1 **Feasibility of formulating ecosystem biogeochemical models from established physical rules**

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

3 [1] Department of climate sciences, earth and environmental sciences area, Lawrence Berkeley

4 National Laboratory, Berkeley, CA, USA.

5 [2] Department of Physical Geography and Bolin Centre for Climate Research, Stockholm

6 University, Stockholm, Sweden.

7 [3] Environmental Engineering, The University of Sydney, 2006 Sydney, NSW, Australia

8

9 Corresponding author: Jinyun Tang (jinyuntang@lbl.gov)

10 **Key Points**:

11 • The popular empirically based modeling approaches are limiting the improvement of

12 existing ecosystem biogeochemical models' predictions.

13 • Physical rules-based approaches will help develop ecosystem biogeochemical models as

14 robust as industrial computer-aided design software.

15 • Inter-disciplinary collaboration can accelerate the development and adoption of the

16 physical rules-based ecosystem biogeochemical models.

17

18

19

20

21

22 **Abstract**

23 To improve the predictive capability of ecosystem biogeochemical models (EBMs), we discuss

24 the feasibility of formulating biogeochemical processes using physical rules that have

25 underpinned the many successes in computational physics and chemistry. We argue that the

26 currently popular empirically based modeling approaches, such as multiplicative empirical

27 response functions and the law of the minimum, will not lead to EBM formulations that can be

28 continuously refined to incorporate improved mechanistic understanding and empirical

29 observations of biogeochemical processes. As an alternative to these empirical models, we

30 propose to formulate EBMs using established physical rules widely used in computational

31 physics and chemistry. Through several examples, we demonstrate how mathematical

32 representations derived from physical rules can improve understanding of relevant

33 biogeochemical processes and enable more effective communication between modelers,

34 observationalists, and experimentalists regarding essential questions, such as what measurements

35 are needed to meaningfully inform models and how can models generate new process-level

36 hypotheses to test in empirical studies?

37 **Plain Language Summary**

38 Robust ecosystem biogeochemical models are needed to provide the humanity with predictions

39 to understand and manage interactions between terrestrial ecosystems and the climate. However,

40 existing models are not sufficient because of their wide use of statistical relationships derived

41 from empirical observations. We argue that wider adoption of physical rules can help us develop

42 better ecosystem biogeochemical models to meet with our needs. This can be achieved by deeper

43 interdisciplinary collaboration between scientists from fields in soils, biology, chemistry, physics

44 and mathematics. And then we will be better positioned 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 Noe et al.*, 2023; *Tang and Zhuang*, 2008]. Despite these many efforts, analyses still find significant uncertainties when model predictions are confronted with field perturbation experiments, including the response to free air CO_2 enrichment, nutrient addition, and warming [*Bouskill et al.*, 2014; *Davies-Barnard et al.*, 2020; *De Kauwe et al.*, 2017; *Todd-Brown et al.*, 2013; *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 [*Ahlstrom et al.*, 2013; *Bouskill et al.*, 2014; *Huntzinger et al.*, 2017; *Tang and 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

68   EBMs have made it difficult, and in some cases nearly impossible, to achieve high-quality

69   predictions. Without a fundamental change in model formulation, this challenge will persist

70   despite efforts to augment processes representations, refine parameter calibration, integrate

71   empirical observations, and employ more accurate numerical schemes.

72   The fundamental fallacies we address below may be considered a specific type of

73   structural error, yet we argue that they possess a unique and crucial character, deserving special

74   attention. This assertion arises from our observation that contemporary EBM formulations

75   heavily rely on combinations of empirical response functions derived from field observations and

76   factorial empirical experiments. However, due to the close coupling between the involved

77   entities, it is expected that the effect of each of these targeted biogeochemical processes only

78   emerges from the interactions among several more basic processes, many of which cannot be

79   orthogonally captured by factorial empirical experiments, nor be discerned from field

80   measurements. For example, microbial respiration in soil is dependent both on the microbial

81   physiological status and the substrate transport in soil, both of which are modulated by soil

82   moisture content and temperature [*Suseela et al.*, 2012; *Zhou et al.*, 2014]. As the transport of

83   heat and moisture is closely coupled [*Milly*, 1982; *Saito et al.*, 2006], changes in one of these two

84   conditions will inevitably change the other, such that the temperature and moisture dependence

85   of microbial respiration is impossible to separate.

86   Meanwhile, with the rapidly changing climate and recent resurgence of fossil fuel use

87   [*Tollefson*, 2022], it is becoming less likely that our society will be able to curb global warming

88   within the 2°C limit set at the Paris Agreement [*Lenton et al.*, 2023]. Rather, we expect that

89   climate adaptation and mitigation measures through active ecosystem management will be

90   increasingly important [*Guan et al.*, 2023; *Obersteiner et al.*, 2010], and therefore society

91   urgently needs predictive models to provide more robust and detailed guidance on how such

92   ecosystem-based measures can be properly executed.

93   Robust EBMs require the underlying mathematical formulations to be either simple (with

94   few well constrained parameters) or well balanced, where processes are described by a complete

95   set of physical rules (see Table 1 for a list of example physical rules that we are referring to in

96   this perspective). Unfortunately, existing EBMs often represent biogeochemical processes

97   without considering the underlying mechanistic details, and thus can only provide limited

98   insights into how ecosystem management can effectively address climate adaptation and

99   mitigation. For instance, existing models usually represent soil organic matter (SOM) as a

100 composite of abstract and unmeasurable pools with predefined turnover times modified by

101 edaphic conditions [*Koven et al.*, 2013; *Tao et al.*, 2023; *Viskari et al.*, 2022]. However, it is the

102 diverse chemical composition of SOM and dynamic physical associations and interactions

103 between SOM, soil particles, microbes, water, and plants that determine SOM storage and

104 decomposition dynamics [*Kleber*, 2022; *Lehmann et al.*, 2020]. Thus, effective management

105 should modulate these interactions holistically for SOM storage to be maintained or even

106 enhanced. Models that account for many of these mechanisms are being developed [*Abramoff et*

107 *al.*, 2022; *Grant et al.*, 2017; *Riley et al.*, 2022; *Wang et al.*, 2022], yet implementing them

108 comprehensively in coupled EBMs is still a far-off goal.

109   Additionally, in most existing EBMs, the plant canopy is at best represented with only

110 two big leaves, one sun-lit and the other shaded (i.e., the two-big-leaf approximation; [e.g., *Dai*

111 *et al.*, 2004]), while fine roots are only included implicitly via parameterizations [*Wang et al.*,

112 2010; *Weng et al.*, 2022; *Zhu et al.*, 2019], such that ecosystem performance associated with

113 different canopy structures and root traits cannot be assessed with these models. Plant models

114   that are more explicit in their representation of plant functional traits and associated processes

115   exist [*Abichou et al.*, 2013; *Kang et al.*, 2012], but these models are often hard to couple with

116   ecosystem-level models due to their high complexity and large number of parameters, while

117   EBMs account for soil and atmospheric processes simplistically. The imbalance between

118   complex plant models and relatively simple EBMs then creates a coupling challenge that calls

119   for special care. Further, even if a biogeochemical or biophysical process is represented in

120   EBMs, its mathematical description may ignore essential physical constraints (as discussed

121   below), resulting in poor long-term predictability and lowering stakeholders' confidence in

122   conducting mitigations based on guidance generated from model predictions [*Luo et al.*, 2015].

123   For example, the emerging enthusiasm in lowering agricultural carbon intensity has stirred

124   interest in applying ecosystem models for measuring, reporting, and verifying soil carbon

125   changes and greenhouse gas emissions due to changes in management practices. However,

126   stakeholders have shown diverging confidence in the models' predictive capability [*Guan et al.*,

127   2023]. This situation stands in stark contrast to the developments in industry, where computer-

128   aided design software has facilitated the production of ships measuring hundreds of meters in

129   length and chip circuits as small as a few nanometers [e.g., *Arrichiello and Gualeni*, 2020; *Seok*

130   *et al.*, 2021], and in weather forecasting, where reasonable weather predictions a week in

131   advance are common [*Bauer et al.*, 2008].

132   The successes of computer-aided design software in industry and numerical models in

133   weather forecasting are founded on mathematical models formulated according to physical rules

134   (see Table 1 for examples of these rules). The performance of these models can be continuously

135   improved by including new essential processes [*Zhou et al.*, 2022], adopting more robust and

136   effective numerical solution strategies [*Candel et al.*, 1999; *Lin and Rood*, 1996; *Liu et al.*,

137 2019], utilizing better constrained parameters [*Kotsuki et al.*, 2018; *Wober et al.*, 2020], and

138 implementing more accurate initial and boundary conditions [*Saredi et al.*, 2021; *Xiao et al.*,

139 2007]. Such a trajectory allows for the realization of "the unreasonable effectiveness of

140 mathematics" wherein simple equations can accurately describe complex real-world phenomena

141 [*Wigner*, 1960].

142   As ecosystem biogeochemistry is heavily influenced by living organisms (spanning

143 micro- to macro-organisms; [*Madigan et al.*, 2009; *Taiz and Zeiger*, 2006]), EBM modelers have

144 avoided physical rules-based mathematical representations (thought to be unfeasible to describe

145 complex living organisms) and have chosen instead empirical representations. Yet, the research

146 community now has access to an unprecedented amount of increasingly detailed observations

147 and inferences of traits [e.g., *Kattge and Sandel*, 2020], and micro-climate data ([e.g., *Kearney et*

148 *al.*, 2014]; Ameriflux). Researchers are also able to design new biological traits by gene editing

149 [*Lam et al.*, 2021; *Saurabh*, 2021], and predict intracellular biochemical rates using proteomic

150 information [*De Falco et al.*, 2022; *Sweetlove and Ratcliffe*, 2011]. All of these are providing us

151 with unprecedented opportunities to do designed interactions with biological organisms. We thus

152 contend that the time is ripe for the development of EBMs with mechanistic representations

153 rooted in physical rules (see **Plate** 1 for definitions of physical rules and mechanistically based

154 process representations). Such models will enable the assimilation of a broader range of

155 empirical data and provide more robust numerical predictions, thus guiding ecosystem

156 management more effectively.

157   We organize the remainder of this perspective paper as follows. First, we delineate the

158 part of ecosystem biogeochemistry that will be discussed in this perspective. Second, we analyze

159 the intrinsic limitations of two popular approaches currently used to formulate EBMs: the

160   multiplier-based empirical response function and the law of the minimum. We highlight that

161   these two approaches are unlikely to result in a model that can be incrementally refined as new

162   theories and observations are developed and integrated. Third, we discuss how adopting physical

163   rules-based approaches can lead to significant progress. In particular, we demonstrate with three

164   examples how physical rules-based models can improve understanding of biogeochemical

165   processes and provide opportunities for engineering biogeochemical responses. Finally, we

166   discuss how the research community can work together to develop comprehensive and coherent

167   EBMs based on physical rules to better realize "the unreasonable effectiveness of mathematics"

168   in ecosystem biogeochemical modeling. We note that the mathematical symbols used in our

169   discussion are defined in Table A.

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, theory, 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. We note that what we call primary rules here have also been termed first, fundamental, and primary principles, among others, in the literature. Derived rules are constructed from primary rules and consistent with abundant observational evidence. For example, the change of momentum of a mass 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, when used for biogeochemical modeling, refer to the construction of a mathematical or computational representation 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. For instance, 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.

---

170   **Plate** 1. Definition of "physical rules" and "Mechanistically based process representations".

171

172

173

174

175 Table 1. **Example physical rules.** There could be more primary rules if subatomic interactions
176 are considered, but the six listed here are proposed to be sufficient to develop biogeochemical
177 models. Additionally, we assume that there is no mass-energy conversion in the biochemical
178 reactions, so that mass and energy balance rules are independent. There are many more derived
179 rules, e.g., Navier-Stokes equation and Darcy's law, each of which can be derived from these
180 primary and secondary rules with proper mathematical approximations.

Name	Domain of application	Reference
Primary rules		
Mass balance	Mass exchange	*Feynman et al.* [2011b]
Energy balance	Energy exchange	*Feynman et al.* [2011b]
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, e.g., 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* [2008]; *Feynman et al.* [2011b]
Advective transport	Convection-driven tracer transport	*Steefel et al.* [2005]
Law of mass action	Chemical reactions	*Koudriavstev et al.* [2001]

181

Terrestrial ecosystem dynamics

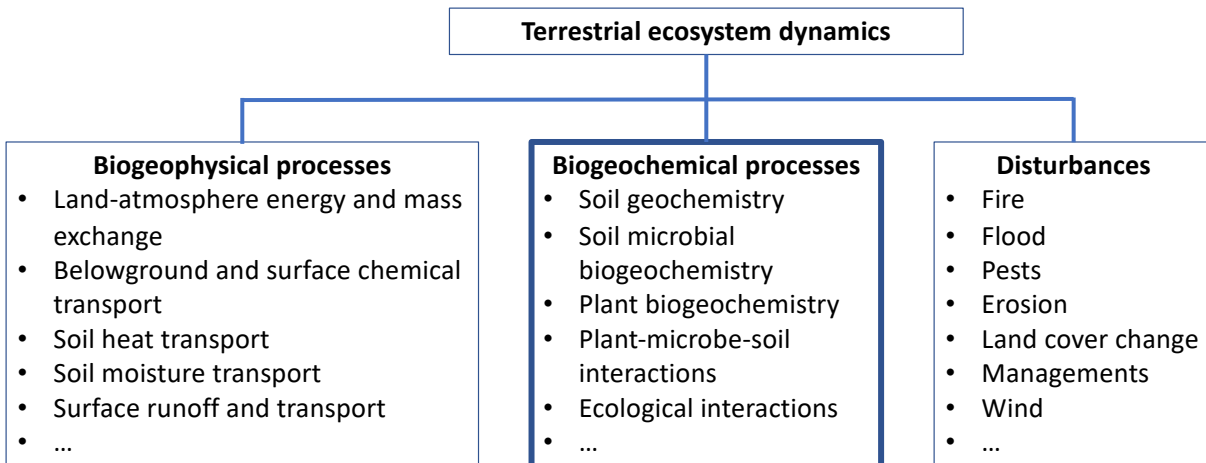| Biogeophysical processes	Biogeochemical processes	Disturbances
• Land-atmosphere energy and mass exchange • Belowground and surface chemical transport • Soil heat transport • Soil moisture transport • Surface runoff and transport • …	• Soil geochemistry • Soil microbial biogeochemistry • Plant biogeochemistry • Plant-microbe-soil interactions • Ecological interactions • …	• Fire • Flood • Pests • Erosion • Land cover change • Managements • Wind • …

182
183 Figure 1. A general (but non-exhaustive) delineation of processes involved in terrestrial
184 ecosystem dynamics. To highlight their importance, plant-microbe-soil interactions are separated
185 from ecological interactions. The biogeochemical processes are the focus of this perspective.

186 **2. Biogeochemical processes in terrestrial ecosystem dynamics and mathematical rules for**

187 **scaling-coherent modeling**

188 To aid our discussion of the difference between empirically based and physical rules-

189 based approaches for formulating EBMs, we first delineate the major biogeochemical processes

190 involved in terrestrial ecosystem dynamics and identify which of them will be within the scope

191 of this perspective (Figure 1). We define biogeochemical processes as those that lead to the

192 production or consumption of chemical species and biomass. Animals are excluded here, even

193 though they are also important in ecosystem nutrient dynamics [e.g., *Atkinson et al.*, 2018].

194 Meanwhile, biogeophysical processes and disturbances (including ecosystem management) are

195 those that affect environmental conditions (e.g., soil and atmospheric temperature and water

196 contents, soil physical and chemical properties) where biogeochemical processes occur [*Robinne*

197 *et al.*, 2020; *Rusu*, 2013]. With this delineation, our discussion in this perspective focuses on how

198 to mathematically represent biogeochemical rates and changes in storage under the influence of

199 environmental and biological factors (e.g., Table 2).

200   We acknowledge that biogeochemical processes are always dependent on the spatial and

201   temporal scales at which they are observed or modeled, but our discussion in this perspective

202   leaves out the challenge of scaling across spatial heterogeneity at the landscape scale. Instead, we

203   discuss scaling issues at a fine scale relevant for process understanding (e.g., from soil pore to

204   core scale; from leaf to canopy scale). Further, we expect that improved physical rules-based

205   modeling will facilitate spatial heterogeneity upscaling through the use of, e.g., remote sensing

206   and machine learning approaches.

207   Moreover, evolutionary processes [*de Vries and Archibald*, 2018; *Greenway*, 1980;

208   *Koonin and Wolf*, 2012; *Tan*, 2022], and processes regulating community and ecosystem

209   assembly [*Higgins*, 2017; *Leibold et al.*, 2017] are also not discussed in this perspective. These

210   processes are linked to the biogeochemical processes discussed here, and, in a first order

211   approximation, can be represented with similar physical rules that describe the movement and

212   transformation of energy and chemical molecules in biogeochemical processes, except that now

213   the functional traits of individual organisms (and their effect on biogeochemical processes) can

214   change through time due to evolution or community-level traits change through time due to

215   variations in community assembly [*Levin*, 1992; *Martiny et al.*, 2023].
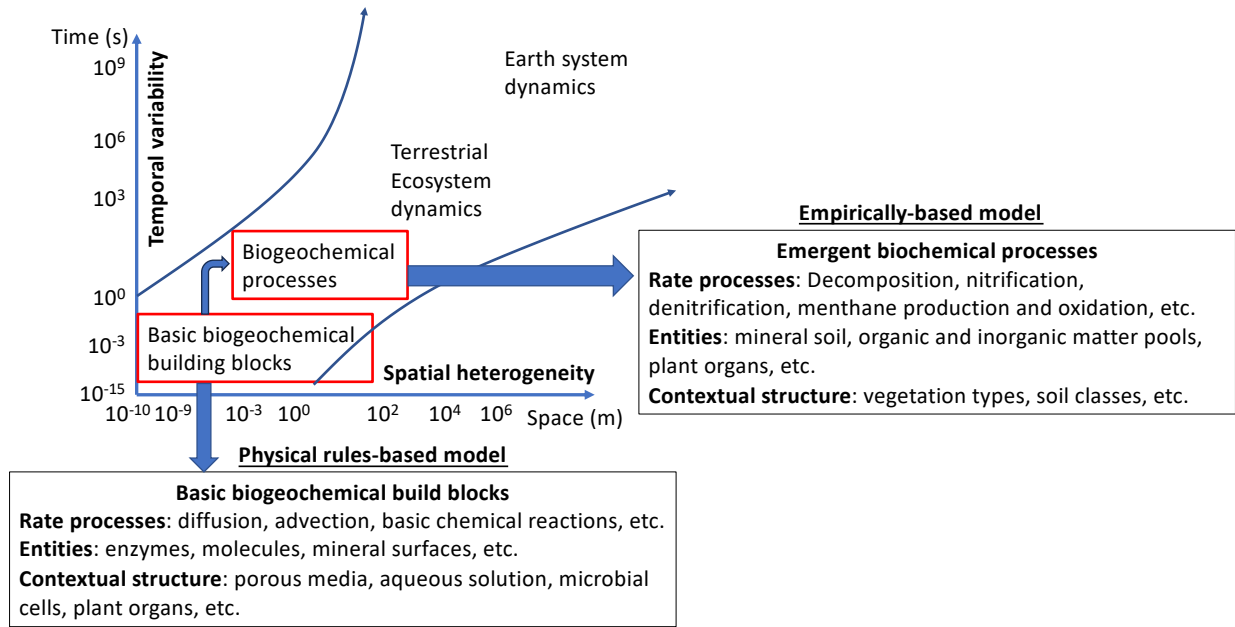
216

Figure 2. A schematic depicting the space-time relationships between basic biogeochemical building blocks and the emergent biogeochemical processes constituting ecosystem 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. We note that heterogeneity increases at greater scales of time and space.

2.1 Modeling biogeochemical processes across scales

One unique feature of natural processes is that their governing equations often change across different spatiotemporal scales (Figure 2). That is, there are qualitative differences between observations of emergent phenomena at fine and 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.*, 2011a; *Thakkar*, 2021], while 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 (10^0 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 and Purcell*,

236   1977; *Feynman et al.*, 2011a; *Koudriavstev et al.*, 2001]). As we further coarsen spatial and

237   temporal scales, the higher levels of organization, nonlinearities, and variability in environmental

238   conditions in space and time become important to biogeochemical rates (due to averaging of

239   nonlinear processes; e.g., *Chakrawal et al.* [2022]; *Wilson and Gerber* [2021]). Thus, to model

240   biogeochemical processes robustly at scales relevant for ecosystem-based climate adaptation and

241   mitigation, model formulations that properly account for these emergent dynamics are needed.

242   We next discuss how development of such formulations could be accomplished and are

243   beneficial.

244   By recognizing that the processes at a particular modeling scale emerge from the

245   processes that occur at finer scales, we expect that there are fundamental relationships between

246   the fine and coarse scales that need to be coherently maintained when model equations are

247   formulated for the coarse scales. Suppose we are to build a model at a coarse spatiotemporal

248   scale designated by index l (Figure 3). The state variables (Y_{n_l}) and their contributing processes

249   (R_{k_l}) (where n_l and k_l are indices for variables and processes, respectively) are constrained by

250   physical rules $\Phi_l\big(Y_{n_l}, R_{k_l}\big) = 0$, where Φ_l is a vector function. Each process R_{k_l} emerges from

251   the interactions between state variables $Y_{m_{l-1}}$ that occur at the fine scale designated by index l-1

252   and state variables Y_{n_l} that occur at the coarse scale l. That is $R_{k_l} = H_{k_l}\big(Y_{n_l}, Y_{m_{l-1}}\big)$. Meanwhile,

253   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

254   involve fine-scale processes $Q_{j_{l-1}}$. Therefore, for a coherent formulation of the parameterization

255   of R_{k_l} at the coarse scale l, one needs to properly maintain the physical constraints

256    $\Phi_{l-1}\left(Y_{m_{l-1}}, Q_{j_{l-1}}\right) = 0$ at the fine scale $l - 1$. Consequently, the extent to which those fine-

257   scale physical constraints are maintained during upscaling determines the quality of the model

258 parameterization at the chosen scale of interest (aka coarse scale l here). Such a coherent scaling

259 approach has been adopted in the transitions between quantum mechanics, Newton's law

260 [*Feynman et al.*, 2011a], Boltzmann's equation [*Boltzmann*, 1964], Chapman-Enskog kinetic gas

261 theory [*Chapman*, 1990], Lattice Boltzmann equation [*Chen and Doolen*, 1998], Navier-stokes

262 equation [*Chen and Doolen*, 1998], Boussinesq equation of surface flow [*Kim et al.*, 2009], and

263 Richards' equation of unsaturated flow [*Bear*, 1972], all of which have contributed to many

264 scientific and engineering successes.

Modeled processes

Scale l

Scale $l-1$

Subscale process (mechanisms)

Upscaled model: $\frac{dY_{n_l}}{dt} = F_{n_l}(R_{k_l})$

Physical constraints: $\Phi_l(Y_{n_l}, R_{k_l}) = 0$

Process representation: $R_{k_l} = H_{k_l}(Y_{n_l}, Y_{m_{l-1}})$

Physical constraints: $\Phi_{l-1}(Y_{m_{l-1}}, Q_{j_{l-1}}) = 0$

Subscale model: $\frac{dY_{m_{l-1}}}{dt} = G_{m_{l-1}}\left(Q_{j_{l-1}}(Y_{m_{l-1}})\right)$
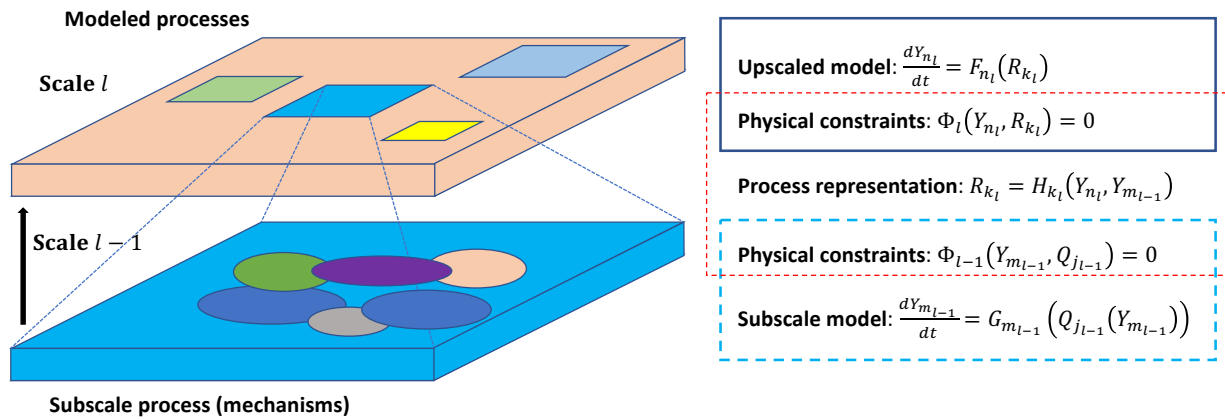
265
266 Figure 3. The relationship between biogeochemical dynamics for an upscaled model (designated
267 by l) and subscale model (designated by $l-1$). Here the subscale model represents processes in
268 a spatial subset of the upscaled model. To indicate that models at scales l and $l-1$ may have
269 different number of state variables and processes, subscripts n_l and k_l are used for scale l, and
270 subscripts m_{l-1} and j_{l-1} are used for scale $l-1$. The parameterization scheme H_{k_l} represents
271 the net effect of process R_{k_l} in the upscaled model, which strives to represent the emergent
272 biochemical effects from the dynamic interactions between state variables Y_{n_l} and $Y_{m_{l-1}}$ at the
273 finer scale. Process $Q_{j_{l-1}}$ is determined by interactions between $Y_{m_{l-1}}$. The potential problem
274 with many existing parameterizations at scale l is that the subscale physical constraints (Φ_{l-1})
275 are ignored, so that R_{k_l} does not include interactions with $Y_{m_{l-1}}$.

276 Conceptually, deriving governing equations for the coarse scale from those of the fine

277 scale can be seen as a lossy data compression problem, where fine-scale details are averaged out

278 while key features are maintained at the coarse scale. Algorithms for such problems have been

279 developed for processing image, video, and audio data [*Hussain et al.*, 2018; *Pan*, 1995; *Poyser*,

280 2021]. Lossy information compression is closely related to machine learning and can be

281   formulated using Bayesian inference [*Cheng*, 2018; *Theodoridis*, 2015]. Additionally, machine

282   learning has also been proposed for efficient upscaling [*Santos et al.*, 2022], while Bayesian

283   inference is frequently used to conduct model-data fusion and estimate model parameters based

284   on observational constraints [*Tang and Zhuang*, 2009; *Vrugt*, 2016]. Therefore, in the following,

285   we use the Bayesian framework [*Jaynes*, 2003] to explain the necessity and benefit of coherent

286   scaling in formulating EBMs.

287   Given N upscaling scheme hypotheses A_j, $j = 1, ..., N$, and a set of measurements B, the

288   Bayesian theorem ranks the merit of A_j by its posterior probability $Pr(A_j|B)$, which is computed

289   as

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

290   where the correlation $Pr(A_j, B) = Pr(B|A_j)Pr(A_j)$ is a product of likelihood probability

291    $Pr(B|A_j)$ and prior probability $Pr(A_j)$, and $Pr(B)$ is the probability of evidence.

292   As formulated above, in the Bayesian inference framework, identification of the best

293   upscaling scheme becomes a model selection problem, where $Pr(A_j)$ represents the prior quality

294   of j-th upscaling scheme, whose posterior merit is $Pr(A_j|B)$ after considering its capability of

295   matching the measurements B. For model-data fusion that aims at parameter estimation of a

296   given model formulation, A_j is the j-th sample of model parameters, whose plausibility is

297    $Pr(A_j|B)$, and the globally optimal parameter set corresponds to the maximum of $Pr(A_j|B)$. For

298   machine learning that uses some kind of numerical approximation, which could be neural

299   networks, polynomials, or regression trees, A_j becomes the coefficients of the approximation

300   method, and $Pr(A_j|B)$ ranks the goodness of model fitting conditioned on the measurements B.

301   Since model selection, parameter estimation, machine learning based upscaling, and model-data

302   fusion all aim to improve EBMs, and share mathematical equivalency, we can explain the

303   necessity and benefit of incorporating physical rules by examining the parameter estimation

304   problem within a model-data fusion framework. This approach, in turn, reinforces the

305   significance of physical rules in all four of these approaches. Specifically, we will show that

306   incorporating physical rules will alleviate the parametric equifinality.

307   To simplify the explanation, we assume that the model-data discrepancy follows the

308   Gaussian distribution, as often assumed in Bayesian inference-based applications of model-data

309   fusion for EBMs [*Tang and Zhuang*, 2009; *Tarantola*, 2005]. Accordingly, the cost function (or

310   loss function termed in machine learning) of model-data discrepancy for a set of model

311   parameter values (i.e., ln $Pr(B|A)$ for parameter set A) can be written in the following plain

312   language form:

Cost function=Observational constraint + scaling coherency rules + prior constraint.

313   Moving towards a more formal definition, the cost function relates processes R_l to the model

314   parameters θ_m through the numerical model, which is constructed based on mechanistic

315   representations, empirical response functions, or neural networks in some machine learning

316   framework [*Tsai*, 2021]. The θ_m and modelled R_l then affect the model goodness of fit (i.e.,

317   observational constraint) and need to satisfy physical rules (i.e., scaling coherency rules). The

318   Bayesian inference seeks the optimal θ_m value that produces the least model-data discrepancy.

319   That is, by identifying the optimal θ_m, we can also obtain the best upscaled equations of EBMs.

320   In mathematical terms, the cost function can be written as

$$J = \frac{1}{2}\Sigma_j\left(Y_j(R_l) - Y_{jo}\right)^T C_j\left(Y_j(R_l) - Y_{jo}\right) + \Sigma_k \frac{\lambda_k}{2}\left(M_k(R_l)\right)^2 + J_0, \qquad (2)$$

321   where vectors $Y_j(R_l)$ are model predicted snapshots of the observed vectors of response variables

322   Y_{jo}, and the corresponding covariance matrix of model-data discrepancy has an inverse specified

323   by C_j.

324   In equation (2), $M_k(R_l)$ represents the residual of the k-th tradeoff or the k-th scaling

325   coherency rule among the processes R_l, which are related to model parameters θ_m. In this term,

326   λ_k is the Lagrangian multiplier for the k-th scaling rule. J_0 is the regularization term from the

327   prior knowledge of θ_m. The scaling coherency rules could be empirical relationships, e.g., the

328   relationship between vapor pressure deficit and stomatal conductance [*Yu et al.*, 2017], the

329   relationship between methane production and pH [*Cao et al.*, 1998], among others. The scaling

330   coherency rules could also be physical rules, e.g., the mass conservation relationship among

331   precipitation, infiltration, surface runoff and ponding. Depending on the specific situation,

332   physical rules may appear as either equality (e.g., for conservation relationship among fluxes) or

333   inequality (e.g., some physical variables like mass or volume should never be negative). For

334   conservation rules, e.g., those of mass or momentum, tradeoff $M_k(R_l)$ should be satisfied

335   exactly, so that considering such rules is equivalent to the incorporation of equality constraint,

336   aka error-free observations encountered in Bayesian inference or data assimilation [*Basir and*

337   *Senocak*, 2022; *Pan and Wood*, 2006]. When the above scaling rules are considered in the

338   physics-based or knowledge-guided machine learning approach [*ElGhawi et al.*, 2023; *Liu et al.*,

339   2022], $M_k(R_l)$ represents the physical knowledge to be incorporated.

340   The identification of optimal parameters based on equation (2) is equivalent to

341   minimizing the cost function J, a process that is related to the first order variation δJ, which can

342   be obtained by applying the chain rule of differentiation to equation (2), such that

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

343   In equation (3), δJ is related to the variation of parameters $\delta\theta_m$ through three types of

344   constraints, with the first from the observations (i.e., $\Sigma_n \left[\Sigma_j C_j \left(Y_j(R_l) - Y_{jO} \right) \frac{\partial Y_j}{\partial R_n} \right] \frac{\partial R_n}{\partial \theta_m} \delta\theta_m$), the

345   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

346   information of the parameters (i.e., $\frac{\partial J_0}{\partial \theta_m} \delta\theta_m$).

347   Equation (3) allows us to make three assertions. First, for two models of the same number

348   of parameters, the lower magnitude of $\frac{\partial R_n}{\partial \theta_m}$ will lead to smaller contributions to the cost function

349   by the first and second types of constraints, so that the cost function is smaller for a given

350   variability of the parameter $\delta\theta_m$. In other words, making the represented process R_n less

351   sensitive to the parameters θ_m leads to a model with lower parametric sensitivity. This is the case

352   when applying physical rules that make the model less sensitive to uncertainty in individual

353   parameters [*Tang and Riley*, 2013; 2021] (also see the example in section 3.2). Second, the

354   tradeoff terms or scaling coherence rules designated by $M_k(R_l)$ act like regularization to the

355   parameter inference processes. When these regularization terms are ignored, posterior models

356   will be less stable and more vulnerable to overfitting, so that the parameters are less well-

357   constrained. The need for regularization is a phenomenon widely observed in machine learning,

358   which is the main driver for the recent surge of interest in physics-guided machine learning

359   [*ElGhawi et al.*, 2023; *Goodfellow et al.*, 2016; *Liu et al.*, 2022]. Third, since δJ measures the

360   resultant model-data discrepancy due to uncertain parameters, it quantifies the severity of

361   parametric equifinality. Therefore, if a model formulation results in lower $\partial J/\partial \theta_m$ (when

362   averaged over the uncertain parameters), the parametric equifinality is reduced and the model's

363   predictive power is improved. Moreover, since explicitly accounting for process tradeoffs by

364    $M_k(R_l)$ often appears as increased model complexity, we contend that more complex models can

365   potentially be more robust, which is at odds with the common criticism of increasing model

366   complexity, i.e., higher model complexity leads to more model parameters, and thereby higher

367   parametric uncertainty.

368   When formulating EBMs, the empirically based approaches work in a top-down manner,

369   where they use regressions to derive response functions (aka $M_k(R_l)$) based on observations of

370   emergent biogeochemical rates, and corresponding environmental factors, such as temperature,

371   moisture, radiation, pH, and soil properties, e.g., some methane production models [*Riley et al.*,

372   2011; *Zhuang et al.*, 2004]. Because the empirically based approaches rely strongly on the

373   amount and context of observations, the resulting response functions can vary significantly from

374   one study to another, so that there is too much uncertainty in the derived scaling coherency rules

375    $M_k(R_l)$. Furthermore, because it is rare and difficult to comprehensively control and measure the

376   variation of all relevant environmental factors and control variables of the biogeochemical rates,

377   the strong context-dependence of the response functions is unlikely to be resolved by further

378   observations. That is, $M_k(R_l)$ at one place or time cannot be transferred confidently to the other

379   place at another time. Such a situation is quite different from measuring the gravitational

380   constant using a pendulum, where the context dependence can be reduced to almost negligible

381   [*Parks and Faller*, 2010]. Moreover, the empirically based regressions of $M_k(R_l)$ generally

382   ignore physical constraints among the subscale processes (i.e., terms as $\Phi_{l-1}\left(Y_{m_{l-1}}, Q_{j_{l-1}}\right) = 0$

383   in Figure 3, which are included by λ_k into equation (3)). Consequently, the resultant

384   parameterizations will unlikely be robust. Such a case has been demonstrated with the superiority

385   of the knowledge guided machine learning model to the pure machine learning based model [*Liu*

386   *et al.*, 2022]. We will provide more analysis on the shortcomings of empirically based

387   approaches in section 2.2.

388   In contrast to the top-down empirically based approaches, physical rules-based

389   approaches work in a bottom-up manner, in which they focus on representing relatively well-

390   understood basic processes and their interactions using well established mathematical constraints

391   and logical inductions. Because these mathematical constraints and logical inductions have been

392   vetted by observations in diverse disciplines, the resultant model constraints are much stronger

393   than can be imposed by limited observed responses (e.g., the term related with Y_{jo} in equation

394   (3)) in a calibration study.

395   Conceptually, the idea adopted by physical rules-based approaches is analogous to

396   building a great variety of lego structures, in which only a few well-designed basic building

397   blocks are used, even though some customized blocks are occasionally needed to knit the pieces

398   together. These customized lego blocks correspond to processes that cannot currently be

399   formulated using known physical rules or are too complex to be computed using known physical

400   rules, but the empirical rules are known to be good enough, thus intuitive or empirical

401   approximations are used instead. For instance, in applying the Richards' equation, we often use

402   the empirical soil water retention curve parameterization that relates soil matric potential and soil

403   water content as a function of soil texture and composition [*Clapp and Hornberger*, 1978; *van*

404   *Genuchten*, 1980]. As another example, in the modeling of plant phenology, empirical rules of

405   plant development are used to guide the plants' temporal development in the model [e.g., *Grant*

406   *et al.*, 2020]. Nevertheless, for EBM modeling, we argue that most biogeochemical processes can

407   be constructed with just a few well-understood basic building blocks. Interestingly, biology

408   seems to work in such a hierarchical way. For example, protein folding can be described by first

409   forming secondary structures from amino acids, then those secondary structures fold into the

410   native state that is able to carry out various biological functions [*Rollins and Dill*, 2014]. In this

411   sense, the physical rules-based approach is explicitly constructing the emergent biogeochemical

412   processes. We will discuss this concept further in section 2.3.

413   Three decades ago, *Agren and Bosatta* [1990] contended that it is impractical to model

414   ecosystems by fitting (sub)systems together piece by piece; instead, they suggested that

415   researchers should focus on "abstract theories" describing ecosystem functioning. We partially

416   agree with them by recognizing that it is impractical to model every function of an ecosystem

417   with physical rules. However, we contend that physical rules will reveal sufficient integrating

418   variables to coherently combine the subsystems, which then will enable better understanding and

419   description of the variability of emergent ecosystem functions. This bottom-up approach

420   contrasts with top-down ecosystem process representations that are based on observational

421   evidence of emergent dynamics. The bottom-up approach allows for prediction in a wide range

422   of conditions, whereas the top-down approach, being constrained by a limited set of

423   observations, might capture current dynamics at some locations, but not necessarily future ones,

424   or at sites with different characteristics. Therefore, EBMs built from physical rules will have

425   much better model-based guidance for ecosystem engineering-based climate mitigation, e.g.,

426   sustainable agriculture intensification [*Pretty and Bharucha*, 2014] and crop plant modification

427   [*Woo et al.*, 2020] for food security.

428

429

430

431

432 Table 2. Example biogeochemical processes used to analyze limitations of the empirically based
433 approaches.

Process	Parameterization	Example references
Soil organic carbon decomposition	$R = R_0 f_1(T) f_2(M) f_3(O)$, where R_0: reference rate, $f_1(T)$: temperature dependence, $f_2(M)$: moisture dependence, $f_3(O)$: 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	*von Caemmerer* [2013]

A_c: carbon-limited rate,
A_j: light-limited rate,
A_p: triosephosphate-limited rate,
R_d: dark respiration.

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]

434

2.2 Limitations of the empirical response function approach

436   Currently, biogeochemical process rates (R) are typically formulated as multiplicative

437   functions of a reference rate (R_0) and 'rate modifiers' (f_j) capturing the effects of environmental

438   conditions (θ_j),

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

439   Alternatively, minimum functions are used under the assumption that the dominant factor

440   constrains the overall rate, such that

$$R = R_0 \min_j f_j(\theta_j). \tag{5}$$

441   These formulations are also used for conductance and resistance that are needed to

442   compute rates. Usually, $f_j(\theta_j)$ is a regression-derived multiplier representing the sensitivity of

443   the rate R to environmental factor θ_j, assuming negligible synergistic or antagonistic interactions

444   among them [e.g., *Jarvis*, 1976]. Occasionally, $f_j(\theta_j)$ has a physical basis. For example, when

445    $f_j(\theta_j)$ represents the dependence of substrate availability in the form of Michaelis-Menten

446   kinetics, it may be argued to be mechanistically based by extrapolating insight obtained at

447   specific conditions. Moreover, for many controlling factors, such as for pH and moisture, $f_j(\theta_j)$

448   is normalized to vary from 0 to 1. The temperature dependence is an exception, where an

449   exponential function without an upper bound may be used, e.g., the Q_{10} function for chemical

450   reactions. We give some example biogeochemical processes that adopt the above formulations in

451   Table 2. In contrast, temperature dependence may have upper bounds in biochemical reactions,

452   with the controlling factor $f_j(\theta_j)$ constrained by physical feasibility [*Pasut et al.*, 2023; *Tang*

453   *and Riley*, 2023b].

454   The multiplicative approach by equation (4) assumes that the different factors θ_j

455   influence the response variable independently [*Jarvis*, 1976]. As this assumption leads to zero

456   covariance between the influence of any two factors, the multiplicative approach is consistent

457   with the multiplicative model in probability theory [*Feller*, 1968; *Hoem*, 1987; *Wermuth*, 1976].

458   Meanwhile, the logic behind the law of the minimum approach of equation (5) is based on the

459   crude empirical observations of crop yield that dates back to the 1820s [*Liebig*, 1840; *Sprengel*,

460   1826]. In some applications, the functional form $f(\theta_j)$ can be argued to be mechanistically

461   based, e.g., the use of Michaelis-Menten or Monod functions [*Liu*, 2007] for the carboxylation

462   process by Rubisco in photosynthesis, or, for biological growth directly related to substrate

463   availability. Despite a mechanistic basis in limited cases, and their mathematical and conceptual

464   simplicity, neither of these approaches provides robust formulations of fluxes and conductance in

465   biogeochemical models, for reasons explained below.

466   In contrast to the assumption of no synergistic or antagonistic interactions underlining the

467   multiplicative approach, first, it is uncommon that in real biogeochemical systems each

468   modulating factor θ_j independently influences the response variable of interest. Instead,

469   interactions between the modulation factors are more common. For example, enzymatic

470   biochemical reactions involve at least two steps: (1) binding of substrate to the enzyme, and (2)

471   new molecule production under enzyme catalysis [*Briggs and Haldane*, 1925]. When one type of

472   enzyme is acting on one type of substrate, this process is often summarized with the Michaelis-

473   Menten kinetics,

$$R = \frac{v_{max}ES}{K+S},$$ (6)

474   where E and S are enzyme and substrate concentrations, K is the affinity parameter (or half

475   saturation constant), and v_{max} is the maximum catalysis rate [*Michaelis and Menten*, 1913].

476   Since biochemical reactions mostly occur in water, and if the unbinding rate is relatively

477   insignificant compared to the forward binding rate (as is usually assumed based on empirical

478   observations), K is approximately proportional to v_{max}/D, with D being the aqueous diffusivity

479   of the substrate with respect to the enzyme [*Tang and Riley*, 2019b; *Zhou*, 1983]. Therefore, in

480   soil, K can be expected to be a function of temperature, moisture, and the type of substrate

481   molecules. Further, in most cases, v_{max} is only a function temperature, while effective

482   concentrations of catalytically active enzyme E and substrate S are functions of soil moisture.

483   Consequently, the temperature and moisture effect on R will only emerge from their influences

484   on K, v_{max} and substrate availability, suggesting that it is highly unlikely that the temperature

485   sensitivity of R is independent of its moisture sensitivity as formulated by the multiplicative

486   approach. For example, *Zhou et al.* [2014] observed that temperature sensitivity of microbial

487   respiration depends strongly on soil moisture status, which cannot be captured by the

488   multiplicative model. On the other hand, if the process is formulated using the law of minimum

489   approach, the model will then predict that once the temperature effect is below a threshold, the

490   moisture effect on increasing the reaction rate will be shielded out. These arguments thus

491   invalidate both equations (4) and (5) for acting as a logically consistent formulation of the

492   biogeochemical rates.

493   Although the above argument does not rule out the feasibility of formulating stomatal

494   conductance with equation (4) or (5) (as one example for their use on representing conductance

495   and resistance), there are sufficient mechanistic reasons to invalidate these two approaches. Like

496   many biochemical processes, stomatal conductance emerges from the interactions between many

497   aspects of plant functioning, so that its response to changes in one influencing factor is

498   dependent on other factors. Moreover, each plant grows by coordinating the traits of all its

499   organs and adjusting to the presence of its neighbors. As a result, stomata behave accordingly.

500   Consequently, there must be a directional information flow (including causality) among a plant's

501   response to its influencing factors, which in turn regulates the behavior of stomatal conductance.

502   Therefore, the effects of different influencing factors on stomatal conductance are unlikely to be

503   simply multiplicative, as this oversimplification neglects the complexity of the plant and its

504   interactions with the environment. A heuristic analogy is the process of putting on one's socks

505   and shoes, in which socks must be put on first even though the selection of socks and shoes may

506   appear to be independent. However, the multiplicative model cannot differentiate the logical

507   order between putting on socks and shoes.

508   In this example of stomatal conductance, the opening of stomata is controlled by the

509   volume and therefore turgor of the guard cells [*Meidner*, 1968] (and epidermal cells as well

510   [*Buckley*, 2019], but including epidermal cells will not influence the conclusion of the following

511   logical induction). The volume of guard cells is an exponential function of their turgor pressure

512   [*Steudle et al.*, 1977], which is linearly related to osmotic pressure inside the guard cells.

513   According to von Hoff's equation [*Atkins and de Paula*, 2006], the osmotic pressure is a linear

514   function of solute concentration inside the guard cell, which depends on the photosynthesis rate

515   (of chloroplasts inside guard cells), and water flux into the guard cell [*Meidner*, 1968]. By

516   applying the mechanical balance in the first order approximation (neglecting xylem cavitation),

517   the osmotic pressure, turgor pressure and the leaf water potential will be linearly related.

518   Therefore, even leaf water potential would be linearly related to soil water potential, soil water

519   potential is affected by photosynthesis rate non-multiplicatively via photosynthesis controls on

520   transpiration. These arguments may partially explain why the empirical Ball-Berry model [*Ball*,

521   1987] and the empirical Leuning model [*Leuning*, 1990; 1995] are not numerically robust in

522   practice, where the effect of soil water stress is applied as a multiplier [*Tang et al.*, 2015].

523   Admitting that the water vapor pressure does not influence stomatal conductance

524   multiplicatively, the model by *Medlyn et al.* [2011] is based on optimality theory [*Cowan*, 1977],

525   which assumes that, within some time period, stomatal conductance adjusts to minimize the

526   marginal water cost for photosynthesis. This assumption results in a deterministic relationship

527   between stomatal conductance and water vapor pressure deficit. In contrast, field data have

528   shown hysteretic relationships between leaf surface vapor pressure deficit and stomatal

529   conductance [*Wang et al.*, 2009]. Coincidentally, when photosynthesis is represented using the

530   model by *Farquhar et al.* [1980], the resultant stomatal conductance is described by a

531   combination of the multiplicative approach and the law of the minimum approach. However,

532   *Walker et al.* [2021] suggested that the law of the minimum approach adopted by the Farquhar

533   model caused significant numerical uncertainty. This evidence indicates that new formulations

534   are needed to capture the rich variability of the response of stomatal conductance to changes in

535   important environmental influencers such as vapor pressure deficit and soil moisture content.

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

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: $\frac{1}{B_V}\frac{dB_R}{dt} = j_{A,max}h_D(s) - \big(\kappa - \mu_D(s)\big)\frac{B_R}{B_V},$

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.

536
537   Plate 2. A brief description of three microbial models that differently treat resource allocation for
538   maintenance and growth.

539   Second, even if the influence from factors θ_j on R can be regarded as mutually

540   independent, $f_j(\theta_j)$ may still be highly context dependent, particularly, when θ_j is dependent on

541   soil conditions. For instance, for soil respiration temperature sensitivity, the literature reports

542   more than 10 different functional forms [*Sierra et al.*, 2015]. Although each of them is able to fit

543   the empirical data used to derive its functional form, none of them is able to extrapolate

544   temperature sensitivity from one experiment to another, and the difference between the diverse

545   functional forms is far from being random and cannot be regarded as noise. When similar

546   procedures are applied to all relevant influencing factors, we should not be surprised to find out

547   that the resultant model lacks accurate spatiotemporal variability when conducting large-scale

548   simulations [*Carvalhais et al.*, 2014; *Todd-Brown et al.*, 2013]. Moreover, because interactions

549   between different modulating factors are not considered with the multiplicative approach of

550   equation (3), the resultant EBMs will be overly sensitive to the parameterization of those

551   factorial dependences [*Exbrayat et al.*, 2013a; *Exbrayat et al.*, 2013b]. As ignoring these

552   interactions corresponds to omitting the scaling rules in equation (2) (aka $M_k(R_l)$), it implies that

553   less constrained relationships are modeled between state variables, and thus the EMBs are of

554   higher parametric equifinality [*Luo et al.*, 2015; *Tang and Zhuang*, 2008]. To help understand

555   this last point, imagine we are describing a dynamic system of three variables, adding a tradeoff

556   between two variables will reduce the three-variable system to a two-variable system, so that the

557   predicted range of variation by the new model is smaller than the original less constrained model.

558   The law of the minimum approach has mostly been applied to biochemical reactions

559   controlled by the supply of multiple substrates. Such processes include photosynthesis [*Farquhar*

560   *et al.*, 1980] and multiple nutrients co-limited biological growth [*Chakrawal et al.*, 2022; *Tang*

561   *and Riley*, 2021; *Wang et al.*, 2010]. It has also been generalized to process rates that are

562   subjected to multiple influencing factors, e.g., the Miami model used it to model net primary

563   productivity as a function of temperature and precipitation [*Lieth*, 1973], and *Noe and Giersch*

564   [2004] used it to model stomatal conductance as a function of light and vapor pressure deficit.

565   Despite its wide adoption in the ecology community, many empirical studies suggest deviations

566   from the law of minimum for both plant and microbial growth [*Egli*, 1991; *Rubio et al.*, 2003].

567   Alternatively, the "multiple limitation hypothesis", which assumes that plants optimize

568   physiologically and morphologically to achieve no wasteful use of nutrients, is not well

569   supported by empirical observations [*Rubio et al.*, 2003]. The "multiple limitation hypothesis" is

570   also unlikely correct for microbial growth due to the often-observed luxury uptake of abundant

571   nutrients while others are in short supply [*Powell et al.*, 2008].

572   Besides the inconclusive empirical support, the law of the minimum approach also

573   introduces numerical difficulty by creating jumps when the predicted limitation shifts between

574   limiting factors. For modeling photosynthesis, such a jump is usually smoothed out by quadratic

575   functions, which involves two hyperparameters without physical meaning (for three co-limiting

576   processes) that are obtained by trial and error [*Collatz et al.*, 1990]. The (arbitrary) choice of

577   these parameters have large effects on simulated plant gross primary productivity [*Walker et al.*,

578   2021]. In the context of model parameter inference, the law of the minimum approach similarly

579   mistreats the tradeoff terms in equation (2), and, as a result, it will lead to significant parametric

580   equifinality, just as the multiplicative approach does.

581   Therefore, due to issues discussed above, the convenient empirically based response

582   function approaches will not facilitate improvement of climate-biogeochemistry feedback

583   predictions, no matter how many more observations are applied as constraints (through model-

584   data fusion). Further, incorporating more biogeochemical processes using the same approach will

585   only increase parametric equifinality and degrade model performance, where the latter has led to

586   the incorrect impression held by the research community that increasing model complexity is

587   usually bad and should be avoided. We will show below (in section 3) that by using physical

588   rules-based approaches, increasing the model's complexity may potentially make it more capable

589   of resolving dynamic variability and be more resilient to parameter uncertainty.

590   **2.3 Feasibility of physical rules-based approaches**

591   As we argued above, the empirically based approaches are unlikely to result in EBMs

592   with high spatiotemporal transferability. We thus turn to the physical rules-based approaches.

593   We argue that physical rules-based approaches are feasible, because biogeochemical processes

594   (the focus of this perspective) can be broadly classified into two groups: (1) transfer of mass and

595   energy, and (2) chemical conversion of molecules. The transfer of mass and energy is achieved

596   through radiation, convection, advection, and diffusion, all of which are well-studied in physics

597   [*Plawsky*, 2020]. The chemical conversion of molecules involves chemical kinetics and

598   thermodynamic control of chemical reaction rates, whose physical rules are also well-studied

599   [*Horn and Jackson*, 1972; *Klotz*, 2008]. These two groups together conceptualize

600   biogeochemical processes as reactive-transport systems. Accordingly, biogeochemistry can be

601   modeled by applying the reactive-transport concept at various spatiotemporal scales with proper

602   mathematical approximations. For aqueous chemistry, this can be done using existing

603   formulations of reactive-transport models [*Molins and Knabner*, 2019; *Steefel and Lasaga*,

604   1994]. Such an approach would in principle be sufficient to model the dynamics of carbon,

605   nutrients, gasses, and solids in soils and water bodies [*Riley et al.*, 2014; *Steefel et al.*, 2015].

606   However, the reaction and transport pathways are not all known, and heterogeneous chemical

607   substrates and complex, biologically-mediated reaction networks make the direct application of

608   reaction-transport model difficult in practice [*Dudal and Gerard*, 2004].

609   Apart from landscape upscaling issues (which we do not consider here), the most difficult

610   part of the physical rules-based approach is to properly deal with biological growth of both

611   micro- and macro-organisms that drive or modulate biogeochemical processes. Biological

612   growth is a phenomenon that emerges from an enormous number of chemical reactions at

613   microscales, for many of which there are robust functional relationships [*Madigan et al.*, 2009;

614   *Young et al.*, 2001]. However, we currently do not have a well-established upscaling scheme to

615   create a bottom-up representation of biological growth—that is, even though we may know the

616   kinetics of each reaction, we do not know their relative importance and their interactions in a

617   living organism. Rather, biological growth could and should currently be modeled using a

618   combination of physical rules and phenomenological formulations. For example,

619   phenomenological rules can be applied to plant growth stages, which are fortunately well-

620   observed for many species, particularly row crops [*Hanway*, 1966; *Miller*, 1992], and have been

621   successfully used to parameterize many natural ecosystem plant types [*Grant*, 2013; *Zhou et al.*,

622   2021]. With a combination of remote sensing data and in-situ measurements [*Dronova and*

623   *Taddeo*, 2022; *Harfenmeister et al.*, 2021; *Xiao*, 2009], it will be possible to obtain plant growth

624   stage parameterizations that are sufficiently robust. Eco-evolutionary approaches can also be

625   used instead of purely phenomenological rules to constrain processes with bounds of ecological

626   and evolutionary feasibility. Compared to phenomenological rules, eco-evolutionary approaches

627   offer a 'built-in' flexibility to deal with varying environmental conditions [*Franklin et al.*, 2020;

628   *Harrison et al.*, 2021].

629   Physical rules can be used to formulate biomass accumulation and translocation in a

630   similar way as for chemical reactions, except that growth is the net result of several related

631   subprocesses. This can be done by dividing growth into several subprocesses, including substrate

632   uptake, internal physiology, and mortality. The substrate uptake process is amenable to relatively

633   well-established physical rules (e.g., law of mass action, which will be discussed with an

634   example in section 3.3). Microbial internal physiology relates how carbon and nutrients,

635   particularly nitrogen and phosphorus, are used for biomass growth, cellular maintenance, and

636   release of extracellular products (e.g., exoenzymes), and has been described using, e.g., the Pirt

637   model [*Pirt*, 1982], the compromise model [*Beeftink et al.*, 1990; *Wang and Post*, 2012], and the

638   dynamic energy budget model [e.g., *Kooijman*, 2009; *Tang and Riley*, 2015; 2023b] (see Plate 2

639   for a brief description).

640   Compared to empirically based models (e.g., the Pirt model, and the compromise model),

641   the dynamic energy budget (DEB) model has a relatively good mechanistic foundation that fully

642   integrates mass and energy trade-offs during metabolic allocation for microbial growth.

643   Importantly, the DEB formulation is structurally compatible with flux-balance models that

644   represent biological growth by considering a large number of intra-cellular chemical processes

645   that include enzymes, metabolites, and genomes [*Antoniewicz*, 2021]. DEB models have also

646   been successfully applied to animals and plants by treating individuals as a population of cells

647   [*Russo et al.*, 2022; *Zonneveld and Kooijman*, 1989]. In particular, it is the only model structure

648   that can reasonably predict the nonlinear relationship between carbon use efficiency (i.e., the

649   ratio of carbon used to grow and carbon from uptake) and growth rate consistent with empirical

650   observations and thermodynamics [*Tang and Riley*, 2023b]. That is, carbon use efficiency will

651   first increase, then plateau and finally decrease with growth rate, whereas the Pirt and

652   compromise models fail to predict the decrease of carbon use efficiency at high growth rate.

653   For plants, the modular nature of plant leaves, branches, stems, and roots, and the

654   associated carbon and nutrient translocation (including both transformation and transport) allows

655 us to model a single plant as a reactive transport system, where each modular part acts as an

656 autonomous individual whose internal physiology follows a DEB-based formulation and the

657 transport of carbon and nutrient between parts follows gradient driven flow with a close coupling

658 with plant hydraulics. A conceptually similar approach has been successfully implemented in the

659 *ecosys* model [*Grant*, 1998] and has been advocated by *Thornley* [1972] to address the balanced

660 growth between plant shoots and roots. We note that this physiological formulation resembles

661 the Pirt's model and does not rigorously follow the DEB framework. However, coupling

662 physiological, plant hydraulics, and morphological parameterizations allows accounting for

663 many geometric and metabolic trade-offs, e.g., the popular pipe-model for plant xylem and

664 canopy development by *Shinozaki* [1964] can emerge naturally from such a treatment [*Grant*,

665 1998].

666 We note that our ideas above are meant to be applied to individual microbes and plants,

667 forming the foundation to model population and community dynamics (e.g., population

668 demography, community and ecosystem assembly). These individual-based formulations can be

669 combined in 'ecosystem demography' models [*Koyen et al.*, 2020; *Ma et al.*, 2022; *Medvigy et*

670 *al.*, 2019], contributing to better modeling of biogeochemistry-climate feedbacks regulated by

671 plant and microbial physiology (a knowledge gap in existing models).

672 **3. Examples to contrast the predictive power of two approaches**

673 In the following, we give three examples to contrast the efficacy of the empirically based

674 and the physical rules-based approaches. Particularly, we demonstrate how the concerns of

675 worsening parameterization equifinality associated with more mechanistic (and usually more

676 complex) models are not universally supported. For more examples, such as for the

677 thermodynamic regulation of biogeochemistry, we refer readers to [*Jin and Bethke*, 2003; *Maggi*

678   *et al.*, 2008], and for nutrient-regulated microbial growth to [*Chakrawal et al.*, 2022; *Tang and*

679   *Riley*, 2023b].

3.1 Soil moisture dependence of substrate affinity parameter

681   This first example addresses the substrate affinity parameter involved in soil carbon and

682   nutrient dynamics, based on the analysis by [*Tang and Riley*, 2019b]. In modeling soil

683   biogeochemistry, we often encounter Michaelis-Menten or Monod-type equations for

684   decomposition rates or nutrient uptake (e.g., equation (6), or variants including microbial

685   biomass instead of enzyme concentrations). The affinity parameter K requires a value to compute

686   the overall reaction rate R. Because K is important for almost every biogeochemical process

687   using a substrate (e.g., Table 2), it often represents a significant fraction of the model calibration

688   effort. Moreover, because K is defined for biogeochemistry in an aqueous environment, and soil

689   moisture is a dynamic variable, the effective K (denoted hereforth by K_w as compared to the

690   intrinsic $K_{w,0}$ for a pure aqueous environment) should be a function of soil moisture, and

691   dimensionally consistent with solutes in the pore space, whose concentration varies with soil

692   moisture (K_w must have a dimension of mass per unit water volume, though).

693   In existing EBMs, the moisture effect on K is often represented using a multiplier

694   function [*Maggi et al.*, 2008; *Riley et al.*, 2011; *Tang et al.*, 2010; *Zhuang et al.*, 2004]. When a

695   dual-Monod formulation is used for biogeochemical reactions involving both gas and solute

696   substrates, a potential double counting of moisture effects may occur. This is because both

697   gaseous and solute substrate concentrations, along with microbial physiology, depend on soil

698   moisture content in different ways, and therefore it is unclear which aspects of the

699   biogeochemical rate's moisture dependence are accounted for by the moisture multiplier.
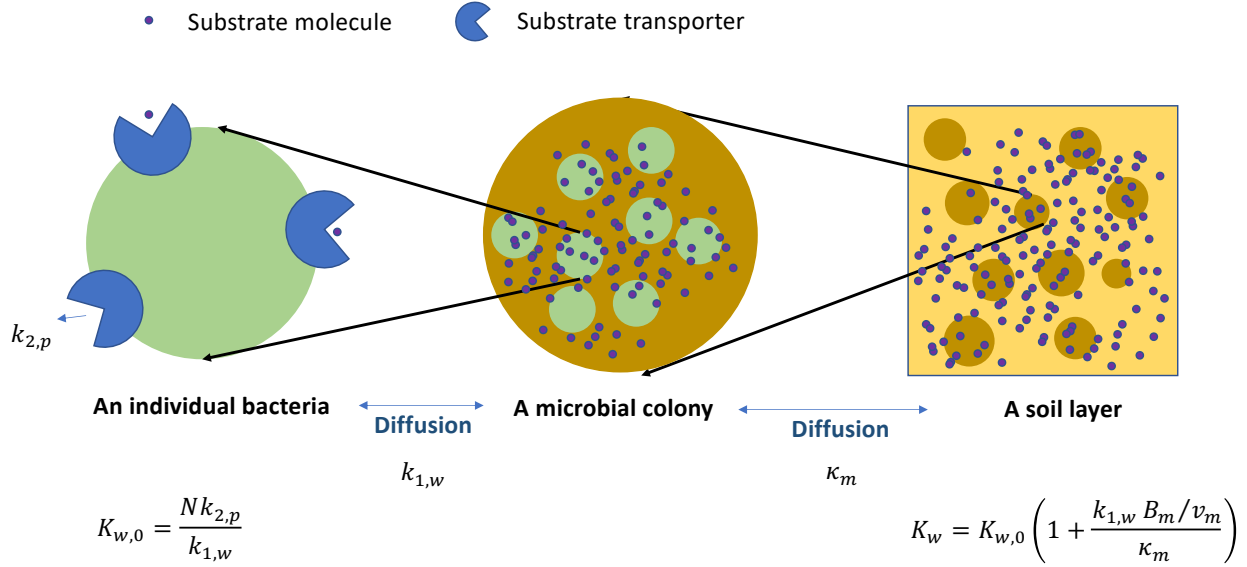
Figure 4. The conceptual model used to derive the moisture dependence of substrate affinity parameter K (i.e., K_w in equation (7)). For simplicity, pumps, channels, and carriers for substrate uptake are all called substrate transporters.

Tang and Riley [2019b] developed a parameterization of K_w by first delineating three

levels of hierarchical structures: (1) an individual microbe (representative of a bacterium), (2) a

colony of microbes (which may represent soil aggregates to some extent), and (3) the soil matrix

(Figure 4). They further assumed that (i) a microbial colony is covered by a water film (whose

thickness is computed as a function of soil suction pressure); (ii) within a colony, microbial cells

are evenly distributed, immersed in water, and compete for a diffusion-limited substrates; and

(iii) microbial colonies are connected to each other by diffusion through the soil matrix. By using

diffusion as the major scaling rule, which is implemented through Smoluchowski's diffusion

theory of chemical reaction [*von Smoluchowski*, 1917] and the Berg-Purcell formula for substrate

interception by a spherical bacterial cell [*Berg and Purcell*, 1977], the affinity parameter for an

aqueous substrate is found as

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

716   where

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

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

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

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

717   and

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

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

719   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}).

3.   *Microbes*: radius of microbial microsite r_m(m), whose volume is $v_m (= 4\pi r_m^3 / 3)$, mean

        microbial biomass density of a microbial microsite B_m / v_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 cm^3. It may also help

734   represent microbial substrate uptake in 1-D vertically resolved reactive-transport based models

735   of soil biogeochemistry [*Dwivedi et al.*, 2019; *Pasut et al.*, 2020; *Riley et al.*, 2014].

736   With its diverse parameters, equation (7) provides insights on how K_w will be modified

737   by soil physical properties (e.g., soil texture, organic matter content), soil moisture content,

738   tracer characteristics, and microbial traits (cf. Figures 2, 3 and 4 in *Tang and Riley* [2019b]). For

739   the results reproduced in Figure 5, K_w for oxygen increases as soils become wetter, following a

740   sigmoidal shape whose inflection point varies with soil texture (Figure 5, left column). K_w for

741   solutes decreases following a nearly exponential decay down to a minimum value upon

742   saturation (Figure 5, center column). When combining these formulations for oxygen and solute

743   affinities in a steady-state microbial respiration model (see [*Tang and Riley*, 2019b] for details),

744   prediction of respiration responses to soil moisture, and their dependence on soil properties,

745   captures observed patterns (Figure 5, right column).

746   Almost all the parameters needed by the above model are routinely measured (e.g., soil

747   characteristics, diffusivities of various molecules [*Cussler*, 2008], tortuosity effect on gas and

748   solute diffusion [*Moldrup et al.*, 2003]), or have been estimated in the literature (e.g., microbial

749   biomass density in a microsite [*Raynaud and Nunan*, 2014]), enabling us to apply equation (7)

750   with very little calibration. *Tang and Riley* [2019b] demonstrated reasonable predictions of the

751   moisture-microbial respiration relationship with typical parameters from the literature and,

752   importantly, without parameter calibration.

753   Additionally, with some modifications, the above model can be adapted to clay particles,

754   fine roots, and fungal hyphae. When this approach is implemented within a reactive-transport

755   based framework of plant shoot-root growth, like that in [*Grant*, 1998], we can obtain new

756   insights on how soil, plant, and microbial traits affect plant-mycorrhizal associations and soil

757   nutrient dynamics.

758   The approach mentioned above has been found to work well for evaluations where

759   microbial traits (e.g., $k_{2,p}$ and N) and microbial biomass (B_m/v_m) are relatively static. However,

760   in dynamic environments, soil moisture fluctuates, and both microbial biomass and traits related

761   to substrate utilization also vary. The next logical step is to couple this framework with equations

762   that describe microbial biomass and trait dynamics, aiming to achieve a mechanistic and

763   ecologically sound soil carbon cycling model. Microbial biomass dynamics are already routinely

764   modeled using more empirical kinetics laws (as discussed in the review by *Chandel et al.*

765   [2023]). Therefore, the challenge lies in how to couple these mechanistic formulations to various

766   aspects of microbial biomass growth (see Plate 2), mortality, maintenance, dormancy, and other

767   functions. We recognize that a mechanistic understanding is not available for some of these

768   functions, but phenomenological or optimization-based approaches can serve as initial

769   approximations to the missing mechanistic representation. During the pursuit of this goal, it

770   would be intriguing to assess to what extent biological (or even ecological) processes are so

771   strongly coupled (or coordinated) to transport processes that they do not need to be modeled

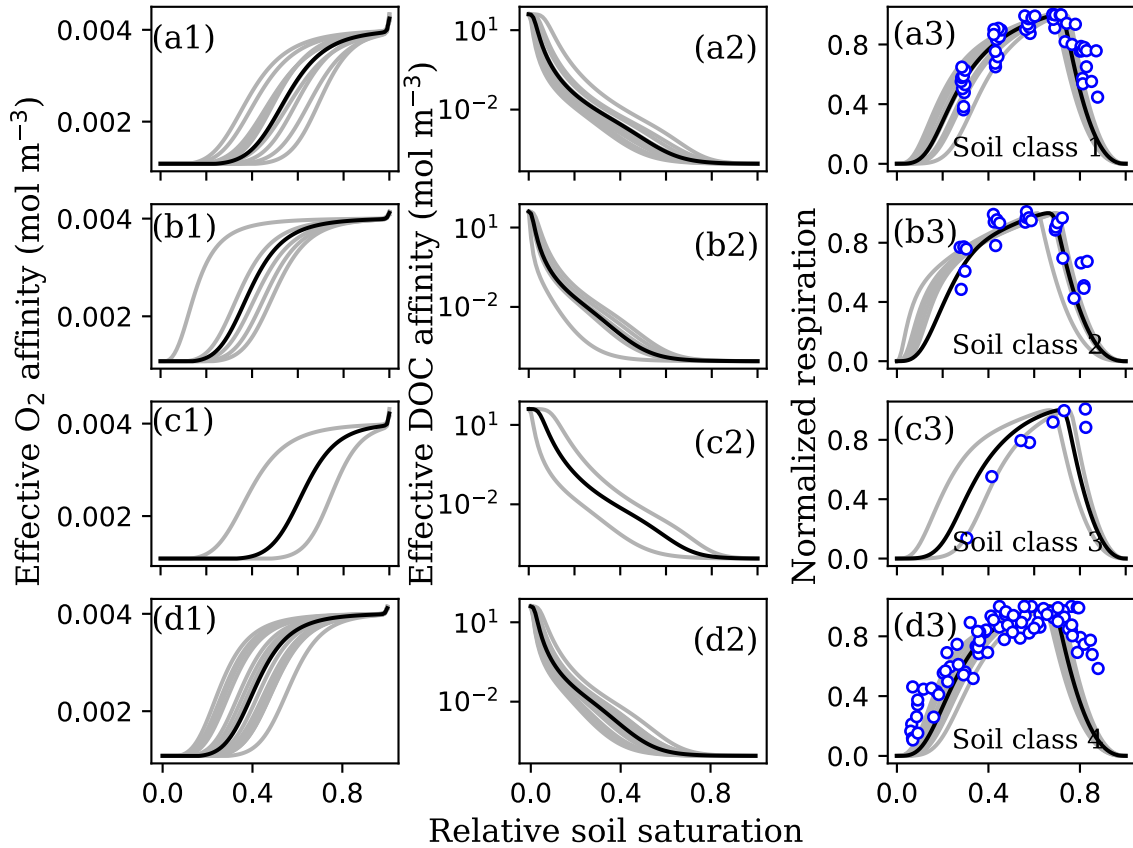772   independently (aka can be lumped through coarse graining).

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.

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

788   microbial cell [*Etienne et al.*, 2020], and trophic networks including producers, consumers, and

789   predators in population ecology [*Barraquand*, 2014; *Buchkowski et al.*, 2022].

790   In the context of biogeochemistry, a network of substrates (S) and consumers (E) can be

791   constructed using the law of mass action, which by aid of the quasi-steady approximation, can be

792   presented in the following form:

$$S_i + \sum_{j=1}^{j=J} X_{ij} = S_{i,T}, \text{ for } i = 1,\cdots,I, \tag{13}$$
$$E_j + \sum_{i=1}^{i=I} X_{ij} = E_{j,T}, \text{ for } j = 1,\cdots,J, \tag{14}$$
$$S_i E_j = K_{ij} X_{ij}, \tag{15}$$
$$\frac{dP_{ij}}{dt} = k_{ij,2} X_{ij}, \tag{16}$$

793   where subscripts i and j indicate different substrates (in total J substrates) and consumers (in total

794   I consumers), and subscript T designates the total concentration of substrate S_i and consumer E_j

795   in the spatial domain of analysis, regardless of their form (free or bound in a complex).

796   Equations (13) and (14) account for the mass balance relationships of substrates and consumers

797   in the system, equation (15) describes the formation of substrate-consumer complex X_{ij}, which is

798   used in equation (16) to compute the production of new materials, denoted by P_{ij}. For a predator-

799   prey network, K_{ij} is related to the handling and attacking rates of a predator on a prey [*Real*,

800   1977], and $k_{ij,2} X_{ij}$ is biomass growth of the predator E_j upon the successful handling of prey S_i.
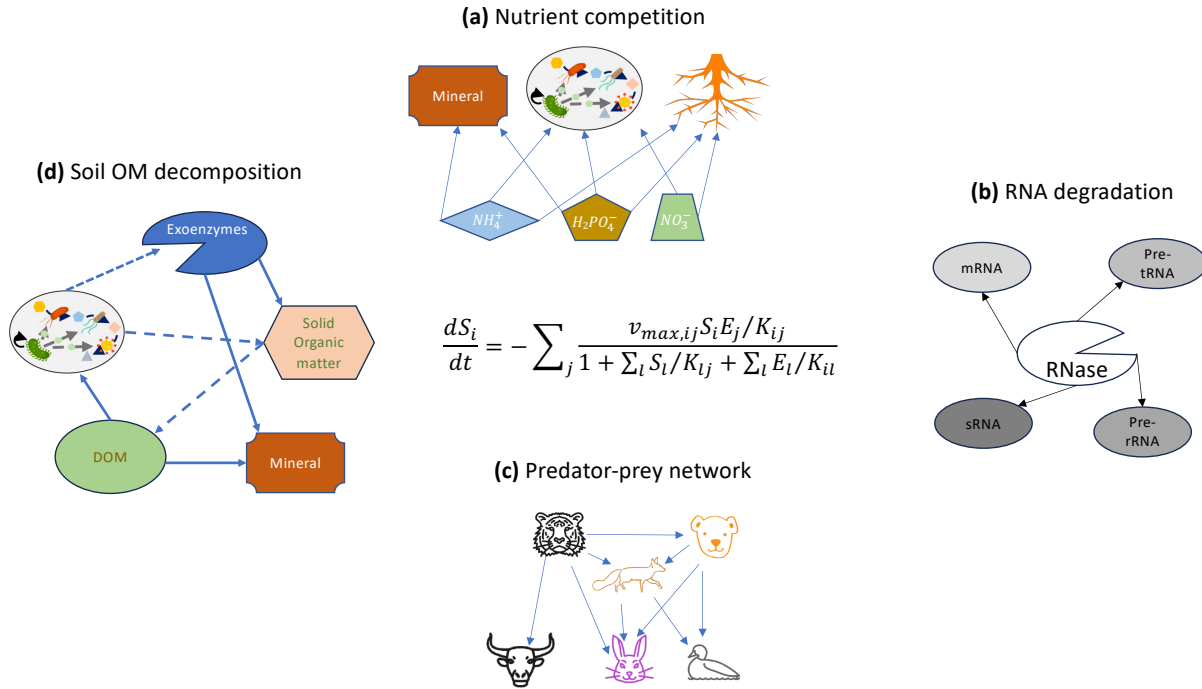
801   The above system as a whole lacks an analytical solution, but it does have a first-order

802   approximation (the Equilibrium Chemistry Approximation, ECA; [*Tang and Riley*, 2013]) as

803   follows:

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

804   Equation (17) can be shown to satisfy the partitioning principle [*Tang and Riley*, 2017], which is

805   critical for developing a theory to coherently upscale from a single chemical reaction to

806   unicellular and multicellular organisms [*Kooijman*, 2009]. Specifically, when S_i are samples

807   from the same substrate S (i.e., $\Sigma_i S_i = S$), and E_j are samples from the same consumer E (i.e.,

808    $\Sigma_j E_j = E$), the sum of X_{ij} will become exactly as the equation that can be obtained by starting

809   with substrate S, and consumer E. That is, by summing over all substrates and consumers in

810   equation (16), we obtain

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

811

(a) Nutrient competition

(d) Soil OM decomposition

$$\frac{dS_i}{dt} = -\sum_j \frac{v_{max,ij} S_i E_j / K_{ij}}{1 + \sum_l S_l / K_{lj} + \sum_l E_l / K_{il}}$$

(b) RNA degradation

(c) Predator-prey network

812

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

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

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

818   In the literature, however, equations (13)-(16) have often been solved with an incomplete

819   consideration of the mass balance constraints imposed by equations (13) and (14). For instance,

820   *Williams* [1973] modeled a system of many consumers competing for a single substrate, whose

821   solution based on ECA is

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

822   where V represents the total consumption rate by predators, with $u_{max,j}$ being the maximum

823   substrate uptake rate by consumer j.

824   However, *Williams* [1973] applied a simple juxtaposition of the empirical Holling's type

825   II predation functions [*Holling*, 1959], and obtained

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

826   where the dependence of individual's predation rate on consumer E_j is not captured (note $E_{j,T}$ is

827   part of $V_{max,j}$ through $V_{max,j} = u_{max,j} E_{j,T}$). Moreover, in models that do include consumer

828   effects on predation rate, the predator competition effect ($\Sigma_j E_j$ in the denominator) is often

829   neglected [*Murdoch*, 1973; *Real*, 1977]. Without these consumer effects, the model could result

830   in incorrect parametric sensitivity when the total substrate is limited [*Tang*, 2015].

831   Additionally, in predator-prey modeling, there has been a long-lasting debate regarding

832   whether the specific predation rate should be dependent on both the density of prey (S_T) and

833   consumers (E_T in our nomenclature), and various formulations have been hypothesized

834   [*Beddington*, 1975; *Berryman*, 1992; *DeAngelis*, 1975; *Ginzburg and Akcakaya*, 1992]. Based on

835   the application of physical rules, the simplest ECA formulation by equation (18) reproduces the

836   Beddington-DeAngelis formulation that is obtained through ad hoc assumptions, while the more

837   general ECA form (equation (17)) has many other applications [*Cheng et al.*, 2019; *Huang et al.*,

838   2018; *Zhu et al.*, 2016].

839   In soil biogeochemical modeling, the simple juxtaposition approach was also used to

840   formulate the decomposition of two pools of soil carbon by a single microbial population, such

841   as in the MIMICS model [*Wang et al.*, 2014], where the growth of microbial biomass is

842   formulated as

$$\frac{dC_b}{dt} = Y_G \frac{C_b v_{max,l} C_l}{C_l + K_l} + Y_G \frac{C_b v_{max,S} C_S}{C_S + K_S}, \tag{22}$$

843   where Y_G is the microbial growth efficiency, assumed the same for both substrates. Note, in

844   equation (22) we have ignored the mortality term to simplify the discussion. Similar as in

845   equation (21), equation (22) predicts that the specific consumption of carbon pool C_l is

846   independent from that of carbon pool C_S.

847   Since there is only one microbial biomass degrading two soil carbon pools, the metabolic

848   effort of the microbial biomass is expected to be divided between the two pools. That is, working

849   on carbon pool C_l has a direct influence on the microbial effort allocated to carbon pool C_S, and

850   vice versa. This subdivision means the formulation by equation (22) predicts the wrong

851   parameter sensitivity, whereas the mechanistically consistent formulation based on the ECA

852   should be

$$\frac{dC_b}{dt} = Y_G \frac{C_b v_{max,l} C_l / K_l}{1 + C_l/K_l + C_S/K_S + \alpha C_b/K_l} + Y_G \frac{C_b v_{max,S} C_S / K_S}{1 + C_l/K_l + C_S/K_S + \alpha C_b/K_S}, \tag{23}$$

853   where α scales the available metabolic effort to the microbial biomass C_b, which, is estimated to

854   be of the order 10^{-4} when substrates are expressed in carbon mass units [*Tang and Riley*,

855   2019a]. Thus, terms multiplied with α can be ignored mostly, but keeping them may prevent

856   runaway microbial biomass growth when applying the model.

857   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

858   lower sensitivity of $\frac{dC_b}{dt}$ to K_l and K_S than equation (22). Further, it can be shown that the

859   parametric sensitivity of $\frac{dC_b}{dt}$ to K_l and K_S are correlated in equation (23), making the resultant

860   model parametrically better constrained, and very likely have less severe parametric equifinality

861   compared to equation (22). This last assertion is consistent with our inference at the beginning of

862 section 2, and supported by the model-data fusion experiment in [*Tang and Riley*, 2013], where

863 the ECA formulation was much more robust than the simple juxtaposition of Holling's type II

864 uptake functions (see comparison of Figures 11 and 12 there). We leave a comprehensive

865 analysis of the new formulations (equation (23) and the corresponding equations of C_l and C_s) on

866 long term soil carbon dynamics for future work.

867 Besides obtaining a more consistent formulation of microbial growth over multiple soil

868 carbon pools, the solution to equations (13)-(16) also leads to a natural incorporation of soil

869 mineral influences on organic carbon decomposition by approximating the organic carbon-

870 mineral interaction with the Langmuir isotherm, leading to a modification of equation (23) as

$$\frac{dC_b}{dt} = Y_G \frac{C_b v_{max,l} C_l / K_l}{1 + C_l/K_l + C_S/K_S + \alpha C_b/K_l + M/K_{l,M}} + Y_G \frac{C_b v_{max,S} C_S / K_S}{1 + C_l/K_l + C_S/K_S + \alpha C_b/K_S + M/K_{S,M}}, \qquad (24)$$

871 where M indicates the total concentration of mineral surfaces available for adsorption and $K_{S,M}$

872 and $K_{l,M}$ are sorption parameters for substrates C_l and C_s.

873 From equation (24), it is inferred that through competitive adsorption of (dissolved) soil

874 organic matter (and exoenzymes; [*Tietjen and Wetzel*, 2003]), microbial decomposition and

875 growth are suppressed by soil minerals. However, if the turnover of exoenzymes is assumed to

876 be positively linked with its catalysis rate, interaction with clay particles could increase the

877 exoenzymes' lifetime by reducing the catalysis rate. Equation (24) then explains that, with

878 increasing soil depth, along with the usual decrease of soil carbon, the specific decomposition

879 rate naturally decreases, lending mechanistic insight to corroborate the attenuation factor in

880 CENTURY-like models, where an exponential attenuation factor is needed to suppress carbon

881 decomposition in order to correctly model the soil carbon profile [*Koven et al.*, 2013]. However,

882 with proper characterization of soil mineral surfaces M and the associated sorption parameters

883 $K_{S,M}$ and $K_{l,M}$, one can more mechanistically characterize the observed vertical heterogeneity

884   than achieved with constant attenuation functions. In particular, the mechanistic model will

885   enable us to evaluate many hypotheses regarding how the interactions between SOM molecule

886   composition, microbial abundance and diversity, soil conditions, and plant input regulate the

887   multiple facet responses of soil respiration and SOM storage to environmental changes.

888   As an example to demonstrate the parametric sensitivity due to different model

889   formulations, we define the specific substrate uptake $F_b = \frac{1}{Y_G C_b} \frac{dC_b}{dt}$ and compute the parametric

890   sensitivity of F_b with respect to V_l and V_S for the Monod kinetics of equation (22),

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

$$\frac{\partial F_b}{\partial v_{max,S}} = \frac{C_S/K_S}{1+C_S/K_S}; \tag{26}$$

891   and, similarly, for the ECA-based equation (24),

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

$$\frac{\partial F_b}{\partial v_{max,S}} = \frac{C_S/K_S}{1+C_l/K_l+C_S/K_S+\alpha C_b/K_S+M/K_{S,M}}. \tag{28}$$

892   From equations (25) and (26), we see that Monod kinetics predicts the parametric sensitivities

893    $\frac{\partial F_b}{\partial v_{max,l}}$ and $\frac{\partial F_b}{\partial v_{max,S}}$ to be independent from the interactions between C_l, C_S, and M, while such

894   dependence is captured by ECA kinetics (equations (27) and (28)). In particular, the Monod

895   kinetics always predicts higher parametric sensitivity than the ECA kinetics (Figure 7), implying

896   that the same parametric uncertainty will lead to higher parametric equifinality for models using
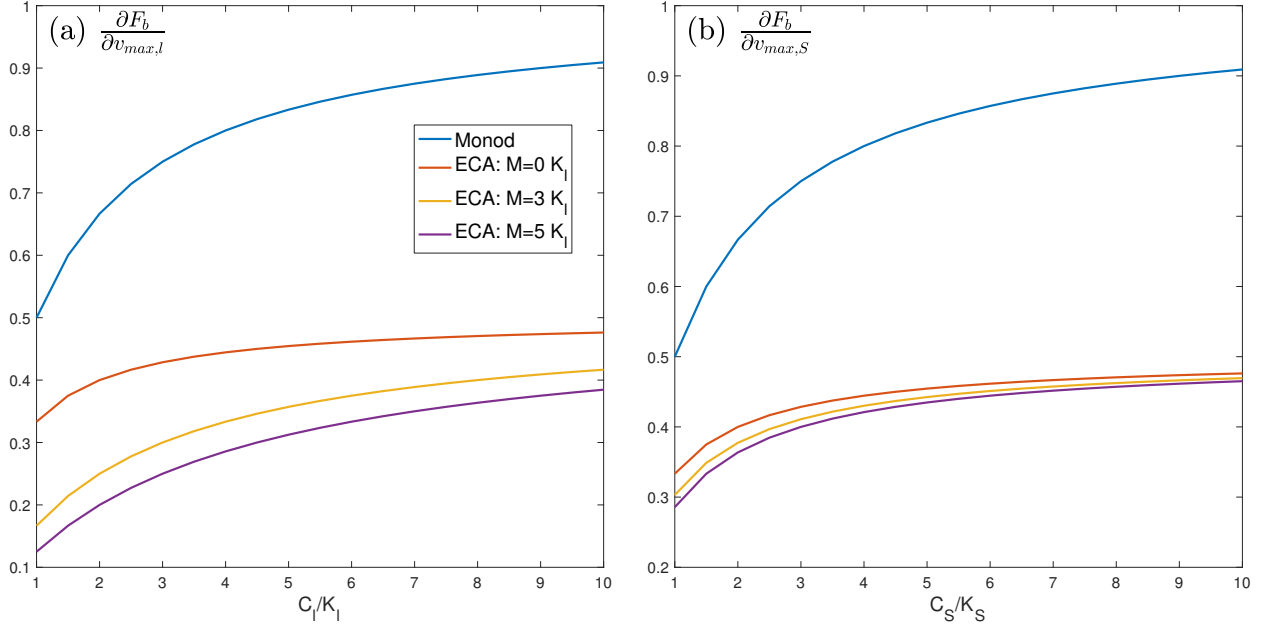
897   the Monod kinetics.

898

899 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

900 Monod kinetics vs the ECA kinetics. For all calculations, it is assumed $K_S = 10K_l$.

3.3 Temperature dependence of enzyme-catalyzed one-substrate reactions

902 In our third example, we discuss the temperature sensitivity of an enzyme-catalyzed one-

903 substrate reaction [*Tang and Riley*, 2013]. Depending on the size contrast between substrate and

904 enzyme molecules, we have three limiting classes of solutions derived from ECA kinetics [*Tang*,

905 2015; *Tang and Riley*, 2019a]:

906 (1) When substrate molecules are much larger than enzymes, or the enzymes are in

907 significant excess of substrate binding surface area (e.g., cellulose during hydrolysis), equations

908 (18) and (19) can be approximated by reverse Michaelis-Menten (MM) kinetics:

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

909 (2) When substrate molecules are much smaller than enzymes (e.g., microbial uptake of

910 glucose) or the system is enzyme limited (i.e., in the typical Michaelis-Menten regime), we have

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

911

912   (3) When substrate and enzyme molecules have similar size (e.g., when fructose is

913   the substrate and invertase is the enzyme), we have the reaction represented using the ECA

914   kinetics

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

915   For all cases, temperature dependence of the maximum reaction rate can be approximated

916   by the transition state theory [*Eyring*, 1935]:

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

917   where ΔH_r is the enthalpy of activation (which is constant), R_g is the universal gas constant, T_0

918   is the reference temperature (K), $v_{max,0}$ is the rate at T_0. By adopting the usual assumption that

919   the unbinding rate is negligible compared to forward binding rate and the relative movement

920   between substrates and enzymes is dominated by diffusion [*Tang et al.*, 2021], the temperature

921   sensitivity of the affinity parameter K_{ES} is determined by the ratio between the temperature

922   sensitivity of v_{max} and that of the aqueous diffusivity D_w (see equation (7) is example 1).

923   According to the Stokes-Einstein equation [*Cussler*, 2008], the aqueous diffusivity for a

924   spherical object of radius a is $D_{w,a} = k_B T/(6\pi\eta a)$, where the dynamic viscosity η has an

925   empirical temperature dependence as $exp(B/T)$ (with $B > 0$; [*Holmes*, 2011]), thus a good

926   approximation is:

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

927   where ΔH_K is the effective enthalpy of K, which is the difference between ΔH_r and the

928   activation enthalpy of the self-diffusion of water, so that $\Delta H_K < \Delta H_r$.

929   In addition to temperature effects on reaction kinetics, temperature also affects the

930   capacity of enzymes to perform the reaction. In fact, enzymes are proteins, and proteins may lose

931   or regain their native structure (and thus functionality) spontaneously. Because this spontaneous

932   transition is taking advantage of the structural perturbations caused by thermal motions in the

933   enzyme solution, the transition between native and unfolded states always occurs for an enzyme

934   that is not irreversibly denatured [*Finkelstein and Ptitsyn*, 2016]. The fraction of active enzymes

935   in native state at a given temperature can be described by the well-established temperature

936   relationship [*Ghosh and Dill*, 2009; *Murphy et al.*, 1990; *Sawle and Ghosh*, 2011]

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

937   where the Gibbs free energy of unfolding ΔG_E is defined as

$$\Delta G_E = \Delta H_E - T\Delta S_E = n_E \Delta C_p[(T - T_H) + Tln(T_S/T)],$$ (35)

938   where n_E is the number of amino acid residues of the enzyme, heat capacity ΔC_p (J K^{-1} (mol

939   amino acid)$^{-1}$) defines the energy required to reorganize the water molecules surrounding the

940   protein, T_H is the temperature at which enthalpy ΔH_E equals to zero, and T_S is the temperature at

941   which entropy ΔS_E equals to zero.

942   Based on the survey by *Silverstein* [2020], ΔC_p seems to be quite consistent among

943   thermophobic, mesophilic, and thermophilic proteins, centering around 60 J (mol amino acid)$^{-1}$,

944   with an increasing variability from thermophobic to thermophilic proteins. Meanwhile, T_H

945   increases from thermophobic to thermophilic proteins, with an increasing difference between T_S

946   and T_H (see Table 6 in [*Silverstein*, 2020]). A comprehensive analysis using data from the

947   Protein Data Bank will be very helpful to gain more insights on the parameterization of equation

948   (35).

949   When the above relationships are applied together to class (1) (reverse MM kinetics), we

950   have

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

951   and when applied to case (2) (MM kinetics), we have

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

952   Therefore, these results show that the overall temperature sensitivity of an enzyme

953 catalyzed one-substrate reaction emerges from three types of temperature functions: (1)

954 Arrhenius equation, (2) the Eyring's transition-state theory, and (3) the thermal stability of native

955 proteins. The equation for case (3) is not reported here because it is just a combination of cases

956 (1) and (2).

957   We note that when the above results are adapted to the substrate affinity parameter of

958 microbial substrate uptake, e.g., the bacterial cells discussed in the first example in Section 3.1,

959 the temperature dependence of the affinity parameter will be more complicated than represented

960 by the Arrhenius-like function, because it will also involve $f_{act}(T)$ through its interaction with

961 the number of transporters distributed over the microbial cells (i.e., parameter N in equations (7)-

962 (9)).

963   To visualize the above relationships, we show some examples of hypothetical enzymes

964 based on mean values of ΔC_p, T_S, and T_H from Table 6 in *Silverstein* [2020] and a typical

965 enzyme length of 290 amino acids of prokaryotes [*Brocchieri and Karlin*, 2005]. In Figure 8a-b,

966 we see that the range of temperatures in which enzymes stay active expands as the Gibbs free

967 energy curves shift from thermophobic to thermophilic enzymes. More interestingly, for case (1)

968 described by equation (36), the normalized reaction rate $R/(v_{max}(T_0)S)$ increases steadily

969 (almost exponentially because f_{act} is close to 1) across most of the biochemically relevant

970 temperatures, with sharp drop-offs at the low and high temperature ends (Figure 8c, where

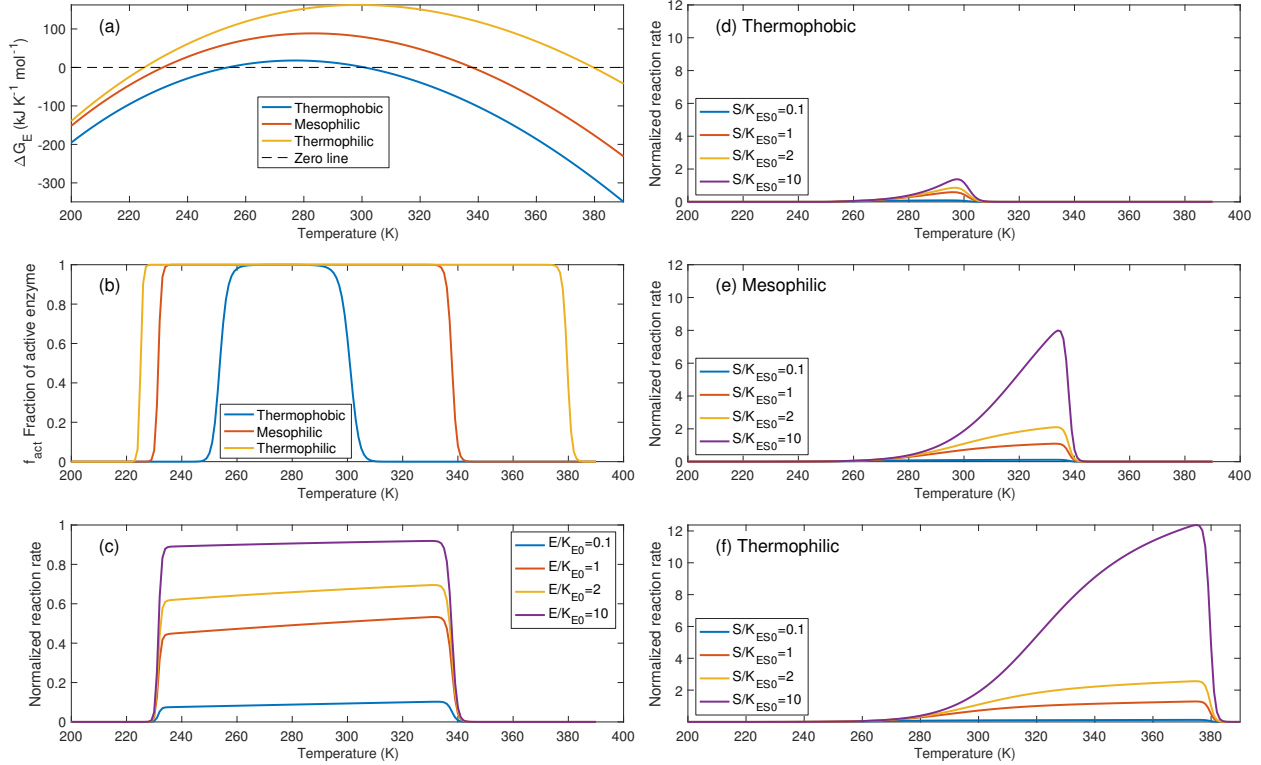971 curves are drawn for the hypothetical mesophilic enzyme).

972

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 based on equation (36) as a function of temperature for a hypothetical mesophilic enzyme for different ratios of enzyme concentration to affinity constant E/K_{E0}; (d) normalized reaction rated based on equation (37) as a function of temperature for a hypothetical mesophilic enzyme; (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 $K(T) = K_0\exp\left(-37300/R_g T\left(1 - \frac{T}{290}\right)\right)$ 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 $v_{max}(T) = \frac{T}{290}\exp\left(-58000/R_g T\left(1 - \frac{T}{290}\right)\right)$.

In contrast to the relationship by equation (36) shown in Figure 8c, when the relationship by equation (37) is illustrated (Figure 8d, e, and f), all cases show the often observed asymmetric temperature response [*Peterson et al.*, 2004; *Ratkowsky et al.*, 2005; *Sharpe and Demichele*, 1977]. In addition to temperature, substrate concentration also plays a role: higher substrate

990 availability increases the reaction rate for a given value of the affinity constant. Therefore, our

991 examples imply that it is problematic to assume that, under high substrate concentrations, the

992 temperature response curve only reflects the temperature-dependence of enzyme catalysis rate. In

993 reality, the temperature response curve also depends on the temperature dependence of the

994 affinity parameter, so that a high substrate concentration cannot ensure equation (37) to derive a

995 temperature response curve that accurately approximates that of v_{max}. As this assumption is the

996 foundation of the macromolecular rate theory by *Hobbs et al.* [2013] that is built off the study by

997 *Peterson et al.* [2004], a comprehensive analysis is presented elsewhere [*Tang and Riley*, 2023a].

998 **4. How to make physical rules-based approaches easily accessible?**

999 With the three examples above, we showed that it is feasible and advantageous to

1000 formulate EBMs using physical rules-based approaches. However, compared to the more

1001 intuitive empirically based approaches, significant efforts are needed to realize these proposed

1002 advantages, at least partially because modeling equations are less intuitive to understand and

1003 apparently contain more parameters to calibrate. We recommend the following steps to achieve

1004 this goal.

1005 First, bringing more expertise and knowledge of mathematical physics into the field of

1006 ecosystem biogeochemistry. This is already done well in the research area of ecosystem

1007 biogeophysics, where physical rules like Ohm's law for resistor networks, transport theories of

1008 diffusion, and advection are used to formulate the exchange and temporal evolution of mass and

1009 energy between soil, water, atmosphere, and other related components [*Shuttleworth and*

1010 *Wallace*, 1985], and textbooks also explain those applications in detail [e.g., *Bonan*, 2019]. For

1011 ecosystem biogeochemistry, we believe constructive effort can be applied through (1) building

1012 long-term and stable collaborations between biogeochemistry empiricists, applied

1013   mathematicians, chemists, and physicists who are keen to model ecosystem biogeochemistry

1014   mechanistically, and (2) enhanced exposure of students in ecosystem biogeochemistry to

1015   concepts in mathematical physics including law of mass action; chemical reaction theories; and

1016   basic reactive transport modeling. From a pedagogic perspective, students could be challenged to

1017   test the classic Michaelis-Menten equation, or a linear model, using datasets where the use of

1018   ECA is necessary. Also, faculties with expertise in biochemistry can team with colleagues in

1019   mathematics and physics to develop a course on mathematical biogeochemistry. This approach

1020   could motivate more young people with interests in physics rules-based ecosystem

1021   biogeochemical modeling.

1022   Second, EBMs formulated using physical rules-based approaches will often be

1023   mathematically more complex, which may be contrary to the heuristic belief that models should

1024   be simple. We fully agree that unnecessary complexity should be avoided by all costs. However,

1025   we contend that the research community should be more open to endorsing the higher

1026   complexity resulting from constructions based on solid mathematical logic and coherency, as

1027   compared to the simpler empirical equations typically derived by regression with context-

1028   dependent measurements. For instance, when the Lagrangian of the standard model of particle

1029   physics is written explicitly term by term [*Shivni*, 2016], the resulting gargantuan equation may

1030   easily fill a whole regular page of a journal paper. Nonetheless, the astonishing success of the

1031   standard model so far does not warrant any omission of its terms, and when the model is

1032   explained term by term, the mechanisms behind are readily interpretable. As we argued

1033   previously, given that ecosystem biogeochemistry encapsulates both living actors and inanimate

1034   matter, which exist in different phases, and interact from very small to very large spatial scales,

1035   the true governing equations of EBM may be as complex as, if not more complex than, the

1036   standard model of particle physics (in terms of length when they are put down onto paper).

1037   Therefore, we should not judge the complexity of EBMs by the number of mathematical terms

1038   involved, rather the complexity should be measured by the basic ideas of physics and

1039   mathematics being incorporated. Notably, even for a very complex system, physical rules will

1040   provide additional constraints to significantly reduce the actual degrees of freedom, so that the

1041   resultant EBM is simpler. As such, we believe that mechanistically more interpretable, and

1042   logically more coherent EBMs could be developed with improved model predictability.

1043   Third, we acknowledge that EBMs formulated using physical rules-based approaches will

1044   usually be computationally more complex and demanding, and therefore may be more difficult to

1045   be applied at large scales. We propose that this scalability issue can be solved in two steps: (1)

1046   creating a numerical library that consists of processes that are formulated using physical rules,

1047   but are provided with user friendly software interfaces to be used in other models [e.g., *Riley et*

1048   *al.*, 2022; *Tang et al.*, 2022], and (2) improving the numerical efficiency of these model

1049   formulations by leveraging new developments in machine learning and artificial intelligence.

1050   The first idea has led to software like OpenFOAM [*Jasak*, 2009] and COMSOL [*Pryor*, 2009]

1051   that are able to solve computational fluid dynamics problems in various configurations. The

1052   second idea is currently used to develop more advanced parameterization schemes, such as

1053   turbulence closure schemes [*Kurz et al.*, 2023] and cloud processes parameterization for

1054   atmospheric models [*Beucler*, 2021]. In ecosystem biochemical modeling, a machine learning

1055   model, when pretrained with a physical rules-based ecosystem biochemistry model, could

1056   conduct spatiotemporal extrapolation more efficiently and even outperform the original EBM,

1057   successfully resolving the challenge of upscaling [*Liu et al.*, 2022].

1058 Last but not least, as can be seen from our three examples, physical rules-based

1059 approaches require substantial comprehensive empirical data support for both forming the

1060 conceptual model and parameterization. Fortunately, much relevant data is available from the

1061 literature, such as solubility and diffusivities of chemical tracers in water and air [*Cussler*, 2008;

1062 *Sander*, 2015]. New biological data that characterize the morphology and rates of biological

1063 organisms, however, are needed. These data should be collected more frequently, together with

1064 macro-chemical data such as carbon and nutrient concentrations. Microbial elemental

1065 stoichiometry, morphology, number or mass density in various soils, and their relative

1066 abundances under various conditions, will be very helpful for formulating physical rules-based

1067 models of soil microbial processes. However, for most of these processes there are no established

1068 measurement methods and novel tracer experiments are only now starting to provide detailed

1069 microbial trait and rate data [e.g., *Canarini et al.*, 2020; *Warren and Manzoni*, 2023]. For plants,

1070 more in-situ phenological data and morphological data (such as leaf sizes, thickness, height, root

1071 architecture, and morphology) will be essential to robustly formulate biogeochemistry using

1072 physical rules, which can also improve the model representation of biogeophysics, such as water

1073 and heat exchange between soils, plants, and atmosphere. On the one hand, existing databases

1074 (e.g., ESS-DIVE (https://ess-dive.lbl.gov), TRY plant trait database [*Kattge and Sandel*, 2020])

1075 can aid making data accessible. On the other hand, physical rules-based approaches can suggest

1076 more specific answers to the question from empiricists to modelers: "what do you want us to

1077 measure?"

1078 **5.   Summary**

1079 Lao Tzu has said that "*the Tao that can be told is not the eternal Tao*", so all ecosystem

1080 biogeochemical models (EBMs) that we develop are inherently limited. Nevertheless, we argue

1081 that the currently popular approach that extensively uses empirically based functions to

1082 formulate biogeochemical processes limits EBMs to incorporate needed improvements. Instead,

1083 by adopting the physical rules-based approach proposed here, more robust and accurate EBMs

1084 can be developed for spatial and temporal extrapolation. Compared with empirical functions, the

1085 primary physical rules are more consistent with our current knowledge of the world, and the

1086 derived physical rules are less context dependent and have more easily quantifiable uncertainty.

1087 Moreover, using physical rules to formulate biogeochemical processes will reveal more detailed

1088 insights about the interactions between the entities involved, which will facilitate the design of

1089 more targeted empirical experiments. To build EBMs that maximally use current knowledge of

1090 physical rules, we advocate more and closer interdisciplinary collaborations in both research and

1091 education between scientists in biogeochemistry, biophysics, soil physics, and mathematics.

1092

1093 **Table A**

1094 Nomenclature. For units, "variable" means the unit is problem formulation dependent.

Symbol	Unit	Meaning and places of use
$f_j(\theta_j)$	None	Effect multiplier from influencer θ_j; Eq. (4) and (5).
$f_{act}(T)$	None	Fraction of enzymes being active; Eq. (34), (36) and (37).
$h_C(s)$	None	Substrate dependency for the compromise model.
$h_D(s)$	None	Substrate dependency for the DEB model.
$h_P(s)$	None	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; Eq. (7) and (8).
$k_{2,p}$	s^{-1}	Maximum substrate uptake rate per site; Eq. (7).
$k_{ij,2}$	s^{-1}	Maximum uptake rate of substrate S_i by enzyme E_j; Eq. (16).
m_P	s^{-1}	Specific microbial maintenance rate.
n_E	None	Number of amino acid residues of the enzyme; Eq. (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; Eq. (8) and (9).
r_m	M	Microbial microsite radius; Eq. (10).
v_m	m^3	Microbial microsite volume; Eq. (10).
$v_{max}(T)$	s^{-1}	Maximum substrate processing rate at temperature T; Eq. (32).
$v_{max,0}$	s^{-1}	Maximum substrate processing rate at temperature T_0; Eq. (32).
$v_{max,E}$	s^{-1}	Maximum substrate processing rate; Eq. (29).
$v_{max,E}(T)$	s^{-1}	Maximum substrate processing rate at temperature T; Eq. (36).
$v_{max,ES}$	s^{-1}	Maximum substrate processing rate; Eq. (30).
$v_{max,ES}(T)$	s^{-1}	Maximum substrate processing rate at temperature T; Eq. (37).
B_m	mol cell m^{-3}	Mean microbial biomass in a microsite; Eq. (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; Eq. (23) and (24).
C_j	None	Inverse of covariance matrix for variable Y_j; Eq. (2) and (3).
C_l	mol C m^{-3}	Fast decaying carbon pool; Eq. (22)-(24), (25), (27), (28).
C_S	mol C m^{-3}	Slow decaying carbon pool; Eq. (22)-(24), (26)-(28).
ΔC_p	J K^{-1} (mol amino acid)$^{-1}$	Heat capacity; Eq. (35)
$D_{g,0}$	$m^2\ s^{-1}$	Gaseous diffusivity; Eq. (11).
$D_{w,0}$	$m^2\ s^{-1}$	Aqueous diffusivity; Eq. (10), (11).
E	mol m-3	Enzyme concentration; Eq. (6)
E_j	Variable	Consumer concentration; Eq. (14), (15).
$E_{j,T}$	Variable	Total consumer concentration; Eq. (14), (17), (18), (20).
F_{n_l}	Variable	Total flux of variable Y_{n_l}; Figure 3.
$G_{m_{l-1}}$	Variable	Total flux of variable $Y_{m_{l-1}}$; Figure 3.
ΔG_E	J mol^{-1}	Gibbs free energy of enzyme unfolding; Eq. (34), (35).
H_{k_l}	Variable	Process function corresponding to R_{k_l}; Figure 3.
ΔH_E	J mol^{-1}	Enthalpy of the enzyme unfolding; Eq. (35).
ΔH_K	J mol^{-1}	Enthalpy of affinity parameter; Eq. (33).
ΔH_r	J mol^{-1}	Enthalpy of enzymatic chemical reaction; Eq. (32).
J_0	None	Cost function contributed by prior information; Eq. (1).
K	mol m^{-3}	Substrate affinity parameter; Eq. (6).
K_0	mol m^{-3}	Substrate affinity parameter at temperature T_0; Eq. (33).
K_E	mol m^{-3}	Substrate affinity parameter; Eq. (29).
K_{ES}	mol m^{-3}	Substrate affinity parameter; Eq. (30).
$K_E(T)$	mol m^{-3}	Substrate affinity parameter at temperature T; Eq. (36).
$K_{ES}(T)$	mol m^{-3}	Substrate affinity parameter at temperature T; Eq. (37)
K_{ij}	Variable	Affinity parameter between substrate S_i and consumer E_j; Eq. (15).

K_l	mol m^{-3}	Microbial affinity parameter to carbon pool C_l; Eq. (22)-(24), (25), (27), (28).
K_S	mol m^{-3}	Microbial affinity parameter to carbon pool C_S; Eq. (22)-(24), (26)-(28).
$K_{l,M}$	mol m^{-3}	Affinity parameter between carbon pool C_l and mineral M. Eq.(24).
$K_{S,M}$	mol m^{-3}	Affinity parameter between carbon pool C_S and mineral M. Eq.(24).
K_w	mol m^{-3}	Effective substrate affinity parameter in soil; Eq. (7).
$K_{w,0}$	mol m^{-3}	Substrate affinity parameter in water; Figure 4.
$M_k(R_l)$	Variable	k-th scaling rule between processes R_l; Eq. (2) and (3).
N	None	Number of transporters per microbial cell; Eq. (8) and (9).
N_A	mol^{-1}	Avogadro number; Eq. (8)
P	Variable	Product concentration from consumption of substrate; Eq. (19).
P_{ij}	Variable	Product concentration by E_j from consuming S_i; Eq. (16).
$Q_{j_{l-1}}$	Variable	Mechanistic interactions between $Y_{m_{l-1}}$; Figure 3.
R	Variable	Rate, or conductance, or resistance; Eq. (4)-(6), (29)-(31), (36), (37).
R_0	Variable	Reference value of R; Eq. (4), (5).
R_g	J mol^{-1} K^{-1}	Universal gas constant; Eq. (32).
R_{k_l}	Variable	Process rate; Figure 3.
S	mol m^{-3}	Substrate concentration; Eq. (6).
ΔS_E	J mol^{-1}	Entropy of enzyme unfolding; Eq. (35).
S_i	Variable	Free concentration i-th Substrate; Eq. (13), (15), (17).
$S_{i,T}$	Variable	Total concentration of i-th substrate; Eq. (13), (17).
T	K	Thermodynamic temperature; Eq. (32)-(37).
T_0	K	Reference thermodynamic temperature; Eq. (32), (33).
T_H	K	Thermodynamic temperature when ΔH_E equals zero; Eq. (35).
T_S	K	Thermodynamic temperature when ΔS_E equals zero; Eq. (35).
V	mol m^{-3} s^{-1}	Total substrate uptake rate; Eq. (20), (21).
$v_{max,l}$	s^{-1}	Specific maximum uptake rate of C_l; Eq. (22)-(25), (27).
$V_{max,j}$	mol m^{-3} s^{-1}	Maximum uptake rate by consumer E_j; Eq. (21).
$v_{max,S}$	s^{-1}	Specific maximum uptake rate of C_S; Eq. (22)-(24), (26), (28).
X_{ij}	Variable	Substrate-consumer complex between S_i and E_j; Eq. (15)-(18).
X	Variable	Total substrate-consumer complex; Eq. (18), (19).
Y_G	None	Biomass yield coefficient; Eq. (22)-(24); Plate 2.
$Y_j(R_l)$	Variable	Generic model variable; Eq. (2), (3).
Y_{jO}	Variable	Observations corresponding to $Y_j(R_l)$; Eq. (2), (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}	None	Reserve biomass yield for the DEB model; Plate 2.
α	None	Scaling parameter from microbial biomass to substrate binding sites; Eq. (23), (24), (27), (28).
λ_k	None	Lagrangian multiplier for k-th scaling rule; Eq. (2), (3).
κ	s⁻¹	Specific reserve biomass mobilization rate; Plate 2.
δ	M	Water film thickness; Eqs. (10), (12).
κ_m	s⁻¹	Specific substrate transfer rate between soil matrix and microbial microsite; Eqs. (7), (10).
θ_m	Variable	Model parameters; Eqs. (2), (3).
$\mu_{max,P}$	s⁻¹	Specific maximum biomass growth rate for the Pirt model; Plate 2.
$\mu_{max,C}$	s⁻¹	Specific maximum biomass growth rate for the compromise model; Plate 2.
$\mu_C(s)$	s⁻¹	Specific biomass growth rate for the compromise model; Plate 2.
$\mu_D(s)$	s⁻¹	Specific biomass growth rate for the DEB model; Plate 2.
$\mu_P(s)$	s⁻¹	Specific biomass growth rate for the Pirt model; Plate 2.
ψ	MPa	Soil matric potential; Eq. (12).
ϕ_g	m³ m⁻³	Air-filled soil porosity; Eq. (11).
ϕ_w	m³ m⁻³	Water-filled soil porosity; Eq. (11).
τ_g	None	Soil tortuosity for gas diffusion; Eq. (11).
τ_w	None	Soil tortuosity for solute diffusion; Eq. (11).
Φ_{l-1}	Variable	Fine-scale physical constraints; Figure 3.
Φ_l	Variable	Coarse-scale physical constraints; Figure 3.

1095

Acknowledgements

1103

1104   **Data Availability Statement**

1105   This study only uses published data. Specifically for Figure 5, data for the four soil classes are

1106   (1) medium to fine texture soils from *Doran et al.* [1990]; (2) coarse texture soil from *Doran et*

1107   *al.* [1990]; (3) other soils from *Doran et al.* [1990]; and (4) soils from *Franzluebbers* [1999],

1108   respectively. Parameters used for the calculation are from *Tang and Riley* [2019b].

1109

1110   **Reference**

1111   Abichou, M., C. Fournier, T. Dornbusch, C. Chambon, R. Baccar, J. Bertheloot, T. Vidal, C.

1112   Robert, D. Gouache, and A. Bruno (2013), Re-parametrisation of Adel-wheat allows reducing

1113   the experimental effort to simulate the 3D development of winter wheat, in *International*

1114   *Conference on Functional Structure Plant Models,*, edited, Saariselkä, Finland.

1115   Abramoff, R. Z., B. Guenet, H. C. Zhang, K. Georgiou, X. F. Xu, R. A. V. Rossel, W. P. Yuan,

1116   and P. Ciais (2022), Improved global-scale predictions of soil carbon stocks with Millennial

1117   Version 2, *Soil Biol Biochem*, *164*, doi:ARTN 108466

1118   10.1016/j.soilbio.2021.108466.

1119   Agren, G. I., and E. Bosatta (1990), Theory and model or art and technology in ecology, *Ecol*

1120   *Model*, *50*(1-3), 213-220, doi:Doi 10.1016/0304-3800(90)90051-H.

1121   Ahlstrom, A., B. Smith, J. Lindstrom, M. Rummukainen, and C. B. Uvo (2013), GCM

1122   characteristics explain the majority of uncertainty in projected 21st century terrestrial ecosystem

1123   carbon balance, *Biogeosciences*, *10*(3), 1517-1528, doi:10.5194/bg-10-1517-2013.

1124   Anderson, P. W. (1972), More is different - broken symmetry and nature of hierarchical structure

1125   of science, *Science*, *177*(4047), 393-396, doi:DOI 10.1126/science.177.4047.393.

1126 Antoniewicz, M. R. (2021), A guide to metabolic flux analysis in metabolic engineering:

1127 Methods, tools and applications, *Metab Eng*, *63*, 2-12, doi:10.1016/j.ymben.2020.11.002.

1128 Arrichiello, V., and P. Gualeni (2020), Systems engineering and digital twin: a vision for the

1129 future of cruise ships design, production and operations, *Int J Interact Des M*, *14*(1), 115-122,

1130 doi:10.1007/s12008-019-00621-3.

1131 Atkins, P., and J. de Paula (2006), *Physical Chemistry*, Eighth Edition ed., W. H. Freeman and

1132 Company, 41 Madison Avenue, New York, NY 10010.

1133 Atkinson, C. L., B. J. Sansom, C. C. Vaughn, and K. J. Forshay (2018), Consumer aggregations

1134 drive nutrient dynamics and ecosystem metabolism in nutrient-limited systems, *Ecosystems*,

1135 *21*(3), 521-535, doi:10.1007/s10021-017-0166-4.

1136 Azizi-Rad, M., G. Guggenberger, Y. M. Ma, and C. A. Sierra (2022), Sensitivity of soil

1137 respiration rate with respect to temperature, moisture and oxygen under freezing and thawing,

1138 *Soil Biol Biochem*, *165*, doi:ARTN 108488

1139 10.1016/j.soilbio.2021.108488.

1140 Baldocchi, D. D., B. A. Hutchison, D. R. Matt, and R. T. Mcmillen (1985), Canopy radiative-

1141 transfer models for spherical and known leaf inclination angle distributions - a test in an Oak

1142 Hickory Forest, *J Appl Ecol*, *22*(2), 539-555, doi:Doi 10.2307/2403184.

1143 Ball, J. T. W., Ian E; Berry, Joseph A (1987), A model predicting stomatal conductance and its

1144 contribution to the control of photosynthesis under different environmental conditions, in

1145 *Progress in Photosynthesis Research: Volume 4 Proceedings of the VIIth International Congress*

1146 *on Photosynthesis Providence, Rhode Island, USA, August 10–15, 1986*, edited by J. Biggins,

1147 Springer Netherlands, Dordrecht, doi:https://doi.org/10.1007/978-94-017-0519-6_48.

1148   Bao, J. L., and D. G. Truhlar (2017), Variational transition state theory: theoretical framework

1149   and recent developments, *Chem Soc Rev*, *46*(24), 7548-7596, doi:10.1039/c7cs00602k.

1150   Barraquand, F. (2014), Functional responses and predator-prey models: a critique of ratio

1151   dependence, *Theor Ecol-Neth*, *7*(1), 3-20, doi:10.1007/s12080-013-0201-9.

1152   Basir, S., and I. Senocak (2022), Physics and equality constrained artificial neural networks:

1153   Application to forward and inverse problems with multi-fidelity data fusion, *J Comput Phys*,

1154   *463*, doi:ARTN 111301 10.1016/j.jcp.2022.111301.

1155   Batchelor, G. K. (1967), *An introduction to fluid dynamics*, Cambridge University Press.

1156   Bauer, J., M. Herbst, J. A. Huisman, L. Weihermuller, and H. Vereecken (2008), Sensitivity of

1157   simulated soil heterotrophic respiration to temperature and moisture reduction functions,

1158   *Geoderma*, *145*(1-2), 17-27, doi:10.1016/j.geoderma.2008.01.026.

1159   Bear, J. (1972), *Dynamics of Fluids in Porous Media*, Dover Publications.

1160   Beddington, J. R. (1975), Mutual interference between parasites or predators and its effect on

1161   searching efficiency, *J Anim Ecol*, *44*(1), 331-340, doi:Doi 10.2307/3866.

1162   Beeftink, H. H., R. T. J. M. Vanderheijden, and J. J. Heijnen (1990), Maintenance requirements -

1163   energy supply from simultaneous endogenous respiration and substrate consumption, *Fems*

1164   *Microbiol Ecol*, *73*(3), 203-209, doi:Doi 10.1016/0378-1097(90)90731-5.

1165   Berg, H. C., and E. M. Purcell (1977), Physics of chemoreception, *Biophys J*, *20*(2), 193-219,

1166   doi:Doi 10.1016/S0006-3495(77)85544-6.

1167   Berryman, A. A. (1992), The origins and evolution of predator prey theory, *Ecology*, *73*(5),

1168   1530-1535, doi:10.2307/1940005.

1169   Beucler, T. G., Ebert-Uphoff, I., Rasp, S., Pritchard, M., Gentine, P. (2021), Machine learning

1170   for clouds and climate, in *Geophys. Monog. Serie*, edited,

1171   doi:https://doi.org/10.1002/essoar.10506925.1.

1172   Blyth, E. M., et al. (2021), Advances in land surface modelling, *Curr Clim Change Rep*, *7*(2),

1173   45-71, doi:10.1007/s40641-021-00171-5.

1174   Boltzmann, L. (1964), *Lectures on Gas Theory*, University of California Press.

1175   Bonan, G. (2019), *Climate Change and Terrestrial Ecosystem Modeling*, Cambridge University

1176   Press, One Liberty Plaza, 20th Floor, New York, NY 10006, USA, doi:10.1017/9781107339217.

1177   Bouskill, N. J., W. J. Riley, and J. Y. Tang (2014), Meta-analysis of high-latitude nitrogen-

1178   addition and warming studies implies ecological mechanisms overlooked by land models,

1179   *Biogeosciences*, *11*(23), 6969-6983, doi:10.5194/bg-11-6969-2014.

1180   Briggs, G. E., and J. B. S. Haldane (1925), A note on the kinetics of enzyme action., *Biochem J*,

1181   *19*(2), 338-339, doi:10.1042/bj0190338.

1182   Brocchieri, L., and S. Karlin (2005), Protein length in eukaryotic and prokaryotic proteomes,

1183   *Nucleic Acids Res*, *33*(10), 3390-3400, doi:10.1093/nar/gki615.

1184   Buchkowski, R. W., J. M. Barel, V. E. J. Jassey, and Z. Lindo (2022), Cannibalism has its limits

1185   in soil food webs, *Soil Biol Biochem*, *172*, doi:ARTN 108773

1186   10.1016/j.soilbio.2022.108773.

1187   Buckley, T. N. (2019), How do stomata respond to water status?, *New Phytol*, *224*(1), 21-36,

1188   doi:10.1111/nph.15899.

1189   Canarini, A., W. Wanek, M. Watzka, T. Sanden, H. Spiegel, J. Santrucek, and J. Schnecker

1190   (2020), Quantifying microbial growth and carbon use efficiency in dry soil environments via O-

1191   18 water vapor equilibration, *Global Change Biol*, *26*(9), 5333-5341, doi:10.1111/gcb.15168.

1192 Candel, S., D. Thevenin, N. Darabiha, and D. Veynante (1999), Progress in numerical

1193 combustion, *Combust Sci Technol*, *149*(1-6), 297-337, doi:Doi 10.1080/00102209908952110.

1194 Cao, M. K., K. Gregson, and S. Marshall (1998), Global methane emission from wetlands and its

1195 sensitivity to climate change, *Atmos Environ*, *32*(19), 3293-3299, doi:Doi 10.1016/S1352-

1196 2310(98)00105-8.

1197 Carvalhais, N., et al. (2014), Global covariation of carbon turnover times with climate in

1198 terrestrial ecosystems, *Nature*, *514*(7521), 213-217, doi:10.1038/nature13731.

1199 Chakrawal, A., S. Calabrese, A. M. Herrmann, and S. Manzoni (2022), Interacting bioenergetic

1200 and stoichiometric controls on microbial growth, *Front Microbiol*, *13*, doi:ARTN 859063

1201 10.3389/fmicb.2022.859063.

1202 Chandel, A. K., L. F. Jiang, and Y. Q. Luo (2023), Microbial models for simulating soil carbon

1203 dynamics: a review, *J Geophys Res-Biogeo*, *128*(8), doi:ARTN e2023JG007436

1204 10.1029/2023JG007436.

1205 Chapman, S. C., T G (1990), *The mathematical theory of non-uniform gases: an account of the*

1206 *kinetic theory of viscosity, thermal conduction and diffusion in gases*, Cambridge University

1207 Press.

1208 Chen, S., and G. D. Doolen (1998), Lattice Boltzmann method for fluid flows, *Annu Rev Fluid*

1209 *Mech*, *30*, 329-364, doi:DOI 10.1146/annurev.fluid.30.1.329.

1210 Cheng, Y. W., N. J. Bouskill, and E. L. Brodie (2019), Model exploration of interactions

1211 between algal functional diversity and productivity in chemostats to represent open ponds

1212 systems across climate gradients, *Ecol Model*, *406*, 121-132,

1213 doi:10.1016/j.ecolmodel.2019.05.007.

1214   Cheng, Z. S., Heming; Takeuchi, Masaru; Katto, Jiro (2018), Deep convolutional autoencoder-

1215   based lossy image compression, in *2018 Picture Coding Symposium (PCS)*, edited,

1216   doi:10.1109/PCS.2018.8456308.

1217   Clapp, R. B., and G. M. Hornberger (1978), Empirical equations for some soil hydraulic-

1218   properties, *Water Resour Res*, *14*(4), 601-604, doi:DOI 10.1029/WR014i004p00601.

1219   Collatz, G. J., J. A. Berry, G. D. Farquhar, and J. Pierce (1990), The relationship between the

1220   rubisco reaction-mechanism and models of photosynthesis, *Plant Cell Environ*, *13*(3), 219-225,

1221   doi:10.1111/j.1365-3040.1990.tb01306.x.

1222   Cowan, I. R. F., G D (1977), Stomatal function in relation to leaf metabolism and environment,

1223   *Symp. Soc. Exp. Biol.*(31), 471-505.

1224   Cussler, E. L. (2008), *Diffusion: Mass Transfer in Fluid Systems*, THIRD EDITION ed.,

1225   Cambridge University Press.

1226   Dai, Y. J., R. E. Dickinson, and Y. P. Wang (2004), A two-big-leaf model for canopy

1227   temperature, photosynthesis, and stomatal conductance, *J Climate*, *17*(12), 2281-2299,

1228   doi:10.1175/1520-0442(2004)017<2281:Atmfct>2.0.Co;2.

1229   Davies-Barnard, T., et al. (2020), Nitrogen cycling in CMIP6 land surface models: progress and

1230   limitations, *Biogeosciences*, *17*(20), 5129-5148, doi:10.5194/bg-17-5129-2020.

1231   De Falco, B., F. Giannino, F. Carteni, S. Mazzoleni, and D. H. Kim (2022), Metabolic flux

1232   analysis: a comprehensive review on sample preparation, analytical techniques, data analysis,

1233   computational modelling, and main application areas, *Rsc Adv*, *12*(39), 25528-25548,

1234   doi:10.1039/d2ra03326g.

1235   De Kauwe, M. G., et al. (2017), Challenging terrestrial biosphere models with data from the

1236   long-term multifactor Prairie Heating and CO2 Enrichment experiment, *Global Change Biol*,

1237   *23*(9), 3623-3645, doi:10.1111/gcb.13643.

1238   de Vries, J., and J. M. Archibald (2018), Plant evolution: landmarks on the path to terrestrial life,

1239   *New Phytol*, *217*(4), 1428-1434, doi:10.1111/nph.14975.

1240   DeAngelis, D. L. G., R A; O'Neill, R V (1975), A model for trophic interaction, *Ecology*, *56*(4),

1241   881-892, doi:10.2307/1936298.

1242   Doran, J. W., L. N. Mielke, and J. F. Power (1990), Microbial activity as regulated by soil water-

1243   filled pore space, in *International Society of Soil Science*, edited.

1244   Dronova, I., and S. Taddeo (2022), Remote sensing of phenology: Towards the comprehensive

1245   indicators of plant community dynamics from species to regional scales, *J Ecol*, *110*(7), 1460-

1246   1484, doi:10.1111/1365-2745.13897.

1247   Dudal, Y., and F. Gerard (2004), Accounting for natural organic matter in aqueous chemical

1248   equilibrium models: a review of the theories and applications, *Earth-Sci Rev*, *66*(3-4), 199-216,

1249   doi:10.1016/j.earscirev.2004.01.002.

1250   Dwivedi, D., J. Y. Tang, N. Bouskill, K. Georgiou, S. S. Chacon, and W. J. Riley (2019), Abiotic

1251   and biotic controls on soil organo-mineral interactions: developing model structures to analyze

1252   why soil organic matter persists, *Rev Mineral Geochem*, *85*, 329-348,

1253   doi:10.2138/rmg.2019.85.11.

1254   Egli, T. (1991), On multiple-nutrient-limited growth of microorganisms, with special reference

1255   to dual limitation by carbon and nitrogen substrates, *Anton Leeuw Int J G*, *60*(3-4), 225-234,

1256   doi:10.1007/Bf00430367.

1257   ElGhawi, R., B. Kraft, C. Reimers, M. Reichstein, M. Korner, P. Gentine, and A. J. Winkler

1258   (2023), Hybrid modeling of evapotranspiration: inferring stomatal and aerodynamic resistances

1259   using combined physics-based and machine learning, *Environ Res Lett*, *18*(3), doi:ARTN 034039

1260   10.1088/1748-9326/acbbe0.

1261   Etienne, T. A., M. Cocaign-Bousquet, and D. Ropers (2020), Competitive effects in bacterial

1262   mRNA decay, *J Theor Biol*, *504*, doi:ARTN 110333 10.1016/j.jtbi.2020.110333.

1263   Exbrayat, J. F., A. J. Pitman, G. Abramowitz, and Y. P. Wang (2013a), Sensitivity of net

1264   ecosystem exchange and heterotrophic respiration to parameterization uncertainty, *J Geophys*

1265   *Res-Atmos*, *118*(4), 1640-1651, doi:10.1029/2012jd018122.

1266   Exbrayat, J. F., A. J. Pitman, Q. Zhang, G. Abramowitz, and Y. P. Wang (2013b), Examining

1267   soil carbon uncertainty in a global model: response of microbial decomposition to temperature,

1268   moisture and nutrient limitation, *Biogeosciences*, *10*(11), 7095-7108, doi:10.5194/bg-10-7095-

1269   2013.

1270   Eyring, H. (1935), The activated complex and the absolute rate of chemical reactions, *Chem Rev*,

1271   *17*(1), 65-77, doi:10.1021/cr60056a006.

1272   Farquhar, G. D., S. V. Caemmerer, and J. A. Berry (1980), A biochemical-model of

1273   photosynthetic CO2 assimilation in leaves of C3 species, *Planta*, *149*(1), 78-90,

1274   doi:10.1007/Bf00386231.

1275   Feller, W. (1968), *An introduction to probability theory and its applications,*, 3rd ed., 509 pp.,

1276   Wiley.

1277   Feynman, R. P., R. B. Leighton, and M. Sands (2011a), *The Feynman Lectures on Physics, Vol*

1278   *III: Quantum Mechanics*, The New Millennium Edition ed., Basic Books; New Millennium ed.

1279   Edition.

1280   Feynman, R. P., R. B. Leighton, and M. Sands (2011b), *The Feynman Lectures on Physics, Vol.*

1281   *I: Mainly Mechanics, Radiation, and Heat*, The New Millennium Edition ed., Basic Books; New

1282   Millennium ed. Edition.

1283   Finkelstein, A. V., and O. Ptitsyn (2016), *Protein Physics: A Course of Lectures*, 2nd edition ed.,

1284   Academic Press.

1285   Franklin, O., et al. (2020), Organizing principles for vegetation dynamics, *Nat Plants*, *6*(5), 444-

1286   453, doi:10.1038/s41477-020-0655-x.

1287   Franzluebbers, A. J. (1999), Microbial activity in response to water-filled pore space of variably

1288   eroded southern Piedmont soils, *Appl Soil Ecol*, *11*(1), 91-101, doi:10.1016/S0929-

1289   1393(98)00128-0.

1290   Ghosh, K., and K. A. Dill (2009), Computing protein stabilities from their chain lengths, *P Natl*

1291   *Acad Sci USA*, *106*(26), 10649-10654, doi:10.1073/pnas.0903995106.

1292   Ginzburg, L. R., and H. R. Akcakaya (1992), Consequences of ratio-dependent predation for

1293   steady-state properties of ecosystems, *Ecology*, *73*(5), 1536-1543, doi:10.2307/1940006.

1294   Goodfellow, I., Y. Bengio, and A. Courville (2016), Regularization for deep learning, *Adapt*

1295   *Comput Mach Le*, 221-265.

1296   Grant, R. F. (1998), Simulation in ecosys of root growth response to contrasting soil water and

1297   nitrogen, *Ecol Model*, *107*(2-3), 237-264, doi:10.1016/S0304-3800(97)00221-4.

1298   Grant, R. F. (2013), Modelling changes in nitrogen cycling to sustain increases in forest

1299   productivity under elevated atmospheric CO2 and contrasting site conditions, *Biogeosciences*,

1300   *10*(11), 7703-7721, doi:10.5194/bg-10-7703-2013.

1301   Grant, R. F., M. Dyck, and D. Puurveen (2020), Nitrogen and phosphorus control carbon

1302   sequestration in agricultural ecosystems: modelling carbon, nitrogen, and phosphorus balances at

1303   the Breton Plots with ecosys under historical and future climates, *Can J Soil Sci*, *100*(4), 408-

1304   429, doi:10.1139/cjss-2019-0132.

1305   Grant, R. F., Z. A. Mekonnen, W. J. Riley, B. Arora, and M. S. Torn (2017), Mathematical

1306   Modelling of Arctic Polygonal Tundra with Ecosys: 2. Microtopography Determines How CO2

1307   and CH4 Exchange Responds to Changes in Temperature and Precipitation, *J Geophys Res-*

1308   *Biogeo*, *122*(12), 3174-3187, doi:10.1002/2017jg004037.

1309   Greenway, H. M., R (1980), Mechanisms of salt tolerance in nonhalophytes, *Annu. Rev. Plant*

1310   *Physiol.*, *31*(1), 149-190, doi:10.1146/annurev.pp.31.060180.001053.

1311   Guan, K., et al. (2023), A scalable framework for quantifying field-level agricultural carbon

1312   outcomes,, *Earth-Science Reviews,*, *243*, doi:https://doi.org/10.1016/j.earscirev.2023.104462.

1313   Hanway, J. (1966), How a corn plant develops*Rep.*, 1–18 pp.

1314   Harfenmeister, K., S. Itzerott, C. Weltzien, and D. Spengler (2021), Detecting phenological

1315   development of winter wheat and winter barley using time series of sentinel-1 and sentinel-2,

1316   *Remote Sens-Basel*, *13*(24), doi:ARTN 503610.3390/rs13245036.

1317   Harrison, S. P., et al. (2021), Eco-evolutionary optimality as a means to improve vegetation and

1318   land-surface models, *New Phytol*, *231*(6), 2125-2141, doi:10.1111/nph.17558.

1319   Higgins, S. I. (2017), Ecosystem assembly: a mission for terrestrial earth system science,

1320   *Ecosystems*, *20*(1), 69-77, doi:10.1007/s10021-016-0054-3.

1321   Hobbs, J. K., W. T. Jiao, A. D. Easter, E. J. Parker, L. A. Schipper, and V. L. Arcus (2013),

1322   Change in heat capacity for enzyme catalysis determines temperature dependence of enzyme

1323   catalyzed rates, *Acs Chem Biol*, *8*(11), 2388-2393, doi:10.1021/cb4005029.

1324   Hoem, J. M. (1987), Statistical-analysis of a multiplicative model and its application to the

1325   standardization of vital-rates - a review, *Int Stat Rev*, *55*(2), 119-152, doi:10.2307/1403190.

1326   Holling, C. S. (1959), Some characteristics of simple types of predation and parasitism, *The*

1327   *Canadian Entomologist*, *XCI*(7), 385-398.

1328   Holmes, M. J. P., N G; Povey, M J W (2011), Temperature dependence of bulk viscosity in

1329   water using acoustic spectroscopy, *J Phys-Condens Mat*, *269*, doi:10.1088/1742-

1330   6596/269/1/012011.

1331   Horn, F., and R. Jackson (1972), General mass action kinetics, *Arch Ration Mech An*, *47*(2), 81-

1332   116.

1333   Houska, T., D. Kraus, R. Kiese, and L. Breuer (2017), Constraining a complex biogeochemical

1334   model for CO2 and N2O emission simulations from various land uses by model-data fusion,

1335   *Biogeosciences*, *14*(14), 3487-3508, doi:10.5194/bg-14-3487-2017.

1336   Huang, Y., B. Guenet, P. Ciais, I. A. Janssens, J. L. Soong, Y. L. Wang, D. Goll, E.

1337   Blagodatskaya, and Y. Y. Huang (2018), ORCHIMIC (v1.0), a microbe-mediated model for soil

1338   organic matter decomposition, *Geosci Model Dev*, *11*(6), 2111-2138, doi:10.5194/gmd-11-2111-

1339   2018.

1340   Huntzinger, D. N., et al. (2017), Uncertainty in the response of terrestrial carbon sink to

1341   environmental drivers undermines carbon-climate feedback predictions, *Sci Rep-Uk*, *7*,

1342   doi:ARTN 4765 10.1038/s41598-017-03818-2.

1343   Hussain, A. J., A. Al-Fayadh, and N. Radi (2018), Image compression techniques: A survey in

1344   lossless and lossy algorithms, *Neurocomputing*, *300*, 44-69, doi:10.1016/j.neucom.2018.02.094.

1345   Jarvis, P. G. (1976), The interpretation of the variations in leaf water potential and stomatal

1346   conductance found in canopies in the field, *Philos. Trans. R. Soc. Lond.*, *273*(927), 593-610,

1347   doi:10.1098/rstb.1976.0035.

1348   Jasak, H. (2009), OpenFOAM: Open source CFD in research and industry, *Int J Nav Arch*

1349   *Ocean*, *1*(2), 89-94, doi:10.3744/Jnaoe.2009.1.2.089.

1350   Jaynes, E. T. (2003), *Probability theory: The logic of science*, Cambridge University Press.

1351   Jin, Q. S., and C. M. Bethke (2003), A new rate law describing microbial respiration, *Appl*

1352   *Environ Microb*, *69*(4), 2340-2348, doi:10.1128/Aem.69.4.2340-2348.2003.

1353   Kang, M. Z., E. Heuvelink, S. M. P. Carvalho, and P. de Reffye (2012), A virtual plant that

1354   responds to the environment like a real one: the case for chrysanthemum, *New Phytol*, *195*(2),

1355   384-395, doi:10.1111/j.1469-8137.2012.04177.x.

1356   Kattge, J., and B. Sandel (2020), TRY plant trait database - Enhanced coverage and open access

1357   (vol 2020, pg 119, 2020), *Global Change Biol*, *26*(9), 5343-5343, doi:10.1111/gcb.15122.

1358   Kearney, M. R., A. P. Isaac, and W. P. Porter (2014), Microclim: Global estimates of hourly

1359   microclimate based on long-term monthly climate averages, *Sci Data*, *1*, doi:ARTN 140006

1360   10.1038/sdata.2014.6.

1361   Keenan, T. F., et al. (2012), Terrestrial biosphere model performance for inter-annual variability

1362   of land-atmosphere CO_2 exchange, *Global Change Biol*, *18*(6), 1971-1987, doi:10.1111/j.1365-

1363   2486.2012.02678.x.

1364   Kim, D. H., P. J. Lynett, and S. A. Socolofsky (2009), A depth-integrated model for weakly

1365   dispersive, turbulent, and rotational fluid flows, *Ocean Model*, *27*(3-4), 198-214,

1366   doi:10.1016/j.ocemod.2009.01.005.

1367   Klausmeier, C. A., E. Litchman, and S. A. Levin (2007), A model of flexible uptake of two

1368   essential resources, *J Theor Biol*, *246*(2), 278-289, doi:10.1016/j.jtbi.2006.12.032.

1369   Kleber, M. L., Adam (2022), *The science and semantics of "soil organic matter stabilization"*,

1370   doi:10.1002/9781119480419.ch2.

1371   Klotz, I. M. R., Robert M (2008), *Chemical Thermodynamics: Basic Concepts and Methods*,

1372   Wiley.

1373   Kondo, J., and N. Saigusa (1994), Modeling the evaporation from bare soil with a formula for

1374   vaporization in the soil pores, *J Meteorol Soc Jpn*, *72*(3), 413-421,

1375   doi:10.2151/jmsj1965.72.3_413.

1376   Kooijman, S. A. L. M. (2009), *Dynamic Energy Budget Theory for Metabolic Organisation*,

1377   Cambridge University Press, Cambridge.

1378   Koonin, E. V., and Y. I. Wolf (2012), Evolution of microbes and viruses: a paradigm shift in

1379   evolutionary biology?, *Front Cell Infect Mi*, *2*, doi:ARTN 119 10.3389/fcimb.2012.00119.

1380   Kotsuki, S., K. Terasaki, H. Yashiro, H. Tomita, M. Satoh, and T. Miyoshi (2018), Online model

1381   parameter estimation with ensemble data assimilation in the Real global atmosphere: A case with

1382   the nonhydrostatic Icosahedral atmospheric model (NICAM) and the global satellite mapping of

1383   precipitation data, *J Geophys Res-Atmos*, *123*(14), 7375-7392, doi:10.1029/2017jd028092.

1384   Koudriavstev, A. B., R. F. Jameson, and W. Linert (2001), *The law of mass action*, Springer,

1385   Berlin Heidelberg, doi:10.1007/978-3-642-56770-4.

1386   Koven, C. D., W. J. Riley, Z. M. Subin, J. Y. Tang, M. S. Torn, W. D. Collins, G. B. Bonan, D.

1387   M. Lawrence, and S. C. Swenson (2013), The effect of vertically resolved soil biogeochemistry

1388   and alternate soil C and N models on C dynamics of CLM4, *Biogeosciences*, *10*(11), 7109-7131,

1389   doi:10.5194/bg-10-7109-2013.

1390   Koyen, C. D., et al. (2020), Benchmarking and parameter sensitivity of physiological and

1391   vegetation dynamics using the Functionally Assembled Terrestrial Ecosystem Simulator

1392   (FATES) at Barro Colorado Island, Panama, *Biogeosciences*, *17*(11), 3017-3044,

1393   doi:10.5194/bg-17-3017-2020.

1394   Kurz, M., P. Offenhauser, and A. Beck (2023), Deep reinforcement learning for turbulence

1395   modeling in large eddy simulations, *Int J Heat Fluid Fl*, *99*, doi:ARTN 109094

1396   10.1016/j.ijheatfluidflow.2022.109094.

1397   Lam, K. N., P. Spanogiannopoulos, P. Soto-Perez, M. Alexander, M. J. Nalley, J. E. Bisanz, R.

1398   R. Nayak, A. M. Weakley, F. Q. B. Yu, and P. J. Turnbaugh (2021), Phage-delivered CRISPR-

1399   Cas9 for strain-specific depletion and genomic deletions in the gut microbiome, *Cell Rep*, *37*(5),

1400   doi:ARTN 109930 10.1016/j.celrep.2021.109930.

1401   Le Noe, J., et al. (2023), Soil organic carbon models need independent time-series validation for

1402   reliable prediction, *Commun Earth Environ*, *4*(1), doi:ARTN 158 10.1038/s43247-023-00830-5.

1403   Lehmann, J., et al. (2020), Persistence of soil organic carbon caused by functional complexity,

1404   *Nat Geosci*, *13*(8), 529-534, doi:10.1038/s41561-020-0612-3.

1405   Leibold, M. A., J. M. Chase, and S. K. M. Ernest (2017), Community assembly and the

1406   functioning of ecosystems: how metacommunity processes alter ecosystems attributes, *Ecology*,

1407   *98*(4), 909-919, doi:10.1002/ecy.1697.

1408   Lenton, T. M., et al. (2023), Quantifying the human cost of global warming, *Nat Sustain*,

1409   doi:10.1038/s41893-023-01132-6.

1410   Leon, J. A., and D. B. Tumpson (1975), Competition between 2 Species for 2 Complementary or

1411   Substitutable Resources, *J Theor Biol*, *50*(1), 185-201, doi:10.1016/0022-5193(75)90032-6.

1412   Leuning, R. (1990), Modeling stomatal behavior and photosynthesis of eucalyptus-grandis, *Aust*

1413   *J Plant Physiol*, *17*(2), 159-175, doi:10.1071/Pp9900159.

1414   Leuning, R. (1995), A critical-appraisal of a combined stomatal-photosynthesis model for C-3

1415   plants, *Plant Cell Environ*, *18*(4), 339-355, doi:10.1111/j.1365-3040.1995.tb00370.x.

1416   Levin, S. A. (1992), The problem of pattern and scale in ecology, *Ecology*, *73*(6), 1943-1967,

1417   doi:Doi 10.2307/1941447.

1418   Li, C. S., J. Aber, F. Stange, K. Butterbach-Bahl, and H. Papen (2000), A process-oriented model

1419   of N2O and NO emissions from forest soils: 1. Model development, *J Geophys Res-Atmos*,

1420   *105*(D4), 4369-4384, doi:10.1029/1999jd900949.

1421   Liebig, J. v. (1840), *Die organische Chemie in ihrer Anwendung auf Agricultur und Physiologie*

1422   *(Organic chemistry in its applications to agriculture and physiology)*, Friedrich Vieweg und

1423   Sohn Publ. Co., Braunschweig, Germany.

1424   Lieth, H. (1973), Primary production: Terrestrial ecosystems, *Human ecology*, *1*(4),

1425   doi:10.1007/BF01536729.

1426   Lin, S. J., and R. B. Rood (1996), Multidimensional flux-form semi-Lagrangian transport

1427   schemes, *Mon Weather Rev*, *124*(9), 2046-2070, doi:10.1175/1520-

1428   0493(1996)124<2046:Mffslt>2.0.Co;2.

1429   Liu, L. C., et al. (2022), KGML-ag: a modeling framework of knowledge-guided machine

1430   learning to simulate agroecosystems: a case study of estimating N2O emission using data from

1431   mesocosm experiments, *Geosci Model Dev*, *15*(7), 2839-2858, doi:10.5194/gmd-15-2839-2022.

1432   Liu, Y. (2007), Overview of some theoretical approaches for derivation of the Monod equation,

1433   *Appl Microbiol Biot*, *73*(6), 1241-1250, doi:10.1007/s00253-006-0717-7.

1434   Liu, Z. G., J. L. Consalvi, and W. J. Kong (2019), An exponential integrator with Schur-Krylov

1435   approximation to accelerate combustion chemistry computation, *Combust Flame*, *203*, 180-189,

1436   doi:10.1016/j.combustflame.2019.01.031.

1437   Luo, Z., E. Wang, H. Zheng, J. A. Baldock, O. J. Sun, and Q. Shao (2015), Convergent

1438   modelling of past soil organic carbon stocks but divergent projections, *Biogeosciences*, *12*(14),

1439   4373-4383, doi:10.5194/bg-12-4373-2015.

1440   Ma, L., et al. (2022), Global evaluation of the Ecosystem Demography model (ED v3.0), *Geosci*

1441   *Model Dev*, *15*(5), 1971-1994, doi:10.5194/gmd-15-1971-2022.

1442   Madigan, M. T., J. M. Martinko, P. V. Dunlap, and D. P. Clark (2009), *Brock biology of*

1443   *microorganisms, twelfth edition*, Pearson Eduction, Inc. , 1301 Sansome Street, San Francisco,

1444   CA 94111.

1445   Maggi, F., C. Gu, W. J. Riley, G. M. Hornberger, R. T. Venterea, T. Xu, N. Spycher, C. Steefel,

1446   N. L. Miller, and C. M. Oldenburg (2008), A mechanistic treatment of the dominant soil nitrogen

1447   cycling processes: Model development, testing, and application, *J Geophys Res-Biogeo*,

1448   *113*(G2), doi:Artn G02016 10.1029/2007jg000578.

1449   Martiny, J. B. H., A. C. Martiny, E. Brodie, A. B. Chase, A. Rodriguez-Verdugo, K. K. Treseder,

1450   and S. D. Allison (2023), Investigating the eco-evolutionary response of microbiomes to

1451   environmental change, *Ecol Lett*, doi:10.1111/ele.14209.

1452   Medlyn, B. E., R. A. Duursma, D. Eamus, D. S. Ellsworth, I. C. Prentice, C. V. M. Barton, K. Y.

1453   Crous, P. de Angelis, M. Freeman, and L. Wingate (2011), Reconciling the optimal and

1454   empirical approaches to modelling stomatal conductance, *Global Change Biol*, *17*(6), 2134-

1455   2144, doi:10.1111/j.1365-2486.2010.02375.x.

1456   Medvigy, D., G. S. Wang, Q. Zhu, W. J. Riley, A. M. Trierweiler, B. G. Waring, X. T. Xu, and J.

1457   S. Powers (2019), Observed variation in soil properties can drive large variation in modelled

1458   forest functioning and composition during tropical forest secondary succession, *New Phytol*,

1459   *223*(4), 1820-1833, doi:10.1111/nph.15848.

1460   Meidner, H. M., Terence Arthur; (1968), *Physiology of stomata*, McGraw-Hill.

1461   Michaelis, L., and M. L. Menten (1913), The kinetics of the inversion effect, *Biochem. Z.*, *49*,

1462   333-369.

1463   Miller, T. D. (1992), Growth Stages of Wheat: Identification and Understanding Improve Crop

1464   Management., *Better Crops*, *76*, 12–17.

1465   Milly, P. C. D. (1982), Moisture and heat-transport in hysteretic, inhomogeneous porous-media -

1466   a matric head-based formulation and a numerical-model, *Water Resour Res*, *18*(3), 489-498.

1467   Moldrup, P., T. Olesen, T. Komatsu, S. Yoshikawa, P. Schjonning, and D. E. Rolston (2003),

1468   Modeling diffusion and reaction in soils: X. A unifying model for solute and gas diffusivity in

1469   unsaturated soil, *Soil Sci*, *168*(5), 321-337, doi:10.1097/00010694-200305000-00002.

1470   Molins, S., and P. Knabner (2019), Multiscale approaches in reactive transport modeling, *Rev*

1471   *Mineral Geochem*, *85*, 27-48, doi:10.2138/rmg.2019.85.2.

1472   Murdoch, W. W. (1973), Functional response of predators, *J Appl Ecol*, *10*(1), 335-342.

1473   Murphy, K. P., P. L. Privalov, and S. J. Gill (1990), Common features of protein unfolding and

1474   dissolution of hydrophobic compounds, *Science*, *247*(4942), 559-561,

1475   doi:10.1126/science.2300815.

1476   Noe, S. M., and C. Giersch (2004), A simple dynamic model of photosynthesis in oak leaves:

1477   coupling leaf conductance and photosynthetic carbon fixation by a variable intracellular CO_2

1478   pool, *Funct Plant Biol*, *31*(12), 1195-1204, doi:10.1071/Fp03251.

1479   Obersteiner, M., H. Bottcher, and Y. Yamagata (2010), Terrestrial ecosystem management for

1480   climate change mitigation, *Curr Opin Env Sust*, *2*(4), 271-276, doi:10.1016/j.cosust.2010.05.006.

1481   Pan, D. (1995), A tutorial on mpeg audio compression, *Ieee Multimedia*, *2*(2), 60-74,

1482   doi:10.1109/93.388209.

1483   Pan, M., and E. F. Wood (2006), Data assimilation for estimating the terrestrial water budget

1484   using a constrained ensemble Kalman filter, *J Hydrometeorol*, *7*(3), 534-547, doi:Doi

1485   10.1175/Jhm495.1.

1486   Parks, H. V., and J. E. Faller (2010), Simple pendulum determination of the gravitational

1487   constant, *Phys Rev Lett*, *105*(11), doi:ARTN 110801 10.1103/PhysRevLett.105.110801.

1488   Pasut, C., F. H. M. Tang, and F. Maggi (2020), A mechanistic analysis of wetland

1489   biogeochemistry in response to temperature, vegetation, and nutrient input changes, *J Geophys*

1490   *Res-Biogeo*, *125*(4), doi:ARTN e2019JG005437 10.1029/2019JG005437.

1491   Pasut, C., F. H. M. Tang, B. Minasny, C. R. Warren, F. A. Dijkstra, W. J. Riley, and F. Maggi

1492   (2023), Seasonal biotic processes vary the carbon turnover by up to one order of magnitude in

1493   wetlands, *Global Biogeochem Cy*, *37*(5), doi:ARTN e2022GB007679 10.1029/2022GB007679.

1494   Pauli, W. (1973), *Statistical Mechanics*, Dover Publications, Inc.

1495   Peterson, M. E., R. Eisenthal, M. J. Danson, A. Spence, and R. M. Daniel (2004), A new

1496   intrinsic thermal parameter for enzymes reveals true temperature optima, *J Biol Chem*, *279*(20),

1497   20717-20722, doi:10.1074/jbc.M309143200.

1498   Pirt, S. J. (1982), Maintenance energy - a general-model for energy-limited and energy-sufficient

1499   growth, *Arch Microbiol*, *133*(4), 300-302, doi:Doi 10.1007/Bf00521294.

1500   Plawsky, J. L. (2020), *Transport Phenomena Fundamentals*, CRC Press.

1501   Powell, N., A. N. Shilton, S. Pratt, and Y. Chisti (2008), Factors influencing luxury uptake of

1502   phosphorus by microalgae in waste stabilization ponds, *Environ Sci Technol*, *42*(16), 5958-5962,

1503   doi:10.1021/es703118s.

1504   Poyser, M. A.-A., Amir; Breckon, Toby P (2021), On the impact of lossy image and video

1505   compression on the performance of deep convolutional neural network architectures, in *2020*

1506   *25th International Conference on Pattern Recognition (ICPR)*, edited,

1507   doi:10.1109/ICPR48806.2021.9412455.

1508   Pretty, J., and Z. P. Bharucha (2014), Sustainable intensification in agricultural systems, *Ann*

1509   *Bot-London*, *114*(8), 1571-1596, doi:10.1093/aob/mcu205.

1510   Pryor, R. W. (2009), *Multiphysics Modeling Using COMSOL®: A First Principles Approach*,

1511   Jones & Bartlett Publishers.

1512   Purcell, E. M. (1977), Life at low Reynolds-number, *Am J Phys*, *45*(1), 3-11,

1513   doi:10.1119/1.10903.

1514   Ratkowsky, D. A., J. Olley, and T. Ross (2005), Unifying temperature effects on the growth rate

1515   of bacteria and the stability of globular proteins, *J Theor Biol*, *233*(3), 351-362,

1516   doi:10.1016/j.jtbi.2004.10.016.

1517   Raynaud, X., and N. Nunan (2014), Spatial ecology of bacteria at the microscale in soil, *Plos*

1518   *One*, *9*(1), doi:ARTN e87217 10.1371/journal.pone.0087217.

1519   Real, L. A. (1977), The kinetics of functional response, *Am. Nat.*, *111*(978), 289-300,

1520   doi:10.1086/283161.

1521   Reinhardt, C., B. Volker, H. J. Martin, J. Kneiseler, and G. F. Fuhrmann (1997), Different

1522   activation energies in glucose uptake in Saccharomyces cerevisiae DFY1 suggest two transport

1523   systems, *Bba-Biomembranes*, *1325*(1), 126-134, doi:10.1016/S0005-2736(96)00252-0.

1524   Riley, W. J., F. Maggi, M. Kleber, M. S. Torn, J. Y. Tang, D. Dwivedi, and N. Guerry (2014),

1525   Long residence times of rapidly decomposable soil organic matter: application of a multi-phase,

1526   multi-component, and vertically resolved model (BAMS1) to soil carbon dynamics, *Geosci*

1527   *Model Dev*, *7*(4), 1335-1355, doi:10.5194/gmd-7-1335-2014.

1528 Riley, W. J., C. Sierra, J. Y. Tang, N. J. Bouskill, Q. Zhu, and R. Abramoff (2022), Next

1529 generation soil biogeochemistry model representations: A proposed community open source

1530 model farm (BeTR-S) in *Multi-scale Biogeochemical Processes in Soil Ecosystems: Critical*

1531 *Reactions and Resilience to Climate Changes*, edited by Y. Yang, M. Keiluweit, N. Senesi and

1532 B. Xing.

1533 Riley, W. J., Z. M. Subin, D. M. Lawrence, S. C. Swenson, M. S. Torn, L. Meng, N. M.

1534 Mahowald, and P. Hess (2011), Barriers to predicting changes in global terrestrial methane

1535 fluxes: analyses using CLM4Me, a methane biogeochemistry model integrated in CESM,

1536 *Biogeosciences*, *8*(7), 1925-1953, doi:10.5194/bg-8-1925-2011.

1537 Robinne, F. N., D. W. Hallema, K. D. Bladon, and J. M. Buttle (2020), Wildfire impacts on

1538 hydrologic ecosystem services in North American high-latitude forests: A scoping review, *J*

1539 *Hydrol*, *581*, doi:ARTN 124360 10.1016/j.jhydrol.2019.124360.

1540 Rollins, G. C., and K. A. Dill (2014), General mechanism of two-state protein folding kinetics, *J*

1541 *Am Chem Soc*, *136*(32), 11420-11427, doi:10.1021/ja5049434.

1542 Ross, J. (1981), *The Radiation Regime and Architecture of Plant Stands*, Dr. W. Junk Publishers,

1543 Norwell, MA.

1544 Rubio, G., J. M. Zhu, and J. P. Lynch (2003), A critical test of the two prevailing theories of

1545 plant response to nutrient availability, *Am J Bot*, *90*(1), 143-152, doi:10.3732/ajb.90.1.143.

1546 Russo, S. E., G. Ledder, E. B. Muller, and R. M. Nisbet (2022), Dynamic Energy Budget

1547 models: fertile ground for understanding resource allocation in plants in a changing world,

1548 *Conserv Physiol*, *10*(1), doi:ARTN coac061 10.1093/conphys/coac061.

1549 Rusu, T. P., Ioan; Dirja, Marcel; Mihai Pacurar, Horea; Oroian, Ioan; Adina Cosma, Smaranda;

1550 Gheres, Marinela (2013), Effect of tillage systems on soil properties, humus and water

1551 conservation, *Agric. Sci. China*, *04*(05), 35-40, doi:10.4236/as.2013.45b007.

1552 Saito, H., J. Simunek, and B. P. Mohanty (2006), Numerical analysis of coupled water, vapor,

1553 and heat transport in the vadose zone, *Vadose Zone J*, *5*(2), 784-800, doi:10.2136/vzj2006.0007.

1554 Sander, R. (2015), Compilation of Henry's law constants (version 4.0) for water as solvent,

1555 *Atmos Chem Phys*, *15*(8), 4399-4981, doi:10.5194/acp-15-4399-2015.

1556 Santos, A., H. F. A. Scanavini, H. Pedrini, D. J. Schiozer, F. P. Munerato, and C. E. A. G.

1557 Barreto (2022), An artificial intelligence method for improving upscaling in complex reservoirs,

1558 *J Petrol Sci Eng*, *211*, doi:ARTN 110071 10.1016/j.petrol.2021.110071.

1559 Saredi, E., N. T. Ramesh, A. Sciacchitano, and F. Scarano (2021), State observer data

1560 assimilation for RANS with time-averaged 3D-PIV data, *Comput Fluids*, *218*, doi:ARTN 104827

1561 10.1016/j.compfluid.2020.104827.

1562 Saurabh, S. (2021), Genome editing: revolutionizing the crop improvement, *Plant Mol Biol Rep*,

1563 *39*(4), 752-772, doi:10.1007/s11105-021-01286-7.

1564 Sawle, L., and K. Ghosh (2011), How do thermophilic proteins and proteomes withstand high

1565 temperature?, *Biophys J*, *101*(1), 217-227, doi:10.1016/j.bpj.2011.05.059.

1566 Seok, M. G., W. T. Cai, and D. Park (2021), Hierarchical aggregation/disaggregation for

1567 adaptive abstraction-level conversion in digital twin-based smart semiconductor manufacturing,

1568 *Ieee Access*, *9*, 71145-71158, doi:10.1109/Access.2021.3073618.

1569 Sharpe, P. J. H., and D. W. Demichele (1977), Reaction-Kinetics of Poikilotherm Development,

1570 *J Theor Biol*, *64*(4), 649-670, doi:10.1016/0022-5193(77)90265-X.

1571   Shinozaki, K. Y., K.; Hozumi, K.; Kira, T. (1964), A quantitative analysis of plant form-the pipe

1572   model theory: I. Basic analyses, *Jap. J. Ecol.*, *14*, 97-105,

1573   doi:https://doi.org/10.18960/seitai.14.3_97.

1574   Shivni, R. (2016), The deconstructed Standard Model equation, edited.

1575   Shuttleworth, W. J., and J. S. Wallace (1985), Evaporation from sparse crops - an energy

1576   combination theory, *Q J Roy Meteor Soc*, *111*(469), 839-855, doi:10.1256/smsqj.46909.

1577   Sierra, C. A., S. E. Trumbore, E. A. Davidson, S. Vicca, and I. Janssens (2015), Sensitivity of

1578   decomposition rates of soil organic matter with respect to simultaneous changes in temperature

1579   and moisture, *J Adv Model Earth Sy*, *7*(1), 335-356, doi:10.1002/2014ms000358.

1580   Silverstein, T. P. (2020), The hydrophobic effect: is water afraid, or just not that interested?,

1581   *Chemtexts*, *6*(4), doi:ARTN 26 10.1007/s40828-020-00117-8.

1582   Simunek, J., and D. L. Suarez (1993), Modeling of carbon-dioxide transport and production in

1583   soil .1. model development, *Water Resour Res*, *29*(2), 487-497, doi:10.1029/92wr02225.

1584   Sollins, P., and J. W. Gregg (2017), Soil organic matter accumulation in relation to changing soil

1585   volume, mass, and structure: Concepts and calculations, *Geoderma*, *301*, 60-71,

1586   doi:10.1016/j.geoderma.2017.04.013.

1587   Sprengel, C. (1826), Ueber Pflanzenhumus, Humussaure und humussaure Salze (About palnt

1588   humus, humic acids and salts of humic acids), *Archiv fur die Gesammte Naturlehre*, *8*, 145-220.

1589   Steefel, C. I., D. J. DePaolo, and P. C. Lichtner (2005), Reactive transport modeling: An

1590   essential tool and a new research approach for the Earth sciences, *Earth Planet Sc Lett*, *240*(3-4),

1591   539-558, doi:10.1016/j.epsl.2005.09.017.

1592   Steefel, C. I., and A. C. Lasaga (1994), A coupled model for transport of multiple chemical-

1593   species and kinetic precipitation dissolution reactions with application to reactive flow in single-

1594   phase hydrothermal systems, *Am J Sci*, *294*(5), 529-592, doi:10.2475/ajs.294.5.529.

1595   Steefel, C. I., S. B. Yabusaki, and K. U. Mayer (2015), Reactive transport benchmarks for

1596   subsurface environmental simulation, *Computat Geosci*, *19*(3), 439-443, doi:10.1007/s10596-

1597   015-9499-2.

1598   Stein, W. (2012), *Transport And Diffusion Across Cell Membranes*, Elsevier.

1599   Steudle, E., U. Zimmermann, and U. Luttge (1977), Effect of turgor pressure and cell-size on

1600   wall elasticity of plant-cells, *Plant Physiol*, *59*(2), 285-289, doi: 10.1104/pp.59.2.285.

1601   Suseela, V., R. T. Conant, M. D. Wallenstein, and J. S. Dukes (2012), Effects of soil moisture on

1602   the temperature sensitivity of heterotrophic respiration vary seasonally in an old-field climate

1603   change experiment, *Global Change Biol*, *18*(1), 336-348, doi:10.1111/j.1365-

1604   2486.2011.02516.x.

1605   Sweetlove, L. J., and R. G. Ratcliffe (2011), Flux-balance modeling of plant metabolism, *Front*

1606   *Plant Sci*, *2*, doi:ARTN 38 10.3389/fpls.2011.00038.

1607   Taiz, L., and E. Zeiger (2006), *Plan physiology, 4th edn*, Sinauer Associates, Inc., Sunderland,

1608   MA 01375 USA.

1609   Tan, Y.-S. Z., Ren-Kuan; Liu, Zhi-Hua; Li, Bing-Zhi; Yuan, Ying-Jin (2022), Microbial

1610   adaptation to enhance stress tolerance, *Front. Microbiol.*, *13*, doi:10.3389/fmicb.2022.888746.

1611   Tang, J., Q. Zhuang, R. D. Shannon, and J. R. White (2010), Quantifying wetland methane

1612   emissions with process-based models of different complexities, *Biogeosciences*, *7*(11), 3817-

1613   3837, doi:10.5194/bg-7-3817-2010.

1614   Tang, J. Y. (2015), On the relationships between the Michaelis-Menten kinetics, reverse

1615   Michaelis-Menten kinetics, equilibrium chemistry approximation kinetics, and quadratic

1616   kinetics, *Geosci Model Dev*, *8*(12), 3823-3835, doi:10.5194/gmd-8-3823-2015.

1617   Tang, J. Y., and W. J. Riley (2013), A total quasi-steady-state formulation of substrate uptake

1618   kinetics in complex networks and an example application to microbial litter decomposition,

1619   *Biogeosciences*, *10*(12), 8329-8351, doi:10.5194/bg-10-8329-2013.

1620   Tang, J. Y., and W. J. Riley (2015), Weaker soil carbon-climate feedbacks resulting from

1621   microbial and abiotic interactions, *Nat Clim Change*, *5*(1), 56-60, doi:10.1038/Nclimate2438.

1622   Tang, J. Y., and W. J. Riley (2017), SUPECA kinetics for scaling redox reactions in networks of

1623   mixed substrates and consumers and an example application to aerobic soil respiration, *Geosci*

1624   *Model Dev*, *10*(9), 3277-3295, doi:10.5194/gmd-10-3277-2017.

1625   Tang, J. Y., and W. J. Riley (2018), Predicted land carbon dynamics are strongly dependent on

1626   the numerical coupling of nitrogen mobilizing and immobilizing processes: a demonstration with

1627   the E3SM land model, *Earth Interact*, *22*(11), doi:ARTN 11 10.1175/EI-D-17-0023.s1.

1628   Tang, J. Y., and W. J. Riley (2019a), Competitor and substrate sizes and diffusion together

1629   define enzymatic depolymerization and microbial substrate uptake rates, *Soil Biol Biochem*, *139*,

1630   doi:ARTN 107624 10.1016/j.soilbio.2019.107624.

1631   Tang, J. Y., and W. J. Riley (2019b), A theory of effective microbial substrate affinity

1632   parameters in variably saturated soils and an example application to aerobic soil heterotrophic

1633   respiration, *J Geophys Res-Biogeo*, *124*(4), 918-940, doi:10.1029/2018jg004779.

1634   Tang, J. Y., and W. J. Riley (2021), Finding Liebig's law of the minimum, *Ecological*

1635   *Applications (In press)*, *31*(8), doi:ARTN e02458 10.1002/eap.2458.

1636   Tang, J. Y., and W. J. Riley (2023a), A reanalysis of the foundations of the macromolecular rate

1637   theory, *Biogeosciences Discussions*, doi:10.5194/bg-2023-77.

1638   Tang, J. Y., and W. J. Riley (2023b), Revising the dynamic energy budget theory with a new

1639   reserve mobilization rule and three example applications to bacterial growth, *Soil Biol Biochem*,

1640   *178*, doi:ARTN 108954 10.1016/j.soilbio.2023.108954.

1641   Tang, J. Y., W. J. Riley, G. L. Marschmann, and E. L. Brodie (2021), Conceptualizing

1642   biogeochemical reactions with an Ohm's law analogy, *J Adv Model Earth Sy*, *13*(10), doi:ARTN

1643   e2021MS002469 10.1029/2021MS002469.

1644   Tang, J. Y., W. J. Riley, and J. Niu (2015), Incorporating root hydraulic redistribution in

1645   CLM4.5: Effects on predicted site and global evapotranspiration, soil moisture, and water

1646   storage, *J Adv Model Earth Sy*, *7*(4), 1828-1848, doi:10.1002/2015ms000484.

1647   Tang, J. Y., W. J. Riley, and Q. Zhu (2022), Supporting hierarchical soil biogeochemical

1648   modeling: version 2 of the Biogeochemical Transport and Reaction model (BeTR-v2), *Geosci*

1649   *Model Dev*, *15*(4), 1619-1632, doi:10.5194/gmd-15-1619-2022.

1650   Tang, J. Y., and Q. L. Zhuang (2008), Equifinality in parameterization of process-based

1651   biogeochemistry models: A significant uncertainty source to the estimation of regional carbon

1652   dynamics, *J Geophys Res-Biogeo*, *113*(G4), doi:Artn G04010 10.1029/2008jg000757.

1653   Tang, J. Y., and Q. L. Zhuang (2009), A global sensitivity analysis and Bayesian inference

1654   framework for improving the parameter estimation and prediction of a process-based Terrestrial

1655   Ecosystem Model, *J Geophys Res-Atmos*, *114*, doi:Artn D15303 10.1029/2009jd011724.

1656   Tao, F., et al. (2023), Microbial carbon use efficiency promotes global soil carbon storage,

1657   *Nature*, *618*(7967), 981-985, doi:10.1038/s41586-023-06042-3.

1658   Tarantola, A. (2005), *Inverse problem theory and methods for model parameter estimation*,

1659   Society for Industrial and Applied Mathematics.

1660   Thakkar, A. J. (2021), *Quantum Chemistry*, Third Edition ed., IOP Publishing, Bristol, UK.

1661   Theodoridis, S. (2015), *Machine learning: a bayesian and optimization perspective*, Academic

1662   Press.

1663   Thornley, J. H. M. (1972), A balanced quantitative model for root: shoot ratios in vegetative

1664   plants, *Ann. Bot.*, *36*(2), 431-441, doi:10.1093/oxfordjournals.aob.a084602.

1665   Tietjen, T., and R. G. Wetzel (2003), Extracellular enzyme-clay mineral complexes: Enzyme

1666   adsorption, alteration of enzyme activity, and protection from photodegradation, *Aquat Ecol*,

1667   *37*(4), 331-339, doi:10.1023/B:AECO.0000007044.52801.6b.

1668   Todd-Brown, K. E. O., J. T. Randerson, W. M. Post, F. M. Hoffman, C. Tarnocai, E. A. G.

1669   Schuur, and S. D. Allison (2013), Causes of variation in soil carbon simulations from CMIP5

1670   Earth system models and comparison with observations, *Biogeosciences*, *10*(3), 1717-1736,

1671   doi:10.5194/bg-10-1717-2013.

1672   Tollefson, J. (2022), Carbon emissions hit new high: warning from COP27, *Nature*,

1673   doi:https://doi.org/10.1038/d41586-022-03657-w.

1674   Tsai, W.-P. F., Dapeng; Pan, Ming; Beck, Hylke; Lawson, Kathryn; Yang, Yuan; Liu, Jiangtao;

1675   Shen, Chaopeng (2021), From calibration to parameter learning: Harnessing the scaling effects

1676   of big data in geoscientific modeling, *Nat. Commun.*, *12*(1), doi:10.1038/s41467-021-26107-z.

1677   van de Griend, A. A., and M. Owe (1994), Bare soil surface-resistance to evaporation by vapor

1678   diffusion under semiarid conditions, *Water Resour Res*, *30*(2), 181-188, doi:10.1029/93wr02747.

1679   van Genuchten, M. T. (1980), A closed-form equation for predicting the hydraulic conductivity

1680   of unsaturated soils, *Soil Sci Soc Am J*, *44*(5), 892-898, doi:

1681   10.2136/sssaj1980.03615995004400050002x.

1682   Viskari, T., J. Pusa, I. Fer, A. Repo, J. Vira, and J. Liski (2022), Calibrating the soil organic

1683   carbon model Yasso20 with multiple datasets, *Geosci Model Dev*, *15*(4), 1735-1752,

1684   doi:10.5194/gmd-15-1735-2022.

1685   von Caemmerer, S. (2013), Steady-state models of photosynthesis, *Plant Cell Environ*, *36*(9),

1686   1617-1630, doi:10.1111/pce.12098.

1687   von Smoluchowski, M. (1917), Versuch einer mathematischen theorie der koagulationkinetik

1688   kolloider loesungen., *Z Phys Chem*, *92*, 129–132.

1689   Vrugt, J. A. (2016), Markov chain Monte Carlo simulation using the DREAM software package:

1690   Theory, concepts, and MATLAB implementation, *Environ Modell Softw*, *75*, 273-316,

1691   doi:10.1016/j.envsoft.2015.08.013.

1692   Walker, A. P., A. L. Johnson, A. Rogers, J. Anderson, R. A. Bridges, R. A. Fisher, D. Lu, D. M.

1693   Ricciuto, S. P. Serbin, and M. Ye (2021), Multi-hypothesis comparison of Farquhar and Collatz

1694   photosynthesis models reveals the unexpected influence of empirical assumptions at leaf and

1695   global scales, *Global Change Biol*, *27*(4), 804-822, doi:10.1111/gcb.15366.

1696   Wang, G. S., Q. Gao, Y. F. Yang, S. E. Hobbie, P. B. Reich, and J. Z. Zhou (2022), Soil enzymes

1697   as indicators of soil function: A step toward greater realism in microbial ecological modeling,

1698   *Global Change Biol*, *28*(5), 1935-1950, doi:10.1111/gcb.16036.

1699   Wang, G. S., and W. M. Post (2012), A theoretical reassessment of microbial maintenance and

1700   implications for microbial ecology modeling, *Fems Microbiol Ecol*, *81*(3), 610-617,

1701   doi:10.1111/j.1574-6941.2012.01389.x.

1702 Wang, S. S., Y. Yang, A. P. Trishchenko, A. G. Barr, T. A. Black, and H. McCaughey (2009),

1703 Modeling the response of canopy stomatal conductance to humidity, *J Hydrometeorol*, *10*(2),

1704 521-532, doi:10.1175/2008jhm1050.1.

1705 Wang, Y. P., B. C. Chen, W. R. Wieder, M. Leite, B. E. Medlyn, M. Rasmussen, M. J. Smith, F.

1706 B. Agusto, F. Hoffman, and Y. Q. Luo (2014), Oscillatory behavior of two nonlinear microbial

1707 models of soil carbon decomposition, *Biogeosciences*, *11*(7), 1817-1831, doi:10.5194/bg-11-

1708 1817-2014.

1709 Wang, Y. P., R. M. Law, and B. Pak (2010), A global model of carbon, nitrogen and phosphorus

1710 cycles for the terrestrial biosphere, *Biogeosciences*, *7*(7), 2261-2282, doi:10.5194/bg-7-2261-

1711 2010.

1712 Warren, C. R., and S. Manzoni (2023), When dry soil is re-wet, trehalose is respired instead of

1713 supporting microbial growth, *Soil Biol Biochem*, *184*, doi:ARTN 109121

1714 10.1016/j.soilbio.2023.109121.

1715 Weng, E. S., et al. (2022), Modeling demographic-driven vegetation dynamics and ecosystem

1716 biogeochemical cycling in NASA GISS's Earth system model (ModelE-BiomeE v.1.0), *Geosci*

1717 *Model Dev*, *15*(22), 8153-8180, doi:10.5194/gmd-15-8153-2022.

1718 Wermuth, N. (1976), Analogies between Multiplicative Models in Contingency-Tables and

1719 Covariance Selection, *Biometrics*, *32*(1), 95-108, doi:10.2307/2529341.

1720 Wieder, W. R., A. S. Grandy, C. M. Kallenbach, and G. B. Bonan (2014), Integrating microbial

1721 physiology and physio-chemical principles in soils with the MIcrobial-MIneral Carbon

1722 Stabilization (MIMICS) model, *Biogeosciences*, *11*(14), 3899-3917, doi:10.5194/bg-11-3899-

1723 2014.

1724 Wigner, E. P. (1960), The Unreasonable Effectiveness of Mathematics in the Natural Sciences,

1725 *Commun Pur Appl Math*, *13*(1), 1-14, doi:10.1002/cpa.3160130102.

1726 Williams, P. J. L. (1973), The validity of the application of simple kinetic analysis to

1727 heterogeneous microbial populations, *Limnol. Oceanogr.*, *18*(1), 159-165,

1728 doi:10.4319/lo.1973.18.1.0159.

1729 Wilson, C. H., and S. Gerber (2021), Theoretical insights from upscaling Michaelis-Menten

1730 microbial dynamics in biogeochemical models: a dimensionless approach, *Biogeosciences*,

1731 *18*(20), 5669-5679, doi:10.5194/bg-18-5669-2021.

1732 Wober, W., G. Novotny, L. Mehnen, and C. Olaverri-Monreal (2020), Autonomous vehicles:

1733 vehicle parameter estimation using variational Bayes and kinematics, *Appl Sci-Basel*, *10*(18),

1734 doi:ARTN 6317 10.3390/app10186317.

1735 Woo, D. K., W. J. Riley, and Y. X. Wu (2020), More fertilizer and impoverished roots required

1736 for improving wheat yields and profits under climate change, *Field Crop Res*, *249*, doi:ARTN

1737 107756 10.1016/j.fcr.2020.107756.

1738 Xiao, Q. N., Y. H. Kuo, J. Z. Sun, W. C. Lee, D. M. Barker, and E. Lim (2007), An approach of

1739 radar reflectivity data assimilation and its assessment with the inland QPF of Typhoon Rusa

1740 (2002) at landfall, *J Appl Meteorol Clim*, *46*(1), 14-22, doi:10.1175/Jam2439.1.

1741 Xiao, X. Z., Junhui; Yan, Huimin; Wu, Weixing; Biradar, Chandrashekhar (2009), Land surface

1742 phenology: convergence of satellite and CO2 eddy flux observations, in *Phenology of Ecosystem*

1743 *Processes*, edited by N. A., pp. 247-270, Springer New York, New York.

1744 Young, I. M., J. W. Crawford, and C. Rappoldt (2001), New methods and models for

1745 characterising structural heterogeneity of soil, *Soil Till Res*, *61*(1-2), 33-45, doi: 10.1016/S0167-

1746 1987(01)00188-X.

1747   Yu, L. Y., H. J. Cai, Z. Zheng, Z. J. Li, and J. Wang (2017), Towards a more flexible

1748   representation of water stress effects in the nonlinear Jarvis model, *J Integr Agr*, *16*(1), 210-220,

1749   doi:10.1016/S2095-3119(15)61307-7.

1750   Zaehle, S., et al. (2014), Evaluation of 11 terrestrial carbon-nitrogen cycle models against

1751   observations from two temperate Free-Air CO2 Enrichment studies, *New Phytol*, *202*(3), 803-

1752   822, doi:10.1111/nph.12697.

1753   Zhou, G. W., M T; Zhou, G Q (1983), Diffusion-controlled reactions of enzymes. An

1754   approximate analytic solution of Chou's model, *Biophys. Chem.*, *18*(2), 125-132.

1755   Zhou, L. J., L. Harris, J. H. Chen, K. Gao, H. Guo, B. Q. Xiang, M. J. Tong, J. J. Huff, and M.

1756   Morin (2022), Improving global weather prediction in GFDL SHiELD through an upgraded

1757   GFDL cloud microphysics scheme, *J Adv Model Earth Sy*, *14*(7), doi:ARTN e2021MS002971

1758   10.1029/2021MS002971.

1759   Zhou, W., K. Y. Guan, B. Peng, J. Y. Tang, Z. N. Jin, C. Y. Jiang, R. Grant, and S. Mezbahuddin

1760   (2021), Quantifying carbon budget, crop yields and their responses to environmental variability

1761   using the ecosys model for US Midwestern agroecosystems, *Agr Forest Meteorol*, *307*,

1762   doi:ARTN 108521 10.1016/j.agrformet.2021.108521.

1763   Zhou, W. P., D. F. Hui, and W. J. Shen (2014), Effects of soil moisture on the temperature

1764   sensitivity of soil heterotrophic respiration: A laboratory incubation study, *Plos One*, *9*(3),

1765   doi:ARTN e92531 10.1371/journal.pone.0092531.

1766   Zhu, Q., W. J. Riley, J. Tang, and C. D. Koven (2016), Multiple soil nutrient competition

1767   between plants, microbes, and mineral surfaces: model development, parameterization, and

1768   example applications in several tropical forests, *Biogeosciences*, *13*(1), 341-363, doi:10.5194/bg-

1769   13-341-2016.

1770   Zhu, Q., W. J. Riley, J. Y. Tang, N. Collier, F. M. Hoffman, X. J. Yang, and G. Bisht (2019),

1771   Representing nitrogen, phosphorus, and carbon interactions in the E3SM land model:

1772   development and dlobal benchmarking, *J Adv Model Earth Sy*, *11*(7), 2238-2258,

1773   doi:10.1029/2018ms001571.

1774   Zhuang, Q., J. M. Melillo, D. W. Kicklighter, R. G. Prinn, A. D. McGuire, P. A. Steudler, B. S.

1775   Felzer, and S. Hu (2004), Methane fluxes between terrestrial ecosystems and the atmosphere at

1776   northern high latitudes during the past century: A retrospective analysis with a process-based

1777   biogeochemistry model, *Global Biogeochem Cy*, *18*(3), doi:Artn Gb3010

1778   10.1029/2004gb002239.

1779   Zonneveld, C., and S. A. L. M. Kooijman (1989), Application of a Dynamic Energy Budget

1780   Model to Lymnaea-Stagnalis (L), *Funct Ecol*, *3*(3), 269-278, doi:10.2307/2389365.

1781