PARAMETER AND UNCERTAINTY ESTIMATION FOR A MODEL OF ATMOSPHERIC CO_2 OBSERVATIONS

ARIANNA KRINOS* AND AIMEE MAURAIS†

Advisor: Matthias Chung

Abstract. In this project, we deduce and analyze a mathematical model for atmospheric carbon dioxide concentrations during the time period from 1958 to 2018, as observed by NOAA at the Mauna Loa Observatory. We approximate atmospheric CO_2 during this period using a linear combination of a constant to represent atmospheric carbon dioxide concentration at the beginning of the modeled period, a sinusoidal function to capture annual seasonal variation in carbon dioxide concentrations, and an exponential component to capture the observed increase in global carbon dioxide concentration in the atmosphere from the Mauna Loa dataset. Using Bayesian inference methods, we estimate parameters for our model via a Markov Chain Monte Carlo method, the Adaptive Metropolis algorithm. We present distributions for each of six important model parameters, and present predictive intervals for projected increases in atmospheric CO_2 concentration for the period from 2018 to 2120. We find that CO_2 concentrations can be predicted reasonably well using our modeling approach, and suggest that our framework be used as an adaptable, extensible method of finding good approximations with low variances for data of this type.

1. Introduction & Background. Since 1957, the National Oceanic and Atmospheric Administration (NOAA) has collected and maintained a record of hourly measurements of the earth's atmospheric carbon dioxide (CO_2) concentration (parts per million, ppm) at its Mauna Loa Observatory in Hawaii. Begun by Charles David Keeling, the CO_2 observation program at Mauna Loa has generated data which serves as quantitative evidence of the earth's seasonal CO_2 cycle and of rising overall average CO_2 levels [2]. The data, known as the Keeling Curve, appears to follow an exponential growth trend and oscillate sinusoidally with a period of about one year (see Figure 1). The dataset has been used extensively to understand natural and human-caused fluctuations in carbon dioxide levels, such as the "fertilization" process catalyzed by shifts in terrestrial vegetation, whereby reduced plant consumption of carbon dioxide due to winter dormancy results in increased atmospheric carbon dioxide concentrations [12, 6]. Relatively long-term carbon dioxide datasets similar to the one collected at Mauna Loa are often used as drivers for modeling studies that aim to develop scenarios for future carbon dioxide concentrations [14].

Mauna Loa CO_2 data have been the subject of many modeling studies, including predator-prey models for the interaction between carbon and living biomass [25], models focusing on human contributions and the atmospheric response [20], and differential equation modeling approaches using least squares regression to fit the data to multiple model types [13]. A Gaussian process modeling approach has also been taken to fit these data, which results in a model largely independent of potentially dangerous assumptions on the data and its parameters (thus also excluding socioeconomic components of atmospheric carbon dioxide), but allowing a final model to be chosen which is not nearly optimal for the underlying dataset [26].

*Division of Computational Modeling and Data Analytics and Departments of Computer Science and Biological Sciences, Virginia Tech, (akrinos@vt.edu).

†Division of Computational Modeling and Data Analytics and Department of Mathematics, Virginia Tech, (amaurais@vt.edu).

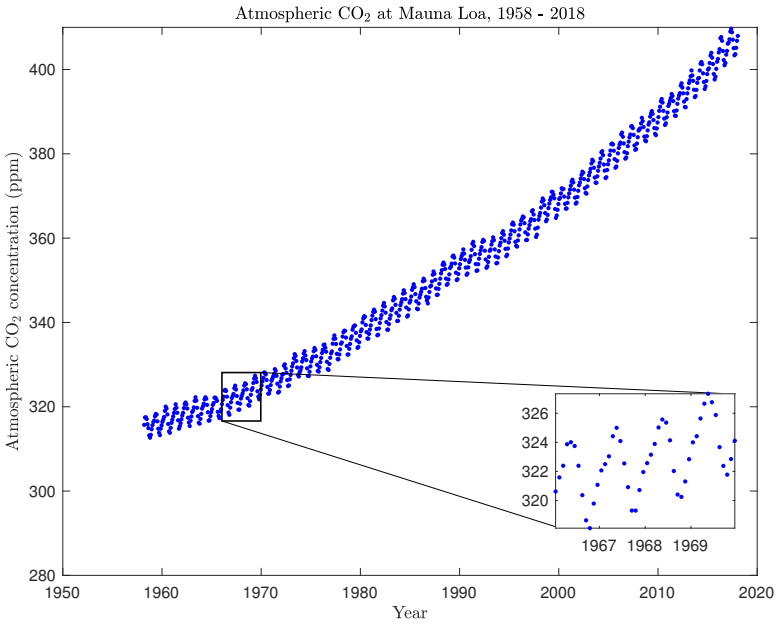

Fig. 1: Monthly mean atmospheric CO_2 concentration as observed at the Mauna Loa Observatory. Inset: four cycles of seasonal CO_2 oscillation [23].

Here, we use data from Mauna Loa to construct and evaluate a mathematical model for global atmospheric CO_2 concentrations. We provide estimates with uncertainties of global atmospheric CO_2 dynamics by tuning our model to the data provided by the Mauna Loa Observatory in Hawaii during the years 1957 to 2018 (Figure 1). We propose a model for CO_2 dynamics based on observed natural patterns, and we use Markov Chain Monte Carlo methods to determine and validate estimates for model parameters.

In Section 2, we provide background on the dataset we use to fit our model, discussing how it is collected and why we consider it reliable. In Section 3, we formulate a three-term mathematical model for atmospheric CO_2 concentration, requiring six model parameters, and discuss the physical meanings of these parameters. In Section 4, we provide an overview of Markov Chain Monte Carlo Methods and the Adaptive Metropolis Algorithm and detail the implementation specific to parameter estimation for our CO_2 model. We then discuss results obtained using these methods to approximate the distributions of our parameter set. In Section 5, we discuss the implications of our findings with respect to future increase in global carbon dioxide concentration relative to the prediction interval we produce. We consider the possibility of future atmospheric change, in the context of literature citing shifts in seasonal cycling of carbon dioxide, which could be well approximated by our modeling procedure and adjusted to make holistic predictions.

Our modeling procedure is unique because it incorporates a standard parameter fitting procedure with the Markov Chain Monte Carlo method, which enables us to produce a tight prediction interval for future atmospheric carbon dioxide concentrations. We can be highly confident in our estimate for modeled atmospheric CO_2 levels

as a consequence of using this approach, which is easily adapted to fit different and more complex mathematical models as well as additional or extended datasets.

2. Mauna Loa Data Collection. Data collected at Mauna Loa is available in raw form or as series of hourly, weekly, monthly, and annual means calculated from multiple hourly measurements made via infrared spectroscopy [23, 22]. Mauna Loa is ideally situated as an atmospheric observatory; due to its elevation and location it is possible to capture "background air" and obtain CO_2 measurements that are not heavily influenced by local vegetation or human activity [22]. However, even though the Mauna Loa Observatory is surrounded by miles of bare volcanic rock and is geographically far removed from large human populations, measurements made there can still be subject to the influences of vegetation on the island due to upslope winds. Hourly averages known to be impacted by upslope winds are flagged in the dataset, as are those characterized by large variability in individual measurements, those that differ significantly from previous hours, and any otherwise unflagged hours surrounded by flagged hours. In 2014, 37.9% of hours were unflagged, 52.3% were flagged, and 9.8% were missing or had no valid measurement [22].

Carbon dioxide levels on Mauna Loa are measured based upon the level of infrared absorption in a cylinder of dry air collected at the observatory. The measurement apparatus is calibrated multiple times per day using a series of reference and test gases (of known CO_2 concentrations), in order to maintain fidelity of the measurements. Although raw data and hourly averages are available, daily and monthly moving averages are also reported and often used for outside research [22].

In this study we use the preprocessed series of monthly average atmospheric CO_2 measurements, provided by NOAA's Earth System Research Laboratory (ESRL) Global Monitoring Division (GMD) and freely available online (see [23]).

3. Model Formulation and Parameter Interpretation.

3.1. Mathematical Model. Based on knowledge that (i) atmospheric CO_2 concentrations are increasing in time, and (ii) atmospheric CO_2 oscillates seasonally due to the vegetation cycle of the Northern Hemisphere [4], we posit the following model for atmospheric CO_2 concentration,

$$\boxed{A(t) = p_1 e^{p_2(t-t_0)} + p_3 \sin(2\pi p_4(t-p_5)) + p_6} \\ = A_{\exp}(t) + A_{\sin}(t) + A_0(t),$$

with explanations and interpretations of each parameter as follows:

I. **Parameter p_1: Scale Factor**. The parameter p_1 scales the rate of growth of the exponential component, $A_{\exp}$, of the model over time.

II. **Parameter p_2: Proportionality Constant**. The parameter p_2 is a proportionality constant between $A_{\exp}$ and $\frac{dA_{\exp}}{dt}$ such that $\frac{dA_{\exp}}{dt} = p_2 A_{\exp}$.

III. **Parameter p_3: Amplitude of Seasonal CO_2 Oscillation**. The parameter p_3 quantifies the magnitude of the seasonal shift in CO_2 production caused by the Northern Hemisphere's yearly vegetation cycle (we expect this to be approximately one year^{-1} [4] in duration). The Northern Hemisphere, being the most concentrated area of earth's vegetation growth, intakes a great deal of carbon dioxide in spring for respiration as leaves are regrowing [5, 11]. This results in an annual maximum in approximately May, just before vegetation has been fully restored for the summer and plants resume high consumption of carbon dioxide [10]. This pattern is both clarified and complicated by the contribution

of global oceans, which sequester carbon dioxide to different degrees seasonally due to increased solubility of gases at low temperatures, as well as changes in ocean pH [15]. Thus, while oceanic contribution is an explanation for seasonal change, it also produces variability which is more difficult to predict and describe. However, ocean contributions are substantially less important in the Northern Hemisphere than in the Southern Hemisphere due to the influence of warm southern tropical oceans, and sometimes other natural phenomena, such as forest fires and volcanic activity, obscure the expected annual trend [15]. We attempted to capture this variation in the amplitude of the seasonal trend due to known ecological phenomena by fitting p_3.

IV. **Parameter p_4: Reciprocated Period of Observed Oscillation**. The parameter p_4, one over the period of observed oscillation, corresponds to the fact that atmospheric carbon dioxide levels correspond closely to a one-year cycle, based on natural phenomena [23]. Because functions of the form $f(t) = \sin(\omega(t-\gamma))$ have period $T = \frac{2\pi}{\omega}$, the oscillatory component of our model,

$$A_{\sin}(t) = p_3 \sin(2\pi p_4(t - p_5)),$$

will have period $\frac{1}{p_4}$. We thus expect p_4 to have a value of about 1 year^{-1}.

V. **Parameter p_5: Shift**. Based on prior work, there is no precedent for the inclusion of a nonzero shift term in this differential equation model [9], but we observe empirically that using a shift close to the model period start year may yield more accurate results [9].

VI. **Parameter p_6: Historic Baseline CO_2 Level**. The parameter p_6 represents the earth's historic atmospheric CO_2 level prior to the period for which Mauna Loa Data is available. We obtained an initial estimate for p_6 by averaging the inferred atmospheric CO_2 levels obtained from Antarctic ice cores for the years 1948-1957, the ten years leading up to the modeled period (see Figure 2) [1]. The addition of p_6 to the model provides a "baseline" CO_2 level in the absence of variability and allows for realistic model behavior as $t \to t_0$.

4. Methods and Results. In the following, we provide a short introduction to Markov Chain Monte Carlo (MCMC) methods and the specific method we used, the Adaptive Metropolis Algorithm (for details, see Robert and Casella [17]), discuss implementation details specific to estimating parameters for our model, and demonstrate the results obtained by using these methods to approximate the sampling distributions of our parameters.

4.1. Markov Chain Monte Carlo Methods. Our goal in using Markov Chain Monte Carlo methods is to approximate the *posterior* density function of our parameter set $\mathbf{p}$ given our data $\mathbf{d}$, which using *Bayes' Theorem* can be expressed as

$$\pi_{\text{post}}(\mathbf{p}|\mathbf{d}) = \frac{\pi_{\text{like}}(\mathbf{d}|\mathbf{p})\pi_{\text{prior}}(\mathbf{p})}{\pi_{\text{marg}}(\mathbf{d})}, \tag{1}$$

where $\mathbf{p} = \begin{bmatrix} p_1 & \dots & p_6 \end{bmatrix}$ is our model's parameter set, $\mathbf{d}$ is the set of observations, π_{like} is the *likelihood* of observing $\mathbf{d}$ under $\mathbf{p}$, and π_{prior} and π_{marg} are the densities of $\mathbf{p}$ and $\mathbf{d}$, respectively [24]. In this setting, π_{post} is unknown, and π_{like}, π_{prior} and π_{marg} are used to estimate it.

MCMC methods utilize Bayes' Theorem to compute the unknown distribution of $\pi_{\text{post}}(\mathbf{p}|\mathbf{d})$ via selective sampling from a chosen proposal distribution. The unknown distribution π_{post} is traversed via a *Markov Chain*, which is a sequence of realizations

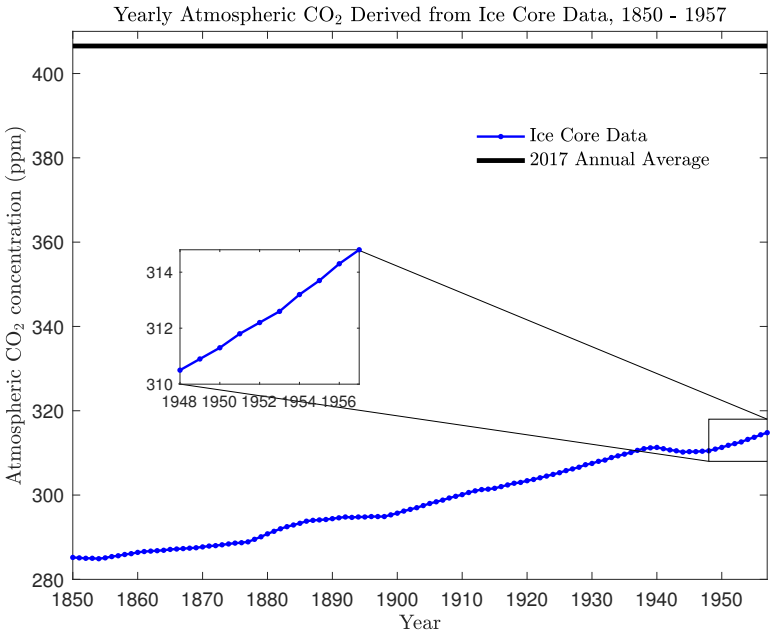

Fig. 2: Atmospheric CO_2 concentration for the years 1850-1957 inferred from Antarctic ice cores, obtained from NASA/NOAA [1].

of the random variable $\mathbf{p} \in \mathbb{R}^m$ possessing the Markov Property, such that on step $i+1$ of the chain,

$$f(\mathbf{p}_{i+1}|\mathbf{p}_1, \ldots, \mathbf{p}_0) = f(\mathbf{p}_{i+1}|\mathbf{p}_i), \tag{2}$$

where f denotes probability density. Thus each realization of the chain depends only on the realization preceding it. On step $i+1$ of the MCMC iteration, we take a sample $\mathbf{p}^*$ from the proposal distribution, such that $\mathbf{p}^*$ only depends on $\mathbf{p}_i$ (in accordance with the Markov Property), but only accept this sample, setting $\mathbf{p}_{i+1} = \mathbf{p}^*$, with probability $p = \min\left\{1, \frac{\pi_{\text{post}}(\mathbf{p}_{i+1})}{\pi_{\text{post}}(\mathbf{p}_i)}\right\}$, where π_{post} is computed using (1) [7]. In this way we approximate the sampling distribution of $\mathbf{p}$ by sampling from the known proposal distribution and using (1) to determine whether our samples are reasonable realizations of $\mathbf{p}$.

Simulations of the chain are performed using randomly generated or preselected initial states $\mathbf{p}_0$, and each subsequent state $\mathbf{p}_i$ is sampled from the proposal distribution with parameter $\mathbf{p}_{i-1}$. As the beginning of the sequence $\{\mathbf{p}_i\}$ is highly dependent on the initial state $\mathbf{p}_0$, the first k samples, where k is an integer selected case-by-case, are sometimes discarded as part of what is known as the *burn-in period*, and only the samples after the burn in period are used to approximate π_{post} [18]. The Markov Chain is assumed to have a stationary distribution, and after many chains are simulated the posterior distribution of $\mathbf{p}$ given $\mathbf{d}$ is inferred from the sample space $\{\mathbf{p}_i\}$ traversed by the Markov Chain [19]. This sample space $\{\mathbf{p}_i\}$ can be used to obtain a *maximum a posteriori* estimate for $\mathbf{p}$, that is, the value of $\mathbf{p}$ of highest density in the posterior distribution, and to perform uncertainty quantification.

The particular flavor of MCMC utilized in this study is the Adaptive Metropolis Algorithm. Like with other MCMC methods, on each iteration of Adaptive Metropolis, we sample $\mathbf{p}^* \sim \mathcal{N}(\mathbf{p}_i, \mathbf{C}_i)$, using a multivariate normal distribution as our proposal distribution, where $\mathbf{C}_i$ is the covariance matrix of the parameter set on step i, and set $\mathbf{p}^* = \mathbf{p}_{i+1}$ with probability $c = \min\left\{1, \frac{\pi_{\text{post}}(\mathbf{p}^*)}{\pi_{\text{post}}(\mathbf{p}_i)}\right\}$, otherwise we set $\mathbf{p}_{i+1} = \mathbf{p}_i$. Adaptive Metropolis differs from the unmodified Metropolis algorithm in that the covariance matrix $\mathbf{C}_i$ is updated every jth iteration, where the value of j is determined by the user, based off the distribution of the previous samples. An update to the covariance matrix $\mathbf{C}$ on step i (where i is a multiple of j) of the iteration will set

$$\mathbf{C}_{i+1} = s_m \text{cov}(\mathbf{p}_0, \ldots, \mathbf{p}_i) + s_m \varepsilon \mathbf{I}_m,$$

where $\mathbf{I}_m$ is the m-dimensional identity matrix, s_m is a "parameter that depends only on dimension" m and $\varepsilon > 0$ is a small constant relative to the magnitude of our samples. In practice, we use an equivalent recursive formula to obtain each realization of $\mathbf{C}_i$, see [7].

Updating $\mathbf{C}_i$ may yield faster convergence, but invalidates the Markov Property (2) of the chain, since the distribution from which the each sample $\mathbf{p}_i$ is drawn has covariance matrix $\mathbf{C}_i$, which is dependent on previous samples $\mathbf{p}_0, \ldots, \mathbf{p}_k$, where k is the largest multiple of j less than i. Haario *et al.* showed in 2001 that despite this, under mild assumptions the Adaptive Metropolis algorithm still possesses the ergodic properties of regular Metropolis, i.e., it will traverse the entire space and converge to the target distribution π_{post} [7].

The Adaptive Metropolis Algorithm is given in Algorithm 1. In the context of our atmospheric CO_2 model, Adaptive Metropolis outperforms regular Metropolis due to the fact that our model parameters are highly correlated (see Figure 3, which depicts the densities of each parameter projected to 2-dimensional space, with the intensity of the color representing measurement frequency and the axes labeled with each parameter). Updating the proposal covariance matrix $\mathbf{C}$ throughout the iterations captures the correlation between parameters in the current sample space, transforming $\mathbf{C}$ to be non-stationary.

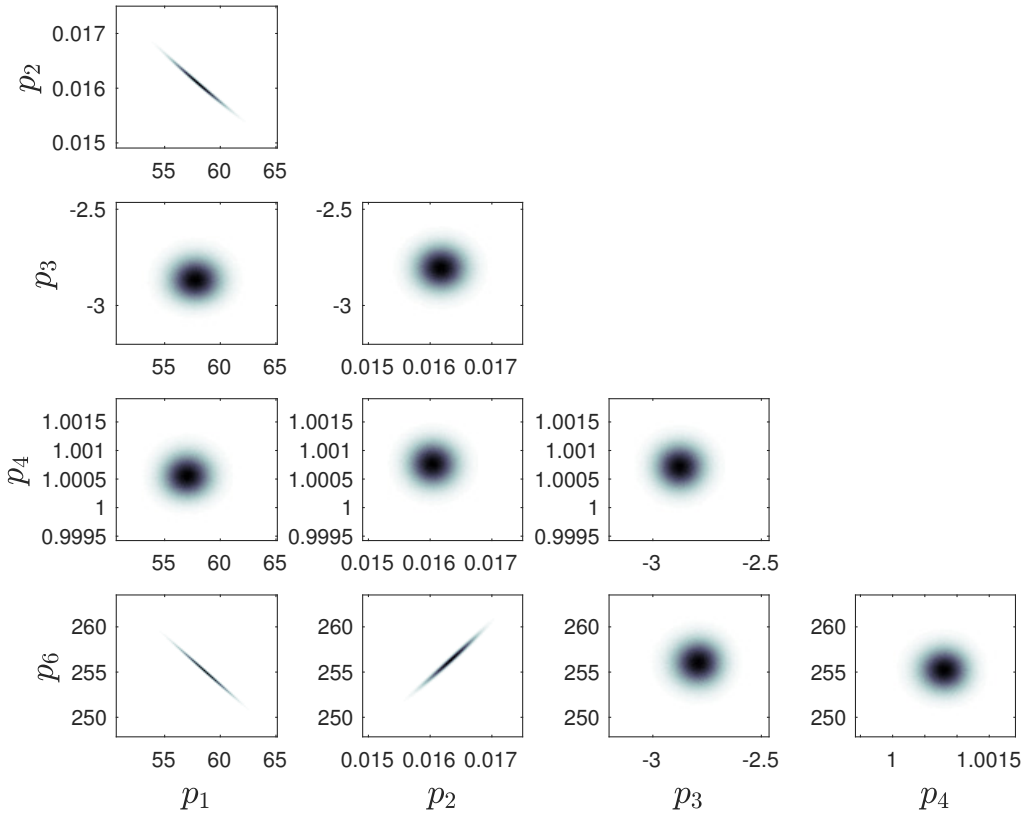

Fig. 3: Two-dimensional marginal densities of parameters in our sample space. Note the strong correlations between p_1 and p_6, p_2 and p_6, and p_1 and p_2. Similarly, a very strong relationship exists between p_4 and p_5, shown in detail in Figure 4.

Algorithm 1 Adaptive Metropolis

Require: π_{like}, π_{prior}, π_{prop}, $\mathbf{d}$, $\mathbf{p}_0$, j
$i = 0$, compute posterior $\pi_{\text{post}}(\mathbf{p}_0|\mathbf{d})$
while not done **do**
  $\mathbf{p}_{\text{prop}} \sim \mathcal{N}(\mathbf{p}_i, \mathbf{C}_i)$
  compute posterior $\pi_{\text{post}}(\mathbf{p}_{\text{prop}}|\mathbf{d})$
  compute $c = \min\left(1, \frac{\pi_{\text{post}}(\mathbf{p}_{\text{prop}}|\mathbf{d})}{\pi_{\text{post}}(\mathbf{p}_i|\mathbf{d})}\right)$
  sample $u \sim \mathcal{U}([0,1])$
  if $u < c$ **then**
    $\mathbf{p}_{i+1} = \mathbf{p}_{\text{prop}}$
  else
    $\mathbf{p}_{i+1} = \mathbf{p}_i$
  end if
  if $\mod(i,j) = 0$ **then**
    update $\mathbf{C}_i \to \mathbf{C}_{i+1}$ using $\mathbf{p}_{i+1}$
  else
    $\mathbf{C}_{i+1} = \mathbf{C}_i$
  end if
  $i = i + 1$
end while
Ensure: $\{\mathbf{p}_i\}_{i=1}^{K}$ samples from posterior

Here $\mathcal{U}([0,1])$ indicates the uniform distribution on $[0,1]$.

4.2. Implementation. In our implementation of Adaptive Metropolis used to obtain samples of our model parameters $\mathbf{p} = \begin{bmatrix} p_1, & p_2, & p_3, & p_4, & p_5, & p_6, \end{bmatrix}^\top \in \mathbb{R}^6$, we used "uninformed priors" for each of the parameters $p_1, \ldots, p_6$, assuming them to be uniformly distributed on wide intervals. Letting $D \subseteq \mathbb{R}^6$ be the set on which $\pi_{\text{prior}}(\cdot)$ is nonzero, we assume that $\mathbf{p}_i \in D$ on every iteration i, and as such π_{prior} is constant for each iteration. This assumption combined with the fact that $\mathbf{d}$ is fixed, gives

$$\frac{\pi_{\text{post}}(\mathbf{p}^*)}{\pi_{\text{post}}(\mathbf{p}_i)} = \frac{\frac{\pi_{\text{like}}(\mathbf{d}|\mathbf{p}^*)\pi_{\text{prior}}(\mathbf{p}^*)}{\pi_{\text{marg}}(\mathbf{d})}}{\frac{\pi_{\text{like}}(\mathbf{d}|\mathbf{p}_i)\pi_{\text{prior}}(\mathbf{p}_i)}{\pi_{\text{marg}}(\mathbf{d})}} = \frac{\pi_{\text{like}}(\mathbf{d}|\mathbf{p}^*)}{\pi_{\text{like}}(\mathbf{d}|\mathbf{p}_i)},$$

and hence $\pi_{\text{marg}}(\mathbf{d})$ and $\pi_{\text{prior}}(\mathbf{p})$ do not impact our calculations.

To obtain an expression for $\pi_{\text{like}}(\mathbf{d}|\mathbf{p})$, we assume that the data $\mathbf{d} \in \mathbb{R}^n$ is a random variable given by

$$\mathbf{d} = \begin{bmatrix} A(t_1;\mathbf{p}), & A(t_2;\mathbf{p}), & \ldots, & A(t_n;\mathbf{p}) \end{bmatrix} + \boldsymbol{\varepsilon} = g(\mathbf{t};\mathbf{p}) + \boldsymbol{\varepsilon},$$

where $\boldsymbol{\varepsilon} \sim \mathcal{N}_n(\mathbf{0}, \sigma^2\mathbf{I}_n)$, and $\sigma^2\mathbf{I}_n$ is the covariance matrix of $\boldsymbol{\varepsilon}$. According to an n-variate normal distribution, under a parameter set $\mathbf{p}$

$$\pi_{\text{like}}(\mathbf{d}|\mathbf{p}) \propto \mathrm{e}^{-\frac{1}{2\sigma^2}\boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon}}.$$

In practice, we work with the negative log likelihood of $\pi_{\text{like}}(\mathbf{d}|\mathbf{p})$ rather than working with $\pi_{\text{like}}(\mathbf{d}|\mathbf{p})$ directly in order to avoid computational difficulties. To analytically obtain the *maximum a posteriori* (MAP) estimate for $\mathbf{p}$, $\mathbf{p}_{\text{map}}$, we maximize the posterior density $\pi_{\text{post}}(\mathbf{p}|\mathbf{d})$, which under our assumptions is proportional to the likelihood $\pi_{\text{like}}(\mathbf{d}|\mathbf{p})$, with respect to $\mathbf{p}$, or equivalently minimize the negative log likelihood

$$\begin{aligned} -\ln(\pi_{\text{like}}(\mathbf{d}|\hat{\mathbf{p}})) &\propto \frac{1}{2\sigma^2}\boldsymbol{\varepsilon}^\top\boldsymbol{\varepsilon} \\ &\propto \frac{1}{2\sigma^2}(g(\mathbf{t};\mathbf{p}) - \mathbf{d})^\top(g(\mathbf{t};\mathbf{p}) - \mathbf{d}) \\ &\propto \frac{1}{2\sigma^2}||g(\mathbf{t};\mathbf{p}) - \mathbf{d}||_2^2, \end{aligned}$$

where $||\cdot||_2$ denotes the Euclidean norm. So then

$$\mathbf{p}_{\text{map}} = \arg\min_{\mathbf{p}} ||g(\mathbf{t};\mathbf{p}) - \mathbf{d}||_2^2,$$

and finding the MAP estimate for $\mathbf{p}$ in this setting is equivalent to finding the least squares estimate for $\mathbf{p}$.

Combining our expressions for π_{prior} and π_{like}, on each iteration of Adaptive Metropolis we compute (using (1))

$$\begin{aligned} -\ln(\pi_{\text{post}}(\mathbf{p}|\mathbf{d})) &\propto -\ln\left(\frac{\pi_{\text{like}}(\mathbf{d}|\mathbf{p})\pi_{\text{prior}}(\mathbf{p})}{\pi_{\text{marg}}(\mathbf{d})}\right) \\ &\propto -\ln(\pi_{\text{like}}(\mathbf{d}|\mathbf{p})) - \ln(\pi_{\text{prior}}(\mathbf{p})) + \ln(\pi_{\text{marg}}(\mathbf{d})) \\ &\propto -\ln(\pi_{\text{like}}(\mathbf{d}|\mathbf{p})) + \ln(\pi_{\text{marg}}(\mathbf{d})) + \begin{cases} C & \mathbf{p} \in D \\ \infty & \mathbf{p} \notin D \end{cases}, \end{aligned}$$

where C is a constant. As discussed above, we assume $\mathbf{p} \in D$ on every iteration. Thus $\ln(\pi_{\text{prior}}(\mathbf{p})) = C$ on every iteration, and $\ln(\pi_{\text{marg}}(\mathbf{d})) + \ln(\pi_{\text{prior}}(\mathbf{p}))$ is effectively a constant. Therefore we have

$$\begin{aligned} -\ln(\pi_{\text{post}}(\mathbf{p}|\mathbf{d})) &\propto -\ln(\pi_{\text{like}}(\mathbf{d}|\mathbf{p})) + \text{Constant} \\ &\propto \frac{1}{2}||g(\mathbf{t};\mathbf{p}) - \mathbf{d}||_2^2 + \text{Constant}. \end{aligned}$$

We use the following procedure to accept $\mathbf{p}_{i+1}$ as the next sample in the chain with probability $c = \min\{1, \frac{\pi_{\text{post}}(\mathbf{p}_{i+1}|\mathbf{d})}{\pi_{\text{post}}(\mathbf{p}_i|\mathbf{d})}\} = \min\{1, \frac{\pi_{\text{like}}(\mathbf{d}|\mathbf{p}_{i+1})}{\pi_{\text{like}}(\mathbf{d}|\mathbf{p}_i)}\}$:

If $||g(\mathbf{t};\mathbf{p}_{i+1}) - \mathbf{d}||_2^2 < ||g(\mathbf{t};\mathbf{p}_i) - \mathbf{d}||_2^2$, then we accept $\mathbf{p}_{i+1}$ as our next sample. If not, then we randomly accept $\mathbf{p}_{i+1}$ with probability $\frac{\pi_{\text{post}}(\mathbf{p}_{i+1})}{\pi_{\text{post}}(\mathbf{p}_i)} = \frac{\pi_{\text{like}}(\mathbf{d}|\mathbf{p}_{i+1})}{\pi_{\text{like}}(\mathbf{d}|\mathbf{p}_i)}$ by sampling a random number $u \sim \mathcal{U}(0,1)$ and accept the sample if

$$u < \frac{\pi_{\text{like}}(\mathbf{d}|\mathbf{p}_{i+1})}{\pi_{\text{like}}(\mathbf{d}|\mathbf{p}_i)}.$$

In terms of the negative log likelihood, this is equivalent to

$$\begin{aligned} -\ln(u) &> -\ln\left(\frac{\pi_{\text{like}}(\mathbf{d}|\mathbf{p}_{i+1})}{\pi_{\text{like}}(\mathbf{d}|\mathbf{p}_i)}\right) \\ -\ln(u) &> -\ln(\pi_{\text{like}}(\mathbf{d}|\mathbf{p}_{i+1})) + \ln(\pi_{\text{like}}(\mathbf{d}|\mathbf{p}_i)). \end{aligned}$$

By our assumption on the data distribution the condition becomes, for $\mathbf{p}_i$ and $\mathbf{p}_{i+1}$ with errors $\boldsymbol{\varepsilon}_i$ and $\boldsymbol{\varepsilon}_{i+1}$,

$$-\ln(u) > \frac{1}{2}||\boldsymbol{\varepsilon}_{i+1}||_2^2 - \frac{1}{2}||\boldsymbol{\varepsilon}_i||_2^2,$$

$$-2\ln(u) > ||\boldsymbol{\varepsilon}_{i+1}||_2^2 - ||\boldsymbol{\varepsilon}_i||_2^2,$$

$$||\boldsymbol{\varepsilon}_{i+1}||_2^2 < ||\boldsymbol{\varepsilon}_i||_2^2 - 2\ln(u).$$

Note that since $u < 1$, $-2\ln(u)$ is a positive additive term.

An initial run of the Adaptive Metropolis algorithm with six independent parameters $p_1, \ldots, p_6$ revealed a very strong linear relationship between parameters p_4 and p_5, the parameters controlling the period and phase shift of the sinusoidal component of our model (Figure 4). We therefore imposed a linear relationship between p_4 and p_5 based upon information from the parameter covariance matrix to restrain and combine their influence in following runs with five independent parameters. While correlations also exist between p_1 and p_6, p_2 and p_6, and p_1 and p_2 (see Figure 3), these correlations are not as strong as the correlation between p_4 and p_5 (Figure 4), and thus the parameters p_1, p_2, p_3, p_4, and p_6 remain independent in our updated model. Because each of these parameters corresponds to an aspect of the real-world system, we retained them to maximize transparency.

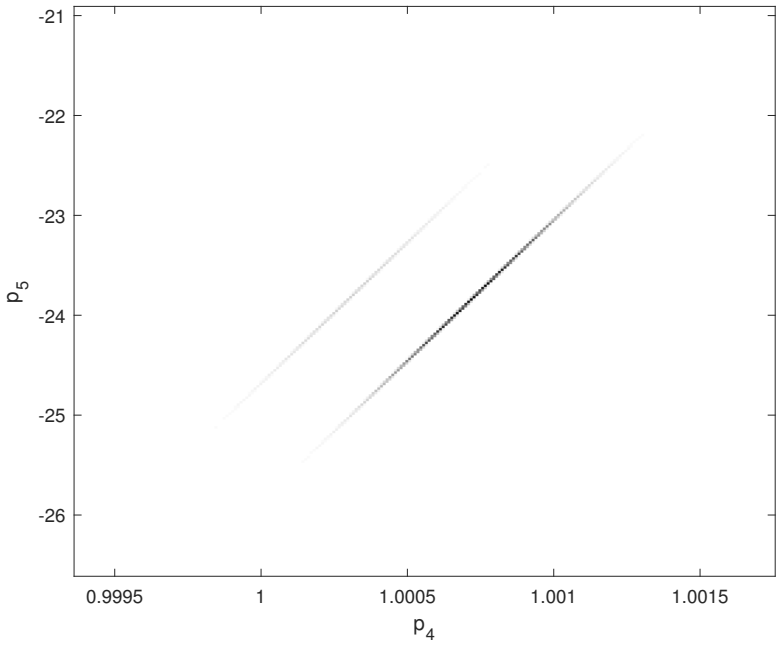

Fig. 4: Two-dimensional density of p_4 and p_5 observed on initial run of Adaptive Metropolis with six independent parameters

Note that there are two distinct linear regions of density in p_4 and p_5; we found that of the 12 parallel chains of Adaptive Metropolis we ran with six independent parameters, ten converged to distributions corresponding to the lower, more dense p_4-p_5 region, while two converged to distributions in the less dense upper region. It is plausible that since p_4 and p_5 control the period and phase shift of the seasonal sinusoidal component of our model, the best fitting values may be different during leap years, which is why we observe bimodality.

When combining the two parameters into one, we used the relationship indicated in the lower (more dense) line in Figure 4, $p_5 = -2032.5 + 2008p_4$, but the difference between the models produced under the two different relationships is negligible (Figure 5). Indeed, the relative error between the data and the model under the denser parameter set is 2.70102×10^{-3}, while the relative error between the data and the model under the secondary parameter set is 2.70099×10^{-3}, a relative difference on the order of 10^{-5}.

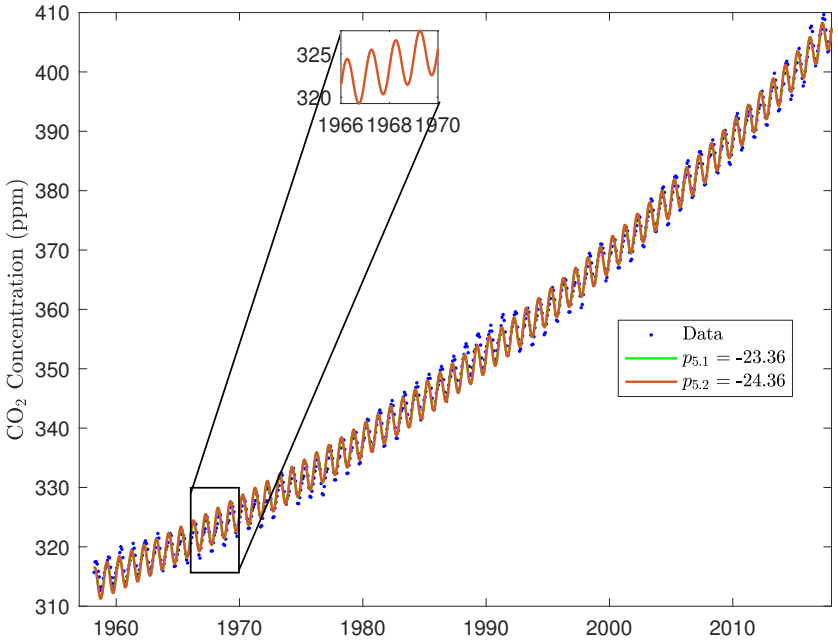

Fig. 5: Mauna Loa Data and model fits under two values of p_4 and p_5; relative error between CO_2 concentration using each parameter value is on the order of 10^-6.

Using a set of five independent parameters, we simulated 12 independent chains of realizations of $\mathbf{p}$ using Adaptive Metropolis in a parallel computer environment, with each chain of length 500,000, to obtain a sample space of 6×10^6 realizations of $\mathbf{p}$. We assumed the chains converged after 500,000 samples, based on little to no change in the parameters well before this number of samples. On each run, we used

$$\mathbf{p}_0 = \left[57.245, \quad 0.0162, \quad -2.842, \quad 1.001, \quad -23.367, \quad 256.38,\right]^\top,$$

obtained through an analytical maximum likelihood estimation, as our starting state. Here, $\mathbf{p}_0$ is the maximum likelihood estimate for $\mathbf{p}$ in that it minimizes $||\mathbf{d} - f(\mathbf{t}; \mathbf{p})||_2^2$ and as noted previously the imposed linear relationship between p_4 and p_5 still holds. We discard the first 5000 samples in each chain as our burn-in period; thus our final sample space for $\mathbf{p}$ contained 5.94×10^6 realizations of $\mathbf{p}$.

4.3. Results. The MAP estimates $\mathbf{p}_{\text{map}}$ obtained using Adaptive Metropolis, those of highest density in our sample space, for our 6 model parameters in the set of 5.94×10^6 realizations of $\mathbf{p}$ are

$$\mathbf{p}_{\text{map}} = \left[57.0155, \quad 0.016246, \quad -2.8426, \quad 1.0005, \quad -23.339, \quad 256.6250\right].$$

We found that the adjusted model for atmospheric CO_2 observations approximated data collected at the Mauna Loa Observatory with relative error $|\delta_i| = \frac{|g(t_i; \mathbf{p}_{\text{map}}) - d_i|}{|d_i|} \leq 0.0081$ over our modeled period from 1958 to 2018 (Figure 6).

Our model arrives at a CO_2 concentration of 407.2968 ppm for the final observation, which was taken at the beginning of 2018, while the data collected at Mauna Loa reflects a concentration of 407.98 ppm, indicating a relatively high degree of concordance between modeled and observed data after using the MCMC procedure to fit parameter values.

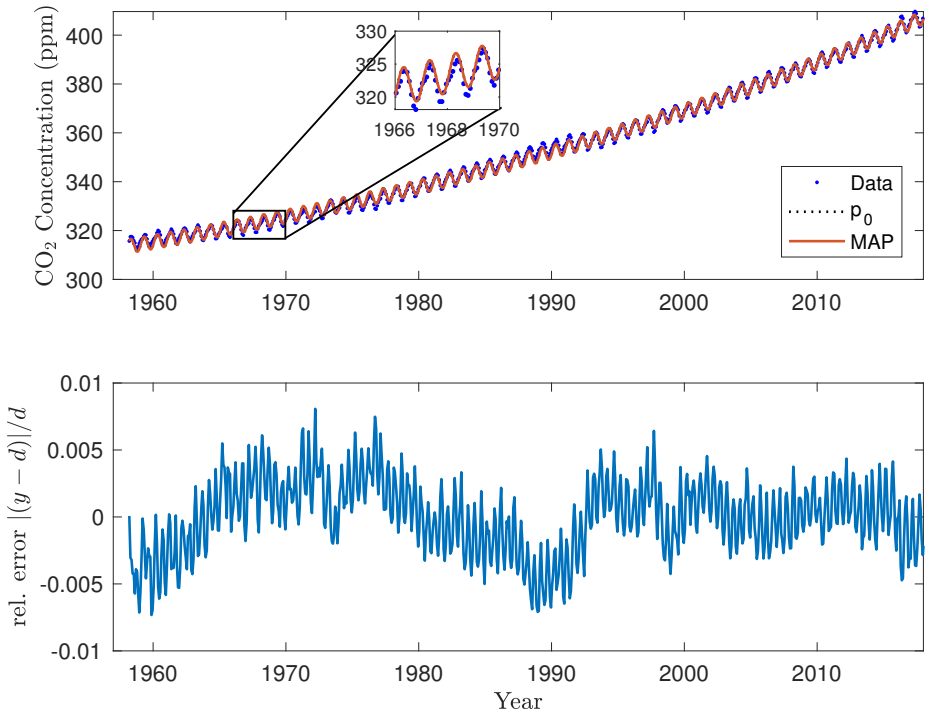

Fig. 6: Modeled CO_2 dynamics from 1958 to 2018 under $\mathbf{p}_0$ and $\mathbf{p}_{\text{map}}$ (top); relative error between CO_2 measurements at Mauna Loa and modeled results (bottom).

We produced marginal density distributions for each of our five independent parameters (p_1 through p_6, excluding p_5), displayed below (Figure 7). All marginal densities appear Gaussian. The 12 processors used to arrive for these estimates for each parameter had very high levels of agreement and demonstrated little uncertainty in their predictions.

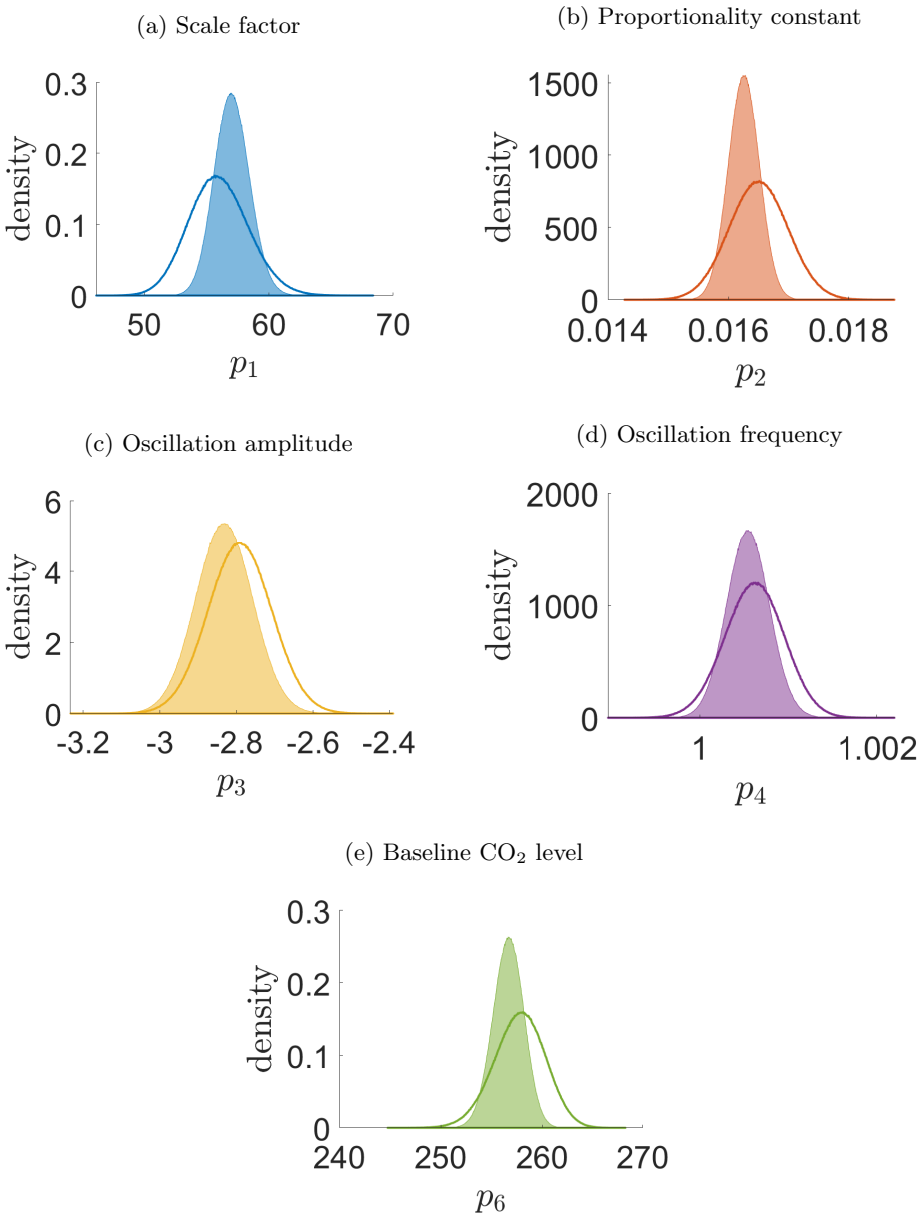

Fig. 7: Marginal densities of p_1, p_2, p_3, p_4, and p_6 resulting from MCMC on the full dataset (1958 - 2018, solid plots) and the reduced dataset (1958 - 2006, lined plots). Recall that our model is given by $A(t) = p_1 e^{p_2(t-t_0)} + p_3 \sin(2\pi p_4(t - p_5)) + p_6$.

We used our modeled relationship to project atmospheric CO_2 concentrations to the year 2120 and include a 95% prediction interval for our model (Figure 8).

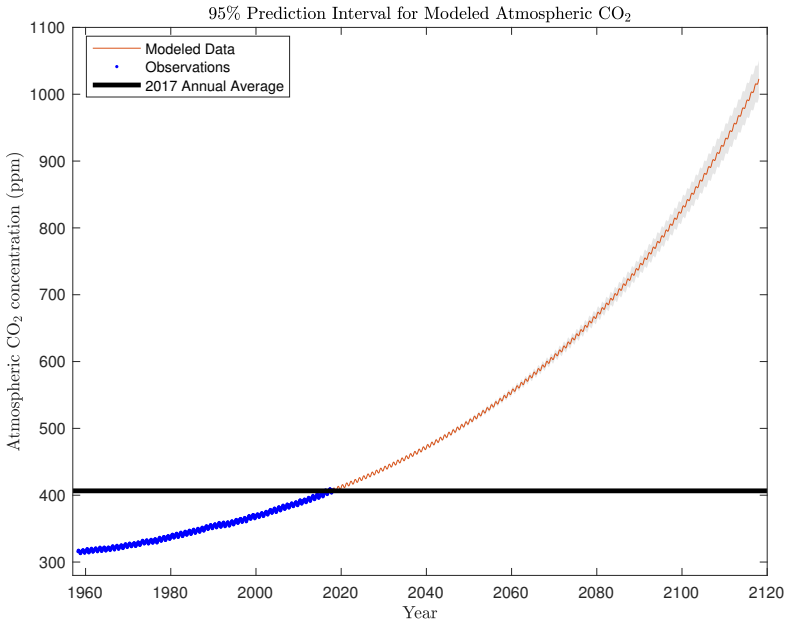

Fig. 8: Modeled CO_2 dynamics from 1958 to 2120, where the gray shaded area depicts the 95% prediction interval for the estimate.

In addition to using Adaptive Metropolis to estimate model parameters informed by our entire dataset, spanning the years 1958—2018, we also performed this parameter estimation procedure based on a reduced dataset only spanning the years 1958 – 2006. The convergence of Adaptive Metropolis in this case was still strong, the MAP estimates were similar (Table 1), and the marginal densities still appeared Gaussian, although they were slightly different from those obtained using the full dataset (Figure 7). In particular, densities obtained using the full dataset demonstrate lower variance than those from the reduced dataset, as would be expected.

	p_1	p_2	p_3	p_4	p_5	p_6
Full dataset	57.026	0.016246	-2.8426	1.0005	-23.339	256.63
Reduced dataset	55.624	0.016544	-2.8022	1.0006	-23.267	258.06
Percent Difference (%)	2.47	1.82	1.43	0.00839	0.311	0.557

Table 1: MAP estimates for each parameter obtained via MCMC on the full dataset and on the reduced dataset along with percent differences.

The most significant difference between reduced and full parameters is for p_1. The value obtained for p_1, which scales the exponential growth in our model, is significantly lower from the reduced dataset than from the full dataset (55.6 versus

57.0), suggesting a perhaps steepening growth trend.

5. Discussion and Conclusions. In this paper, we present a method of fitting data to a mathematical model via MCMC methods, and our algorithm of choice, Adaptive Metropolis. Our results indicate that carbon dioxide data collected at Mauna Loa are well-described by the model we propose [4, 9]. We base this conclusion on the confidence we have in the parameters we attained using MCMC, which we infer from the convergence of each chain of Adaptive Metropolis to the same parameter set, including arrival at parameters that would be expected based on physical reality, such as the 1-year period of the annual oscillation in CO_2. Using the selection condition described in Section 4.2, and a chain length of 500,000 for each of 12 chains, we can claim reasonable confidence in the parameters estimated via our approach. The appropriateness of our model and MCMC parameter selection is further supported by results obtained using a reduced dataset, which were in line with those obtained from the full dataset. We developed realistic parameter estimates, as evidenced by our Gaussian-shaped density distributions (Figure 7), and found these distributions to be relatively tight around expected parameter values.

We acknowledge that the seasonal rhythm observed over the course of a year is not perfectly described by a sinusoidal function. We suggest that further thought be invested into whether a better model exists, but simultaneously note that an improved model might not lend improved insight. We compare our model to approaches using modeling alone [13] and Bayesian inference alone [24], and find that our model produces a tighter prediction interval for future change in carbon dioxide concentrations while being heavily based on the observed data. These advantages create an adaptable modeling framework fit to an MCMC method that can be used to make predictions and enable management decisions. We suggest our model be used as an estimate for future carbon dioxide concentrations in the absence of social action or major environmental change, and note that the framework may be extended or modified easily to new CO_2 concentration scenarios, as will be necessary as developed and developing countries modify their carbon strategies [3, 8].

We note that caution should be used when interpreting our results in terms of the environmental impact of CO_2 increases, in particular because other studies address potential mechanisms, such as ocean thermal expansion, that are not addressed in our model, which relies upon historic "baseline" conditions [21]. However, we note that this study provides a revised look at the current course of atmospheric increase in carbon dioxide, as has long been predicted based on fossil fuel consumption [16], and posit that our parameters are modular enough to accommodate developments in our collective understanding of atmospheric regulatory mechanisms, which are to date incomplete [21].

REFERENCES

[1] *Global mean CO2 mixing ratios (ppm): Observations*. Available at https://data.giss.nasa.gov/modelforce/ghgases/Fig1A.ext.txt.

[2] *The early Keeling curve*, 2017, http://scrippsco2.ucsd.edu/history_legacy/early_{K}eeling_curve.

[3] W. S. Broecker, *Co2 arithmetic*, 2007.

[4] W. S. Cleveland, A. E. Freeny, and T. E. Graedel, *The seasonal component of atmospheric CO2: Information from new approaches to the decomposition of seasonal time series*, Journal of Geophysical Research: Oceans, 88 (1983), pp. 10934–10946.

[5] J. R. Ehleringer, T. Cerling, and M. D. Dearing, *A History of Atmospheric CO2 and Its Effects on Plants, Animals, and Ecosystems*, Ecological Studies, Springer New York, 2006, https://books.google.com/books?id=q7O7tycPzBgC.

[6] D. A. Graybill and S. B. Idso, *Detecting the aerial fertilization effect of atmospheric CO2 enrichment in tree-ring chronologies*, Global Biogeochemical Cycles, 7 (1993), pp. 81–95.

[7] H. Haario, E. Saksman, J. Tamminen, et al., *An adaptive Metropolis algorithm*, Bernoulli, 7 (2001), pp. 223–242.

[8] J. Hansen, D. Johnson, A. Lacis, S. Lebedeff, P. Lee, D. Rind, and G. Russell, *Climate impact of increasing atmospheric carbon dioxide*, Science, 213 (1981), pp. 957–966.

[9] C. D. Keeling, R. B. Bacastow, A. Carter, S. C. Piper, T. P. Whorf, M. Heimann, W. G. Mook, and H. Roeloffzen, *A three-dimensional model of atmospheric CO2 transport based on observed winds: 1. analysis of observational data*, Aspects of Climate Variability in the Pacific and the Western Americas, (1989), pp. 165–236.

[10] C. D. Keeling, J. Chin, and T. Whorf, *Increased activity of northern vegetation inferred from atmospheric CO2 measurements*, Nature, 382 (1996), p. 146.

[11] G. W. Koch and J. Roy, *Carbon Dioxide and Terrestrial Ecosystems*, Physiological Ecology, Elsevier Science, 1995, https://books.google.com/books?id=ecIAlOBbeqgC.

[12] G. H. Kohlmaier, E. Siré, A. Janecek, C. D. Keeling, S. C. Piper, and R. Revelle, *Modeling the seasonal contribution of a CO2 fertilization effect of the terrestrial vegetation to the amplitude increase in atmospheric CO2 at Mauna Loa observatory*, Tellus B, 41B, pp. 487–510, https://doi.org/10.1111/j.1600-0889.1989.tb00137.x, https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1600-0889.1989.tb00137.x, https://arxiv.org/abs/https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1600-0889.1989.tb00137.x.

[13] W. E. Long and J. Ernstberger, *Modeling atmospheric carbon dioxide over the United States*, SIAM Undergraduate Research Online, 6 (2013), pp. 200–207.

[14] R. H. Moss, J. A. Edmonds, K. A. Hibbard, M. R. Manning, S. K. Rose, D. P. Van Vuuren, T. R. Carter, S. Emori, M. Kainuma, T. Kram, et al., *The next generation of scenarios for climate change research and assessment*, Nature, 463 (2010), p. 747.

[15] C. D. Nevison, N. M. Mahowald, S. C. Doney, I. D. Lima, G. R. Van der Werf, J. T. Randerson, D. F. Baker, P. Kasibhatla, and G. A. McKinley, *Contribution of ocean, fossil fuel, land biosphere, and biomass burning carbon fluxes to seasonal and interannual variability in atmospheric CO2*, Journal of Geophysical Research: Biogeosciences, 113 (2008).

[16] R. Revelle and H. E. Suess, *Carbon dioxide exchange between atmosphere and ocean and the question of an increase of atmospheric co2 during the past decades*, Tellus, 9 (1957), pp. 18–27.

[17] C. Robert and G. Casella, *Monte Carlo statistical methods*, Springer Science & Business Media, 2013.

[18] C. P. Robert and G. Casella, *The Metropolis–Hastings algorithm*, in Monte Carlo Statistical Methods, Springer, 1999, pp. 231–283.

[19] C. P. Robert and S. Richardson, *Markov Chain Monte Carlo Methods*, Springer New York, New York, NY, 1998, pp. 1–25, https://doi.org/10.1007/978-1-4612-1716-9_1, https://doi.org/10.1007/978-1-4612-1716-9_1.

[20] B. W. Rust, *A mathematical model of atmospheric retention of man-made CO2 emissions*, Mathematics and Computers in Simulation, 81 (2011), pp. 2326 – 2336, https://doi.org/https://doi.org/10.1016/j.matcom.2010.12.019, http://www.sciencedirect.com/science/article/pii/S0378475410004209. MAMERN 2009: 3rd International Conference on Approximation Methods and Numerical Modeling in Environment and Natural Resources.

[21] S. Solomon, G.-K. Plattner, R. Knutti, and P. Friedlingstein, *Irreversible climate change due to carbon dioxide emissions*, Proceedings of the national academy of sciences, 106 (2009), pp. 1704–1709.

[22] P. Tans and K. Thoning, *How we measure background CO2 levels on Mauna Loa*, September

2008, https://www.esrl.noaa.gov/gmd/ccgg/about/co2_measurements.html.
[23] K. W. Thoning, D. R. Kitzis, and A. Crotwell, *Atmospheric carbon dioxide dry air mole fractions from quasi-continuous measurements at Mauna Loa, Hawaii*. http://dx.doi.org/10.7289/V54X55RG, 2016.
[24] R. B. Torrence, *Bayesian parameter estimation on three models of influenza*, master's thesis, Virginia Tech, 2017.
[25] L. A. Trevisan and F. Meira de Moura Luz, *Prey-predator modeling of CO2 atmospheric concentration*, ArXiv e-prints, (2008), https://arxiv.org/abs/0805.0819.
[26] C. K. Williams and C. E. Rasmussen, *Gaussian processes for machine learning*, The MIT Press, 2 (2006), p. 4.