Running Head: Drivers of annual global temperature increase

Title: Urban Temperature Trends: Exploring the Annual Temperature Increase in Major Cities Around the World Over the Past 100 Years

Ines Yang[1]

[1]Department of Environmental Science, Policy, and Management, University of California, Berkeley, 219 Wellman Hall, Berkeley, CA, USA, 94720

Abstract

Global climate change has been a widely contested topic, since it holds the potential to severely impact our environmental, social, and economic spheres. Since major changes to the Earth's surface happened around the late 1900s during the Industrial Revolution, I have conducted a time series analysis on annual urban temperature data from the past 100 years from the 1900s to 2013. I looked at 15 different cities and tested three separate hypotheses in increasing order of complexity and found that the first and simplest model had the best fit to analyze the annual temperature over a hundred years. Using the first model, the MARSS results concluded that cities are unevenly experiencing temperature increase. Every city experienced a significant positive increase in slope. The most prominent results came from Cali, Columbia and Toronto, Canada. Cali, an equatorial city with a latitude of 3.45N, showed a very small increase in temperature, whereas Toronto, a temperate city with a latitude of 43.65N had the largest range of temperatures and showed significant increase. While not every city exhibited a very drastic change in temperature, all showed a positive increase. Further research can be done to increase the model's complexity.

Key words: global warming, annual average temperature, time series, climate change, equator, urban cities

Introduction

Large populated cities have been known to be major contributors of greenhouse gasses and have significant carbon footprints. Because cities are known to be densely populated and are hubs for human activity, their effects on the climate are magnified in these areas. According to the Environmental Protection Agency (EPA), this phenomenon is known as the "urban heat island effect," where cities experience much higher temperatures than their outlying areas. The implications of the urban heat island effect around the world pose many hazards. With the warming of cities come problems such as deaths due to heat waves, power outages, and infrastructure failure such as melted asphalt, railroads, and airplane runways (Leighton, 2019). Climate change and its implications on urban cities pose a severe issue that requires awareness and action. The purpose of this analysis is to see how rapidly cities are warming up, and which cities are experiencing the most changes.

I plan on analyzing the annual mean temperature across different major cities, and see if there are general global trends that are present throughout all cities. I think it would be particularly interesting to know if there are certain cities that have experienced a more drastic change in temperature over the years.

The model I am using in this paper is the Multivariate Autoregressive State-Space (MARSS) model. The MARSS model is well-suited for handling multivariate data, in this case multiple cities, and analyzing correlation in each city over time. The main parameters that I worked with in this project were the Z, B, U, R, Q parameters. The Z matrix describes how the input data is grouped. In this case, I worked with the Z matrices to see how I could group the cities based on certain characteristics they had, such as climatic zones. The B matrix describes variable interaction. The U variable represents drift and random walk. The Q matrix describes the nature of the process error across all species. It can be set to diagonal and unequal, which implies that process error varies across variables, whereas diagonal and equal assumes the same process error in all variables. Lastly, the R matrix represents the observation error across all variables, and has similar characteristics to the Q matrix.

I tested 3 different models to see if climatic zones would improve the model's fit. I hypothesized that cities that were further from the equator would be more resilient and have little to no significant change since they are already at an angle where the sun does not intensely hit the Earth's surface. Overall, I figured that cities that are close in distance would experience warming at similar rates, compared to cities at a higher or lower latitude. Unfortunately, even after 2000 iterations of running this model, the complex model did not merge. However, the simpler model converged within 343 iterations.

Methods

Data Collection and Libraries Used

I obtained my data through a Kaggle search for cleaned data on annual temperature data. The data I am using has been repackaged by the Lawrence Berkeley National Library, called the Climate Change: Earth Surface Temperature Data. The global temperatures data includes monthly temperature data on major cities from 1849 to 2013. While they also contain data such as "GlobalLandTemperatureByCountry.csv" ,"GlobalLandTemperatureByCity" and "GlobalLandTemperatureByState", the dataset that I used was "GlobalLandTemperaturesByMajorCity.csv" . The libraries I used in this paper were tidyverse and dplyr (for data manipulations and data visualizations), patchwork (plot composition), and MARSS (time series analysis).

Data Cleaning and Variable Selection

Since the original dataset contained monthly annual data from 1849 to 2013, I grouped the data by year to account for temporal fluctuations due to seasonality. Additionally, I filtered out the table by only looking at data from 1900 to 2013 because most of the relevant timeline for climate change occurred starting from the late 1950s.

I choose the 15 cities by looking at the data completeness of each city, meaning no missing data. Lastly, I choose variables based on their latitude and selected cities that were spread apart to get a more accurate representation of major cities on a global scale. After selecting the cities, I used the "*pivot_wider"* function with *names_from* Year, and *values_from* the mean annual temperature. The resulting y-matrix was a 15x113 matrix with 15 rows for each city, and 113 rows for each year, starting from 1900 to 2013.

Model Creation

<u>First model</u>: All 15 cities were treated as individually independent. The parameters of the model: Z: 'identity', B: 'identity', U: 'zero', Q: 'diagonal and equal', R: 'diagonal and equal', A: 'zero'

<u>Second Model</u>: Cities were divided up into "Tropical and Non-Tropical". The threshold cutoff for a "Tropical City" was that its latitude had to fall between the range of (-10N, 10N). The cities that fell in this range were Cali, Nairobi, Bogota, Singapore, Fortaleza. The "Non-tropical" cities were Los Angeles, Ankara, Nanjing, Madrid, and Toronto. Originally, I only had 10 cities. The only difference in parameters is the Z matrix, which was 15x2, where each city was given a 0 or 1 based on their tropical characteristics. For example, Nairobi would have 1 for "Tropical" and 0 for "Non-Tropical".

Third Model: Cities were categorized into four climatic regions: Equatorial (0N), Tropical (-45N, 0), Temperate (0 - 45N), Subpolar (45N-90N).The Equatorial cities were Cali, Nairobi, Bogota, Singapore, and Fortaleza. The "Tropical" cities were Rio De Janeiro, São Paulo, and Cape Town. The "Temperate" cities were Los Angeles, Ankara, Nanjing, Toronto, and Madrid. The "Subpolar" cities were Saint Petersburg and Paris. The selection of the cities were also limited by the dataset. Specifically, even though there might have been other more suitable cities that fall into "Subpolar", because there were only a limited number of subpolar cities in the dataset, these two were the most viable cities. The only difference in parameters is the Z matrix, which was 15x4, where each column represented its climatic region.

Running MARSS on the models

After creating the parameters for each model, I created a list of matrices called 'mod' which contained each parameter. The model is essentially a list of parameter matrices that gets passed into the MARSS model to be fitted on. Then I used the *MARSS* function, with the y-matrix passed in as *data*, and mod as *model*, with a *maxit* of 1000 iterations. After running the model on all three models, I compared their AICc (Akaike's Information Corrected Criterion) and whether or not the model converged completely. I chose my model based on which one had the lower AICc (better fit) and complete convergence.

Interpreting Results Based on Confidence Intervals and U Estimate

After choosing my model, I looked at the drift estimates to see their slope (change in annual temperature in Celsius over 113 years) and the 95% confidence intervals ranges. To identify a "significant positive increase,' I checked whether both the lower and upper bounds of the 95 percent confidence interval were above zero. A zero slope indicates no change in temperature since the 1900s.

Data Visualizations

Lastly, I created some visualizations using *ggplot* to better illustrate my data. Some notable visualizations I created were individual line plots of certain cities' changes in temperature over 113 years using *geom_lineplot()*, and a side-by-side boxplot to illustrate the range of temperatures that each city has using *geom_boxplot()*.

Results

Note: The slope estimates show the change in annual temperature over 113 years.

Major City	**Slope (U.X.$Y_{1...n}$)**	**Lower CI**	**Upper CI**
Ankara, Turkey	0.00936	0.00176	0.01697
Bogota, Brazil	0.00997	0.00239	0.01756
Cali, Columbia	0.01007	0.00249	0.01765
Cape Town, South Africa	0.01049	0.00291	0.01808
Fortaleza, Brazil	0.01047	0.00288	0.01806
Los Angeles, USA	0.01135	0.00380	0.01890
Madrid, Spain	0.01204	0.00444	0.01963
Nairobi, Kenya	0.00903	00.0144	0.01662
Nanjing, China	0.01137	0.00378	0.01896
Paris, France	0.01168	0.00409	0.01927
Rio De Janeiro, Brazil	0.01128	0.00369	0.01886
Saint Petersburg, Russia	0.01329	0.00568	0.02089

Singapore, Singapore	0.01028	0.00270	0.01786
São Paulo, Brazil	0.01187	0.00428	0.01945
Toronto, Canada	0.01518	0.00758	0.02277

First Model AICc: 2550.204

Second Model AICc: 6310.256 (abstol convergence only)

Third Model AICc: 9277.048 (abstol convergence only)

Discussion

The results of my analysis show that every city showed a significant increase in temperature. This supports the evidence that urban cities have been heating up in the last 100 years. From the data, we can see that not all cities are heating up at the same rate. Something to note is that cities closer to the equator seem to experience less change than those far from the equator (Figure 1). This result was the inverse of my original hypothesis, which can actually be explained by "polar amplification" which states that any change in the net radiation of the Earth seems to affect the poles more severely than the rest of the regions (Turton, 2021).

On the less severe side were Nairobi, Ankara, Bogota, and Cali (in increasing order – Figure 2) with a slope estimate of 0.00903, 0.00936, 0.00997, 0.01007 respectively – meaning that they had an average increase of around 0.9 to 1 degree Celsius increase in the past 113 years. On the more extreme side was Sao Paulo, Madrid, Saint Petersburg, and Toronto, with the slope estimates of 0.01187, 0.01204, 0.01329, 0.01518 respectively – with a range of around 1.18 to 1.51 degrees Celsius increase in the past 113 years.

Surprisingly, the more complex models did not do very well. Both the second and third model had AICc's that were much higher than the first model's AICc (2550.204). Additionally, the two models also failed to converge completely. Additionally, the second and third model took a very long time to run, while the first model took very little time. This means that the model's output is uncertain, and therefore the results should not be used. A couple of reasons for this could be that the climatic zones are either not a good feature to group cities by, there wasn't enough data, or that the other parameters could have been better adjusted for this model. It is concerning to see that many cities that are closer to the pole (Paris, Saint Petersburg, and Toronto) have experienced much more significant increases in temperature.

Acknowledgements

I thank the many people who helped advise me on the project – Albert Ruhi and Kyle Leathers for initial project brainstorming and providing guidance in how to develop the time series analysis. Thank you to the Lawrence Berkeley National Library for providing comprehensive temperature data.

Literature Cited

Environmental Protection Agency. "Heat Island Effect." *Environmental Protection Agency*, www.epa.gov/heatislands.

Leighton, Heather. "How Climate Change is Going to Affect Cities and Urban Spaces.'" *Rice University Kinder Institute for Urban Research*, Kinder Institute for Urban Research, 3 Sept. 2019, www.kinder.rice.edu/urbanedge/how-climate-change-going-affect-cities-urban-spaces#:~:text=According%20to%20a%202018%20study,National%20Climate%20Assessment%20concluded%3A%20%E2%80%9CHeavy."

Turton, Steve. "Climate Explained: Why Is the Arctic Warming Faster than Other Parts of the World?" *International Science Council*, 8 Dec. 2021, council.science/current/blog/climate-explained-why-is-the-arctic-warming-faster-than-other-parts-of-the-world/#:~:text=This%20warming%20differential%20between%20the,Arctic%20(or%20polar)%20amplification.&text=It%20occurs%20whenever%20there%20is,poles%20than%20the%20global%20average.

Tables and Figures

Tables

City	Lower.CI	Upper.CI	Slope
ANK	0.00176	0.01697	0.00936
BOG	0.00239	0.01756	0.00997
CALI	0.00249	0.01765	0.01007
CAPE	0.00291	0.01808	0.01049
FOR	0.00288	0.01806	0.01047
LA	0.00380	0.01890	0.01135
MAD	0.00444	0.01963	0.01204
NAI	0.00144	0.01662	0.00903
NAN	0.00378	0.01896	0.01137
PAR	0.00409	0.01927	0.01168
RIO	0.00369	0.01886	0.01128
STP	0.00568	0.02089	0.01329
SIN	0.00270	0.01786	0.01028
SAO	0.00428	0.01945	0.01187
TOR	0.00758	0.02277	0.01518

Table 1: Table that contains the MARSS analysis on the estimated slope and the upper 95% confidence interval and lower 95% confidence interval.

Figures

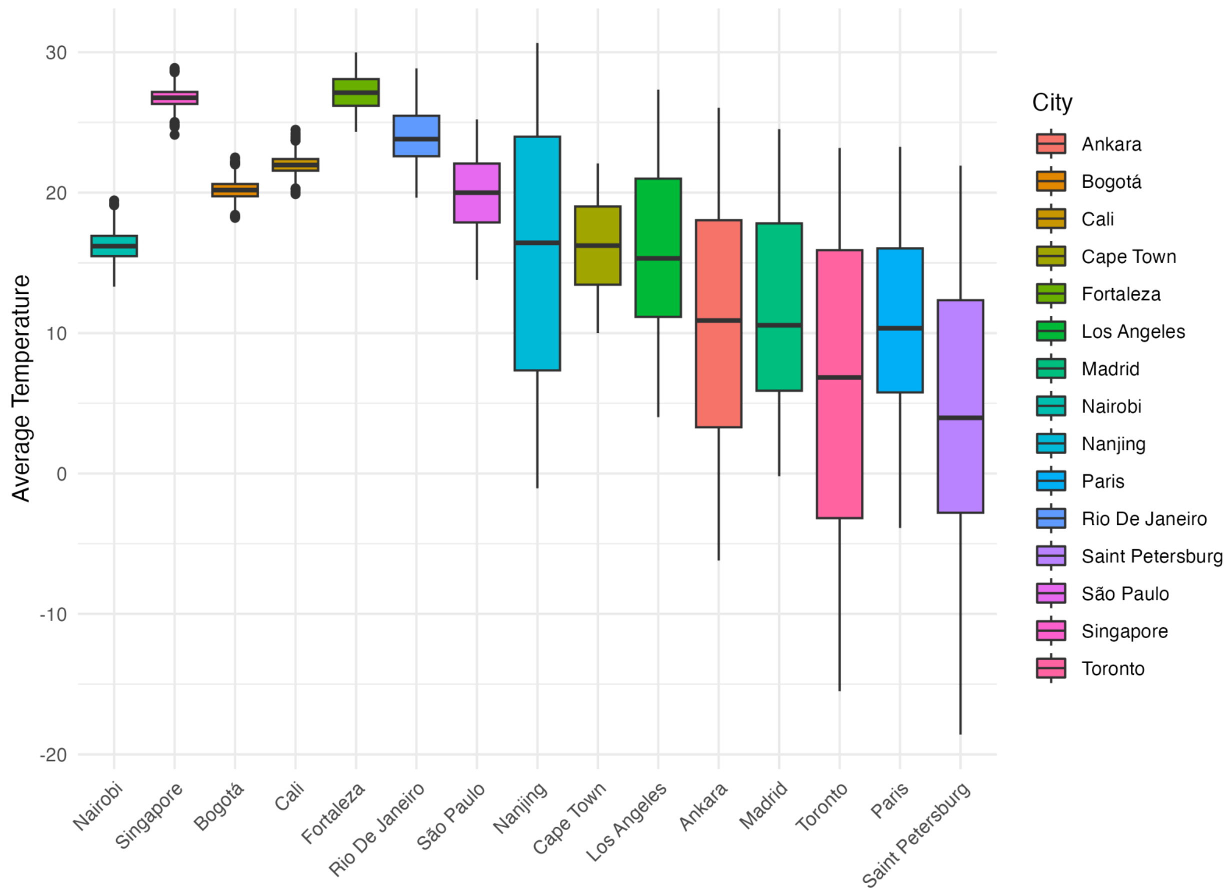

Figure 1: Boxplot that shows the range of annual temperature (in Celsius) of a particular city in the past 113 years from 1900 to 2013. The x-axis is ordered by distance from the equator (Nairobi is closest, while Saint Petersburg is the farthest)

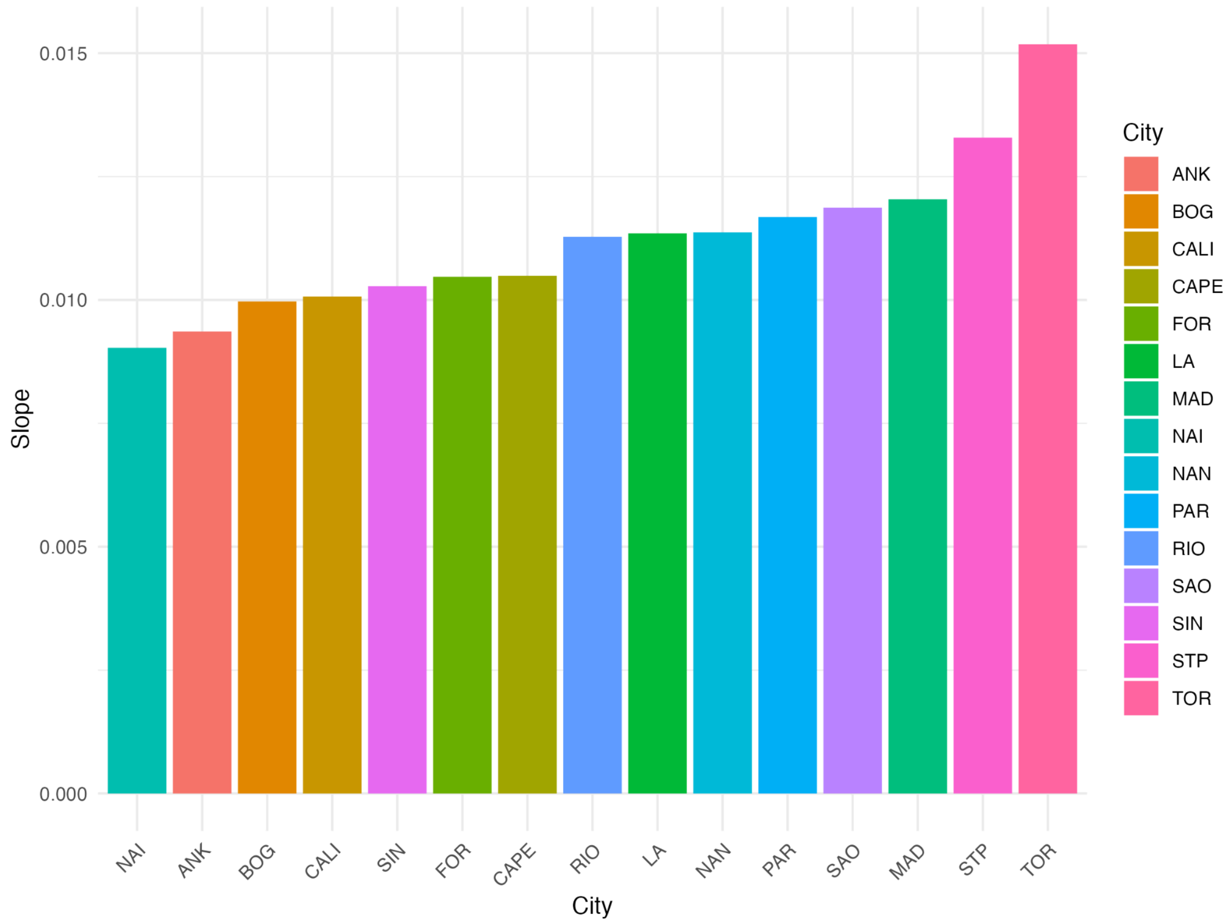

Figure 2: Bar chart showing the slopes of each city in ascending order.

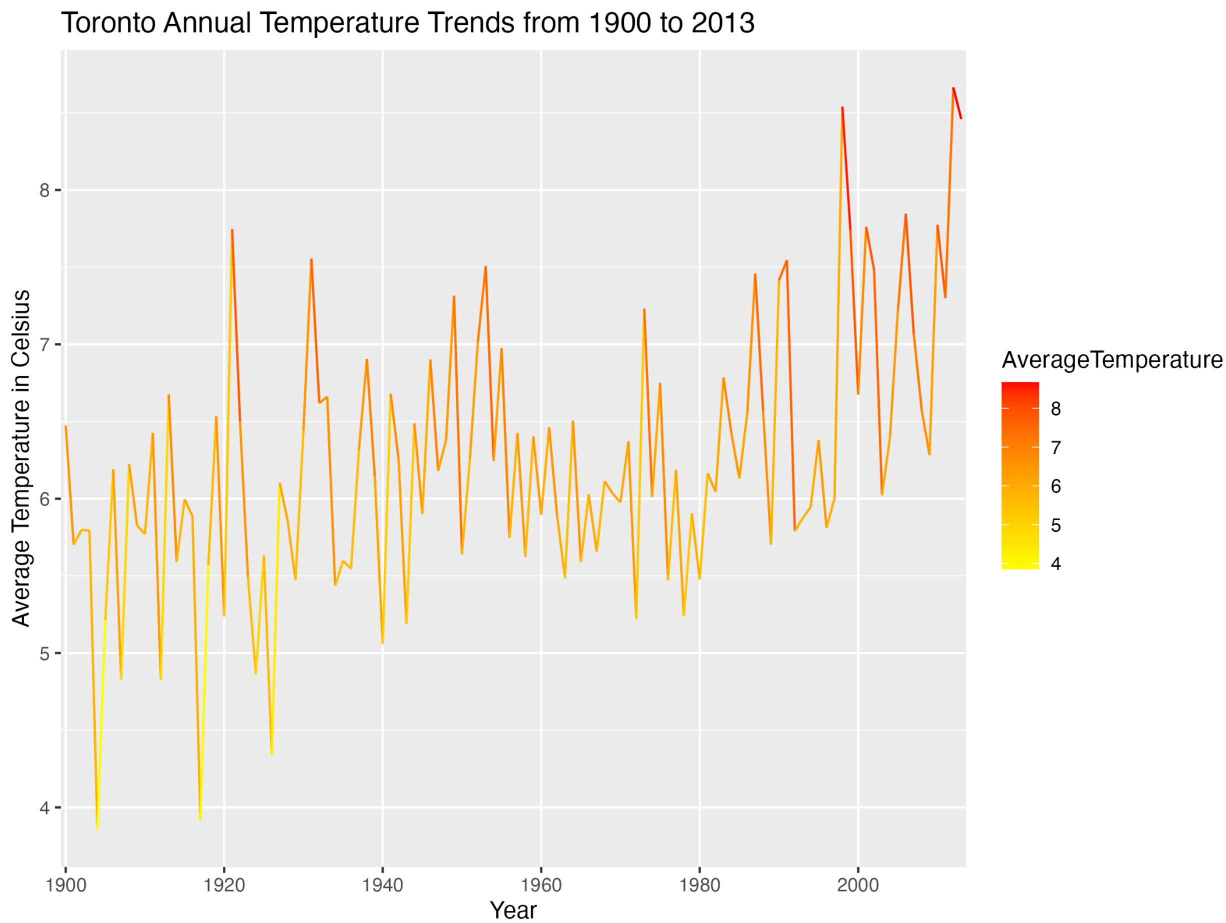

Figure 3: Line plot of Toronto's annual temperature trends from 1900 to 2013. Toronto was chosen since it shows the most drastic increase in temperature.