1. The Great Africas Cup: ECMWF vs MOGREPS ensemble precipitation forecasts

Dataset: One rainy season (8.5 months) worth of ensemble forecasts from the ECMWF 51 member ensemble (50 members + control forecast), the ECMWF deterministic forecast, and the UK Met Office MOGREPS 24 member ensemble (23 members + control). Forecasts are for 24h accumulations, 12 UTC to 12 UTC from day 1 to day 10 (ECMEF), and day 1 to 6 (UK Met Office). All ensembles are presented matched to all available precipitation observations from the WMO Global Telecommunications System for stations in 6 Eastern Africa countries: Ethiopia, Burundi, Kenya, Tanzania, Rwanda and Uganda (Although I don’t think there is any Ethiopian data in the database). Station lat/long locations are included in case there is an interest in plotting a map of the stations included in the study

The dataset is fairly large and multidimensional: There are many forecasts for each station, each forecast projection. Verification results are normally obtained separately dor the different forecast projections, at least in short- and medium range forecasts, because the quality of the forecasts changes substantially with forecast projection. Therefore, it will be best to reconfigure the data so that forecast projections are separate. For the purposes of the exercise it is probably best to group all stations together for each projection, though the specific station numbers are included with each record so that the data can easily be filtered by station.

As a starting point for the students, I suggest the following:

-One file for each forecast projection time, with all events included. Each event (row) to consist of FCDATE, STEP (projection in hr), LAT, LON, STN\_ID, VALID TIME, OBS, ECDETFC, ECCNTRLFC, ECEnsemble forecast M1 to M50, then UKMETCNTRLFC, UKMET Ensemble forecasts M1 to M23. In other words, for each of the ECMWF forecast projection, day 1 to 6 only, station number and valid time, find the corresponding event in the UKMET dataset and add it on at the end of the row. Discard any ECMWF or UKMET events that can’t be matched with the corresponding event from the other model.

-Set up files for forecast projection times day 1 to 6 only even though EC has 10 day forecasts.

-include only the events and stations for which there is an observation and both ensembles are available.

-It would also be good to keep files of the data in this original form, but with only 2 files, one for ECMWF and one for UKMET. This just means compositing the data from each model in the existing format over all the months.

The dataset is raw ensemble output; the students may wish to evaluate the ensembles this way, but they will also need help setting up simple routines to estimate probabilities of precipitation over specific thresholds, as is normally done in ensemble forecasting, calculating the ensemble mean and spread.