
Problem Set Four

Jake Goh Si Yuan

October 16, 2016

1. FITTING TRENDS OVER THE ENTIRE DATA SERIES

1A.

The average increase of CO_2 concentration per year from the linear regression model of CO_2 concentration with time (with 1958 as origin) is 1.520, and the average increase per decade is 15.20.

1B.

Taking the time data, with origin 0 set as year 1958 and with year 2016 represented as 58, we will arrive at a nominal start-point of 307ppm and endpoint of 395.16ppm

The start-point of the data is 315.71ppm or 8.71ppm higher than trend-predicted, and the end-point is 402.25ppm or 7.19ppm higher.

1C.

The standard deviation of the data-misfit residual, or the residual standard error, is 3.871 on 700 degrees of freedom. The F variance ratio, or statistic, is 30880.

1D.

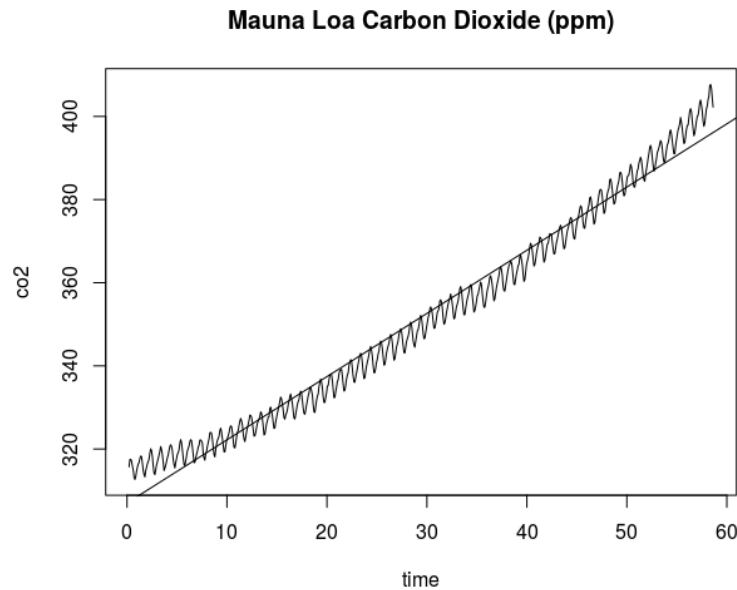


Figure 1.1: Linear regression of time(with 1958 as origin) and Mauna Loa Carbon Dioxide

The data does not seem to follow strictly to a straight line, in fact, it appears to be have an exponential trend with a small enough exponent that it takes on pseudo-linear characteristics.

Also, the linear prediction fails to take into account the saw-tooth motion that the data is taking, or the trend within the trend.

1E.

The predicted start-point for the quadratic model is 314.2ppm, as compared to the actual start-point of 315.71ppm, with a difference of -1.51ppm.

The predicted end-point is 401.89ppm, as compared to the actual end-point of 402.25ppm, with a difference of -0.36ppm.

1F.

The predicted yearly increase at the start-point of the quadratic model is 0.8055ppm, and at the end-point it is 2.235ppm.

1G.

The standard deviation of the data-misfit residual, or the residual standard error, is 2.209 on 699 degrees of freedom. The F variance ratio, or statistic, is 48160.

1H.

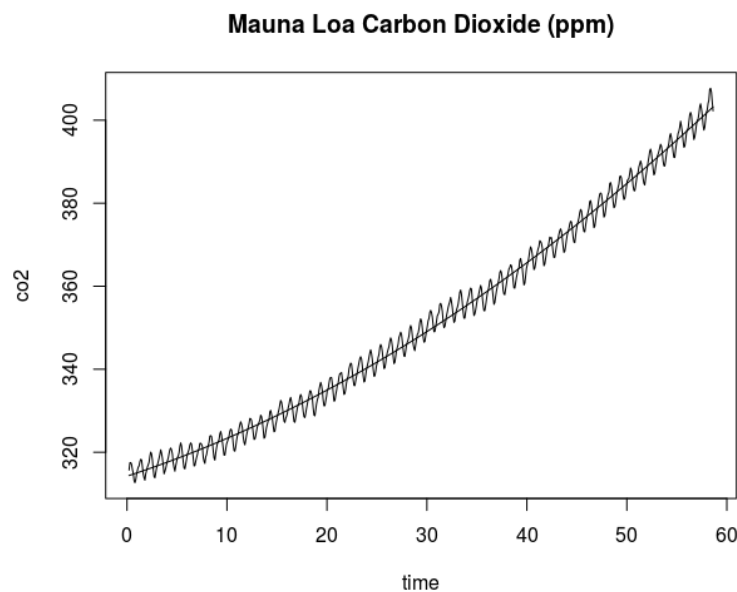


Figure 1.2: Quadratic regression of time(with 1958 as origin) and Mauna Loa Carbon Dioxide

The quadratic model offers a better fit as compared to the linear model. We can look at the comparisons between predicted start-point/endpoint and actual start-point/endpoint to see that the error has become much smaller.

However, the big inherent problem of not being able to take into account of the saw-tooth motion is still present.

Although the quadratic model is able to help us reject a linear model, we do not know if the global trend is indeed quadratic-like, or if it is only localized to the 58 years frame.

1I.

The standard deviation of the data-misfit residual, or the residual standard error, is 2.21 on 699 degrees of freedom.

The significance of the cubic parameter is 0.642, which means that it is not very significant, and the null hypothesis of not having a cubic parameter cannot be rejected.

2. FITTING TRENDS AND CYCLES OVER THE ENTIRE DATA SERIES

2A.

The amplitude of the annual cycle in CO_2 is 2.833 ± 0.06967 ppm.

2B.

The annual maximum occurs at 0.3095 of the cycle, or approximately towards the end of April. A quick sanity check at a zoomed in slice roughly confirms the assertion.

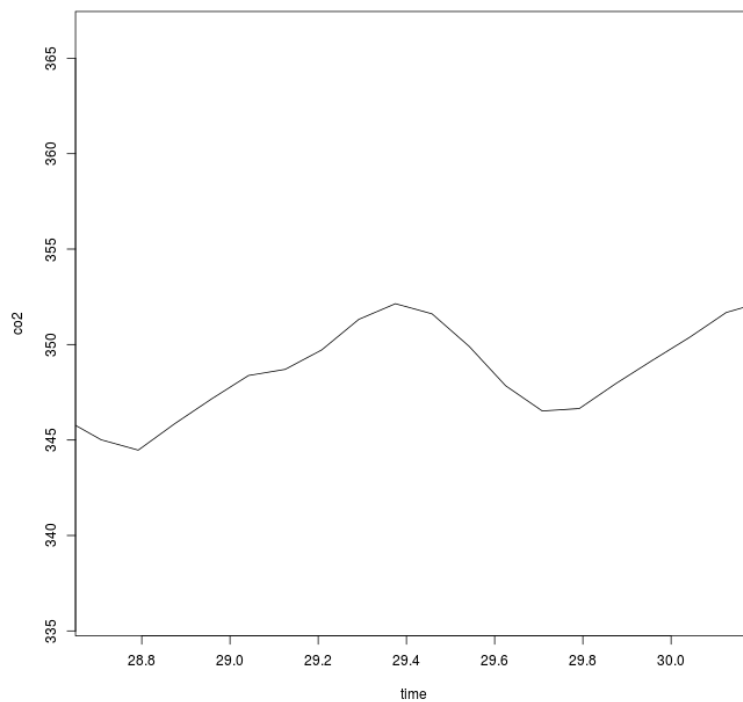


Figure 2.1: Zoomed in slice of time(with 1958 as origin) and Mauna Loa Carbon Dioxide

2C.

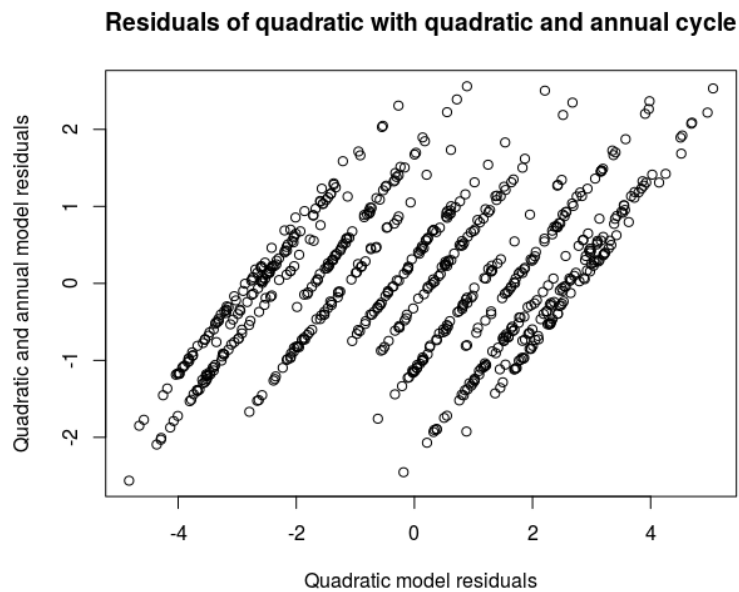


Figure 2.2: Comparing quadratic model residuals with quadratic and annual cycle residuals

The standard deviation of the quadratic model only is 2.206 whilst the standard deviation of the quadratic model and annual cycle residuals is 2.064.

2D.

The amplitude of the semi-annual cycle of CO_2 is 0.7814 ± 0.055805 ppm.

The inclusion of the semi-annual cycle gives the model a sharper characteristic in terms of the reverse in trend in the annual cycle. In other words, when maxima or minima is arrived at during the annual cycle, the first derivative changes direction much faster.

The amplitudes of maximum and minimum also increase by approximately 10%.

2E.

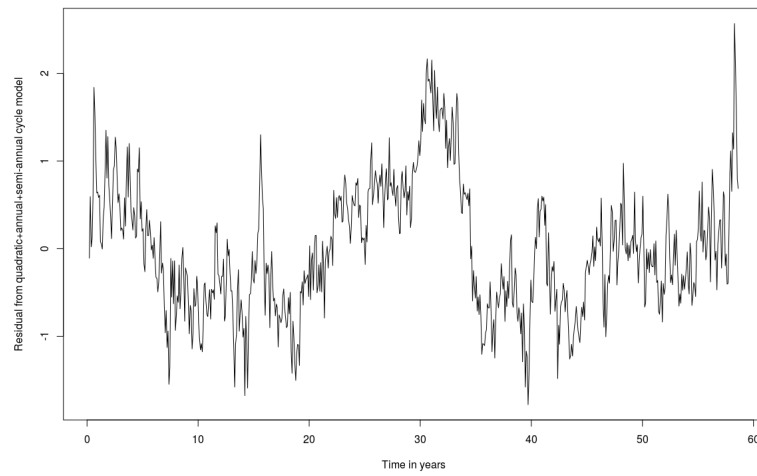


Figure 2.3: Residuals from quadratic, semi-annual and annual model with time

The standard deviation is 0.7361.

3. FITTING TRENDS AND CYCLES OVER 10-YEAR DATA SEGMENTS

3A.

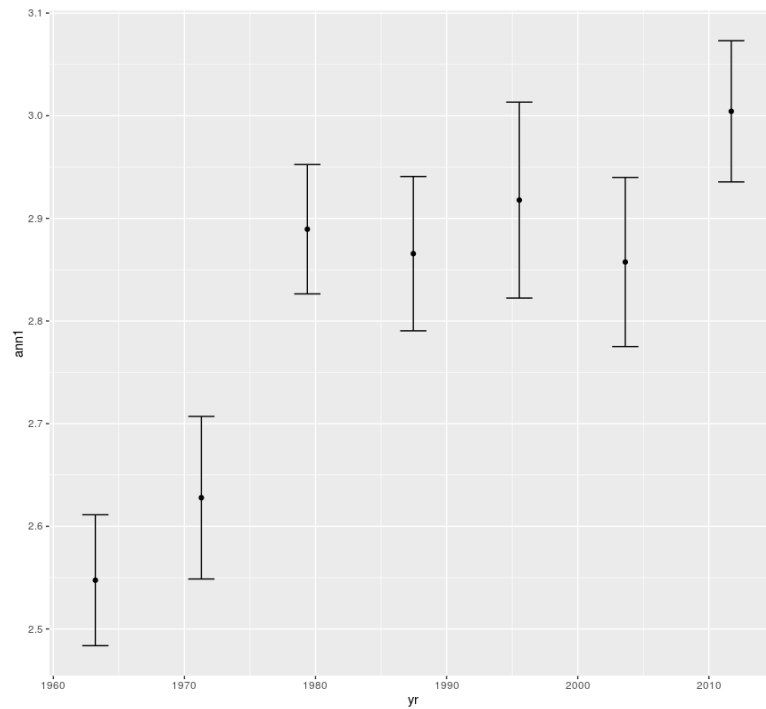


Figure 3.1: Annual-cycle amplitudes for decade-long segments of Mauna Loa CO₂ time series

There is definitely an upward trend in the annual-cycle amplitudes. However, this trend at first glance does not appear to follow a simple model (linear, quadratic or otherwise).

The trend can be best described with two characteristics:

1. A sharp transitional increase from 1970 - 1980.
2. Slow but steady increase between 1980-2010.

3B.

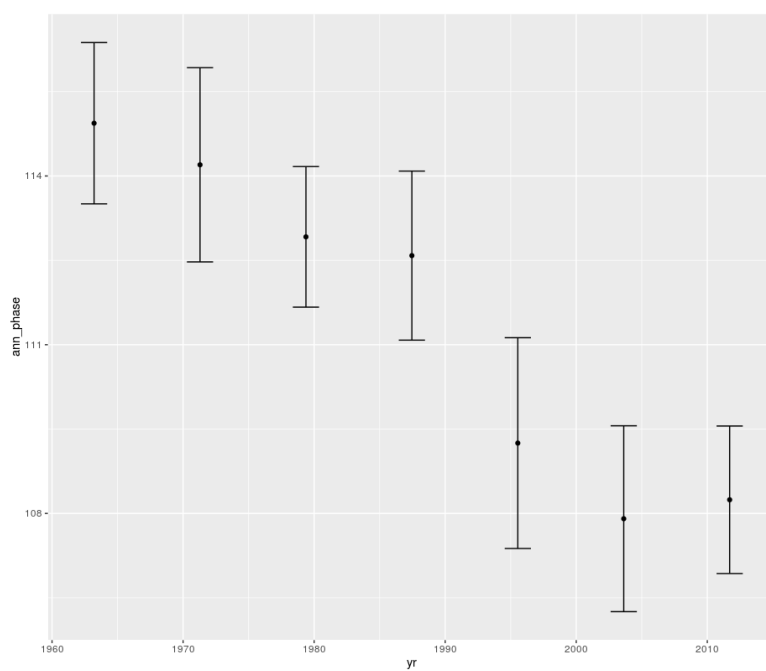


Figure 3.2: Annual-cycle phases for decade-long segments of Mauna Loa CO_2 time series

The days have shifted by approximately 6.695 ± 2.745 days.

3c.

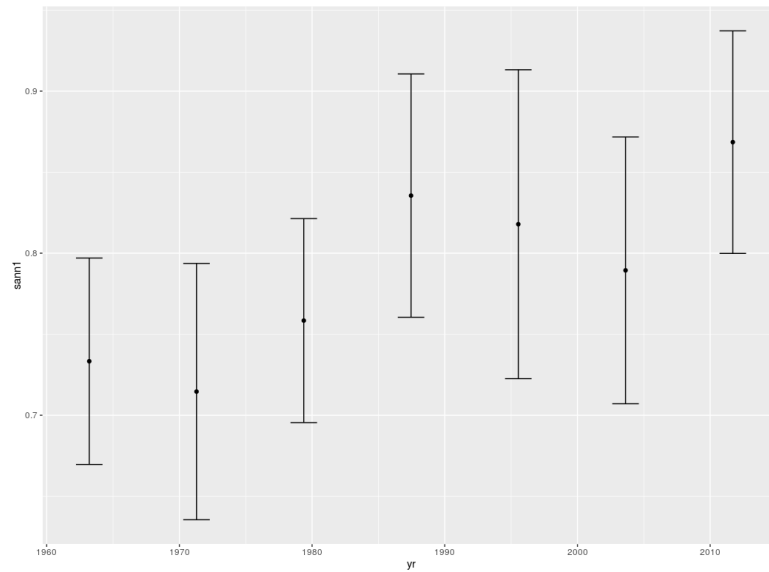


Figure 3.3: Semi-annual-cycle amplitudes for decade-long segments of Mauna Loa CO_2 time series

The semi-annual amplitudes follow a slow increasing trend from 1958 to 2010, from a starting point of 0.733 to a final maximum endpoint of 0.8685.

Interestingly, the starting point is not the minima of the graph, although that can be easily explained by the high error bars, and the negligible distance (in y and x axis) between the minimum and the starting point.

Just like in the annual trend, there is a sharper increase between 1970-1980, although in this case it goes on further towards 1990, before taking a small decline until the mid of 2000 and 2010, before reversing in trend and making a sharp increase to 2010 and onwards.

4. FITTING AN EXPONENTIAL TREND OVER THE FULL DATA SERIES

4A.

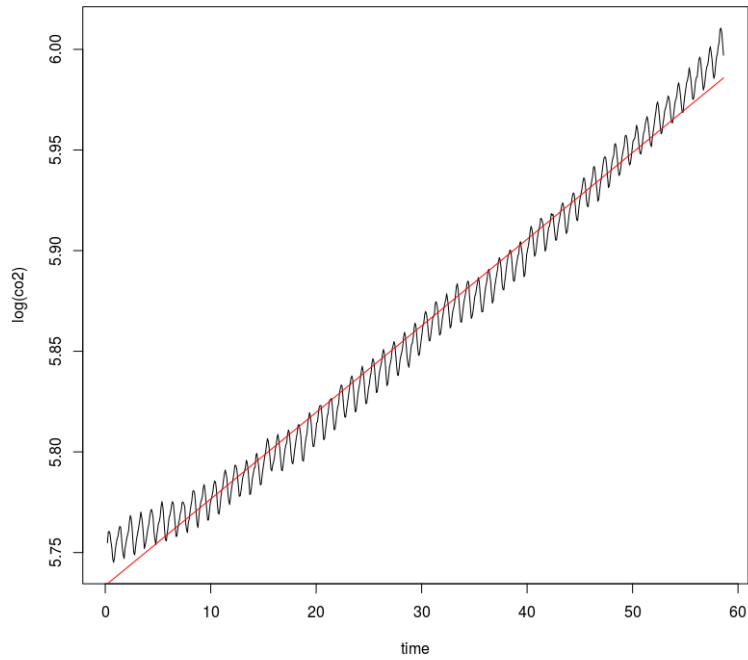


Figure 4.1: Logarithmic version of exponential model of Mauna Loa

4B.

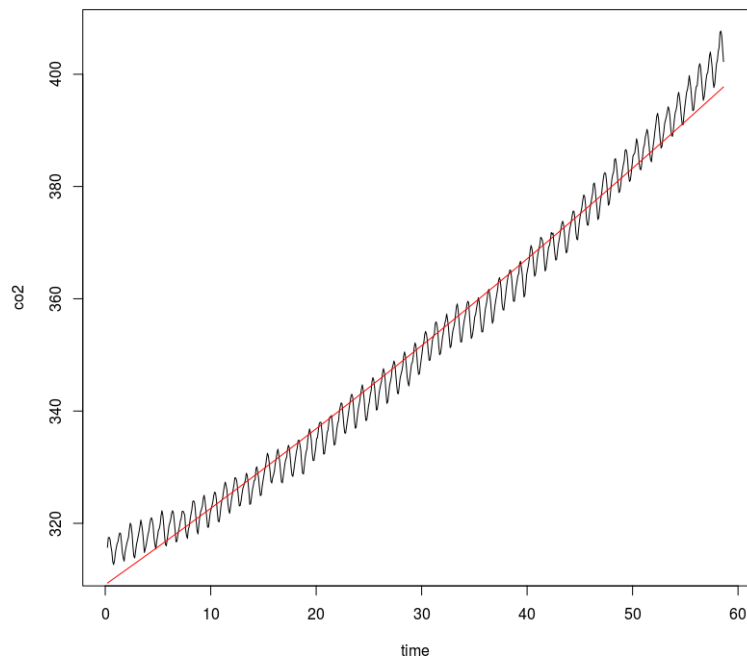


Figure 4.2: Exponential model of Mauna Loa

The value at 1958 is 309.36 and the value at 2016 is 397.74.

The data fit fails to take into account the semi-annual and annual cycle, and provides a seemingly poor prediction for the start and end-points.

For an exponential model, this appears very linear, and it is no surprise given that the method employed begins with a linear modelling through the logarithmic before converting back by exponentiating.

4C.

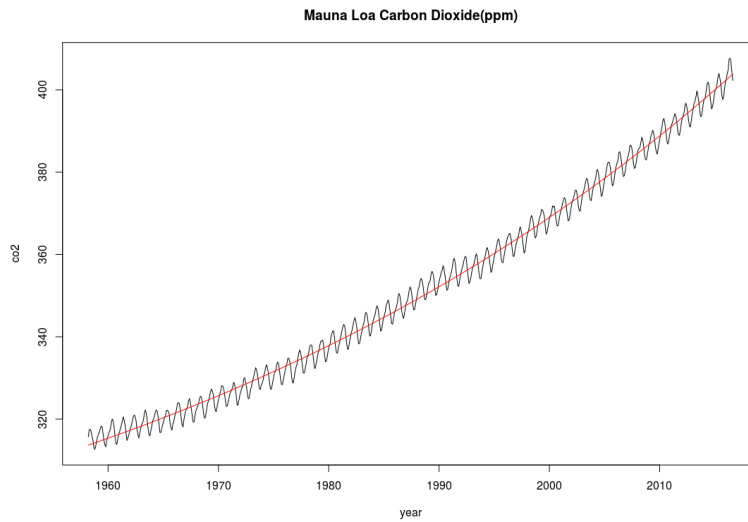


Figure 4.3: Exponential model of Mauna Loa with nls

The value for A is 256.9 ± 2.542 , the value for B is 56.61 ± 2.357 and the value for C is 0.01627 ± 0.0004390 .

The standard error of residual is 2.225 on 699 degrees of freedom.

This model bears a remarkable similarity to the quadratic-polynomial solution.

4D.

The standard deviation of 280 from A is 9.087.

We get an extrapolation of 260.46ppm.