

Examining Linear Correlation Between Rising Global Temperatures and Armed Conflict and Violence in Africa from 2000-2015

Zach Fisher

Hypothesis & Examining Conflict Data

The dataset used in this analysis comes from ACLED (Armed Conflict Location & Event Data Project) and the data provided by ACLED only covers Africa and some parts of Asia. This limitation is okay for what this analysis hopes to accomplish. Internet users in the comment sections of community sites like Reddit(https://www.reddit.com/r/videos/comments/557v99/earths_c02_levels_just_crossed_the_400_parts_per/d88o9zb/) have proposed that rising global temperatures have caused a lot of the unrest in these regions, namely the “Arab Spring” which took place in several countries in Africa.

I will base my hypothesis off of this speculation and state my hypothesis as such:

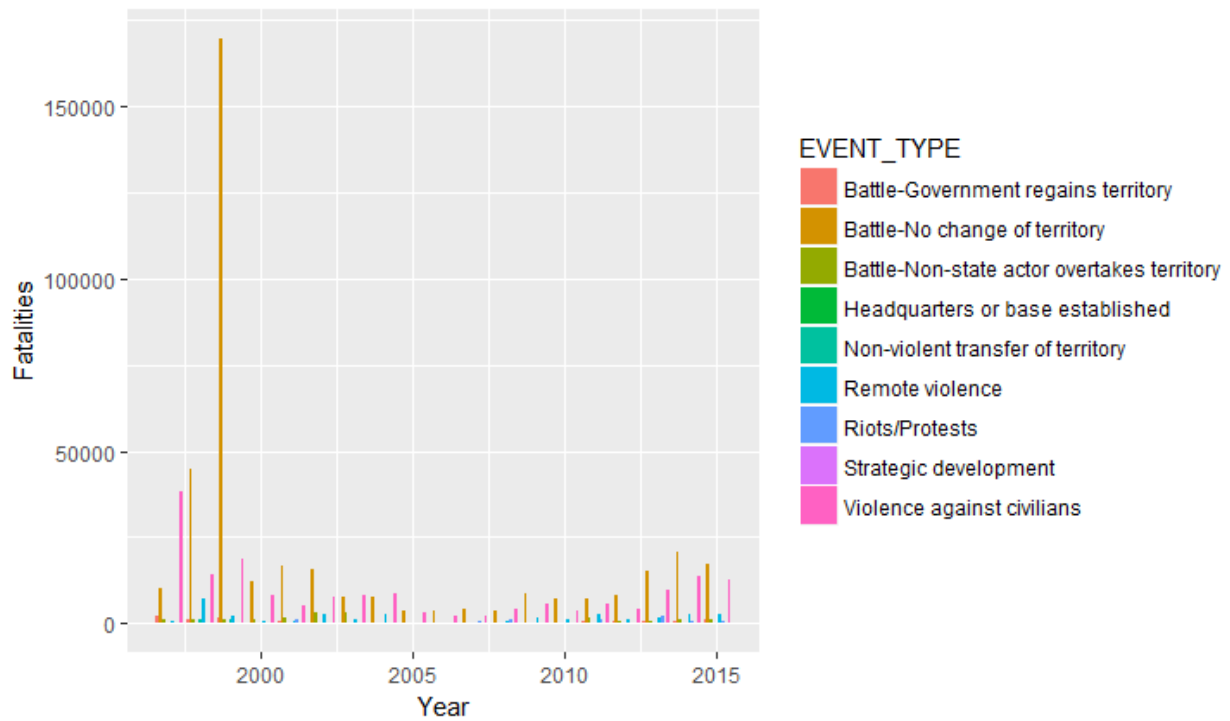
H_A = “The rising amount of violence in Africa is linearly related to rising global temperatures.”

H_0 = “The rising amount of violence in Africa is not linearly related to rising global temperatures.”

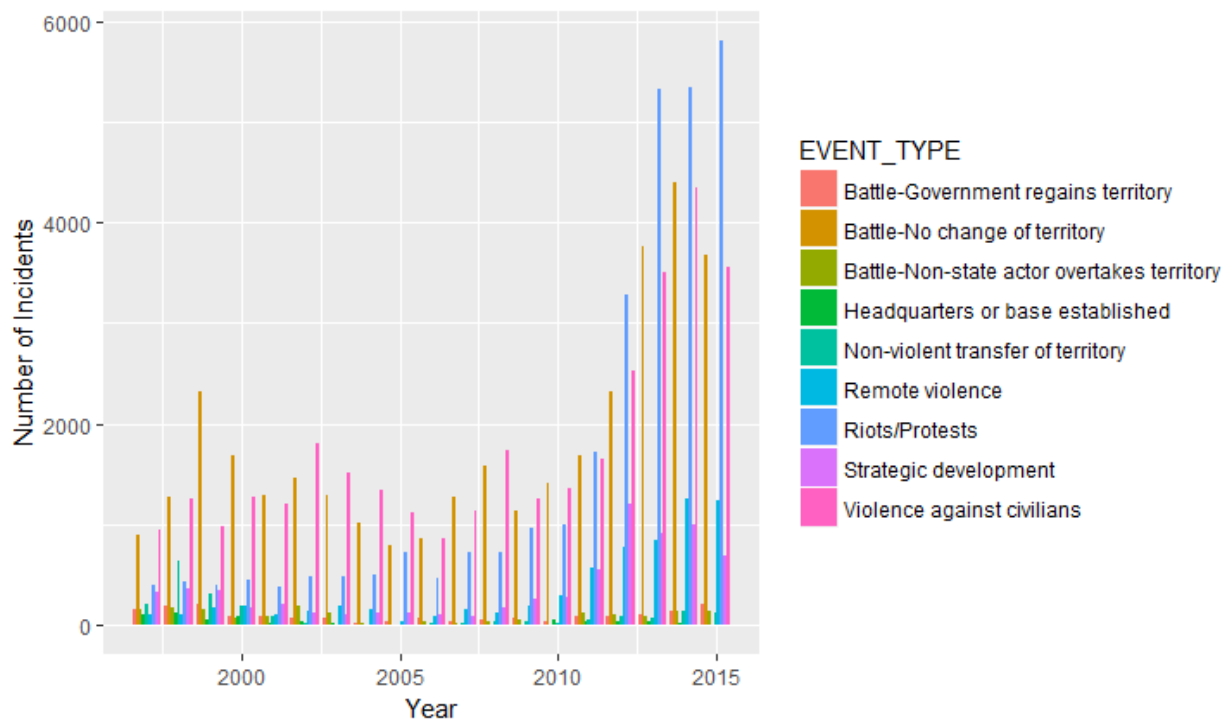
My alpha level will be 0.05.

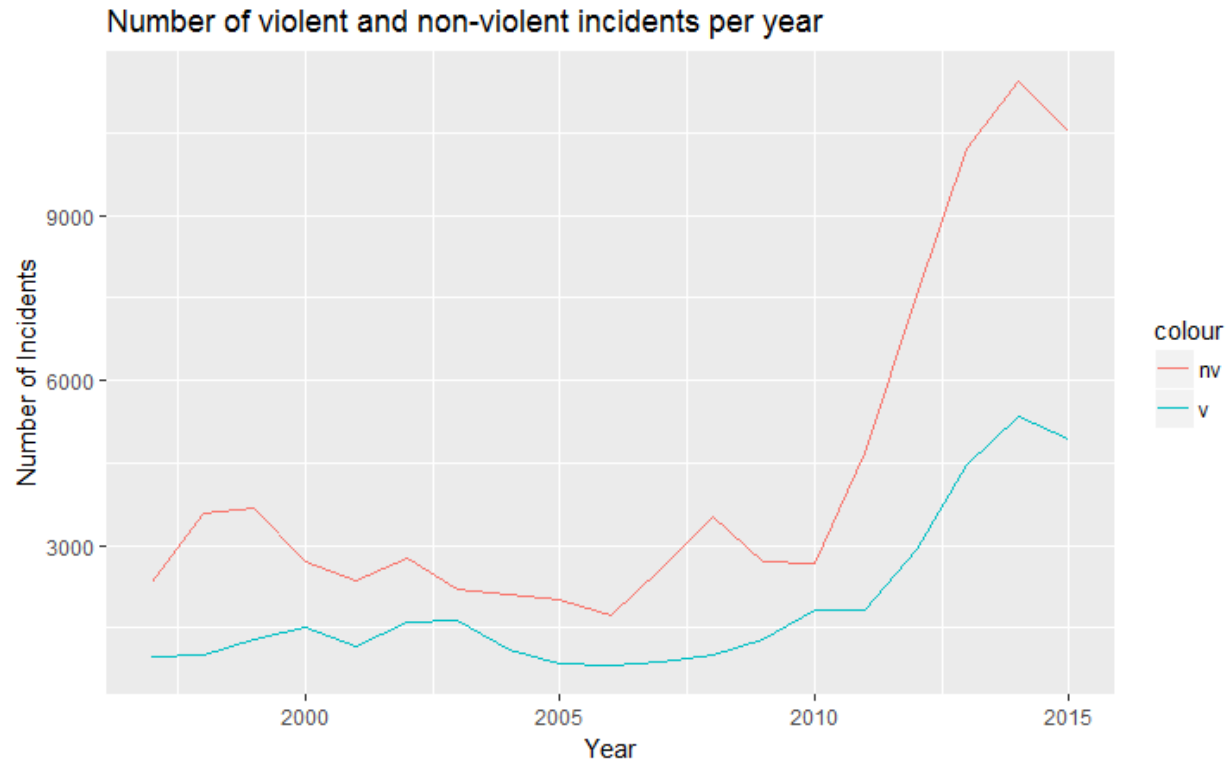
The conflict data is a collection of news reports from the Associated Press involving violence or armed conflict. Each reported incident is an entry in the dataset. Preliminary examination of the conflict data set shows that the number of fatalities from these incidents has sharply declined since the civil wars of the late 1990's. However, the number of incidents reported has significantly increased since that time. The number of violent and non-violent incidents reported is also examined.

Number of fatalities per year by event type



Number of violent incidents per year by event type





Climate Data Exploration

Climate data comes from NASA's GISS database which provides global average temperature data as well as zonal average temperature data. The country specific climate data was too large for the NOAA database to serve over their free portal so the zonal and global climate data will be used for the analysis.

NASA's GISS database provides two separate datasets, one containing land and sea data, and one containing specifically land data. Because the ocean has a significant cooling effect on the surface temperature, the land and sea data set tends to underrepresent the actual trend while the land only data tends to over-represent the actual climate trends. The dataset chosen for this analysis was the land and sea data.

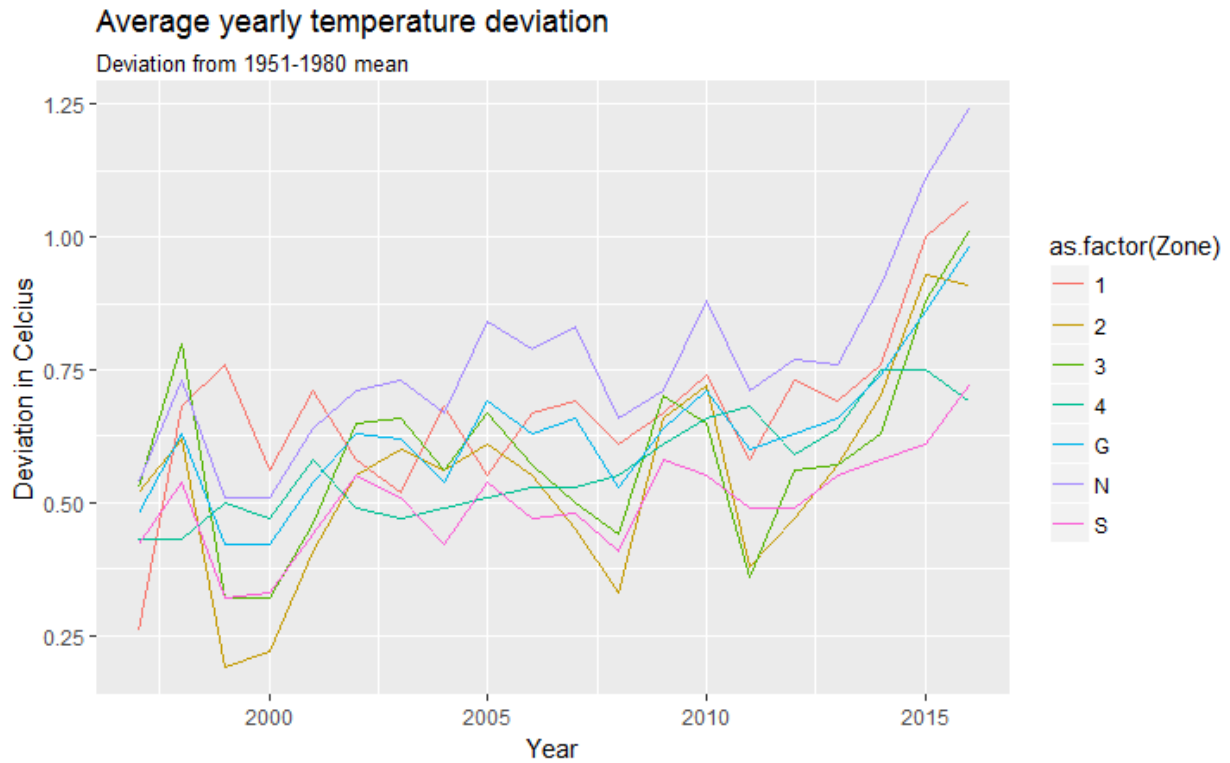
The zones given in the dataset were Global, Northern Hemisphere, Southern Hemisphere, and several latitudinal zones. Using Google Maps, the dataset was pruned to only include the zones which included some portion of the continent of Africa. Manual work was done in Excel to associate countries with their corresponding zones. In the event that a country overlapped two zones, the zone which contained a majority of the country's major cities was chosen. This is a source of bias in this experiment, however, given the aggregated nature of the data, the skew was decided to be insignificant.

Zone 1: 44N – 24N

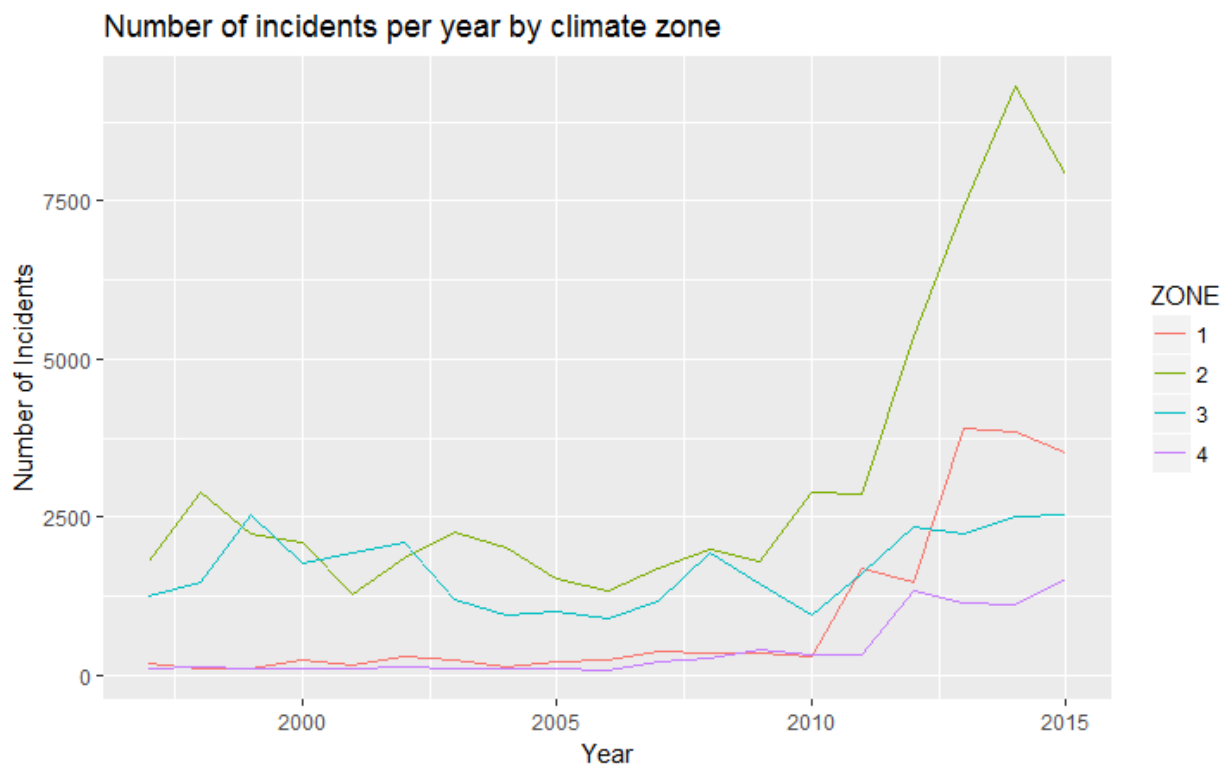
Zone 2: 24N – EQU

Zone 3: EQU – 24S

Zone 4: 24S – 44S



Once countries were segregated into their zones, the violence per zone could be examined. Of note, zone 2 include Sudan, Nigeria, and the Chad Basin where Boko Haram operates.



Linear Modeling

Using linear modeling in R, I first examined the relationship between the number of incidents and the annual temperature data on a zone-by-zone basis. The summaries of the models are as follows:

Zone 1

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-1961	1329	-1.476	0.1583
TEMP	4421	1985	2.227	0.0398 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1206 on 17 degrees of freedom

Multiple R-squared: 0.2258, Adjusted R-squared: 0.1803

F-statistic: 4.959 on 1 and 17 DF, p-value: 0.03976

Zone 2

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-311.6	1621.2	-0.192	0.8498
TEMP	6618.3	2917.0	2.269	0.0366 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2189 on 17 degrees of freedom

Multiple R-squared: 0.2324, Adjusted R-squared: 0.1873

F-statistic: 5.148 on 1 and 17 DF, p-value: 0.03659

Zone 3

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1863.5	552.1	3.375	0.0036 **
TEMP	-328.9	938.3	-0.351	0.7303

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 597.3 on 17 degrees of freedom

Multiple R-squared: 0.007175, Adjusted R-squared: -0.05123

F-statistic: 0.1229 on 1 and 17 DF, p-value: 0.7303

Zone 4

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-1622.2	443.1	-3.661	0.001935 **
TEMP	3615.0	778.5	4.643	0.000233 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 325.3 on 17 degrees of freedom

Multiple R-squared: 0.5591, Adjusted R-squared: 0.5332

F-statistic: 21.56 on 1 and 17 DF, p-value: 0.0002325

The interesting part here is that overall, the two variables seem to be loosely correlated. Two zones have p-values low enough to reject the null despite only explaining 18% of the variance. The model for Zone 3 gives a p-value that doesn't let me reject the null and Zone 4 provides the best model of them all. There are only 3 countries in Zone 4 including South Africa which has been a generally stable country since the Nelson Mandela era. The low number of incidents combined with the mostly stable southern climate helps to explain the quality of the model.

In order to reject the null hypothesis, the relationship between global temperature averages and African violence as a whole will need to be examined. A linear model with global climate data and aggregated violence data for the entire continent was run using R. The summary of that model is as follows:

Global

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-8216	5290	-1.553	0.1388
TEMP	23553	8518	2.765	0.0132 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3892 on 17 degrees of freedom
 Multiple R-squared: 0.3102, Adjusted R-squared: 0.2696
 F-statistic: 7.645 on 1 and 17 DF, p-value: 0.01324

Conclusion

The p-value for the temperature variable is moderately significant and the F-statistic and p-value is lower than 0.05, therefore I can reject H_0 and accept H_A .

Supplemental Investigation

However, this model is only explaining ~27% of the variance and spurious regression could be possible here. Data around food and water availability or farming production in Africa would need to be collected to relate climate to violence. This data was not collected nor included in this experiment. However, data around internet penetration and population growth was collected. It is well known that Twitter was an essential ingredient in the Arab Spring movement and young people were some of the main characters in many of the revolutions.

Temperature + Internet Penetration

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	5.511e+03	3.966e+03	1.390	0.188
TEMP	-6.615e+03	6.988e+03	-0.947	0.361
USERS	8.511e-05	1.186e-05	7.179	7.17e-06 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1923 on 13 degrees of freedom
 Multiple R-squared: 0.8659, Adjusted R-squared: 0.8453
 F-statistic: 41.98 on 2 and 13 DF, p-value: 2.127e-06

Temperature + Internet Penetration + Population Growth

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	3.656e+04	1.052e+04	3.474	0.00460 **
TEMP	-3.396e+03	5.530e+03	-0.614	0.55067
USERS	1.548e-04	2.440e-05	6.345	3.69e-05 ***
POP	-3.783e-05	1.226e-05	-3.086	0.00944 **

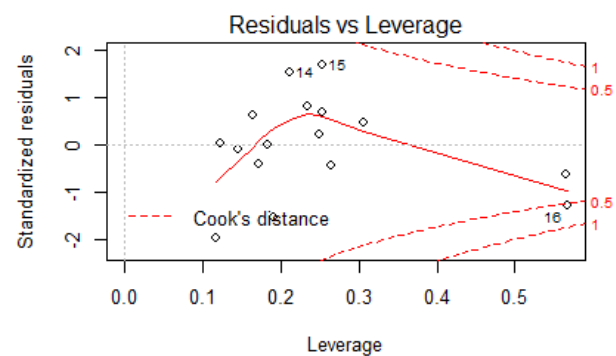
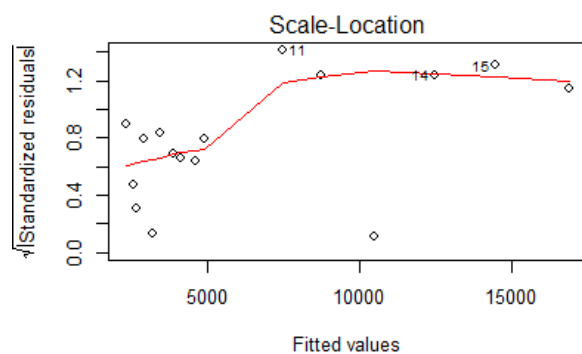
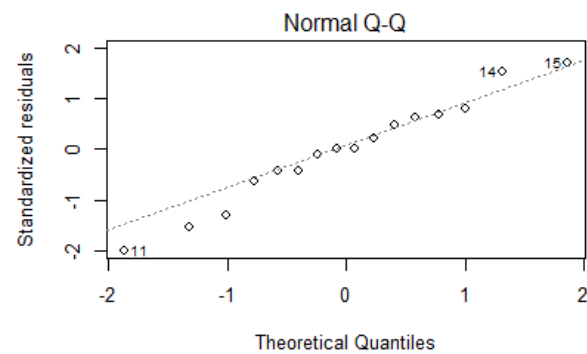
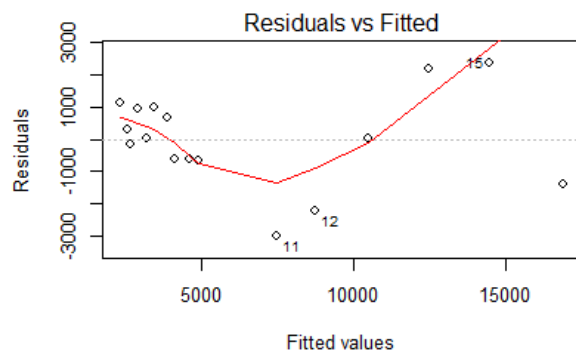
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1495 on 12 degrees of freedom

Multiple R-squared: 0.9252, Adjusted R-squared: 0.9065

F-statistic: 49.5 on 3 and 12 DF, p-value: 4.954e-07

These models are significantly better at explaining the rise in violence and armed conflict than just using temperature alone. In fact, as the additional variables are added to the linear model, the significance of the temperature data dwindles.



References

Internet Data - Accessed: 5/26/2017

ITU ICT: <http://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx>

Temperature Data - Accessed: 5/26/2017

NASA GISS: <https://data.giss.nasa.gov/gistemp/>

Population Data - Accessed: 5/26/2017

UN ESA PD: <https://esa.un.org/unpd/wpp/Download/Standard/Population/>

Violence Data - Accessed: 4/21/2017

ACLED Africa: <http://www.acleddata.com/data/>