Application and Validation of a Seasonal Ensemble Prediction System Using a Dynamic Malaria Model

ANNE E. JONES AND ANDREW P. MORSE

School of Environmental Science, University of Liverpool, Liverpool, United Kingdom

(Manuscript received 16 April 2009, in final form 29 January 2010)

ABSTRACT

Seasonal multimodel forecasts from the Development of a European Multimodel Ensemble System for Seasonal-to-Interannual Prediction (DEMETER) project are used to drive a malaria model and create reforecasts of malaria incidence for Botswana, in southern Africa, in a unique integration of a fully dynamic, process-based malaria model with an ensemble forecasting system. The forecasts are verified against a 20-yr malaria index and compared against reference simulations obtained by driving the malaria model with data from the 40-yr European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-40). Performance assessment reveals skill in the DEMETER-driven malaria forecasts for prediction of low (below the lower tercile), above-average (above the median), and high (above the upper tercile) malaria events, with the best results obtained for low malaria events [relative operating characteristics (ROC) area = 0.84, 95% confidence interval = 0.63–1.0]. For high malaria events, the DEMETER-driven malaria forecasts are skillful, but the forecasting system performs poorly for those years that it predicts the highest probabilities of a high malaria event. Potential economic value analysis demonstrates the potential value for the DEMETER-driven malaria forecasts over a wide range of user cost-loss ratios, which is primarily due to the ability of the system to save on the cost of action in low malaria years.

1. Introduction

Recent studies have shown that state-of-the-art multimodel ensemble prediction systems can make skillful seasonal forecasts of a number of variables, including the El Niño-Southern Oscillation (ENSO) sea surface temperature (SST) anomalies in the tropical Pacific (Palmer et al. 2004; Krishnamurti et al. 2006) known to influence climate variability in regions of Africa (Nicholson and Kim 1997; Camberlin et al. 2001). One objective of the Development of a European Multimodel Ensemble System for Seasonal-to-Interannual Prediction (DEMETER) project (Palmer et al. 2004; available online at http://www.ecmwf.int/research/demeter), continuing on in its successor ENSEMBLES (available online at http://ensembles-eu.metoffice.com), has been to assess the utility of such forecasts for prediction of a range of climate impacts, from optimal energy production to human disease control. This paper presents the final results of a study where a dynamic, process-based model of malaria transmission (Hoshen and Morse 2004) was used to carry out validation of the DEMETER multimodel reforecasts for seasonal prediction of malaria in Botswana, demonstrating the first integration of a fully dynamic, process-based model malaria model with an ensemble prediction system (EPS).

It has been proposed that the best strategy for the forecasting of climate impacts is to take a closely integrated modeling approach, whereby users can learn to take optimal advantage of skill available in the climate forecasts and climate models can begin to be improved in the areas most useful for impact modeling (Buizer et al. 2000). A key component of this approach is the assessment of climate forecast quality by user-oriented verification strategies (Doblas-Reyes et al. 2006). Previous authors have described a three-tiered approach to forecast verification for impacts (Morse et al. 2005), of which the first and simplest is ''tier 1,'' where the relevant climate variables are verified against observations of those variables. For complex impacts—where processes may be nonlinear—better user-oriented performance assessment can be achieved by driving an impact model with seasonal climate forecasts and verifying the

Corresponding author address: Anne Jones, School of Environmental Science, Roxby Building, University of Liverpool, Liverpool, L69 7ZT, United Kingdom.
E-mail: anne.jones@liverpool.ac.uk

DOI: 10.1175/2010JCLI3208.1

resulting forecasts of the impact variable. The reference for verification may be either simulations obtained by driving the impact model with observations (tier 2) or observations of the impact itself (tier 3). This paper presents an example of performance assessment at all three tiers. Verification of an integrated forecasting system has been carried out for prediction of malaria in Botswana, in southern Africa. Seasonal reforecasts of malaria have been compared against reference simulations obtained by driving the malaria model with data from the the 40-yr European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-40; Uppala et al. 2005), and both simulations and reforecasts have been verified against a published index of the observed countrywide anomalies of the disease (Thomson et al. 2005).

Malaria is caused in humans by infection with one of four protozoan species belonging to the genus *Plasmodium,* which is transmitted between humans by the *Anopheles spp.* vector. The disease results in over one million deaths annually, with over 80% of these fatalities occurring in sub-Saharan Africa (World Health Organization 2005). Epidemics of malaria can be triggered by factors affecting human, vector, or parasite populations including abnormal meteorological conditions (such as increased rainfall and/or temperatures), changes in antimalarial programs, population movement, and environmental changes (Nájera et al. 1998). Although malaria is preventable and treatable, epidemics in sub-Saharan Africa are often detected too late for effective interventions to be implemented (World Health Organization 2001), and malaria remains the biggest cause of deaths for children under five (Lee 2006). Part of the Roll Back Malaria (RBM) Partnership global strategic plan for 2005–15 includes the establishment and maintenance of early warning systems for epidemics, to meet the target of 60% of malaria epidemics detected within two weeks and responded to within two weeks of detection (Roll Back Malaria Partnership 2005). By employing models of the relationships between malaria and climate and driving such models by observations, predictive lead times of climate-related epidemic risk forecasts can theoretically be increased by several weeks. The availability of seasonal climate forecasts provides the potential to extend this to several months.

Climate affects malaria via life cycle controls on both the Plasmodium spp. parasite and the Anopheles spp. mosquito vector. Because *An. gambiae*, the mosquito complex that includes the most important vectors of malaria in Africa (Service and Townson 2002), lay their eggs in standing water, sufficient rainfall is a major factor influencing malaria transmission. A second factor is temperature, which affects the development and survival of larval and adult mosquitoes (Lindsay and Martens 1998; Bayoh and Lindsay 2003, 2004). Temperature also impacts disease transmission through an increase in biting frequency with temperature (Githeko et al. 2000), and in the rate at which the parasite multiplies within the mosquito (Detinova et al. 1962). Humidity is a further climate factor that has been linked with malaria through its influence on the mosquito biting–laying cycle and mortality (Detinova et al. 1962; Molineaux 1988).

Previous work on the application of seasonal forecasts in Africa has focused primarily on agriculture, with impact model forecast assessment carried out for South Africa (Martin et al. 2000; Bezuidenhout and Schulze 2006), Kenya (Hansen and Indeje 2004), and Zimbabwe (Cane et al. 1994; Philips et al. 1998; Martin et al. 2000). The potential for seasonal hydrological forecasting has also been explored (Landman et al. 2001; Eldaw et al. 2003). Previous verification of seasonal forecasts using health application models in Africa has been limited to malaria; Thomson et al. (2006) used a statistical-empirical model (Thomson et al. 2005), linking rainfall and malaria in Botswana with DEMETER multimodel forecasts, to skillfully predict anomalies in malaria for the period 1982–2002. The present study continues the work of Morse et al. (2005), who carried out preliminary tier-2 verification of DEMETER-driven malaria model forecasts for four grid points in southern Africa, using a dynamic, process-based malaria model (Hoshen and Morse 2004) formulated to represent the climate-driven biological mechanisms associated with transmission of the disease.

Previous authors have discussed reasons for preferring process-based models over statistical-empirical models in the context of assessing the impact of climate change on vector-borne disease; namely, the insight such models can give us into the relationship between climate and biological processes and the potential for quantitative assessment of the effect of possible human interventions (Rogers and Randolph 2006) such as indoor residual spraying (for which models have already been developed, e.g., Worrall et al. 2007). Process-based models are, however, limited by the understanding of the biological mechanisms involved and by the availability of data for model validation. Furthermore, the practicalities of integrating such a model with the output from a dynamic multimodel EPS are not straightforward, and the performance of forecasts obtained by driving a process-based disease model with climate model output will depend not only on the skill of the driving forecasts but also the sensitivity of the disease model to particular characteristics of the driving data.

This study aims to use a dynamic–dynamic model integration to answer a number of questions concerning both the performance of the forecasting system and the integration of impact models with seasonal climate forecasts.

Can the DEMETER forecasts be used with a complex, dynamical, process-based impact model to produce skillful seasonal reforecasts of malaria in Botswana? How should users of complex impact models such as these best make use of the seasonal climate forecasts? What information about model performance can we give as feedback to the climate forecast providers? A brief overview of the malaria model is given in section 2. In section 3, the study region is described. The reanalysis and forecasts datasets are described in section 4, along with the preprocessing techniques used to attempt to correct for seasonal forecast bias. In section 5, the procedure used to create multimodel ensemble forecasts of malaria is outlined, along with the techniques employed to assess forecast skill. The results are presented in section 6, and issues highlighted by the study, along with its limitations, are discussed in section 7.

2. Description of the malaria model

The Liverpool malaria model (LMM) is a process-based dynamic model of malaria, consisting of two climate-driven components: a malaria transmission model, derived from the theoretical approach to malaria modeling described by Anderson and May (1991), and a dynamic mosquito population growth model. The malaria life cycle, along with details of the malaria model, have been described elsewhere (Hoshen and Morse 2004; Morse et al. 2005), as have the dynamics of malaria transmission by Anopheles mosquitoes (Smith and Ellis McKenzie 2004). Here we will focus on a description of how the two climate drivers, daily temperature and rainfall, are incorporated into the model.

Temperature drives both the development rate of the malaria parasite within the mosquito (the sporogonic cycle) and the biting–laying cycle of the mosquito itself (the gonotrophic cycle, governed by the rate at which eggs can be produced). The sporogonic development rate is arguably the most important of these constraints for rainy season temperatures typically found in Botswana, and in LMM this rate is taken to be linearly proportional to the number of "degree-days" above a threshold temperature. Parameters are taken from the literature (Detinova et al. 1962) such that the sporogonic cycle takes 111 degree-days above a threshold of 18°C. Below this temperature no development takes place. A similar relationship governs the rate of progression of the gonotrophic cycle, with a total cycle length in humid conditions of 37 degree-days above a threshold of 7.7°C. Adult mosquito mortality is also dependent on temperature, and LMM uses a quadratic form for the survival probability function, based on mosquito survival probabilities of 0.82, 0.90, and 0.04 at temperatures of 9°, 20°, and 40°C, respectively, as reported by Martens et al. (1995). Rainfall is the source of growth of the modeled mosquito population via the availability of breeding sites, modeled by setting the number of eggs laid per mosquito per gonotrophic cycle to be linearly proportional to the previous 10 days (dekadal) rainfall. An upper limit is applied to the number of laying mosquitoes to simulate competition for resources. The survival of mosquito larvae also depends on rainfall, with the daily larval survival probability varying between 0.5 at zero rainfall and 1.0 for high rainfall, simulating pressure on the larval population if breeding sites begin to dry out. Unlike the mosquito population, development of the disease in the human host population is not directly climate driven, and infected humans remain in a latent incubation period of 15 days before becoming able to transmit malaria gametocytes to mosquitoes. The human hosts have a natural clear-up rate of infection of 3% of the population per day, corresponding to 90% clear up after 80 days, and the model has no human mortality. LMM does not model immunity, and as a result is only applicable in epidemic areas where climate conditions are marginally suitable for transmission, resulting in high interannual variability and no opportunity for the host population to acquire immunity.

3. Study region

Malaria is ranked as a major public health problem in Botswana (Thomson et al. 2005) despite a national malaria control program dating back to the 1970s. Since the 1980s malaria epidemics in the countries of southern Africa have become more frequent and severe due to a combination of factors including environmental change, drug resistance, and social issues affecting the efficacy of control measures (Mabaso et al. 2004). For example, in 1996, widespread and severe epidemics affected the whole region (le Sueur et al. 1996). Rainfall variability has been found to be a strong driver of malaria variability in Botswana; Thomson et al. (2005) used records of malaria from two sources along with information about intervention policy changes to derive a standardized malaria index for the whole country for 1982–2001. They also found that a quadratic relationship existed between countrywide rainfall anomalies derived from a gridded rainfall dataset and their malaria index.

The potential for seasonal prediction of rainfall in Africa arises from the association between interannual variability of rainfall over the continent and large-scale SST variability in tropical regions. Central southern African rainfall is mostly strongly linked to the ENSO, with positive ENSO years usually being associated with drier conditions during the months November–July (Camberlin

et al. 2001), although correlations have also been demonstrated between an east–west dipole in the Indian Ocean and above normal rainfall in the region (Behera and Yamagata 2001). The use of observed SSTs in seasonal forecasting restricts lead times to 1–2 months; longer leads times can be obtained by predicting ENSO itself using coupled atmospheric oceanic general circulation models (AOGCMs). Seasonal multimodel ensemble forecasts from AOGCMs have been shown to be capable of predicting the ENSO signal as skillfully as dedicated ENSO models and much better than atmospheric general circulation models driven by persisted SSTs (Palmer et al. 2004). Seasonal forecasts have been used by the health community in southern Africa for a number of years (DaSilva et al. 2004), and more recently a malaria early warning system (MEWS), developed by the Roll Back Malaria partnership has been prototyped in southern Africa and tested extensively in Botswana (Connor et al. 2007). This system includes, along with vulnerability assessment, environmental monitoring and case surveillance, and the use of tailored seasonal climate forecasts to plan, prepare, and respond to epidemics.

4. Data

The seasonal forecasts assessed in this study are the product of the European Union DEMETER project (Palmer et al. 2004). Here, the term ''reforecasts'' refers to forecasts for past time periods, and the DEMETER seasonal reforecasts consist of output from a multimodel ensemble of seven different coupled AOGCMs, each run with nine different sets of initial conditions. (Three of the DEMETER models have subsequently been used in the EUROSIP project to produce operational multimodel seasonal forecasts, available online at http://www.ecmwf.int/products/forecasts/seasonal/documentation/eurosip/.) DEMETER forecasts are available for four start dates each year: February, May, August, and November, each extending to 180 days from the forecast origin. Of the seven models, three are available for the period 1958–2001 and all seven for 1980–2001. Long and reliable daily time series of meteorological station data are difficult to obtain for many countries in Africa. Here, reference data were used in the form of the ERA-40 reanalysis dataset (Uppala et al. 2005), which consists of 44 yr of in situ and remotely sensed data assimilated to form a set of global gridded analyses from 1958 to 2001.

ERA-40 and DEMETER data for daily accumulated precipitation and 2-m temperature were extracted from the Meteorological Archival and Retrieval System (MARS; available online at http://www.ecmwf.int/publications/manuals/mars) database at ECMWF. Experience during the DEMETER project resulted in the choice of the maximum of the four 6-hourly temperatures issued each day, minus 5°C, that is used as a proxy for the daily mean temperature experienced by the mosquito (Morse et al. 2005). The data region covering Botswana consisted of 25 grid points of a 5×5 grid at 2.5° resolution, from 20° to 30°E and 17.5° to 27.5°S. ERA-40 reanalysis along with multimodel forecasts from all seven DEMETER models were extracted for the years 1982–2001, with the November DEMETER forecast origin selected to correspond to the timing of the start of the rainy season in Botswana. Observed malaria data for Botswana were obtained in the form of the 20-yr time series of standardized malaria anomalies previously published by Thomson et al. (2005).

The suitability of ERA-40 rainfall as a reference dataset was first assessed (for Botswana) by comparing the seasonal cycle and interannual variability of seasonal totals averaged over the 25 grid points, with those derived from the monthly Climate Prediction Center (CPC) Merged Analysis of Precipitation (CMAP; Xie and Arkin 1997), which combines gauge and satellite estimates of rainfall. Although ERA-40 tends to overestimate rainfall in Botswana compared to CMAP, the seasonal cycles are in broad agreement, and interannual variability is highly correlated [Pearson correlation coefficient $r(20) = 0.90$, $p < 0.01$ for 1982–2001]. Comparison of daily ERA-40 rainfall with a limited amount of daily rain gauge data for Botswana (not shown) also revealed ERA-40 possesses rainfall frequency and intensity characteristics similar to observations. ERA-40 was therefore deemed a suitable reference, or pseudo-observational dataset, for the remainder of the analysis, and used to provide both a target for climate model bias correction (see below) and measurement of potential predictability of the combined climate–malaria modeling system (see section 5a).

One component of error in climate models is a systematic difference between the long-term climatology of the model and reality (Lazar et al. 2005). As a result of this inherent bias, the mean state of the seasonal forecasts tends to drift toward the individual GCM climatology as the lead time of the forecast increases, resulting in discrepancies between long-term averages of observed and forecast data. The magnitude of the drift is model dependent and therefore must be corrected before multiple model outputs are combined in a multimodel ensemble forecast. The usual approach with climate variables is to consider monthly anomalies (Stockdale 1997), but when impact models require realistic daily time series as input, and combine multiple variables in a nonlinear manner, correction of biases can be difficult. Here, bias correction was carried out for the temperature forecasts from each of the seven DEMETER models by applying a low-pass filter to the long-term daily averages

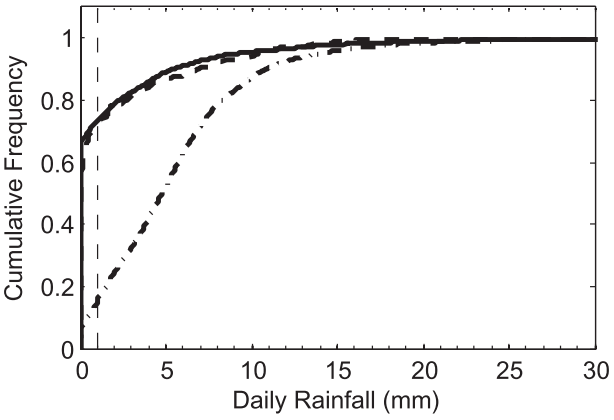

FIG. 1. Example comparison of ERA-40 and DEMETER rainfall distributions. ERA-40 daily rainfall (solid line), DEMETER Météo France model raw daily rainfall (dot–dashed line), and same DEMETER data corrected to match ERA-40 frequency and intensity (dashed line). Rainfall for November only, for grid point 20°S, 22.5°E. The vertical dashed line indicates the position of daily rainfall = 1 mm, the threshold used to discern rainfall–nonrainfall days.

for each grid point, thereby obtaining a set of seven smoothed model-specific seasonal cycles. The same procedure was carried out for the ERA-40 data, and the differences between the ERA-40 cycle and each DEMETER model cycle were applied as offsets to the DEMETER forecasts for the corresponding model. For rainfall, bias correction was carried out using a frequency–intensity method similar to the ''local scaling'' approach of previous authors (Ines and Hansen 2006; Schmidli et al. 2006; Moron et al. 2008), where the frequency of rain days and the rainfall intensity on rain days are corrected separately. In this case, ERA-40 rainfall was used as a reference for the correction of DEMETER rainfall (Fig. 1). The correction procedure is applied as follows. First, a rainfall offset is calculated so that when the offset is applied to the model rainfall, the mean frequency of rain days (days with rainfall >1 mm) matches that of the reference rainfall from ERA-40. For example, in Fig. 1, an offset of approximately -6.5 mm is applied to the DEMETER rainfall to shift the distribution left toward that of ERA-40. (Note this method can only lead to reduced rainfall frequency because there is no way to generate new rain days.) In the second part of the procedure, a rainfall intensity scaling factor is computed from the ratio of the mean rainfall intensity of reference rain days to the mean rainfall intensity on the model rain days. The offset and scaling factors are then applied to the model forecasts to obtain the bias-corrected rainfall. Rainfall correction was carried out separately for each DEMETER model, using a moving 30-day window to calculate the seasonally varying parameters in each case. For both the bias correction of temperature and rainfall, a ''leave one out'' method was employed, in which the long-term averages were computed using all years excluding the year to be corrected.

5. Methods

a. Production of malaria reforecasts

Malaria reforecasts were created by driving the LMM with temperature and rainfall from each DEMETER ensemble member, resulting in a 63-member ensemble reforecast of malaria transmission, starting in November and extending out to 180 days for each of the 20 yr from 1982 to 2001. Each LMM run was carried out for a single grid point. In each forecast instance, the malaria model was initialized by running with ERA-40 data for a ''spinup'' period of the 12 months prior to the DEMETER forecast origin. To investigate the impact of bias correction on the skill of the malaria forecasts, three combinations of input data were used to drive the model: 1) uncorrected rainfall and uncorrected temperature, 2) uncorrected rainfall and corrected temperature, and 3) corrected rainfall and corrected temperature.

In addition to the DEMETER-driven model runs, two other types of LMM forecasts, both starting from the same malaria model state at the forecast origin, were created to provide a baseline against which to measure DEMETER skill. The first type consisted of ''potential predictability'' or ''reference'' runs, obtained by driving the malaria model with the correct ERA-40 data for both spinup and the duration of the forecast period. The second consisted of ''best-guess'' runs (referred to henceforth as control runs), obtained using a similar methodology to ensemble streamflow prediction (ESP) used in operational hydrology to predict likely future flows (Carpenter and Georgakakos 2004; Wood et al. 2005), and analogous to the ''persistence'' forecasts often used as a baseline in seasonal forecasting—where a model is driven by observed conditions up to the forecast origin and then constant conditions such as persisted SST anomalies for the forecast period (e.g., Graham et al. 2005). Here, the malaria model was driven with the correct spinup data for a given year, followed by data for the forecast period taken in turn from ERA-40 for each of the other 19 (wrong) yr. The resulting set of forecasts was treated as an ensemble. This second type of baseline forecast would allow assessment of the impact of spinup data on forecast skill.

Modeled malaria transmission was averaged over the 25 grid points to give a set of areal average malaria reforecasts for Botswana, for three different integration periods: months 2–4 [December–February (DJF)] to coincide with the start and the peak of the rainfall season, 4–6 [February–April (FMA)] to coincide with the peak

and the end of the rainfall season, and finally the entire 6-month forecast period (November–April).

Forecast performance assessment was carried out using a number of standard techniques applied to the prediction of events, defined by three thresholds in the distribution of the observed or forecast variables: the lower tercile, the median, and the upper tercile. The 20 yr of malaria anomalies published by Thomson et al. (2005) were used to create three time series of binary events: low malaria years (observed malaria anomalies below the lower tercile), above-average malaria years (observed malaria anomalies above the median), and high malaria years (observed malaria anomalies above the upper tercile). The corresponding thresholds were also calculated for the distribution of DEMETER-driven malaria forecasts (separately for each DEMETER model) and for the ERA-40-driven runs. For the DEMETER forecasts, the probability of low, above-average, and high malaria events were calculated as the proportion of the 63 ensemble member forecast values falling into the appropriate category. For the ERA-40 reference runs, the simulated malaria totals were used as a direct indicator of malaria risk (Mason 2003), thereby avoiding any penalization that could arise by reducing deterministic forecasts to binary forecasts.

b. *Performance assessment*

Forecast performance was assessed by calculating relative operating characteristics (ROC) areas (Mason 2003) for each event and forecast integration window. The area under the ROC curve gives the probability that, given an event and a nonevent, the forecasting system will correctly distinguish the two (Mason and Graham 2002). A ROC area of 1.0 indicates the forecasting system can perfectly distinguish events and nonevents, whereas an area of 0.5 indicates a forecast no better than climatology. Here, the ROC areas for the DEMETER-driven LMM forecasts were calculated for each category of event relative to both the Thomson et al. (2005) malaria index (tier-3 verification) and the ERA-40 driven simulations (tier-2 verification). Tier-1 skill (the performance of the driving variables with respect to ERA-40, independent of the malaria model) was also determined by calculating the ROC areas for low, above-average, and high categories for integration period averages of rainfall and degree-days above 18°C, the latter reflecting the critical minimum temperature for development of the sporogonic cycle in malaria. To give an estimate of the uncertainty due to the sample size of 20 yr, 95% confidence intervals were calculated for all the ROC areas using bootstrap resampling (Efron 1981). In addition, for the purposes of comparing forecast performance, confidence intervals for relative ROC areas were calculated, also using bootstrap resampling (in a similar manner to the estimation of uncertainty in the relative ignorance skill score by Clarke et al. 2004).

As a further step toward user-oriented verification, the potential economic value V of the forecasts as a function of the cost/loss ratio (C/L) was calculated for the DEMETER and ERA-40 reference forecasts. The principle behind this approach is the cost/loss model (Murphy 1977), which is a simple decision model that assumes decision makers can respond to a forecast event by taking action at a cost C. The potential loss if no action is taken and the event occurs is L. The economic value V (Richardson 2003) is a measure of the reduction in mean expense $E_{\text{clim}} - E_{\text{forecast}}$ obtained by using the forecasting system over climatology, relative to the reduction $E_{\text{clim}} - E_{\text{perfect}}$ that would have been obtained by using ''perfect'' forecasts:

$$V = \frac{E_{\text{clim}} - E_{\text{forecast}}}{E_{\text{clim}} - E_{\text{perfect}}}. \tag{1}$$

Within this framework, it is assumed that the strategy of ''climatology'' is to always act if the cost-loss ratio is less than the climatological probability s of the event in question, because it is cheaper to act every year than to take the losses associated with events occurring at that frequency. Conversely, the strategy is to never act if the cost-loss ratio is higher than the climatological probability of the event, because it is cheaper to take the losses associated with the events occurring at this frequency than pay to take action every year. The mean expense of climatology is then equal to either the cost of acting or the loss multiplied by the number of events, depending on the value of the cost-loss ratio:

$$E_{\text{clim}} = \min(C, sL). \tag{2}$$

A perfect forecast, on the other hand, will only incur an expense when an event occurs:

$$E_{\text{perfect}} = sC. \tag{3}$$

For probabilistic forecasts, the economic value represents the maximum value available to a user by selecting the optimum probability decision threshold for their cost/loss ratio.

6. Results

a. *Seasonal cycles*

Box and whisker plots of mean seasonal cycles over the period 1982–2001 (Fig. 2) reveal that the LMM-simulated malaria season for Botswana runs from approximately February to June, lagged by approximately

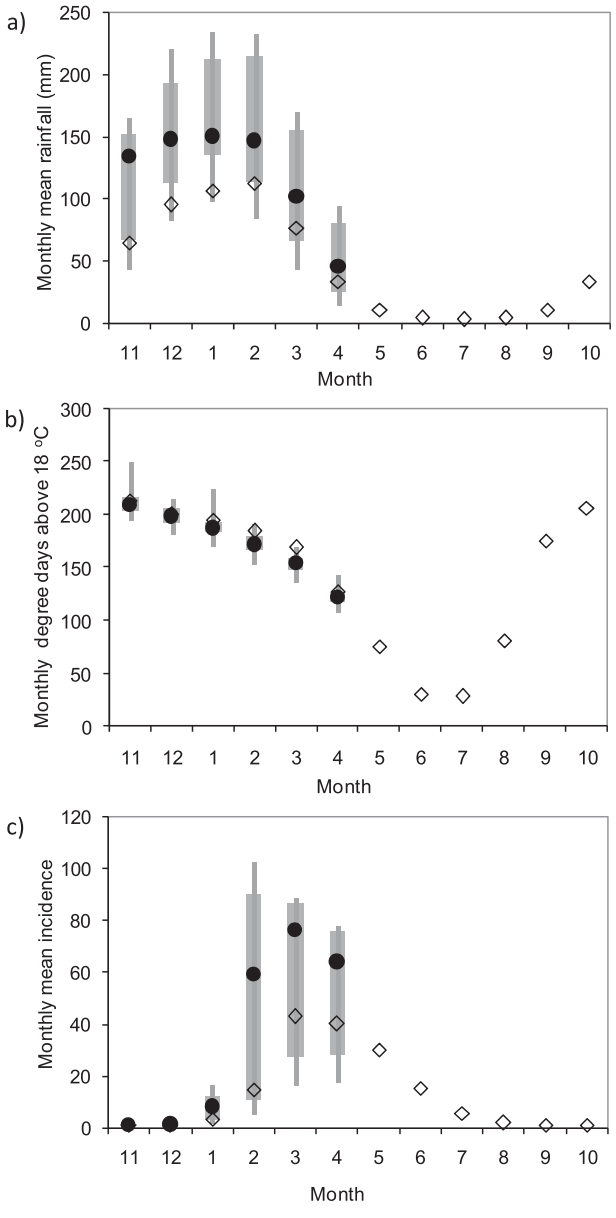

FIG. 2. Box plots showing seasonal cycle in malaria-relevant climate variables according to ERA-40 and DEMETER and seasonal cycle of malaria incidence as simulated by LMM. Long-term averages for 1982–2001 of ERA-40 (open diamonds), DEMETER median (filled circles), interquartile range (gray boxes), and full range (gray whiskers) (a) uncorrected rainfall, (b) bias-corrected degree-days above 18°C, and (c) LMM monthly mean malaria incidence per 100 people, obtained using uncorrected rainfall and corrected temperature.

2–3 months behind the ERA-40 rainfall season. The ERA-40-driven peak incidence occurs in March. The median of the DEMETER ensemble members also peaks in March, although some members peak as early as February. The uncorrected DEMETER rainfall (Fig. 2a) has a positive bias compared to ERA-40, and a positive bias is also seen in the malaria incidence obtained by driving the model with uncorrected rainfall and corrected temperature (Fig. 2c). The ERA-40 driven incidence falls within the interquartile range of the DEMETER ensemble members for all months. Initial inspection of the results for tier-3 performance assessment over different integration periods (not shown) revealed the highest skill was obtained by using malaria incidence during forecast months 4–6 (FMA), and so the tier-3 verification presented here will focus on this integration period.

b. Impact of bias correction on malaria forecast skill

A striking result of verification at tier 3 (Table 1) is that the most skillful forecasts are obtained by driving the malaria model with bias-corrected temperature but uncorrected rainfall. These forecasts are skillful compared to climatology (ROC area greater than 0.5) for low and above-average malaria (ROC area = 0.780 for above average and 0.841 for low) and significantly more skillful than climatology at 95% confidence (lower bound of the confidence interval >0.5). For high malaria, although the ROC area itself is skillful (0.670), this skill is not significant at 95% confidence [confidence interval (CI) = 0.412–0.929]. Forecasts made using raw DEMETER-driving data are skillful for the low category (but not significantly more so than climatology at 95% confidence), and unskillful for the high and above-average malaria categories. Those forecasts made using corrected temperature and corrected rainfall have skillful ROC areas for all three categories, but only the low malaria category is skillful at 95% confidence (ROC area = 0.769; 95% CI = 0.536–0.980). Bootstrapped confidence intervals for the relative performance of different forecasts (not tabulated) indicate that the improvements in skill obtained by bias correcting both temperature and rainfall (compared to the skill obtained with raw DEMETER drivers) are not significantly different from zero at 95% confidence (e.g., increase in ROC area = 0.242, 95% CI = −0.094–0.560 for the low malaria category. However, for forecasts made using corrected temperature and uncorrected rainfall the improvements in skill over the raw DEMETER-driven forecasts are significant for all categories (e.g., increase in ROC area = 0.314; CI = 0.027–0.616 for the low malaria category). The surprising result that bias correction of rainfall decreases malaria forecast skill will be further examined in section 7.

For the remainder of this section the discussion will focus on the results obtained by driving the model with uncorrected rainfall and corrected temperature. A box and whisker plot for this combination of driving data (Fig. 3) shows the observed malaria anomalies lie within the range of the DEMETER ensemble members for all years except 1988, and within the interquartile range of the DEMETER ensemble members for 8 out of the

TABLE 1. Tier-3 performance of DEMETER- and ERA-40-driven LMM malaria simulations. ROC areas for three categories of malaria events: low (below the lower tercile), above average (above the median), and high (above the upper tercile). November forecast months 4–6 (FMA): validation against observed malaria in Botswana for 1982–2001. LMM was driven by ERA-40 reanalysis for a spinup period of 12 months prior to the forecast start in November. The malaria model was subsequently driven by the five different forecast categories as listed in the table: ''ERA-40'' (ERA-40 reanalysis); ''ERA-40 control'' (an ensemble of the incorrect years of ERA-40 data); ''DEMETER raw'' (uncorrected DEMETER data); ''DEMETER bias corrected'' (bias-corrected temperature and bias-corrected rainfall); and ''DEMETER bias-corrected *T* only'' (corrected temperature and uncorrected rainfall). Range of values in brackets represent 95% confidence intervals calculated with 999 bootstrap samples. ROC areas above 0.5 indicate skill relative to climatology. Skillful ROC areas are highlighted in bold.

Event	ERA-40	ERA-40 control	DEMETER Raw	DEMETER Bias corrected	DEMETER Bias-corrected *T* only
Low	**0.714** (0.438–0.938)	0.412 (0.164–0.714)	**0.527** (0.202–0.843)	**0.769** (0.536–0.980)	**0.841** (0.627–1.0)
Above average	**0.820** (0.615–0.969)	0.305 (0.094–0.571)	0.445 (0.187–0.722)	**0.705** (0.458–0.919)	**0.780** (0.544–0.949)
High	**0.879** (0.640–1.0)	0.368 (0.117–0.654)	0.374 (0.141–0.620)	**0.648** (0.364–0.905)	**0.670** (0.412–0.929)

20 yr. The ROC diagrams further illustrating the results in Table 1 are given in Fig. 4.

c. *Comparison of DEMETER skill with skill of reference forecasts*

For all events, the ERA-40 control-run forecasts are unskillful (Table 1), demonstrating that the skill of the DEMETER-driven malaria forecasts must be due to the DEMETER forecasts, and not due to the ERA-40 data used to spinup the model prior to the forecast origin. Considering the malaria forecasts obtained using reanalysis data to drive the model, as expected, the ERA-40-driven simulations for above-average and high malaria have skill higher than that of any of the DEMETER forecasts for the same categories. For low malaria, the bias-corrected DEMETER-driven forecasts have higher skill than those driven by ERA-40 (0.769 and 0.841 for DEMETER; 0.714 for ERA-40). However, the bootstrapped confidence intervals for the difference in ROC areas (not tabulated) indicate that none of the differences between ERA-40- and DEMETER-corrected forecasts are significantly different from zero at 95% confidence (e.g., increase in ROC area from ERA-40 to DEMETER with corrected temperature only = 0.127, 95% CI = −0.095–0.393 for the low malaria category).

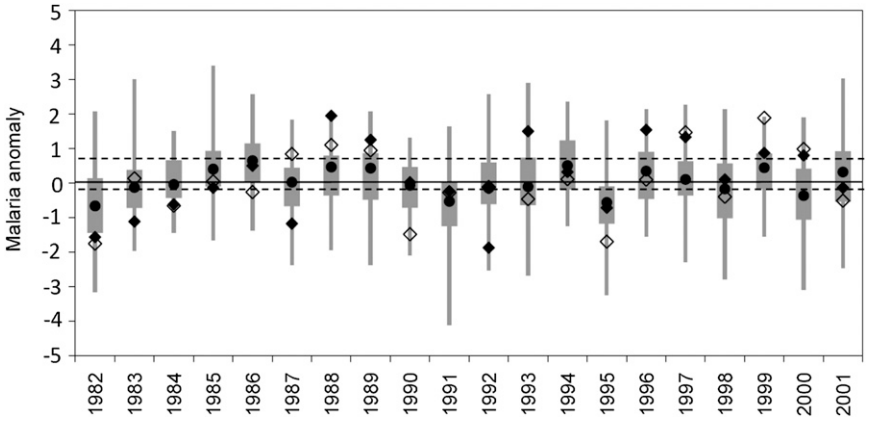

FIG. 3. Box plot of forecast malaria anomalies for 1982–2001 compared to simulated and observed anomalies. Model malaria anomalies for November forecast months 4–6 (FMA) driven by ERA-40 (open diamonds); 63-member DEMETER multimodel driven by uncorrected rainfall and bias-corrected temperature: median (filled circles), interquartile range (gray boxes), and full range (gray whiskers); and malaria anomalies from Thomson et al. (2005) (filled diamonds). Dashed lines show upper and lower tercile of observed anomalies. Forecast and simulated malaria anomalies were calculated separately for ERA-40 and for each DEMETER model.

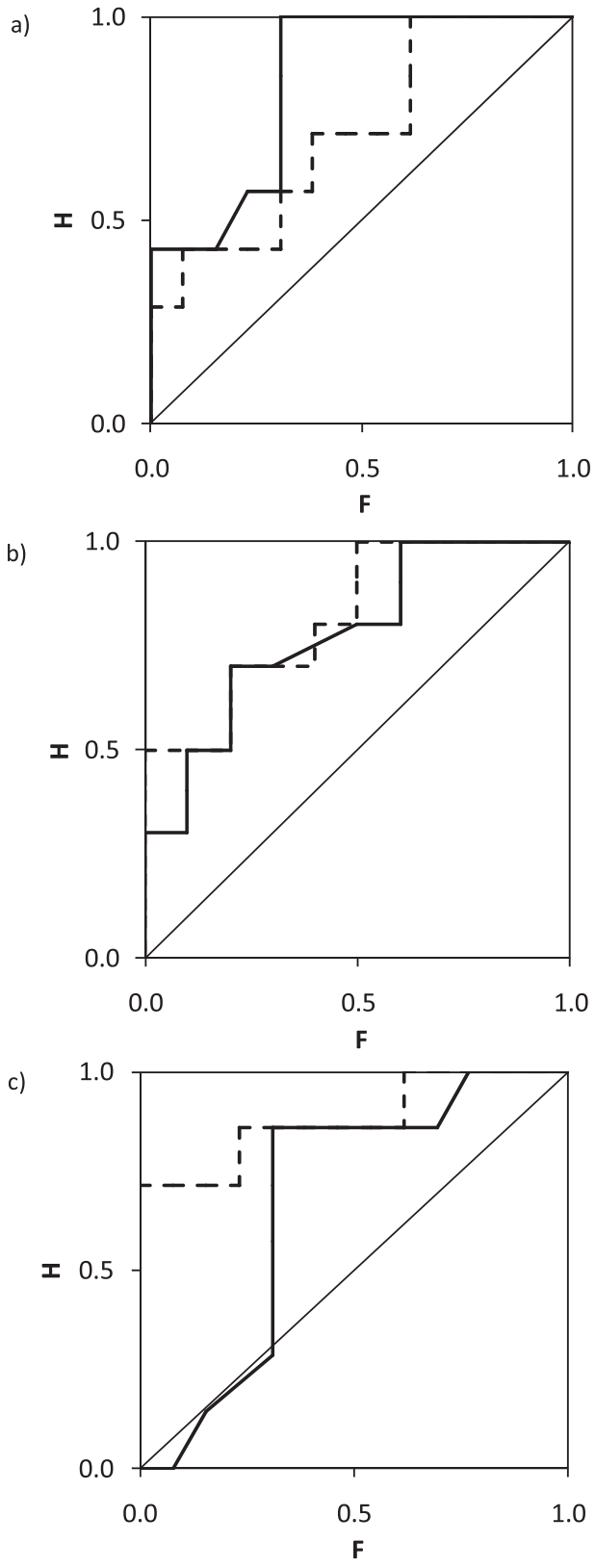

FIG. 4. ROC diagrams showing performance of DEMETER- and ERA-40-driven LMM forecasts for November forecast months 4–6 (FMA). DEMETER-driven malaria forecasts (solid line) and ERA-40-driven malaria simulations (dashed line). The diagonal line at $H = F$ corresponds to the skill of climatology. Validation is at tier 3 against observed countrywide malaria anomalies in Botswana for 1982–2001: (a) below the lower-tercile event, (b) above the median event, and (c) above upper-tercile event.

d. Tier-1 and tier-2 verification

The ROC areas obtained for performance assessment at tier 1 and tier 2 (Table 2) enable two types of comparison to be made: a comparison between the skill of the driving data and skill in the application model forecast, and a comparison between skill measured at tier 2 and tier 3.

At tier 1, the ROC areas for both rainfall and degree-days above 18°C are skillful for all thresholds and both integration periods, with the highest skill occurring for low rainfall (0.835 for FMA) and high degree-days (0.786 for FMA). Comparing with the ROC areas for the malaria model driven by uncorrected rainfall and corrected temperature in Table 1, the tier-3 ROC areas for incidence are consistent with the tier-1 skill in rainfall for DJF; and although the skill in incidence for low malaria (0.841) is actually higher than the skill in rainfall for the corresponding category (0.808), this difference is not significantly greater than zero at 95% (difference in ROC area = 0.033, 95% CI = −0.225–0.372). The tier-2 ROC areas for incidence during FMA are generally lower than those at tier 3, although again these differences are not significant at 95% confidence (not tabulated).

e. Potential economic value analysis

Calculation of the tier-3 potential economic value of the multimodel forecast (Fig. 5) reveals positive value for the DEMETER-driven malaria forecasts across the full range of cost/loss ratios for the low and above-average malaria events, but positive value only up to a cost/loss ratio of 0.6 for high malaria events. Where DEMETER performs least well is in predicting high malaria events, and the shape of the value curve for DEMETER in Fig. 5c reveals that for cost/loss ratios above 0.6, the value of the forecasts falls to zero because the probability decision threshold cannot be set sufficiently high for no false alarms to occur. In fact, examination of the individual DEMETER forecast probabilities (Fig. 6) reveals that the two years predicted as high malaria years by DEMETER with the highest forecast probabilities (1985–86) are not high malaria years according to the malaria index. For high malaria, both DEMETER value and ROC area (although still skillful) are much lower than

TABLE 2. Performance of DEMETER at tier 1 and tier 2: ROC areas for bias-corrected degree-days above 18°C and uncorrected rainfall against the same variables in ERA-40 (tier 1), and November forecast malaria anomalies compared to ERA-40-driven anomalies for the same period (tier 2). Values are given for forecast months 2–4 (DJF) and 4–6 (FMA). Range of values in brackets represent 95% confidence intervals calculated with 999 bootstrap samples. ROC areas above 0.5 indicate skill relative to climatology.

	Rainfall		Tmax-5 DD > 18°C		Incidence	
Malaria Event	DJF	FMA	DJF	FMA	DJF	FMA
Low	0.808 (0.556–0.969)	0.835 (0.608–1)	0.764 (0.510–0.947)	0.659 (0.405–0.906)	0.874 (0.688–1)	0.802 (0.586–0.967)
Above average	0.780 (0.522–0.976)	0.655 (0.354–0.886)	0.750 (0.494–0.945)	0.655 (0.365–0.904)	0.860 (0.670–1)	0.730 (0.475–0.949)
High	0.698 (0.429–0.929)	0.511 (0.260–0.774)	0.835 (0.560–1)	0.786 (0.493–1)	0.703 (0.420–0.929)	0.533 (0.255–0.797)

those of the ERA-40 simulations, which have positive value for the full range of cost-loss ratios and a very high ROC area of 0.879.

This study integrated a dynamic, process-based model of malaria with DEMETER multimodel seasonal forecasts and carried out verification of the combined modeling system for the prediction of observed malaria in Botswana. Tier-3 performance assessment of categorical events revealed the highest skill was found for forecasts driven by corrected temperature and uncorrected rainfall; forecasts were skillful for all three thresholds, with the best results obtained for low malaria events. For all events, the skill of the DEMETER-driven forecasts was higher than a ''best-guess'' control run, demonstrating the benefit of using the forecasts over using only the observed data available in November. For high (above the upper tercile) malaria events, although the ROC area was skillful, it was not significantly greater than climatology at 95% confidence. For low malaria events, the ERA-40-driven reference malaria simulations were less skillful than the best-performing DEMETER-driven forecasts, although this difference was not significant at 95% confidence. Potential economic value analysis of the DEMETER uncorrected rainfall and corrected temperature-driven forecasts revealed that they performed poorly for years they predicted with the highest probabilities of high malaria. Nevertheless, a positive value was found over a wide range of user cost-loss ratios.

7. Discussion and conclusions

The results described in section 6b concerning the impact of bias correction on forecast performance suggest that, for the particular datasets and application model used in this study, although correction of the driving temperatures is essential for skillful predictions, correction of rainfall is not. A significant improvement (over using raw data) was only obtaining data if rainfall was left uncorrected. This result is particularly surprising because rainfall has previously been found to be a major driver of malaria in Botswana; Thomson et al. (2006) successfully predicted seasonal malaria anomalies based on rainfall alone. An explanation may lie in a combination of the nature of the biases present in the daily DEMETER data and the sensitivity of the malaria model. In comparison with ERA-40, the uncorrected DEMETER models tend to contain large negative temperature biases over southern Africa: up to 60 or even 100 degree-days month^{-1}, corresponding to 2°–3.3° day^{-1}, sufficient to take some ensemble members below the 18°C threshold for transmission employed in LMM (not shown). Bias correction of temperature restores the temperatures to values at which transmission is not limited (Fig. 2b). For rainfall, biases in the raw data are positive: 50–70 mm month^{-1} during the main transmission months (Fig. 2a). A study of the cumulative distributions of rainfall for DEMETER (Fig. 1) reveals that, compared to ERA-40, DEMETER contains too few days with zero rainfall. This is a common problem in climate models, where the output is considered to be the average over a large grid box (Goddard et al. 2001). In correcting the rainfall with the frequency–intensity method, a large number of low rainfall days are lost (in the case of Fig. 1 all days with rainfall below approximately 6.5 mm). The poorer malaria forecast performance for corrected driving rainfall could therefore be due to a combination of two factors: 1) sensitivity of the malaria model to the interannual variability of some features of the rainfall that is removed, and 2) systematic errors in the malaria model that may favor more continuous (i.e., unrealistic) rainfall patterns in certain circumstances. The second factor seems less likely to be the case for the above-average and high malaria categories, where the more realistic ERA-40-driven simulations perform better than those driven by DEMETER; however, it may help to explain the result that, for low malaria years, DEMETER-driven forecast skill was higher than that of the ERA-40-driven simulations. Errors in the ERA-40 dataset may have also contributed to this result; for example, for 1990, ERA-40-driven malaria was below the lower tercile but observed

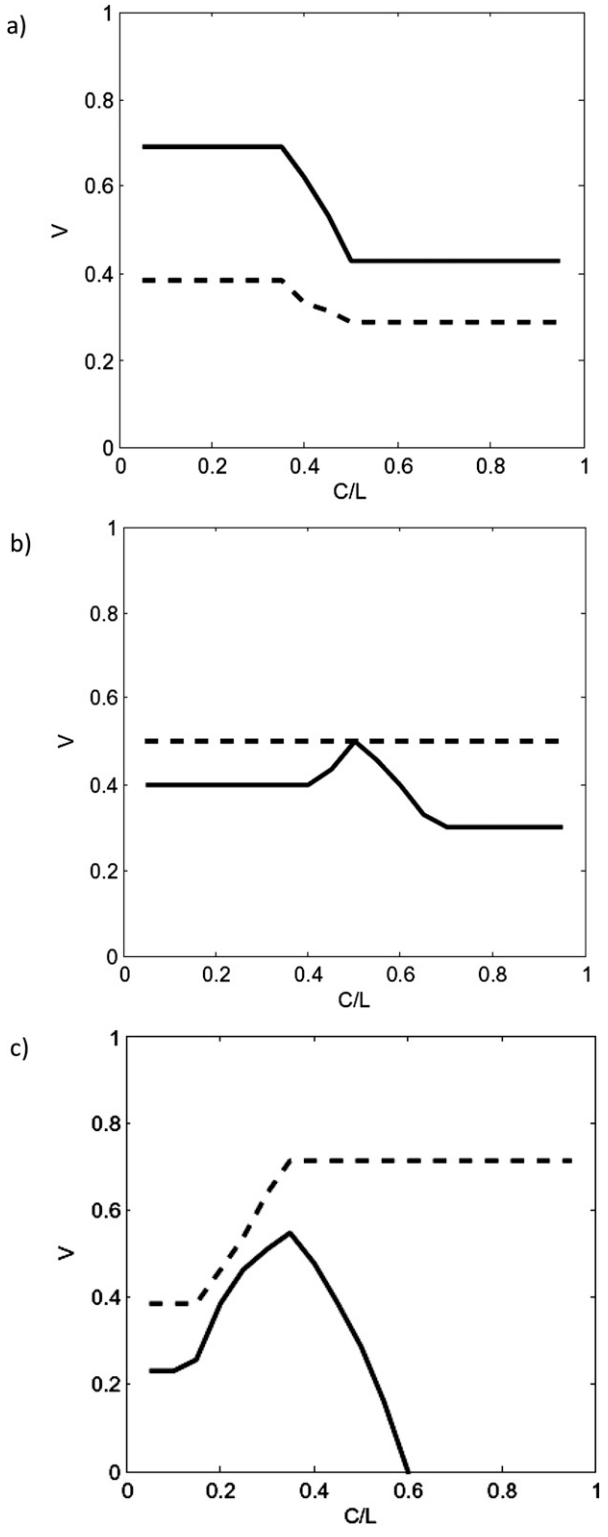

FIG. 5. Potential economic value as a function of user cost/loss ratio for FMA DEMETER-driven malaria forecasts (solid line) compared to value of ERA-40 driven malaria simulations (dashed line). Validation is at tier 3 against observed countrywide malaria anomalies in Botswana for 1982–2001: (a) low malaria event (below the lower tercile), (b) above-average malaria event (above the median), and (c) high malaria event (above the upper tercile).

malaria was only just below the median, whereas DEMETER-driven malaria was consistent with the observations (see Fig. 2); ERA-40 rainfall for that year was close to the lower tercile, whereas CMAP rainfall was above the median (not shown). That the aforementioned differences in performance between DEMETER and the ERA-40 reference were not found to be statistically significant at 95% confidence illustrates the difficulty in drawing robust conclusions based on the relatively "short" datasets available for seasonal forecast validation, where individual years can have an impact on the measured skill.

Although it was not the aim of this study to carry out operational validation of the integrated modeling system for malaria early warning, the results obtained here have provided an insight into the potential usefulness of seasonal climate forecasts for malaria early warning when coupled with a dynamic, process-based disease model. The cost-loss decision model employed here has been shown to be applicable to a number of "real-life" decision making cases (Wilks 2006), and although real user cost/loss ratios are difficult to obtain for health applications, quoted figures for other applications tend to be low, around 0.1 or less (Richardson 2003). In the potential economic value curves (Fig. 5), these cost/loss ratios correspond to regions of positive value in the DEMETER-driven LMM forecasts—value that is obtained by setting very low probability decision thresholds so that no events are missed. This implies DEMETER-driven LMM malaria forecasts could have value in saving expense by not acting when very low forecast probabilities are issued. However, the situation for malaria control will inevitably be more complicated than this simple cost–loss model; assessment in an operational context would be required to give a clearer indication of the real value of the forecasts. In fact, truly integrated modeling would need to include not only the application model but also the decision model itself.

Several adjustments would need to be made to the assessment carried out here to provide a validation of an operational system. For example, to derive robust model climatologies, bias correction of the DEMETER forecasts made use of both past and future data relative to the year being corrected. Operationally, model biases would need to be corrected using historical data only. Furthermore, both past and future data were used to define the events by which forecast skill was measured; operational assessment would require the forecasts each year to define events relative to past climatologies only. Other limitations include the degree to which model simulations can hope to represent "real-world" malaria; for example, here, gridded climate datasets have been used, where a single grid square represents a large area over which transmission may in reality be highly variable.

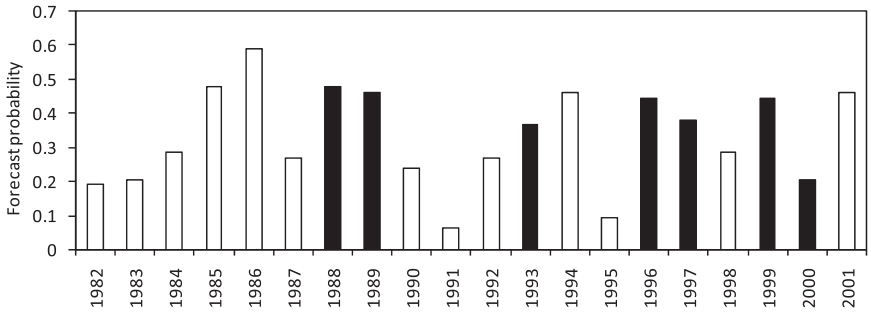

FIG. 6. DEMETER-driven LMM forecast probabilities of above the upper-tercile malaria in Botswana for 1982–2001 obtained from forecast for November months 4–6 (FMA). Above the upper-tercile years according to the published index of malaria anomalies are shaded black, other years are shown with open bars.

The LMM, although complex and based on well-understood components of the malaria life cycle, makes several simplifications that limit the model's ability to simulate real, observed malaria. Hence, the types of malaria forecast that have been used here should really only be treated as simple indicators of malaria risk, and if used operationally, they should only form part of a toolset available to decision makers in a malaria early warning system.

The results obtained in this study for Botswana demonstrated more skill for the November forecast than found in a previous assessment of DEMETER-driven malaria forecasts for grid points in Zimbabwe (Morse et al. 2005). In Botswana, the study of Thomson et al. (2006), using an empirical malaria model fitted to observed monthly rainfall anomalies and applied to a subset of three DEMETER models, found similar skill for low (below the lower quartile) malaria events to that found here but higher skill for high malaria events (0.80 with the empirical model compared to 0.67 with LMM). Because of the advantages of process-based models such as LMM over empirical ones, including the potential to model human interventions to disease transmission, it is very encouraging to find some seasonal predictive skill here with a process-based model.

Because the simple bias correction procedure used here had a negative impact on forecast skill, an important area for future work highlighted by this study is the preprocessing of the rainfall forecasts used to drive the model. Better bias correction may be achieved using statistical downscaling techniques (e.g., Moron et al. 2008), including, for example, the use of a weather generator (Racsko et al. 1991) to produce a time series of rainfall by combining monthly anomalies from the seasonal forecasts with the correct local statistical rainfall properties obtained from observations. This approach has already been used for seasonal crop forecasting (Hansen et al. 2006). Alternatively, a statistical model could be employed to link a set of GCM forecast variables to the required input application model variables by regression on historical data (e.g., Paeth et al. 2006). The use of high-resolution regional climate models, where improved representations of rainfall-generating processes should result in more realistic and more skillful rainfall forecasts, may be the ultimate solution, and these are already being applied to African locations (Afiesimama et al. 2006; Herceg et al. 2007).

The issue of coupling impact models with seasonal forecasts is not restricted to malaria modeling, and future work will proceed as part of a community of users requiring seasonal application model forecasts for climate-sensitive decisions, in particular crop modeling (e.g., Challinor et al. 2005; Hansen et al. 2006) and hydrology (e.g., Bader et al. 2006; Chowdhury and Ward 2007). The results of the present study suggest other mosquito-borne diseases, which have been linked to climate such as Dengue fever and Rift Valley fever (Gagnon et al. 2001; Indeje et al. 2006), could potentially be simulated and successfully predicted using a similar modeling approach.

In section 1, it was stated that integrated impact forecasting should provide a framework by which the user could learn to make best use of the climate forecasts while also providing feedback on areas of improvement to the forecast provider. This study integrated a dynamic, process-based malaria model with forecasts of climatic variables to create and verify a climate-driven seasonal malaria-forecasting model. The results show that, at least for Botswana, where rainfall appears to be a major driver of malaria, the skill of DEMETER-driven malaria forecasts is consistent with the tier-1 skill of the rainfall forecasts used to drive the model. The results also indicate that, even though the patterns of rainfall in the forecasts may be unrealistic, a dynamic process-based impact model can make use of skillful elements of variability within the rainfall and temperature forecasts to make skillful predictions of malaria. In terms of

feedback to the forecast provider, it can be concluded that for Botswana better rainfall forecasts (and possibly forecasts with more realistic rainfall patterns) should result in better malaria predictions. In particular, these results indicate a key area of improvement is the forecasting of high rainfall and, in turn, high malaria events.

Acknowledgments. Moshe Hoshen is acknowledged for original development of the malaria model. This research was funded by the European Commission as part of the DEMETER project (Contract EVK2-CT-1999-00024), and its successor ENSEMBLES (Contract GOCE-CT-2003-505539). AEJ received support from a Natural Environmental Council E-science studentship. We thank all the reviewers for their helpful comments.

REFERENCES

Afiesimama, E. A., J. S. Pal, B. J. Abiodun, W. J. Gutowski, and A. Adedoyin, 2006: Simulation of West African monsoon using the RegCM3. Part I: Model validation and interannual variability. *Theor. Appl. Climatol.,* **86,** 23–37.

Anderson, R. M., and R. M. May, 1991: *Infectious Disease of Humans: Dynamics and Control*. Oxford University Press, 757 pp.

Bader, J. C., J. P. Piedelievre, and J. P. Lamagat, 2006: Seasonal forecasting of the flood volume of the Senegal River, based on results of the ARPEGE Climate model. *Hydrol. Sci. J.,* **51,** 406–417.

Bayoh, M. N., and S. W. Lindsay, 2003: Effect of temperature on the development of the aquatic stage of Anopheles gambiae sensu stricto (Diptera: Culicidae). *Bull. Entomol. Res.,* **93,** 375–381.

——, and ——, 2004: Temperature-related duration of aquatic stages of the Afrotropical malaria vector mosquito Anopheles gambiae in the laboratory. *Med. Vet. Entomol.,* **18,** 174–179.

Behera, S. K., and T. Yamagata, 2001: Subtropical SST dipole events in the southern Indian Ocean. *Geophys. Res. Lett.,* **28,** 327–330.

Bezuidenhout, C. N., and R. E. Schulze, 2006: Application of seasonal climate outlooks to forecast sugarcane production in South Africa. *Climate Res.,* **30,** 239–246.

Buizer, J. L., J. Foster, and D. Lund, 2000: Global impacts and regional actions: Preparing for the 1997–98 El Niño. *Bull. Amer. Meteor. Soc.,* **81,** 2121–2139.

Camberlin, P., S. Janicot, and I. Poccard, 2001: Seasonality and atmospheric dynamics of the teleconnection between African rainfall and tropical sea surface temperature: Atlantic vs. ENSO. *Int. J. Climatol.,* **21,** 973–1005.

Cane, M. A., G. Eshel, and R. W. Buckland, 1994: Forecasting Zimbabwean maize yield using eastern equatorial Pacific sea surface temperature. *Nature,* **370,** 204–205.

Carpenter, T. M., and K. P. Georgakakos, 2004: Impacts of parametric and radar rainfall uncertainty on the ensemble streamflow simulations of a distributed hydrologic model. *J. Hydrol.,* **298,** 202–221.

Challinor, A. J., J. M. Slingo, T. R. Wheeler, and F. J. Doblas-Reyes, 2005: Probabilistic simulations of crop yield over western India using the DEMETER hindcast ensembles. *Tellus,* **57A,** 498–512.

Chowdhury, M. R., and M. N. Ward, 2007: Seasonal flooding in Bangladesh—Variability and predictability. *Hydrol. Processes,* **21,** 335–347.

Clarke, L., D. Kilminster, and L. A. Smith, 2004: Do multi-model ensemble forecasts yield added value? *First THORPEX Int. Science Symp.,* Montreal, QC, Canada, National Oceanic and Atmospheric Administration. [Available online at http://www.lse.ac.uk/collections/cats/papersPDFs/LC_do%20multi%20model%20ensemble_2005.pdf.]

Connor, S. J., J. DaSilva, and S. Katikiti, 2007: Malaria control in southern Africa. *Climate Risk Management in Africa: Learning From Practice,* M. E. Hellmuth et al., Eds., International Research Institute for Climate and Society, 45–57.

DaSilva, J., B. Garanganga, V. Teveredzi, S. M. Marx, S. Mason, and S. J. Connor, 2004: Improving epidemic malaria planning, preparedness and response in southern Africa. *Malar. J.,* **3,** 37, doi:10.1186/1475-2875-3-37.

Detinova, T. S., W. N. Beklemishev, and D. S. Bertram, 1962: Age grouping methods in Diptera of Medical Importance. *World Health Organization Monograph,* No. 47, World Health Organization, 216 pp.

Doblas-Reyes, F. J., R. Hagedorn, and T. N. Palmer, 2006: Developments in dynamical seasonal forecasting relevant to agricultural management. *Climate Res.,* **33,** 19–26.

Efron, B., 1981: Nonparametric estimates of standard error: The jackknife, the bootstrap, and other methods. *Biometrika,* **68,** 589–599.

Eldaw, A. K., J. D. Salas, and L. A. Garcia, 2003: Long-range forecasting of the Nile River flows using climatic forcing. *J. Appl. Meteor.,* **42,** 890–904.

Gagnon, A. S., A. B. G. Bush, and K. E. Smoyer-Tomic, 2001: Dengue epidemics and the El Niño–Southern Oscillation. *Climate Res.,* **19,** 35–43.

Githeko, A. K., S. W. Lindsay, U. E. Confalonieri, and J. A. Patz, 2000: Climate change and vector-borne diseases: A regional analysis. *Bull. W. H. O.,* **78,** 1136–1147.

Goddard, L., S. J. Mason, S. E. Zebiak, C. F. Ropelewski, R. Basher, and M. A. Cane, 2001: Current approaches to seasonal-to-interannual climate predictions. *Int. J. Climatol.,* **21,** 1111–1152.

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.

Hansen, J. W., and M. Indeje, 2004: Linking dynamic seasonal climate forecasts with crop simulation for maize yield prediction in semi-arid Kenya. *Agric. For. Meteor.,* **125,** 143–157.

——, A. Challinor, A. Ines, T. Wheeler, and V. Moron, 2006: Translating climate forecasts into agricultural terms: Advances and challenges. *Climate Res.,* **33,** 27–41.

Herceg, D., A. H. Sobel, and L. Q. Sun, 2007: Regional modeling of decadal rainfall variability over the Sahel. *Climate Dyn.,* **29,** 89–99.

Hoshen, M. B., and A. P. Morse, 2004: A weather-driven model of malaria transmission. *Malar. J.,* **3,** doi:10.1186/1475-2875-3-32.

Indeje, M., M. N. Ward, L. J. Ogallo, G. Davies, M. Dilley, and A. Anyamba, 2006: Predictability of the normalized difference vegetation index in Kenya and potential applications as an indicator of Rift Valley Fever outbreaks in the Greater Horn of Africa. *J. Climate,* **19,** 1673–1687.

Ines, A. V. M., and J. W. Hansen, 2006: Bias correction of daily GCM rainfall for crop simulation studies. *Agric. For. Meteor.,* **138,** 44–53.

Krishnamurti, T. N., A. Chakraborty, R. Krishnamurti, W. K. Dewar, and C. A. Clayson, 2006: Seasonal prediction of sea surface

temperature anomalies using a suite of 13 coupled atmosphere–ocean models. *J. Climate,* **19,** 6069–6088.

Landman, W. A., S. J. Mason, P. D. Tyson, and W. J. Tennant, 2001: Statistical downscaling of GCM simulations to streamflow. *J. Hydrol.,* **252,** 221–236.

Lazar, A., A. Vintzileos, F. J. Doblas-Reyes, P. Rogel, and P. Delecluse, 2005: Seasonal forecast of tropical climate with coupled ocean–atmosphere general circulation models: On the respective role of the atmosphere and the ocean components in the drift of the surface temperature error. *Tellus,* **57A,** 387–397.

Lee, J. W., 2006: Report by the director general to the 59th World General Assembly, 22 May 2006. WHO. [Available online at http://www.who.int/mediacentre/events/2006/wha59/dr_lee_report/en/index.html.]

le Sueur, D., B. L. Sharp, E. Gouws, and S. Ngxongo, 1996: Malaria in South Africa. *S. Afr. Med. J.,* **86,** 936–939.

Lindsay, S. W., and W. J. M. Martens, 1998: Malaria in the African highlands: Past, present, and future. *Bull. W. H. O.,* **76,** 33–45.

Mabaso, M. L. H., B. Sharp, and C. Lengeler, 2004: Historical review of malarial control in southern African with emphasis on the use of indoor residual house-spraying. *Trop. Med. Int. Health,* **9,** 846–856.

Martens, W. J. M., T. H. Jettens, J. Rotmans, and L. W. Niessen, 1995: Climate change and vector-borne diseases—A global modeling perspective. *Glob. Environ. Change,* **5,** 195–209.

Martin, R. V., R. Washington, and T. E. Downing, 2000: Seasonal maize forecasting for South Africa and Zimbabwe derived from an agroclimatological model. *J. Appl. Meteor.,* **39,** 1473–1479.

Mason, I. B., 2003: Binary events. *Forecast Verification: A Practitioner's Guide in Atmospheric Science,* I. T. Jolliffe and D. B. Stephenson, Eds., John Wiley and Sons, 37–73.

Mason, S. J., and N. E. Graham, 2002: Areas beneath the relative operating characteristics (ROC) and relative operating levels (ROL) curves: Statistical significance and interpretation. *Quart. J. Roy. Meteor. Soc.,* **128,** 2145–2166.

Molineaux, L., 1988: The epidemiology of human malaria as an explanation of its distribution, including some implications for its control. *Malaria: Principles and Practice of Malariology,* W. H. Wernsdorfer and I. McGregor, Eds., Churchill Livingstone, 913–998.

Moron, V., A. W. Robertson, M. N. Ward, and O. Ndiaye, 2008: Weather types and rainfall over Senegal. Part II: Downscaling of GCM simulations. *J. Climate,* **21,** 288–307.

Morse, A. P., F. J. Doblas-Reyes, M. B. Hoshen, R. Hagedorn, and T. N. Palmer, 2005: A forecast quality assessment of an end-to-end probabilistic multimodel seasonal forecast system using a malaria model. *Tellus,* **57A,** 464–475.

Murphy, A. H., 1977: The value of climatological, categorical, and probabilistic forecasts in the cost-loss ratio situation. *Mon. Wea. Rev.,* **105,** 803–816.

Nájera, J. A., R. L. Kouznetsov, and C. Delacollette, 1998: Malaria epidemics: Detection and control, forecasting, and prevention. World Health Organization, WHO/MAL/98.1084. [Available online at http://www.rollbackmalaria.org/docs/najera_epidemics/naj_toc.htm.]

Nicholson, S. E., and E. Kim, 1997: The relationship of the El Niño–Southern Oscillation to Africa rainfall. *Int. J. Climatol.,* **17,** 117–135.

Paeth, H., R. Girmes, G. Menz, and A. Hense, 2006: Improving seasonal forecasting in the low latitudes. *Mon. Wea. Rev.,* **134,** 1859–1879.

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.

Philips, J. G., M. A. Cane, and C. Rosenzweig, 1998: ENSO, seasonal rainfall patterns and simulated maize yield variability in Zimbabwe. *Agric. For. Meteor.,* **90,** 39–50.

Racsko, P., L. Szeidl, and M. A. Semenov, 1991: A serial approach to local stochastic weather models. *Ecol. Modell.,* **57,** 27–41.

Richardson, D. S., 2003: Economic value and skill. *Forecast Verification: A Practitioner's Guide in Atmospheric Science,* I. T. Jolliffe and D. B. Stephenson, Eds., John Wiley and Sons, 165–187.

Rogers, D. J., and S. E. Randolph, 2006: Climate change and vector-borne disease. *Adv. Parasitol.,* **62,** 345–381.

Roll Back Malaria Partnership, 2005: Global strategic plan 2005–2015. WHO. 43 pp. [Available online at http://www.rollbackmalaria.org/forumV/docs/gsp_en.pdf.]

Schmidli, J., C. Frei, and P. L. Vidale, 2006: Downscaling from GCM precipitation: a benchmark for dynamical and statistical downscaling methods. *Int. J. Climatol.,* **26,** 679–689.

Service, M. W., and H. Townson, 2002: The Anopheles vector. *Essential Malariology,* D. A. Warrell and H. M. Gilles, Eds., Arnold, 59–84.

Smith, D. L., and F. Ellis McKenzie, 2004: Statics and dynamics of malaria infection in Anopheles mosquitoes. *Malar. J.,* **3,** doi:10.1186/1475-2875-3-13.

Stockdale, T. N., 1997: Coupled ocean–atmosphere forecasts in the presence of climate drift. *Mon. Wea. Rev.,* **125,** 809–818.

Thomson, M. C., S. J. Mason, T. Phindela, and S. J. Connor, 2005: Use of rainfall and sea surface temperature monitoring for malaria early warning in Botswana. *Amer. J. Trop. Med. Hyg.,* **73,** 214–221.

——, F. J. Doblas-Reyes, S. J. Mason, R. Hagedorn, S. J. Connor, T. Phindela, A. P. Morse, and T. N. Palmer, 2006: Malaria early warnings based on seasonal climate forecasts from multimodel ensembles. *Nature,* **439,** 576–579.

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

Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences*. Academic Press, 627 pp.

Wood, A. W., A. Kumar, and D. P. Lettenmaier, 2005: A retrospective assessment of National Centers for Environmental Prediction climate model-based ensemble hydrologic forecasting in the western United States. *J. Geophys. Res.,* **110,** D04105, doi:10.1029/2004JD004508.

World Health Organization, 2001: Malaria early warning systems: concepts, indicators, and partners—A framework for field research in Africa. WHO, WHO/CDS/RBM/2001.32, 80 pp. [Available online at http://helid.desastres.net/en/d/Js2893e/.]

——, 2005: World malaria report 2005. WHO, 294 pp. [Available online at http://rbm.who.int/wmr2005/html/toc.htm.]

Worrall, E., S. J. Connor, and M. C. Thomson, 2007: A model to simulate the impact of timing, coverage and transmission intensity on the effectiveness of indoor residual spraying (IRS) for malaria control. *Trop. Med. Int. Health,* **12,** 75–88.

Xie, P., and P. A. Arkin, 1997: Global precipitation: A 17-yr monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bull. Amer. Meteor. Soc.,* **78,** 2539–2558.