One-tiered vs. Two-tiered Forecasting of South African Seasonal Rainfall

Willem A. Landman[1], Dave DeWitt[2] and Daleen Lötter[3]

1: Council for Scientific and Industrial Research; WALandman@csir.co.za
2: International Research Institute for Climate and Society; Daved@iri.columbia.edu
3: Council for Scientific and Industrial Research; DLotter@csir.co.za

ABSTRACT

A forecast system with all the components of the boundary between atmosphere, ocean, land and ice known to be of importance to atmospheric interannual variability modelled as fully interacting is called a fully coupled model system. Forecast performance by such systems predicting seasonal rainfall totals over South Africa is compared with forecasts produced by a computationally less demanding two-tiered system where prescribed sea-surface temperature (SST) anomalies are used to force the atmospheric general circulation model. Two coupled models and one two-tiered model are considered here, and they are respectively the ECHAM4.5-MOM3-DC2, the ECHAM4.5-GML-cfsSST, and the ECHAM4.5 atmospheric model that is forced with SST anomalies predicted by a statistical model. The 850 hPa geopotential height forecast fields of the three systems are statistically downscaled to South African Weather Service district rainfall by retro-actively predicting 3-month seasonal rainfall totals over a 14-year retro-active test period from the 1995/96 to the 2008/09 rainfall season. Forecasts are made for lead-times of up to 4 months and probabilistic forecast performance is evaluated for three categories with the outer two categories respectively defined by the 25th and 75th percentile values of the climatological record. The resulting forecast skill levels are also compared with levels obtained by downscaling forecasts produced by forcing the atmospheric model with observed SST in order to produce a reference forecast set. Forecasts produced by the coupled systems are generally outperforming the forecasts produced by the two-tiered system, but neither one of the two systems outscore the reference forecasts, suggesting that further improvement in operational seasonal rainfall forecast skill for South Africa is still achievable. Forecast verification results further supports the notion that predicting for the middle category has very limited skill notwithstanding the fact that in this study the middle category is defined by 50% of the climatological record.

1. INTRODUCTION

Sea-surface temperature (SST) anomalies, themselves a result of coherent atmosphere-ocean interactions, have already been found as a probable cause of low-frequency variability in the atmosphere. Moreover, SST anomalies are arguably of greatest significance on the seasonal to interannual time scales and their slow evolution influence seasonal mean weather conditions (Goddard and Mason, 2002). Therefore, estimation of the evolution of SST anomalies, which are often relatively predictable, and subsequently employing them in atmospheric general circulation models (AGCMs), potentially provides means of generating forecasts of seasonal-average weather (Graham et al. 2000). Such a so-called two-tiered procedure to predict the outcome of the rainfall season has been employed in South Africa for a number of years already.

The advent of fully coupled ocean-atmosphere models (e.g. Stockdale et al, 1998), or one-tiered systems, promised improved seasonal forecasts since in theory coupled models should eventually outperform two-tiered systems because the former is able to describe the feedback between ocean and atmosphere while the latter assumes that the atmosphere responds to SST but does not in turn affect the oceans (Copsey et al., 2006). This notion will be tested here by comparing the seasonal rainfall forecast performance of a two-tiered system with forecasts from fully coupled systems. For both two-tiered and fully coupled systems the same AGCM will be used.

2. DATA AND GLOBAL MODELS

The 3-month seasonal rainfall data used for the downscaling are calculated from the district rainfall data set of the South African Weather

Service, and comprises of 94 evenly distributed locations across South Africa. This data set consists of monthly data from 1951 to 2009.

All of the global model data are obtained from the data library of the International Research Institute for Climate and Society (IRI). The AGCM data used is produced by the ECHAM4.5 (Roeckner et al., 1996) and consists of two sets. The first set (available from January 1950 to present) is produced by forcing the ECHAM4.5 with observed SST and consists of 24 ensemble members, and the second set (available from January 1957 to July 2008), also consisting of 24 ensemble members, is produced by forcing the model with SST anomalies that are forecast using constructed analogue SST (Van den Dool, 2007). Forecast data from two coupled models are also used and their ocean models are respectively the MOM3 (Pacanowski and Griffies, 1998) directly coupled to the ECHAM4.5 (DeWitt, 2005), and a slab mixed layer (denoted ECHAM4.5-GML). Each of these forecast sets consist of 12 ensemble members, and the data are available from January 1982 to present.

There are four forecast lead-times considered. For the two-tiered and the ECHAM4.5-GML systems, forecasts are produced near the beginning of the month, and for the ECHAM4.5-MOM3 system near the end of the month. A 1-month lead-time for the former two models implies that there are about three weeks from the issuance of the forecast to the beginning of the forecast season. For example, a 1-month lead-time forecast for the December-January-February (DJF) season is produced at the beginning of November, 2-month lead-time forecasts are produced early October, 3-month lead-time forecasts early September, and 4-month lead-time forecasts early August. For the ECHAM4.5-MOM3 system, there are at least 4 weeks between the production of the forecast and the first month of the forecast season. For example, DJF forecasts at a 1-month lead-time is produced near the end of October, 2-month lead-time forecasts at the end of September, 3-month lead-time forecasts at the end of August, and 4-month lead-time forecasts at the end of July.

3. MODEL OUPUT STATISTICS

Model output statistics (MOS; Wilks, 2006) equations are developed here because they can compensate for systematic deficiencies in the global models directly in the regression equations. The reason why these model errors can be overcome is because MOS uses predictor values from the global models in both the development and forecast stages. Notwithstanding, the selection of the appropriate model field require careful consideration: Raw model forecast of rainfall that is a result of, for example, the interaction between atmospheric circulation and topography is poorly resolved, and may therefore not be a good predictor of rainfall observed at ground level. Rainfall fields, even when totalled over a season, are noisy, and normally contain structures on spatial scales well below those resolved by the models. However, variables such as large-scale circulation are more accurately simulated by models than rainfall and should therefore be used instead in a MOS system to predict seasonal rainfall totals (Landman and Goddard, 2002).

The MOS equations are developed by using the canonical correlation analysis (CCA) option of the Climate Predictability Tool (CPT). This tool was developed at the IRI (http://iri.columbia.edu). The forecast fields from each global model used in the MOS are restricted over a domain that covers an area between the Equator and 45°S, and 15°W to 60°E. Empirical orthogonal function (EOF) analysis is performed on both the predictor (global model fields) and predictand sets (district rainfall) prior to CCA, and the number of EOF and CCA modes to be retained in the CPT's CCA procedure is determined using cross-validation skill sensitivity tests.

In order to minimize artificial inflation of forecast skill, the downscaled forecast performance of the individual models should be verified over a test period that is independent of the training period and should involve evaluation of predictions compared to their matching observations excluding any information following the forecast year. Such a system mimics a true operational forecasting environment where no prior knowledge of the coming season is available. For the example of DJF rainfall, the models are first trained with information from 1982/83 and leading up to and including 1994/95. The seasonal rainfall of the next year (1995/96) is subsequently predicted using the trained models. The various MOS sets of equations are subsequently retrained using information leading up to and including 1995/96 to predict for 1996/97 conditions. This procedure is continued until the 2008/09 DJF rainfall is predicted using MOS systems trained with data from 1982/83 to 2007/08, resulting in 14 years (1995/96 – 2008/09) of independent forecast data. In estimating the skill in predicting seasonal rainfall totals over South Africa, the observed and predicted fields are separated into three categories defining above-normal, near-normal and below-normal seasonal rainfall totals. However, these categories are not equi-probable here since the above- and below-normal threshold values respectively represent the 75th and 25th percentile values of the climatological record.

The distribution of individual ensemble members is supposed to be able to indicate forecast uncertainty. However, only a finite ensemble is available (12 or 24 members depending on the available global model data) suggesting that the forecast distribution may be poorly sampled or differently sampled owing to the difference in the available ensemble sizes – and so the uncertainty associated with the forecasts has to be estimated. Probabilistic MOS forecasts for each of the 14 retro-active years are obtained here from the error variance of the cross-validated predictions using the ensemble mean (Troccoli *et al*., 2008) for each of the various training periods. Cross-validation is performed using a (large) 5-year-out window, which means that 2 years on either side of the predicted year is omitted, in order to minimize the chance of obtaining biased results.

Seasonal climate is inherently probabilistic, and so seasonal forecasts should be judged probabilistically. The forecast verification measure presented here is the relative operating characteristic (ROC; Mason and Graham, 2002). ROC applied to probabilistic forecasts indicates whether the forecast probability was higher when an event such as a flood season occurred compared to when it did not occur, and therefore identifies whether a set of forecasts has the attribute of discrimination.

4. RESULTS

A ROC graph is made by plotting the forecast hit rates against the false alarm rates. The area beneath the ROC curve is used as a measure of discrimination and is referred to as a ROC score. If the area would be ≤ 0.5 the forecasts have no skill, and for a maximum ROC score of 1.0, perfect discrimination has been obtained. Figure 1 shows the ROC scores for the three forecast categories for DJF rainfall totals for each of the individual downscaled models as calculated over the 14-year test period. On the figure the ROC scores for the two coupled models and AGCM is shown for the three categories and for the four lead-times, together with the ROC scores from the simulations that used observed SSTs to force the AGCM (the reference scores).

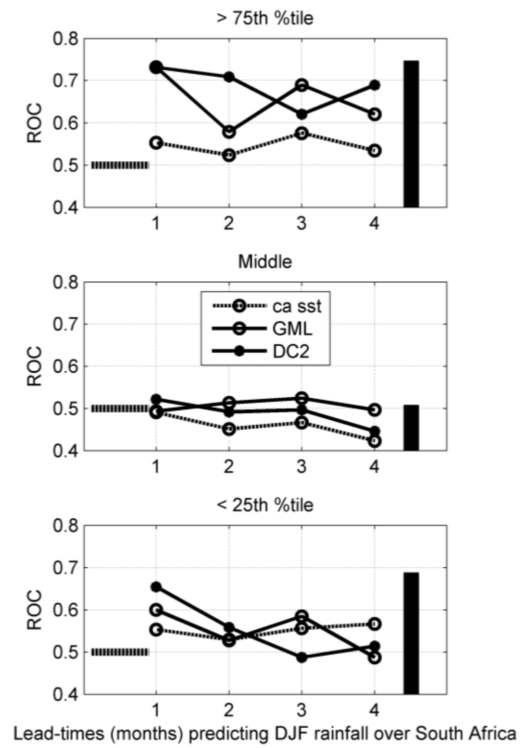

Figure 1. ROC scores for the prediction of DJF rainfall totals over South Africa. The scores of the fully coupled and AGCM downscaled forecasts are shown for each lead-time (solid and dashed lines), as well as the scores for the AGCM simulation runs (black bar).

For the most part ROC scores associated with the coupled models are the highest, especially for the prediction of wet conditions over South Africa during DJF. Notwithstanding, neither one of the coupled or two-tiered systems outscore the reference forecasts, suggesting that further improvement in operational seasonal rainfall forecast skill for South Africa should still be achievable. The middle panel shows scores when predicting for the near-normal category defined by 50% of the climatological record. These verification results support the notion that predicting for the middle category has limited skill since ROC scores for most of the lead-times are near or below 0.5

5. DISCUSSION AND CONCLUSION

Centres producing operational seasonal forecasts for South Africa need to know whether or not modelling research should be directed towards more expensive coupled models as opposed to more generally used two-tiered operational forecasting systems. This will be the case when a more demanding (in a computational and resource based sense) coupled system outscores a two-tiered system, which has been shown to be the case here for DJF rainfall. However, when skilful SST forecasts are used two-tiered systems may perform at least equally well as coupled systems as has been demonstrated by the simulation case when the AGCM was forced with observed SST. In conclusion, coupled models perform skillfully over South Africa and may even be as skilful as an AGCM forced with perfect SST. This paper has therefore demonstrated that it certainly is feasible to direct some of the available research and modelling funds as well as effort towards the development of operational seasonal forecasting systems that incorporate fully coupled models.

6. ACKNOWLEDGEMENTS

The district rainfall data set was kindly provided free of charge by the South African Weather Service.

7. REFERENCES

Copsey, D., Sutton, R. and Knight, J.R. (2006) Recent trends in sea level pressure in the Indian Ocean region. *Geophysical Research Letters*, **33**, L19712, doi:10.1029/2006GL027175.

DeWitt, D.G. (2005) Retrospective forecasts of interannual sea surface temperature anomalies from 1982 to present using a directly coupled atmosphere-ocean general circulation model. *Monthly Weather Review*, **133**, 2972-2995.

Goddard, L. and Mason, S.J. (2002) Sensitivity of seasonal climate forecasts to persisted SST anomalies. *Climate Dynamics*, **19**, 619-631.

Graham, R.J., Evans, A.D.L., Mylne, K.R., Harrison, M.S.J. and Robertson, K.B. (2000) An assessment of seasonal predictability using atmospheric general circulation models. *Quarterly Journal of the Royal Meteorological Society*, **126**, 2211-2240.

Landman, W.A. and Goddard, L. (2002) Statistical recalibration of GCM forecasts over southern Africa using model output statistics. *Journal of Climate*, **15**, 2038-2055.

Mason, S.J. and Graham N.E. (2002) Areas beneath the relative operating characteristics (ROC) and levels (ROL) curves: Statistical significance and interpretation. *Quarterly Journal of the Royal Meteorological Society*, **128**, 2145-2166.

Pacanowski, R.C. and Griffies, S.M. (1998) *MOM 3.0 manual*. NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, NJ, 608pp.

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

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

Troccoli A., Harrison, M., Anderson, D.L.T. and Mason S.J. (2008) *Seasonal Climate: Forecasting and managing risk*. Springer, NATO Science Series. Earth and Environmental Sciences, Volume **82**, 467pp.

Van den Dool, H. M. (2007) *Empirical Methods in Short-Term Climate Prediction*. Oxford University Press, 215pp.

Wilks, D.S. (2006) *Statistical Methods in the Atmospheric Sciences, 2nd Edition*. Academic Press, International Geophysics Series, Volume **91**, 627pp.