# Combining ENSO Forecasts: A Feasibility Study

SWEN METZGER

*Max Planck Institute for Chemistry, Mainz, Germany*

MOJIB LATIF

*Max Planck Institute for Meteorology, Hamburg, Germany*

KLAUS FRAEDRICH

*Meteorological Institute, University Hamburg, Hamburg, Germany*

## ABSTRACT

Numerical and statistical predictions of simplified models are linearly combined in a sensitivity study to identify the crucial parameters that are needed to enhance predictability of the El Niño–Southern Oscillation (ENSO) phenomenon. The results indicate that the ENSO prediction skill of the simplified models can be improved. The profit achieved strongly depends on the phase information that is utilized by the forecast combination and is inherent in predictions of a quasi-periodic process such as ENSO. The simplest forecast combination that still yields useful forecasts at longer lead times of about half of the ENSO period (18–24 months) is the combination of two persistence forecast schemes. For the prediction period 1982–2003, that is the persistence of a sea surface temperature anomaly (SSTA) index area at 60°S, 180°W and the Niño-3 index SSTA.

The level of skill improvement critically depends on the prediction schemes and prediction period, as well as on the period from which the combination weights are derived. Differences between combination forecast and hindcast are minimized if the statistical weights are derived from a time period that is characterized by an ENSO statistic that is close to the prediction period. In this study the prediction period has simply been halved for the sake of simplicity to derive the statistical weights, which is sufficient for predicting the 1980s and 1990s (with each other), but not for predicting, for example, the 1970s. The suppressed 1976/77 El Niño event makes the periodic occurrence less regular compared to the other decades. However, this forecast combination technique can be applied in a much more elaborate way.

## 1. Introduction

One of the most important modes of global climate variability is the El Niño–Southern Oscillation (ENSO) phenomenon. It connects the oceanic El Niño phenomenon with the atmospheric phenomenon of the Southern Oscillation and results from the interaction between the ocean and the atmosphere on interannual time scales (Bjerknes 1969). For a more detailed introduction see, for example, Philander (1990) or Neelin et al. (1994). Because the interannual occurrence of ENSO—in average every four years—affects life and life quality in the Tropics and midlatitudes, its predictability has evoked great public and scientific interest. Therefore, under the World Climate Research Program's (WCRP) Tropical Ocean Global Atmosphere (TOGA) Project,

various numerical models for the prediction of ENSO have been developed in the past 15 years. These range from modern statistical techniques (e.g., Xu and von Storch 1990; Barnston and Ropelewski 1992; Penland and Magorian 1993; Xue et al. 1994; Van den Dool and Barnston 1995; Tangang et al. 1997; Johnson et al. 2000a,b) to climate models of different complexity (e.g., Cane et al. 1986; Zebiak and Cane 1987; Blumenthal 1991; Goswami and Shukla 1991a,b; Latif and Flügel 1991; Barnett et al. 1993; Latif et al. 1993; Balmaseda et al. 1994; Oberhuber et al. 1998; Stockdale et al. 1998; Ji et al. 1998; Grötzner et al. 1999; Mason et al. 1999; Wang et al. 2002). Reviews of the state-of-the-art ENSO prediction can be found in Barnston et al. (1994), Latif et al. (1994), Latif et al. (1998), and O'Brien et al. (1998).

The wide range of ENSO prediction models, which perform quite differently, suggests that a combination of forecast schemes may be useful in order to improve the prediction skill. For instance, simple statistical models, such as Markov-type prediction models, are more

*Corresponding author address:* Swen Metzger, Dept. of Atmospheric Chemistry, Max Planck Institute for Chemistry, J. J. Becherweg 27, D-55128 Mainz, Germany.
E-mail: metzger@mpch-mainz.mpg.de

likely to show higher prediction skills at short lead times, that is, up to 3 months, compared to coupled ocean–atmosphere models that may achieve useful skills for lead times up to 1 yr (e.g., Latif et al. 1998). In fact, Thompson (1977) has shown for a theoretical "weather prediction" case that an optimal combination of independent forecasts can reduce the error variance by about 20%. Fraedrich and Leslie (1987) and Fraedrich and Smith (1989) have demonstrated that the combination of short-term forecasts of a stochastic and a numerical weather prediction model significantly enhances prediction skill. This was further demonstrated by a combination of long-range forecasts of the SST anomaly (SSTA) of the tropical Pacific up to a 3-month lead time. Metzger (1996, 1997) has applied the same forecast combination scheme to the prediction of tropical Pacific SST anomalies, exceeding the 3-month lead time. More recently, Krishnamurti et al. (2000) showed that multimodel superensemble forecasts for weather and seasonal climate are superior to the individual ensemble forecasts, and Rajagopalan et al. (2002) demonstrated that categorical climate forecasts can be improved through regularization and optimal combination of multiple GCM ensembles. Furthermore, Peng et al. (2002) presented an analysis of multimodel ensemble predictions of seasonal climate analysis, and Coelho et al. (2004) discussed forecast calibration and combination for ENSO based on a Bayesian approach.

In this paper, we focus on the combination of numerical and statistical models and on the combination of purely statistical models by extending the study of Metzger (1996, 1997). To understand and identify the crucial parameters that are needed to enhance predictability of ENSO, we combine here only on simplified models.

## 2. ENSO prediction schemes

The following prediction models are used: (i) persistence; (ii) second-order autoregressive prediction model [AR(2)] (e.g., Honerkamp 1990); (iii) principal oscillation patterns (POPs) (Xu and von Storch 1990; Penland and Magorian 1993; Johnson et al. 2000a,b); and (iv) a hybrid coupled ocean–atmosphere model (HCM) (Flügel 1994).

### a. Persistence prediction model

The simplest forecast that can be made is the persistence forecast; for example, the sea surface temperature anomaly $\hat{\phi}$ observed at the time ($t$) is held constant for the lead time ($\tau$). For a given lead time, the lag $\tau$ persistence forecast is expressed as

$$\phi_{\text{Per}}(t + \tau) = \hat{\phi}(t). \tag{1}$$

The skill of the persistence prediction model (which is a univariate, zero-order Markov-type process) of the Niño-3 index (as defined below) simply reflects the pe-

riodic nature of ENSO, with a maximum negative correlation value at a certain lead time that can be identified at the half of the mean ENSO period. This 180° phase shift occurs for the period January 1950 to December 1990 at a lead time of 2 yr (Metzger 1996; 1997). The correlation attains a value of about −0.5 for a smoothed time series, indicating a dominant 4-yr period for these decades (smoothed with a 13-month running mean to remove the annual cycle). Therefore, an autocorrelation of the SSTA contains some useful information about its future evolution.

### b. Principal oscillation pattern prediction model

A more sophisticated set of forecasts can be derived from a statistical investigation of the space–time structure of ENSO with a multivariate, first-order autoregressive prediction model, that is, the method of POPs. The POPs were developed to extract the dominant modes of variability from higher-dimensional datasets (e.g., Hasselmann 1988; von Storch et al. 1988). A detailed discussion of POP prediction models can be found in Xu and von Storch (1990), Penland and Magorian (1993), and Johnson et al. (2000a,b). The POPs are the eigenvectors of the system matrix obtained by fitting the data to a multivariate, first-order Markov process in which the residual forcing is minimized:

$$\phi_{\text{POP}}(t + \tau) = \mathbf{A}\hat{\phi}(t) + \eta(t). \tag{2}$$

POPs are, in general, complex. The corresponding complex coefficient times series satisfies the standard damped harmonic oscillator equation, so that the evolution of the system in the two-dimensional POP space can be interpreted as a cyclic sequence of spatial patterns. The characteristic period to complete a full cycle will be referred to as the rotation period, and the *e*-folding time for exponential decay as damping time. Both are estimated as part of the POP analysis. The POPs can be regarded as the eigenmodes of the coupled ocean–atmosphere system if we consider, for instance, simultaneously atmospheric and oceanic quantities.

In our forecast experiments, the system matrix of the POP prediction model is based on an SSTA time series from January 1950 to December 1978 of the equatorial Pacific (170°E–90°W; 10° zonal resolution). The leading POP made explaining almost a quarter of the total variance has a rotation period of 40 months and a damping time of 29 months (not shown); only this leading mode is used in the prediction scheme.

### c. Second-order autoregressive prediction model

Another set of statistical forecasts was generated with the AR(2). The AR(2) prediction model (which describes the dynamics of a stochastically driven harmonic oscillator) has two time-independent parameters $\alpha_{1,2}$, so that the forecasts can be expressed as

$$\phi_{\text{AR(2)}}(t + \tau) = \alpha_1\hat{\phi}(t) + \alpha_2\hat{\phi}(t - \tau), \tag{3}$$

satisfying the standard damped harmonic oscillator equation [for a more detailed description see, e.g., Honerkamp (1990)]. The statistical parameters of the POP and AR(2) prediction models are derived from the smoothed SSTA time series (1950–78) and are equatorially averaged (170°E–90°W) for fitting the model parameters of the AR(2) process. Note that we choose an AR(2) process, as this process already yields reasonable forecasts compared to higher-order fits; a result of a sensitivity study that is not shown.

### d. Hybrid coupled model

The numerical predictions were generated with a hybrid coupled ocean–atmosphere model (HCM) described in Flügel (1994). The HCM consists of a nonlinear oceanic general circulation model (Latif 1987) coupled to a simple linear atmospheric feedback model without internal dynamics. The ocean model is based on the primitive equations with the following approximations: equatorial $\beta$ plane, Boussinesq approximation, constant salinity, and vertical mixing coefficients dependent on the Richardson number. The atmospheric feedback model is based on a coupling matrix that is derived from a linear regression between the observed wind stress and the SST anomaly fields. Additionally, a correction matrix is applied during the coupled integration to account for the systematic differences between observed and simulated SSTA fields (see e.g., Barnett et al. 1993; Flügel 1994). The numerical SSTA forecasts

$$\phi_{\text{HCM}}(t + \tau) = M[\hat{\phi}(t), \tau] \tag{4}$$

are derived for lead times $\tau$ from the model integration $M(\tau)$ initialized at time $t$. The initial states for the HCM forecasts are obtained from an uncoupled run in which the oceanic circulation model was driven with observed Florida State University (FSU) wind stresses (Goldenberg and O'Brien 1981) from January 1961 to July 1994. The statistical atmosphere model was derived from the period 1961–94.

### e. Forecast combination scheme

To combine the forecasts $\phi_k = \phi_k(t + \tau) = \phi_k(1 + \tau), \ldots, \phi_k(n + \tau)$, $t = 1, \ldots, n$, of the previously described prediction schemes $\phi_k = \phi_{\text{Per}}, \phi_{\text{POP}}, \phi_{\text{AR}(2)}, \phi_{\text{HCM}}$, $k = 1, \ldots, m$, to a new forecast

$$\phi^* = \sum_{k=1}^{m} \alpha_k \cdot \phi_k, \tag{5}$$

the statistical weights $\alpha_k$ of the $k$th prediction scheme $\phi_k$ have to be determined by minimizing the mean-square error of the combination forecast $E^* = \hat{\phi} - \phi^*$.

Both the observation time series $\hat{\phi} = \hat{\phi}(1), \ldots, \hat{\phi}(n)$ and the predictions $\phi_k = \phi_k(1 + \tau), \ldots, \phi_k(n + \tau)$ can be normalized to zero mean and unit variance to combine forecasts of different physical properties, for example, SSTA with wind or pressure anomalies. For a two-forecast combination, the statistical weights $\alpha_k$, $k = 1, 2$ can be expressed in terms of correlation coefficients:

$$\alpha_1 = (R_1 - r \cdot R_2)/(1 - r^2), \tag{6}$$

with changing indices for $\alpha_2$. The coefficients $R_1$, $R_2$ describe the correlation of each prediction scheme $\phi$ with the observation $\hat{\phi}$, while $r$ gives the cross correlation between the prediction schemes. The combination method is described in greater detail in Fraedrich and Leslie (1987) and Fraedrich and Smith (1989).

All schemes provide regionally averaged STTA ensemble forecasts (the areas are mentioned above) that are verified against the Niño-3 index; the anomalous sea surface temperature of the eastern tropical Pacific averaged over the area 5°N–5°S, 150°W–90°W. The Niño-3 index is chosen here only exemplarily as predictand for our combination experiments, as this index is a widely accepted indicator for extreme ENSO events (e.g., Neelin et al. 1994). The forecasts are initialized monthly. The model parameters of the statistical prediction schemes [AR(2), POP] have been derived from a time span independent of the prediction periods. Note that persistence does not require an independent time period to fit model parameters, the main disadvantage of the other schemes when applied to short time series. Therefore, we restrict to persistence for the sensitivity study.

## 3. Combination experiments

In the following, forecasts of the sea surface temperature anomaly averaged over the Niño-3 region (Niño-3 index) are shown. First, the numerical forecasts of the HCM model are combined with the statistical forecasts of the POP and AR(2) models and then with persistence. Second, two statistical forecasts (persistence) are combined with two combination forecasts (based on persistence). The HCM–persistence combination and persistence–persistence combination will be analyzed in greater detail to investigate the dependence of the combination skill on the prediction period, the ENSO index, the seasonality, and the time period the statistical weights are derived from. We emphasize, however, that the prediction models are chosen only exemplarily to identify the crucial parameters that are needed to enhance predictability of the ENSO phenomenon and not to achieve the maximum skill that might be achievable with this combination method.

### a. HCM–POP and HCM–AR(2) combinations

For the combination of the numerical HCM forecasts with the statistical forecasts of the POP model, a 10-yr prediction period is used (1979–89) to allow an estimation of the POP system matrix from an independent time period (1950–78). The results are presented in Fig. 1.
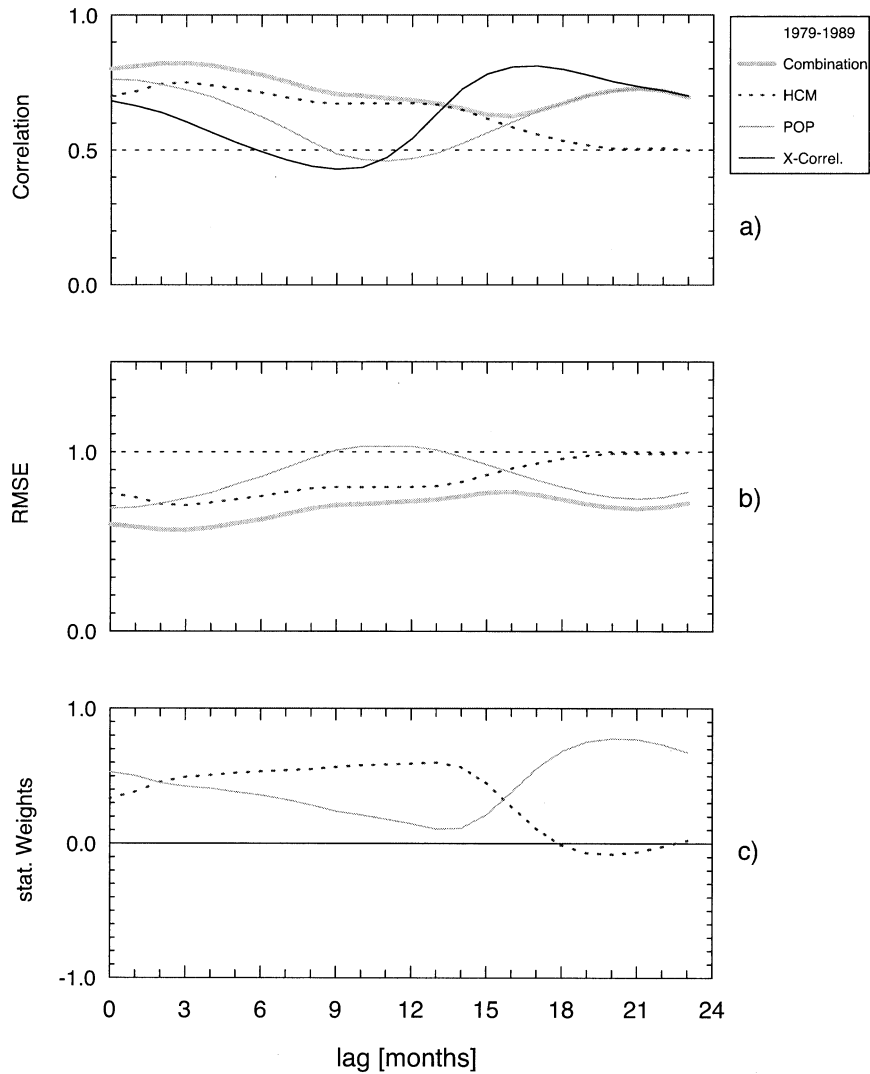
FIG. 1. (a) The correlation skill, (b) rmse, and (c) statistical weights of the HCM (dotted) and POP (thin gray) predictions of the Niño-3 index and their combination (thick gray) for the prediction period Jan 1979–Dec 1988. The cross correlation between the HCM and POP predictions is shown in (a) (thin solid line).

While the correlation skill of the combination shows only a slight improvement for lead times up to 1 yr (Fig. 1a), the root-mean-square error (rmse) (Fig. 1b) is reduced for the entire prediction and smaller than the standard deviation of the observation period. Note that the rmse is more sensitive to the error in amplitude, while the correlation skill is a more sensitive to the error in phase. The statistical weights (Fig. 1c) further demonstrate that the POP prediction model achieves a higher prediction skill at longer lead times than the HCM model, since the POP model reflects the phase information that is inherent in the coupled ocean–atmosphere system.

Figure 2 shows (similar to Fig. 1) the results of the forecast combination of the HCM model now combined with the forecasts of the AR(2) model for the same period. In contrast to the HCM–POP combination, the correlation skill of this combination is initially higher because of the high initial correlation of the AR(2) model (Fig. 2a versus Fig. 1a). The HCM–AR(2) combination also yields a larger reduction of the rmse relative to the two individual prediction schemes at higher lead times (Fig. 2b versus Fig. 1b). The statistical weights of the individual forecasts (Fig. 2c) now show higher initial weights of the AR(2) model, while both schemes have comparable weights at lead times longer than 1 yr, in opposition to the HCM–POP combination (Fig. 1c).

Various lag prediction errors derived from both combinations, HCM–POP and HCM–AR(2), are shown in Fig. 3 together with prediction errors of the HCM itself. For the prediction period 1979–89 the prediction errors for lags 3, 12, 15, and 21 (lead time in months) are
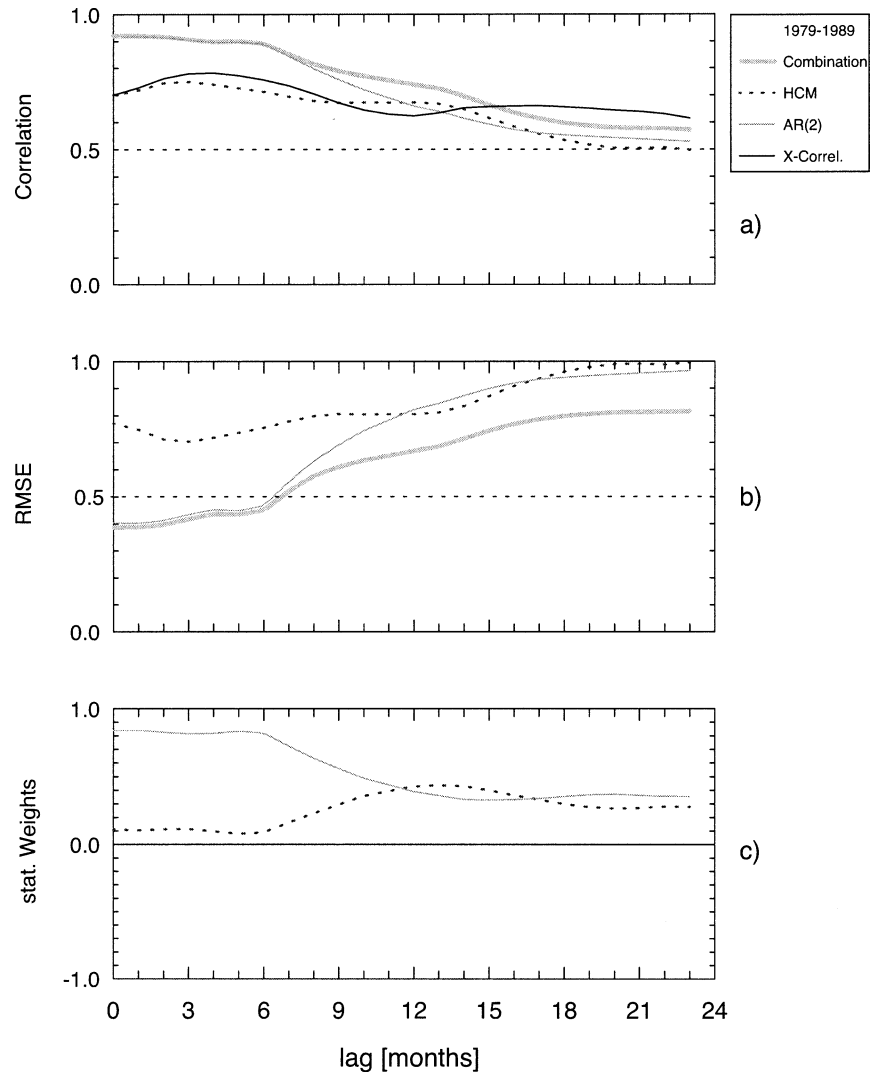
FIG. 2. (a) The correlation skill, (b) rmse, and (c) statistical weights of the HCM (dotted) and AR(2) (thin gray) predictions of the Niño-3 index and their combination (thick gray) for the prediction period Jan 1979–Dec 1988. The cross correlation between the HCM and AR(2) predictions is shown in (a) (thin solid line).

shown. In general, the two combinations perform rather similarly. For example, the prediction errors of both combinations are smaller for all lags than those of the HCM forecasts if we consider the 1984/85 cold phase. However, the forecasts of both El Niño events (1982/83 and 1986/87) are not considerably improved. In particular, the intensity of the extraordinary 1982/83 El Niño is to the same extent underestimated by both combinations as by the HCM, with a prediction error in the range of the magnitude of the internal variability.

The cross correlations that correspond to the combinations are also shown in Figs. 1a and 2a, respectively. Both cross correlations show relatively high values (above 0.5) for all lead times. Such high values indicate

that numerical and statistical forecasts are not independent from each other. This implies that only relative little skill improvement can be expected with these combinations. The fact that the numerical and statistical forecasts are somehow dependent results from the fact that each prediction scheme is based on the quasi periodicity of ENSO by describing approximately the same mean period: The hybrid coupled ocean–atmosphere model is based on a concept that is in accord with the delayed action oscillator theory (Schopf and Suarez 1988; Latif and Flügel 1991; Flügel 1994) while, similarly, the POP and AR(2) models exploit the quasi-periodic nature contained in the observations.

TABLE 1. Forecast skills of the HCM–persistence (Niño-3 index) combination based on statistical weights derived from the entire prediction period (hindcast mode, combination only); $R$ is correlation skill.

| Forecast scheme | Lag (months) | | | | Prediction period |
|---|---|---|---|---|---|
| | 3 | 6 | 12 | 18 | |
| Combination | 0.79/0.48 | 0.56/0.65 | 0.27/0.77 | 0.31/0.75 | 1962–94 |
| ($R$/rmse) | 0.78/0.51 | 0.40/0.72 | 0.37/0.69 | 0.44/0.63 | 1962–72 |
| | 0.80/0.46 | 0.55/0.68 | 0.20/0.75 | 0.36/0.69 | 1972–82 |
| HCM | 0.67/0.65 | 0.50/0.80 | 0.24/0.97 | 0.16/0.97 | 1962–94 |
| ($R$/rmse) | 0.50/0.82 | 0.25/0.99 | 0.20/1.04 | 0.22/0.93 | 1962–72 |
| | 0.71/0.63 | 0.53/0.76 | −0.09/1.05 | −0.26/1.03 | 1972–82 |
| Persistence | 0.75/0.55 | 0.42/0.82 | −0.16/1.17 | −0.31/1.30 | 1962–94 |
| ($R$/rmse) | 0.74/0.55 | 0.34/0.84 | −0.35/1.14 | −0.44/1.18 | 1962–72 |
| | 0.71/0.56 | 0.29/0.89 | −0.18/1.13 | −0.13/1.11 | 1972–82 |

### b. HCM–persistence combination

The forecasts of the hybrid coupled ocean–atmosphere model are now combined with those of the persistence of the Niño-3 index SSTA (persistence 1) for the same period (1979–88). Relative to the HCM–POP and HCM–AR(2) combinations, the correlation of this combination is superior at short lead times because of the initial skill of the autocorrelation (Fig. 4a). However, beyond the 3-month lead time, the combination skill approaches the HCM prediction skill, and the persistence loses weight (Fig. 4c). Nevertheless, the rmse shows a slight improvement for all lags (Fig. 4b). The cross correlation (Fig. 4a) of the two individual prediction models decreases with lead time, which reflects the additional phase information of the persistence scheme; however, the combination skill is not significantly enhanced by this persistence at longer lead times because of the relatively high skill of the HCM model.

#### 1) SENSITIVITY TO THE PREDICTION PERIOD

To study the sensitivity of the forecast combination to the prediction period, we investigate first the dependence on the verification period using the combination of the HCM model with persistence 1. The prediction periods are 1) 1962–72, 2) 1972–82, and 3) 1962–94. The results are compared to those of the 1979–88 period.

Table 1 shows, according to Fig. 3, correlation skill $R$ and rmse for the lead times 3, 6, 12, and 18 months. Comparing the results of the prediction of the 1960s (second row) and the 1970s (third row) with the previously discussed 1980s (Fig. 4), we find a larger profit of the combination for the 1960s and 1970s, while the overall skills are considerably lower compared to the 1980s.

The variations in skill are associated with the variability of ENSO on a decadal scale that determine its predictability. Clearly, a regular occurrence of ENSO (within a decade) is easier to predict. The periodicity of ENSO is indicated by the lag-dependent correlation minimum of the autocorrelation that occurs at the half of the mean ENSO period (as mentioned in section 2a).

From Table 1 (and Fig. 4), the dominant ENSO period can be derived as follows: for the 1960s, a mean ENSO period of 36 months with a correlation value of −0.44; for the 1970s, a mean period of 24 months with a weaker minimum correlation value of −0.18; and for the 1980s, a mean period of 48 months with a relatively strong minimum correlation of −0.45 (Fig. 4a). The mean period of the 33-yr prediction period is, however, closer to a 3-yr period, as indicated by the correlation minimum at lag 18 months with a value of −0.31.

The 1980s thus represent more than the 1970s and the 1960s a ''periodic'' occurrence of ENSO which results in a relatively successful prediction in terms of correlation skill. The 1970s differ most from the periodicity since this decade was influenced by the 1972/73 El Niño event and the 1976/77 event that was more or less ''suppressed.'' However, the combination attains more profit in comparison with the 1960s and 1980s because the forecasts of the HCM and persistence are more independent from each other: The mean ENSO period that the HCM model captures does not change in contrast to the mean ENSO period that is reflected by the persistence. This additional phase information enables skill improvement—a prerequisite for a successful combination. However, the 30-yr prediction period is overall ''noisier,'' thus with a less dominant mean ENSO period and, therefore, less long-term predictability compared to the 1980s or 1960s.

#### 2) SENSITIVITY TO THE ENSO INDEX

To investigate the influence of the ENSO index used in the combination, we have combined the HCM in an extended study (not shown) additionally with (i) the persistence of the Southern Oscillation index (SOI), an index highly anticorrelated with the Niño-3 index, and (ii) the persistence of an SSTA equatorially averaged over the region 170°E–90°W (persistence 2). The prediction period chosen ranges from 1962 to 1991, and the forecasts are verified against the Niño-3 SSTA index.

The HCM–persistence 2 combination outperforms the HCM–persistence 1 combination at longer lead times
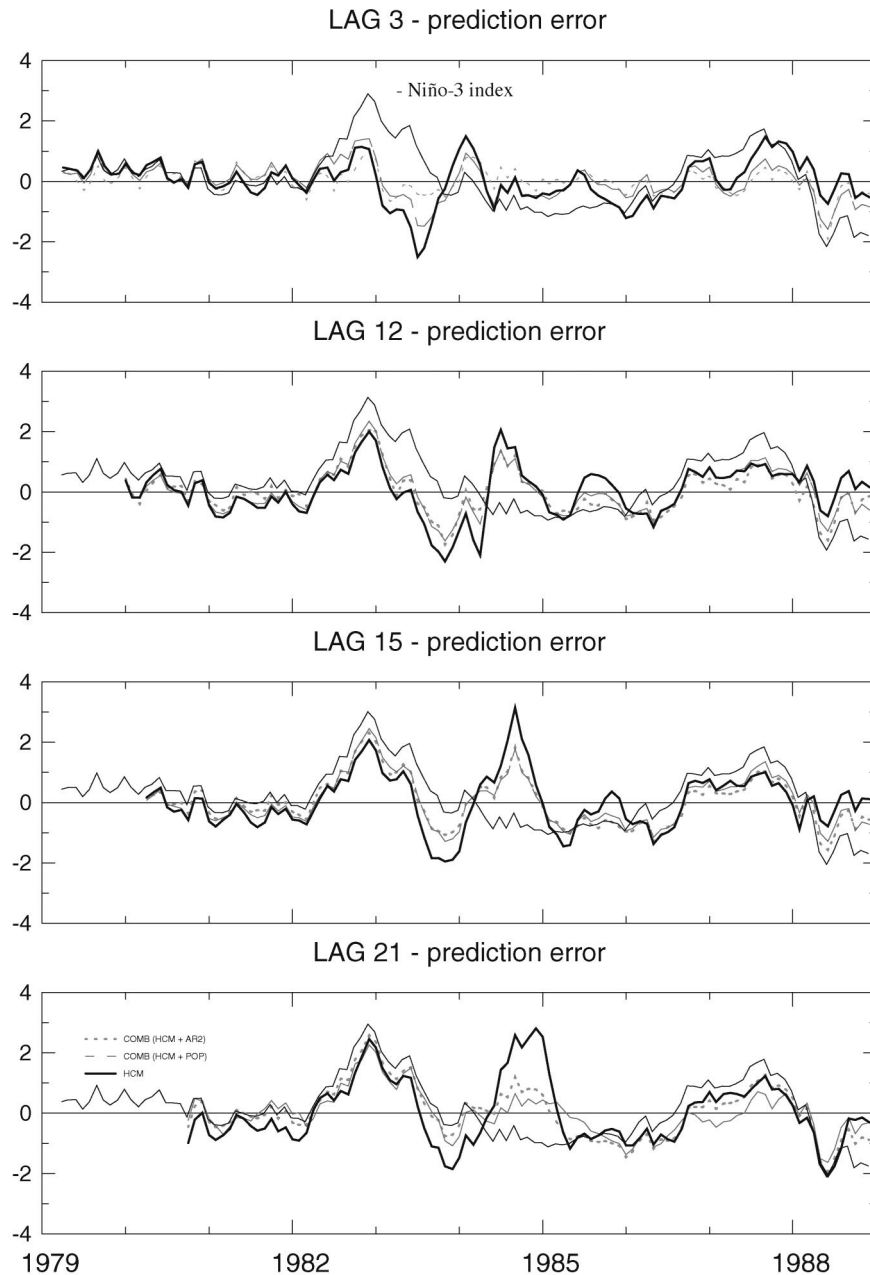
LAG 3 - prediction error

LAG 12 - prediction error

LAG 15 - prediction error

LAG 21 - prediction error

FIG. 3. Niño-3 index prediction errors of the HCM–POP combination (dashed) and HCM–AR(2) combination (dotted) vs HCM (thick solid line) and observation (thin solid line) for different lead times (3, 12, 15, and 21 months) and the prediction period Jan 1979–Dec 1988.

(Fig. 5a versus Fig. 4a), mainly because of the weaker correlation of the HCM forecast at higher lead times, which are compensated in this case by the relatively high anticorrelation of the persistence 2 at these lead times (i.e., approximately −0.5 at a lead time of 21 months; Fig. 5a). Specifically, the SSTA of the western Pacific basin is in average less noisy compared to the Niño-3 index because of the relatively deep mixed layer in this region. Therefore the anomalous temperature of the subsurface ocean as well as the upper-ocean heat content are good candidates for the combination technique; both are less affected by atmospheric noise. In contrast, the use of the SOI results in a lower profit of the forecast combination (results are shown in Metzger 1996, 1997). This indicates that, in general, the noisier atmospheric ENSO indices are less suitable for this type of forecast combination compared to oceanographic ones.
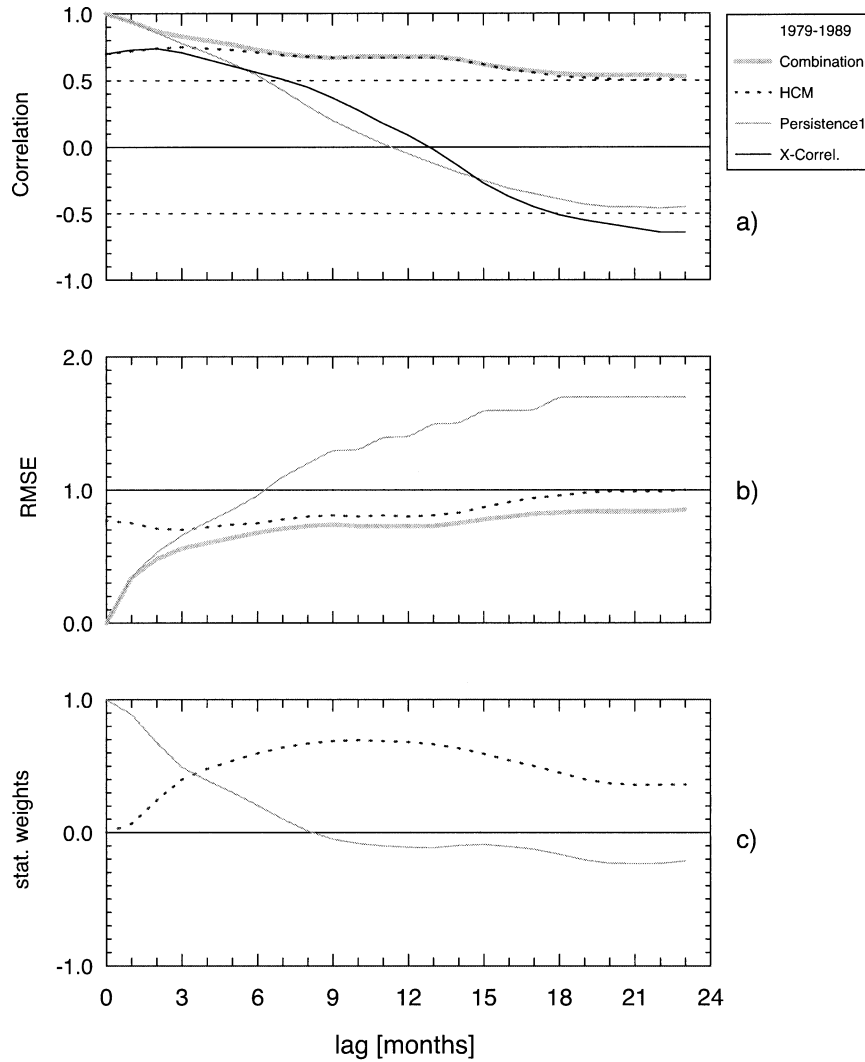
FIG. 4. (a) The correlation skill, (b) rmse, and (c) statistical weights of the HCM Niño-3 index predictions (dotted); persistence 1, i.e., of Niño-3 index (thin gray); and their combination (thick gray) for the prediction period Jan 1979–Dec 1988. The cross correlation between HCM predictions and persistence is shown in (a) (thin solid line).

### 3) SEASONALITY

Both persistence of the SOI and SSTA exhibit a strong seasonality (Wright 1985). This is also the case for the prediction skill of the HCM model (Latif and Flügel 1991). Therefore, we investigate the influence of the seasonality on the forecast combination. Analogous to the previous combination experiments, we combine the HCM forecasts with persistence 1 (of the Niño-3 index) for the prediction period 1962–91. The forecasts are initialized monthly and averaged over seasons: March–April–May (MAM) in spring, June–July–August (JJA) in summer, September–October–November (SON) in autumn, and December–January–February (DJF) in winter, which yields 90 individual forecasts for each set.

The seasonal correlation skills of the HCM, persistence, and combination forecasts are shown in Fig. 6. Consistent with Wright (1985), Latif and Flügel (1991), and Balmaseda et al. (1995), the same seasonal behavior of the forecasts is found: a drop in predictability around March for forecast initialization in MAM and JJA at lead times of 12 and 9 months, respectively (Figs. 6a,b), with a weaker drop around March for the forecasts initialized in SON and DJF at 6- and 3-month lead times, respectively (Figs. 6c,d). The combination, however, achieves the largest overall profit at higher lead times from the forecast initialization in MAM (Fig. 6a). In this case, the SSTA is more persistent for the rest of the prediction period, resulting in a higher autocorrelation from which the combination skill profits.
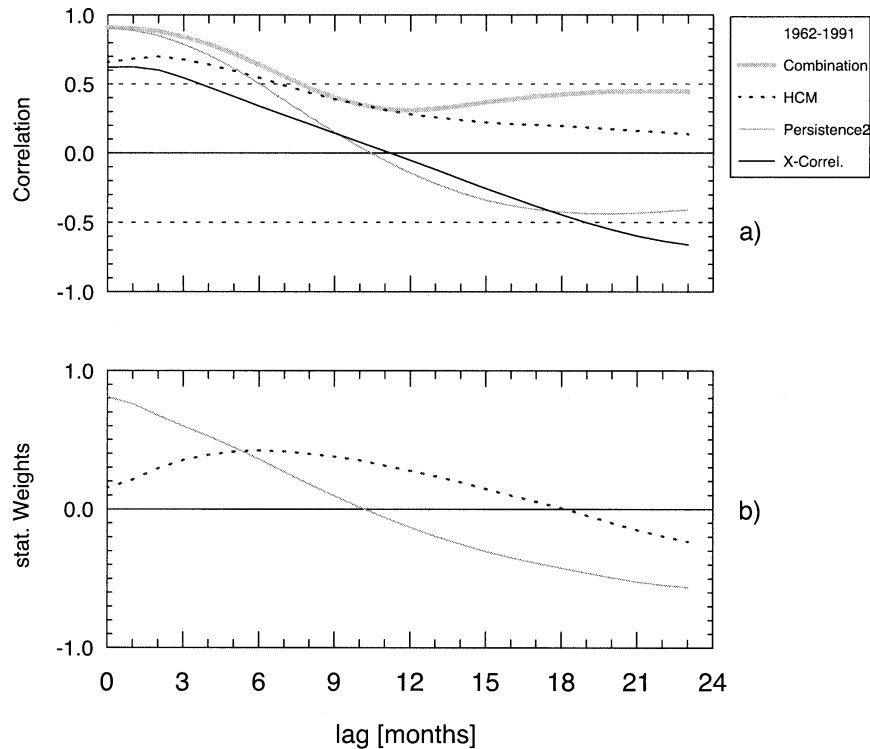
FIG. 5. (a) The correlation skill and (b) statistical weights of the HCM Niño-3 index predictions (dotted); persistence 2, i.e., of equatorially averaged (170°E–90°W) SSTA (thin gray); and their combination (thick gray) for the prediction period Jan 1962–Dec 1991. The cross correlation between HCM predictions and persistence is shown in (a) (thin solid line).

#### 4) ESTIMATION TECHNIQUE OF THE STATISTICAL WEIGHTS

For the previously shown combination forecasts, the statistical weights (that are used to combine the forecast ensembles) were derived from the entire prediction period. To derive the statistical weights from a period independent of the prediction period, we divide, for the sake of simplicity, the prediction period into halves, deriving the statistical weights from one half while predicting the other and averaging over both forecast sets. This procedure has been applied to the previous combinations (Table 1, Fig. 4). The results of the combination are shown in Table 2. In comparison to Table 1 and Fig. 4 the prediction skills now depend more strongly on the prediction period. In terms of the correlation skill, the profit of the combination forecast in predicting the phase of ENSO often vanishes completely, while in terms of the rmse some profit is still achieved in predicting the amplitude, especially at longer lead times. Regarding the prediction of the phase, the correlation values demonstrate that the predictions improve as ENSO has been evolved more regularly in the entire time period. In other words, the ENSO statistics should be similar for the prediction period and the period from which the statistical weights are derived to maximize the profit of this combination technique, especially if applied to noise-reduced time series (see the note by the

discussion of artificial skill). For instance, the 1970s, because of the suppressed 1976/77 El Niño event, are not so well predicted by the combination forecasts because the statistical weights have been derived from a period where ENSO evolved more regularly [see section 3b(1)]. For same reason, it does not make sense to derive the statistical weights from the 1970s and predict the 1980s. However, this phase information inherent in the coupled atmosphere–ocean system can be utilized by this combination technique.

#### c. Persistence–persistence combination

To give an example of how the phase information inherent in the coupled atmosphere–ocean system can be utilized by this combination technique, we focus now on a forecast combination of two persistence schemes. To include the latest ENSO events, we focus on the prediction period January 1982 to December 2002 for which we use the SST of Reynolds et al. (2002) (available online at http://dss.ucar.edu/datasets/ds277.0/data/oiv2/). Besides constructing the SST anomalies, the data are not preprocessed (smoothed), thus including the annual cycle.

#### 1) UTILIZING THE ENSO DYNAMICS

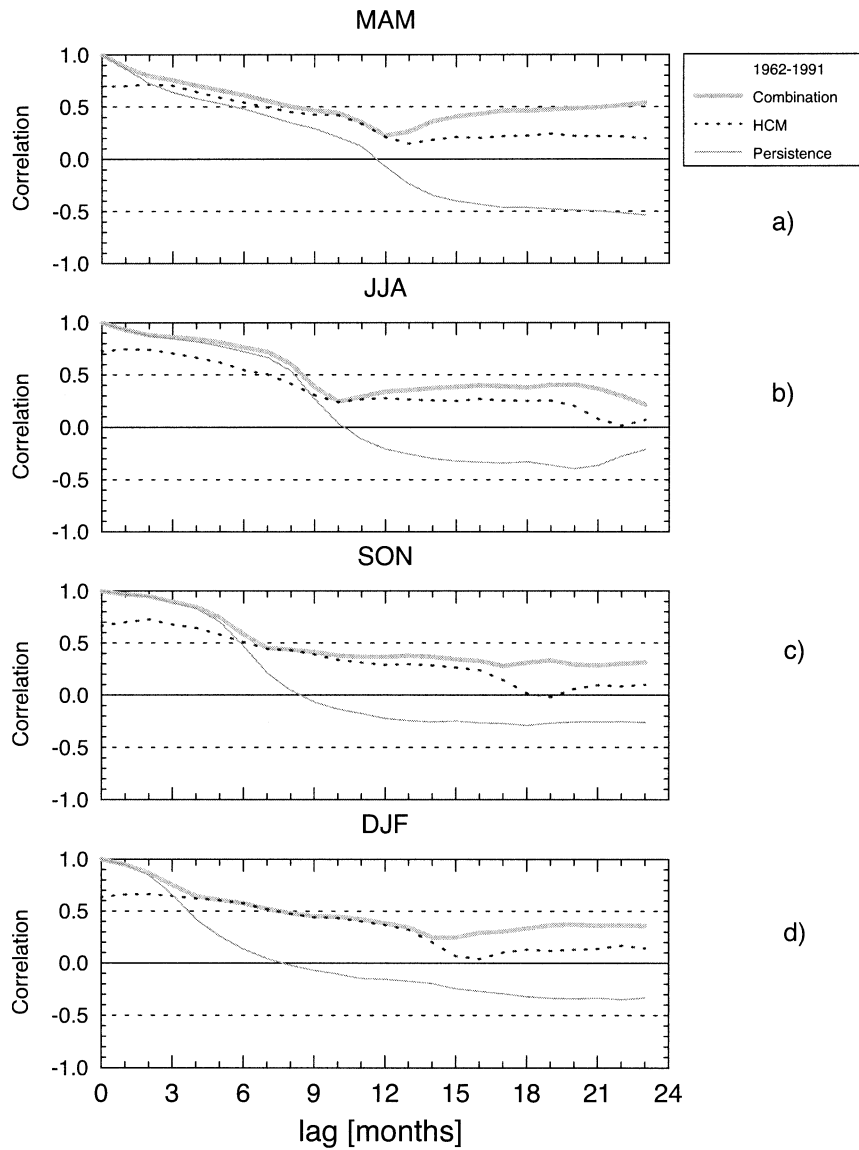Figure 7 shows skill of four forecast combinations: three combinations of the persistence of the Niño-3 in-

FIG. 6. The correlation skill of the HCM Niño-3 index predictions (dotted), persistence (Niño-3 index) (thin gray), and their combination (thick gray) are shown for the prediction period Jan 1962–Dec 1991 and stratified according to the seasons (a) spring (MAM), (b) summer (JJA), (c) autumn (SON), and (d) winter (DJF).

TABLE 2. Forecast skills of the HCM–persistence (Niño-3 index) combination based on statistical weights independently derived (forecast mode); $R$ is correlation skill.

| Forecast scheme | Lag (months) | | | | Prediction period |
|---|---|---|---|---|---|
| | 3 | 6 | 12 | 18 | |
| Combination | 0.79/0.49 | 0.55/0.66 | 0.10/0.84 | 0.19/0.77 | 1962–94 |
| ($R$/rmse) | 0.77/0.51 | 0.29/0.79 | −0.18/0.82 | 0.33/0.68 | 1962–72 |
| | 0.53/0.71 | −0.13/0.85 | −0.01/0.72 | 0.15/0.59 | 1972–82 |
| | 0.78/0.49 | 0.63/0.60 | 0.54/0.66 | 0.49/0.67 | 1979–88 |
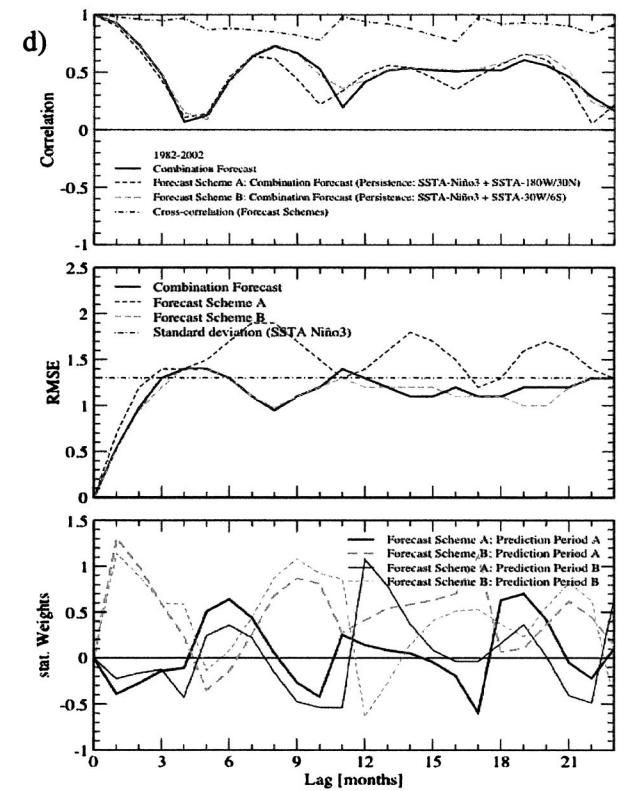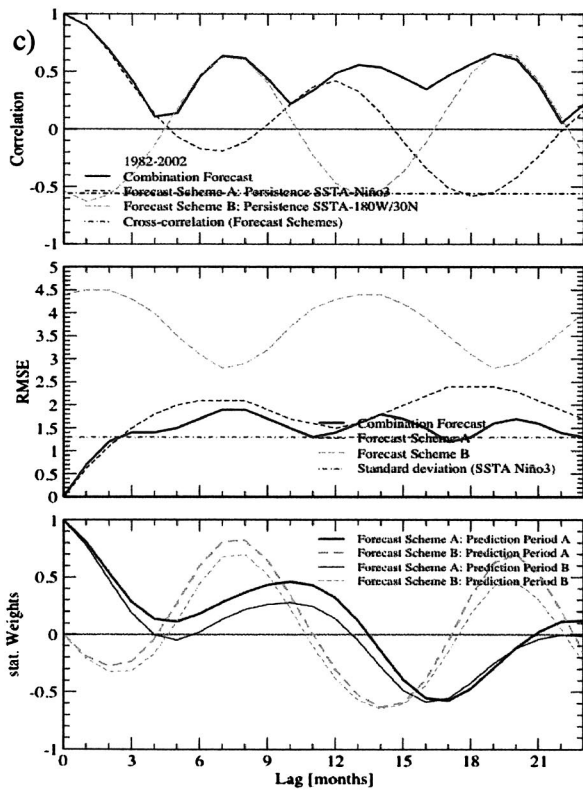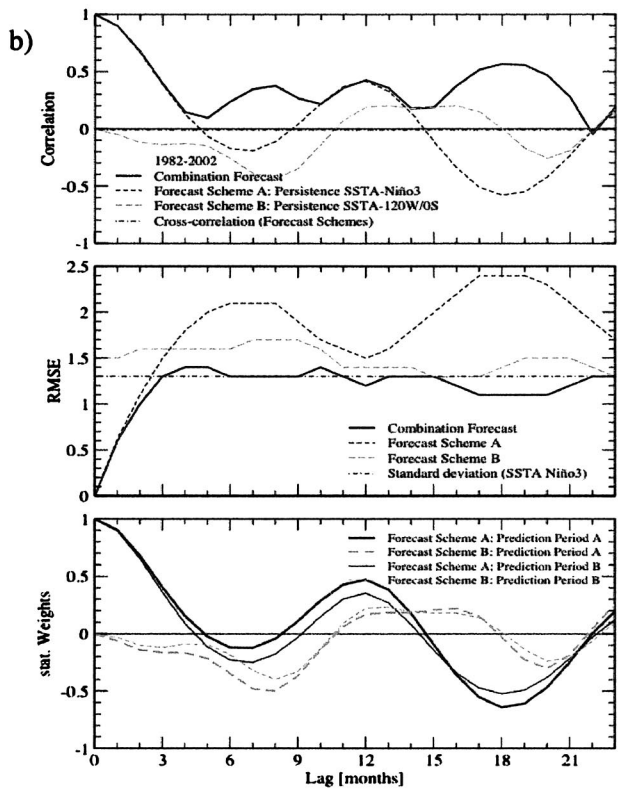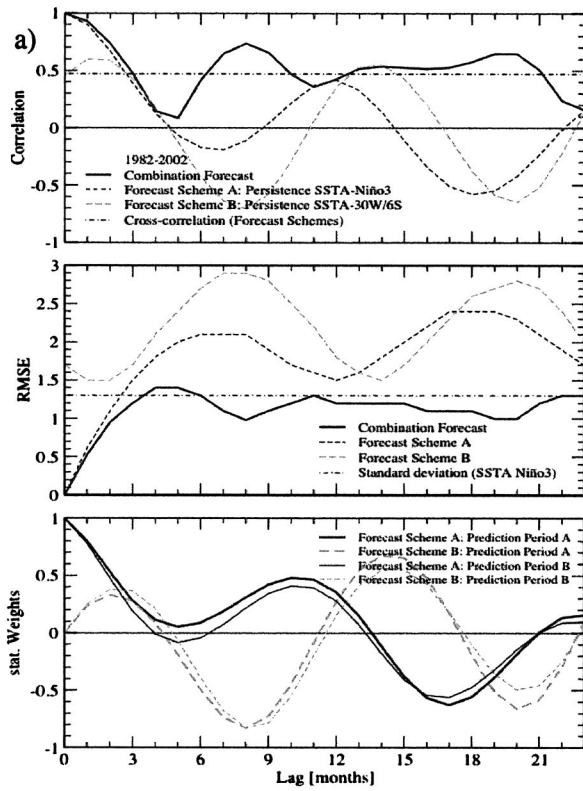
dex SSTA (persistence 1) with persistence of an index SSTA at (a) 6°S, 30°W, (b) at 0°S, 120°W, and (c) at 30°N, 180° (the region averaged is given in the figure caption). Set (d) shows the skill of the combination forecast based on the combination forecasts of (a) and (c). The statistical weights have been derived independently (as described above). Note that the locations have been chosen only exemplarily (based on some sensitivity experiments); other locations might be even more suited.

The results show that in general the prediction skill is improved by the forecast combination of the two persistence forecasts. While the largest profit is achieved for set (a), the combination of two combination forecasts [set (a) and (c); shown in Fig. 7d] shows no additional profit. Interestingly, the SSTA persistence of the remote location 6°S, 30°W (a spot in the Atlantic at the Brazilian coast) contains some additional phase information with respect to the Niño-3 index SSTA—a consequence of the coupled atmosphere–ocean system. Remarkably, even the most simple case of a forecast combination scheme almost yields comparable results relative to the more sophisticated schemes shown before if the phase information of the coupled atmosphere–ocean system is utilized. However, the combinations of the other two persistence schemes (one is located in the equatorial west Pacific and one in the North Pacific) yield less profit for this time period—also reflecting the ENSO dynamics.

Note that the dynamics are indicated by the phase of the persistence (with respect to the Niño-3 index), that is, the correlation value at lag zero. The persistence of the Niño-3 index is simply the autocorrelation; the correlation therefore starts at lag zero with a value of one (short dashed lines in Figs. 7a–c, top panels). In contrast, the persistence of the other three SSTA locations indicate, with the lag zero correlation value, the phase shift relative to the Niño-3 index (long dashed lines): This phase shift increases from set (a) to (c) for this prediction period. While the SSTA of the Niño-3 index is leading by 1–2 months compared to the SSTA of the Brazilian coast (Fig. 7a), the SSTA in the North Pacific is leading by approximately 90° (Fig. 7c). However, at other prediction periods this might be different. Then, these SSTA series might contain more (or less) profit for the combination prediction of the Niño-3 index, as previously indicated by Figs. 4 and 5. Another example of the utilization of the ENSO dynamics is shown in Fig. 8a (which is rather similar to Fig. 7a), to be discussed in the following.

## 2) ARTIFICIAL SKILL

First, a comment on artificial skill is in order. The previously shown numerical and statistical combination forecasts are to some extent hindcast because the statistical weights (of the forecast combination) were derived from the entire prediction period. Furthermore, these combinations are based on "smoothed" time series (to remove noise, as mentioned earlier), which introduces extra dependence (of the order of 12 months) into the series that can boost the short-term skill artificially. Note that the amount of artificial skill resulting from the estimation of the coupling matrix of the HCM is negligible (M. Flügel 1995, personal communication).

We therefore distinguish in the following between so-called hindcast combinations and forecast combinations in that sense all previously discussed combinations are hindcast combinations, except for those shown in Fig. 7. Note that in contrast to a hindcast combination, statistical weights are derived for a forecast combination from a period independent of the prediction period (as described above). As long as the ENSO statistics do not change much within the prediction period, the artificial skill introduced by the hindcast mode is negligible compared to the forecast mode. Otherwise, differences may arise and the combination skill may be even lower than those of the individual schemes (Table 2 versus Table 1).

To quantify the artificial skill, we focus on a combination of the persistence of the Niño-3 index SSTA (persistence 1) and a persistence of an index SSTA at 60°S, 180°. The dependencies are illustrated in Fig. 8. Set (a) shows skill and weights of the forecast combination and the individual schemes for the forecast mode and "raw" SSTA data (similar to Figs. 7a–c), set (b) shows the same for "smoothed" data, sets (c) and (d) show the same as (a) and (b) for the hindcast mode. While differences between the forecast and hindcast mode are small for correlation skill and rmse of the combinations based on the raw SSTA data, differences are in particular larger for the correlation skill of the combinations based on the smoothed SSTA data. Nevertheless, both correlation skill and rmse are improved by the combinations for the forecast mode almost to the same extent as compared to the hindcast mode. This is also true for longer lead times and the smoothed data [sets (b) and (d)]. Interestingly, the combinations based on the raw data have even an higher skill at lead times around the half of the ENSO period (about 19 months for the raw and 23 months for the smoothed data) than

←

FIG. 7. (a)–(d) The prediction skill of forecast combinations based on raw SSTA data (that include the annual cycle): combination forecasts (thick gray) of the persistence of the Niño-3 SSTA index (short dashed) and persistence of an index SSTA at 6°S, 30°W (averaged over 6°–9°S, 30°–33°W) (long dashed); persistence of an index SSTA at 0°S, 120°W (averaged over 5°S–5°N, 120°–130°W) (long dashed); and persistence of an index SSTA at 30°N, 180° (averaged over 28°–32°N, 175°–185°W) (long dashed). Set (d) shows the skill of the combination forecast based on the combination forecasts (a) and (c). Each set shows (top) the correlation skill (and cross correlation), (middle) the rmse, and (bottom) the statistical weights of two forecast schemes. The prediction period is Jan 1982–Dec 2002.
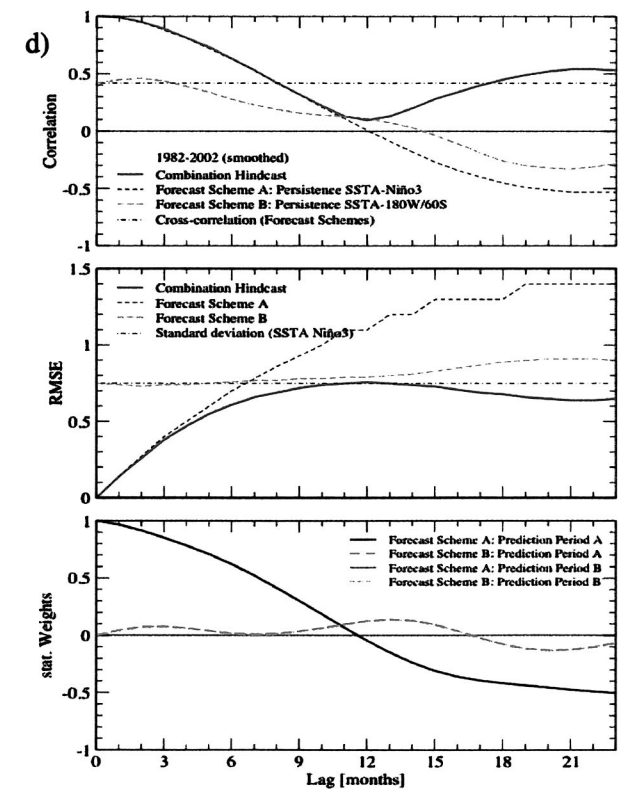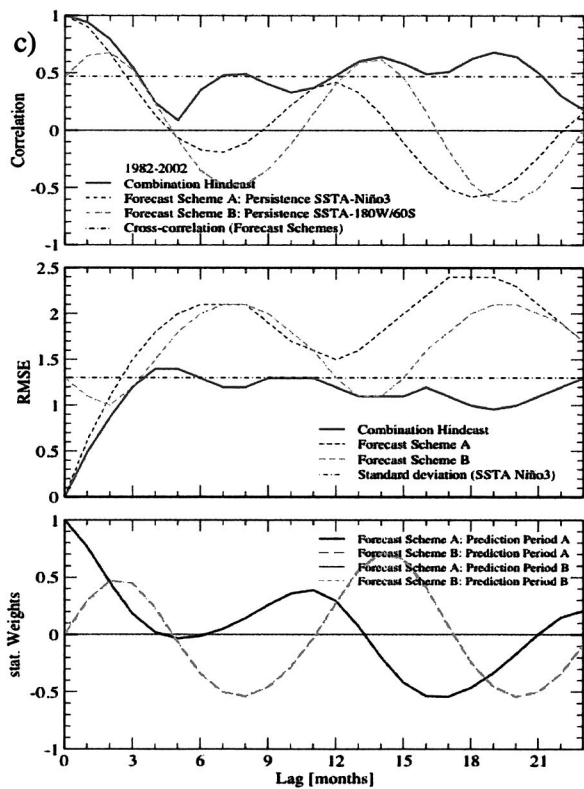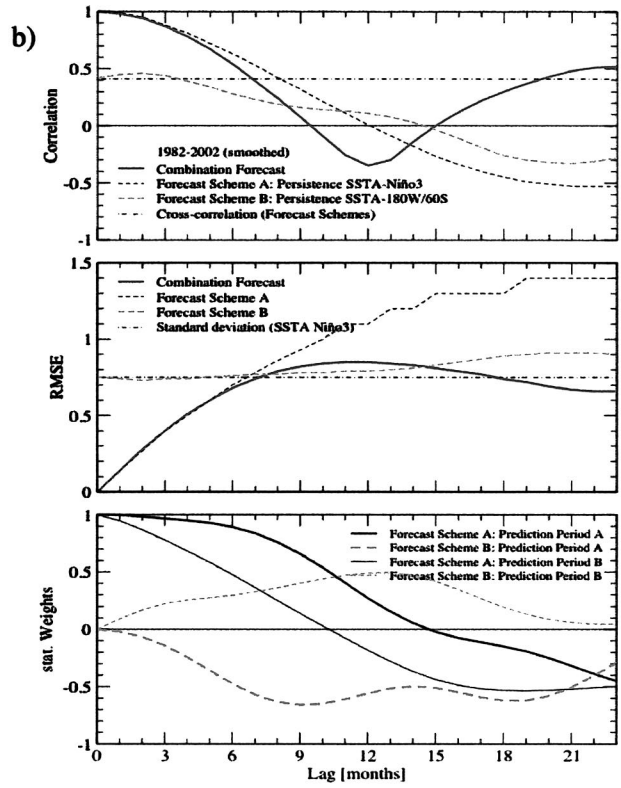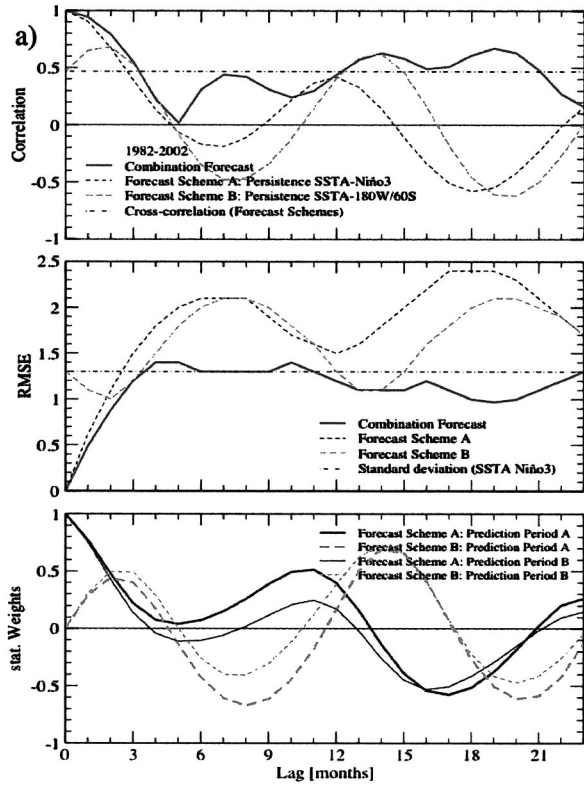
their counterparts that are based on the smoothed data. This is true for both the forecast and hindcast mode, although the annual cycle has not been removed. Note that the statistical weights of the forecast mode equal each other and the one of the hindcast mode more (for each scheme) compared to the ones that are based on the smoothed data. For the latter case the persistence of the index SSTA at 60°S, 180° has less statistical weight for all lead times in contrast to the raw data case. This indicates that the annual cycle at this particular location (this spot lies within the Antarctic Circumpolar Gyre) also contains useful phase information for the prediction of the Niño-3 index SSTA at this prediction period that can be utilized by the forecast combination.

### 3) FORECAST SKILL

To illustrate the differences between combination forecast and hindcast, we show in Fig. 9 the individual combination forecasts of the Niño-3 index SSTA for the prediction period January 1982 to December 2002. Shown are the lag 0–23 combination forecasts based on the index SSTA at 60°S, 180° and the lag 23 combination forecasts and hindcasts, all based on smoothed data (according to Figs. 8b and 8d). Additionally, the lag 19 (this lag yields the highest correlation skill at longer lead times) combination forecasts, which are based on the raw data (according to Fig. 8a), are included. Also shown are the smoothed and raw Niño-3 index SSTA. While the latter time series is much more noisy (because of the annual cycle), the former highlights the ENSO events. Interestingly, the prediction of the phase and amplitude is much better for the noisier time series compared to the smoothed SSTA, although the lead time of 19 months is far beyond the predictive capability of a persistence forecast that usually keeps memory (i.e., a high autocorrelation) only up to 3 months. Nevertheless, the combination forecast of the smoothed SSTA is also still rather good for the phase at this long lead time, although the amplitude of ENSO extremes of the 1997/98 event is pretty much underestimated, while the onset of the ongoing event (2002/03) is somewhat overestimated.

More important is the fact that the skill of the forecast mode is almost equal to the one of the hindcast mode, which indicates that in this case only little artificial skill is introduced by the hindcast mode. Even more important is the fact that obviously the most simple forecast combination has some predictive capability that can be used for predicting the phase and amplitude of ENSO up to 2 yr in advance. It should be noted, however, that the predictions of the other lags are in less agreement, as indicated by the lag 0–23 ensemble. For instance, lag 5 yields no predictive capability at all (nor for the phase or amplitude of the annual cycle) for the individual forecast of this combination forecast if based on the raw data (not shown), which is in accord with the zero correlation shown in Fig. 8a. Note further that the verification period ends with December 2002, so that a forecast of the future is given until mid-2004 for the raw SSTA (based on lag 19); the smoothed forecast is approximately a half year shorter because of to the filtering with a 13-month running mean.

## 4. Summary and conclusions

Numerical and statistical predictions of simplified models are linearly combined to enhance predictability of the El Niño–Southern Oscillation (ENSO) phenomenon. The crucial parameters that determine the maximum profit achievable with this combination technique are identified. The combination experiments are based for the sake of simplicity only on simplified numerical and statistical prediction models: the hybrid coupled model (HCM) and the POP, AR(2), and persistence scheme. For the latter we have used various ENSO indices: (i) anomalous sea surface temperature (SSTA) of the Niño-3 index, (ii) an equatorially averaged SSTA of the tropical Pacific, index SSTAs (regional averages) at 6°S, 30°W; 0°S, 120°W; 30°N, 180°; and 60°S, 180°, as well as the Southern Oscillation index (SOI). The predictand was the Niño-3 index SSTA. The results can be summarized as follows.

In general, the variance of the prediction error can be reduced by an error-minimizing forecast combination, and the correlation skill can be improved. The linear combination achieves the largest profit if the forecasts of the individual schemes contain a different phase information that is inherent in predictions of a quasi-periodic process, such as ENSO. It is this phase information that is utilized by the forecast combination.

The error of predicting the phase and amplitude is not consistently reduced by the combination. It depends on the type of prediction schemes used for the combination (numerical, statistical, and the type of ENSO index) as well as on the prediction period and the determination of the statistical weights for the forecast combination. For instance, the combinations of numerical

←

FIG. 8. (a) The prediction skill of the forecast combination for raw SSTA data (that include the annual cycle), (b) same as (a), but for the smoothed data (with a 13-month running mean), (c), (d) same as (a) and (b), respectively, but for a hindcast combination (with statistical weights derived from the entire prediction period). Each set shows (top) the correlation skill (and cross correlation), (middle) the rmse, and (bottom) the statistical weights of two forecast schemes: the persistence of the Niño-3 SSTA index (short dashed) and the persistence of an index SSTA at 60°S, 180° (averaged over 58°–62°S, 175°–185°W) (long dashed), and their combination (thick), for the prediction period Jan 1982–Dec 2002. Note that in contrast to the hindcast combination, the prediction period was halved to derive the statistical weights for the forecast combination. The first half is marked with thick lines, the second half with thin lines [sets (a) and (b), bottom panels].
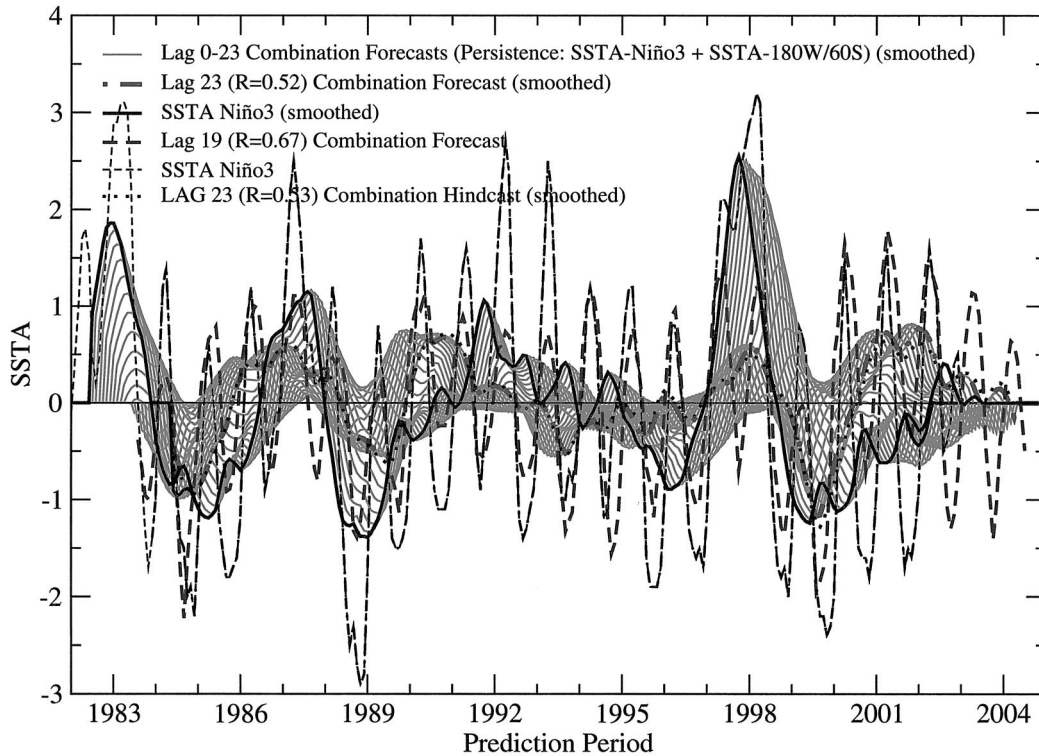
FIG. 9. Niño-3 index SSTA forecasts for the prediction period Jan 1982–Dec 2003. Shown are the raw observation, Reynolds SSTA, (thin dashed line), the smoothed observation (13-month running mean) (thick solid line), the lag 0–23 combination forecasts of the persistence of the Niño-3 SSTA index, and the persistence of the index SSTA at 60°S, 180° corresponding to Fig. 7 (thin solid lines). Additionally, the lag 23 combination forecast (thick dotted–dashed line) and hindcast (thick dotted line), which are based on the smoothed data, are shown, as well as the lag 19 combination forecast (thick dashed line), which is based on the raw data. Note that the lag 19 combination forecast and lag 23 combination forecast and hindcast correspond to the highest (at higher lead times) correlation values that are shown in Figs. 8a, 8b, and 8d, respectively, and that the combination forecast and hindcast are more or less on top of each other.

HCM forecasts with the statistical POP forecasts showed modest skill improvements at lead times up to one year for the correlation skill, while the rmse is reduced for the entire 23-month lead time (Fig. 1). For the HCM–AR(2) combination, the skill is improved only beyond the 6-month lead time, with a larger reduction of the error in amplitude (rmse) compared to the phase error (Fig. 2). The HCM–persistence (Niño-3 index SSTA) combination achieves for the 1980s prediction period the lowest skill profit because of the relatively high prediction skill of the HCM forecast (Fig. 4). Nevertheless, the prediction errors of the individual combination forecast are shown to be lower compared to those of the HCM model (Fig. 3).

In contrast to the combination of the HCM forecasts with the persistence of the Niño-3 index SSTA, the combination with the persistence of the equatorially averaged SSTA of the Pacific results in an improved combination skill for a 30-yr prediction period because of the phase information inherent in this persistence (Fig. 5).

This phase information is also seasonally dependent.

HCM–persistence forecast combinations initialized in spring showed the largest profit at higher lead times since the Niño-3 index SSTA is then more persistent for the rest of the prediction period, resulting in a higher autocorrelation from which the combination skill profits (Fig. 6).

The combination of the HCM predictions with persistence of the Niño-3 index SSTA appeared to be quite sensitive to the verification period (Table 1) since the autocorrelation exhibits a strong variability, also on a decadal scale. Furthermore, the persistence of atmospheric anomalies like SOI are less suited for a forecast combination relative to oceanic ones because of higher noise levels. Likewise, the Niño-3 index is on average noisier than the SSTA averaged along the Pacific equator, and forecast combinations using the equatorially averaged SSTA are hence somewhat superior to those using the persistence of Niño-3 index SSTA.

In the coupled atmosphere–ocean system, different regions contain a different phase information with respect to the Niño-3 index SSTA for a certain prediction period. For instance, forecast combinations of the Niño-

3 index SSTA persistence with the persistence of index SSTAs at various locations showed that the most simple forecast combination has some predictive capability if the phase information of the coupled atmosphere–ocean system is utilized (Fig. 7). The combination with the persistence of an index SSTA at 60°S, 180° even yields for the prediction period January 1982–December 2002 some useful individual forecasts for up to 2 yr in advance for both the phase and amplitude of ENSO and the annual cycle (Figs. 8 and 9).

The amount of artificial skill that is introduced if the combination weights are derived from the entire prediction period depends on the prediction period. As long as the ENSO statistics do not change much within the prediction period, the artificial skill introduced by the hindcast mode is negligible compared to the forecast mode (Fig. 8). Otherwise, differences may arise and the combination skills may even be lower than those of the individual schemes (Table 2 versus Table 1). Furthermore, if the combinations are based on ''smoothed'' time series, some extra dependence (of the order of 12 months) is introduced into the series that can alter the ENSO statistics of the prediction period to increase the hindcast skill artificially, even if hindcast skill and forecast skill of the ''raw'' data are almost the same (Fig. 8).

Finally, the forecast combination technique can be applied in a much more elaborate way. One can imagine a more complex forecast combination model (to be discussed separately) that makes use of different (physical) ENSO indices that are variable in time and space depending on the prediction period to optimally utilize the phase information that is inherent in the coupled atmosphere–ocean system.

## REFERENCES

Balmaseda, M. A., D. L. T. Anderson, and M. K. Davey, 1994: ENSO prediction using a dynamical ocean model coupled to statistical atmospheres. *Tellus,* **46A,** 497–451.
——, M. K. Davey, and D. L. T. Anderson, 1995: Decadal and seasonal dependence of ENSO prediction skill. *J. Climate,* **8,** 2705–2715.
Barnett, T. P., M. Latif, N. Graham, M. Flügel, S. Pazan, and W. White, 1993: ENSO and ENSO-related predictability. Part I: Prediction of equatorial Pacific sea surface temperature with a hybrid coupled ocean–atmosphere model. *J. Climate,* **6,** 1545–1566.
Barnston, A. G., and C. F. Ropelewski, 1992: Prediction of ENSO episodes using canonical correlation analysis. *J. Climate,* **5,** 1316–1345.
Blumenthal, M. B., 1991: Predictability of a coupled ocean–atmosphere model. *J. Climate,* **4,** 766–784.
Bjerknes, J., 1969: Atmospheric teleconnections from the equatorial Pacific. *Mon. Wea. Rev.,* **97,** 163–172.

Cane, M. A., S. E. Zebiak, and S. C. Dolan, 1986: Experimental forecasts of El Niño. *Nature,* **321,** 827–832.
Coelho, C. A. S., S. Pezzulli, M. Balmaseda, F. J. Doblas-Reyes, and D. B. Stephenson, 2004: Forecast calibration and combination: A simple Bayesian approach for ENSO. *J. Climate,* in press.
Flügel, M., 1994: Untersuchungen zu Mechanismus und Vorhersagbarkeit von ENSO miteinem vereinfachten gekoppelten Ozean–Atmosphären Modell. Ph.D. dissertation, Universität Hamburg, 65 pp. [Available from Max-Planck-Institut für Meteorologie, Bundesstraße 55, D-20146 Hamburg, Germany.]
Fraedrich, K., and L. M. Leslie, 1987: Combining predictive schemes in short-term forecasting. *Mon. Wea. Rev.,* **115,** 1640–1644.
——, and N. R. Smith, 1989: Combining predictive schemes in long-range forecasting. *J. Climate,* **2,** 291–294.
Goldenberg, S., and J. O'Brien, 1981: Time and space variability of the tropical Pacific wind stress. *Mon. Wea. Rev.,* **109,** 1190–1207.
Goswami, B., and J. Shukla, 1991a: Predictability of a coupled ocean–atmosphere model. *J. Climate,* **4,** 3–22.
——, and ——, 1991b: Predictability and variability of a coupled ocean–atmosphere model. *J. Mar. Syst.,* **1,** 217–228.
Grötzner, A., M. Latif, A. Timmermann, and R. Voss, 1999: Interannual to decadal predictability in a coupled ocean–atmosphere general circulation model. *J. Climate,* **12,** 2607–2624.
Hasselmann, K., 1988: PIPs and POPs: The reduction of complex dynamical systems using principal interactions and oscillation patterns. *J. Geophys. Res.,* **83,** 11 015–11 021.
Honerkamp, J., 1990: *Stochastische Dynamische Systeme: Konzepte numerische Methoden, Datenanalysen.* Weinheim, 340 pp.
Ji, M., D. W. Behringer, and A. Leetmaa, 1998: An improved coupled model for ENSO prediction and implications for ocean initialization. Part II: The coupled model. *Mon. Wea. Rev.,* **126,** 1022–1034.
Johnson, S. D., D. S. Battisti, and E. S. Sarachik, 2000a: Empirically derived Markov models and prediction of tropical Pacific sea surface temperature anomalies. *J. Climate,* **13,** 3–17.
——, ——, and ——, 2000b: Seasonality in an empirically derived Markov model of tropical Pacific sea surface temperature anomalies. *J. Climate,* **13,** 3327–3335.
Krishnamurti, T. N., C. M. Kishtawal, T. LaRow, D. Bachiochi, Z. Zhang, C. E. Williford, S. Gadgil, and S. Surendran, 2000: Multimodel ensemble forecasts for weather and seasonal climate. *J. Climate,* **13,** 4196–4216.
Latif, M., 1987: Tropical ocean circulation experiments. *J. Phys. Oceanogr.,* **17,** 246–263.
——, and M. Flügel, 1991: An investigation of short-range climate predictability in the tropical Pacific. *J. Geophys. Res.,* **96,** 2661–2673.
——, A. Sterl, E. Maier-Reimer, and M. M. Junge, 1993: Structure and predictability of the El Niño/Southern Oscillation phenomenon in a coupled ocean–atmosphere general circulation model. *J. Climate,* **6,** 6701–6708.
——, T. P. Barnett, M. A. Cane, M. Flügel, N. E. Graham, H. von Storch, J. S. Xu, and S. E. Zebiak, 1994: A review of ENSO prediction studies. *Climate Dyn.,* **9,** 167–179.
——, and Coauthors, 1998: TOGA review paper: Predictability and prediction. *J. Geophys. Res.,* **103** (C7), 14 375–14 393.
Mason, S. J., L. Goddard, N. E. Graham, E. Yulaeva, L. Sun, and P. A. Arkin, 1999: The IRI seasonal climate prediction system and the 1997/98 El Niño event. *Bull. Amer. Meteor. Soc.,* **80,** 1853–1873.
Metzger, S. M., 1996: Kombination von Ensemble-Klimavorhersagen am Beispiel ENSO. Diploma thesis, Max-Planck-Institut für Meteorologie, Examensarbeit 39, 75 pp. [Available from Max-Planck-Institut für Meteorologie, Bundesstraße 55, D-20146 Hamburg, Germany.]
——, 1997: Combination of ensemble climate predictions using the example of ENSO. *Oceanogr. Lit. Rev.,* **44,** 22.

Neelin, J. D., M. Latif, and F. F. Jin, 1994: Dynamics of coupled ocean–atmosphere models: The tropical problem. *Annu. Rev. Fluid Mech.,* **26,** 617–659.

Oberhuber, J. M., E. Roeckner, M. Christoph, M. Esch, and M. Latif, 1998: Predicting the 1997 El Niño event with a global climate model. *Geophys. Res. Lett.,* **25,** 2273–2276.

O'Brien, J., A. Rosati, and E. Schneider, 1998: TOGA review paper: Predictability and prediction. *J. Geophys. Res.,* **103C,** 14 375–14 393.

Peng, P., A. Kumar, H. Van den Dool, and A. Barnston, 2002: An analysis of multimodel ensemble predictions for seasonal climate anomalies. *J. Geophys. Res.,* **107**, 4710, doi.10.1029/2002JD002712.

Penland, C., and T. Magorian, 1993: Prediction of Niño 3 sea surface temperature using linear inverse modeling. *J. Climate,* **6,** 1067–1076.

Philander, S. G. H., 1990: *El Niño, La Niña, and the Southern Oscillation.* Academic Press, 289 pp.

Rajagopalan, B., U. Lall, and S. E. Zebiak, 2002: Categorical climate forecasts through regularization and optimal combination of multiple GCM ensembles. *Mon. Wea. Rev.,* **130,** 1792–1811.

Reynolds, R. W., N. A. Rayner, T. M. Smith, D. C. Stokes, and W. Wang, 2002: An improved in situ and satellite SST analysis for climate. *J. Climate,* **15,** 1609–1625.

Schopf, P., and M. Suarez, 1988: Vacillations in a coupled ocean–atmosphere model. *J. Atmos. Sci.,* **45,** 549–566.

Stockdale, T. N., D. L. T. Anderson, J. O. S. Alves, and M. A. Balmaseda, 1998: Global seasonal rainfall forecasts using a coupled ocean–atmosphere model. *Nature,* **392,** 370–373.

von Storch, H., T. Bruns, I. Fischer-Bruns, and K. Hasselmann, 1988: Principal oscillation analysis of the 30- to 60-day oscillation in a GCM. *J. Geophys. Res.,* **93** (D9), 11 022–11 036.

Tangang, F. T., W. W. Hsieh, and B. Tang, 1997: Forecasting the equatorial Pacific sea surface temperature by neural network models. *Climate Dyn.,* **13,** 135–147.

Thompson, P. D., 1977: How to improve accuracy by combining independent forecasts. *Mon. Wea. Rev.,* **105,** 228–229.

Van den Dool, H. M., and A. G. Barnston, 1995: Forecasts of global sea surface temperature out to a year using the constructed analogue method. *Proc. 19th Annual Climate Diagnostics Workshop,* College Park, MD, NOAA, 416–419.

Wang, G., R. Kleeman, N. R. Smith, and F. Tseitkin, 2002: The BMRC coupled general circulation model ENSO forecast system. *Mon. Wea. Rev.,* **130,** 975–991.

Wright, P. B., 1985: The Southern Oscillation: An ocean–atmosphere feedback system? *Bull. Amer. Meteor. Soc.,* **66,** 398–412.

Xu, J. S., and H. von Storch, 1990: Principal oscillation patterns—Prediction of the state of ENSO. *J. Climate,* **3,** 1316–1329.

Xue, Y., M. A. Cane, S. E. Zebiak, and M. B. Blumenthal, 1994: On the prediction of ENSO: A study with a low-order Markov model. *Tellus,* **46A,** 512–528.

Zebiak, S. E., and M. A. Cane, 1987: A model El Niño–Southern Oscillation. *Mon. Wea. Rev.,* **115,** 2262–2278.