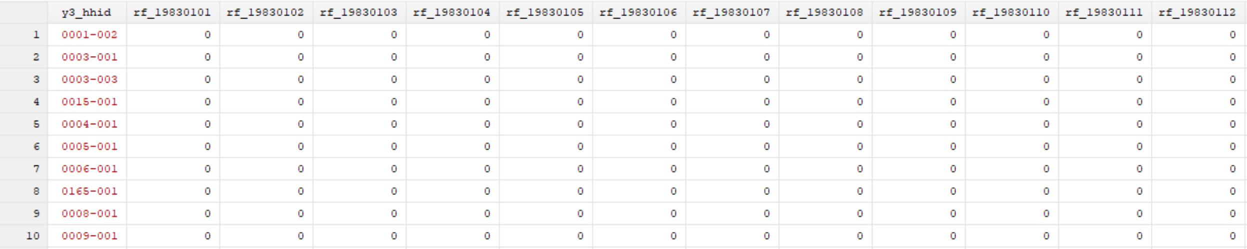
Explanation weather command

Data

The command can be used with either rainfall or temperature data from any source. The only requirements are that:

1. The data is "wide," meaning each location is a row and each column is rainfall or temperature reading from a different day.
2. The data is measured daily.
3. The variable names for each column contain yyyymmdd. The variable names can have any prefix but must contain the year, month, and day.

An example of how the data set should be set-up is below.



	y3_hhid	rf_19830101	rf_19830102	rf_19830103	rf_19830104	rf_19830105	rf_19830106	rf_19830107	rf_19830108	rf_19830109	rf_19830110	rf_19830111	rf_19830112
1	0001-002	0	0	0	0	0	0	0	0	0	0	0	0
2	0003-001	0	0	0	0	0	0	0	0	0	0	0	0
3	0003-003	0	0	0	0	0	0	0	0	0	0	0	0
4	0015-001	0	0	0	0	0	0	0	0	0	0	0	0
5	0004-001	0	0	0	0	0	0	0	0	0	0	0	0
6	0005-001	0	0	0	0	0	0	0	0	0	0	0	0
7	0006-001	0	0	0	0	0	0	0	0	0	0	0	0
8	0165-001	0	0	0	0	0	0	0	0	0	0	0	0
9	0008-001	0	0	0	0	0	0	0	0	0	0	0	0
10	0009-001	0	0	0	0	0	0	0	0	0	0	0	0

Command: set up

Using data sets as defined above, the weather command creates useful statistics in the same fashion for all years. The command together with an example do file is in the zipped folder *weather_command*. Place the file *weather.ado* into wherever Stata stores your .ado files. The file *weather_wrapper.do* provides an example of the syntax and how to run the command.

The general syntax of the command is as follows.

i. After the command name, one has to define what variables contain the rain/temperature information. I provide examples for CHIRPS and ECMWF. For the CHIRPS datasets, the prefix on the weather variables is pic_ while in the case of ECMWF the prefix is y_.

ii. Next, one needs to tell the command whether the data is *rain_data* or *temperature_data*.

iii. One then has to select a season to study using the options *ini_month(*number*)*, *fin_month(*number*)* and *day_month(*number*)*, if not specified, the default is the first day of the month. For example, in the example .do file I chose a season from the middle of March to the middle of June. At this moment, we must be careful with seasons that contain months that contain two different years, such as November to February. In this case some renaming of the variables is necessary. (Sorry about that…)

iv. The option *keep* tells the command to keep the variables it creates plus some of the original variables in order to match them with other datasets.

v. The save of option tells the program to save the dataset in a given location with a given name.

Command: Results

i. Rainfall variables

When the rainfall option is chosen, the command generates the following variable for each season:

• Mean daily rainfall:	mean_season_`year’
• Median daily rainfall:	median_season_`year’
• Standard deviation of daily rainfall:	sd_season_`year’
• Cumulative rainfall for the season:	total_season_`year’
• Skew of daily rainfall:	skew_season_`year’
• Number of days with less than 1mm of rainfall:	norain_`year’
• Number of days with more than 1mm of rainfall:	raindays_`year’
• Percentage of days in a season with more than 1mm rainfall:	raindays_percent_`year’
• Longest intra-season period with less than 1mm of rainfall:	dry_`year’

In addition to these basic statistics, the command also calculates the long-term averages for total_season, norain, raindays, and raindays_percent. It then uses these to generate variables that measure each seasons deviation from the long-term average and the deviation measured as a z-score.

ii. Temperature variables

When the temperature option is chosen, the command generates the following variable for each season:

• Mean daily temperature:	mean_season_`year’
• Median daily temperature:	median_season_`year’
• Standard deviation of daily temperature:	sd_season_`year’
• Skew of daily temperature:	skew_season_`year’
• Maximum daily temperature:	max_season_`year’
• Growing degree days for the season:	gdd_`year’

Growing degree days are calculated using the options *growbase_low(*number*)* and *growbase_high(*number*)* to determine the number of days where the temperature was between that range. As with the rainfall option, the temperature option also generates deviations in GDD from the long-term average and the deviation measured as a z-score.

Following Schlenker/Roberts, the command also calculates temperature bins as the percentage of days that fall in each temperature quintile during the season. It then generates the variables tempbin`quintile’`year'.