Empirical Model: ENSO, Solar, Volcanic Aero, Anthro

Revised 7-23-2015

On time scales of 10 to 50 years (and longer) decadal climate forecasts are difficult to make with general circulation climate models due to their many uncertainties [IPCC, 2007]. We will use the methodology first used by Schonwiese and Bayer in "Some statistical aspects of anthropogenic and natural forced global temperature change", 1995. We will model global temperature by with a multiforced lagged regression by combining ENSO, Volcanic Aerosols, Irradiance, and the effects of Anthropogenic Influence.

Analysis

Using the most recently available characterizations of ENSO, E, volcanic aerosols, V, solar irradiance, S, and anthropogenic influences, A, we perform multiple linear regression analysis to decompose monthly mean surface temperature anomalies since 1980 into four components.
Monthly mean surface temperature anomalies ΔT_{MS} are reconstructed as:

$$\Delta T_{MS}(t) = co + c_E \cdot E\left(t - \Delta t_E\right) + c_V \cdot V\left(t - \Delta t_V\right) + c_S \cdot S\left(t - t_S\right) + c_A \cdot A\left(t - t_A\right)$$

Where E, V, S and A are the time series and the lags (in months) are $\Delta t_E = 7$, $\Delta t_V = 8$, and $\Delta t_S = 2$ and $\Delta t_A = 17$ years. The lags are chosen to maximize the proportion of global variability that the statistical model captures and are spatially invariant (although a geographical dependence is expected). The fitted coefficients, c_o. . ., are obtained by multiple linear regression against the instrumental surface temperature record (HadCRUT4). The Temp data was median smooth within a 7 month period.

The multivariate ENSO index, E, is a weighted average of the main ENSO features contained in sea-level pressure, surface wind, surface sea and air temperature, and cloudiness [Wolter and Timlin, 1988]. Volcanic aerosols, V, in the stratosphere are compiled by Sato et al. [1993] since 1850, updated from giss.nasa.gov to 1999 and extended to the present with zero values. Although some volcanic activity occurred between 2006 and 2008, it is difficult to calculate the aerosol optical depth because of the lack of direct quantitative space-based observations. Solar irradiance, S, is estimated as the competing effects of sunspots and facular, identified in observations made by space-based radiometers [Lean et al., 2005]. The anthropogenic influence, A, is the Forcing Effect of the concentration (ppm) of CO_2.

Climate Forecasting:

Using global and regional surface temperature responses to the four individual influences parameterized by regression against the observations from 1980 to 2008, we forecast change from 2009 to 2020 by adopting the best estimate of how each influence will change in the future. The anthropogenic forcing in the past 40 years is well represented by a linear trend that we extrapolate into the future.

We assume that future solar irradiance cycles replicate cycle 23, with cycle 24
(See: **http://www.leapcad.com/Climate_Analysis/Climate_Data-Proxies_and_Reconstructions.pdf**
 pg. 21) commencing at the beginning of 2009. Although solar activity (as indicted by sunspot numbers) was less in cycle 23 than in cycles 21 and 22, the total irradiance amplitude (near 0.1%) is similar in the three past cycles since it is the net effect of sunspot darkening and facular brightening, both of which are altered by solar activity. Since ENSO fluctuations and volcanic eruptions are not predictable on decadal time scales, we estimate their maximum likely future impact with a scenario that includes a Pinatubo-like eruption with peak impact in 2014 and a super ENSO with maximum impact in 2019, mimicking a similar sequence that occurred from 1992 to 1997 (Figure Stratospheric Optical Depths).

The Data: Temperature, ENSO Index, Volcanic Aerosols, Anthropogenic (CO2 ppm)

HadCrut Temperature and CO2 ppm

http://www.cru.uea.ac.uk/cru/data/temperature/hadcrut3vgl.txt   Monthly Temp Data 1850 to 2009

http://www.cru.uea.ac.uk/cru/data/temperature/HadCRUT4-gl.dat   Monthly Temp Data 1850 to May 2013

Read data from http://www.esrl.noaa.gov/gmd/ccgg/trends/co2_data_mlo.

MLCO2 Columns: Year, Month, Decimal Year, Average, Interpolated, Trend

$MLCO2 := READPRN("co2\_mm\_mlo6-15.txt")$

$$MLCO2_{262,2} = 1980.042$$

$Date := MLCO2^{\langle 2 \rangle}$ $CO2_{ML} := MLCO2^{\langle 4 \rangle}$ $TrendCO2 := MLCO2^{\langle 5 \rangle}$ $RD := rows(Date)$

Get CO2 Trend Line from 1990 to 2015, then project to 2020

$Date_{1990} := submatrix(Date, 383, RD - 1, 0, 0)$ $L_{co2} := line(Date_{1990}, submatrix(TrendCO2, 383, RD - 1, 0, 0))$

$Trend_{co2}(Year) := L_{co2_0} + L_{co2_1} \cdot Year$ $Co := 340$ $m := 0..12 \cdot 10$ $Yr2020_m := 2010 + \dfrac{m}{12}$

$$Keeling(yr) := 1.054 \cdot 10^{-2}(yr - 1960)^2 + 9 \cdot 10^{-1} \cdot (yr - 1960) + 315.5$$

$SipleCO2 := READPRN("Friedli Siple CO2 1986.TXT")$ $IceCO2 := READPRN("CO2 Ice Core Data.txt")$

http://www.wasserplanet.becsoft.de/180CO2/CO2tot1812-2007.txt
Column C: CO2 total 1812-1961 corrected. annual averages from raw data.
   HaCrutE4 is HadCrut4 with the repeated rows deleted

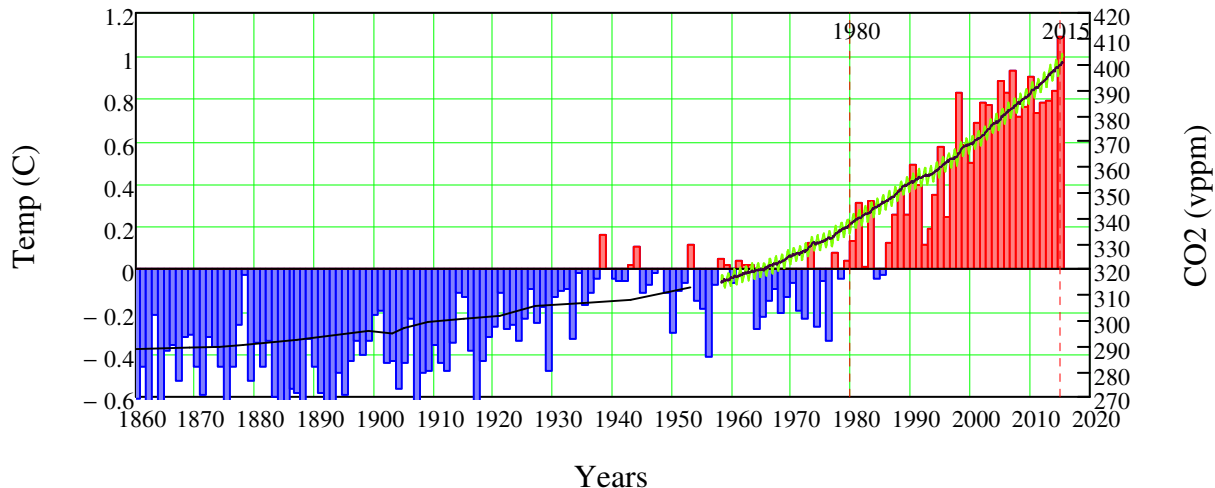$HadCrutz := READPRN("hadcrut3vgl.txt")$ $rows(HadCrutz) = 320$ $cols(HadCrutz) = 14$ $nz := 0..159$

$HadCrutE4 := READPRN("HadCrutE4-2015.txt")$ $rows(HadCrutE4) = 165$ $cols(HadCrutE4) = 14$

$HadCrut := READPRN("HadCrutE4-2015.txt")$ $Time_{crut} := HadCrut^{\langle 0 \rangle}$

$$TCrutz_{nz} := \sum_{m=1}^{12}\left(HadCrutz_{2 \cdot nz, m} \cdot \frac{1}{12}\right) \quad TCrut := HadCrut^{\langle 13 \rangle} \quad TCrutPlus := \overrightarrow{(\Phi(TCrut) \cdot TCrut)}$$

Note: For this plot, years 1980 - 2014 show a solid plus warming.

Global Temp (Bars) & CO2 Levels (Black Line) vs. Time

$n := 0..159$

ENSO Index Monthly Data from 1950 to 2015

http://www.esrl.noaa.gov/psd/people/klaus.wolter/MEI/mei.html Monthly ENSO Data 1850 to 2009

http://www.esrl.noaa.gov/psd/data/climateindices/list/   Month

$MEIM := READPRN("censo.data-2015.txt")$ $\qquad R := rows(MEIM)$ $\quad rr := 0..(R-1)\cdot 12 + 11$

$MEIx := submatrix(MEIM, 0, R-1, 1, 12)$ $\qquad\qquad MEIM_{32,0} = 1980$

$MEI_{rr} := MEIx_{floor\left(\frac{rr}{12}\right), mod(rr,12)}$ $\qquad MEID_{rr} := \left(MEIM^{\langle 0 \rangle}\right)_{floor\left(\frac{rr}{12}\right)} + \frac{mod(rr,12)}{12}$
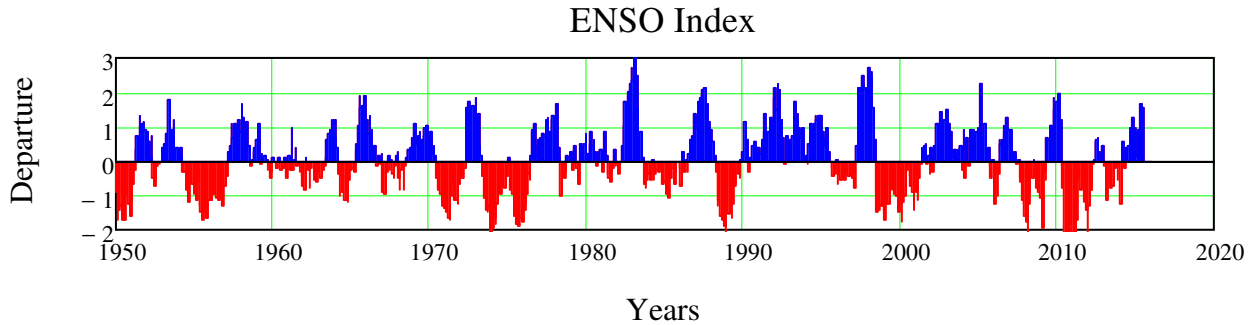
ENSO Index

Pacific Decadal Oscillation (PDOI) Index Monthly from 1900 to 2015
        Note: ENSO and PDO are not statistically independent. They have a 47% correlation.

http://jisao.washington.edu/data/pdo/   Year, Jan to Dec

$PDO := READPRN("PDO\_Index.dat")$

$PDO := READPRN("pdo.data-2015.txt")$ $\qquad R := rows(PDO) = 68$ $\quad rr := 0..(R-1)\cdot 12 + 11$

$PDOI_{rr} := PDOx_{floor\left(\frac{rr}{12}\right), mod(rr,12)}$ $\qquad PDOYr_{rr} := \left(PDO^{\langle 0 \rangle}\right)_{floor\left(\frac{rr}{12}\right)} + \frac{mod(rr,12)}{12}$

$PDO_{80,0} = \blacksquare$

PDO Index
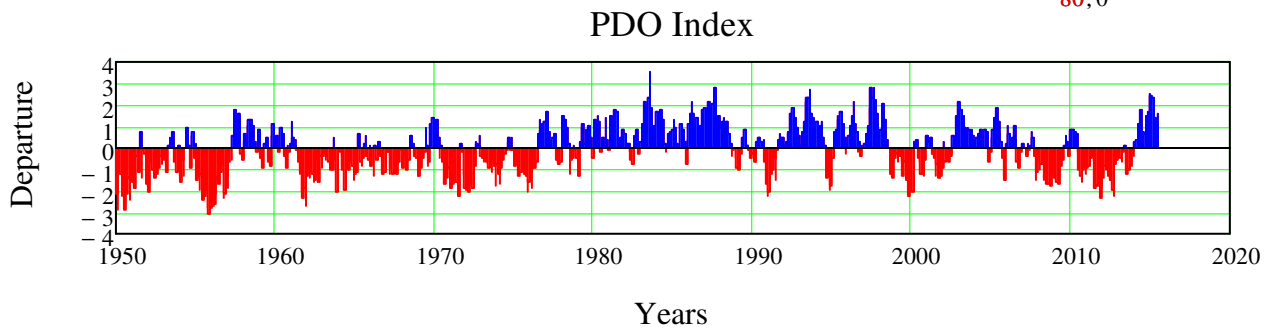
Stratospheric Volcanic Aerosols
        Note: Volcanic Aerosols and ENSO are not statistically independent. Corr = 40%
        http://data.giss.nasa.gov/modelforce/strataer/   Data: Global, NH, SH

$VA := READPRN("Aerosols-Monthly-Mean Optical Thickness\_tau\_line-2012-2015-Repeated.txt")$

$VA_{1560,0} = 1980.042$

Stratospheric Aerosol Optical Depths - Sato
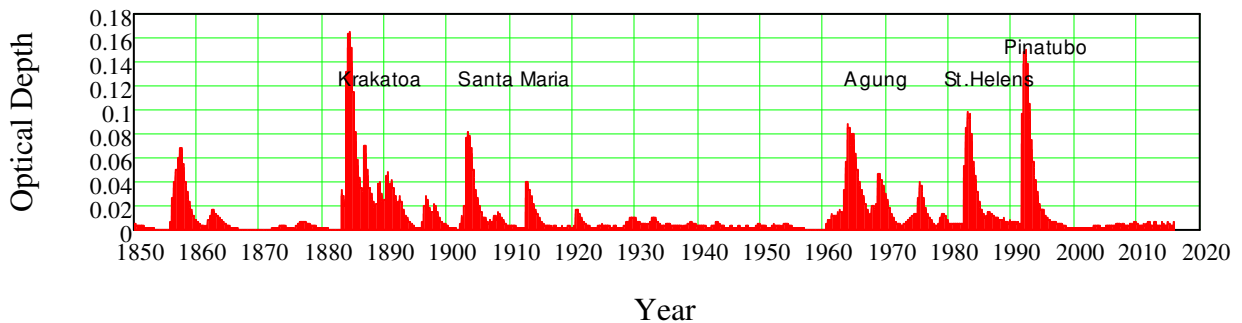
PMOD Satelite Daily Solar Irradiance (from 1980) and IPCC AR5 Monthly Reconstruction from 1610

$\text{TSI}_{\text{PMOD}} := \text{READPRN}(\text{"TSI.PMODF\_15.TXT"})$

$\text{TSI}_{\text{PMOD}_{410,0}} = 800101$

TAM, Convert Daily to Monthly Average

TSI 2014 is Yearly

$\text{TSIMA06} := \text{READPRN}(\text{"TSIMonthAvg-2006Data.txt"})$

$\text{TSI}_{\text{smm}} := \text{READPRN}(\text{"TSIsm8.txt"})$

$\text{TSIMon} := \text{DMAvg}(\text{TSI}_{\text{PMOD}})$

PMOD Total Solar Irradiance (Red), Smoothed (Yellow), Monthly (Blue)
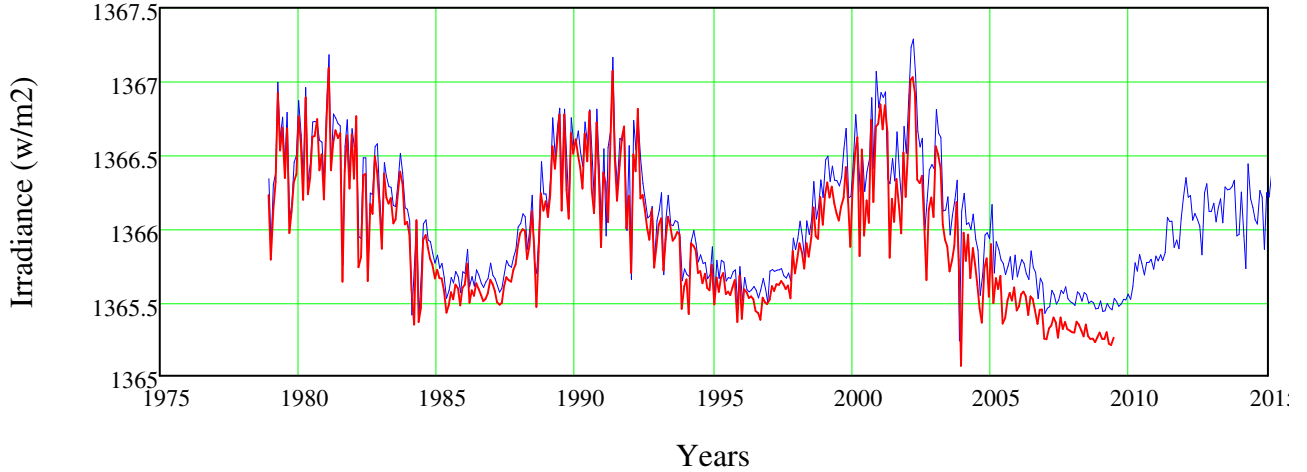
Reconstruction of Solar Irradiance since 1610, Lean 1995 (1600-1995
ftp://ftp.ncdc.noaa.gov/pub/data/paleo/contributions_by_author/lean1995/

$\text{TSD}_{\text{lean}} := \text{READPRN}(\text{"TSDLeanFilled.txt"})$

$\text{TSDF}_{\text{lean}} := \text{READPRN}(\text{"lean1995data.txt"})$

Solar Irradiance Correlates with U.S. Temp Anomaly

To even comparison of PMOD (Green Dash) and Lean Reconstructed on same graph, add 1.5 to PMOD

Lean Reconstructed Solar Irradiance (Red), PMOD +1.5 (Green), and Temp Anomaly (Blue)
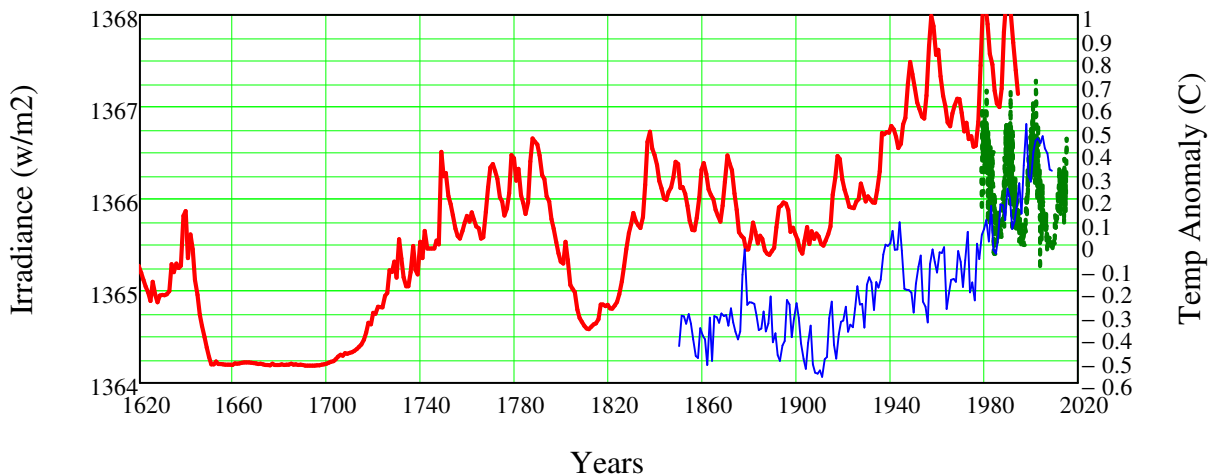
Zonal Variation of Irradiance

http://data.giss.nasa.gov/cgi-bin/cdrar/effij.py
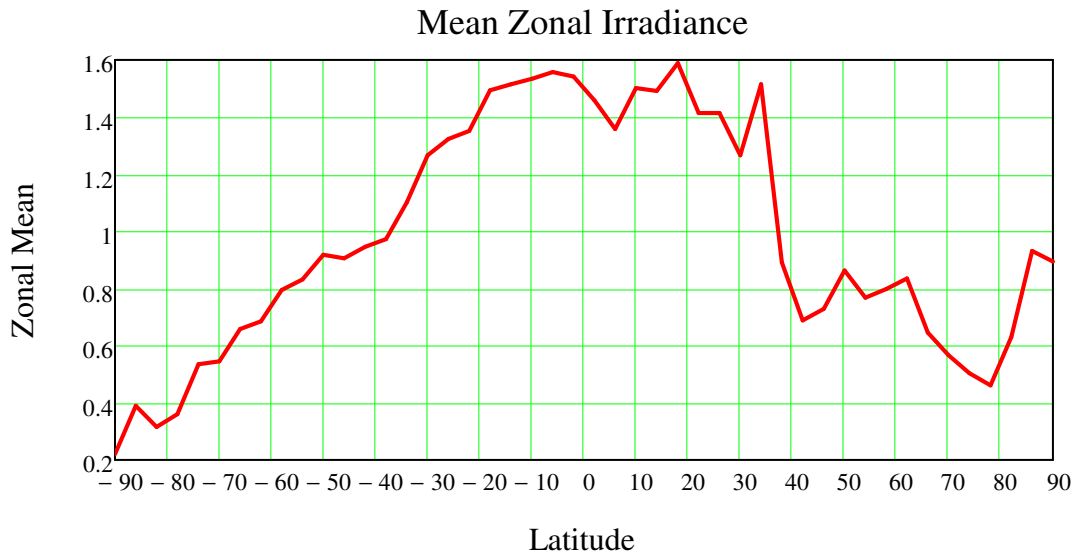
$IrradZonal := READPRN("Irradiance zonal01.dat")$

Mean Zonal Irradiance

Monthly Time Series Matrices with Optimized Lags for Period 1980 to 2006

Let Y be the Temp and X1, X2, X3, and X4 be the delayed matrices for ENSO, Irradiance, Volcanic Aerosols, and Anthropogenic Influence.

HadCrut Temperature Data 2015: HCrutE4

$HadCrut_{129,0} = 1980$ $2015 - 1980 = 35$ $YX := 35$ $r25 := 0 .. YX$ $YrEven_{r25} := 129 + r25$ $n := 0 .. 12$

$HadCrutDat_{r25,n} := HadCrut_{129+r25,n}$ $HadCrutMDat$ $R := rows(HadCrutDat)$ $rr := 0 .. (R - 1) \cdot 12 + 11$

$HCTime_{rr} := \left(HadCrutDat^{\langle 0 \rangle}\right)_{floor\left(\frac{rr}{12}\right)} + \frac{mod(rr, 12)}{12}$ $HCTempX_{rr} := HadCrutDat_{floor\left(\frac{rr}{12}\right), mod(rr, 12)+1}$
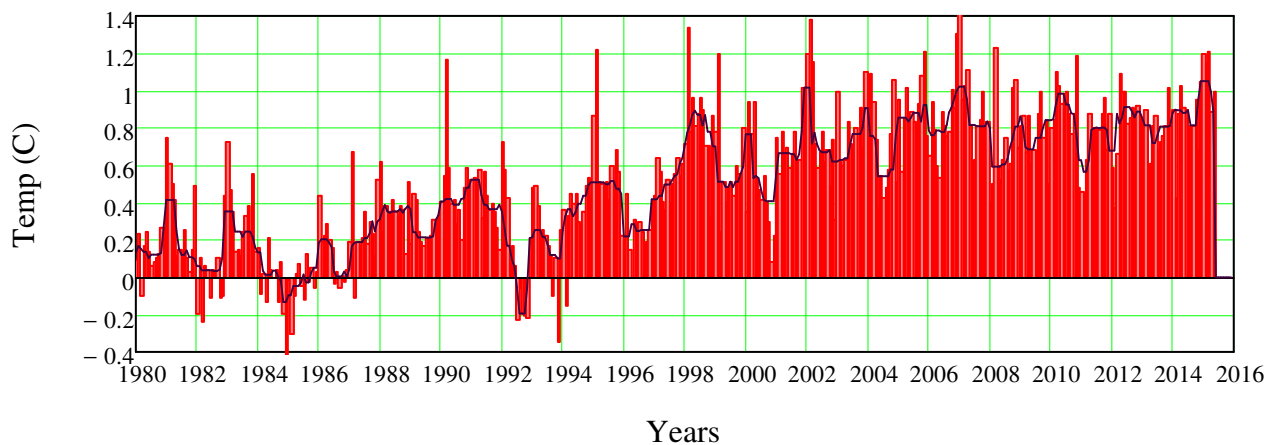
Smooth the Temp Data for Regression $HCTemp := medsmooth(HCTempX, 7)$ $rows(HCTemp) = 432$

$Y := submatrix(HCTemp, 0, 12 \cdot YX, 0, 0)$ $Year_Y := submatrix(MEID, 384, 384 + 12 \cdot YX, 0, 0)$ $MEID_{384} = 1980$

HadCrut Monthly Global Temp (Bars) vs. Time

Empirical Component Data for 1980 to 2006, X1, X2, X3, X4, Given Optimized Lags (Months), Δt:
The Optimization Procedure to Find the Optimized Delays (Months): 3, 6, 0, 17*12, follows on pg. 7

Lag Operator Δt(X, NStart, #Years,τ): $\Delta t\left(X, N_{Start}, Yrs, \tau\right) := submatrix\left(X, N_{Start} - \tau, N_{Start} - \tau + 12 \cdot Yrs, 0, 0\right)$

Define Forcing Parameters: X are normal forcings, Xs are Gasussian Smoothed

X1 = ENSO (E): 1980 through 2005 ΔtE = 7, $\quad x1 := \Delta t(MEI, 384, YX, 7)\quad MEID_{384} = 1980 \quad rows(x1) = 42$

$\mu1 := mean(x1) \qquad \sigma1 := stderr\left(Year_Y, x1\right) \qquad z1 := (x1 - \mu1) \cdot \sigma1^{-1} \qquad xs1 := ksmooth\left(Year_Y, z1, 12\right)$

X2 = Monthly Volcanic Aerosols (V), τV = 8, $\qquad\qquad\qquad VA_{1560,0} = 1980.042$

$\underset{\wedge\wedge\wedge}{m} := 0 .. 100 \qquad Zeros_m := 0 \qquad VAZ := stack\left(VA^{\langle1\rangle}, Zeros\right) \qquad x2 := \Delta t(VAZ, 1560, YX, 8)$

$\mu2 := mean(x2) \qquad \sigma2 := stderr\left(Year_Y, x2\right) \qquad z2 := (x2 - \mu2) \cdot \sigma2^{-1} \qquad xs2 := ksmooth\left(Year_Y, z2, 2\right)$

X3 = Solar Irradiance (S), τS = 2

$x3 := \Delta t\left(TSI_{PMOD}^{\langle2\rangle}, 13, YX, 4\right) \quad \mu3 := mean(x3) \qquad \sigma3 := stderr\left(Year_Y, x3\right) \qquad z3 := (x3 - \mu3) \cdot \sigma3^{-1}$
$xs3 := ksmooth\left(Year_Y, z3, 2\right)$

X4 = Anthropogenic Forcing, ΔF, of CO2 ppm (A), τA = 10*17 (17 Yr Delay)

Assume total radiative forcing (includes CH4) is proportional to radiative forcing due to carbon dioxide.

$\Delta F(C) := 4.841 \ln\left(\dfrac{C}{Co}\right) + 0.0906 \cdot \left(\sqrt{C} - \sqrt{Co}\right) \qquad \Delta F_{IPCC}(C) := 6.3 \ln\left(\dfrac{C}{Co}\right)$

$CO2_{ppm} := \Delta t(TrendCO2, 261, YX, 17 \cdot 12) \quad x4 := \overrightarrow{\Delta F_{IPCC}\left(CO2_{ppm}\right)} \qquad Date_{262} = 1980.042$

$\mu4 := mean(x4) \qquad \sigma4 := stderr\left(Year_Y, x4\right) \qquad z4 := (x4 - \mu4) \cdot \sigma4^{-1} \qquad xs4 := ksmooth\left(Year_Y, z4, 2\right)$

X5 = Pacific Decadal Oscillation (PDO) Index (PDO), τD = 10*17)

$x5 := \Delta t(PDOI, 384, YX,) \quad \mu5 := mean(x5) \qquad \sigma5 := stderr\left(Year_Y, x5\right) \qquad z5 := (x5 - \mu5) \cdot \sigma5^{-1}$

Multi-Variate Component Construction and Design Matrix (Optimized Lags):

$$Y_i = \beta_0 + \beta_1 \cdot X1_i + \beta_2 \cdot X2_i + \beta_3 \cdot X3_i + \beta_4 \cdot X4_i + \varepsilon_i$$

The design matrix for our Temp Stats data can be constructed with the components $\qquad i := 0 .. 313 - 1$

	ENSO	**Volcanic Aerosols**	**Solar**	**Anthro - Effects CO2**

$ONE_i := 1$

$X^{\langle0\rangle} := ONE \qquad X^{\langle1\rangle} := z1 \qquad X^{\langle2\rangle} := z2 \qquad X^{\langle3\rangle} := z3 \qquad X^{\langle4\rangle} := z4$

$Xs^{\langle0\rangle} := ONE \qquad Xs^{\langle1\rangle} := xs1 \qquad Xs^{\langle2\rangle} := xs2 \qquad Xs^{\langle3\rangle} := xs3 \qquad Xs^{\langle4\rangle} := xs4$

$b := \left(X^T \cdot X\right)^{-1} \cdot \left(X^T \cdot Y\right) \qquad\qquad b^T = (0.58935 \quad 0.0421 \quad -0.0605 \quad 0.00848 \quad 0.03128)$

$bs := \left(Xs^T \cdot Xs\right)^{-1} \cdot \left(Xs^T \cdot Y\right) \qquad bs^T = (0.62585 \quad -0.96799 \quad -0.05185 \quad 0.02604 \quad 0.01722)$

Save Model β Coefficient Results: $\qquad WRITEPRN("Emp\_ESVA\_Coefficients.txt") := b$

$\Delta T := b_0 + b_1 \cdot X^{\langle1\rangle} + b_2 \cdot X^{\langle2\rangle} + b_3 \cdot X^{\langle3\rangle} + b_4 \cdot X^{\langle4\rangle} \quad \Delta Ts := bs_0 + bs_1 \cdot xs1 + bs_2 \cdot xs2 + bs_3 \cdot xs3 + bs_4 \cdot xs4$

$corr(Y, \Delta T) = 0.9038 \qquad RSquare := corr(Y, \Delta T)^2 = 0.81685 \qquad\qquad corr(Y, \Delta Ts)^2 = 0.72544$

Optimization Procedure and Results:

Determine Time Lags to Maximize Correlation (R^2) of Regression Model to Global Temperature
Evaluate lags of 0 to 12 Months for ENSO, Aerosols, and Irradiance and 5 to 20 yrs for Effects of CO2 ppm

$$
\text{OptLags(Y)} := \left| \begin{array}{l}
X^{\langle 0 \rangle} \leftarrow \text{ONE} \\
\text{col} \leftarrow \text{MaxCorr} \leftarrow 0 \\
\text{for } \text{te} \in 5..12 \\
\quad \text{for } \text{ta} \in 0..12 \\
\qquad \text{for } \text{ti} \in 0..5 \\
\qquad\quad \text{for } \text{tc}_{\text{yrs}} \in 15..21 \\
\qquad\qquad \left| \begin{array}{l}
X^{\langle 1 \rangle} \leftarrow \Delta t(\text{MEI}, 384, \text{YX}, \text{te}) \\
X^{\langle 2 \rangle} \leftarrow \Delta t(\text{VAZ}, 1560, \text{YX}, \text{ta}) \\
X^{\langle 3 \rangle} \leftarrow \Delta t\left(\text{TSI}_{\text{PMOD}}^{\langle 2 \rangle}, 13, \text{YX}, \text{ti}\right) \\
\text{co2ppm} \leftarrow \Delta t\left(\text{TrendCO2}, 261, \text{YX}, \text{tc}_{\text{yrs}} \cdot 12\right) \\
X^{\langle 4 \rangle} \leftarrow \overrightarrow{\Delta F_{\text{IPCC}}(\text{co2ppm})} \\
b \leftarrow \left(X^T \cdot X\right)^{-1} \cdot \left(X^T \cdot Y\right) \\
\Delta T \leftarrow b_0 + b_1 \cdot X^{\langle 1 \rangle} + b_2 \cdot X^{\langle 2 \rangle} + b_3 \cdot X^{\langle 3 \rangle} + b_4 \cdot X^{\langle 4 \rangle} \\
\text{Corr} \leftarrow \text{corr}(Y, \Delta T)^2 \\
\text{if } \text{Corr} > \text{MaxCorr} \\
\quad \left| \begin{array}{l} \text{Opt}_{1,0} \leftarrow \text{col} \\ \text{MaxCorr} \leftarrow \text{Corr} \end{array} \right. \\
\text{Opt}^{\langle \text{col} \rangle} \leftarrow \left(\text{Corr} \quad \text{te} \quad \text{ta} \quad \text{ti} \quad \text{tc}_{\text{yrs}}\right)^T \\
\text{col} \leftarrow \text{col} + 1
\end{array} \right. \\
\text{Opt}
\end{array} \right.
$$

Run Opt Routine and Print Opt Data

$$\text{OptDat} := \text{OptLags(Y)}$$

Optimum: R^2 **Δt$_E$ Δt$_V$ Δt$_S$ Δt$_A$**

$$\text{OptDat}^{\left\langle \text{OptDat}_{1,0} \right\rangle T} = (0.8483 \quad 7 \quad 8 \quad 4 \quad 17)$$

<div style="background:yellow">

Optimization Results
Optimum E, V, S, A Lags:
3, 6, 0, 17 (yrs), respectively.
R^2 = 0.85

</div>

The regression and estimation results are:
The numbers shown in parentheses below the regression coefficients are the magnitudes of their t-ratios; i.e. the coefficients divided by the standard deviation of the regression coefficient. All but the coefficient for Sunspot number are significantly different from zero at the 95 percent level of confidence and they are of the right sign.

Shown below is a comparison of the observed temperature change and the temperature change predicted by the regression equation. The observations are shown in red and the estimations from the regression equation are shown in blue.

Another way of viewing the comparison is in the scatter diagram below of the actual and regression predicted temperature changes.

The t-ratios for the variables included in the regression equation are significant. They **explain 76 percent of the variation in the year-to-year temperature change**. The insolation and CO2 ppm both a 76% correlation. Also the effect of the CO2 in the equation includes the effects of all variables influencing temperature change which are correlated with the general trend on CO2 concentration but are not in the equation. These would i**nclude the effects of anthropogenic water vapor and anthropogenic cloudiness.**
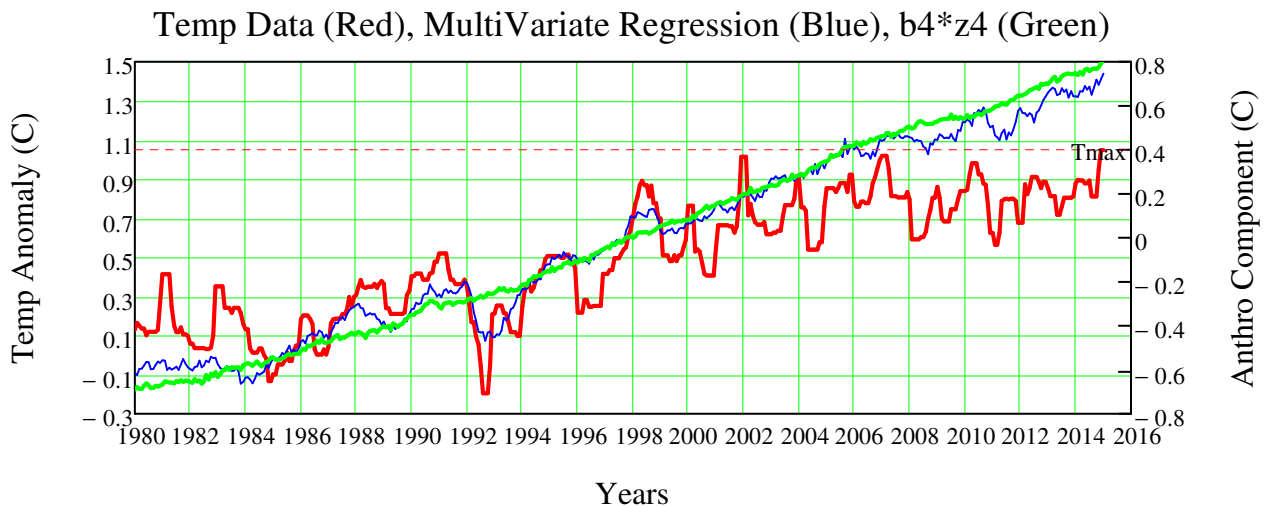
Compare Anthropogenic Forcing Component - (17 Year Lag) = b4*z4 (Green) of ΔT to Data
b4*z4 is the Optimized Match of Effects of Antro Forcing to Global Temperature Data

Note: The Global Temp data (red) flattens out, while the regression (blue) and CO2 levels rise linearly.

$$\text{Tmax} := \max(Y) = 1.05$$

IPCC CO_2 Log Forcing Function used in Regression Model. Co = 340 ppm.

$$Co := 340 \qquad \Delta F_{IPCC}(C) := 6.3 \ln\left(\frac{C}{Co}\right)$$

Temp Data (Red), MultiVariate Regression (Blue), b4*z4 (Green)
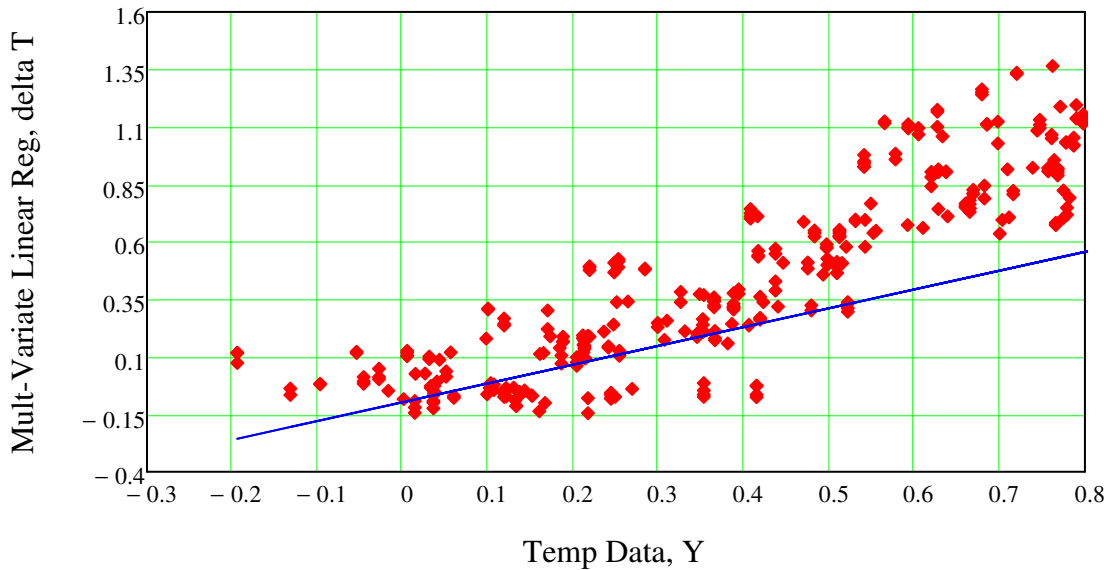

Temp Anomaly (C) — Years — Anthro Component (C)

Statistical Analysis:

Four Factors - ENSO, Volcanic Aerosols, Insolation, and CO2 ppm explain 3/4 (76%) of the temperature variation.

$$\text{int} := \text{intercept}(Y, \Delta T) \qquad s := \text{RSquare}$$

$$\text{int} = -0.09626 \qquad \boxed{s = 0.81685} \qquad \delta T := \text{int} + s \cdot Y$$

Scatter Plot: Linear Regression delta T vs Temp Data, Y

Test for Possible Regression

By extending this test to include p slope parameters

H_0: $\beta_1 = \beta_2 = \beta_3 = ... = \beta_p = 0$

we have the equivalent test for the possibility of a multiple regression,

H_0: no multiple regression relationship

As in simple linear regression, we can associate each data value with three types of deviations, specifically, the residual error, e

$$\text{yhat} := X \cdot b \qquad e := Y - \text{yhat}$$

Sum of Squares

We can also use matrices to calculate the sum of squares for residual error,

$$\text{SSE} := e^T \cdot e \qquad \text{SSE} = 15.99085$$

as well as for regression,

$$\text{SSR} := \text{yhat} - \text{mean}(Y)^T \cdot \left(\text{yhat} - \text{mean}(Y)\right)$$

The total sum of squares equals

$$n := \text{rows}(Y) \qquad p := 4$$

$$\text{SST} := \text{SSE} + \text{SSR} \qquad \text{DF\_REG} := p \qquad \text{DF\_RESID} := n - (p + 1) \qquad \text{DF\_RESID} = 416$$

Mean Squares

Again, as in simple linear regression, dividing each sum of squares by the corresponding degrees of freedom provides us with variance estimates. The mean square for residual error

$$\text{MSE} := \frac{\text{SSE}}{\text{DF\_RESID}} \qquad \text{MSR} := \frac{\text{SSR}}{\text{DF\_REG}} \qquad \text{DF\_TOTAL} := n - 1$$

F Test

The final entry in the table is the F statistic and corresponding p-value for the significance of an overall multiple regression. Under the null hypothesis of

H_0: no regression relationship

the test statistic $F := \dfrac{MSR}{MSE}$ $\qquad Rsq := \dfrac{SSR}{SST}$ $\qquad Rsq = 0.73748$

has an F distribution with $\qquad n1 := DF\_REG \qquad n2 := DF\_RESID$

degrees of freedom. The p-value of the test, then, is given by
$p\_val := 1 - pF(F, n1, n2) \qquad p\_val = 0$

Summary: Analysis of Variance Table
Summarizing the above for our example,

DF	SS	MS	F
DF_REG = 4	SSR = 44.92168	MSR = 11.23042	F = 292.15792
DF_RESID = 416	SSE = 15.99085	MSE = 0.03844	p-value
DF_TOTAL = 420	SST = 60.91254		p_val = 0

The amount of variability explained by the linear regression (MSR) is greater than the amount due to residual error (MSE). The difference is large enough (the p-value is, in fact, close to 0) to strongly reject the null hypothesis,

Correlations between each pair of variables
in the model can be displayed in matrix form as

$$FLEX := augment\left(X^{\langle 1 \rangle}, X^{\langle 2 \rangle}, X^{\langle 3 \rangle}, X^{\langle 4 \rangle}, Y\right)$$

$$j := 0 .. p \qquad k := 0 .. p$$

$$CORR_{j,k} := corr\left(FLEX^{\langle j \rangle}, FLEX^{\langle k \rangle}\right)$$

STATISTICAL CORRELATIONS TO COMPONENTS
Correlation between "independent" variables
35% correlation between x1 (ENSO) and x2 (Aerosols)

Correlation to Global Temp between:
Y (ΔTemp) and x1 (ENSO)   = - 3.5%
Y (ΔTemp) and x2 (Volcanic Aero) = - 49%
Y (ΔTemp) and x3 (Solar)   = - 8.5%
Y (ΔTemp) and x4 (Anthropogenic) = 87%

$$CORR = \begin{pmatrix} 1 & 0.35061 & -0.11702 & -0.23532 & -0.05022 \\ 0.35061 & 1 & 0.01393 & -0.3596 & -0.48887 \\ -0.11702 & 0.01393 & 1 & 0.12867 & 0.08534 \\ -0.23532 & -0.3596 & 0.12867 & 1 & 0.87369 \\ -0.05022 & -0.48887 & 0.08534 & 0.87369 & 1 \end{pmatrix} \begin{matrix} x1 \\ x2 \\ x3 \\ x4 \\ y \end{matrix} \begin{pmatrix} ENSO \\ Aerosols \\ Solar \\ CO2 \\ \Delta Temp \end{pmatrix}$$

$$\text{Var\_Covar\_b} := \left(X^T \cdot X\right)^{-1} \cdot \text{MSE}$$

$$k := 0 .. p$$

$$\text{se\_b}_k := \sqrt{\overrightarrow{\text{Var\_Covar\_b}_{k,k}}}$$

$$\text{se\_b}^T = (\,0.01208 \quad 0.01011 \quad 0.01015 \quad 0.00964 \quad 0.00078\,)$$

t tests

$$t := \frac{\overrightarrow{b}}{\text{se\_b}}$$

$$t^T = (\,48.80282 \quad 4.16426 \quad -5.9618 \quad 0.87993 \quad 40.34912\,)$$

2010 to 2020 Climate Forecasting:

Empirical Component Data and Forecast: X1, X2, X3, X4

X1 = ENSO: 1980 through 2010, then mimic 1992 to 1997 twice, Δtε = 3

X2 = Monthly Volcanic Aerosols, repeat Pinatubo eruption with a peak in 2014, Δtᵥ = 6,

X3 = Irradiance: 1980 through 2009, then replicate cycle 23, Δts = 0:

X4 = Anthropogenic, Δtₐ = 17*12 (17 Yr Delay), then maintain trend.

Multi-Variate Component Forecast Model (βs Determined from Previous Regression):

$$Y_i = \beta_0 + \beta_1 \cdot X1_i + \beta_2 \cdot X2_i + \beta_3 \cdot X3_i + \beta_4 \cdot X4_i + \varepsilon_i$$

The design matrix for our Temp Stats data can be constructed with the statements

$$X^{\langle 1 \rangle} := x1 \qquad X^{\langle 2 \rangle} := x2 \qquad X^{\langle 3 \rangle} := x3 \qquad X^{\langle 4 \rangle} := x4$$

$$b := \text{READPRN}(\text{"Emp\_ESVA\_Coefficients.txt"})$$

$$\Delta T_{\text{forecast}} := b_0 + b_1 \cdot x1 + b_2 \cdot x2 + b_3 \cdot x3 + b_4 \cdot x4$$

The regression and estimation forecast:

Empirical Multi-Variate Regression Model Forecast: Global Temp Anomaly

Years