**Price Data**

Original Dataset (Local only, Haiti removed)

Data Points: 263452

Date range: 2002-01-31 - 2017-12-31

Data Points for 3 Commodities in 3 Countries in abridged date ranges (Sorghum, Millet, Maize. Chad, Mauritania, Nigeria) : 31,315

Note: These are the *deflated* prices (def\_value)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Country** | **Type** | **Min Price** | **Max Price** | **Mean Price** | **Min Date** | **Max Date** |
| Chad | Maize | 92.035124 | 270.28656 | 200.84216 | 2012-05-01 | 2017-11-01 |
| Chad | Millet | 100.798723 | 383.47968 | 207.41094 | 2011-02-01 | 2017-11-01 |
| Chad | Sorghum | 67.352403 | 267.10515 | 143.79524 | 2009-03-01 | 2017-09-01 |
| Mauritania | Maize | 5.809351 | 19.60656 | 11.29252 | 2012-05-01 | 2017-12-01 |
| Mauritania | Millet | 4.699543 | 26.11826 | 11.38352 | 2011-04-01 | 2017-12-01 |
| Mauritania | Sorghum | 5.377839 | 31.53937 | 13.17781 | 2009-04-01 | 2017-09-01 |
| Nigeria | Maize | 18.182416 | 579.95596 | 61.28171 | 2012-06-01 | 2017-11-01 |
| Nigeria | Millet | 18.329048 | 180.70759 | 72.19422 | 2011-02-01 | 2017-12-01 |
| Nigeria | Sorghum | 5.889729 | 273.06924 | 69.9122 | 2009-04-01 | 2017-09-01 |

FEWS Reporting Bias:

Are the countries that only grow one type of commodity actually not growing the other three?

Djibouti only shows sorghum data in FEWS dataset, and FAO does not show either maize or millet data for this country.

Angola & Malawi show Maize data in FEWS set, but no Sorghum or Millet. FAO numbers for 2017 state that they do produce the other two crops in smaller quantities:

A close up of a white background

Description automatically generated

These countries were left out of the analysis for now.

Price Distribution: see next page

A close up of text on a white background

Description automatically generated

**Precipitation Data**

Average rainfall (in mm)

A picture containing scoreboard

Description automatically generated

Precipitation distributions

A close up of a map

Description automatically generated

Markets

Full list of markets is on the next page

A screen shot of a computer

Description automatically generated

Spatial distribution of markets analyzed

A close up of a map

Description automatically generated

**Analysis**

Regressions were run based on 5 variables: Commodity, Buffer, Time Window, Aggregation, and Zero/No Zeros.

Commodities

* Sorghum
* Millet
* Maize

Buffers (radial distance from city center):

* 25km Note that I also ran the analysis for 50km and 75km but the results were identical to the 25km buffer.
* 100km
* 200km

Time Windows

* P\_sow: sowing season (varies for each country)
* P\_grow: growing season (varies for each country)
* P\_harv: harvest season (varies for each country)
* P\_sup: supply chain period (all time from harvest season to the date of sale)
* P\_onemonth: one month of precipitation (one month before the date of sale)

Aggregation

* Average: average rainfall for the window of time
* Accumulated: accumulated rainfall for the given window

Zero/No Zeros

* Zero: includes data points where rainfall was 0
* No Zero: does not include data points with less than 1mm of rainfall

**Model Selection + Plots**

Each set of data was fit with a log\_linear, log\_log, quadratic, and cubic function.

The best model was selected based on a few criteria:

* All coefficients must be significant (pvalue less than 0.05) therefore likely to be a meaningful addition to the model
* All model AIC values were compared. The lowest value was selected *unless…*
* If the difference between the lowest and second lowest AIC value was less than 4, the simplest model of the two was selected. (Method taken from Burnham and Anderson, 2004)

The data was then bootstrapped with the ideal model. Plotted is the median estimate across bootstraps, and the errors are the estimates at the 95th and 5th percentile.

Color of plot coordinates to the buffer size.

Red = 25km

Yellow = 100km

Blue = 200km

To make decision about which of the aggregation/zeros/buffers to use:

Buffers

* If all three are the same model
  + lowest AIC value
* If they are different
  + Check to see if the model difference comes from sparse data at the extremes. If so, still lowest AIC