Recommendation: Embrace the Possibility of Multi-Model Ensemble Approaches to Soil Carbon and GHG Emissions

Dave Gustafson
Conservation Technology Information Center
gustafson@ctic.org

2024 International SWAT Conference, Strasbourg

11 July 2024

Outline

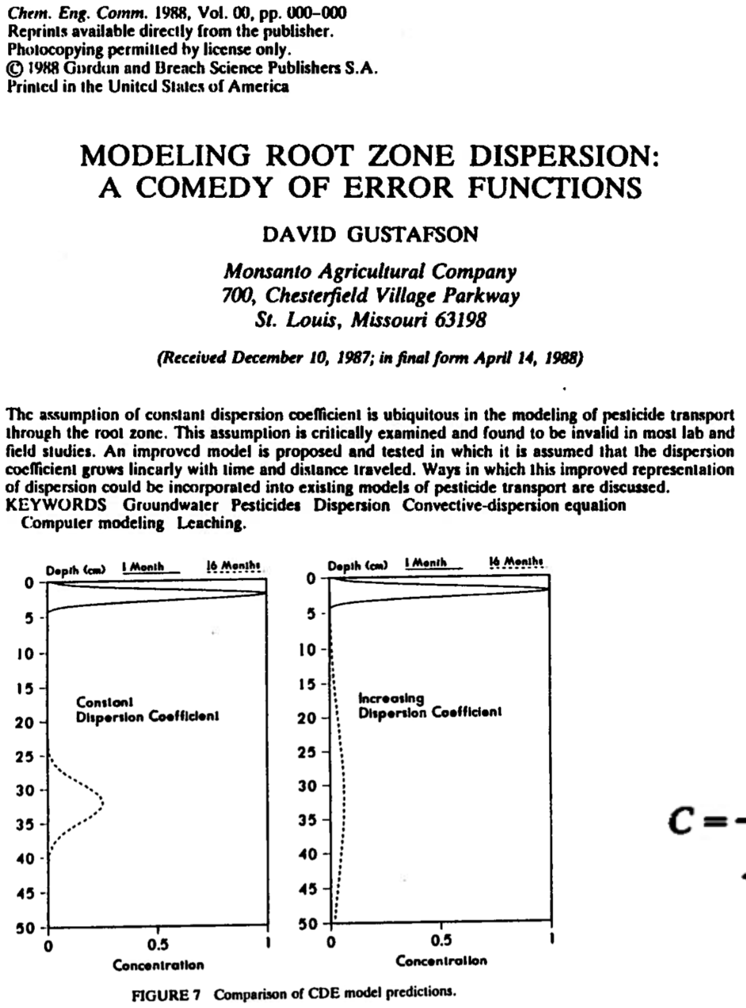
- A bit of personal history and background, with a focus on what should be relevant for further development of SWAT
- Outcomes from recent exploration of MME approaches to Soil Carbon
- Closing thoughts and recommendations

Personal History

- I began writing computer models in 1981, to simulate performance of a controlled release systemic deer repellent tablet in forest soils
- After a brief stint in Formulations at Shell, I joined Monsanto in 1985, where I was promptly placed on “special assignment” to understand EPA modeling of pesticides in ground & surface water

CDE-k Model for Dispersion

- This model (developed in 1987) assumes that the dispersion coefficient increases linearly with time and distance traveled, resulting in much better fits to observed leaching data



Chem. Eng. Comm. 1988, Vol. 00, pp. 000–000
Reprints available directly from the publisher.
Photocopying permitted by license only.
© 1988 Gordon and Breach Science Publishers S.A.
Printed in the United States of America

MODELING ROOT ZONE DISPERSION:
A COMEDY OF ERROR FUNCTIONS

DAVID GUSTAFSON

Monsanto Agricultural Company
700, Chesterfield Village Parkway
St. Louis, Missouri 63198

(Received December 10, 1987; in final form April 14, 1988)

The assumption of constant dispersion coefficient is ubiquitous in the modeling of pesticide transport through the root zone. This assumption is critically examined and found to be invalid in most lab and field studies. An improved model is proposed and tested in which it is assumed that the dispersion coefficient grows linearly with time and distance traveled. Ways in which this improved representation of dispersion could be incorporated into existing models of pesticide transport are discussed.

KEYWORDS Groundwater Pesticides Dispersion Convective-dispersion equation Computer modeling Leaching.

FIGURE 7 Comparison of CDE model predictions.

MODELING ROOT ZONE DISPERSION

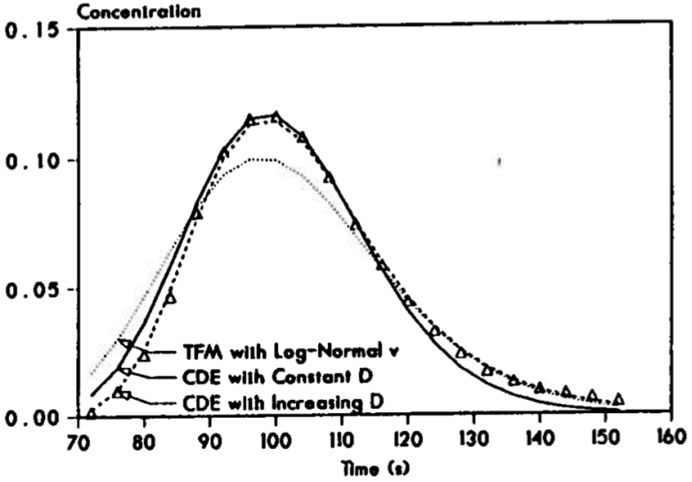
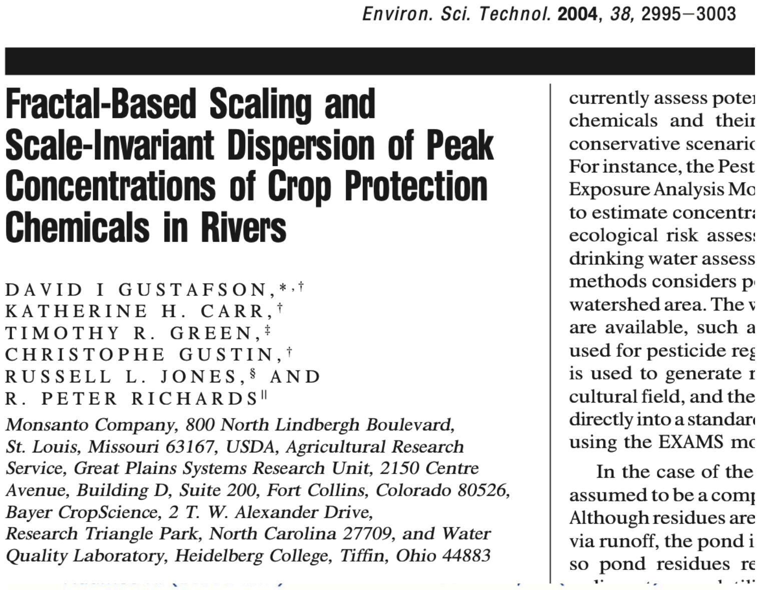


FIGURE 1 Comparison of Kan and Greenfield data with three models.

$$C=\frac{e^{-(z-vt)^2/2kv_wvt^2}}{\sqrt{\frac{\pi k v_w v t^2}{2}}\left(\operatorname{erf}\left(\frac{1}{\sqrt{2k}}\right)+1\right)}\quad z\geq 0$$

The <u>same</u> CDE-k model also fits observed watershed-scale dispersion

- In the early 2000’s, we discovered that the very same dispersion model fits observed pesticide data in surface waters, an apparent consequence of fractal behavior



Environ. Sci. Technol. **2004**, *38*, 2995–3003

Fractal-Based Scaling and Scale-Invariant Dispersion of Peak Concentrations of Crop Protection Chemicals in Rivers

DAVID I GUSTAFSON,*,†
KATHERINE H. CARR,†
TIMOTHY R. GREEN,‡
CHRISTOPHE GUSTIN,†
RUSSELL L. JONES,§ AND
R. PETER RICHARDS‖

Monsanto Company, 800 North Lindbergh Boulevard, St. Louis, Missouri 63167, USDA, Agricultural Research Service, Great Plains Systems Research Unit, 2150 Centre Avenue, Building D, Suite 200, Fort Collins, Colorado 80526, Bayer CropScience, 2 T. W. Alexander Drive, Research Triangle Park, North Carolina 27709, and Water Quality Laboratory, Heidelberg College, Tiffin, Ohio 44883

currently assess pote
chemicals and thei
conservative scenari
For instance, the Pest
Exposure Analysis Mo
to estimate concentr
ecological risk asses
drinking water assess
methods considers p
watershed area. The
are available, such a
used for pesticide reg
is used to generate
cultural field, and the
directly into a standar
using the EXAMS m

In the case of the
assumed to be a com
Although residues are
via runoff, the pond i
so pond residues r

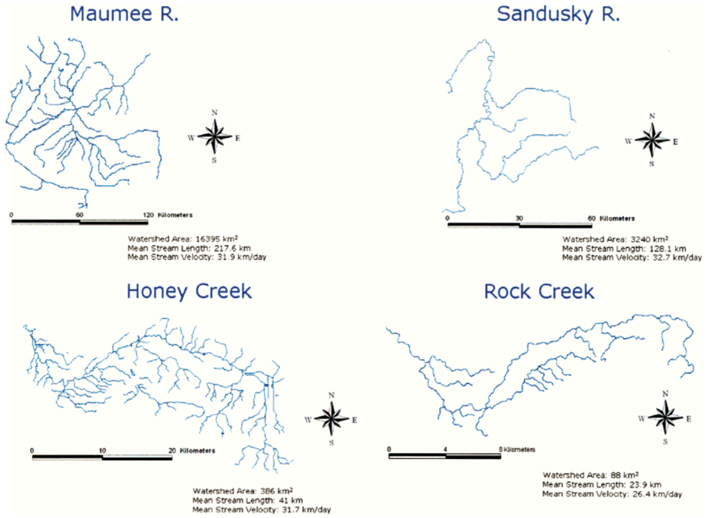


FIGURE 2. Maps show the stream networks upstream of four points being monitored by the Heidelberg College Water Quality Laboratory. The given mean stream length is the average of each branch shown here.

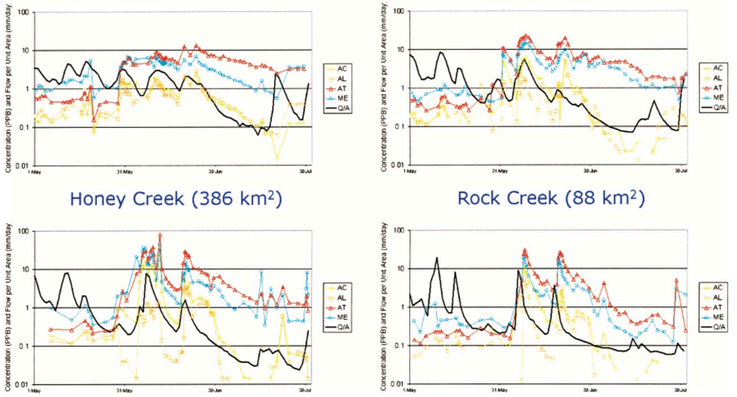


FIGURE 3. Surface water monitoring results from the Water Quality Laboratory. Each plot shows daily streamflow per unit area (Q/A) and concentrations of four herbicides: acetochlor (AC), alachlor (AL), atrazine (AT), and metolachlor (ME) during 1996, a high runoff year.

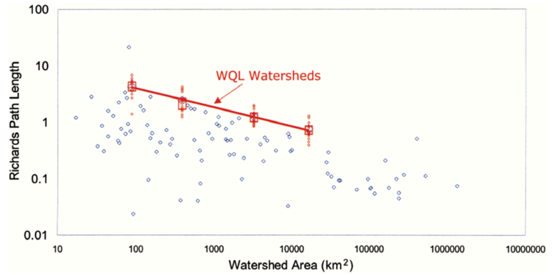


FIGURE 5. Richards path length (L_R) calculated for 1 year of streamflow data at the 89 USGS monitoring sites and during 1984–2000 at the four WQL watersheds. Quartile boxes, means, and individual annual values are plotted for the WQL data. The line connects the four mean values.

And the <u>exact same</u> CDE-k model fit nation-scale COVID outbreak data!

- In the spring of 2020, I discovered that the exact same model fit the observed patterns of national-scale COVID-19 outbreak mortality data, giving accurate projections of the first wave of the epidemic

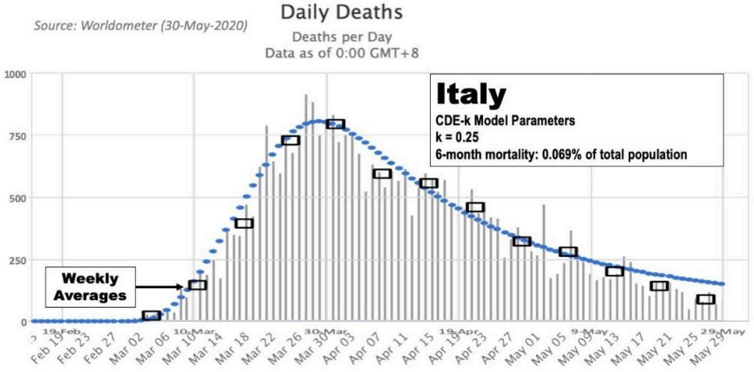
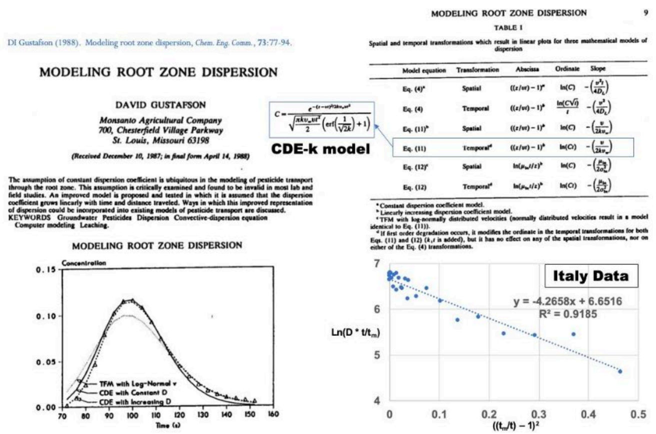


Figure 1. Application of the **CDE-k** model to observed mortality data in Italy.



DI Gustafson (1988). Modeling root zone dispersion, *Chem. Eng. Comm.*, 73:77-94.

MODELING ROOT ZONE DISPERSION

DAVID GUSTAFSON

Monsanto Agricultural Company
700, Chesterfield Village Parkway
St. Louis, Missouri 63198

(Received December 10, 1987; in final form April 14, 1988)

The assumption of constant dispersion coefficient is ubiquitous in the modeling of pesticide transport through the root zone. This assumption is critically examined and found to be invalid in most lab and field studies. An improved model is proposed and tested in which it is assumed that the dispersion coefficient grows linearly with time and distance traveled. Ways in which this improved representation of dispersion could be incorporated into existing models of pesticide transport are discussed.
KEYWORDS Groundwater Pesticides Dispersion Convective-dispersion equation Computer modeling Leaching.

$$C = \frac{e^{-(z-vt)^2/2kv_m vt^2}}{\sqrt{\frac{\pi k v_m vt^2}{2}}\left(\mathrm{erf}\left(\frac{1}{\sqrt{2k}}\right)+1\right)}$$

CDE-k model

MODELING ROOT ZONE DISPERSION 9

TABLE I

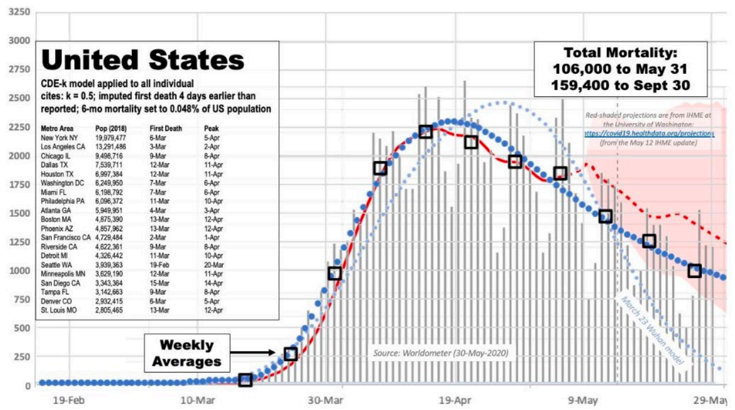
Spatial and temporal transformations which result in linear plots for three mathematical models of dispersion

Model equation	Transformation	Abscissa	Ordinate	Slope
Eq. (4)[a]	Spatial	$((z/vt)-1)^a$	$\ln(C)$	$-\left(\frac{v^2t}{4D_L}\right)$
Eq. (4)	Temporal	$((z/vt)-1)^b$	$\frac{\ln(C\sqrt{t})}{t}$	$-\left(\frac{v^2}{4D_L}\right)$
Eq. (11)[b]	Spatial	$((z/vt)-1)^a$	$\ln(C)$	$-\left(\frac{v}{2kv_m}\right)$
Eq. (11)	Temporal[d]	$((z/vt)-1)^a$	$\ln(Ct)$	$-\left(\frac{v}{2kv_m}\right)$
Eq. (12)[c]	Spatial	$\ln(\mu_m t/z)^b$	$\ln(C)$	$-\left(\frac{\mu_m}{2\sigma_m^2}\right)$
Eq. (12)	Temporal[d]	$\ln(\mu_m t/z)^b$	$\ln(Ct)$	$-\left(\frac{\mu_m}{2\sigma_m^2}\right)$

[a] Constant dispersion coefficient model.
[b] Linearly increasing dispersion coefficient model.
[c] TFM with log-normally distributed velocities (normally distributed velocities result in a model identical to Eq. (11)).
[d] If first order degradation occurs, it modifies the ordinate in the temporal transformations for both Eqs. (11) and (12) ($k_1 t$ is added), but it has no effect on any of the spatial transformations, nor on either of the Eq. (4) transformations.

MODELING ROOT ZONE DISPERSION

Figure 2. Source document (Gustafson 1988) for the **CDE-k** model.



United States

CDE-k model applied to all individual cites: k = 0.5; imputed first death 4 days earlier than reported; 6-mo mortality set to 0.048% of US population

Metro Area	Pop (2018)	First Death	Peak
New York NY	19,979,477	6-Mar	5-Apr
Los Angeles CA	13,291,486	3-Mar	2-Apr
Chicago IL	9,498,716	9-Mar	8-Apr
Dallas TX	7,539,711	12-Mar	11-Apr
Houston TX	6,997,384	12-Mar	11-Apr
Washington DC	6,249,950	7-Mar	6-Apr
Miami FL	6,198,792	7-Mar	6-Apr
Philadelphia PA	6,096,372	11-Mar	10-Apr
Atlanta GA	5,949,951	4-Mar	3-Apr
Boston MA	4,875,390	13-Mar	12-Apr
Phoenix AZ	4,857,962	13-Mar	12-Apr
San Francisco CA	4,729,484	2-Mar	1-Apr
Riverside CA	4,622,361	9-Mar	8-Apr
Detroit MI	4,326,442	11-Mar	10-Apr
Seattle WA	3,939,363	19-Feb	20-Mar
Minneapolis MN	3,629,190	12-Mar	11-Apr
San Diego CA	3,343,364	15-Mar	14-Apr
Tampa FL	3,142,663	9-Mar	8-Apr
Denver CO	2,932,415	6-Mar	5-Apr
St. Louis MO	2,805,465	13-Mar	12-Apr

Total Mortality:
106,000 to May 31
159,400 to Sept 30

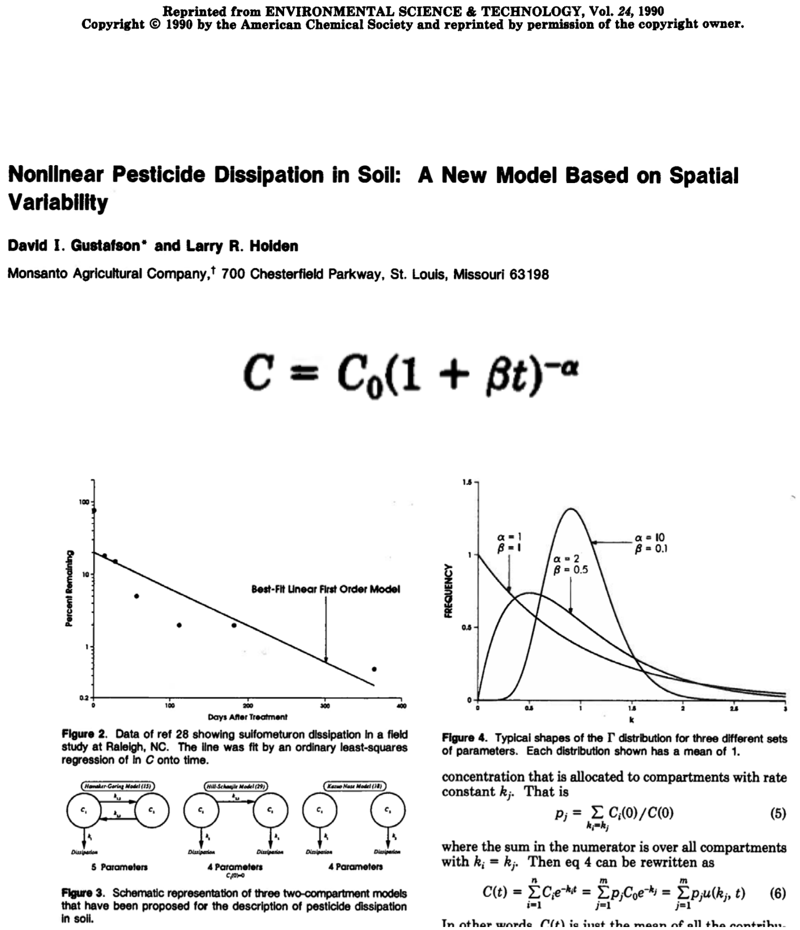
Red-shaded projections are from IHME at the University of Washington: https://covid19.healthdata.org/projections (from the May 12 IHME update)

Source: Worldometer (30-May-2020)

Figure 3. Application of the March 23 Wuhan model (light blue) and the **CDE-k** model (dark blue) to the US, overlaid by Worldometer data (gray bars, black boxes), as well as May 12 data (solid red line) and projections (red dashed line and shading) from IHME.

Nonlinear Dissipation Model

- Developed in 1989, this model is based on the idea that the traditional linear first-order dissipation constant is spatially-variable, following a 2-parameter Gamma Distribution. It fits pesticide dissipation data extremely well and has direct applications for SOC modeling.



Reprinted from ENVIRONMENTAL SCIENCE & TECHNOLOGY, Vol. 24, 1990
Copyright © 1990 by the American Chemical Society and reprinted by permission of the copyright owner.

Nonlinear Pesticide Dissipation in Soil: A New Model Based on Spatial Variability

David I. Gustafson* and Larry R. Holden

Monsanto Agricultural Company,† 700 Chesterfield Parkway, St. Louis, Missouri 63198

$$C = C_0(1 + \beta t)^{-\alpha}$$

Figure 2. Data of ref 28 showing sulfometuron dissipation in a field study at Raleigh, NC. The line was fit by an ordinary least-squares regression of ln *C* onto time.

Figure 3. Schematic representation of three two-compartment models that have been proposed for the description of pesticide dissipation in soil.

Figure 4. Typical shapes of the Γ distribution for three different sets of parameters. Each distribution shown has a mean of 1.

concentration that is allocated to compartments with rate constant k_j. That is

$$p_j = \sum_{k_i=k_j} C_i(0)/C(0) \qquad (5)$$

where the sum in the numerator is over all compartments with $k_i = k_j$. Then eq 4 can be rewritten as

$$C(t) = \sum_{i=1}^{n} C_i e^{-k_i t} = \sum_{j=1}^{m} p_j C_0 e^{-k_j} = \sum_{j=1}^{m} p_j u(k_j, t) \qquad (6)$$

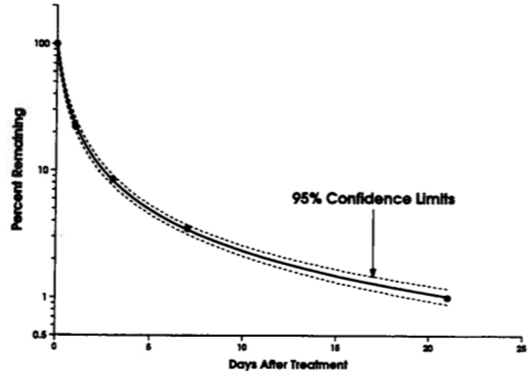


Figure 5. Data of ref 27 showing fluazifop butyl dissipation in an aerobic laboratory study. The new nonlinear model described in this work was fit to the data by the methods described herein.

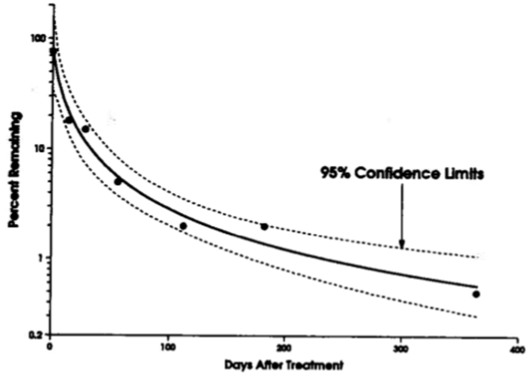
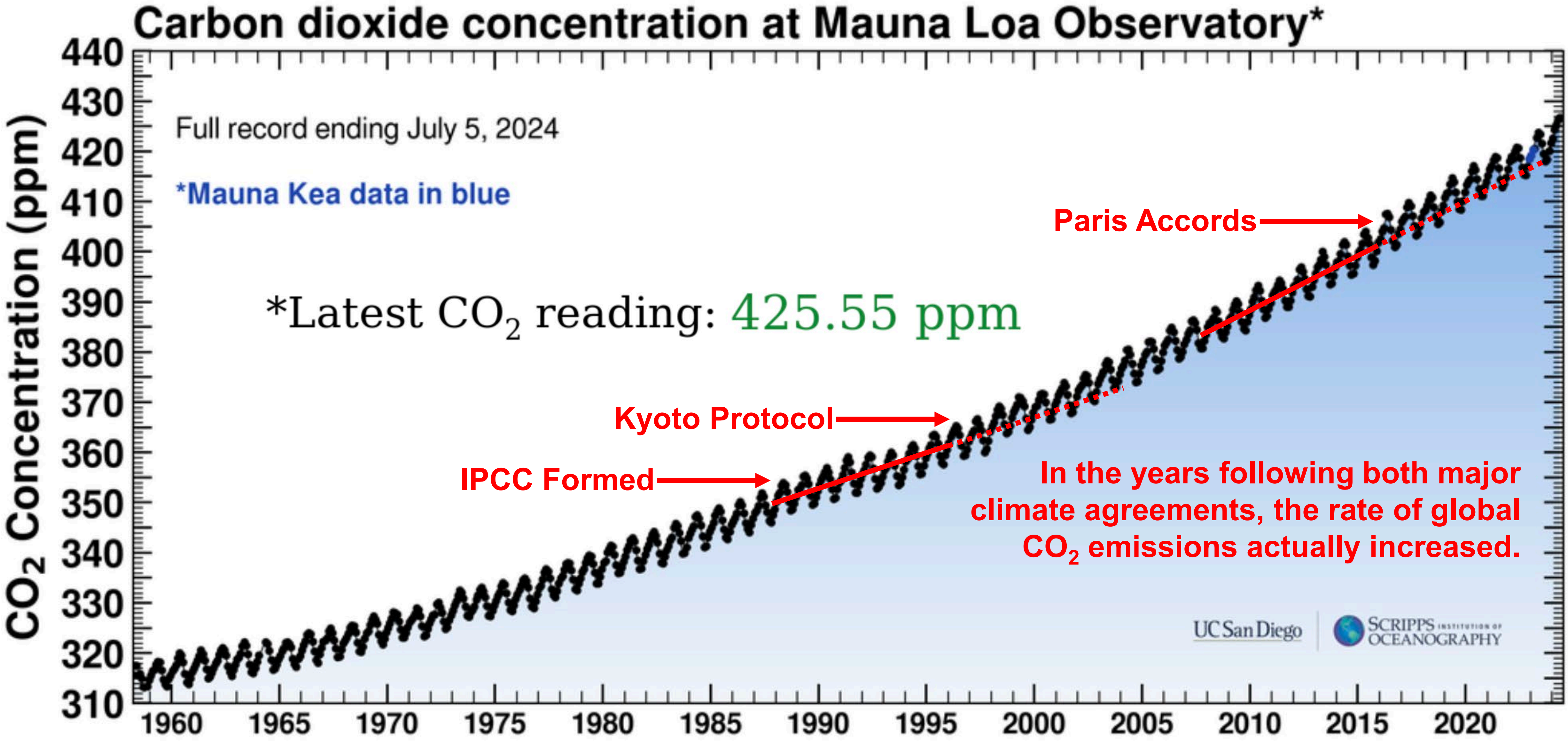
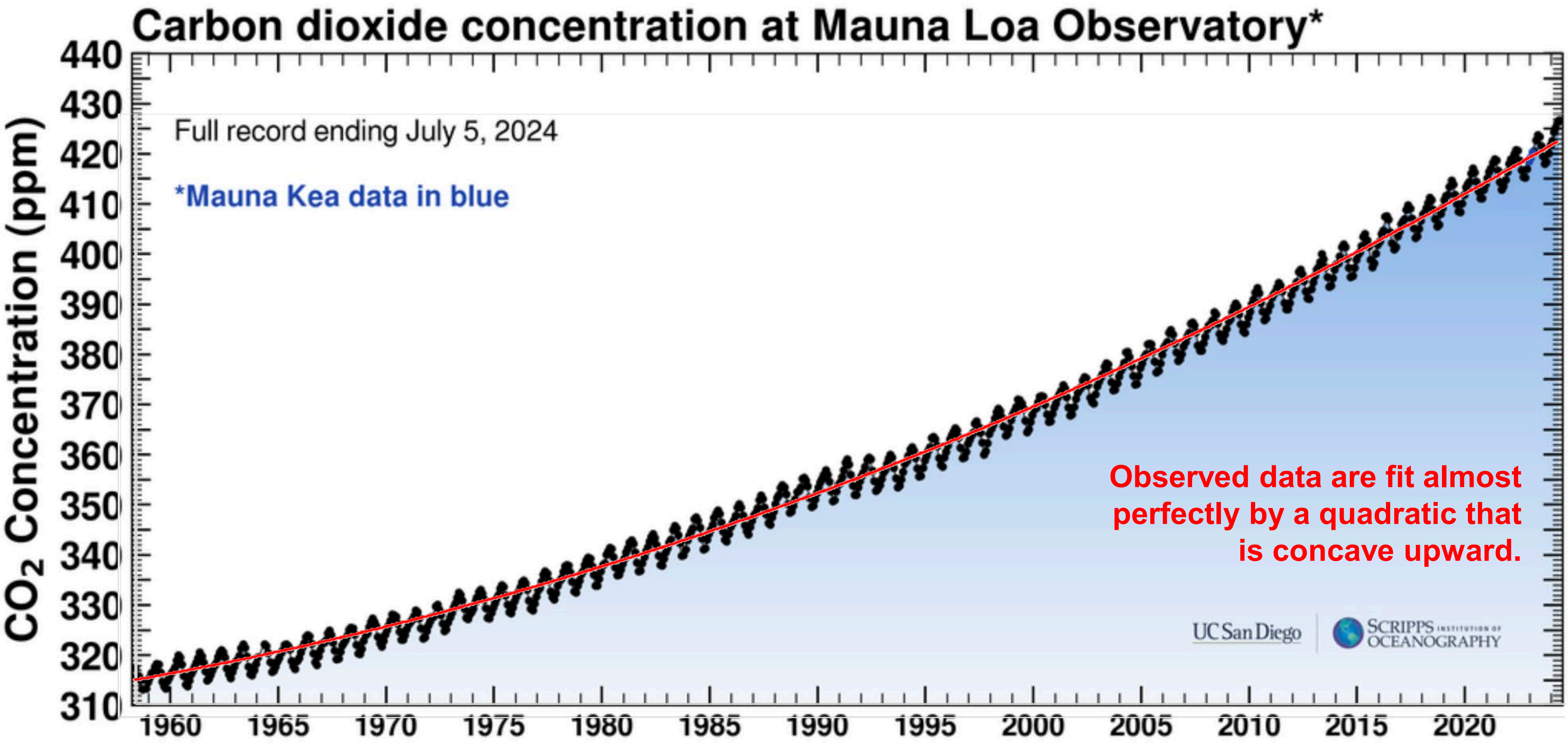


Figure 6. Data of ref 28 showing sulfometuron dissipation in a field study at Raleigh, NC. The new nonlinear model described in this work was fit to the data by the methods described herein.



In the years following both major climate agreements, the rate of global CO_2 emissions actually increased.

source: [Keeling Curve, UCSD (2024)](https://keelingcurve.ucsd.edu).



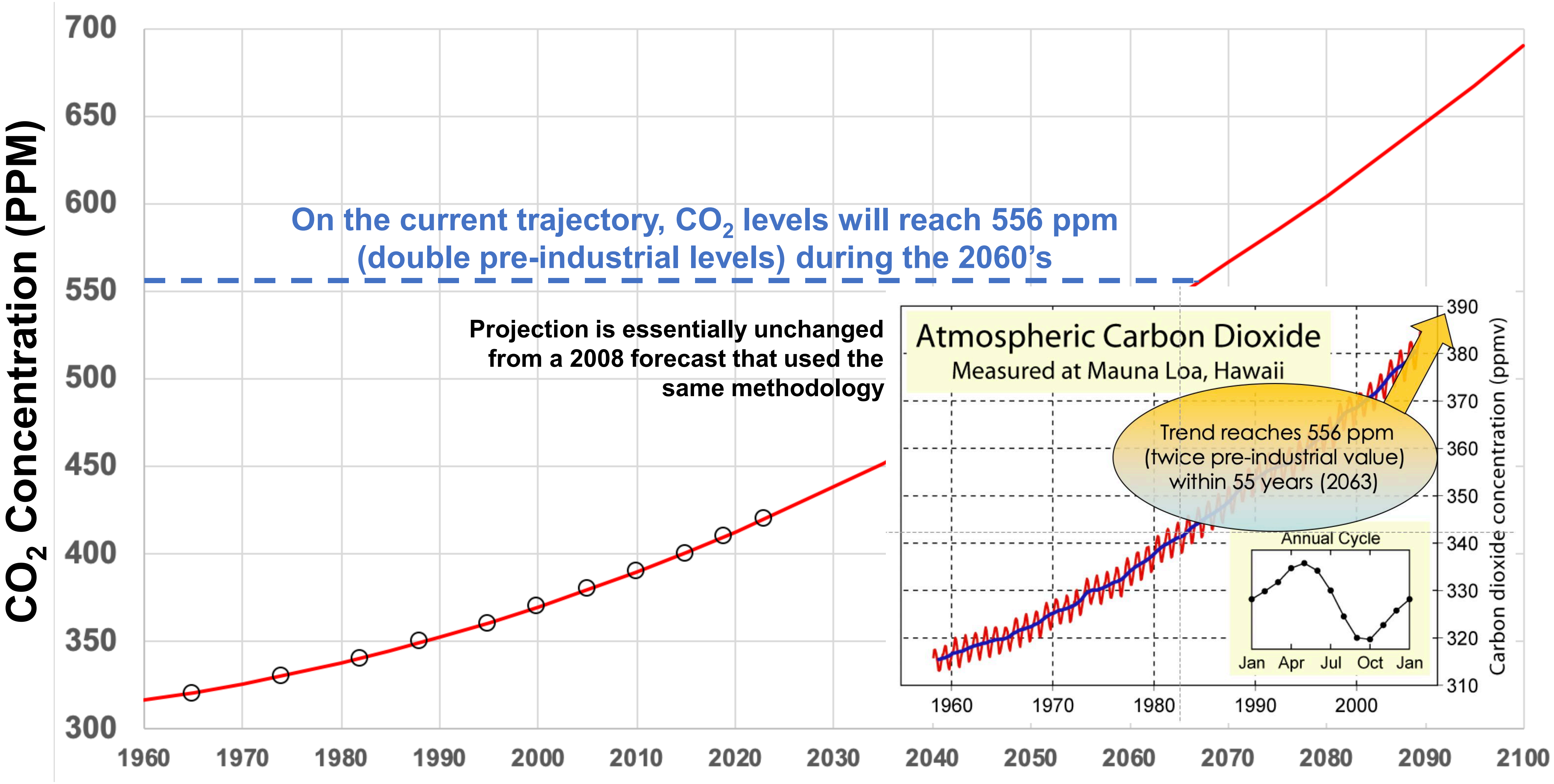
Carbon dioxide concentration at Mauna Loa Observatory*

Full record ending July 5, 2024

*Mauna Kea data in blue

Observed data are fit almost perfectly by a quadratic that is concave upward.

source: Keeling Curve, UCSD (2024).



On the current trajectory, CO$_2$ levels will reach 556 ppm (double pre-industrial levels) during the 2060's

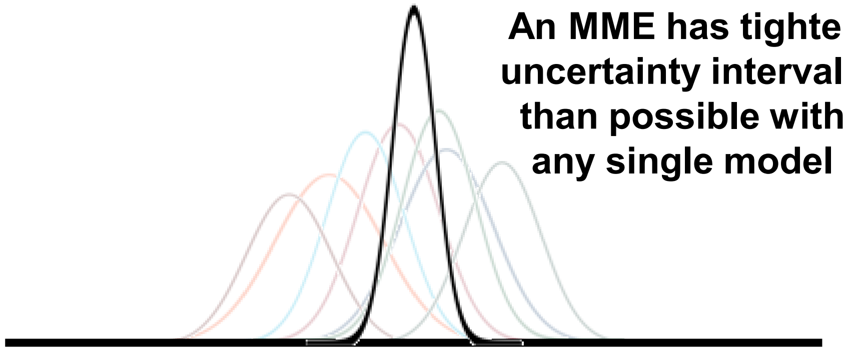
Projection is essentially unchanged from a 2008 forecast that used the same methodology

Trend reaches 556 ppm (twice pre-industrial value) within 55 years (2063)

MME Soil C Workgroup

- In early 2023, CTIC & Field to Market co-launched a workgroup to explore the feasibility of developing a multi-model ensemble (MME) approach to soil carbon
- Such approaches have ample precedent in modeling complex processes (e.g., climate, crop yields, weather, etc.)
- As demonstrated by [AgMIP]() and others in multiple [peer-reviewed studies](), the median of an MME gives better predictions than any single model (e.g., [Riggers, et al., 2019]())

Image sources: National Hurricane Center and CTIC.



An MME has tighter uncertainty intervals than possible with any single model

Image source: Gustafson (2023), [AgClimate.Net]().

Target Deliverables

- A peer-reviewed article in a first-tier journal showing the benefits of the MME approach
- An API (free to anyone) that allows any interested party to deploy the MME approach

Phase 1 Workgroup Activities

- Hybrid kickoff workshop (March, MSP)
- Poster presentation (April, EGU23)
- Panel presentation at Field to Market (June, St. Louis)
- Public comments on USDA MMRV Strategy (August, filed by CTIC)
- Field to Market Science Team presentations to Board (October) and Metrics committee (December)
- Continued discussions with Bruno Basso, at AgMIP9 (June) and AGU23 (December)

Phase 1 Workgroup Findings

- Bruno Basso has made excellent progress on a viable MME approach, which he plans to publish
- He has proposed a process for development of an API based on publicly-available models and free to use
- Important to ensure the API is fully interoperable with the FieldPrint platform and other relevant tools & datasets (e.g., the National Calibration Dataset)
- There is a continued role for the MME-Soil-C workgroup to ensure the API will meet user needs

Concerns with the Proposed API

- Is it reasonable to use uncalibrated models?
- Should individual models remain unidentified?
- Are these the right models? What about DNDC, other “real” biogeochemical models, and additional modern approaches (e.g., models based on ML, etc.)?
- Should API development and implementation be left within a single academic institution vs. a partnership involving a “real” software developer?
- Timelines have already been slipping and getting funding for API development will likely bring further delay, all at a high environmental cost. Humanity generates 0.14 Gt CO_2e each day. It takes ~400M acres of cover crops to capture that much C in one year. For mitigation to be effective, it must be fast. *Given all this urgency, is an API the next step?*
  - Related concern: The far bigger mitigation opportunity in croplands is N_2O, not C. *Is an API that doesn’t handle N_2O worth the effort?*

Phase 1 Report Conclusions

- Despite some challenging feedback, there is support for Dr. Basso's proposed prototype MME-API to proceed at MSU
- Report encourages USDA/NRCS to fund this work, in parallel with his planned publication (given the urgency)
- Workgroup to continue in some form to ensure user needs are being met (e.g., help define MME-API specifications for input data requirements, for interoperability, etc.)
- Issues to receive strong consideration as the prototype is developed and certainly before it becomes operational
  - Involve a commercial software vendor outside of academia
  - Include additional biogeochemical models
  - Model calibration, anonymity, N_2O

https://www.ctic.org/media/web/1706541941_MME-Soil-C-Phase-1-Report-DRAFT.pdf

Thoughts & Recommendations

- Though not discussed here, there are **HUGE** differences in the predictions of leading SOC/GHG models
- All forms of environmental modeling are likely to be radically transformed by AI and MME-based approaches
- As the SWAT model is enhanced to address SOC/GHG outcomes, an object-oriented modeling approach should be taken that allows for the easy inclusion of an MME-based API, once it has been made available

THANK YOU!

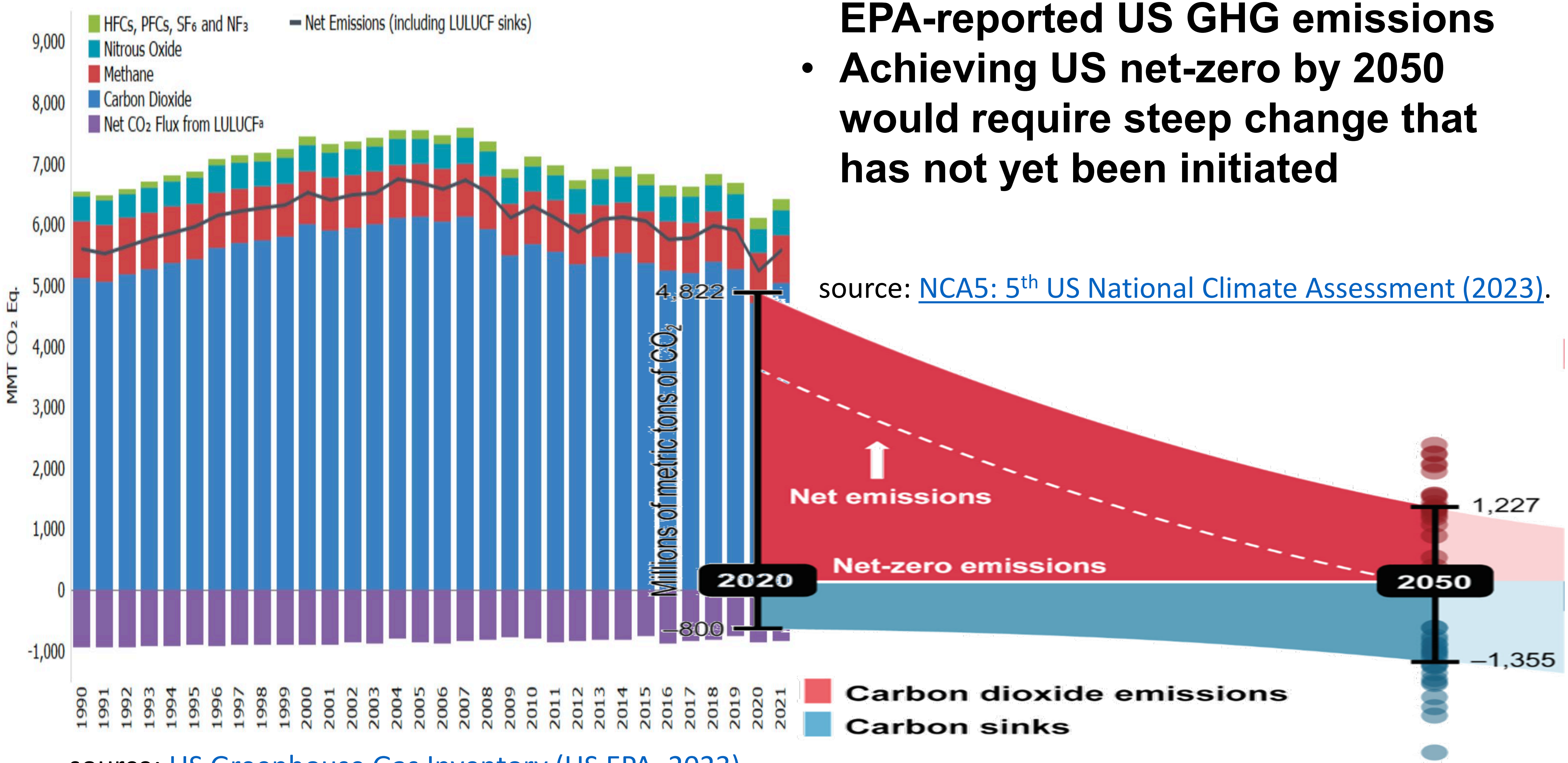
Dave Gustafson

+1-314-409-7123

gustafson@ctic.org

- **Disconnect between NCA5 and EPA-reported US GHG emissions**
- **Achieving US net-zero by 2050 would require steep change that has not yet been initiated**

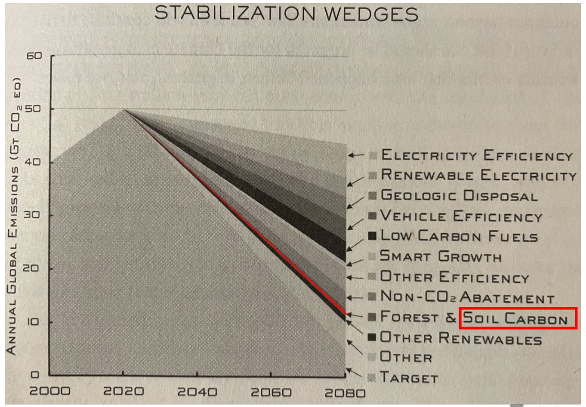
Figure ES-1: U.S. Greenhouse Gas Emissions and Sinks by Gas



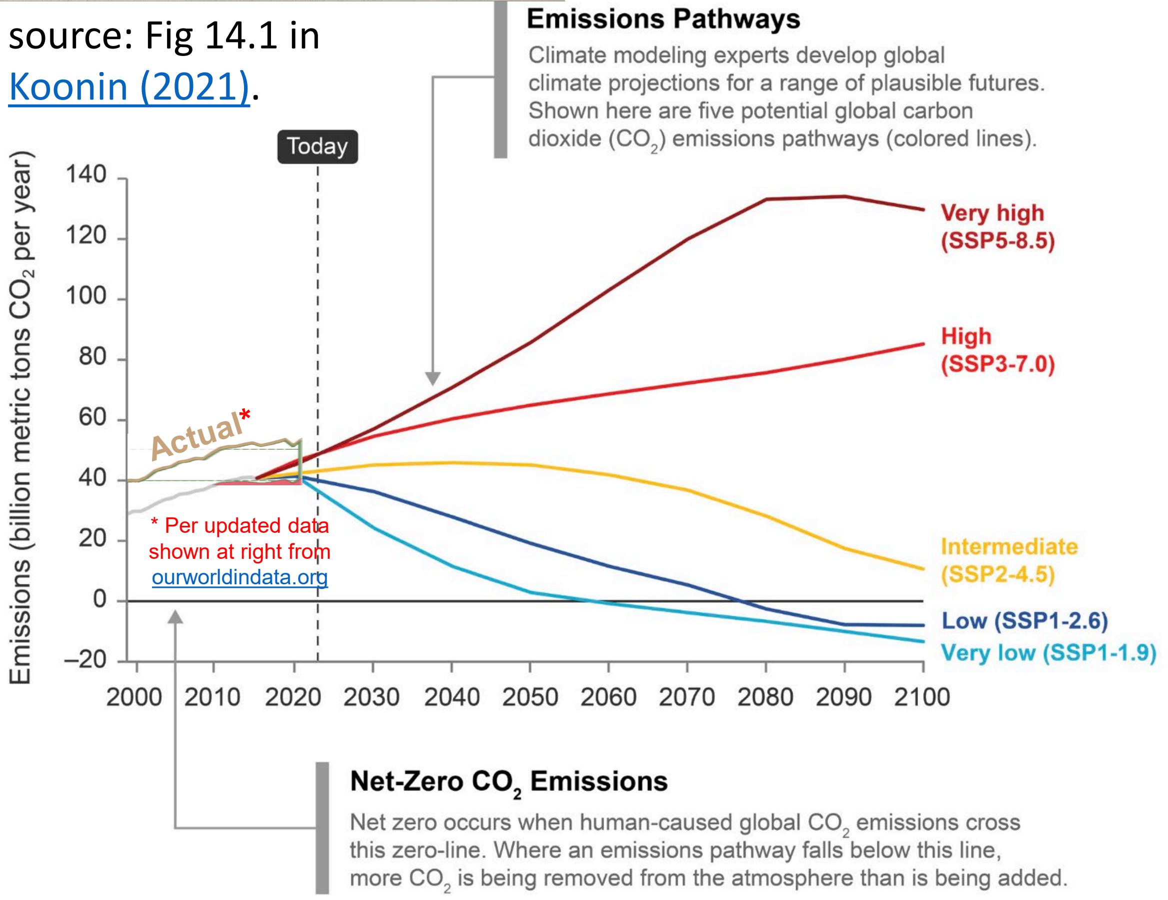
source: NCA5: 5th US National Climate Assessment (2023).

source: US Greenhouse Gas Inventory (US EPA, 2023).

- **Disconnect between NCA5 and reported global GHG emissions (now 50 Gt CO_2e per year or 0.14 Gt per day)**
- **If effective mitigation ever begins, soil carbon will play a helpful but relatively minor role; 400 M acres of cover crops (at 0.33 t/ac) require one year to offset the 0.14 Gt per day**



source: Fig 14.1 in Koonin (2021).



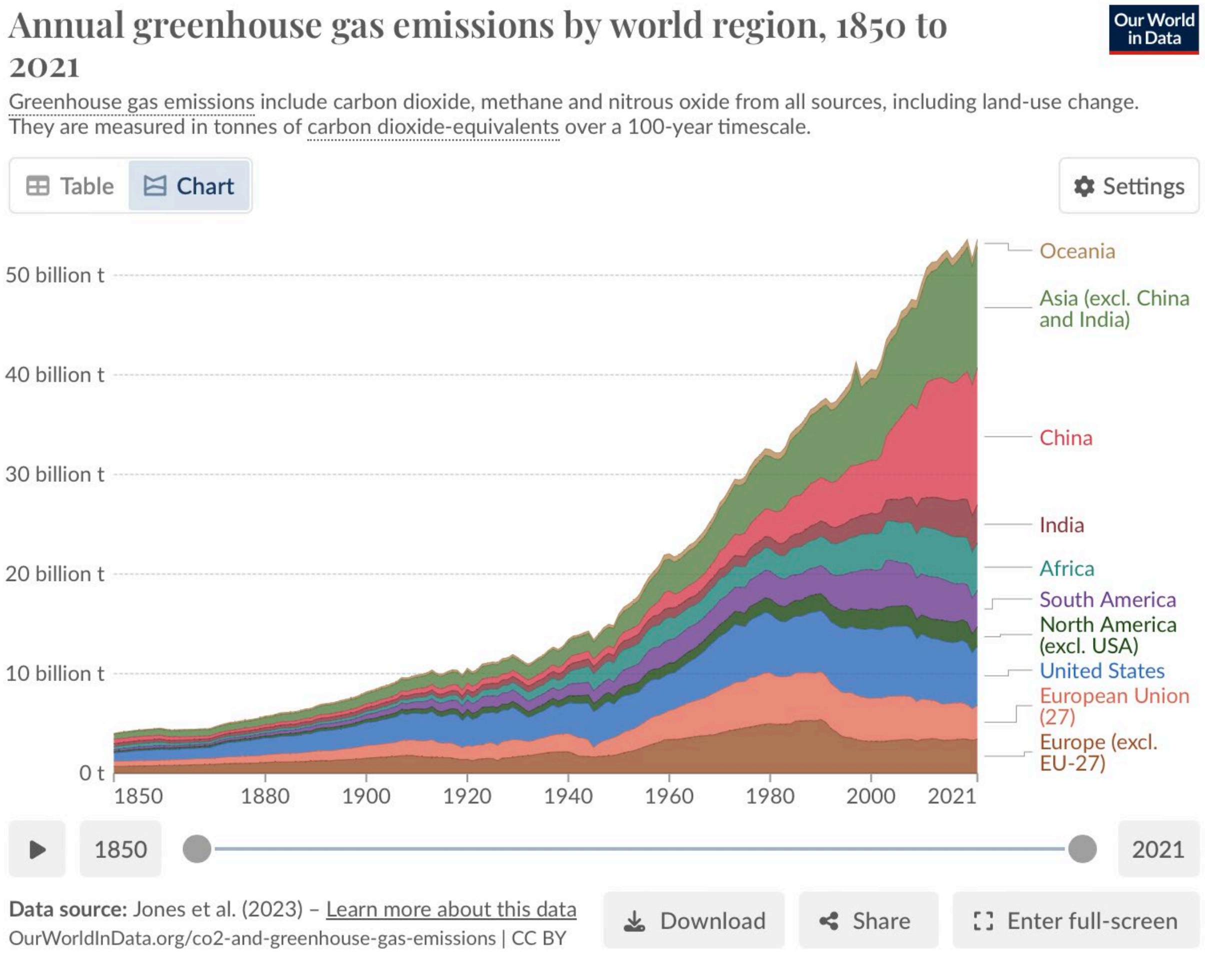
Emissions Pathways

Climate modeling experts develop global climate projections for a range of plausible futures. Shown here are five potential global carbon dioxide (CO_2) emissions pathways (colored lines).

* Per updated data shown at right from ourworldindata.org

Net-Zero CO_2 Emissions

Net zero occurs when human-caused global CO_2 emissions cross this zero-line. Where an emissions pathway falls below this line, more CO_2 is being removed from the atmosphere than is being added.



Annual greenhouse gas emissions by world region, 1850 to 2021

Greenhouse gas emissions include carbon dioxide, methane and nitrous oxide from all sources, including land-use change. They are measured in tonnes of carbon dioxide-equivalents over a 100-year timescale.

Data source: Jones et al. (2023) – Learn more about this data
OurWorldInData.org/co2-and-greenhouse-gas-emissions | CC BY

source: NCA5: 5th US National Climate Assessment (2023). source: https://ourworldindata.org/co2-and-greenhouse-gas-emissions

Workgroup Guiding Principles

- Initially co-led by Dave Gustafson (CTIC) & Paul Hishmeh (Field to Market)
- MME to be based on models that are each "publicly available, documented transparently, and based in peer-reviewed literature" (USDA Climate Smart PCSC language)
- Invite all private- and public-sector modeling teams
- IP for the API to be governed by requirements of the funders
- API itself should not be owned by a for-profit, commercial entity
- API will require no more data than running a single model

MME Workgroup Focus

- As a new quantitative measure of Soil C for the Fieldprint® Platform
- To quantify Soil C changes in USDA Climate Smart Commodity projects

Applying the best available science to this topic will ultimately result in greater accuracy, tighter confidence intervals, and higher payments for producers.

Both applications have a US-only focus – thus not initially intended for use in global initiatives (e.g., SBTi, GHG Protocol, carbon registries). However, the workgroup will include sufficient global representation such that consistency and the possibility of future applications of the approach in such domains are both maintained.

Although it is not a target of the workgroup at this time, the same modeling approach will eventually be expanded to include methane and nitrous oxide.

Kickoff Meeting Outcomes

- Meeting was well-attended, including leading public- and private-sector teams, e.g., Indigo, Regrow, Nori, and HabiTerre
- Workgroup deliverables were broadly endorsed as being valuable and achievable
- Bruno Basso (wearing his MSU/AgMIP hat) reported that he has a shareable tool which runs 7 leading soil carbon models
- Inputs currently required by the Fieldprint Platform should be sufficient to use as input to the proposed API
- Open invitation for others to join the core team, which included Dave, Paul, Jeff Lail (Syngenta), Ross Bricklemyer (Bayer) & Ellen Herbert (Ducks Unlimited)

Questions Posed at Close-out

What changes in functionality are needed for the API proposed by Michigan State U?

What changes in the proposed implementation process are needed for the API?

Is there a conditional consensus to support development of the API, as modified?

Stakeholder Feedback (Dec-2023)

- Concerns about the GHG Protocol, SBTi and related initiatives, if MME will never be adoptable in those contexts, why bother?
  - *Is this indeed a killer issue? We had previously said it shouldn’t stop us.*
- Developing an API is the “trivial” and “easy” part. *Is it worth pursuing an MME approach to improved MMRV, when there are other more pressing issues?* Here were some of the issues listed:
  - Test and improve individual models
  - Engage companies to collate soil sampling data
  - Leverage existing remote sensing data to fill data gaps
- Despite the above questions, there is support for developing a prototype of the proposed API now, subject to concerns listed on next slide

Abstract (1 of 2)

In early 2023, the Conservation Technology Information Center (CTIC), in partnership with Field to Market, launched a new workgroup to begin exploring the feasibility of developing and using an appropriate multi-model ensemble (MME) approach to modeling soil carbon in agricultural systems. The workgroup has included the research community, participants in the emerging agricultural carbon marketplace, policy-makers, foundations, and other relevant stakeholders – all led by a small core team who began meeting on a biweekly basis in February 2023. We believe that applying the best available science to this topic will ultimately result in greater accuracy, tighter confidence intervals, and higher payments for producers. Although it has not been the initial target of our workgroup, the same modeling approach should eventually be expanded to include methane and nitrous oxide. The MME approach is initially intended for two specific purposes: (1) as a future quantitative measure of soil carbon for the Fieldprint Platform; and (2) to be available as an alternative method for quantifying soil carbon changes in USDA Climate-Smart Commodity projects.

Abstract (2 of 2)

The workgroup presented its Phase 1 findings in a report published in February 2024. Despite some of the challenging feedback received from certain workgroup members, there was consensus support for the development of a prototype MME-API at Michigan State University, as proposed by Dr. Bruno Basso. The workgroup made specific suggestions for Phase 2 of the effort around model calibration, model anonymity, model calibration, and on which models to include – including SWAT+. The Phase 2 work is currently on-hold, pending the outcome of current efforts to secure the funding needed to support development work. However, whether in Dr. Basso's lab or elsewhere, it seems inevitable that the demonstrated technical advantages of an MME approach will eventually result in the availability of the tools needed to implement them. Accordingly, as the SWAT+ team continues to pursue the addition of robust soil carbon and GHG emission capabilities, it would seem wise to design its code in a way that it will be able to fully embrace such approaches once they become available.