Title
Feasibility of Formulating Ecosystem Biogeochemical Models From Established Physical Rules

Journal
Journal of Geophysical Research Biogeosciences, 129(6)

ISSN
2169-8953

Authors
Tang, Jinyun
Riley, William J
Manzoni, Stefano
et al.

Publication Date
2024-06-01

DOI
10.1029/2023jg007674

JGR Biogeosciences

INVITED ARTICLE
10.1029/2023JG007674

Special Collection:
Celebrating 20 years at JGR: Biogeosciences

Feasibility of Formulating Ecosystem Biogeochemical Models From Established Physical Rules

Jinyun Tang[1], William J. Riley[1], Stefano Manzoni[2], and Federico Maggi[3]

[1]Department of Climate Sciences, Earth and Environmental Sciences area, Lawrence Berkeley National Laboratory, Berkeley, CA, USA, [2]Department of Physical Geography and Bolin Centre for Climate Research, Stockholm University, Stockholm, Sweden, [3]Environmental Engineering, School of Civil Engineering, The University of Sydney, Sydney, NSW, Australia

Key Points:

- The popular empirically based modeling approaches limit improvement of existing ecosystem biogeochemical model predictions
- Physical rules-based approaches will help develop scaling consistent ecosystem biogeochemical models
- Inter-disciplinary collaboration can accelerate development and adoption of physical rules-based ecosystem biogeochemical models

Correspondence to:
J. Tang,
jinyuntang@lbl.gov

Citation:
Tang, J., Riley, W. J., Manzoni, S., & Maggi, F. (2024). Feasibility of formulating ecosystem biogeochemical models from established physical rules. *Journal of Geophysical Research: Biogeosciences*, *129*, e2023JG007674. https://doi.org/10.1029/2023JG007674

Received 18 SEP 2023
Accepted 15 MAY 2024

Author Contributions:
Conceptualization: Jinyun Tang, William J. Riley, Stefano Manzoni, Federico Maggi
Formal analysis: Jinyun Tang
Funding acquisition: Jinyun Tang, William J. Riley, Stefano Manzoni
Investigation: Jinyun Tang
Methodology: Jinyun Tang
Software: Jinyun Tang
Validation: Jinyun Tang
Visualization: Jinyun Tang
Writing – original draft: Jinyun Tang
Writing – review & editing: Jinyun Tang, William J. Riley, Stefano Manzoni, Federico Maggi

Abstract To improve the predictive capability of ecosystem biogeochemical models (EBMs), we discuss the feasibility of formulating biogeochemical processes using physical rules that have underpinned the many successes in computational physics and chemistry. We argue that the currently popular empirically based approaches, such as multiplicative empirical response functions and the law of the minimum, will not lead to EBM formulations that can be continuously refined to incorporate improved mechanistic understanding and empirical observations of biogeochemical processes. Instead, we propose that EBM parameterizations, as a lossy data compression problem, can be better formulated using established physical rules widely used in computational physics and chemistry, and different biogeochemical processes can be more robustly integrated within a reactive-transport framework. Through several examples, we demonstrate how mathematical representations derived from physical rules can improve understanding of relevant biogeochemical processes and enable more effective communication between modelers, observationalists, and experimentalists regarding essential questions, such as what measurements are needed to meaningfully inform models and how can models generate new process-level hypotheses to test in empirical studies. Finally, while empirical models with more parameters are often less robust, physical rules-based models can be more robust and show lower predictive equifinality, stemming from their enhanced consistency in representations of processes, interactions and spatial scaling.

Plain Language Summary Robust ecosystem biogeochemical models are needed to provide humanity with predictions to understand and manage interactions between terrestrial ecosystems and the climate. However, existing models do not fully achieve this target because of their wide use of statistical relationships derived from empirical observations. We argue that wider adoption of physical rules can help develop better ecosystem biogeochemical models to meet with society's needs. This can be achieved by deeper interdisciplinary collaboration between scientists from fields in soils, biology, chemistry, physics, and mathematics. Development of improved biogeochemical models will better position society to adapt to climate change.

1. Introduction

Biogeochemistry plays important roles in modulating greenhouse gas and energy exchanges between ecosystems and the Earth's atmosphere. Thus, it is imperative to develop ecosystem biogeochemical models (EBMs) that can deliver high quality predictions to improve understanding and management of biogeochemistry-climate feedbacks. Indeed, taking land biogeochemical models as an example, after decades of research, their representations in climate models (which are now called earth system models) have evolved from simple mathematical formulations focusing on surface energy balance to considering interactions between energy, water, carbon, and nutrient dynamics (Zhu et al., 2019), and even human management of land use and land cover (Blyth et al., 2021). Meanwhile, to reduce prediction uncertainty, more observations are collected and model-data fusion techniques are employed to constrain the parameters and process representations in these models (Houska et al., 2017; Keenan et al., 2012; Le Noë et al., 2023; Tang & Zhuang, 2008). Despite these many efforts, analyses still find significant uncertainties when model predictions are confronted with field perturbation experiments, including ecosystem responses to free air CO_2 enrichment, nutrient addition, and warming (Bouskill et al., 2014; Davies-Barnard et al., 2020; De Kauwe et al., 2017; Keenan et al., 2023; Qiu et al., 2023; Todd-Brown et al., 2013; Varney et al., 2022; Zaehle et al., 2014).

These large modeling uncertainties have been attributed to uncertain model parameters, missing or inaccurate process representations, inaccurate initial and boundary conditions, and poor numerical implementations (Ahlström et al., 2013; Bouskill et al., 2014; Huntzinger et al., 2017; Tang & Riley, 2018). We note that, in actual model applications, these four types of uncertainties often are compounded and hard to disentangle. Nonetheless, in this perspective, we infer that some fundamental fallacies in the currently popular approaches used to formulate EBMs have made it difficult, and in some cases nearly impossible, to achieve high-quality predictions. Without a fundamental change in model formulation, this challenge will persist despite efforts to augment processes representations, refine parameter calibration, integrate empirical observations, and employ more accurate numerical schemes.

The fundamental fallacies addressed below may be considered a specific type of structural error, yet they possess a unique and crucial character, deserving special attention. This assertion arises from our observation that contemporary EBM formulations heavily rely on combinations of empirical response functions derived from field observations and factorial empirical experiments. However, due to the close coupling between the involved entities, it is expected that the effect of each of these targeted biogeochemical processes only emerges from the interactions among several more basic processes, many of which cannot be orthogonally captured by factorial empirical experiments, nor be discerned from field measurements. For example, soil microbial respiration is dependent both on the microbial physiological status and substrate transport in soil, both of which are modulated by soil moisture content and temperature (Suseela et al., 2012; Zhou et al., 2014). As the transport of heat and moisture are closely coupled (Milly, 1982; Saito et al., 2006; Wu et al., 2021), changes in one of these two conditions will inevitably change the other. Moreover, both temperature and moisture alter the rates of substrate consumption, and affect substrate availability. Thus, the effects of temperature and moisture on microbial respiration are impossible to separate from field data and from laboratory experiments where substrate availability was not controlled. Consequently, using independently derived temperature and moisture response functions together could result in double counting of temperature and moisture effects on soil microbial respiration.

Meanwhile, with the rapidly changing climate and recent resurgence of fossil fuel use (Tollefson, 2022), it is becoming less likely that our society will be able to curb global warming within the 2°C limit set at the Paris Agreement (Lenton et al., 2023). We therefore expect that climate adaptation and mitigation measures through active ecosystem management will be increasingly important (Guan et al., 2023; Obersteiner et al., 2010), and our society urgently needs predictive models to provide more robust and detailed guidance on how and under what conditions such ecosystem-based measures can be properly executed.

Robust EBMs require the underlying mathematical formulations to be either simple (with fewer but well-constrained parameters) or well balanced (with less inconsistency among interacting entities), where processes are described by a complete set of physical rules (see Table 1 for a list of example physical rules that we are referring to in this perspective). Unfortunately, existing EBMs often represent biogeochemical processes without considering the underlying mechanistic details, and thus can only provide limited insights into how ecosystem management can effectively address climate adaptation and mitigation. For instance, existing models usually represent soil organic matter (SOM) as a composite of abstract and unmeasurable pools with predefined turnover times modified by edaphic conditions (Koven et al., 2013; Tao et al., 2023; Viskari et al., 2022). However, it is the diverse chemical composition of SOM and dynamic physical associations and interactions between SOM, soil particles, microbes, water, and plants that determine SOM storage and decomposition dynamics (Kleber & Adam, 2022; Lehmann et al., 2020). Thus, effective management should modulate these interactions holistically for SOM storage to be maintained or even enhanced. Models that account for many of these mechanisms are being developed (Abramoff et al., 2022; Grant et al., 2017; Riley et al., 2022; Wang et al., 2022), yet implementing them comprehensively in coupled EBMs is still a far-off goal.

Additionally, in most existing EBMs, plant canopies are at best represented with only two big leaves, one sun-lit and the other shaded (i.e., the two-big-leaf approximation; (e.g., Dai et al., 2004), while fine roots are only included implicitly via parameterizations (Wang et al., 2010; Weng et al., 2022; Zhu et al., 2019). As a result, ecosystem performance associated with different canopy structures and root traits cannot be assessed with these models. Plant models that are more explicit in their representation of plant functional traits and associated processes exist (Abichou et al., 2013; Kang et al., 2012), but their high complexity and large number of parameters make it hard to couple with ecosystem-level models that account for soil and atmospheric processes simplistically. This imbalance between complex plant models and relatively simple ecosystem models then creates a

Table 1
Example Physical Rules

Name	Domain of application	Reference
Primary rules		
Mass balance	Mass exchange	Feynman et al. (2011c)
Energy balance	Energy exchange	Feynman et al. (2011c)
Charge balance	Chemical reactions	Atkins and de Paula (2006)
Volume balance	Freeze-thaw, SOM accumulation, transpiration-induced transport, incompressible flow	Simunek and Suarez (1993); Sollins and Gregg (2017)
Momentum balance	Pressure driven mass flow	Batchelor (1967)
Entropy balance	Chemical reaction and transport	Atkins and de Paula (2006)
Derived rules		
Newton's laws of motion	Mechanic processes, for example, bacterial movement	Purcell (1977)
Maxwell's theory of electromagnetism	Radiation processes	Baldocchi et al. (1985); Ross (1981)
Quantum Mechanics	Chemical reactions	Bao and Truhlar (2017); Eyring (1935)
Thermodynamic laws	Equilibrium thermodynamic processes for chemical reactions and other processes	Atkins and de Paula (2006)
Gradient driven transport	Diffusion of mass and energy, energy dissipation	Cussler (2009); Feynman et al. (2011c)
Advective transport	Convection-driven tracer transport	Steefel et al. (2005)
Law of mass action	Chemical reactions	Kudryavtsev et al. (2001)

Note. There could be more primary rules if subatomic interactions are considered, but the six listed here are proposed to be sufficient to develop biogeochemical models. Additionally, we assume that there is no mass-energy conversion in the biochemical reactions, so that mass and energy balance rules are independent. There are many more derived rules, for example, Navier-Stokes equation and Darcy's law, each of which can be derived from these primary and secondary rules with proper mathematical approximations.

coupling challenge that calls for special care. (Just imagine how much mess a jet engine bike could make in a normal city.) Further, even if a biogeochemical or biophysical process is represented in EBMs, its mathematical description may ignore essential physical constraints (as discussed below), resulting in poor long-term predictability and lowering stakeholders' confidence in conducting mitigations based on guidance generated from model predictions (Luo et al., 2015). (In the ecological predictability framework discussed in Dietze (Dietze, 2017), such a situation is equivalent to a poor representation of the model's stability.) For example, the emerging enthusiasm in sustainable agriculture ecosystem management has stirred interest in applying ecosystem models for measuring, reporting, and verifying soil carbon changes and greenhouse gas emissions due to changes in management practices. However, stakeholders have shown diverging confidence in the models' predictive capability (Guan et al., 2023). This situation stands in stark contrast to the developments in industry, where computer-aided design software has facilitated the production of ships measuring hundreds of meters in length and chip circuits as small as a few nanometers (e.g., Arrichiello & Gualeni, 2020; Seok et al., 2021), and in weather forecasting, where reasonable weather predictions a week in advance are common (Bauer et al., 2008).

The successes of computer-aided design software in industry and numerical models in weather forecasting are founded on mathematical models formulated according to physical rules (see Table 1 for examples of these rules). The performance of these models can be continuously improved by including new essential processes (Zhou et al., 2022), adopting more robust and effective numerical solution strategies (Candel et al., 1999; Lin & Rood, 1996; Liu et al., 2019), utilizing better constrained parameters (Kotsuki et al., 2018; Wöber et al., 2020), implementing more accurate initial and boundary conditions (Saredi et al., 2021; Xiao et al., 2007), and increasing spatiotemporal resolution with more computational resources (Caldwell et al., 2021). Such a trajectory allows for the realization of "the unreasonable effectiveness of mathematics" wherein simple equations can accurately describe complex real-world phenomena (Wigner, 1960).

As ecosystem biogeochemistry is heavily influenced by living organisms (spanning micro-to macro-organisms (Madigan et al., 2009; Taiz & Zeiger, 2006), EBM modelers have avoided physical rules-based mathematical

representations (thought to be unfeasible to describe complex living organisms) and have chosen instead empirical representations. Yet, the research community now has access to an unprecedented amount of increasingly detailed observations and inferences of traits (e.g., Kattge & Sandel, 2020), and micro-climate data (e.g., Kearney et al., 2014; Novick et al., 2018). Researchers are also able to design new biological traits by gene editing (Lam et al., 2021; Saurabh, 2021), and predict intracellular biochemical rates using proteomic information (De Falco et al., 2022; Sweetlove & Ratcliffe, 2011). All of these advances are providing us with unprecedented opportunities to manage interactions with biological organisms. We thus contend that the time is ripe for developing EBMs with mechanistic representations rooted in physical rules (see Box 1 for definitions of physical rules and mechanistically based process representations). Such models will enable the assimilation of a broader range of empirical data and provide more robust numerical predictions, thus guiding ecosystem management more effectively.

The remainder of this perspective paper is organized as follows. First, we delineate the part of ecosystem biogeochemistry that will be discussed in this perspective. Second, we analyze the intrinsic limitations of two popular approaches currently used to formulate EBMs: the multiplier-based empirical response function and the law of the minimum. We highlight that these two approaches are unlikely to result in a model that can be incrementally refined as new theories and observations are developed and integrated. Third, we discuss how adopting physical rules-based approaches can lead to significant progress. In particular, we demonstrate with three examples how physical rules-based models can improve understanding of biogeochemical processes and provide opportunities for engineering biogeochemical responses. Finally, we discuss how the research community can work together to develop comprehensive and coherent EBMs based on physical rules to better realize "the unreasonable effectiveness of mathematics" in ecosystem biogeochemical modeling. We note that the mathematical symbols used in our discussion are defined in Nomenclature section.

Box 1. Definition of "Physical Rules" and "Mechanistically Based Process Representations"

Physical rules refer to fundamental principles or laws that govern the behavior and interactions of physical systems in the natural world. These rules are derived from scientific observations, experiments, theories, and mathematical models that describe the fundamental properties of matter, energy, and forces. Here, we categorize them as primary (e.g., conservation) and derived (e.g., Newton's Second Law) rules. Primary rules have also been referred to as first, fundamental, and primary principles, among other terms, in the literature. Derived rules are constructed from primary rules with known domain of applications, and are consistent with abundant observational evidence. For example, the change of momentum of a particle is proportional to the force applied, and the rate of heat conduction between two locations in space is proportional to the temperature gradient between these locations.

Mechanistically based process representations are mathematical or computational descriptions of the target process based on detailed knowledge about the underlying biological and physical mechanisms. As such, the mathematical description of a given process explicitly considers the involved entities and logical understanding of their interactions within the environment. By accounting for these interactions, they emphasize the emergent nature of the response of a process or variable to environmental changes. In application, mechanistically based representations of microbial substrate uptake could consider sub-processes including substrate transport, capture, and assimilation. The environmental dependence of each of these sub-processes can be separately described for example, by physical rules.

2. Biogeochemical Processes in Terrestrial Ecosystem Dynamics and Mathematical Rules for Scaling-Coherent Modeling

To aid our discussion of the difference between empirically based and physical rules-based approaches for formulating EBMs, we first delineate the major biogeochemical processes involved in terrestrial ecosystem dynamics and identify which of them will be within the scope of this perspective (Figure 1). We define biogeochemical processes as those that lead to the production or consumption of chemical species and biomass. Animals are excluded here, even though they likely play important roles in ecosystem nutrient dynamics (e.g., Atkinson et al., 2018). Meanwhile, biogeophysical processes and disturbances (including ecosystem management) are those

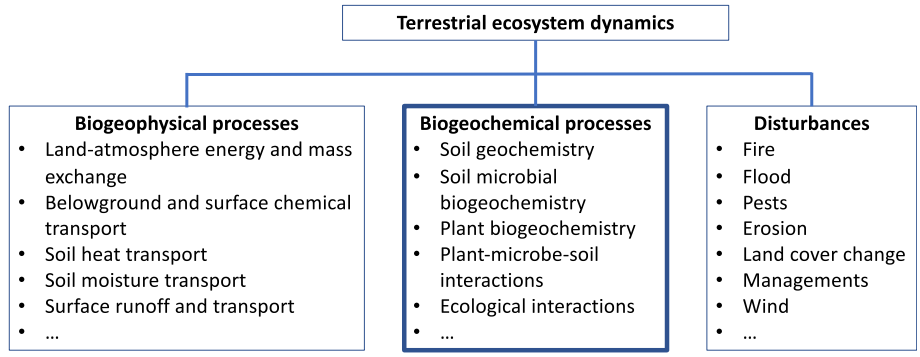

Figure 1. A general (but non-exhaustive) delineation of processes involved in terrestrial ecosystem dynamics. To highlight their importance, plant-microbe-soil interactions are separated from ecological interactions. The biogeochemical processes are the focus of this perspective.

that affect environmental conditions (e.g., soil and atmospheric temperature and water contents, soil physical and chemical properties) where biogeochemical processes occur (Robinne et al., 2020; Rusu et al., 2013). With this delineation, our discussion in this perspective focuses on how to mathematically represent biogeochemical rates and changes in storage under the influence of environmental and biological factors (e.g., Table 2).

Biogeochemical processes are always dependent on the spatial and temporal scales at which they are observed or modeled, but our discussion in this perspective leaves out the challenge of scaling across spatial heterogeneity at the landscape scale. Instead, we discuss scaling issues at a fine scale relevant for process understanding (e.g., from

Table 2
Examples of Biogeochemical Processes Used to Analyze Limitations of the Empirically Based Approaches

Process	Parameterization	Example references
Soil organic carbon decomposition	$R = R_0 f_1(T) f_2(M) f_3(O_2)$, where R_0: reference rate, $f_1(T)$: temperature dependence, $f_2(M)$: moisture dependence, $f_3(O_2)$: oxygen dependence	Azizi-Rad et al. (2022); Bauer et al. (2008)
Methane production	$R = R_0 f_1(T) f_2(pH) f_3(pE)$, where R_0: reference rate, $f_1(T)$: temperature dependence, $f_2(pH)$: pH dependence, $f_3(pE)$: redox dependence.	Riley et al. (2011); Zhuang et al. (2004)
Methane consumption	$R = R_0 f_1(CH_4) f_2(O_2) f_3(T) f_4(M)$, where R_0: reference rate, $f_1(CH_4)$: CH_4 availability dependence, $f_2(O_2)$: O_2 availability dependence, $f_3(T)$: temperature dependence, $f_4(M)$: moisture dependence.	Riley et al. (2011); Zhuang et al. (2004)
Nitrification	$R = R_0 f_1(NH_4^+) f_2(T) f_3(M) f_4(pH)$, where R_0: microbial biomass dependent reference nitrification rate, $f_1(NH_4^+)$: NH_4^+ availability dependence, $f_2(T)$: soil temperature dependence, $f_3(M)$: soil moisture dependence, $f_4(pH)$: pH dependence.	Li et al. (2000)
Denitrification	$R = R_0 f_1(T) f_2(M) f_3(pH) f_4(clay)$, where R_0: reference rate as a function NO_3^-, NO_2^-, and NO availability, $f_1(T)$: temperature dependence, $f_2(M)$: moisture dependence, $f_3(pH)$: pH dependence, $f_4(clay)$: clay content dependence.	Li et al. (2000)
Photosynthesis	$R = min(A_c, A_j, A_p) - R_d$, where A_c: carbon-limited rate, A_j: light-limited rate, A_p: triosephosphate-limited rate, R_d: dark respiration.	von Caemmerer (2013)
Stomatal conductance	$R = R_0(PAR) f_1(VPD) f_2(T_a) f_3(C_a) f_4(\psi)$, where $R_0(PAR)$: reference conductance depending on photosynthetically active radiation, $f_1(VPD)$: vapor pressure deficit dependence, $f_2(T_a)$: air temperature dependence, $f_3(C_a)$: atmospheric CO_2 dependence, $f_4(\psi)$: leaf water potential dependence.	Jarvis (1976); Yu et al. (2017)
Soil hydraulic resistance	$R = R_0 F(\theta_1)$, where R_0: reference resistance, $F(\theta_1)$: regression equation of topsoil moisture θ_1.	Kondo and Saigusa (1994); van de Griend and Owe (1994)
Microbial growth	$R = R_{max} g_1(pH)\, g_2(T)\, g_3(M) \prod_j f_j(S_j)$, or $R = R_{max} g_1(pH) g_2(T) g_3(M) min\{f_j(S_j)\}$, where R_{max}: maximum growth rate, $g_1(pH)$: pH dependence, $g_2(T)$: temperature dependence, $g_3(M)$: moisture dependence, $f_j(S_j)$: dependence of nutrient S_j.	Klausmeier et al. (2007); Leon and Tumpson (1975); Maggi et al. (2008)

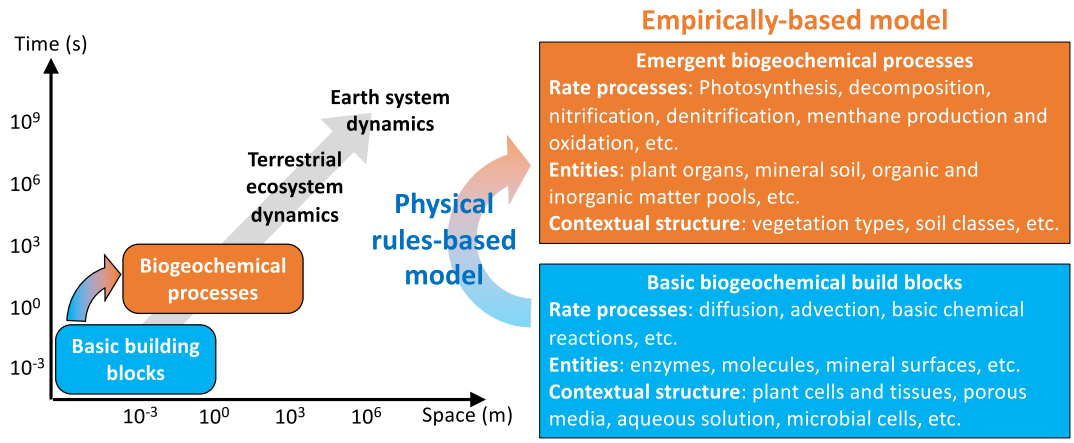

Figure 2. Space-time relationships between basic biogeochemical building blocks and the emergent biogeochemical processes constituting ecosystem and Earth system dynamics. Biogeochemical processes are represented in existing empirically based models without accounting for underlying mechanisms, while in the proposed physical rules-based models, biogeochemical processes are represented through logical combinations of basic biogeochemical building blocks that are formulated with physical rules.

soil pore to core scale; from leaf to canopy scale). At this scale, the legitimacy of a representative volume that assumes local homogeneity and equilibrium is likely ensured, as it is essential to the development of non-equilibrium thermodynamics (which includes most derived rules in Table 1) that underpin current theoretical modeling of transport, mechanics, and chemical reactions (de Groor & Mazur, 1984). We expect that improved physical rules-based modeling will facilitate better spatial heterogeneity upscaling when combined with the use of, for example, remote sensing and machine learning approaches.

Evolutionary processes (de Vries & Archibald, 2018; Greenway & Munns, 1980; Koonin & Wolf, 2012; Tan et al., 2022), and processes regulating community and ecosystem assembly (Higgins, 2017; Leibold et al., 2017) are not discussed in this perspective. These processes are linked to the biogeochemical processes discussed here, and, in a first order approximation, can be represented with similar physical rules that describe the movement and transformation of energy and chemical molecules in biogeochemical processes, except that now the functional traits of individual organisms (and their effect on biogeochemical processes) can change through time due to evolution or community-level traits change through time due to variations in community assembly (Levin, 1992; Martiny et al., 2023).

2.1. Modeling Biogeochemical Processes Across Scales

One unique feature of natural processes is that their governing equations often change across spatiotemporal scales (Figure 2). That is, there are qualitative differences between observations of emergent phenomena from fine to coarse scales, a concept termed "more is different" by Anderson (1972). For instance, the electron and charge exchange that give rise to chemical reactions at (the fine) angstrom (10^{-10} m) and femtosecond (10^{-15} s) scale are well-described by quantum mechanics (Feynman et al., 2011b; Thakkar, 2021), while chemical reactions facilitated by collisions between molecules at (the coarse) nanometer (10^{-9} m) and millisecond (10^{-3} s) scales are well-described by Newton's laws (Boltzmann, 1964; Pauli, 1973). At the micrometer (10^{-3} m) and second (10° s) scale, particle transport laws and the law of mass action are appropriate governing equations, which can be derived from Newton's law and quantum mechanics (Berg & Purcell, 1977; Feynman et al., 2011b; Kudryavtsev et al., 2001). As we further coarsen spatial and temporal scales, the higher levels of organization, nonlinearities, and variability in environmental conditions in space and time become important to biogeochemical rates (due to averaging of nonlinear processes; e.g., Chakrawal et al., 2022; Wilson & Gerber, 2021). Thus, to model biogeochemical processes robustly at scales relevant to, for example, ecosystem carbon cycling and climate adaptation, model formulations that properly account for these emergent dynamics are needed.

By recognizing that processes at a particular modeling scale emerge from processes that occur at finer scales, we expect that there are fundamental relationships between the fine and coarse scales that need to be coherently maintained when model equations are developed for the coarse scales. Suppose we are to build a model at a coarse

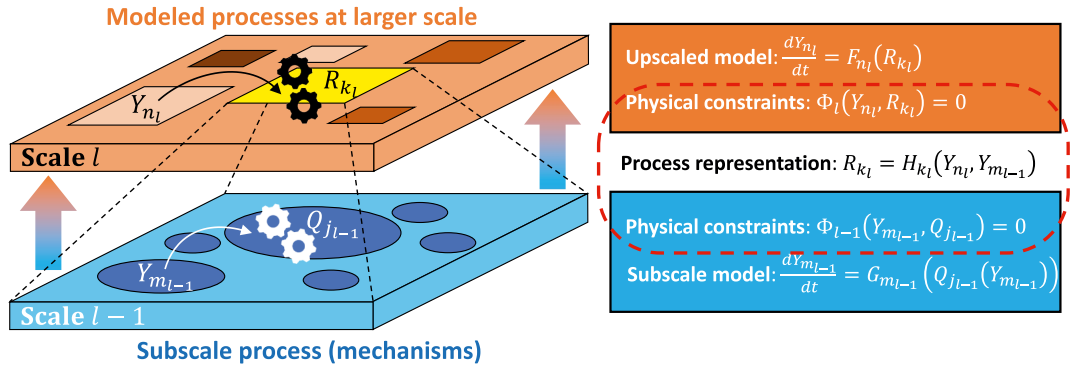

Figure 3. The relationship between biogeochemical dynamics for an upscaled model (at the scale designated by l) and subscale model (at the scale $l-1$). Here the subscale model represents processes in a spatial subset of the upscaled model. To indicate that models at scales l and $l-1$ may have different number of state variables and processes (due to lossy information compression from scale $l-1$ to scale l), subscripts n_l and k_l are used to indicate variables and processes at scale l, while subscripts $m_{l\text{-}1}$ and $j_{l\text{-}1}$ are used at scale l-1. The parameterization scheme H_{k_l} represents the net effect of process R_{k_l} in the upscaled model, which strives to represent the emergent biochemical effects from the dynamic interactions between state variables Y_{n_l} and $Y_{m_{l-1}}$ at the finer scale. Process $Q_{j_{l-1}}$ is determined by interactions between $Y_{m_{l-1}}$. The potential problem with many existing parameterizations at scale l is that the subscale physical constraints ($\Phi_{l\text{-}1}$) are ignored, so that R_{k_l} does not include interactions with $Y_{m_{l-1}}$. For instance, examples in Table 2 suffer from this problem. Another example is law of the minimum as discussed in Tang and Riley (2021).

spatiotemporal scale designated by index l (Figure 3). The state variables (Y_{n_l}) and their contributing processes (R_{k_l}) (where n_l and k_l are indices for variables and processes, respectively) are constrained by physical rules $\Phi_l\,(Y_{n_l}, R_{k_l}) = 0$, where Φ_l is a vector function. (The second law of thermodynamics is an exception, which may require $\Phi_l \geq 0$. However, in application, this can be approximated as $\Phi_l = 0$ with a proper use of phenomenological rules (de Groor & Mazur, 1984)). Each process R_{k_l} emerges from the interactions between state variables $Y_{m_{l-1}}$ that occur at the fine scale designated by index l-1 and state variables Y_{n_l} that occur at the coarse scale l. That is $R_{k_l} = H_{k_l}(Y_{n_l}, Y_{m_{l-1}})$. Meanwhile, the fine scale variables ($Y_{m_{l-1}}$) are subject to physical constraints $\Phi_{l-1}\left(Y_{m_{l-1}}, Q_{j_{l-1}}\right) = 0$ that involve fine-scale processes $Q_{j_{l-1}}$. Therefore, for a coherent formulation of the parameterization of R_{k_l} at the coarse scale l, one needs to properly maintain the physical constraints $\Phi_{l-1}\left(Y_{m_{l-1}}, Q_{j_{l-1}}\right) = 0$ at the fine scale l-1. Consequently, the extent to which those fine-scale physical constraints are maintained during upscaling determines the quality of the model parameterization at the chosen scale of interest (i.e., the coarse scale l here). Such a coherent scaling approach has been adopted in the transitions between quantum mechanics, Newton's law (Feynman et al., 2011b), Boltzmann's equation (Boltzmann, 1964), Chapman-Enskog kinetic gas theory (Chapman & Cowling, 1990), Lattice Boltzmann equation (Chen & Doolen, 1998), Navier-stokes equation (Chen & Doolen, 1998), Boussinesq equation of surface flow (Kim et al., 2009), and Richards' equation of unsaturated flow (Bear, 1972), all of which have contributed to many scientific and engineering successes. This idea is also essential to the development of nonequilibrium thermodynamics, where the macroscale equilibrium required by classical thermodynamics is allowed to emerge from local equilibriums (de Groor & Mazur, 1984).

Conceptually, deriving governing equations for the coarse scale from those of the fine scale can be seen as a lossy data (or information) compression problem, where fine-scale details are averaged out while key features are maintained at the coarse scale. (This is well illustrated in the transition from quantum mechanics to Newtonian mechanics, where the probabilistic interpretation becomes deterministic (Shankar, 1994).) Algorithms for such problems have been developed for processing image, video, and audio data (Hussain et al., 2018; Pan, 1995; Poyser et al., 2021). Methods such as multi-scale perturbation analysis and renormalization group theory, which adopt a similar spirit, have provided significant insights in the field of mathematical physics (Chen et al., 1994; Muraki et al., 1999; Roberts, 2015; Zhou, 2010). Lossy information compression is closely related to machine learning (note that renormalization group theory is closely related to deep learning (Mehta & Schwab, 2014)) and can be formulated using Bayesian inference (Cheng et al., 2018; Theodoridis, 2015). Additionally, machine learning has also been proposed for efficient upscaling (Santos et al., 2022), while Bayesian inference is frequently used to conduct model-data fusion and estimate model parameters based on observational constraints

(Tang & Zhuang, 2009; Vrugt, 2016). Therefore, in the following, the Bayesian framework (Jaynes, 2003) is used to explain the necessity and benefit of coherent scaling in formulating EBMs.

Given N upscaling scheme hypotheses A_j, $j = 1,\ldots,N$, and a set of measurements B, the Bayesian theorem ranks the merit of A_j by its posterior probability $Pr(A_j|B)$, which is

$$Pr(A_j|B) = Pr(A_j,B)/Pr(B) = Pr(B|A_j)\,Pr(A_j)/Pr(B) \quad (1)$$

where the correlation $Pr(A_j,B) = Pr(B|A_j)Pr(A_j)$ is a product of likelihood probability $Pr(B|A_j)$ and prior probability $Pr(A_j)$, and $Pr(B)$ is the probability of evidence.

As formulated in the Bayesian inference framework, identification of the best upscaling scheme becomes a model selection problem, where $Pr(A_j)$ represents the prior quality of jth upscaling scheme, whose posterior merit is $Pr(A_j|B)$ after considering its capability of matching the measurements B. For model-data fusion that aims at parameter estimation of a given model formulation, A_j is the jth sample of model parameters, whose plausibility is $Pr(A_j|B)$, and the globally optimal parameter set corresponds to the maximum of $Pr(A_j|B)$. For machine learning that uses some kind of numerical approximation, which could be neural networks, polynomials, or regression trees, A_j becomes the coefficients of the approximation method, and $Pr(A_j|B)$ ranks the goodness of model fitting conditioned on the measurements B.

Since model selection, parameter estimation, machine learning based upscaling, and model-data fusion all aim to improve EBMs, and share mathematical equivalency, we can explain the necessity and benefit of incorporating physical rules by examining the parameter estimation problem within a model-data fusion framework. This approach, in turn, reinforces the significance of physical rules in all four of these approaches. Specifically, we will show that incorporating physical rules can alleviate the predictive equifinality, which refers to the phenomenon where different model parameters, due to uncertainty in calibration, result in diverging model predictions. In a sense, the predictive equifinality is one part of the ecological predictability discussed in Dietze (2017). It differs from calibration equifinality (i.e., the conventional parametric equifinality) discussed by Beven and Freer (2001), which indicates an equally good model-data agreement obtained by different sets of model parameters during calibration.

To simplify the explanation, we assume that the model-data discrepancy follows the Gaussian distribution, as often assumed in Bayesian inference-based model-data fusion for EBMs (Tang & Zhuang, 2009; Tarantola, 2005). Accordingly, the cost function (or loss function as termed in machine learning) of model-data discrepancy for a set of model parameter values (i.e., ln $Pr(B|A)$ for parameter set A) can be written in the following plain language form:

$$\text{Cost function} = \text{Observational constraint} + \text{scaling coherency rules} + \text{prior constraint.}$$

In a more formal definition, the cost function relates processes R_l to model parameters θ_m through the numerical model, which is constructed based on mechanistic representations, empirical response functions, or neural networks in some machine learning framework (Tsai et al., 2021). θ_m and the modeled R_l then affect the model's goodness of fit (i.e., observational constraint) and need to satisfy scaling coherency rules. The Bayesian inference seeks the optimal θ_m value that produces the least model-data discrepancy. That is, for a given model formulation, by identifying the optimal θ_m, we can also obtain the best upscaled equations of EBMs.

In mathematical terms, the cost function can be written as

$$J = \frac{1}{2}\Sigma_j\,(Y_j(R_l) - Y_{jO})^T C_j\,(Y_j(R_l) - Y_{jO}) + \Sigma_k \frac{\lambda_k}{2}(M_k(R_l))^2 + J_{0,} \quad (2)$$

where vectors $Y_j(R_l)$ are model predicted snapshots of the observed vectors of response variables Y_{jO}, and the corresponding covariance matrix of model-data discrepancy has an inverse specified by C_j.

In Equation 2, $M_k(R_l)$ represents the residual of the kth scaling coherency rule among the processes R_l, which are related to model parameters θ_m. In this term, λ_k is the Lagrangian multiplier for the kth scaling rule. J_0 is the regularization term from the prior knowledge of θ_m. The scaling coherency rules could be empirical relationships,

for example, the relationship between vapor pressure deficit and stomatal conductance (Yu et al., 2017), or the relationship between methane production and pH (Cao et al., 1998), among others. The scaling coherency rules could also be physical rules, such as the primary rule (Table 1) of mass conservation linking precipitation, infiltration, evapotranspiration, surface runoff, and ponding (when doing data assimilation for hydrological models (Tian et al., 2021)), or the derived rule (Table 1) of the gradient-driven CO_2 flux that controls photosynthesis, or the "common sense" (i.e., ecological and dynamical constraints) as formulated in Bloom and Williams (2015). Depending on the specific situation, scaling coherency rules may appear as either equalities (e.g., for a conservation relationship among fluxes) or inequalities (e.g., some physical variables like mass or volume should never be negative). For conservation rules, for example, those of mass or momentum, tradeoff $M_k(R_l)$ should be satisfied exactly, so that considering such rules is equivalent to incorporating an equality constraint, that is, error-free observations encountered in Bayesian inference or data assimilation (Basir & Senocak, 2022; Pan & Wood, 2006). When the above scaling rules are considered in the physics-based or knowledge-guided machine learning approach (ElGhawi et al., 2023; Liu et al., 2022, 2024), $M_k(R_l)$ represents the physical knowledge to be incorporated.

The identification of optimal parameters based on Equation 2 is equivalent to minimizing the cost function J, a process that is related to the first order variation δJ, which can be obtained by applying the chain rule of differentiation to Equation 2, such that

$$\delta j = \Sigma_n \left[\Sigma_j C_j (Y_j(R_l) - Y_{jO}) \frac{\partial Y_j}{\partial R_n} \right] \frac{\partial R_n}{\partial \theta_m} \delta\theta_m + \Sigma_n \left[\Sigma_k \lambda_k M_k(R_l) \frac{\partial M_K}{\partial R_n} \right] \frac{\partial R_n}{\partial \theta_m} \delta\theta_m + \frac{\partial J_0}{\partial \theta_m} \delta\theta_m. \quad (3)$$

In Equation 3, δJ is related to the variation of parameters $\delta\theta_m$ through three types of constraints, with the first from the observations (i.e., $\Sigma_n \left[\Sigma_j C_j (Y_j(R_l) - Y_{jO}) \frac{\partial Y_j}{\partial R_n} \right] \frac{\partial R_n}{\partial \theta_m} \delta\theta_m$), the second from physical rules (i.e., $\Sigma_n \left[\Sigma_k \lambda_k M_k(R_l) \frac{\partial M_k}{\partial R_n} \right] \frac{\partial R_n}{\partial \theta_m} \delta\theta_m$), and the third from the prior information of the parameters (i.e., $\frac{\partial J_0}{\partial \theta_m} \delta\theta_m$).

Equation 3 allows us to make three assertions. First, for two models of the same number of parameters, a lower magnitude of $\frac{\partial R_n}{\partial \theta_m}$ will lead to smaller contributions to the cost function by the first and second types of constraints, so that the cost function is less sensitive to a given variability of the parameter $\delta\theta_m$. In other words, making the represented process R_n less sensitive to parameters θ_m leads to a model with lower parametric sensitivity. This is the case where applying (scaling coherency rules in the form of) physical rules makes the model more immune to uncertainty in individual parameters (Tang & Riley, 2013, 2021) (also see the example in Section 3.2). Second, the scaling coherence rules designated by $M_k(R_l)$ act like regularization to the parameter inference processes. When these regularization terms are ignored, posterior models will be prone to overfitting because parameters are less well-constrained. The need for regularization is a phenomenon widely observed in machine learning, which is the main driver for the recent surge of interest in physics-guided machine learning (ElGhawi et al., 2023; Goodfellow et al., 2016; Liu et al., 2022). This regularization also underlies the power of "common sense" in Bloom and Williams (2015). However, such "common sense" should emerge naturally in EBMs formulated based on physical rules. Third, by applying Equation 2 to ensemble predictions and treating Y_{jO} as the mean of the ensemble simulations, δJ quantifies the divergence of the simulations resulting from equally good parameters derived from calibration. In this sense, it is one of the terms in the ecological predictability framework by (Dietze, 2017). That is, for model parameterizations that lead to calibration equifinality, those with weaker inter correlations among parameters will result in greater predictive equifinality for the same level of parametric uncertainty (Tang & Riley, 2013). Therefore, if a model formulation results in low $\partial J/\partial\theta_m$ when averaged over the uncertain parameters, the predictive equifinality is reduced and the model's predictive power (after the calibration period) is improved. Moreover, since explicitly accounting for process tradeoffs by $M_k(R_l)$ often appears as increased model complexity (by involving more equations), we contend that more complex models can potentially have lower predictive equifinality due to stronger correlations among the processes as enforced by physical rules. This lower predictive equifinality is at odds with the common criticism of increasing model complexity, that is, higher model complexity leads to more model parameters, and thereby higher predictive uncertainty.

Empirically based approaches formulated within EBMs work in a top-down manner, where regressions describe response functions (encoded as scaling coherency rules $M_k(R_l)$) derived from observations of emergent biogeochemical rates, and corresponding environmental factors, such as temperature, moisture, radiation, pH, and

soil properties such as in some methane production models (Riley et al., 2011; Zhuang et al., 2004). Because the empirically based approaches rely strongly on the amount and context of observations, the resulting response functions can vary significantly from one study to another, so that there is substantial uncertainty in the derived scaling coherency rules $M_k(R_l)$. Furthermore, because it is rare and difficult to comprehensively control and measure the variation of all relevant environmental factors and control variables of the biogeochemical rates, the strong context-dependence of the response functions is unlikely to be resolved by further observations. That is, $M_k(R_l)$ at one place or time cannot be transferred confidently to another place at another time. Such a situation is quite different from measuring the gravitational constant using a pendulum, where the context dependence can be reduced to be almost negligible (Parks & Faller, 2010). Moreover, $M_k(R_l)$ derived by empirically based regressions generally ignore physical constraints among the subscale processes (i.e., terms as $\Phi_{l-1}\left(Y_{m_{l-1}}, Q_{j\,l-1}\right) = 0$ in Figure 3, which are included by λ_k into Equation 3), and very likely place process interactions at wrong locations (e.g., by representing nonmultiplicative interactions with multiplicative functions (Tang & Riley, 2024)). Consequently, the resulting parameterizations will unlikely be robust. Section 2.2 provides more analysis on the shortcomings of empirically based approaches.

In contrast to the top-down empirically based approaches, physical rules-based approaches work in a bottom-up manner. They focus on representing relatively well-understood basic processes and related interactions using established mathematical constraints and logical inductions. Because these mathematical constraints and logical inductions have been vetted by observations in diverse disciplines, the resultant model constraints are much stronger than can be imposed by limited observed responses (e.g., the term related with Y_{jO} in Equation 3) in a calibration study. In this sense, physical rules-based approaches provide a meaningful way to alleviate the problem of limited observational data in ecosystem biogeochemical modeling.

Conceptually, the idea adopted by physical rules-based approaches is analogous to building a great variety of lego structures, in which only a few well-designed basic building blocks are used, even though some customized blocks are occasionally needed to knit the pieces together. These customized lego blocks correspond to processes that cannot currently be formulated using known physical rules or are too complex to be computed using known physical rules, but the empirical rules are known to be good enough, thus intuitive or empirical approximations are used instead. For instance, in applying the Richards' equation, the empirical soil water retention curve function is often used to relate soil matric potential and soil water content with soil texture (Clapp & Hornberger, 1978; van Genuchten, 1980). As another example, in modeling plant phenology, empirical rules are used to guide the plants' temporal development in the model (Grant et al., 2020). Nevertheless, for EBM modeling, we argue that most biogeochemical processes can be constructed with just a few well-understood basic building blocks. Interestingly, biology seems to work in such a hierarchical way. For example, protein folding can be described by first forming secondary structures from amino acids, then those secondary structures fold into the native state that is able to carry out various biological functions (and these processes can be represented with simple functions derived from thermodynamics (Rollins & Dill, 2014)). In this sense, physical rules-based approaches are explicitly constructing the emergent biogeochemical processes, and potentially avoiding logical inconsistencies that may be introduced by building EBMs mostly with empirical functions. Section 2.3 will further discuss this concept.

2.2. Limitations of the Empirical Response Function Approach

Currently, biogeochemical process rates (R) are typically formulated as multiplicative functions of a reference rate (R_0) and "rate modifiers" (f_j) capturing the effects of environmental conditions (θ_j),

$$R = R_0 \prod_{j=1}^{j=N} f_j(\theta_j). \tag{4}$$

Alternatively, minimum functions are used under the assumption that the dominant factor constrains the overall rate (i.e., the law of the minimum), such that

$$R = R_0 \min_j \left[f_j(\theta_j)\right]. \tag{5}$$

These formulations are also used for conductance and resistance needed to compute rates of different processes. Usually, $f_j(\theta_j)$ is a regression-derived multiplier representing the sensitivity of the rate R to environmental factor θ_j, assuming negligible synergistic or antagonistic interactions among θ_j (Jarvis, 1976). Occasionally, $f_j(\theta_j)$ has a physical basis. For example, when $f_j(\theta_j)$ represents the dependence of substrate availability in the form of

Michaelis-Menten kinetics, it may be argued to be mechanistically based by extrapolating insight obtained at specific conditions. Moreover, for many controlling factors, such as for pH and moisture, $f_j(\theta_j)$ is normalized to vary from 0 to 1. The temperature dependence is an exception, where an exponential function without an upper bound is often used, for example, the Q_{10} function for chemical reactions. In contrast, temperature dependence has an upper bound in biochemical reactions, with the controlling factor $f_j(\theta_j)$ constrained by physical feasibility (Pasut et al., 2023; Tang & Riley, 2024). We give some example biogeochemical processes that adopt the above formulations in Table 2.

The multiplicative approach of Equation 4 assumes that the different factors θ_j influence the response variable independently (Jarvis, 1976). As this assumption leads to zero covariance between the influence of any two factors, the multiplicative approach is equivalent to the multiplicative model without correlation in probability theory (Feller, 1967; Hoem, 1987; Wermuth, 1976). Meanwhile, the logic behind the law of the minimum approach of Equation 5 is based on observations of crop yield that dates back to the 1820s (Liebig, 1840; Sprengel, 1826). In some applications, the functional form $f(\theta_j)$ can be argued to be mechanistically based, for example, the use of Michaelis-Menten or Monod functions (Liu, 2007) for the carboxylation process by Rubisco in photosynthesis, or for biological growth directly related to substrate availability. Despite a mechanistic basis in limited cases, and their mathematical and conceptual simplicity, neither of these approaches provides robust formulations of fluxes and conductance in biogeochemical models, for reasons explained below.

In contrast to the assumption of no synergistic or antagonistic interactions underlining the multiplicative approach, it is uncommon that in real biogeochemical systems each modulating factor θ_j independently influences the response variable of interest. Instead, interactions between the modulation factors are more common. For example, enzymatic biochemical reactions involve at least two steps: (a) binding of substrate to the enzyme, and (b) new molecule production under enzyme catalysis (Briggs & Haldane, 1925). When one type of enzyme is acting on one type of substrate, by the law of mass action, this process is often summarized with the Michaelis-Menten kinetics,

$$R = \frac{v_{max} E S}{K + S}, \tag{6}$$

where E and S are enzyme and substrate concentrations, K is the affinity parameter (or half saturation constant), and v_{max} is the maximum catalysis rate (Michaelis & Menten, 1913). Since biochemical reactions mostly occur in water, and if the unbinding rate is relatively insignificant compared to the forward binding rate (as is usually assumed based on empirical observations), K is proportional to v_{max}/D, with D being the aqueous diffusivity of the substrate with respect to the enzyme (Tang & Riley, 2019b; Zhou et al., 1983). Therefore, in soil, K can be expected to be a function of temperature, moisture, and the type of substrate molecules. Further, in most cases, v_{max} is only a function temperature, while effective concentrations of catalytically active enzyme E and substrate S are functions of soil moisture. Consequently, the temperature and moisture effect on R will only emerge from their combined influences on K, v_{max}, and substrate availability, suggesting that the temperature sensitivity of R is not independent from its moisture sensitivity contrarily to as formulated by the multiplicative approach. Indeed, nonmultiplicative relationships between the temperature and moisture sensitivity of soil organic matter decomposition have been observed (Zhou et al., 2014). On the other hand, if a process is formulated using the law of minimum, the model will predict that once the temperature effect is below a threshold, moisture will have no effect on the reaction rate. These arguments thus invalidate both Equations 4 and 5 for acting as a logically consistent formulation of the biogeochemical rates.

Similarly, considering stomatal conductance as an example, there are sufficient mechanistic reasons to invalidate the approaches in Equation 4 or 5. Like many biochemical processes, stomatal conductance emerges from the interactions between many aspects of plant functioning, so that its response to changes in one influencing factor is dependent on other factors. Moreover, each plant grows by coordinating the traits of all its organs and adjusting to the presence of its neighbors. Consequently, there must be a directional information flow (including causality) among a plant's response to its influencing factors, which in turn regulates the behavior of stomatal conductance. Therefore, the effects of different influencing factors on stomatal conductance are unlikely to be simply multiplicative, as this oversimplification neglects the complexity of the plant and its interactions with the environment. A heuristic analogy is the process of putting on one's socks and shoes, in which socks must be put on first even though the selection of socks and shoes may appear to be independent. However, the multiplicative model cannot differentiate the logical order between putting on socks and shoes.

In the example of stomatal conductance, mechanistically, the opening of stomata is controlled by the volume and therefore turgor of the guard cells (Meidner & Mansfield, 1968) and epidermal cells (Buckley, 2019), but including epidermal cells will not influence the conclusion of the following logical induction. The volume of guard cells is an exponential function of their turgor pressure (Steudle et al., 1977), which is linearly related to osmotic pressure inside the guard cells. According to von Hoff's equation (Atkins & de Paula, 2006), the osmotic pressure is a linear function of solute concentration inside the guard cell, which depends on the photosynthesis rate (of chloroplasts inside guard cells), and water flux into the guard cell (Meidner & Mansfield, 1968). By applying the mechanical balance in the first order approximation (neglecting xylem cavitation), the osmotic pressure, turgor pressure, and leaf water potential will be linearly related. Therefore, even if leaf water potential would be linearly related to soil water potential, soil water potential is related to photosynthesis rate non-multiplicatively via photosynthetic controls on transpiration. These arguments may partially explain why the empirical Ball-Berry (Ball et al., 1987) and Leuning models (Leuning, 1990, 1995) are not numerically robust in practice when the effect of soil water stress is applied as a multiplier (Tang et al., 2015).

Acknowledging that water vapor pressure does not influence stomatal conductance multiplicatively, the model by Medlyn et al. (2011), based on optimality theory (Cowan & Farquhar, 1997), assumes that, within some time period, stomatal conductance adjusts to minimize the marginal water cost for photosynthesis. This assumption results in a deterministic relationship between stomatal conductance and water vapor pressure deficit. In contrast, field data have shown hysteretic relationships between leaf surface vapor pressure deficit and stomatal conductance (Wang et al., 2009). Coincidentally, when photosynthesis is represented using the model by (Farquhar et al., 1980), the resultant stomatal conductance is described by a combination of the multiplicative approach and the law of the minimum approach. However, Walker et al. (2021) suggested that the law of the minimum approach adopted by the Farquhar model caused significant numerical uncertainty. This evidence indicates that new formulations are needed to capture the rich variability of the response of stomatal conductance to changes in important environmental influencers such as vapor pressure deficit and soil moisture content.

Box 2. A Brief Description of Three Microbial Growth Models That Treat Resource Allocation for Maintenance and Growth Differently

Pirt Model

Specific growth rate: $\frac{1}{B_V}\frac{dB_V}{dt} = \mu_P(s) = \mu_{\max,P} h_P(s)$,

Specific substrate uptake rate: $\frac{1}{B_V}\frac{ds}{dt} = q_P(s) = \frac{\mu_P(s)}{Y_G} + m_P$.

In this model, specific biomass (and population) growth μ_P is non-negative and increases with substrate (s) availability, while cellular maintenance m_P is only provided by substrate taken up. Y_G is growth yield of biomass B_V from the substrate assimilation.

Compromise model

Specific growth rate: $\frac{1}{B_V}\frac{dB_V}{dt} = \mu_C(s) = \mu_{\max,C} h_C(s) - m_q(1 - h_C(s))$,

Specific substrate uptake rate: $\frac{1}{B_V}\frac{ds}{dt} = q_C(s) = \mu_{\max,C}\frac{h_C(s)}{Y_G} + m_q\frac{h_C(s)}{Y_G}$.

In this model, specific biomass (and population) growth μ_C increases from negative values under low substrate availability to positive values at high substrate availability, while the cost of maintenance $m_q\left(1 + \left(\frac{1}{Y_G} - 1\right)h_C(s)\right)$ B_V is paid by both biomass and substrate taken up. Y_G is growth yield from the substrate assimilation.

Dynamic energy budget (DEB) model

Specific reserve biomass growth: $\frac{1}{B_V}\frac{dB_R}{dt} = j_{A,\max} h_D(s) - \left(\kappa - \mu_D(s)\right)\frac{B_R}{B_V}$,

Specific structural biomass growth: $\frac{1}{B_V}\frac{dB_V}{dt} = \mu_D(s) = \frac{\kappa B_R Y_{RV} - m_D B_V}{B_V + B_R Y_{RV}}$,

Specific substrate uptake rate: $\frac{1}{B_V}\frac{ds}{dt} = q_D(s) = j_{A,\max}\frac{h_D(s)}{Y_{sR}}$.

In this model, substrate (s) first drives the increase of reserve biomass B_R, whose turnover flux κB_R drives the growth of structural biomass $\mu_D B_V$ after subtracting the cost of structural biomass maintenance ($m_D B_V$). When reserve biomass turnover is lower than the cost of maintenance, deficit will lead to the decrease of structural biomass or cell lysis (Tolla et al., 2007). Y_{sR} is the reserve biomass yield from substrate assimilation, while Y_{RV} is the structural biomass yield from mobilizing the reserve biomass.

Further, even if the influence from factors θ_j on R can be regarded as mutually independent, $f_j(\theta_j)$ may still be highly context dependent, particularly, when θ_j is dependent on soil conditions. For instance, for soil respiration temperature sensitivity, the literature reports more than 10 different functional forms (Exbrayat, Pitman, Abramowitz, & Wang, 2013; Sierra et al., 2015). Although each of them was able to fit the empirical data used to derive its functional form, none of them could be able to extrapolate temperature sensitivity from one experiment to another, and the difference between the diverse functional forms is far from being random and cannot be regarded as noise. When similar procedures are applied to all relevant influencing factors, we should not be surprised to find out that the resultant model cannot accurately capture spatiotemporal variability when conducting large-scale simulations (Carvalhais et al., 2014; Todd-Brown et al., 2013).

Moreover, because interactions between different modulating factors are not considered with the multiplicative approach of Equation 3, the resultant EBMs will be overly sensitive to the parameterization of those factorial dependences (Exbrayat, Pitman, Abramowitz, & Wang, 2013; Exbrayat, Pitman, Zhang, et al., 2013). As ignoring these interactions corresponds to omitting the scaling rules in Equation 2 (i.e., $M_k(R_l)$), less constrained relationships will be modeled between state variables, and thus the EMBs will have higher predictive equifinality (Luo et al., 2015; Tang & Zhuang, 2008). To help understand this last point, imagine we are putting together a bolt and a nut. Without matching their shape and size, they can be duct-taped together in various ways, but only to form a very vulnerable pair. Alternatively, carefully matching them in shape and size ensures that the resulting pair will last for a long time.

The law of the minimum approach has mostly been applied to biochemical reactions controlled by the supply of multiple substrates. Such processes include photosynthesis (Farquhar et al., 1980) and multiple nutrients co-limited biological growth (Chakrawal et al., 2022; Tang & Riley, 2021; Wang et al., 2010). It has also been generalized to process rates that are subjected to multiple influencing factors, for example, the Miami model used it to estimate net primary productivity as a function of temperature and precipitation (Lieth, 1973), and Noe and Giersch (2004) used it to estimate stomatal conductance as a function of light and vapor pressure deficit. Despite its wide adoption in ecology, many empirical studies suggest deviations from the law of minimum for both plant and microbial growth (Egli, 1991; Rubio et al., 2003). Alternatively, the "multiple limitation hypothesis," which assumes that plants optimize physiologically and morphologically to achieve no wasteful use of nutrients, is also not well supported by empirical observations (Rubio et al., 2003). The "multiple limitation hypothesis" is also unlikely correct for microbial growth due to the often-observed luxury uptake of abundant nutrients while others are in short supply (Powell et al., 2008).

Besides the inconclusive empirical support, the law of the minimum approach also introduces numerical difficulty by creating jumps when the predicted limitation shifts between limiting factors. For modeling photosynthesis, such a jump is usually smoothed out by quadratic functions, which involves two hyperparameters without physical meaning (for three co-limiting processes) that are obtained by trial and error (Collatz et al., 1990). The (arbitrary) choice of these parameters and the quadratic functional form have large effects on simulated plant gross primary productivity (Walker et al., 2021). In the context of model parameter inference, the law of the minimum approach similarly mistreats the scaling coherency rules in Equation 2, and, as a result, can lead to significant predictive equifinality, just as the multiplicative approach does.

Therefore, due to issues discussed above, the convenient empirically based response function approaches will not facilitate improvement of climate-biogeochemistry feedback predictions, no matter how many more observations are applied as constraints (through model-data fusion). Further, incorporating more biogeochemical processes using the same approach will only increase predictive equifinality and degrade model performance, where the latter has led to the incorrect impression held by the research community that increasing model complexity is usually bad and should be avoided.

2.3. Feasibility of Physical Rules-Based Approaches

Since the empirically based approaches are unlikely to result in EBMs with high spatiotemporal transferability, we thus turn to physical rules-based approaches. We argue that physical rules-based approaches are feasible, because biogeochemical processes (the focus of this perspective) can be broadly classified into two groups: (a) transfer of mass and energy, and (b) chemical conversion of molecules. The transfer of mass and energy is achieved through radiation, convection, advection, and diffusion, all of which are well-studied in physics (de Groor & Mazur, 1984; Plawsky, 2019). The chemical conversion of molecules involves chemical kinetics and

thermodynamic control of chemical reaction rates, whose physical rules are also well-studied (Horn & Jackson, 1972; Klotz & Rosenberg, 2008). These two groups together conceptualize biogeochemical processes as reactive-transport systems. Accordingly, different biogeochemical processes can be integrated by applying the reactive-transport concept at various spatiotemporal scales with proper mathematical approximations. (It is noted that the first law of thermodynamics for nonequilibrium systems can be written in the reactive-transport form, highlighting the potential to use it as a framework of process integration (de Groor & Mazur, 1984).) For aqueous chemistry, its modeling can be done using existing formulations of reactive-transport models (Molins & Knabner, 2019; Steefel & Lasaga, 1994). Such an approach would also in principle be sufficient to model the dynamics of carbon, nutrients, gasses, and solids in soils and water bodies (Riley et al., 2014; Steefel et al., 2015). However, the reaction and transport pathways are not all known, and heterogeneous chemical substrates and complex, biologically mediated reaction networks make the direct application of reaction-transport models difficult in practice (Dudal & Gérard, 2004).

Apart from landscape upscaling issues (which are not considered here), the most difficult part of applying physical rules-based approaches is to properly deal with biological growth of micro- and macro-organisms that drive or modulate biogeochemical processes. Biological growth is a phenomenon that emerges from an enormous number of chemical reactions at microscales, for many of which there are known functional relationships (Madigan et al., 2009). However, we currently do not have a well-established upscaling scheme to create a bottom-up representation of biological growth—that is, even though we may know the kinetics of each reaction, we do not know their relative importance and their interactions in a living organism. Rather, biological growth could and should currently be modeled using a combination of physical rules and empirical phenomenological formulations. For example, phenomenological rules can be applied to plant growth stages, which are fortunately well-observed for many species, particularly crops (Hanway, 1966; Miller, 1992), and have been successfully used to parameterize many natural ecosystem plant types (Grant, 2013; Zhou et al., 2021). A combination of remote sensing data and in-situ measurements (Dronova & Taddeo, 2022; Harfenmeister et al., 2021; Xiao et al., 2009) will facilitate us to obtain plant growth stage parameterizations that are sufficiently robust. Eco-evolutionary approaches can also be used instead of purely phenomenological rules to constrain processes with bounds of ecological and evolutionary feasibility. Compared to phenomenological rules, eco-evolutionary approaches offer flexibility to deal with varying environmental conditions (Franklin et al., 2020; Harrison et al., 2021).

Physical rules can be used to formulate biomass accumulation and translocation in a similar way as for chemical reactions, except that biomass growth is the net result of several related subprocesses. The modeling of biomass growth can be done by dividing growth into several subprocesses, including substrate uptake, internal physiology, and mortality. The substrate uptake process is amenable to relatively well-established physical rules (e.g., law of mass action, which will be discussed with an example in Section 3.3). Microbial internal physiology relates how nutrients are used for biomass growth, cellular maintenance, and release of extracellular products (e.g., exoenzymes), and has been described using, for example, the Pirt model (Pirt, 1982), the compromise model (Beeftink et al., 1990; Wang & Post, 2012), and the dynamic energy budget model (Kooijman, 2009; Tang & Riley, 2015, 2023) (see Box 2 for a brief description).

Compared to empirically based models (e.g., the Pirt and the compromise models), DEB has a relatively good mechanistic foundation that fully integrates mass and energy trade-offs during metabolic allocation for microbial growth. Importantly, the DEB formulation is structurally compatible with flux-balance models that represent biological growth by considering a large number of intra-cellular chemical processes that include enzymes, metabolites, and genomes (Antoniewicz, 2021). DEB models have also been successfully applied to animals and plants by treating individuals as a population of cells (Russo et al., 2022; Zonneveld & Kooijman, 1989). In particular, it is the only model structure that can reasonably predict the nonlinear relationship between carbon use efficiency (i.e., the ratio of carbon used to grow structural biomass and carbon uptake) and structural growth rate consistent with empirical observations and thermodynamics (Tang & Riley, 2023). That is, carbon use efficiency will first increase, then plateau and finally decrease with growth rate, whereas the Pirt and compromise models fail to predict the decrease of carbon use efficiency at high growth rate.

For plants, the modular nature of plant leaves, branches, stems, and roots, and the associated carbon and nutrient translocation (including both transformation and transport) allows one to model a single plant as a reactive transport system, where each modular part acts as an autonomous organ whose internal physiology follows a DEB-based formulation and the transport of carbon and nutrient between organs follows gradient driven flow

with a close coupling with plant hydraulics (a simple example was shown in (Kooijman, 2009) and a more comprehensive example is in Russo et al. (2022)). A conceptually similar approach has been advocated by (Thornley, 1972) and has been successfully implemented in the *ecosys* model (Grant, 1998) to address the balanced growth between plant shoots and roots. Analyses with *ecosys* showed that coupling plant physiological, hydraulic, and morphological parameterizations allows accounting for many geometric and metabolic trade-offs, for example, the popular pipe-model for plant xylem and canopy development by (Shinozaki et al., 1964) can emerge naturally from such a treatment (Grant, 1998).

The ideas above are meant to be applied to individual microbes and plants (which can be representative of a population), forming the foundation to model population and community dynamics (e.g., population demography, community and ecosystem assembly). These individual-based formulations, when combined in "ecosystem demography" models (Koven et al., 2020; Ma et al., 2022; Medvigy et al., 2019), will contribute to better modeling of biogeochemistry-climate feedbacks regulated by plant and microbial physiology (a knowledge gap in existing models).

2.4. A Simple Recipe for Physical Rules-Based Approaches

While applying physical rules-based approaches to parameterize ecosystem biogeochemistry will require significant effort, we provide an algorithm-like recipe in Box 3. By recognizing the hierarchical multi-scale nature of real-world problems, the recipe involves recursive application of the approach at different hierarchical levels. When formulating the interactions among entities, the essence is to break down any compound process into simple steps that can be formulated by physical rules. Meanwhile, when mathematically linking the entities by those interactions at each scale, it is essential to ensure the primary rules are satisfied (with known accuracy determined by the solution method). The examples presented in Section 3 were by and large developed by following this recipe. To readers interested in applying physical rules to real-world problems (including biogeochemistry), we recommend the three volumes of "*The Feynman Lectures on Physics*" (Feynman et al., 2011a, 2011b, 2011c), *Physical Biology of the Cell* (Phillips et al., 2012), and *Aquatic chemistry: chemical equilibria and rates in natural waters* (Stumm & Morgan, 1996), among others.

Box 3. An Algorithm-Like Recipe for Developing Physical Rules-Based Ecosystem Biogeochemistry Parameterization. The Algorithm Is Also Motivated by the Solution Strategy to the Very Large Scale Integration Layout Problem in Chip Design (Russell & Norvig, 2010)

PROCEDURE Physical_Rules_based_Process_Formulation(*process*)

STEP 1: Identify the number of *entities* involved in the *process*, set it to *N*.

STEP 2: LOOP: *entity L* = 1 to *N*

STEP 3: Identify the spatiotemporal context of *entity L*.

STEP 4: IF (there is hierarchical structure within *entity L*) THEN

Define *entity L* as a *new_process*,

RETURN, **Physical_Rules_based_Process_Formulation**(*new_process*)

END IF

END LOOP

STEP 5: Determine how the *entities* interact.

(A) Is the interaction through transport, chemical reaction, or mechanics?

(B) What are the conservation laws for the entities?

Step 6: Form a verbal description of the problem that details each interaction.

Step 7: Mathematize the verbal description.

Step 8: RETURN the mathematical formulation.

END PROCEDURE

3. Examples to Contrast the Predictive Power of Two Approaches

In the following, we give three examples to contrast the efficacy of the empirically based and the physical rules-based approaches. Particularly, we demonstrate how the concerns of worsening prediction equifinality associated with more mechanistic models (which usually appear as being more complex) are not universally supported. While the three examples are described mostly in the context of soil biogeochemistry, the second example can also be used to model plant respiration upon multiple carbon substrates (Wendering & Nikoloski, 2023), and the third example can help improve the temperature dependence of plant autotrophic respiration (Liang et al., 2018). For more examples, such as for the thermodynamic regulation of biogeochemical processes, we refer readers to (Jin & Bethke, 2003; Maggi et al., 2008), and for nutrient-regulated microbial and plant growth to (Chakrawal et al., 2022; Tang & Riley, 2021).

3.1. Soil Moisture Dependence of Substrate Affinity Parameter

This first example addresses the substrate affinity parameter involved in soil carbon and nutrient dynamics, based on the analysis by Tang and Riley (2019b). In modeling soil biogeochemistry, we often encounter Michaelis-Menten or Monod-type equations for decomposition rates or nutrient uptake (e.g., Equation 6, or variants including microbial biomass instead of enzyme concentrations). The affinity parameter K requires a value to compute the overall reaction rate R. Because K is important for almost every biogeochemical process using a substrate (e.g., Table 2), it often represents a significant fraction of the model calibration effort. Moreover, as K is defined for biogeochemistry in an aqueous environment, and soil moisture is a dynamic variable, the effective K (denoted hereforth by K_w as compared to the intrinsic $K_{w,0}$ for a pure aqueous environment) should be a function of soil moisture, and dimensionally consistent with solutes in the pore space, whose concentration varies with soil moisture (K_w must have a dimension of mass per unit water volume).

In existing EBMs, the moisture effect on K is often represented using a multiplier function (Maggi et al., 2008; Riley et al., 2011; Tang et al., 2010; Zhuang et al., 2004) but it is unclear which aspects of the biogeochemical rate's moisture dependence are accounted for by this multiplier.

Tang and Riley (2019b) developed a parameterization of K_w by first delineating three levels of hierarchical structures: (a) an individual microbe (representative of a bacterial cell), (b) a colony of microbes (which may represent soil aggregates to some extent), and (c) the soil matrix (Figure 4). They further assumed that (a) a microbial colony is covered by a water film (whose thickness is computed as a function of soil suction pressure); (b) within a colony, microbial cells are evenly distributed, immersed in water, and compete for diffusion-limited substrates; and (c) microbial colonies are connected to each other by diffusion through the soil matrix. By assuming that diffusion is the major physical process linking these scales, and modeling diffusion through von Smoluchowski's diffusion theory of chemical reaction (von Smoluchowski, 1917) and the Berg-Purcell formula for substrate interception by a spherical bacterial cell (Berg & Purcell, 1977), the affinity parameter for an aqueous substrate is found as

$$K_w = \frac{Nk_{2,p}}{k_{1,w}}\left(1 + \frac{k_{1,w}B_m/\nu_m}{\kappa_m}\right), \tag{7}$$

where

$$k_{1,w} = 4\pi N_A D_{w,0} r_C \frac{Nr_p}{Nr_p + \pi r_C f_A}, \tag{8}$$

$$f_A = 1 - \frac{N}{4}\left(\frac{r_p}{r_C}\right)^2, \tag{9}$$

$$\frac{1}{\kappa_m} = \frac{\nu_m \delta}{4\pi D_{w,0} r_m (r_m + \delta)} + \frac{\nu_m}{4\pi D (r_m + \delta)}, \tag{10}$$

$$D = D_{w,0}\tau_w \phi_w + D_{g,0}\tau_g \frac{\phi_g}{\alpha}, \tag{11}$$

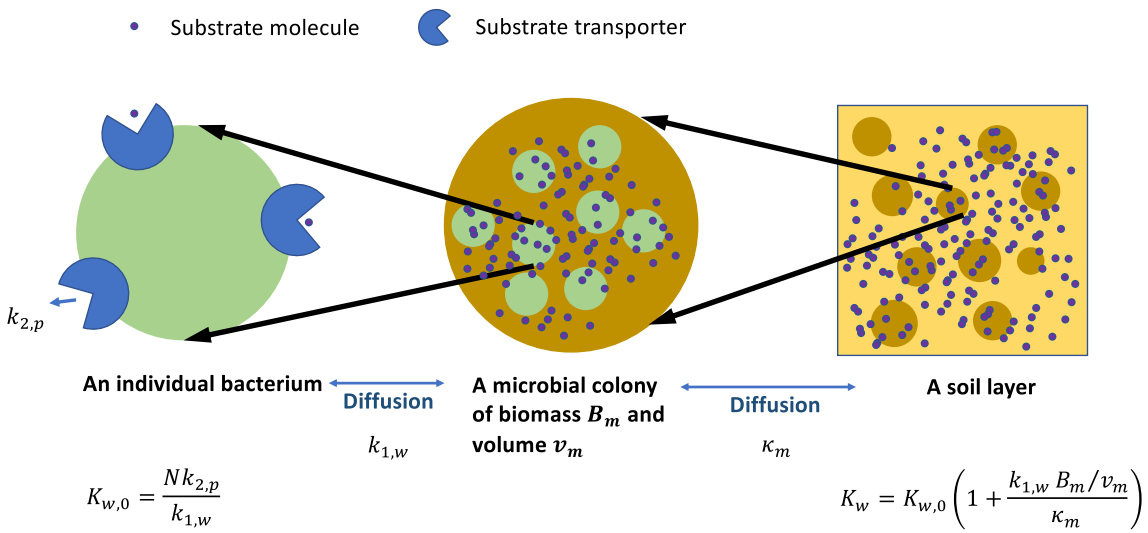

Figure 4. The hierarchical conceptual model used to derive the moisture dependence of substrate affinity parameter K (i.e., K_w in Equation 7). Diffusion is the major interaction that links the entities involved in the problem. For simplicity, pumps, channels, and carriers for substrate uptake are all called substrate transporters.

and

$$\delta = \max\left(10^{-8}, \exp\left(-13.65–0.857 \log(-\psi)\right)\right). \quad (12)$$

From the above, we can see that K_w (mol m^{-3} water) is influenced by the following four groups of input parameters:

1. *Soil*: soil matric water potential ψ (MPa), tortuosity of aqueous tracer τ_w (m m^{-1}), tortuosity of gaseous tracer τ_g (m m^{-1}), water-filled porosity ϕ_w (m^3 m^{-3}), air-filled porosity ϕ_g (m^3 m^{-3}), and water film thickness δ (m).
2. *Tracers*: aqueous tracer diffusivity $D_{w,0}$ (m^2 s^{-1}), gaseous tracer diffusivity $D_{g,0}$ (m^2 s^{-1}), and Bunsen solubility for gas tracer α (mol mol^{-1}). Note for solutes, $D_{g,0} = 0$ removes the contribution of gaseous phase.
3. *Microbes*: radius of microbial microsite r_m(m), whose volume is $\nu_m (= 4\pi r_m^3/3)$, mean microbial biomass density of a microbial microsite B_m/ν_m, number of substrate uptake sites per microbial cell N (sites per cell), mean microbial cell radius r_C (m), mean radius of microbial substrate uptake site r_p (m), the maximum substrate processing rate per uptake site $k_{2,p}$ (s^{-1}), and the production rate of the given substrate in the microsite p_C (mol m^{-3}).
4. *Universal constants*: Avogadro number N_A (mol^{-1}), and π.

The above formulation of K_w allows one to describe the moisture control of substrate uptake for a biogeochemical process in a soil volume that is of the order of 1 cm^3. It may also help represent microbial substrate uptake in 1-D vertically resolved reactive-transport based models of soil biogeochemistry (Dwivedi et al., 2019; Pasut et al., 2020; Riley et al., 2014).

With its diverse input parameters, Equation 7 provides insights into how K_w will be affected by soil physical properties (e.g., soil texture, organic matter content), soil moisture content, solute characteristics, and microbial traits (cf. Figures 2–4 in Tang and Riley (2019b)). For the results reproduced in Figure 5, K_w for oxygen increases as soils become wetter, following a sigmoidal shape whose inflection point varies with soil texture (Figure 5, left column). K_w for solutes decreases following a nearly exponential decay to a minimum value upon saturation (Figure 5, center column). When combining these formulations for oxygen and solute affinities in a steady-state microbial respiration model (see Tang & Riley, 2019b for details), prediction of respiration responses to soil moisture, and their dependence on soil properties, captures observed patterns (Figure 5, right column).

Almost all the parameters needed by the above model are either routinely measured (e.g., soil characteristics, diffusivities of various molecules (Cussler, 2009), tortuosity effect on gas and solute diffusion (Moldrup et al., 2003)), or have been estimated for specific case studies (e.g., microbial biomass density in a microsite (Raynaud & Nunan, 2014)), enabling us to apply Equation 7 with very little calibration at least in some contexts.

Tang and Riley (2019b) demonstrated reasonable predictions of the moisture-microbial respiration relationship with typical parameters from the literature and, importantly, without parameter calibration.

With modest modifications, the above model can be adapted to substrate kinetics of clay particles, fine roots, and fungal hyphae. When the resultant substrate kinetic parameterizations are implemented within a reactive-transport based framework of plant shoot-root growth, like that in *ecosys* (Grant, 1998), we can obtain new insights on how soil, plant, and microbial traits affect the dynamic nutrient coupling between plants and microbes in soil.

The approach above works well when microbial traits (e.g., $k_{2,p}$ and N) and microbial biomass (B_m/ν_m) are relatively static. However, in dynamic soil moisture environments, microbial biomass and traits related to substrate utilization also vary. The next logical step is to couple this framework with equations that describe microbial biomass and trait dynamics, aiming to achieve a mechanistic and ecologically sound soil carbon cycling model. Microbial biomass dynamics are already routinely modeled using empirical kinetics laws (as discussed in the review by Chandel et al. (2023)), but could also be modeled in a more mechanistic ways as discussed in Section 3.2. Therefore, the challenge lies in how to couple these mechanistic formulations to various aspects of microbial biomass growth (see Box 2), mortality, maintenance, dormancy, and other functions. We recognize that a mechanistic understanding is not available for some of these functions, but phenomenological or eco-evolutionary approaches can serve as initial approximations to the missing mechanistic representation. During the pursuit of this goal, it would be intriguing to assess to what extent biological (or even ecological) processes are so strongly coupled (or coordinated) to transport processes that they do not need to be modeled independently (i.e., they can be lumped through coarse graining).

3.2. A Network of Multiple Substrates and Consumers

The second example of physical rules-based approaches is for competitive interactions in a network of substrates and consumers, which are relevant in various contexts of biogeochemistry and ecology (Figure 6). These interactions include soil organic carbon decomposition by microbes (Wieder et al., 2014), nutrient competition between plants and microbes (Zhu et al., 2016), interactions between enzymes and substrates in the cytoplasm of a microbial cell (Etienne et al., 2020), and trophic networks including producers, consumers, and predators in population ecology (Barraquand, 2014; Buchkowski et al., 2022).

In the context of biogeochemistry, a network of substrates (S) and consumers (E) can be constructed using the law of mass action, which, by aid of the quasi-steady approximation, can be presented in the following form:

$$S_i + \sum_{j=1}^{j=J} X_{ij} = S_{i,T}, \text{for } i = 1, \cdots, I, \tag{13}$$

$$E_j + \sum_{i=1}^{i=I} X_{ij} = E_{j,T}, \text{for } j = 1, \cdots, J, \tag{14}$$

$$S_i E_j = K_{ij} X_{ij}, \tag{15}$$

$$\frac{dP_{ij}}{dt} = k_{ij,2} X_{ij}, \tag{16}$$

where subscripts i and j indicate different substrates (in total J substrates) and consumers (in total I consumers), and subscript T designates the total concentration of substrate S_i and consumer E_j in the spatial domain of analysis, regardless of their form (free or bound in a complex). Equations 13 and 14 account for the mass balance relationships of substrates and consumers in the system, Equation 15 describes the formation of substrate-consumer complex X_{ij}, which is used in Equation 16 to compute the production of new materials, denoted by P_{ij}. For a predator-prey network, K_{ij} is related to the handling and attacking rates of a predator on a prey (Real, 1977), and $k_{ij,2}X_{ij}$ is biomass growth of the predator E_j upon the successful handling of prey S_i.

The above system as a whole lacks an analytical solution, but it does have a first-order approximation (the Equilibrium Chemistry Approximation, ECA (Tang & Riley, 2013)) as follows:

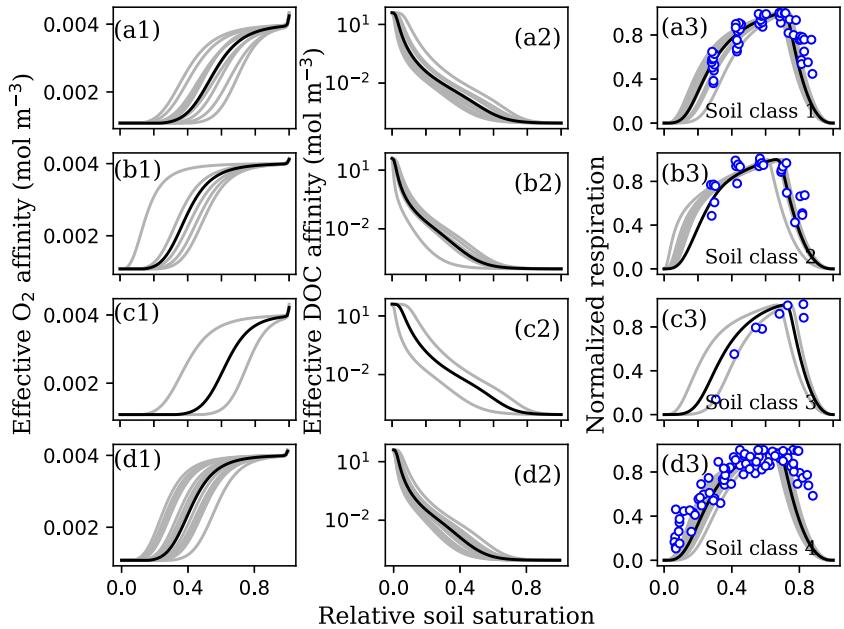

Figure 5. Example application of Equations 7–12 for affinity parameters of gaseous O_2 (panels a1, a2, a3, and a4), and dissolved organic carbon (DOC; panels b1, b2, b3, and b4) as a function of soil moisture for 32 soils in 4 classes. The four soil classes are (1) medium to fine texture soils from (Doran et al., 1990); (2) coarse texture soil from (Doran et al., 1990); (3) other soils from (Doran et al., 1990); and (4) soils from Franzluebbers (1999). The rightmost panels are correspondingly predicted respiration-moisture relationships using the synthesizing unit model. Same parameters are used from Tang and Riley (2019b). Gray lines are for different soils, black lines are computed from mean soil texture of each soil class, blue circles are measurements.

$$X_{ij} = \frac{S_{i,T}E_{j,T}/K_{ij}}{1 + \sum_{l=1}^{l=I} S_{l,T}/K_{l,j} + \sum_{l=1}^{l=J} E_{l,T}/K_{i,l}}, \tag{17}$$

Equation 17 can be shown to satisfy the partitioning principle (Tang & Riley, 2017), which is critical for developing a theory to coherently scale up from a single chemical reaction to unicellular and multicellular organisms (Kooijman, 2009). Specifically, when S_i are samples from the same substrate S (i.e., $\Sigma_i S_i = S$), and E_j are samples from the same consumer E (i.e., $\Sigma_j E_j = E$), the sum of X_{ij} will equal X obtained using substrate S, and consumer E. That is, by summing over all substrates and consumers in Equation 16, we obtain

$$X = \Sigma_{ij} X_{ij} = \frac{S_T E_T/K}{1 + S_T/K + E_T/K}. \tag{18}$$

Corresponding to Equation 18, the total production rate of new material ($P = \Sigma_{ij} P_{ij}$) is

$$\frac{dP}{dt} = k_2 X, \tag{19}$$

In the literature, however, Equations 13–16 have often been solved with an incomplete consideration of the mass balance constraints imposed by Equations 13 and 14. For instance, (Williams, 1973) modeled a system of many consumers competing for a single substrate, whose solution based on ECA is

$$V = \sum_{j=1}^{j=J} \frac{S_T v_{max,j} E_{j,T}/K_j}{1 + \Sigma_{j=1}^{j=J} S_T/K_j + E_{j,T}/K_j,}, \tag{20}$$

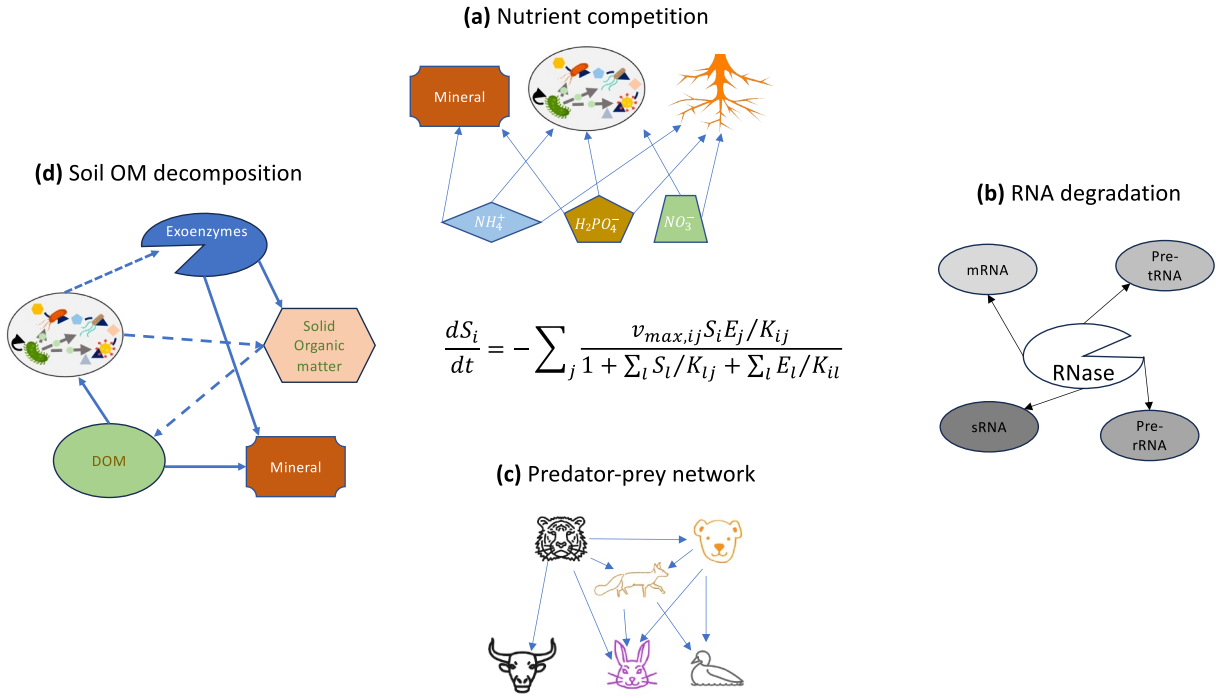

Figure 6. Examples of substrate-consumer networks that can be approximated by the Equilibrium Chemistry Approximation kinetics. Here substrate S_i is consumed by consumer E_j as specified with kinetic parameters K_{ij} and $v_{max,ij}$. It is assumed that the units of S_i and E_j have been properly converted for the equation shown in the figure to hold for various problems.

where V represents the total consumption rate by predators, with $v_{max,j}$ being the maximum substrate uptake rate by consumer j.

However (Williams, 1973), applied a simple juxtaposition of the empirical Holling's type II predation functions (Holling, 1959), and obtained

$$V = \sum_{j=1}^{j=J} \frac{V_{max,j} S}{S + K_j}, \tag{21}$$

where the dependence of individual's predation rate on consumer E_j is not captured (note $E_{j,T}$ is part of $V_{max,j}$ through $V_{max,j} = v_{max,j} E_{j,T}$). Moreover, in models that do include consumer effects on predation rate, the predator competition effect ($\Sigma_j E_j$ in the denominator) is often neglected (Murdoch, 1973; Real, 1977). Without these consumer effects, the model could result in incorrect parametric sensitivity when the total substrate is limited (Tang, 2015).

Additionally, in predator-prey modeling, there has been a long-lasting debate regarding whether the specific predation rate should be dependent on both the density of prey (S_T) and consumers (E_T in our nomenclature), and various formulations have been hypothesized (Beddington, 1975; Berryman, 1992; DeAngelis et al., 1975; Ginzburg & Akcakaya, 1992). Based on the application of physical rules, the simplest ECA formulation by Equation 18 reproduces the Beddington-DeAngelis formulation that is obtained through ad hoc assumptions, while the more general ECA form (Equation 17) has many other applications (Cheng et al., 2019; Huang et al., 2018; Zhu et al., 2016).

In soil biogeochemical modeling, the simple juxtaposition approach was also used to formulate the decomposition of two pools of soil carbon by a single microbial population, such as in the MIMICS model (Wang et al., 2014), where the growth of microbial biomass is formulated as

$$\frac{dC_b}{dt} = Y_G \frac{C_b v_{max,l} C_l}{C_l + K_l} + Y_G \frac{C_b v_{max,s} C_s}{C_s + K_s}, \tag{22}$$

where Y_G is the microbial growth efficiency, assumed the same for both substrates. Note that in Equation 22 the mortality term is ignored to simplify the discussion. Similar to Equations 21, 22 predicts that the specific consumption of carbon pool C_l is independent from that of carbon pool C_s.

Since there is only one microbial population degrading two soil carbon pools, the metabolic effort of the microbial population is expected to be divided between the two pools. That is, working on carbon pool C_l has a direct influence on the microbial effort allocated to carbon pool C_s, and vice versa. This subdivision means the formulation of Equation 22 predicts the wrong parameter sensitivity, whereas the mechanistically consistent formulation based on the ECA should be

$$\frac{dC_b}{dt} = Y_G \frac{C_b v_{max,l} C_l/K_l}{1 + C_l/K_l + C_s/K_s + \alpha C_b/K_l} + Y_G \frac{C_b v_{max,s} C_s/K_s}{1 + C_l/K_l + C_s/K_s + \alpha C_b/K_s}, \tag{23}$$

where α scales the available metabolic effort to the microbial biomass C_b, which is estimated to be of the order 10^{-4} when substrates are expressed in carbon mass units (Tang & Riley, 2019a). Thus, terms multiplied with α can be ignored mostly, but keeping them may prevent runaway microbial biomass growth when applying the model.

Because $1+C_l/K_l + C_s/K_s > max(1+C_l/K_l, 1+C_s/K_s)$, Equation 23 then predicts lower sensitivity of $\frac{dC_b}{dt}$ to K_l and K_s than Equation 22. Further, it can be shown that the parametric sensitivity of $\frac{dC_b}{dt}$ to K_l and K_s are correlated in Equation 23, making the resultant model parametrically better constrained, and very likely have less severe predictive equifinality compared to Equation 22. This last assertion is consistent with our inference at the beginning of Section 2, and supported by the model-data fusion experiment in (Tang & Riley, 2013), where the ECA formulation was much more robust than the simple juxtaposition of Holling's type II uptake functions (see comparison of Figures 11 and 12 there). We leave a comprehensive analysis of the new formulations (Equation 23 and the corresponding equations of C_l and C_s) on long term soil carbon dynamics for future work.

Besides obtaining a more consistent formulation of microbial growth over multiple soil carbon pools, the solution to Equations 13–16 also leads to a natural incorporation of soil mineral influences on organic carbon decomposition by approximating the organic carbon-mineral interaction with the Langmuir isotherm, leading to a modification of Equation 23 as

$$\frac{dC_b}{dt} = Y_G \frac{C_b v_{max,l} C_l/K_l}{1 + C_l/K_l + C_s/K_s + \alpha C_b/K_l + M/K_{l,M}} + Y_G \frac{C_b v_{max,s} C_s/K_s}{1 + C_l/K_l + C_s/K_s + \alpha C_b/K_s + M/K_{s,M}}, \tag{24}$$

where M indicates the total concentration of mineral surfaces available for adsorption and $K_{s,M}$ and $K_{l,M}$ are sorption parameters for substrates C_l and C_s.

From Equation 24, it is inferred that through competitive adsorption of (dissolved) soil organic matter (and exoenzymes Tietjen and Wetzel (2003)), microbial decomposition and growth are suppressed by soil minerals (Tang & Riley, 2015). However, if the turnover of exoenzymes is assumed to be positively linked with its catalysis rate, interaction with clay particles could increase the exoenzymes' lifetime by reducing the catalysis rate. Equation 24 then explains that, with increasing soil depth, along with the usual decrease of soil carbon, the specific decomposition rate naturally decreases, lending mechanistic insight to corroborate the attenuation factor in CENTURY-like models, where an exponential attenuation factor is needed to suppress carbon decomposition in order to correctly model the soil carbon profile (Koven et al., 2013). However, with proper characterization of soil mineral surfaces M and the associated sorption parameters $K_{S,M}$ and $K_{l,M}$, one can characterize the observed vertical heterogeneity more mechanistically than achieved with exponentially decreasing attenuation functions (Dwivedi et al., 2017). In particular, the mechanistic model will enable us to evaluate many hypotheses regarding how the interactions between SOM molecule composition, microbial abundance and diversity, soil conditions, and plant input regulate the multiple facet responses of soil respiration and SOM storage to environmental changes.

As an example to demonstrate the parametric sensitivity due to different model formulations, we define the specific substrate uptake $F_b = \frac{1}{Y_G C_b} \frac{dC_b}{dt}$ and compute the parametric sensitivity of F_b with respect to V_l and V_S for the Monod kinetics of Equation 22,

$$\frac{\partial F_b}{\partial v_{max,l}} = \frac{C_l/K_l}{1 + C_l/K_l}, \tag{25}$$

$$\frac{\partial F_b}{\partial v_{max,s}} = \frac{C_s/K_s}{1 + C_s/K_s}; \tag{26}$$

and, similarly, for the ECA-based Equation 24,

$$\frac{\partial F_b}{\partial v_{max,l}} = \frac{C_l/K_l}{1 + C_l/K_l + C_s/K_s + \alpha C_b/K_l + M/K_{l,M}}, \tag{27}$$

$$\frac{\partial F_b}{\partial v_{max,s}} = \frac{C_s/K_s}{1 + C_l/K_l + C_s/K_s + \alpha C_b/K_s + M/K_{s,M}}. \tag{28}$$

From Equations 25 and 26, we see that Monod kinetics predicts the parametric sensitivities $\frac{\partial F_b}{\partial v_{max,l}}$ and $\frac{\partial F_b}{\partial v_{max,s}}$ to be independent from the interactions between C_l, C_s, and M, while such dependence is captured by ECA kinetics (Equations 27 and 28). In particular, the Monod kinetics always predicts higher parametric sensitivity than the ECA kinetics (Figure 7), implying that the same parametric uncertainty will lead to higher predictive equifinality for models using the Monod kinetics.

3.3. Temperature Dependence of Enzyme-Catalyzed One-Substrate Reactions

In our third example, we discuss the temperature sensitivity of an enzyme-catalyzed one-substrate reaction (Tang & Riley, 2024). In this problem, the two entities—enzymes and substrates—are linked with a diffusion limited chemical reaction. The enzymes involve a sub-scale problem where they are partitioned between active and inactive conformation states as a function of temperature. Depending on the size contrast between substrate and enzyme molecules, we have three limiting solutions derived from ECA kinetics (Tang, 2015; Tang & Riley, 2019a):

1. When substrate molecules are much larger than enzymes, or the enzymes are in significant excess of substrate binding surface area (e.g., cellulose during hydrolysis), Equations 18 and 19 can be approximated by reverse Michaelis-Menten (MM) kinetics:

$$R = v_{max,E}\frac{ES}{K_E + E}. \tag{29}$$

2. When substrate molecules are much smaller than enzymes (e.g., microbial uptake of glucose) or the system is enzyme limited, we have the typical Michaelis-Menten kinetics,

$$R = v_{max,ES}\frac{ES}{K_{ES} + S}. \tag{30}$$

3. When substrate and enzyme molecules have similar size (e.g., when fructose is the substrate and invertase is the enzyme), we have the reaction represented using the ECA kinetics

$$R = v_{max,ES}\frac{ES}{K_{ES} + S + E}. \tag{31}$$

For all cases, temperature dependence of the maximum reaction rate can be approximated by the transition state theory (Eyring, 1935):

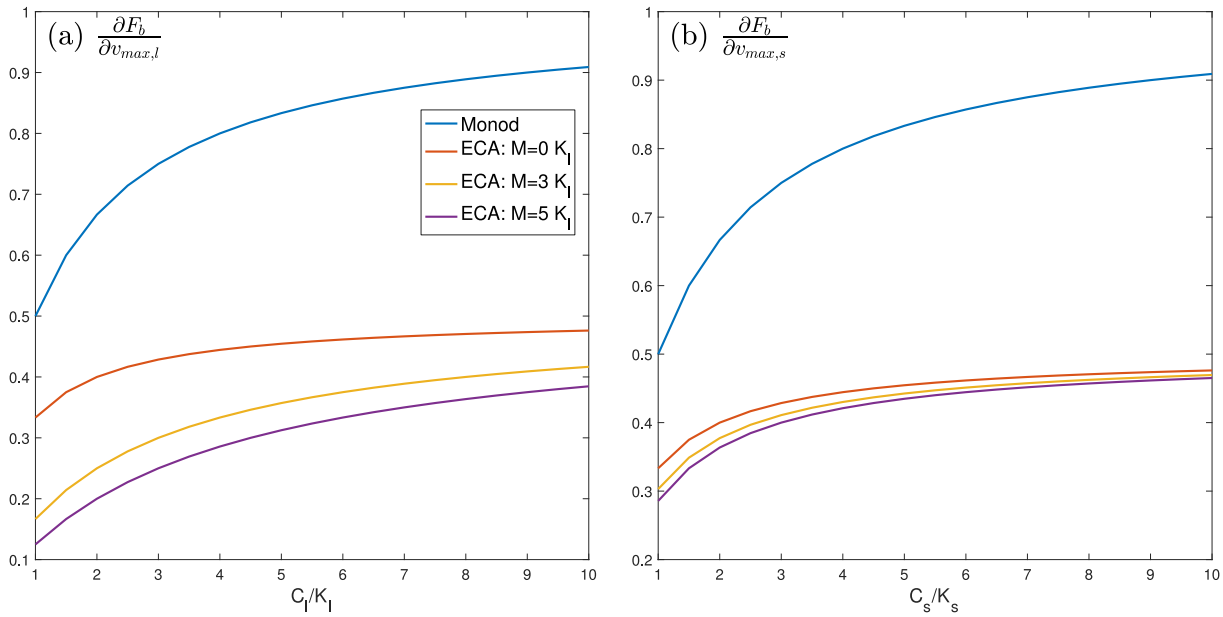

Figure 7. Comparison of parametric sensitivity for $\frac{\partial F_b}{\partial v_{max,l}}$ and $\frac{\partial F_b}{\partial v_{max,s}}$ when computed using the Monod kinetics versus the ECA kinetics. For all calculations, it is assumed $K_s = 10K_l$.

$$v_{max}(T) = v_{max,0}\frac{T}{T_0}\exp\left(-\frac{\Delta H_r}{R_g T}\left(1-\frac{T}{T_0}\right)\right), \tag{32}$$

where ΔH_r is the enthalpy of activation (which is constant), R_g is the universal gas constant, T_0 is the reference temperature (K), $v_{max,0}$ is the rate at T_0. By adopting the usual assumption that the unbinding rate is negligible compared to forward binding rate and the relative movement between substrates and enzymes is dominated by diffusion (Tang et al., 2021), the temperature sensitivity of the affinity parameter K_{ES} is determined by the ratio between the temperature sensitivity of v_{max} and that of the aqueous diffusivity D_w (see Equation 7 in example 1). According to the Stokes-Einstein equation (Cussler, 2009), the aqueous diffusivity for a spherical object of radius a is $D_{w,a} = k_B T/(6\pi\eta a)$, where the dynamic viscosity η has an empirical temperature dependence as $exp(B/T)$ (with $B > 0$; (Holmes et al., 2011)), thus a good approximation is:

$$K(T) = K_0\exp\left(-\frac{\Delta H_K}{R_g T}\left(1-\frac{T}{T_0}\right)\right), \tag{33}$$

where ΔH_K is the effective enthalpy of K, which is the difference between ΔH_r and the activation enthalpy of the self-diffusion of water, so that $\Delta H_K < \Delta H_r$.

In addition to temperature effects on reaction kinetics, temperature also affects the capacity of enzymes to perform the reaction (Murphy et al., 1990; Ratkowsky et al., 2005). In fact, enzymes are proteins, and proteins may lose or regain their native structure (and thus functionality) spontaneously. Because this spontaneous transition is taking advantage of the structural perturbations caused by thermal motions in the enzyme solution, the transition between native and unfolded states always occurs for an enzyme that is not irreversibly denatured (Finkelstein & Ptitsyn, 2016). The fraction of active enzymes in native state at a given temperature can be described by the well-established temperature relationship (Ghosh & Dill, 2009; Murphy et al., 1990; Sawle & Ghosh, 2011)

$$f_{act}(T) = \frac{1}{1+\exp\left(-\frac{\Delta G_E}{R_g T}\right)}, \tag{34}$$

where the Gibbs free energy of unfolding ΔG_E is defined as

$$\Delta G_E = \Delta H_E - T\Delta S_E = n_E \Delta C_p \left[(T - T_H) + T \ln\left(\frac{T_S}{T}\right) \right], \tag{35}$$

where n_E is the number of amino acid residues of the enzyme, heat capacity ΔC_p (J K^{-1} (mol amino acid)$^{-1}$) defines the energy required to reorganize the water molecules surrounding the protein, T_H is the temperature at which enthalpy ΔH_E equals to zero, and T_S is the temperature at which entropy ΔS_E equals to zero.

Based on the survey by Silverstein (2020), ΔC_p seems to be quite consistent among thermophobic, mesophilic, and thermophilic proteins, centering around 60 J (mol amino acid)$^{-1}$, with an increasing variability from thermophobic to thermophilic proteins. Meanwhile, T_H increases from thermophobic to thermophilic proteins, with an increasing difference between T_S and T_H (see Table 6 in Silverstein (2020)). A comprehensive analysis using data from the Protein Data Bank will be very helpful to gain more insights on the parameterization of Equation 35.

When the above relationships are applied together in the reverse MM kinetics (Equation 29), we have

$$R = v_{max,E}(T) \frac{f_{act}(T)\, ES}{K_E(T) + f_{act}(T)\, E}, \tag{36}$$

and when applied to MM kinetics (Equation 30), we have

$$R = v_{max,ES}(T) \frac{f_{act}(T)\, ES}{K_{ES}(T) + S}. \tag{37}$$

Therefore, these results show that the overall temperature sensitivity of an enzyme catalyzed one-substrate reaction emerges from three types of temperature functions: (a) Arrhenius equation, (b) the Eyring's transition-state theory, and (c) the thermal stability of native proteins. The equation for ECA kinetics is not reported here because it is just a combination of reverse and forward MM kinetics.

When the above results are adapted to the substrate affinity parameter of microbial substrate uptake, for example, the bacterial cells discussed in the first example in Section 3.1, the temperature dependence of the affinity parameter will be more complicated than represented by the Arrhenius-like function, because it will also involve $f_{act}(T)$ through its interaction with the number of transporters distributed over the microbial cells (i.e., parameter N in Equations 7–9).

To visualize the above relationships, we show some examples of hypothetical enzymes based on mean values of ΔC_p, T_S, and T_H from Table 6 in Silverstein (2020) and a typical enzyme length of 290 amino acids of prokaryotes (Brocchieri & Karlin, 2005). In Figures 8a and 8b, we see that the range of temperatures in which enzymes stay active expands as the Gibbs free energy curves shift from thermophobic to thermophilic enzymes. More interestingly, for the reverse MM kinetics described by Equation 36, the normalized reaction rate $R/(v_{max}(T_0)S)$ increases steadily (almost exponentially because f_{act} is close to 1) across most of the biochemically relevant temperatures, with sharp drop-offs at the low and high temperature ends (Figure 8c, where curves are drawn for the hypothetical mesophilic enzyme).

In contrast to the relationship by Equation 36 shown in Figure 8c, when the relationship by Equation 37 is illustrated (Figures 8d–8f), all kinetics show the often observed asymmetric temperature response (Peterson et al., 2004; Ratkowsky et al., 2005; Sharpe & Demichele, 1977). In addition to temperature, substrate concentration also plays a role: higher substrate availability increases the reaction rate for a given value of the affinity constant. Therefore, our examples imply that it is problematic to assume that, under high substrate concentrations, the temperature response curve only reflects the temperature-dependence of enzyme catalysis rate. In reality, the temperature response curve also depends on the temperature dependence of the affinity parameter, so that a high substrate concentration cannot ensure Equation 37 to derive a temperature response curve that accurately approximates that of v_{max}. As this assumption is the foundation of the macromolecular rate theory by Hobbs et al. (2013) that is built off the study by Peterson et al. (2004), a comprehensive analysis is presented elsewhere (Tang & Riley, 2024).

4. How to Make Physical Rules-Based Approaches Easily Accessible?

With the three examples above, we showed that it is feasible and advantageous to formulate EBMs using physical rules-based approaches. These examples also clearly reveal that empirically based functions can result in inconsistent influences of processes, for example, replacing the temperature-dependent parameterization in the third example with a single multiplier function will cause the temperature to improperly affect the relevant subprocesses. In another example shown by Tang and Riley (2021), it was found that, to obtain reasonable model fitting with observations, the law of the minimum model had to use incorrect parameter values. Nevertheless, compared to the more intuitive empirically based approaches, significant efforts are needed to realize these proposed advantages, at least partially because modeling equations are less intuitive to understand and may contain more parameters to constrain, either by theory or calibration. We therefore recommend the following steps to achieve this goal.

First, bring more expertise and knowledge of mathematical physics into the field of ecosystem biogeochemistry. This is already done well in the research area of ecosystem biogeophysics, where physical rules like Ohm's law for resistor networks, transport theories of diffusion and advection are used to formulate the exchange and temporal evolution of mass and energy between soil, water, atmosphere, and other related components (Shuttleworth & Wallace, 1985), and textbooks also explain those applications in detail (e.g., Bonan, 2019). For ecosystem biogeochemistry, constructive effort can be applied through (a) building long-term and stable collaborations between biogeochemistry empiricists, applied mathematicians, chemists, and physicists who are keen to model ecosystem biogeochemistry mechanistically, and (b) enhanced exposure of students in ecosystem biogeochemistry to concepts in mathematical physics including law of mass action, chemical reaction theories, and basic reactive transport modeling. From a pedagogic perspective, students could be challenged to test the classic Michaelis-Menten equation, or a linear model, using data sets where the use of equilibrium chemistry approximation is necessary. Also, faculty with expertise in biochemistry can team with colleagues in mathematics and physics to develop a course on mathematical biogeochemistry. This approach could motivate the next generation of geoscientists to engage in developing physical rules-based ecosystem biogeochemical modeling.

Second, EBMs formulated using physical rules-based approaches will often be mathematically more complex, which may be contrary to the heuristic belief that models should be simple (in appearance). We fully agree that unnecessary complexity should be avoided. However, we contend that the research community should be more open to endorsing the higher complexity resulting from constructions based on solid mathematical and physics-based logic, as compared to the simpler empirical equations typically derived by regression with context-dependent measurements. For instance, when the Lagrangian of the standard model of particle physics is written explicitly term by term (Shivni, 2016), the resulting gargantuan equation may easily fill a whole regular page of a journal paper. Nonetheless, the astonishing success of the standard model so far does not warrant any omission of its terms, and when the model is explained term by term, the mechanisms are readily interpretable. As we argued previously, given that ecosystem biogeochemistry encapsulates both living organisms and inanimate matter, which exist in different phases, and interact from very small to very large spatial scales, the true governing equations of an EBM may be as complex as, if not more complex than, the standard model of particle physics (in terms of length when they are put down onto paper). Therefore, the complexity of EBMs should not be judged by the number of mathematical terms involved, rather the complexity should be measured by the basic ideas of physics and mathematics being incorporated. Notably, even for a very complex system, physical rules will provide additional constraints to significantly reduce the actual degrees of freedom, so that the resultant EBM may be de facto simpler. As such, mechanistically more interpretable, and logically more coherent EBMs could be developed with improved model predictability.

Third, by being less intuitive than the empirically based approaches, physical rules-based approaches may appear too challenging to be integrated into the current training cycle of graduate students and postdoctoral researchers. Additionally, funding agencies often favor short-term grants over long-term support. To address this resource limitation, we contend that the research community should acknowledge the inadequacy of empirically based EBMs in providing specific guidance for precision ecosystem managements, which are urgently needed for mitigating climate change. Funding agencies and the research community can collaborate through workshops to identify strategies for effectively allocating resources to develop physical rules-based EBMs and collecting relevant empirical data. On the other hand, we observe that the development of physical rules-based parameterization can be modular for many processes (see Box 3). Consequently, a graduate student could in principle

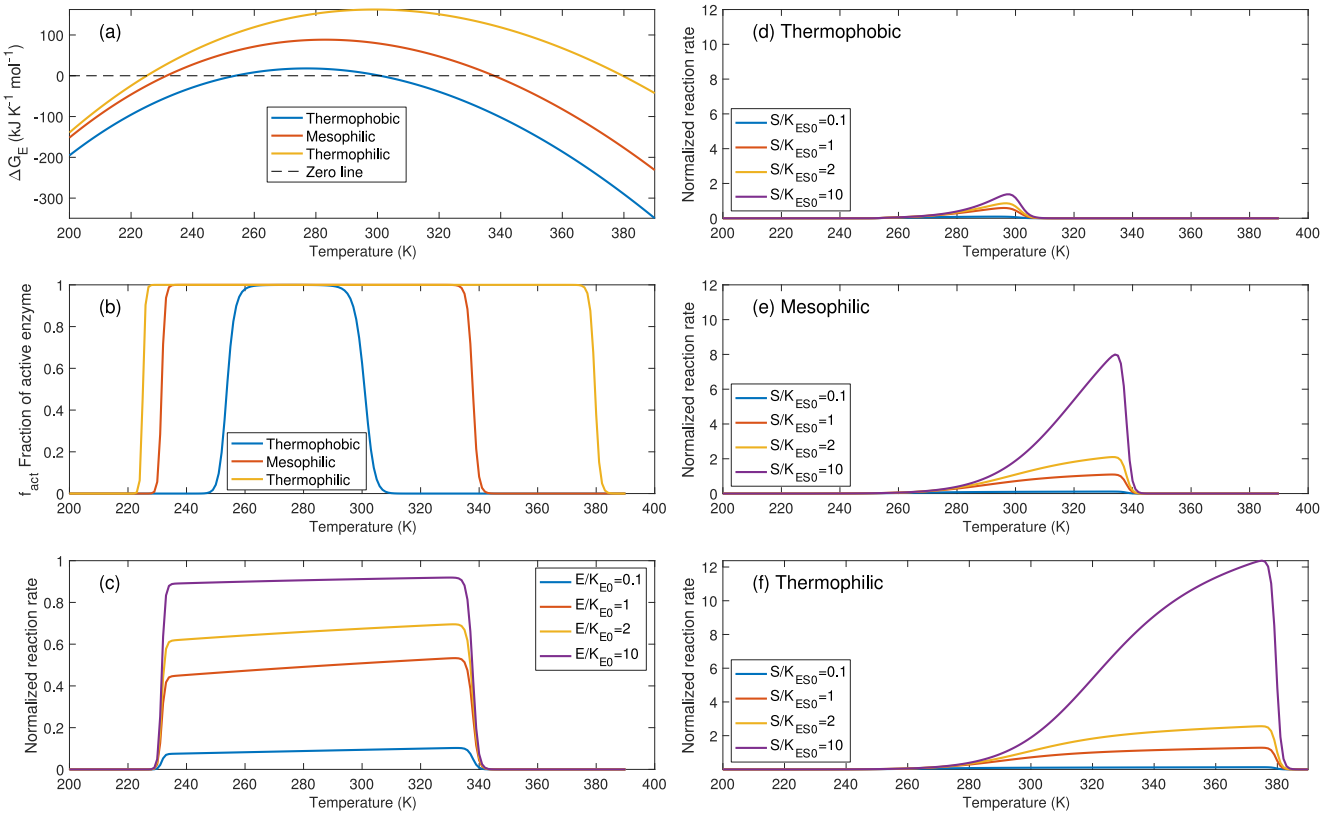

Figure 8. (a) Examples of unfolding Gibbs free energy ΔG_E as a function of temperature based on Equation 35; (b) fraction of active enzymes under different temperatures based on Equation 34; (c) normalized reaction rate as a function of temperature for a hypothetical mesophilic enzyme for different ratios of enzyme concentration to affinity constant E/K_{E0}, based on Equation 36; (d) normalized reaction rated as a function of temperature for a hypothetical mesophilic enzyme, based on Equation 37; (e) and (f) are the same as (d) but for hypothetic mesophilic and thermophilic enzymes, respectively, and at different ratios of substrate concentration to affinity constant S/K_{ES0}. For both K_E and K_{ES}, we use $K(T) = K_0 \exp\left(-\frac{37300}{R_g T}\left(1 - \frac{T}{290}\right)\right)$ computed from the activation energy of glucose uptake (58 kJ mol^{-1}) reported by Reinhardt et al. (1997), and the activation energy of diffusion (20.7 kJ mol^{-1}) reported in Table 2.3 by Stein (2012). Accordingly, for both $v_{max,E}$ and $v_{max,ES}$, we use $v_{max}(T) = \frac{T}{290}\exp\left(-\frac{58000}{R_g T}\left(1 - \frac{T}{290}\right)\right)$.

focus on an individual process or module and test it. In their Ph.D. thesis, the final chapter could be dedicated to integrating the developed module into a larger, more complex soil or ecosystem model. Such educational experience would be very valuable, allowing students to delve into specific aspects of broader problems, while developing their module(s) and gaining an understanding of the overall context during integration. Such an approach is scalable depending on student math/physics skills, interest, data availability etc. With more students trained in this manner, more postdoctoral researchers will be prepared to tackle the challenging task of developing physical rules-based EBMs.

Fourth, we acknowledge that EBMs formulated using physical rules-based approaches will usually be computationally more complex and demanding, and therefore may be more difficult to apply at large scales. This scalability issue could be solved in two steps: (a) creating an open-source numerical library that consists of processes that are formulated using physical rules, but are provided with user friendly software interfaces to be used in other models (Riley et al., 2022; Tang et al., 2022), and (b) improving the numerical efficiency of these model formulations by leveraging new developments in machine learning and artificial intelligence. The first idea has led to software like OpenFOAM (Jasak, 2009) and COMSOL (Pryor, 2012) that are able to solve computational fluid dynamics problems in various configurations. The second idea is currently used to develop more advanced parameterization schemes, such as turbulence closure schemes (Kurz et al., 2023) and cloud processes parameterization for atmospheric models (Beucler et al., 2023). In ecosystem biogeochemical modeling, a machine learning model, when pretrained with a physical rules-based ecosystem biogeochemistry model, could

conduct spatiotemporal extrapolation more efficiently and even outperform the original EBM, successfully resolving the challenge of upscaling (Liu et al., 2022, 2024).

Last but not least, as can be seen from our three examples, physical rules-based approaches require substantial comprehensive empirical data support for both forming the conceptual model and parameterization. Fortunately, much relevant data is available from the literature, such as solubility and diffusivities of chemical tracers in water and air (Cussler, 2009; Sander, 2023). New biological data that characterize the morphology and metabolic rates of organisms, as well as intra- and inter-specific interactions among organisms, however, are needed. These data should be collected more frequently, together with macro-chemical data such as carbon and nutrient concentrations and soil physical properties. Microbial elemental stoichiometry, morphology, number or mass density in various soils, and their relative abundances and activity level under various conditions, will be very helpful for formulating physical rules-based models of soil microbial processes. However, for most of these quantities there are no established measurement methods and novel tracer experiments are only now starting to provide detailed microbial trait and rate data (e.g., Canarini et al., 2020; Warren & Manzoni, 2023). For plants, more in-situ phenological data and morphological data (such as leaf sizes, thickness, height, root architecture, and morphology) will be essential to robustly formulate biogeochemistry using physical rules, which can also improve the model representation of biogeophysics, such as water and heat exchange between soils, plants, and atmosphere. On the one hand, existing databases (e.g., ESS-DIVE (https://ess-dive.lbl.gov), TRY plant trait database (Kattge & Sandel, 2020)) can aid making data accessible. On the other hand, physical rules-based approaches can suggest more specific answers to the question from empiricists to modelers: "what do you want us to measure?"

5. Summary

Lao Tzu said, "*the Tao that can be told is not the eternal Tao,*" which we interpret 2600 years later to mean that all ecosystem biogeochemical models (EBMs) are inherently limited. Nevertheless, we argue that the currently popular approach, which extensively uses empirically based functions to formulate biogeochemical processes, limits EBM predictability and the community's attempts to incorporate needed improvements. Instead, by treating EBM parameterization as a lossy data compression problem and solving it with physical rules-based approaches proposed here, more robust and accurate EBMs can be developed. Compared with empirical functions, the primary physical rules are more consistent with current knowledge of the world, and the derived physical rules are less context dependent and have more easily quantifiable uncertainty. Consequently, physical rules-based EBMs will be more robust and less plagued by predictive equifinality, even though they appear to be more complex than empirical function-based EBMs. Moreover, using physical rules to formulate biogeochemical processes will reveal more detailed insights about the interactions between the entities involved, which will facilitate the design of more targeted empirical experiments. To build EBMs that maximally use current knowledge of physical rules, we advocate more and closer interdisciplinary collaborations in both research and education between scientists in biogeochemistry, biophysics, soil physics, and mathematics.

Nomenclature. For Units, "Variable" Means the Unit Is Problem Formulation Dependent

Symbol	Unit Meaning and places of use
$f_j(\theta_j)$	Effect multiplier from influencer θ_j; Equations 4 and 5.
$f_{act}(T)$	Fraction of enzymes being active; Equations 34, 36 and 37.
$h_C(s)$	Substrate dependency for the compromise model.
$h_D(s)$	Substrate dependency for the DEB model.
$h_P(s)$	Substrate dependency for the Pirt model.
$j_{A,max}$ (s^{-1})	Maximum substrate uptake rate for the DEB model.
$k_{1,w}$ (m mol^{-1} s^{-1})	Microbe-substrate forward binding rate; Equations 7 and 8.
$k_{2,p}$ (s^{-1})	Maximum substrate uptake rate per site; Equation 7.
$k_{ij,2}$ (s^{-1})	Maximum uptake rate of substrate S_i by enzyme E_j; Equation 16.
m_P (s^{-1})	Specific microbial maintenance rate.
n_E	Number of amino acid residues of the enzyme; Equation 35.
$q_C(s)$ (s^{-1})	Specific substrate uptake rate for the compromise model.

$q_D(s)$ (s^{-1})	Specific substrate uptake rate for the DEB model.
$q_P(s)$ (s^{-1})	Specific substrate uptake rate for the Pirt model.
r_C (M)	Bacteria cell radius; Equations 8 and 9.
r_m (M)	Microbial microsite radius; Equation 10.
v_m (m^3)	Microbial microsite volume; Equation 10.
$v_{max}(T)$ (s^{-1})	Maximum substrate processing rate at temperature T; Equation 32.
$v_{max,E}$ (s^{-1})	Maximum substrate processing rate; Equation 29.
$v_{max,E}(T)$ (s^{-1})	Maximum substrate processing rate at temperature T; Equation 36.
$v_{max,ES}$ (s^{-1})	Maximum substrate processing rate; Equation 30.
$v_{max,ES}(T)$ (s^{-1})	Maximum substrate processing rate at temperature T; Equation 37.
$v_{max,l}$ (s^{-1})	Specific maximum uptake rate of C_l; Equations 22–25, 27.
$v_{max,S}$ (s^{-1})	Specific maximum uptake rate of C_S; Equations 22–24, 26, 28.
B_m (mol cell m^{-3})	Mean microbial biomass in a microsite; Equation 7.
B_R (mol C m^{-3})	Reverse microbial biomass in DEB model.
B_V (mol C m^{-3})	Structural microbial biomass in DEB model.
C_b (mol C m^{-3})	Microbial biomass; Equations 23 and 24.
C_j	Inverse of covariance matrix for variable Y_j; Equations 2 and 3.
C_l (mol C m^{-3})	Fast decaying carbon pool; Equations 22–25, 27, 28.
C_s (mol C m^{-3})	mol C m^{-3} Slow decaying carbon pool; Equations 22–24, 26–28.
ΔC_p (J K^{-1} (mol amino acid)$^{-1}$)	Heat capacity; Equation 35
$D_{g,0}$ (m^2 s^{-1})	Gaseous diffusivity; Equation 11.
$D_{w,0}$ (m^2 s^{-1})	Aqueous diffusivity; Equations 10 and 11.
E (mol m^{-3})	Enzyme concentration; Equation 6
E_j (Variable)	Consumer concentration; Equations 14 and 15.
$E_{j,T}$ (Variable)	Total consumer concentration; Equations 14, 17, 18 and 20.
F_{n_l} (Variable)	Variable Total flux of variable Y_{n_l}; Figure 3.
ΔG_E (J mol^{-1})	Gibbs free energy of enzyme unfolding; Equations 34 and 35.
H_{k_l} (Variable)	Process function corresponding to R_{k_l}; Figure 3.
ΔH_E (J mol^{-1})	Enthalpy of the enzyme unfolding; Equation 35.
ΔH_K (J mol^{-1})	Enthalpy of affinity parameter; Equation 33.
ΔH_r (J mol^{-1})	Enthalpy of enzymatic chemical reaction; Equation 32.
J_0	Cost function contributed by prior information; Equation 1.
K (mol m^{-3})	Substrate affinity parameter; Equation 6.
K_0 (mol m^{-3})	Substrate affinity parameter at temperature T_0; Equation 33.
K_E (mol m^{-3})	Substrate affinity parameter; Equation 29.
K_{ES} (mol m^{-3})	Substrate affinity parameter; Equation 30.
$K_E(T)$ (mol m^{-3})	Substrate affinity parameter at temperature T; Equation 36.
$K_{ES}(T)$ (mol m^{-3})	Substrate affinity parameter at temperature T; Equation 37
K_{ij} (Variable)	Affinity parameter between substrate S_i and consumer E_j; Equation 15.
K_l (mol m^{-3})	Microbial affinity parameter to carbon pool C_l; Equations 22–25, 27, 28.
K_s (mol m^{-3})	Microbial affinity parameter to carbon pool C_s; Equations 22–24, 26–28.
$K_{l,M}$ (mol m^{-3})	Affinity parameter between carbon pool C_l and mineral M. Equation 24.
K_w (mol m^{-3})	Effective substrate affinity parameter in soil; Equation 7.
$K_{w,0}$ (mol m^{-3})	Substrate affinity parameter in water; Figure 4.
$M_k(R_l)$ (Variable)	kth scaling rule between processes R_l; Equations 2 and 3.
N	Number of transporters per microbial cell; Equations 8 and 9.
N_A (mol^{-1})	Avogadro number; Equation 8
P (Variable)	Product concentration from consumption of substrate; Equation 19.
P_{ij} (Variable)	Product concentration by E_j from consuming S_i; Equation 16.
$Q_{j_{l-1}}$ (Variable)	Mechanistic interactions between $Y_{m_{l-1}}$; Figure 3.
R (Variable)	Rate, or conductance, or resistance; Equations 4–6, 29–31, 36, 37.
R_0 (Variable)	Reference value of R; Equations 4 and 5.
R_g (J mol^{-1} K^{-1})	Universal gas constant; Equation 32.

R_{k_l} (Variable)	Process rate; Figure 3.
S (mol m^{-3})	Substrate concentration; Equation 6.
ΔS_E (J mol^{-1})	Entropy of enzyme unfolding; Equation 35.
S_i (Variable)	Free concentration ith Substrate; Equations 13, 15 and 17.
$S_{i,T}$ (Variable)	Total concentration of ith substrate; Equations 13 and 17.
T (K)	Thermodynamic temperature; Equations 32–37.
T_0 (K)	Reference thermodynamic temperature; Equations 32 and 33.
T_H (K)	Thermodynamic temperature when ΔH_E equals zero; Equation 35.
V (mol m^{-3} s^{-1})	Total substrate uptake rate; Equations 20 and 21.
$V_{max,j}$ (mol m^{-3} s^{-1})	Maximum uptake rate by consumer E_j; Equation 21.
X_{ij} (Variable)	Substrate-consumer complex between S_i and E_j; Equations 15–18.
X (Variable)	Total substrate-consumer complex; Equations 18 and 19.
Y_G	Biomass yield coefficient; Equations 22–24; Box 2.
$Y_j(R_l)$ (Variable)	Generic model variable; Equations 2 and 3.
Y_{jO} (Variable)	Observations corresponding to $Y_j(R_l)$; Equations 2 and 3.
$Y_{m_{l-1}}$ (Variable)	State variable at the fine scale; Figure 2.
Y_{n_l} (Variable)	State variable at the coarse scale; Figure 2.
Y_{sR}	Reserve biomass yield for the DEB model; Box 2.
α	Scaling parameter from microbial biomass to substrate binding sites; Equations 23, 24, 27 and 28.
λ_k	Lagrangian multiplier for kth scaling rule; Equations 2 and 3.
κ (s^{-1})	Specific reserve biomass mobilization rate; Box 2.
δ (m)	Water film thickness; Equations 10 and 12.
κ_m (s^{-1})	Specific substrate transfer rate between soil matrix and microbial microsite; Equations 7 and 10.
θ_m (Variable)	Model parameters; Equations 2 and 3.
$\mu_{max,P}$ (s^{-1})	Specific maximum biomass growth rate for the Pirt model; Box 2.
$\mu_{max,C}$ (s^{-1})	Specific maximum biomass growth rate for the compromise model; Box 2.
$\mu_C(s)$ (s^{-1})	Specific biomass growth rate for the compromise model; Box 2.
$\mu_D(s)$ (s^{-1})	Specific biomass growth rate for the DEB model; Box 2.
$\mu_P(s)$ (s^{-1})	Specific biomass growth rate for the Pirt model; Box 2.
ψ (MPa)	Soil matric potential; Equation 12.
ϕ_g (m^3 m$^{-3)}$	Air-filled soil porosity; Equation 11.
ϕ_w (m^3 m^{-3})	Water-filled soil porosity; Equation 11.
τ_g	Soil tortuosity for gas diffusion; Equation 11.
τ_w	Soil tortuosity for solute diffusion; Equation 11.
$\Phi_{l\text{-}1}$ (Variable)	Variable Fine-scale physical constraints; Figure 3.
Φ_l (Variable)	Coarse-scale physical constraints; Figure 3.

Acknowledgments
JYT and WJR are supported by the Director, Office of Science, Office of Biological and Environmental Research of the US Department of Energy under contract no. DE-AC02-05CH11231 as part of the Belowground Biogeochemistry Science Focus Area and the Reducing Uncertainties in Biogeochemical Interactions through Synthesis and Computation (RUBISCO) Scientific Focus Area. SM has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Programme (Grant 101001608). We sincerely appreciate the constructive comments from the editor, associate editor, and two anonymous reviewers. They helped us more effectively delivering our thoughts through this manuscript.

Data Availability Statement

This study only uses published data. Specifically for Figure 5, data for the four soil classes are (a) medium to fine texture soils from (Doran et al., 1990); (b) coarse texture soil from (Doran et al., 1990); (c) other soils from (Doran et al., 1990); and (d) soils from Franzluebbers (1999), respectively. Parameters used for the calculation are from Tang and Riley (2019b).

References

Abichou, M., Fournier, C., Dornbusch, T., Chambon, C., Baccar, R., Bertheloot, J., et al. (2013). Re-parametrisation of Adel-wheat allows reducing the experimental effort to simulate the 3D development of winter wheat. In *International conference on functional structure plant models*.

Abramoff, R. Z., Guenet, B., Zhang, H. C., Georgiou, K., Xu, X. F., Rossel, R. A. V., et al. (2022). Improved global-scale predictions of soil carbon stocks with millennial version 2. *Soil Biology and Biochemistry*, *164*, 108466. https://doi.org/10.1016/j.soilbio.2021.108466

Ahlström, A., Smith, B., Lindström, J., Rummukainen, M., & Uvo, C. B. (2013). GCM characteristics explain the majority of uncertainty in projected 21st century terrestrial ecosystem carbon balance. *Biogeosciences*, *10*(3), 1517–1528. https://doi.org/10.5194/bg-10-1517-2013

Anderson, P. W. (1972). More is different - Broken symmetry and nature of hierarchical structure of science. *Science*, *177*(4047), 393. https://doi.org/10.1126/science.177.4047.393

Antoniewicz, M. R. (2021). A guide to metabolic flux analysis in metabolic engineering: Methods, tools and applications. *Metabolic Engineering*, *63*, 2–12. https://doi.org/10.1016/j.ymben.2020.11.002

Arrichiello, V., & Gualeni, P. (2020). Systems engineering and digital twin: A vision for the future of cruise ships design, production and operations. *International Journal on Interactive Design and Manufacturing*, *14*(1), 115–122. https://doi.org/10.1007/s12008-019-00621-3

Atkins, P., & de Paula, J. (2006). *Physical chemistry* (8th Edn, Vol. 41). W. H. Freeman and Company.

Atkinson, C. L., Sansom, B. J., Vaughn, C. C., & Forshay, K. J. (2018). Consumer aggregations drive nutrient dynamics and ecosystem metabolism in nutrient-limited systems. *Ecosystems*, *21*(3), 521–535. https://doi.org/10.1007/s10021-017-0166-4

Azizi-Rad, M., Guggenberger, G., Ma, Y. M., & Sierra, C. A. (2022). Sensitivity of soil respiration rate with respect to temperature, moisture and oxygen under freezing and thawing. *Soil Biology and Biochemistry*, *165*, 108488. https://doi.org/10.1016/j.soilbio.2021.108488

Baldocchi, D. D., Hutchison, B. A., Matt, D. R., & Mcmillen, R. T. (1985). Canopy radiative-transfer models for spherical and known leaf inclination angle distributions - A test in an Oak Hickory forest. *Journal of Applied Ecology*, *22*(2), 539–555. https://doi.org/10.2307/2403184

Ball, J. T., Woodrow, I. E., & Berry, J. A. (1987). *A model predicting stomatal conductance and its contribution to the control of photosynthesis under different environmental conditions. paper presented at the VIIth International Congress on progress in photosynthesis research*. Springer.

Bao, J. L., & Truhlar, D. G. (2017). Variational transition state theory: Theoretical framework and recent developments. *Chemical Society Reviews*, *46*(24), 7548–7596. https://doi.org/10.1039/c7cs00602k

Barraquand, F. (2014). Functional responses and predator-prey models: A critique of ratio dependence. *Theoretical Ecology*, *7*(1), 3–20. https://doi.org/10.1007/s12080-013-0201-9

Basir, S., & Senocak, I. (2022). Physics and equality constrained artificial neural networks: Application to forward and inverse problems with multi-fidelity data fusion. *Journal of Computational Physics*, *463*, 111301. https://doi.org/10.1016/j.jcp.2022.111301

Batchelor, G. K. (1967). *An introduction to fluid dynamics*. (Vol. xviii, p. 615). Cambrige, .

Bauer, J., Herbst, M., Huisman, J. A., Weihermüller, L., & Vereecken, H. (2008). Sensitivity of simulated soil heterotrophic respiration to temperature and moisture reduction functions. *Geoderma*, *145*(1–2), 17–27. https://doi.org/10.1016/j.geoderma.2008.01.026

Bear, J. (1972). *Dynamics of fluids in porous media*. (Vol. xvii, p. 764). American Elsevier Pub. Co., .

Beddington, J. R. (1975). Mutual interference between parasites or predators and its effect on searching efficiency. *Journal of Animal Ecology*, *44*(1), 331–340. https://doi.org/10.2307/3866

Beeftink, H. H., Vanderheijden, R. T. J. M., & Heijnen, J. J. (1990). Maintenance requirements - Energy supply from simultaneous endogenous respiration and substrate consumption. *FEMS Microbiology Ecology*, *73*(3), 203–209. https://doi.org/10.1016/0378-1097(90)90731-5

Berg, H. C., & Purcell, E. M. (1977). Physics of chemoreception. *Biophysical Journal*, *20*(2), 193–219. https://doi.org/10.1016/S0006-3495(77)85544-6

Berryman, A. A. (1992). The origins and evolution of predator prey theory. *Ecology*, *73*(5), 1530–1535. https://doi.org/10.2307/1940005

Beucler, T. G., Ebert-Uphoff, I., Rasp, S., Pritchard, M., & Gentine, P. (2023). Machine learning for clouds and climate. In *Clouds and their climatic impacts: Radiation, circulation, and precipitation*. S. C. Sullivan, & H. Hoose, (Eds.). https://doi.org/10.1002/9781119700357.ch16

Beven, K., & Freer, J. (2001). Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *Journal of Hydrology*, *249*(1–4), 11–29. https://doi.org/10.1016/S0022-1694(01)00421-8

Bloom, A. A., & Williams, M. (2015). Constraining ecosystem carbon dynamics in a data-limited world: Integrating ecological "common sense" in a model-data fusion framework. *Biogeosciences*, *12*(5), 1299–1315. https://doi.org/10.5194/bg-12-1299-2015

Blyth, E. M., Arora, V. K., Clark, D. B., Dadson, S. J., De Kauwe, M. G., Lawrence, D. M., et al. (2021). Advances in land surface modelling. *Current Climate Change Reports*, *7*(2), 45–71. https://doi.org/10.1007/s40641-021-00171-5

Boltzmann, L. (1964). *Lectures on gas theory*. (Vol. ix, p. 490). University of California Press, .

Bonan, G. B. (2019). *Climate change and terrestrial ecosystem modeling*. Cambridge University Press.

Bouskill, N. J., Riley, W. J., & Tang, J. Y. (2014). Meta-analysis of high-latitude nitrogen-addition and warming studies implies ecological mechanisms overlooked by land models. *Biogeosciences*, *11*(23), 6969–6983. https://doi.org/10.5194/bg-11-6969-2014

Briggs, G. E., & Haldane, J. B. S. (1925). A note on the kinetics of enzyme action. *Biochemical Journal*, *19*(2), 338–339. https://doi.org/10.1042/bj0190338

Brocchieri, L., & Karlin, S. (2005). Protein length in eukaryotic and prokaryotic proteomes. *Nucleic Acids Research*, *33*(10), 3390–3400. https://doi.org/10.1093/nar/gki615

Buchkowski, R. W., Barel, J. M., Jassey, V. E. J., & Lindo, Z. (2022). Cannibalism has its limits in soil food webs. *Soil Biology and Biochemistry*, *172*, 108773. https://doi.org/10.1016/j.soilbio.2022.108773

Buckley, T. N. (2019). How do stomata respond to water status? *New Phytologist*, *224*(1), 21–36. https://doi.org/10.1111/nph.15899

Caldwell, P. M., Terai, C. R., Hillman, B., Keen, N. D., Bogenschutz, P., Lin, W., et al. (2021). Convection-permitting simulations with the E3SM global atmosphere model. *Journal of Advances in Modeling Earth Systems*, *13*(11), e2021MS002544. https://doi.org/10.1029/2021MS002544

Canarini, A., Wanek, W., Watzka, M., Sandén, T., Spiegel, H., Santrucek, J., & Schnecker, J. (2020). Quantifying microbial growth and carbon use efficiency in dry soil environments via ^{18}O water vapor equilibration. *Global Change Biology*, *26*(9), 5333–5341. https://doi.org/10.1111/gcb.15168

Candel, S., Thévenin, D., Darabiha, N., & Veynante, D. (1999). Progress in numerical combustion. *Combustion Science and Technology*, *149*(1–6), 297–337. https://doi.org/10.1080/00102209908952110

Cao, M. K., Gregson, K., & Marshall, S. (1998). Global methane emission from wetlands and its sensitivity to climate change. *Atmospheric Environment*, *32*(19), 3293–3299. https://doi.org/10.1016/S1352-2310(98)00105-8

Carvalhais, N., Forkel, M., Khomik, M., Bellarby, J., Jung, M., Migliavacca, M., et al. (2014). Global covariation of carbon turnover times with climate in terrestrial ecosystems. *Nature*, *514*(7521), 213–217. https://doi.org/10.1038/nature13731

Chakrawal, A., Calabrese, S., Herrmann, A. M., & Manzoni, S. (2022). Interacting bioenergetic and stoichiometric controls on microbial growth. *Frontiers in Microbiology*, *13*. https://doi.org/10.3389/fmicb.2022.859063

Chandel, A. K., Jiang, L. F., & Luo, Y. Q. (2023). Microbial models for simulating soil carbon dynamics: A review. *J Geophys Res-Biogeo*, *128*(8), e2023JG007436. https://doi.org/10.1029/2023jg007436

Chapman, S., & Cowling, T. G. (1990). *The mathematical theory of non-uniform gases: An account of the kinetic theory of viscosity, thermal conduction, and diffusion in gases*. (3rd ed., Vol. xxiv, p. 422). Cambridge University Press.

Chen, L. Y., Goldenfeld, N., & Oono, Y. (1994). Renormalization-group theory and variational calculations for propagating fronts. *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, *49*(5), 4502–4511. https://doi.org/10.1103/PhysRevE.49.4502

Chen, S., & Doolen, G. D. (1998). Lattice Boltzmann method for fluid flows. *Annual Review of Fluid Mechanics*, *30*(1), 329–364. https://doi.org/10.1146/annurev.fluid.30.1.329

Cheng, Y. W., Bouskill, N. J., & Brodie, E. L. (2019). Model exploration of interactions between algal functional diversity and productivity in chemostats to represent open ponds systems across climate gradients. *Ecological Modelling*, *406*, 121–132. https://doi.org/10.1016/j.ecolmodel.2019.05.007

Cheng, Z. X., Sun, H. M., Takeuchi, M., & Katto, J. (2018). Deep convolutional autoencoder-based lossy image compression. *Pict Cod Symp*, 253–257. https://doi.org/10.1109/pcs.2018.8456308

Clapp, R. B., & Hornberger, G. M. (1978). Empirical equations for some soil hydraulic-properties. *Water Resources Research*, *14*(4), 601–604. https://doi.org/10.1029/WR014i004p00601

Collatz, G. J., Berry, J. A., Farquhar, G. D., & Pierce, J. (1990). The relationship between the rubisco reaction-mechanism and models of photosynthesis. *Plant, Cell and Environment*, *13*(3), 219–225. https://doi.org/10.1111/j.1365-3040.1990.tb01306.x

Cowan, I. R., & Farquhar, G. D. (1997). Stomatal function in relation to leaf metabolism and environment. In D. H. Jennings (Ed.), *Integration of activity in the higher plant* (pp. 471–505). Cambridge University Press.

Cussler, E. L. (2009). *Diffusion, mass transfer in fluid systems* (3rd ed., p. 647). Cambridge University Press.

Dai, Y. J., Dickinson, R. E., & Wang, Y. P. (2004). A two-big-leaf model for canopy temperature, photosynthesis, and stomatal conductance. *Journal of Climate*, *17*(12), 2281–2299. https://doi.org/10.1175/1520-0442(2004)017<2281:Atmfct>2.0.Co;2

Davies-Barnard, T., Meyerholt, J., Zaehle, S., Friedlingstein, P., Brovkin, V., Fan, Y., et al. (2020). Nitrogen cycling in CMIP6 land surface models: Progress and limitations. *Biogeosciences*, *17*(20), 5129–5148. https://doi.org/10.5194/bg-17-5129-2020

DeAngelis, D. L., Goldstein, R. A., & O'Neill, R. V. (1975). A model for trophic interaction. *Ecology*, *56*(4), 881–892. https://doi.org/10.2307/1936298

De Falco, B., Giannino, F., Carteni, F., Mazzoleni, S., & Kim, D. H. (2022). Metabolic flux analysis: A comprehensive review on sample preparation, analytical techniques, data analysis, computational modelling, and main application areas. *RSC Advances*, *12*(39), 25528–25548. https://doi.org/10.1039/d2ra03326g

de Groor, S. R., & Mazur, P. (1984). *Non-equilbrium thermodynamics*. Dover publications, Inc.

De Kauwe, M. G., Medlyn, B. E., Walker, A. P., Zaehle, S., Asao, S., Guenet, B., et al. (2017). Challenging terrestrial biosphere models with data from the long-term multifactor Prairie Heating and CO_2 Enrichment experiment. *Global Change Biology*, *23*(9), 3623–3645. https://doi.org/10.1111/gcb.13643

de Vries, J., & Archibald, J. M. (2018). Plant evolution: Landmarks on the path to terrestrial life. *New Phytologist*, *217*(4), 1428–1434. https://doi.org/10.1111/nph.14975

Dietze, M. C. (2017). Prediction in ecology: A first-principles framework. *Ecological Applications*, *27*(7), 2048–2060. https://doi.org/10.1002/eap.1589

Doran, J. W., Mielke, L. N., & Power, J. F. (1990). Microbial activity as regulated by soil water filled pore space. In *The 14th International congress of soil science*. (pp. 94–99).

Dronova, I., & Taddeo, S. (2022). Remote sensing of phenology: Towards the comprehensive indicators of plant community dynamics from species to regional scales. *Journal of Ecology*, *110*(7), 1460–1484. https://doi.org/10.1111/1365-2745.13897

Dudal, Y., & Gérard, F. (2004). Accounting for natural organic matter in aqueous chemical equilibrium models: A review of the theories and applications. *Earth-Science Reviews*, *66*(3–4), 199–216. https://doi.org/10.1016/j.earscirev.2004.01.002

Dwivedi, D., Riley, W. J., Torn, M. S., Spycher, N., Maggi, F., & Tang, J. Y. (2017). Mineral properties, microbes, transport, and plant-input profiles control vertical distribution and age of soil carbon stocks. *Soil Biology and Biochemistry*, *107*, 244–259. https://doi.org/10.1016/j.soilbio.2016.12.019

Dwivedi, D., Tang, J. Y., Bouskill, N., Georgiou, K., Chacon, S. S., & Riley, W. J. (2019). Abiotic and biotic controls on soil organo-mineral interactions: Developing model structures to analyze why soil organic matter persists. *Reviews in Mineralogy and Geochemistry*, *85*(1), 329–348. https://doi.org/10.2138/rmg.2019.85.11

Egli, T. (1991). On multiple-nutrient-limited growth of microorganisms, with special reference to dual limitation by carbon and nitrogen substrates. *Antonie van Leeuwenhoek*, *60*(3–4), 225–234. https://doi.org/10.1007/Bf00430367

ElGhawi, R., Kraft, B., Reimers, C., Reichstein, M., Körner, M., Gentine, P., & WinklerWinkler, A. J. (2023). Hybrid modeling of evapotranspiration: Inferring stomatal and aerodynamic resistances using combined physics-based and machine learning. *Environmental Research Letters*, *18*(3), 034039. https://doi.org/10.1088/1748-9326/acbbe0

Etienne, T. A., Cocaign-Bousquet, M., & Ropers, D. (2020). Competitive effects in bacterial mRNA decay. *Journal of Theoretical Biology*, *504*, 110333. https://doi.org/10.1016/j.jtbi.2020.110333

Exbrayat, J. F., Pitman, A. J., Abramowitz, G., & Wang, Y. P. (2013). Sensitivity of net ecosystem exchange and heterotrophic respiration to parameterization uncertainty. *Journal of Geophysical Research-Atmosphere*, *118*(4), 1640–1651. https://doi.org/10.1029/2012jd018122

Exbrayat, J. F., Pitman, A. J., Zhang, Q., Abramowitz, G., & Wang, Y. P. (2013). Examining soil carbon uncertainty in a global model: Response of microbial decomposition to temperature, moisture and nutrient limitation. *Biogeosciences*, *10*(11), 7095–7108. https://doi.org/10.5194/bg-10-7095-2013

Eyring, H. (1935). The activated complex and the absolute rate of chemical reactions. *Chemistry Review*, *17*(1), 65–77. https://doi.org/10.1021/cr60056a006

Farquhar, G. D., Von Caemmerer, S., & Berry, J. A. (1980). A biochemical model of photosynthetic CO_2 assimilation in leaves of C_3 species. *Planta*, *149*(1), 78–90. https://doi.org/10.1007/BF00386231

Feller, W. (1967). *An introduction to probability theory and its applications*, 3d ed., v. pp., Wiley, .

Feynman, R. P., Leighton, R. B., & Sands, M. (2011a). *The Feynman lectures on physics* (The New Millennium Edition). Basic Books.

Feynman, R. P., Leighton, R. B., & Sands, M. (2011b). *The Feynman Lectures on physics* (The New Millennium Edition). Basic Books.

Feynman, R. P., Leighton, R. B., & Sands, M. (2011c). *The Feynman Lectures on physics* (The New Millennium Edition). Basic Books.

Finkelstein, A. V., & Ptitsyn, O. (2016). *Protein physics: A course of Lectures* (2nd ed.). Academic Press.

Franklin, O., Harrison, S. P., Dewar, R., Farrior, C. E., Brännström, Å., Dieckmann, U., et al. (2020). Organizing principles for vegetation dynamics. *Nature Plants*, *6*(5), 444–453. https://doi.org/10.1038/s41477-020-0655-x

Franzluebbers, A. J. (1999). Microbial activity in response to water-filled pore space of variably eroded southern Piedmont soils. *Applied Soil Ecology*, *11*(1), 91–101. https://doi.org/10.1016/S0929-1393(98)00128-0

Ghosh, K., & Dill, K. A. (2009). Computing protein stabilities from their chain lengths. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(26), 10649–10654. https://doi.org/10.1073/pnas.0903995106

Ginzburg, L. R., & Akcakaya, H. R. (1992). Consequences of ratio-dependent predation for steady-state properties of ecosystems. *Ecology*, *73*(5), 1536–1543. https://doi.org/10.2307/1940006

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Regularization for deep learning. In *Adaptive Computation and Machine Learning series*. (pp. 221–265). MIT Press.

Grant, R. F. (1998). Simulation in ecosys of root growth response to contrasting soil water and nitrogen. *Ecological Modelling*, *107*(2–3), 237–264. https://doi.org/10.1016/S0304-3800(97)00221-4

Grant, R. F. (2013). Modelling changes in nitrogen cycling to sustain increases in forest productivity under elevated atmospheric CO_2 and contrasting site conditions. *Biogeosciences*, *10*(11), 7703–7721. https://doi.org/10.5194/bg-10-7703-2013

Grant, R. F., Dyck, M., & Puurveen, D. (2020). Nitrogen and phosphorus control carbon sequestration in agricultural ecosystems: Modelling carbon, nitrogen, and phosphorus balances at the Breton plots with ecosys under historical and future climates. *Canadian Journal of Soil Science*, *100*(4), 408–429. https://doi.org/10.1139/cjss-2019-0132

Grant, R. F., Mekonnen, Z. A., Riley, W. J., Arora, B., & Torn, M. S. (2017). Mathematical modelling of arctic polygonal tundra with ecosys: 2. Microtopography determines how CO_2 and CH_4 exchange responds to changes in temperature and precipitation. *Journal of Geophysical Research-Biogeoscience*, *122*(12), 3174–3187. https://doi.org/10.1002/2017jg004037

Greenway, H., & Munns, R. (1980). Mechanisms of salt tolerance in nonhalophytes. *Annual Review of Plant Physiology*, *31*(1), 149–190. https://doi.org/10.1146/annurev.pp.31.060180.001053

Guan, K. Y., Jin, Z., Peng, B., Tang, J., DeLucia, E. H., West, P. C., et al. (2023). A scalable framework for quantifying field-level agricultural carbon outcomes. *Earth-Science Reviews*, *243*, 104462. https://doi.org/10.1016/j.earscirev.2023.104462

Hanway, J. (1966). How a corn plant develops. *Crop Production*, *1–18*.

Harfenmeister, K., Itzerott, S., Weltzien, C., & Spengler, D. (2021). Detecting phenological development of winter wheat and winter barley using time series of Sentinel-1 and Sentinel-2. *Remote Sensing-Basel*, *13*(24), 5036. https://doi.org/10.3390/rs13245036

Harrison, S. P., Cramer, W., Franklin, O., Prentice, I. C., Wang, H., Brännström, Å., et al. (2021). Eco-evolutionary optimality as a means to improve vegetation and land-surface models. *New Phytologist*, *231*(6), 2125–2141. https://doi.org/10.1111/nph.17558

Higgins, S. I. (2017). Ecosystem assembly: A mission for terrestrial earth system science. *Ecosystems*, *20*(1), 69–77. https://doi.org/10.1007/s10021-016-0054-3

Hobbs, J. K., Jiao, W. T., Easter, A. D., Parker, E. J., Schipper, L. A., & Arcus, V. L. (2013). Change in heat capacity for enzyme catalysis determines temperature dependence of enzyme catalyzed rates. *ACS Chemical Biology*, *8*(11), 2388–2393. https://doi.org/10.1021/cb4005029

Hoem, J. M. (1987). Statistical-analysis of a multiplicative model and its application to the standardization of vital-rates - A review. *International Statistical Review*, *55*(2), 119–152. https://doi.org/10.2307/1403190

Holling, C. S. (1959). Some characteristics of simple types of predation and parasitism. *The Canadian Entomologist*, *91*(7), 385–398. https://doi.org/10.4039/ent91385-7

Holmes, M. J., Parker, N. G., & Povey, M. J. W. (2011). Temperature dependence of bulk viscosity in water using acoustic spectroscopy. *Journal of Physics: Conf. Ser*, *269*, 012011. https://doi.org/10.1088/1742-6596/269/1/012011

Horn, F., & Jackson, R. (1972). General mass action kinetics. *Archive for Rational Mechanics and Analysis*, *47*(2), 81–116. https://doi.org/10.1007/BF00251225

Houska, T., Kraus, D., Kiese, R., & Breuer, L. (2017). Constraining a complex biogeochemical model for CO2 and N2O emission simulations from various land uses by model-data fusion. *Biogeosciences*, *14*(14), 3487–3508. https://doi.org/10.5194/bg-14-3487-2017

Huang, Y., Guenet, B., Ciais, P., Janssens, I. A., Soong, J. L., Wang, Y. L., et al. (2018). ORCHIMIC (v1.0), a microbe-mediated model for soil organic matter decomposition. *Geoscientific Model Development*, *11*(6), 2111–2138. https://doi.org/10.5194/gmd-11-2111-2018

Huntzinger, D. N., Michalak, A. M., Schwalm, C., Ciais, P., King, A. W., Fang, Y., et al. (2017). Uncertainty in the response of terrestrial carbon sink to environmental drivers undermines carbon-climate feedback predictions. *Scientific Reports*, *7*(1), 4765. https://doi.org/10.1038/s41598-017-03818-2

Hussain, A. J., Al-Fayadh, A., & Radi, N. (2018). Image compression techniques: A survey in lossless and lossy algorithms. *Neurocomputing*, *300*, 44–69. https://doi.org/10.1016/j.neucom.2018.02.094

Jarvis, P. G. (1976). The interpretation of the variations in leaf water potential and stomatal conductance found in canopies in the field. *Philosophical Transactions of the Royal Society of London*, *273*(927), 593–610. https://doi.org/10.1098/rstb.1976.0035

Jasak, H. (2009). OpenFOAM: Open source CFD in research and industry [Software]. *International Journal of Naval Architecture and Ocean*, *1*, 89–94. https://doi.org/10.3744/Jnaoe.2009.1.2.089

Jaynes, E. T. (2003). *Probability theory: The logic of science*. Cambridge University Press.

Jin, Q. S., & Bethke, C. M. (2003). A new rate law describing microbial respiration. *Applied and Environmental Microbiology*, *69*(4), 2340–2348. https://doi.org/10.1128/Aem.69.4.2340-2348.2003

Kang, M. Z., Heuvelink, E., Carvalho, S. M. P., & de Reffye, P. (2012). A virtual plant that responds to the environment like a real one: The case for chrysanthemum. *New Phytologist*, *195*(2), 384–395. https://doi.org/10.1111/j.1469-8137.2012.04177.x

Kattge, J., & Sandel, B. (2020). TRY plant trait database - Enhanced coverage and open access [Dataset]. *Global Change Biology*, *26*, 119. https://doi.org/10.1111/gcb.15122

Kearney, M. R., Isaac, A. P., & Porter, W. P. (2014). microclim: Global estimates of hourly microclimate based on long-term monthly climate averages [Dataset]. *Scientific Data*, *1*, 140006. https://doi.org/10.1038/sdata.2014.6

Keenan, T. F., Baker, I., Barr, A., Ciais, P., Davis, K., Dietze, M., et al. (2012). Terrestrial biosphere model performance for inter-annual variability of land-atmosphere CO_2 exchange. *Global Change Biology*, *18*(6), 1971–1987. https://doi.org/10.1111/j.1365-2486.2012.02678.x

Keenan, T. F., Luo, X., Stocker, B. D., De Kauwe, M. G., Medlyn, B. E., Prentice, I. C., et al. (2023). A constraint on historic growth in global photosynthesis due to rising CO_2. *Nature Climate Change*, *13*(12), 1376–1381. https://doi.org/10.1038/s41558-023-01867-2

Kim, D. H., Lynett, P. J., & Socolofsky, S. A. (2009). A depth-integrated model for weakly dispersive, turbulent, and rotational fluid flows. *Ocean Modelling*, *27*(3–4), 198–214. https://doi.org/10.1016/j.ocemod.2009.01.005

Klausmeier, C. A., Litchman, E., & Levin, S. A. (2007). A model of flexible uptake of two essential resources. *Journal of Theoretical Biology*, *246*(2), 278–289. https://doi.org/10.1016/j.jtbi.2006.12.032

Kleber, M., & Adam, L. (2022). The science and semantics of "soil organic matter stabilization. In Y. Yang, M. Keiluweit, N. Senesi, & B. Xing (Eds.), *Multi-scale biogeochemical processes in soil ecosystems: Critical reactions and resilience to climate changes*. John Wiley & Sons, Inc https://doi.org/10.1002/9781119480419.ch2

Klotz, I. M., & Rosenberg, R. M. (2008). *Chemical thermodynamics: Basic concepts and methods* (7th ed.). Wiley-Interscience.

Kondo, J., & Saigusa, N. (1994). Modeling the evaporation from bare soil with a formula for vaporization in the soil pores. *Journal of the Meteorological Society of Japan*, *72*(3), 413–421. https://doi.org/10.2151/jmsj1965.72.3_413

Kooijman, S. A. L. M. (2009). *Dynamic energy budget theory for metabolic organisation*. Cambridge University Press. https://doi.org/10.1017/CBO9780511805400

Koonin, E. V., & Wolf, Y. I. (2012). Evolution of microbes and viruses: A paradigm shift in evolutionary biology? *Frontiers in Cellular and Infection Microbiology*, *2*. https://doi.org/10.3389/fcimb.2012.00119

Kotsuki, S., Terasaki, K., Yashiro, H., Tomita, H., Satoh, M., & Miyoshi, T. (2018). Online model parameter estimation with ensemble data assimilation in the real global atmosphere: A case with the nonhydrostatic icosahedral atmospheric model (NICAM) and the global satellite mapping of precipitation data. *Journal of Geophysical Research-Atmosphere*, *123*(14), 7375–7392. https://doi.org/10.1029/2017jd028092

Koven, C. D., Knox, R. G., Fisher, R. A., Chambers, J. Q., Christoffersen, B. O., Davies, S. J., et al. (2020). Benchmarking and parameter sensitivity of physiological and vegetation dynamics using the functionally assembled terrestrial ecosystem simulator (FATES) at Barro Colorado Island, Panama. *Biogeosciences*, *17*(11), 3017–3044. https://doi.org/10.5194/bg-17-3017-2020

Koven, C. D., Riley, W. J., Subin, Z. M., Tang, J. Y., Torn, M. S., Collins, W. D., et al. (2013). The effect of vertically resolved soil biogeochemistry and alternate soil C and N models on C dynamics of CLM4. *Biogeosciences*, *10*(11), 7109–7131. https://doi.org/10.5194/bg-10-7109-2013

Kudryavtsev, A. B., Jameson, R. F., & Linert, W. (2001). *The law of mass action*. (xiv, p. 328). Springer, .

Kurz, M., Offenhäuser, P., & Beck, A. (2023). Deep reinforcement learning for turbulence modeling in large eddy simulations. *International Journal of Heat and Fluid Flow*, *99*, 109094. https://doi.org/10.1016/j.ijheatfluidflow.2022.109094

Lam, K. N., Spanogiannopoulos, P., Soto-Perez, P., Alexander, M., Nalley, M. J., Bisanz, J. E., et al. (2021). Phage-delivered CRISPR-Cas9 for strain-specific depletion and genomic deletions in the gut microbiome. *Cell Reports*, *37*(5), 109930. https://doi.org/10.1016/j.celrep.2021.109930

Lehmann, J., Hansel, C. M., Kaiser, C., Kleber, M., Maher, K., Manzoni, S., et al. (2020). Persistence of soil organic carbon caused by functional complexity. *Nature Geoscience*, *13*(8), 529–534. https://doi.org/10.1038/s41561-020-0612-3

Leibold, M. A., Chase, J. M., & Ernest, S. K. M. (2017). Community assembly and the functioning of ecosystems: How metacommunity processes alter ecosystems attributes. *Ecology*, *98*(4), 909–919. https://doi.org/10.1002/ecy.1697

Le Noë, J., Manzoni, S., Abramoff, R., Bölscher, T., Bruni, E., Cardinael, R., et al. (2023). Soil organic carbon models need independent time-series validation for reliable prediction. *Communications Earth & Environment*, *4*(1), 158. https://doi.org/10.1038/s43247-023-00830-5

Lenton, T. M., Xu, C., Abrams, J. F., Ghadiali, A., Loriani, S., Sakschewski, B., et al. (2023). Quantifying the human cost of global warming. *Nature Sustainability*, *6*(10), 1237–1247. https://doi.org/10.1038/s41893-023-01132-6

Leon, J. A., & Tumpson, D. B. (1975). Competition between 2 species for 2 complementary or substitutable resources. *Journal of Theoretical Biology*, *50*(1), 185–201. https://doi.org/10.1016/0022-5193(75)90032-6

Leuning, R. (1990). Modeling stomatal behavior and photosynthesis of *Eucalyptus-grandis*. *Australian Journal of Plant Physiology*, *17*(2), 159–175. https://doi.org/10.1071/Pp9900159

Leuning, R. (1995). A critical-appraisal of a combined stomatal-photosynthesis model for C-3 plants. *Plant, Cell and Environment*, *18*(4), 339–355. https://doi.org/10.1111/j.1365-3040.1995.tb00370.x

Levin, S. A. (1992). The problem of pattern and scale in ecology. *Ecology*, *73*(6), 1943–1967. https://doi.org/10.2307/1941447

Li, C. S., Aber, J., Stange, F., Butterbach-Bahl, K., & Papen, H. (2000). A process-oriented model of N_2O and NO emissions from forest soils: 1. Model development. *Journal of Geophysical Research*, *105*(D4), 4369–4384. https://doi.org/10.1029/1999jd900949

Liang, L. L., Arcus, V. L., Heskel, M. A., O'Sullivan, O. S., Weerasinghe, L. K., Creek, D., et al. (2018). Macromolecular rate theory (MMRT) provides a thermodynamics rationale to underpin the convergent temperature response in plant leaf respiration. *Global Change Biology*, *24*(4), 1538–1547. https://doi.org/10.1111/gcb.13936

Liebig, J. F. v. (1840). *Die organische Chemie in ihrer Anwendung auf Agricultur und Physiologie (Organic chemistry in its applications to agriculture and physiology)*. Friedrich Vieweg und Sohn Publ. Co.

Lieth, H. (1973). Primary production: Terrestrial ecosystems. *Human Ecology*, *1*(4), 303–332. https://doi.org/10.1007/BF01536729

Lin, S. J., & Rood, R. B. (1996). Multidimensional flux-form semi-Lagrangian transport schemes. *Monthly Weather Review*, *124*(9), 2046–2070. https://doi.org/10.1175/1520-0493(1996)124<2046:Mffslt>2.0.Co;2

Liu, L. C., Xu, S., Tang, J., Guan, K., Griffis, T. J., Erickson, M. D., et al. (2022). KGML-Ag: A modeling framework of knowledge-guided machine learning to simulate agroecosystems: A case study of estimating N2O emission using data from mesocosm experiments. *Geoscientific Model Development*, *15*(7), 2839–2858. https://doi.org/10.5194/gmd-15-2839-2022

Liu, L. C., Zhou, W., Guan, K., Peng, B., Xu, S., Tang, J., et al. (2024). Knowledge-guided machine learning can improve carbon cycle quantification in agroecosystems. *Nature Communications*, *15*(1), 357. https://doi.org/10.1038/s41467-023-43860-5

Liu, Y. (2007). Overview of some theoretical approaches for derivation of the Monod equation. *Applied Microbiology and Biotechnology*, *73*(6), 1241–1250. https://doi.org/10.1007/s00253-006-0717-7

Liu, Z. G., Consalvi, J. L., & Kong, W. J. (2019). An exponential integrator with Schur-Krylov approximation to accelerate combustion chemistry computation. *Combustion and Flame*, *203*, 180–189. https://doi.org/10.1016/j.combustflame.2019.01.031

Luo, Z., Wang, E., Zheng, H., Baldock, J. A., Sun, O. J., & Shao, Q. (2015). Convergent modelling of past soil organic carbon stocks but divergent projections. *Biogeosciences*, *12*(14), 4373–4383. https://doi.org/10.5194/bg-12-4373-2015

Ma, L., Hurtt, G., Ott, L., Sahajpal, R., Fisk, J., Lamb, R., et al. (2022). Global evaluation of the Ecosystem Demography model (ED v3.0). *Geoscientific Model Development*, *15*(5), 1971–1994. https://doi.org/10.5194/gmd-15-1971-2022

Madigan, M. T., Martinko, J. M., Dunlap, P. V., & Clark, D. P. (2009). *Brock biology of microorganisms*, (12th ed., xxviii, p. 1061). Pearson/Benjamin Cummings.

Maggi, F., Gu, C., Riley, W. J., Hornberger, G. M., Venterea, R. T., Xu, T., et al. (2008). A mechanistic treatment of the dominant soil nitrogen cycling processes: Model development, testing, and application. *Journal of Geophysical Research-Biogeosciences*, *113*(G2). https://doi.org/10.1029/2007jg000578

Martiny, J. B. H., Martiny, A. C., Brodie, E., Chase, A. B., Rodriguez-Verdugo, A., Treseder, K. K., & Allison, S. D. (2023). Investigating the eco-evolutionary response of microbiomes to environmental change. *Ecology Letters*, *26*(S1). https://doi.org/10.1111/ele.14209

Medlyn, B. E., Duursma, R. A., Eamus, D., Ellsworth, D. S., Prentice, I. C., Barton, C. V. M., et al. (2011). Reconciling the optimal and empirical approaches to modelling stomatal conductance. *Global Change Biology*, *17*(6), 2134–2144. https://doi.org/10.1111/j.1365-2486.2010.02375.x

Medvigy, D., Wang, G. S., Zhu, Q., Riley, W. J., Trierweiler, A. M., Waring, B. G., et al. (2019). Observed variation in soil properties can drive large variation in modelled forest functioning and composition during tropical forest secondary succession. *New Phytologist*, *223*(4), 1820–1833. https://doi.org/10.1111/nph.15848

Mehta, P., & Schwab, D. J. (2014). An exact mapping between the variational renormalization group and deep learning, arXiv:1410.3831.

Meidner, H., & Mansfield, T. A. (1968). *Physiology of stomata*, (xiii, p. 178). McGraw-Hill, .

Michaelis, L., & Menten, M. L. (1913). The kinetics of the inversion effect. *Biochemische Zeitschrift*, *49*, 333–369.

Miller, T. D. (1992). Growth stages of wheat: Identification and understanding improve crop management. *Better Crops*, *76*, 12–17.

Milly, P. C. D. (1982). Moisture and heat-transport in hysteretic, inhomogeneous porous-media - A matric head-based formulation and a numerical-model. *Water Resources Research*, *18*(3), 489–498. https://doi.org/10.1029/WR018i003p00489

Moldrup, P., Olesen, T., Komatsu, T., Yoshikawa, S., Schjonning, P., & Rolston, D. E. (2003). Modeling diffusion and reaction in soils: X. A unifying model for solute and gas diffusivity in unsaturated soil. *Soil Science*, *168*(5), 321–337. https://doi.org/10.1097/00010694-200305000-00002

Molins, S., & Knabner, P. (2019). Multiscale approaches in reactive transport modeling. *Reviews in Mineralogy and Geochemistry*, *85*(1), 27–48. https://doi.org/10.2138/rmg.2019.85.2

Muraki, D. J., Snyder, C., & Rotunno, R. (1999). The next-order corrections to quasigeostrophic theory. *Journal of the Atmospheric Sciences*, *56*(11), 1547–1560. https://doi.org/10.1175/1520-0469(1999)056<1547:Tnoctq>2.0.Co;2

Murdoch, W. W. (1973). Functional response of predators. *Journal of Applied Ecology*, *10*(1), 335–342.

Murphy, K. P., Privalov, P. L., & Gill, S. J. (1990). Common features of protein unfolding and dissolution of hydrophobic compounds. *Science*, *247*(4942), 559–561. https://doi.org/10.1126/science.2300815

Noe, S. M., & Giersch, C. (2004). A simple dynamic model of photosynthesis in oak leaves:: Coupling leaf conductance and photosynthetic carbon fixation by a variable intracellular CO_2 pool. *Functional Plant Biology*, *31*(12), 1195–1204. https://doi.org/10.1071/Fp03251

Novick, K. A., Biederman, J. A., Desai, A. R., Litvak, M. E., Moore, D. J. P., Scott, R. L., & Torn, M. S. (2018). The AmeriFlux network: A coalition of the willing. *Agricultural and Forest Meteorology*, *249*, 444–456. https://doi.org/10.1016/j.agrformet.2017.10.009

Obersteiner, M., Böttcher, H., & Yamagata, Y. (2010). Terrestrial ecosystem management for climate change mitigation. *Current Opinion in Environmental Sustainability*, *2*(4), 271–276. https://doi.org/10.1016/j.cosust.2010.05.006

Pan, D. (1995). A tutorial on mpeg audio compression. *Ieee Multimedia*, *2*(2), 60–74. https://doi.org/10.1109/93.388209

Pan, M., & Wood, E. F. (2006). Data assimilation for estimating the terrestrial water budget using a constrained ensemble Kalman filter. *Journal of Hydrometeorology*, *7*(3), 534–547. https://doi.org/10.1175/Jhm495.1

Parks, H. V., & Faller, J. E. (2010). Simple pendulum determination of the gravitational constant. *Physical Review Letters*, *105*(11), 110801. https://doi.org/10.1103/PhysRevLett.105.110801

Pasut, C., Tang, F. H. M., & Maggi, F. (2020). A mechanistic analysis of wetland biogeochemistry in response to temperature, vegetation, and nutrient Input changes. *J Geophys Res-Biogeo*, *125*(4), e2019JG005437. https://doi.org/10.1029/2019JG005437

Pasut, C., Tang, F. H. M., Minasny, B., Warren, C. R., Dijkstra, F. A., Riley, W. J., & Maggi, F. (2023). Seasonal biotic processes vary the carbon turnover by up to one rrder of magnitude in wetlands. *Global Biogeochemical Cycles*, *37*(5), e2022GB007679. https://doi.org/10.1029/2022gb007679

Pauli, W. (1973). *Statistical mechanics*. (x, p. 121). MIT Press, .,.

Peterson, M. E., Eisenthal, R., Danson, M. J., Spence, A., & Daniel, R. M. (2004). A new intrinsic thermal parameter for enzymes reveals true temperature optima. *Journal of Biological Chemistry*, *279*(20), 20717–20722. https://doi.org/10.1074/jbc.M309143200

Phillips, R., Kondev, J., Theriot, J., & Garcia, H. (2012). *Physical biology of the cell* (2nd ed., p. 1088). Garland Science.

Pirt, S. J. (1982). Maintenance energy - A general-model for energy-limited and energy-sufficient growth. *Archives of Microbiology*, *133*(4), 300–302. https://doi.org/10.1007/Bf00521294

Plawsky, J. L. (2019). Transport phenomena fundamentals. In *Chemical industries, edited, p. 1 online resource*. CRC Press.

Powell, N., Shilton, A. N., Pratt, S., & Chisti, Y. (2008). Factors influencing luxury uptake of phosphorus by microalgae in waste stabilization ponds. *Environmental Science & Technology*, *42*(16), 5958–5962. https://doi.org/10.1021/es703118s

Poyser, M., Atapour-Abarghouei, A., & Breckon, T. P. (2021). On the impact of lossy image and video compression on the performance of deep convolutional neural network architectures. *International C Pattern Recognition*, 2830–2837. https://doi.org/10.1109/Icpr48806.2021.9412455

Pryor, R. W. (2012). Multiphysics modeling using COMSOL 4: A first principles approach [Software]. *Mercury Learning and Information*. xviii, 553. Retrieved from http://dl.poweren.ir/downloads/PowerEn/Book/2019/Mar/28%20%DA%A9%D8%A7%D9%85%D9%84%20%DA%86%D8%A7%D9%85%D8%B9%20%D8%A2%D9%85%D9%88%D8%B2%D8%B4%20%DA%A9%D8%AA%D8%A7%D8%A8%20PowerEn.ir%29.pdf

Purcell, E. M. (1977). Life at low Reynolds-number. *American Journal of Physics*, *45*(1), 3–11. https://doi.org/10.1119/1.10903

Qiu, H., Hao, D. L., Zeng, Y. L., Zhang, X. S., & Chen, M. (2023). Global and northern-high-latitude net ecosystem production in the 21st century from CMIP6 experiments. *Earth System Dynamics*, *14*(1), 1–16. https://doi.org/10.5194/esd-14-1-2023

Ratkowsky, D. A., Olley, J., & Ross, T. (2005). Unifying temperature effects on the growth rate of bacteria and the stability of globular proteins. *Journal of Theoretical Biology*, *233*(3), 351–362. https://doi.org/10.1016/j.jtbi.2004.10.016

Raynaud, X., & Nunan, N. (2014). Spatial ecology of bacteria at the microscale in soil. *PLoS One*, *9*(1), e87217. https://doi.org/10.1371/journal.pone.0087217

Real, L. A. (1977). The kinetics of functional response. *The American Naturalist*, *111*(978), 289–300. https://doi.org/10.1086/283161

Reinhardt, C., Volker, B., Martin, H. J., Kneiseler, J., & Fuhrmann, G. F. (1997). Different activation energies in glucose uptake in Saccharomyces cerevisiae DFY1 suggest two transport systems. *Bba-Biomembranes*, *1325*(1), 126–134. https://doi.org/10.1016/S0005-2736(96)00252-0

Riley, W. J., Maggi, F., Kleber, M., Torn, M. S., Tang, J. Y., Dwivedi, D., & Guerry, N. (2014). Long residence times of rapidly decomposable soil organic matter: Application of a multi-phase, multi-component, and vertically resolved model (BAMS1) to soil carbon dynamics. *Geoscientific Model Development*, *7*(4), 1335–1355. https://doi.org/10.5194/gmd-7-1335-2014

Riley, W. J., Sierra, C., Tang, J. Y., Bouskill, N. J., Zhu, Q., & Abramoff, R. (2022). Next generation soil biogeochemistry model representations: A proposed community open source model farm (BeTR-S). In Y. Yang, M. Keiluweit, N. Senesi, & B. Xing (Eds.), *Multi-scale biogeochemical processes in soil ecosystems: Critical reactions and resilience to climate changes*. John Wiley & Sons, Inc. https://doi.org/10.1002/9781119480419.ch11

Riley, W. J., Subin, Z. M., Lawrence, D. M., Swenson, S. C., Torn, M. S., Meng, L., et al. (2011). Barriers to predicting changes in global terrestrial methane fluxes: Analyses using CLM4Me, a methane biogeochemistry model integrated in CESM. *Biogeosciences*, *8*(7), 1925–1953. https://doi.org/10.5194/bg-8-1925-2011

Roberts, A. J. (2015). *Model emergent dynamics in complex systems*, (xii, p. 748). Society for Industrial and Applied Mathematics, .

Robinne, F. N., Hallema, D. W., Bladon, K. D., & Buttle, J. M. (2020). Wildfire impacts on hydrologic ecosystem services in North American high-latitude forests: A scoping review. *Journal of Hydrology*, *581*, 124360. https://doi.org/10.1016/j.jhydrol.2019.124360

Rollins, G. C., & Dill, K. A. (2014). General mechanism of two-state protein folding kinetics. *Journal of the American Chemical Society*, *136*(32), 11420–11427. https://doi.org/10.1021/ja5049434

Ross, I. U. (1981). *The radiation regime and architecture of plant stands*, (Vol. xxvii, p. 391). Springer. https://doi.org/10.1007/978-94-009-8647-3

Rubio, G., Zhu, J. M., & Lynch, J. P. (2003). A critical test of the two prevailing theories of plant response to nutrient availability. *American Journal of Botany*, *90*(1), 143–152. https://doi.org/10.3732/ajb.90.1.143

Russell, S. J., & Norvig, P. (2010). *Artificial intelligence, A modern approach* (3rd ed.). Pearson.
Russo, S. E., Ledder, G., Muller, E. B., & Nisbet, R. M. (2022). Dynamic energy budget models: Fertile ground for understanding resource allocation in plants in a changing world. *Conserv Physiol*, *10*(1). ARTN coac061. https://doi.org/10.1093/conphys/coac061
Rusu, T., Pacurar, I., Dirja, M., Pacurar, H. M., Oroian, I., Cosma, S. A., & Gheres, M. (2013). Effect of tillage systems on soil properties, humus and water conservation. *Agricultural Sciences in China*, *4*(5B), 35–40. https://doi.org/10.4236/as.2013.45B007
Saito, H., Simunek, J., & Mohanty, B. P. (2006). Numerical analysis of coupled water, vapor, and heat transport in the vadose zone. *Vadose Zone Journal*, *5*(2), 784–800. https://doi.org/10.2136/vzj2006.0007
Sander, R. (2023). Compilation of Henry's law constants (version 5.0.0) for water as solvent. *Atmospheric Chemistry and Physics*, *23*(19), 10901–12440. https://doi.org/10.5194/acp-23-10901-2023
Santos, A., Scanavini, H. F. A., Pedrini, H., Schiozer, D. J., Munerato, F. P., & Barreto, C. E. A. G. (2022). An artificial intelligence method for improving upscaling in complex reservoirs. *Journal of Petroleum Science and Engineering*, *211*, 110071. https://doi.org/10.1016/j.petrol.2021.110071
Saredi, E., Ramesh, N. T., Sciacchitano, A., & Scarano, F. (2021). State observer data assimilation for RANS with time-averaged 3D-PIV data. *Computers & Fluids*, *218*, 104827. https://doi.org/10.1016/j.compfluid.2020.104827
Saurabh, S. (2021). Genome editing: Revolutionizing the crop improvement. *Plant Molecular Biology Reporter*, *39*(4), 752–772. https://doi.org/10.1007/s11105-021-01286-7
Sawle, L., & Ghosh, K. (2011). How do thermophilic proteins and proteomes withstand high temperature? *Biophysical Journal*, *101*(1), 217–227. https://doi.org/10.1016/j.bpj.2011.05.059
Seok, M. G., Cai, W. T., & Park, D. (2021). Hierarchical aggregation/disaggregation for adaptive abstraction-level conversion in digital twin-based smart semiconductor manufacturing. *IEEE Access*, *9*, 71145–71158. https://doi.org/10.1109/Access.2021.3073618
Shankar, R. (1994). *Principles of quantum mechanics* (2nd ed. ed.). Plenum Press.
Sharpe, P. J. H., & Demichele, D. W. (1977). Reaction-kinetics of poikilotherm development. *Journal of Theoretical Biology*, *64*(4), 649–670. https://doi.org/10.1016/0022-5193(77)90265-X
Shinozaki, K. Y. K., Hozumi, K., & Kira, T. (1964). A quantitative analysis of plant form-the pipe model theory: I. Basic analyses. *Jap. J. Ecol.*, *14*, 97–105. https://doi.org/10.18960/seitai.14.3_97
Shivni, R. (2016). The deconstructed Standard Model equation.
Shuttleworth, W. J., & Wallace, J. S. (1985). Evaporation from sparse crops - An energy combination theory. *The Quarterly Journal of the Royal Meteorological Society*, *111*(469), 839–855. https://doi.org/10.1256/smsqj.46909
Sierra, C. A., Trumbore, S. E., Davidson, E. A., Vicca, S., & Janssens, I. (2015). Sensitivity of decomposition rates of soil organic matter with respect to simultaneous changes in temperature and moisture. *Journal of Advances in Modeling Earth Systems*, *7*(1), 335–356. https://doi.org/10.1002/2014ms000358
Silverstein, T. P. (2020). The hydrophobic effect: Is water afraid, or just not that interested? *Chemtexts*, *6*(4), 26. https://doi.org/10.1007/s40828-020-00117-8
Simunek, J., & Suarez, D. L. (1993). Modeling of carbon-dioxide transport and production in soil .1. model development. *Water Resources Research*, *29*(2), 487–497. https://doi.org/10.1029/92wr02225
Sollins, P., & Gregg, J. W. (2017). Soil organic matter accumulation in relation to changing soil volume, mass, and structure: Concepts and calculations. *Geoderma*, *301*, 60–71. https://doi.org/10.1016/j.geoderma.2017.04.013
Sprengel, C. (1826). Ueber Pflanzenhumus, Humussaure und humussaure Salze (About palnt humus, humic acids and salts of humic acids). *Archiv fur die Gesammte Naturlehre*, *8*, 145–220.
Steefel, C. I., DePaolo, D. J., & Lichtner, P. C. (2005). Reactive transport modeling: An essential tool and a new research approach for the Earth sciences. *Earth and Planetary Science Letters*, *240*(3–4), 539–558. https://doi.org/10.1016/j.epsl.2005.09.017
Steefel, C. I., & Lasaga, A. C. (1994). A coupled model for transport of multiple chemical-species and kinetic precipitation dissolution reactions with application to reactive flow in single-phase hydrothermal systems. *American Journal of Science*, *294*(5), 529–592. https://doi.org/10.2475/ajs.294.5.529
Steefel, C. I., Yabusaki, S. B., & Mayer, K. U. (2015). Reactive transport benchmarks for subsurface environmental simulation. *Computational Geosciences*, *19*(3), 439–443. https://doi.org/10.1007/s10596-015-9499-2
Stein, W. D. (2012). *Transport and diffusion across cell membranes* (p. 704). Academic Press.
Steudle, E., Zimmermann, U., & Luttge, U. (1977). Effect of turgor pressure and cell-size on wall elasticity of plant-cells. *Plant Physiology*, *59*(2), 285–289. https://doi.org/10.1104/pp.59.2.285
Stumm, W., & Morgan, J. J. (1996). *Aquatic chemistry: Chemical equilibria and rates in natural waters* (3rd ed.). Wiley-Interscience.
Suseela, V., Conant, R. T., Wallenstein, M. D., & Dukes, J. S. (2012). Effects of soil moisture on the temperature sensitivity of heterotrophic respiration vary seasonally in an old-field climate change experiment. *Global Change Biology*, *18*(1), 336–348. https://doi.org/10.1111/j.1365-2486.2011.02516.x
Sweetlove, L. J., & Ratcliffe, R. G. (2011). Flux-balance modeling of plant metabolism. *Frontiers in Plant Science*, *2*. https://doi.org/10.3389/fpls.2011.00038
Taiz, L., & Zeiger, E. (2006). *Plant physiology* (4th ed.). Sinauer Associates, Inc.
Tan, Y. S., Zhang, R. K., Liu, Z. H., Li, B. Z., & Yuan, Y. J. (2022). Microbial adaptation to enhance stress tolerance. *Frontiers in Microbiology*, *13*, 888746. https://doi.org/10.3389/fmicb.2022.888746
Tang, J., Zhuang, Q., Shannon, R. D., & White, J. R. (2010). Quantifying wetland methane emissions with process-based models of different complexities. *Biogeosciences*, *7*(11), 3817–3837. https://doi.org/10.5194/bg-7-3817-2010
Tang, J. Y. (2015). On the relationships between the Michaelis-Menten kinetics, reverse Michaelis-Menten kinetics, equilibrium chemistry approximation kinetics, and quadratic kinetics. *Geoscientific Model Development*, *8*(12), 3823–3835. https://doi.org/10.5194/gmd-8-3823-2015
Tang, J. Y., & Riley, W. J. (2013). A total quasi-steady-state formulation of substrate uptake kinetics in complex networks and an example application to microbial litter decomposition. *Biogeosciences*, *10*(12), 8329–8351. https://doi.org/10.5194/bg-10-8329-2013
Tang, J. Y., & Riley, W. J. (2015). Weaker soil carbon-climate feedbacks resulting from microbial and abiotic interactions. *Nature Climate Change*, *5*(1), 56–60. https://doi.org/10.1038/Nclimate2438
Tang, J. Y., & Riley, W. J. (2017). SUPECA kinetics for scaling redox reactions in networks of mixed substrates and consumers and an example application to aerobic soil respiration. *Geoscientific Model Development*, *10*(9), 3277–3295. https://doi.org/10.5194/gmd-10-3277-2017
Tang, J. Y., & Riley, W. J. (2018). Predicted land carbon dynamics are strongly dependent on the numerical coupling of nitrogen mobilizing and immobilizing processes: A demonstration with the E3SM land model. *Earth Interactions*, *22*(11), 1–18. https://doi.org/10.1175/EI-D-17-0023.s1

Tang, J. Y., & Riley, W. J. (2019a). Competitor and substrate sizes and diffusion together define enzymatic depolymerization and microbial substrate uptake rates. *Soil Biology and Biochemistry*, *139*, 107624. https://doi.org/10.1016/j.soilbio.2019.107624
Tang, J. Y., & Riley, W. J. (2019b). A theory of effective microbial substrate affinity parameters in variably saturated soils and an example application to aerobic soil heterotrophic respiration. *J Geophys Res-Biogeo*, *124*(4), 918–940. https://doi.org/10.1029/2018jg004779
Tang, J. Y., & Riley, W. J. (2021). Finding Liebig's law of the minimum. *Ecological Applications*, *31*(8), e02458. https://doi.org/10.1002/eap.2458
Tang, J. Y., & Riley, W. J. (2023). Revising the dynamic energy budget theory with a new reserve mobilization rule and three example applications to bacterial growth. *Soil Biology and Biochemistry*, *178*, 108954. https://doi.org/10.1016/j.soilbio.2023.108954
Tang, J. Y., & Riley, W. J. (2024). A chemical kinetics theory for interpreting the non-monotonic temperature dependence of enzymatic reactions. *Biogeosciences*, *21*(5), 1061–1070. https://doi.org/10.5194/bg-21-1061-2024
Tang, J. Y., Riley, W. J., Marschmann, G. L., & Brodie, E. L. (2021). Conceptualizing biogeochemical reactions with an Ohm's law analogy. *Journal of Advances in Modeling Earth Systems*, *13*(10), e2021MS002469. https://doi.org/10.1029/2021MS002469
Tang, J. Y., Riley, W. J., & Niu, J. (2015). Incorporating root hydraulic redistribution in CLM4.5: Effects on predicted site and global evapotranspiration, soil moisture, and water storage. *Journal of Advances in Modeling Earth Systems*, *7*(4), 1828–1848. https://doi.org/10.1002/2015ms000484
Tang, J. Y., Riley, W. J., & Zhu, Q. (2022). Supporting hierarchical soil biogeochemical modeling: Version 2 of the biogeochemical transport and reaction model (BeTR-v2). *Geoscientific Model Development*, *15*(4), 1619–1632. https://doi.org/10.5194/gmd-15-1619-2022
Tang, J. Y., & Zhuang, Q. L. (2008). Equifinality in parameterization of process-based biogeochemistry models: A significant uncertainty source to the estimation of regional carbon dynamics. *Journal of Geophysical Research-Biogeoscience*, *113*(G4), G04010. https://doi.org/10.1029/2008jg000757
Tang, J. Y., & Zhuang, Q. L. (2009). A global sensitivity analysis and Bayesian inference framework for improving the parameter estimation and prediction of a process-based Terrestrial Ecosystem Model. *Journal of Geophysical Research*, *114*(D15), D15303. https://doi.org/10.1029/2009jd011724
Tao, F., Huang, Y., Hungate, B. A., Manzoni, S., Frey, S. D., Schmidt, M. W. I., et al. (2023). Microbial carbon use efficiency promotes global soil carbon storage. *Nature*, *618*(7967), 981–985. https://doi.org/10.1038/s41586-023-06042-3
Tarantola, A. (2005). *Inverse problem theory and methods for model parameter estimation*, (Vol. xii, p. 342). Society for Industrial and Applied Mathematics, .
Thakkar, A. J. (2021). *Quantum Chemistry (A concise introduction for students of physics, chemistry, biochemistry and materials science)*, 3rd ed., IOP Publishing, ., https://doi.org/10.1088/978-0-7503-3827-1
Theodoridis, S. (2015). *Machine learning: A Bayesian and optimization perspective* (1st ed.). Academic Press.
Thornley, J. H. M. (1972). A balanced quantitative model for root: Shoot ratios in vegetative plants. *Annals of Botany*, *36*(2), 431–441. https://doi.org/10.1093/oxfordjournals.aob.a084602
Tian, S. Y., Renzullo, L. J., Pipunic, R. C., Lerat, J., Sharples, W., & Donnelly, C. (2021). Satellite soil moisture data assimilation for improved operational continental water balance prediction. *Hydrology and Earth System Sciences*, *25*(8), 4567–4584. https://doi.org/10.5194/hess-25-4567-2021
Tietjen, T., & Wetzel, R. G. (2003). Extracellular enzyme-clay mineral complexes: Enzyme adsorption, alteration of enzyme activity, and protection from photodegradation. *Aquatic Ecology*, *37*(4), 331–339. https://doi.org/10.1023/B:AECO.0000007044.52801.6b
Todd-Brown, K. E. O., Randerson, J. T., Post, W. M., Hoffman, F. M., Tarnocai, C., Schuur, E. A. G., & Allison, S. D. (2013). Causes of variation in soil carbon simulations from CMIP5 Earth system models and comparison with observations. *Biogeosciences*, *10*(3), 1717–1736. https://doi.org/10.5194/bg-10-1717-2013
Tolla, C., Kooijman, S. A. L. M., & Poggiale, J. C. (2007). A kinetic inhibition mechanism for maintenance. *Journal of Theoretical Biology*, *244*(4), 576–587. https://doi.org/10.1016/j.jtbi.2006.09.001
Tollefson, J. (2022). Carbon emissions hit new high: Warning from COP27. *Nature News*. https://doi.org/10.1038/d41586-022-03657-w
Tsai, W. P., Feng, D. P., Pan, M., Beck, H., Lawson, K., Yang, Y., et al. (2021). From calibration to parameter learning: Harnessing the scaling effects of big data in geoscientific modeling. *Nature Communications*, *12*(1), 5988. https://doi.org/10.1038/s41467-021-26107-z
van de Griend, A. A., & Owe, M. (1994). Bare soil surface-resistance to evaporation by vapor diffusion under semiarid conditions. *Water Resources Research*, *30*(2), 181–188. https://doi.org/10.1029/93wr02747
van Genuchten, M. T. (1980). A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Science Society of America Journal*, *44*(5), 892–898. https://doi.org/10.2136/sssaj1980.03615995004400050002x
Varney, R. M., Chadburn, S. E., Burke, E. J., & Cox, P. M. (2022). Evaluation of soil carbon simulation in CMIP6 Earth system models. *Biogeosciences*, *19*(19), 4671–4704. https://doi.org/10.5194/bg-19-4671-2022
Viskari, T., Pusa, J., Fer, I., Repo, A., Vira, J., & Liski, J. (2022). Calibrating the soil organic carbon model Yasso20 with multiple datasets. *Geoscientific Model Development*, *15*(4), 1735–1752. https://doi.org/10.5194/gmd-15-1735-2022
von Caemmerer, S. (2013). Steady-state models of photosynthesis. *Plant, Cell and Environment*, *36*(9), 1617–1630. https://doi.org/10.1111/pce.12098
von Smoluchowski, M. (1917). Versuch einer mathematischen theorie der koagulationkinetik kolloider loesungen. *Zeitschrift für Physikalische Chemie*(92), 129–132. https://doi.org/10.1515/zpch-1918-9209
Vrugt, J. A. (2016). Markov chain Monte Carlo simulation using the DREAM software package: Theory, concepts, and MATLAB implementation. *Environmental Modelling & Software*, *75*, 273–316. https://doi.org/10.1016/j.envsoft.2015.08.013
Walker, A. P., Johnson, A. L., Rogers, A., Anderson, J., Bridges, R. A., Fisher, R. A., et al. (2021). Multi-hypothesis comparison of Farquhar and Collatz photosynthesis models reveals the unexpected influence of empirical assumptions at leaf and global scales. *Global Change Biology*, *27*(4), 804–822. https://doi.org/10.1111/gcb.15366
Wang, G. S., Gao, Q., Yang, Y. F., Hobbie, S. E., Reich, P. B., & Zhou, J. Z. (2022). Soil enzymes as indicators of soil function: A step toward greater realism in microbial ecological modeling. *Global Change Biology*, *28*(5), 1935–1950. https://doi.org/10.1111/gcb.16036
Wang, G. S., & Post, W. M. (2012). A theoretical reassessment of microbial maintenance and implications for microbial ecology modeling. *FEMS Microbiology Ecology*, *81*(3), 610–617. https://doi.org/10.1111/j.1574-6941.2012.01389.x
Wang, S. S., Yang, Y., Trishchenko, A. P., Barr, A. G., Black, T. A., & McCaughey, H. (2009). Modeling the response of canopy stomatal conductance to humidity. *Journal of Hydrometeorology*, *10*(2), 521–532. https://doi.org/10.1175/2008jhm1050.1
Wang, Y. P., Chen, B. C., Wieder, W. R., Leite, M., Medlyn, B. E., Rasmussen, M., et al. (2014). Oscillatory behavior of two nonlinear microbial models of soil carbon decomposition. *Biogeosciences*, *11*(7), 1817–1831. https://doi.org/10.5194/bg-11-1817-2014
Wang, Y. P., Law, R. M., & Pak, B. (2010). A global model of carbon, nitrogen and phosphorus cycles for the terrestrial biosphere. *Biogeosciences*, *7*(7), 2261–2282. https://doi.org/10.5194/bg-7-2261-2010

Warren, C. R., & Manzoni, S. (2023). When dry soil is re-wet, trehalose is respired instead of supporting microbial growth. *Soil Biology and Biochemistry*, *184*, 109121. https://doi.org/10.1016/j.soilbio.2023.109121

Wendering, P., & Nikoloski, Z. (2023). Toward mechanistic modeling and rational engineering of plant respiration. *Plant Physiology*, *191*(4), 2150–2166. https://doi.org/10.1093/plphys/kiad054

Weng, E. S., Aleinov, I., Singh, R., Puma, M. J., McDermid, S. S., Kiang, N. Y., et al. (2022). Modeling demographic-driven vegetation dynamics and ecosystem biogeochemical cycling in NASA GISS's Earth system model (ModelE-BiomeE v.1.0). *Geoscientific Model Development*, *15*(22), 8153–8180. https://doi.org/10.5194/gmd-15-8153-2022

Wermuth, N. (1976). Analogies between multiplicative models in contingency-tables and covariance selection. *Biometrics*, *32*(1), 95–108. https://doi.org/10.2307/2529341

Wieder, W. R., Grandy, A. S., Kallenbach, C. M., & Bonan, G. B. (2014). Integrating microbial physiology and physio-chemical principles in soils with the MIcrobial-MIneral Carbon Stabilization (MIMICS) model. *Biogeosciences*, *11*(14), 3899–3917. https://doi.org/10.5194/bg-11-3899-2014

Wigner, E. P. (1960). The unreasonable effectiveness of mathematics in the natural sciences. *Communications on Pure and Applied Mathematics*, *13*(1), 1–14. https://doi.org/10.1002/cpa.3160130102

Williams, P. J. L. (1973). The validity of the application of simple kinetic analysis to heterogeneous microbial populations. *Limnology & Oceanography*, *18*(1), 159–165. https://doi.org/10.4319/lo.1973.18.1.0159

Wilson, C. H., & Gerber, S. (2021). Theoretical insights from upscaling Michaelis-Menten microbial dynamics in biogeochemical models: A dimensionless approach. *Biogeosciences*, *18*(20), 5669–5679. https://doi.org/10.5194/bg-18-5669-2021

Wöber, W., Novotny, G., Mehnen, L., & Olaverri-Monreal, C. (2020). Autonomous vehicles: Vehicle parameter estimation using variational Bayes and kinematics. *Appl Sci-Basel*, *10*(18), 6317. https://doi.org/10.3390/app10186317

Wu, M. S., Zhao, Q., Jansson, P. E., Wu, J., Tan, X., Duan, Z., et al. (2021). Improved soil hydrological modeling with the implementation of salt-induced freezing point depression in CoupModel: Model calibration and validation. *Journal of Hydrology*, *596*, 125693. https://doi.org/10.1016/j.jhydrol.2020.125693

Xiao, Q. N., Kuo, Y. H., Sun, J. Z., Lee, W. C., Barker, D. M., & Lim, E. (2007). An approach of radar reflectivity data assimilation and its assessment with the inland QPF of Typhoon Rusa (2002) at landfall. *Journal of Applied Meteorology and Climatology*, *46*(1), 14–22. https://doi.org/10.1175/Jam2439.1

Xiao, X., Zhang, J., Yan, H., Wu, W., & Chandrashekhar, B. (2009). Land surface phenology: Convergence of satellite and CO_2 eddy flux observations. In A. Noormets (Ed.), *Phenology of ecosystem processes*. Springer Science+Business Media, LLC. https://doi.org/10.1007/978-1-4419-0026-5_11

Yu, L. Y., Cai, H. J., Zheng, Z., Li, Z. J., & Wang, J. (2017). Towards a more flexible representation of water stress effects in the nonlinear Jarvis model. *Journal of Integrative Agriculture*, *16*(1), 210–220. https://doi.org/10.1016/S2095-3119(15)61307-7

Zaehle, S., Medlyn, B. E., De Kauwe, M. G., Walker, A. P., Dietze, M. C., Hickler, T., et al. (2014). Evaluation of 11 terrestrial carbon-nitrogen cycle models against observations from two temperate Free-Air CO_2 Enrichment studies. *New Phytologist*, *202*(3), 803–822. https://doi.org/10.1111/nph.12697

Zhou, G., Wong, M. T., & Zhou, G. Q. (1983). Diffusion-controlled reactions of enzymes an approximate analytic solution of Chous Model. *Biophysical Chemistry*, *18*(2), 125–132. https://doi.org/10.1016/0301-4622(83)85006-6

Zhou, L. J., Harris, L., Chen, J. H., Gao, K., Guo, H., Xiang, B. Q., et al. (2022). Improving global weather prediction in GFDL SHiELD through an upgraded GFDL cloud microphysics scheme. *Journal of Advances in Modeling Earth Systems*, *14*(7), e2021MS002971. https://doi.org/10.1029/2021MS002971

Zhou, W., Guan, K. Y., Peng, B., Tang, J. Y., Ji, Z. N., Jiang, C. Y., et al. (2021). Quantifying carbon budget, crop yields and their responses to environmental variability using the model for US Midwestern agroecosystems. *Agricultural and Forest Meteorology*, *307*, 108521. https://doi.org/10.1016/j.agrformet.2021.108521

Zhou, W. P., Hui, D. F., & Shen, W. J. (2014). Effects of soil moisture on the temperature sensitivity of soil heterotrophic respiration: A laboratory incubation study. *PLoS One*, *9*(3), e92531. https://doi.org/10.1371/journal.pone.0092531

Zhou, Y. (2010). Renormalization group theory for fluid and plasma turbulence. *Physics Reports*, *488*(1), 1–49. https://doi.org/10.1016/j.physrep.2009.04.004

Zhu, Q., Riley, W. J., Tang, J., & Koven, C. D. (2016). Multiple soil nutrient competition between plants, microbes, and mineral surfaces: Model development, parameterization, and example applications in several tropical forests. *Biogeosciences*, *13*(1), 341–363. https://doi.org/10.5194/bg-13-341-2016

Zhu, Q., Riley, W. J., Tang, J. Y., Collier, N., Hoffman, F. M., Yang, X. J., & Bisht, G. (2019). Representing nitrogen, phosphorus, and carbon interactions in the E3SM Land Model: Development and global benchmarking. *Journal of Advances in Modeling Earth Systems*, *11*(7), 2238–2258. https://doi.org/10.1029/2018ms001571

Zhuang, Q., Melillo, J. M., Kicklighter, D. W., Prinn, R. G., McGuire, A. D., Steudler, P. A., et al. (2004). Methane fluxes between terrestrial ecosystems and the atmosphere at northern high latitudes during the past century: A retrospective analysis with a process-based biogeochemistry model. *Global Biogeochemical Cycles*, *18*(3). https://doi.org/10.1029/2004gb002239

Zonneveld, C., & Kooijman, S. A. L. M. (1989). Application of a dynamic energy budget model to Lymnaea-Stagnalis (L). *Functional Ecology*, *3*(3), 269–278. https://doi.org/10.2307/2389365