Forecast skill

Worked example 13.1

Table of Contents

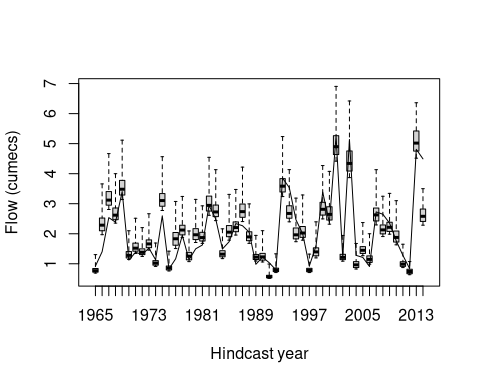
# Loading the Data

In this example we will be calculating the *skill* of the Ensemble Streamflow Prediction (ESP) forecast method for the Lambourn catchment at Shaw gauging station in southern England (catchment area of 234.1 km2). The dataset consists of monthly-averaged river discharge observations in the second column and corresponding ESP forecasts initialised on 1st January each year for a one month lead time for 51 ensemble members, in columns 3 to 53. The dataset contains 50 observation-forecast pairs across the 1965 to 2014 hindcast periods as shown in column 1. The units for both observations and forecasts are m3s-1. More detail on how these data were generated can be found in Harrigan et al. (2018).

First we load the dataset and define the observations and forecast variables:

# Load the data   
library(hydroDrought)  
rmarkdown::paged\_table(fcst\_dataset, options = list(cols.print = 5))

# Subset observations vector   
obs <- fcst\_dataset$Observed\_flow  
  
# Subset ensemble forecast matrix  
fcst\_ens <- fcst\_dataset[, 3:53]  
  
# Plot observed and ensemble forecast  
boxplot(t(fcst\_ens), xlab = "Hindcast year", ylab = "Flow (cumecs)",   
 xaxt = "n", range = 0)  
axis(side = 1, at = 1:50, labels = fcst\_dataset$Year)  
lines(obs)



# Calculate long-term low flow threshold from observations

In this example we want to evaluate the ability of the ESP forecast to predict the probability of river discharge events below the long-term Q90 low flow threshold. We need to first calculate the long-term Q90 threshold from observations using the lfquantile() [[1]](#footnote-22) function:

obs\_Q90\_threshold <- lfquantile(obs, exc.freq = 0.9)  
obs\_Q90\_threshold

## Q90   
## 0.9278767

# Calculate the Brier Score (BS)

The forecast evaluation metric we will use is the Brier score. The Brier score is essentially the mean squared error of the probability forecasts when the observation *obs1* = 1 if the event occurs, and *obs2* = 0 if the event does not occur. The score averages the squared differences between pairs of forecast probabilities and the subsequent binary observations,

where the index *k* denotes numbering of the *n* forecast-event pairs. In our example *n* = 50. Perfect forecasts will have *BS* = 0. The steps to calculate the BS are as follows:

* Create vector of binary observations *obs* with 1 when an an event is observed (i.e. below the long-term observed Q90 threshold for that year) and 0 if not
* Create the vector *fcst* of forecast probabilities (i.e. number of ensemble members below the long-term observed Q90 threshold for that year, expressed as a probability [0,1])
* In order to calculate the skill, we need a *benchmark* forecast against which to compare our ESP forecast A common benchmark forecast is the long-term climatology. In any given year the probability of the flow being below the Q90 threshold is simply 0.1 (i.e. the climatological probability). Therefore create the vector *bench* of climatological probabilities for each forecast (i.e. 0.1)
* The Brier score is then calculated using the brier() function from the verification R package

library(verification)  
  
# Vector of binary observations   
obs\_vec <- ifelse(obs < obs\_Q90\_threshold, 1, 0)  
  
# Vector of forecast probabilities (i.e. number of ensemble members below   
# threshold / total number of ensemble members)  
fcst\_prob\_vec <- apply(fcst\_ens < obs\_Q90\_threshold, MARGIN = 1, FUN = mean)  
  
# Vector of climatological probabilities for benchmark forecast, simply a   
# vector 0.1 (i.e. below Q90 threshold 10% of the time by chance)  
clim\_vec <- rep(0.1, 50)  
  
# Calculate Brier Score for ESP forecast  
BS\_esp <- brier(obs = obs\_vec, pred = fcst\_prob\_vec, bins = FALSE)$bs  
BS\_esp

## [1] 0.06038447

# Calculate Brier Score for climatology benchmark forecast  
BS\_bench <- brier(obs = obs\_vec, pred = clim\_vec, bins = FALSE)$bs  
BS\_bench

## [1] 0.09

# Calculate the Brier Skill Score (BSS)

We can compute the **Brier Skill Score** of the ESP forecast for predicting Q90 low flow events compared to a simple climatology benchmark forecast using the the generic skill score formula introduced in Section 12.3.8, with the following variables:

* *Afc* = BS\_esp
* *Abench* = BS\_bench
* *Aperf* = 0

# Brier Skill Score   
BSS <- (BS\_esp - BS\_bench) / (0 - BS\_bench)  
BSS

## [1] 0.3290615

1. Tobias: would use the function lfquantile() from the package. The names of the vector are easier to interpret. [↑](#footnote-ref-22)