Workflow

# Prepare Data

* CHIRPS (daily) and Hobbins RefET (dekadal) tifs wrangled into country stacks for entire historical time period.
* CHIRPS summed to dekadal totals
* CHIRPS and RefET converted to table form, rows are cell index, dekad, year, rain and ETo
* -9999 values set to NA
* AI calculated as Rain/ETo

# Seasonality from LT averages (CLIM.LT)

* Seasons classified using long-term averages only (both rain and AI)
* If AI<0.5 at start of season this is ignored in logic (F is converted to T)
* Sequences are merged if they meet at the year boundary (dekad 36 & dekad 1)
* Where AI is met in all 36 dekads then we choose the first instance where the rainfall rules are met in the three wettest months as onset

## Plotting of LTavg analysis:

* Min season length allowed = 1 dekad (i.e. any value allowed)
* Minimum seasonal rainfall allowed = 0mm (i.e. any value allowed)
* Where there are more than one seasons in a cell the a minimum gap length between seasons for bimodality is set (SeasonGapT). If gap between seasons is > SeasonGapT then seasonality is considered to be bimodal, if <= SeasonGapT then it considered to be a split season.
* Determine the order of seasons by the length of the distance between them, this can be used to determine how bimodal are merged.
* Assign seasons to first or second season by country:
  + Using all SOS data from all seasons and grid cells we use multimode function to determine seasons to determine the peaks of the SOS distribution for the country. In countries with >1 season, season membership of dekads is determined using the mid-points between SOS modes.
  + This doesn’t always work well, so for Kenya, South Africa and Mozambique dekadal season membership is assigned manually.
* Manually tidy up countries which are largely unimodal but have a few pixels that split the seasons
* Classify seasonality systems
  + Two seasons = bimodal (2 WS)
  + Two seasons close together = (1 WS – Split)
  + One season = 1 WS
  + The three season classes are then intersection with LGP to highlight areas where LGP is short (<=6 dekads) or long (>=30 dekads)

# Seasonal SOS/EOS Calculations

* 1. Find wettest 3 months:
     1. Rolling 3 month average taken over entire time series n:(n+8) then averaged by month
     2. 3 month sequence with most rain identified
     3. A second sequence of rainfall is then searched for, this is separated by at least 6 dekads from the first season (so padding in 3.b.ii does not overlap between season 1 and 2).
     4. Dekads in wettest months assigned labels
  2. Identify SOS and EOS and assign sequences
     1. Wet months padded by 3 dekads (15 in total)
     2. Wet months converted to new field of unique sequences
     3. Remove first sequence of first year and last sequence of last year to avoid boundary issues at the edge of the dataset timeframe
     4. Create logic columns for rainfall and AI requirements met or failed
        1. Logic for calculation of AI as long-term average or annually included, LT average used.
     5. Assign sequences within wet months using rainfall and AI logic columns, multiple or fragmented sequences can be generated within a season
  3. Merge sequences (optional):
     1. If 1 sequence present nothing happens
     2. Sequences separated by 1 dekad are merged
     3. If the first sequence in a series (i.e. within a padded wet period) is 1 dekad and separated from the next sequence by >1 dekad (i.e. it would not be merged with the next sequence) it is considered a false start and removed.
  4. Statistic are derived from sequences (i.e. rainy seasons):
     1. SOS is first dekad of a sequence
     2. EOS is last dekad of a sequence
     3. EOS-SOS = LGP (takes into account year end boundary)
     4. Where EOS = SOS, sequences are removed.
     5. Last sequence in last year of data is removed (sequence might extend beyond available data)
     6. Total rainfall summed for each season

## Cleaning of Seasonal SOS->EOS Sequences:

* 1. Roll back SOS where SOS is fixed (e.g. where growing season length is very long and padding is insufficient to capture the starting dekad)
     1. If any one SOS dekad accounts for more than 95% of SOS and failed seasons are <20% of data:
        1. Scenario 1 - SOS fixed and one season present. Here we simple increase the backward padding of the rainy months by a factor of 2 (i.e. 6 dekads instead of 3).
        2. Scenario 2 - SOS fixed and two seasons present.:
           1. Calculate season separations
           2. Assign leading/following seasons based on gap lengths between seasons
           3. Flag which seasons are fixed
           4. Scenario 3: Seasons are adjacent (min separation of 1 dekad) & leading season is not fixed then do nothing.
           5. Scenario 4: Seasons are adjacent (1 dekad apart) & leading season is fixed (i.e. date need adjusting back) then we adjust the start of the season window for both season 1 and season 2 (both shifted by another 3 dekads backwards). This should help balance the lengths of the two seasons where the rainy season is long enough to accommodate two growing seasons.
           6. Scenario 5: Seasons are not adjacent (min separation of >1 dekad) we adjust the rainy season window only for the fixed season but not beyond EOS of other season (non-fixed).
           7. I don’t think we have a scenario where there were two non-adjacent fixed seasons in a cell.
     2. After adjusting the season padding we rerun step 3 for the adjusted indices (cells).
  2. Remove second seasons that are present for less than a threshold proportion years. I arbitrarily used a third as the cut-off.
  3. I have added additional logic to look for a 3rd season in the gap between season 1 and season 2, I haven’t really done anything with these data as the conditions are rarely met.

1. Summary statistics of seasonal data per cell per season:
   1. Mean, median, mode, min, max, sd calculated for EOS, SOS, LGP per cell.
   2. Total seasons (where minimum onset conditions were met)