

Extrapolation of Climate Data using Numerical Methods

Philippe Nadon*

March 29, 2019

Faculty of Science, University of Alberta
Camrose, Alberta, Canada

Abstract

Climate change is currently a major concern for society, due to the irreversible damage predicted by our models if we continue our current trend of excessive emissions and pollution. In this paper various methods of data-fitting and extrapolation are explored, and compared to see how accurately they represent the data provided.

Keywords: Climate, Numerical Methods, Extrapolation

*Corresponding author: nadon@ualberta.ca

1 Introduction

Data is obtained from Berkeley Earth, the table representing the "Global Average Temperature Anomaly with Sea Ice Temperature Inferred from Air Temperatures". This data spans from 1850 to 2018, and shows us how global temperatures have changed in the last 168. Using this data, we can use the method of least squares to fit various equations onto this data, and then extrapolate the data's trend into the future. These include a linear, quadratic, and cubic fit, which are approximated by solving an $n \times n$ matrix, where n is the degree of the polynomial + 1. Afterwards, two more fits are used, an exponential fit and finally 2 different combinations of a polynomial and exponential. These last three are solved using Matlab's *lsqcurvefit* function. Although modern techniques are far more sophisticated, these results will still give us a general understanding of what will happen if we do not change our habits which impact the planet's climate.

2 Methods

Multiple equations were used to fit the data so we could compare their respective accuracies and predictions. They are listed below:

$$linear = \alpha + \beta x \tag{1}$$

$$quadratic = \alpha + \beta x + \gamma x^2 \tag{2}$$

$$cubic = \alpha + \beta x + \gamma x^2 + \delta x^3 \tag{3}$$

$$exponential = \alpha + e^{\beta x} \tag{4}$$

$$mixed_1 = \alpha + (\gamma + \delta x)e^{\beta x} \tag{5}$$

$$mixed_2 = \alpha + (\gamma + \delta x + \epsilon x^2)e^{\beta x} \tag{6}$$

2.1 Linear, Quadratic, Cubic Fits

The following code was used to calculate the best fit for the data, using equations (1), (2), and (3). The method of least squares can be solved by forming the problem into a system of linear equations in the form $Ax = B$, and solving for x .

```
function res = extrapolatePolyToYear( t, data, degree)
    N = length(data);
    rows = degree + 1;
    A = zeros( rows, rows);
    B = zeros( rows, 1);

    for n = 1:N
        for i = 1:rows
            % for A:
            % The diagonal entries:
            A(i, i) = A(i, i) + t(n) ^ (2 * ( i - 1));
            % The upper / lower entries (mirror along diagonal)
            for j = 1:(i - 1)
                A( i, j) = A(i, j) + t(n) ^ ( i + j - 2);
                A( j, i) = A(i, j);
            end

            % for B:
            B(i) = B(i) + data(n) * t(n) ^ ( i - 1);
        end
    end
    disp(A);
    disp(B);

    res = linsolve( A, B);
end
```

This function receives an array containing t values, an array of data, and the degree of polynomial you wish to fit onto the data. It then returns an array containing the coefficients of the polynomial that is to be fit onto the data.

2.2 Exponential and Mixed Fits

Equations (4), (5), and (6) were simpler to solve, but involved modifying the initial guesses until the plot fit well. The plot was done using the Matlab function *lsqcurve*. The initial guesses are as follows:

Exponential:

$$\alpha = -0.4, \beta = 0.02, \gamma = 0.1$$

Exponential and Linear:

$$\alpha = -0.5, \beta = 0.02, \gamma = -0.4, \delta = -0.002$$

Exponential and Quadratic:

$$\alpha = -0.5, \beta = 0.02, \gamma = -0.4, \delta = -0.002, \epsilon = 0.03$$

3 Results

3.1 Linear, Quadratic, and Cubic Fits

After computing the data for a linear, quadratic, and cubic fit, the following results were obtained:

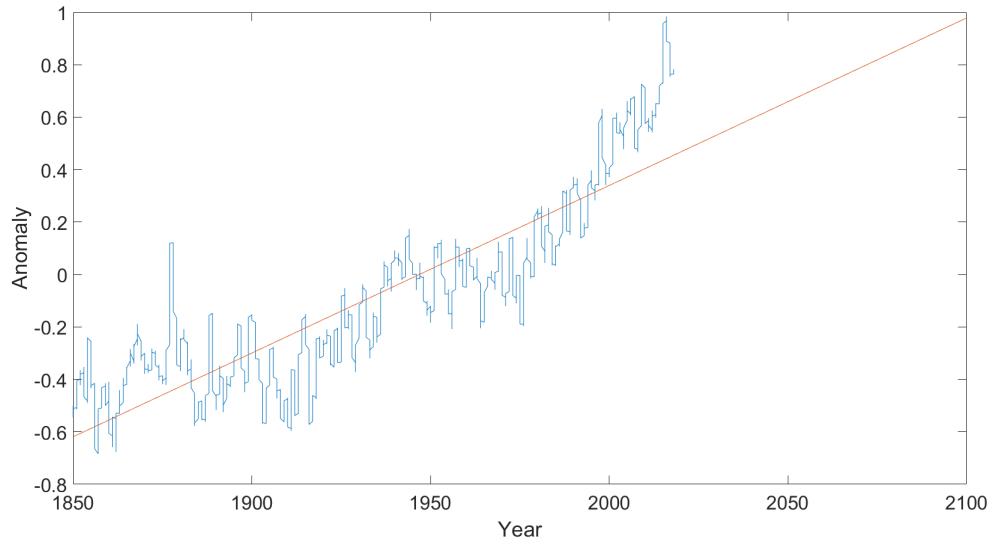


Figure 1: Linear extrapolation of climate anomalies. The linear fit predicted an anomaly value of 9.79×10^{-1} in year 2100, with a cumulative error of 5.71×10^1 calculated between the years 1850 and 2018.

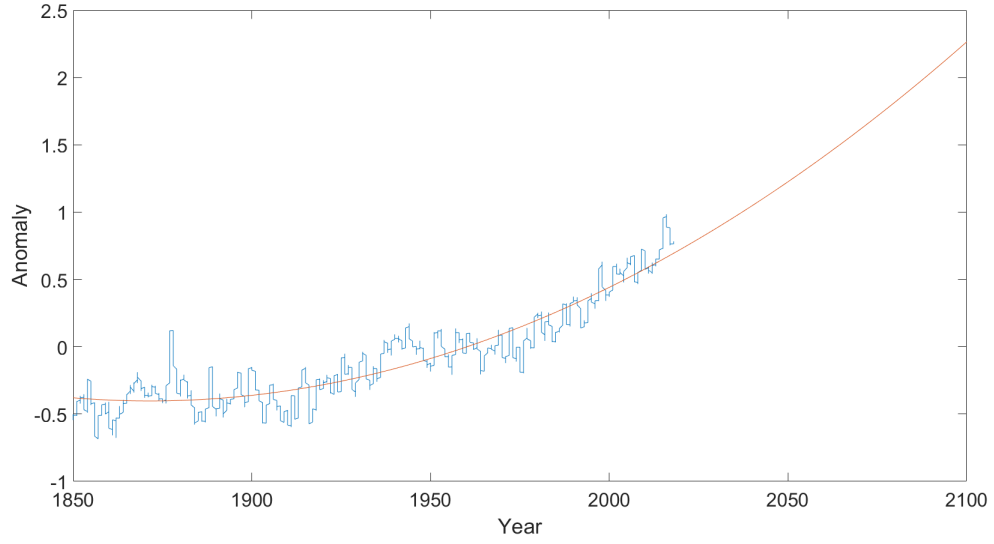


Figure 2: Quadratic extrapolation of climate anomalies. The quadratic fit predicted an anomaly value of 2.26×10^0 in year 2100, with a cumulative error of 3.38×10^1 calculated between the years 1850 and 2018.

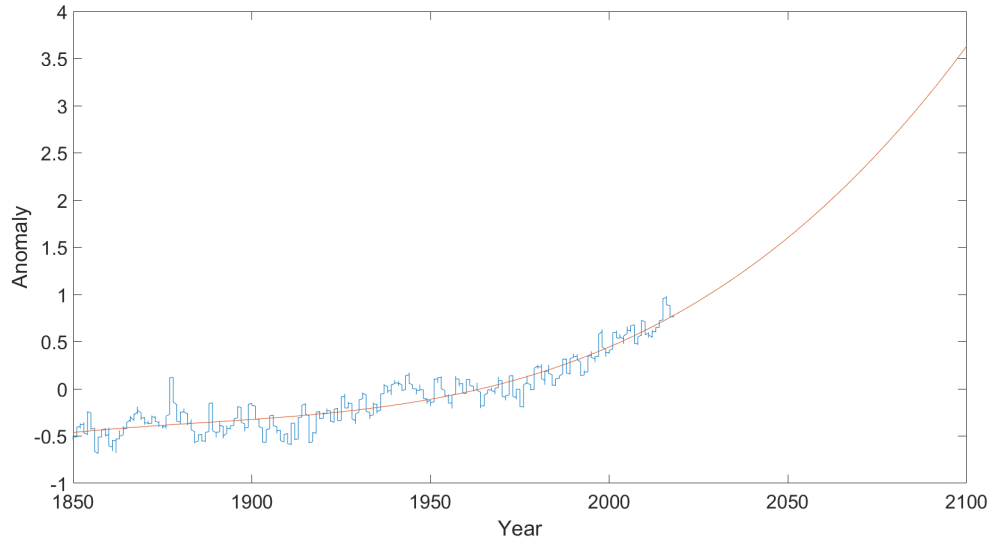


Figure 3: Cubic extrapolation of climate anomalies. The cubic fit predicted an anomaly value of 3.63×10^0 in year 2100, with a cumulative error of 3.17×10^1 calculated between the years 1850 and 2018.

3.2 Exponential and Mixed Fits

After computing the data for an exponential fit, and two different mixed fits, the following results were obtained:

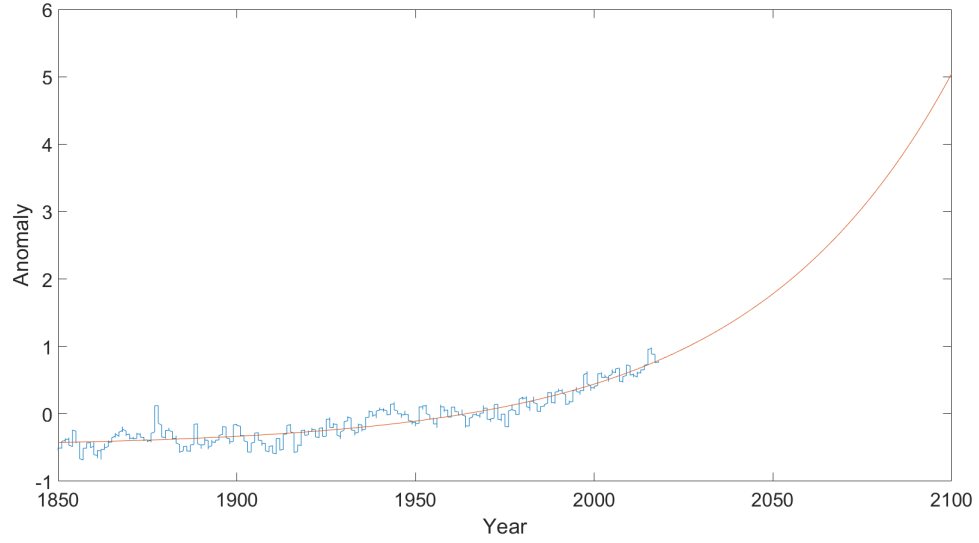


Figure 4: Exponential extrapolation of climate anomalies. The exponential fit predicted an anomaly value of $5.04 * 10^0$ in year 2100, with a cumulative error of $3.14 * 10^2$ calculated between the years 1850 and 2018.

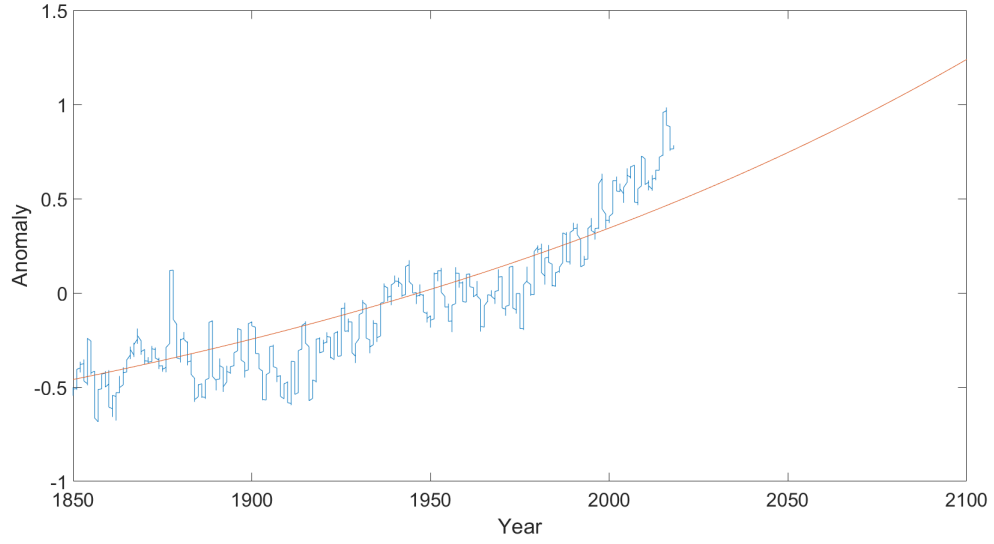


Figure 5: Mixed exponential and linear extrapolation of climate anomalies. The mixed fit predicted an anomaly value of $1.32 \cdot 10^0$ in year 2100, with a cumulative error of $4.89 \cdot 10^1$ calculated between the years 1850 and 2018.

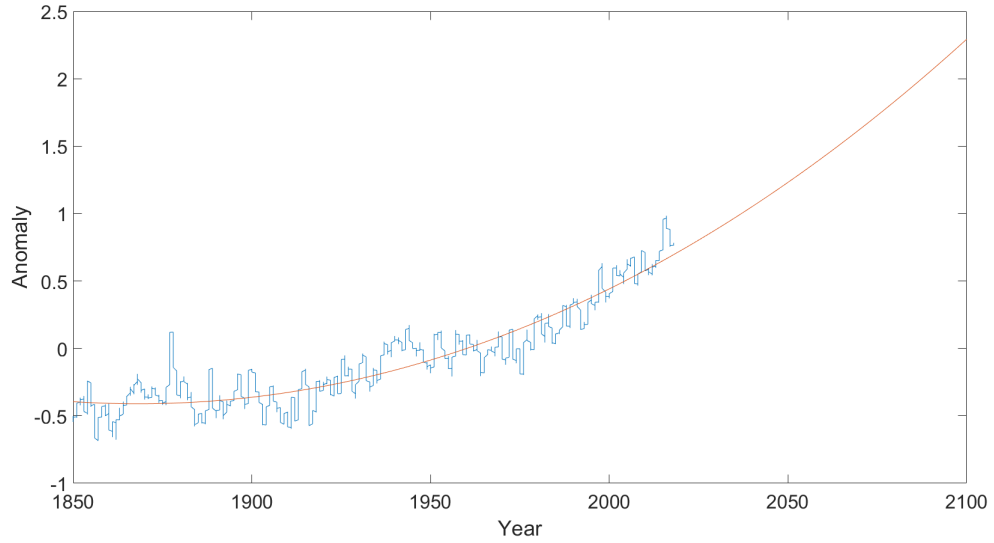


Figure 6: Mixed exponential and quadratic extrapolation of climate anomalies. The mixed fit predicted an anomaly value of $2.29 \cdot 10^0$ in year 2100, with a cumulative error of $3.37 \cdot 10^1$ calculated between the years 1850 and 2018.

4 Analysis

4.1 Accuracy of Fit

We can see that the fit becomes progressively more accurate as the degree of the polynomial increases, for both the mixed and purely polynomial fits. Initially, equation (6) returned a nonsensical plot since it predicted a sharp decline in anomalies just past the last data point, but increasing the maximum iterations for *lsqcurve*, we obtained a much more precise and logical fit. The interesting analysis we can make is that although equations (5) and (6) should have *at least* the same amount of accuracy as (4) since we can simply set the coefficients of the polynomial portion to zero, it turns out due to computational limits that more error is introduced than eliminated. Matlab itself limits the number of computations allowed before it detects that the minimum difference of results between iterations has been reached and stops computing. Although the mixed results offer better fits than their purely polynomial counter parts, the cubic and exponential fits offer the most accurate fits according to the sum of the differences between the data points and the equation.

The fit with the lowest error appears to be the exponential fit, followed by the cubic fit. It is therefore likely that these two will offer the most accurate interpretation of the data.

4.2 Interpretation of Results

Both the exponential and cubic fit predict a far more aggressive increase in the anomaly value than any other fit, which is concerning. The exponential fit predicts an anomaly of +31.4 in the year 2100, while the cubic fit predicts an anomaly of +31.7. Considering the value has remained around 0.5 to 1 in 2018, for climate anomalies to triple or even quintuple is surely not a good sign. What this value means essentially is that we should expect to see more aggressive, unpredictable weather changes in the future if our impact on the environment does not change. Tornadoes, lightning storms, flash floods, and summer-time blizzards will become more and more common, which will impose a much higher cost in infrastructure maintenance and repair, and drastically increase the rate of injuries and

deaths caused by weather.

5 Conclusion and Future Considerations

This project has helped demonstrate the utility of numerical analysis for predicting trends in areas such as climate, and also prompts the reader to consider the impact that our society has on the climate of our planet. For future improvements, one would potentially include more than just data from a single source, and improve the fit in areas such as extending computation time, using data types which can hold more precision for calculations, and using different equations altogether, such as a sin wave, as it appears from the graphs that there tends to be a trend where the anomaly value fluctuates in the shape of a wave. Reflecting on the trends at the time, such as vehicle and factory efficiency, and general consumerism would also help to improve accuracy, as those two areas are huge contributors to climate change.