Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, 1992: *Numerical Recipes in Fortran*. 2d ed. Cambridge University Press, 963 pp.
- Stephenson, D. B., and F. J. Doblas-Reyes, 2000: Statistical methods for interpreting Monte Carlo ensemble forecasts. *Tellus*, **52A**, 300–322.
- Toth, Z., and E. Kalnay, 1993: Ensemble forecasting at NMC: The generation of perturbations. *Bull. Amer. Meteor. Soc.*, **74**, 2317–2330.
- , and —, 1997: Ensemble forecasting at NCEP and the breeding method. *Mon. Wea. Rev.*, **125**, 3297–3319.
- Zwiers, F. W., 1996: Interannual variability and predictability in an ensemble of AMIP climate simulations conducted with the CCC GCM2. *Climate Dyn.*, **12**, 825–847.