MAT 375 Mini Project 3: Modeling Rainfall and Temperature in Kouma-Konda

Jacob Englert, Joey Stevens, and Austin Campbell

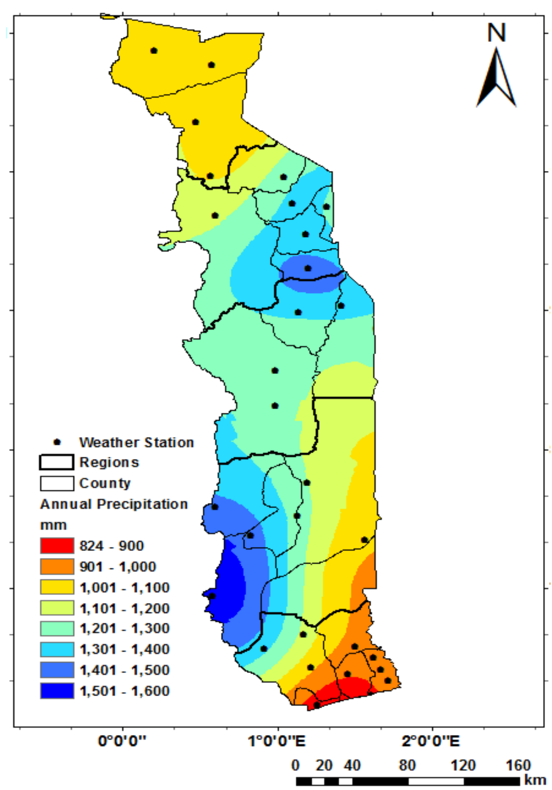
Introduction

Kouma-Konda, Togo

Kouma-Konda is an African mountain village located in the lush forest region of Togo. It is near Lake Volta, close to the Togo-Ghana border. The village of Kouma-Konda was established in 1920 following family disputes that divided part of the Kuma-Adamé population. In past years, dense mountain forest covered most of the Kouma-Konda township. More recently, poverty has pushed some of the villagers to exploit the forest resources. With bush fires, cutting down trees for charcoal, as well as clearing to create new agricultural parcels has all attributed to a declining forest. Tourism has also led to increased deforestation. The climate is relatively uniform year-round. From May to June and from August to September, rains are abundant and even torrential at times. Kouma-Konda is situated in the intensive agricultural region which is likely the case for this torrential rainfall.

Problem

Temperatures in Togo as a whole, not just the city of Kouma-Konda, are being studied to determine whether significant changes are occurring over time. Preliminary work has concluded about the city of Kouma-Konda that average monthly and average annual minimum temperatures are indeed increasing. If this coincides with a change in rainfall, the agricultural scene in Kouma-Konda could be affected greatly. The purpose of this research is to further examine the temperature data, as well as to explore and discover any changes in newly acquired rainfall data over the course of time. Changes in rainfall can lead to public health, agricultural, and many other issues if not accounted for by the people of Togo and the proper authorities. If rainfall ends up showing interesting behavior, it could be useful in examining the temperature changes as well.



Summary of Earlier Work

Previous research concerning modeling climate for the Togolese city of Kouma-Konda is detailed in two reports: *Mini Project 1* (Dufek and May) and *Evaluation of Average Monthly Surface Temperature in Kouma-Konda, Togo* (Compton and Ficke).

The goal of Dufek and May's research was to determine whether or not annual minimum and/or annual maximum temperatures were rising in the city over time. They chose to fit simple linear regression models to both sets of temperature data. Evidence was found to support the hypothesis that average minimum annual temperatures have significantly increased since 1961 in Kouma-Konda (p-value < 0.000, R^2 = 67.8%). A simple linear regression was also used to fit the maximum temperature data, however there was not enough evidence to support the hypothesis that average maximum annual temperatures have increased since 1961 (p-value = 0.832).

Compton and Ficke conducted similar analyses, except the data used was further refined into monthly averages. Using ANOVA tests, it was found that the month was a significant predictor for both minimum and maximum temperatures. Thus, models were constructed with the interpolated data. These models utilized sine and cosine functions with periods of one year to account for the seasonal variation in temperatures. The best model predicting minimum temperatures from time was the linear model incorporating sine and cosine oscillations (R^2 = 57.36%) This model indicates an increase in average minimum temperatures over time. The best model predicting maximum temperatures from time was the quadratic model incorporating sine and cosine terms (R^2 = 84.48%). However, Compton and Ficke concluded this model did not show evidence of any significant change in average monthly maximum temperatures over time.

In summary, preliminary work on modeling temperatures in Kouma-Konda, Togo have shown significant increases since 1961 in average minimum temperatures and insignificant changes since 1961 in average maximum temperatures.

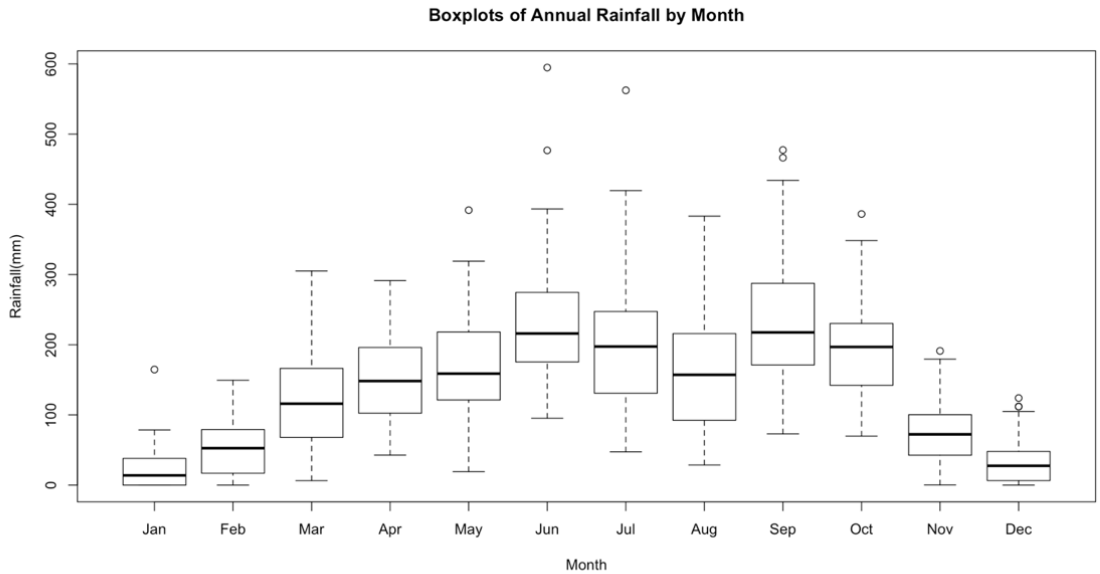
Data

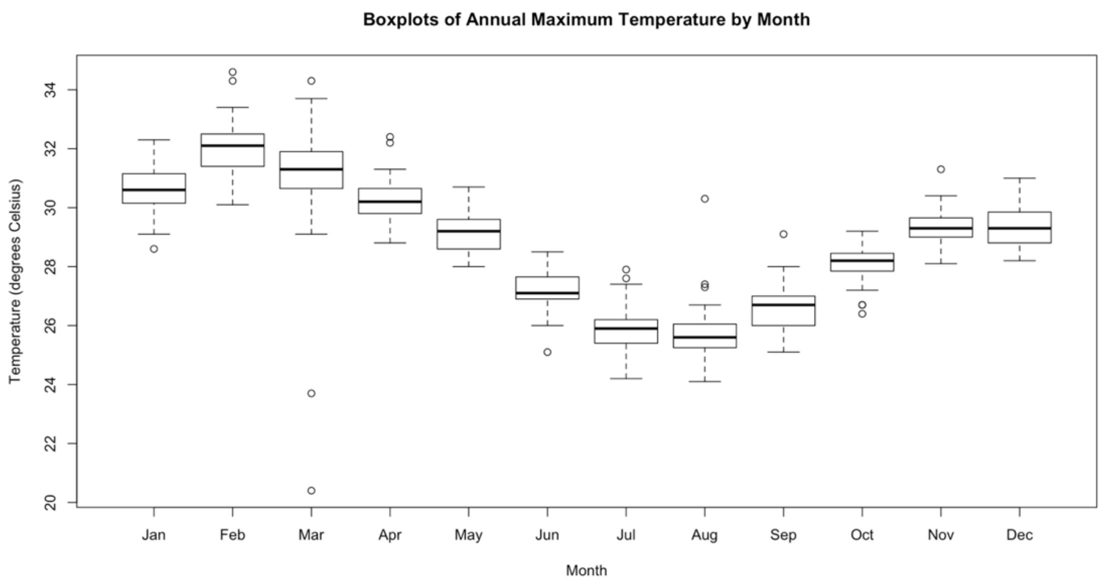
Discussion of Data and Data Integrity

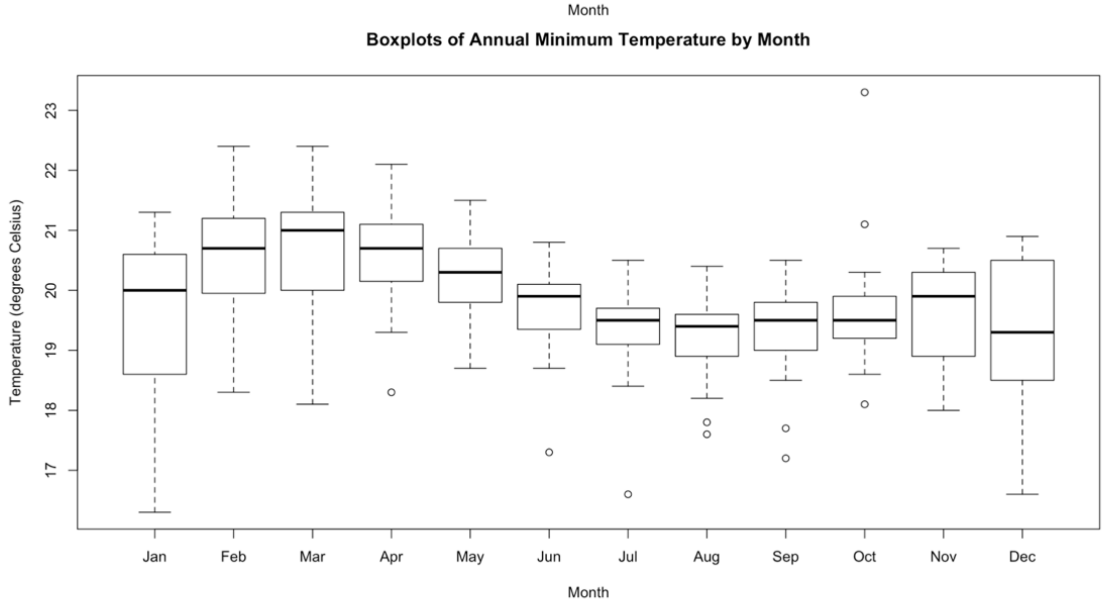
The dataset we have been provided includes monthly rainfall and temperature observations beginning with January 1961 and ending with December 2015. Prior to beginning our analyses and modeling, we examined the data and noticed several potential problem observations. These issues will be discussed individually in the following four subsections:

Outliers

On the following page are boxplots of each of the datasets for temperature and the dataset for rainfall. Outliers on these plots (denoted by points beyond the whiskers on either side of the corresponding month's box) were removed prior to the model building phase when necessary. (i.e. models including temperature exclude observations with outliers in either the maximum or minimum temperature datasets).


Boxplots of Annual Rainfall by Month
Rainfall(mm)
0
100
200
300
400
500
600
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Month


Boxplots of Annual Maximum Temperature by Month
Temperature (degrees Celsius)
20
22
24
26
28
30
32
34
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Month


Boxplots of Annual Minimum Temperature by Month
Temperature (degrees Celsius)
17
18
19
20
21
22
23
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Month

Duplicates

The integrity of the data is questionable at best with several instances of duplicates in both the sets of temperature observations. The observations are monthly averages, so we would expect matching months to have similar observations. Even if temperatures have been shown to be increasing over time, successive years may still yield similar observations for the same months. However, when the same number (recorded to a tenth of a degree Celsius) is the observation for several consecutive months and/or several consecutive years for the same months, we consider this highly unlikely to have been the case. In particular it appears that much of the years 1974 and 1975 have matching values in both sets of temperature data. This issue should be brought up to the Togolese, however it will not exclude these observations from our analyses. Other suspect observations that were noticed can be seen in Appendices A and B.

Missing Observations

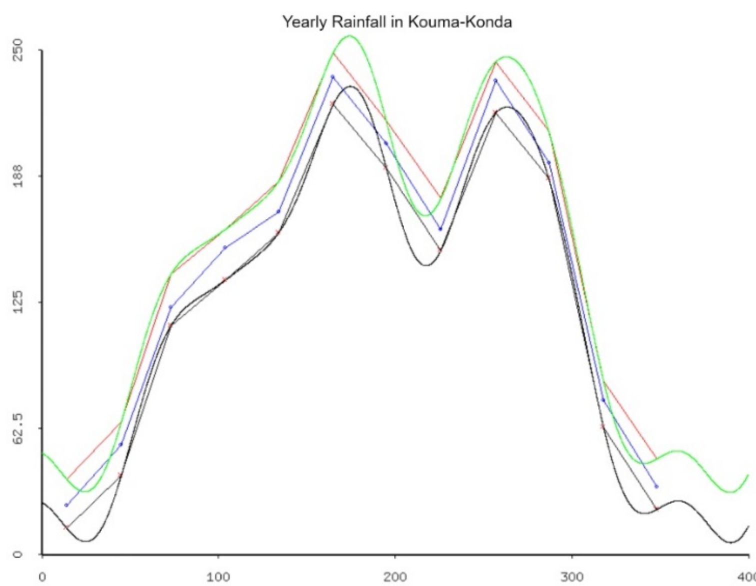
In the rainfall dataset, there are 20 missing observations. Most alarmingly, there is a twelve month stretch where no observation was recorded (July 1971 – Jun 1972). Due to the model building in this report using Rainfall as the dependent variable, all of these observations were omitted from the analyses. See Appendix C for all missing observations.

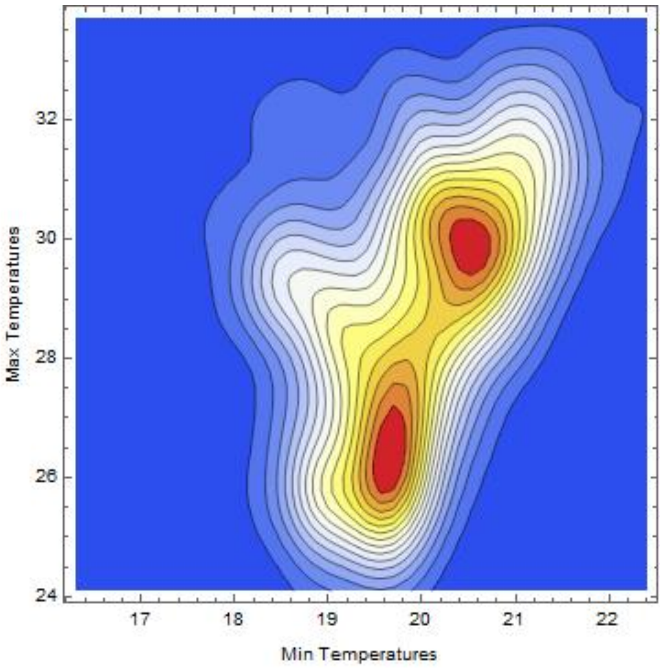
Nonsense Observations

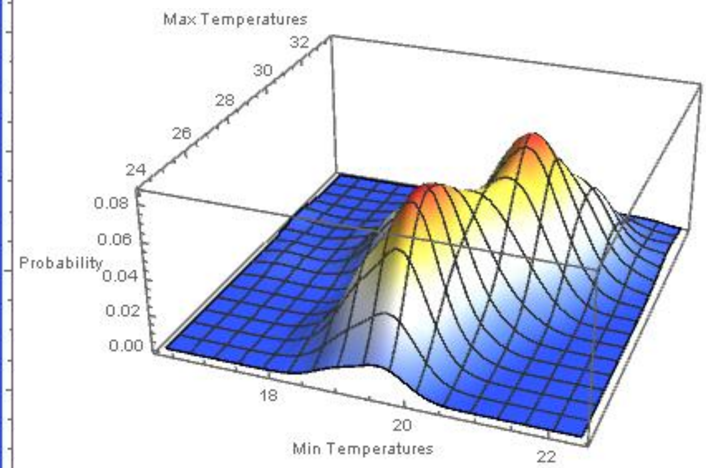
In the maximum temperature dataset, there appears to be one obviously nonsensical observation. In March 2005, the average maximum temperature was recorded as 20.4°C while the corresponding minimum was 21.3°C. The maximum temperature cannot be lower than the minimum temperature for any given month, so there is clearly an error here. This observation is excluded from our analyses.

Visualizations

The graph located to the right shows the monthly average rainfall in Kouma-Konda, and our model, for a one-year period. Along the x-axis are the days of the year, and along the y-axis is the rainfall measured in millimeters. This graphic allows us to look for when the “rainy season” occurs in Kouma-Konda. As you can see, there appears to be two separate peaks in rainfall: near the months of May/June and during the months of September/October. We found this to be very interesting information because there are usually not two separate periods of intense rainfall for a certain region. This poses the question of whether temperature is a factor in explaining these peaks.

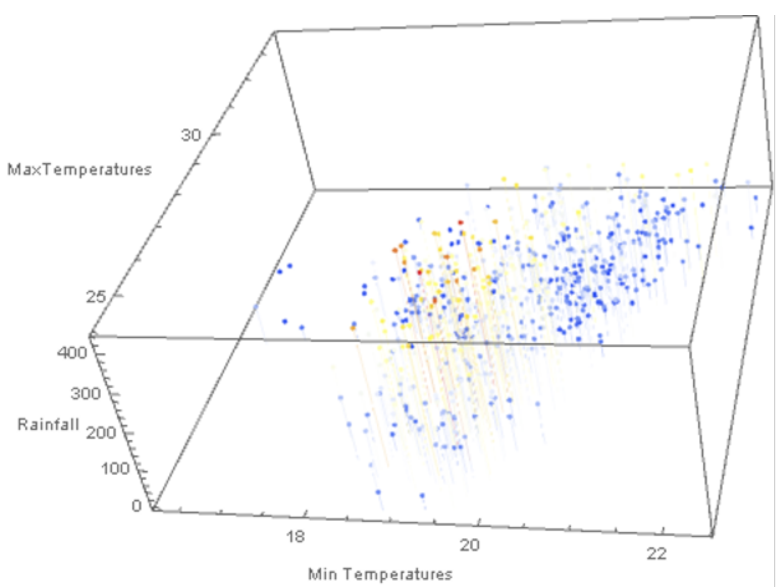






The two plots above illustrate the probability density function of the monthly minimum and maximum temperatures. The two peaks represent the highest frequency combinations of maximum and minimum temperatures. We can also see that the monthly minimum temperatures are less variable in general than the maximum temperatures.

The plot below displays the rainfall for each month given the maximum and minimum temperatures. There appears to be a pattern that indicates larger rainfall observations occur at lower minimum temperatures (which often coincide with lower maximum temperatures). This suggests that as the monthly minimum temperature increases, a behavior that has been modeled in the past, a general decrease in the amount of monthly rainfall is observed.

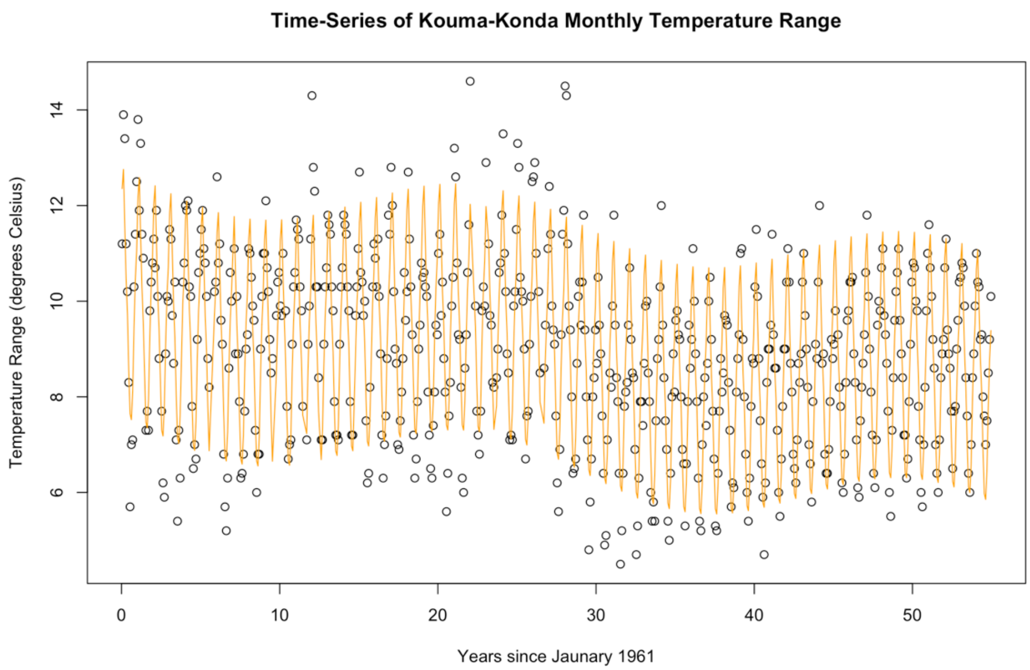


Model Building

All prior research regarding modeling the change in temperature over time for Kouma-Konda utilized average minima and average maxima temperatures. Instead of attempting to reinvent the wheel, we chose to explore a different model in this realm. Rather than predicting minimum or maximum temperature, we shifted our focus to the temperature range. The temperature range for a given month can be defined as the difference between the average monthly maximum and average monthly minimum for the month. Using linear regression, we were able to fit a model that uses temperature range (°C) as a function of time (years), including several periods representing sub-year oscillations to capture seasonal fluctuation. This model provides a different perspective to the data as we are now focusing on maximum variability in temperature as opposed to the temperature themselves.

We also opted to model the newly acquired rainfall data. This data is in millimeters and was recorded over the same period of time as the temperature data. Using time (in years) as the explanatory variable, as well as various periods of sub-year oscillations to capture seasonal fluctuation in rainfall, the dependent variable rainfall is modeled.

Temperature Range Model



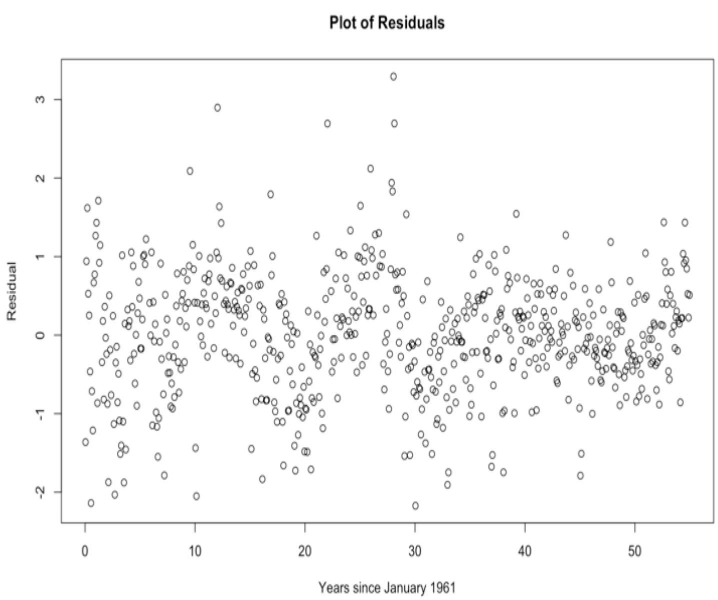
Prior research, as indicated above, has shown that maximum temperature in Kouma-Konda are relatively stationary over time, whereas minimum temperatures are growing. Because of this, we would expect to see the general range in monthly temperatures to be shrinking as well. This is exactly what we see in this model.

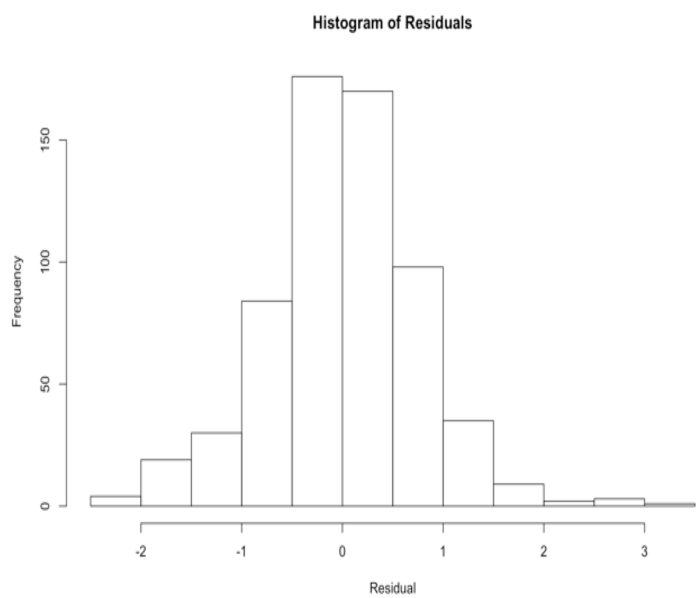
Diagnostics

	Estimate	Standard Error	t-Statistic	P-Value	95% CI
Constant	10.137824	0.095874	105.742	< 2e-16***	(9.949548, 10.326101)
t (years)	-0.061768	0.007896	-7.823	2.23e-14 ***	(-0.077274, -0.046263)
t^2	0.000475	0.000138	3.435	0.000632 ***	(0.000203, 0.000746)
Cos(2πt)	1.960436	0.042952	45.643	< 2e-16***	(1.876087, 2.044785)
Sin(2πt)	1.297885	0.043514	29.827	< 2e-16***	(1.212433, 1.383337)
Cos(2πt*4)	-0.142408	0.043340	-3.286	0.001074 **	(-0.227519, -0.057297)
Sin(2πt*4)	0.069560	0.043089	1.614	0.106962	(-0.015058, 0.154177)
Cos(2πt*3)	-0.133832	0.043213	-3.097	0.002043 **	(-0.218694, -0.048970)
Sin(2πt*3)	0.247911	0.043224	5.736	1.52e-08 ***	(0.163029, 0.332794)
Cos(2πt/28)	0.128754	0.044694	2.881	0.004104 **	(0.040983, 0.216524)
Sin(2πt/28)	-0.581099	0.046458	-12.508	< 2e-16 ***	(-0.672334, -0.489864)

**p-value < 0.01 ***p-value < 0.001

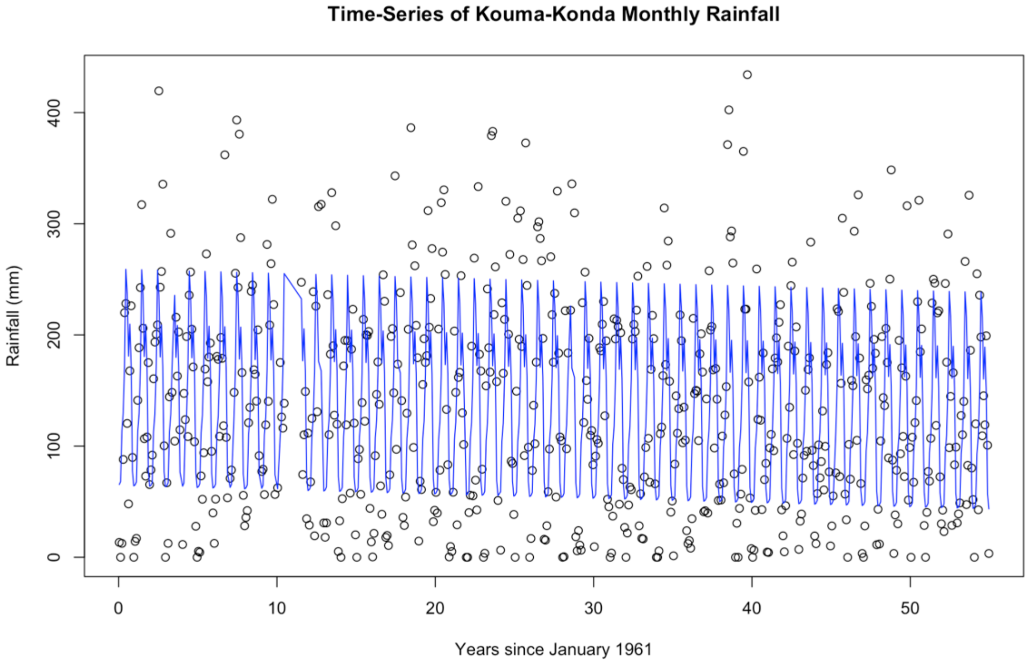
$R^2 = 0.846$ – 84.6% of the variation in Kouma-Konda monthly temperature range is explained by the quadratic oscillating model relating temperature range to time. This is a strong relationship.





The plot of residuals appears to be a random scatter about a mean of 0 and the histogram of residuals indicate the errors follow a normal distribution with a mean of 0. Thus, we conclude that the assumptions of constant variance and normality are met.

Rainfall Model

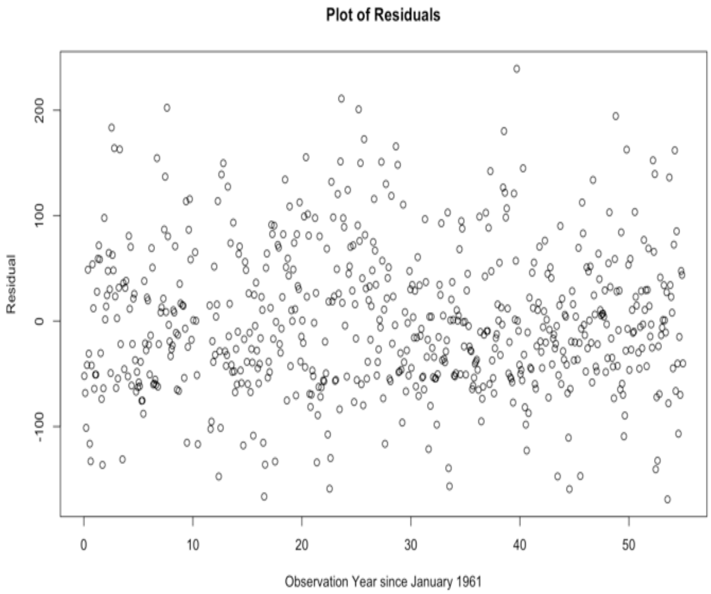


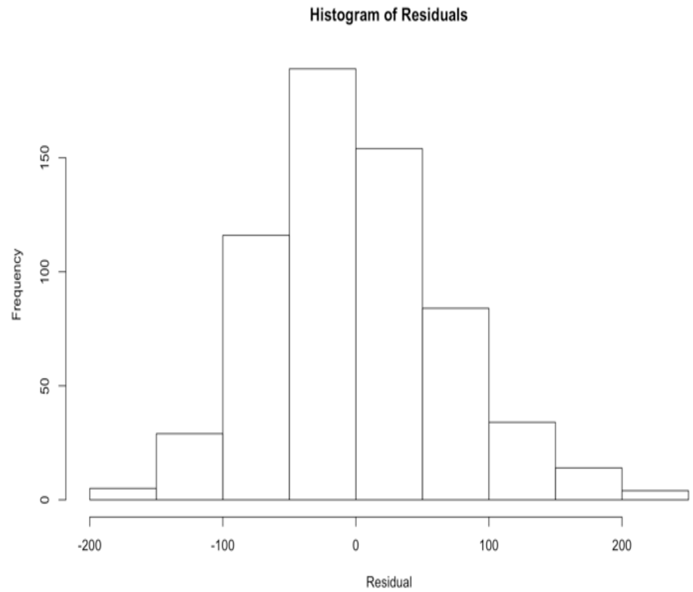
Diagnostics

	Estimate	Standard Error	t-Statistic	P-Value	95% CI
Constant	146.052	5.550	26.315	< 2e-16***	(135.152457, 156.950764)
t (years)	-0.383	0.172	-2.226	0.02635 *	(-0.720822, -0.045175)
Cos(2πt)	-85.617	3.866	-22.145	< 2e-16***	(-93.209124, -78.024361)
Sin(2πt)	-23.751	3.850	-6.169	1.24e-09 ***	(-31.312159, -16.190275)
Cos(2πt*4)	21.351	3.866	5.523	4.90e-08 ***	(13.759828, 28.942765)
Sin(2πt*4)	-6.143	3.851	-1.595	0.11123	(-13.705680, 1.420378)
Cos(2πt*3)	-12.645	3.858	-3.278	0.00111 **	(-20.221460, -5.068846)
Sin(2πt*3)	16.846	3.858	4.366	1.48e-05 ***	(9.269415, 24.422254)

*p-value < 0.05 **p-value < 0.01 ***p-value < 0.001

R^2 = 0.4884 – 48.84% of the variation in Kouma-Konda monthly Rainfall is explained by the linear oscillating model relating rainfall (mm) to time. This is a moderate relationship.





The plot of residuals appears to be a mostly random scatter about a mean of 0 and the histogram of residuals indicate the errors follow a normal distribution with a mean of 0. Thus, we conclude that the assumptions of constant variance and normality are met.

Conclusion

In conclusion, two models were developed for Kouma-Konda, Togo: one predicting monthly temperature range as a function of time (in years), and another predicting monthly rainfall as a function of time (in years). Both models included various sub-year period oscillations to account for seasonal variation in the response, and the first included a 28-year period to account for what appear to be long-term fluctuations in temperature range.

After accounting for these fluctuations, both models found an overall decrease in their respective responses. So, it appears that as the variation, as measured by range, in monthly temperatures has been decreasing over time, and so has the monthly rainfall. This idea is supported from a meteorological perspective because the largest periods of rainfall typically occur when conflicting warm and cool air masses collide. Thus, when the “conflict”, or difference, between warm and cool air masses is less prominent, less rainfall occurs.

Suggestions for Future Work

The measure for monthly temperature variation used in this research was the range between a monthly high and a monthly low. However, the range is not necessarily the best statistic to use as the monthly high likely occurred at a different point in the month than the monthly low. Despite this issue we believed the range to be a reasonable estimate of monthly temperature variability, although It would be interesting to see how models predicting other measures of variation (i.e. standard deviation) would perform.

We still have a few questions for the Togolese. Firstly, we would like to know how the data was collected and if they have any explanation for the apparent duplicate values in the temperature data. Most notably, we would like to enquire about the apparent issue with the 1974-1975 temperatures being nearly identical. We are also interested in determining why there is a twelve month stretch of missing rainfall data. Lastly, we are curious as to why there appears to be a 28-year oscillatory period in our temperature range model. What, if anything, would cause this term to be significant in predicting temperature data?

Interesting sources for information on Kouma-Konda:

http://www.netmaps.net/Netmaps/vegetation-vector-map/

https://eros.usgs.gov/westafrica/ecoregions-and-topography/ecoregions-and-topography-togo

http://medcraveonline.com/IJH/IJH-01-00019.php

http://www.sitamag.com/html/05_afrique/06_renc_africaines/08_kouma_konda_chateau_viale.html

http://urgenceafrique.org/fr/mission-togo-kouma-konda