Global Warming

Effrosyni Simou

École Polytechnique Fédérale de Lausanne

effrosyni.simou@epfl.ch

January 27, 2017

Overview

Aim of the Project

- Since the 2016 Presidential Elections in the USA, the interest of people with regards to climate change and the correct environmental policy has reached an all-time high.
- Use a dataset with temperature data from 1750 to 2015 [1] and check whether global warming is a fact or a speculation.
- The dataset is nicely packaged and allows for slicing into interesting subsets (by country, by city, global temperatures e.t.c.). It was put together by Berkeley Earth.

Data Acquisition

Cleaning the Data

There is a need to clean our data:

- There are missing data.
- There are duplicates in our data.
- Older measurements are less reliable.

Missing Data

	dt	AverageTemperature	AverageTemperatureUncertainty	State	Country
0	1855-05-01	25.544	1.171	Acre	Brazil
1	1855-06-01	24.228	1.103	Acre	Brazil
2	1855-07-01	24.371	1.044	Acre	Brazil
3	1855-08-01	25.427	1.073	Acre	Brazil
4	1855-09-01	25.675	1.014	Acre	Brazil
5	1855-10-01	25.442	1.179	Acre	Brazil
6	1855-11-01	25.400	1.064	Acre	Brazil
7	1855-12-01	24.100	1.718	Acre	Brazil
8	1856-01-01	25.814	1.159	Acre	Brazil
9	1856-02-01	24.658	1.147	Acre	Brazil
10	1856-03-01	24.659	1.547	Acre	Brazil
11	1856-04-01	24.907	1.186	Acre	Brazil
12	1856-05-01	24.418	1.168	Acre	Brazil
13	1856-06-01	24.930	1.355	Acre	Brazil
14	1856-07-01	NaN	NaN	Acre	Brazil
15	1856-08-01	24.615	1.053	Acre	Brazil
16	1856-09-01	25.150	1.093	Acre	Brazil
17	1856-10-01	25.196	1.048	Acre	Brazil
18	1856-11-01	25.441	1.102	Acre	Brazil
19	1856-12-01	25.362	1.382	Acre	Brazil
20	1857-01-01	NaN	NaN	Acre	Brazil
21	1857-02-01	NaN	NaN	Acre	Brazil
22	1857-03-01	NaN	NaN	Acre	Brazil
23	1857-04-01	NaN	NaN	Acre	Brazil
24	1857-05-01	NaN	NaN	Acre	Brazil
25	1857-06-01	NaN	NaN	Acre	Brazil
26	1857-07-01	NaN	NaN	Acre	Brazil
27	1857-08-01	NaN	NaN	Acre	Brazil
28	1857-09-01	NaN	NaN	Acre	Brazil
29	1857-10-01	NaN	NaN	Acre	Brazil
30	1857-11-01	NaN	NaN	Acre	Brazil
31	1857-12-01	NaN	NaN	Acre	Brazil
32	1858-01-01	NaN	NaN	Acre	Brazil
33	1858-02-01	NaN	NaN	Acre	Brazil
34	1858-03-01	NaN	NaN	Acre	Brazil
35	1858-04-01	NaN	NaN	Acre	Brazil
36	1858-05-01	NaN	NaN	Acre	Brazil
37	1858-06-01	NaN	NaN	Acre	Brazil
38	1858-07-01	NaN	NaN	Acre	Brazil
39	1858-08-01	NaN	NaN	Acre	Brazil
40	1858-09-01	NaN	NaN	Acre	Brazil
41	1858-10-01	NaN	NaN	Acre	Brazil
42	1858-11-01	NaN	NaN	Acre	Brazil
43	1858-12-01	NaN	NaN	Acre	Brazil
44	1859-01-01	NaN	NaN	Acre	Brazil
45	1859-02-01	NaN	NaN	Acre	Brazil
46	1859-03-01	NaN	NaN	Acre	Brazil
47	1859-04-01	NaN	NaN	Acre	Brazil
48	1859-05-01	NaN	NaN	Acre	Brazil
49	1859-06-01	NaN	NaN	Acre	Brazil
50	1859-07-01	NaN	NaN	Acre	Brazil

	dt	LandAverageTemperature	LandAverageTemperatureUncertainty	LandMaxTemperature	LandMaxTemperatureUncertainty	LandMinTemperature	LandMinTemperatureUncertainty	LandAndOceanAverageTemperature	LandAndOceanAverageTemperatureUncertainty
0	1750-01-01	3.034	3.574	NaN	NaN	NaN	NaN	NaN	NaN
1	1750-02-01	3.083	3.702	NaN	NaN	NaN	NaN	NaN	NaN
2	1750-03-01	5.626	3.076	NaN	NaN	NaN	NaN	NaN	NaN
3	1750-04-01	8.490	2.451	NaN	NaN	NaN	NaN	NaN	NaN
4	1750-05-01	11.573	2.072	NaN	NaN	NaN	NaN	NaN	NaN
5	1750-06-01	12.937	1.724	NaN	NaN	NaN	NaN	NaN	NaN
6	1750-07-01	15.868	1.911	NaN	NaN	NaN	NaN	NaN	NaN
7	1750-08-01	14.750	2.231	NaN	NaN	NaN	NaN	NaN	NaN
8	1750-09-01	11.413	2.637	NaN	NaN	NaN	NaN	NaN	NaN
9	1750-10-01	6.367	2.668	NaN	NaN	NaN	NaN	NaN	NaN
10	1750-11-01	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
11	1750-12-01	2.772	2.970	NaN	NaN	NaN	NaN	NaN	NaN
12	1751-01-01	2.495	3.469	NaN	NaN	NaN	NaN	NaN	NaN
13	1751-02-01	0.963	3.827	NaN	NaN	NaN	NaN	NaN	NaN
14	1751-03-01	5.800	3.051	NaN	NaN	NaN	NaN	NaN	NaN
15	1751-04-01	7.670	2.368	NaN	NaN	NaN	NaN	NaN	NaN
16	1751-05-01	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
17	1751-06-01	13.827	1.801	NaN	NaN	NaN	NaN	NaN	NaN
18	1751-07-01	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
19	1751-08-01	14.405	2.296	NaN	NaN	NaN	NaN	NaN	NaN
20	1751-09-01	10.673	2.656	NaN	NaN	NaN	NaN	NaN	NaN
21	1751-10-01	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
22	1751-11-01	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
23	1751-12-01	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
24	1752-01-01	0.348	3.789	NaN	NaN	NaN	NaN	NaN	NaN
25	1752-02-01	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
26	1752-03-01	5.806	3.192	NaN	NaN	NaN	NaN	NaN	NaN
27	1752-04-01	8.265	2.398	NaN	NaN	NaN	NaN	NaN	NaN
28	1752-05-01	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
29	1752-06-01	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
30	1752-07-01	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
31	1752-08-01	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
32	1752-09-01	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
33	1752-10-01	7.839	2.754	NaN	NaN	NaN	NaN	NaN	NaN
34	1752-11-01	7.335	2.747	NaN	NaN	NaN	NaN	NaN	NaN
35	1752-12-01	5.086	2.982	NaN	NaN	NaN	NaN	NaN	NaN
36	1753-01-01	2.039	3.478	NaN	NaN	NaN	NaN	NaN	NaN
37	1753-02-01	0.715	6.341	NaN	NaN	NaN	NaN	NaN	NaN
38	1753-03-01	5.424	2.929	NaN	NaN	NaN	NaN	NaN	NaN
39	1753-04-01	8.272	3.177	NaN	NaN	NaN	NaN	NaN	NaN
40	1753-05-01	12.186	3.111	NaN	NaN	NaN	NaN	NaN	NaN
41	1753-06-01	14.585	2.590	NaN	NaN	NaN	NaN	NaN	NaN
42	1753-07-01	15.092	1.929	NaN	NaN	NaN	NaN	NaN	NaN
43	1753-08-01	14.148	2.453	NaN	NaN	NaN	NaN	NaN	NaN
44	1753-09-01	13.322	2.538	NaN	NaN	NaN	NaN	NaN	NaN
45	1753-10-01	8.962	2.585	NaN	NaN	NaN	NaN	NaN	NaN
46	1753-11-01	5.353	2.807	NaN	NaN	NaN	NaN	NaN	NaN
47	1753-12-01	0.559	4.174	NaN	NaN	NaN	NaN	NaN	NaN
48	1754-01-01	1.574	5.020	NaN	NaN	NaN	NaN	NaN	NaN
49	1754-02-01	-1.249	7.492	NaN	NaN	NaN	NaN	NaN	NaN
50	1754-03-01	3.677	3.336	NaN	NaN	NaN	NaN	NaN	NaN

	dt	AverageTemperature	AverageTemperatureUncertainty	City	Country	Latitude	Longitude
0	1743-11-01	6.068	1.737	Ã…rhus	Denmark	57.05N	10.33E
1	1743-12-01	NaN	NaN	Ã…rhus	Denmark	57.05N	10.33E
2	1744-01-01	NaN	NaN	Ã…rhus	Denmark	57.05N	10.33E
3	1744-02-01	NaN	NaN	Ã…rhus	Denmark	57.05N	10.33E
4	1744-03-01	NaN	NaN	Ã…rhus	Denmark	57.05N	10.33E
5	1744-04-01	5.788	3.624	Ã…rhus	Denmark	57.05N	10.33E
6	1744-05-01	10.644	1.283	Ã…rhus	Denmark	57.05N	10.33E
7	1744-06-01	14.051	1.347	Ã…rhus	Denmark	57.05N	10.33E
8	1744-07-01	16.082	1.396	Ã…rhus	Denmark	57.05N	10.33E
9	1744-08-01	NaN	NaN	Ã…rhus	Denmark	57.05N	10.33E
10	1744-09-01	12.781	1.454	Ã…rhus	Denmark	57.05N	10.33E
11	1744-10-01	7.950	1.630	Ã…rhus	Denmark	57.05N	10.33E
12	1744-11-01	4.639	1.302	Ã…rhus	Denmark	57.05N	10.33E
13	1744-12-01	0.122	1.756	Ã…rhus	Denmark	57.05N	10.33E
14	1745-01-01	-1.333	1.642	Ã…rhus	Denmark	57.05N	10.33E
15	1745-02-01	-2.732	1.358	Ã…rhus	Denmark	57.05N	10.33E
16	1745-03-01	0.129	1.088	Ã…rhus	Denmark	57.05N	10.33E
17	1745-04-01	4.042	1.138	Ã…rhus	Denmark	57.05N	10.33E
18	1745-05-01	NaN	NaN	Ã…rhus	Denmark	57.05N	10.33E
19	1745-06-01	NaN	NaN	Ã…rhus	Denmark	57.05N	10.33E
20	1745-07-01	NaN	NaN	Ã…rhus	Denmark	57.05N	10.33E
21	1745-08-01	NaN	NaN	Ã…rhus	Denmark	57.05N	10.33E

	dt	AverageTemperature	AverageTemperatureUncertainty	Country
0	1743-11-01	4.384	2.294	Ã…land
1	1743-12-01	NaN	NaN	Ã…land
2	1744-01-01	NaN	NaN	Ã…land
3	1744-02-01	NaN	NaN	Ã…land
4	1744-03-01	NaN	NaN	Ã…land
5	1744-04-01	1.530	4.680	Ã…land
6	1744-05-01	6.702	1.789	Ã…land
7	1744-06-01	11.609	1.577	Ã…land
8	1744-07-01	15.342	1.410	Ã…land
9	1744-08-01	NaN	NaN	Ã…land
10	1744-09-01	11.702	1.517	Ã…land
11	1744-10-01	5.477	1.862	Ã…land
12	1744-11-01	3.407	1.425	Ã…land
13	1744-12-01	-2.181	1.641	Ã…land
14	1745-01-01	-3.850	1.841	Ã…land
15	1745-02-01	-6.575	1.360	Ã…land
16	1745-03-01	-4.195	1.213	Ã…land
17	1745-04-01	-0.966	1.172	Ã…land
18	1745-05-01	NaN	NaN	Ã…land
19	1745-06-01	NaN	NaN	Ã…land
20	1745-07-01	NaN	NaN	Ã…land
21	1745-08-01	NaN	NaN	Ã…land
22	1745-09-01	NaN	NaN	Ã…land
23	1745-10-01	NaN	NaN	Ã…land
24	1745-11-01	NaN	NaN	Ã…land
25	1745-12-01	NaN	NaN	Ã…land
26	1746-01-01	NaN	NaN	Ã…land
27	1746-02-01	NaN	NaN	Ã…land
28	1746-03-01	NaN	NaN	Ã…land
29	1746-04-01	NaN	NaN	Ã…land
30	1746-05-01	NaN	NaN	Ã…land
31	1746-06-01	NaN	NaN	Ã…land

Working with the missing data

As far as the missing data is concerned we can chose to either:

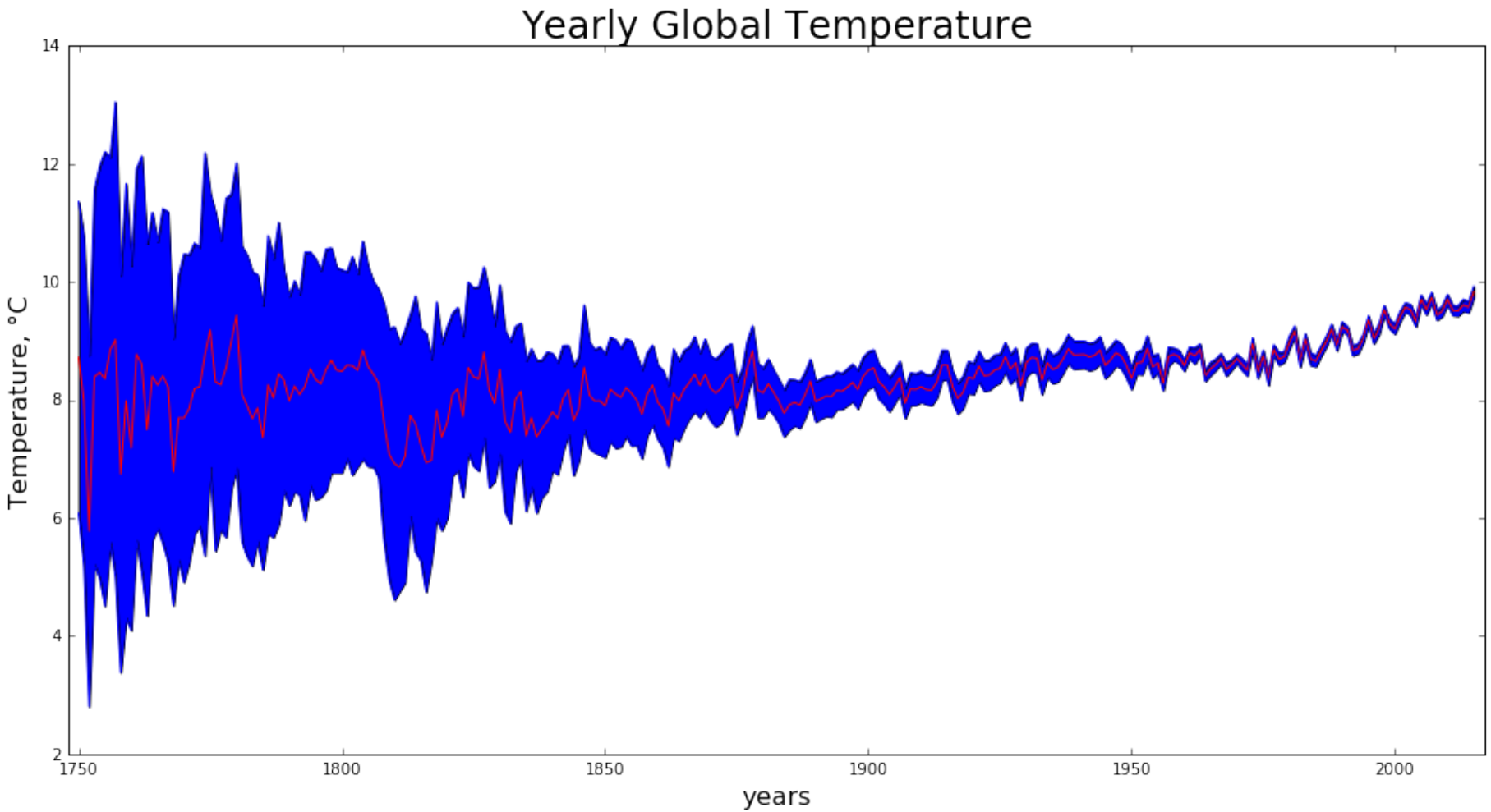
- Ignore the missing values
- Use the values we have in order to fill in the missing values (e.g. pad, interpolate e.t.c.).

For the purposes of this project we chose to ignore the missing values.

Removing duplicates

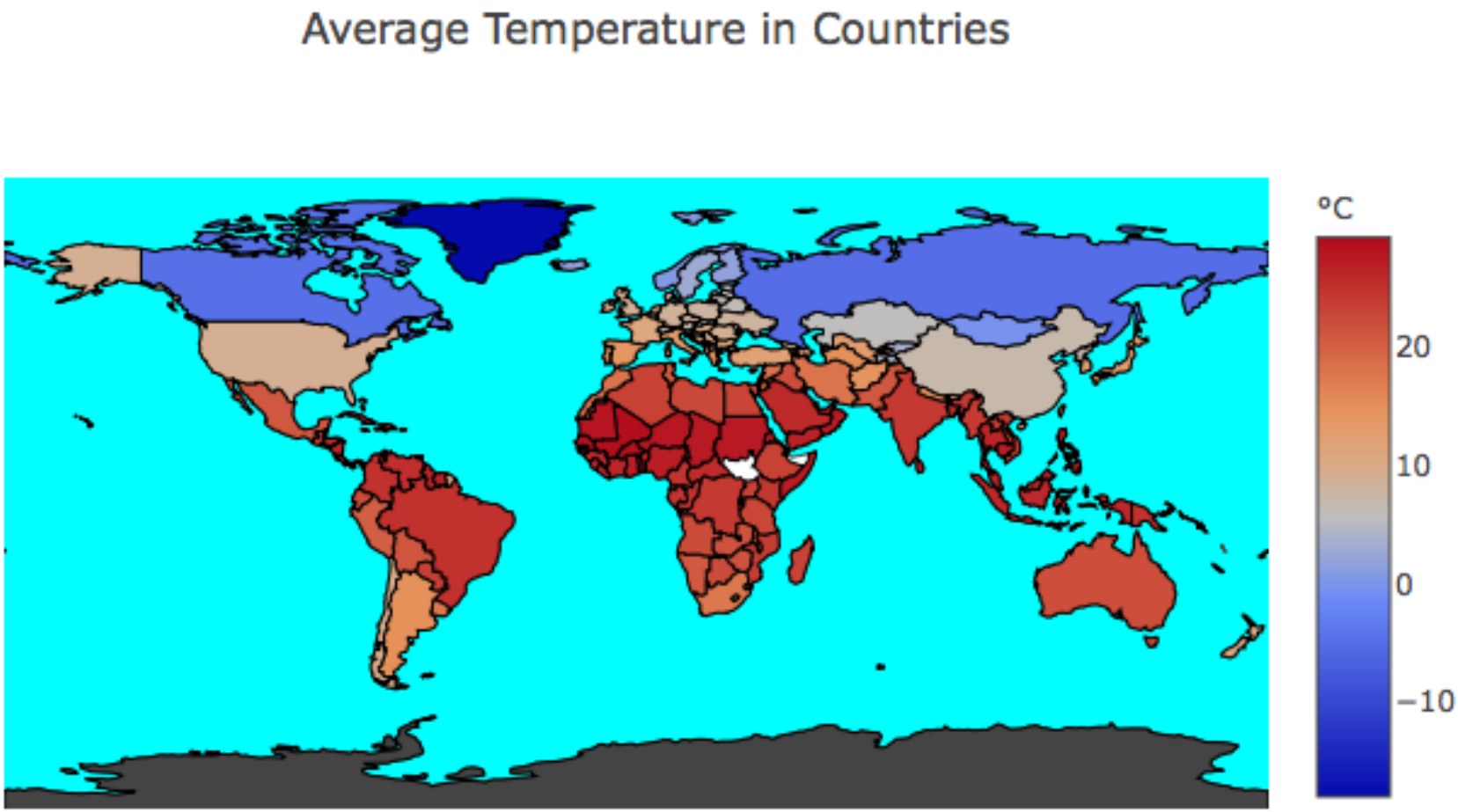
- The dataset was created by combining 16 pre-existing archives.
- In the case of temperatures by country the temperatures for Denmark, France, Netherlands and United Kingdom are duplicate.
- In the case of temperatures by city the temperatures for Guatemala City are duplicate.

Uncertainty of Measurements With Time

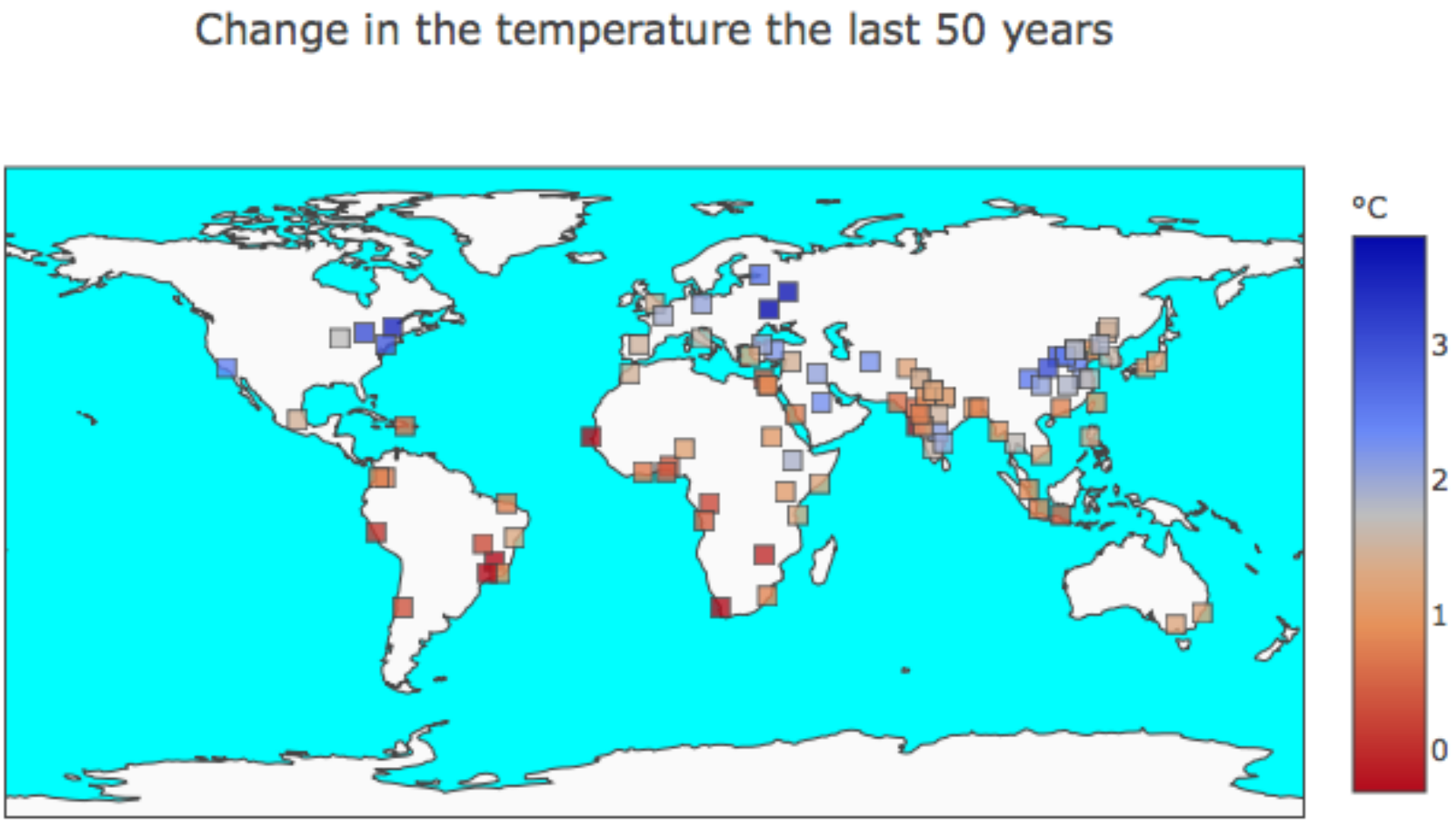


Data Exploration

Which countries are warmer?



Which cities have experienced the biggest change of temperature the last 50 years?

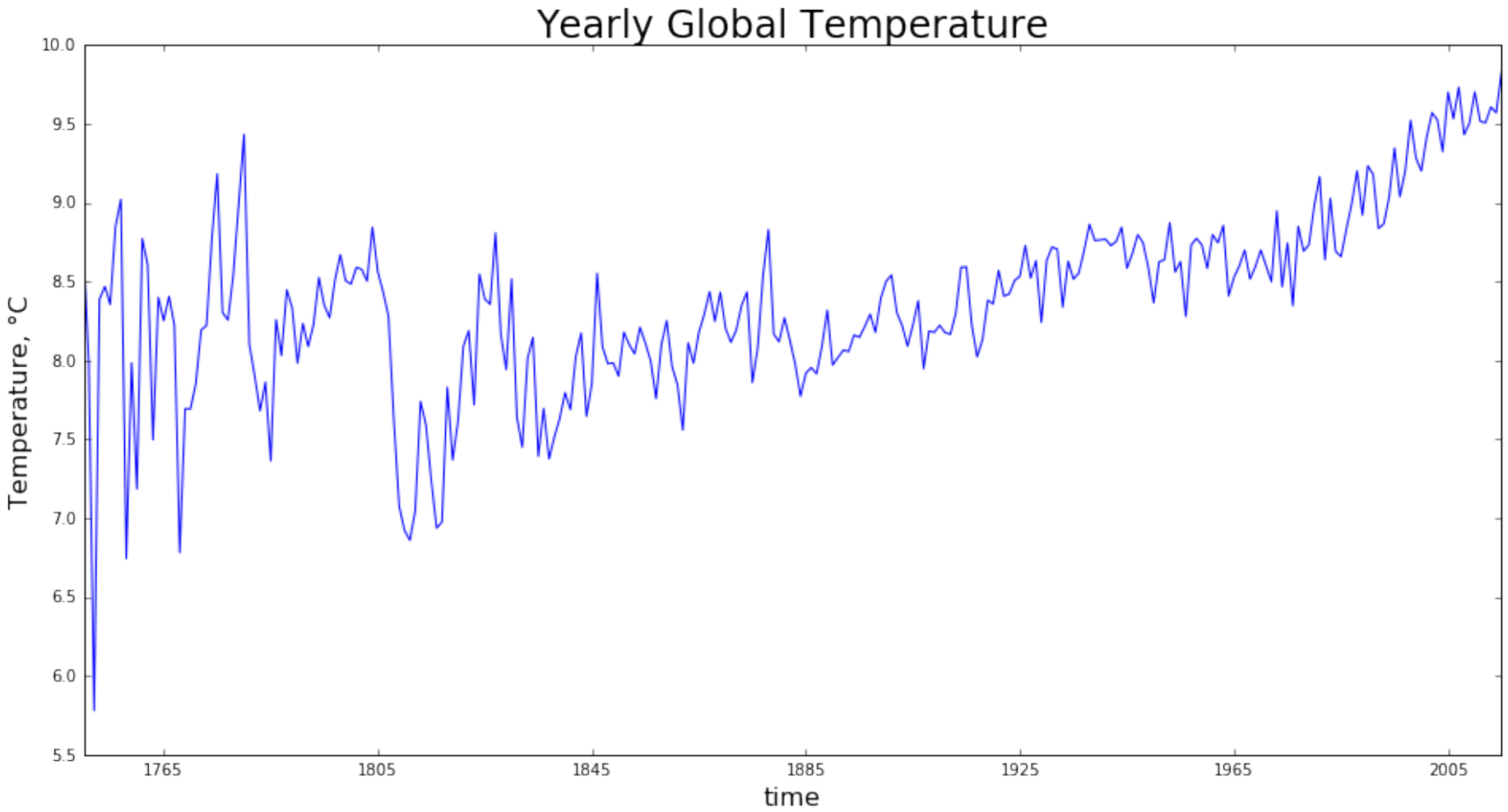


Data Exploitation

Data Partitioning

- *Train set*: Temperatures from 1900 to 2000
- *Test set*: Temperatures from 2001 to 2015

Non-stationary process



Making the process a stationary process

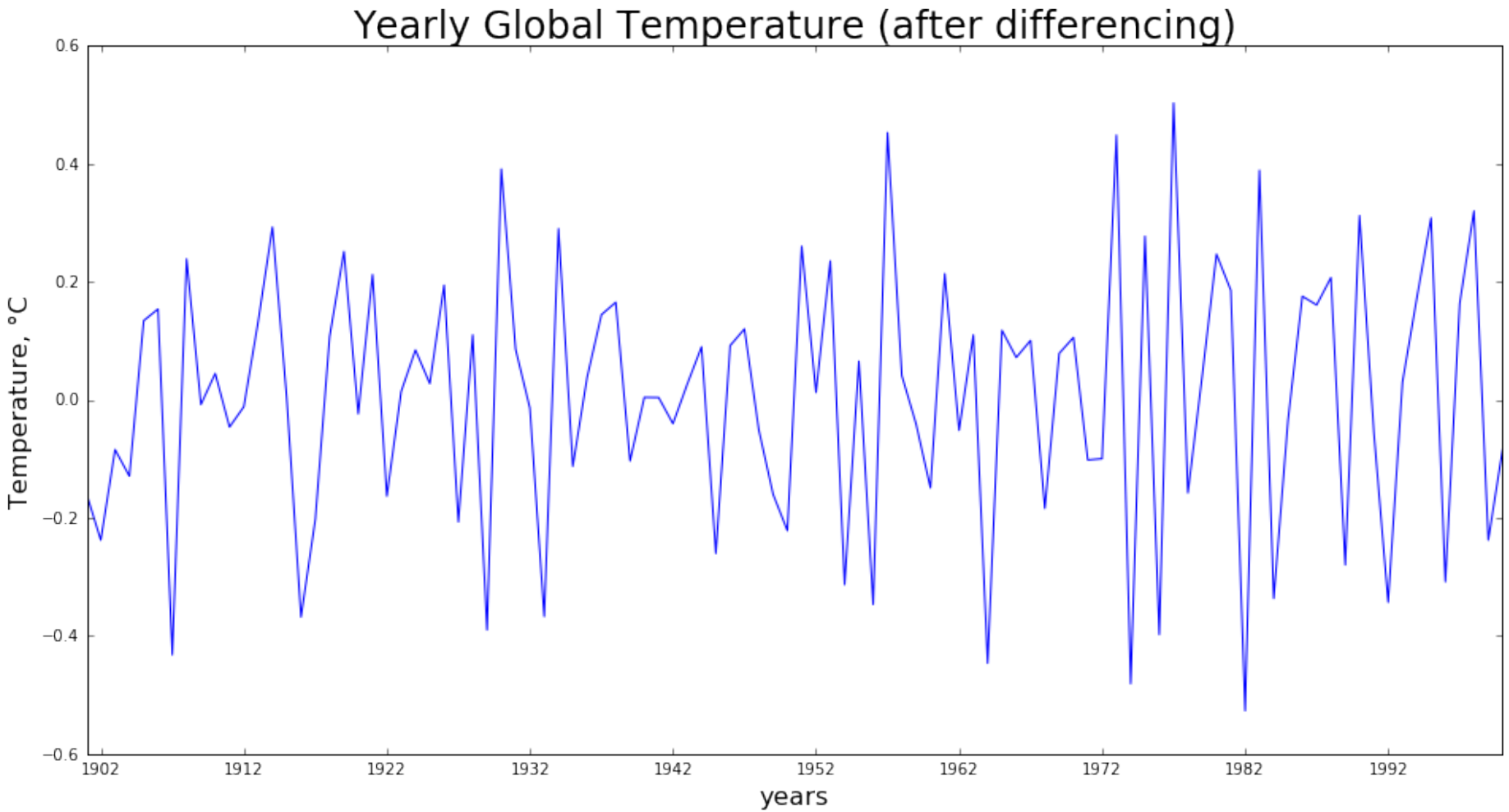
- Differencing
- Modeling the trend and then subtracting it from the data

Differencing

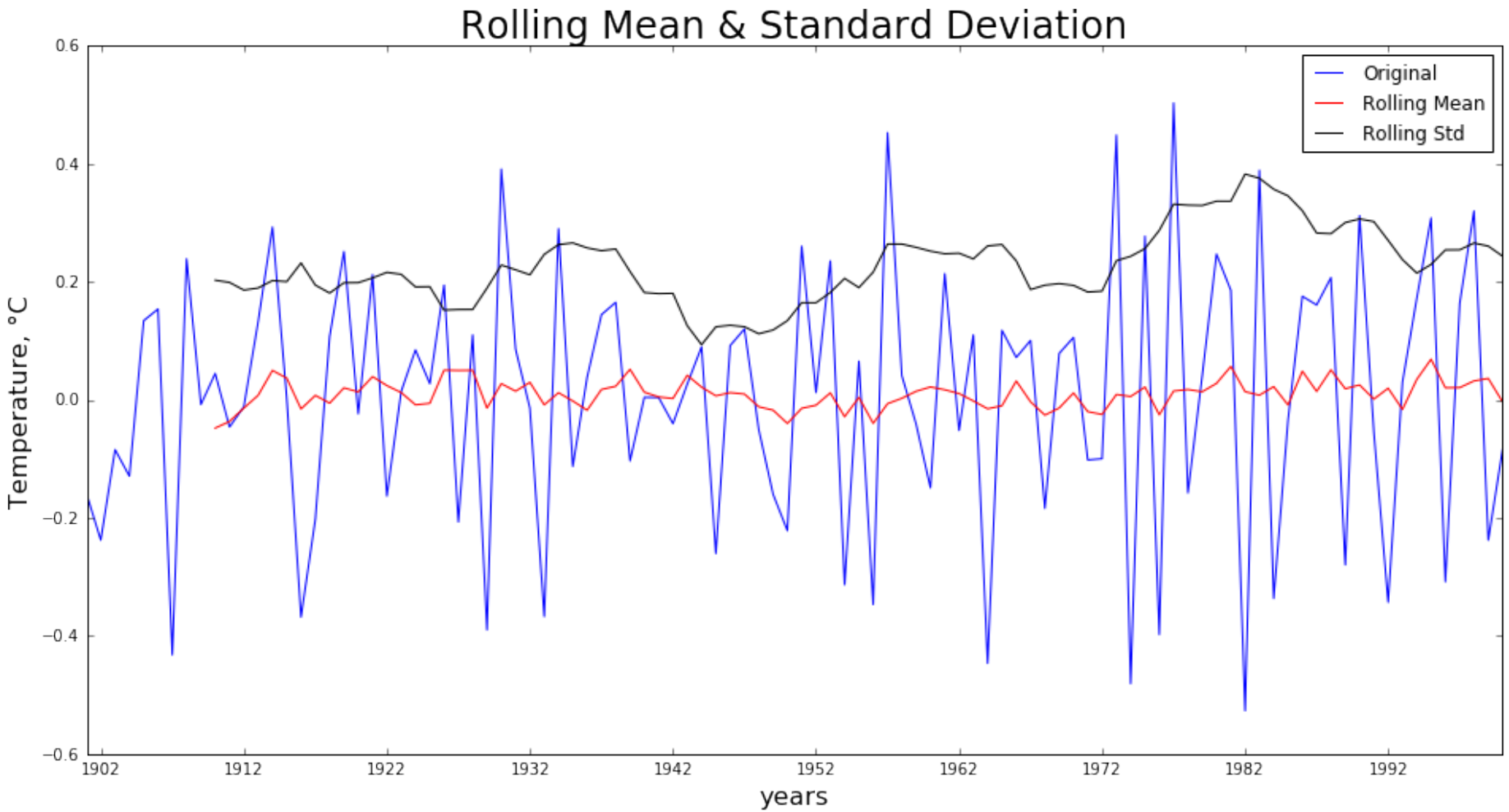
Given the non-stationary time series X, its corresponding time series after differencing X_{diff} is:

$$X_{diff}(i) = X(i) - X(i-1)$$

Differencing



Stationarity Test After Differencing



Dickey-Fuller Test

Hypothesis Testing-Critical Value Approach

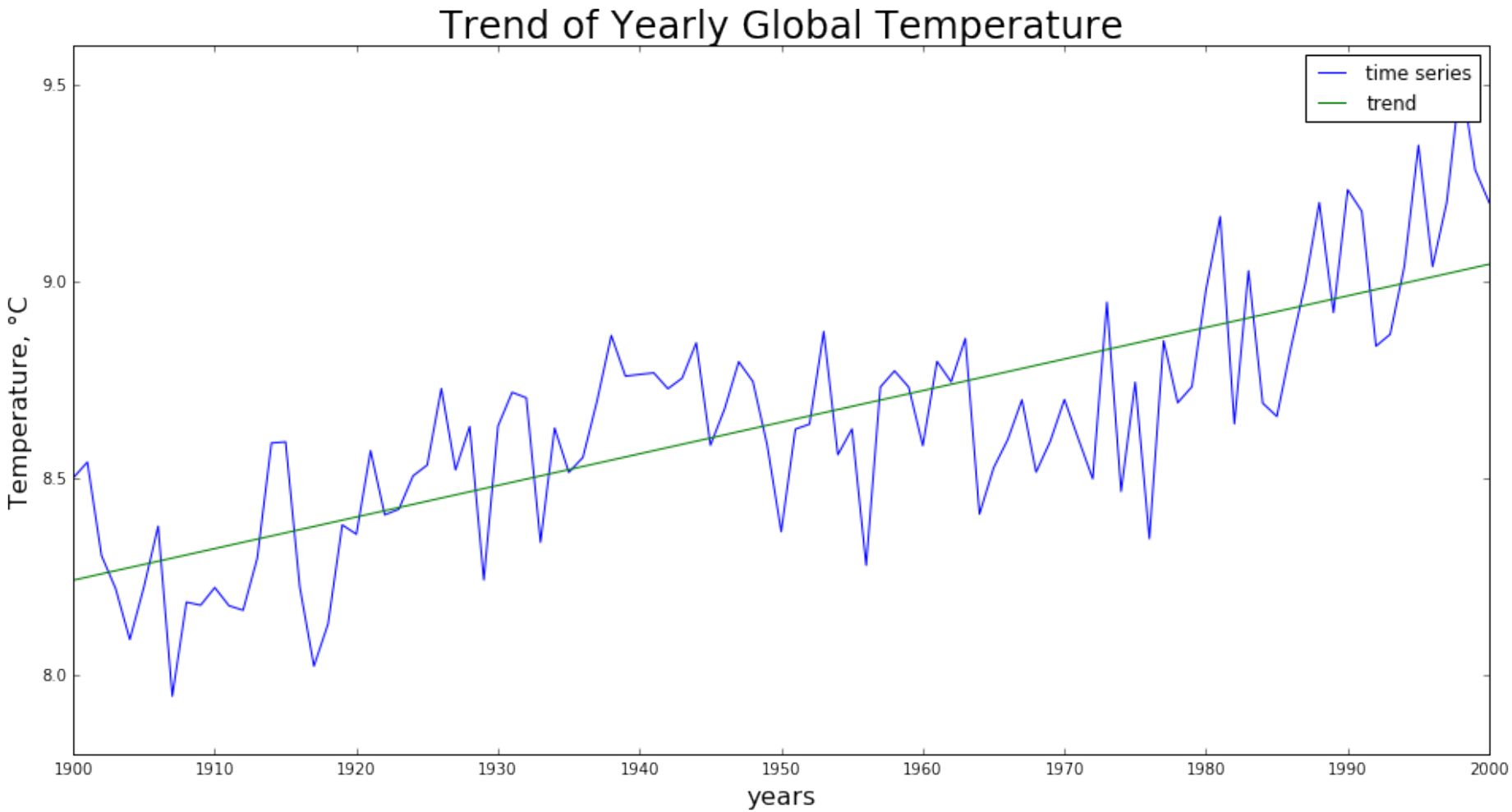
- Null hypothesis: the model has a unit root (equivalent to the model is non-stationary)
- If the test statistic is smaller than the critical value we can reject the null hypothesis for the alternative one

Results of Dickey-Fuller Test

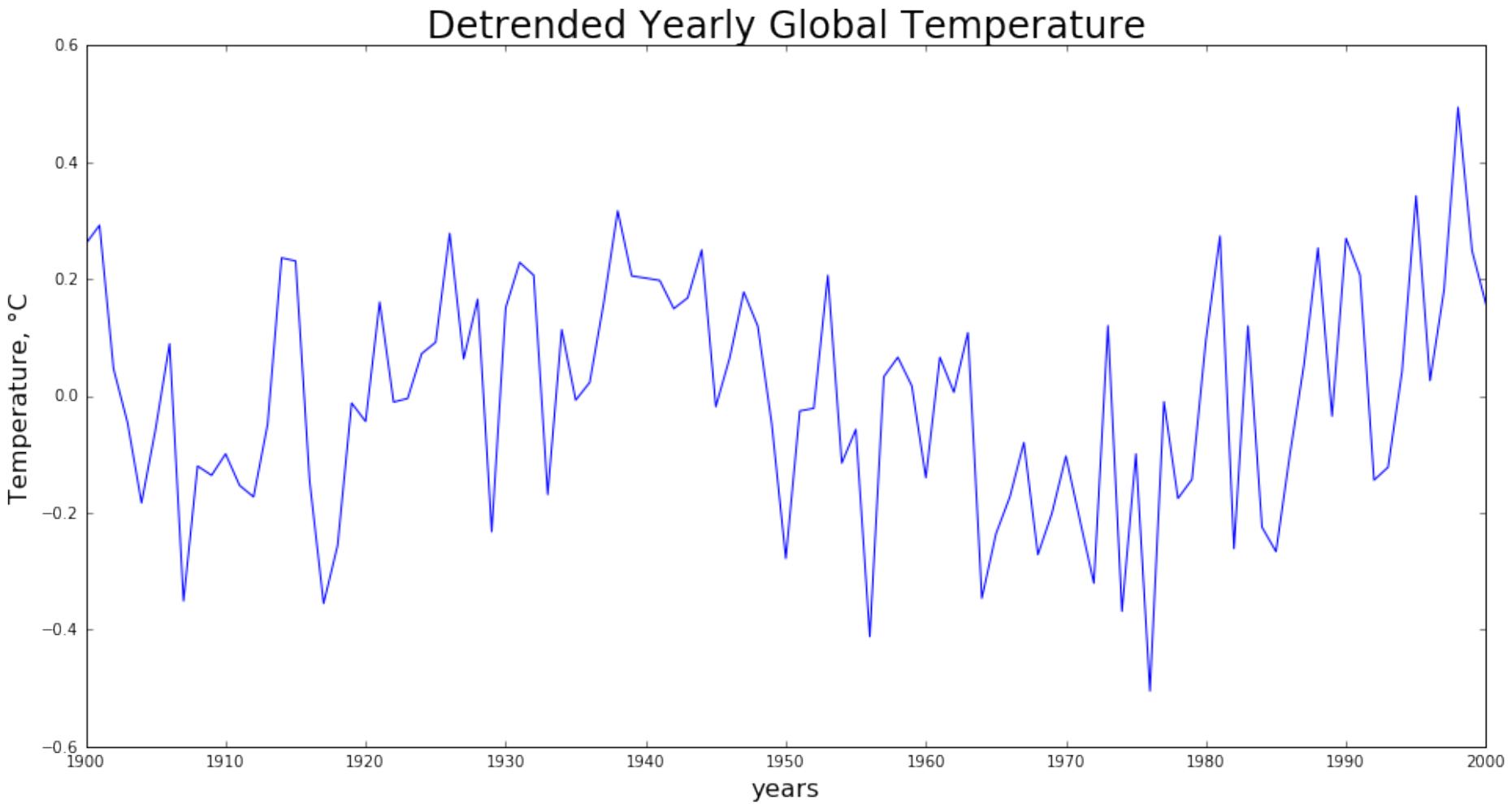
Test Statistic	-5.370426
Critical Value (1%)	-3.504343
Critical Value (5%)	-2.893866
Critical Value (10%)	-2.584015

The alternative hypothesis is accepted at the 1% level.

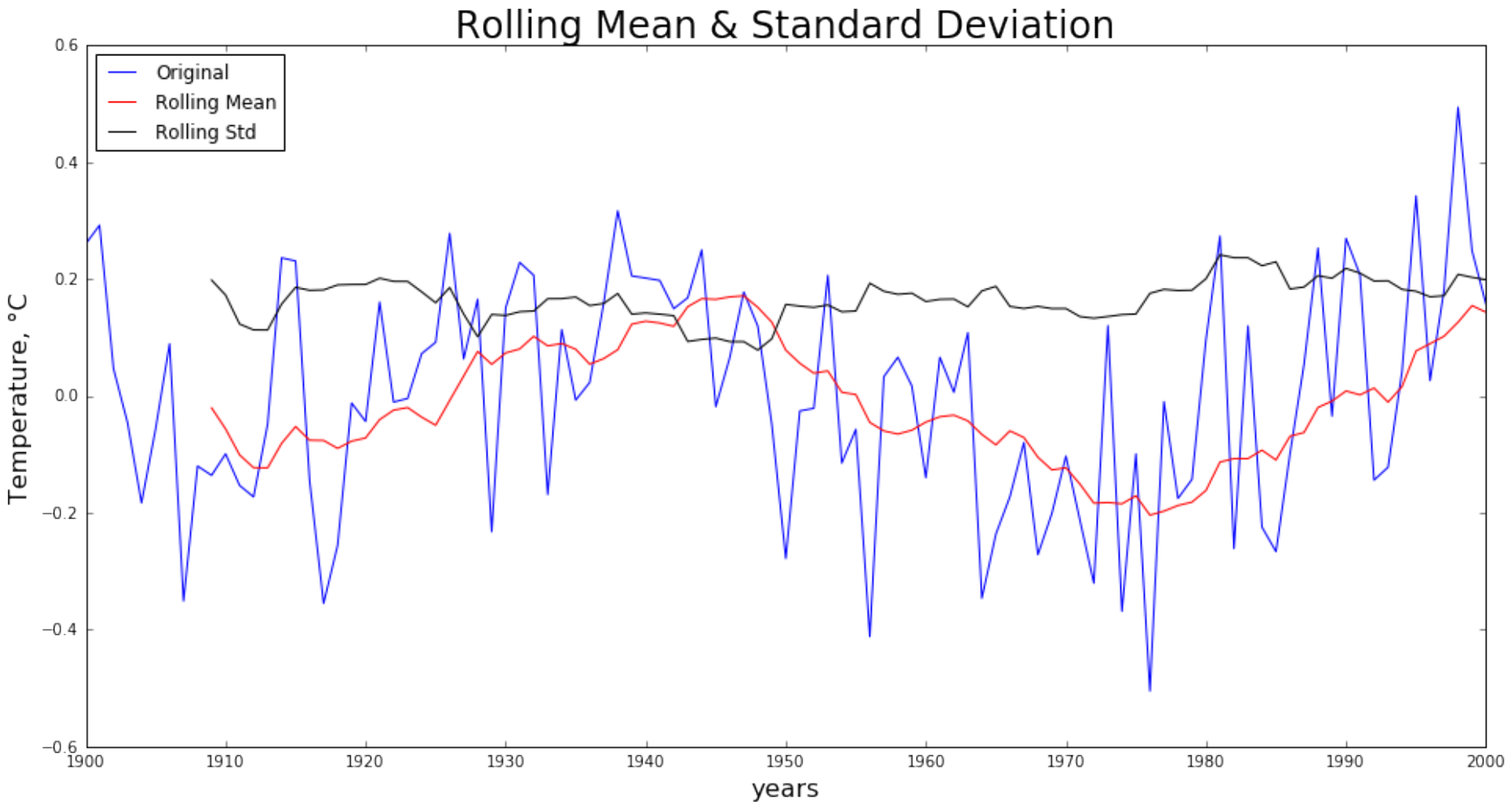
Model Fitting



Detrended Process



Stationarity Test After Modeling the Trend



Results of Dickey-Fuller Test

Test Statistic	–2.876054
Critical Value (1%)	-3.499637
Critical Value (5%)	-2.891831
Critical Value (10%)	–2.582928

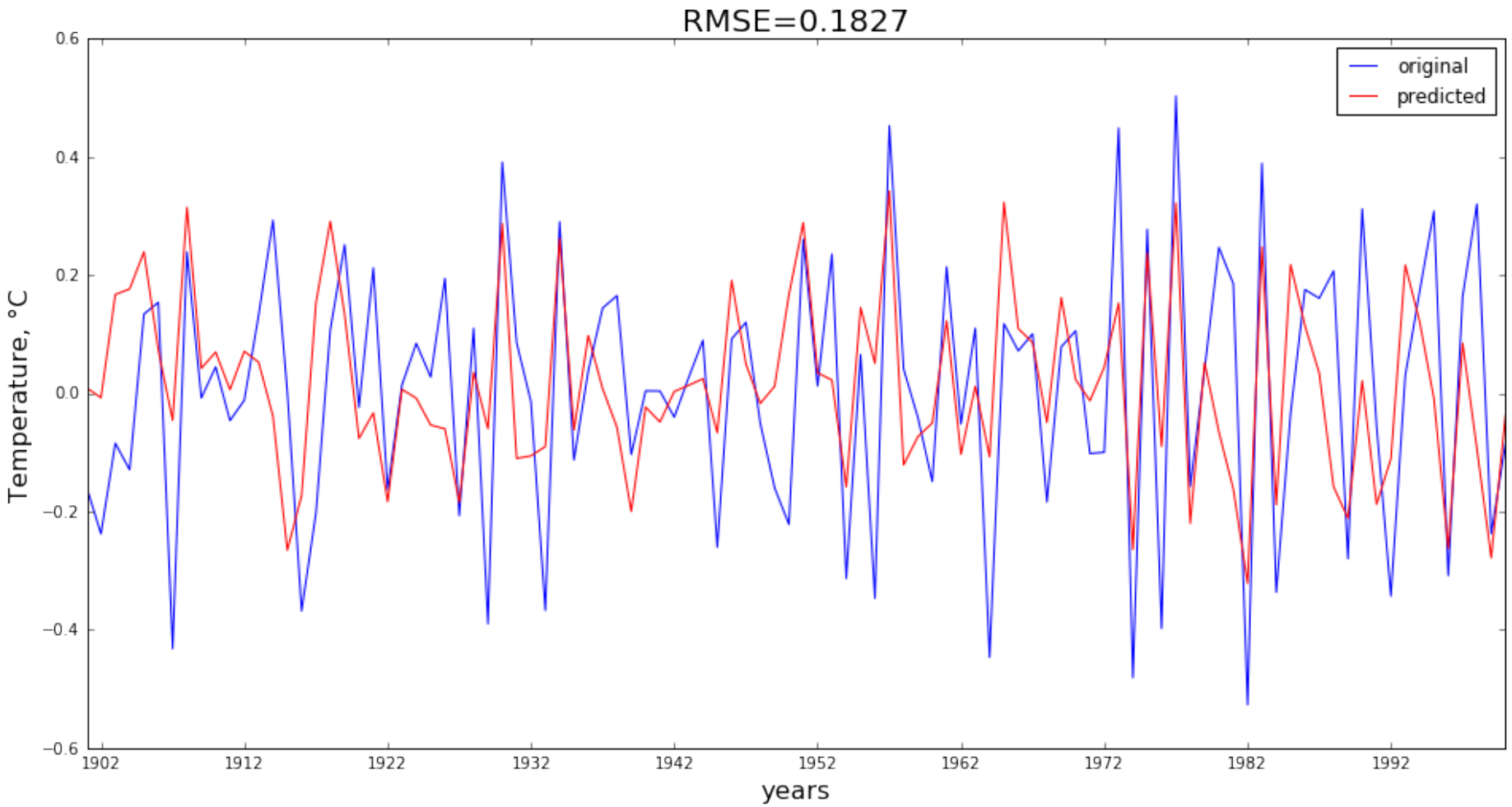
The alternative hypothesis is accepted at the 10% level.

Modeling

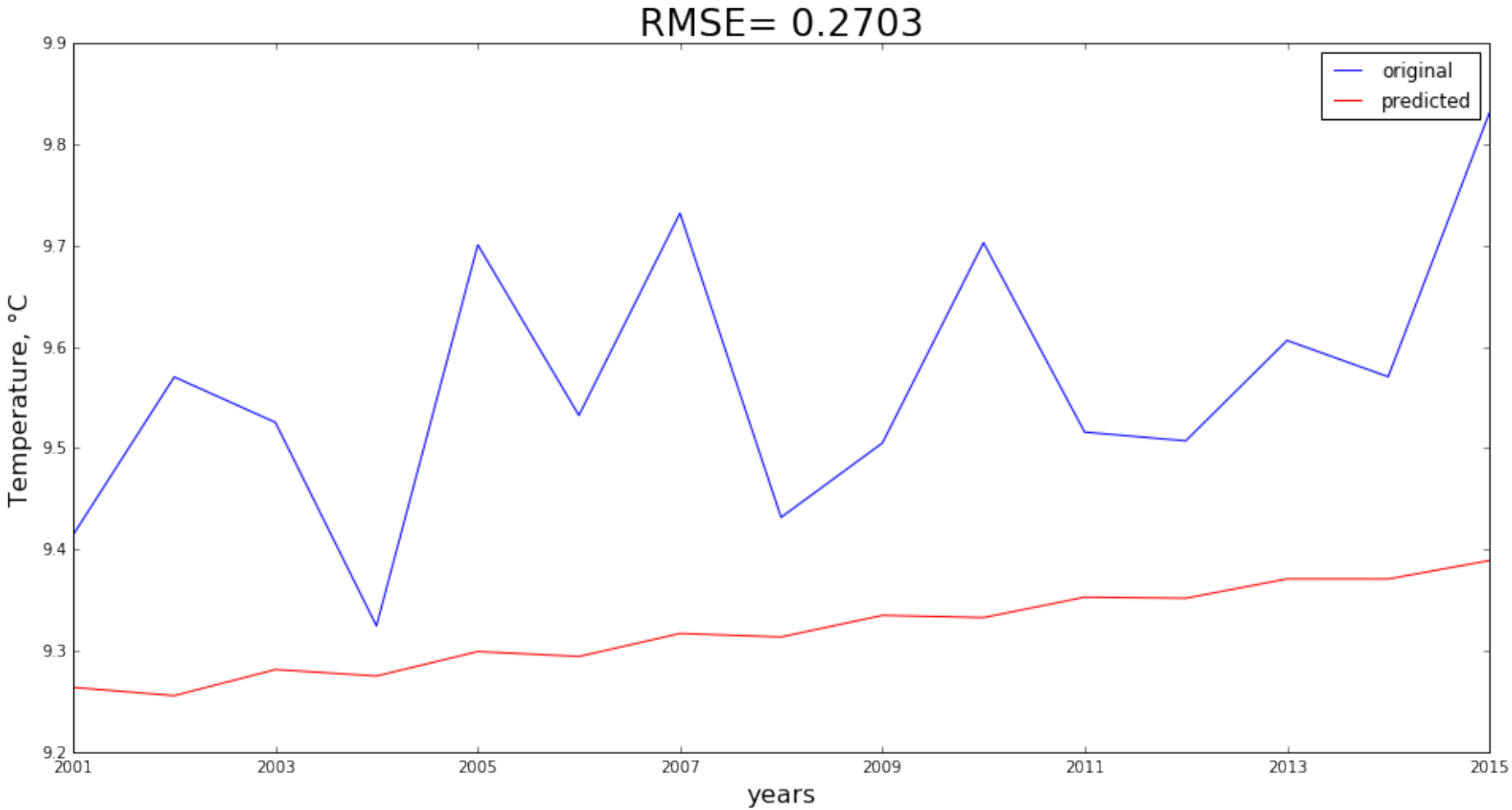
- Auto-Regressive Integrated Moving Averages (ARIMA) model
- Appropriate models when the data show evidence of non-stationarity
- $ARIMA(p, q, d)$ where:
  p is the order of the auto-regressive model
  d is the degree of differencing
  q is the order of the moving average model
- We chose $p = 1$, $d = 1$, $q = 2$

Evaluation

In-sample Performance



Out-of-sample Performance



Conclusion

Global warming is a fact.
Nations must work towards a systematic and organized fight against climate change.

References

Climate Change: Earth Surface Temperature Data. Exploring global temperatures since 1750. Available from World Wide Web: (https://www.kaggle.com/berkeleyearth/climate-change-earth-surface-temperature-data).

Thank you!