Assessing Potential Seasonal Predictability with an Ensemble of Multidecadal GCM Simulations

DAVID P. ROWELL

Hadley Centre for Climate Prediction and Research, Meteorological Office, Bracknell, United Kingdom

(Manuscript received 19 September 1996, in final form 28 May 1997)

ABSTRACT

A global search for areas where seasonal prediction may be feasible has attracted scientific interest for many years. This contribution is based primarily on data from a six-member ensemble of 45-yr climate runs, each of which is forced by observed sea surface temperatures (SSTs) and sea-ice extents and are unique only in their initial atmospheric conditions. The potentially predictable component of atmospheric interannual variability is assumed to be that due to oceanic forcing, and using ''analysis of variance,'' this is separated from the unpredictable internal component; potential predictability is measured as the ratio of ocean-forced variance to total variance. Significance levels and confidence intervals are calculated, both of which are essential for a meaningful interpretation of predictability estimates; the latter show there is low susceptibility to sampling problems here because of the large data source available.

Global maps of potential predictability, for simulated seasonal mean precipitation and mean sea level pressure (MSLP), are shown for both the solstitial and equinoctial seasons. In most regions, this model-based predictability estimate has large variations through the annual cycle. Not surprisingly, the highest predictability occurs over the tropical oceans, particularly the Atlantic and Pacific, for which a better knowledge of the influence of SSTs on diabatic heating is important for understanding the variability of teleconnected regions. Land areas displaying high predictability tend to support existing empirical studies, although over Australia and parts of Africa the model's response to SSTs seems erroneously weak. In the midlatitude Northern Hemisphere, a winter–spring peak of predictability is confirmed, but a notable autumnal minima of predictability is also proposed. At polar latitudes, there is a small but significant influence of SSTs on spring MSLP, and in some localities a moderate influence on precipitation. Further work is required with observational data to properly assess these findings.

1. Introduction

It is commonly accepted that the atmosphere behaves as a chaotic system (e.g., Lorenz 1963, 1993; Palmer 1993), and as a consequence, deterministic predictability is limited. However, it is also well known that at many locations, predictability of seasonal weather statistics is also possible (see reviews by Palmer and Anderson 1994; Hastenrath 1995). This arises from what may be termed ''external'' factors that alter the likelihood of residence in atmospheric attractors (cf. Palmer 1993), enabling probabilistic forecasts to be made of the seasonal mean state, on condition that the external forcing is itself predictable. The primary source of such external forcing at seasonal timescales arises from anomalous sea surface temperature (SST) patterns; these can indeed be predicted, either using coupled dynamical models (e.g., Cane et al. 1986; Barnston et al. 1994; Ji et al. 1994) or statistical models (e.g., Ward et al. 1993; Barnston et al. 1994). Further potential sources of predictive skill, for example, land surface anomalies, are believed to be much less important and are neglected through much of this paper.

It is clearly important to be able to assess where on the globe atmospheric variations are sufficiently affected by oceanic forcing to enable practical seasonal prediction. This requires measurements of atmospheric *potential predictability* (''potential'' indicates that this also depends on predictions of anomalous oceanic forcing), whose definition and mapping have been topics of ongoing research in the climate community.

To date, three basic approaches have been taken. The first involves estimating (at each point) the random component of monthly or seasonal variance from a spectrum of daily data and comparing this with the actual interannual variance of monthly or seasonal means (e.g., Madden 1976; Shukla and Gutzler 1983; Trenberth 1985; Zwiers 1987; hereafter, MSGT or MSGTZ). The ratio of these variances defines a measure of potential predictability that can be plotted globally. This method has the advantage that it can be based purely on observational data, since it requires only a single realization of atmospheric evolution (although Zwiers uses a single model run). Also, it includes secondary sources of predictability, not just that due to the ocean surface.

Corresponding author address: Dr. David P. Rowell, Hadley Centre for Climate Prediction and Research, U.K. Meteorological Office, London Road, Bracknell, Berkshire RG12 2SY, United Kingdom.
E-mail: dprowell@meto.gov.uk

On the other hand, its assumptions (reviewed by Zwiers) include a supposition that the predictable signal is constant through each season; this will often be untrue. Furthermore, Madden also found that heterogeneous observational methods led to some erroneous predictability estimates; this may also have affected the analyses of Shukla and Gutzler (1983) and Trenberth (1985).

The second approach, adopted by Lau (1985) and Chervin (1988), compares the interannual variance of seasonal means in two runs of a general circulation model (GCM). The first run is forced by climatological SSTs in every year and used to estimate the variance due to random effects. The second run includes realistic interannual variations of SST and provides an estimate of total interannual variance. Again, the ratio of these quantities measures potential predictability and can be plotted as global maps. This approach requires none of the more doubtful assumptions of MSGTZ, but it does rely entirely on the skill of a GCM. In particular, the models used by Lau and Chervin were less sophisticated than many available today. Furthermore, the method assumes that the interannual variance of seasonal-mean random anomalies is identical over a smooth climatological SST field to that over a real SST field that has systematically sharper gradients; this assumption has yet to be assessed, though it may not be greatly in error.

More recently, potential predictability has been measured using an ensemble of climate simulations, where all are forced by the same observed interannually varying SSTs but started from different initial atmospheric conditions (Dix and Hunt 1995; Harzallah and Sadourny 1995; Kumar and Hoerling 1995; Rowell et al. 1995; Stern and Miyakoda 1995; Zwiers 1996). Since these simulations provide temporally and spatially complete climate data, they are also valuable in other areas of climate research. For predictability studies, the philosophy is that sensitivity to initial atmospheric conditions can be used to quantify the random component of interannual variability, whereas the relative similarity (or lack of it) between ensemble members can be used to quantify the potentially predictable component of variance. The standard statistical tool for this kind of problem is ''analysis of variance'' (ANOVA), detailed in the regional study of Rowell et al. (1995), used slightly incorrectly by Harzallah and Sadourny (1995) and Kumar and Hoerling (1995) (cf. Rowell et al. 1995), and used in an appropriately different form by MSGTZ and Zwiers (1996). Alternative, less formal techniques were used by Dix and Hunt (1995) and Stern and Miyakoda (1995). A particular advantage of the ensemble approach is that it is more powerful at detecting weak influences of SST (see the appendix, this paper), but, like the approach of Lau (1985) and Chervin (1988), it has the disadvantage of relying primarily on a model's climate skill. Arguably, this may be preferable to the observational approach of MSGT (for which some of the assumptions are not well satisfied; see above) but nevertheless necessitates that findings should later, and as far as possible, be assessed against observational data.

In this paper, the ensemble approach and ANOVA are used together to provide a global assessment of potential seasonal predictability. Three principal advances are made over previous ensemble studies: a rigorous statistical analysis, the use of much longer (45-yr) simulations to reduce sampling error, and a region-by-region analysis of predictability in each season of the year.

In section 2, the statistical methodology and assumptions are presented for climate applications, including significance testing, the estimation of confidence intervals, and a technique for assessing the impact of random atmospheric variations on model interannual skill. Section 3 describes the model and the experiments. In section 4, the maps of potential seasonal predictability are described, both globally and regionally, and compared with regional observational studies where available. The value of these plots for climate variability studies is also emphasized. Finally, section 5 provides a summary and concluding remarks.

2. Statistical approach

As already noted, the prime statistical tool employed in this paper is ANOVA. It is used with data from an ensemble of climate simulations to separate the total atmospheric variance (σ^2_{TOT}) of some time-averaged quantity into two components, one due to oceanic (SST and sea ice) forcing (σ^2_{SST}), and the other due to random internal variability (σ^2_{INT}). Potential predictability is then measured as the ratio of ocean-forced variance to total variance ($\sigma^2_{\mathrm{SST}}/\sigma^2_{\mathrm{TOT}}$), having an intuitive scale of 0%–100%. The main advantage of ANOVA is that having been used extensively in other scientific applications, it is well understood and also amenable to extensions beyond its basic application. Good references with further statistical discussion are Scheffe (1959) and Searle et al. (1992).

First, we describe the conceptual ''model'' on which the technique is based. Imagine that data are available from an ensemble of n climate simulations, each of N years in length, each forced by the same varying SSTs and sea ice, and each one differing from the others only by its initial atmospheric conditions. Consider a characteristic, X, which can be quantified for each individual season, for example, the time mean of some variable at a point, a higher moment statistic, a regional index, or an EOF time coefficient. For each simulated season, this seasonal characteristic can be modeled as the sum of two independent components:

$$x_{ij} = \mu_i + \varepsilon_{ij}, \tag{1}$$

where x_{ij} is the simulated value of X; $i = 1, \ldots, N$ is the year; $j = 1, \ldots, n$ is the ensemble member; μ_i is the component of x_{ij} due to oceanic forcing (the variance of the μ_i's is σ^2_{SST}); and ε_{ij} is the anomalous component of x_{ij} due to internal variability (the variance of the ε_{ij}'s is σ^2_{INT}). (Note that in this paper, ε_{ij} includes the role of land-

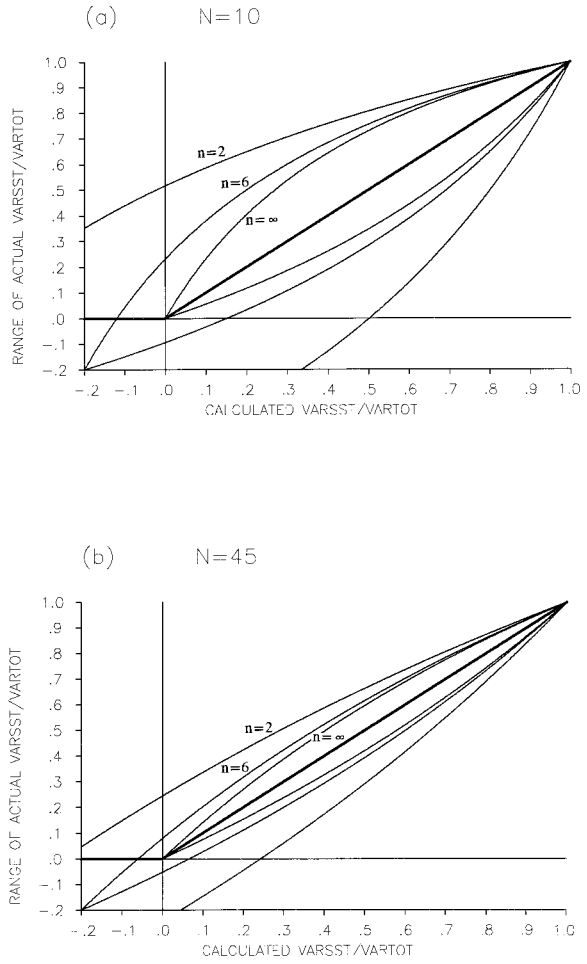

FIG. 1. Confidence limits for ρ, for (a) 10-yr experiments and (b) 45-yr experiments. Thick line represents the best estimate of ρ. Thin lines, computed for various-sized ensembles, define the limits between which the true value of ρ lies with 90% confidence.

surface forcing; see section 3a.) This is often described as a ''one-way random-effects model.'' Note that this model can be viewed as a simplification of that described by Zwiers (1996); his additional terms, neglected here, represent a ''configuration effect,'' (arising from his use of a different computer for one of the simulations), an ''initial conditions effect,'' (necessary if the choice of initial conditions could cause different mean climate states, i.e., multiple equilibria), and an ''internal sources effect'' (necessary if supraannual internal modes are possible). The first of these is unnecessary here, whereas a notable influence from the two latter terms seems unlikely (see next paragraph and Lorenz 1990) and, as a subsequent paper will show, they are in any case statistically insignificant in the model atmosphere.

Two assumptions are made in the application of the conceptual model described above (see also Scheffe 1959, 222–223). The first is that the ε_{ij}'s are all randomly and independently selected from the same population. The aspect of randomness is satisfied so long as regime-residence times are much less than a year. Periods of atmospheric intransitivity (persistent unforced anomalies) lasting a year or more seem unlikely, not the least because of the large changes of forcing associated with the passage of the annual cycle (Lorenz 1990). The phrase, ''independently selected from the same population,'' is used because the variance within each year of the ε_{ij}'s (σ_i^2) must be identical in all years. It can be shown that heterogeneity of σ_i^2 is not unduly large in the data used in this paper (Rowell 1996a), and furthermore, Scheffe (1959, chap. 10) notes that for a constant ensemble size these heterogeneities have little impact on ANOVA. The second assumption is that the μ_i's are independently and identically distributed random variables from some defined population. This is valid, except where μ_i has substantial serial correlation, in which case potential predictability is slightly underestimated (by up to a quarter of its 90% confidence range, for the ensemble size used in section 4). A further qualification for this second assumption is that it is important to identify the ''population.'' In this paper it is defined as late twentieth-century climate. It is interesting to note that this may, however, differ from twenty-first century climate because of anthropogenic influences; this is a pertinent issue in seasonal forecasting, since the current statistics of oceanic forcing and atmospheric response may or may not remain intact.

Based on this conceptual model, Rowell et al. (1995) show how the simulation data (the x_{ij}'s) can be used to obtain unbiased estimates of σ^2_{INT} and σ^2_{SST}, and that the sum of these variance components provides an unbiased estimate of the total variance, σ^2_{TOT}. The variance ratio used to measure potential predictability is then

$$\hat{\rho} = \frac{\hat{\sigma}^2_{\text{SST}}}{\hat{\sigma}^2_{\text{TOT}}}, \tag{2}$$

where $\hat{}$ denotes the best estimator of a population quantity.[1] This is the predictability measure plotted in section 4. It is important to realize that ρ should always be viewed as some kind of ''average'' impact of oceanic forcing over *all* years.

It is also important to assess the statistical significance of potential predictability calculations, and this is done by testing the null hypothesis that $\sigma^2_{\text{SST}} = 0$ (or $\rho = 0$). A rejection criterion is derived in the appendix [Eq. (A8)] and utilized in section 4.

Also of importance when assessing predictability is an appreciation of the confidence intervals on ρ. These are displayed in Fig. 1 for various-sized ensembles, for

[1] Note that $\hat{\sigma}^2_{\text{SST}}/\hat{\sigma}^2_{\text{TOT}}$ is a slightly biased estimate of ρ, since in general $E(a/b) \neq E(a)/E(b)$, where $E()$ is the expectation of a population quantity. Donoghue and Collins (1990) show how to correct this bias, and for the data to be analyzed here ($N = 45$, $n = 6$), it can be shown to be considerably less than the sampling error. To simplify the mathematics of the appendices, we equate $\hat{\rho}$ to $\hat{\sigma}^2_{\text{SST}}/\hat{\sigma}^2_{\text{TOT}}$ and also note that this has only a minor impact on the plotted estimates of ρ in section 4.

10-yr experiments such as AMIP (Atmospheric Model Intercomparison Project; Gates 1992), and for 45-yr experiments, such as those used in section 4 of this paper. Their derivation is described in the appendix, along with the assumptions involved, and an interpretation of the negative values plotted. It is not surprising that as the ensemble size increases, the accuracy of estimates of ρ are enhanced. However, even in the limiting case of an infinite ensemble, a finite interval remains, which becomes more pronounced with fewer years. This is because the accuracy with which σ^2_{SST} can be estimated is limited by the length of the integrations. Thus, data from multidecadal simulations, such as those available here, are generally much less susceptible to sampling problems than data from shorter integrations.

Finally, we consider the relationship between potential predictability and GCM temporal correlation skill (assessed against an observed time series), which is needed for regional predictability and verification studies (e.g., Davies et al. 1997). Even if a model behaves perfectly, it can be envisaged that only the SST-forced component of its time series will be skillfully simulated, and the presence of random variations will inevitably lower its overall skill. It can easily be shown that for a perfect model, the correlation skill for a given index equals its potential predictability, ρ (see the appendix). For a nonperfect model, this correlation skill is, in general, *limited* by ρ. An extension of this idea can be applied to an ensemble of realizations. If each run is again correlated with an observed time series, then the average of these correlations is again limited by ρ. [Note that in a similar vein, Dix and Hunt (1995) use an average of correlations between all possible pairs of runs as an estimate of ''areas where a systematic relationship [with SSTs] can be expected''; since the average of these correlations must estimate ρ, then their diagnostic is equivalent to the ANOVA approach described here.] Another approach for ensembles is to correlate the observed time series with the ensemble mean data rather than with individual runs. In this case, the model's random component of variability is reduced, and a higher limit is imposed on correlation skill [see the appendix, Eq. (A22)]. Whatever approach is used, the application of these relationships should help distinguish between skill variations arising from regional model failures and those arising simply from geographic variations of the relative influence of SSTs.

3. Model and experimental design

a. Model formulation

The GCM used in this paper is known as HADAM1, the first version of the Hadley Centre Atmospheric Model, and is identical to that submitted to AMIP in 1993 (Gates 1992). It is a gridpoint model with a horizontal resolution of 2.5° lat $\times$ 3.75° long and 19 hybrid levels in the vertical. Physical parameterizations include a gravity wave drag scheme; a radiation scheme that computes fluxes in four longwave bands and six shortwave bands, and responds to prognostic cloud variables (large-scale cloud amount, cloud liquid-water and ice content, and convective cloud amount); a penetrative convection scheme with stability-dependent closure, which represents both updrafts and downdrafts; boundary layer mixing in up to five of the lowest model layers; and a land–surface scheme that includes a vegetation canopy model, a four-layer soil model for heat conduction, and a single-layer soil model for moisture storage including surface and subsurface runoff. The GCM's chemistry involves seasonally and meridionally varying ozone profiles and a fixed carbon dioxide concentration (321 ppmv). The background behind the model development is given by Cullen (1993), and a detailed summary of the physics and dynamics of the precise version used here is provided by Phillips (1994) and references therein.

It is important to briefly appraise the suitability of HADAM1 for assessing atmospheric seasonal predictability. Initially (but see below), the practical approach taken here for global studies is to simply record that HADAM1 can be regarded as ''state of the art,'' having been ranked highly among those models submitted to AMIP in 1993–94 (M. Fiorino 1995, personal communication), and then proceed to use ANOVA to compute its seasonal predictability estimates. [Note also that HADAM1's first- and second-moment seasonal statistics have been further discussed by Rowell (1996a).] In reality, however, a model's suitability depends on many aspects of its behavior, and since these should be verified before one is convinced by its regional results, it is vital in section 4 to at least retain the important caveat that unsupported findings must later be assessed by detailed comparisons with available observational data and experiments with other models.

Finally, a feature of the model, of particular relevance to the interpretation of the ANOVA results presented in section 4, is that its heat and moisture contents over the land surface interact freely with the model atmosphere, and therefore the interannual variability of the land surface (and its effect on the atmosphere) is incorporated within the internal component of variance. In future studies, it might also be desired to consider the influence of the land surface as *an agent that contributes to the forcing* of atmospheric variability, rather than merely as *an internal component of a (partially) forced system*. In the former case, the model should be set up with prescribed heat and moisture fluxes over the global land surface, and ANOVA techniques could then be applied following section 2 to separate land-surface forced variability and internal atmospheric variability.

b. Experimental design

The philosophy behind an ensemble approach to GCM experimental design was discussed in section 1. The particular suite of experiments utilized in this paper consists of an ensemble of six 45-yr integrations, forced throughout by observed SST and sea-ice data. In the notation of sec-

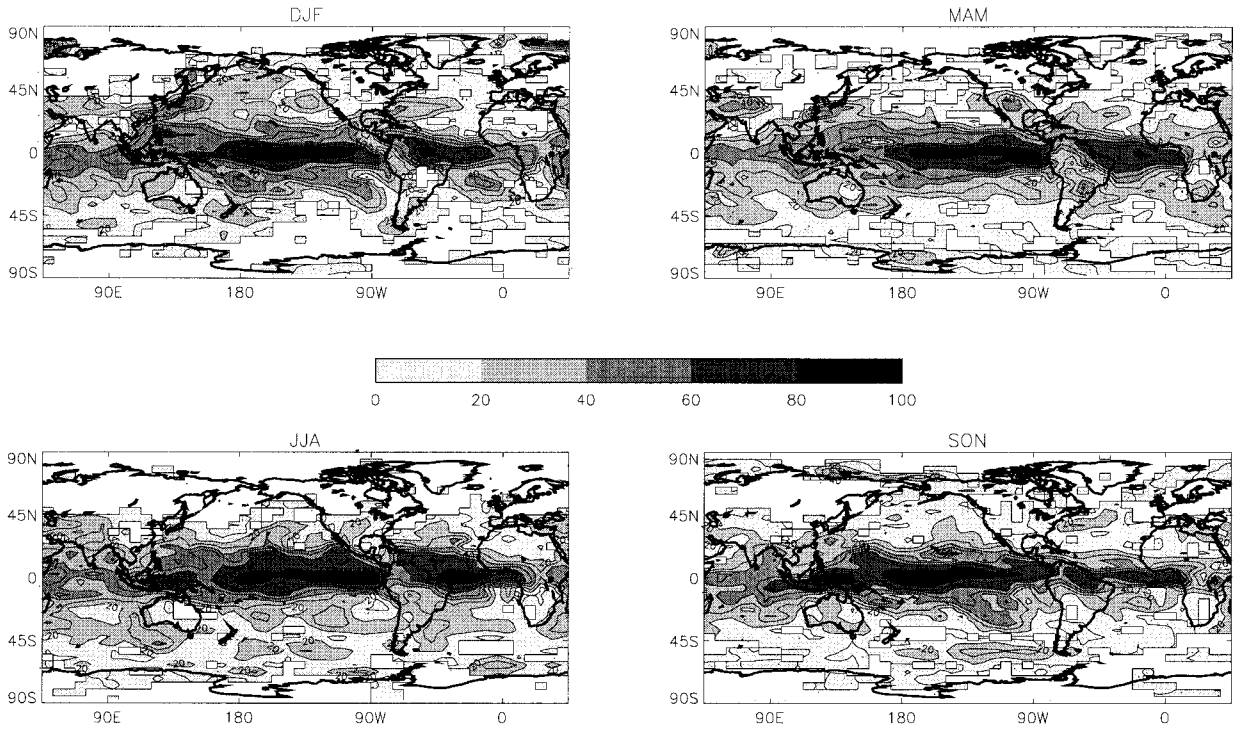

FIG. 2. Percentage variance of seasonal mean precipitation due to oceanic forcing, computed from the ensemble of six 1949–93 runs. Data were first interpolated to a 5° × 7.5° grid. Results for each of four standard seasons are shown. The contour interval is 10%, and white areas show where values do not significantly exceed zero at the 5% significance level (tested using log precipitation; see the appendix).

tion 2, we thus have $N = 45$ and $n = 6$. Each integration runs from 0000 UTC 1 October 1948 to 0000 UTC 1 December 1993, with the first 2 months of each run being discarded because of possible spinup effects as the model atmosphere evolves from its initial state to one in quasi-equilibrium with the imposed SSTs. The initial atmospheric conditions were all chosen arbitrarily from recent U.K. Meteorological Office operational analyses, but for a date close to 1 October: 1 October 1991, 1 October 1992, 2 October 1992, 3 October 1992, 1 October 1993, and 2 October 1993. For all runs, soil moisture and snow depth were initialized from an adaptation of Willmott et al.'s (1985) climatologies, soil surface temperature was initialized from the same operational analyses as the atmosphere, and deep soil temperatures were initialized using equations taken from Warrilow et al. (1986).

The observed ocean surface data used to force the GCM is the first version (1.1) of the specifically designed GISST (global sea ice and sea surface temperature) dataset, described by Parker et al. (1995). The raw GISST data, available as monthly means on a 1° grid, were interpolated to the model grid as follows: model grid boxes in the open ocean and those with up to 50% ice cover were set as sea points with a temperature determined by the area-weighted average of nonice 1° squares; model grid boxes with greater than 50% ice cover were specified as ice points. The data were then enhanced temporally, by linear interpolation to 5-day means, so as to obtain a smoother evolution of boundary forcing.

4. Global potential predictability assessment

As an application of the techniques described in section 2, maps of seasonal potential predictability, ρ, are presented in Figs. 2 and 3 (hereafter, the terms ''predictability'' and ''potential predictability'' are used interchangeably). These have been computed by applying, at each grid point, the analysis of variance to the six 45-yr runs and taking seasonal-mean precipitation or seasonal-mean MSLP (mean sea level pressure) as the ''seasonal characteristic'' whose variability is being assessed (i.e., X in section 2). Recall that they represent potential predictability under the assumption that this derives entirely from oceanic forcing.

Not surprisingly, the overall impression of these plots exhibits a strong impact of SSTs in the Tropics (cf. Charney and Shukla 1981; Palmer and Anderson 1994) and conversely much greater chaotic variability in the extratropics. It is striking that oceanic forcing has a statistically significant, and therefore predictable, influence on interannual variability over 80%–90% of the globe (note this is partly a function of the number of degrees of freedom), and that the only 15° latitude bands where less than 5% of points achieve significance are 60°–75°N for June–August (JJA) precipitation and 75°–90°N/S for December–February (DJF) MSLP. It is also noteworthy that many regions have a substantial annual cycle of predictability, which emphasizes the need to seasonally stratify any analysis of this nature.

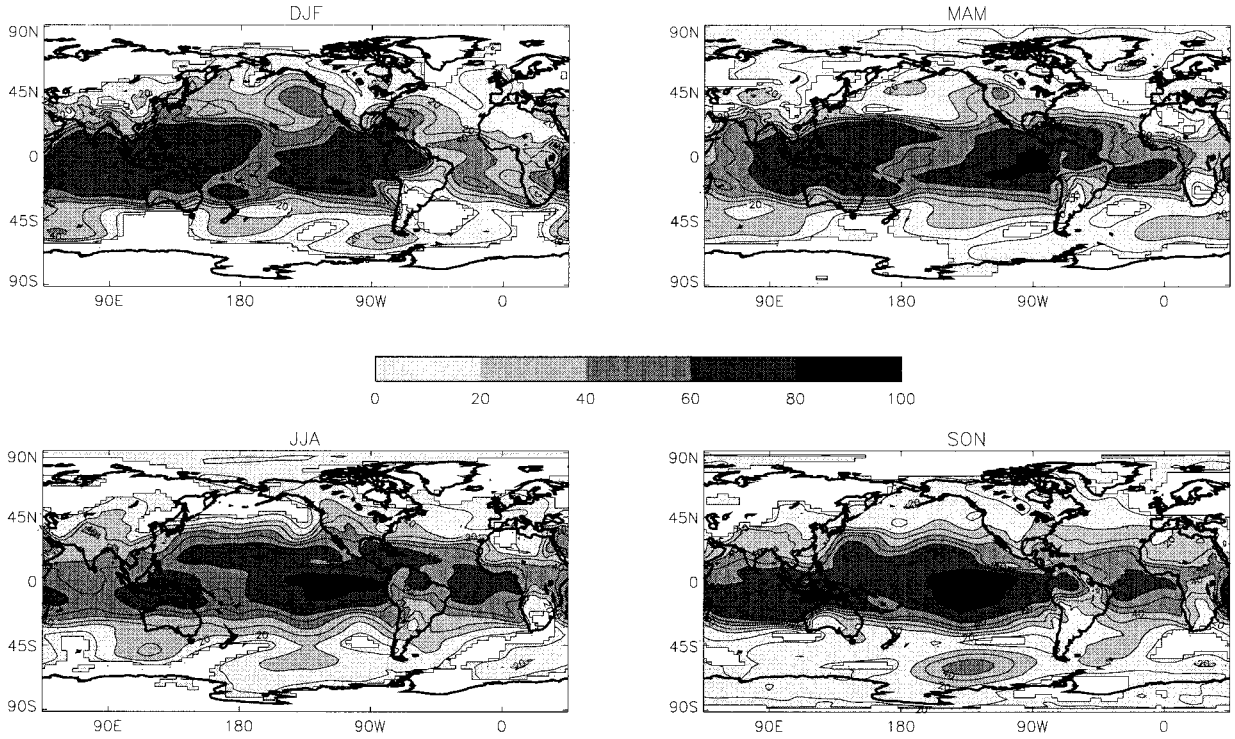

FIG. 3. Percentage variance of seasonal mean MSLP due to oceanic forcing computed from the ensemble of six 1949–93 runs. Results for each of four standard seasons are shown. The contour interval is 10%, and white areas show where values do not significantly exceed zero at the 5% significance level.

Regional patterns of the model's predictability are discussed in the following sections, taking each latitude sector in turn.

a. Tropics

First, consider the predictability of precipitation over tropical oceans, which is important because of the impact that deep convective heating has on predictability in teleconnected regions. Note, however, that since data here is either rather sparse or record length is rather short, these results lack observational support and so should be treated with prudence. Over the equatorial Pacific, the annual cycle and pattern of ρ in Fig. 2 seems partly linked to the typical evolution of SST anomalies during ENSO events (cf. Rasmusson and Carpenter 1982), perhaps because larger anomalies can have a greater impact on local convection relative to random internal variability. Over the tropical Atlantic, HADAM1's predictability is also very high [despite a lower variance of local SSTs; e.g., Bottomley et al. (1990)] and exhibits a marked annual cycle with a minimum in September–November (SON). This seasonal variation may be due to seasonalities in the influence of nonlocal forcings, such as those from the Pacific (cf. Curtis and Hastenrath 1995), but further verification and investigation is clearly required. Over the Indian Ocean, the predictability suggested by the model is somewhat lower, and again further analysis is necessary, and in this case is under way.

It is over land that predictability of seasonal precipitation is of greatest societal importance. First, over northern South America, the model's suggestion of large areas of high predictability is partly supported by regional observational studies (e.g., Aceituno 1988; Hastenrath 1990; Ward and Folland 1991), whereas other parts, which are at present unsupported, indicate that further study is necessary and could also be worthwhile. Farther north, over the Caribbean and Central America, the model's moderate predictability also finds some support in observational studies (e.g., Hastenrath 1976, 1984; Ropelweski and Halpert 1987, 1989; Kiladis and Diaz 1989) but again indicates that further work is necessary to properly verify its geographic and seasonal variations. Over the Indonesian archipelago, the model reflects the work of Nicholls (1981) and Ropelweski and Halpert (1987, 1989), for example, which also indicates high predictability, probably peaking in the early wet season, SON. Over Africa and Australia, however, values of ρ are more disappointing, given the expectations raised by other studies (e.g., Nicholson and Entekhabi 1987; Farmer 1988; Drosdowsky 1993; Rowell et al. 1995; and many others). Note that interannual variability of the North African summer monsoon is already known to be a difficult modeling problem (see discussion by Rowell 1996b; Sud and Lau 1996). Finally, the lower

predictability for the wet season of India and Southeast Asia is consistent with the work and ideas of Goswami (1994), Palmer (1994), and Brankovic and Palmer (1997).

For MSLP, the spatial and seasonal patterns of ρ (Fig. 3) have many interesting differences from those for precipitation. Most striking is the larger latitudinal range of high MSLP predictability, presumably reflecting the larger spatial coherence of MSLP and the impact of *random* isolated showers on rainfall totals in the arid subtropics. Reasons for other differences are less obvious and in any case require support from observational data.

b. Midlatitudes

In the extratropics, we focus attention mainly on MSLP (Fig. 3), which seems more responsive to SSTs than precipitation and is also more valuable for a large-scale understanding of extratropical climate variability. In the Northern Hemisphere, most locations have their highest model predictability during winter or spring; from 40° to 60°N, the fractional areas where $\hat{\rho}$ exceeds 20% are 25%, 17%, 13%, and 4% for DJF, March–May (MAM), JJA, and SON, respectively. A partial explanation for this suggested seasonality lies in the work of Opsteegh and Van den Dool (1980) and Webster (1982) (the latter compares only the solstitial seasons), whose simple linear models show that only when the westerly flow is far enough south can the Rossby waves communicate predictable signals from the Tropics to the extratropics (cf. also Brankovic et al. 1994). The resulting winter–spring peak of midlatitude predictability is now supported by this large sample of GCM data. However, the autumnal minima of MSLP predictability suggested by Fig. 3 and by Brankovic et al.'s small sample of GCM runs is contrary to Opsteegh and Van den Dool's study. This could be partly due to lower predictability of modeled diabatic heating over the tropical Atlantic in SON (suggested by Fig. 2 and section 4a) but probably also involves the tropical Pacific, where simulated convective heating is as predictable in SON as it is in other seasons. Therefore, communication with the extratropics may also be weaker in SON, which suggests that some important feature may be present in HADAM1 but is perhaps missing from Opsteegh and Van den Dool's model. However, further work is clearly required, including an assessment of each of the above ideas against available observational data.

This seasonality is particularly apparent over the United States, where the existence of often moderate predictability is well known (e.g., Kiladis and Diaz 1989; Barnston 1994). Over Europe, Fig. 3 shows that simulated predictability is much lower than over North America, perhaps in part due to a weaker influence of ENSO (see Fraedrich 1994), and this would explain the less rapid development of operational seasonal forecasting here. For further analysis of this sector with HADAM1 data, see Davies et al. (1997). Over Asia, north of 30°N, HADAM1's seasonal predictability is also low. Exceptions are precipitation over southwest Asia during spring and summer, and over Japan in winter (Fig. 2), though at least the latter is probably erroneous since it fails to match the experience of long-range forecasters (M. Sugi 1995, personal communication).

In Southern Hemisphere midlatitudes, predictability appears moderate in all seasons. In spring (SON), a strong model maximum around 60°S, 135°W relates well to observed decadal MSLP anomalies arising from variations of the semiannual oscillation (Hurrell and Van Loon 1994), and in winter (JJA), high predictability south of western Australia seems to match the observed ENSO response described by Karoly (1989). In summer (DJF), however, the model maxima around 55°S, 55°E and 60°S, 80°W correspond better to the wider predictability results of Trenberth (1985) (using the first of the three methods described in section 1) than to Karoly's composite, suggesting they may have causes other than ENSO.

c. High latitudes

Poleward of about 60°N/S, a marked annual cycle of simulated predictability is apparent in both hemispheres, whereby oceanic forcing affects interannual variability of MSLP only in spring, that is, MAM in the Northern Hemisphere and SON in the Southern Hemisphere. However, this is so far unsupported in observational studies, since few, if any, of these have included the equinoctial seasons. Nevertheless, in the model world at least, we can speculate that the cause of this seasonality may be enhanced teleconnections from the Tropics in spring, when potential vorticity gradients should be stronger than in summer, but internal variability is probably weaker than in winter (cf. Brankovic et al. 1994 and section 4b). This influence from the Tropics is supported by performing ANOVA on only the 20 most extreme Southern Oscillation index (SOI) years, in which case the high-latitude aspects of Fig. 3 are enhanced (not shown), whereas if only weak and neutral SOI years are used, no spring predictability is found at high latitudes.

Although these experiments do not allow a separate examination of the influence of sea-ice anomalies, it seems that their effects may sometimes be apparent in the smaller-scale variables, such as surface air temperature (not shown) or precipitation (Fig. 2). This matches the expectation of observational studies such as Rogers (1978), Carleton (1981), and Overland and Pease (1982). The most striking example in the model is over the Barents Sea in winter, where snowfall appears to respond strongly to the (ice/water) state of the sea surface. This is also apparent in the all-season results of the Commonwealth Scientific and Industrial Research Organisation (CSIRO) model (Dix and Hunt 1995, their Fig. 5a). The effects of ice-edge variability on large-scale fields such as MSLP are much harder to infer, although the preceding paragraph suggests that in HADAM1 they are negligible in at least most years and seasons. This suggestion that ice extents sometimes influ-

ence local weather but rarely influence the large-scale flow is supported by the synthesis and review of Walsh (1993).

5. Conclusions and discussion

The potential skill of seasonal forecasts in a particular region depends primarily on the balance between oceanic forcing and the random effects of chaotic weather events. These define predictable and unpredictable components, respectively, and it has been shown how ANOVA can be used to separate these. None of the statistical assumptions involved are seriously violated. Significance levels and confidence intervals have been computed, both of which are vital for a meaningful interpretation of geographical or seasonal variations of predictability. Finally, the relationship between the fraction of forced variability and the skill of model-observed comparisons has been derived.

One of the underlying assumptions of the approach is that atmospheric predictability is derived entirely from oceanic anomalies. Although this should lead to a reasonable first-order estimate of predictability, further analysis and experimentation is required to assess the importance of sources neglected here. Three of these can be highlighted: (a) preseason land-surface anomalies—these may sometimes play a role in maintaining summer droughts, for example (e.g., Ratcliffe 1976; Trenberth and Branstator 1992); (b) anthropogenically and naturally induced changes to atmospheric constituents—ideally these should be correctly specified in predictive GCMs, since they have probably already had a small effect on climate (e.g., Santer et al. 1996; Hegerl et al. 1996) and hence on seasonal anomalies; and (c) the preseason atmospheric state—the issue here is whether internal atmospheric anomalies can persist for a season or more without external forcing (e.g., Shukla 1985; James and James 1989), or whether the transitions between internal modes occur too frequently to allow such behavior.

ANOVA was applied here to a six-member 45-yr ensemble of simulations, with a state-of-the-art climate model, to provide worldwide estimates of potential seasonal predictability in both the solstitial and equinoctial seasons. The sampling error is low (owing to the large data quantity), thus yielding more accurate estimates of potential predictability within the bounds of model reliability (see below). Some of the main conclusions are as follows.

- Seasonal and spatial variations of the relative influence of SSTs on simulated deep convective heating over the tropical oceans have been examined. This is important for understanding the predictability and climate variability of teleconnected regions.
- A winter/spring peak of Northern Hemisphere midlatitude predictability has been further confirmed, and an autumnal minima of predictability has been suggested. The latter may be due to weaker teleconnections from the tropical Pacific and reduced diabatic forcing from the tropical Atlantic; further experimentation and analysis of observational data is needed.
- At high latitudes, spring MSLP exhibited a small but discernable impact from tropical SSTs, whereas in other seasons, interannual variability of MSLP was purely random. Also, in some polar regions, precipitation experienced a moderate influence from the local sea state close to the ice edge.

To a large extent, these conclusions and the mechanistic speculations in section 4 rely on the integrity of HADAM1's climate skill. This includes its local response to SST anomalies, its regional and planetary-scale teleconnections, its regional responses to these teleconnections, and the behavior of its internal modes. For some regions, support for model predictability estimates has been gained from comparisons with available observational studies, whereas for some areas model problems have been revealed, and for other regions (including some of those noted above) the results are as yet unsupported. Further evidence must come first from more comparisons of model behavior with observational data. Davies et al. (1997) is one such study for HADAM1; others are in progress. Second, the agreement (or lack of it) with predictability estimates from other state-of-the-art models will also help determine confidence in the results presented here (cf. initial comparisons by Zwiers et al. 1995).

A more complete study of seasonal predictability should then incorporate further additional components. These would include analysis of other parameters, such as temperature, height, etc., and other statistics, for example, intraseasonal variance, EOF time coefficients, blocking frequency, etc. An analysis of year-to-year variations of predictability is also required, which might involve their detection using large multidecadal ensembles (cf. Rowell 1996a) and then comparisons between individual years, either using nonparametric statistics (cf. Anderson and Stern 1996) or, where appropriate, the usually more powerful parametric statistics. Furthermore, it is also important to assess the sensitivity of predictability to the user's requirements for temporal and spatial averaging (or lack of it); this depends on differences between the spectral and covariance structures of the ''signal'' versus the ''noise'' (e.g., Hasselmann 1979). Finally, one of the most challenging avenues for future research is to understand the mechanisms by which ocean surface anomalies affect seasonal climate variability and predictability. Simulation data, such as the data utilized here, along with the output of idealized experiments and available observational data, provide excellent material to pursue these issues.

Acknowledgments. The advice of John Rowell has been much appreciated, and also discussions with Chris Folland, Gil Ross, Ian Jolliffe, Masato Sugi, and Francis Zwiers. Thanks are due to David Sexton, Alison Ren-

shaw, and John Davies for running the model and developing some of the software. Computer time was provided by the U.K. Department of Environment, under Contract PECD7/12/37, with further partial support under CEC Contract EV5V-CT92-0121 (''Medium Term Climate Variability'').

APPENDIX

Derivation of ANOVA-Related Statistics

a. Significance testing

Here we show how to test the null hypothesis that oceanic forcing has no impact on simulated interannual climate variability (at a particular point or for a particular index):

$$H_0\colon \sigma^2_{\mathrm{SST}} = 0, \quad \text{or} \quad \rho = 0 \tag{A1}$$

(also equivalent to a hypothesis that the population means for all years are identical, i.e., $\mu_i = \mu$, for all i). Notation is defined in section 2.

An additional assumption now needed is that the anomalies due to internal variability (ε_{ij}) are derived from a Gaussian distribution (note that elsewhere, except in section b below, no distributional assumptions are needed). The validity of this assumption has been tested on the data used in section 4, with a Kolmogorov–Smirnov test (e.g., Hoel 1971; Press et al. 1992) and a χ^2 test. For seasonal MSLP, the entire data were found to be near-Gaussian. For precipitation, however, arid regions are non-Gaussian, and so in section 4, $\sigma^2_{\mathrm{SST}} = 0$ was tested using log precipitation, for which much or all of the data is near-Gaussian.

To determine a criterion at which the null hypothesis in Eq. (A1) can be rejected (at a given level of significance), two equations involving the total variance, valid under the null hypothesis, are used. First,

$$\sigma^2_{\mathrm{TOT}} = \sigma^2_{\mathrm{INT}} \qquad (\text{under } H_0,\ \sigma^2_{\mathrm{SST}} = 0), \tag{A2}$$

and second (using the equation for the variance of sample means),

$$\sigma^2_{\mathrm{EM}} = \frac{\sigma^2_{\mathrm{TOT}}}{n}. \tag{A3}$$

Combining Eqs. (A2) and (A3) gives

$$n\,\sigma^2_{\mathrm{EM}} = \sigma^2_{\mathrm{INT}}, \qquad \text{under } H_0. \tag{A4}$$

Estimates of each of these variances can be obtained from the experimental data, and then the possibility that they derive from the same population (i.e., the null hypothesis is true) can be assessed using Fisher's F test. Under H_0 the ratio of the two estimates follows a central F distribution. Thus,

$$\Pr\left\{\frac{n\hat{\sigma}^2_{\mathrm{EM}}}{\hat{\sigma}^2_{\mathrm{INT}}} > F_{[N-1,N(n-1)],\alpha}\right\} = \alpha, \quad \text{under } H_0, \tag{A5}$$

where $F_{[\nu_1,\nu_2],\alpha}$ is the F statistic at the α significance level, with degrees of freedom ν_1 and ν_2. (Note that under the null hypothesis, serial correlation is estimated as zero and so has no effect on the test.) It is the use of this F statistic that requires that the ε_{ij}'s be derived from a Gaussian distribution.

The criterion for rejecting H_0 at the α significance level is then

$$\frac{n\hat{\sigma}^2_{\mathrm{EM}}}{\hat{\sigma}^2_{\mathrm{INT}}} > F_{[N-1,N(n-1)],\alpha}. \tag{A6}$$

Note that this is a one-sided test since it is known that if H_0 is rejected, $n\sigma^2_{\mathrm{EM}} > \sigma^2_{\mathrm{INT}}$.

In section 4, this rejection criterion is required in terms of $\hat{\rho}$. To reformulate Eq. (A6), we use

$$\frac{n\hat{\sigma}^2_{\mathrm{EM}}}{\hat{\sigma}^2_{\mathrm{INT}}} = 1 + n(\hat{\rho}^{-1} - 1)^{-1}, \tag{A7}$$

which is easily derived from Eq. (2) (section 2) and Eqs. (A6) and (A8) of Rowell et al. (1995). So, from Eqs. (A6) and (A7), the criterion for rejecting H_0 at the α significance level is alternatively

$$\hat{\rho} > [1 + n(F_{[N-1,N(n-1)],\alpha} - 1)^{-1}]^{-1}. \tag{A8}$$

b. Confidence limits

We now derive the confidence limits on ρ, plotted in section 2, and defined as $\hat{\rho}_a$ and $\hat{\rho}_b$ such that

$$\hat{\rho}_a < \rho < \hat{\rho}_b, \quad \text{with probability } 1 - \alpha. \tag{A9}$$

Additional assumptions required are again that the ε_{ij}'s are Gaussian (see above) and also that the ensemble means ($\overline{x_i}$'s) are Gaussian. For large ensembles, the latter is valid under the central limit theorem, but for small ensembles its validity should be tested. Using a Kolmogorov–Smirnov test and a χ^2 test, the MSLP ensemble means used in this paper can be shown to be Gaussian. For precipitation, however, the former of the above assumptions remains most problematic, and so it is suggested simply that confidence limits over arid regions are viewed with caution.

Consider first the left-hand side of the confidence interval [Eq. (A9)], which can also be written

$$\Pr\{\hat{\rho}_a > \rho\} = \alpha/2. \tag{A10}$$

Now the distribution of $\hat{\rho}$ is unknown, so we choose instead to work with variables whose distribution is known and can later be related to ρ and $\hat{\rho}$. Consider

$$\frac{\hat{\sigma}^2_{\mathrm{EM}}/\sigma^2_{\mathrm{EM}}}{\hat{\sigma}^2_{\mathrm{INT}}/\sigma^2_{\mathrm{INT}}} = F_{N-1,N(n-1)}. \tag{A11}$$

To obtain a relationship in same form as Eq. (A10), we find the point on this F distribution, which is exceeded with probability $\alpha/2$:

$$\Pr\left\{\frac{\hat{\sigma}^2_{\mathrm{EM}}/\sigma^2_{\mathrm{EM}}}{\hat{\sigma}^2_{\mathrm{INT}}/\sigma^2_{\mathrm{INT}}} > F_{[N-1,N(n-1)],\alpha/2}\right\} = \alpha/2, \tag{A12}$$

and then, using Eq. (A7), some straightforward manipulation gives

$$\Pr\{\langle n\{F_{[N-1,N(n-1)],\alpha/2}[1 + n(\rho^{-1} - 1)^{-1}] - 1\}^{-1} + 1\rangle^{-1} > \rho\} = \alpha/2. \quad \text{(A13)}$$

So by comparing Eqs. (A10) and (A13),

$$\hat{\rho}_a = \langle n\{F^{-1}_{[N-1,N(n-1)],\alpha/2}[1 + n(\hat{\rho}^{-1} - 1)^{-1}] - 1\}^{-1} + 1\rangle^{-1}. \quad \text{(A14)}$$

The value of $\hat{\rho}_b$ is obtained by a similar process, except that the inequalities in Eqs. (A10), (A12), and (A13) are reversed, and hence, $F_{[\upsilon_1,\upsilon_2],\alpha/2}$ is replaced by $F^{-1}_{[\upsilon_2,\upsilon_1],\alpha/2}$:

$$\hat{\rho}_b = \langle n\{F_{[N(n-1),N-1],\alpha/2}[1 + n(\hat{\rho}^{-1} - 1)^{-1}] - 1\}^{-1} + 1\rangle^{-1}. \quad \text{(A15)}$$

Three further points should be noted.

1) When computing $\hat{\rho}_a$ and $\hat{\rho}_b$, the value of $\hat{\rho}$ should be obtained *without* resetting negative values of $\hat{\sigma}^2_{\text{SST}}$ to zero [i.e., contrary to the normal recommendation of Rowell et al. (1995)]. This is reflected in the axes of Fig. 1, although consistent with Rowell et al., when $\hat{\rho} < 0$, the "best estimate" of ρ is plotted as 0. This philosophy follows that of Scheffe (1959), who notes that "most users of confidence intervals have a more or less conscious feeling that the length of a two-sided confidence interval is a measure of the error of some point estimate of the parameter," and "if the interval is considerably shortened by deleting the part, if any, to the left of the origin [i.e., below the origin in Fig. 1], a misleading impression of accuracy of the estimation may result." Hence, he suggests that negative values of $\hat{\rho}_a$ or $\hat{\rho}_b$ are *not* modified, and so they appear uncorrected in Fig. 1. Of course when the upper end of the interval is less than zero, this indicates that the associated value of $\hat{\rho}$ is actually rather unlikely to occur in practice. Taken to the extreme, an absolute lower limit for $\hat{\rho}$ is evident from Eq. (A7); since $\hat{\sigma}^2_{\text{EM}}/\hat{\sigma}^2_{\text{INT}} \geq 0$, this limit is $\hat{\rho} \geq (1 - n)^{-1}$.

2) If the data have significant serial correlation, then the confidence interval is widened. This effect can be computed either by replacing the number of years with an "effective" number of years or by using a spectral approach to be described in a later paper focusing on decadal variability.

3) This section also hints at a test against an alternative null hypothesis: $\rho < \rho_0$. The rejection criterion at the $\alpha/2$ significance level can be seen to be $\hat{\rho}_a > \rho_0$. This may be useful if, for example, a seasonal prediction scheme requires that a particular level of potential SST-induced skill is achievable before the scheme can be considered viable.

c. Relationship with correlation skill

The interannual correlation between a perfect model and observations can be expected to be related to their potential predictability, ρ (see section 2). Two such relationships are derived below, using only the assumptions of section 2.

First, consider the availability of a single-model run. Since the observed and model data are assumed to have the same behavior, they can each be treated as realizations of the same system and labeled, for example, as $j = 1$ and $j = 2$. This "perfect model (PM) correlation skill" can then be written as

$$r_{\text{PM}} = \frac{\sum_{i=1}^{N} [x_{i1} - (x)_1][x_{i2} - (x)_2]}{\left\{\sum_{i=1}^{N} [x_{i1} - (x)_1]^2 \sum_{i=1}^{N} [x_{i2} - (x)_2]^2\right\}^{1/2}}, \quad \text{(A16)}$$

where

$$(x)_j = \frac{1}{N}\sum_{i=1}^{N} x_{ij}. \quad \text{(A17)}$$

Substituting Eq. (1) into the numerator of Eq. (A16), expanding this, and multiplying the denominator by a factor $(N - 1)/(N - 1)$, we obtain

$$r_{\text{PM}} = \frac{\sum_{i=1}^{N} (\mu_i - \mu)^2 + \sum_{i=1}^{N} (\mu_i - \mu)[\varepsilon_{i1} - (\varepsilon)_1 + \varepsilon_{i2} - (\varepsilon)_2] + \sum_{i=1}^{N} [\varepsilon_{i1} - (\varepsilon)_1][\varepsilon_{i2} - (\varepsilon)_2]}{(N-1)\left\{\frac{1}{N-1}\sum_{i=1}^{N} [x_{i1} - (x)_1]^2\right\}^{1/2}\left\{\frac{1}{N-1}\sum_{i=1}^{N} [x_{i2} - (x)_2]^2\right\}^{1/2}}. \quad \text{(A18)}$$

The first term of the numerator is now simply $(N - 1)\hat{\sigma}^2_{\text{SST}}$, the second term is estimated by zero since the signal due to oceanic forcing is uncorrelated with the internal noise (first assumption, section 2), and the third term is also estimated by zero since internal variations are uncorrelated between realizations (also first assumption, section 2). The denominator can be simplified by noting that each of the two main "terms" are unbiased estimates of σ^2_{TOT}. Hence,

$$r_{\text{PM}} = \frac{\hat{\sigma}^2_{\text{SST}}}{\hat{\sigma}^2_{\text{TOT}}} = \hat{\rho}. \quad \text{(A19)}$$

Second, consider instead a situation where an ensemble of model runs is available, and we choose to cor-

relate the observed time series with the ensemble means. To derive the ''ensemble mean perfect model correlation skill,'' the observed time series is viewed as an additional realization, labeled ''A,'' which is compared with an existing ensemble of n realizations. The correlation skill is then

$$r_{\mathrm{EMPM}} = \frac{\sum_{i=1}^{N} [x_{iA} - (x)_A](\overline{x}_i - \overline{\overline{x}})}{\left\{\sum_{i=1}^{N} [x_{iA} - (x)_A]^2 \sum_{i=1}^{N} (\overline{x}_i - \overline{\overline{x}})^2\right\}^{1/2}}. \quad \text{(A20)}$$

After substituting Eq. (1), an expansion of the numerator can be shown to give the same result as before, but the denominator now differs:

$$r_{\mathrm{EMPM}} = \frac{\hat{\sigma}^2_{\mathrm{SST}}}{\hat{\sigma}_{\mathrm{TOT}}\hat{\sigma}_{\mathrm{EM}}}. \quad \text{(A21)}$$

Using Eqs. (A6) and (A8) of Rowell et al. (1995), this can also be written

$$r_{\mathrm{EM\,PM}} = \hat{\rho} n^{1/2}[\hat{\rho}(n - 1) + 1]^{-1/2}. \quad \text{(A22)}$$

As expected, $r_{\mathrm{EM\,PM}} > r_{\mathrm{PM}}$ for $n > 1$ (unless $\hat{\rho} = 1$), and $r_{\mathrm{EM\,PM}} = r_{\mathrm{PM}}$ for $n = 1$. Also note that as $n \to \infty$, so $r_{\mathrm{EM\,PM}} \to \hat{\rho}^{1/2}$ (this is the well-known relationship between correlation and explained variance, where x_{ij} and its component μ_i are correlated).

Finally, one can also envisage using these relationships as part of a skill score, which, by comparing the actual correlation with the perfect model correlation, has the property of being unaffected by spatial heterogeneities of the impact of internal variability. An important caveat, however, is that such a score would inherently assume that the model estimate of ρ is a realistic estimate of ρ in the real atmosphere; this will not always be true.

REFERENCES

Aceituno, P., 1988: On the functioning of the Southern Oscillation in the South American sector. Part I: Surface climate. *Mon. Wea. Rev.,* **116,** 505–524.

Anderson, J. L., and W. F. Stern, 1996: Evaluating the potential predictive utility of ensemble forecasts. *J. Climate,* **9,** 260–269.

Barnston, A. G., 1994: Linear statistical short-term climate predictive skill in the Northern Hemisphere. *J. Climate,* **7,** 1513–1564.

——, and Coauthors, 1994: Long-lead seasonal forecasts—Where do we stand? *Bull. Amer. Meteor. Soc.*, **75,** 2097–2114.

Bottomley, M., C. K. Folland, J. Hsiung, R. E. Newell, and D. E. Parker, 1990: *Global Ocean Surface Temperature Atlas.* Joint Met. Office/MIT Project, 20 pp. and 313 plates.

Brankovic, C., and T. N. Palmer, 1997: Atmospheric seasonal predictability and estimates of ensemble size. *Mon. Wea. Rev.,* **125,** 859–874.

——, ——, and L. Ferranti, 1994: Predictability of seasonal atmospheric variations. *J. Climate,* **7,** 217–237.

Cane, M. A., S. E. Zebiak, and S. C. Dolan, 1986: Experimental forecasts of El Niño. *Nature,* **321,** 827–832.

Carleton, A. M., 1981: Ice–ocean–atmosphere interactions at high southern latitudes in winter from satellite observations. *Aust. Meteor. Mag.,* **29,** 183–195.

Charney, J. G., and J. Shukla, 1981: Predictability of monsoons. *Monsoon Dynamics,* J. Lighthill and R. P. Pearce, Eds., Cambridge University Press, 99–109.

Chervin, R. M., 1988: Predictability of time-averaged atmospheric states. *Physically-Based Modelling and Simulation of Climate and Climate Change—Part II,* M. E. Schlesinger, Ed., Kluwer Academic, 983–1008.

Cullen, M. J. P., 1993: The unified forecast/climate model. *Meteor. Mag.,* **122,** 81–94.

Curtis, S., and S. Hastenrath, 1995: Forcing of anomalous sea surface temperature evolution in the tropical Atlantic during Pacific warm events. *J. Geophys. Res.,* **100** (C8), 15 835–15 847.

Davies, J. R., D. P. Rowell, and C. K. Folland, 1997: North Atlantic and European seasonal predictability using an ensemble of multidecadal AGCM simulations. *Int. J. Climatol.,* **17,** 1263–1284.

Dix, M. R., and B. G. Hunt, 1995: Chaotic influences and the problem of deterministic seasonal predictions. *Int. J. Climatol.,* **15,** 729–752.

Donoghue, J. R., and L. M. Collins, 1990: A note on the unbiased estimate of the intraclass correlation. *Psychometrika,* **55,** 159–164.

Drosdowsky, W., 1993: An analysis of Australian seasonal rainfall anomalies: 1950–87. II: Temporal variability and teleconnection patterns. *Int. J. Climatol.,* **13,** 111–149.

Farmer, G., 1988: Seasonal forecasting of the Kenya short rains, 1901–84. *Int. J. Climatol.,* **8,** 489–497.

Fraedrich, K., 1994: An ENSO impact on Europe? A review. *Tellus,* **46A,** 541–552.

Gates, W. L., 1992: AMIP: The atmospheric model intercomparison project. *Bull. Amer. Meteor. Soc.,* **73,** 1962–1970.

Goswami, B. N., 1994: Dynamical predictability of seasonal monsoon rainfall: Problems and prospects. *Proc. Indian Natl. Sci. Acad.,* **60A,** 101–120.

Harzallah, A., and R. Sadourny, 1995: Internal versus SST-forced atmospheric variability as simulated by an atmospheric general circulation model. *J. Climate,* **8,** 474–495.

Hasselmann, K., 1979: On the signal-to-noise problem in atmospheric response studies. *Meteorology over the Tropical Oceans,* D. B. Shaw, Ed., Royal Meteorological Society, 251–259.

Hastenrath, S., 1976: Variations in low-latitude circulation and extreme climatic events in the tropical Americas. *J. Atmos. Sci.,* **33,** 202–215.

——, 1984: Interannual variability and annual cycle: Mechanisms of circulation and climate in the tropical Atlantic sector. *Mon. Wea. Rev.,* **112,** 1097–1107.

——, 1990: Predictability of anomalous river discharge in Guyana. *Nature,* **345,** 53–54.

——, 1995: Recent advances in tropical climate prediction. *J. Climate,* **8,** 1519–1532.

Hegerl, G. C., H. von Storch, K. Hasselmann, B. D. Santer, U. Cubasch, and P. D. Jones, 1996: Detecting greenhouse gas-induced climate change with an optimal fingerprint method. *J. Climate,* **9,** 2281–2306.

Hoel, P. G., 1971: *Introduction to Mathematical Statistics.* John Wiley, 409 pp.

Hurrell, J. W., and H. van Loon, 1994: A modulation of the atmospheric annual cycle in the Southern Hemisphere. *Tellus,* **46A,** 325–338.

James, I. N., and P. M. James, 1989: Ultra-low-frequency variability in a simple atmospheric circulation model. *Nature,* **342,** 53–55.

Ji, M., A. Kumar, and A. Leetmaa, 1994: A multiseason climate forecast system at the National Meteorological Center. *Bull. Amer. Meteor. Soc.,* **75,** 569–577.

Karoly, D. J., 1989: Southern Hemisphere circulation features associated with El Niño–Southern Oscillation events. *J. Climate,* **2,** 1239–1252.

Kiladis, G. N., and H. F. Diaz, 1989: Global climatic anomalies associated with extremes in the Southern Oscillation. *J. Climate,* **2,** 1069–1090.

Kumar, A., and M. P. Hoerling, 1995: Prospects and limitations of

seasonal atmospheric GCM predictions. *Bull. Amer. Meteor. Soc.*, **76,** 335–345.
Lau, N.-C., 1985: Modeling the seasonal dependence of the atmospheric response to observed El Niños in 1962–76. *Mon. Wea. Rev.*, **113,** 1970–1996.
Lorenz, E. N., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.*, **20,** 130–141.
——, 1990: Can chaos and intransitivity lead to interannual variability? *Tellus*, **42A,** 378–389.
——, 1993: *The Essence of Chaos.* University of Washington Press, 227 pp.
Madden, R. A., 1976: Estimates of natural variability of time-averaged sea level pressure. *Mon. Wea. Rev.*, **104,** 942–952.
Nicholls, N., 1981: Air–sea interaction and the possibility of long-range weather prediction in the Indonesian Archipelago. *Mon. Wea. Rev.*, **109,** 2435–2443.
Nicholson, S. E., and D. Entekhabi, 1987: Rainfall variability in equatorial and southern Africa: Relationships with sea surface temperature along the southwestern coast of Africa. *J. Climate Appl. Meteor.*, **26,** 561–578.
Opsteegh, J. D., and H. M. Van den Dool, 1980: Seasonal differences in the stationary response of a linearized primitive equation model: Prospects for long-range weather forecasting? *J. Atmos. Sci.*, **37,** 2169–2185.
Overland, J. E., and C. H. Pease, 1982: Cyclone climatology of the Bering Sea and its relation to sea-ice extent. *Mon. Wea. Rev.*, **110,** 5–13.
Palmer, T. N., 1993: Extended-range atmospheric prediction and the Lorenz model. *Bull. Amer. Meteor. Soc.*, **74,** 49–65.
——, 1994: Chaos and predictability in forecasting the monsoons. *Proc. Indian Natl. Sci. Acad.*, **60A,** 57–66.
——, and D. L. T. Anderson, 1994: The prospects for seasonal forecasting—A review paper. *Quart. J. Roy. Meteor. Soc.*, **120,** 755–793.
Parker, D. E., C. K. Folland, A. Bevan, M. N. Ward, M. Jackson, and K. Maskell, 1995: Marine surface data for analyses of climatic fluctuations on interannual to century time scales. *Natural Climate Variability on Decade to Century Time Scales*, D. G. Martinson, K. Bryan, M. Ghil, M. M. Hall, T. R. Karl, E. S. Sarachik, S. Sorooshian, and L. D. Talley, Eds., National Academy Press, 241–250.
Phillips, T. J., 1994: A summary documentation of the AMIP models. PCMDI Rep. 18, Univ. of California, Lawrence Livermore National Laboratory, Livermore, CA, 343 pp. [Available from Program for Model Diagnosis and Intercomparison, Lawrence Livermore National Laboratory, Livermore, CA 94550.]
Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, 1992: *Numerical Recipes in FORTRAN.* Cambridge University Press, 963 pp.
Rasmusson, E. M., and T. H. Carpenter, 1982: Variations in tropical sea surface temperature and surface wind fields associated with the Southern Oscillation/El Niño. *Mon. Wea. Rev.*, **110,** 354–384.
Ratcliffe, R. A., 1976: The hot spell of late June–early July 1976. *Weather*, **31,** 355–357.
Rogers, J. C., 1978: Meteorological factors affecting the interannual variability of summertime ice extent in the Beaufort Sea. *Mon. Wea. Rev.*, **106,** 890–897.
Ropelewski, C. F., and M. S. Halpert, 1987: Global and regional scale precipitation patterns associated with the El Niño/Southern Oscillation. *Mon. Wea. Rev.*, **115,** 1606–1625.
——, and ——, 1989: Precipitation patterns associated with the high index phase of the Southern Oscillation. *J. Climate*, **2,** 268–284.
Rowell, D. P., 1996a: Using an ensemble of multidecadal GCM simulations to assess potential seasonal predictability. Climate Research Tech. Note 69, 40 pp. [Available from Hadley Centre, U.K. Meteorological Office, Bracknell, Berkshire RG12 2SY, United Kingdom.]
——, 1996b: Reply to comments by Y. C. Sud and W. K.-M. Lau on 'Variability of summer rainfall over tropical North Africa (1906–92): Obervations and modelling' by D. P. Rowell, C. K. Folland, K. Maskell, and M. N. Ward (April 4, 1995, 121, 669–704): Further analysis of simulated interdecadal and interannual variability of summer rainfall over tropical North Africa. *Quart. J. Roy. Meteor. Soc.*, **122,** 1007–1013.
——, C. K. Folland, K. Maskell, and M. N. Ward, 1995: Variability of summer rainfall over tropical North Africa (1906–92): Observations and modeling. *Quart. J. Roy. Meteor. Soc.*, **121,** 669–704.
Santer, B. D., T. M. L. Wigley, T. P. Barnett, and E. Anyamba, 1996: Detection of climate change and attribution of causes. *Climate Change 1995: The Science of Climate Change. The Second Assessment Report of the Intergovernmental Panel on Climate Change*, J. J. Houghton, L. G. Meira Filho, B. A. Callander, N. Harris, A. Kattenberg, and K. Maskell, Eds., Cambridge University Press, 407–443.
Scheffe, H., 1959: *The Analysis of Variance.* John Wiley and Sons, 477 pp.
Searle, S. R., G. Casella, and C. E. McCulloch, 1992: *Variance Components.* John Wiley and Sons, 501 pp.
Shukla, J., 1985: Predictability. *Advances in Geophysics*, Vol. 28B, Academic Press, 87–122.
——, and D. S. Gutzler, 1983: Interannual variability and predictability of 500-mb geopotential heights over the Northern Hemisphere. *Mon. Wea. Rev.*, **111,** 1273–1279.
Stern, W., and K. Miyakoda, 1995: The feasibility of seasonal forecasts speculated from multiple GCM simulations. *J. Climate*, **8,** 1071–1085.
Sud, Y. C., and W. K.-M. Lau, 1996: Comments on 'Variability of summer rainfall over tropical North Africa (1906–92): Observations and modelling'. *Quart. J. Roy Meteor. Soc.*, **122,** 1001–1006.
Trenberth, K. E., 1985: Potential predictability of geopotential heights over the Southern Hemisphere. *Mon. Wea. Rev.*, **113,** 54–64.
——, and G. W. Branstator, 1992: Issues in establishing causes of the 1988 drought over North America. *J. Climate*, **5,** 159–172.
Walsh, J. E., 1993: Observational and modelling studies of the influence of sea ice anomalies on atmospheric circulation. *Prediction of Interannual Climate Variations*, J. Shukla, Ed., Springer–Verlag, 71–88.
Ward, M. N., and C. K. Folland, 1991: Prediction of seasonal rainfall in the North Nordeste of Brazil using eigenvectors of sea surface temperatures. *Int. J. Climatol.*, **11,** 711–743.
——, C. K. Folland, K. Maskell, A. Colman, D. P. Rowell, and K. Lane, 1993: Experimental seasonal forecasting of tropical rainfall at the UK Meteorological Office. *Prediction of Interannual Climate Variations*, Springer–Verlag, 197–216.
Warrilow, D. A., A. B. Sangster, and A. Slingo, 1986: Modelling of land surface processes and their influence on European climate. Dyn. Climatol. Tech. Note, 92 pp. [Available from Hadley Centre, U.K. Meteorological Office, Bracknell, Berkshire RG12 2SY, United Kingdom.]
Webster, P. J., 1982: Seasonality in the local and remote atmospheric response to sea surface temperature anomalies. *J. Atmos. Sci.*, **39,** 41–52.
Willmott, C. J., C. M. Rowe, and Y. Mintz, 1985: A global archive of land cover and soils data for use in general circulation climate models. *Int. J. Climatol.*, **5,** 119–143.
Zwiers, F. W., 1987: A potential predictability study conducted with an atmospheric general circulation model. *Mon. Wea. Rev.*, **115,** 2957–2974.
——, 1996: Interannual variability and predictability in an ensemble of AMIP climate simulations conducted with the CCC GCM2. *Climate Dyn.*, **12,** 825–848.
——, D. P. Rowell, A. Hense, J. R. Davies, and M. Christoph, 1995: AGCM intercomparison diagnostics: SST-forced variability vs internal variability. *Workshop on Simulations of the Climate of the Twentieth Century Using GISST*, Bracknell, United Kingdom, UKMO, 35–39.