Seasonal temperature prediction skill over southern Africa and human health

Melissa J. Lazenby
University of Sussex, Department of Geography, Brighton, United Kingdom

Willem A. Landman
Council for Scientific and Industrial Research, Natural Resources and the Environment, and Department of Geography, Geoinformatics and Meteorology, University of Pretoria, South Africa

Rebecca M. Garland
Council for Scientific and Industrial Research, Natural Resources and the Environment, Pretoria, South Africa

David G. DeWitt
International Research Institute for Climate and Society, Lamont-Doherty Earth Observatory of Columbia University, Palisades, New York, USA

ABSTRACT

An assessment of probabilistic prediction skill of seasonal temperature extremes over southern African is presented. Verification results are presented for six run-on seasons; September to November, October to December, November to January, December to February, January to March and February to April over a 15-year retroactive period. Comparisons are drawn between downscaled seasonal 850 hPa geopotential height field forecasts of a two-tiered system versus downscaled height forecasts from a coupled ocean-atmosphere system. The ECHAM4.5 atmospheric general circulation model is used for both systems; in the one-tiered system the ECHAM4.5 is directly coupled to the ocean model MOM3, and the two-tiered system the ECHAM4.5 is forced with Van den Dool SST hindcasts. Model output statistic equations are developed using canonical correlation analysis to reduce system deficiencies. Probabilistic verification is conducted using the relative operating characteristic (ROC) and reliability diagram. The coupled model performs best in capturing seasonal maximum temperature extremes. Seasons demonstrating the highest ROC scores coincide with the period of highest seasonal temperatures found over southern Africa. The above-normal category of the one-tiered system indicates the highest skill in predicting maximum temperature extremes, implying the coupled model skilfully predicts when there is a high likelihood of experiencing extremely high seasonal maximum temperatures during mid to late summer. The downscaled coupled maximum temperature hindcasts are additionally evaluated in terms of their monetary value and quality to the general public. The seasonal forecast system presented here should be able to reduce risks in decision making by the health industry in southern Africa.

1. Introduction

The dramatic improvement in supercomputer power as well as an increase in the understanding of atmospheric processes have resulted in the progression of seasonal forecasts experiencing a vast improvement over the past decade (Cane *et al*., 1994; Hastenrath *et al*., 1995; Barnston and Smith 1996; Hunt 1997; Mason *et al*., 1996, 1999; Jury 1996; Mason 1998; Mattes and Mason 1998; Makarau and Jury 1997; Jury *et al*., 1999; Landman and Mason 1999b; Landman and Tennant 2000; Landman *et al*., 2001a; Landman and Goddard, 2002). A forecastat any time-scale can never be 100% accurate due to the atmosphere having inherent internal variability as a key characteristic, (Doblas-Reyes *et al*., 2000) therefore requiring seasonal climate simulations to be probabilistically represented (e.g. Mason *et al*., 1999; Goddard *et al*., 2001; Goddard and Mason, 2002). General circulation model (GCM) ensembles, if correctly configured, allow for such probabilistic forecasts, due to ensemble forecasting providing a viable way to approximate the probability spread of atmospheric states (Brankovic and Palmer, 2000; Landman and Beraki, 2012). Improving the skill of these forecasts is also made possible by expanding the understanding of local climate systems and their respective processes, in this case over southern Africa (Klopper *et al*., 1998).

Dynamic numerical models that include the interactions between the atmosphere, ocean, and land should theoretically provide superior seasonal forecasts than purely statistical approaches, asthey have the capability of handling a vast range of both linear and nonlinear interactions and their probable flexibility against a changing climate (Barnston *et al*., 1999).Therefore the choice of using dynamical models in seasonal forecasting has become a significantly popular one (Stockdale *et al*., 1998; Landman *et al*., 2009b). However, even with such improvements, model errors are still a substantial source of problems (Latif *et al*., 2001; Palmer *et al*., 2004), and it is still unclear as to what degree the current generation of numerical forecast models still in use are able to challenge and improve upon existing empirical methods in seasonal forecasting (Van Oldenborgh *et al*., 2004). Barnston *et al*. (1999) concluded that dynamical models were not able to capture the 1997/1998 El Niño event and following La Niña event better than statistical models (Van Oldenborgh *et al*., 2004).

Skill has been established by GCMs for a variety of scales with resolutions of approximately 100–300 km (Landman and Goddard, 2002; Landman *et al*., 2012; Landman and Beraki, 2012). However, GCMs do not have the ability to accurately capture local smaller-scale features, and a common consequence is overestimating certain variables such as rainfall over southern Africa (Joubert and Hewitson, 1997; Mason and Joubert, 1997; Landman *et al*., 2009b). The representation of rainfall at mid-to-high latitudes is highly complex and more often than not poorly predicted (Graham *et al*., 2000; Goddard and Mason, 2002). Such systematic biases have resulted in statistical recalibration and downscaling of GCM simulations becoming a necessity, particularly over places such as southern Africa (Landman and Beraki, 2012). A viable method to make the necessary corrections for these biases is to employ a model output statistics (MOS) approach (e.g., Landman and Goddard, 2002; Landman and Beraki, 2012). Previous research has been conducted on forecasting seasonal rainfall over southern Africa using global models that have been downscaled using a MOS approach and was found to perform better during El Niño and La Niña seasons than during neutral years (neither an El Niño nor a La Niña event) or years when there was no strong El Niño or La Niña influence (Landman and Beraki, 2012).

Seasonal temperature forecasts over southern Africa however, have been neglected since the seminal paper completed by Klopper *et al* in 1998, where a deterministic assessment of forecast skill was investigated. A more recent study of temperature variation over Africa shows that a significant increase in temperatures has been seen over Africa and that it is not exclusively due to variations in the El Niño-Southern Oscillation but rather due to other natural variability of the climate and/or human activity (Collins, 2011). This current paper presents a probabilistic assessment of forecast skill of temperature, including both minimum and maximum seasonal temperatures by using state-of-the-art GCM configurations.

Due to the seasonal progression of accurate sea surface temperature (SST) anomaly predictions, it has been made possible to produce seasonal-average weather forecasts by incorporating them in atmospheric GCMs (Graham *et al*., 2000; Goddard and Mason, 2002). Using such a modelling system (two-tiered) to predict a particular regions seasonal characteristics has been operational in South Africa for numerous years (e.g., Landman *et al*., 2001). The one-tiered system, which followed that of the two-tiered system, is known as fully coupled ocean-atmosphere model (e.g., Stockdale *et al*., 1998; Saha *et al*., 2006; Weisheimer *et al*., 2009). Coupled models fundamental feature is their ability to describe interactions between the atmosphere and the ocean, and therefore aim to yield more reliable seasonal forecasts, whereas two-tiered systems exclude this interaction, however do include the atmospheres response to SSTs (Copsey *et al*., 2006; Troccoli *et al*., 2008). Coupled ocean-atmosphere models are at the top of the modelling ladder in terms of complexity and computational expense (DeWitt, 2005). Theoretically fully coupled models should outperform the two-tiered forecasting system, as they should capture reality more accurately and therefore produce more accurate forecasts (DeWitt, 2005).

Global models are employed in this study to predict seasonal minimum and maximum temperature extremes over southern Africa during the austral summer period from September through April. Winter has been omitted due to minimal forecast skill found in the austral winter forecasts (Mason *et al*., 1996; Landman *et al*., 2012). Moreover, extremes in minimum and maximum temperatures have a direct impact on human health (particularly maximum temperatures); therefore an application of this study may include the potential of health hazards associated with extremely hot summer seasons over the region.

Exposure to high temperatures can impact human health in many different ways. Such exposure to high temperatures can lead to symptoms such as fatigue, dizziness and cramps, as well as heat illnesses such as heat exhaustion and heatstroke. High ambient temperatures, including those experienced during a heat wave, have been associated with increases in mortality. These increases in mortality are not just from heatstroke, but also from cardiovascular, cerebrovascular and respiratory diseases and have also been seen in all-cause mortality (Smoyer-Tomic *et al*., 2001; Vaneckova *et al*., 2011; Baccini *et al*., 2008; Rocklöv *et al*., 2011; Ballester *et al*., 2011; Medina-Ramón and Schwartz 2007; Diaz *et al*., 2002). The heat-health relationship varies for different locations and for different population, and thus it is critical to develop these relationships based on local health and meteorological data. Those most vulnerable to heat have generally been found to be the elderly, people living in urban areas, people with pre-existing cardiovascular and respiratory disease and those with compromised coping capacities (Basu and Samet, 2002; Kovats and Hajat, 2008; Naughton *et al*., 2002; Semenza *et al*., 1996). In addition to increased mortality with increased ambient temperatures, there is some evidence of increases in hospital and emergency admissions for specific heat related illnesses (such as respiratory diseases), particularly amongst vulnerable groups (Wichmann, *et al*., 2011; Michelozzi *et al*., 2009, Green *et al*., 2010). However, there are far fewer studies on non-fatal illnesses and their relationship to high ambient temperature, and thus less of a consensus on the impact.

In a study in London, the increases in admissions during high ambient temperatures were not at the same magnitude as increases in mortality, which led the authors to suggest that many heat related deaths may occur before they receive medical attention (Kovats *et al*., 2004). Analyses from the 1995 and the 1999 Chicago heat wave determined factors that isolated people such as living alone, not leaving home every day and being confined to bed were the strongest risk factors for heat-related death (Naughton *et al*., 2002; Semenza *et al*., 1996). Thus, in order to protect public health from high temperatures, it is critical to both empower the public, through education and alerts, to be able to effectively protect their own health as well as increase assistance and prevention measures, particularly focused on those who are vulnerable and isolated.

Planning measures, such as heat warnings and heat-health plans, have been implemented in many countries to aid in the prevention of increases in morbidity and mortality from high temperatures events (e.g., Kalkstein *et al*., 1996; Tan *et al*., 2003; Sheridan and Kalkstein, 2004; Smoyer-Tomic and Rainham, 2001; Pascal *et al*., 2006). Some weather bureaus call for a heat warning when temperatures are forecasted to reach above a certain threshold. Some cities have more elaborate intervention plans that are activated by such warnings. For example, in Philadelphia, Pennsylvania, USA, there are 10 activities that are enacted once a warning has been issued by the U.S. National Weather Service ranging from media announcements, to halting suspensions of utility services, to increases in emergency medical service staffing (Sheridan and Kalkstein, 2004; Kalkstein *et al*., 1996). An evaluation of Philadelphia heat watch/warning system concluded that for 1995-1998, issuing a warning saved 2.6 lives on average with high benefits and low costs (Ebi *et al*., 2004).

After the August 2003 heat wave in France where the excess mortalities during the heat wave was 14,800, the French government developed a Heat Health Watch Warning System (Pascal *et al*., 2006). In July 2006, France experienced another severe heat wave, though not as intense as the 2003 heat wave, and the system was utilized. While there still were 2100 excess deaths recorded for that period, it was ~1/3 the predicted excess deaths from a health model for the same period (Fouillet *et al*., 2008). While it is difficult to elucidate the exact reason for these decreases, it may be due to increased public awareness, preventative measures and the heat warning system developed by the government (Fouillet *et al*., 2008). In general, evaluations of heat warning systems are difficult, as there are many factors that might influence the changes in mortality. However, as the impact of high temperature on mortality is well-documented, and temperatures and heat waves are expected to increase in the future due to climate change, the need for effective heat-health warnings and plans are increasing in order to prevent large public health impacts from increasing temperatures.

Most of the research on increases in mortality with high temperatures has been in temperate regions. Thus, there is a lack of knowledge in general of how populations living in more tropical environments might be impacted by increasing temperatures, though the limited research does suggest that increases in temperature do impact mortality rates (Vaneckova *et al*., 2011; Kynast-Wolf *et al*., 2010). In southern Africa little is known on how public health is impacted by high ambient temperatures, with only one study in Cape Town that investigated the heat-mortality relationship (McMichael *et al*., 2008). Previous work has focused on occupational health and heat stress, with a focus on the mining industry (e.g. Wyndham, 1965, Mathee *et al*., 2010). And, as the interior regions of southern Africa are projected to experience increases in temperature as great as 4-6˚C under the A2 emission scenario by the end of the century, the occurrence of heat events is also projected to increase (Engelbrecht *et al*., 2011).

Increased temperatures will not only directly impact health through increased heat-related illnesses, as many health impacts are also impacted by climate factors (e.g., malaria, cholera, dengue fever and malnutrition.). There is a need for accurate seasonal forecasts of climate variables in order to effectively plan for health impacts in advance. For heat impacts on health, much of the focus internationally has been on short-term forecasting, however longer term forecasting will aid in planning (McGregor *et al*., 2004; McGregor *et al*., 2006). Little research has been performed on applying seasonal forecasting of extreme events to health impacts, though seasonal forecasts have been utilized in developing early warning systems for infectious diseases (e.g., Kuhn *et al*., 2004; Thompson *et al*., 2006; Kelly-Hope and Thomson, 2008; Connor *et al*., 2008). Seasonal forecasts can provide information with long lead times that will allow the health community with enough time for planning. Myers *et al*. (2000) suggests that for epidemic forecasting, when it is critical to give health workers enough time to plan for unexpectedly high (or low) number of cases, a lead time of 2-6 months is the most useful. A longer lead time is also needed for decisions such as increasing the health care sector budget (Myers *et al*., 2000). Such seasonal forecasts can then be followed-up by shorter range forecasts, which can provide more precise spatial and temporal information for more focused response measures. Such a combination provides information across decision timescales in order to effectively prepare to mitigate negative health impacts.

In order to develop effective plans on both the seasonal and shorter timescales, adequate skill in forecasting is critical. Coelho and Costa (2010) describe the challenges in integrating seasonal forecasts into application in areas such as agriculture and health. The first two challenges that they identify are adequate skill in seasonal forecasts of information that is useful to application models, and downscaling this modelling spatially and temporally. This paper highlights the skill in forecasting extremely high temperatures over southern Africa on a seasonal timescale.

2. Methodology

2.1 Data

2.1.1 Temperature data

The temperature data have been obtained in three-month datasets of 2 m minimum and maximum temperatures from the Climatic Research Unit (CRU) TS3.1 (Mitchell and Jones, 2005; Harris *et al.*, 2013). These are on high-resolution 0.5˚ X 0.5˚ grids and allow for the comparison of variations in climate with variations in other phenomena (Harris *et al.*, 2013). The seasons of interest for which data have been obtained were the six run-on seasons September-October-November (SON), October-November-December (OND), November-December-January (NDJ), December-January-February (DJF), January-February-March (JFM) and February-March-April (FMA), which includes the period when southern Africa is mainly controlled by influences from the tropics therefore being a reasonably high predictability time and hence ideal for seasonal predictability studies over the region (Landman and Beraki, 2012; Landman *et al*., 2012). The data were extracted in a format which is compatible with the Climate Predictability Tool (CPT). CRU TS3.1 minimum and maximum temperature data are available from January 1901 to June 2009, however were only extracted from 1982/1983-2008/2009 due to the global models data availability restrictions.

2.1.2 Atmospheric general circulation model data

Comparisons are drawn between the statistically downscaled seasonal forecasts of an atmospheric general circulation model (AGCM) versus the statistically downscaled forecasts from a fully coupled system. The ECHAM4.5 (Roeckner *et al*., 1996) AGCM is used for both the coupled and two-tiered systems. The AGCM hindcast set is available from January 1957 to November 2012 and is obtained by forcing the model with SST anomalies that are produced using constructed analogue SST's (Van den Dool, 2007) and consists of 24 ensemble members. Three-month data of 2 m temperatures and the 850 hPa geopotential height fields have been extracted for the same six run-on seasons SON, OND, NDJ, DJF, JFM and FMA with three-month lead-times for each season for the time period from 1982/1983 until 2008/2009 supplying 27 years of available hindcast data. The largest number of years possible were extracted, as longer records of archived data improve the chance of creating more robust empirical downscaling equations as opposed to shorter ones (Landman and Beraki, 2012). These data are acquired from the archive data library of the IRI in a format compatible with the CPT and seasonal means of the model data are used in this study. For this two-tiered system, forecasts are produced near the beginning of the month, therefore for a one month lead-time there are roughly three weeks from dissemination of the forecast, to the start of the forecasting season. For example a one month lead-time forecast for DJF is produced in the first week of November and for a two month lead-time forecasts are produced beginning October and so forth.

2.1.3 Coupled ocean-atmosphere model data

The ocean model, which is the Geophysical Fluid Dynamics Laboratory (GFDL) Modular Ocean Model version three (MOM3) (Pacanowski and Griffies, 1998), is directly coupled to the ECHAM4.5 (DeWitt, 2005) using the Ocean Atmosphere Sea Ice Soil (OASIS) coupling software (Terray et al. 1999) provided by the European Center for Research and Advanced Training in Scientific Computation (CERFACS). The model consists of 12 ensemble members. Further explanation of the coupled models configuration can be found in DeWitt (2005). The ocean-atmosphere model used in this research is therefore the ECHAM4.5-MOM3 and data for this fully coupled forecasting system are only available from 1982 until July 2012. The model data obtained are the 2 m temperatures and the 850 hPa geopotential height field data. Three-month seasonal averages have been extracted for the six run-on seasons SON, OND, NDJ, DJF, JFM and FMA with three-month lead-times for each season from 1982/1983 to 2008/2009. These data were acquired from the archive data library of the IRI and were also extracted in a format compatible with the CPT. Seasonal means are extracted for this one-tiered system with forecasts also being produced near the beginning of the month; therefore the same process applies as for the two-tiered forecasting system.

2.2 Methodology

The hindcasts from both the one-tiered and two-tiered forecasting systems are statistically downscaled to southern Africa seasonal minimum and maximum temperatures for the six run-on seasons SON, OND, NDJ, DJF, JFM and FMA. The forecasts are downscaled from an approximate 2.5˚ X 2.5˚ resolution to a 0.5˚ X 0.5˚ resolution. Due to the coarse spatial resolution of coupled models (Palmer *et al*., 2004) downscaling global model output to a higher resolution is essential to fulfil the needs of end-users and to also further improve upon the forecasts where possible (Landman and Goddard, 2002) by fixing systematic deficiencies found in the global models. To address this requirement, MOS equations (Wilks, 2006) are employed to adjust for any deficiencies found within the global models directly in the regression equations (Wilks, 2006; Landman and Beraki, 2012; Landman and Goddard, 2002). The technique used to overcome these model errors is by the MOS algorithm using predictor values from the global models in both the development and forecast stages (Landman *et al*., 2012).

Choosing the appropriate predictor variable or model field is very important and requires much attention (Landman *et al*., 2012). Raw model forecast of temperatures that are influenced by topography and other external factors such as soil moisture are poorly resolved and therefore may not be a good predictor of seasonal temperatures at ground level. Temperature fields may be complex and contain structures on spatial scales well below those resolved by models. However, variables such as large-scale circulation may be better simulated by models than temperature, and should therefore be considered instead in a MOS system to predict seasonal temperatures. In this study 2 m temperatures were initially used as the predictor fields, however it was found that using the large-scale circulation pattern in the form of 850 hPa geopotential height fields, proved a better predictor and produced more skill in capturing these seasonal minimum and maximum temperatures.

Canonical correlation analysis (CCA) is an option of the CPT whereby the MOS equations are established (Barnston and Smith, 1996). This tool was developed at the International Research Institute for Climate and Society (IRI: http://iri.columbia.edu). The forecast fields (predictors) for each of the global models that were used in the MOS are confined over a domain that covers the area between the equator and 40˚S, and 15˚W to 60˚E. The minimum and maximum temperature data (predictand fields) cover a smaller domain of 12˚S to 35˚S and 11˚W to 41˚E. The predictor fields cover a larger area than the predictand fields so that any surrounding large-scale circulation patterns that could potentially influence the smaller domain are included. The MOS process begins by performing empirical orthogonal function (EOF) analysis on both the predictor and predictand fields; in this case the model forecast fields (850 hPa geopotential height fields) and the CRU temperature data respectively, followed by the CCA (Landman and Beraki, 2012). A choice is made as to how many EOF and CCA modes are retained by cross-validation skill sensitivity tests using the CPT's CCA tool (Landman and Beraki, 2012). In this study four CCA modes were chosen to be retained.

Verification is the final stage of analysis and is executed over a test period that is completely distinct from the training period. Forecast skill that may be artificially overstated is kept to a minimum via this method and includes evaluations of the predictions against their observations that exclude any information following the forecast year (Landman *et al*., 2012). A true operational forecasting environment is emulated where there is no information available of the approaching season (Landman *et al*., 2012). This process is known as retroactive forecasting (e.g. Landman and Beraki, 2012). The process utilized is similar to the one explained in the Landman *et al* (2012) paper, however, instead of predicting rainfall, temperatures (minimum and maximum) are used. Using the DJF maximum temperatures as an example here, the models are first trained with information from 1982/83 up to 1994/95, providing 12 years of trained data. The maximum DJF temperatures for the following year, 1995/96, are predicted using these trained models. Then the MOS equations are re-trained using information up to and including 1995/96 in order to predict the 1996/97 maximum temperatures. This process is repeated up until 2008/09, therefore resulting in 15 years of independent forecast data.

The skill of the six run-on seasons of minimum and maximum temperature forecasts are evaluated by placing the observed and predicted fields into three categories being defined as above-normal, near-normal and below-normal (e.g., Landman *et al*., 2012). These three categories are not equally divided due to the above- and below-normal threshold values, representing the 75th and 25th percentile values of the climatological record assessed in this study (i.e., 27 years climatology) respectively (Landman *et al*., 2012). This tests the models ability to predict extremes in seasonal minimum and maximum temperatures.

Due to seasonal climate having an inherently probabilistic nature, it therefore needs to be judged probabilistically. The key attributes of probabilistic forecasts include the reliability: defined by the confidence level communicated and if appropriate and if there are systematic biases (Landman *et al*., 2012), the resolution: showing if any information is useful, discrimination: indicates if events are distinguished from non-events, and lastly sharpness: accounts for the confidence level of the forecasts (Troccoli *et al*., 2008; Wilks 2006). Forecast verification in this study is evaluated by using the relative operating characteristic (ROC) (Mason and Graham, 2002) and the reliability diagram (Hamill, 1997, Wilks, 2006). When a high seasonal minimum or maximum temperature event occurred opposed to when it did not occur the ROC scores will indicate a higher probability by means of a higher score. Therefore ROC detects whether a set of forecasts has the attribute of discrimination (Mason and Graham, 2002). If there is consistency between predicted probabilities and observed frequencies of an event, the forecasts would be accepted as reliable (Hamill, 1997; Wilks, 2006).

3. Results

3.1 Retro-active forecast verification

ROC is the ability of a forecasting system to discriminate events from non-events, and ROC graphs are created by plotting the forecast hit rates against the forecast false alarm rates (Wilks, 2006). The ROC score is represented by the area beneath the ROC curve and is used as a gauge of discrimination between the events versus the non-events. If the area beneath the ROC curve (i.e., ROC score) is ≤ 0.5 the forecasts would be classified as having no skill but above 0.5 there would be increasing skill to the perfect discrimination of a ROC score of 1.0 (Landman *et al*., 2012). The ROC score here can be interpreted as the probability of the forecasting system successfully discriminating above- or below-normal seasons from other seasons.

Figures 1a and 1b are graphical representations of the ROC scores for the above- and below-normal categories for the five run-on seasons OND, NDJ, DJF, JFM and FMA of the maximum and minimum temperatures for both forecasting systems as calculated for the 15-year retroactive test period respectively. Only five out of the six seasons investigated were taken into consideration due to the fact that the spring season (SON) indicated very little to no skill in capturing both minimum and maximum temperatures and therefore was omitted.

Figure 1a shows encouraging results in terms of the skill captured when predicting seasonal maximum temperature extremes. The coupled model or one-tiered forecasting system produces the highest overall skill in predicting seasonal maximum temperature extremes shown by the highest amount of

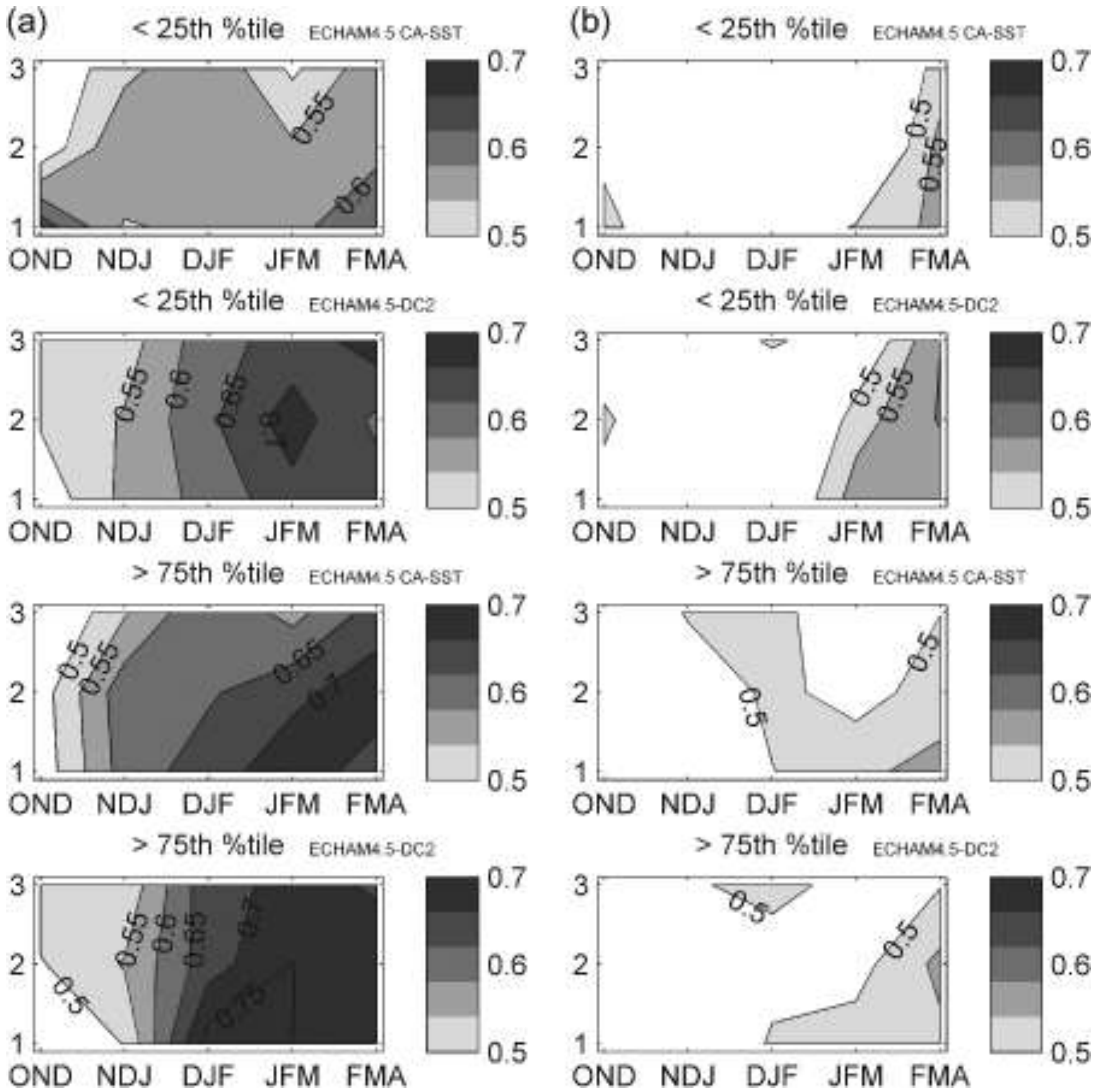

Figure 1: (a) ROC scores of maximum temperatures of above (75^{th}) and below (25^{th}) normal percentiles for both the one- and two-tiered systems for the seasons OND, NDJ, DJF, JFM and FMA with lead-times up to three months. 1 (b) Same as (a) but for the minimum temperatures.

dark grey and black in the second and fourth images in Figure 1a. Moreover, the seasons that showed the highest ROC values are DJF, JFM and FMA, which coincide with the period of highest seasonal temperatures normally found over southern Africa. The above-normal category of the one-tiered system indicates the highest overall skill, which implies that the coupled model is able to skilfully predict when there is a high likelihood of experiencing extremely high seasonal maximum temperatures during mid-summer. The coupled system is also useful in predicting the likelihood of experiencing extremely high maximum temperatures during the second half of the summer season, i.e., January to March. It was also interesting to note that even though skill was the highest at a lead-time of one month, even up to three-month lead-times skill was found to be high particularly for late summer when using the coupled model.

Figure 1b shows that skill in discriminating seasonal minimum temperature extremes from non-extremes is only found from mid-summer onwards where the ROC scores reached above the 0.5 ROC score threshold. Neither the one- or two-tiered forecasting system definitively outperformed the other in capturing these minimum temperature extremes, with both above- and below-normal categories performing equally poorly.

Both the one- and two-tiered forecasting systems have been considered in predicting seasonal minimum and maximum temperature extremes, however it has been shown that the coupled model performs best at predicting seasonal maximum temperature extremes for mid and late summer, and therefore has the most potential to support decisions to be made by the health sector. It is for these reasons above that this paper will focus on only the coupled model forecasts, as well as only the prediction of seasonal maximum temperature extremes from hereon in.

The ROC score is sometimes criticized as a measure of forecasting performance due to its insensitivity to reliability (Troccoli *et al*., 2008), hence the inclusion of reliability diagrams here. Reliability diagrams are created by plotting the observed relative frequencies against the forecast probabilities and therefore provide the attribute of consistency between the two (Troccoli *et al*., 2008). The diagonal lines in Figure 2 indicate perfect reliability between the observations and forecast probabilities. These reliability diagrams also indicate whether over-forecasting or under-forecasting occurs. When the reliability graph (solid black line: above-normal or dotted black line: below-normal) lies above the diagonal, under-forecasting occurs due to observations of the specific event occurring more frequently than the event being forecast. When the reliability graph lies below the diagonal, over-forecasting occurs due to the forecast probability exceeding the observations.

Figure 2 shows three reliability diagrams with their accompanying frequency histograms of the coupled model for the seasons DJF, JFM and FMA at a lead-time of two months. The DJF maximum temperature reliability diagram shows good reliability for the above-normal category, as the solid black line follows the diagonal fairly consistently but does drop below the diagonal at around the 60% forecast probability threshold and there tends to be a lot of over forecasting occurring for the higher forecast probabilities. The JFM and FMA reliability diagrams both indicate very strong reliability for the above-normal categories as the solid black lines almost follow the diagonal exactly but do tend to over-forecast towards the very high forecast probabilities. There is less consistency in capturing the below-normal seasonal maximum temperature extremes for all three images, with over-forecasting dominating as the dotted reliability graph lies below the diagonal more often than not.

The JFM season shows the highest reliability overall, indicating that there is a high level of consistency in capturing seasonal maximum temperature extremes for this late summer season. These results correspond well to the high ROC scores found in this season over southern Africa. These results have shown that the coupled model has skill for certain configurations and certain seasons i.e. JFM, but how can one tell how economically feasible these forecasts are and if it is possible for a policymaker or health practitioner to understand and measure their value? Such questions can be

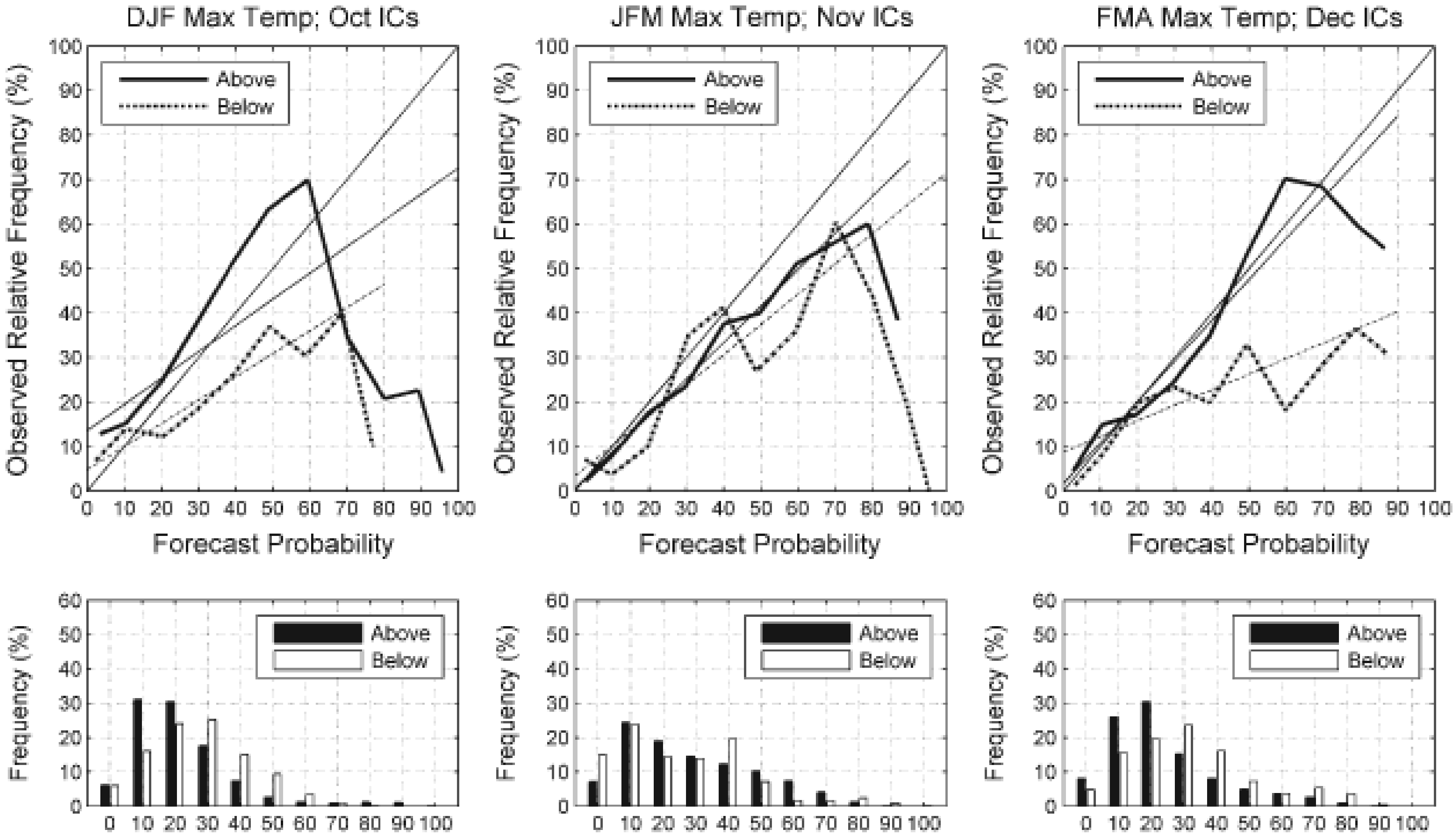

Figure 2: Reliability diagrams and frequency histograms of seasonal maximum temperature extremes for above (75th percentile) and below (25th percentile) categories for the one-tiered system for DJF, JFM and FMA with two month lead-times i.e. DJF maximum temperatures are forecasting using initial conditions (ICs) from the beginning of October. The thick (dotted) black curve and the black (white) bars represent the above (below) category. The thin black lines represent the least squares regression line in the reliability diagrams.

addressed through the following two skill scores; the cumulative profits score and the two-alternative forced choice (2AFC) score.

3.2 Forecast performance for decision makers

The cumulative profit analysis of the coupled model hindcasts presented here may be used to communicate the monetary value of forecasts. This analysis is described in detail in Hagedorn and Smith (2008) and can be summarised as follows. A cumulative profit scoring parameter is used to ensure that the focus is not solely on the general skill of the forecast, but also represents various aspects of the potential economic value of such a forecasting system. The cumulative profits parameter provides an intuitive way to communicate the skill of probabilistic forecasts to both experts and more particularly to non-experts. It is of upmost importance that research completed by scientists is presented in such a way that end users including the general public may make use of the findings effectively. This parameter evaluates probabilistic forecasts by means of quantifying the skill of the forecast using an effective daily interest rate. How it works is that one invests capital into a specific probabilistic forecast or forecasting system and depending on the outcome of how well the forecast performs one will obtain a return on the investment. Therefore the higher the amount of capital placed on a forecast that is correct, the higher the profit or return will be due to the influence of the effective daily interest rate, indicating to customers, such as a forecast user, that the forecast is worth the money spent on it.

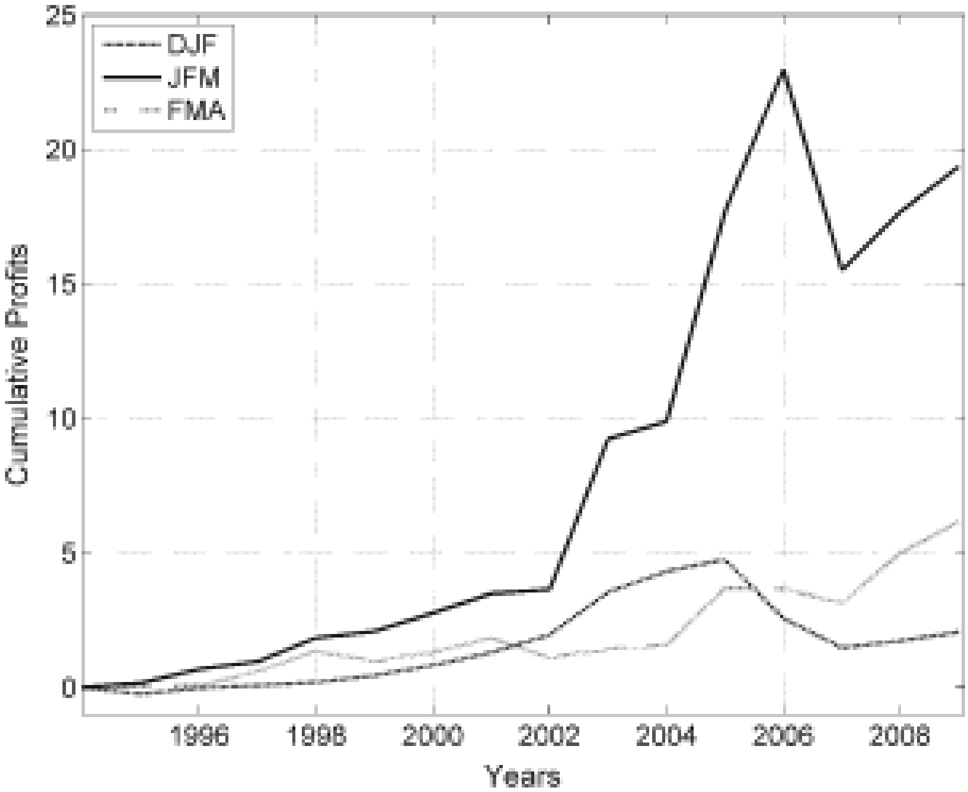

Figure 3: Cumulative Profits Graph of maximum temperatures for the one-tiered system showing the mid to late summer seasons DJF, JFM and FMA from 1995 to 2009.

Figure 3 represents a graph showing the three run-on summer seasons cumulative profits from 1995 to 2009. From the image above one can clearly note that JFM is the season with the highest cumulative profit percentage for a forecast user (around 20%) towards the end of the test period, which means that if one was to invest in the forecast of seasonal maximum temperature extremes in JFM over southern Africa, the return on the investment would be the highest, indicating to a non-expert that the skill of the forecast is good in JFM. A possible reason for the dramatic increase from 2002 for JFM could be due to improved seasonal forecasts that occurred during this period and then in 2007 a slight drop in cumulative profits due to an underestimation of an extremely warm season (cf. Figure 5), therefore the cumulative profits parameter is directly affected by the accuracy of the seasonal forecast. This provides non-experts with more understanding and confidence in the JFM seasonal maximum temperature extreme forecasts by representing the monetary value of the forecasting systems skill. DJF and FMA however do not have such a high return on investments resulting in profit, but have smaller cumulative profit percentages that reach approximately 5% from 2004 and onwards. It should be noted that, however, the return is still positive.

The final score analysed for these forecasts is a verification score known as the two-alternative forced choice (2AFC; Mason and Weigel, 2009) test. This particular score was chosen due to its usefulness in terms of its administrative purpose. The 2AFC score provides an indication of a forecasts quality to the general public as well as communicating or transferring changes in forecast quality to officials. Therefore this score is a very useful and informative score to a variety of stakeholders and not only atmospheric scientists.

To calculate this 2AFC score, one starts off with a set of two forecast-observation pairs and then determines if the forecasts can be used to discriminate between the observations. An example would be if there were two seasons, one where the temperature was warmer than the other, whether the warmer season could be identified through the forecasting result. The main objective of this score is to compare all the sets of two forecast-observation pairs available and ask the same question every time, and from that information, calculate the proportion of time the question is correctly answered. The proportion is known as the probability of a correct decision, and the question asked is the 2AFC test. If the observations are assumed to be distinguishable, the chance of choosing the correct forecast season, provided unskilful forecasts is 50%; therefore any value above 50% for this score indicates a skilful forecast, as it is better than purely guessing.

In Figure 4 one can see that the highest 2AFC scores are found in DJF and JFM over southern Africa. This is the mid-late summer season over this region and therefore it corresponds well as to when the seasonal maximum temperate extremes will be at their highest. The north-eastern regions are the areas with the highest 2AFC scores which means that the highest amount of discrimination can be found over these regions. There are a large percentage of the forecasts being correctly discriminated over the Gauteng Province and adjacent regions, which is one of the most highly populated areas of southern Africa. This essentially means that the general public can have confidence in the forecasts over this region for seasonal maximum temperature extremes, specifically in DJF and JFM. The south-western parts of southern Africa tend to have less useful forecasts of seasonal maximum temperature extremes as produced by the forecast system described here.

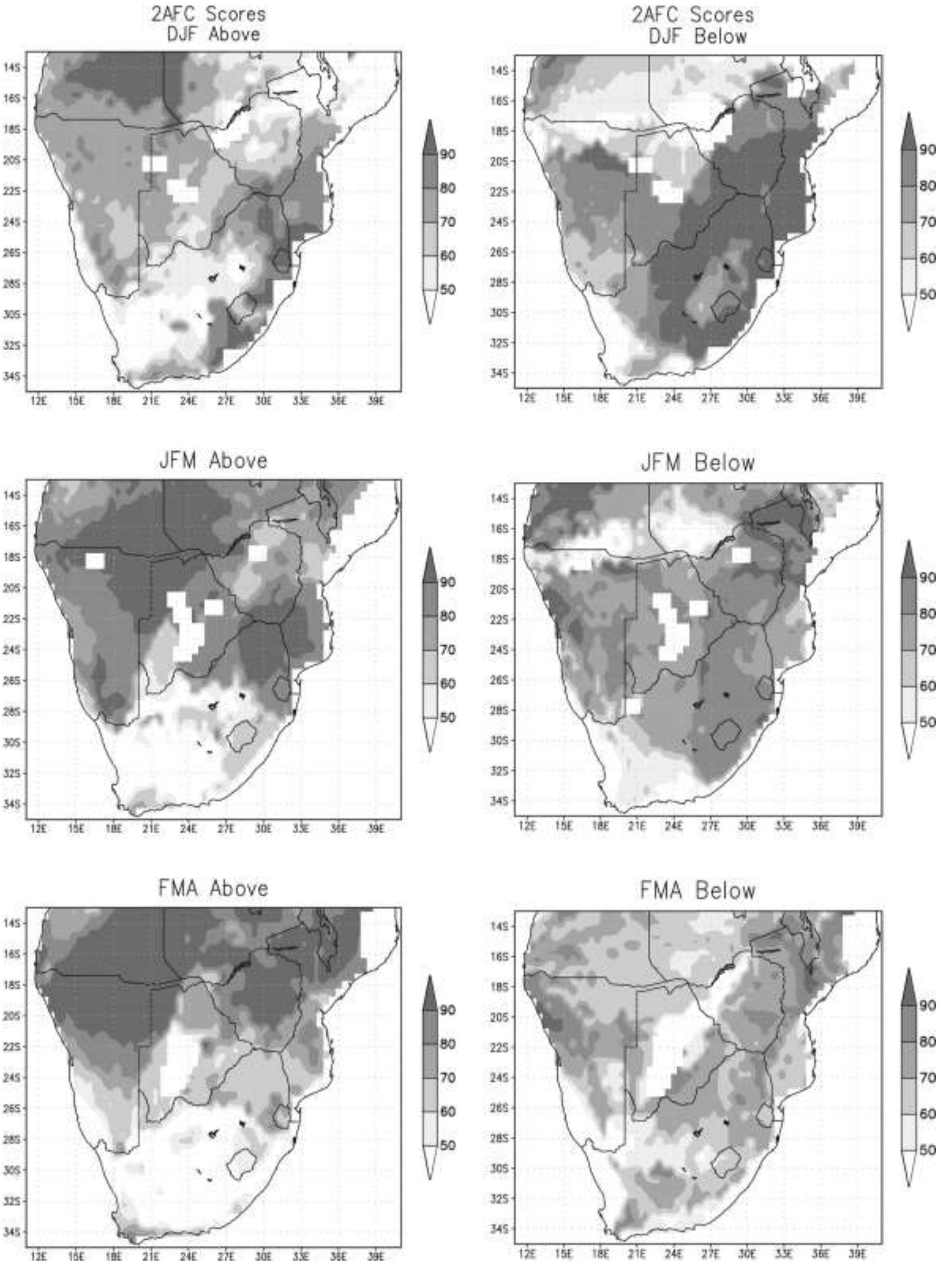

Figure 4: 2AFC Scores of maximum temperatures of above (75th) and below (25th) normal percentiles for the one-tiered system showing the mid to late summer seasons DJF, JFM and FMA from 1995 to 2009.

To determine whether the coupled maximum temperature system is a practicable one, it would make sense to focus in on a large metropolis such as Pretoria to assess the predictions that have been made over a recent 15 year period, which would be a good indication of the actual forecast if it were issued operationally. The reason for choosing this particular location is due to the fact that it is a large urban centre with a large population of approximately three million people and has the potential to be adversely affected by these maximum temperature extremes. Also this area falls within the high predictability region and therefore forecasts can be made with more confidence over this region. The exact location of Pretoria is 25.7256° S, 28.2439° E.

In Figure 5a, which is a retro-active deterministic forecast of maximum temperatures for Pretoria, the observations follow the forecasts relatively closely as well as remaining within the upper and lower one standard deviation confidence levels for 12 out of the 15 cases, therefore the majority of the time. The most obvious outliers of the forecast compared to the observations in Figure 5a were in 2000 and 2007 when the observations were much smaller and larger than the forecasts respectively. However if one considers figure 5b, the probabilistic maximum temperature hindcasts for Pretoria, one can see that in 2000 there was an enhanced probability that extremely low maximum temperatures were likely for that year, which means that the probabilistic forecast was able to capture that lower maximum temperature possibility better than the deterministic forecast. The same applies for 2007 but just that the probabilistic forecast captured the higher probability of an extremely high maximum temperature for that JFM season. By having knowledge of the probability of a predicted category occurring, additional forecast value is acquired (Mason and Graham, 1999), due to probabilistic forecasts exhibiting reliability considerably in excess of that achieved by corresponding deterministic forecasts (Murphy, 1998). If one considers the 1999 forecast it was found that a lower JFM maximum temperature was predicted than actually occurred. In fact 1999 was the third hottest season over the test period, even though it was a La Niña year. During La Niña years, southern Africa tends to be cooler and wetter than usual, which was reflected in the temperature forecast presented here as well as in a rainfall forecast previously determined (Landman *et al*., 2012), as most seasonal forecast models are significantly swayed by ENSO’s influence (Landman and Beraki, 2012).

Figures 5a and 5b indicate that the deterministic forecast captured the nature of the seasonal maximum extreme better than the probabilistic forecast in only a limited number of cases (e.g. 1995, 2002, 2003, and 2004). If one considers how maximum temperature prediction for a large metropolis such as Pretoria would have been produced over a recent 15 year period, one could say from the results presented that it would have been a good idea for decision makers to make better use of probabilistic forecasts as opposed to deterministic forecasts.

4. Discussion and Conclusions

When considering the seasonal maximum temperature extremes, a large amount of skill is found in capturing these extremes, specifically when using the coupled model. The one-month lead-time indicates the highest overall skill, however when using the coupled model, predictions of seasonal maximum temperature extremes are possible up to three months lead-time with almost the same amount of skill as at a lead-time of one month. Therefore forecasts can be skilfully made approximately three months in advance when using the fully coupled model for seasonal maximum

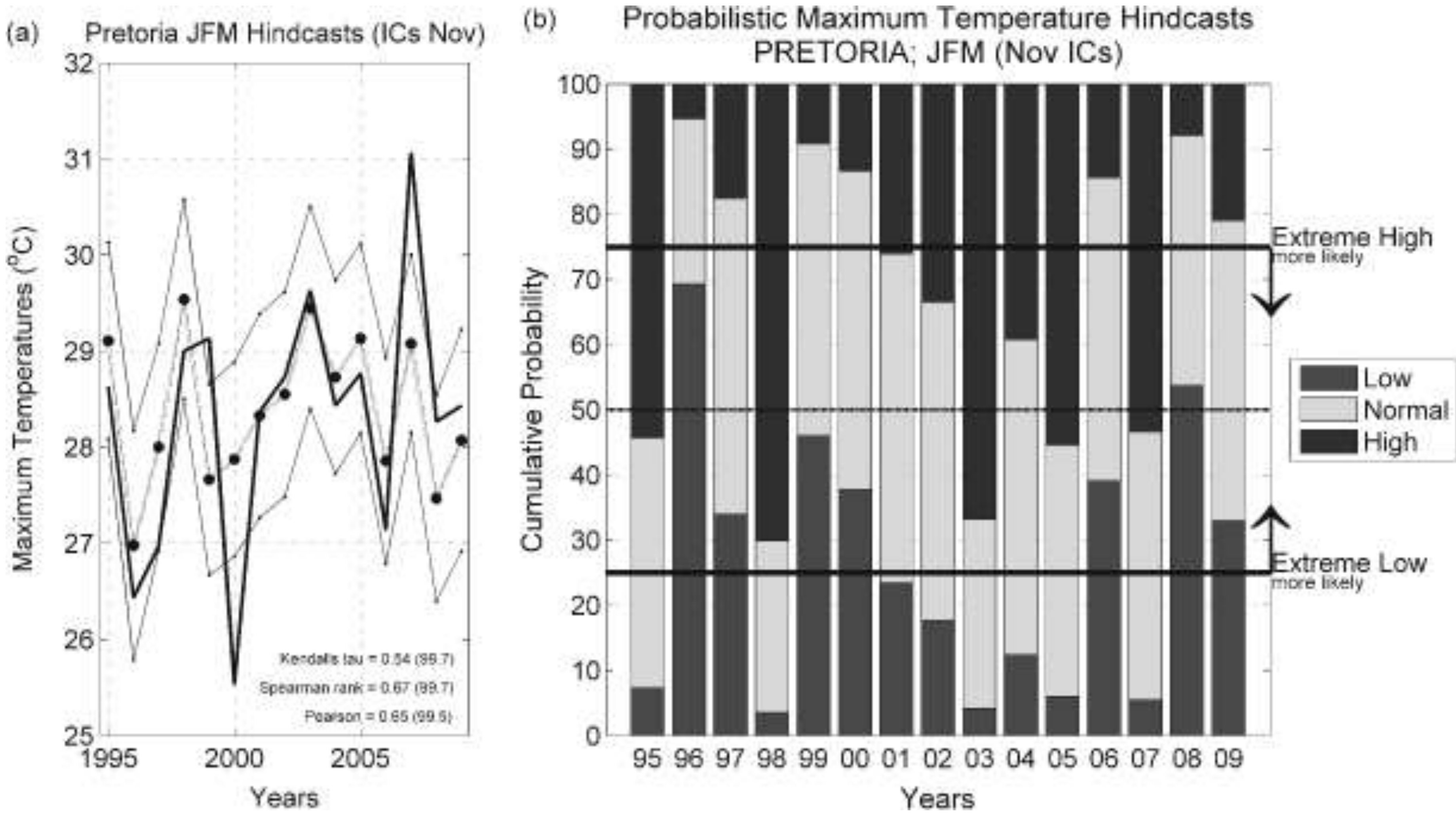

Figure 5: (a) Line graph of Pretoria's observed and forecast maximum temperatures for the one-tiered system for JFM with a two month lead-time. Observations (solid dark black line), forecasts (grey lines with large black dots) and upper and lower one standard deviation confidence levels (thin black line with small black dots). 5(b) Bar graph of Pretoria's cumulative probability probabilistic maximum temperature hindcasts using the couple model for JFM with a two month lead-time. Both images are for the time period 1995 to 2009.

temperatures. The above-normal category shows the highest overall skill, implying that the forecast system is able to capture when there was a high likelihood of experiencing high seasonal maximum temperatures for that particular season. The coupled model proves useful in predicting the likelihood of experiencing extremely high seasonal maximum temperatures during the second half of summer from around January to March with spring and winter indicating very little to no skill. The seasons with the highest overall skill are those of DJF, JFM and FMA, which coincide with the period of the highest peak of the seasonal temperature annual cycle over southern Africa. The seasonal minimum temperatures were found to have very little skill and were therefore excluded for further analysis.

The coupled model also proves to be the most reliable in predicting seasonal maximum temperature extremes over southern Africa with the above-normal category producing the largest amount of confidence, again implying that when using the coupled model it can be said with a large amount of confidence when there will be a high likelihood of experiencing extremely high seasonal maximum temperatures for a particular season. The most reliable season was JFM with a close to perfect reliability at a two month lead-time. Again JFM coincides with the highest seasonal maximum temperatures found over southern Africa. Most of the forecasts did however; tend to over-forecast, especially for the high probability forecasts. There was however an apparent discrepancy between the ROC and 2AFC scores with the below normal category performing better according to the 2AFC score and the above normal category performing better in the ROC scores. The reason for this discrepancy may be due to the fact that the ROC and reliability diagrams are calculated over the entire southern African region, whereas the 2AFC score indicates a geographical distribution of skill over the chosen domain. Therefore making it difficult to directly compare these parameters. However if one focuses on the Pretoria area, the above normal category performs best in both the ROC and 2AFC scores.

Most forecasts are verified using scientific processes and outcomes, whereas less attention is given to scores that may be of use to a non-expert audience such as government officials and the general public (Mason and Weigel, 2009). The scientific verification processes used here were the ROC and reliability diagrams, which provided insight into which forecasting system and which season were most skilful and reliable. In addition to those scores, the cumulative profits and 2AFC scores were analysed, which are more intuitive and provide better understanding of the forecasts to end users. These two scores both indicated that JFM would be the most profitable season to invest in when considering seasonal maximum temperature extremes, which corresponds well to the results found when assessing the ROC and reliability verification parameters.

A further analysis over Pretoria established that there was a large amount of predictability of these seasonal maximum temperatures when using the probabilistic forecast over this location. The impact of this is that Pretoria is part of a large urban centre, where many millions of people reside in houses that are not equipped with air-conditioning, thus making it a community potentially vulnerable to extreme temperatures. This paper has shown that by using a coupled probabilistic forecasting system, these seasonal maximum temperatures which may adversely affect such a city centre, can be predicted well in advance with high skill and confidence, as well as rendering profitable and useful information to non-experts.

The health industry is directly affected by temperature extremes, due to the fact that with increased maximum temperatures there are more associated health problems. Therefore with the skilful prediction of maximum temperature extremes, the health sector may be better prepared for the health risks associated with the increase of maximum temperatures and can therefore be more capable of dealing with these problems such as heat stress and respiratory problems. Thus it is critical for South Africa, and Africa as a whole, to develop Heat Health Watch Warning Systems; this is highlighted as a priority of the South African government in a recently released Climate Change Response White Paper (Government of RSA, 2011).

One key component in developing such a system is the need for forecasting ability of indicators, such as maximum temperature, that can be related to health outcomes. This paper describes the ability to forecast, on a seasonal timescale, maximum temperatures. This is a critical first step in developing and refining the needed forecasting ability. Such forecasts will be useful for alerting the South African government and public health stakeholders to summers that are expected to have above normal temperatures with a 3-month lead time. This will allow for sufficient preparation to be ready to enact heat health alerts and prevention measures. After this, forecasting of critical indicators on a timescale of a few days will be needed in order to activate the alerts and the prevention plans.

In addition to the ability to forecast key indicators, the knowledge of how South Africans respond to heat is needed to determine when an alert should be activated, and information on vulnerability of populations to heat is needed to know where and for which populations the alert should be focused on. The temperature where health is impacted differs across countries and population (McMicheal *et al*., 2008; Basu and Samet 2002; Baccini *et al*., 2008; Medina-Ramón and Schwartz, 2007), as do the best meteorological indicators for predicting health outcomes (Vaneckova *et al*., 2011). For example, in the French system, a combination of forecasted minimum and maximum temperatures is used. This was chosen both for the suitability for forecasting and for the performance of the indicator to determine excess mortality (Pascal *et al*., 2004). Indices that combine multiple meteorological parameters (e.g., temperature, relative humidity and wind speed) have been created in order to describe how hot it feels. One such example is apparent temperature (AT) and has been used by countries to develop heat alerts and has been found to be related to excess mortalities (Steadman, 1979; Smoyer-Tomic and Rainham, 2001; Ballester *et al*., 2011; Watts and Kalkstein, 2004). However, a study in Brisbane, Australia found that average temperature measurements performed similarly to composite indices (such as apparent temperature) in detecting excess mortality days. There is not agreement on what the best meteorological indictor is as it is strongly dependent upon forecasting ability and the population's response to heat.

Thus, before a heat-health plan could be developed in South Africa, research is needed in collecting historical health and meteorological data to investigate the relationship between heat and mortality and morbidity across South Africa. In addition, the relationship needs to be probed across multiple meteorological indicators. The forecasting work in this paper must be continued to investigate, and improve upon where needed, the ability for these various meteorological indicators to be forecasted seamlessly from seasonal to weather timescales. And then, by forecasters and health researchers working collaboratively, health alert and prevention plans can be developed.

In conclusion, one-tiered forecasting systems or coupled models outperform two-tiered systems when predicting seasonal maximum temperature extremes over southern Africa. This paper shows that coupled models, when analysed probabilistically, exhibit skill in capturing maximum temperature extremes over southern Africa for the mid-late summer season. Therefore this modelling contribution demonstrates that it is definitely feasible to direct some of the available research and modelling funds, as well as effort towards the development and implementation of operational seasonal forecasting systems that incorporate fully coupled models, as well as health alert and prevention strategies.

5. Acknowledgements

This work was partly supported financially by the National Research Foundation (NRF) of South Africa, the Applied Centre for Climate and Earth System Science (www.access.ac.za), a CSIR grant and by the Peter Carpenter Scholarship for African Climate Change. The computing time used to produce the retrospective forecasts from IRI was largely provided by grants from the multiagency Climate Simulation Laboratory (CSL) program, with Dave DeWitt as PI.

6. References:

Baccini M, Biggeri A, Accetta G, Kosatsky, Katsouyanni K, Analitis A, Anderson HR, Bisanti L, D'Ippoliti D, Danova J, Forsberg B, Medina S, Paldy A, Rabczenko, Schindler C, Michelozzi P. 2008. Heat effects on mortality in 15 European cities. *Int.J.Epidemiol*. **19**(5): 711-719.

Ballester J, Robine JM, Herrmann FR, Rodó X. 2011. Long-term projections and acclimatization scenarios of temperature-related mortality in Europe. *Nat. Comms.***2**:358.

Barnston AG, Smith TM. 1996. Specification and prediction of global surface temperature and precipitation from global SST using CCA. *J. Climate.* **9:**2660–2697.

Barnston AG, He Y, Glantz MH. 1999. Predictive skill of statistical and dynamical climatemodels in SST forecasts during the 1997-98 El Niño episode and the 1998 La Niña onset. *Bull. Amer. Meteor. Soc*. **80**: 217-244.

Basu R, Samet JM.2002. Relation between elevated ambient temperature and mortality: A review of epidemiologic evidence. *Epidemiol. Rev*. **24**(2): 190-202.

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

Brankovic C, Palmer TN. 2000. Seasonal skill and predictability of ECMWF PROVOST ensembles. *Q. J.Roy.Meteor. Soc.***126**: 2035-2067.

Cane MA, Eshel G, Buckland RW. 1994. Forecasting Zimbabwean maize yield using eastern Pacific sea surface temperatures. *Nature*. **370**: 204–206.

Coelho CAS, Costa SMS. 2010. Challenges for integrating seasonal climate forecasts in user applications. *Current Opinion in Environmental Sustainability*. **2**:317-325.

Collins JM. 2011. Temperature Variability over Africa. *Amer. Meteor. Soc*. DOI: 10.1175/2011JCLI3753.1.

Connor SJ, Thomson MC, Menne BE. 2008.Seasonal climate forecasting for health in MC Thomson *et al*. (eds) *Seasonal forecasts, climatic change and human health*. Springer: pp 232.

Copsey D, Sutton R, Knight JR. 2006. Recent trends in sea level pressure in the Indian Ocean region. *Geophys. Res. Lett.***33**. L19712, DOI:10.1029/2006GL027175.

DeWitt DG. 2005. Retrospective forecasts of interannual sea surface temperature circulation model. *Mon. Wea. Rev.***133**: 2972-2995.

Diaz J, Jordán A, Garcia R, López C, Alberdi JC, Hernández E, Otero A. 2002.Heat waves in Madrid 1986-1997: effects on the health of the elderly. *Int.Arch.Occup. Environ. Health*.**75**: 163-170.

Doblas-Reyes F.J, Deque M, Piedelievre JP. 2000. Multi-model spread and probabilistic seasonal forecasts in PROVOST. *Q. J. Roy. Meteor. Soc.***126**: 2069-2087.

Ebi KL, Teisberg TJ, Kalkstein LS, Robinson L,Weiher RF. 2004. Heat watch/warning systems save lives. *Bull. Amer. Meteor. Soc*. 1067-1073.

Engelbrech F, Bopape MJ. 2011. High resolution projected climate futures for southern Africa. *27th Annual conference of South African Society for Atmospheric Sciences: the Interdependent Atmosphere, Land and Ocean, 22-23 September 2011*, Hartbeespoort, North-West Province, South Africa.

Fouillet A, Rey G, Wagner V, Laaidi K, Empereur-Bissonnet P, Le Tertre A, Frayssinet P, Bessemoulin P, Laurent F, De Crouy-Chanel P, JouglaE,Hémon D. 2008. Has the impact of heat waves on mortality changes in France since the European heat wave of summer 2003? A study of 2006 heat wave. *Int.J.Epidemiol*. **37**(2): 309-317.

Goddard L, Mason SJ, Zebiak SE, Ropelewski CF, Basher R, Cane MA. 2001. Current approaches to seasonal to interannual climate predictions. *Int J Climatol.* **21**: 1111–1152

Goddard, L, Mason SJ. 2002. Sensitivity of seasonal climate forecasts to persisted SST anomalies. *Clim. Dynam.***19**. 619-631.

Graham RJ, Evans ADL, Mylne KR, Harrison MSJ, Robertson KB. 2000. An assessment of seasonal predictability using atmospheric general circulation models. *Q. J.Roy. Meteor.Soc*.**126**: 2211-2240.

Green RS, BasuR, Malig B, Broadwin R, Kim JJ, Ostro B. 2010. The effect of temperature on hospital admissions in nine California counties. *Int. J. Public Health*.**55**:113-121.

Hadedorn R, Smith LA. 2009. Communicating the value of probabilistic forecasts with weather roulette. *Meteorol. Appl.***16** (2): 143-155.

Hamill TM. 1997. Reliability diagram for multicategory probabilistic forecasts. *Wea. Forecasting*.**12**: 736-741.

Harris, I., Jones, P.D., Osborn, T.J., and Lister, D.H., 2013. Updated high-resolution grids of monthly climatic observations. In press, Int. J. Climatol., Doi: 10.1002/joc.3711

Hastenrath, S., L. Greischar, and J. van Heerden, 1995: Prediction of summer rainfall over South Africa. *J. Climate*, **8:**1511–1518.

Hunt BG. 1997. Prospects and problems for multi-seasonal predictions. Some issues arising from a study of 1992. *Int. J. Climatol.* **17:**137–154.

Joubert AM, Hewitson. 1997. Simulating present and future climates of southern Africausing general circulation models. *Progress in Phys. Geo.* **21**:51. DOI:10.1177/030913339702100104.

Jury MR. 1996. Regional teleconnection patterns associated with summer rainfall over South Africa, Namibia and Botswana. *Int.J. Climatol.,* **16:**135–153.

Jury H, Mulenga M, Mason SJ. 1999. Exploratory long-range models to estimate summer climate variability over southern Africa. *J. Climate.***12:**1892–1899.

Kalkstein LS, Jamason PF, Greene S, Libby J, Robinson L. 1996. The Philadelphia Hot Weather-Health Watch/Warning System: Development and Application, Summer 1995. *Bull. Amer. Meteor. Soc*. **77**(7):1519-1528.

Kelly-Hope L, Thomson MC. 2008. Climate and infectious disease in MC Thomson *et al*. (eds) *Seasonal forecasts, climatic change and human health*. Springer: pp 232.

Klopper E, Landman WA, Van Heerden J. 1998. The predictability of seasonal maximum temperature in South Africa. *Int. J.Climatol*.**18**: 741-758.

Kovats RS, Hajat S, Wilkinson P. 2004. Contrasting patterns of mortality and hospital admissions during hot weather and heat waves in Greater London, UK. *Occup. Environ. Med*.**61**: 893-898.

Kovats RS, Hajat S. 2008. Heat stress and public health: A critical review. *The Annual Review of Public Health*.**29**:9.1-9.15.

Kuhn K, Campbell-Lendrum D, Haines A, Cox J. 2005. Using climate to predict infectious disease epidemics. *WHO*: Geneva. pp 54.

Kynast-Wolf G, Preuss M, Sié A, Kouyaté B, Becher H. 2010. Seasonal patterns of cardiovascular disease mortality of adults in Burkino Faso, West Africa. *Trop. Med. Int. Health*. **15**(9): 1082-1089.

Landman, WA, Mason SJ. 1999a. Change in the association between Indian Ocean sea-surface temperatures and summer rainfall over South Africa and Namibia. *Int. J. Climatol.***19:** 1477–1492.

Landman, WA, Tennant WJ. 2000. Statistical downscaling of monthlyforecasts. *Int. J. Climatol.***20:**1521–1532.

Landman WA, Mason SJ, Tyson PD, Tennant WJ. 2001a. Retroactive skill of multi-tiered forecasts of summer rainfall over southern Africa. *Int. J. Climatol.* **21:**1–19.

Landman WA, Beraki A. 2012. Multi-model forecast skill for midsummer rainfall over southern Africa. *Int. J. Climatol*. DOI: 10.1002/joc.2273.

Landman WA, DeWitt D, Lee DE, Beraki A, Lötter D. 2012. Seasonal Rainfall Prediction Skill over South Africa: One- versus Two-Tiered Forecasting Systems. *Wea. Forecasting*. **27**: 489–501. DOI: http://dx.doi.org/10.1175/WAF-D-11-00078.1.

Landman WA, Goddard L. 2002. Statistical recalibration of GCM forecasts over southern Africa using model output statistics. *J. Climate.***15**: 2038-2055.

Landman WA, Mason SJ, Tyson PD, Tennant WJ. 2001. Retro-active skill of multi-tiered forecasts of summer rainfall over southern Africa. . *Int. J. Climatol.* **21**: 1-19.

Landman WA, MJ Kgatuke, M Mbedzi, A Beraki, A Bartman and A Du Piesanie. 2009b. Performance comparison of some dynamical and empirical downscaling methods for South Africa from a seasonal climate modelling perspective. *Int. J. Climatol*.**29**: 1535–1549.

Latif M and Coauthors. 2001. ENSIP: The El Nino simulation intercomparison project. *Climate Dyn*.**18**: 255-276.

Makarau A, Jury MR. 1997. Predictability of Zimbabwe summer rainfall. *Int. J. Climatol*. **17:**1421–1432.

Mason AM, Joubert C, Cosijn SJ, Crimp. 1996. Review ofseasonal forecasting techniques and their applicability of southernAfrica. *Water SA.***22:**203–209.

Mason SJ, Joubert AM. 1997. Simulated changes in extreme rainfall over southern Africa. *Roy. Meteor. Soc*. Vol**17.** Issue 3.pp 291-301.

Mason SJ. 1998. Seasonal forecasting of South African rainfall using a non-linear discriminant analysis model. *Int. J. Climatol.***18:** 147–164.

Mason SJ, Goddard L, Graham NE, Yulaeva E, Sun, L, Arkin, PA. 1999. The IRI seasonal climate prediction system and the 1997/98 El Niño event. *Bull. Amer. Meteor. Soc*. **80**(9): 1853-1874

Mason SJ, Graham NE. 1999. Conditional probabilities, relative operating characteristics, and relative operating levels. *Wea. Forecasting.***14**: 713–725.

Mason SJ, Graham NE. 2002. Areas beneath the relative operating characteristics (ROC) and levels (ROL) curves: Statistical significance and interpretation. *Q. J. Roy. Meteor. Soc*. **128**: 2145-2166.

Mason SJ, Weigel AP. 2009. A Generic Forecast Verification Framework for Administrative Purposes. *Mon. Wea. Rev.***137**. 331–349. doi: http://dx.doi.org/10.1175/2008MWR2553.1.

Mathee A, Oba J, Rose A. 2010. Climate change impacts of working people (the HOTHAPS initiative): findings of the South African pilot study. *Global Health Action***. 3**.

Mattes M, Mason SJ. 1998. Evaluation of seasonal forecasting for Namibian rainfall. *S. Afr. J. Sci.***94:**183–185.

McGregor GR, Watkin HA, Cox M. 2004. Relationships between the seasonality of temperature and ischaemic heart disease mortality: implications for climate based health forecasting. *Climate Res.* **25**:253-263.

McGregor GR, Cox M, Cui Y, Cui Z, Davey MK, Grahan RF, Brookshaw A. 2006. Winter-season climate prediction for the UK health sector. *J. Appl. Meteor.Climatol*. **45**:1782-1792.

McMichael AJ, Wilkinson P, Kovats RS, Pattenden S, Hajat S, Armstrong B, Vajanapoom N, Niciu EM, Mahomed H, Kingkeow C, Kosnik M, O'Neill MS, Romieu I, Ramirez-Aguilar, Barreto ML, Gouvela N, Nikiforov B. 2008. International study of temperature, heat and urban mortality: The 'ISOTHURM' project. *Int.J.Epidemiol.* **37**:1121-1131.

Medina-Ramón M, Schwartz J. 2007. Temperature, temperature extremes, and mortality: a study of acclimatisation and effect modification in 50 US cities. *Occup. Environ. Med*. **64**: 827-833.

Michelozzi P, Accetta G, De Sario M, D'Ippoliti D, Marino C, Baccini M, Biggeri A, Anderson HR, Katsouyanni K, Ballester F, Bisanti L, Cadum E, Forsberg B, Forastiere F, Goodman PG, Hojs A, KirchmayerU, Medina S, Paldy A, Schindler C, SunyerJ,Perucci CA. 2009. High temperature and hospitalizations for cardiovascular and respiratory causes in 12 European cities. *Amer.J.Repir.Crit. Care. Med*. **179**: 383-389.

Mitchell T, Jones PD. 2005. An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *Int. J. Climatol*. **25**: 693–712. DOI: 10.1002/joc.1181.

Murphy AH. 1998. The early history of probability forecasts: Some extensions and clarifications. *Wea. Forecasting.* **13:**5–15.

Myers MF, Rogers DJ, Cox J, Flahault A, Hay SI. 2000. Forecasting disease risk for increased epidemic preparedness in public health. *Adv.Parasitol*. **47**:309-330.

Naughton MP, Henderson A, Mirabelli MC, Kaiser R, Wilhelm JL, Kieszak SM, Rubin CH, McGeehin MA. 2002. Heat-related mortality during a 1999 heat wave in Chicago. *Amer. J. Prev. Med*. **22**(4): 221-227.

Pacanowski RC, Griffies SM. 1998. MOM3.0 manual. *NOAA/Geophysical Fluid Dynamics Laboratory*. Princeton. NJ. 608pp.

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

Pascal M, Laaidi K, Ledrans M, Baffert E, Caserio-Schönemann, Le Terte A, Manach J, Medina S, Rudant J, Empereur-Bissonnet P. 2006. France's heat health watch warning system. *Int. J. Biometeorol***. 50**:144-153.

Rocklöv J, Ebi, K, Forsburg B. 2011. Mortality related to temperature and persistent extreme temperature: a study of cause-specific and age-stratified mortality. *Occup. Environ. Med*. **68**: 531-536.

Roeckner E and Coauthors. 1996. The atmospheric general circulation model ECHAM4: Model description and simulation of present-day climate. *Max-Planck-Institutfur Meteorologie Rep*. **218**. Hamburg, Germany, 90pp.

Saha S, Nadiga S, Thiaw C, Wang J, Wang W, Zhang Q, van denDool HM, Pan HL, Moorthi S, Behringer D, Stokes D, Pe˜na P,Lord S, White G, Ebisuzaki W, Peng P, Xie P. 2006. The NCEP climate forecast system. *J. Clim.***19**: 3483–3517.

Semenza JC, Rubin CH, Falter KH, Selanikio JD, Flanders WD, Howe HL, Wilhelm JL. 1996. Heat-related deaths during the July 1995 heat wave in Chicago. *N. Engl. J. Med.* **335**(2):84-90.

Sheridan SC, Kalkstein LS. 2004. Progress in heat watch-warning system technology. *Bull. Amer. Meteor. Soc*. 1931-1941.

Smoyer-Tomic KE, Rainham DGC. 2001. Beating the heat: development and evaluation of a Canadian hot weather health-response plan. *Environ. Health Perspectives*. **109** (12): 1241- 1248.

Steadman RG. 1979. The assessment of sultriness. Part I: A temperature-humidity index based on human physiology and clothing science. *J. Appl. Meteor*. **18**: 861-873.

Stockdale TN, Anderson DLT, Alves JOS, Balmaseda MA. 1998. Global seasonal rainfall forecasts using a coupled ocean atmosphere model. *Nature*. **392**: 370-373.

Tan J, Kalkstein LS, Huang J, Lin S, Yin H, Shao D. 2004. An operational heat/health warning system in Shanghai. *Int.J.Biometerol*. **48**: 157-162.

Terray L, Piacentini A and Valcke S. 1999. OASIS 2.3, Ocean Atmosphere Sea Ice Soil: User $(B!G (Bs guide. CERFACS Tech. Rep. TR/CMGC/99/37, Toulouse, France, 82 pp. [Available online at www.cerfacs.fr/globc/publication.html.].

The Government of the Republic of South Africa. 2011. *National climate change response white paper*.

Thomson MC, Doblas-Reyes FJ, Mason SJ, Hagedorn R, Connor CJ, Phindela T, Morse AP, Palmer TN. 2006. Malaria early warnings based on seasonal climate forecasts from multi-model ensembles. *Nature*. **439**.

Troccoli A, Harrison M, Coughlan M, Williams JB. 2008. Seasonal Forecasts in Decision Making. *NATO Sci. Ser*. **82**: 1341. DOI: 10.1007/978-1-4020-6992-5_2.

Vaneckova P, Neville G, Tippett V, Aitken P, Fitzgerald G, Tong S. 2011. Do biometeorological indices improve modeling outcomes of heat-related mortality? *J. Appl. Meteor. Climatol*. **50**: 1165-1176.

Van den Dool HM. 2007. Empirical Methods in Short-Term Climate Prediction. *Oxford University press*. 215pp.

Van Oldenborgh GJ, Balmaseda MA, Ferranti L, Stockdale TN, Anderson DLT. 2004. Did the ECMWF Seasonal Forecast Model Outperform Statistical ENSO Forecast Models over the Last 15 Years? *J. Climate*.**18**: 3240-3249.

Watts JD, Kalkstein LS. 2004. The development of a warm-weather relative stress index for environmental application. *Bull. Amer. Meteor. Soc*. 503-513.

Weisheimer A, Dobals-Reyes FJ, Palmer TN, Alessandri A, Arribas A,D´equ´e M, Keenlyside N, MacVean M, Navarra A, Rogel P. 2009.ENSEMBLES: A new multi-model ensemble for seasonal-to-annual predictions – Skill and progress beyond DEMETER in forecasting tropical Pacific SSTs. *Geophys. Res. Lett.* **36**: L21711,

DOI:10.1029/2009GL040896. Wichmann J, Andersen Z, Ketzel M, Ellerman T, Loft S. 2011. Apparent temperature and cause-specific emergency hospital admissions in greater Copenhagen. Denmark. *PLoS One*. **6**(7).

Wilks DS. 2006. Statistical Methods in the Atmospheric Sciences, 2nd Edition. Academic Press. *Int.Geophys. Ser*. Volume 91; 627pp.

Wyndham CH. 1965. A survey of causal factors in heat stroke and of their prevention in the gold mining industry. *J. South African Inst Mining Metallurgy*. **66**:125-156.