Predictability of Cold Spring Seasons in Europe

Mxolisi E. Shongwe*

Royal Netherlands Meteorological Institute, De Bilt, Netherlands

Christopher A. T. Ferro

Walker Institute, Department of Meteorology, University of Reading, Reading, United Kingdom

Caio A. S. Coelho

Centro de Previsão de Tempo e Estudos Climàticos, Instituto Nacional de Pesquisas Espaciais, Cachoeira Paulista, Brazil

Geert Jan van Oldenborgh

Royal Netherlands Meteorological Institute, De Bilt, Netherlands

(Manuscript received 17 November 2006, in final form 28 February 2007)

ABSTRACT

The seasonal predictability of cold spring seasons (March–May) in Europe from hindcasts/forecasts of three operational coupled general circulation models (CGCMs) is investigated. The models used in the investigation are the Met Office Global Seasonal Forecast System (GloSea), the ECMWF System-2 (S2), and the NCEP Climate Forecast System (CFS). Using the relative operating characteristic score and the Brier skill score the long-term prediction skill for spring 2-m temperature in the lower quintile (20%) is assessed. Over much of central and eastern Europe the predictive skill is found to be high. The skill of the Met Office GloSea and ECMWF S2 models significantly surpasses that of damped persistence over much of Europe but the NCEP CFS model outperforms this reference forecast only over a small area. The higher potential predictability of cold spring seasons in eastern relative to southwestern Europe can be attributed to snow effects as areas of high skill closely correspond with the climatological snow line, and snow is shown in this paper to be linked to cold spring 2-m temperatures in eastern Europe. The ability of the models to represent snow cover during the melt season is also investigated. The Met Office GloSea and the ECMWF S2 models are able to accurately mimic the observed pattern of monthly snow-cover interannual variability, but the NCEP CFS model predicts too short a snow season. Improvements in the snow analysis and land surface parameterizations could increase the skill of seasonal forecasts for cold spring temperatures.

1. Introduction

The seasonal prediction of near-surface temperatures over many parts of the globe has received considerable attention. In fact, together with precipitation, seasonal prediction of 2-m temperature has a wide application. Considerable effort has gone into predictions of mean temperatures over many parts of the globe including Europe (e.g., Barnston and Smith 1996). However, it is extreme temperatures such as heat waves and cold outbreaks that have a larger effect on human society. Despite the worldwide notion of global warming and its impacts, the absence of a clear trend toward fewer extreme cold events in Europe (Klein Tank et al. 2002) has prompted an investigation of their potential predictability in this study. These extreme events typically occur in winter and spring possibly due to snow effects, and have huge impacts on transportation systems, energy supply, ecology, agriculture, winter tourism, the clothing industry, etc.

The predictable component of atmospheric variability at seasonal to interannual time scales is that forced

* Additional affiliation: Swaziland National Meteorological Service, Mbabane, Swaziland.

Corresponding author address: Mxolisi E. Shongwe, KNMI, P.O. Box 201, De Bilt 3730 AE, Netherlands.
E-mail: mxolisi.shongwe@knmi.nl

TABLE 1. Coupled GCMs used in this study. Model resolution is given as wavenumber of spectral truncation (T) and number of vertical layers (L). The Met Office GloSea model horizontal resolution is given as a latitude–longitude grid.

Model	Atmospheric resolution	Initialization	Hindcasts from	Ensemble size	Forecasts from	Ensemble size	References
ECMWF S2	T95L40	ERA-15	1987	5	2001	40	Anderson et al. (2003); van Oldenborgh et al. (2005a,b)
NCEP CFS	T62L64	NCEP	1981	15	2004	60	Saha et al. (2006)
Met Office GloSea	2.5° × 3.75°, L19	ERA-40	1987	15	2004	40	Graham et al. (2005)

by changes in boundary conditions (e.g., Barnston et al. 2005). Walsh et al. (2001) related extreme winter and spring cold outbreaks to negative North Atlantic Oscillation (NAO) signatures consistent with a "blocking" of westerly airflow into Europe. However, at smaller spatial scales the influence of large-scale circulation anomalies such as those related to SST variability is modulated by interactions and feedbacks between the atmosphere and the land surface. Among other land surface processes, observational and modeling studies have revealed that snow impacts near-surface temperature variability in the Northern Hemisphere extratropics during the cold season (Walsh et al. 1982, 1985; Yang et al. 2001; Kumar and Yang 2003).

In Europe near-surface temperature variability has been attributed to three main factors: large-scale flow, which determines the origin and tracks of air masses (e.g., van Oldenborgh and van Ulden 2003), radiation balance (as determined by cloudiness among other factors; Lenderink et al. 2006), and local lower boundary conditions (Ferranti and Viterbo 2006). Palmer et al. (2004) have shown that seasonal predictability of large-scale flow in Europe is low. The effect of cloudiness offers some limited predictability. Skill from SST persistence has been shown to be confined mainly along coastal areas (Van den Dool and Nap 1985).

This paper aims to investigate the predictability of cold spring seasons in Europe. We investigate the hypothesis of snow accumulated during the preceding winter being a source of skill in predicting cold spring seasons in Europe. Snow increases the surface albedo, thereby altering the terrestrial heat balance and providing a positive feedback mechanism that modulates atmospheric variability. In the first part of the paper, we assess the coupled general circulation models (CGCMs) below-median, lower-, and upper-quintile (i.e., coldest and warmest 20% of the climatological records) 2-m temperature seasonal predictive skill over Europe in spring [March–May (MAM)]. Near-surface temperature forecasts obtained from damped persistence of low temperatures from the previous winter and early February snow depth are used as a baseline to

judge the performance of the CGCMs. We then relate the 2-m temperature to snow water equivalent (SWE) and snow cover to explain the physical basis of the model skill.

2. Data and methods

a. CGCM predictions

The CGCMs used in this study are the Met Office Global Seasonal Forecast System (GloSea), the European Centre for Medium-Range Weather Forecasts System-2 (ECMWF S2), and the National Centers for Environmental Prediction Climate Forecast System (NCEP CFS). For more details on these models see Table 1. The model data are accessible through the Royal Netherlands Meteorological Institute (KNMI) Climate Explorer (see online at climexp.knmi.nl; van Oldenborgh and Burgers 2005). The ECMWF S2 consists of five members in each ensemble for the hindcast period (1987–2001) and 40 members thereafter. Only five members in the forecast period have been used to match the hindcast period. The NCEP CFS and GloSea models have 15 ensemble members each in hindcast mode, which are used in this study. However, since 2004, the NCEP CFS model produces twice-daily operational forecasts resulting in about 60 ensemble members. GloSea produces 40 ensemble members for the operational forecasts since 2004. The February-start hindcasts/forecasts have been verified against MAM 2-m temperature, implying a 1-month lead time (but half-month from dissemination time, usually the 15th of each month). An extreme event is defined whenever 2-m temperatures fall within the lower quintile of the climatological records, otherwise it is a nonevent. Events were defined for model predictions and observations independently (i.e., based on their own model predictions and observed distributions). When using this approach model biases are inherently corrected. Model ensemble predictions are converted into probabilistic forecasts by expressing, as a percentage, the fraction of the ensemble predicting the extreme event.
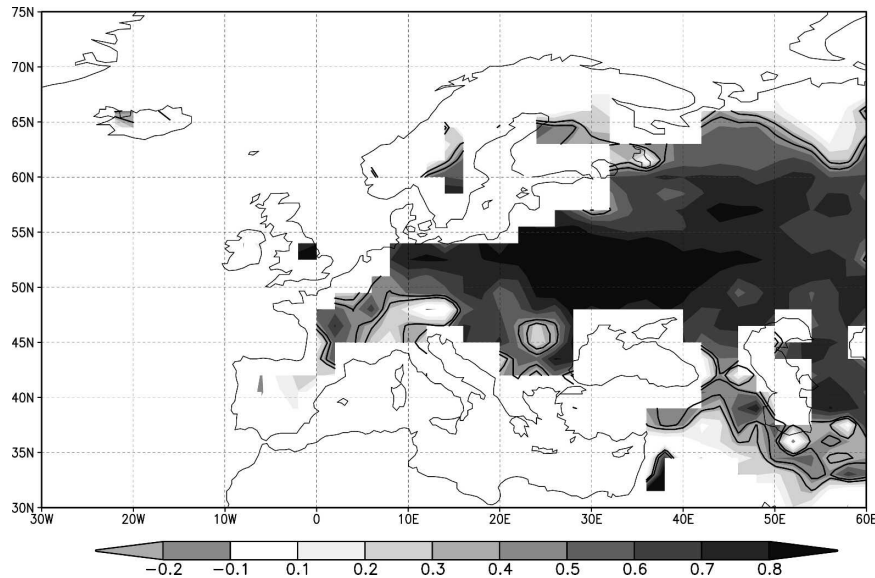
FIG. 1. Correlation plot between FMA NOAA/NESDIS and ERA-40-derived snow cover. A correlation of 0.34 (0.44) is found to be statistically significant at the 5% (1%) level, and the corresponding contours are plotted.

For instance if 9 members out of 15 are fall into the lowest quintile, and the rest fall outside this category, then the probability for the event occurring (not occurring) is 60% (40%).

b. Observations

The snow-depth data (expressed as SWE) used in this study are a combination of data from the 40-yr ECMWF Re-Analysis (ERA-40) project (Uppala et al. 2005) up to 2002, and the ECMWF snow analysis (White 2003) from 2003 to present. Weekly snow-cover data spanning the 1972–2006 period, obtained from National Oceanic and Atmospheric Administration/National Environmental Satellite Data and Information Service (NOAA/NESDIS) datasets (Robinson et al. 1993), are also used. The weekly snow cover has been averaged into monthly data. To compare these two datasets, snow extent has been generated for each grid by assigning a one if the ECMWF SWE on a given day is at least 3 mm, and 0 otherwise. The 3-mm SWE threshold yielded the best agreement with the NOAA/NESDIS dataset. Monthly averages are generated based on these derived variables. Given that the effect of the extent of snow cover on 2-m temperature is felt during and after its existence (e.g., Groisman et al. 1994) and because the atmosphere adjusts rapidly in response to anomalous land surface forcing, only short (1 month) lead–lag relationships are considered whereby February–April (FMA) mean snow is related to MAM mean 2-m temperature. The 2-m temperature

data have been obtained from the recently developed high-resolution ($0.5° \times 0.5°$ latitude–longitude grid) gridded monthly dataset of Fan and Van den Dool (2007), which combines the Global Historical Climatology Network (GHCN), version 2, and the Climate Anomaly Monitoring System (CAMS) datasets.

Pointwise correlations between FMA NOAA/NESDIS and ERA-40-derived FMA mean snow cover calculated over the period 1972–2005 are shown in Fig. 1. Statistical significance at the 5% (1%) level (as obtained from a one-sided two-sample t test with 32 degrees of freedom) corresponds to a correlation coefficient of 0.34 (0.44). The high correspondence between the two datasets is evident over a wide area extending eastward from 10°E. Areas with zero standard deviation during FMA (e.g., northern Scandinavia and Africa in or close to the Sahara Desert) are masked out in the plot. Biases in remotely sensed data such as NOAA/NESDIS satellite data exist (Robinson et al. 1993; Foster et al. 1996). The ERA-40 snow-depth analysis assimilates some observations, otherwise it relies completely on the model snow precipitation and melt parameterizations, which are known to have biases. These limitations notwithstanding, the high correlation between the remotely sensed and reanalysis data over Europe is encouraging. Because it is quite unlikely that both datasets would have the same biases, both snow datasets should be reliable enough for our purposes.

c. Forecast verification

Skill measures used to assess the CGCMs probabilistic temperature forecasts are the relative operating characteristic (ROC) score and the Brier skill score (BSS). These metrics and other forecast verification measures are computed by the KNMI Climate Explorer using statistical software developed by the R Software for Climate Analysis (RCLIM) initiative (see online at http://www.secam.ex.ac.uk/index.php?nav=696). The ROC score measures the ability of the forecast system to correctly detect events or nonevents. The reference ROC score of 0.5 implies no skill in the forecasts, and the score increases to a maximum of 1.0 for perfect forecasts. Scores less than 0.5 are indicative of negative skill (i.e., worse than chance). The numerical formulation and comprehensive discussion of the ROC score can be found in Mason and Graham (1999, 2002). For a set of n forecasts, if n_1 (n_2) is the number of cases in which the event occurs (does not occur), Mason and Graham (2002) define the ROC score for probabilistic forecasts from an ensemble of size m as

$$A_{m,n} = 1 - \frac{1}{n_1 n_2} \sum_{\substack{1 \le i \le n_1 \\ 1 \le j \le n_2}} I(P_{1i} < P_{2j})$$

$$- \frac{1}{2n_1 n_2} \sum_{\substack{1 \le i \le n_1 \\ 1 \le j \le n_2}} I(P_{1i} = P_{2j}), \quad (1)$$

where P_{1i} (P_{2j}) are the forecasts issued prior to an event (nonevent), and $I(B) = 1$ if condition B holds, 0 otherwise. Suppose a subset K_{1i} (K_{2j}) of the ensemble predict the event (nonevent). Given an ensemble of size m, C. A. T. Ferro (2007, unpublished manuscript) has shown that an unbiased estimator for the expected ROC score that would be obtained by an ensemble of size $M \le m$ is

$$\hat{A}_{M,n} = 1 - \frac{1}{n_1 n_2} \sum_{\substack{1 \le i \le n_1 \\ 1 \le j \le n_2}} \left\{ \sum_{(k,l) \in L_{ij}} \binom{K_{1i}}{l} \binom{m - K_{1i}}{M - l} \right.$$

$$\times \binom{K_{2j}}{k} \binom{m - K_{2j}}{M - k} + \frac{1}{2} \sum_{k \in \kappa_{ij}} \binom{K_{1i}}{k}$$

$$\left. \times \binom{m - K_{1i}}{M - k} \binom{K_{2j}}{k} \binom{m - K_{2j}}{M - k} \right\} / \binom{m}{M}^2, \quad (2)$$

where L_{ij} is the subset of $\{(k, l): 0 \le l < k \le M\}$, such that the combinations in the first summation exist, and κ_{ij} is the subset of $\{0, \dots, M\}$ such that the combinations in the second summation exist.

Aware of the effect of ensemble size on ROC scores

across the three CGCMs used in this research, the expected ROC scores have been estimated for each model for 50% (for below-median 2-m temperature predictions) and 20% (lower quintile) thresholds using $M = 5$ corresponding to the ECMWF S2 ensemble size before 2002 (see Table 1). This gives us confidence that the intermodel-ROC-score differences shown below reflect true CGGMs' individual strengths and weaknesses. Critical values are calculated from a normal approximation to the distribution of the ROC score under a null hypothesis of random forecasts, with issued probabilities distributed uniformly on $M + 1$ distinct values. Differences in the ROC scores have been computed and their statistical significance tested using the method discussed by C. A. T. Ferro (2007, unpublished manuscript).

The relative improvement of numerical model predictions of cold seasons over an empirical approach is assessed using Brier skill scores with reference to forecasts from a simple statistical model. The BSS measures the forecast system's improvement over a reference forecast strategy (Wilks 1995).

The reference empirical model for spring temperature (T'_{MAM}) is a linear model based on three predictors: global warming, temperature persistence from winter into spring (T'_{NDJ}), and snow depth at analysis time (H_{1Feb}):

$$T'_{MAM}(y) = \alpha(y)T'_{NDJ}(y - 1) + \beta(y)H_{1Feb}(y) + \eta(y), \quad (3)$$

$$T'_s(y) = T_s(y) - \frac{1}{10} \sum_{y'=y-11}^{y-1} T_s(y'), \quad (4)$$

with y the year being forecast. Global warming is taken into account as a running mean of the local observed temperature 10 yr prior to the forecast time as shown in Eq. (4) for a given season. In many areas this "optimal normal correction" gives most skill to seasonal forecasts (Huang et al. 1996). The persistence predictor is the November–January-averaged temperature anomaly relative to this climatology. Temperature persistence from winter into spring is known to be significant in Europe. The third predictor is the ERA-40 snow-depth analysis at analysis time, 1 February. The coefficients $\alpha(y)$ and $\beta(y)$ are determined by linear regression of all years except the ones being forecast (jackknife). Winter temperature and snow depth are correlated ($r \approx 0.5$ in eastern Europe) as a thick snow layer in eastern Europe on 1 February is often produced by a cold winter preceding it. Because of this, statistical forecast models based on persistence only ($\beta = 0$), snow depth only ($\alpha = 0$), and a combination of the two have very similar skill in forecasting the spring temperature.

To compute probabilistic skill scores for the empirical forecast model, an ensemble of forecasts is required. This ensemble, of equal size N as the CGCM ensemble, is generated by sampling the distribution of the residuals $\eta(y)$ at the quantiles $i/(N + 1)$, $(i = 1, \ldots, N)$.

d. Near-surface temperature–snow relationships

Scatter diagrams are plotted to show relationships between spatial averages of observed FMA mean snow and MAM mean 2-m temperature over the 1972–2005 period. Canonical correlation analysis (CCA) is then used to diagnose the spatial extent of the linear relationship between the two fields. CCA is a multivariate statistical technique that seeks to identify a sequence of pairs of patterns in the data whose time evolution is optimally correlated. This technique has been widely applied in seasonal climate prediction studies (e.g., Landman and Mason 2001; Shongwe et al. 2006), and its procedure in the context of climate data analysis can be found in Wilks (1995) and von Storch and Zwiers (1999). Prior to the CCA, the spatial degrees of freedom were reduced by projecting original data onto their empirical orthogonal functions (EOFs; Barnett and Preisendorfer 1987). Performing the CCA on the EOF space minimizes large sampling errors (Bretherton et al. 1992) at the expense of possible losses of useful information in the original data from EOF truncation. The EOF space is truncated at 70% of the average characteristic root, which is the Guttman–Kaiser criterion (Jackson 1991) modified after Jolliffe (1972). The length of the sequence of successive pairs of canonical variates (CCA modes) is limited to be the minimum of the number of principal components (EOF modes) retained in the EOF analysis. The optimal combination of EOF and CCA modes producing the best fit are then determined from cross-validated sensitivity tests.

3. Skill of MAM 2-m temperature forecasts

a. Model ROC scores

Maps of ROC scores calculated for CGCM probabilistic predictions of MAM 2-m temperatures over the region 30°–75°N, 30°W–60°E are shown in Figs. 2–3. Figure 2 compares the ROC scores for the GloSea model for median (Fig. 2a) and lower-quintile (Fig. 2b) 2-m temperature. Differences between scores for model predictions of cold springs and below-median temperatures (representing normal seasons) are shown in Fig. 2c. Areas where the null hypothesis of no difference in ROC scores for the different thresholds (here, 20% and 50%) could be rejected at the 10%

error level are shaded in Fig. 2c. Corresponding comparisons are made for the ECMWF S2 (Figs. 2d–f) and NCEP CFS (Figs. 2g–i) systems.

ROC scores in excess of 0.6 for below-median 2-m temperature are confined to the area around the Baltic Sea extending from Poland northward into Scandinavia in GloSea (Fig. 2a) and ECMWF S2 (Fig. 2d) systems. At the 10% significance level, the critical ROC scores for rejection of the null hypothesis of unskillful forecasts are 0.7 for GloSea and ECMWF S2, and 0.68 for the NCEP CFS models. For below-median forecasts the models' ROC scores are not statistically significant over much of Europe. The NCEP CFS forecasts show little or no skill (Fig. 2g).

ECMWF S2 and GloSea models show elevated ROC scores (>0.7) for predictions of 2-m temperature in the lowest quintile. This is particularly the case over eastern Europe east of 10°E. GloSea and ECMWF S2 attain ROC scores in excess of 0.8 for the cold extremes (Figs. 2b,e). Widespread positive differences between scores for the 20% and 50% thresholds are evident over much of Europe except the southwest, although only statistically significant differences are shaded in the figures. Remarkably, for GloSea and ECMWF S2 systems, statistically significant differences at the 10% level are found over a wide area extending from the North Sea through central toward eastern Europe (Figs. 2b,f). The CGCMs show either no skill or negative skill over southwestern Europe. Albeit least skillful overall, the NCEP CFS model also can provide skilful forecasts over eastern Scandinavia and western Russia.

The high ROC scores and differences for below-median and lower-quintile 2-m temperature over much of central and eastern Europe warrants an investigation of their physical basis. Interestingly, the lowest-quintile 2-m temperature forecasts are more skillful than those of the upper quintile (≥80% of the distribution), notably in the ECMWF S2 model (Fig. 3). We hypothesize that elevated ROC scores could be due to a successful representation of snow persisting in spring and influencing near-surface temperatures (Cohen and Rind 1991). This mechanism naturally gives rise to higher skill for cold extreme forecasts than below-median or upper-quintile 2-m temperature forecasts.

b. Model BSSs

The CGCMs advantage over a statistical model assessed using BSS are shown in Fig. 4. For below-median 2-m temperature predictions GloSea outscores the statistical model forecasts over much of Europe west of 40°E (Fig. 4a). The ECMWF S2 and NCEP models perform slightly better than damped persistence forecasts only over patchy areas in Europe, otherwise the
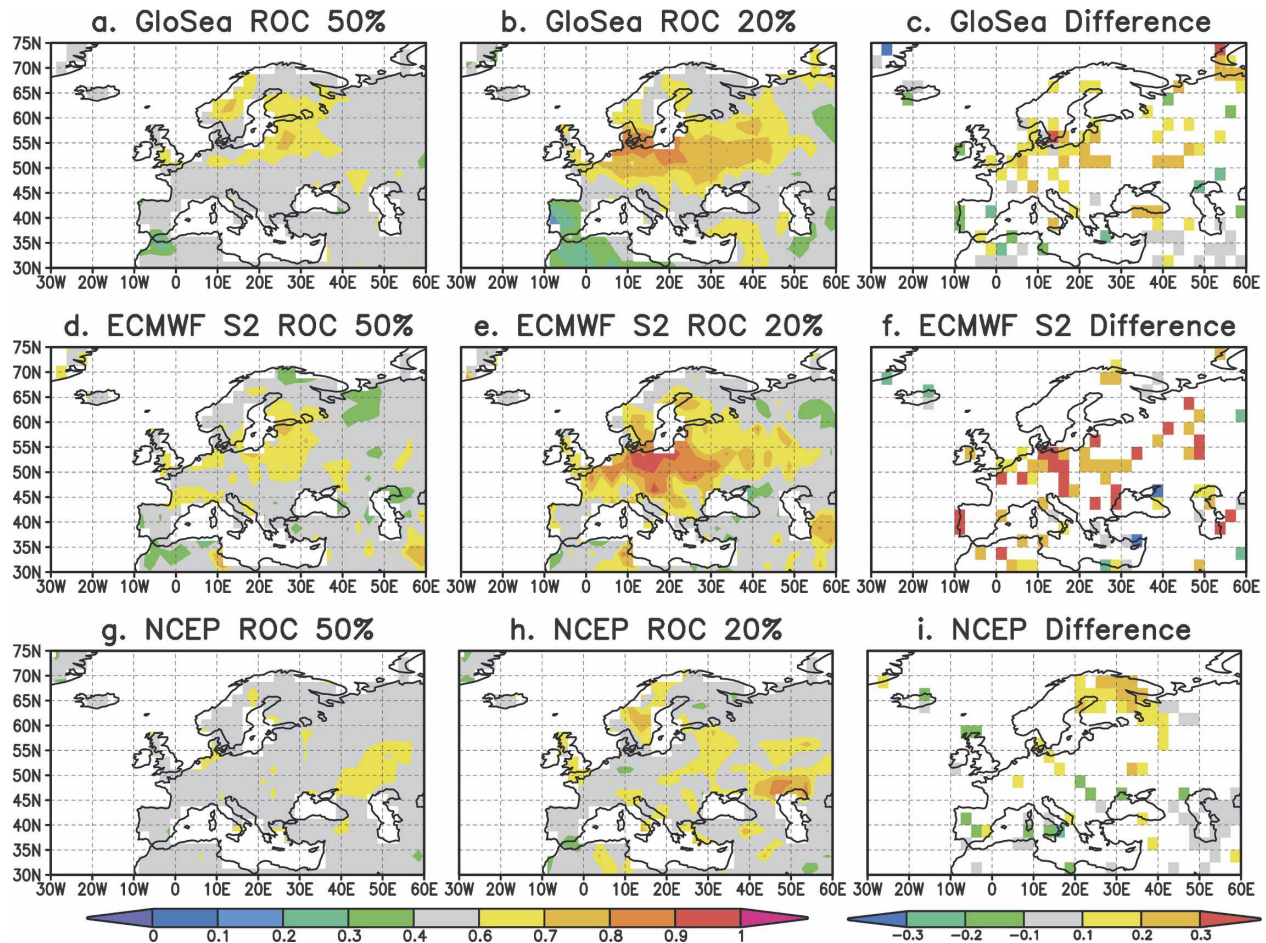
FIG. 2. Geographical distribution of ROC scores for GloSea predictions of MAM 2-m temperature: ROC scores for (a) below-median and (b) lower-quintile 2-m temperature predictions. ROC scores in excess of 0.7 are statistically significant at the 10% level. (c) Differences in the ROC scores (i.e., $\text{ROC}_{20} - \text{ROC}_{50}$). Areas where the differences are statistically significant at the 10% level are colored. (d), (e), (f) Same as in (a), (b), (c), but for ECMWF S2. (g), (h), (i) Same as in (a), (b), (c), but for NCEP CFS. ROC scores greater than 0.68 are statistically significant at the 10% level.

models attain skill equal to or less than that of the reference forecast over a wide area in Europe (Figs. 4b,c).

For lower-quintile 2-m temperature forecasts, Glo-Sea and ECMWF S2 models surpass damped persistence over a wide area extending from the North and Baltic Seas through central Europe toward the Mediterranean area. Negative skill scores are confined to a narrow area from about 15° to 20°E in GloSea. The BSS provide further evidence that there is more potential for forecasting cold spring seasons than below-median seasons; this is particularly noticeable in the case of ECMWF S2, which attains more widespread positive skill scores in Europe (Fig. 4e). However, in southwestern Europe, the models perform worse than damped persistence. The NCEP CFS outperforms the reference forecasts over isolated areas (Fig. 4f). At the

short lead time considered in this paper, Brier skill scores show that statistical models provide competitive predictions of lower-quintile 2-m temperatures over certain regions (e.g., southwestern Europe and western Russia).

c. Ensemble 2-m temperature predictions in eastern Europe

Bias-corrected ensemble model predictions for March and spring 2-m temperature, spatially averaged over eastern Europe (45°–55°N, 20°–30°E) for the period 1987–2005 (common to all models) are shown in Fig. 5. The number of members in each model ensemble are as shown in Table 1. The spatial averages are taken over the area characterized by high skill for lowest-quintile 2-m temperature forecasts in spring (as shown in Figs. 2 and 4). Prior to the area averaging, the
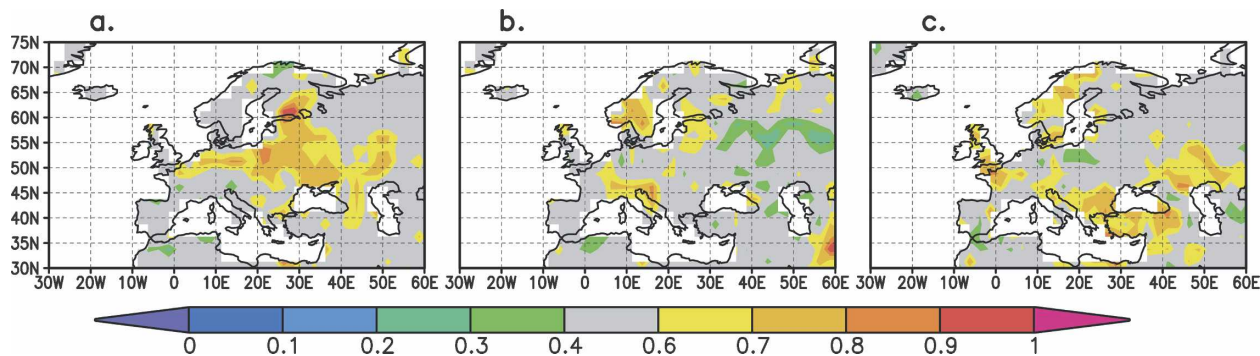
FIG. 3. Geographical distribution of ROC scores for CGCMs predictions of MAM in the upper quintile: (a) GloSea, (b) ECMWF S2, and (c) NCEP CFS.

CGCM predictions were bias corrected for the mean in a cross-validation mode. Anomalies are defined as departures from the 1971–2000 climatology. Observed 2-m temperature anomalies for the month and season in a given year are denoted by squares. The climatological lower-quintile category is shown by the gray shaded region. SWE anomalies for February, March, and April averaged over the same area are plotted in Fig. 6.

In Fig. 5b, years characterized by cold spring 2-m temperatures common to all models are 1987, 1996, 2005, and 1997, although in some cases not all 3 months in each season fell in the lowest quintile. Noting that March exhibits the highest skill of all spring months (mainly from snow signal; map not shown), attention is placed on the distribution of the ensemble members during March of each cold spring (leaving out the warmer March 1997; Fig. 5a). As shown in Fig. 6, a thick snowpack occurred from February to March preceding each coldest event, an exception being 1997.

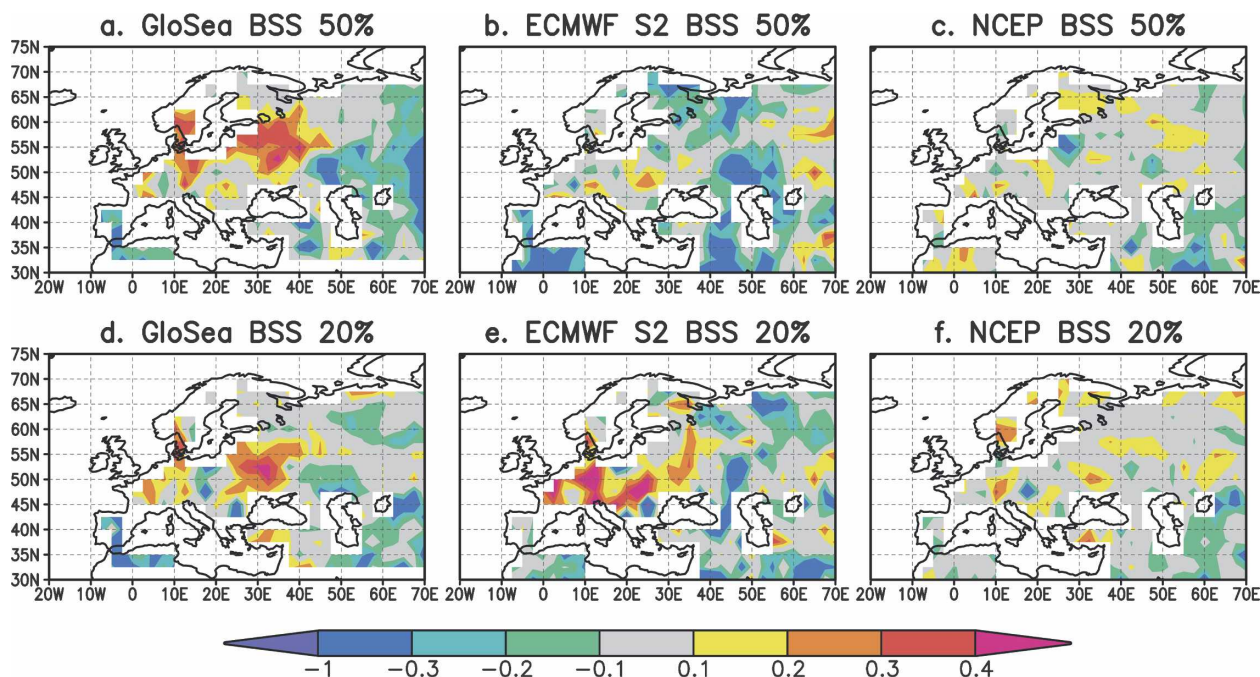In the prediction for March 1987, GloSea had 12 ensemble members (80%) falling within the lowest quin-

FIG. 4. Evaluation of the CGCM MAM 2-m temperature forecasts' BSS with reference to damped persistence. Skill scores for (a), (b), (c) below-median 2-m temperature forecasts and (d), (e), (f) lower-quintile scores. Skill scores are for (a), (d) GloSea; (b), (e) ECMWF S2; and (c), (f) NCEP CFS.

a. March 2–m temperature ensemble predictions

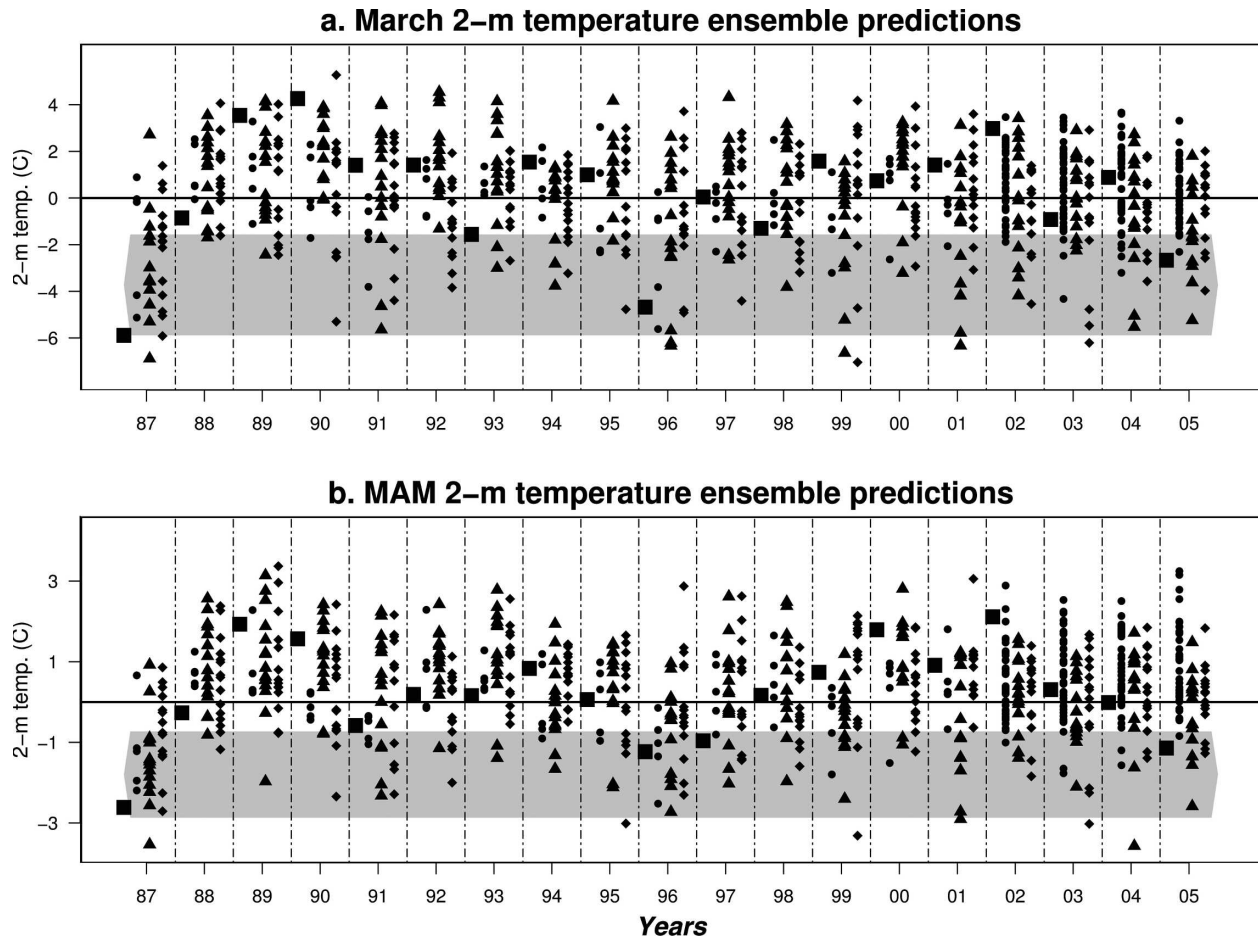## b. MAM 2–m temperature ensemble predictions

Years

FIG. 5. CGCM ensemble predictions for (a) March and (b) spring. In both (a) and (b) and for a given year (left) the squares show the observed anomalies, (middle left) the circles show the ECMWF S2 ensemble predictions, (middle right) the triangles show the GloSea ensemble predictions, and (right) the diamonds show the NCEP CFS ensemble predictions. The observations/predictions have been averaged over eastern Europe (45°–55°N, 20°–30°E). The gray shaded region shows the lower-quintile category defined from the 1971–2000 climatology.

tile, while 9 NCEP CFS model members (60%) predicted the extreme event. The ECMWF S2 model predicted the event with 40% probability (two members). In 1996, GloSea had seven of its members (≈47%) correctly predicting the event. ECMWF S2 anticipated the event with 40% probability, while the NCEP CFS model indicated about 27% (four members) chance. March 2005 cold event was predicted with 60% probability (nine members) by GloSea, NCEP CFS with about 27% while the ECMWF S2 system indicated only a 5% chance of 2-m temperatures falling within the lowest quintile.

Considering all spring months (MAM), the 1987 cold event was predicted very well by the GloSea and ECMWF S2 systems with each indicating at least 80% probability (12 and 4 members, respectively) of lowest-quintile 2-m temperatures. The NCEP CFS model pre-

dicted this event with about 53% probability. The ECMWF S2 system predicted the 1996 cold event with high probability (60%), with GloSea and NCEP CFS models indicating a 47% and 27% likelihood, respectively. The 1997 and 2005 events were least skillfully predicted with all models predicting the 1997 cold spring with 20% probability, equivalent to the lower-quintile prior probability or climatological forecasts. The 2005 event was predicted with low probability of about 33% and 20% by GloSea and NCEP CFS models, respectively, while the ECMWF S2 model almost completely missed (≈3% probability) the 2005 cold spring season in eastern Europe. As shown in Fig. 6, the 1997 cold event was preceded by negative SWE anomalies from February to March. It is therefore not surprising that this "atypical" event was either underforecast or completely missed by the CGCMs.
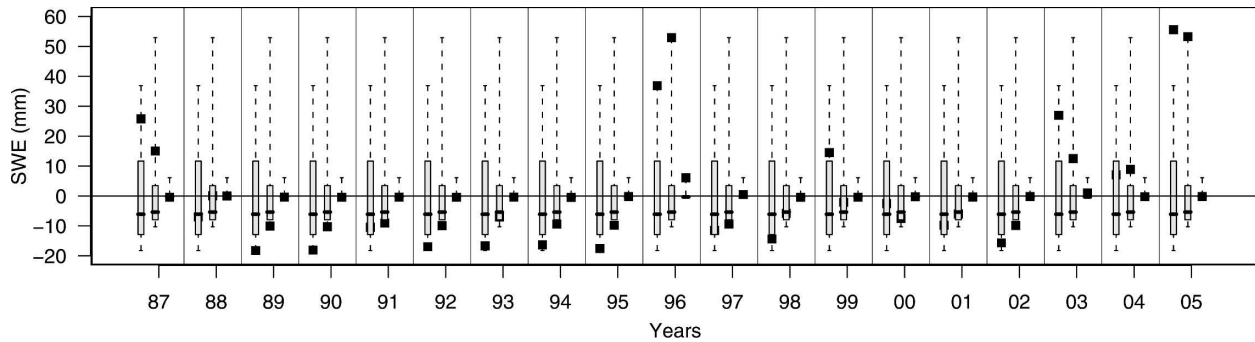
FIG. 6. Area-averaged FMA SWE in eastern Europe (same geographical area as in Fig. 5). Boxplots show the climatological range based on the 1971–2000 period, while the black squares indicate the observed SWE in a given month. For each year, the first series shows February SWE, the second series shows March, and the third series shows April SWE.

4. Snow influence

Snow-cover variability for FMA expressed as standard deviations of the extent of snow cover on the ground is shown in Fig. 7. The figure shows that areas in central and eastern Europe are characterized by the greatest interannual fluctuations of early spring snow cover. The other neighboring regions (e.g., southwest Europe and parts of Scandinavia) have a high frequency of either 0% or 100% snow cover during early spring months. There is a close correspondence in geographical area between the regions where this land surface condition is most variable from year to year, and where ROC scores for cold springs are highest (in GloSea and ECMWF S2 models; Fig. 2).

a. 2-m temperature–snow relationships in observations

1) SNOW AND NEAR-SURFACE TEMPERATURE RELATIONSHIPS IN EASTERN EUROPE

To illustrate the influence of changes in local land surface forcing associated with snow on near-surface temperature, spatial averages of SWE, snow cover, and 2-m temperature were calculated for a rectangular area (45°–55°N, 20°–30°E) and plotted in Fig. 8. As shown in Fig. 7, the area averages are taken over a region with highest interannual snow-cover variability in FMA and highest skill for cold spring forecasts (Fig. 2). In Fig. 8a, MAM 2-m temperature is plotted against February
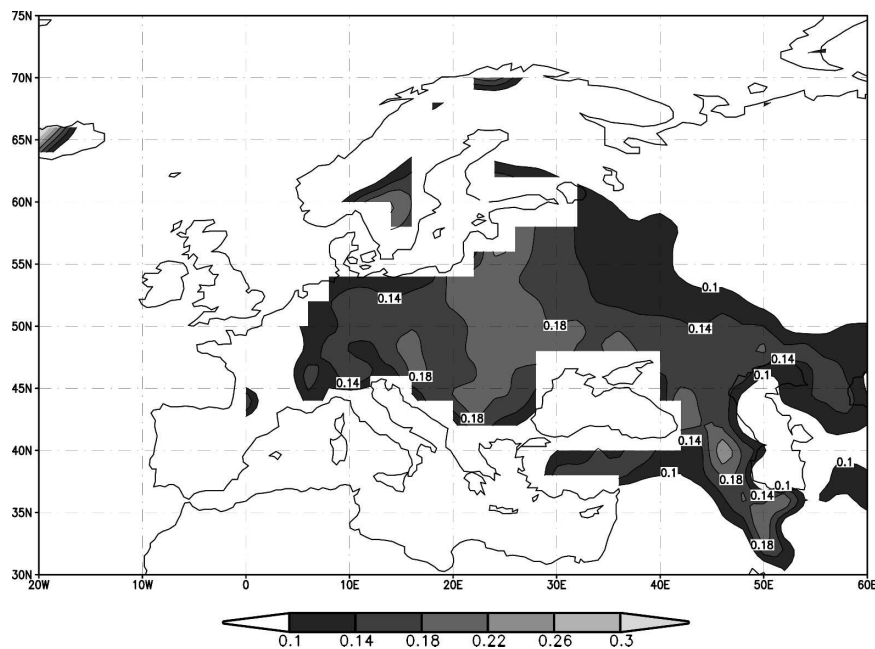

FIG. 7. FMA snow-cover variability in Europe expressed as standard deviations of the extent of snow on the ground.
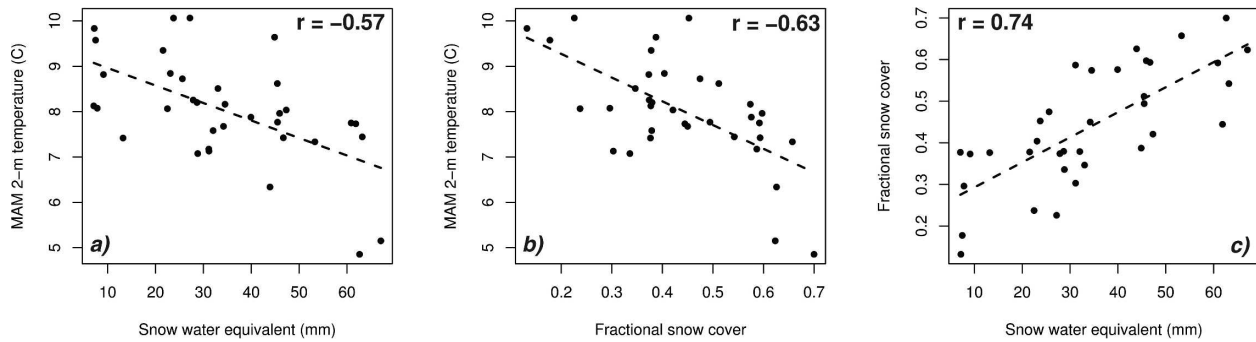
FIG. 8. Scatterplots of spatially averaged (i.e., averaged from 45°–55°N to 20°–30°E) (a) MAM 2-m temperature (°C) vs 1 February SWE (mm), (b) MAM 2-m temperature (°C) vs FMA fractional snow cover, and (c) FMA fractional snow cover vs 1 February SWE (mm).

SWE. The plot shows that very cold temperatures occurred often following high SWE at the beginning of February, and vice versa for the warm extreme. The correlation is −0.57 (statistical significance at the 1% level). The plot for MAM 2-m temperature against FMA snow-cover extent (Fig. 8b) is consistent with the results from SWE (obtained from ERA-40), despite the fact that the snow datasets are obtained from independent sources. The plot shows that the coldest MAMs over eastern Europe have occurred following extensive snow cover from February to April (correlation = −0.63). The plots make physical sense. Snow influences near-surface temperatures through its high reflectivity to incident radiation, low thermal conductivity, which inhibits sensible heat flux from the ground into the overlying air, and by acting as a latent heat sink during the melting process. Even after melting, the resulting soil moisture influence near-surface air temperature through its effect on the surface heat balance in the form of alterations in partition of latent and sensible heat fluxes. These results are in qualitative agreement with Groisman et al. (1994), who classified eastern Europe as a temperature-sensitive zone from December to March. The scatterplot for FMA fractional snow extent against February SWE is presented in Fig. 8c. The strong correlation of 0.74 indicates that a thick snowpack on 1 February on average translates into longer lingering snow in spring, despite the well-known observations that a snowpack can be melted in surprisingly short times by, for example, heavy precipitation at high temperatures.

2) CCA DIAGNOSTICS

To the extent that there exists some linear relationship between snow cover and 2-m temperature as suggested by the scatterplots, the use of a linear statistical technique such as CCA to diagnose the spatial patterns

of covariability between the two fields is justified. Cross-validated sensitivity tests suggested that a subspace consisting of 11 observed snow-cover EOFs (70% cumulative variance) and six 2-m temperature EOFs (72% cumulative variance) could optimally be used in conjuction with four CCA modes (successive pairs of canonical variates). The spatial patterns for the first canonical mode (whose canonical correlation is highest) and the corresponding time series are shown in Fig. 9. In Fig. 9a, the first canonical eigenvectors of the snow-cover field leading spring 2-m temperature by 1 month are plotted. The corresponding canonical eigenvectors for spring 2-m temperature are presented in Fig. 9b. The magnitude of the loadings (elements of the canonical eigenvectors) indicate the relative contribution of each grid to the corresponding canonical pattern. The time series for this pattern is shown in Fig. 9c.

The first canonical mode represents a pattern of covariability that is located in eastern Europe and western Russia. The snow cover and 2-m temperature patterns for this mode explain about 15% and 24% of the total variance of each field, respectively. This mode shows that years with excessive (deficient) snow cover over eastern Europe from February to March are those with anomalously cold (warm) spring 2-m temperature over the area stretching from about 15°E in eastern Europe, western Russia, and parts of Scandinavia. Indeed there is close correspondence between the geographical area of anomalous land surface forcing from snow (highest loadings in Fig. 9a) on the lowest atmosphere as suggested by this CCA mode and the area exhibiting highest FMA snow-cover interannual variability (Fig. 7). In addition to providing some empirical evidence in support of snow-radiation (e.g., increasing ground reflectance) and snow-hydrological (e.g., energy used for melting snow and evaporating the resulting water) effects on 2-m temperature, the maps also show the spa-
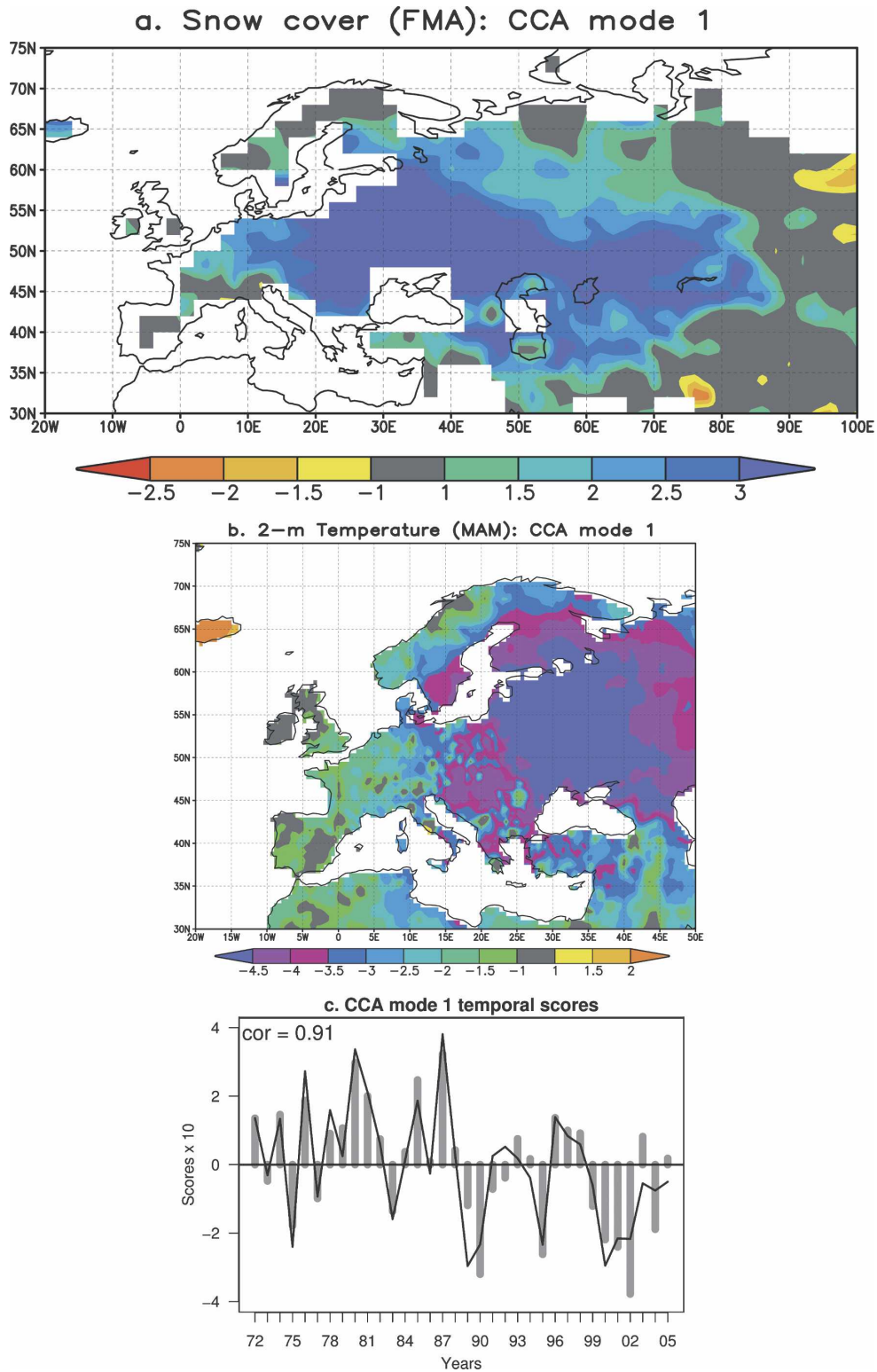
FIG. 9. CCA diagnostics: (a) FMA snow-cover CCA loadings for mode 1, (b) the corresponding CCA loadings for MAM 2-m temperature, and (c) the canonical component time series for snow (gray bars) and 2-m temperature (line).

tial extent of the snow influence. The results show that the impact of anomalous snow-cover forcing could occur through local feedback mechanisms and possible changes in atmospheric circulation. The effect of snow–albedo–temperature feedbacks is more localized (e.g., Dewey 1977), which explains the eastern Europe case. The impact of snow on atmospheric circulation can extend much deeper into the middle or even the upper troposphere (e.g., Zhang et al. 2004) thus exerting a remote influence on the weather/short-term climate.

The canonical component time series for this mode whose correlation is 0.91 show that this pattern of anomalously extensive snow cover and the associated colder 2-m temperatures in eastern Europe occurred more frequently from the 1970s to the late 1980s (Fig. 9c). The frequency declined since the 1990s. Possible mechanisms forcing this pattern of temporal variation have not been investigated in this paper. A statistically significant reduction in western Eurasia (40°–60°N, 20°–90°E) spring (particularly April) snow-cover extent (SCE) has been shown elsewhere (Brown 2000) using reconstructed snow-cover data. Superimposed to the significant downward SCE trend shown in Brown (2000) are decadal–multidecadal fluctuations. Consistent with our findings, the decade from the late 1970s to the late 1980s fell above the trend line (Brown's Fig. 14) with a more rapid decline in the early 1990s. A link between snow cover and concurrent NAO cannot be ruled out. The correlation between the snow-cover canonical time series for this mode and the FMA-averaged NAO–Gibraltar index (Jones et al. 1997) is −0.34 (statistical significance at the 5% level).

b. Model snow predictions

1) FORECAST–OBSERVED SNOW CORRELATION

To determine the performance of the CGCMs in terms of snow extent predictions, the FMA model snow predictions are validated against NOAA/NESDIS snow data using correlation analysis. The FMA CGCMs snow predictions issued in February are considered, implying zero lead time. Snow extent is defined as the area covered by at least 3-mm thickness of SWE for the model data, which have been interpolated to a 2° × 2° latitude–longitude grid consistent with the validation dataset.

The spatial pattern of correlation between model FMA snow conditions and observed snow cover is shown in Fig. 10. At the 5% (1%) level, a correlation coefficient of 0.44 (0.54) is found to be statistically significant. Remarkably, in agreement with results from 2-m temperature forecast verification, statistically sig-
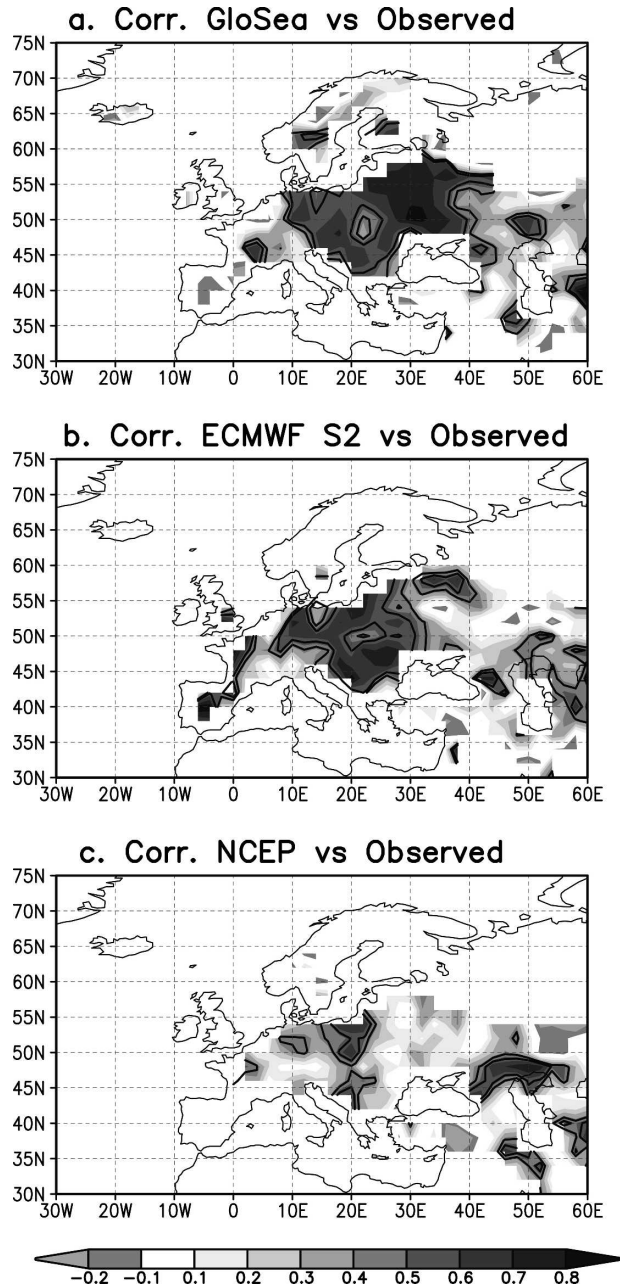

FIG. 10. Spatial pattern of correlation between modeled FMA snow cover and NOAA/NESDIS observations: the correlation for the (a) GloSea model, (b) ECMWF S2, and (c) NCEP CFS. The critical values for statistical significance at the 5% (1%) error level are 0.44 (0.54), and are shown by contours in the figure.

nificant correlations are found over much of central and eastern Europe in GloSea and ECMWF S2 systems (Figs. 10a,b). For the NCEP CFS system, statistically significant correlations are only confined to a limited area in eastern Europe and around the Caucasus (Fig. 10c).
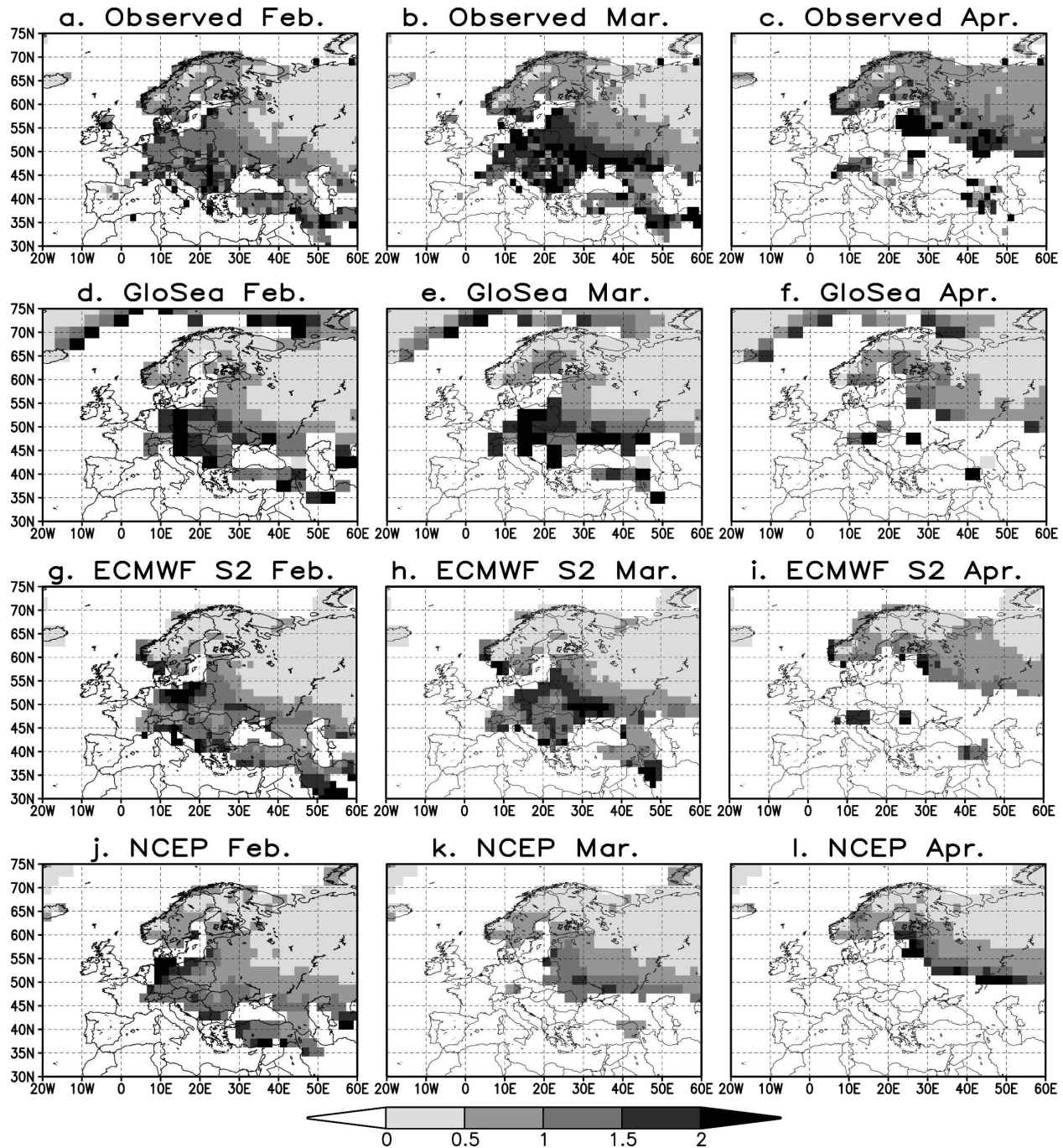
FIG. 11. Snow-depth-normalized standard deviation (coefficient of variation) over Europe from February to March: (a), (b), (c) the observed SWE variation; (d), (e), (f) GloSea; (g), (h), (i) ECMWF S2; and (j), (k), (l) NCEP CFS. Coefficients are shown for (a), (d), (g), (j) February; (b), (e), (h), (k) March; and (c), (f), (i), (l) April.

2) MODEL REPRESENTATION OF MONTHLY SNOW VARIATION

To assess the discrepancies that exist between modeled and observed SWE interannual variations, a nor-malized standard deviation is calculated for February to April based on the 1987–2005 period and shown in Fig. 11. Assessing the models' ability to capture the snow migration at monthly time scales reveal more informa-tion than would have been obtained from the seasonal

(FMA) patterns considered in foregoing sections. The dimensionless coefficient used to make the assessment is the ratio of the standard deviation to the monthly mean SWE. In the figure, each column shows the spatial distribution of the statistic in the same month. These plots allow a visual comparison of the ability of the models to capture the year-to-year monthly snow variations expressed as a proportion of average monthly snow depletion. The monthly snow line can be inferred from the plots.

In February, the observed snow line stretches from the Alps in southeastern France (about 5°E) southward into Bulgaria and northward through Germany toward Scandinavia (Fig. 11a). The year-to-year February SWE variation is highest in eastern Europe where the standard deviation either equals or exceeds the mean. All the models satisfactorily capture the broad pattern of February SWE variability. GloSea and ECMWF S2 also capture the mean SWE amounts adequately (map not shown). The NCEP CFS model, while locating the snow boundary close to the observed, underestimates the thickness of the snowpack over eastern Europe. Such an underestimate in snow depth would result in a subsequent underestimate of snow properties such as its albedo, which is one factor that influences the short-term climate, particularly near-surface temperatures.

By March the snow line in the observations does not retreat significantly with the westmost boundary still located over the French and Swiss Alps (Fig. 11b). Remarkably, year-to-year SWE in eastern Europe exhibits the highest variation during this month. In this snow-transient region, interannual standard deviations as high as twice the monthly mean are found. Indeed, the thickness and area extent of the snowpack during this month has a direct impact on spring near-surface temperature in the lowest quintile. The observed magnitude of March SWE variability in eastern Europe is adequately represented in the GloSea and ECMWF S2 models (Figs. 11e,h). This accuracy shows that in years when snow occurred in eastern Europe in March, GloSea and ECMWF S2 hindcasted/predicted the thickness of the snowpack successfully. As noted in (Slater et al. 2001), snow albedo changes in models' snow-albedo schemes as a function of its age and depth (up to a certain limit after which is becomes constant). In their discussion, most models attain a maximum albedo of about 85% for the visible, and about 65% for the near-infrared portions of the spectrum at SWE values less than 50 mm. Climate models that predict snow thickness at these "crucial" levels with some skill should commensurately represent the magnitude of the land surface forcing on the lower atmosphere reasonably well, thereby predicting cold spring seasons with

some degree of accuracy. As shown in Fig. 11k, the NCEP CFS system hardly predicts any snow in eastern Europe in March. In this model the snow boundary migrates too rapidly so it is situated much farther east (western Russia) than the climatological snow line. This serves as a clear illustration of differences in model snow simulation during the melt season. Discrepancies in model snow simulations during transition seasons have been noted elsewhere (e.g., Foster et al. 1996) and are in part attributable to differences in parameterization of snow processes in the models' land surface schemes (Slater et al. 2001).

Except for patches of snow over the Swiss Alps and the Caucasus, there is no observed or modeled snow over Europe in April (Figs. 11c,f,i,l). This is in response to the seasonal increase in available solar energy that forces a rapid migration of the climatological snow line. The snow patches in the Swiss Alps and the Caucasus are well predicted in the ECMWF S2 and GloSea models. In extreme cases (e.g., April 1996) some remnants of seasonal snow still survive during this month over much of eastern Europe and western Russia resulting in a delayed onset of seasonal warming. During this extreme case, the GloSea model accurately captured the delayed migration of the seasonal snow line (map not shown). The ECMWF S2 and NCEP CFS models on the other hand located the snow boundary much farther east along the western Russian border.

5. Discussion and conclusions

Predicting land temperatures more than 14 days ahead is a difficult task. One of the few extratropical areas with high skill scores in the current generation of operational seasonal forecast models is central and eastern Europe in spring. Forecasts started on 1 February show some skill as evidenced by ROC scores for the below-median MAM-averaged temperature exceeding 0.6. The skill for forecasting cold extremes is even higher (reaching 0.8) and statistically significant. The models' skill surpasses that of damped persistence forecasts.

Comparisons of skill measures for below-median and colder spring (2-m temperature in the lower quintile) for each CGCM suggests that there could be more predictive potential for cold extremes than "normal" seasons at the short lead time considered here. This is indeed the case over snow transient regions in central and eastern Europe. A notable difference in the skill of the CGCMs used in this research is evident. The GloSea and ECMWF S2 models attain the highest overall skill in predicting spring cold extremes during the verification period. The NCEP CFS model on the other hand performs the least well.

Noting that cold spring forecasts have higher skill than below-median seasons prompted a further investigation of the physical basis of the higher skill in predicting 2-m temperatures in the lowest quintile. The correspondence in geographical area between the high-skill area and the climatological snow line hinted at land surface processes as a possible source. Both remotely sensed and model-generated snow fields have been used to test our hypothesis. Both datasets inherently have their own biases. However, the close agreement between the two datasets over the study area suggested that they could reliably be used for the purpose of this paper.

Correlation analysis has shown a statistically significant relationship between modeled snow cover and the NOAA/NESDIS datasets over eastern Europe. The highest correlations (>0.5) were obtained for the GloSea system. Albeit of a lesser magnitude, correlations attained by the ECMWF S2 model were statistically significant over a large area. The NCEP CFS model had the lowest overall correlations. It is remarkable that the snow predictions from GloSea and ECMWF S2 are in closer agreement with observations and that the same models show higher skill in predicting cold spring seasons. This consistency provides evidence that the CGCMs' different ability to accurately model snow cover is a physical basis for the skill differences shown in the previous sections.

The abilities of the models to mimic the monthly migration of the climatological snow line and interannual variability of monthly SWE during the melt season further substantiates the snow-being-the-source hypothesis. Notable discrepancies in the model's ability to accurately locate the monthly climatological snow boundary and the year-to-year variations in monthly snow thickness have been shown. Again, the monthly snow line and the interannual variations of snow thickness in the GloSea and ECMWF S2 CGCMs models are very close to what is observed. In February, the NCEP CFS system predicts an almost realistic snow boundary but underestimates the thickness of the snowpack in eastern Europe. As snow albedo varies with snow depth (up to a certain limit) and age, underestimating the SWE in February would imply lower-than-observed albedo, and higher surface heat fluxes and hence warmer temperatures. In March, snow-depth variability is highest in eastern Europe. The NCEP CFS model misses the pattern of the interannual variability in March because the snow line migrates too rapidly in this model. In April all models faithfully represent the observed snow distribution.

The ability of GloSea and ECMWF S2 models to capture the observed pattern of interannual variability of February–April snow thickness suggests that these models adequately represent snow processes during the melt season. The NCEP CFS model on the other hand simulates a shorter snow season in eastern Europe, thus underestimating the amount of snow on the ground in February and March. The overall poorer performance shown by the NCEP CFS model in predicting cold springs in Europe could be attributable to the earlier melting of snow.

Patterns of covariability between FMA snow cover and MAM 2-m temperature have shown a link between snow cover in eastern Europe and western Russia and 2-m temperature, with temperature lagging a month behind. The patterns show that extensive snow cover from February to March has often preceded cold temperatures in spring. The underlying physics has been discussed extensively in the literature. Snow alters surface radiant energy fluxes through its high albedo and low thermal conductivity. Energy is extracted from the air during the melting process. Soil moisture from melted snow leads to a release of surface latent heat flux from the wet ground.

The initial snow conditions are important for cold spring season forecasts over Europe except over the western areas where temperatures are modulated by nearby oceans. High snow depth in eastern Europe and western Russia at the beginning of February (model initialization), persisting into early spring is closely related to cold springs in eastern Europe. Strong correlation shows that a thick snowpack early in February generally precedes lingering snow in spring. However, the dependence of snowpack duration on the ground on several factors complicates modeling of the snow fields during late winter and/or early spring. Over and above the seasonal increase in solar radiation, the frequency and strength of ablation events determines the snow durability. For instance, a large (or a sequence of) rainstorm(s) would provide large amounts of energy for melting the snowpack. Such rapid snowmelt would diminish the chances of an extremely cold spring. Therefore snow initial conditions would lead to skillful cold spring seasons' forecast if the CGCMs temperature and precipitation schemes were more realistic.

The skill of the Development of a European Multimodel Ensemble System for Seasonal-to-Interannual Prediction (DEMETER) versions of the ECMWF S2 and GloSea models (Palmer et al. 2004) in predicting lowest-quintile spring 2-m temperatures over 1958/59–2001 has also been assessed. The Met Office model attains higher ROC scores over a wide area in eastern Europe extending into Scandinavia (map not shown). No attempt has been made to investigate the strengths and weaknesses of the individual land parameterization

schemes used in the model. In a models' land surface schemes intercomparison study (Slater et al. 2001), it was noted that specific snow properties such as its albedo and thermal conductivity vary from model to model. Also the ablation rates were found to differ, which apparently is the case in the models used here as evidenced by the differing lengths of the snow season. It is clear that the predictive skill of spring cold spells in numerical models is tied to their accuracy in predicting snow fields.

REFERENCES

Anderson, D. L. T., and Coauthors, 2003: Comparison of the ECMWF seasonal forecast systems 1 and 2, including the relative performance for the 1997/8 El Niño. Tech. Memo. 404, ECMWF, Shinfield Park, Reading, United Kingdom, 93 pp.

Barnett, T. P., and R. Preisendorfer, 1987: Origins and levels of monthly and seasonal forecast skill for U.S. surface air temperatures determined by canonical correlation analysis. *Mon. Wea. Rev.,* **115,** 1825–1850.

Barnston, A. G., and T. M. Smith, 1996: Specification and prediction of global surface temperatures and precipitation from global SST using CCA. *J. Climate,* **9,** 2660–2697.

——, A. Kumar, L. Goddard, and M. P. Hoerling, 2005: Improving seasonal climate prediction practices through attribution of climate variability. *Bull. Amer. Meteor. Soc.,* **86,** 59–72.

Bretherton, C., C. Smith, and J. M. Wallace, 1992: An intercomparison of methods for finding coupled patterns in climate data. *J. Climate,* **5,** 541–560.

Brown, R. S., 2000: Nothern Hemisphere snow cover variability and change, 1915–97. *J. Climate,* **13,** 2339–2355.

Cohen, J., and D. Rind, 1991: The effect of snow on climate. *J. Climate,* **4,** 689–706.

Dewey, K. F., 1977: Daily maximum and minimum temperature forecasts and the influence of snow cover. *Mon. Wea. Rev.,* **105,** 1594–1597.

Fan, Y., and H. van den Dool, 2007: A global monthly land surface air temperature analysis for 1948–present. *J. Geophys. Res.,* in press.

Ferranti, L., and P. Viterbo, 2006: The European summer of 2003: Sensitivity to soil water initial conditions. *J. Climate,* **19,** 3659–3680.

Foster, J., and Coauthors, 1996: Snow cover and snow mass intercomparisons of general circulation models and remotely sensed datasets. *J. Climate,* **9,** 409–426.

Graham, R. J., M. Gordon, P. J. Mclean, S. Ineson, M. R. Huddleston, M. K. Davey, A. Brookshaw, and R. T. H. Barnes, 2005: A performance comparison of coupled and uncoupled versions of the Met Office seasonal prediction general circulation model. *Tellus,* **57A,** 320–339.

Groisman, P. Ya, T. R. Karl, and R. W. Knight, 1994: Changes of snow cover, temperature, and radiative heat balance over the Northern Hemisphere. *J. Climate,* **7,** 1633–1656.

Huang, J., H. M. van den Dool, and A. G. Barnston, 1996: Long-lead seasonal temperature prediction using optimal climate normals. *J. Climate,* **9,** 809–817.

Jackson, J. E., 1991: *A User's Guide to Principal Components.* Wiley-InterScience, 569 pp.

Jolliffe, I. T., 1972: Discarding variables in principal component analysis. II: Real data. *Appl. Stat.,* **21,** 160–173.

Jones, P. D., T. Jónsson, and D. Wheeler, 1997: Extension to the North Atlantic Oscillation using early instrumental pressure observations from Gibraltar and south-west Iceland. *Int. J. Climatol.,* **17,** 1433–1450.

Klein Tank, A., and Coauthors, 2002: Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment. *Int. J. Climatol.,* **22,** 1441–1453.

Kumar, A., and F. Yang, 2003: Comparative influence of snow and SST variability on extratropical climate in northern winter. *J. Climate,* **16,** 2248–2261.

Landman, W. A., and S. J. Mason, 2001: Forecasts of near-global sea surface temperatures using canonical correlation analysis. *J. Climate,* **14,** 3819–3833.

Lenderink, G., A. P. van Ulden, B. van den Hurk, and E. van Meijgaard, 2007: Summertime inter-annual temperature variability in an ensemble of regional model simulations: Analysis of the surface energy budget. *Climatic Change,* **81,** 233–247.

Mason, S. J., and N. E. Graham, 1999: Conditional probabilities, relative operating characteristics, and relative operating levels. *Wea. Forecasting,* **14,** 713–725.

——, and ——, 2002: Areas beneath the relative operating characteristics (ROC) and relative operating levels (ROL) curves: Statistical significance and intepretation. *Quart. J. Roy. Meteor. Soc.,* **128,** 2145–2166.

Palmer, T. N., and Coauthors, 2004: Development of a European Multimodel Ensemble System for Seasonal to Interannual Prediction (DEMETER). *Bull. Amer. Meteor. Soc.,* **85,** 853–872.

Robinson, D. A., K. F. Dewey, and R. R. J. Heim, 1993: Global snow cover monitoring: An update. *Bull. Amer. Meteor. Soc.,* **74,** 1689–1696.

Saha, S., and Coauthors, 2006: The NCEP climate forecast system. *J. Climate,* **19,** 3483–3517.

Shongwe, M. E., W. A. Landman, and S. J. Mason, 2006: Performance of recalibration systems for GCM forecasts for southern Africa. *Int. J. Climatol.,* **17,** 1567–1585.

Slater, A. G., and Coauthors, 2001: The representation of snow in land surface schemes: Results from PILPS 2(d). *J. Hydrometeor.,* **2,** 7–25.

Uppala, S. M., and Coauthors, 2005: The ERA-40 re-analysis. *Quart. J. Roy. Meteor. Soc.,* **131,** 2961–3012.

van den Dool, H. M., and J. L. Nap, 1985: Short and long range air temperature forecasts near an ocean. *Mon. Wea. Rev.,* **113,** 878–886.

van Oldenborgh, G. J., and A. van Ulden, 2003: On the relationship between global warming in the Netherlands and changes in circulation in the 20th century. *Int. J. Climatol.,* **23,** 1711–1724.

——, and G. Burgers, 2005: Searching for decadal variations in ENSO precipitation teleconnections. *Geophys. Res. Lett.,* **32,** L15701, doi:10.1029/2005GL023110.

——, M. A. Balmaseda, L. Ferranti, T. N. Stockdale, and D. L. T. Anderson, 2005a: Did the ECMWF seasonal forecast model outperform statistical ENSO forecast models over the last 15 years? *J. Climate,* **18,** 3240–3249.

——, ——, ——, ——, and ——, 2005b: Evaluation of atmospheric fields from the ECMWF seasonal forecasts over a 15-year period. *J. Climate,* **18,** 3250–3269.

von Storch, H., and F. W. Zwiers, 1999: *Statistical Analysis in Climate Research.* Cambridge University Press, 484 pp.

Walsh, J. E., D. R. Tucek, and M. R. Petersonvan, 1982: Seasonal snow cover and short-term climatic fluctuations over the United States. *Mon. Wea. Rev.,* **110,** 1474–1486.

——, W. H. Jasperson, and B. Ross, 1985: Influences of snow cover and soil moisture on monthly air temperature. *Mon. Wea. Rev.,* **113,** 756–768.

——, A. S. Phillips, D. H. Ports, and W. L. Chapman, 2001: Extreme cold outbreaks in the United States and Europe, 1948–99. *J. Climate,* **14,** 2642–2658.

White, P. W., 2003: IFS documentation cycle. Part II: Data assimiliation. Tech. Rep. CY23r4, ECMWF, Reading, United Kingdom, 126 pp.

Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Sciences: An Introduction.* Academic Press, 464 pp.

Yang, F., A. Kumar, W. Wang, H.-M. H. Juang, and M. Kanamitsu, 2001: Snow-albedo feedback and seasonal climate variability over North America. *J. Climate,* **14,** 4245–4248.

Zhang, Y., T. Li, and B. Wang, 2004: Decadal change of the spring snow depth over the Tibetan Plateau: The associated circulation and influence on the East Asian summer monsoon. *J. Climate,* **17,** 2780–2793.