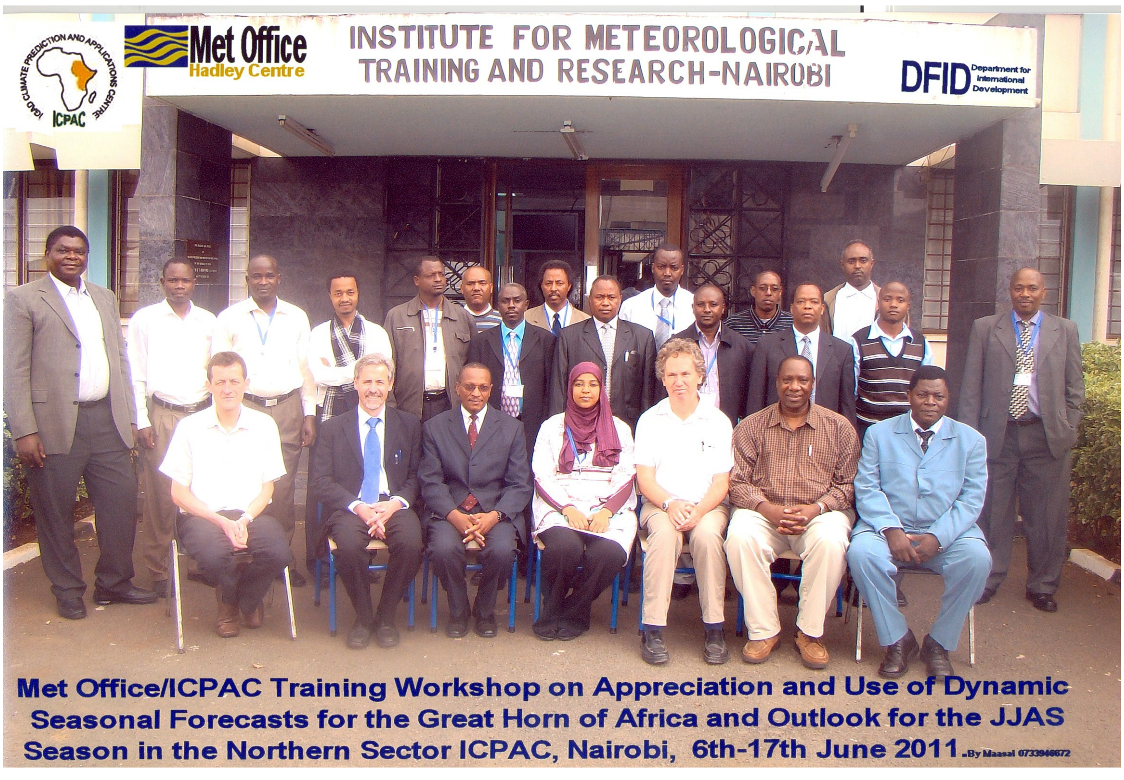
The ICPAC - Met Office CSRP training workshop, 6-17 June 2011, ICPAC, Nairobi

Appreciation and use of dynamical seasonal forecasts for the Greater Horn of Africa: summary of workshop aims, outcomes and training materials



The ICPAC-Met Office Hadley Centre CSRP training workshop, June 2011

Introduction

As part of the knowledge management component of the CSRP programme, a 2-week workshop entitled: "Training workshop on appreciation and use of dynamical seasonal forecasts for the Greater Horn of Africa and outlook for the July-August-September season in the northern sector" was held at the IGAD Climate Prediction and Application Centre (ICPAC), Nairobi 6-17 June 2011. The workshop, which included a 1-day user forum, was planned and delivered jointly with ICPAC, with additional trainers from the International Research Institute for Climate and Society (IRI) and the Universities of North Carolina and Connecticut.

The workshop had 5 key aims:

1. To take stock of current seasonal forecasting methods used by countries in the region.
2. To strengthen understanding of seasonal climate variability and its predictability and build appreciation of dynamical (i.e. climate-model based) seasonal forecasts of such variability including assessment of their strengths, weaknesses and of optimum ways of using the model information – with a view to increased and more informed use of these relatively new forecast outputs in country and regional-level seasonal outlooks.
3. To generate new knowledge on the performance of the seasonal prediction systems of 4 Global Producing Centres (GPCs) for 3 rainy seasons over the Greater Horn of Africa (GHA); March-May (MAM), July-September (JAS) and September-December (SOND). The 4 GPC systems investigated were those run by the Met Office, ECMWF, Météo-France and the NOAA NCEP Climate Prediction Centre (CPC) (the CFS-1 system).
4. To make recommendations on procedures for integrating dynamical forecast products into seasonal forecasting activities for the region.
5. To prepare a consensus forecast for the GHA for the July-August-September (JAS) season and to present it to stakeholders at a user forum (the 29th Greater Horn of Africa Climate Outlook Forum, GHACOF29). This first ever consensus outlook for the JAS season was included in response to a strong stakeholder demand that ICPAC extend prediction services to this season.

The focus of the workshop's training - appreciation and use of dynamical seasonal forecasts - was selected for the following reasons:

- Seasonal prediction methods used in the region are at present based primarily on statistical methods. There is a recognised need to develop use of seasonal forecast products from dynamical climate models, the availability of which has increased in the last few years following WMO initiatives (see below).
- A good understanding of the strengths and weaknesses of dynamical forecast products is needed to make informed use of them. Given the relative recent availability of these products, such understanding is not well developed in many regions.
- There are now 12 centres providing dynamical seasonal forecasts through the auspices of the World Meteorological Organisation (WMO) - the WMO-designated Global Producing Centres of long-range forecasts (GPCs). There is a need for detailed information on the relative performance of the GPC systems over the GHA.

Workshop format and training style

The workshop comprised mainly hands on interactive exercises to give participants maximum exposure to the principles and forecasting techniques introduced. Lectures were restricted to essential underpinning science material and introductory sessions on the exercises.

Workshop participants

There were 15 participants in all, one each from the National Meteorological Services of Sudan, Eritrea, Djibouti, Ethiopia, Uganda, Burundi, Rwanda and Tanzania, 3 from Kenya Meteorological Department and 4 Research Assistants from ICPAC who both participated and assisted the other participants.

Training materials

Materials developed included a set of pre-prepared exercises delivered using an interactive spreadsheet format. The spreadsheets contained precipitation hindcasts for 9 regions of the Greater Horn of Africa from 4 seasonal forecasting centres: Met Office, ECMWF system3 (spreadsheets have since been updated with system4 hindcasts), Météo-France and NOAA CPC CFS-1, and observed precipitation data from GPCP and NCEP. The 9 regions (Fig. 1) were selected to make the analysis exercises over the GHA region manageable in the time available, and to broadly represent countries and geographic regions – they are not intended to replace any existing climatological subdivision of the region. The spreadsheets were accompanied by written exercises for participants to work through. Examples of figures generated interactively by participants using spreadsheet manipulation of the hindcast data are provided below and cover the following topics:

- uncertainties in observed rainfall data for the regions (through comparison of two gridded rainfall datasets) (Fig. 2);
- properties of the ensemble mean forecast and comparison of the skill of the ensemble mean for the 4 GPCs (Fig. 3);
- The concept of ensemble spread (Fig. 4);
- Calculation of forecast probabilities for quantile categories (Fig. 6);
- The concept of reliability in probability forecasting and construction of reliability diagrams (Fig. 5).
- The concept of hits and false alarms and calculation of the Relative Operating Characteristic (ROC) diagram and ROC scores (area under the ROC curve) (Fig. 6);

Advanced statistical post-processing of dynamical model seasonal forecasts

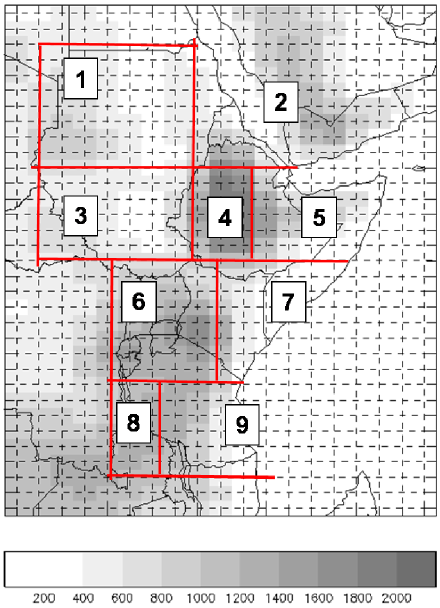
In the second week of the workshop statistical techniques for advanced calibration of dynamical model output were introduced. Calibration involves use of the performance 'track record' of the forecasts as obtained from retrospective forecasts and observations. Calibration has potential to improve forecast skill by reducing systematic biases present in the 'raw' model output. Three methods were considered: multiple linear regression, principle component regression and canonical correlation analysis. The Climate Predictability Tool (CPT) a software package developed by the International Research Institute for Climate and Society (IRI) was used to apply these techniques.
http://portal.iri.columbia.edu/portal/server.pt?open=512&objID=697&PageID=7264&mode=2

Workshop recommendations

- It was recommended that digital seasonal forecast data be supplied to all the participants ahead of future GHACOF meetings – enabling them to do some preliminary analysis of forecast signals using techniques developed at the workshop. Results of these analyses would then be brought to the meeting and would help to enrich discussions on the consensus forecast;
- The workshop noted that the dynamical models have good skill in predicting El Niño/La Niña conditions during SOND from as far ahead as the preceding May and that this suggested potential for issuing a preliminary, long-lead forecast for the SOND 'short rains' season in the preceding June (as an adjunct to the July-September GHACOF meeting). It was recommended that a long-lead forecast of the short-rains season be considered as an output of the next June meeting for the July-September 2012 season;
- Participants requested the opportunity to keep in contact with their training group and with future training activities.

Concluding remarks

Feedback from the participants of the workshop indicated that their understanding and interpretation of dynamical seasonal forecasts had been increased by the training materials. Nine out of the ten participants who responded to a feedback questionnaire stated that they felt more confident in using dynamical model seasonal forecasts to make their national and regional predictions; all ten stated that they had advanced their understanding of dynamical models and verification methods. Trials in other regions of Africa and further feedback would be useful in helping to refine and develop these training aids for wider use.



1. **Northern Sudan**
2. **Eritrea, Northeast Sudan, N. Ethiopia**
3. **South Sudan, West Ethiopia**
4. **Ethiopia highlands**
5. **East Ethiopia, North Somalia, Djibouti**
6. **Uganda, Rwanda, Burundi, West Kenya, North Tanzania**
7. **East Kenya, South Somalia**
8. **Southwest Tanzania**
9. **East Tanzania**

Fig. 1: The 9 regions defined for the purpose of the exercises. Shading indicates orographic height (m) as represented by the GloSea4 system

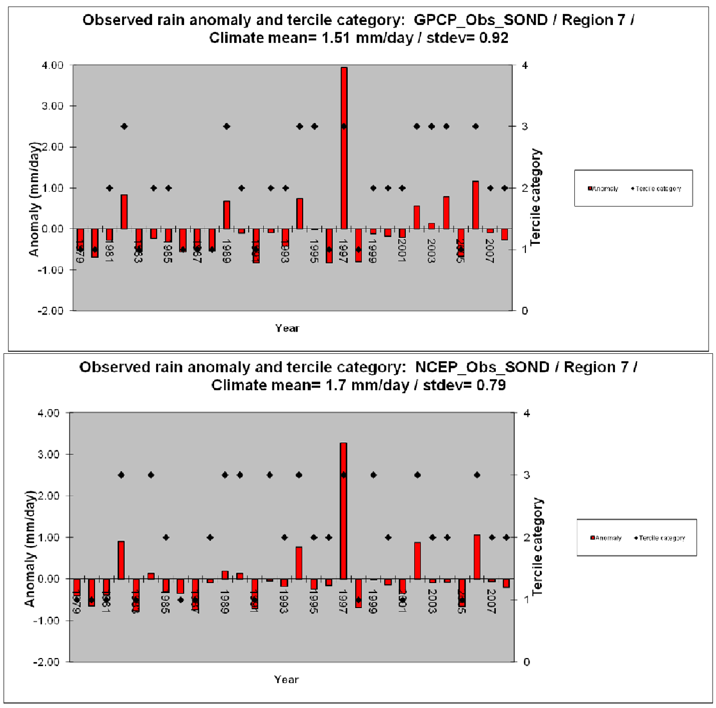


Fig. 2: SOND season rainfall for region 7, 1979-2008 from, the GPCP dataset (top) and NCEP dataset (bottom). Red bars show rainfall anomalies, black diamonds show the corresponding tercile categories relative to the 1979-2008 climatology (1=below; 2=average; 3=above, right-hand axis). Note substantial differences in the observed anomalies of the two datasets in some years (e.g. 2004) leading to difference in assignment of tercile category. Such differences lead to different forecast verification results when different observation datasets are used as ‘truth’. In this case the assigned tercile category differs in 11 out of the 30 years, though never by more than one tercile category.

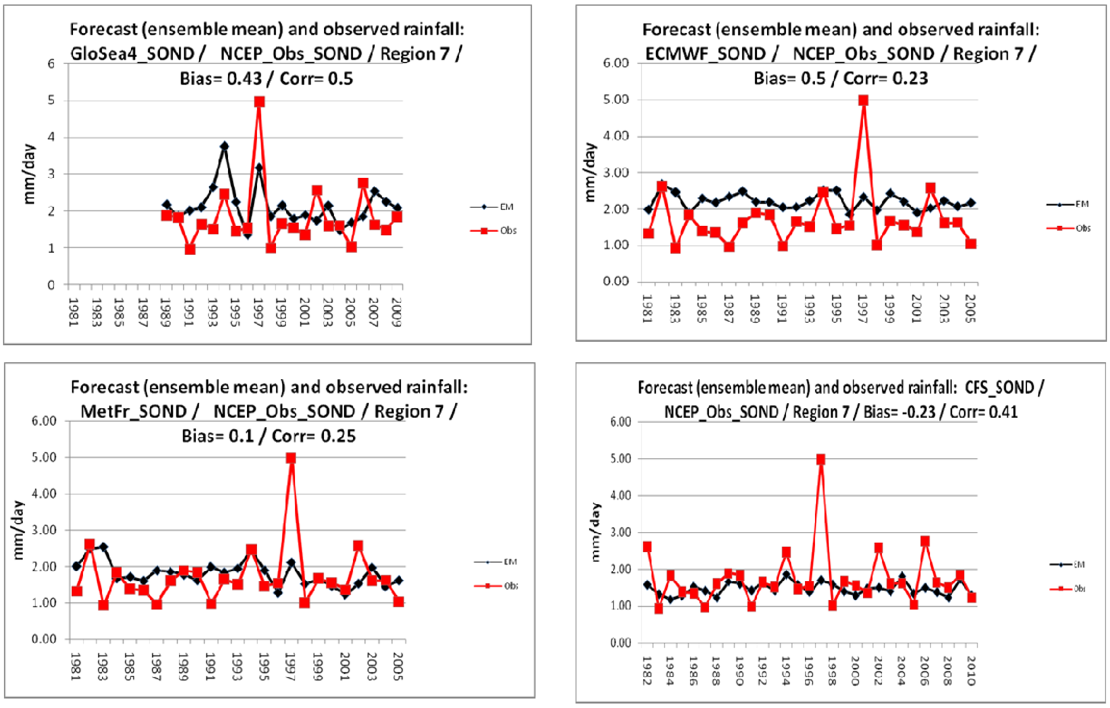


Fig. 3: Ensemble mean (black) and observed SOND rainfall (red) for region 7 for 4 GPCs over their respective hindcast periods. Bias and correlation scores are shown. For this region and season GloSea4 has the highest correlation (0.5) and is the only model to show a substantial wet signal for the 1997 season.

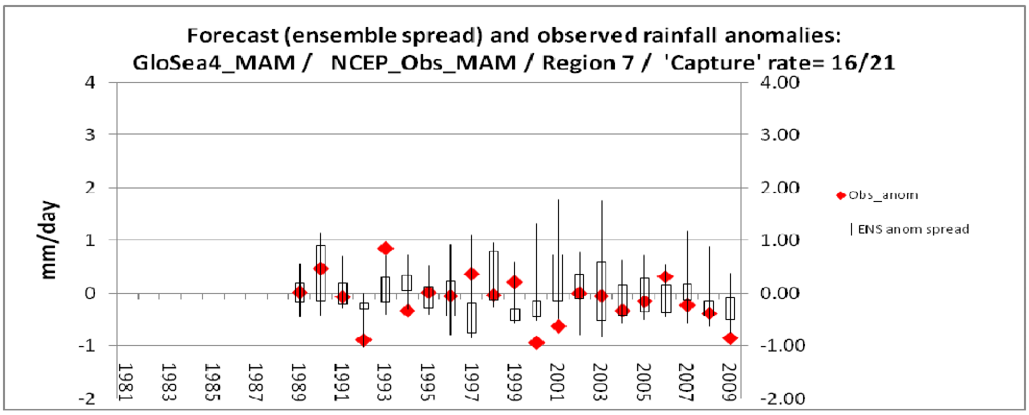


Fig. 4: GloSea4 ensemble hindcast and observed (red diamonds) MAM rainfall anomalies, 1989-2009. The ensemble predictions are shown in box and whisker format where the whiskers extend to the maximum and minimum values in the ensemble and the box shows the interquartile range (the whiskers cover driest and wettest 25% of ensemble members, the box covers the remaining (central) 50% of members). The capture rate refers to the proportion of years when the observed anomaly lies within the ensemble range. The better the ensemble represents the forecast uncertainty, the more frequently the observed value will lie within the range of values predicted by the ensemble. In this case the observed value lies within the predicted range in 16 of 21 years. To illustrate the interpretation of the box and whiskers, in each of the 3 driest observed years, 1992, 2000 and 2009 more than 75% of the ensemble members indicated negative rainfall anomalies. In 1992 some members predicted values as dry as or drier than the observed value, in 2009 some members came close to predicting the severity of the dry conditions while in 2000 no members anticipated the severity of the dry conditions.

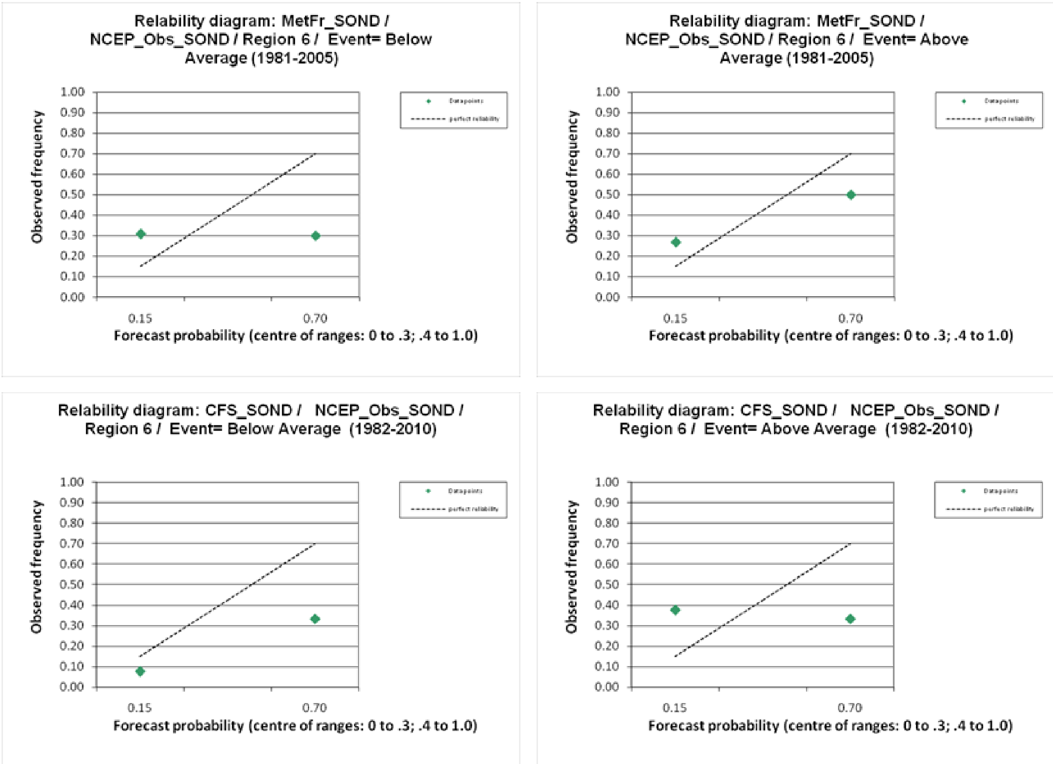


Fig. 5: Reliability diagrams for below normal (left) and above normal (right) SOND rainfall for region 6, obtained with the Météo France system (top) and the NOAA CFS-1 system (bottom). The diagrams show the observed frequency of the event (vertical axis) conditional on the forecast probability lying in two broad ranges: 0 to 0.3 and 0.4 to 1.0. Only 2 ranges were used in order to maximise sample size. The Météo France system's predictions of above normal rainfall show some reliability: i.e. the event is more frequently observed to occur when the forecast probability lies in the higher (0.4 to 1.0) range than when the forecast probability is in the lower (0 to 0.3) range. However, forecasts of below normal rainfall show no reliability: i.e. the likelihood of the event occurring is approximately equal to the climate chance (0.33) regardless of the forecast probability. In contrast the CFS system shows some reliability for the below normal event (though probabilities are biased low) and no reliability for the above normal event.

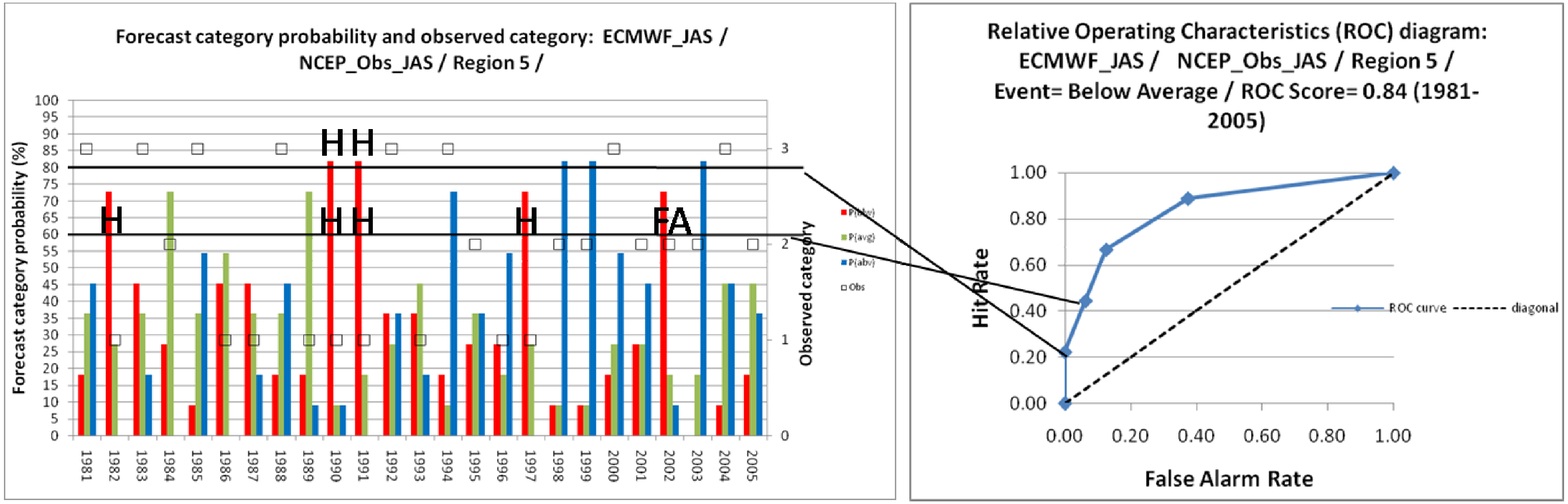


Fig. 6: Left: ECMWF system3 hindcast probabilities for the below (red bars), average (green bars) and above (blue bars) tercile categories of rainfall and the observed category (black squares: 1=below; 2=average; 3=above – right-hand axis) for the JAS season, region 5, 1981-2005. Right: corresponding Relative Operating Characteristic diagram for the below normal category, plotted points are (from right to left) for the probability thresholds of 0%, 20%, 40%, 60%, 80% and 100%. Annotations illustrate how the ROC curve is constructed. If a warning of below normal rains is issued when the forecast probability of this event exceeds 80% there are two 'hits' (H) - years 1990 and 1991, when below normal was observed, and no false alarms (FA) when below normal was not observed. The hit rate is defined as the number of hits divided by the total number of times in the series when below normal was observed. The number of times below normal was observed is 9 (i.e. number of black squares aligned with '1' on the right-hand axis). Thus the hit rate is 2/9=0.22. The false alarm rate is the number of false alarms divided by the number of times the event did not occur over the time series (i.e. 30-9=21). Thus the false alarm rate is 0/21=0. Thus the point (0.0,0.22) is plotted on the ROC diagram for the 80% probability threshold. Therefore, in this sample, when the forecast probability of below normal was equal to or greater than 80% (2 occasions), below normal always occurred (note, though, that there would have been 7 misses (1982, 1986, 1987, 1989, 1993, 1996 and 1997) - occasions when below normal occurred and a warning was not issued). If a lower probability threshold (e.g 60%) is used, 5 warnings are issued of which 4 are hits (1982, 1990, 1991 and 1997) and 1 is a false alarm (2002), when average (not below average) was observed – giving a hit rate of 4/9=0.44 and a false alarm rate of 1/21=0.05 (giving the coordinates of the point corresponding to the 60% threshold). For a skilful system, hit rates must exceed false alarm rates and thus the greater the skill of the forecast system the more the ROC curve will bow up to the top left-hand corner of the diagram. The area under the ROC curve (in this case 0.84) gives a summary measure of the skill. Perfect deterministic forecasts would achieve a score of 1. Forecasts that are unable to discriminate the observed category (i.e. forecasts with no skill) achieve a ROC score of 0.5. In the 'no-skill' case the ROC curve lies along the diagonal (hit rates equal false alarm rates).