CSE 397: Validation & UQ R. Moser Group Problem 3

Shreyas Gaikwad Shane McQuarrie Caleb Phillips Bassel Saleh

This problem is intended to illustrate more of the challenges of real-world imperfect models and their calibration against real-world data with its challenging nuances. Here we consider Climate Change—a much more complex real-world modeling problem than the dropped spheres—which we model with a global *Energy Balance Model* (EBM). Such models, which seem to be the first introduced in efforts to model global climate, have been superseded (or augmented) far beyond their humble origins of at least 50 years ago. Still, they offer a viable starting point for our objectives.

As with most models, this is constructed from a highly reliable theory: global conservation of energy, considering the solar energy incident on earth and that radiated away. However, this reliability is compromised almost from the outset by embedded models, which are obviously necessary because there are no full-physics, first-principles representations available for so complex a system as our entire planet.

The variant we emply here is based on the simplified panetary view in Figure 1. The incident solar radiation $S = 1370\text{W/m}^2$ heats the projected area of the earth (radius $R = 6.38 \times 10^3\text{km}$). Some fraction of this is relected, as measured by the effective global albedo $\alpha \in [0, 1]$. What is not reflected is balanced by the net radiation P from the planet. Hence, the conservation expression is

$$(1 - \alpha)S\pi R^2 = P. \tag{1}$$

At this stage, α is the only model parameter. Given the well-known inhomogeneity of the earth's surface, α is obviously some sort of average of a very complex net reflectivity. Otherwise, the model is exact (energy is conserved). To be useful, we need to understand how P depends upon the earth's temperature T. This too is rather complex.

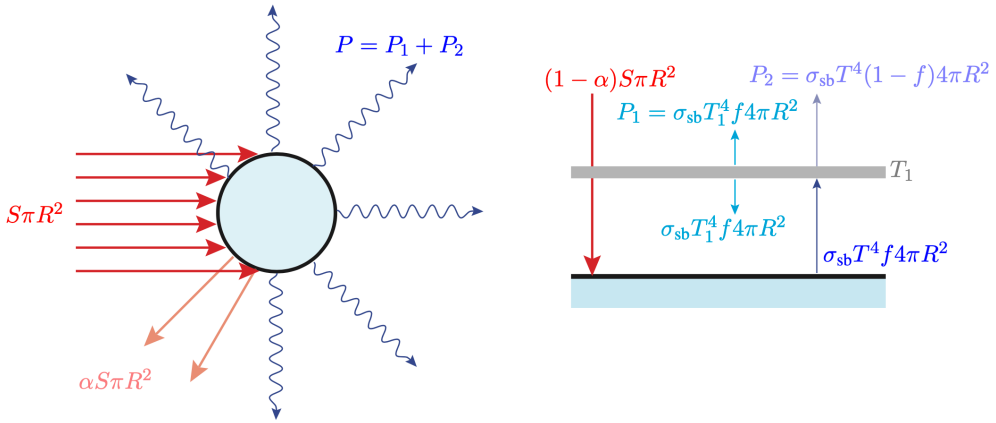


Figure 1: Schematic.

The greenhouse effect arises because the solar radiation is short wavelength (because of its high temperature), in a frequency range for which the atmosphere is relatively transparent, so much of the energy reaches the surface of the planet. However, the radiation from the much cooler earth at temperature T is in a portion of the spectrum for which the atmosphere can be relatively opaque (see Figure 1). This opacity depends on the concentration of greenhouse gases in the atmosphere. Both CO_2 and CH_4 can, in low concentrations, absorb significant amounts of this transmitted radiation, blocking this "atmospheric window" (see Figure 2). Note that H_2O is also a greenhouse gas. Thus, the radiation from the surface $\sigma_{\rm sb}T^4$, where $\sigma_{\rm sb} = 5.67 \times 10^{-8}{\rm W/m^2K^4}$ is the Stephan-Boltzmann constant, is somewhat absorbed by the atmosphere, with non-dimensional opacity f.

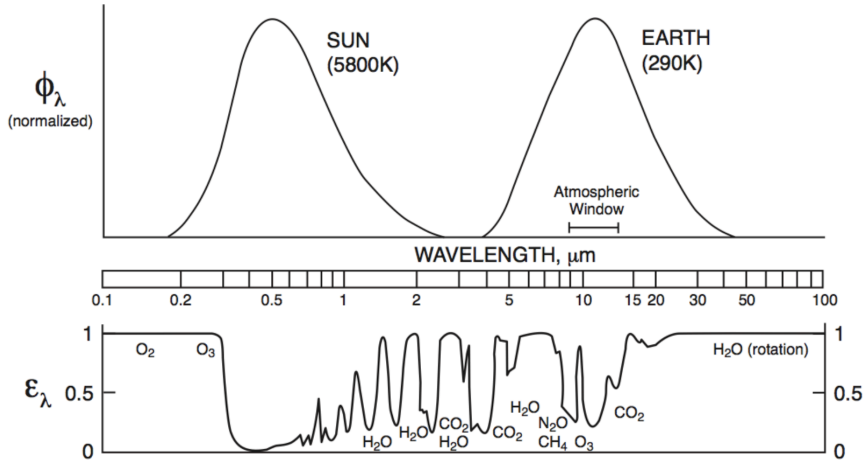


Figure 2: Atmospheric absorption from Jacob (1999).

In this very simple model, we assume that the atmosphere is at a single temperature T_1. We also assume that it radiates equally outward and inward, each with energy flux $f f \sigma_{\rm sb} T_1^4$. Thus,

$$P = 4\pi R^2 \left(f\sigma_{\rm sb} T_1^4 + (1-f)\sigma_{\rm sb} T^4 \right). \tag{2}$$

The atmosphere is also assumed to be in energy equilibrium

$$2f\sigma_{\rm sb} T_1^4 = f\sigma_{\rm sb} T^4, \tag{3}$$

so $T_1^4 = T^4/2$. Using this in (2) to eliminate T_1 and substituting into (1) yields a model for the earth's nominal surface temperature:

$$T = \left[\frac{\frac{S}{4}(1-\alpha)}{\sigma_{\rm sb}\left(1 - \frac{f}{2}\right)} \right]^{1/4}. \tag{4}$$

We know that f depends on the concentration of CO_2 in the atmosphere. We assume the simplest possible linear model to describe this dependence:

$$f = f_0 + f_1 Y_{CO_2}. \tag{5}$$

Overall, this model should be viewed as a stand-in for more reasoned but far more complex climate models. Still, it is obviously grounded in the same laws expressed in any such model.

To summarize, our model has three parameters: the effective global albedo $\alpha \in [0, 1]$, and the linear model parameters $f_0 \in [0, 1)$ (since we know it must be absorbing, the actual absorption will increase with increasing Y_{CO_2}), and f_1. The final parameter f_1 depends on the units used to measure Y_{CO_2} (it is dimensional), though it is constrained such that $f \in [0, 1]$.

We have three data sets for calibrating our experiments, which we plot in Figure 3.

- `T-range.dat`: Global mean temperature from 1900–2014 in Kelvins.
- `CO2-mauna-loa.dat`: Mauna Loa atmospheric CO_2 levels in ppm for the years 1956–2014.
- `CO2-sipple-ice.dat`: Sipple Station ice core inferred CO2 levels in ppm for the years 1734–1983.

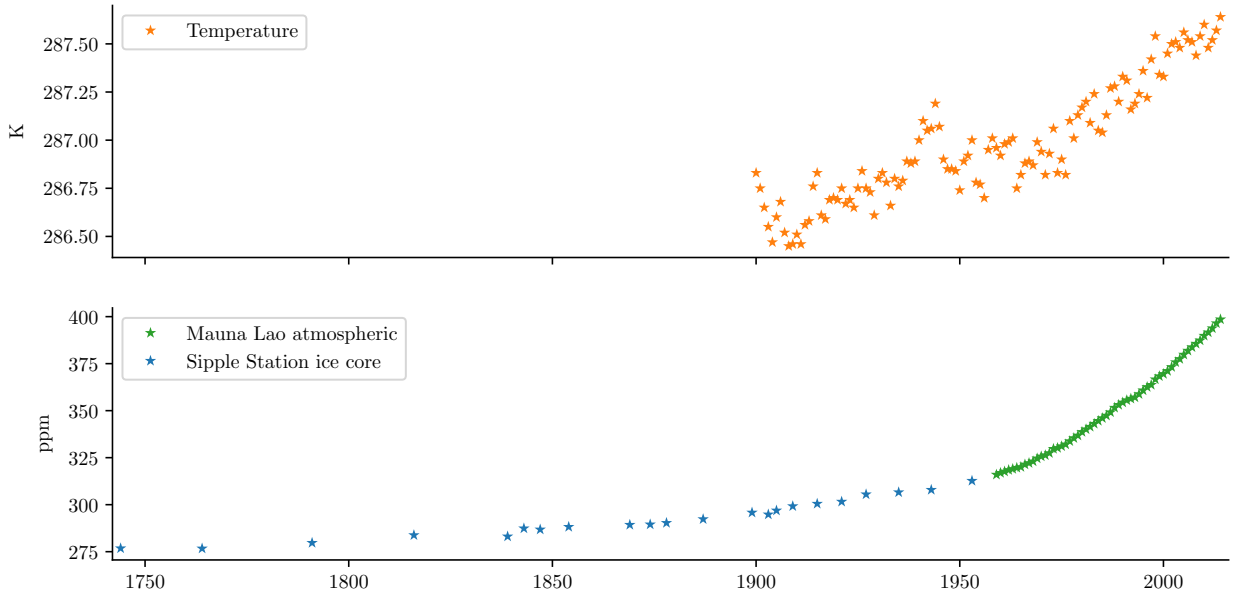


Figure 3: Raw data temperature and CO_2 data.

1. (a) Suppose the albedo of non-ice-covered land and ocean is $\alpha \approx 0.3$, and the transitivity of the atmosphere is $f \approx 0.6$. Using (4),

$$T = \left[\frac{\frac{S}{4}(1-\alpha)}{\sigma_{\mathrm{sb}}\left(1-\frac{f}{2}\right)}\right]^{1/4} \approx \left[\frac{\frac{S}{4}(1-0.3)}{\sigma_{\mathrm{sb}}\left(1-\frac{0.6}{2}\right)}\right]^{1/4} \approx 278.78\mathrm{K} = 42.13^{\circ}\mathrm{F}.$$

The mean temperature from the data is 286.98K = 56.89°F, so these choices for α and f aren't spot on, but they aren't astronomically bad either.

(b) For climate change, we are concerned with variations of only a degree or so, so it seems prudent to calibrate the unknown parameters against available data. To get an idea of the parameter values implied by the data, we perform a least-squares fit to the 1900–2014 data. Since the ice core data does not provide estimates for every year, we must either ignore the years with no CO_2 data, or compute estimates for the missing CO_2 levels. Since the ice core data follows a nearly linear trend, we do a linear regression on data after 1900, and use that linear model to fill in the missing data; see Figure 4.

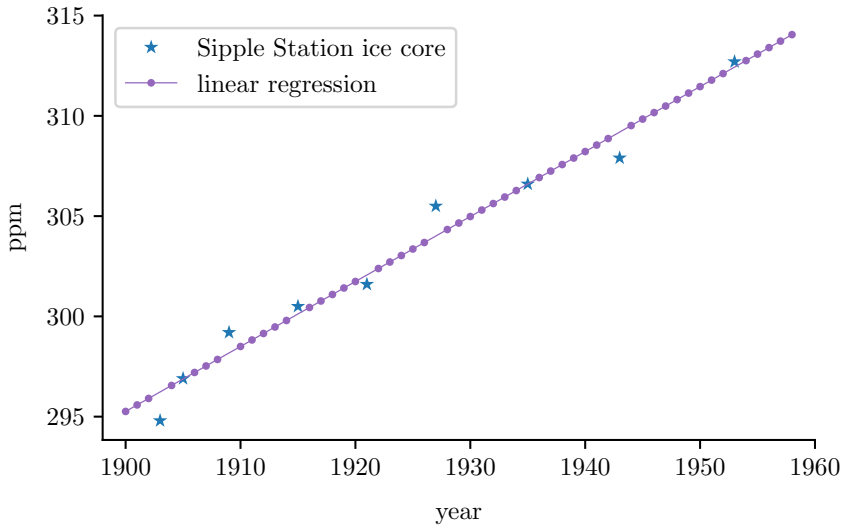


Figure 4: Using a linear regression to fill in missing CO_2 data.

Next, we want to compute the parameters α, f_0, and f_1 that minimize the sum of squares of the data misfits. For convenience, we write $Y = Y_{CO_2}$ and denote (T_k, Y_k) for a pair of temperature and CO_2 estimates. Then the least-squares problem is

$$\min_{\alpha,f_0,f_1} \sum_{i=k}^{N} \left| T_k - \left[\frac{\frac{S}{4}(1-\alpha)}{\sigma_{\text{sb}}\left(1 - \frac{f_0+f_1 Y_k}{2}\right)} \right]^{\frac{1}{4}} \right|^2 \qquad \text{s.t. } \alpha \in [0,1], f_0 \in [0,1). \tag{6}$$

We can quickly solve this problem with an out-of-the-box nonlinear least-squares solver, `scipy.optimize.least_squares()`, which gives the results $\alpha \approx 0.2618$, $f_0 \approx 0.6272$, and $f_1 \approx 0.0001769$. From our data, this gives a mean value of $f \approx 0.6851$, and these values of α and f predict a global mean temperature of 286.97K = 56.88°F (which, of course, matches our data).

(c) To calibrate the model (4) from data, we aim to determine the joint distribution of the scenario parameters,

$$p(\alpha, f_0, f_1 | T, Y, X) = \frac{p(T|Y, \alpha, f_0, f_1, X) p(\alpha, f_0, f_1 | Y, X)}{p(T|Y, X)}.$$

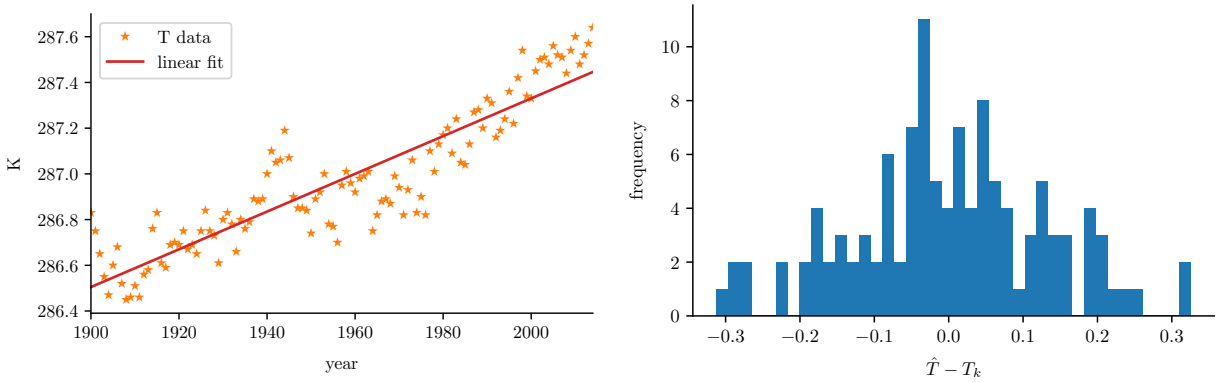
To compute the likelihood $p(T|Y, \alpha, f_0, f_1, X)$, we assume a Gaussian distribution for the difference between the model prediction and the observation. That is, given an observation Y_k and values of α, f_0, and f_1, we calculate $\overline{T}_k = T(Y_k, \alpha, f_0, f_1)$ from (4) and assume $\delta T_k = (T - \overline{T}_k) \sim \mathcal{N}(0, \sigma^2)$ (we will specify σ shortly). Then $T \sim \mathcal{N}(\overline{T}_k, \sigma^2)$, so the evaluation of the likelihood for a pair of observations (T_k, Y_k) is given by

$$p(T_k|Y_k, \alpha, f_0, f_1, X) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(T_k - \overline{T}_k)^2}{2\sigma^2}}.$$

Similarly, for N (independent) observations $\{(T_k, Y_k)\}_{k=1}^N$, we have

$$p(T_1, \ldots, T_N|Y_1, \ldots, Y_N, \alpha, f_0, f_1, X) = \prod_{k=1}^{N} p(T_k|Y_k, \alpha, f_0, f_1, X) = \prod_{k=1}^{N} \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(T_k - \bar{T}_k)^2}{2\sigma^2}}.$$

The uncertainties in the model and the data are characterized in the distribution by σ, which should at least account for deviations of T_k from its expected value. Since T appears to follows a trend by year (see Figure 3), we fit the temperature to a linear curve and calculate a standard deviation of $\sigma \approx 0.1307$ from this curve.



(a) Linear regression for the temperature data to determine an appropriate σ.

(b) The distribution of deviations of data T_k and the linear fit $\hat{T}$.

Figure 5: Determining a standard deviation σ for temperature.

Next, we choose prior distributions for the parameters α, f_0, and f_1. There are two obvious choices: a uniform distribution over a valid range, or a Gaussian distribution centered at a reasonable value.

- Since the earth's albedo has had only small variations in satellite observations over the 30 years with measurements of $0.29\% \pm 2\%$, we choose a Gaussian with 95% confidence between 0.24 and 0.34, i.e., $p(\alpha|X) = \mathcal{N}(0.29, 0.05^2)$.
- Since $f_0 \in [0, 1)$, we simply set $p(f_0|X) = \mathcal{U}(0, 1)$.

- For f_1, we have the relation

$$0 \leq f = f_0 + f_1 Y \leq 1 \qquad \Longrightarrow \qquad -\frac{f_0}{Y} \leq f_1 \leq \frac{1 - f_0}{Y}.$$

In order to guarantee that f_1 satisfies these bounds, we look at the extremum of f_0. If $f_0 = 0$, then we get

$$-\frac{0}{Y} \leq f_1 \leq \frac{1-0}{Y}. \qquad \Longrightarrow \qquad 0 \leq f_1 \leq \frac{1}{Y}.$$

If $f_0 = 1$, then we get the condition,

$$-\frac{1}{Y} \leq f_1 \leq \frac{1-1}{Y}. \qquad \Longrightarrow \qquad -\frac{1}{Y} \leq f_1 \leq \frac{0}{Y}.$$

but since the temperature grows with increasing CO_2, we require that $f_1 \geq 0$. We don't require f_1 to live completely in this condition, but it just needs to be able to fit in this range. Since $f_1 = 0$ is valid when $f_0 = 1$, this is sufficient. Combining this and the other two conditions and taking the maximum range of Y values, we get that

$$0 \leq f_1 \leq \frac{1}{Y_{\max}}.$$

Since we wish to use the model to predict the temperature for $Y = 717$ (see part (e), we set $Y_{\max} = 720$.

(d) To calibrate the model analytically, we begin by simplifying the expression for the posterior:

$$\begin{aligned} p(\alpha, f_0, f_1 | T, Y, X) &= \frac{p(T|Y, \alpha, f_0, f_1, X) p(\alpha, f_0, f_1 | Y, X)}{p(T|Y, X)} \\ &= \frac{p(T|Y, \alpha, f_0, f_1, X) \frac{p(\alpha, f_0, f_1, Y|X)}{p(Y|X)}}{\frac{p(T,Y|X)}{p(Y|X)}} \\ &= \frac{p(T|Y, \alpha, f_0, f_1, X) p(\alpha|X) p(f_0, f_1|X) \cancel{p(Y|X)}}{p(T|X) \cancel{p(Y|X)}} \\ &= \frac{p(T|Y, \alpha, f_0, f_1, X) p(\alpha|X) p(f_0|X) p(f_1|X)}{p(T|X)}. \end{aligned}$$

For the likelihood, we linearize (4) about the points $\bar{\alpha} = .29$ (the mean of our prior for α), $\bar{f}_0 \approx 0.6272$, and $\bar{f}_1 \approx 0.0001769$ (the values from the least squares fit in part (b)). That is, we set

$$\bar{T}_k = T(Y_k, \alpha, f_0, f_1) \approx T(Y_k, \bar{\alpha}, \bar{f}_0, \bar{f}_1) + \frac{\partial T_k}{\partial \alpha}(\alpha - \bar{\alpha}) + \frac{\partial T_k}{\partial f_0}(f_0 - \bar{f}_0) + \frac{\partial T_k}{\partial f_1}(f_1 - \bar{f}_1),$$

where

$$T(Y_k, \bar{\alpha}, \bar{f}_0, \bar{f}_1) = \left[\frac{\frac{S}{4}(1-\bar{\alpha})}{\sigma_{\text{sb}}\left(1 - \frac{\bar{f}_0 + \bar{f}_1 Y_k}{2}\right)} \right]^{1/4},$$

$$\frac{\partial T_k}{\partial \alpha} = \frac{\partial T}{\partial \alpha}(Y_k, \bar{\alpha}, \bar{f}_0, \bar{f}_1) = -\frac{1}{4}\left[\frac{\frac{S}{4}(1-\bar{\alpha})^{-3}}{\sigma_{\text{sb}}\left(1 - \frac{\bar{f}_0 + \bar{f}_1 Y_k}{2}\right)} \right]^{1/4},$$

$$\frac{\partial T_k}{\partial f_0} = \frac{\partial T}{\partial f_0}(Y_k, \bar{\alpha}, \bar{f}_0, \bar{f}_1) = \frac{1}{8}\left[\frac{\frac{S}{4}(1-\bar{\alpha})}{\sigma_{\text{sb}}} \right]^{1/4} \left(1 - \frac{\bar{f}_0 + \bar{f}_1 Y_k}{2}\right)^{-5/4},$$

$$\frac{\partial T_k}{\partial f_1} = \frac{\partial T}{\partial f_1}(Y_k, \bar{\alpha}, \bar{f}_0, \bar{f}_1) = \frac{Y_k}{8}\left[\frac{\frac{S}{4}(1-\bar{\alpha})}{\sigma_{\text{sb}}} \right]^{1/4} \left(1 - \frac{\bar{f}_0 + \bar{f}_1 Y_k}{2}\right)^{-5/4}.$$

Note that $T(Y_k, \bar{\alpha}, \bar{f}_0, \bar{f}_1)$ and each of the evaluated partial derivatives are constants. Substituting this approximation for $\bar{T}_k$ into the likelihood function, the posterior becomes

$$\begin{aligned} p(\alpha, f_0, f_1 | T, Y, X) &= \frac{p(T|Y, \alpha, f_0, f_1, X) p(\alpha|X) p(f_0|X) p(f_1|X)}{p(T|X)} \\ &\propto \left(\prod_{k=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(T_k - \bar{T}_k)^2}{2\sigma^2}} \right) \mathcal{N}(\bar{\alpha}, \sigma_\alpha^2) \\ &\propto \exp\left(-\frac{1}{2\sigma^2} \sum_{k=1}^{n} \left[T_k - T(Y_k, \bar{\alpha}, \bar{f}_0, \bar{f}_1) - \frac{\partial T_k}{\partial \alpha}\delta\alpha - \frac{\partial T_k}{\partial f_0}\delta f_0 - \frac{\partial T_k}{\partial f_1}\delta f_1 \right]^2 - \frac{1}{2\sigma_\alpha^2}\delta\alpha^2 \right). \end{aligned}$$

To simplify this expression, let $\delta\mathbf{T} \in \mathbb{R}^n$ be the vector with entries $\delta T_k = T_k - T(Y_k, \bar{\alpha}, \bar{f}_0, \bar{f}_1)$, so that $\frac{\partial \mathbf{T}}{\partial \alpha} \in \mathbb{R}^n$ is the vector with entries $-\frac{\partial T}{\partial \alpha}(Y_k, \bar{\alpha}, \bar{f}_0, \bar{f}_1)$, with $\frac{\partial \mathbf{T}}{\partial f_0} \in \mathbb{R}^n$ and $\frac{\partial \mathbf{T}}{\partial f_1} \in \mathbb{R}^n$ defined similarly. Then

$$\sum_{k=1}^{n} \left[T_k - T(Y_k, \bar{\alpha}, \bar{f}_0, \bar{f}_1) - \frac{\partial T_k}{\partial \alpha}\delta\alpha - \frac{\partial T_k}{\partial f_0}\delta f_0 - \frac{\partial T_k}{\partial f_1}\delta f_1 \right]^2 = \left\| \delta\mathbf{T} + \delta\alpha\frac{\partial \mathbf{T}}{\partial \alpha} + \delta f_0 \frac{\partial \mathbf{T}}{\partial f_0} + \delta f_1 \frac{\partial \mathbf{T}}{\partial f_1} \right\|^2.$$

Letting $A = \frac{1}{\sigma}\left[\begin{array}{c|c|c} \frac{\partial \mathbf{T}}{\partial \alpha} & \frac{\partial \mathbf{T}}{\partial f_0} & \frac{\partial \mathbf{T}}{\partial f_1} \end{array} \right] \in \mathbb{R}^{n \times 3}$ and $\mathbf{z}(\alpha, f_0, f_1) = \begin{bmatrix} \delta\alpha & \delta f_0 & \delta f_1 \end{bmatrix}^\mathsf{T} \in \mathbb{R}^3$,

$$\frac{1}{\sigma^2}\left\| \delta\mathbf{T} + \delta\alpha\frac{\partial \mathbf{T}}{\partial \alpha} + \delta f_0 \frac{\partial \mathbf{T}}{\partial f_0} + \delta f_1 \frac{\partial \mathbf{T}}{\partial f_1} \right\|^2 = \|A\mathbf{z} + \delta\mathbf{T}\|^2 = \mathbf{z}^\mathsf{T} A^\mathsf{T} A\mathbf{z} + 2\delta\mathbf{T}^\mathsf{T} A\mathbf{z} + \delta\mathbf{T}^\mathsf{T}\delta\mathbf{T}.$$

Now define $B \in \mathbb{R}^{3\times 3}$ as the zero matrix except for $B_{11} = \frac{1}{\sigma_\alpha^2}$ so that $\mathbf{z}^\mathsf{T} B\mathbf{z} = \frac{1}{\sigma_\alpha^2}\delta\alpha^2$ and

$$\begin{aligned} p(\alpha, f_0, f_1 | T, Y, X) &\propto \exp\left(-\frac{1}{2}\mathbf{z}^\mathsf{T} A^\mathsf{T} A\mathbf{z} - \delta\mathbf{T}^\mathsf{T} A\mathbf{z} - \frac{1}{2}\mathbf{z}^\mathsf{T} B\mathbf{z} \right) \\ &= \exp\left(-\frac{1}{2}\mathbf{z}^\mathsf{T}(A^\mathsf{T} A + B)\mathbf{z} - \delta\mathbf{T}^\mathsf{T} A\mathbf{z} \right). \end{aligned}$$

Defining $\bar{\mathbf{z}} = (A^T A + B)^{-1} A^T \delta\mathbf{T}$ and completing the square,

$$p(\alpha, f_0, f_1 | T, Y, X) \propto \exp\left(-\frac{1}{2}(\mathbf{z} - \bar{\mathbf{z}})^T (A^T A + B)(\mathbf{z} - \bar{\mathbf{z}})\right).$$

@e can rewrite this in terms of the parameter vector $\mathbf{x} = \begin{bmatrix}\alpha & f_0 & f_1\end{bmatrix}^\mathsf{T} \in \mathbb{R}^3$ as

$$p(\alpha, f_0, f_1 | T, Y, X) \propto \exp\left(-\frac{1}{2}(\mathbf{x} - \bar{\mathbf{x}})^T (A^T A + B)(\mathbf{x} - \bar{\mathbf{x}})\right),$$

where $\bar{\mathbf{x}} = \begin{bmatrix}\bar{\alpha} & \bar{f_0} & \bar{f_1}\end{bmatrix}^\mathsf{T} + \bar{\mathbf{z}}$. We now have a symmetric positive definite covariance matrix Σ satisfying $\Sigma^{-1} = A^T A + B$ and a mean $\bar{\mathbf{x}}$. Numerically, these values are

$$\bar{\mathbf{x}} \approx \begin{bmatrix} 0.2900000 \\ 0.6202013 \\ 0.0001778 \end{bmatrix}, \quad \Sigma \approx \begin{bmatrix} 0.00250000000 & 0.00483369835 & -0.00000062295 \\ 0.00483369835 & 0.00935233895 & -0.00000122411 \\ -0.00000062295 & -0.00000122411 & 0.00000000022 \end{bmatrix}.$$

The joint distributions of any two variables can be found by integrating out the third variable. Since the posterior is Gaussian, the resulting marginalized distribution is also Gaussian[1] with mean $\bar{\mathbf{w}} \in \mathbb{R}^2$ and covariance matrix $S \in \mathbb{R}^{2\times 2}$ where $\bar{\mathbf{w}}$ contains the entries of $\bar{\mathbf{x}}$ corresponding to the remaining free variables and S contains the corresponding rows and columns of Σ. Thus,

$$\begin{aligned}
p(\alpha, f_0 | T, Y, X) &= \int_{f_1} p(\alpha, f_0, f_1 | T, Y, X)\, df_1 \\
&\propto \mathcal{N}\left(\begin{bmatrix} 0.2900000 \\ 0.6202013 \end{bmatrix}, \begin{bmatrix} 0.00250000000 & 0.00483369835 \\ 0.00483369835 & 0.00935233895 \end{bmatrix}\right), \\
p(\alpha, f_1 | T, Y, X) &= \int_{f_0} p(\alpha, f_0, f_1 | T, Y, X)\, df_0 \\
&\propto \mathcal{N}\left(\begin{bmatrix} 0.2900000 \\ 0.0001778 \end{bmatrix}, \begin{bmatrix} 0.00250000000 & -0.00000062295 \\ -0.00000062295 & 0.00000000022 \end{bmatrix}\right), \\
p(f_0, f_1 | T, Y, X) &= \int_{\alpha} p(\alpha, f_0, f_1 | T, Y, X)\, d\alpha \\
&\propto \mathcal{N}\left(\begin{bmatrix} 0.6202013 \\ 0.0001778 \end{bmatrix}, \begin{bmatrix} 0.00935233895 & -0.00000122411 \\ -0.00000122411 & 0.00000000022 \end{bmatrix}\right).
\end{aligned}$$

We can also easily get the marginal distributions for α, f_0, and f_1 alone. See Figure 6.

[1] `https://en.wikipedia.org/wiki/Multivariate_normal_distribution#Marginal_distributions`

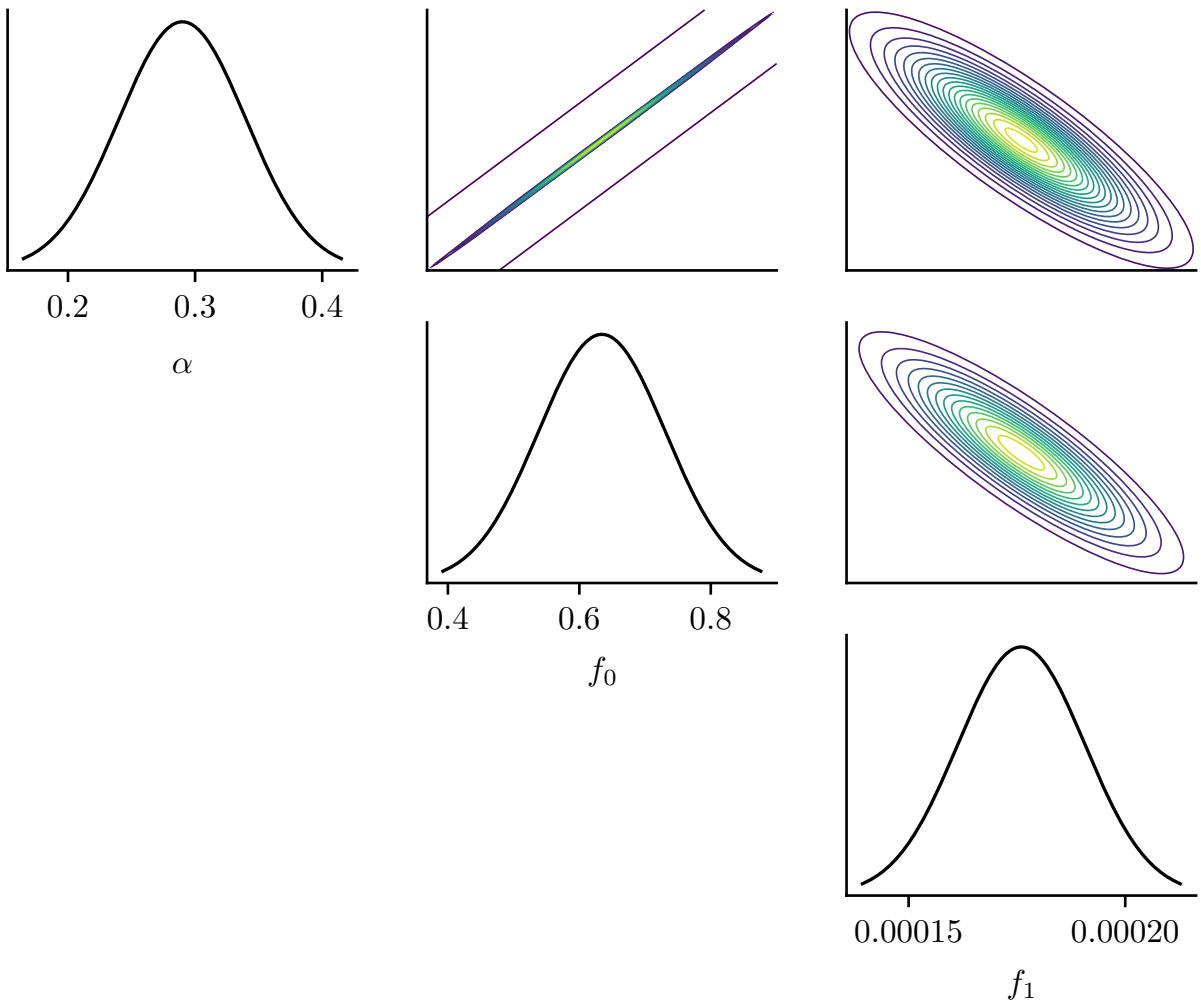


Figure 6: Joint and marginal posteriors from the linearized approximation of (4).

(e) From the posterior above, we pick the values of α, f_0, and f_1 from $\bar{\mathbf{x}}$. Assuming a CO_2 level of $Y_{CO_2} = 717.0$ in the year 2100, we estimate the temperature as

$$T = \left[\frac{\frac{S}{4}(1-\alpha)}{\sigma_{\text{sb}} \left(1 - \frac{f_0 + f_1 Y_{CO_2}}{2}\right)} \right]^{1/4} \approx 288.4159\text{K} = 59.4787^\circ\text{F}.$$

To account for the uncertainty in this prediction, we simply reuse the estimated standard deviation $\sigma \approx 0.1307$ for T; see Figure 7. Alternatively, we could reuse the linearization of T, treating α, f_0, and f_1 as independent Gaussian distributions with the means and variances computed in part (d). In this case, T is Gaussian with mean $T(\bar{\alpha}, \bar{f}_0, \bar{f}_1)$, and

the variance is

$$\begin{aligned}\mathbb{V}\text{ar}[T] &= \mathbb{V}\text{ar}\left[T(Y_k, \bar{\alpha}, \bar{f}_0, \bar{f}_1) + \frac{\partial T_k}{\partial \alpha}(\alpha - \bar{\alpha}) + \frac{\partial T_k}{\partial f_0}(f_0 - \bar{f}_0) + \frac{\partial T_k}{\partial f_1}(f_1 - \bar{f}_1)\right] \\ &= \left(\frac{\partial T_k}{\partial \alpha}\right)^2 \mathbb{V}\text{ar}\left[\alpha\right] + \left(\frac{\partial T_k}{\partial f_0}\right) \mathbb{V}\text{ar}\left[f_0\right] + \left(\frac{\partial T_k}{\partial f_1}\right) \mathbb{V}\text{ar}\left[f_1\right].\end{aligned}$$

This value, however, ends up being very large. In other words, directly propagating the uncertainties in α, f_0, and f_1 leads to a highly uncertain prediction.

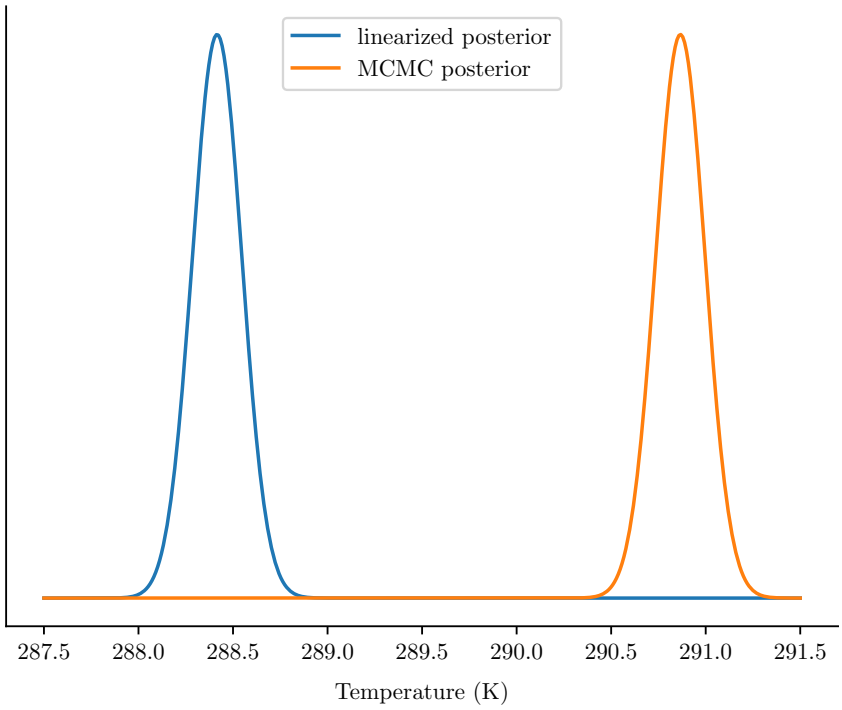


Figure 7: Temperature predictions for the year 2100 with by the linearized model and MCMC sampling.

(f) Now we recalibrate the model using MCMC instead of linearization. In this case, we can be a little more specific with our priors. The posterior is

$$p(\alpha, f_0, f_1 | T, Y, X) \propto p(T | Y, \alpha, f_0, f_1, X) p(\alpha | X) p(f_0, f_1 | X).$$

The likelihood is given by

$$p(T | Y, \alpha, f_0, f_1, X) = \prod_{k=1}^{n} \mathcal{N}(\bar{T}_k, \sigma^2), \qquad \bar{T}_k(Y_k, \alpha, f_0, f_1) = \left[\frac{\frac{S}{4}(1-\alpha)}{\sigma_{\text{sb}}\left(1 - \frac{f_0 + f_1 Y_k}{2}\right)}\right]^{1/4}.$$

We leave $\bar{T}_k$ in this form and do not linearize it. For the prior on α we again use a Gaussian distribution, but we also ensure that $\alpha \in [0, 1]$ by setting $p(\alpha | X) \propto \mathcal{N}(\bar{\alpha}, \sigma_\alpha^2)\mathcal{U}(0, 1)$.

Similarly, we choose a joint uniform prior for f_0 and f_1 that ensures that the opacity $f = f_0 + f_1 Y \in [0, 1]$, i.e,

$$p(f_0, f_1|X) \propto \begin{cases} 1 & \text{if } f_0 + f_1 Y_{\max} \in [0, 1], \\ 0 & \text{else.} \end{cases}$$

The results are computed with the `emcee` package in Python. See Figure 8 and Table 1.

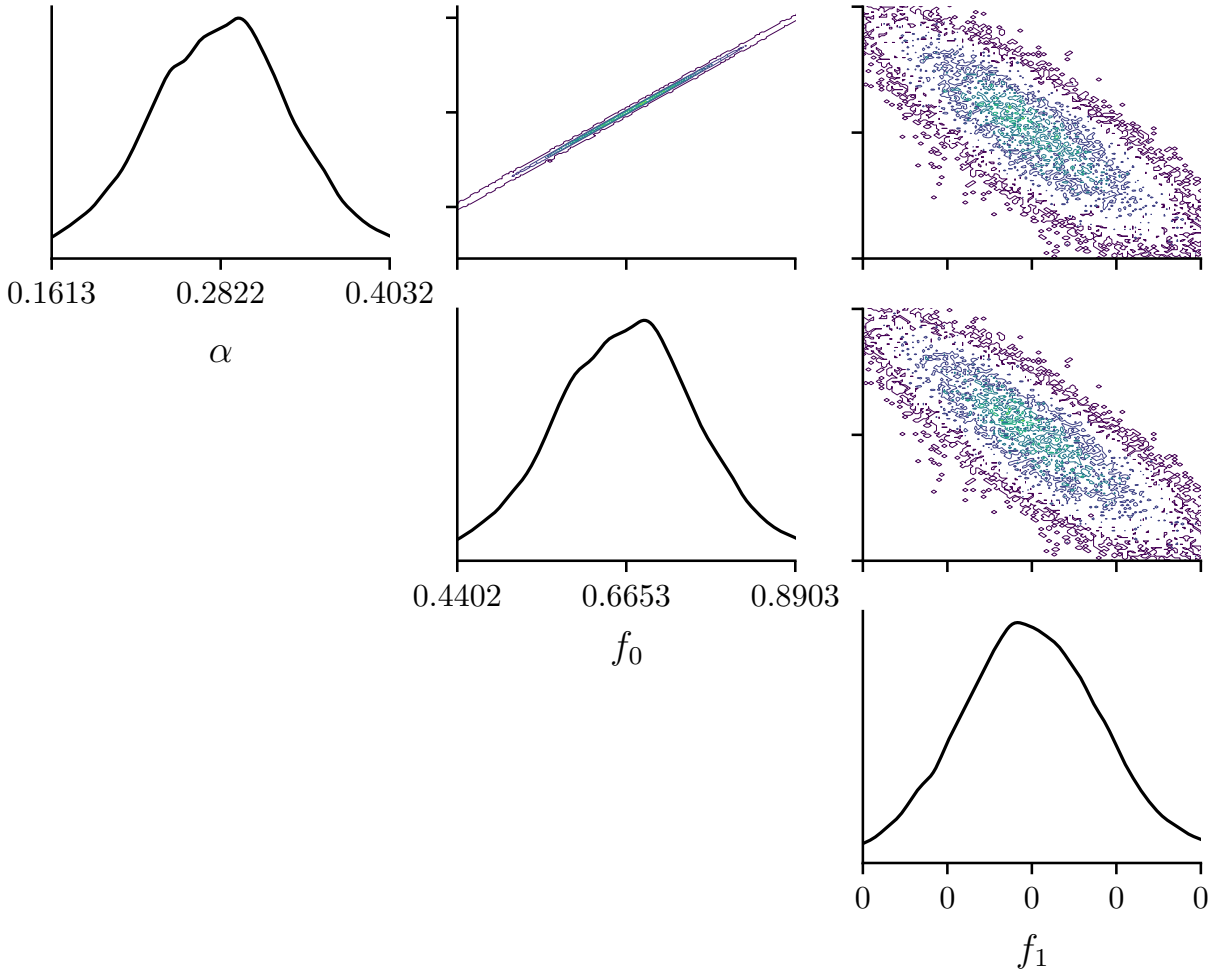


Figure 8: Joint posteriors from MCMC sampling with a burn-in of 500 steps and 1000 samples. Compare with Figure 6.

	min	max	mean	MAP	σ
α	0.0945431	0.4181664	0.2822273	0.2945716	0.0483785
f_0	0.3134436	0.9199450	0.6652612	0.6883163	0.0900085
f_1	0.0001232	0.0002306	0.0001720	0.0001690	0.0000139

Table 1: Statistics of MCMC posterior distributions.

Again assuming $Y_{\mathrm{CO}_2} = 717.0$ in the year 2100, we now estimate the temperature as

$$T = \left[\frac{\frac{S}{4}(1-\bar{\alpha})}{\sigma_{\mathrm{sb}} \left(1 - \frac{\bar{f}_0 + \bar{f}_1 Y_{\mathrm{CO}_2}}{2}\right)} \right]^{1/4} \approx 290.8662\mathrm{K} = 63.8891^{\circ}\mathrm{F}.$$

See Figure 7 again. Note that this prediction is not super close to that of the linearized model, but that they both predict alarming rises in the average global temperature. In fact

(g) We now discuss the differences in the linearization and MCMC approaches.

i. The linearized model discussed in (d) is simple, but it may only naïvely account for nuanced, nonlinear interactions between the parameters that occur in the complete model (4). The results from the linearization also heavily depend on the choice of $\bar{\alpha}$, $\bar{f}_0$, and $\bar{f}_1$ (the points we linearize about), since the resulting calibrated means are very close to these initial guesses. However, even with such limitations, the posterior distributions for the parameters calculated via the linearized model are fairly similar to those calculated via MCMC (i.e., Figures 6 and 8 are comparable). This means we may be able to use the linearized model as a sort of surrogate for the complete nonlinear model in some cases (for example, in the next problem we will attempt to use it as a reference distribution for importance sampling).

ii. It is notable that the temperature predictions for $Y = 717$ are not the same between the posterior via linearization and the posterior via MCMC (though they are only a few degrees off). Here there are two main factors that combine to produce the discrepancy: the limitations of linearization, as discussed above, and the relatively large sensitivity of predictions to small changes in the parameters (see the discussion at the end of part (e)). For example, the mean of f_0 from linearization and the mean of f_0 from MCMC are within a standard deviation of each other, but the mean temperature predictions are several standard deviations away.

iii. We have taken little thought about the uncertainties in the CO_2 data, since the variations in CO_2 are very small compared to the variations in the the temperature data. But, as Figure 9 shows, CO_2 levels are not guaranteed to continue growing as they have, depending on the actions we take to reduce CO_2 emissions. In other words, our predictions should only be considered as possibilities for what may happen if we continue on our present course with CO_2 emissions.

In addition, since (4) is only a zeroeth-order model for the global temperature, there are epistemic uncertainties in the model that we are not able to fully reconcile. These include—but are not limited to—local variations of the parameters α, f_0, and f_1 (for example, the albedo α varies wildly over the earth's surface); the presence of

greenhouse gasses besides CO_2; time-dependent evolution of the parameters; the relationship between CO_2 and the opacity f (which we explore a little in the next problem), and so on. Furthermore, even if we did account for some of these factors, any model for a system as complex as the climate is inevitably unable to represent the underlying phenomena (and we aren't even using a differential equation!).

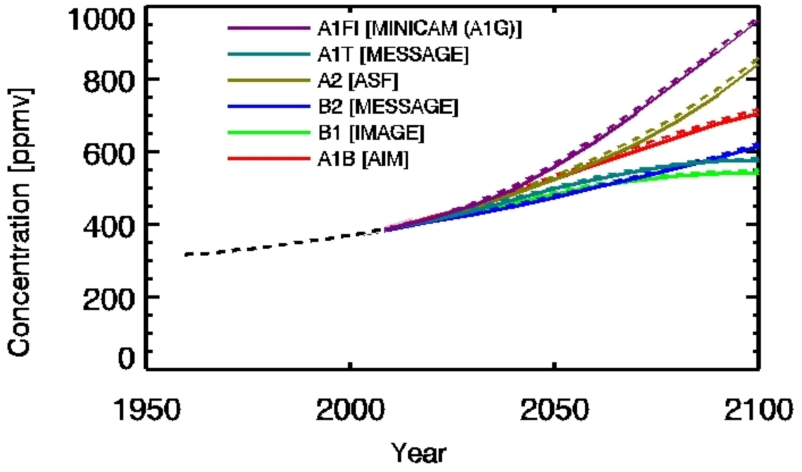


Figure 9: Atmospheric CO_2 concentrations as observed at Mauna Loa from 1958 to 2008 (black dashed line) and projected under the 6 SRES marker and illustrative scenarios. Two carbon cycle models (see Box 3.7 in IPCC, 2001) are used for each scenario: BERN (solid lines) and ISAM (dashed). [figure and caption from `http://www.ipcc-data.org/observ/ddc_co2.html`].

2. Clearly, a critical component of the climate model—and indeed the one that controls the dependence of the global temperature on the CO_2 concentration—is the model for f. Consider the following three proposed f models:

   - $f = f_0$,
   - $f = f_0 + f_1 Y_{CO_2}$, and
   - $f = f_0 + f_1 Y_{CO_2} + f_2 Y_{CO_2}^2$.

   We will use Bayesian model selection to determine relative probabilities of these three forms.

   (a) The evidence in the Bayesian formulation of the calibration problem is given by

   $$p(T|X) = \iiint_{\alpha,f_0,f_1} p(T|Y,\alpha,f_0,f_1,X)p(\alpha|X)p(f_0,f_1|X)d\alpha\, df_0\, df_1$$

   This evidence plays the role of the likelihood in the Bayesian model selection problem,

   $$P(M_i|T,X) = \frac{P(T|M_i,X)P(M_i|X)}{\sum_{i=1}^m P(T|M_i,X)P(M_i|X)}.$$

Unfortunately, the MCMC sampling algorithm used to sample the posterior distribution of the parameters does not allow the direct computation of the evidence. Indeed, MCMC can only give us samples of the posterior distribution, but this does not provide us a way to directly estimate the entire volume of the posterior, since there is no information about how that point should be weighted in a numerical quadrature scheme.

(b) To compute the evidence for each model, we use an importance sampling scheme. This requires choosing an easy-to-sample-from reference distribution q that is similar to the posterior distribution. There are two good options for q:

- Use a linearization of the model (4) as in the previous exercise. This results in a multivariate Gaussian which is easy to sample from, and which appears to represent the true posterior somewhat faithfully based on our MCMC experiments.
- The joint and marginal posteriors in Figure 8 appear very close to Gaussian, so we could also take empirical statistics from the MCMC process (means and the covariance matrix) and construct a multivariate Gaussian with the same statistics.

We will use both of these strategies and compare them.

Let $\pi(\alpha, f_0, f_1)$ be the true posterior and let $q(\alpha, f_0, f_1)$ be the reference distribution (e.g., $q \sim \mathcal{N}(\bar{\mathbf{x}}, \Sigma)$ if we are using the linearized model from 1(d)). To compute the evidence with N samples, we collect N samples $q_i = (\alpha^{(i)}, f_0^{(i)}, f_1^{(i)})$, $i = 1, \ldots, N$, from q, then compute

$$\begin{aligned} p(T|X) &= \iiint_{\alpha,f_0,f_1} p(T|Y,\alpha,f_0,f_1,X)p(\alpha|X)p(f_0,f_1|X)d\alpha\, df_0\, df_1 \\ &= \iiint_{\alpha,f_0,f_1} \frac{\pi(\alpha,f_0,f_1)}{q(\alpha,f_0,f_1)} q(\alpha,f_0,f_1)d\alpha\, df_0\, df_1 \\ &\approx \frac{1}{N}\sum_{i=1}^{N} w_i, \qquad \text{where} \qquad w_i = \frac{\pi(\alpha^{(i)}, f_0^{(i)}, f_1^{(i)})}{q(\alpha^{(i)}, f_0^{(i)}, f_1^{(i)})}. \end{aligned}$$

As mentioned, to choose q for the constant or quadratic models for f, we make a slight modification of the linearized posterior from 1(d). First, for general f we have

$$\begin{aligned} T(\bar{\alpha}, \bar{f}) &= \left[\frac{\frac{S}{4}(1-\bar{\alpha})}{\sigma_{\mathrm{SB}}(1-\frac{\bar{f}}{2})}\right]^{\frac{1}{4}} \\ \frac{\partial T}{\partial \alpha}(\bar{\alpha}, \bar{f}) &= -\frac{1}{4}\left[\frac{\frac{S}{4}(1-\bar{\alpha})^{-3}}{\sigma_{\mathrm{SB}}(1-\frac{\bar{f}}{2})}\right]^{\frac{1}{4}} = -\frac{T(\bar{\alpha},\bar{f})}{4(1-\alpha)} \\ \frac{\partial T}{\partial f}(\bar{\alpha}, \bar{f}) &= -\frac{1}{4(f-2)} T(\bar{\alpha}, \bar{f}). \end{aligned}$$

Now we can consider each model for f separately. For $f = f_0$,

$$\begin{aligned}\frac{\partial T}{\partial f_0} &= \frac{\partial T}{\partial f}\frac{df}{df_0} = -\frac{1}{4(f-2)}T \\ \mathbf{x} &= \begin{bmatrix}\alpha & f_0\end{bmatrix}^\mathsf{T} \\ A &= \frac{1}{\sigma}\begin{bmatrix}\frac{\partial T}{\partial \alpha} & \frac{\partial T}{\partial f_0}\end{bmatrix}.\end{aligned}$$

For $f = f_0 + f_1 Y_{\mathrm{CO}_2}$,

$$\begin{aligned}\frac{\partial T}{\partial f_1} &= \frac{\partial T}{\partial f}\frac{df}{df_1} = -\frac{1}{4(f-2)}TY_{\mathrm{CO}_2} \\ \mathbf{x} &= \begin{bmatrix}\alpha & f_0 & f_1\end{bmatrix}^\mathsf{T} \\ A &= \frac{1}{\sigma}\begin{bmatrix}\frac{\partial T}{\partial \alpha} & \frac{\partial T}{\partial f_0} & \frac{\partial T}{\partial f_1}\end{bmatrix}.\end{aligned}$$

For $f = f_0 + f_1 Y_{\mathrm{CO}_2} + f_2 Y_{\mathrm{CO}_2}^2$,

$$\begin{aligned}\frac{\partial T}{\partial f_2} &= \frac{\partial T}{\partial f}\frac{df}{df_2} = -\frac{1}{4(f-2)}TY_{\mathrm{CO}_2}^2 \\ \mathbf{x} &= \begin{bmatrix}\alpha & f_0 & f_1 & f_2\end{bmatrix}^\mathsf{T} \\ A &= \frac{1}{\sigma}\begin{bmatrix}\frac{\partial T}{\partial \alpha} & \frac{\partial T}{\partial f_0} & \frac{\partial T}{\partial f_1} & \frac{\partial T}{\partial f_2}\end{bmatrix}.\end{aligned}$$

With these modifications, the formulation

$$p(\alpha, f_0, f_1|T, Y, X) \propto \exp\left(-\frac{1}{2}(\mathbf{x}-\bar{\mathbf{x}})^T(A^TA + B)(\mathbf{x}-\bar{\mathbf{x}})\right),$$

is valid for each case, and we can use these linearized posteriors as reference distributions for importance sampling.

Alternatively, we can use MCMC to estimate the statistics of the posterior. The only difference from the linear case discussed in 1(f) is that we specify priors

$$p(f_0|X) \propto \begin{cases}1 & \text{if } f_0 \in [0,1], \\ 0 & \text{else,}\end{cases} \qquad p(f_0, f_1|X) \propto \begin{cases}1 & \text{if } f_0 + f_1 Y_{\max} \in [0,1], \\ 0 & \text{else,}\end{cases}$$

or

$$p(f_0, f_1, f_2|X) \propto \begin{cases}1 & \text{if } f_0 + f_1 Y_{\max} + f_2 Y_{\max}^2 \in [0,1], \\ 0 & \text{else.}\end{cases}$$

We also modify the likelihood to account for the form of f, then easily adapt the `emcee` code to sampling the posteriors for each model.

(c) Finally, we compute the evidence and posterior probabilities for each model. Let M_0 indicate the constant model $f = f_0$, M_1 indicate the linear model, and M_2 the quadratic

model. Recall that the posterior probability for M_i is given by

$$P(M_i|T,X) = \frac{P(T|M_i,X)P(M_i|X)}{\sum_{i=0}^{2} P(T|M_i,X)P(M_i|X)}.$$

The prior probabilities $P(M_i|X)$ turn out not to matter too much in this case (spoilers!), but we can justify picking them in a few ways. For example, we could say $P(M_0|X) = P(M_1|X) = P(M_2|X) = \frac{1}{3}$, meaning we have no reason to believe that any model is more likely than another. However, the constant model is highly suspicious (it assumes that the CO_2 level has no effect on the opacity of the atmosphere), so we might inject the problem with some doubt about this model by setting, e.g., $P(M_0|X) = 0.2$ and $P(M_1|X) = P(M_2|X) = 0.4$. There are many other justifiable ways to pick these prior probabilities, but we'll stick with these last numbers for now.

As discussed in part (b), there are (at least) two ways to pick the reference distribution q for each model, and it turns out to have a large impact. First, if we use Gaussians derived from linearizations of each model, then each evidence evaluates to essentially zero (using 5000 samples to compute the integrals). This is an indication[2] that these choices of q are not quite appropriate. On the other hand, if we use Gaussians with statistics derived from MCMC estimates of each model, then we obtain some very large evidences,

$$P(T|M_0,X) \approx 0, \qquad P(T|M_1,X) \approx 2.14 \times 10^{33}, \qquad P(T|M_2,X) \approx 2.81 \times 10^{28}.$$

The hugeness of these numbers makes us wonder if there is some numerical instability that is distorting our results, but if we discount that possibility, then we obtain posterior probabilities for each model as follows:

$$P(M_0|T,X) \approx 0, \qquad P(M_1|T,X) \approx 99.998687\%, \qquad P(M_2|T,X) \approx 0.001313\%.$$

This is believable, and supported by the fact that a least squares fit to the data assuming the model M_2 yields a very small coefficient f_2 for the quadratic term, so it seems that the linear model M_1 for f is much more likely.

[2]Unless, of course, there is a big bug in the code... :)

Appendix A: Code

See pretty much all of our code and some figures in this Google Colab notebook: `https://colab.research.google.com/drive/1JjxovRT4-oum2YHFbZFA3H9_-Xbzipqv`.

Appendix B: Additional Figures

Figure 10: Oof.