

11-14-2017

Trend and Acceleration: A Multi-model Approach to Key West Sea Level Rise

John Tenenholtz
jtenenho@fiu.edu

DOI: 10.25148/etd.FIDC004007

Follow this and additional works at: <https://digitalcommons.fiu.edu/etd>

 Part of the [Applied Statistics Commons](#), [Climate Commons](#), and the [Oceanography Commons](#)

Recommended Citation

Tenenholtz, John, "Trend and Acceleration: A Multi-model Approach to Key West Sea Level Rise" (2017). *FIU Electronic Theses and Dissertations*. 3552.

<https://digitalcommons.fiu.edu/etd/3552>

This work is brought to you for free and open access by the University Graduate School at FIU Digital Commons. It has been accepted for inclusion in FIU Electronic Theses and Dissertations by an authorized administrator of FIU Digital Commons. For more information, please contact dcc@fiu.edu.

FLORIDA INTERNATIONAL UNIVERSITY
Miami, Florida

TREND AND ACCELERATION:
A MULTI-MODEL APPROACH TO
KEY WEST SEA LEVEL RISE

A thesis submitted in partial fulfillment of the
requirements for the degree of
MASTER OF SCIENCE
in
STATISTICS
by
John S. Tenenholtz

2017

To: Dean Michael R. Heithaus
College of Arts, Sciences and Education

This thesis, written by John S. Tenenholtz and entitled Trend and Acceleration: A Multi-model Approach to Key West Sea Level Rise having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this thesis and recommended that it be approved.

B.M. Golam Kibria

Robert Burgman

Sneh Gulati, Major Professor

Date of Defense: November 14, 2017

The thesis of John S. Tenenholtz is approved.

Dean Michael R. Heithaus
College of Arts, Sciences and Education

Andrés G. Gil
Vice President for Research and Economic Development
and Dean of the University Graduate School

Florida International University, 2017

ABSTRACT OF THE THESIS
TREND AND ACCELERATION:
A MULTI-MODEL APPROACH TO KEY WEST SEA LEVEL RISE

by

John S. Tenenholtz

Florida International University, 2017

Miami, Florida

Professor Sneh Gulati, Major Professor

Sea level rise (SLR) varies depending on location. It is therefore important to local residents, businesses and government to analyze SLR locally. Further, because of increasing ice melt and other effects of climate change, rates of SLR may change. It is therefore also important to evaluate rates of change of SLR, which we call sea level acceleration (SLA) or deceleration.

The present thesis will review the annual average sea level data compiled at the Key West tidal gauge in Key West, Florida. We use a multi-model approach that compares the results of various models on that data set. The goal is to determine if there is a consistent result that can be ascertained from the various models.

Generally, all the models reveal a clear upward trend of SLR. Further, the models provide evidence that the trend has increased over the last 8-10 years, i.e., that there is SLA.

CONTENTS

| CHAPTER | PAGE |
|---|------|
| 1 MULTI-MODEL APPROACH TO SLR AND SLA. | 1 |
| 1.1 Defining the Trend. | 1 |
| 1.2 Defining acceleration. | 2 |
| 1.3 Methodology. | 4 |
| 2 SLR IS A LOCAL PROBLEM. | 4 |
| 3 SLR AND SLA FROM RADIATIVE FORCING AND CLIMATE CHANGE. | 5 |
| 3.1 Causes of and evidence for warming. | 6 |
| 3.2 How warming causes SLR and SLA. | 6 |
| 3.3 Historical SLR: SLR during the common era. | 8 |
| 3.4 Projections of future global SLR. | 8 |
| 4 DATA USED IN THE ANALYSIS. | 10 |
| 4.1 Sea level data from PSMSL. | 11 |
| 4.2 Additional data sets used in multiple regression. | 13 |
| 5 EXPLORATORY ANALYSIS OF THE DATA. | 14 |
| 5.1 Exploring the full time series 1913-2016. | 14 |
| 5.2 The short series 1959-2016 (used for multiple regression analysis.) | 16 |
| 6 MULTIPLE REGRESSION ANALYSIS. | 18 |
| 6.1 Reduced model from step regression and subset regression. | 20 |
| 6.2 Ridge regression. | 20 |
| 6.3 Principal component regression. | 21 |

| | | |
|-----|--|----|
| 6.4 | Comparison of multiple regression models. | 23 |
| 7 | TIME SERIES ANALYSIS. | 24 |
| 7.1 | Cyclic behavior. | 24 |
| 7.2 | Upward trend. | 25 |
| 7.3 | ARIMA analysis for a stochastic trend. | 26 |
| 7.4 | Forecast errors for ARIMA models. | 30 |
| 7.5 | Dynamic regression analysis for a deterministic trend. | 30 |
| 7.6 | Comparison of ARIMA and dynamic regression models. | 31 |
| 8 | REGRESSION MODELS (LINEAR, QUADRATIC, CUBIC). | 32 |
| 9 | NONPARAMETRIC ANALYSIS: LOESS SMOOTH. | 35 |
| 10 | CHANGEPOINT ANALYSIS. | 37 |
| 11 | DISCUSSION. | 40 |
| | REFERENCES | 43 |

LIST OF TABLES

| TABLE | | PAGE |
|-------|---|------|
| 1 | Reconstructed historical sea level - 0-2000 C.E. | 8 |
| 2 | Correlation Matrix - Multiple Regression | 17 |
| 3 | Multiple Regression Forecast errors: Four models | 23 |
| 4 | ARIMA Model Comparison (models with drift) | 28 |
| 5 | ARIMA Model Comparison (models without drift) | 28 |
| 6 | ARIMA Model Coefficients | 29 |
| 7 | Forecast errors - ARIMA models | 30 |
| 8 | Forecast and model errors Best ARIMA v. Dynamic Linear Regression | 31 |
| 9 | Comparison of OLS models on the training set | 33 |
| 10 | Comparison of OLS models on the full set | 34 |

LIST OF FIGURES

| FIGURE | PAGE |
|--|------|
| 1 SLR Projected 2100-AR5 v. Deconto/Pollard (revised Antarctic melt) | 10 |
| 2 Spectral Periodogram for Key West monthly data | 13 |
| 3 Key West yearly tidal gauge data (1913-present) | 14 |
| 4 Boxplots Key West Sea Level Eight year segments | 15 |
| 5 Boxplots Key West Sea Level Thirteen year segments | 15 |
| 6 Key West tidal gauge data Yearly (1959-present) | 16 |
| 7 Mauna Loa CO2 data Yearly (1959-present) | 16 |
| 8 Ocean Heat Content Yearly (1959-present) | 17 |
| 9 HADCRUT4 Yearly Global Mean Surface Temperature (1959-present) | 17 |
| 10 Scatterplot matrix - Sea level, OHC, CO2, Time, GMST | 18 |
| 11 Plot of coefficients vs. lambda values | 21 |
| 12 GCV of Ridge Regression | 21 |
| 13 Forecasts and prediction intervals for 4 multiple regression models . . | 24 |
| 14 Spectral Periodogram Key west Annual | 25 |
| 15 Key West Data - Differenced | 27 |
| 16 Forecasts: Best three ARIMA models plus Dynamic Regression . . . | 32 |
| 17 Linear, Quadratic and Cubic Regression Through 2002 | 34 |
| 18 Linear, Quadratic and Cubic Regression Through 2016 | 35 |
| 19 Key West annual data with Loess smooth (1913-present) | 36 |
| 20 Slope of Loess Smooth (velocity) SLR Key West - (1913-2016) | 37 |
| 21 Acceleration of SLR Key West - (1913-2016) | 37 |
| 22 Chow test results - Key West Data | 39 |
| 23 Changepoint analysis - R package “Segmented” | 40 |

| | | |
|----|--|----|
| 24 | Changepoint analysis - R package “Changepoint” | 40 |
|----|--|----|

ABBREVIATIONS AND ACRONYMS

- ACF - Autocorrelation function
- ADF - Augmented Dickey-Fuller
- AIC - Akaike information criteria
- AICc - Corrected Akaike information criteria
- AR - Autoregressive
- AR5 - Assessment Report 5
- ARIMA - Autoregressive integrated moving average
- ARMA - Autoregressive moving average
- BIC - Bayesian information criteria
- GCM - General Circulation Model (a.k.a., General Climate Model.)
- GCN - Global Core Network
- GHG - Greenhouse gases
- GIA - Glacial isostatic adjustment
- GLOSS - Global Sea Level Observing system
- GMSL - Global mean sea level
- GMST - Global mean surface temperature
- IPCC - International panel on climate change
- MA - Moving average
- MAD - Mean absolute deviation
- NOAA - National Oceanic and Atmospheric Administration
- OHC - Ocean heat content
- PACF - Partial autocorrelation function
- PC - Principal component
- PSMSL - Permanent Service for Mean Sea Level
- RCP - Representative concentration pathway

RLR - Revised local reference
RMSE - Root mean squared error
SLA - Sea level acceleration/deceleration
SLR - Sea level rise
SS - Sum of squares
VIF - Variance inflation factor
WMO - World Meteorological Organization

1 MULTI-MODEL APPROACH TO SLR AND SLA.

There are many methods for extracting an SLR trend and determining if SLA exists. There is, however, no clear “best method.” Each has its own advantages and disadvantages, as well as underlying assumptions. Each uses the information from the sample differently and each conveys different information about the sample. The present thesis will therefore use a multi-model approach to determine if SLR and SLA exist. It will use a variety of different types of models, each with its own set of advantages and disadvantages, and will examine their results to determine whether there is a consistent finding across the models.

1.1 Defining the Trend.

We are concerned with sea level trends and acceleration. Acceleration is the derivative of the trend curve. Therefore, the first goal is to find that trend curve. A general definition of a trend is a “long-term temporal variation in the statistical properties of a process, where ‘long-term’ depends on the application.” (Chandler & Scott, 2011, at p.5) So, for example, if a process can be described as having a time-dependent mean μ_t for $t = 1, 2, \dots, T$ where μ_t is changing over time, the process can be said to have a trend. As a practical matter, most statistical methods for extracting a climatological trend attempt to find a smooth trend which is meant to represent climate, while the residuals represent natural variability. (Visser et al., 2015) We will therefore use the following practical definition of trend: “a smooth signal where the smoothness is chosen to filter out shorter-term (decadal) natural fluctuations.” (Visser et al., 2015, at p.3874). There are numerous mathematical/numerical methods for extracting such a smooth signal. In addition, there are other methods that rely on visual interpretation to infer that a trend exists. We can split the methods into the following five broad categories:

1. Exploratory analysis. This includes visual inspection, simple statistical analysis such as boxplots and expert judgment.
2. Parametric trend estimation. Parametric models precisely define the trend as a function of time and other variables. These relationships are often unchanging and are good for describing **deterministic** processes. They are problematic when the relationship to time or other variables is changing during the relevant time period.
3. Nonparametric trend estimation. Unlike parametric models, nonparametric models are not expressed as a function of time (or other predictors). Instead, they are derived from the data values, typically data values that are close in time. These models include various smoothing techniques, such as moving averages, loess and exponential smoothing.
4. Stochastic (probabilistic) trend models. Stochastic trends are presumed to originate from a noise process. These are not deterministic in nature, though they can represent the residual process after a deterministic process has been removed.
5. Miscellaneous models. These can involve a mixture of the above methods, such as a parametric trend estimation on shifting windows, or can be altogether unrelated to the above methods.

(Chandler & Scott 2011), (Visser et al. 2015).

1.2 Defining acceleration.

The instantaneous slope of the trend line at a given time is the rate of change of sea level at that time. When that slope is increasing, there is sea level acceleration

(SLA). When it is decreasing, there is deceleration. It should be noted that SLA is not necessary for continued SLR. Even if the acceleration is zero or negative, sea level can still be rising, with damaging consequences to coastal areas.

Sometimes, finding the instantaneous slope of the trend line can be simple. For example, for a linear trend derived by simple linear regression, the slope is the coefficient of the linear term. Finding the slope can also be more difficult, such as in the case of a smooth, but irregular, trend derived via a smoothing technique. When the rate of change of the slope is zero, or is otherwise constant, it is easy to predict future sea levels. When the rate of change is increasing (i.e., if there is SLA), prediction becomes more difficult - and the potential adverse consequences of the sea level become greater. Several methods have been used to describe SLA, including:

1. The second derivative of the trend line derived by least squares regression. In the case of a simple linear regression, the second derivative (and therefore the SLA) is zero. If the regression contains a quadratic term, the SLA will be constant. If it contains a cubic term, the SLA will be linearly time dependent - and so on.
2. Comparisons of the slope of regressions taken over different intervals. If the slope of later intervals becomes monotonically greater than that of earlier ones, this is evidence of SLA.
3. Trend concavity. If the trend function is locally “concave up,” it can be interpreted as having a positive acceleration. By concave up, we mean that if a line segment is drawn through two points on the curve, the points of the curve lie below that line segment. Whether a curve is concave up is easily determined in the case of a polynomial trend (i.e., $f''(x) > 0$). It is harder to determine in the case of less regular trends, such as those derived by smoothing techniques.

1.3 Methodology.

Most papers that deal with trends or acceleration in sea level pick a model and use it to determine the extent of SLR or SLA. This thesis will instead use a multi-model approach that compares the results of various models on a single data set. The goal will be to determine whether there is a consistent result that can be ascertained from the various models.

We will loosely follow the “good modeling practices” set forth in Visser et al. (2015). In particular, we will:

- i) Use a multimodel approach with a discussion of the sensitivity of the models to variation of the parameters;
- ii) Where possible, make determinations whether parameter estimates are statistically significant.
- iii) Where possible, provide uncertainty information for μ_t .
- iv) Where possible, provide uncertainty information for extrapolations. (Long term predictions should not be based on extrapolation.)
- v) Where possible, use models that incorporate additional information into the analysis (e.g., multiple regression).¹

2 SLR IS A LOCAL PROBLEM.

Because there are multiple factors that contribute to sea level, and the contributions of these factors vary with location, rates of SLR also vary with location. The

¹The Visser paper suggests an additional modeling practice, i.e., that extrapolation should be based the application of emissions scenarios and general circulation models, incorporating physical mechanisms. Because this thesis is purely a historical statistical analysis and is not based on the physics associated with the various models and emissions scenarios, this practice is beyond its scope.

thesis will explore techniques that can be applied on a local basis and will apply those techniques to the measurements of annual mean sea level derived from the tidal gauge data set from Key West, Florida.

While SLR is often studied on a global basis, regional and local applications are often more important because the effects of SLR on a particular location depend directly on SLR at that location, not on a global measure. This is critical because much of the world's population lives on coastal plains abutting rising bodies of water. An understanding of the local rates of SLR and SLA is therefore critical for the prediction and mitigation of SLR and its associated risks.

3 SLR AND SLA FROM RADIATIVE FORCING AND CLIMATE CHANGE.

Carbon dioxide and other greenhouse gases (GHGs) released into the atmosphere from the burning of fossil fuels and other human activities are causing the atmosphere and the seas to warm. (See, e.g., Hartmann, 2016, Ch. 13) The warming creates a combination of effects that result in SLR. Among these are (i) thermal expansion of sea water, (ii) melting of land-based glaciers and ice-sheets, and (iii) alteration of ocean currents.

This section will discuss the following background issues regarding climate change induced SLR:

1. Causes of and evidence for warming.
2. How warming causes SLR and SLA.
3. Historical SLR (during the common era.)
4. Projections of future global SLR.

3.1 Causes of and evidence for warming.

The GHG induced greenhouse effect currently causes an imbalance of approximately 3.027 Wm^{-2} between the energy coming from the Sun and that returning to space. (Butler & Montzka, 2016) This energy imbalance is referred to as a radiative forcing. The additional energy heats the atmosphere and the oceans. Multiple lines of evidence confirm that the earth is indeed warming. For example, warming is evidenced by increases in surface temperatures, increases in extreme high temperatures and decreases in extreme lows, increases in ocean heat content, increases in specific humidity, decreases in glacier and ice sheet mass, decreases in arctic sea ice extent and thickness, decreases in winter snow cover, changes in air and water circulation features, changes in ocean salinity and ocean acidification. (IPCC AR5, WG I Technical Summary.)

3.2 How warming causes SLR and SLA.

The following eight factors, all of which can be altered by global warming, have been shown to affect sea level at a particular location:

1. Astronomical tides, which are caused by the gravity of the sun and the moon.
2. Vertical land movement from glacial isostatic adjustment (GIA) or other causes. GIA results from the massive ice sheets that covered the upper latitudes during the last ice age. This mass compressed the land under it, which began to rise when the ice receded. The land is still rising, even 10,000 years after the end of the ice age. On the other hand, land around the edges of the ice sheet bulged up during the ice age (the forebulge) and is still settling (forebulge collapse). These adjustments are collectively known as GIA. Other vertical land movements can be caused by seismic activity, and by the removal of ground

water or hydrocarbons. GIA is not a major issue in southern latitudes such as Florida.²

3. Onshore-offshore wind component, i.e., an increase in sea level ahead of winds as they head onshore or offshore.
4. Longshore wind component, i.e., change in sea level due to winds running parallel to the coast.
5. Thermal expansion or contraction of ocean water due to temperature changes. Increases in ocean temperature will cause ocean waters to expand, thus increasing sea level.
6. Changes in ocean currents. Intense ocean currents (such as the Florida current/Gulf Stream) decrease coastal sea level, more so the closer the currents are to shore. For example, the Florida current sustains a sea level difference of more than one meter between Florida and the Bahamas. (Domingues et al. 2017). It is anticipated that warming will cause a substantial decrease in the strength of the Florida current. (Park & Sweet 2015).
7. River discharge can raise sea level near river deltas.
8. Melting of land based ice, which includes glaciers and ice sheets such as Greenland and Antarctica increases total ocean volume, thereby increasing sea level.

(Parker 1992) At least three of these eight factors will likely be critical to SLR and SLA in Florida: (i) thermal expansion from temperature increases, which will occur globally (ii) increased water volume from melting land-based ice, which will also be a

²There can be vertical land movement due to other factors such as wetland loss, groundwater depletion, etc.

global effect and (iii) changes in the Gulf Stream and the Florida current, which will be local to Florida. (Parker, 1992)(Kopp et al., 2016, p. 4).

3.3 Historical SLR: SLR during the common era.

Historical sea levels have been reconstructed by researchers using a variety of data. For example, Kopp, et al. (2016) analyzed global SLR using satellite, tide gauge and proxy data and found the following rates of SLR during the past two thousand years:

Table 1: Reconstructed historical sea level - 0-2000 C.E.

| Period (years C.E.) | Rate of SLR (cm/yr) |
|-----------------------|---------------------|
| 0 – 700 | +0.01 ± 0.01 |
| 700 – 1000 | stable |
| 1000 – 1400 | -0.02 ± 0.02 |
| 1400 – 1600 | +0.03 ± 0.04 |
| 1600 – 1800 | -0.03 ± 0.03 |
| 1860 – 1900 | +0.04 ± 0.05 |
| 1900 – <i>present</i> | +0.14 ± 0.02 |

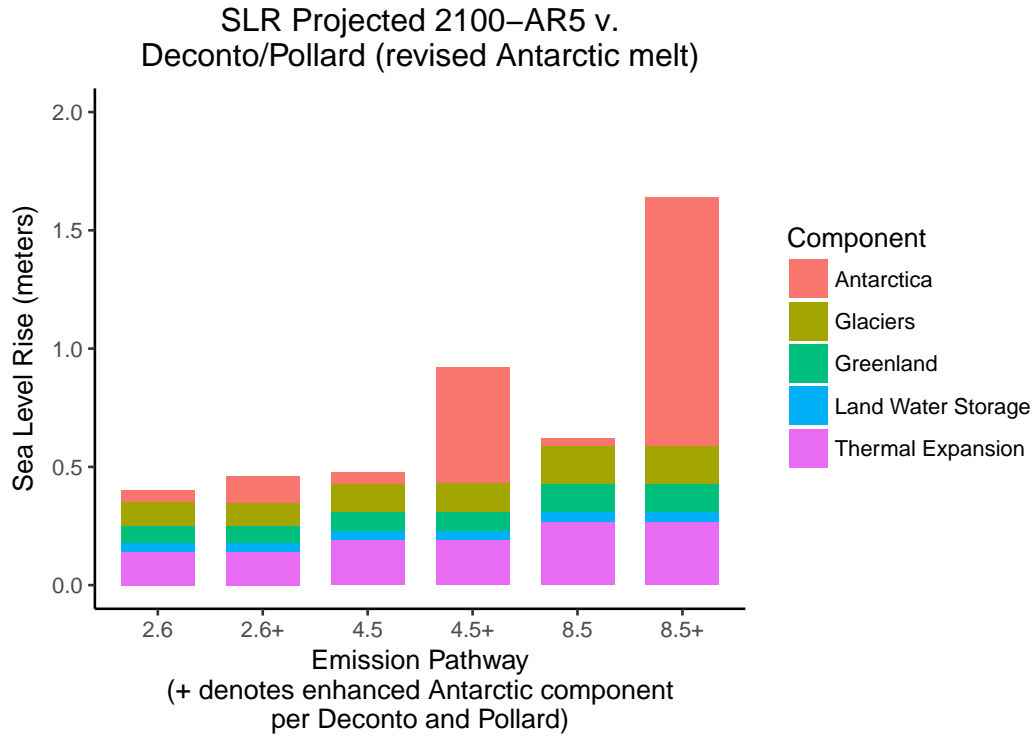
We can see from these data that sea level was fairly stable up until the start of the twentieth century and increased substantially thereafter.

3.4 Projections of future global SLR.

Even if we are able to ascertain historical sea levels, it is important for purposes of planning and remediation to be able to make reliable projections of future SLR. As discussed above, long term projections should be determined by the physics of SLR, not statistical extrapolation. The Intergovernmental Panel on Climate Change (IPCC) periodically assesses the state of climate research, which includes forecasting climate variables. The IPCC’s most recent Assessment Report 5 (AR5) uses several “representative concentration pathways” (RCPs) to denote possible GHG emission patterns going forward. The numerical portion of the RCP denotes the anticipated

radiative forcing (in Wm^{-2}) under the RCP as of 2100. For example, RCP 8.5 means that the anticipated radiative forcing at 2100 will be $8.5Wm^{-2}$. The RCP 8.5 is often referred to as the “business as usual” RCP because GHG emissions continue to increase unabated in accordance with current trends, resulting in substantially increased forcing. AR5’s sea level forecasts do not include large contributions from the melting of the Antarctic ice sheet. In doing so, it notes that “[o]nly the collapse of the marine-based sectors of the Antarctic ice sheet, if initiated, could cause GMSL [global mean sea level] to rise substantially above the likely range during the 21st century.” Many researchers currently believe that the AR5 estimates with respect to Antarctic ice melt were too conservative. For example, DeConto & Pollard (2016) suggests mechanisms for accelerated ice melt in Antarctica based on the interaction of sea water with the Antarctic ice sheets. These estimates are taken from sea levels from the Pliocene period (about 3 million years ago) which is the last time that GHG levels were as high as today, and gives several estimates of the increase in Antarctic ice melt depending on the estimate of SLR during the Pliocene period (DeConto & Pollard (2016), Fig. 5b,d). In the figure below, we compare the AR5 projections for each RCP with the DeConto/Pollard projections (based on Pliocene SLR of 10-20m). Note that the columns of the figure without a “+” show the SLR projections under AR5 and those with a “+” show the DeConto/Pollard projections. We see that the AR5 results vary from about 0.4 to 0.6 meters of SLR by 2100. DeConto and Pollard’s estimates range from about 0.6 to 1.6 meters.

Fig. 1



Other researchers have made even more extreme projections of SLR. Hansen et al. (2016) also postulates increased ice melt and posits possible doubling of Antarctic and Greenland ice mass loss in 5, 10 or 20 year increments. The doubling behavior results in exponential SLR with 5 meter SLR by approximately 2060, 2090 or 2160 for the three doubling intervals. The results under the Hansen paper are frankly catastrophic, especially in the case of 5 year doubling.

4 DATA USED IN THE ANALYSIS.

The current thesis uses annual mean sea level data from the Key West sea level station, which is part of a Global Core Network (GCN) of 290 sea level stations used for long term climate change and oceanographic sea level monitoring. The GCN is in turn part of the Global Sea Level Observing system (GLOSS), an international program

conducted under the auspices of the World Meteorological Organization (WMO). The GCN is “designed to provide an approximately evenly-distributed sampling of global coastal sea level variations.” (GLOSS 2014).

4.1 Sea level data from PSMSL.

The Key West sea level station came into service in 1913 and has been more or less reliably operational since then. It is located at $24.555^{\circ}N$, $278.193^{\circ}E$. The land on which it is situated has a vertical velocity (i.e., GIA) of -0.75 mm/year as of 2013. (Peltier & Drummond 2012). It uses an acoustic tidal gauge with a pressure gauge backup. (NOAA 2013, at p.57).³ Measurements are taken at 6 minute intervals. (GLOSS 2014). The data are maintained by NOAA and are also transmitted to the Permanent Service for Mean Sea Level (PSMSL) in the United Kingdom. The “PSMSL is the global data bank for long term sea level change information from tide gauges and bottom pressure recorders.” (PSMSL Homepage, 2017) The PSMSL creates the monthly and annual data sets from the data transmitted to it by NOAA. NOAA then incorporates these data sets into its own. (NCEI, 2015).

In order to reduce the data from all stations worldwide to a common data set, the PSMSL converts the data to its Revised Local Reference (RLR) system, which it describes as follows:

The RLR datum at each station is defined to be approximately 7000mm below mean sea level, with this arbitrary choice made many years ago in order to avoid negative numbers in the resulting RLR monthly and annual mean values. The detailed relationships at each site between RLR datum,

³Tidal gauges are generally located in stilling chambers that isolate the gauge from wave activity. Acoustic tidal gauges “send an audio signal down a 1/2 inch wide ‘sounding tube’ and measure the time it takes for the reflected signal to travel back from the water’s surface.” (National Ocean Service, What is a tide gauge?)

benchmark heights, tide gauge zero etc. are not normally required by analysts of the data set, but is [sic] available for most station [sic] from the individual station pages.

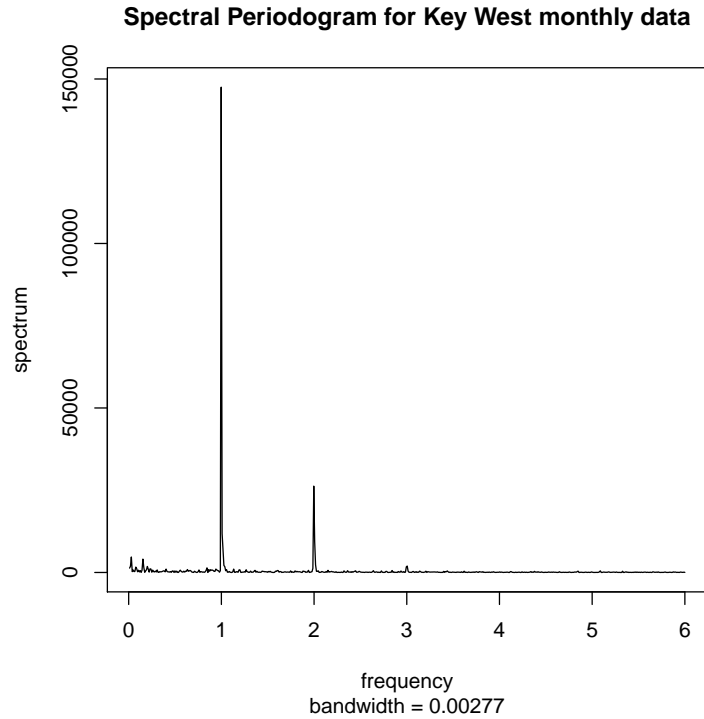
(PSMSL RLR Definition). There is no absolute zero point with respect to tidal gauge stations and sea levels are not coordinated between stations. Tidal gauge measurements are sometimes referred to as measurements of relative sea level (Baart et al. 2012, at 311) as opposed to absolute sea level, which is measured by satellites. Satellite measurements are made with respect to a geoid that is superimposed on the globe.⁴

We will use annual data herein, mainly because it eliminates a strong annual cycle which is irrelevant to the long term behavior at issue for SLR and SLA.⁵ Below, we see a spectral periodogram for the monthly Key West data. It shows a clear annual cycle, with periods of one year and 1/2 year (and perhaps further periods of 1/n). We could attempt to model this trend with a second or higher order Fourier series, but since annual patterns are not our concern in this thesis, it is easier to eliminate the annual cycle by using annual data. For a discussion of this issue, see (Foster & Brown 2015).

⁴This thesis will not use satellite data because (i) it is a short data set, only 24 years long, and (ii) while it provides resolution to a regional level, it does not resolve down to a local level.

⁵While this thesis is concerned with mean trends, the nature of the annual cycle is still critical for determining susceptibility to damage, since it is positive variations from the mean that will cause the worst damage.

Fig. 2



There is one missing data point in our data set (at 1953), which we fill using interpolation. For purposes of the time series analysis and the changepoint analysis, we use a 90 year training set, and a 14 year test set.

4.2 Additional data sets used in multiple regression.

For multiple regression, we added the following three data sets to the analysis: (i) Annual mean CO₂ concentration at Mauna Loa, Hawaii (available for 1959-present), (ii) Global Ocean Heat Content (OHC) for 0-700 meters (available for 1955-present), and (iii) the HadCRUT4 global mean surface temperature data (GMST) (available for 1850-present). The HadCRUT4 data are from the Hadley Centre in the United Kingdom, and the CO₂ and OHC data are both from NOAA. Because the Mauna Loa CO₂ series is the shortest series, starting in 1959, we limit the multiple regression analysis to the period 1959-2016, with a 50 year training set and an 8 year test set.

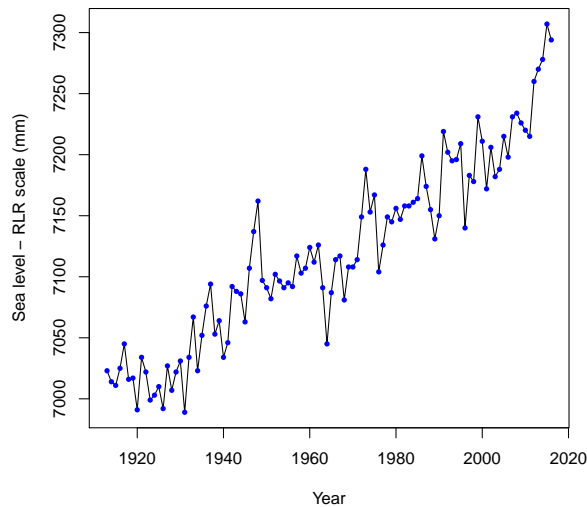
5 EXPLORATORY ANALYSIS OF THE DATA.

5.1 Exploring the full time series 1913-2016.

The full Key West time series from 1913 to 2016 exhibits an apparently linear upward trend with some periods of more extreme variation. In addition, there appears to be an abrupt upward movement starting around 2010. There is no indication that this abrupt movement has ceased or reversed. There were other jumps in the series before, but they tended to be of short duration - no more than 3 years. The current jump is 5 years long so far.

Fig. 3

**Key West yearly tidal gauge data
(1913–present)**



Another way to visualize the trend is to split it into segments and examine boxplots for each segment. Below, there are two sets of boxplots. The first splits the series into 13 eight year segments and the second splits it into 8 thirteen year segments. The line in the middle of each boxplot represents the median. We can see that it is steadily rising, which can be said to represent an upward trend, i.e., a steady increase of the central tendency as time moves forward. It is notable that the last boxplot in

the first graphic appears to represent a larger upward move than all of the previous moves.

Fig. 4

**Boxplots Key West Sea Level
Eight year segments**

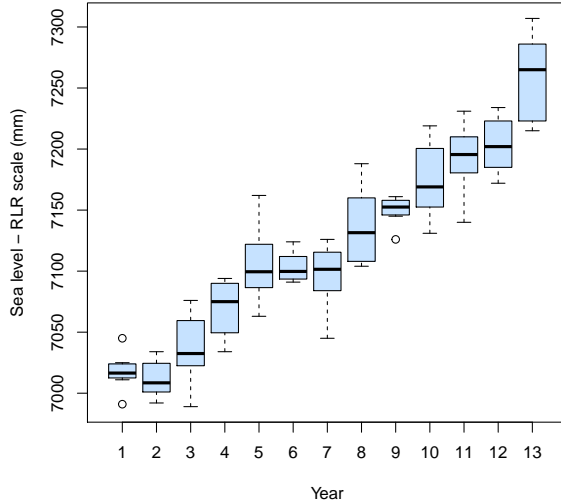
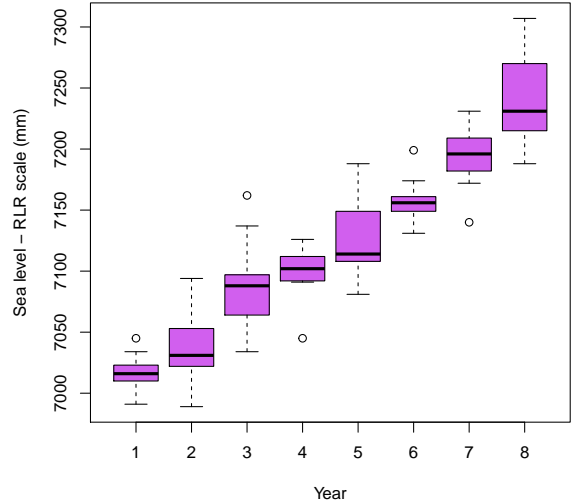


Fig. 5

**Boxplots Key West Sea Level
Thirteen year segments**



The numbers confirm this appearance. In the first graphic, the interquartile range (IQR)⁶ for the last boxplot is 63 mm, which is 11 mm greater than the previous largest IQR of 52 mm. In addition, the difference between the 13th median and the 12th is 63 mm, which is 20.5 mm more than the previous highest gap between medians. The second graphic (dealing with longer, 13 year, intervals) tells a similar story, though not quite so severe. The IQR for the last boxplot is 55 mm, which is 14 mm greater than the previous largest IQR. On the other hand, the difference between the 8th and 7th medians is only 35 mm, which is not as severe as several of the other differences. Together, these results seem to suggest that if an extraordinary movement (i.e., SLA) has occurred, it more likely to have occurred over the smaller time frame of the last eight years - rather than the larger 13 year time frame.

⁶The IQR is the difference between the upper quartile and the lower quartile and is represented graphically by the colored boxes in the boxplots.

5.2 The short series 1959-2016 (used for multiple regression analysis.)

We did multiple regression on a shorter time series, from 1959 to 2016, because data for all the regressors were not available before 1959. Plots for the regressors and the response variable are shown below. Each exhibits a similar upward trend. There are some differences, though. The plot of CO₂ is very smooth, while the other three are jagged. Temperature appears to be the most jagged, which is to be expected because air loses and gains heat more easily than water. The OHC is less jagged than sea level, which may be because it results primarily from a single cause, the retention of heat by the oceans, whereas sea level has multiple causes (as discussed above.)

The OHC and GMST also show a pronounced upturn in the past several years, much like sea level.

Fig. 6
Key West tidal gauge data
Yearly (1959–present)

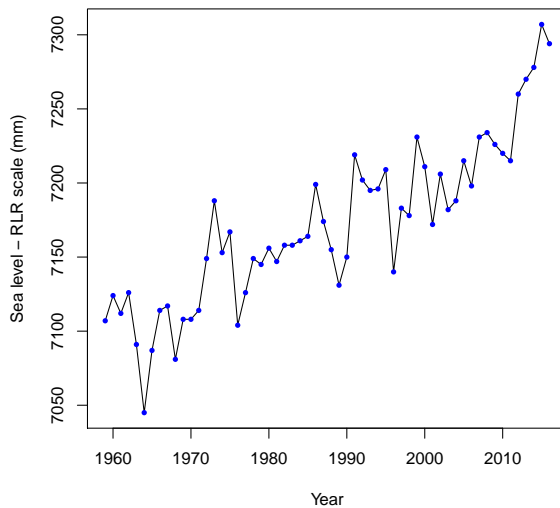


Fig. 7
Mauna Loa CO2 data
Yearly (1959–present)

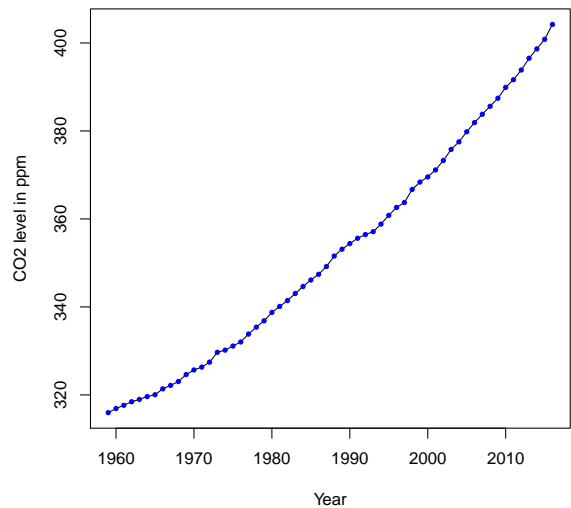


Fig. 8
Ocean Heat Content
Yearly (1959–present)

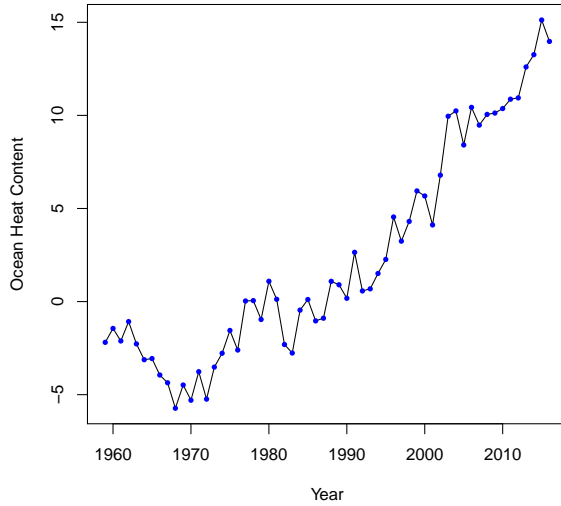
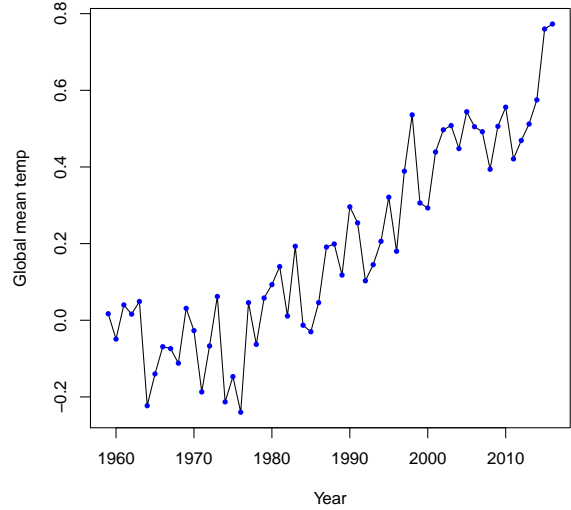


Fig. 9
HadCRUT4 Yearly
Global Mean Surface Temperature (1959–present)



The similar trending nature of the four data sets gives rise to questions of correlation and multicollinearity. An examination of the correlation matrix shows that all of the data sets exhibit pairwise correlation coefficients in excess of 0.82.

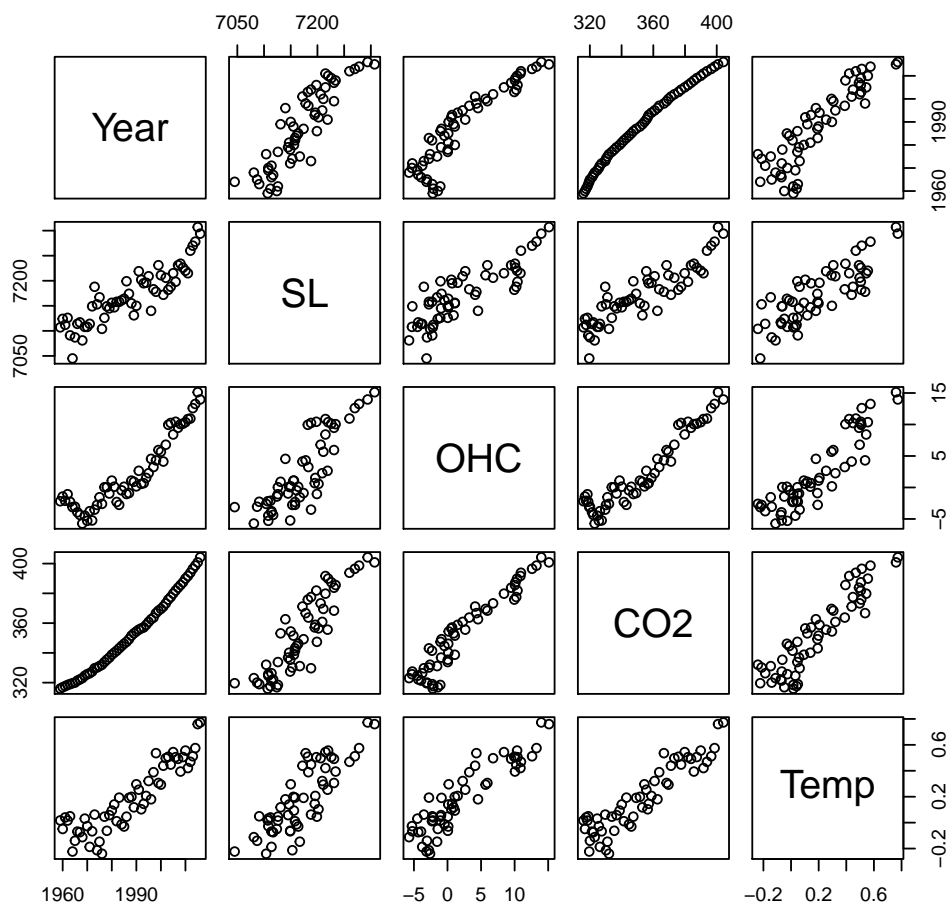
Table 2: Correlation Matrix - Multiple Regression

| | Year | SL | OHC | CO2 | Temp |
|------|-------|-------|-------|-------|-------|
| Year | 1.000 | 0.885 | 0.920 | 0.993 | 0.898 |
| SL | 0.885 | 1.000 | 0.846 | 0.894 | 0.828 |
| OHC | 0.920 | 0.846 | 1.000 | 0.952 | 0.912 |
| CO2 | 0.993 | 0.894 | 0.952 | 1.000 | 0.920 |
| Temp | 0.898 | 0.828 | 0.912 | 0.920 | 1.000 |

The scatterplot matrix below also seem to support the suspected multicollinearity, with all of the scatterplots exhibiting fairly clear linear relationships:

Fig. 10

**Scatterplot Matrix –
Sea level, OHC, CO₂, Time, GMST**



6 MULTIPLE REGRESSION ANALYSIS.

The multiple regression was done with time, GMST, OHC and CO₂ as the regressors and sea level as the response variable. To start, we use a full model with sea level regressed on all four regressors.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \varepsilon$$

where y is sea level, x_1 is time, x_2 is CO₂ concentration, x_3 is OHC and x_4 is GMST. The multiple regression calculation generates the following regression equation:

$$y = 9197 + 3.278x_1 - 1.599x_2 - 1.995x_3 + 6.219x_4 + \varepsilon$$

Statistical analysis of the model shows that the regression itself is highly significant. Further, the coefficients of determination (R^2 and adjusted R^2) are very high.⁷ On the other hand, none of the regressors are statistically significant.⁸ The mismatch is a hallmark of multicollinearity. An analysis of the correlation matrix, variance inflation factors (VIFs) and eigenvalues confirmed that there is severe multicollinearity.⁹

There are several ways to reduce or eliminate multicollinearity. Among these are (i) variable elimination through step regression and subset regression, (ii) ridge regression and (iii) principal component (PC) regression. The following three subsections develop models for the three approaches.

⁷The null hypothesis for the test of overall significance is $H_0 : \beta_0 = \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$. The p-value is very small, 2.2×10^{-16} , so the regression is statistically significant. Further, $R^2 = 0.8026$ and $R_{adj}^2 = 0.7877$, meaning that approximately 80% of the variation is explained by the regression.

⁸The p-values for the individual regression coefficients are all greater than 0.05, except for the intercept. This means that, at level of significance $\alpha = 0.05$, we cannot say that any of the regressor coefficients differ from 0.

⁹As discussed above, the correlation matrix shows pairwise correlation in excess of 0.82 for all of the variable pairs. The VIFs range from 7.457 for temperature to 223.833 for CO₂. Generally, a VIF in excess of 5 indicates multicollinearity. Finally, eigensystem analysis gives a condition number $k = 1366$. A condition number in excess of 1000 indicates severe multicollinearity. Separately and together, these tests indicate severe multicollinearity.

6.1 Reduced model from step regression and subset regression.

The step regression and subset regression both resulted in a model with only one predictor variable - time.¹⁰ The resulting reduced model is:

$$\text{SL} = 2162.836 + 2.518x_1$$

where x_1 is time. The reduction to a single variable, time, suggests that time is an adequate proxy for the other three predictors, which makes sense because (i) CO₂ concentration is nearly perfectly correlated with time ($r = 0.993$) and (ii) the physics of climate change suggests that increased carbon dioxide concentration in the atmosphere is a direct cause of all three of the other phenomena.

6.2 Ridge regression.

In normal linear regression, the least squares estimators are calculated as

$$\hat{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}.$$

If $|\mathbf{X}^T \mathbf{X}| \approx 0$, i.e., if $\mathbf{X}^T \mathbf{X}$ is ill-conditioned,¹¹ then the least squares estimators will be bad. Ridge regression substitutes $\mathbf{X}^T \mathbf{X} + \lambda \mathbf{I}_p$ for $\mathbf{X}^T \mathbf{X}$ (with $\lambda \geq 0$) to get the following regression equation:

$$\hat{\beta}_R = (\mathbf{X}^T \mathbf{X} + \lambda \mathbf{I}_p)^{-1} \mathbf{X}^T \mathbf{y}$$

¹⁰We conducted (i) a step regression in both directions and (ii) a subsets regression analysis, both on the training set (the first 50 years of the series.) In both cases, the best result is a regression of the response (sea level) on a single predictor - time. For step regression, this model gives the best AIC score, i.e., 320.93. Subsets regression compares BIC, Cp and adjusted R^2 scores for all possible subsets. Once again, using time as the single predictor gave the best score for all three statistics.

¹¹A matrix is considered to be ill-conditioned if it is “close to” singular. This will mean that small changes in the data matrix will cause large changes in the coefficient estimates.

The purpose of the λ adjustment is to adjust $\mathbf{X}^T \mathbf{X}$ to a value where the coefficient estimates stabilize.

The key to ridge regression is to find the best value for λ . We do this by finding the smallest value of λ where the ill-conditioning goes away. The plots below reveal that the coefficients stabilize somewhere between $\lambda = 3$ and $\lambda = 4$.

Fig. 11

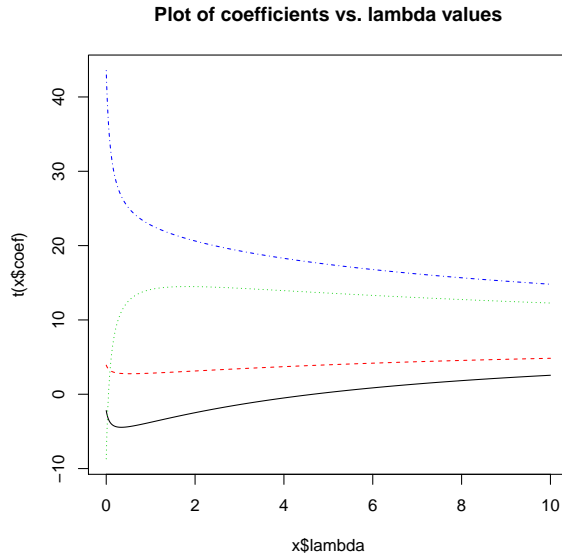
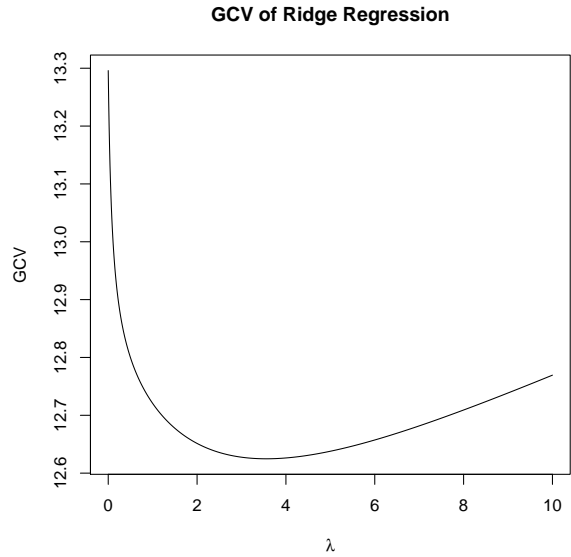


Fig. 12



The final value is $\lambda = 3.56$. The ridge regression model winds up being:

$$y = 4410.61 + 1.296x_1 + 0.6755x_2 - 0.196x_3 + 16.35x_4$$

where y is sea level, x_1 is time, x_2 is CO₂ concentration, x_3 is OHC and x_4 is GMST.

6.3 Principal component regression.

Because the measurement scales for the various predictors are quite different, it is appropriate to first standardize the variables and conduct principal component analysis on the correlation matrix. The i th principal component for the standardized

variables $\mathbf{Z}' = [Z_1, Z_2, \dots, Z_p]$ with $\text{Cov}(\mathbf{Z}) = \boldsymbol{\rho}$ is given by:

$$Y_i = \mathbf{e}_i \mathbf{Z} = \mathbf{e}_i' (\mathbf{V}^{1/2})^{-1} (\mathbf{X} - \boldsymbol{\mu}), \quad i = 1, 2, \dots, p$$

where $\mathbf{V}^{1/2} = \text{diag}(\sqrt{\sigma_{11}}, \sqrt{\sigma_{22}}, \dots, \sqrt{\sigma_{pp}})$ and $(\lambda_i, \mathbf{e}_i)$ are the eigenvalue-eigenvector pairs of $\boldsymbol{\rho}$ with $\lambda_1 \geq \lambda_2 \geq \dots \lambda_p \geq 0$.

When we perform the principal component analysis on the centered and scaled data, we see that approximately 92% of the variance is accounted for by the first PC, with about 4.5% and 3% accounted for by the second and third PCs, respectively.¹²

Then, we conduct the PC regression. We anticipate from the variance analysis above that the first PC will have far more statistical significance than the others. In fact, we see that PC1 is highly significant, PC2 is barely so, and PC3 and PC4 are not statistically significant.¹³ Because the PCs are orthogonal to one another, we should be able to drop the insignificant PCs. To test this, we ran an ANOVA comparison of the two models. The null hypothesis is that the two extra regressors are in fact zero. Our p-value is quite high, so we can use the reduced model (i.e., the model with only PC1 and PC2). The final model (based on centered data) is:

$$y = 18.82x_1 + 15.90x_2 + 5.86x_3 - 4.189x_4$$

with the variables as set forth above. Note that the coefficient assigned to the time variable x_1 is the largest coefficient. When the regression is performed on centered and scaled data, larger coefficients signify greater importance. The large coefficient assigned to time supports the notion that time may be a good proxy for the other

¹²The proportion of variance is 0.9242 for PC1, 0.04402 for PC2, 0.03127 for PC3, and 0.00050 for PC4.

¹³The p-values for the coefficients are: 5.78×10^{-13} for PC1, 0.0425 for PC2, 0.1659 for PC3, and 0.6688 for PC4.

variables. In the end, we will find that pure time series analysis provides at least as good a result as the multiple regression analysis.

6.4 Comparison of multiple regression models.

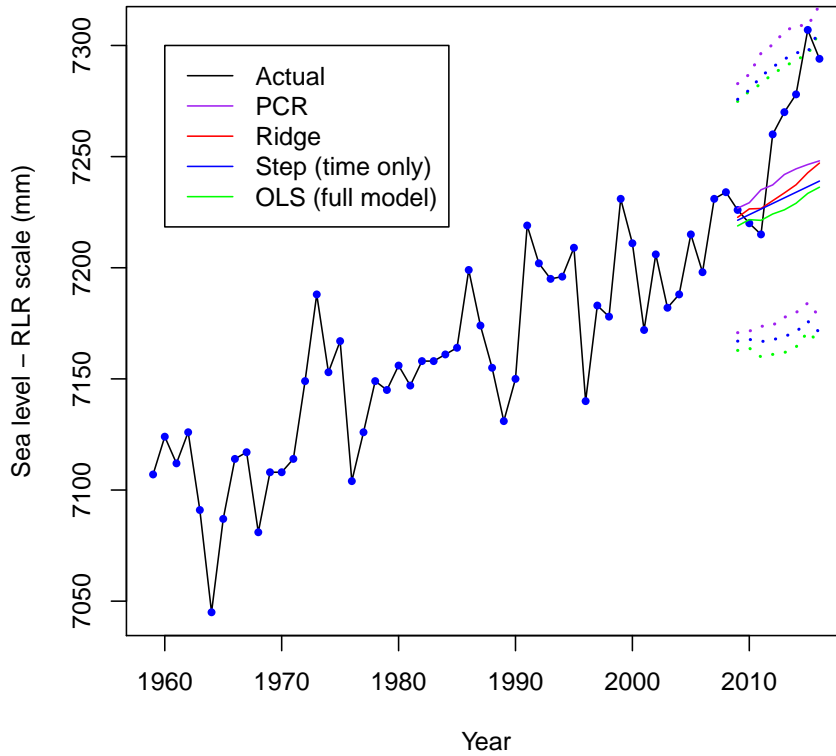
When we compare the step regression, ridge regression and PC regression, we find that they all give fairly similar results, with PC regression being the best. Below is a table showing the root mean square error (RMSE) and the mean absolute deviation (MAD) for the forecasts (with the best results in bold type.) Below that is a plot of the forecasts on the eight year test set for the 4 models (along with prediction intervals for all but ridge regression (which does not have generally accepted prediction intervals.)) We can see from both that the PC regression model is slightly better than the rest. The difference between the models, however, is not as significant as the difference between the modeled data and the actual data. As we have noted before, this can be taken as evidence that the trend has changed, i.e., that there has been SLA during the test period (2009-2016).

Table 3: Multiple Regression Forecast errors: Four models

| Model | Root mean squared error | Mean absolute deviation | Mean absolute percentage error |
|---|-------------------------|-------------------------|--------------------------------|
| Full model (ohc,time,temp,co2) | 42.627 | 34.473 | 0.4734 |
| Reduced model (time only) | 39.621 | 32.392 | 0.4449 |
| Ridge Regression Model | 36.112 | 29.987 | 0.4119 |
| Principal Component Regression Model | 32.933 | 27.590 | 0.3791 |

Fig. 13

**Forecasts and prediction intervals
for 4 multiple regression models**



7 TIME SERIES ANALYSIS.

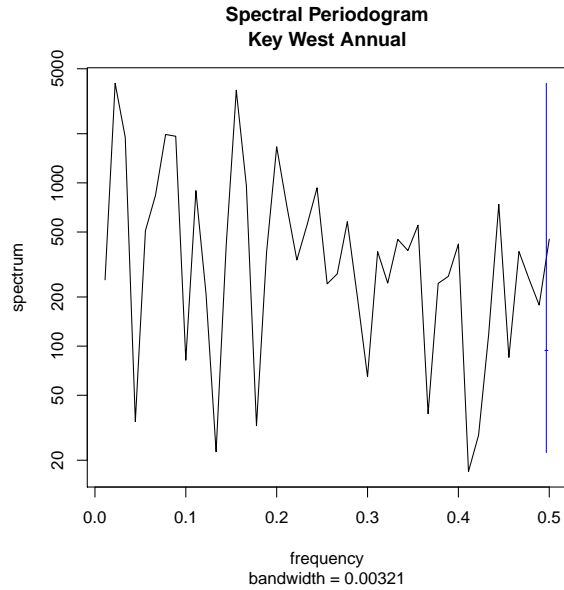
Time series analysis treats time as the sole predictor variable. We have seen above that because of the mutual causal relationship between CO_2 , OHC, and GMST, time can perhaps stand as an adequate proxy for the others. In fact, we will see that the best time only models perform very closely to the best multiple regression model.

7.1 Cyclic behavior.

We selected annual data instead of monthly data to eliminate the annual cycle from the data - since we are really interested in long term behavior. Visual inspection

of the annual data did not show any other apparent cyclical behavior. We can run a spectral periodogram to confirm this.

Fig. 14



The spectral periodogram does not indicate that there is cyclical behavior in addition to the annual cycle that we eliminated with annual data.¹⁴

7.2 Upward trend.

Our visual inspection reveals a clear upward trend. An upward trend can indicate either non-stationarity¹⁵ or trend stationarity¹⁶. We test for these using an augmented

¹⁴There is in fact a heated controversy about whether there is a multidecadal (perhaps 60 year) cycle in sea level. Some claim that such a cycle exists. Others claim that there is not. See (Visser 2015, at pp. 3873-4) for a discussion of the controversy. We do not see any evidence for such a cycle in the current data and will leave it up to others to resolve this issue.

¹⁵A time series is stationary if its properties do not depend upon the time, or more precisely, for all s , the distribution of (y_t, \dots, y_{t+s}) does not depend on t . (Hyndman & Athanasopoulos 2014, at p. 207)

¹⁶A series is trend stationary if it takes the form $y_t = \mu_t + \varepsilon_t$, with μ_t being a deterministic mean trend and ε_t being a stationary stochastic process. (Hyndman & Athanasopoulos (2014), at p. 259.)

Dickey-Fuller (ADF) test.¹⁷ The null hypothesis for the ADF test is that the series is non-stationary (i.e., it has a unit root). The resulting p-value in our case is 0.05263, which means that we cannot determine definitively whether the series is stationary, nonstationary or trend stationary. For this reason, we will conduct two analyses: (i) an autoregressive integrated moving-average (ARIMA) analysis which treats the trend as stochastic, and (ii) a dynamic linear regression, which treats the trend as deterministic.

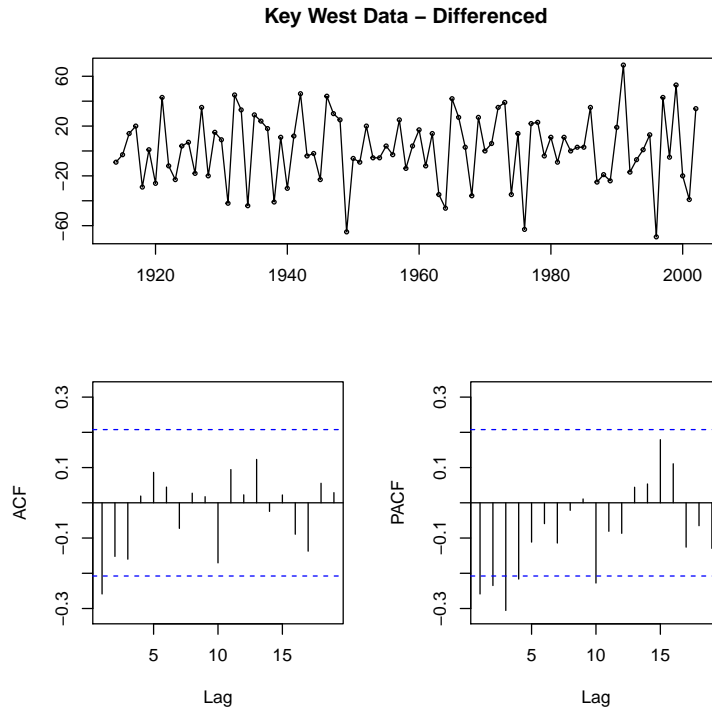
7.3 ARIMA analysis for a stochastic trend.

The ARIMA analysis assumes a stochastic trend and uses differencing to remove that trend. Once the trend is removed, a stationary autoregressive moving-average (ARMA) process should remain.

Initially, we do a first difference of the training data. The plot of the differenced series is below along with plots of the autocorrelation function (ACF) and partial autocorrelation function (PACF). A visual inspection of this plot indicates that the differenced series is stationary. We test for stationarity with an ADF test. The ADF test gives a p-value of 0.01, confirming that the differenced series is indeed stationary.

¹⁷We use the ADF test from the package `tseries` in R. This test automatically tests for difference stationarity, i.e., the series becomes stationary after differencing. If the series is difference stationary, the series is first differenced and then the test is applied. If the series is trend stationary, the test will automatically detrend the series and test for stationarity on the residuals.

Fig. 15



The ACF shows a significant value for lag 1 and cuts off after that. An ACF that cuts off at a given lag is indicative of a moving average process with that order. Initially, this looks like it might be an MA(1) process, which means that the original (undifferenced) time series would be an ARIMA(0,1,1)¹⁸ process. The PACF appears to show exponential decay or perhaps a mildly sinusoidal pattern. This also indicates an MA process, further supporting the suspicion.

Using this as a starting point, we can test several models, each of them with and without stochastic drift. They are:

1. ARIMA(0, 1, 1) with and without drift;
2. ARIMA(0, 1, 2) with and without drift;

¹⁸The standard notation for ARIMA models is ARIMA(p,d,q), where d is the number of differences needed to make the differenced model stationary, p is the lag of the autoregressive (AR) component, and q is the lag of the moving-average (MA) component.

3. ARIMA(1, 1, 0) with and without drift;
4. ARIMA(2, 1, 0) with and without drift; and
5. ARIMA(1, 1, 1) with and without drift.

The models with stochastic drift uniformly tested better than those without it, which was to be expected because the series drifts upwards. Also, the MA coefficients had much smaller p-values than the AR coefficients, which was also expected because the ACF and the PACF gave no real indication that the model is an AR only model. The MA only models performed much better, as did the ARIMA(1, 1, 1) models. Below are tables comparing various evaluating criteria for the models.

Table 4: ARIMA Model Comparison (models with drift)

| | (0,1,1) | (0,1,2) | (1,1,0) | (2,1,0) | (1,1,1) |
|----------------|---------|----------------|---------|---------|----------------|
| RMSE | 24.912 | 23.081 | 26.709 | 25.960 | 22.998 |
| MAD | 19.01 | 17.526 | 20.497 | 20.001 | 17.250 |
| AIC | 832.64 | 824.24 | 844.3 | 841.34 | 823.27 |
| AICc | 832.92 | 824.72 | 844.58 | 841.82 | 823.74 |
| BIC | 840.1 | 834.2 | 851.76 | 851.3 | 833.22 |
| Log Likelihood | -413.32 | -408.12 | -419.15 | -416.67 | -407.63 |

Table 5: ARIMA Model Comparison (models without drift)

| | (0,1,1) | (0,1,2) | (1,1,0) | (2,1,0) | (1,1,1) |
|----------------|---------|----------------|---------|---------|---------------|
| RMSE | 25.644 | 24.970 | 26.826 | 26.150 | 25.049 |
| MAD | 19.507 | 19.508 | 20.738 | 20.390 | 19.489 |
| AIC | 835.39 | 832.67 | 843.08 | 840.62 | 833.2 |
| AICc | 835.53 | 832.95 | 843.21 | 840.91 | 833.49 |
| BIC | 840.37 | 840.14 | 848.05 | 848.09 | 840.67 |
| Log Likelihood | -415.7 | -413.34 | -419.54 | -417.31 | -413.6 |

The clear best model was ARIMA(1, 1, 1) with drift. ARIMA(0, 1, 2) with drift was fairly close behind. Finally, ARIMA(0, 1, 1) with drift came in third place (but was not very close.) The coefficients for these three models were:

Table 6: ARIMA Model Coefficients

| Model | Drift Coefficient | AR(1) Coefficient | MA(1) Coefficient | MA(2) Coefficient |
|---------------------------|-------------------|-------------------|-------------------|-------------------|
| ARIMA(1, 1, 1) with drift | 2.2346 | 0.3884 | -1.000 | NA |
| ARIMA(0, 1, 2) with drift | 2.2413 | NA | -0.6419 | -0.3581 |
| ARIMA(0, 1, 1) with drift | 2.1103 | NA | -0.7476 | NA |

We can calculate the actual model from the coefficients in the above table. The general form for an ARIMA model is:

$$\phi_p(B)\nabla^d z_t = \theta_0 + \theta_q(B)\varepsilon_t$$

where θ_0 is the drift coefficient, B is the backward shift operator¹⁹, the ϕ values are the AR coefficients, the θ values are the MA coefficients, and:

$$\phi_p(B) = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)$$

$$\theta_q(B) = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q)$$

$$\nabla^d = (1 - B)^d$$

We will calculate the best model, i.e., ARIMA(1, 1, 1) with drift. It turns out to be:

$$(1 - 0.3884B)(1 - B)z_t = 2.2346 + (1 + B)\varepsilon_t$$

$$(1 - 1.3884B + 0.3884B^2)z_t = 2.2346 + \varepsilon_t + \varepsilon_{t-1}$$

$$z_t = 2.2346 + 1.3884z_{t-1} - 0.3884z_{t-2} + \varepsilon_t + \varepsilon_{t-1}$$

¹⁹The backward shift operator, when applied to a time subscripted variable, shifts that variable back one time period. For example, $B(z_t) = z_{t-1}$.

7.4 Forecast errors for ARIMA models.

The following table compares the forecast errors for various ARIMA models. We see that the best ARIMA model is ARIMA(1,1,1) with drift - confirming the result we got for the modeling errors:

Table 7: Forecast errors - ARIMA models

| | RMSE | MAD |
|----------------------|---------------|---------------|
| Models with Drift | | |
| (0,1,1) | 36.185 | 28.050 |
| (0,1,2) | 35.255 | 27.532 |
| (1,1,0) | 37.740 | 29.140 |
| (2,1,0) | 36.794 | 28.599 |
| (1,1,1) | 35.293 | 27.479 |
| Models without Drift | | |
| (0,1,1) | 54.685 | 43.203 |
| (0,1,2) | 54.890 | 43.626 |
| (1,1,0) | 53.370 | 41.998 |
| (2,1,0) | 52.941 | 41.822 |
| (1,1,1) | 55.498 | 44.180 |

7.5 Dynamic regression analysis for a deterministic trend.

Dynamic linear regression analysis uses the model form $y_t = \beta_0 + \beta_1 x_t + n_t$, where n_t is an ARIMA process having the usual ARIMA form: $y_t = c + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \theta_1 e_{t-1} + \dots + \theta_q e_{t-q} + e_t$.²⁰ In our case, n_t turns out to be an AR(1) process.²¹ The

²⁰The dynamic linear regression process differs from simple linear regression on the time variable because the errors are not assumed to be uncorrelated as they are in the general regression context. Instead, only the noise process e_t is considered to be uncorrelated, and estimation proceeds by minimizing the sum of squares of e_t instead of the sum of squares of n_t .

²¹An AR(1) process is the same as an ARIMA(1,0,0) process or an ARMA(1,0) process.

final model turns out to be:

$$z_t = 6999.95 + 2.2351(t - 1913) + n_t$$

$$n_t = 0.373n_{t-1} + \varepsilon_t$$

$$\varepsilon_t \sim NID(0, 546.5)$$

7.6 Comparison of ARIMA and dynamic regression models.

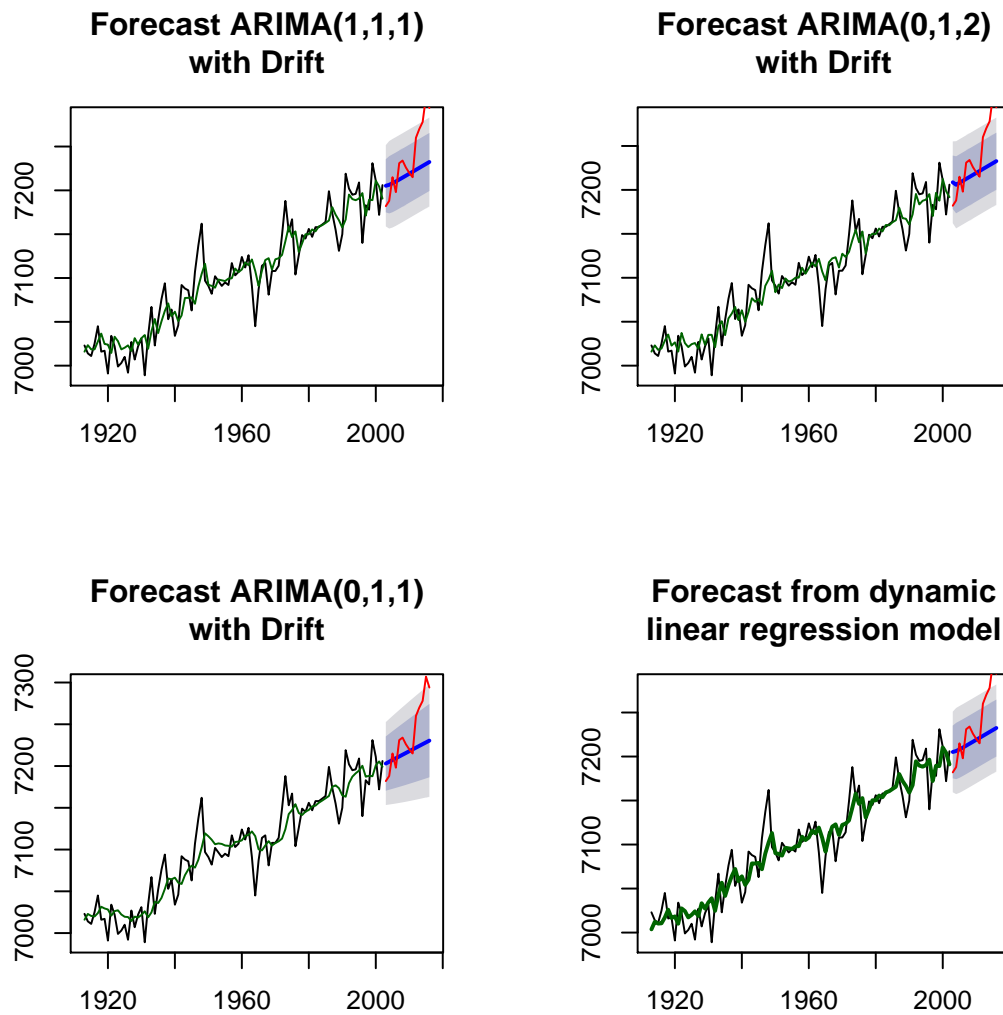
The dynamic linear regression models gave similar results, i.e., the results were similar regardless of whether the trend was considered to be stochastic or deterministic. The model errors and the forecast errors for the both models were extremely close. The table immediately below compares the RMSE and the MAD for the best ARIMA model with the dynamic linear regression model.

Table 8: Forecast and model errors
Best ARIMA v. Dynamic Linear Regression

| Model | Model Errors | | Forecast Errors | |
|---|--------------|-------|-----------------|--------|
| | RMSE | MAD | RMSE | MAD |
| Dynamic Linear Regression | 22.985 | 17.01 | 35.280 | 27.466 |
| Best ARIMA Model ARIMA(1,1,1) with drift | 22.998 | 17.25 | 35.293 | 27.479 |

We can also see the similarity of the models in the plots below, which show the models and the forecasts for the three best ARIMA models and the dynamic linear regression model. They are all very similar. We can also see that the forecasts on the test set are biased substantially low when compared to the actual results - and even go outside the prediction intervals. This is a further indication that the trend increased during the test period, i.e., that there is SLA.

Fig. 16: Forecasts: Best three ARIMA models plus Dynamic Regression



8 REGRESSION MODELS (LINEAR, QUADRATIC, CUBIC).

Time series can also be analyzed using ordinary least squares models and assuming that the residuals will take on ARIMA form. As discussed above in relation to dynamic linear regression models, ordinary regression models on time series violate one of the fundamental assumptions of regression, i.e., that the errors are uncorrelated. The effect of this is that forecasts will still be unbiased, but will have larger prediction intervals than in the uncorrelated case. (Hyndman & Athanasopoulos 2014, at p.103)

Since we are more concerned here with trend detection than with forecasting, we proceeded with the analysis. We began by conducting three different linear regressions on the 90 year training set:

1. Linear regression;
2. Quadratic regression (regression with a quadratic term); and
3. Cubic regression (regression with quadratic and cubic terms).

On the 90 year training set, the linear term is significant on all three models, but the quadratic and cubic terms are not. The p-value for the quadratic term on the quadratic regression was 0.999, and the p-values for the quadratic and cubic terms on the cubic regression were 0.8691 and 0.8671, respectively. This suggests that the quadratic and cubic models do not add much to the analysis of the training set.

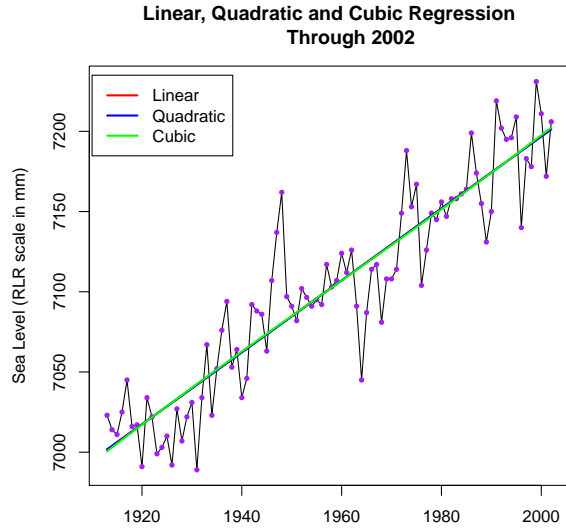
The table below contains the coefficients of determination and residual analyses for the three models. The error values and the coefficients of determination (R^2) are very similar for all three models. Below the table is a plot of the training series with the three regression curves superimposed. The three curves basically sit on top of one another. All of these things indicate that the quadratic and cubic models probably do not add anything to the linear model - for the training set.²²

Table 9: Comparison of OLS models on the training set

| Model | RMSE | MAD | R^2 | R^2_{adj} |
|-----------|--------|--------|-------|-------------|
| Linear | 24.799 | 18.748 | 0.847 | 0.845 |
| Quadratic | 24.799 | 18.749 | 0.847 | 0.843 |
| Cubic | 24.795 | 18.784 | 0.847 | 0.841 |

²²Note also, as discussed above, the ARIMA and dynamic linear regression models give better performance than the pure OLS models.

Fig. 17



Things change dramatically, however, if we run the same models on the full 104 year time series. In that case, the quadratic term becomes significant (p-value=0.0433) in the quadratic model, and slightly less significant in the cubic model (p-value=0.0574). This is a big change from the models on the training set, where the quadratic term p-values were 0.999 and 0.869, respectively. Also, the cubic term becomes significant (p-value=0.0236). The residual analysis for these models (in the table below) indicates that the quadratic and cubic models now add something to the analysis. We can see that the cubic and quadratic models not only differ from the linear model, but they seem to track the actual data better.

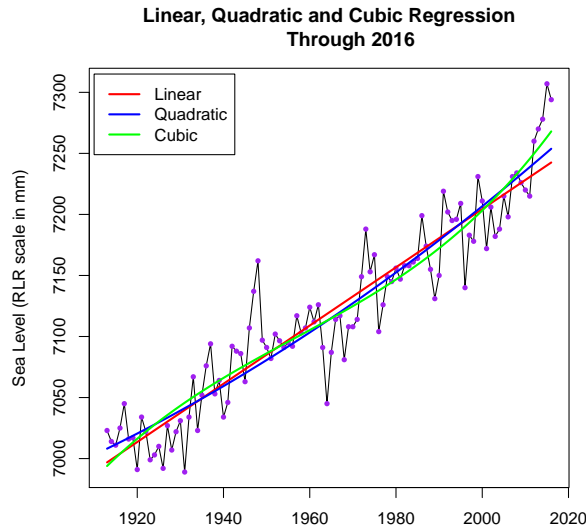
Table 10: Comparison of OLS models on the full set

| Model | RMSE | MAD | R^2 | R^2_{adj} |
|-----------|--------|--------|-------|-------------|
| Linear | 25.936 | 19.829 | 0.884 | 0.883 |
| Quadratic | 25.414 | 19.336 | 0.889 | 0.886 |
| Cubic | 24.768 | 19.353 | 0.894 | 0.891 |

The plot of the full series makes this more clear. Below is the plot with the three regression curves overlaid. Recall that with the shorter series the regression curves

for all three models were right on top of each other. With the longer series, they diverge substantially, with the cubic model seeming to track the deviations in the past few years best. In addition, both the quadratic and cubic models are concave up, providing further evidence of SLA.

Fig. 18



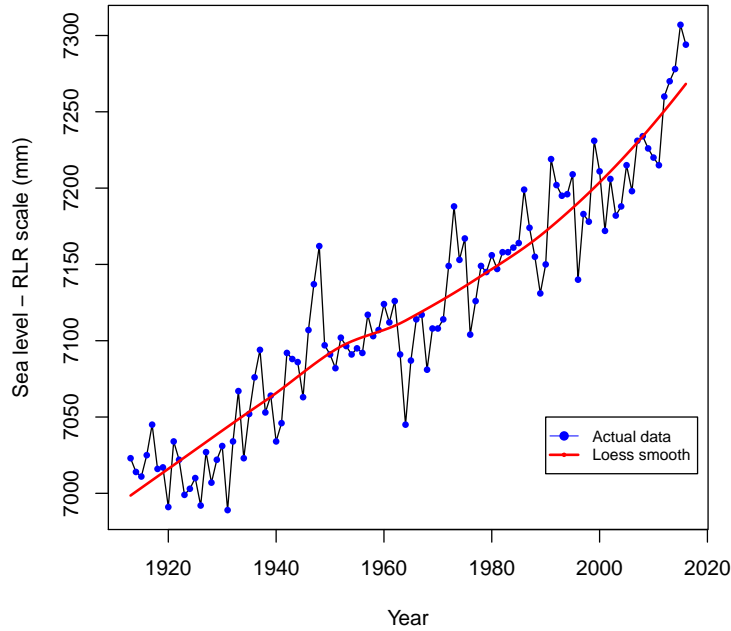
9 NONPARAMETRIC ANALYSIS: LOESS SMOOTH.

We can also use nonparametric techniques to remove a trend. This generally involves some kind of local averaging to ascertain the change in level of the series. An example of this is the Loess smooth (a nonparametric locally weighted technique). Loess has the advantage of smoothing without edge effects.²³ Below is the full Key West series with a Loess estimated trend in red:²⁴

²³An edge effect occurs when there is not enough data at the edges of the series to calculate the trend function. For example, a 12 period moving average cannot calculate a value for the first 6 or the final 6 values of the predictor.

²⁴Note that the loess curve has an appearance similar to the cubic regression curve above and is concave down before about 1960 and concave up afterwards.

Fig. 19
Key West annual data with Loess smooth
(1913–present)



Nonparametric techniques cannot be used for extrapolation of a trend because the trend does not take a predictable functional form. However, the modeled trend can be used to infer rates of change and acceleration. For example, based on the Loess smoothed data, a simple estimator of the instantaneous rate of change of sea level for each year might be the difference between the smoothed sea level value in two consecutive years divided by the time difference - as follows:

$$\dot{x} \approx \frac{y_t - y_{t-1}}{x_t - x_{t-1}}$$

where $y_t - y_{t-1}$ is the difference in sea level between two consecutive years and $x_t - x_{t-1}$ is the difference in time between 2 years, i.e., one year. Then, we can do the same thing a second time to yield an estimate for acceleration.

The results for our example are below. The figure on the left shows the estimate for the rate of change (“velocity”) of the trend curve. It shows a steady increase in

the rate of change of SLR since about 1960.²⁵ The figure on the right shows the acceleration. It shows a more or less increasing positive acceleration since the 1960's which appears to have stabilized at a (more or less) constant level recently. These plots provide further evidence that SLA is taking place.

Fig. 20
Slope of Loess Smooth (velocity)
SLR Key West - (1913-2016)

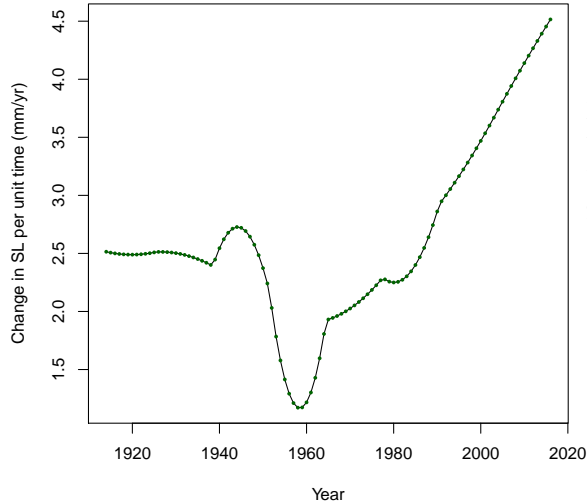
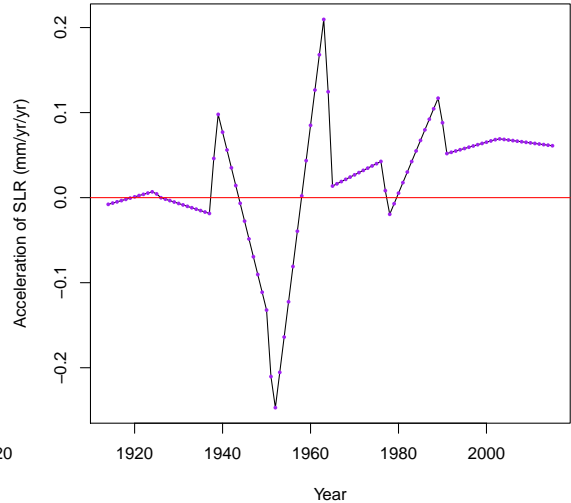


Fig. 21
Acceleration of SLR
Key West - (1913-2016)



10 CHANGEPOINT ANALYSIS.

We have seen that the most interesting feature of the Key West time series comes at the end of the series, where it seems to very upwards off the trend. The trend for all the forecasted models substantially underestimates sea level at the end of the test period. While this could be the result of a faulty model or a random fluctuation, it could also be the result of a change in the trend over the last 10 years or so. (In fact, most climate researchers expect such a change to occur at some point because of (i) increased melting of land based ice sheets and the flow of the melt water into

²⁵Note that the scales of vertical axes of the three plots are substantially different. If one is not careful, this can be misleading as to the magnitude of the rates of change.

the oceans, (ii) thermal expansion of sea water as it warms, and (iii) changes in the Florida Current.)

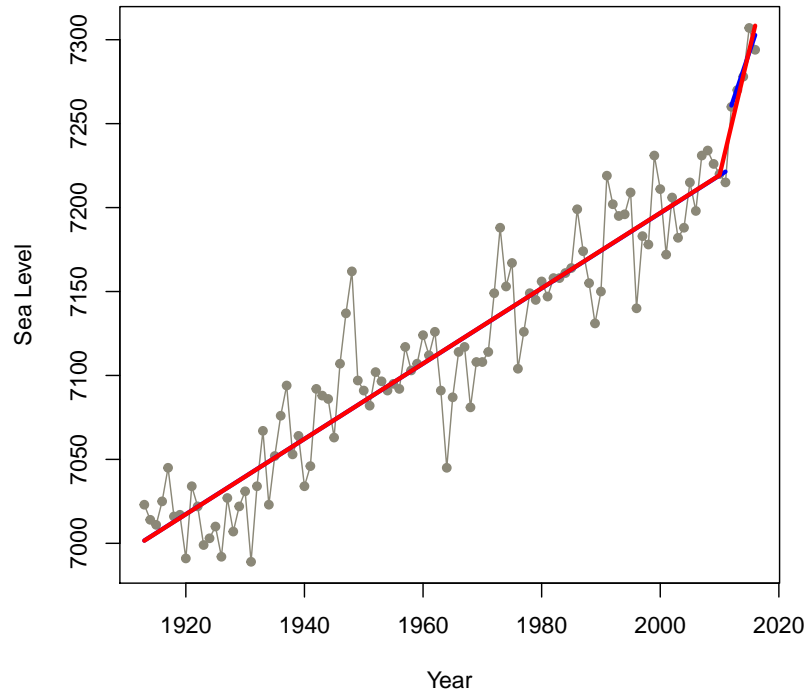
Changepoint analysis is a way to test for this. There are several tests for change-point. The first test was based on the Chow test. The Chow test statistic is as follows:

$$\frac{(SS_C - (SS_1 + SS_2))/k}{(SS_1 + SS_2)/(n_1 + n_2 - 2k)}$$

where SS_C is the SS_{RES} for the full series, SS_1 is the SS_{RES} for the model based on the series up to the breakpoint, SS_2 is the SS_{RES} for the model based on the series after the breakpoint, and k is the number of parameters. The test is run on all potential breakpoints and the one with the lowest p-value is chosen. The result shows a change in trend at 2010. In addition, the RMSE and MAD for the last 14 years were 13.85 and 12.20, respectively, which is a substantial improvement over the forecast errors for the other models.²⁶

²⁶The algorithm used for the Chow changepoint analysis was formulated in (Foster 2017).

Fig. 22
Chow test results – Key West Data



There are also dedicated R packages that conduct changepoint analysis. Below on the right is the result from package “Segmented.” It is very similar to the Chow result. Below on the right is the result from the package “Changepoint,” which is run on the residual series, but gives a similar result - i.e., a change in the residual series over the last several years.

Fig. 23

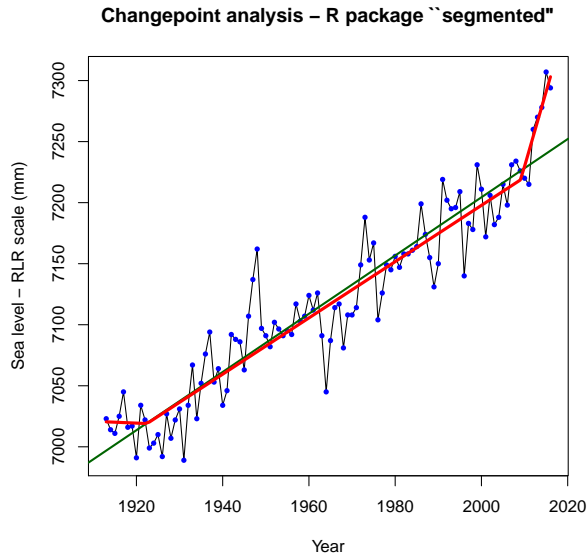
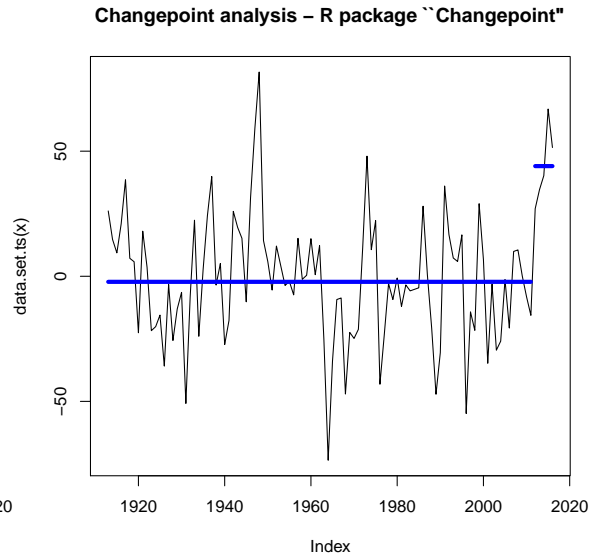


Fig. 24



Each of these tests give further evidence that the trend has increased over the past few years, providing further evidence that SLA has occurred.

11 DISCUSSION.

We have looked at a variety of models, each of which has its own characteristics, its own disadvantages and advantages. Some of the models explore the relationship of sea level to a single independent variable, time. Others add additional non-temporal variables such as OHC, GMST, and CO_2 . All cases revealed an obvious upward trend. Further, all the methods revealed evidence for a change in the trend (acceleration) over the past several years.

The initial visual exploration of the sea level data reveals a clear upward trend through the entire series. In addition, the trend appears to have quickly become more steep at the very end. Boxplot analysis supports the idea that the trend has changed over a period of less than thirteen years, closer to eight.

Parametric techniques were represented by various types of multiple regression analysis. Multiple regression analysis was undertaken on time plus three additional

variables (OHC, GMST and CO₂). This analysis initially revealed a high degree of multicollinearity, which can cause the regression to be unreliable, so it was necessary to take steps to ameliorate the multicollinearity. We tried several different approaches. Ultimately, we wound up with four different multiple regression models, ranked as follows from best to worst: (i) principal component regression, (ii) ridge regression, (iii) step regression/subset regression, which gave us a model based solely on time vs. sea level (similar to dynamic linear regression discussed below), and (iv) the full model, with its accompanying problem of multicollinearity. All of the multiple regression models revealed a clear upward trend. In addition, all of these models gave severe forecast errors over the training set, which indicates that the trend changed during the training set, i.e., the last eight years.

Pure time series analysis compares sea level to time alone. This was done two ways, one of which assumes that the trend is deterministic in nature (dynamic linear regression) and the other (ARIMA analysis) which assumes that the trend is stochastic in nature.²⁷ The ARIMA analysis and the dynamic linear regression present similar upward trends. In addition, the forecast on the test set gives values that are substantially less than the actual time series values. This implies that the trend increased during the test set, which is evidence of acceleration.

Because of residual autocorrelation, ordinary least squares analysis (another parametric technique) is not optimal for forecasting time series, but it does deliver unbiased trend estimates. We looked at linear, quadratic and cubic ordinary least squares analysis of the series on two different length windows: (i) the shorter 90 year training series (1913-2002), and (ii) the full 104 year series (1913-2016). The trend curves for the shorter window were all very close to each other, indicating that the quadratic and cubic terms did not add much to the analysis. On the other hand, when applied

²⁷While we would generally assume that the trend results from a natural cause (or causes) and is therefore deterministic, our statistical tests could not clearly verify that this is the case.

to the longer series, the three models began to differ from each other, indicating that the slope of the trend line was changing. Because the trend line became concave up, this indicates that SLA may be present.

Nonparametric techniques were represented by a Loess smooth. Once again, the extracted trend was clearly upward. Nonparametric techniques do not allow us to make forecasts in the same fashion as parametric techniques do, because the trend is locally based and it we cannot simply extend the trend as we normally would in the parametric case. On the other hand, we can analyze the local Loess trend for rates of change over time. When we do this, we can see a clearly increasing rate of change since 1960, as well as something resembling a positive acceleration over the same period. This provides further support for the existence of acceleration.

Finally, we looked at changepoint analysis, which uses numerical methods to find the points where a trend changes. We tried several changepoint techniques. Each revealed a significant trend break over the past 8-10 years, all indicating an increase in trend, i.e., an acceleration.

REFERENCES

- Baart, F., van Koningsveld M., & Stive, M.J.F. (2012) Trends in Sea-Level Trend Analysis. *Journal of Coastal Research*, 28(2), 311-315. West Palm Beach, FL.
- Butler, J. & Montzka, S. (2016). NOAA Annual Greenhouse Gas Index. Retrieved from: <https://www.esrl.noaa.gov/gmd/aggi/aggi.html>.
- Chandler, R., Scott, E.M. 2011. *Statistical Methods for Trend Detection and Analysis in the Environmental Sciences*: West Sussex, United Kingdom: John Wiley & Sons, Ltd.
- DeConto, Robert M. & Pollard, David (2016). Contribution of Antarctica to past and future sea-level rise. *Nature*, 531, 591-597.
- Domingues, R., Baringer, M., Goni, G. (2017) Recent efforts from NOAA/AOML towards monitoring and improving the understanding of changes in the Florida current: relationships with sea level along the east U.S. coast. Retrieved from: <http://www.ces.fau.edu/gulfstream/pdfs/domingues.pdf>.
- Foster, G., & Brown, P. T. (2015). Time and tide: Analysis of sea level time series. *Climate Dynamics*, 45(1-2), 291-308.
- Foster, G. (2017). Climate Trend Change? Do it Right. Retrieved from: <https://tamino.wordpress.com/2017/07/11/climate-trend-change-do-it-right/>.
- GLOSS. Global Sea Level Observing System. 2014. Retrieved from: <http://www.gloss-sealevel.org>.
- Hansen, J., Sato, M., Hearty, P., Ruedy, R., Kelley, M., Masson-Delmotte, V., Russell, G., Tselioudis, G., Cao, J., Rignot, E., Velicogna, I., Tormey, B., Donovan, B., Kandiano, E., von Schuckmann, K., Kharecha, P., Legrande, A.N., Bauer, M., Lo, K (2016). Ice melt, sea level rise and superstorms: evidence from paleoclimate data, climate modeling, and modern observations that 2 C global warming could be dangerous. *Atmospheric Chemistry and Physics*, 16, 3761-3812.
- Hartmann, D. 2016. *Global Physical Climatology*. 2d Ed. Waltham, MA: Elsevier.
- Hyndman, R., Athanasopoulos, G. 2014 *Forecasting: Principles and Practice*. 1st Ed. Otexts.com.

IPCC AR5, WG I Technical Summary - Stocker, T.F., D. Qin, G.K. Plattner, L.V. Alexander, S.K. Allen, N.L. Bindoff, F.M. Brion, J.A. Church, U. Cubasch, S. Emori, P. Forster, P. Friedlingstein, N. Gillett, J.M. Gregory, D.L. Hartmann, E. Jansen, B. Kirtman, R. Knutti, K. Krishna Kumar, P. Lemke, J. Marotzke, V. Masson-Delmotte, G.A. Meehl, I.I. Mokhov, S. Piao, V. Ramaswamy, D. Randall, M. Rhein, M. Rojas, C. Sabine, D. Shindell, L.D. Talley, D.G. Vaughan and S.P. Xie, 2013: Technical Summary. In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Stocker, T.F., D. Qin, G.K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 33-115, doi:10.1017/ CBO9781107415324.005.

IPCC AR5, WG I Chapter 5 - Masson-Delmotte, V., M. Schulz, A. Abe-Ouchi, J. Beer, A. Ganopolski, J.F. Gonzalez Rouco, E. Jansen, K. Lambeck, J. Luterbacher, T. Naish, T. Osborn, B. Otto-Bliesner, T. Quinn, R. Ramesh, M. Rojas, X. Shao and A. Timmermann, 2013: Information from Paleoclimate Archives. In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 383-464, doi:10.1017/CBO9781107415324.013.

IPCC AR5, WG I Chapter 13 - Church, J.A., P.U. Clark, A. Cazenave, J.M. Gregory, S. Jevrejeva, A. Levermann, M.A. Merrifield, G.A. Milne, R.S. Nerem, P.D. Nunn, A.J. Payne, W.T. Pfeffer, D. Stammer and A.S. Unnikrishnan, 2013: Sea Level Change. In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 1137-1216, doi:10.1017/ CBO9781107415324.026.

Kopp, Robert E.; Kemp, Andrew C.; Bittermann, Klaus; Horton, Benjamin P.; Donnelly, Jeffrey P.; Gehrels, W. Roland; Hay, Carling C.; Mitrovica, Jerry X.; Morrow, Eric D.; Rahmstorf, Stefan (2016). Temperature-driven global sea-level variability in the Common Era. *Proceedings of the National Academy of Sciences of the United States of America*. Vol. 113 Issue 11, pE1434-E1441. DOI: 10.1073/pnas.1517056113.

National Centers for Environmental Information (NCEI)- Sea Level Data . October 8, 2015. Retrieved from: <https://www.nodc.noaa.gov/General/sealevel.html>.

National Ocean Service. What is a tide gauge? National Oceanic and Atmospheric Administration (NOAA). Retrieved from: <https://oceanservice.noaa.gov/facts/tide-gauge.html>.

National Oceanic and Atmospheric Administration (NOAA), Red Sismica de Puerto Rico, Global Sea Level Observing System (GLOSS), U. South Florida, U. Hawaii Sea Level Center. 2013. U.S. National Sea Level Report. Retrieved from: http://www.gloss-sealevel.org/publications/documents/us_gexiii2013.pdf.

Park, J., Sweet, W. (2015). Accelerated sea level rise and Florida current transport. *Ocean Sci.* 11, 607-615, doi:10.5194/os-11-607-2015.

Parker, Bruce B. (1992). Sea Level as an Indicator of Climate and Global Change. Retrieved from: <https://tidesandcurrents.noaa.gov/sltrends/mtsparker.htm>.

Peltier, R., Drummond, R. Permanent Service for Mean Sea Level (PSMSL). (2012) Peltier GIA data sets. Retrieved from: http://www.psmsl.org/train_and_info/geo_signals/gia/peltier/.

Permanent Service for Mean Sea Level (PSMSL). Home page. August 21, 2017. Retrieved from: <http://www.psmsl.org>.

Permanent Service for Mean Sea Level (PSMSL). Revised Local Reference (RLR) Definition. Retrieved from: <http://www.psmsl.org/data/obtaining/rlr.php>.

Rahmstorf, S., & Vermeer, M., 2011. Discussion of: Houston, J.R. and Dean, R.G., 2011. Sea-Level Acceleration Based on U.S. Tide Gauges and Extensions of Previous Global-Gauge Analyses. *Journal of Coastal Research*, 27(3), 409-417. *Journal of Coastal Research*, 27(4), 784-787. West Palm Beach, Florida.

Visser, H., Dangendorf, S., Petersen, A (2015) A review of trend models applied to sea level data with reference to the "acceleration-deceleration debate," *J. Geophys. Res. Oceans*, 120, 3873-3895, doi: 10.1002/2015JC010716.

Woodworth, P. L., Menendez, M., & Gehrels, W. R. (2011). Evidence for Century-Timescale Acceleration in Mean Sea Levels and for Recent Changes in Extreme Sea Levels. *Surv Geophys*, 32, 603-618.