Met Office
Hadley Centre

Forecast of East African rainfall for October-December 2020 using statistical and dynamical methods

Issued 2 October 2020

FORECAST OF EAST AFRICAN RAINFALL FOR OCTOBER-DECEMBER 2020 USING DYNAMICAL AND STATISTICAL METHODS

Andrew Colman
Met Office Hadley Centre, Exeter, UK

PART 1. FORECAST

Summary

There are elevated probabilities for below average rainfall for the forthcoming October-December season.

1.1 Whole Region

Statistical and dynamical forecasts of October-November-December (OND) rainfall for the East Africa region, defined as 5°N-15°S, 30°E to Indian Ocean Coast, are presented in Figure 1.1. The statistical forecast is produced by discriminant analysis using indices of worldwide Sea Surface Temperature (SST) as predictors. The dynamical forecast is produced using the Met Office GloSea5 dynamical seasonal prediction system. Information on GloSea5 and the methodology used to generate the statistical and dynamical forecasts is described in part 2.

Observed SST for July-September 2020 is used to produce the statistical forecast. For the dynamical forecast, GloSea5 is run as an ensemble of 42 predictions initialised with atmosphere and oceanic conditions dating from between 21 August and 10 September 2020.

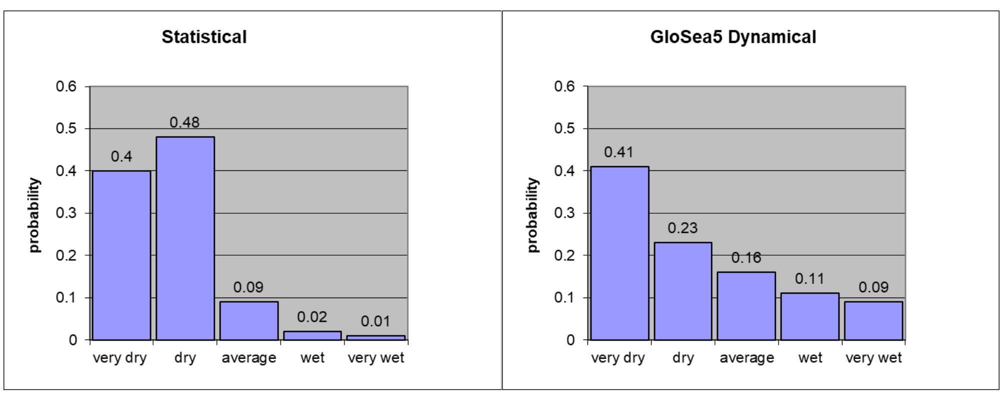


Figure 1.1: Predicted probabilities for October-November-December 2020 rainfall from the statistical and dynamical forecasts. Probabilities are for 5 quintile categories referred to as: very dry, dry, average, wet and very wet. The baseline (1961-1990) climatological probability for each category is by definition 0.2 (20%).

Highest probabilities from the statistical forecast are for the dry and very dry categories (0.48 and 0.4 respectively). The probability of the very dry category is also strongly enhanced in the GloSea5 dynamical forecast (0.41), compared with the chance level of 0.2. Correspondingly, probabilities of the wet and very wet categories are well below the chance level in both the statistical and dynamical forecasts.

1.2 Grid Box forecasts

Note: Predictability on small (grid box) scales is generally lower than on larger scales therefore small (grid box) scale spatial variations in the forecast should be regarded with caution.

Statistical tercile category probability forecasts (Figure 1.2a)

Figure 1.2a consists of 6 statistical forecast probability maps. The three maps in the upper row show probabilities for the three tercile categories for all grid boxes for which there are sufficient data to make statistical models. The second row of three maps (labelled “with skill mask”) is a repeat of the first row but only probabilities for grid boxes where there is significant correlation skill according to an independent test are shown. Details of the skill test used are in part 2.

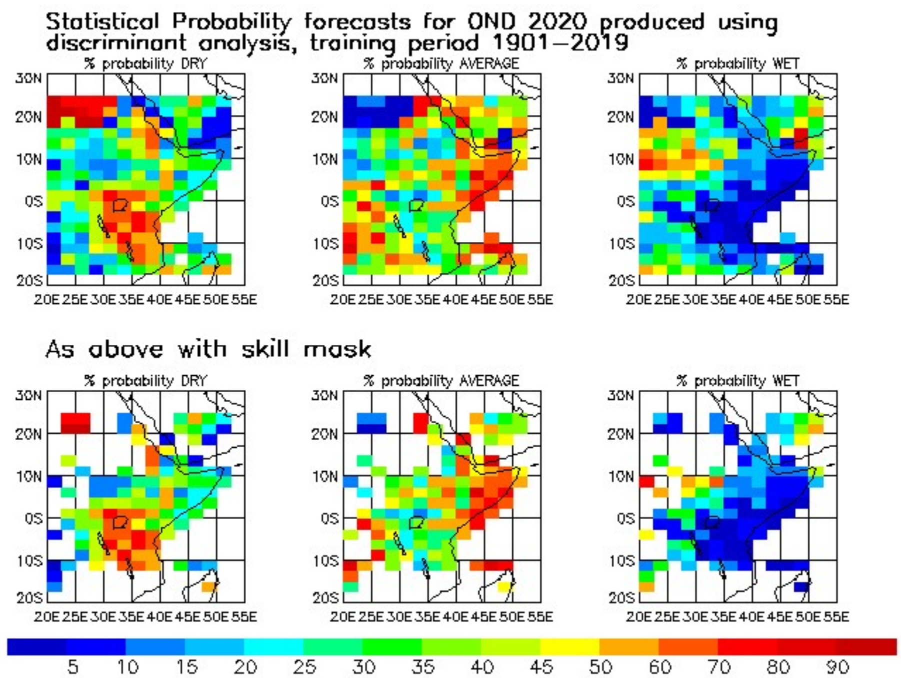


Figure 1.2a: Probability forecasts for the tercile categories obtained using the statistical method. The legend shows % probability, baseline for terciles is 1961-1990.

The dry category is predominantly the most probable from the equator southward between 30˚ and 40˚E. Elsewhere the average category is predominantly the most probable except for a band around 10˚N between 20˚ and 30˚E where the wet category is favoured.

Statistical forecasts of extremes and change from last year (figure 1.2b)

Figure 1.2b is similar to figure 1.2a but probabilities of a wetter season than last year and of 10-year extremes are presented. A wetter season than 2019 is predicted to be generally very unlikely. This is not surprising given that 2019 was a record wet year in many locations. With reference to the skill-masked figures, probabilities are less than 5% in nearly all locations (fig. 1.2a, bottom left). Similarly, the probability of a season wetter than any of the last 10 years (fig. 1.2a, bottom centre) is not enhanced above the chance level (about 10%). In contrast the probability of a season drier than any of the last 10 years (fig. 1.2a, bottom right) is enhanced above the chance level, with probabilities typically around 20 to 45% over much of the region

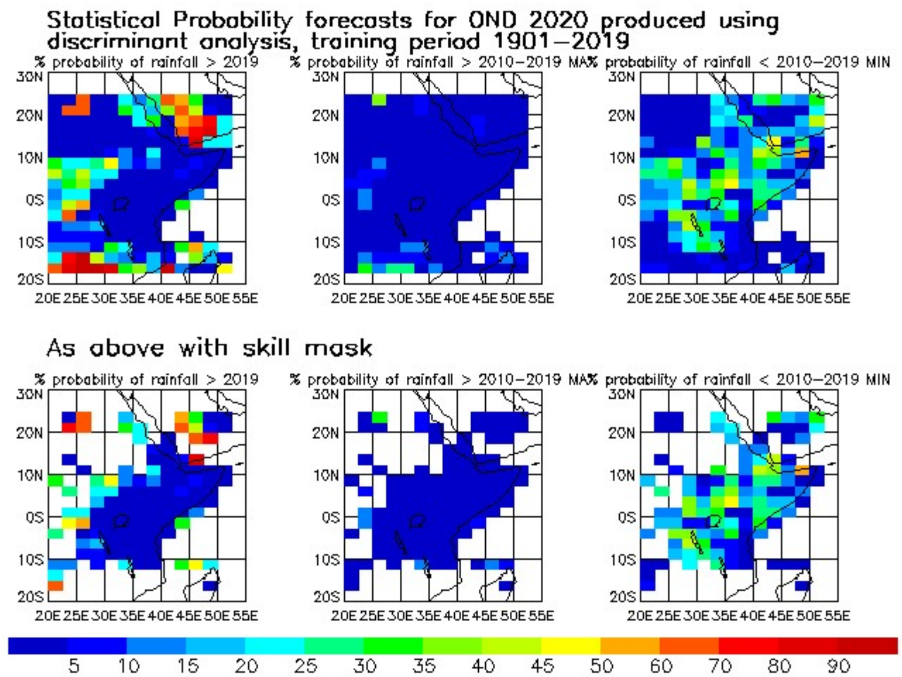


Figure 1.2b: Top: probability forecasts for a wetter season than last year (left); a wetter season than any in the last 10 years (centre) and a drier season than any in the last 10 years (right) obtained using the statistical method. Bottom: as top, but with skill mask (the skill mask is the same as in figure 1.2a).

CCA-Calibrated Dynamical (GloSea5) tercile category probability forecasts (Figure 1.3)

Direct output from GloSea5 dynamical forecasts for the region can be viewed at http://www.metoffice.gov.uk/research/climate/seasonal-to-decadal/gpc-outlooks/glob-seas-prob. Here we present forecasts generated by calibrating the GloSea5 output against historical observed rainfall using Canonical Correlation Analysis (CCA). For more details about CCA calibration see part 2.

Tercile category probability forecasts from the CCA-calibrated GloSea5 forecasts are presented in figure 1.3. Consistent with results of the statistical forecast and (quintile) dynamical forecast presented in section 1.1, the dry category is most probable over most of the region except for an area around 10^{o}N, 20^{o}-30^{o}E and also in the far north where it is usually dry at this time of year.

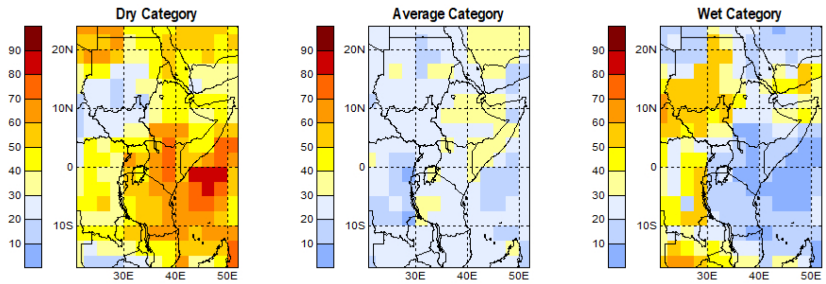


Figure 1.3: Probability forecasts for tercile categories obtained using the CCA-calibrated GloSea5. Baseline is 1981-2010

1.3 Concluding remarks

This year the statistical and dynamical forecasts are in good agreement in favouring the dry and very dry quintile categories for the region as a whole. On a grid point scale, both statistical and dynamical forecasts agree on favouring the dry category from the equator southward between 30 and 40 E. Further north and east the statistical forecasts favour the average category whilst the dynamical forecast favours the dry category.

PART 2. FURTHER INFORMATION

1. Introduction

The Met Office uses SST-based statistical methods and dynamical models of the global ocean-atmosphere system to make seasonal predictions of tropical rainfall. Forecasts have been made for October-December rainfall over tropical east Africa (the 'short rains') since 1994.

Forecasts are presented for (a) the “whole region” (5°N-15 °S, 30°E to Indian Ocean Coast) using both statistical and dynamical (GloSea5) methods and (b) for 2.5° latitude x 2.5° longitude grid-boxes covering the region using statistical and CCA-calibrated GloSea5 dynamical forecasts.

2. Forecast Methodology

Statistical forecasts

The statistical forecasts are made by using linear regression and discriminant analysis techniques, with 3 indices of global sea surface temperature (SST) anomaly patterns and 2 indices of mean SST anomaly for 2 box regions known to be linked to East African rainfall (all predictors are defined in the Appendix). The statistical forecast models are derived from historical rainfall and SST information over the period 1901-2019.

The statistical forecast probabilities for the “whole region” are calculated using discriminant analysis. Statistical gridbox forecast probabilities are generated using multiple linear regression The regression predictions are converted to probabilities by assuming a normal distribution and a standard error based on the performance of the regression forecasts

Dynamical Forecasts

The dynamical forecasts are produced using GloSea5, the latest version of the Met Office Global Seasonal forecast system which uses the HadGEM3 coupled ocean-atmosphere model modified for seasonal forecasting, for more about GloSea5 see http://www.metoffice.gov.uk/research/modelling-systems/unified-model/climate-models/glosea5.

In order to account for uncertainty in the model forecasts and initial conditions, the GloSea5 forecasts are run as ensembles, i.e. the models are run multiple times with slightly differing initial conditions and different model parameterisation to reflect the uncertainty in these factors.

Calibration is used to correct systematic errors in model output. Dynamical grid box probabilities are evaluated using a statistical calibration technique, Canonical Correlation Analysis (CCA, Von Storch and Zwiers (1999)). The CCA method used is similar to the MOS method described by Ndiaye et al (2011).

CCA works by identifying pairs of patterns in the predictor (model output) and predictand data (observations) that correlate well through time. Dynamical models can often predict larger scale (e.g. ocean-wide) circulation patterns better than variations for smaller regions. To apply CCA to East Africa region forecasts, model precipitation output for the domain 40°N to 40°S, 60°W to 100°E which covers most of the tropical Atlantic, Africa and the Indian ocean is used as the predictor data and the predictand data is 2.5° degree gridded GPCP for the region in figures 1.3 and 2.2. Rainfall patterns over the larger region reflect circulation anomalies which are known to be related to East African rainfall. Higher GloSea5 skill for these large-scale patterns can be used to enhance skill for East Africa. The CCA forecasts are produced using the IRI CPT package, http://iri.columbia.edu/our-expertise/climate/tools/cpt/.

As with the statistical linear regression forecasts, CCA forecasts are converted to probabilities by assuming the forecasts have a normal distribution and a standard error calculated from the performance of CCA calibrated GloSea5 hindcasts for 1993-2016.

CCA is used in a similar way to calculate the “GloSea dynamical large area” probabilities displayed in figure 1.1. (i.e. model precipitation output for the same domain is used)

3. Forecast Formats

For “whole region” forecasts, the observed historical rainfall record is divided into 5 equi-probable categories. The category boundaries are based on the 1961-1990 climatology and are given below as percentages of mean rainfall:

Very Dry/Dry	Dry/Average	Average/Wet	Wet /Very Wet
74%	86%	102%	124%

Grid box forecasts produced using the statistical method are expressed as probabilities of tercile categories which are climatologically equi-probable over 1961-1990. The tercile format is compatible with that used for the GHACOF (Greater Horn of Africa Climate Outlook Forum) consensus forecast. Grid box forecasts produced using the dynamical forecasts are expressed as probabilities of tercile categories which are climatologically equi-probable over 1981-2010. The latter period is used as it is more consistent with the dynamical hindcast period (1993-2015).

Grid box forecasts are also presented as probabilities of a wetter season than experienced in the previous year and as probabilities of an extreme season where “extreme” is defined as wetter or drier than any of the previous 10 years.

4. Forecast skill

Performance of retrospective grid box forecasts

The ROC skill (Broecker 2012, http://www.bom.gov.au/wmo/lrfvs/roc.shtml) of forecasts for grid box scale rainfall averages using the statistical method is presented in figure 2.2 for the lower and upper tercile categories. A ROC score of 50% indicates forecasts are no better than guesswork whilst a score of 100% would be achieved by perfect deterministic forecasts. ROC scores are generally highest between 30^{o} and 40^{o}E and between 5^{o}N and 10^{o}S but are below 50% for a some northern areas. In common with many seasonal forecast assessments, skill is higher for the outer categories than for the average category (not shown). Forecasts are assessed against the GPCP dataset (ftp://precip.gsfc.nasa.gov/pub/gpcp-v2.2/doc/V2.2_doc.pdf).

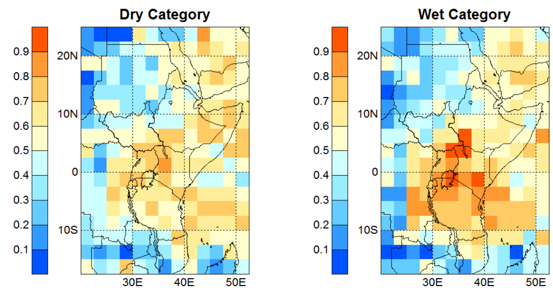


Figure 2.1: ROC skill (%) of statistical forecasts of gridded OND rainfall below the lower tercile (left) and above the upper tercile (right) evaluated over 1993-2015. Verification data=GPCP

Similar ROC skill plots of CCA spatial bias corrected GloSea5 dynamical forecasts are presented in figure 2.2. CCA calibration enhances skill (relative to direct model output) with highest ROC scores again (around 0.8) located between 30° and 50°E and between 5°N and 10°S for both the dry and wet category.

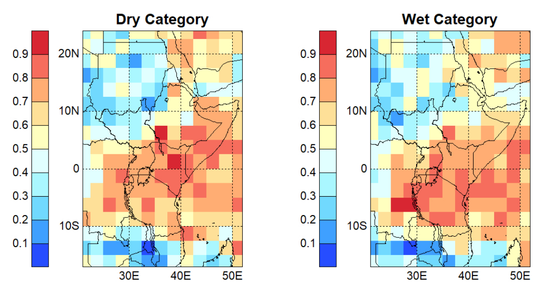


Figure 2.2: As Figure 2.1 but for CCA calibrated dynamical (GloSea5) forecasts of gridded rainfall.

Skill masking

A skill mask is applied to the statistical forecast maps. To be included in the “with skill mask” map, linear regression hindcasts for the box must pass at least one of these two tests:

1) Pearson correlation between hindcasts and observations over 1901-2019 are significant at the 95% level.
2) Pearson correlation between hindcasts of the same tercile category as predicted for this year and observations during 1901-2019 are significant at the 95% level.

Note: The skill mask test is applied to each of the 3 predictions described in the Appendix. Sometimes just one or two of the three predictor sets will pass this test which can result in the masked and unmasked probabilities being different for the same grid box and category.

Statistical forecasts have been produced for this region since 1994. Since 1997 our issued best estimate forecasts have been based on combined statistical and dynamical forecast information. The issued 'best estimate' forecasts are produced by forecaster interpretation of the available prediction information. These issued forecast categories for the whole region are compared with best estimate forecasts based on the statistical method alone (no forecaster interpretation or dynamical input) and with the corresponding observed category in figure 2.3. Over the 26 years, the correlation between issued best estimate forecast and observed category is 0.60 and the correlation between statistical best estimate forecast and observed category is 0.71, both of which are significant at the 95% level. Note that the issued forecast sample includes forecasts made with input from earlier, less skilful, versions of the dynamical system.

Note: Prior to 1999, forecast categories were based on the 1951-1980 climatology. From 1999 onwards, the 1961-1990 climatology has been used as it is the accepted WMO standard climatology period. The 1961-1990 rainfall average is 104% of the 1951-1980 average.

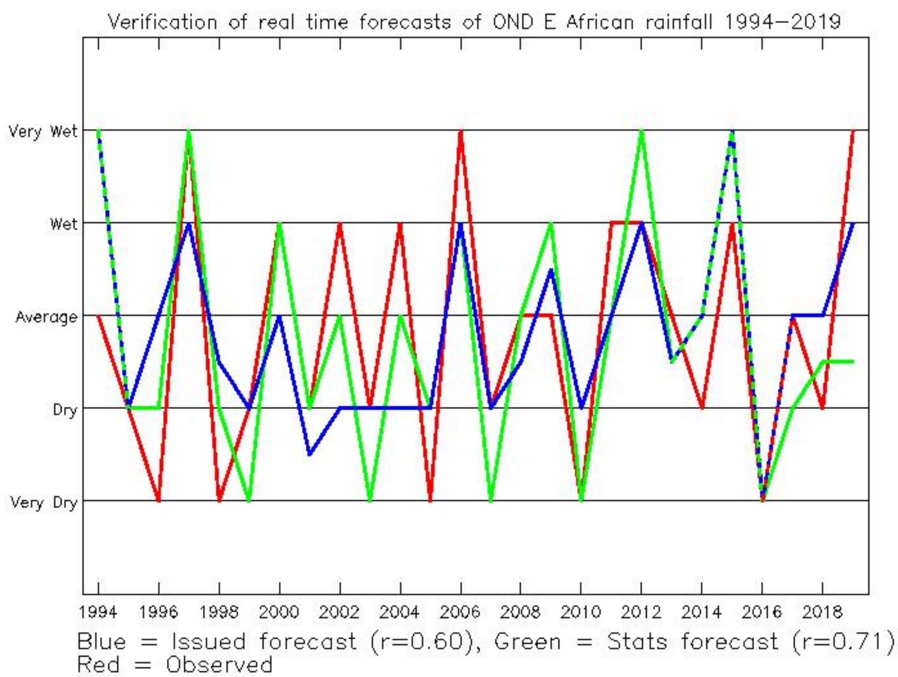


Figure 2.3: Verification of real time “best estimate” forecasts of October-December East African rainfall 1994-2019

A plotted point between 2 categories represents a best-estimate forecast that could not distinguish highest likelihood between two adjacent categories. For example, the issued forecast for 2008 was for Average or Dry and was plotted between the average and dry category. (Note that the statistics forecast is the same as the issued forecast for 1994,1995, 2013, 2014, 2015 and 2016, and in these cases the blue “issued forecast” line obscures the green “stats forecast” line)

Finally, to maintain a consistent record of forecasts for verifications such as Fig 2.3, we provide a best estimate forecast   The issued “best estimate” forecast for 2020 is for the “very dry” quintile category which is the most probable category (probability =0.44) when statistical and dynamical probabilities are averaged

REFERENCES:

Broecker J. Probability Forecasts. In Forecast Verification: A Practitioners Guide in Atmospheric Science, 2nd edition, eds. Jolliffe IT, Stephenson DB. Wiley, Chichester. 119-140 (2012).

Mutai, C.C., Ward, M.N and Colman, A.W. Towards the prediction of East Africa Short Rains based on Sea-Surface Temperature- Atmosphere coupling. I.J.Climatol.18 975-997 (1998).

Ndiaye, O, Ward, M.N., and Thiaw, W.M. Predictability of Seasonal Sahel Rainfall Using GCMs and Lead-Time Improvements Through the Use of a Coupled Model. J. Climate, 24, 1931–1949 (2011).

Von Storch, H. and Zwiers, F. W. *Statistical Analysis in Climate Research*. Cambridge University Press, Cambridge, 494 pp. (1999)

APPENDIX:

The figures show the rotated global SST EOF patterns used as predictors for the statistical forecast. Recent SST anomalies are projected onto these patterns to produce the statistical forecasts. For all 3 patterns, positive SST anomalies in regions with positive weights and negative SST anomalies in regions with negative weights favour above average rainfall and vice-versa.

The discriminant and regression forecasts are the average of 3 predictions from the following 3 sets of predictors:

1) Rotated EOF2, EOF 4 and EOF 5
2) Rotated EOF2, EOF 4 and EOF 5 + NW Indian Ocean box region
3) Rotated EOF2, EOF 4 and EOF 5 + NW Indian Ocean box region + Pacific box region.

The 2 box regions are

1) "NW Indian ocean" region 0-10N 50-60E + 0-10S,40-60E
2) "Pacific" region 0-20N, 150E-150W

The predictors are weighted as follows

Rotated EOF2, EOF4 and EOF 5 all 25% , NW Indian Ocean region 16.7%, Pacific region 8.3%.

The pattern shown in figure A3 (rotated EOF 5) is the most important predictor contributing to over 50% of the forecast variance.

Figure A1:

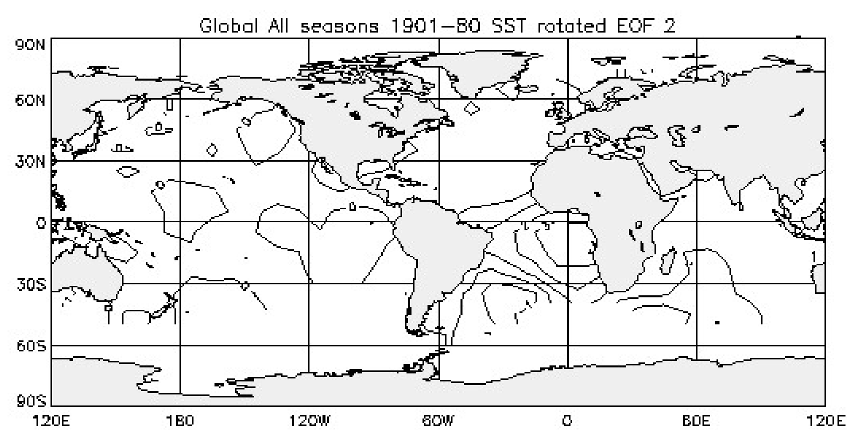


Figure A2:

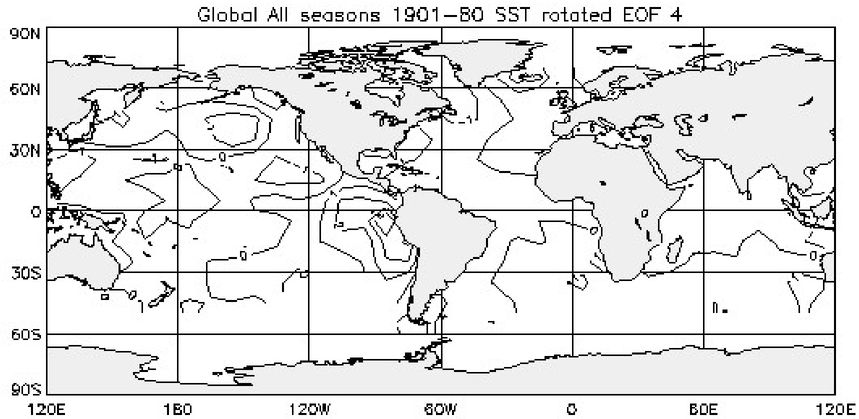
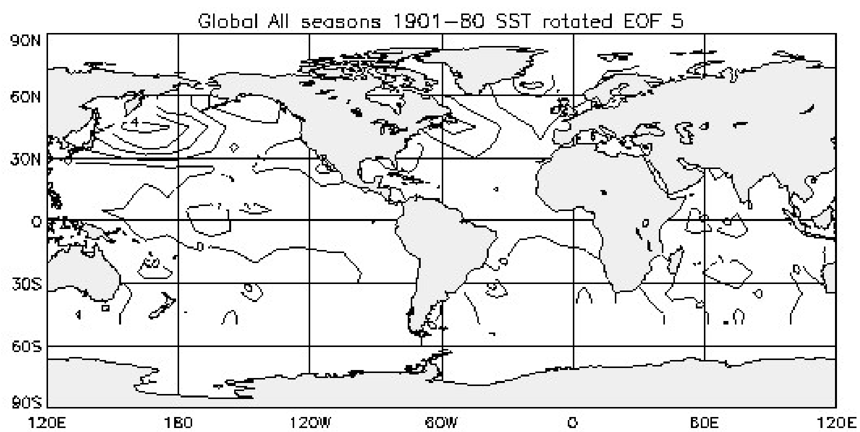


Figure A3:



Met Office:
FitzRoy Road, Exeter
Devon, EX1 3PB
UK

Tel: +44(0)1392 885680
Fax: +44(0)1392 885681
enquiries@metoffice.gov.uk
www.metoffice.gov.uk