OPEN

Estimating the short-time rate of change in the trend of the Keeling curve

Sven Nordebo^{1✉}, Muhammad Farhan Naeem¹ & Pieter Tans²

What exactly is the short-time rate of change (growth rate) in the trend of CO₂ data such as the Keeling curve? The answer to this question will obviously depend very much on the duration in time over which the trend has been defined, as well as the smoothing technique that has been used. As an estimate of the short-time rate of change we propose to employ a very simple and robust definition of the trend based on a centered 1-year sliding data window for averaging and a corresponding centered 1-year difference (2-year data window) to estimate its rate of change. In this paper, we show that this simple strategy applied to weekly data of the Keeling curve (1974–2020) gives an estimated rate of change which is perfectly consistent with a more sophisticated regression analysis technique based on Taylor and Fourier series expansions. From a statistical analysis of the regression model and by using the Cramér–Rao lower bound, it is demonstrated that the relative error in the estimated rate of change is less than 5 %. As an illustration, the estimates are finally compared to some other publicly available data regarding anthropogenic CO₂ emissions and natural phenomena such as the El Niño.

Are global events such as the natural climate phenomena, the major economic recessions or the COVID-19 global crisis and economic shutdown visible in the continuous monitoring of carbon dioxide in the atmosphere? This is a timely question that probably can not be given a definite answer, but which nevertheless is worthwhile to study and to discuss. This query also leads directly to the follow up question regarding how the short-time rate of change in the trend of CO₂ data can be defined and analysed.

The continuous monitoring of the atmospheric greenhouse gases as well as their historical, pre-industrial levels continues to provide an irreplaceable source of data, information and analyses that are becoming increasingly important for policymakers in their quest to mitigate the worst scenarios concerning the global warming, see e.g.^{1–5}. To this end, continuous monitoring was started by David Keeling at NOAA's Mauna Loa Observatory (MLO) in Hawaii, which gives access to well-mixed clean air at 3.5–4 km altitude. These measurements later became known as the “Keeling curve”, see e.g.⁶.

Trends and analyses of local as well as of global carbon dioxide cycles at various time-scales are continuously monitored and analysed by the National Oceanic and Atmospheric Administration (NOAA), see e.g. the curve fitting methods described in^{1,7}. In this paper, we will revisit some of the basic regression analysis techniques and perform a statistical analysis based on weekly data from the Keeling curve covering the period from 1974-05-19 to 2020-03-29, cf.⁸ as well as the master's thesis⁹. The purpose of this study is to validate a simple short-time estimate of the rate of change (or growth rate) in the trend of the CO₂ data. This simple, intuitive method is based on a centered 1-year data window to estimate the trend and a centered 1-year difference to estimate its rate of change. In fact, the latter can conveniently be interpreted as an estimation based on a differencing digital filter based on a 2-year data window. It is emphasized that the 1-year term is fundamental due to the annual periodicity of the natural global carbon dioxide cycle. It should be noted, however, that even though the simple digital filter approach takes full advantage of the 1-year term to create very effective spectral nulls, it also implies a limitation of the method in comparison to the more flexible regression techniques which can readily be adapted to any window length.

The simple digital filter approach turns out to be very stable and robust and could hence be used as a readily accessible standard definition of the short-time (1-year) trend and its rate of change. As a reference, the simple approach can always be validated by a comparison with any of the more sophisticated linear regression techniques that are available. In this study, we consider a well-established linear regression analysis based on 3 polynomial coefficients and 8 Fourier series coefficients (4 yearly harmonics) followed by smoothing (low-pass filtering) of the residuals, as is employed in⁷. The difference to the NOAA approach, however, is that we

¹Department of Physics and Electrical Engineering, Linnæus University, 351 95 Växjö, Sweden. ²NOAA Global Monitoring Laboratory, Boulder, CO 80305, USA. ✉email: sven.nordebo@lnu.se

are using here a short-time data window of only 2 years and no postfiltering of the residuals. This enables us to perform a statistical analysis of the residual errors and to establish the corresponding estimation accuracy. We will demonstrate that the resulting residual errors based on a 2-year data window are in good agreement with the statistical assumption that the weekly data from 1974 to 2020 are corrupted by additive, uncorrelated Gaussian noise. Indeed, most of the daily variability in this data is caused by weather systems that bring different air masses to MLO so that there is memory from day to day, but not from week to week. There can also be variability on a seasonal scale, caused for example by droughts, or an early spring, etc., and which will then be seen as local changes in the short-time trend. We will show that the regression analysis that is considered here is perfectly consistent (within a standard deviation of ± 0.02 ppm/year) with the simple estimation approach that has been described above. As a confirmation of our approach we have furthermore compared our results with the trend analyses performed by NOAA and found very good agreement provided that the smoothing of the residual data is adequately chosen.

The question about the smallest useful window length in the short-time estimate of the rate of change is an interesting, subtle issue. It is noted that the task to estimate the trend of the Keeling curve by averaging of continuous-time data is a well-posed problem. However, the task to differentiate the data is an ill-posed problem in the sense that the derivative does not depend continuously on the given data¹⁰. Hence, a regularization, or smoothing, is necessary. In a discrete-time linear regression analysis with a linear least squares (pseudo-inverse) solution, this ill-posedness will manifest itself as an increasingly ill-conditioned system matrix when the observation interval decreases. Obviously, the ill-conditioning also increases with an increasing number of regression parameters (such as a third derivative, etc). To obtain error estimates in such inverse problems, it is common to employ the Singular Value Decomposition (SVD) as well as very effective L-curve techniques where the residual error is plotted against some regularizing parameter^{10–12}. Due to the Gaussian assumption referred to above, we can take a more direct approach here and consider the Maximum Likelihood (ML) estimation and the associated Cramér–Rao Lower Bound (CRLB)¹³ of the pertinent parameter as the objective for the L-curve analysis. The subsequent analysis will show that a stable estimate of the rate of change can be obtained with the 2-year window yielding a relative estimation error of less than 5 %, see also⁹. Considering the long term data used in this study (1974–2020), it is finally worth noting that the error estimates that have been presented here are very much coherent with the 1.42 ± 0.02 ppm/year that was reported for the period 1974–1985, see¹.

A simple estimation strategy

Let $f_{\text{CO}_2}(j)$ denote the sequence of weekly data of the Keeling curve. We define the following estimate of the short-time trend

$$\tilde{f}_{\text{CO}_2}(i) = \frac{1}{53} \sum_{j=i-26}^{i+26} f_{\text{CO}_2}(j), \quad (1)$$

corresponding to a centered 1-year sliding window. It is noted that (1) can be interpreted as an approximation of the constant term in a Fourier series expansion of the 1-year data, or as the constant term in the corresponding 53-point Discrete Fourier Transform (DFT). Notice that the 53 week indices cover a time span of exactly 52 weeks (or 53 weeks in the sense of approximating an integral), which is a good approximation of 1 year. It is also desirable to employ an odd number of week indices to facilitate a symmetric window. The corresponding short-time rate of change can similarly be estimated from the following centered 1-year differences

$$\Delta_t \tilde{f}_{\text{CO}_2}(i) = \tilde{f}_{\text{CO}_2}(i+26) - \tilde{f}_{\text{CO}_2}(i-26) = \frac{1}{53} \left[\sum_{j=i}^{i+52} f_{\text{CO}_2}(j) - \sum_{j=i-52}^i f_{\text{CO}_2}(j) \right], \quad (2)$$

which covers data over an interval of exactly 104 weeks (approximately 2 years).

From our empirical numerical experiments we have experienced that it is crucial to employ a symmetric data window centered in the middle of an observation interval in order to obtain reliable and consistent results⁹. Hence, the averaging (1) can be interpreted as a discrete-time (digital) filtering $f_{\text{CO}_2}(i) = h(i) * f_{\text{CO}_2}(i)$ based on the symmetric (and non-causal) impulse response function

$$h(i) = \begin{cases} 1/53 & |i| \leq 26, \\ 0 & \text{otherwise,} \end{cases} \quad (3)$$

with the corresponding frequency response function

$$H(v) = \sum_i h(i) e^{-j2\pi vi} = \frac{1}{53} \frac{\sin \pi v 53}{\sin(\pi v)}, \quad (4)$$

and where v is the frequency in weeks^{−1}. Note that (4) is the Discrete Time Fourier Transform (DTFT) defined for continuous frequencies¹⁴, p. 248, in contrast to the Discrete Fourier Transform (DFT) which is defined for uniformly sampled discrete frequencies¹⁴, p. 456. The DFT is usually computed by using an algorithm referred to as the Fast Fourier Transform (FFT). Similarly, the differencing (2) corresponds to a digital filtering $\Delta_t f_{\text{CO}_2}(i) = g(i) * f_{\text{CO}_2}(i)$ with the anti-symmetric impulse response

Figure 1. Magnitudes of the frequency responses of the digital averaging and differencing filters, $H(\nu)$ and $G(\nu)$, respectively. The frequency of the yearly harmonics for $n = 1, \dots, 5$ are indicated with the vertical dashed lines.

$$g(i) = h(i + 26) - h(i - 26) = \begin{cases} 1/53 & 1 \leq i \leq 52, \\ -1/53 & -52 \leq i \leq -1, \\ 0 & \text{otherwise,} \end{cases} \quad (5)$$

and frequency response function

$$G(\nu) = \sum_i g(i) e^{-j2\pi \nu i} = H(\nu) e^{j2\pi \nu 26} - H(\nu) e^{-j2\pi \nu 26} = j2H(\nu) \sin(\pi \nu 52). \quad (6)$$

It is noted that $H(\nu)$ is a low-pass filter with spectral zeros at $\nu = m/53 \text{ weeks}^{-1}$ for integer $m \neq 0$. Similarly, $G(\nu)$ is a differentiating high-pass filter with additional spectral zeros at $\nu = m/52 \text{ weeks}^{-1}$ for all integers m . In Fig. 1 is plotted the magnitudes of the frequency responses $H(\nu)$ and $G(\nu)$. As can be seen in this figure, the ability of the filters to smooth and to remove the periodic annual variability in the data stems from the deep nulls that are created close to the yearly harmonics at $n/52.1775 \text{ weeks}^{-1}$, and where 1 year corresponds approximately to 52.1775 weeks. Note that the double zeros of $G(\nu)$ are at either side of these yearly harmonics. To validate the simple estimation strategy proposed in (1) and (2) above, a statistical regression model will be defined and analysed below, see also^{1,7,9}.

Statistical modeling

Define a time window of ΔT years covering exactly M weeks ($M + 1$ week-points) and where M is an even integer. Let the time window be centered at time t_i where i is a week-index and define the temporary time variable $\tau = t - t_i \in [-\Delta T/2, \Delta T/2]$ where t is the decimal time in years (e.g. starting at 0 some 2020 years ago). Let $f_{\text{CO}_2}(\tau_n)$ denote the correspondingly sampled weekly data of the Keeling curve at (translated) time $\tau_n \in [-\Delta T/2, \Delta T/2]$ for $n = 1, \dots, M + 1$.

Consider now the following statistical model

$$f_{\text{CO}_2}(\tau_n) = a_0 + a_1 \tau_n + \frac{a_2}{2} \tau_n^2 + \sum_{k=1}^4 \{b_k \cos(2\pi k \tau_n) + c_k \sin(2\pi k \tau_n)\} + w(\tau_n), \quad (7)$$

where $n = 1, \dots, M + 1$ and where a_j , for $j = 0, 1, 2$, are Taylor series coefficients defining the trend and its first and second derivative, and b_k and c_k the Fourier series coefficients modeling the annual changes. The remaining model error $w(\tau_n)$ is assumed to be zero mean uncorrelated Gaussian noise with variance σ^2 . The statistical model (7) is a linear regression model which is conveniently written in matrix form as

$$\mathbf{x} = \mathbf{A}\boldsymbol{\theta} + \mathbf{w}, \quad (8)$$

where $\mathbf{x}$ is an $(M + 1) \times 1$ vector representing the Keeling curve CO_2 data, $\mathbf{A}$ is an $(M + 1) \times P$ matrix representing the basis functions defined in (7) and $\boldsymbol{\theta}$ is a $P \times 1$ vector representing the model parameters $a_0, a_1, a_2, b_1, \dots, b_4, c_1, \dots, c_4$. Here, we are using $P = 11$ regression parameters as in⁷, but other model orders can easily be incorporated and investigated in a similar way. The covariance matrix of the noise is given by $\mathbf{C} = E\{\mathbf{w}\mathbf{w}^T\} = \sigma^2 \mathbf{I}$ where $E\{\cdot\}$ denotes the expectation operator, $(\cdot)^T$ the transpose and $\mathbf{I}$ the identity matrix.

The Fisher Information Matrix (FIM) and the Maximum-Likelihood (ML) estimate for this situation are given by

$$\mathcal{J} = \frac{1}{\sigma^2} \mathbf{A}^T \mathbf{A}, \quad (9)$$

and

Figure 2. Estimating the trend of the Keeling curve. The estimate $\hat{a}_0$ is based on (7) and (10) for a 2-year window and the estimate $\tilde{f}_{\text{CO}_2}$ on a 1-year window (1). The left and right endpoints of the 2-year curve fitting are indicated with the circle and the square, respectively. The estimate $\hat{a}_0$ at the central point of the window is indicated by the asterisk.

$$\hat{\theta}_{\text{ML}}(\mathbf{x}) = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{x}, \quad (10)$$

respectively, see e.g.¹³. The corresponding Cramér–Rao Lower Bound (CRLB) is given by

$$E\{(\hat{\theta}(\mathbf{x}) - \theta)(\hat{\theta}(\mathbf{x}) - \theta)^T\} \geq \mathcal{J}^{-1} = \sigma^2 (\mathbf{A}^T \mathbf{A})^{-1}, \quad (11)$$

which is valid for any unbiased estimator $\hat{\theta}(\mathbf{x})$ of θ . Since the statistical model (8) is linear and the noise is assumed to be Gaussian, the bound in (11) is in fact achieved and the ML-estimate (10) is said to be efficient, see¹³. Here, we are particularly interested in the Cramér–Rao lower bound of the two parameters a_0 (the trend of the Keeling curve) and its derivative a_1 , given by

$$\text{CRB}\{a_0\} = [\mathcal{J}^{-1}]_{11} = \sigma^2 \left[(\mathbf{A}^T \mathbf{A})^{-1} \right]_{11}, \quad (12)$$

$$\text{CRB}\{a_1\} = [\mathcal{J}^{-1}]_{22} = \sigma^2 \left[(\mathbf{A}^T \mathbf{A})^{-1} \right]_{22}, \quad (13)$$

i.e. the first and second diagonal elements of $\mathcal{J}^{-1}$, respectively.

Define the residual error

$$\boldsymbol{\varepsilon} = \mathbf{x} - \mathbf{A} \hat{\theta}_{\text{ML}}(\mathbf{x}) = \mathbf{P} \mathbf{x}, \quad (14)$$

where $\mathbf{P} \mathbf{x}$ is the projection onto the nullspace of $\mathbf{A}^T$, and the projection matrix is $\mathbf{P} = \mathbf{I} - \mathbf{A}(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$. Based on (8), it is readily seen that $E\{\boldsymbol{\varepsilon}\} = \mathbf{0}$, and

$$\mathbf{C}_{\boldsymbol{\varepsilon}} = E\{\boldsymbol{\varepsilon} \boldsymbol{\varepsilon}^T\} = \sigma^2 \mathbf{P}^2, \quad (15)$$

where we have employed that $E\{\mathbf{x}\} = \mathbf{A} \theta$ and $\mathbf{P} \mathbf{A} = \mathbf{0}$. Based on (15), we see that the variance of the noise can be determined as

$$\sigma^2 = \frac{\text{tr}\{\mathbf{C}_{\boldsymbol{\varepsilon}}\}}{\text{tr}\{\mathbf{P}^2\}}, \quad (16)$$

where $\text{tr}\{\cdot\}$ denotes the trace of a matrix. It is finally emphasized that the results (9) through (13) depend on the Gaussian assumption, whereas the only requirement for (15) and (16) is that the least squares method (10) is employed and that the noise has zero mean.

Evaluation of the statistical model

Comparison of estimates. The estimates (10) can now be repeated based on an $M + 1$ point sliding window defined by a sequence of weekly sampled central points t_i , as described above. In Fig. 2 is illustrated the Keeling curve based on noisy weekly data, together with the trend estimates $\hat{a}_0$ using (10) for a $M = 104$ week (2 year) window. The 1-year short-time trend $\tilde{f}_{\text{CO}_2}$ defined in (1) is also included for comparison. The two estimates $\hat{a}_0$ and $\tilde{f}_{\text{CO}_2}$ are indistinguishable in this plot. It is noticed, however, that the estimate $\tilde{f}_{\text{CO}_2}$ extends to half a year from the endpoint of the given data.

In Fig. 3 is plotted the two estimates of the rate of change $\hat{a}_1$ using (10) for the $M = 104$ week (2-year) window and $\Delta_t \tilde{f}_{\text{CO}_2}$ defined in (2). As can be seen, the two estimates are indistinguishable in this plot. The estimated

Figure 3. Estimates of the rate of change in the Keeling curve. Here, $\hat{a}_1$ is based on the regression analysis (7) and (10), and $\Delta_t \tilde{f}_{CO_2}$ is defined by the averaging (1) and differencing (2). These two estimates are indistinguishable in this plot. The red dashed plot (NOAA) is based on a long term trend and a 1.4 year smoothing chosen to fit our short-time estimates.

Figure 4. Estimated and theoretical values of the residual covariance matrix C_ϵ . The theoretical values are calculated as $\sigma^2 \mathbf{P}^2$.

standard deviation in the error between the two estimates for the whole period 1974–2020, is 0.02 ppm/year. In Fig. 3 we have also included a comparison to the trend analysis performed by NOAA, and which is described in detail in¹⁷. The NOAA estimate shown here is based on a long-term trend and where a smoothing of the residual data has been employed with a Full Width at Half Maximum (FWHM) of 1.4 year to match the amplitude of the short-time estimates.

Validation of the Gaussian assumption. To validate the Gaussian assumption above, we perform a statistical analysis of the residual error ϵ defined in (14). The 2-year (104 week) window is used once again and the covariance matrix C_ϵ defined in (15) is estimated by the sample covariance matrix

$$\hat{C}_\epsilon = \frac{1}{N} \sum_{j=1}^N \epsilon_j \epsilon_j^T, \quad (17)$$

where the sequence ϵ_j consists of residual vectors of dimension $(M + 1) \times 1$ sampled over $N = 23$ non-overlapping 2-year intervals in the period 1974 to 2020 and where $M = 104$. Based on (16), the variance of the noise can now be estimated as

$$\hat{\sigma}^2 = \frac{\text{tr}\{\hat{C}_\epsilon\}}{\text{tr}\{\mathbf{P}^2\}}, \quad (18)$$

giving the numerical value $\hat{\sigma}^2 = 0.145 \text{ ppm}^2$. The resulting estimates are illustrated in Fig. 4a showing the values of $\hat{C}_\epsilon$ in a comparison to the theoretical correlation values properly scaled as $\hat{\sigma}^2 \mathbf{P}^2 \sim C_\epsilon$ and which are plotted in

Figure 5. Probability density function for residual samples. The histogram is created by using 23 2-year (104 week) non-overlapping intervals from the period 1974–2020. The matching to the Gaussian density has been obtained with an estimated variance $\sigma_\varepsilon^2 = 0.130 \text{ ppm}^2$.

Figure 6. The relative error in the estimation of the rate of change in the trend of the Keeling curve (parameter a_1), plotted here as an L-curve with respect to the window length M .

Fig. 4b. The two plots in Fig. 4a,b are plotted in the same color range with max and min values according to $\hat{\mathbf{C}}_\varepsilon$. As can be seen in these figures, the estimated as well as the theoretical covariance matrices have almost Toeplitz structures, indicating that the elements of ε can be considered to be drawn from a stochastic process that is almost weekly stationary. The covariance matrices furthermore have a pronounced diagonal dominance indicating that the elements of ε are almost (but not quite) uncorrelated, in accordance with (15).

Finally, we evaluate the Gaussian assumption by stacking all the $(M + 1) \times 1$ non-overlapping residual vectors ε_j together for $j = 1, \dots, N$ where $N = 23$ and $M = 104$, and create the corresponding histogram. The corresponding samples are not completely uncorrelated, but they can be treated as being approximately uncorrelated due to the study illustrated in Fig. 4 above. We can furthermore assume that the residual samples are approximately identically distributed having the same variance, cf. the almost constant diagonal terms seen in Fig. 4. A normalized histogram over all of these residual samples are shown in Fig. 5 together with the theoretical Gaussian density corresponding to an estimated variance obtained here as $\sigma_\varepsilon^2 = 0.130 \text{ ppm}^2$. For $M = 104$ weeks we have $\text{tr}\{\mathbf{P}^2\} = 94$, and following (18) we obtain here $\hat{\sigma}^2 = \hat{\sigma}_\varepsilon^2(M + 1)/\text{tr}\{\mathbf{P}^2\} = 0.145 \text{ ppm}^2$, and which agrees perfectly with the previously obtained value.

Sensitivity analysis. To obtain a unitless measure of the estimation errors we consider the following normalized Cramér–Rao lower bounds (relative errors)

$$\bar{\sigma}_{\text{opt}}(a_0) = \frac{\sqrt{\text{CRB}\{a_0\}}}{\text{mean}\{\hat{a}_0\}}, \quad (19)$$

$$\bar{\sigma}_{\text{opt}}(a_1) = \frac{\sqrt{\text{CRB}\{a_1\}}}{\text{mean}\{\hat{a}_1\}}, \quad (20)$$

where $\text{CRB}\{a_0\}$ and $\text{CRB}\{a_1\}$ are given by (12) and (13), and where $\text{mean}\{\hat{a}_0\}$ and $\text{mean}\{\hat{a}_1\}$ are the respective mean values of the estimated parameters for the whole time interval of interest. In Fig. 6 is shown the relative error $\bar{\sigma}_{\text{opt}}(a_1)$ plotted as an L-curve with respect to the window length M . All the weekly data between 1974 and 2020 have been used to obtain the mean value, and as the variance of the noise we have employed the estimate $\hat{\sigma}^2 = 0.145 \text{ ppm}^2$.

Figure 7. Estimated rate of change $\hat{a}_1$ for data windows with length $M = 60$, $M = 80$ and $M = 104$ weeks. Notice the unstable behavior when the window length approaches the “knee” of the L-curve around $M = 60$, as seen in Fig. 6.

Figure 8. The Keeling curve f_{CO_2} and the estimated trend $\tilde{f}_{CO_2}$ (in ppm on the left y -axis), in a comparison with global CO_2 emissions from fossil-fuel burning according to CDIAC (in GtonC/year on the right y -axis).

The relative error in the estimate of the rate of change (the derivative) a_1 is more than 2 orders of magnitude ($100\times$) larger than the relative error in the estimate of the trend itself, a_0 , which is not shown here, cf. also⁹. Nevertheless, provided that the window length M (the regularizing parameter) is large enough, we can indeed have a stable estimate of the derivative a_1 . The relative errors shown in Fig. 6 agree rather well with the qualitative behavior of the estimate $\hat{a}_1$, as is illustrated in Fig. 7. At $M = 60$ weeks or less, the “knee” of the L-curve is approached and the estimation becomes unstable. With a short-time estimate of the rate of change based on a 2-year window ($M = 104$ weeks) we have a stable estimate with a relative error less than 5%. It is noted that the average slope is about 1.8 ppm/year, so that 5% corresponds to about 0.09 ppm/year. For even larger window lengths, the estimates get more and more smoothed out, as would be expected.

The short-time rate of change in context

It may be of interest to compare or to correlate the estimated rate of change in the trend of the Keeling curve with natural phenomena as well as the major anthropogenic activities. Of course, any conclusions based on such a comparison must be treated with great care considering the vast complexity of the Earth system. Nevertheless, to demonstrate the application of the estimation of trends in CO_2 data, it is worthwhile to put some of these global events together and to discuss their possible interpretations.

In Fig. 8 is shown the Keeling curve and its estimated trend over the period 1974–2020 plotted in ppm CO_2 on the left y -axis, together with estimated values of global CO_2 emissions from fossil-fuel burning according to CDIAC¹⁵ which is presented here in units of GtonC/year (Gigaton carbon per year) on the right y -axis. It is noted that the last update of CDIAC is from 2014¹⁵. The values for 2015–2018 have been extrapolated from BP review by multiplying the increase factors from BP review to the carbon numbers in CDIAC. The values for 2019–2020 are assumed to be constant. The plot in Fig. 8 can serve to illustrate the present day non-sustainability of fossil-fuel burning, which by an overwhelming consensus among climate experts in the world is the main cause of global warming and climate change⁵. Many people are confused about emissions and CO_2 . When emissions become smaller they expect CO_2 to go down. Chemical destruction in the atmosphere occurs for many chemicals on different time scales but for CO_2 it is very close to zero. Therefore, the CO_2 level in the atmosphere is the integral of all past emissions and removals¹⁶. That is why the trend of CO_2 is so smooth compared to the emissions history,

Figure 9. The estimate $\Delta_t \tilde{f}_{\text{CO}_2}$ of the rate of change in the trend of the Keeling curve in a comparison with year-to-year differences dCDIAC of the CDIAC data rescaled to the units of ppm/year shown on the left y-axis. The El Niño related SSTs are plotted as ΔT_s in $^{\circ}\text{C}$ shown on the right y-axis. The time of the Pinatubo eruption is indicated with an arrow.

why the current emissions slowdown is so hard to see, and why the diagnosis of changes of sources/sinks depends on the reliable detection of very small spatial/temporal differences of atmospheric CO_2 .

In Fig. 9 is shown the estimated rate of change in the trend of the Keeling curve $\Delta_t \tilde{f}_{\text{CO}_2}$ defined in (2) in a comparison with some major global events. The growth rates are plotted in ppm/year shown on the left y-axis. The NOAA growth rates of Fig. 3 are not shown here as they are almost indistinguishable from the estimates $\Delta_t \tilde{f}_{\text{CO}_2}$ in this plot. The historical El Niño/La Niña episodes are shown as Sea Surface Temperature (SST) anomalies in the Niño 3.4 region (5°N – 5°S , 120° – 170°W), see¹⁷. These SSTs are plotted here as ΔT_s in $^{\circ}\text{C}$ shown on the right y-axis of the figure. In the same figure, we have also plotted year-to-year differences (dCDIAC) of the CDIAC data that was shown in Fig. 8, and which has been rescaled here to units of ppm/year. It takes 2.124 GtonC to change the entire atmosphere by 1 ppm^{18,19}. If the emission change is for 1 year, then there has not been enough time for the anomaly to spread uniformly through the atmosphere. The Southern Hemisphere (SH) has only partially caught up with the NH where the emissions change originated, nor has the stratosphere. So, in 1 year the emissions are effectively diluted into 70–80 % of the atmosphere, which means that 1.5–1.7 GtonC will cause a 1 ppm change in the NH troposphere. Hence, we have normalized the year-to-year differences of CDIAC data with a factor of 1.6 to yield the differences in units of ppm/year. Finally, we have indicated the time of the Pinatubo eruption in June 1991, which is believed to have caused a climate anomaly lasting for about 2 years. It should be emphasized that the El Niño and the Pinatubo climate anomalies have been studied extensively by many researchers, and that they are considered here merely as a demonstration and a confirmation that the estimated growth rates are reliable.

The changes seen in dCDIAC seem to correlate somewhat with the major economic recessions in 1980–1983, 1990–1993, 2001–2002 and 2008–2009. However, the corresponding changes in dCDIAC are too small to be reliably connected to the rather strong fluctuations that are seen in the estimated growth rates $\Delta_t \tilde{f}_{\text{CO}_2}$ over this period. On the other hand, the timings of the El Niño SST anomalies are clearly correlated with the observed changes in the estimated growth rates, except for the times around the climate anomaly followed by the Pinatubo eruption. The years 1992 and 1993 following Pinatubo were relatively cool, which is likely to have reduced respiration from plants and soils, thus lowering atmospheric CO_2 . Furthermore, the increased portion of diffuse (scattered) light enhanced photosynthesis by reducing self-shading in plant canopies, enhancing CO_2 removal from the atmosphere²⁰. It may also be noted that the eruption of Pinatubo was estimated to produce some extra 50 Mtons of CO_2 into the atmosphere over a very short time. However, this corresponds to a change of only $0.05 \times 12/44/1.6 = 0.0085$ ppm in the atmosphere, which is too small to be observable (the ratio 12/44 corresponds to the molecular weights of carbon to carbon dioxide).

It is observed that the statistical analysis presented in this paper is restricted to one station only, the MLO in Hawaii. The main reason for choosing this particular station is the uniqueness of the extensive record of weekly data that is available since 1974, and which have made it possible to assess the Gaussian statistics with very high precision. Nevertheless, the simple estimation strategy expressed in (1) and (2) is readily applicable to globally averaged multistation data such as the daily data that is available from the NOAA/GML baseline observatories in Barrow (Alaska), Mauna Loa (Hawaii), American Samoa and the South Pole (Antarctica) with data from 2010 to 2020, cf.²¹. In Fig. 10 is shown a comparison between the weekly averages of MLO data f_{CO_2} and the corresponding globally (and weekly) averaged data from the four NOAA/GML baseline observatories which is denoted here by g_{CO_2} . In Fig. 11 is shown the corresponding growth rate estimates (2) denoted $\Delta_t \tilde{f}_{\text{CO}_2}$ and $\Delta_t \tilde{g}_{\text{CO}_2}$, respectively, and where the El Niño related SSTs are also shown in the same figure. As can be seen in Fig. 11, the local (MLO) and the global growth rates behave very similar and deviate maximally with about 0.5 ppm/year over this period. In Fig. 10 it can furthermore be seen from the timing of the annual variability (seasonal curves) that the global cycle is somewhat ahead of the local cycle at MLO. More interestingly, however, is that in Fig. 11 we can see that the growth rate of the global trend seems to be slightly delayed with respect to the growth rate of the local trend at MLO. This makes sense considering that the time-constants associated with the response to the El Niño stimulated phenomena is expected to be larger for global growth rates in comparison to the local responses.

Figure 10. Weekly averaged data f_{CO_2} from the Mauna Loa station (the Keeling curve) in a comparison with globally (and weekly) averaged data g_{CO_2} from the four NOAA/GML baseline observatories.

Figure 11. A comparison between the estimated growth rates $\Delta_t \tilde{f}_{\text{CO}_2}$ and $\Delta_t \tilde{g}_{\text{CO}_2}$ based on the Mauna Loa site and the global data, respectively. The El Niño related SSTs are plotted as ΔT_s shown on the right y-axis. Notice the slight delay of the global response $\Delta_t \tilde{g}_{\text{CO}_2}$ in comparison to the local response $\Delta_t \tilde{f}_{\text{CO}_2}$.

It may finally be of interest to consider the potential impact that the COVID-19 crisis may have on the present and future CO_2 readings. According to the International Energy Agency IEA, the global CO_2 emissions are expected to decline to reach 30.6 Gton CO_2 for 2020, almost 8 % lower than in 2019. This means an abrupt change of 2.66 Gton less CO_2 in comparison to the previous year, or 0.45 ppm using the conversion factors 3.667 for $\text{C} \rightarrow \text{CO}_2$ and 1.6 for $\text{ppm} \rightarrow \text{GtonC}$. In principle, a 0.45 ppm/year drop in the growth rate should be clearly visible in the Keeling curve from the point of view of the 5 % estimation accuracy that is presented here. However, as illustrated in Figs. 9 and 11 one must also consider the variability in the El Niño or other climate related phenomena that may cause much stronger fluctuations in the estimated growth rate. In any case, due to the centered 2-year data window which is required in order to obtain a reliable estimate of the rate of change, we can conclude that we will have to wait until 2021 to see any conclusive effects of the COVID-19 outbreak in the atmospheric CO_2 data.

Summary and conclusion

We have proposed a simple strategy for estimating the short-time rate of change in the trend of the Keeling curve. This estimate is based on a centered 1-year sliding average to estimate the trend, and a corresponding centered 2-year sliding data window with differencing to determine its rate of change. To validate this estimator, we have compared it to a more sophisticated regression analysis based on a combined Taylor and Fourier series expansion and found a very good agreement based on 3 Taylor coefficients, 8 Fourier series coefficients (4 yearly harmonics) and a 2-year data window to determine the parameters. A statistical analysis based on weekly data for the years 1974–2020 has shown that the model errors can be considered to be uncorrelated and identically Gaussian distributed. The Gaussian assumption justifies the use of standard formulas for Maximum Likelihood (ML) estimation, Fisher information and the related Cramér–Rao Lower Bounds (CRLB). An L-curve technique based on the CRLB has been used to study the accuracy and stability of the regression analysis. With the regression model that has been studied here, it is found that the limit of stable inversion for the rate of change in the trend of the Keeling curve is about 50–60 weeks, and that the 2-year window of 104 weeks yields a relative error less than 5 %. Hence, to obtain a stable, reliable and readily accessible estimate, we propose to use the simple strategy based on a 2-year data window with differencing, as described above. As a confirmation of the method we have found a very good agreement with the trend analyses performed by NOAA. As an illustration of the method, we have

demonstrated how the growth rates are visibly correlated with the El Niño climate phenomenon. And finally, as an interesting topic for future research, we would propose an in-depth study of multistation data in order to determine the statistical accuracy for the joint estimation of global growth rates.

Received: 16 June 2020; Accepted: 18 November 2020

Published online: 04 December 2020

References

- Thoning, K. W., Tans, P. P. & Komhyr, W. D. Atmospheric carbon dioxide at Manua Loa observatory 2. Analysis of the NOAA GMCC data, 1974–1985. *J. Geophys. Res.* **94**, 8549–8565. <https://doi.org/10.1029/JD094iD06p08549> (1989).
- Lüthi, D. *et al.* High-resolution carbon dioxide concentration record 650,000–800,000 years before present. *Nature* **453**, 379–382. <https://doi.org/10.1038/nature06949> (2008).
- Etheridge, D. M. *et al.* Natural and anthropogenic changes in atmospheric CO₂ over the last 1000 years from air in Antarctic ice and firn. *J. Geophys. Res. Atmos.* **101**, 4115–4128. <https://doi.org/10.1029/95JD03410> (1996).
- Rubino, M. *et al.* A revised 1000 year atmospheric $\delta^{13}\text{C}$ -CO₂ record from Law Dome and South Pole, Antarctica. *J. Geophys. Res. Atmos.* **118**, 8482–8499. <https://doi.org/10.1002/jgrd.50668> (2013).
- Stocker, T. *et al.* IPCC, 2013: *Climate Change 2013: The Physical Science. Basis Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* 1535 (Cambridge University Press, Cambridge, New York, 2013).
- Harris, D. C. Charles David Keeling and the story of atmospheric CO₂ measurements. *Anal. Chem.* **82**, 7865–7870. <https://doi.org/10.1021/ac1001492> (2010).
- Curve Fitting Methods Applied to Time Series in NOAA/ESRL/GMD. (Accessed 10 May 2020); <https://www.esrl.noaa.gov/gmd/ccgg/mb/cvfit/>.
- Dr. Pieter Tans, NOAA/ESRL (www.esrl.noaa.gov/gmd/ccgg/trends/) and Dr. Ralph Keeling, Scripps Institution of Oceanography (www.scrippsco2.ucsd.edu/). Accessed 10 May 2020.
- Naeem, M. F. *Analysis of an Ill-posed Problem of Estimating the Trend Derivative Using Maximum Likelihood Estimation and the Cramér-Rao Lower Bound*. Master's thesis, Linnæus University, 351 95 Växjö, Sweden. (2020). URN: urn:nbn:se:lnu:diva-95163.
- Kirsch, A. *An Introduction to the Mathematical Theory of Inverse Problems* (Springer, New York, 1996).
- Hansen, P. C. Analysis of discrete ill-posed problems by means of the L-curve. *SIAM Rev.* **34**, 561–580. <https://doi.org/10.1137/1034115> (1992).
- Hansen, P. C. *Discrete Inverse Problems: Insight and Algorithms* (SIAM-Society for Industrial and Applied Mathematics, Philadelphia, 2010).
- Kay, S. M. *Fundamentals of Statistical Signal Processing, Estimation Theory* (Prentice-Hall Inc, Upper Saddle River, 1993).
- Proakis, J. G. & Manolakis, D. G. *Digital Signal Processing* 4th edn. (Pearson Education, Inc., London, 2007).
- Boden, T., Marland, G. & Andres, R. *Global, Regional, and National Fossil-Fuel CO₂ Emissions* (Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, 2017).
- Tans, P. An accounting of the observed increase in oceanic and atmospheric CO₂ and an outlook for the future. *Oceanography* **22**, 26–35 (2009).
- Historical El Niño/La Niña episodes (1950–present). United States Climate Prediction Center (accessed 13 May, 2020). https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php.
- Carbon Dioxide Information Analysis Center—Conversion Tables. (Accessed 10 May 2020); <https://cdiac.ess-dive.lbl.gov/pns/convert.html>.
- Glossary: Carbon Dioxide and Climate. Environmental Sciences Division Publication No. 3532 (Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, 1990).
- Gu, L. *et al.* Response of a deciduous forest to the Mount Pinatubo eruption: enhanced photosynthesis. *Science* **299**, 2035–2038. <https://doi.org/10.1126/science.1078366> (2003).
- Ed Dlugokencky and Pieter Tans, NOAA/GML. (Accessed 10 May 2020); www.esrl.noaa.gov/gmd/ccgg/trends/.

Author contributions

S.N. wrote this paper as a reformulation and an extension of a masters thesis work authored by M.-F.N. Furthermore, P.T. has contributed with comments and suggestions regarding the Earth and environmental sciences as well as with complementing trend analyses based on the Keeling curve data. All authors have contributed and given approval to the final version.

Funding

Open Access funding provided by Linnaeus University.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to S.N.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020