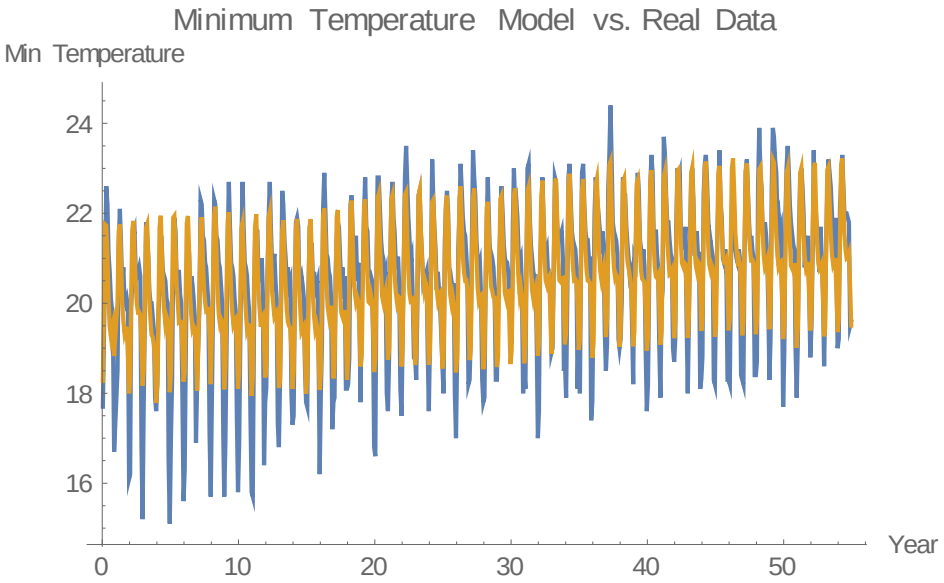
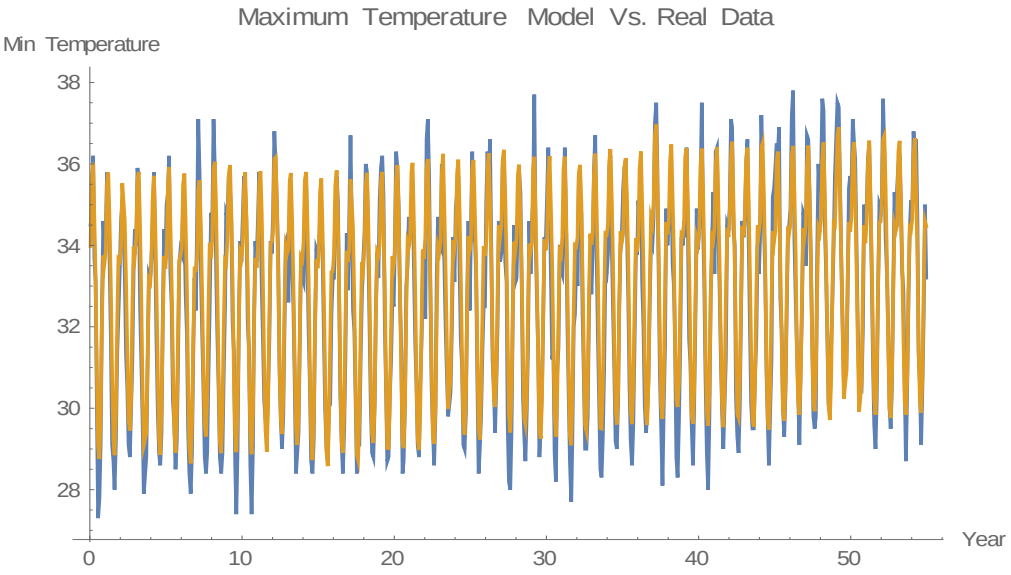
Section 1: Comparing Togo Model to Real Data

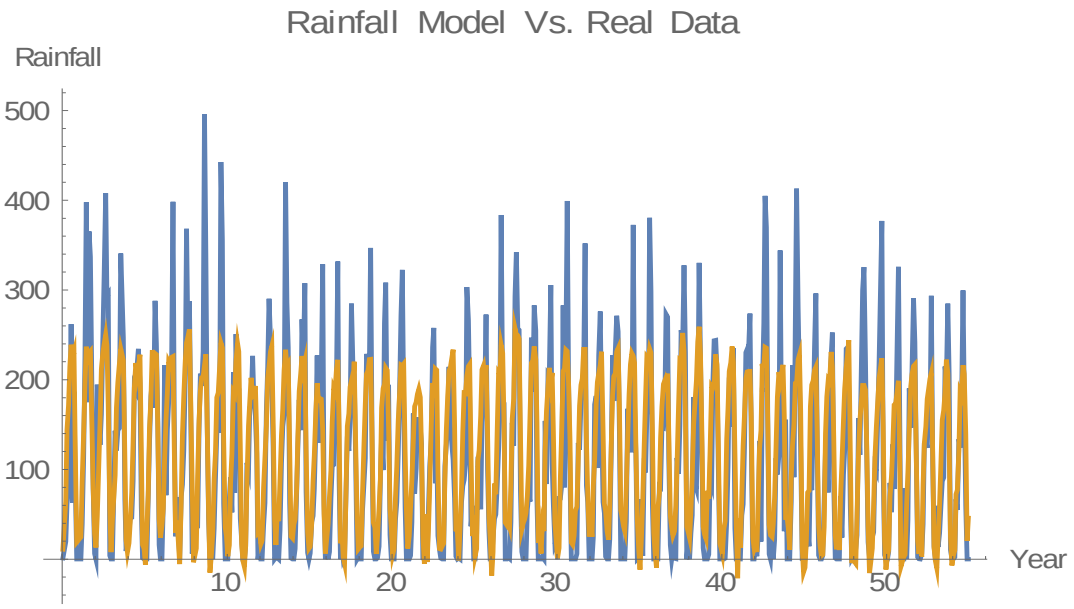
To begin the evaluation of the fit of the global Togo model against Sokode, I imposed the model (using Sokode's data) against the actual data of Togo. I did this for minimum temperature data, maximum temperature data, and rainfall data.



Looking at this graph of the real minimum data graphed over the model, the model does a somewhat decent job of modeling the real data. The real data has several points that are pretty significantly lower than the model goes. As expected, the model is much smoother than the real data. The real data clearly has an increasing trend to it, and the model captures that aspect of it, which is important.



The plot of the maximum real data graphed over the model looks better than the minimum temperature plot. The maximum data is increasing (although not as drastically as the minimum temperature) and the model captures that.

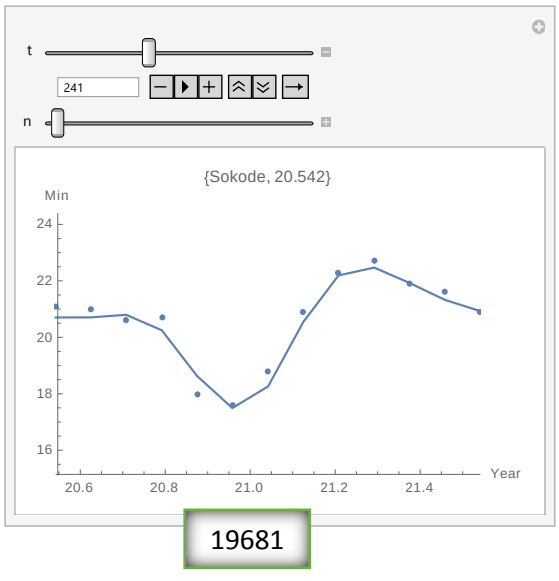


The graph of the real rainfall data graphed over the rainfall model does not look super great. The model does not capture the higher values of rainfall. As discussed in class, the model does predict some negative rainfall, but this is expected since we had so many values of 0 in our rainfall data.

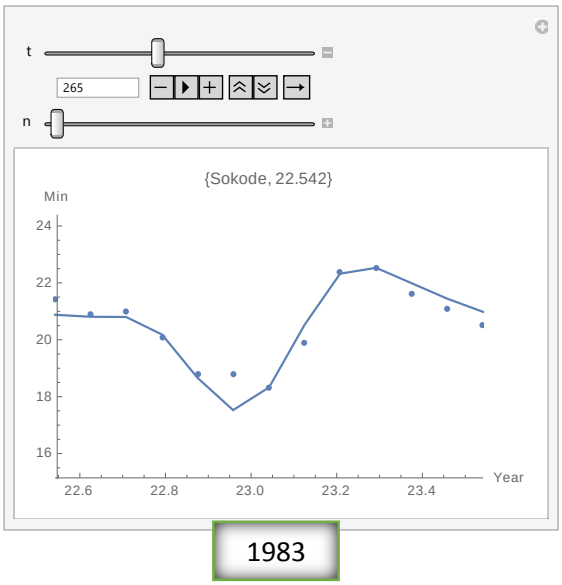
Section 1.1: Animated Visualization of the Model Versus Real Data

The graphs above are helpful to some extent. But since there are so many data points, it can be hard to see what is really going on between the model and the real data. I created animated visualizations in Mathematica using the Manipulate command. The interactive visualizations are in the Mathematica file, if you're interested in experimenting with them.

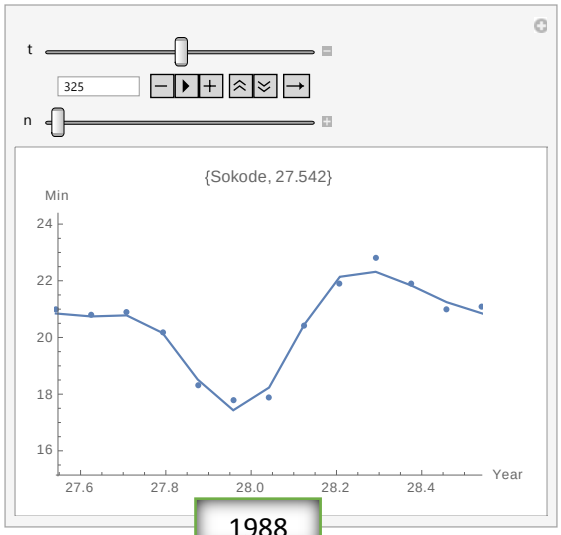
For the **minimum temperature**, I think the model does the best job between the years 1981 and 1989. For the rest of the years, the model does a pretty good job, but I think 1981 – 1989 are the best years of the model. The graphs below shows the monthly minimum values for 1981, 1983, 1988, and 1989 against the minimum model for Sokode:



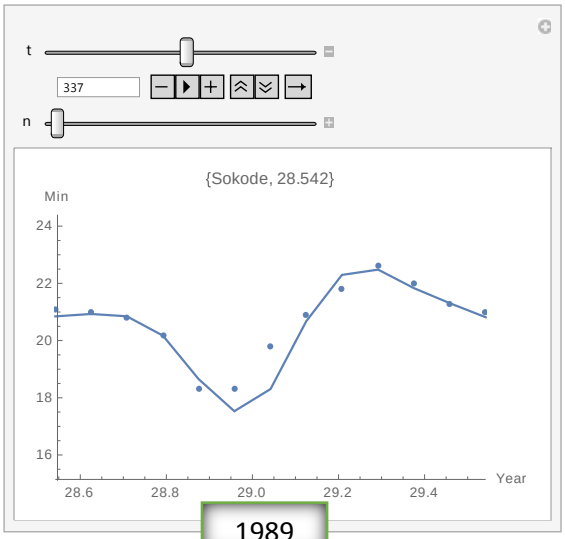
19681



1983

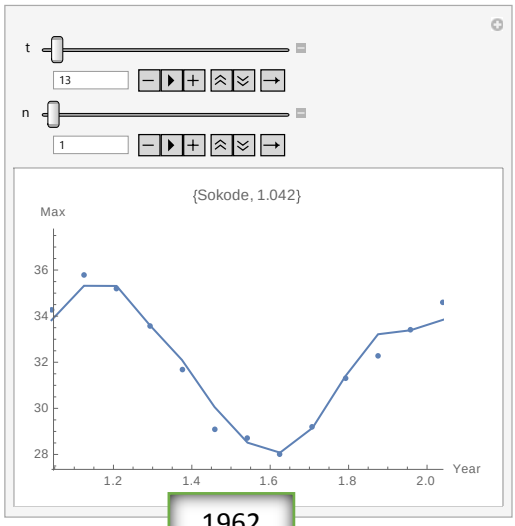


1988

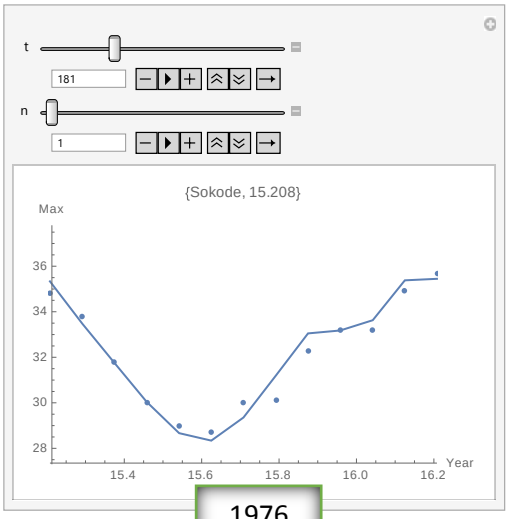


1989

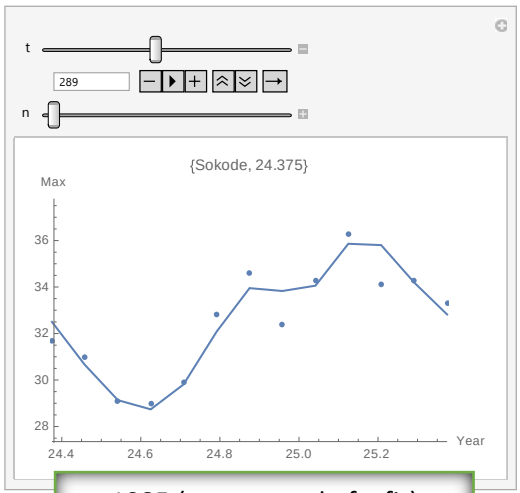
For the **maximum temperature**, the model does very well from 1961 to around 1984. Then, while it is still pretty good, it does a somewhat poorer job between 1985 and 1994. Then, around 1995, it begins to be closer to the real data again. Below is a sample of some of those years:



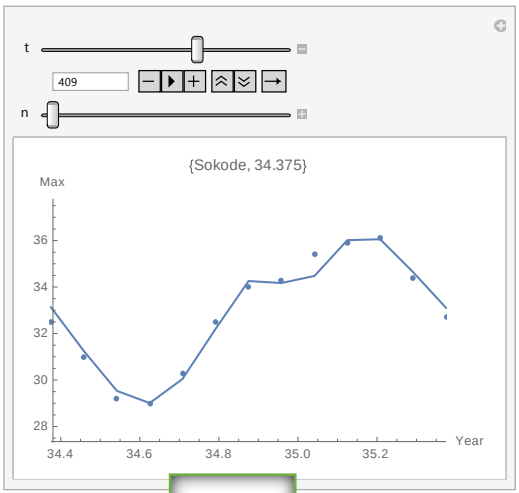
1962



1976

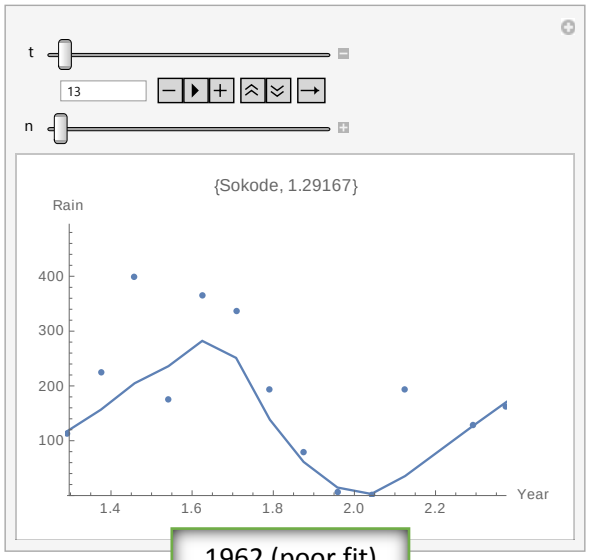


1985 (not as good of a fit)

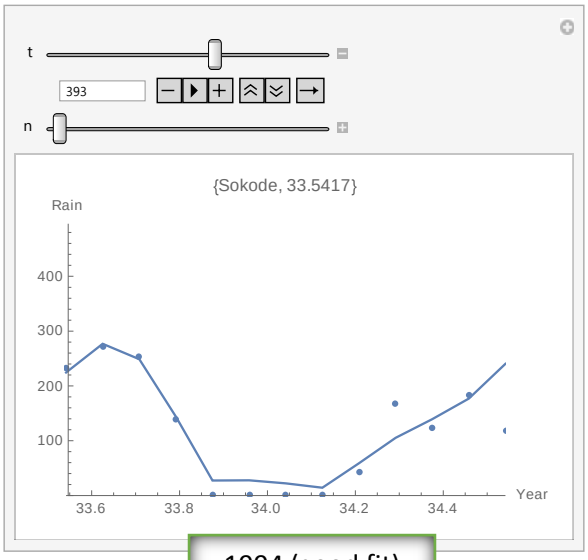


1995

The **rainfall model** has certain years where the model does a really good job with the real data. But there are also some years where it is not very good. For the first couple years (1961 – 1963), the model does a pretty poor job with the data. After that, I think it does a pretty decent job (overall). I did not notice a run of years where the model did particularly bad or particularly well. There are some years where the model is pretty off, but that can be expected since rainfall varies a lot and is very unpredictable. Below is an example of a pretty poor fit and a pretty good fit:



1962 (poor fit)



1994 (good fit)

Section 2.0: Testing Togo Model Terms against Sokode Data

Section 2.1: Minimum Model

I used the terms from the global Togo model with the minimum Sokode Data to create a model to test how well the terms from the Global Model fit to the individual city of Sokode. After creating this, I got the following parameter table:

Parameter Table for Minimum Model

	Estimate	Standard Error	t-Statistic	P-Value
1	2.37191	0.235467	10.0732	3.09195×10^{-22}
x	0.0323706	0.00162945	19.8659	1.26583×10^{-68}
Cos[4 π x]	−1.11068	0.0337407	−32.9181	2.96472×10^{-139}
Sin[4 π x]	0.276965	0.0528661	5.239	2.20154×10^{-7}
Cos[2 π x]	−1.57286	0.0595173	−26.4269	2.77371×10^{-104}
Sin[2 π x]	0.807843	0.102965	7.84581	1.83774×10^{-14}
Cos[6 π x]	−0.21793	0.0320155	−6.80702	2.31664×10^{-11}
Sin[6 π x]	0.0781154	0.0320233	2.43933	0.0149882
$\cos\left[\frac{2\pi x}{13}\right]$	0.0499095	0.0341586	1.46111	0.144481
$\sin\left[\frac{2\pi x}{13}\right]$	−0.0589158	0.0330122	−1.78467	0.0747944
$\cos\left[\frac{\pi x}{10}\right]$	0.0957291	0.0336096	2.84826	0.00453859
$\sin\left[\frac{\pi x}{10}\right]$	−0.0563365	0.0342438	−1.64516	0.100433
la long	0.228943	0.0227279	10.0732	3.09195×10^{-22}
la^2	0.0293143	0.00291012	10.0732	3.09195×10^{-22}
$long^2$	1.78803	0.177503	10.0732	3.09195×10^{-22}
la	0.263687	0.0261771	10.0732	3.09195×10^{-22}
long	2.05938	0.204441	10.0732	3.09195×10^{-22}
elev	0.00612897	0.000608442	10.0732	3.09195×10^{-22}
sstt	0.112509	0.0604637	1.86077	0.0632405
ensoo	−0.00779711	0.00222065	−3.51118	0.000477828

The terms that have a period of 1 year, 1/2 year, 1/3 year, and 20 years (for the 20 years period, only the cosine term is significant, but we would still keep both terms since the sine term is significant) are all significant. The linear (decimal year), latitude, longitude, elevation, and ENSO term are also significant. However, the SST and the sine and cosine terms that have a period of 13 years are not significant (their p-value is greater than .05). When we searched for significant periods for Sokode in the first evaluation, we did not find any periods greater than 7 years to be significant for minimum temperature so it is not too surprising that a period of 13 years was not significant for Sokode.

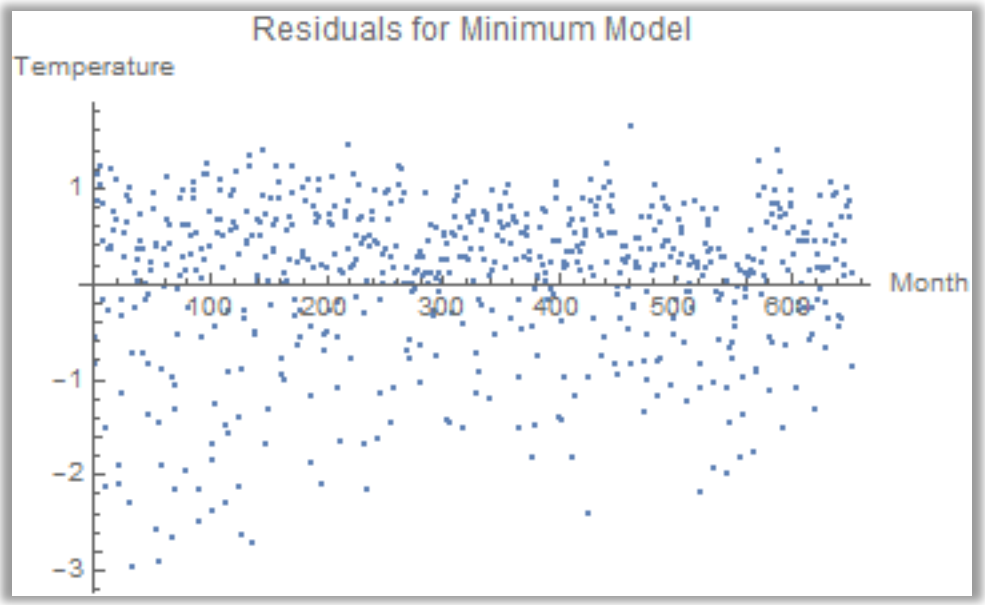
The Parameter Confidence Intervals for the Minimum Model are below:

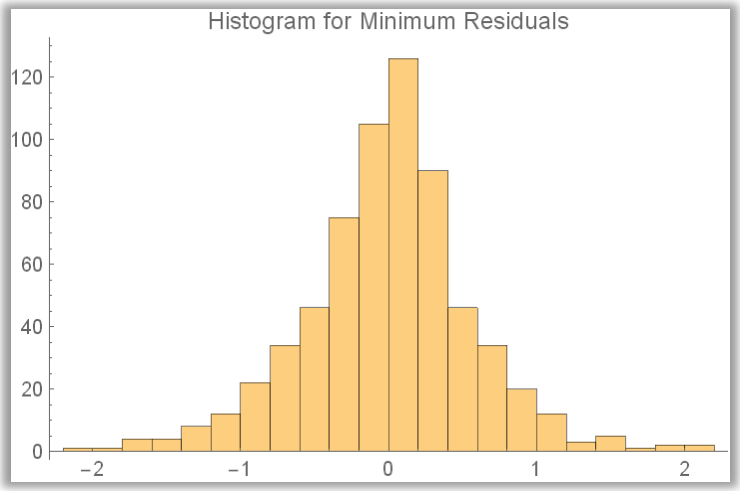
Confidence Intervals for Minimum Model

```
{{1.90952, 2.8343}, {0.0291708, 0.0355703}, {-1.17694, -1.04442},
 {0.173151, 0.38078}, {-1.68973, -1.45598}, {0.605649, 1.01004},
 {-0.2808, -0.155061}, {0.0152306, 0.141}, {-0.0171683, 0.116987},
 {-0.123742, 0.00591095}, {0.0297292, 0.161729}, {-0.123582, 0.0109087},
 {0.184312, 0.273574}, {0.0235997, 0.035029}, {1.43946, 2.1366},
 {0.212283, 0.315092}, {1.65792, 2.46085}, {0.00493416, 0.00732378},
 {-0.00622483, 0.231242}, {-0.0121578, -0.00343638}}
```

The coefficients given by the Togo model were within the confidence intervals for the Cos[2π/(1/3)], Sin[2π/13], Cos[2π/20], and ENSO terms. There were some other terms that were very close though. For example, the coefficient for the Decimal Year term is .02739 and the confidence interval for that term is {.0291708, .0355703}. The SST term was also pretty close. The SST coefficient is .29451 and the confidence interval for SST is {-.00622483, .231242}. I would like to see more coefficients within the confidence intervals.

The residuals for the minimum model look pretty good. You can detect some patterns in the residuals—you can see some sinusoidal behavior at the top of the graph and there are some upside down smiley faces at the bottom of the graph. The residuals are not completely independently and identically distributed. Also, we can see from the histogram that the residuals are not normal, which is concerning. Overall, I am not too worried about these residuals, but there are some aspects that are cause for concern.





The R-Squared value (using the Mathematica function) for this model is around .887.

The "Our-Squared" value (using the formula given in class) for this model is around .761. This "Our-Squared" value is pretty decent. Of course, we would like it to be higher, but I think that the model does a decent job with the minimum data (overall).

Note: The formula for "Our-Squared" is:

$$1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

Section 2.2: Maximum Model

Using the terms from the global Togo model with the maximum Sokode data, I created a model to test how well the terms from the global model fit to the individual city of Sokode. After creating this model, I got the following parameter table:

Parameter Table for Maximum Model

	Estimate	Standard Error	t-Statistic	P-Value
1	15.7475	3.98146	3.95522	0.0000851306
x	0.0197634	0.00177311	11.1462	1.83654×10^{-26}
Sin[2 π x]	1.31047	0.1183	11.0776	3.49418×10^{-26}
Cos[2 π x]	2.15381	0.0668409	32.223	2.00588×10^{-135}
Sin[4 π x]	−0.159034	0.0596604	−2.66565	0.00788071
Cos[4 π x]	−0.419897	0.0450459	−9.32153	1.89792×10^{-19}
Sin[6 π x]	−0.0232688	0.0397627	−0.585192	0.558628
Cos[6 π x]	−0.533602	0.041087	−12.9871	2.39272×10^{-34}
$\sin\left[\frac{2\pi x}{13}\right]$	−0.103937	0.0368857	−2.81781	0.00498672
$\cos\left[\frac{2\pi x}{13}\right]$	0.0565283	0.0374045	1.51127	0.13122
Sin[8 π x]	−0.114289	0.0357988	−3.19253	0.00148042
Cos[8 π x]	−0.116641	0.037394	−3.11925	0.00189573
la lon	0.00120155	0.00163431	0.7352	0.462491
lon^2	0.000407067	0.000166168	2.44973	0.0145669
la^2	0.0596956	0.0217374	2.74622	0.00620056
la	−2.77263	1.17661	−2.35646	0.018755
lon	−0.0362486	0.0453705	−0.798946	0.424622
elv	13.6726	3.45685	3.95522	0.0000851306
sstt	1.75067	0.442622	3.95522	0.0000851306
ensoo	0.0406913	0.010288	3.95522	0.0000851306

The Sin and Cosine terms that have a period of 1 year, 1/2 year, 1/3 year (for the period of 1/3 year, the sine term is not significant but we would keep both terms since the cosine term is significant), 1/4 year, 13 year (except for the Cosine term, but since the Sine term is significant, we would keep both terms) are all significant. The linear term, elevation, SST, ENSO, latitude, longitude^2, and latitude^2 term are also significant (their p-values are less than .05). The longitude and latitude*longitude terms are not significant (their p-values are greater than .05) which is concerning.

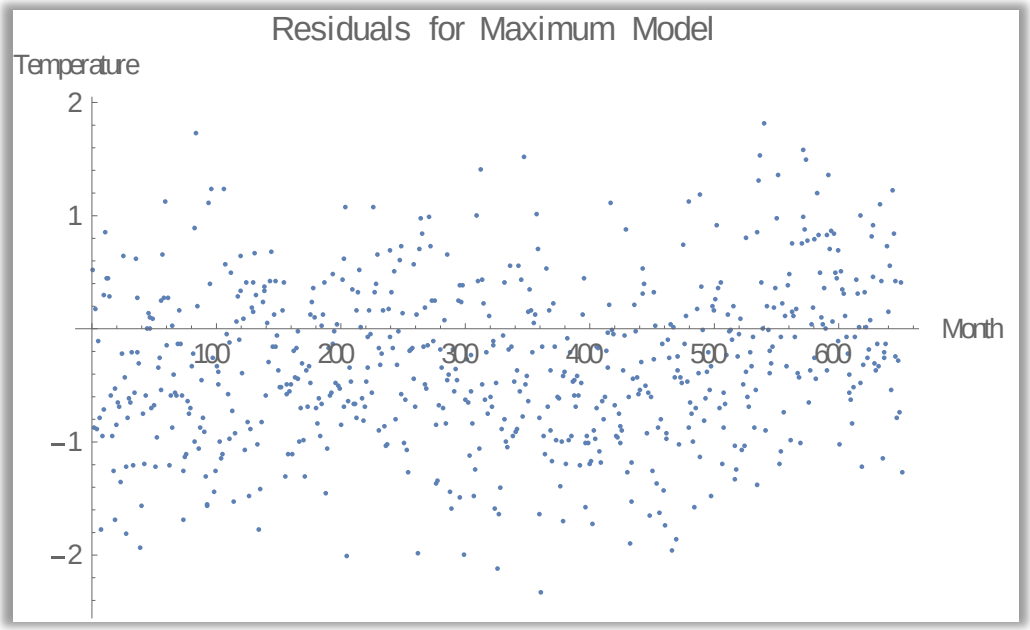
The parameter confidence intervals for the maximum model for Sokode are:

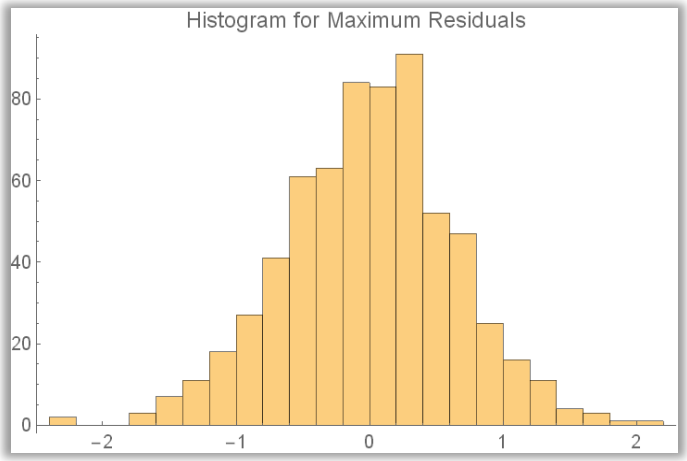
Confidence Intervals for Maximum Model

```
{{7.92904, 23.5661}, {0.0162815, 0.0232453}, {1.07816, 1.54278},
 {2.02256, 2.28507}, {-0.276191, -0.0418767}, {-0.508355, -0.331439},
 {-0.101352, 0.0548144}, {-0.614285, -0.452918}, {-0.17637, -0.0315031},
 {-0.016924, 0.129981}, {-0.184588, -0.0439896}, {-0.190073, -0.0432096},
 {-0.0020078, 0.00441089}, {0.0000807578, 0.000733375},
 {0.0170093, 0.102382}, {-5.08317, -0.462083}, {-0.125344, 0.0528469},
 {6.88428, 20.4609}, {0.881477, 2.61986}, {0.0204885, 0.0608942}}
```

The coefficients given by the Togo model were within the confidence intervals for the Sin[2 π x], Sin[2 π x / (1/2)], Sin [2π x / (1/3)], Sin[2 π x / (1/4)], Cos [2 π x/ (1/4)], Sin [2 π x / 13], and Cos [2 π x / 13] terms. It was very close for the linear term—the coefficient was .01284 and the confidence interval was {.0162815, .0232453}. I am happy to see that so many of the coefficients from the Togo model were within the confidence intervals.

The residuals for the maximum model have some smiley faces and some upside down smiley faces. Other than those patterns, the residuals look pretty randomly distributed (although not perfectly so). The histogram of the residuals are clearly not normally distributed, which is concerning.





The R-Squared value (calculated by Mathematica) is around .936 which is a high R-Squared value.

The "Our-Squared" value (calculated by the formula given in class) is around .906 which is a high value. I am happy with this value.

Section 2.3: Rainfall Model

Using the terms from the global Togo rainfall model with the rainfall Sokode data, I created a model to test how well the terms from the global model fit to the individual city of Sokode. After creating this model, I got the following parameter table:

Parameter Table for Rainfall Model

	Estimate	Standard Error	t-Statistic	P-Value
1	24.7833	23.0407	1.07563	0.282506
x	−0.2281	0.186893	−1.22049	0.222741
Sin[2 π x]	−41.9777	10.1019	−4.15545	0.0000369898
Cos[2 π x]	−66.3522	10.0102	−6.62848	7.33569×10^{-11}
Sin[4 π x]	38.3193	4.62233	8.29004	6.93887×10^{-16}
Cos[4 π x]	−18.9487	5.02234	−3.77288	0.000176719
Sin[6 π x]	−0.825161	2.82303	−0.292296	0.770158
Cos[6 π x]	8.00776	3.24618	2.46683	0.0138986
Sin[8 π x]	−4.58269	2.82996	−1.61935	0.105876
Cos[8 π x]	2.78433	2.86325	0.972436	0.33121
lat	2.75518	2.56145	1.07563	0.282506
lat long	2.39215	2.22394	1.07563	0.282506
lat^2	0.306295	0.284758	1.07563	0.282506
long	21.5178	20.0047	1.07563	0.282506
$long^2$	18.6825	17.3688	1.07563	0.282506
elev	0.0640395	0.0595366	1.07563	0.282506
mint	4.72319	3.45789	1.36592	0.172456
maxt	−22.9148	3.03952	−7.53896	1.6734×10^{-13}
ensoo	0.269616	0.188301	1.43183	0.152691
sstt	21.5714	5.24384	4.11365	0.0000441567

The sine and cosine terms of 1 year, 1/2 year, and 1/3 year (except the Sine term but the Cosine term is significant so we would keep both terms) are all significant. The SST term is also significant. The constant term is not significant. The linear, latitude, latitude*longitude, latitude^2, longitude, longitude^2, longitude, longitude, longitude^2, elevation, minimum temperature, and ENSO terms are all not significant either. It is concerning that so many terms are not significant for the rainfall model.

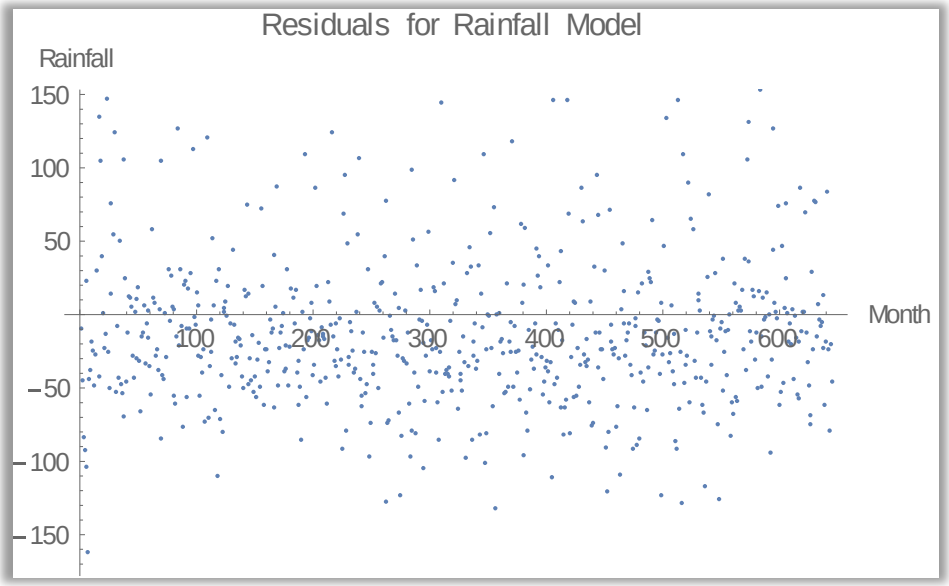
The parameter confidence intervals for the rainfall model are below:

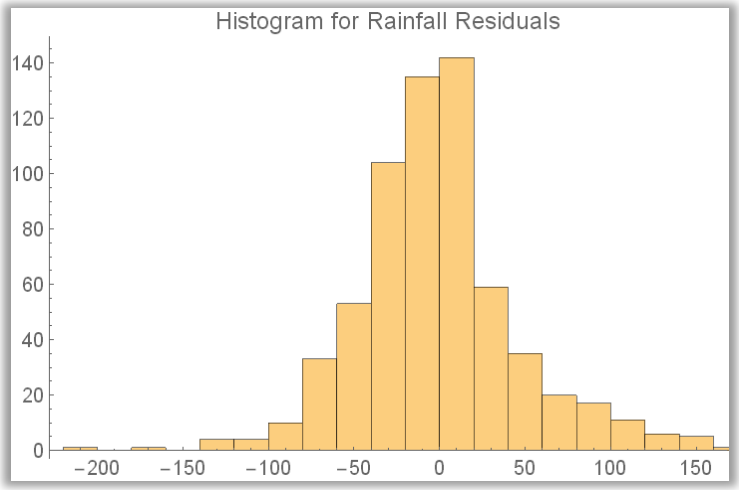
Confidence Intervals for Rainfall Model

```
{{-20.4632, 70.0298}, {-0.595115, 0.138914},
 {-61.8154, -22.14}, {-86.0099, -46.6946}, {29.2421, 47.3965},
 {-28.8114, -9.08599}, {-6.36894, 4.71862}, {1.63303, 14.3825},
 {-10.1401, 0.974681}, {-2.83843, 8.40709}, {-2.27491, 7.78526},
 {-1.97516, 6.75945}, {-0.252903, 0.865494}, {-17.7669, 60.8024},
 {-15.4259, 52.7909}, {-0.0528764, 0.180955}, {-2.0673, 11.5137},
 {-28.8837, -16.9459}, {-0.100164, 0.639396}, {11.2737, 31.8691}}
```

The coefficients given by the global Togo model are within the confidence intervals for the linear (decimal years), SST, ENSO, maximum temperature, minimum temperature, Sin[2 π x], Cos [2 π x], Cos [2 π / (1/2)], Sin [2 π x / (1/4)], and the Cos [2 π x / (1/4)] terms. It is good that there are several coefficients within the confidence intervals, although I am concerned that none of the latitude or longitude coefficients were within the confidence intervals.

The residuals for the rainfall model look somewhat randomly distributed. As with temperature residuals, I can detect some patterns, such as sinusoidal behavior and "smiley faces". The histogram is not normal. The values on the histogram that are around -200 and -150 are very concerning.





The R-Squared value (as calculated by Mathematica) is around .782.

The "Our-Squared" value (calculated from the formula given in class) is around .722. While the "Our-Squared" is not as high as it was for temperature data, it is still pretty good. Of course, we would like a model that yields a higher value but I am not terribly upset by the "Our-Squared" that the model produced.

Section 3.0: Summary

Sokode, the second largest city in Togo, is located approximately halfway between the Gulf of Guinea and the northernmost part of Togo and has an elevation of around 387 meters. The minimum model performs well for Sokode's data and we detected no major fit problems. Sokode's minimum temperature clearly has an upward trend which is accurately captured by the model—the model predicts a .032 degree increase in temperature for every year. The maximum model does a very good job of fitting Sokode's data, as well as the data for the other cities in Togo. The maximum temperature in Sokode has an upward trend to it. The model predicts a .0197 degree increase in temperature for every year. The rainfall model does not perform as well for Sokode as the temperature models did because of the variation in the rainfall data—there are numerous values of 0 mm as well as many values that are over 200 mm. Because of the wide variation and unpredictability in the rainfall data, the rainfall model suffered some fit problems for Sokode's data. Even though the rainfall model did not fit the data as well as the temperature data did, we believe that the model still offers valuable insight. The model predicts a decrease in rainfall of around .2281 mm for every year.