Sally Dufek, Connor Edwards

Are Temperatures Rising in Togo?
Sotouboua

Introduction

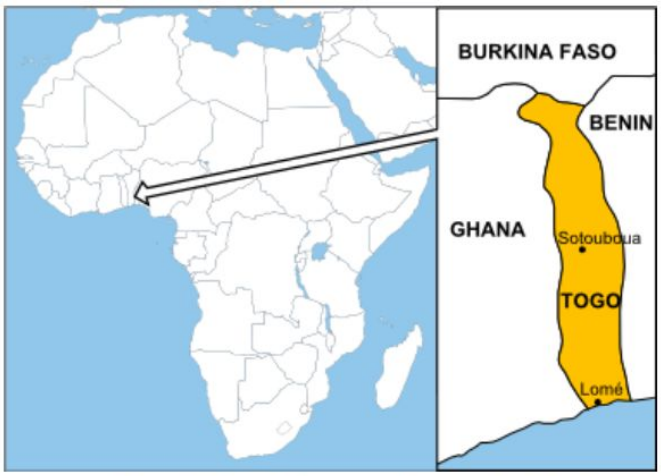


Figure 1: Map of Togo and Surrounding Nations

According to Jacob Englert and Adam May, "Sotouboua is a town located in Togo, a Western African nation that gained its independence from France in April 1960. Geographically, Sotouboua is near the center of Togo both in terms of longitude and latitude. It is considered a mid-sized town at a population of approximately 21,000. Sotouboua is relatively flat, with an elevation of 650 feet above sea level, and is about average when it comes to rainfall among its Togolese counterparts, receiving an annual average of 1,318.5 mm of rain."

Figure 2: Contour Map of Togo

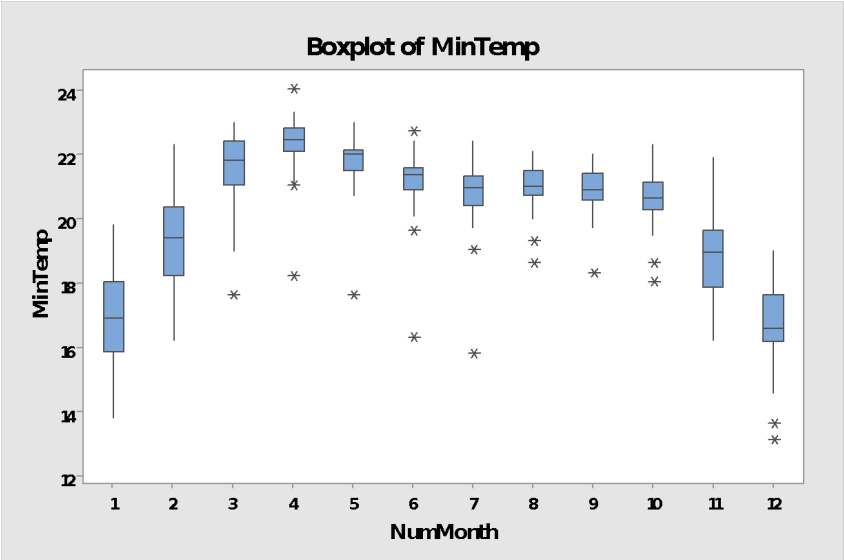
The Problem

We have been tasked with the job of determining whether Togolese Rainfall have experienced a significant change over time, specifically in the town of Sotouboua. We are also interested in the relationship between temperature and rainfall. Large changes in rainfall have the potential to not only wreck a nation's economy and ecosystems, but also to cause long-lasting climate problems.

Data Preparation

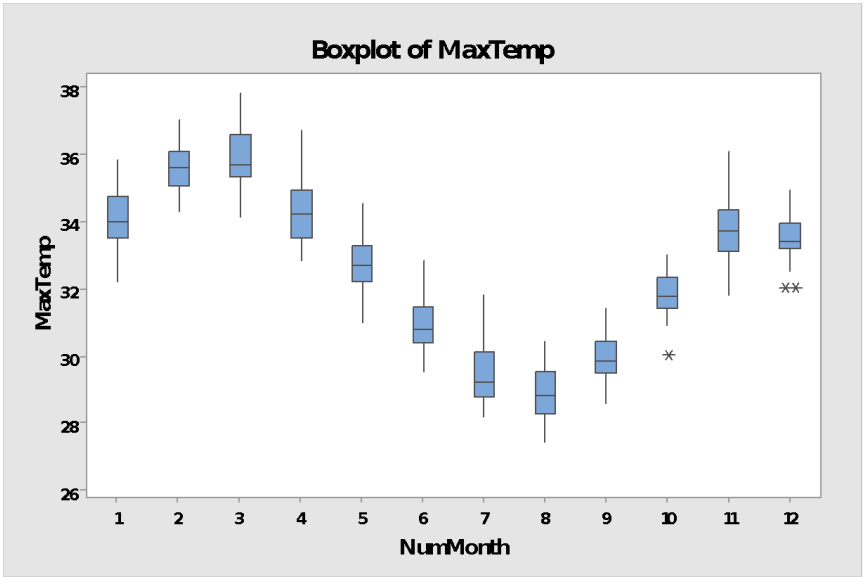
The temperature and rainfall data is quite messy and potentially unreliable. To identify potential outliers we

created boxplots of the three datasets by month, as recommended by the previous group. We then classified observations as outliers if they were less than the first quartile of that month - 1.5*IQR of that month or if they were more than their third quartile of that month + 1.5*IQR of that month. If an observations was identified as an outlier in any of the datasets, it was subsequently removed before analysis.



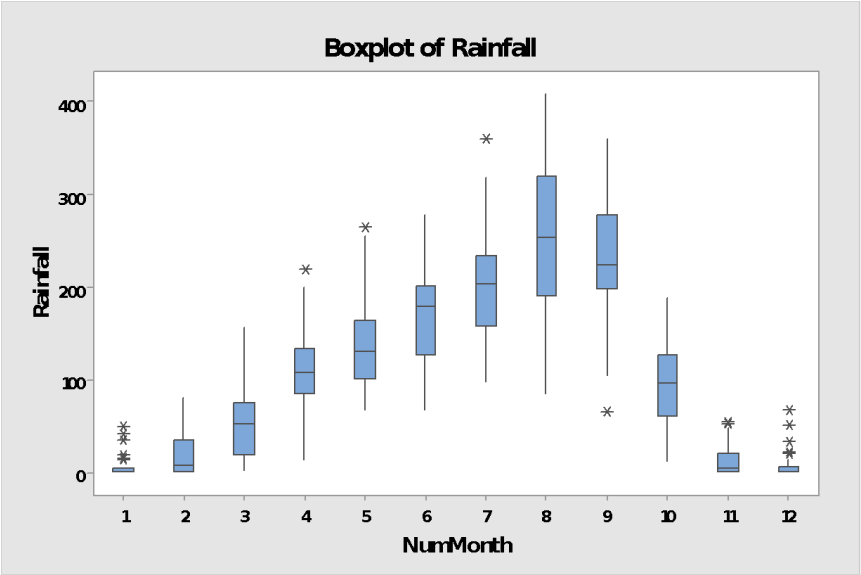
Monthly Min Temp Outliers

Year	Month	MinTemp
1986	Aug	19.3
1990	Oct	18.6
1998	Apr	24
2003	Apr	21
2003	Jun	19.6
2003	Jul	19
2003	Aug	18.6
2003	Sep	18.3
2003	Oct	18
2003	Dec	13.1
2004	Mar	17.6
2004	Apr	18.2
2004	May	17.6
2004	Jun	16.3
2004	Jul	15.8
2006	Dec	13.6
2014	Jun	22.7



Monthly Max Temp Outliers

Year	Month	MaxTemp
1986	Dec	32
2002	Oct	30
2004	Dec	32



There were 39 outliers in total removed from the dataset. In addition, we are missing the 2013 January and February minimum and maximum temperature data. Therefore, these two time points were also removed from the rainfall data.

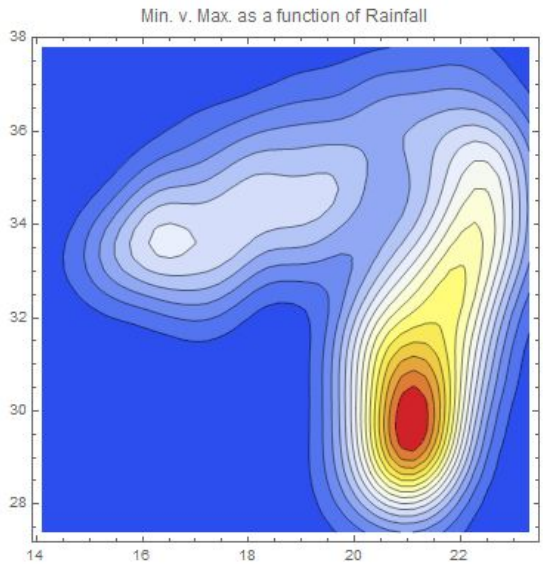
Monthly Rainfall Outliers

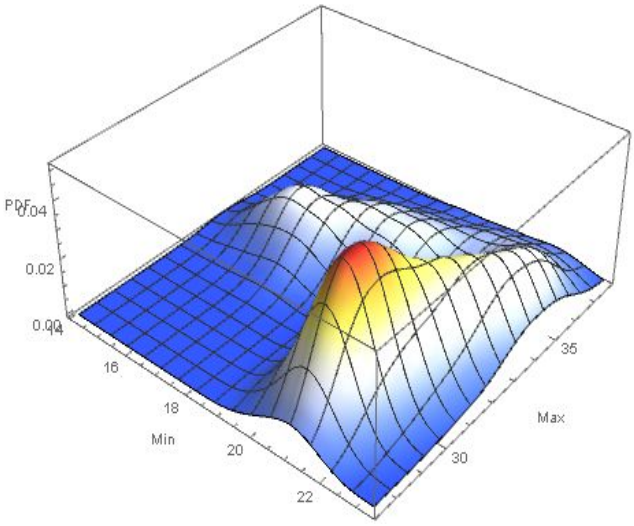
Year	Month	Rainfall
1983	Sep	65.1
1985	Jan	33.3
1986	Nov	53.2
1987	Jan	40.4
1988	Jan	34.1
1988	Dec	19.4
1990	Dec	66.3
1992	Nov	53
1993	Jul	358.9
1994	May	264
1995	Apr	217.9
1995	Dec	31.6
1998	Jan	13.3
2004	Jan	17.5
2004	Nov	51.3
2006	Jan	49
2007	Dec	49.1
2012	Dec	21
2014	Jan	14.7

Analysis

Summary Plots

Using the code provided by Dr. Long for Mathematica, a contour plot of the minimum and maximum temperatures as a function of rainfall was created, which was subsequently used to create a PDF of the data.





Rainfall seems negatively correlated with the Maximum Temperature, meaning there was more rainfall in the months with lower Maximum Temperatures. There did not appear to be a clear relationship with Rainfall and Minimum Temperature.

Rainfall Model with Oscillations

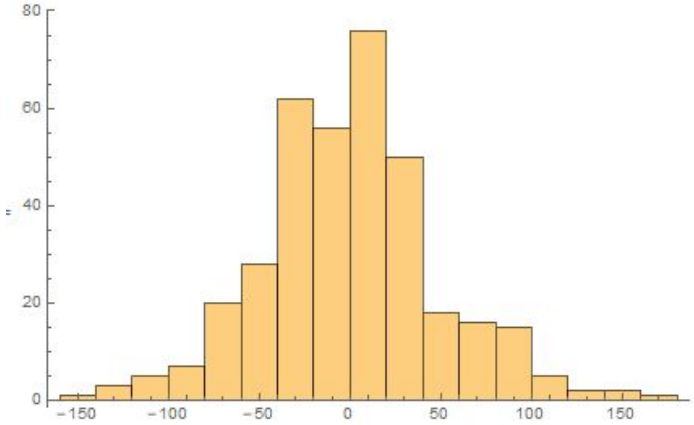
$$Rainfall(\boldsymbol{t}) = 96.345+0.000631631\boldsymbol{t} - 35.8606\ Sin(0.0172024\boldsymbol{t}) - 111.749\ Cos(0.0172024\boldsymbol{t})$$

$$+ 31.5053\ Sin(0.0344048\boldsymbol{t}) - 15.4743\ Cos(0.0344048\boldsymbol{t})$$

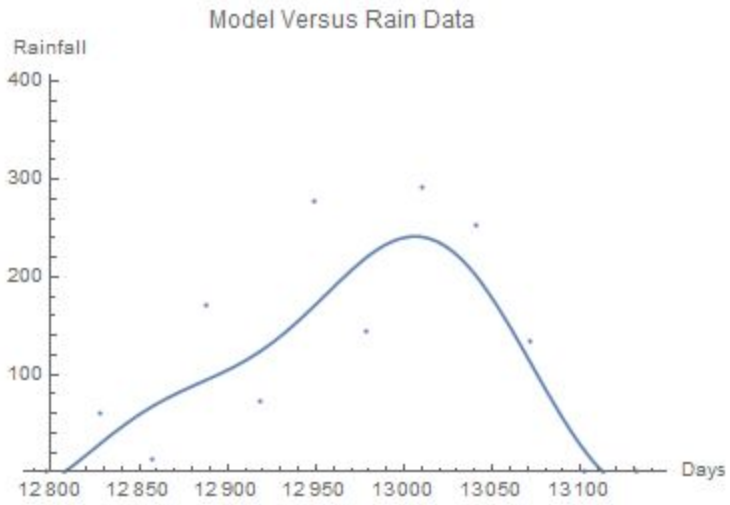
In Mathematica, using Julian dates, a linear model of rainfall over time was constructed, this model uses an annual and a semi-annual period to account for the periodicity of the data. Additional terms that were considered but not ultimately significant at the 0.05 level include Sine and Cosine terms with a quarterly period, linear time, and quadratic time. **The R^2 is 75.1%**.

	Estimate	Standard Error	t- Statistic	P- Value
1	105.088	2.57539	40.8046	1.77759×10^{-137}
Sin[0.0172024 x]	−36.0003	3.57512	−10.0697	3.44028×10^{-21}
Cos[0.0172024 x]	−111.744	3.70712	−30.1429	9.89655×10^{-101}
Sin[0.0344048 x]	31.4681	3.5891	8.76768	7.21386×10^{-17}
Cos[0.0344048 x]	−15.5237	3.69368	−4.20278	0.0000332521

There is no obvious pattern in the residual versus fits plot.



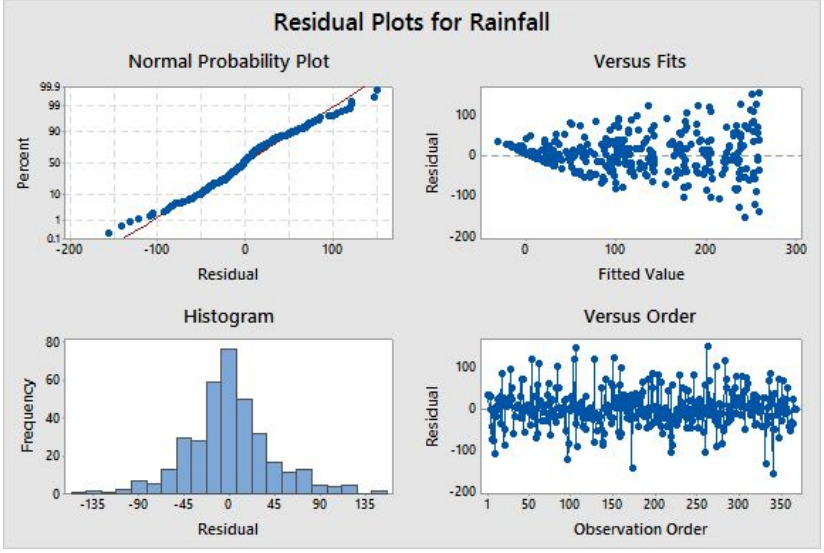
Our histogram shows a mostly normal distribution.



Section of model of year 1987 in Julian days.

Rainfall Model with Categorical Month

Alternatively, Minitab was used to build a model to predict rainfall. Initially the following variables were considered: MaxTemp, MinTemp, Year, Month, Year^2, MaxTemp^2, MinTemp^2, and MinTemp*MaxTemp. All of these terms with the exception of MinTemp^2 were significant at the 0.05 level, although many of the higher ordered terms had incredibly small coefficients and were not contributing very much to the overall model. For the sake of simplicity, the variable TempRange was created which was the range of the MinTemp and MaxTemp. The following model was then created.



The normality assumption seems reasonably met through the QQ-plot and histogram of residuals. There is also no clear pattern in the residuals versus order plot. However, there is a very dramatic megaphone-shape in the residuals versus fits graph. An attempt to rectify this was made by using various transformations on rainfall but no improvement seemed to be made. Therefore no transformation of the data was done in the end.

Analysis of Variance

Source	DF	Seq SS	Contribution	Adj SS	Adj MS	F-Value	P-Value
Regression	15	2782228	78.97%	2782228	185482	87.87	0.000
TempRange	1	2382997	67.64%	6194	6194	2.93	0.088
Year	1	618	0.02%	16063	16063	7.61	0.006
NumMonth	11	371691	10.55%	318548	28959	13.72	0.000
TempRange*TempRange	1	10877	0.31%	11262	11262	5.34	0.021
Year*Year	1	16045	0.46%	16045	16045	7.60	0.006
Error	351	740889	21.03%	740889	2111		
Total	366	3523117	100.00%				

Regression Equation

Rainfall = -307572 + 17.3 TempRange + 308 Year + 0.0 NumMonth_1 + 9.5 NumMonth_2
+ 29.3 NumMonth_3 + 66.7 NumMonth_4 + 91.4 NumMonth_5 + 128.0 NumMonth_6
+ 157.8 NumMonth_7 + 211.9 NumMonth_8 + 188.7 NumMonth_9 + 51.8 NumMonth_10
- 12.5 NumMonth_11 - 1.9 NumMonth_12 - 0.850 TempRange*TempRange
- 0.0769 Year*Year

This model performed well, with a **R^2 value of 78.97%**. Even after penalizing for the number of terms, the adjusted R^2 value is 78.07%.

Model Summary

S	R-sq	R-sq(adj)	PRESS	R-sq(pred)
45.9434	78.97%	78.07%	810283	77.00%

Conclusions

In the rainfall model with oscillations there was no significant change in rainfall over time found. The alternative model with month as a categorical variable did find a significant change over time, however the issues with the extremely megaphone-shaped residuals throws the entire model into question. For this reason, and the similar R^2 values, the oscillatory rainfall model is the more appropriate choice.

We have the same questions for the Togolese meteorologists as the last group. We identified 39 outliers relative to month. Were extreme weather conditions occurring during these months? Or is it possible there was error in measurement? It would be interesting to see the outliers identified in another method as well, such as if a standardized residual is more than 4.