Circumventing kinetics in biogeochemical modeling

Stilianos Louca[a,b,1], Mary I. Scranton[c], Gordon T. Taylor[c], Yrene M. Astor[d,e], Sean A. Crowe[f,g], and Michael Doebeli[h,i,j]

[a]Institute of Ecology and Evolution, University of Oregon, Eugene, OR 97403; [b]Department of Biology, University of Oregon, Eugene, OR 97403; [c]School of Marine and Atmospheric Sciences, Stony Brook University, Stony Brook, NY 11794; [d]Estación de Investigaciones Marinas de Margarita, Fundación La Salle de Ciencias Naturales, Punta de Piedras, Estado Nueva Esparta, Venezuela; [e]Institute for Marine Remote Sensing, University of South Florida, Tampa, FL 33701; [f]Department of Microbiology and Immunology, University of British Columbia, Vancouver, BC V6T 1Z3, Canada; [g]Department of Earth, Ocean, and Atmospheric Sciences, University of British Columbia, Vancouver, BC V6T 1Z4, Canada; [h]Biodiversity Research Centre, University of British Columbia, Vancouver, BC V6T 1Z4, Canada; [i]Department of Mathematics, University of British Columbia, Vancouver, BC V6T 1Z2, Canada; and [j]Department of Zoology, University of British Columbia, Vancouver, BC V6T 1Z4, Canada

Edited by Katherine H. Freeman, Pennsylvania State University, University Park, PA, and approved April 19, 2019 (received for review November 20, 2018)

Microbial metabolism drives biogeochemical fluxes in virtually every ecosystem. Modeling these fluxes is challenged by the incredible diversity of microorganisms, whose kinetic parameters are largely unknown. In poorly mixed systems, such as stagnant water columns or sediments, however, long-term bulk microbial metabolism may become limited by physical transport rates of substrates across space. Here we mathematically show that under these conditions, biogeochemical fluxes are largely predictable based on the system's transport properties, chemical boundary conditions, and the stoichiometry of metabolic pathways, regardless of the precise kinetics of the resident microorganisms. We formalize these considerations into a predictive modeling framework and demonstrate its use for the Cariaco Basin subeuphotic zone, one of the largest anoxic marine basins worldwide. Using chemical concentration data solely from the upper boundary (depth 180 m) and lower boundary (depth 900 m), but without a priori knowledge of metabolite fluxes, chemical depth profiles, kinetic parameters, or microbial species composition, we predict the concentrations and vertical fluxes of biologically important substances, including oxygen, nitrate, hydrogen sulfide, and ammonium, across the entire considered depth range (180–900 m). Our predictions largely agree with concentration measurements over a period of 14 years (R^2 = 0.78–0.92) and become particularly accurate during a period where the system was near biogeochemical steady state (years 2007–2009, R^2 = 0.86–0.95). Our work enables geobiological predictions for a large class of ecosystems without knowledge of kinetic parameters or geochemical depth profiles. Conceptually, our work provides a possible explanation for the decoupling between microbial species composition and bulk metabolic function, observed in various ecosystems.

geobiology | reaction kinetics | microbial system | redox gradient | marine anoxic region

Understanding the factors that determine microbial metabolic activity at ecosystem scales is essential for deciphering the processes shaping modern ecosystems and for determining Earth's past and future biogeochemical trajectory. Microbial population dynamics and metabolic activity influence and are influenced by abiotic processes, such as the diffusive transport of electron donors and acceptors across redox gradients. This coupling between biotic and abiotic processes is well illustrated in stagnant water columns or sediments, where redox gradients and spatially structured microbial communities develop across depth and result from the interplay between microbial metabolism, population dynamics, and slow physical mixing (1–3). Mechanistic models aiming to predict biogeochemical reaction rates and fluxes typically require knowledge of the kinetic properties of the resident microbial communities, such as bulk rate coefficients, microbial growth rates, or substrate affinities, in addition to the system's physical characteristics (e.g., diffusion coefficients) and boundary conditions (e.g., substrate concentrations at both ends of a water column) (3–8). When kinetic parameters are unknown, as is often the case, additional data such as geochemical depth profiles are needed for model fitting (3, 9, 10). Over several decades, thousands of experiments have been performed to determine reaction- and bio-kinetic parameters of specific microorganisms or specific systems, with the ultimate goal of using these parameters for modeling (11, 12). However, the vast majority of microorganisms have not been, and may never be, kinetically characterized (13, 14). Even if all kinetic/physiological parameters were known for all extant microorganisms, biogeochemical predictions for natural ecosystems would require knowledge of current microbial community composition, and this knowledge is typically lacking. There is thus a need to identify the key mechanisms actually constraining biogeochemical flux rates at ecosystem scales and the minimal set of parameters needed to describe these mechanisms, to construct appropriately streamlined models. As we explain below, basic ecological principles and physical arguments provide a previously unrecognized avenue toward a type of substantially streamlined models applicable to a broad class of ecosystems.

In the presence of ample substrates for energy, microbial population growth driven by the consumption of these substrates typically leads to an acceleration of consumption rates, until at least one essential substrate becomes scarce and limiting. When substrates are supplied at a finite rate and in the presence of competing populations, Tilman's classical ecological theory predicts that, at steady state, the concentration of

Significance

Predicting biochemical processes driven by microbes in the environment remains challenging, because the "kinetic" parameters conventionally used to predict reaction rates are usually poorly known. Here we mathematically show that in poorly mixed systems, such as stagnant waters, bulk biochemical reaction rates can become limited by the slow transport of substrates across space and essentially independent of kinetic parameters. We demonstrate our arguments for a large and heavily studied ocean basin, where we accurately predict the microbially driven fluxes of various substances, across a depth range of hundreds of meters. Our work opens up avenues for predicting ecosystem processes without knowledge of kinetic parameters and without laborious chemical profile measurements.

Author contributions: S.L. designed research; S.L., M.I.S., G.T.T., Y.M.A., S.A.C., and M.D. performed research; S.L. analyzed data; S.L., M.I.S., G.T.T., Y.M.A., S.A.C., and M.D. wrote the paper; and M.I.S., G.T.T., and Y.M.A. organized the Cariaco Time Series program.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. Email: louca.research@gmail.com.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1819883116/-/DCSupplemental.

Published online May 16, 2019.

a limiting substrate will eventually drop to the lowest concentration at which any population could possibly survive (15). In such systems selection for efficient substrate gathering tends to reduce long-term residual concentrations of limiting substrates to a low threshold, below which net population growth becomes impossible (16). As local substrate concentrations approach this threshold, the long-term average bulk consumption rate of a substrate becomes constrained by the rate at which it can be transported across space, from its site of production to its site of consumption. For example, in some anoxic marine regions hydrogen sulfide (H_2S) produced in the underlying sediments slowly diffuses upward, eventually reaching an oxic/anoxic transition zone where it is completely oxidized using oxidants such as oxygen, transported downward from the euphotic zone (17, 18). In these cases, simple diffusion models predict that steady-state H_2S oxidation rates are limited by the maximum possible rate of diffusion from the sediments to the oxic/anoxic transition zone, which in turn depends on the mixing properties of the system and the H_2S concentration at the sediment/water interface. Similar considerations also apply to many other systems with relatively slow mixing rates and pronounced geochemical gradients (19). In these systems, biogeochemical flux rates could become entirely determined (and thus predictable) by the system's transport properties and chemical boundary conditions. While the importance of transport limitation to microbial activity is generally appreciated (20) and geochemical models sometimes assume a separation of time scales between reaction and transport rates (21, 22), these considerations have not been systematically used to streamline conventional geobiological models.

Here we propose a streamlined model framework for reconstructing or predicting steady-state bulk biogeochemical fluxes in microbial systems with pronounced geochemical gradients, when fluxes are limited by transport rates across space. We focus on essentially 1D systems, that is, where all gradients are aligned along a single axis (e.g., depth), such as stagnant marine water columns (3, 23–27), meromictic lakes (28–30), sediments (1, 31, 32), and biofilms (33). As we elaborate below, our "spatial metabolic flux" (SMF) framework differs strongly from conventional approaches in geobiology. Notably, contrary to conventional reaction-transport models, our framework requires only knowledge of chemical concentrations at the system's boundaries (e.g., at the top and bottom of a water column), but no reaction- or bio-kinetic parameters, or any thermodynamic parameters, or any a priori knowledge of metabolite flux rates. The SMF framework also does not require chemical depth profile data, for example for parameter fitting, a common requirement of conventional inverse modeling methods. To showcase the predictive power of the framework, we predict the distribution and fluxes of several biologically important chemical substances (henceforth "metabolites") in the Cariaco Basin subeuphotic water column, one of the largest anoxic marine basins worldwide (34, 35). The model's predictions are largely consistent with chemical depth profiles measured over the course of 14 years, previously estimated metabolite fluxes, and microbiological observations.

A Framework for Spatial Metabolic Flux Analysis

Our spatial metabolic flux framework is based on a minimal set of intuitive but idealizing assumptions, the first three of which are commonly encountered in conventional models:

i) Metabolite sources/sinks: Metabolites can be produced and/or consumed by metabolic reactions (whose rates can differ between locations) and may be transported across the system's boundaries.

ii) Physical metabolite transport: Metabolites are transported across space according to a diffusion and/or advection process. Examples of diffusive transport include molecular diffusion in sediments (36), turbulent mixing (eddy diffusion) in water columns (37), or (in some cases) bioturbation in sediments (38). Examples of advective transport include the sinking of organic particles (39), the settling of minerals in a water column (40), or pore water movement in sediments (41).

iii) Steady state (flux balancing): Metabolite concentrations and reaction rates are near steady state. This implies in particular that metabolite fluxes across space and metabolite production/consumption rates are balanced. Note that in practice slow changes are permitted, as long as any changes in the system's transport properties and boundary conditions occur very slowly compared with the time scales at which microbial communities respond and chemical gradients stabilize.

iv) Efficient use of metabolic niches: At every location, every modeled reaction is limited by the availability of at least one substrate; that is, no metabolic niche is left unutilized (42). Competition for limiting substrates and selection for efficient substrate gathering lead to the concentration of those substrates to approach zero (Fig. 1). While here we focus on biologically driven growth-inducing reactions, this assumption could also apply to rapidly occurring abiotic reactions.

In 1D systems assumptions *i–iii* translate to a standard reaction-transport equation, which describes the spatial distribution of any given metabolite m,

$$0 = \sum_r S_{mr} R_r + \frac{\partial}{\partial z}\left[D_m \cdot \frac{\partial C_m}{\partial z} - \mathrm{v}_m C_m\right], \qquad [1]$$

where C_m (z) is the metabolite's concentration (e.g., mol per volume) at any given depth z, D_m is the applicable diffusion coefficient for the metabolite, r iterates over all reactions, S_{mr} is the stoichiometric coefficient of metabolite m in reaction r, v_m is

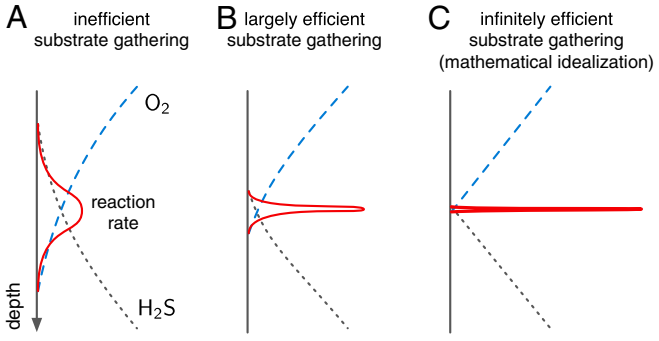

Fig. 1. Sharp chemical transition zones as hotspots. Shown is an illustration of the interpretation of hotspots—a mathematical idealization of chemical transition zones where microorganisms consume coupled substrates diffusing from opposite ends. The aerobic oxidation of hydrogen sulfide (diffusing from the bottom) using oxygen (diffusing from the top) along a stagnant water column is used as an example ($H_2S + 2O_2 \rightarrow SO_4^{2-} + 2H^+$). (*A–C*) Different hypothetical scenarios for the steady-state concentration depth profiles of H_2S and O_2 (dashed curves) and the corresponding reaction rate profile (red solid curves), assuming no other reaction occurs. (*A*) Substrate gathering is rather inefficient, since both substrates co-occur at substantial concentrations. (*B*) Substrate gathering is largely efficient and hence the reaction is confined to a narrow transition zone, as commonly observed in poorly mixed systems. The overall (depth-integrated) rate of the reaction is largely limited by substrate transport rates to the transition zone. (*C*) Mathematical idealization of *B*, whereby substrate gathering is infinitely efficient and the transition zone collapses into a single point ("hotspot"). For a more detailed discussion of how the transition zone's width in reality depends on the microbial kinetics, see *SI Appendix*, Fig. S1.

the advective velocity, R_r is the rate of the rth reaction (e.g., mol per volume per time), and $\partial/\partial z$ denotes the derivative along the spatial coordinate z. In some cases, such as in the model introduced below for Cariaco Basin, the advection term v_m may be zero, while in other (advection-dominated) systems D_m may be zero. The first term in Eq. **1** corresponds to the net production or consumption of metabolites at any given location, while the remaining terms describe metabolite transport across space. The equality to zero specifies that fluxes must be balanced such that C_m is at steady state (SMF assumption *iii*).

The differential Eq. **1** allows the calculation of steady-state metabolite concentrations, provided that the reaction rates R_r and the boundary conditions (e.g., the values or gradients of C_m at the system's boundary) are known (see *SI Appendix*, sections S.1 and S.2 for details). Eq. **1** is widely used to describe the distribution of compounds in stagnant water columns, sediments, and biofilms resulting from some underlying reaction rates (3, 5, 8, 10, 43–45). We emphasize that neither assumptions *i–iii* nor Eq. **1** make any statement about what the reaction rates R_r should be, as this requires additional biological reasoning. Conventional models either require a priori knowledge of reaction kinetics and/or microbial population dynamics to predict the rates R_r or estimate unknown kinetic parameters and/or unknown rates by fitting the model to measured geochemical depth profiles (3, 10, 20, 43, 45–49). Here we circumvent these requirements by assuming that microbial communities eventually occupy every available metabolic niche and become infinitely efficient at gathering available substrates (assumption *iv*). We make no assumptions about which or how many different microbial species perform each reaction. Rather, assumption *iv* states only that each reaction runs at the maximum rate possible, as permitted by substrate transport across space and supply by other reactions. A similar idea was introduced by Bouldin (ref. 50, model III therein), who assumed an instantaneous reaction rate to calculate the speed at which a single oxidant/reductant interface (e.g., between oxygen and ferrous iron in sediments) would move downward over time, if the speed was limited entirely by diffusion. The assumption of instantaneous reactions is also used in geochemical equilibrium–reaction-transport models, where chemical reactions are assumed to locally instantaneously converge to thermodynamic equilibrium between subsequent time steps (22, 51–53). The seemingly simple assumption of infinitely efficient substrate gathering has profound consequences. In principle, predicting reaction rates under the SMF framework translates to finding those R_r that, when inserted into Eq. **1**, would result in all metabolite profiles C_m satisfying assumption *iv* as close as possible. Implementing this approach in practice is challenging due to the large number of unknown variables (R_r for each r and at each location). Fortunately, as we explain below, in many cases the number of unknown variables can be substantially reduced.

A high efficiency in substrate gathering and limited mixing rates imply that reactions requiring at least two coupled substrates produced at distinct locations (as is the case for typical redox reactions in systems with pronounced redox gradients) will take place only within narrow chemical transition zones, within which inflowing coupled substrates exhibit just sufficient overlap and outside (on either side) of which at least one of the substrates is essentially absent (Fig. 1*B* and *SI Appendix*, Fig. S1). The same reasoning also applies to abiotic reactions that, in the presence of ample reactants, occur very rapidly compared with reactant transport rates (22) (further discussion in *SI Appendix*, section S.4). In the theoretical limit of infinitely efficient substrate gathering (assumption *iv*) or, equivalently, infinitely fast reactions, such transition zones become increasingly narrow and eventually collapse into single points, henceforth referred to as "hotspots" (Fig. 1*C*). At these hotspots the concentration of at least one substrate approaches zero, and substrate influx and outflux rates are balanced by production and consumption rates (Fig. 2). The expectation that transition zones become narrower as reaction kinetics become faster is straightforward to confirm using existing reaction-transport models (54). The emergence of narrow chemical transition zones is well documented in virtually all systems with low mixing rates, such as stagnant marine water columns (3, 23–27), meromictic lakes (28–30), sediments (1, 31, 32), and biofilms (33). The flux of oxidants and reductants into these transition zones fuels abiotic and metabolic reactions and microbial biomass production (3, 24, 32, 55–57) (*SI Appendix*, Fig. S4*B*), and relatively high cell abundances are often found in such zones (58–61). In the geochemical literature, hotspots are

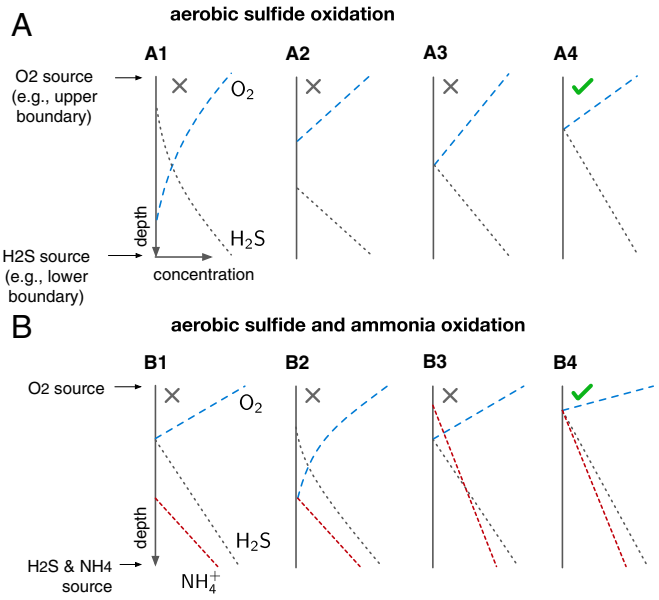

Fig. 2. Constraints on the intensity and location of hotspots. (*A* and *B*) Schematic illustration of the constraints that physical transport processes, chemical boundary conditions, and reaction stoichiometry impose on the location of chemical transition zones (idealized as hotspots) and reaction rates therein under steady state. (*A*) The aerobic oxidation of hydrogen sulfide (diffusing from the bottom) using oxygen (diffusing from the top) through a water column is used as an example, similar to Fig. 1. *A1–A4* show different hypothetical scenarios for the steady-state depth profiles of H_2S and O_2, assuming no other reaction occurs and assuming a constant diffusion coefficient across depth. Only scenario *A4* is physically possible and consistent with the SMF framework. (*A1*) O_2 and H_2S coexist at substantial concentrations in an extended depth interval, contradicting the assumption of efficient substrate gathering. (*A2*) The location of O_2 consumption is separated from the location of H_2S consumption, and hence H_2S cannot possibly be directly oxidized using O_2. (*A3*) O_2 and H_2S are consumed at the same hotspot and at a rate that is sufficiently high to minimize the overlap of O_2 and H_2S. However, the consumption rates of O_2 and H_2S are equal (since their profiles have equal slopes), which is inconsistent with the reaction's stoichiometry ($O_2 : H_2S = 2 : 1$). (*A4*) O_2 and H_2S are both consumed at the same hotspot at sufficient rates, and their flux rate ratio is consistent with the reaction's stoichiometry. See the main text for a mathematical discussion. (*B*) Two competing electron donors diffusing from the bottom are considered (H_2S and NH_4^+) and assumed to be oxidized using O_2 diffusing from the top. Following similar arguments to those in *A*, at steady state only scenario *B4* is physically possible and consistent with the SMF framework. Observe that in *B4* the hotspot is closer to the O_2 source compared with scenario *A4*, to sustain the higher O_2 fluxes needed to oxidize both electron donors. In both *A* and *B*, the location of the hotspot and metabolite fluxes into the hotspot are completely determined by the system's transport properties and chemical boundary concentrations and thus predictable regardless of microbial kinetics and without depth profile measurements. Similar arguments can be made for a greater number of reactions, although the constraints on hotspot locations and reaction rates become increasingly complex.

known as sharp reaction fronts that form when reactions are fast in comparison with transport processes (22, 62).

It is important to note that the locations of hotspots and reaction rates therein are strongly constrained by a system's transport properties, boundary conditions, and the stoichiometry of reactions. This is because at steady state stoichiometric ratios of substrates used by reactions in a hotspot need to match the substrate flux ratios into the hotspot, which in turn depend on a hotspot's distance to substrate sources (such as the system's boundaries). For illustration, consider the hypothetical scenario where sulfide diffuses upward from the sediments through a water column and into an oxic/sulfidic transition zone (modeled as a hotspot), where it is entirely oxidized using downward diffusing oxygen according to a stoichiometric ratio of $O_2 : H_2S = 2 : 1$ (Fig. 2*A*). If we assume for simplicity that the diffusivity D is the same for both O_2 and H_2S and constant across depths, then at steady state the oxygen and sulfide fluxes into the hotspot will be $f_{O_2} = D \cdot C_{O_2}(z_t)/(z_h - z_t)$ and $f_{H_2S} = D \cdot C_{H_2S}(z_b)/(z_b - z_h)$, respectively, where z_t and z_b are the depths of the top and bottom boundaries of the system and z_h is the (a priori unknown) hotspot depth. Stoichiometric balancing implies that $f_{O_2}/f_{H_2S} = 2$, which after some algebraic reordering leads to

$$z_h = \left[2z_t + z_b \frac{C_{O_2}(z_t)}{C_{H_2S}(z_b)}\right] \cdot \left[2 + \frac{C_{O_2}(z_t)}{C_{H_2S}(z_b)}\right]^{-1}. \qquad [2]$$

Thus, if the boundary concentrations $C_{O_2}(z_t)$ and $C_{H_2S}(z_b)$ are specified, the hotspot location, the vertical fluxes, and thus the reaction rate can be determined regardless of the underlying kinetics (see Fig. 2*A* for illustration and Fig. 2*B* for a more complex example). Similar algebraic arguments can in fact also be made for more complex scenarios, for example in the presence of multiple hotspots and reactions or in cases where D varies across depth or between metabolites; however, explicit algebraic solutions become increasingly difficult to retrieve. A more flexible and scalable numerical approach for determining hotspot locations and reaction rates based on the same principles is thus presented below.

Following the above considerations, the differential Eq. **1** can be simplified into an algebraic equation for each steady-state concentration C_m,

$$C_m(z) = C_m^o(z) + \sum_h \sum_r S_{mr} R_{hr} G_m(z_h, z), \qquad [3]$$

where h iterates over all hotspots, z_h is the location of hotspot h, R_{hr} is the integrated rate of reaction r at hotspot h (e.g., mol per m^2 per time), C_m^o is the steady-state concentration of metabolite m in the absence of any reactions, and G_m is the "Green's function" (63) of the differential Eq. **1**. Specifically, $G_m(z_h, z)$ is the metabolite's hypothetical concentration at location z, if a single point source of unit rate were present at location z_h. Green's functions are widely used in geophysics, meteorology, ecology, and sedimentology to describe local or nonlocal transport of matter or energy (63–65), such as molecular diffusion or bioturbation in sediments (5, 66), as an alternative to using differential equations. Note that $C_m^o(z)$ and $G_m(z_h, z)$ can be precalculated independent of the reaction rates for any arbitrary z_h and z (details in *SI Appendix*, section S.1), and thus the only free variables in Eq. **3** are the hotspot locations z_h and the reaction rates R_{hr} at each hotspot. In principle, z_h and R_{hr} can be predicted by finding those values for which the corresponding steady-state metabolite concentrations (i.e., calculated using Eq. **3**) satisfy SMF assumption *iv*; that is, at any location outside of hotspots and for every reaction at least one substrate has concentration zero.

In practice, for complex models there may not exist a combination of hotspot locations and reaction rates satisfying assumption *iv* exactly. Hence, an approximate solution may be sought by choosing a combination of z_h and R_{hr} that minimizes a suitable "stress" function (denoted f), which quantifies the deviations from assumption *iv*. This stress function will be zero if and only if assumption *iv* is exactly satisfied; in the sulfide/oxygen example discussed above, solving for $f = 0$ would yield the same hotspot location as derived from stoichiometric arguments (Eq. **2** and Fig. 2 *A4*). The advantage of using a stress function is that even when no exact solution exists (e.g., if the problem is overdetermined), an approximate solution may still be obtained by minimizing f. This approach of finding an approximate solution is heuristic and should be verified using real systems (see below) because, strictly speaking, the SMF framework cannot predict exactly how real microbial systems behave if they deviate from the framework's idealizing assumptions. We point out that minimizing a model's stress function differs fundamentally from classical model fitting, where the deviations from available data are minimized to determine unknown kinetic parameters or unknown reaction rates (3, 10, 45). Indeed, a stress function as introduced above measures only deviations from the model's own assumption *iv*, but not deviations from any data.

Comparison with Existing Approaches and Implications

The SMF framework differs from biogeochemical models that predict instantaneous reaction rates at each location based on current local chemical conditions (e.g., using first-order kinetics) and which, more recently, also account for microbial population dynamics (3, 8, 20, 43, 46–48, 51, 67–69); these include "free boundary" models where boundary locations are a priori unknown variables representing penetration depths of metabolites, spatial bounds of a model's validity, or transition points between distinct kinetic/transport regimes (4, 6, 7, 50, 69–75) (reaction rates at these free boundaries are often assumed to be zero, and thus these boundaries should not be confused with hotspots). Virtually all of these models require reaction-kinetic parameters such as rate constants and/or biological parameters, such as microbial growth yields and substrate affinities, to explicitly model the dependency of reaction rates on local physicochemical conditions and/or community state. While such models are usually more realistic (and potentially more accurate) than SMF models, they often require dozens of poorly quantified kinetic parameters and may be overly and needlessly complex. The SMF framework shortcuts all reaction- and biokinetic dynamics by assuming that physical transport and reaction stoichiometry eventually dictate large-scale biogeochemical flux rates. While this assumption clearly does not apply to all systems, it has profound implications in those systems where it does. In such systems, the SMF framework provides a substantially simplified approach to predicting biogeochemical fluxes and metabolic activity. Even in systems far from steady state, SMF predictions could provide a reasonable first-order approximation or help identify a system's eventual convergence point, deviations from which correspond to transient processes.

The SMF framework differs from conventional "reconstructive" approaches where measured geochemical depth profiles are used to reconstruct the underlying fluxes or reaction rates generating those profiles, for example based on concentration gradients (18, 27, 76–78) or via inverse transport modeling (79–82). Similarly, the SMF framework differs from approaches that fit unknown kinetic parameters or unknown reaction rates to geochemical depth profiles (3, 9, 10, 45, 54, 68, 83, 84), for example via least squares. All of these approaches require knowledge of the geochemical profiles that resulted from the very processes to be reconstructed; that is, they require that "nature already took its course" and the outcome was subsequently measured in

sufficient detail. In contrast, the SMF framework relies only on the chemical conditions at a system's boundary, for example sulfate concentrations at the water–sediment boundary (~28 mM in most modern marine systems), or on oxygen concentrations at the water–atmosphere boundary (typically near saturation), to predict the resulting process rates in the system's interior. Hence, SMF models could be used to predict future process rates that have not yet occurred, requiring only knowledge of boundary conditions. For example, SMF models may help predict transitions in the biogeochemistry of expanding ocean oxygen minimum zones (85, 86), especially since the future microbial community composition of these systems (and thus their reaction kinetics) is unknown. Reciprocally, SMF models could be used to reconstruct past unknown geochemical boundary conditions based on minimal geological data, such as the likely location of reaction hotspots as indicated by the sedimentary rock record.

The SMF framework exhibits conceptual similarities to flux balance analysis (FBA), a popular framework for predicting the metabolism of single cells at steady state based on the reactions encoded in their genome (87). In FBA, the production, consumption, uptake, and export rates of each metabolite are assumed to be balanced so that the intracellular concentration is constant over time; reaction rates are predicted under these constraints by assuming that a cell regulates reactions such that it achieves the highest possible growth rate. By analogy, in SMF the production/consumption rates of metabolites within hotspots are assumed to be balanced by fluxes across the system's boundaries and transport across space; the hotspot locations and reaction rates are predicted under these constraints by assuming that microbial populations are infinitely efficient at substrate gathering. As in FBA, in SMF the need for kinetic parameters is eliminated based on some principle of optimality. FBA allows predicting the effects of hypothetically added or removed reactions on cell metabolism without knowledge of cell regulation, thus facilitating drug discovery and identification of gene functions (88). Analogously, SMF allows predicting the effects of hypothetically or putatively occurring reactions in an ecosystem without knowledge of kinetics, thus facilitating ecosystem engineering and assessing the potential role of metabolic pathways at ecosystem scales. For example, SMF could provide valuable predictions of maximum feasible biochemical flux rates during the design of spatially structured bioreactors (89), regardless of the eventual inhabiting microorganisms.

Biogeochemical Fluxes in the Cariaco Basin

To demonstrate the applicability of the SMF framework to real ecosystems, we constructed an SMF model for the Cariaco Basin subeuphotic water column (depths 180–900 m, roughly spanning from hypoxia to anoxia, with O_2 concentrations ~ 0–70 μM) during the years 2001–2014. Within this time interval and depth range, water exchange with the open ocean was limited (34), eddy diffusion was the main mode of vertical transport (27, 77, 90), and microbial productivity was likely largely fueled by the supply of inorganic reductants (especially sulfide) from the sediments (18, 35, 82). While this system appeared at times to be near biogeochemical steady state, substantial fluctuations and decadal dynamics have also been observed, especially with regard to sulfide fluxes (18, 34, 82, 91) (Fig. 3*C*). This system thus provides an opportunity to test the robustness of the SMF framework under steady-state as well as non–steady-state conditions. We considered major inorganic electron donors and acceptors (oxygen, nitrate, nitrite, hydrogen sulfide, ammonium), diffusing downward from the overlying layers or diffusing upward from the basin bottom or sediments (27, 77, 82), and redox pathways using or

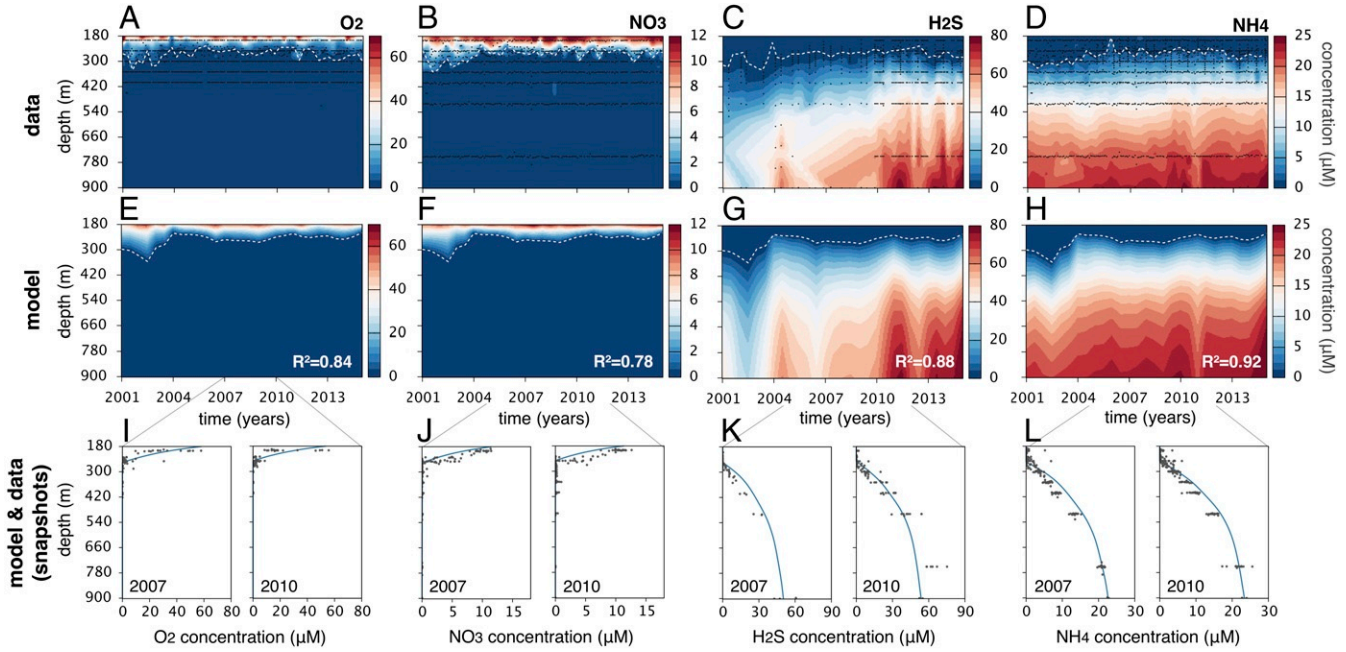

Fig. 3. Chemical concentration profiles in the Cariaco Basin (data vs. predictions). (*A*–*D*) Concentrations of oxygen (O_2, *A*), nitrate (NO_3^-, *B*), hydrogen sulfide (H_2S, *C*), and ammonium (NH_4^+, *D*), measured in the Cariaco Basin over depth and time. Black dots indicate original data points. White dashed curves indicate the deepest point of oxygen penetration (*A*, $[O_2] \leq 2$ μM), the deepest point of nitrate penetration (*B*, $[NO_3^-] \leq 1$ μM), the shallowest point of sulfide penetration (*C*, $[H_2S] \leq 1$ μM), and the shallowest point of ammonium penetration (*D*, $[NH_4^+] \leq 1$ μM). (*E*–*H*) Concentration profiles of the same metabolites as in *A*–*D*, predicted solely based on the boundary values at 180 m and 900 m using an SMF model with a single hotspot. Fractions of variance explained by the model (R^2) are written inside the plots. White dashed curves indicate the location of the hotspot predicted by the model. For a closer comparison between predicted hotspot locations and chemical zone boundaries, see Fig. 4*A*. (*I*–*L*) Predicted metabolite concentrations at specific time points (January 1, 2007 or January 1, 2010, blue curves), compared with concentrations measured around that time (within ±9 mo, gray dots).

producing these metabolites and believed to be of importance to biogeochemical cycling in Cariaco Basin (18, 27, 57, 92–94): aerobic oxidation of sulfide, ammonium, or nitrite; anaerobic oxidation of sulfide using nitrate (producing nitrite); anaerobic oxidation of sulfide using nitrite (producing N_2); and anaerobic ammonium oxidation using nitrite (anammox). We ignored carbon remineralization, the bulk of which likely occurs outside of the considered depth range (mostly in the euphotic zone and sediments, but potentially including the bottom of the basin, depths 900–1,400 m), based on measured heterotrophy rate profiles (*SI Appendix*, Fig. S4*A*) and based on relatively low estimated net in situ sulfate reduction rates (82). To parameterize and validate our model, we used geochemical data generated by the CARIACO Ocean Time Series program at a single station (34, 35) and previously estimated eddy diffusion coefficients (82). We used chemical concentrations, measured at depths 180 m and 900 m, as boundary values for the SMF model, and predicted the hotspot locations and reaction rates between those boundaries at various time points, by minimizing a stress function as described above (Eq. **14** in *Methods*).

We found that a single hotspot was sufficient to largely reproduce the concentration profiles of considered metabolites; the fraction of variance in the profiles that could be explained by the model (R^2) was between 0.78 and 0.92, depending on the metabolite (Fig. 3). The adequacy of a single hotspot to largely reproduce the observed geochemical profiles is consistent with the observation that during this period oxygen, nitrate, hydrogen sulfide, and ammonium were typically consumed in close spatial proximity (Fig. 4*A*), within a zone sometimes referred to as the "redoxcline" (27, 34, 82, 94). The hotspot depths predicted by our model (mean 255 ± 32 m SD, depending on time point) fall within the typical range of the redoxcline (200–350 m). Measured prokaryotic cell densities and dark carbon assimilation (DCA) rates typically peak within the redoxcline (*SI Appendix*, Fig. S4), and previous studies found elevated densities of putative sulfide oxidizers (95) and ammonia oxidation genes (94) therein. Recall that the predicted hotspot location is the location at which reductant and oxidant influxes from opposite sides are stoichiometrically balanced at steady state, taking into account variations of the diffusivity across depth. Because in Cariaco Basin diffusivity at depth is much greater than diffusivity near the top (*SI Appendix*, Fig. S2*C*), the hotspot is located closer to the top boundary. Our model thus provides a simple explanation for the location of the redoxcline.

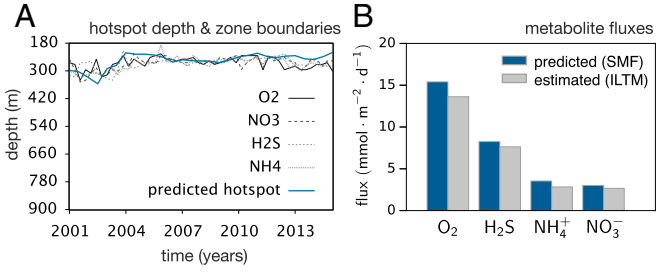

Fig. 4. Hotspot depth and metabolite fluxes in the Cariaco Basin. (*A*) Hotspot depths predicted by the Cariaco SMF model at various time points (blue solid curve), compared with various chemical zone boundaries, including the deepest point of oxygen penetration ($[O_2] \leq 2$ μM, following ref. 34), the deepest point of nitrate penetration ($[NO_3^-] \leq 1$ μM), the shallowest point of sulfide penetration ($[H_2S] \leq 1$ μM, following ref. 34), and the shallowest point of ammonium penetration ($[NH_4^+] \leq 1$ μM). Observe that during years 2007–2009, where Cariaco was close to biogeochemical steady state, the penetration depths of all metabolites are in close proximity to each other and to the predicted hotspot location. (*B*) Net metabolite flux rates into the Cariaco Basin redoxcline during the years 2007–2009, averaged over time. Blue bars show SMF-predicted net fluxes into the hotspot, solely based on the concentrations at the top and bottom boundaries. Gray bars show depth-integrated consumption rates between depths 200 m and 400 m and for the same time period, previously independently estimated via inverse linear transport modeling using geochemical depth profiles (82). All flux rates are expressed as area-specific densities, normalized with respect to the basin area at 150 m depth.

On a finer scale, predicted hotspot depths are often shallower than peak depths of measured prokaryote cell densities (280 ± 77 m SD; *SI Appendix*, Fig. S4*C*), peak depths of measured DCA rates (307 ± 32 m SD; *SI Appendix*, Fig. S4*B*), and peak depths of previously estimated sulfide consumption rates (303 m on average) (82). One potential explanation could be that manganese and iron act as redox shuttles between sulfide and oxygen, whereby dissolved divalent manganese (Mn^{2+}) and iron (Fe^{2+}) are oxidized using oxygen, and sinking particulate manganese and iron (oxyhydr)oxides (MnO_2 and FeOOH) are reduced using sulfide at depth (40, 57, 96). To test the plausibility of this interpretation, we examined an extended SMF model with two hotspots that incorporates putative manganese and iron oxidation and reduction pathways ($MnO_2 \leftrightarrow Mn^{2+}$ and $FeOOH \leftrightarrow Fe^{2+}$; schematic illustration in *SI Appendix*, Fig. S5). This extended model predicts the emergence of sharp MnO_2 and FeOOH concentration maxima within the redoxcline (*SI Appendix*, Fig. S6 *F* and *H*), consistent with previous observations (40, 97, 98). This model also predicts a fine spatial separation between the oxygen and sulfide fronts (~23 m on average; *SI Appendix*, Fig. S7), consistent with frequent observations (98, 99), with sulfide oxidation predicted at somewhat deeper depths (271 ± 35 m SD) than in the original single-hotspot model. These revised predictions, however, only partly alleviate the aforementioned discrepancies with measured DCA rates and previously estimated peak sulfide consumption depths. A more important cause of these discrepancies is likely the fact that the system is not at perfect biogeochemical steady state. Indeed, sulfide concentrations and influxes at depth have been increasing appreciably during most of the considered time period (*SI Appendix*, Fig. S8 and refs. 18 and 82), leading to a mismatch between a transiently deep zone of sulfide consumption on the one hand and a shallower predicted steady-state hotspot location on the other hand. The latter corresponds to the hypothetical eventual point of convergence, if boundary conditions stabilized.

To further assess the model's performance under steady-state conditions, we subsequently restricted our analysis to a time period where sulfide concentrations at depth appeared relatively stable (years 2007–2009; *SI Appendix*, Fig. S8). During this period, the agreement of the model with the profile data increased noticeably ($R^2 = 0.86$–0.95), and the predicted hotspot depths (256 ± 4 m SD) more closely matched the depths of peak measured prokaryotic cell densities (252 ± 38 m SD), peak measured DCA rates (296 ± 25 m SD), and peak estimated sulfide consumption rates (280 m on average) (82), as well as the depths where measured oxygen, nitrate, sulfide, and ammonium concentrations approach zero (~255 m, Fig. 4*A*). Remaining discrepancies may be partly due to uncertainties in the eddy diffusion coefficients and uncertainties in chemical boundary concentrations, as well as the coarse resolution of measured DCA profiles and estimated sulfide consumption profiles (82). Considering the simplicity of the model and the fact that it covers a depth range of 720 m, the model's predictions appear remarkably accurate. Metabolite fluxes into the hotspot, predicted by the SMF model during the years 2007–2009 (Fig. 4*B*), are also highly consistent with depth-integrated metabolite consumption rates within the Cariaco Basin redoxcline, previously estimated from the full geochemical depth profiles (82). This supports our hypothesis that, in the system at hand, large-scale biogeochemical gradients and bulk microbially driven metabolite fluxes are largely determined by the system's transport properties and chemical concentrations at the top and bottom boundaries and largely

independent of the precise kinetics and physiological properties of resident species.

Limitations

The simplicity of SMF models does not come without a number of tradeoffs. First, strictly speaking SMF models apply only to steady state, although in practice the assumption of steady state may not always be crucial. Second, while SMF models may adequately describe bulk biogeochemical fluxes, by design they make no predictions about the population dynamics of microbial communities or about the partitioning of metabolic activity across microbial species. Such predictions may be important in cases where a mechanistic understanding of the underlying ecology is important or when available microbial composition data could otherwise be used to calibrate a model. Third, the SMF variant presented here, whereby reactions are assumed to occur only at discrete hotspots, requires that all reactions be limited by transport rates of at least one substrate across space. In some situations, some reactions may not meet this criterion, in which case the rate of these reactions must be modeled separately, e.g., using kinetic parameters. A notable example is the hydrolytic and fermentative degradation of organic matter in sediments, which even at steady state, and even in the presence of ample organic carbon, can proceed very slowly relative to physical transport (5, 100). This process may ultimately be limited by the number of enzyme-secreting microorganisms that could possibly colonize the outer surface of an organic matter particle, the reactivity of their secreted hydrolytic enzymes, and diffusion rates at the scale of single particles (101). To accommodate such scenarios, our SMF code allows for explicitly specifying input rate profiles of organic matter degradation products (e.g., depth profiles of methane production rates) based on kinetic models and/or empirical rate profiles (5, 102); however, currently these products may themselves not be limiting substrates of reactions confined to hotspots at depths where substantial organic matter degradation occurs. Future SMF variants, allowing for reactions outside of discrete hotspots while still satisfying SMF assumption *iv*, may resolve this limitation.

Conclusions

In general, microbial activity and geochemical gradients are bidirectionally coupled dynamic aspects of ecosystems, continuously influencing each other over time and together leading to the emergence of interwoven microbial and chemical spatial structures (94). In stagnant water columns and sediments, the eventually inevitable utilization of available metabolic niches by microorganisms (42), the self-amplifying nature of microbial metabolism, and selection for efficient substrate gathering (15) can render physical transport processes the rate-limiting step for bulk microbial metabolism at ecosystem scales. Similar arguments also hold for rapidly occurring abiotic reactions. Consequently, biogeochemical fluxes may become largely determined (and as we showed, largely predictable) by a system's physical transport properties, chemical boundary conditions, and the stoichiometry of metabolic pathways. In contrast, variations in the species composition of microbial communities and potential variations in their kinetics may have little net effect on large-scale biogeochemical fluxes in these systems (19). Here we formalized these considerations into a scalable predictive framework and demonstrated that such a framework can yield accurate steady-state predictions for systems with multiple reactions, with heterogeneous diffusion rates, and spanning hundreds of meters. Indeed, our model for the Cariaco Basin reproduced the system's broad geochemical structure remarkably well. The model even reproduced major changes in sulfide and ammonium profiles occurring over time scales of months to years, indicating that exact steady state is not an essential prerequisite for obtaining reasonable predictions. Thus, the SMF framework may also explain the broad-brush biogeochemical structure in other similar systems (25, 29, 30, 32).

Reciprocally, while factors such as boundary conditions and large-scale mixing properties can strongly constrain bulk biogeochemical process rates, they may have limited influence on the precise outcome of community dynamics such as competition, antibiotic warfare, or predation (19). Our work thus provides an explanation for the apparent decoupling between taxonomic community composition and bulk metabolic function frequently observed in microbial systems, whereby stable function can coincide with high species turnover over space or time (19, 103–107). Note that the SMF framework makes no assertion as to whether and how this species turnover affects the kinetic properties of the resident communities; our framework merely illustrates how these kinetic properties can become largely irrelevant to bulk biogeochemical flux rates. Our framework may also be applicable to large-scale biogeochemical processes at geological time scales, where microbial kinetic parameters are often unknown, thus enabling a better understanding of Earth's past and future biogeochemical trajectory.

Materials and Methods

Cariaco Basin Data. Chemical and physical data were downloaded from the CARIACO Ocean Time Series project website (http://www.imars.usf.edu/cariaco) on April 28, 2018. Methods of data collection have been described previously (34, 35). Some of the hydrogen sulfide concentration data, recently published in ref. 35, were obtained directly from the authors. To obtain boundary conditions at the top and bottom and to compare model predictions with measurements, chemical concentration data were bilinearly interpolated onto a regular spatiotemporal grid (Fig. 3 *A*–*D*).

Cariaco Basin SMF Model. The Cariaco Basin SMF model considers the depth range of 180–900 m during the years 2001–2014 (see *SI Appendix*, section S.5 for justifications). The single-hotspot model (which is the main model discussed in the article) considers the following metabolites: nitrate (NO_3^-), nitrite (NO_2^-), ammonium (NH_4^+), hydrogen sulfide (H_2S), and oxygen (O_2). The model considers energy-yielding redox pathways using or producing the above metabolites, which are thought to be particularly important in Cariaco Basin (18, 27, 34, 57, 92–94):

- ASOS, aerobic sulfide oxidation to sulfate:

$$H_2S + 2O_2 \rightarrow SO_4^{2-} + 2H^+. \quad [4]$$

- oxH2SrNO3, oxidation of sulfide to sulfate, coupled to the reduction of nitrate to nitrite:

$$H_2S + 4NO_3^- \rightarrow SO_4^{2-} + 4NO_2^- + 2H^+. \quad [5]$$

- oxH2SrNO2, oxidation of sulfide to sulfate, coupled to the reduction of nitrite to N_2:

$$3H_2S + 8NO_2^- + 2H^+ \rightarrow 3SO_4^{2-} + 4N_2 + 4H_2O. \quad [6]$$

- AMO, aerobic ammonium oxidation to nitrite:

$$2NH_4^+ + 3O_2 \rightarrow 2H_2O + 2NO_2^- + 4H^+. \quad [7]$$

- NXR, aerobic nitrite oxidation to nitrate:

$$2NO_2^- + O_2 \rightarrow 2NO_3^-. \quad [8]$$

- anammox, anaerobic ammonium oxidation using nitrite:

$$NH_4^+ + NO_2^- \rightarrow N_2 + 2H_2O. \quad [9]$$

The extended two-hotspot model, incorporating metal redox shuttles, also considered dissolved manganese (Mn^{2+}), particulate manganese (MnO_2),

dissolved iron (Fe^{2+}), and particulate iron (FeOOH) concentrations and also included the following pathways:

- MnOx, aerobic oxidation of Mn^{2+}:

$$2Mn^{2+} + O_2 + 2H_2O \rightarrow 2MnO_2 + 4H^+. \quad [10]$$

- MnRHS, reduction of manganese oxides using H_2S:

$$4MnO_2 + H_2S + 6H^+ \rightarrow 4Mn^{2+} + SO_4^{2-} + 4H_2O. \quad [11]$$

- FeOx, aerobic oxidation of Fe^{2+}:

$$4Fe^{2+} + O_2 + 6H_2O \rightarrow 4FeOOH + 8H^+. \quad [12]$$

- FeRHS, reduction of FeOOH using H_2S:

$$8FeOOH + H_2S + 14H^+ \rightarrow 8Fe^{2+} + SO_4^{2-} + 12H_2O. \quad [13]$$

In addition to transport via eddy diffusion, particulate metals were also subject to sinking at speeds $v_{MnO_2} = 5.3\ m \cdot d^{-1}$ and $v_{FeOOH} = 3.2\ m \cdot d^{-1}$, based on the formula by Yakushev (48) and typical concentrations in the order of $[MnO_2] \sim 50$ nM (40) and $[FeOOH] \sim 25$ nM (97). Note that sulfide-driven denitrification was split into two sequential steps ($NO_3^- \rightarrow NO_2^-$ and $NO_2^- \rightarrow N_2$), to account for a potential leakage of NO_2^- that may be fueling other NO_2^--consuming reactions such as anammox (3). Nitrite rarely accumulates in the considered depth range and is frequently below the detection limit; however, it was included in the SMF model because it is an important (and potentially limiting) intermediate metabolite. Metabolites formally appearing in the reactions but not expected to limit any reaction (e.g., H^+ or H_2O) were not included in the SMF model.

The hotspot locations z_h and reaction rates R_{hr} in the Cariaco SMF model were predicted at 29 regularly spaced discrete time points during the years 2001–2014, according to the SMF assumptions, as follows. For any given hotspot locations z_h (position along the water column) and reaction rates R_{hr} ($mol \cdot m^{-2} \cdot d^{-1}$), the corresponding steady-state metabolite concentration profiles (C_m) were calculated for the depth interval 180–900 m and using fixed-concentration (also known as "Dirichlet") boundary conditions. Our model accounted for geometric dilution effects due to variation of the basin area with depth, by using an appropriately modified diffusion–advection equation (*SI Appendix*, section S.3) and basin area estimates reported by Samodurov et al. (90) (*SI Appendix*, Fig. S3). For each time point, boundary values of C_m were taken from the chemical concentration time series at the top (180 m) and bottom (900 m) boundaries. Explicit formulas for the Green's functions G_m and reactionless profiles C_m^o, introduced in Eq. **3**, are provided in *SI Appendix*, section S.1. Integrals over depth were calculated using the trapezoid rule, after discretizing the depth range into a regular grid of 200 points. Green's functions were calculated for each point on that grid using explicit formulas (*SI Appendix*, section S.1); for hotspot locations (z_h) between grid points, Green's functions were linearly interpolated.

For any given time point, hotspot locations z_h and reaction rates R_{hr} were determined by minimizing the following stress function:

$$f((z_h)_h, (R_{hr})_{hr}) = \mathbb{E}_r \left[\mathbb{E}_z \bigotimes_{m \in L_r} C_m^2(z) \right]^{\frac{1}{2}}. \quad [14]$$

Here, h iterates over all hotspots, r iterates over all reactions, $\mathbb{E}_r$ denotes the arithmetic average over all reactions, $\mathbb{E}_z$ denotes the arithmetic average over all depths, L_r is the set of substrates (reactants) for reaction r, $C_m(z)$ are the predicted steady-state metabolite concentrations (in μM) and $\bigotimes$ denotes the geometric mean over a reaction's substrates. The right-hand side of Eq. **14** penalizes nonzero concentrations of limiting substrates for each reaction; the term is equal to zero if at every location and for each reaction at least one substrate of the reaction has zero concentration. Note that this stress function is in units of μM and roughly corresponds to the geometric-mean concentration of limiting substrates for each reaction, averaged over all depths and all reactions. This stress function depends on all substrates and is differentiable with respect to z_h and R_{hr}, thus allowing the use of standard numeric minimization algorithms. We mention that in other systems where concentrations vary by several orders of magnitude between metabolites, the concentrations C_m may need to be appropriately rescaled within the stress function to ensure that terms corresponding to different metabolites are all within comparable scales.

Separately for each considered time point, we minimized the stress function f by iteratively varying the z_h and R_{hr} under the constraint that all C_m must be nonnegative, using the MATLAB function `fmincon` (108). To avoid nonglobal local minima, for each time point we repeated the minimization process 500 times while randomly varying the starting values for z_h and R_{hr} (each time obtaining a separate prediction for the z_h and R_{hr}) and kept the prediction from the repeat achieving the lowest stress. The average stress value across all time points was 0.24 μM, which is well below the average oxygen, sulfide, nitrate, and ammonium concentrations in the system, meaning that SMF assumption *iv* was almost exactly satisfied.

Code Availability. MATLAB code implementing the SMF framework, as well as demonstration code for the Cariaco Basin and other scenarios, is available at http://www.loucalab.com/archive/CariacoMetabolic. The code can handle an arbitrary number of metabolites, reactions, and hotspots; separate diffusion coefficients and advection speeds for each metabolite (each of which can vary with depth); and optional constraints on hotspot locations and/or individual reaction rates; as well as various combinations of boundary conditions (fixed concentration and/or fixed flux). The code can account for variation of the lateral (cross-sectional) area with depth and can optionally take into account available depth-profile data to improve predictive accuracy (the latter functionality was not applied in the present study, since the profile data were used for a posteriori model validation).

Data Availability. Raw data used in this article are publicly available at the Cariaco Basin Time Series project website (http://www.imars.usf.edu/cariaco). Alternative sources of Cariaco time series data are the NOAA National Centers for Environmental Information (NCEI), the Ocean Carbon Data System, the US Biological and Chemical Oceanography Data Management Office (BCO-DMO), and the NASA SeaBASS database.

ACKNOWLEDGMENTS. We thank Laura W. Parfrey for useful comments. S.L. was supported by a postdoctoral fellowship from the Biodiversity Research Center, University of British Columbia, and by a startup grant from the University of Oregon. M.D. was supported by NSERC (Natural Sciences and Engineering Research Council of Canada) Discovery Grant 219930. M.I.S., G.T.T., and Y.M.A. were supported by NSF grants as part of the CARIACO Ocean Time Series program. Y.M.A. was also supported by grants from the Venezuelan Fondo Nacional de Ciencia y Tecnología (FONACIT). S.A.C. was supported by NSERC Discovery Grant 0487. We thank all participants and funders of the CARIACO Ocean Time Series program, including the US National Science Foundation, for long-term support and in particular through NSF Grant OCE-1259110 (to M.I.S. and G.T.T.), the Venezuelan Consejo Nacional de Investigaciones Cienti cas y Tecnológicas, and FONACIT.

1. Canfield DE, Thamdrup B (2009) Towards a consistent classification scheme for geochemical environments, or, why we wish the term 'suboxic' would go away. *Geobiology* 7:385–392.
2. Beyenal H, Babauta JT (2012) Microscale gradients and their role in electron-transfer mechanisms in biofilms. *Biochem Soc Trans* 40:1315–1318.
3. Louca S, et al. (2016) Integrating biogeochemistry with multiomic sequence information in a model oxygen minimum zone. *Proc Natl Acad Sci USA* 113:E5925–E5933.
4. Boudreau BP, Westrich JT (1984) The dependence of bacterial sulfate reduction on sulfate concentration in marine sediments. *Geochim Cosmochim Acta* 48:2503–2516.
5. Boudreau BP (1997) *Diagenetic Models and Their Implementation* (Springer, Heidelberg).
6. Di Toro DM, Paquin PR, Subburamu K, Gruber DA (1990) Sediment oxygen demand model: Methane and ammonia oxidation. *J Environ Eng* 116:945–986.
7. Soetaert K, Herman PM (2008) *A Practical Guide to Ecological Modelling: Using R as a Simulation Platform* (Springer Science & Business Media, Dordrecht, The Netherlands).
8. Reed DC, Algar CK, Huber JA, Dick GJ (2014) Gene-centric approach to integrating environmental genomics and biogeochemical models. *Proc Natl Acad Sci USA* 111:1879–1884.
9. Spencer DW, Brewer PG (1971) Vertical advection diffusion and redox potentials as controls on the distribution of manganese and other trace metals dissolved in waters of the Black Sea. *J Geophys Res* 76:5877–5892.
10. Crowe SA, et al. (2008) The biogeochemistry of tropical lakes: A case study from Lake Matano, Indonesia. *Limnol Oceanogr* 53:319–331.
11. Jin Q, Bethke CM (2005) Predicting the rate of microbial respiration in geochemical environments. *Geochim Cosmochim Acta* 69:1133–1143.

12. Jin Q, Roden EE, Giska JR (2013) Geomicrobial kinetics: Extrapolating laboratory studies to natural environments. *Geomicrobiol J* 30:173–185.
13. Nichols D (2007) Cultivation gives context to the microbial ecologist. *FEMS Microbiol Ecol* 60:351–357.
14. Lloyd KG, Steen AD, Ladau J, Yin J, Crosby L (2018) Phylogenetically novel uncultured microbial cells dominate Earth microbiomes. *mSystems* 3:e00055-18.
15. Tilman D (1982) *Resource Competition and Community Structure* (Princeton Univ Press, Princeton).
16. Bristow L, et al. (2016) Ammonium and nitrite oxidation at nanomolar oxygen concentrations in oxygen minimum zone waters. *Proc Natl Acad Sci USA* 113:10601–10606.
17. Ulloa O, Canfield DE, DeLong EF, Letelier RM, Stewart FJ (2012) Microbial oceanography of anoxic oxygen minimum zones. *Proc Natl Acad Sci USA* 109:15996–16003.
18. Taylor GT, et al. (2018) Temporal shifts in dominant sulfur-oxidizing chemoautotrophic populations across the Cariaco Basin's redoxcline. *Deep Sea Res Part II Topical Stud Oceanogr* 156:80–96.
19. Louca S, et al. (2018) Function and functional redundancy in microbial systems. *Nat Ecol Evol* 2:936–943.
20. Hunter KS, Wang Y, Van Cappellen P (1998) Kinetic modeling of microbially-driven redox chemistry of subsurface environments: Coupling transport, microbial metabolism and geochemistry. *J Hydrol* 209:53–80.
21. Rubin J (1983) Transport of reacting solutes in porous media: Relation between mathematical nature of problem formulation and chemical nature of reactions. *Water Resour Res* 19:1231–1252.
22. Neretnieks I (1993) Diffusion, flow and fast reactions in porous media. *Migration and Fate of Pollutants in Soils and Subsoils*, eds Petruzzelli D, Helfferich FG (Springer, Berlin), pp 329–354.
23. Murray JW, et al. (1989) Unexpected changes in the oxic/anoxic interface in the Black Sea. *Nature* 338:411–413.
24. Brettar I, Rheinheimer G (1991) Denitrification in the Central Baltic: Evidence for H_2S-oxidation as motor of denitrification at the oxic-anoxic interface. *Mar Ecol Prog Ser* 77:157–169.
25. Lam P, et al. (2007) Linking crenarchaeal and bacterial nitrification to anammox in the Black Sea. *Proc Natl Acad Sci USA* 104:7104–7109.
26. Hannig M, et al. (2007) Shift from denitrification to anammox after inflow events in the central Baltic Sea. *Limnol Oceanogr* 52:1336–1345.
27. Li XN, Taylor GT, Astor Y, Varela R, Scranton MI (2012) The conundrum between chemoautotrophic production and reductant and oxidant supply: A case study from the Cariaco Basin. *Deep Sea Res Part I Oceanogr Res Pap* 61:1–10.
28. Sorokin DY, Foti M, Pinkart HC, Muyzer G (2007) Sulfur-oxidizing bacteria in Soap Lake (Washington State), a meromictic, haloalkaline lake with an unprecedented high sulfide content. *Appl Environ Microbiol* 73:451–455.
29. Gies EA (2015) Illuminating phylogeny and function of microbial dark matter in Sakinaw Lake. PhD thesis (University of British Columbia, Vancouver, BC, Canada).
30. İnceoğlu O, et al. (2015) Vertical distribution of functional potential and active microbial communities in meromictic Lake Kivu. *Microb Ecol* 70:596–611.
31. Iversen N, Jorgensen BB (1985) Anaerobic methane oxidation rates at the sulfate-methane transition in marine sediments from Kattegat and Skagerrak (Denmark). *Limnol Oceanogr* 30:944–955.
32. Treude T, et al. (2005) Anaerobic oxidation of methane and sulfate reduction along the Chilean continental margin. *Geochim Cosmochim Acta* 69:2767–2779.
33. Stewart PS, Franklin MJ (2008) Physiological heterogeneity in biofilms. *Nat Rev Microbiol* 6:199–210.
34. Scranton MI, et al. (2014) Interannual and subdecadal variability in the nutrient geochemistry of the Cariaco Basin. *Oceanography* 27:148–159.
35. Muller-Karger FE, et al. (2019) The scientific legacy of the CARIACO ocean time-series program. *Annu Rev Mar Sci* 11:5.1–5.25.
36. Yuan-Hui L, Gregory S (1974) Diffusion of ions in sea water and in deep-sea sediments. *Geochim Cosmochim Acta* 38:703–714.
37. Gargett A (1984) Vertical eddy diffusivity in the ocean interior. *J Mar Res* 42: 359–393.
38. Boudreau BP (1986) Mathematics of tracer mixing in sediments: I. Spatially-dependent, diffusive mixing. *Am J Sci* 286:161–198.
39. Alldredge AL, Silver MW (1988) Characteristics, dynamics and significance of marine snow. *Prog Oceanogr* 20:41–82.
40. Percy D, Li X, Taylor GT, Astor Y, Scranton MI (2008) Controls on iron, manganese and intermediate oxidation state sulfur compounds in the Cariaco Basin. *Mar Chem* 111:47–62.
41. Bredehoeft JD, Papaopulos IS (1965) Rates of vertical groundwater movement estimated from the Earth's thermal profile. *Water Resour Res* 1:325–328.
42. Finlay BJ, Maberly SC, Cooper JI (1997) Microbial diversity and ecosystem function. *Oikos* 80:209–213.
43. Wang Y, Van Cappellen P (1996) A multicomponent reactive transport model of early diagenesis: Application to redox cycling in coastal marine sediments. *Geochim Cosmochim Acta* 60:2993–3014.
44. Stewart PS (2003) Diffusion in biofilms. *J Bacteriol* 185:1485–1491.
45. Crowe SA, et al. (2014) Sulfate was a trace constituent of Archean seawater. *Science* 346:735–739.
46. Taillefert M, Gaillard JF (2002) Reactive transport modeling of trace elements in the water column of a stratified lake: Iron cycling and metal scavenging. *J Hydrol* 256: 16–34.
47. Konovalov S, Samodurov A, Oguz T, Ivanov L (2004) Parameterization of iron and manganese cycling in the Black Sea suboxic and anoxic environment. *Deep Sea Res Part I Oceanogr Res Pap* 51:2027–2045.
48. Yakushev EV (2013) RedOx layer model: A tool for analysis of the water column oxic/anoxic interface processes. *The Handbook of Environmental Chemistry*, ed Yakushev EV (Springer, Berlin), Vol 22, pp 203–233.
49. Louca S, Doebeli M (2016) Reaction-centric modeling of microbial ecosystems. *Ecol Model* 335:74–86.
50. Bouldin DR (1968) Models for describing the diffusion of oxygen and other mobile constituents across the mud-water interface. *J Ecol* 56:1–77.
51. Berner R (1980) *Early Diagenesis: A Theoretical Approach*, Princeton Series in Geochemistry (Princeton Univ Press, Princeton).
52. Herzer J, Kinzelbach W (1989) Coupling of transport and chemical processes in numerical transport models. *Geoderma* 44:115–127.
53. Walter A, Frind E, Blowes D, Ptacek C, Molson J (1994) Modeling of multicomponent reactive transport in groundwater: 1. Model development and evaluation. *Water Resour Res* 30:3137–3148.
54. Boudreau BP (1991) Modelling the sulfide-oxygen reaction and associated pH gradients in porewaters. *Geochim Cosmochim Acta* 55:145–159.
55. Indrebo G, Pengerud B, Dundas I (1979) Microbial activities in a permanently stratified estuary. II. Microbial activities at the oxic-anoxic interface. *Mar Biol* 51: 305–309.
56. Zubkov M, Sazhin A, Flint M (1992) The microplankton organisms at the oxic-anoxic interface in the pelagial of the Black Sea. *FEMS Microbiol Lett* 101:245—250.
57. Taylor GT, et al. (2001) Chemoautotrophy in the redox transition zone of the Cariaco Basin: A significant midwater source of organic carbon production. *Limnol Oceanogr* 46:148–163.
58. Lehours AC, Bardot C, Thenot A, Debroas D, Fonty G (2005) Anaerobic microbial communities in Lake Pavin, a unique meromictic lake in France. *Appl Environ Microbiol* 71:7389–7400.
59. Grote J, Jost G, Labrenz M, Herndl GJ, Jürgens K (2008) Epsilonproteobacteria represent the major portion of chemoautotrophic bacteria in sulfidic waters of pelagic redoxclines of the Baltic and Black Seas. *Appl Environ Microbiol* 74:7546–7551.
60. Zaikova E, et al. (2010) Microbial community dynamics in a seasonally anoxic fjord: Saanich inlet, British Columbia. *Environ Microbiol* 12:172–191.
61. Wenk CB, et al. (2013) Anaerobic ammonium oxidation (anammox) bacteria and sulfide-dependent denitrifiers coexist in the water column of a meromictic south-alpine lake. *Limnol Oceanogr* 58:1–12.
62. Lichtner PC (1985) Continuum model for simultaneous chemical reactions and mass transport in hydrothermal systems. *Geochim Cosmochim Acta* 49:779–800.
63. Duffy D (2001) *Green's Functions with Applications*, Applied Mathematics (Chapman & Hall/CRC, New York).
64. Nabighian M, Corbett J (1988) *Electromagnetic Methods in Applied Geophysics: Theory*, Electromagnetic Methods in Applied Geophysics (Soc Exploration Geophysics, Tulsa, OK).
65. Steinkamp K (2011) Inverse modeling of the sources and sinks of atmospheric CO2: Joint constraints from the ocean and atmosphere. PhD thesis (ETH Zürich, Zürich).
66. Boudreau BP (1986) Mathematics of tracer mixing in sediments; II, nonlocal mixing and biological conveyor-belt phenomena. *Am J Sci* 286:199–238.
67. Berner RA (1983) Kinetics of weathering and diagenesis. *Rev Mineral* 8:111–133.
68. Mackin JE, Aller RC (1984) Diagenesis of dissolved aluminum in organic-rich estuarine sediments. *Geochim Cosmochim Acta* 48:299–313.
69. Boudreau BP (2000) The mathematics of early diagenesis: From worms to waves. *Rev Geophys* 38:389–416.
70. Jahnke RA, Emerson SR, Murray JW (1982) A model of oxygen reduction, denitrification, and organic matter mineralization in marine sediments. *Limnol Oceanogr* 27:610–623.
71. Jahnke R (1985) A model of microenvironments in deep-sea sediments: Formation and effects on porewater profiles. *Limnol Oceanogr* 30:966–971.
72. Mackin JE, Swider KT (1989) Organic matter decomposition pathways and oxygen consumption in coastal marine sediments. *J Mar Res* 47:681–716.
73. Wilkinson M, Dampier MD (1990) The rate of growth of sandstone-hosted calcite concretions. *Geochim Cosmochim Acta* 54:3391–3399.
74. Cai WJ, Sayles FL (1996) Oxygen penetration depths and fluxes in marine sediments. *Mar Chem* 52:123–131.
75. Chernyavsky BM, Wortmann UG (2007) Remap: A reaction transport model for isotope ratio calculations in porous media. *Geochem Geophys Geosystems* 8:Q02009.
76. Aller RC, et al. (1985) Early chemical diagenesis, sediment-water solute exchange, and storage of reactive organic matter near the mouth of the Changjiang, East China Sea. *Cont Shelf Res* 4:227–251.
77. Scranton MI, Sayles FL, Bacon MP, Brewer PG (1987) Temporal changes in the hydrography and chemistry of the Cariaco Trench. *Deep Sea Res Part A Oceanogr Res Pap* 34:945–963.
78. Aller RC (1994) The sedimentary Mn cycle in Long Island Sound: Its role as intermediate oxidant and the influence of bioturbation, O2, and Corg flux on diagenetic reaction balances. *J Mar Res* 52:259–295.
79. Berg P, Risgaard-Petersen N, Rysgaard S (1998) Interpretation of measured concentration profiles in sediment pore water. *Limnol Oceanogr* 43:1500–1510.
80. Lam P, et al. (2011) Origin and fate of the secondary nitrite maximum in the Arabian Sea. *Biogeosciences* 8:1565–1577.
81. Lettmann KA, et al. (2012) Estimation of biogeochemical rates from concentration profiles: A novel inverse method. *Estuarine Coastal Shelf Sci* 100:26–37.
82. Louca S, Astor YM, Doebeli M, Taylor GT, Scranton MI (May 2, 2019) Microbial metabolite fluxes in a model marine anoxic ecosystem, 10.1101/625087.
83. Mackin JE, Aller RC (1984) Processes affecting the behavior of dissolved aluminum in estuarine waters. *Mar Chem* 14:213–232.

84. Soetaert K, et al. (1996) Modeling ^{210}Pb-derived mixing activity in ocean margin sediments: Diffusive versus nonlocal mixing. *J Mar Res* 54:1207–1227.
85. Wright JJ, Konwar KM, Hallam SJ (2012) Microbial ecology of expanding oxygen minimum zones. *Nat Rev Microbiol* 10:381–394.
86. Breitburg D, et al. (2018) Declining oxygen in the global ocean and coastal waters. *Science* 359:eaam7240.
87. Orth JD, Thiele I, Palsson BO (2010) What is flux balance analysis? *Nat Biotechnol* 28:245–248.
88. Gianchandani EP, Chavali AK, Papin JA (2010) The application of flux balance analysis in systems biology. *Wiley Interdiscip Rev Syst Biol Med* 2:372–382.
89. Smith DP, et al. (2008) Redox control bioreactor: A unique biological water processor. *Biotechnol Bioeng* 99:830–845.
90. Samodurov AS, et al. (2013) Modeling vertical exchange of heat, salt, and other dissolved substances in the Cariaco Basin. *Deep Sea Res Part I Oceanogr Res Pap* 71:61–72.
91. Taylor G, et al. (2012) Ecosystem response to global climate change in the southern Caribbean Sea. *Proc Natl Acad Sci USA* 109:19315–19320.
92. Madrid VM, Taylor GT, Scranton MI, Chistoserdov AY (2001) Phylogenetic diversity of bacterial and archaeal communities in the anoxic zone of the Cariaco Basin. *Appl Environ Microbiol* 67:1663–1674.
93. Lin X, et al. (2006) Comparison of vertical distributions of prokaryotic assemblages in the anoxic Cariaco Basin and Black Sea by use of fluorescence in situ hybridization. *Appl Environ Microbiol* 72:2679–2690.
94. Cernadas-Martín S, Suter EA, Scranton MI, Astor Y, Taylor GT (2017) Aerobic and anaerobic ammonium oxidizers in the Cariaco Basin: Distributions of major taxa and nitrogen species across the redoxcline. *Aquat Microb Ecol* 79:31–48.
95. Rodriguez-Mora MJ, Scranton MI, Taylor GT, Chistoserdov AY (2015) The dynamics of the bacterial diversity in the redox transition and anoxic zones of the Cariaco Basin assessed by parallel tag sequencing. *FEMS Microbiol Ecol* 91: fiv088.
96. Lewis B, Landing W (1991) The biogeochemistry of manganese and iron in the Black Sea. *Deep Sea Res Part A Oceanogr Res Pap* 38:S773–S803.
97. Bacon MP, Brewer PG, Spencer DW, Murray JW, Goddard J (1980) Lead-210, polonium-210, manganese and iron in the Cariaco trench. *Deep Sea Res Part A Oceanogr Res Pap* 27:119–135.
98. Ho TY, et al. (2004) Vertical and temporal variability of redox zonation in the water column of the Cariaco Basin: Implications for organic carbon oxidation pathways. *Mar Chem* 86:89–104.
99. Li X, Taylor GT, Astor Y, Scranton MI (2008) Relationship of sulfur speciation to hydrographic conditions and chemoautotrophic production in the Cariaco Basin. *Mar Chem* 112:53–64.
100. Middelburg JJ (1989) A simple rate model for organic matter decomposition in marine sediments. *Geochim Cosmochim Acta* 53:1577–1581.
101. Enke TN, Leventhal GE, Metzger M, Saavedra J, Cordero OX (2018) Microscale ecology regulates particulate organic matter turnover in model marine microbial communities. *Nat Commun* 9:2743.
102. Andersson JH, et al. (2004) Respiration patterns in the deep ocean. *Geophys Res Lett* 31:L03304.
103. Fernández A, et al. (1999) How stable is stable? Function versus community composition. *Appl Environ Microbiol* 65:3697–3704.
104. Turnbaugh PJ, et al. (2009) A core gut microbiome in obese and lean twins. *Nature* 457:480–484.
105. Louca S, et al. (2016) High taxonomic variability despite stable functional structure across microbial communities. *Nat Ecol Evol* 1:0015.
106. Louca S, Parfrey LW, Doebeli M (2016) Decoupling function and taxonomy in the global ocean microbiome. *Science* 353:1272–1277.
107. Tully B, Wheat CG, Glazer BT, Huber J (2017) A dynamic microbial community with high functional redundancy inhabits the cold, oxic subseafloor aquifer. *ISME J* 12: 1–16.
108. MATLAB (2015) MATLAB R2015b version 8.6.0 (The MathWorks Inc., Natick, MA).