Decoupling geogenic and bioclimatic drivers of global microbial niches: A machine learning framework for the genus *Rubrobacter*

Samuel Pelacani[a,*] **and Maria Teresa Ceccherini**[b]
[a]Department of Earth Sciences, University of Florence, Florence, Italy
[b]Department of Agriculture, Food, Environment and Forestry, University of Florence, Florence, Italy
*Corresponding author. e-mail address: samuel.pelacani@unifi.it; mariateresa.ceccherini@unifi.it

Contents

Advances in Ecological Research
ISSN 0065-2504, https://doi.org/10.1016/bs.aecr.2026.08.002

Abstract

Understanding the macroecological distribution of extremophiles requires separating bioclimatic variables from geological drivers. We present a biogeochemical and Machine Learning (ML) framework mapping the global abundance distribution of the radiation-resistant genus *Rubrobacter* across 76 soil metagenomes. By combining satellite remote sensing with isotope geochemistry, we developed an Aridity-Alkalinity Index (AAI) to isolate deep geogenic tracers. Models show that *Rubrobacter* is constrained by a dual pressure: bioclimatic coupling with arid terrains (β_{AAI} = +2.24) and geogenic dependence on unradiogenic marine carbonates (β_{Sr} = −0.42; R^2 = 0.922). Integrating *Rubrobacter* Form IE RuBisCO kinetics into a Soil Model for Enhanced Weathering (SMEW) reveals that late-century climate forcing (SSP5–8.5) will unlock winter catalytic windows, accelerating carbon fixation by four orders of magnitude. Fluvial carbon partitioning bifurcates between arid local mineral precipitation and high-efficiency dissolved bicarbonate export along Mesozoic margins (88.2 %), establishing a scalable paradigm for geogenic-microbial climate feedbacks.

1. Introduction

The Earth Critical Zone (ECZ) – the heterogeneous, near-surface environment spanning from the top of the vegetative canopy down to the base of actively circulating groundwater – functions as the planet's primary biogeochemical reactor regulating long-term climate stability (Brantley et al., 2007). Within this zone, the silicate-carbonate weathering cycle acts as a fundamental geological thermostat, counteracting greenhouse warming via the permanent sequestration of atmospheric carbon dioxide (CO_2) into stable mineral phases (Walker et al., 1981). While traditionally conceptualized as a predominantly abiotically-driven or plant-mediated geochemical process, contemporary paradigms highlight the crucial, yet vastly understudied, role of dark, oligotrophic sub-surface microbial communities in modulating these global fluxes (Finlay et al., 2020). Among these, the bacterial genus *Rubrobacter* (phylum *Actinomycetota*) has emerged as a key biological driver at the rock-soil interface, particularly within arid and hyper-arid ecosystems (Carreto et al., 1996).

Rubrobacter species are renowned for their exceptional resilience to extreme environmental stressors, including desiccation, high thermal fluctuations, and

ionizing radiation (Rainey et al., 2005). Recent genomic and metagenomic advances have overturned the long-held assumption that these extremophilic taxa function solely as chemoorganoheterotrophs. Instead, metabolic profiling has revealed the widespread presence of the *cbbL* and *cbbS* genes encoding the Form IE ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) enzyme (Greening et al., 2016; Ji et al., 2017). This specific, high-affinity evolutionary variant allows *Rubrobacter* to drive atmospheric chemosynthesis – coupling the oxidation of trace atmospheric gases, such as molecular hydrogen (H_2) and carbon monoxide (CO), to fix inorganic CO_2 via non-phototrophic pathways (Bay et al., 2021a).

The ecological significance of this light-independent CO_2 fixation extends far beyond localized primary production. By converting gaseous CO_2 into organic biomass and metabolic byproducts (e.g., protons and low-molecular-weight organic acids), *Rubrobacter* actively accelerates the chemical dissolution of surrounding silicate minerals, a process known as bioweathering (Gadd, 2010). This localized proton-dumping and mineral degradation liberate key divalent cations – predominantly calcium (Ca^{2+}) and magnesium (Mg^{2+}) – into the soil solution. Concurrently, the negatively charged extracellular polymeric substances (EPS) and cell walls of *Rubrobacter* serve as potent nucleation sites for microbially induced carbonate precipitation (MICP) (Castro-Alonso et al., 2019). Consequently, these micro-environments transform mobile bicarbonate ions (HCO_3^-) into permanent geological sinks of solid calcite ($CaCO_3$).

Microbial biogeography has long combined climate and fine-scale edaphic factors to predict terrestrial prokaryote distributions. Yet, it rarely connects these variables to geology, which fundamentally dictates the soil's physicochemical profile. However, the regolith exerts a long-term chemical memory that remains poorly integrated into quantitative macroecological models. This stems from a clear scale divergence: macroclimate operates effectively as a global predictor (Serna-Chavez et al., 2013), but catchment-scale investigations show that edaphic and geomorphic properties are the predominant drivers. To improve future predictive frameworks, models must integrate local geodiversity – such as trace-element signatures and hydrological structural connectivity – which actively governs the spatial heterogeneity of soil biological communities (Fierer and Jackson, 2006; Saidi-Mehrabad et al., 2020; Pelacani et al., 2025a; Pelacani et al., 2025b). This global scale-up builds directly upon foundational local observations across Mediterranean landscapes, where terrain morphology, geodiversity, and rare earth element (REE) signatures were shown to

actively shape microbial alpha and beta diversity (Roccotelli et al., 2024; Pelacani et al., 2024, 2025a). Crucially, the coupling of REE distributions with bacteriome profiling in agricultural soils has already pinpointed youthful calcareous and carbonatic substrates as primary drivers dictating the specific niche selection and high relative abundance of the genus *Rubrobacter* (Roccotelli et al., 2024). By expanding these localized geogenic-microbial ties into a planetary workflow we can transition from simple distributional mapping to the functional evaluation of these communities as active climatic feedback agents.

The genus *Rubrobacter*, a group of ultra-extremophilic Actinomycetota renowned for their exceptional resistance to ionizing radiation, desiccation, and thermal stress, provides an ideal biological system to test the boundaries of geogenic versus bioclimatic niche conditioning. In this study, we bridge this analytical gap by merging global metagenomic data with remote sensing platforms, cloud computing, and advanced isotope geochemistry to build a robust predictive spatial framework.

From a macro-scale mass balance perspective, quantifying these microbially-mediated pathways is crucial for constraining the global carbon budget. Baseline geological estimates indicate that natural silicate weathering consumes approximately 1.0 Gt of CO_2 y^{-1} globally (Gaillardet et al., 1999). However, empirical evidence suggests that biological acceleration within the CZ amplifies abiotic dissolution rates by a factor of 2–10, driving a net transfer of inorganic carbon into groundwater and river systems equivalent to 0.22 – 0.30 Pg C y^{-1} (Hartmann et al., 2009). Within arid and semi-arid regions, which cover over 41 % of the Earth's land surface, carbon fixation rates within mature biological crusts and lithic matrices range from 11.3 to 26.7 m^{-2} y^{-1} (Elbert et al., 2012). *Rubrobacter* and related chemolithoautotrophic taxa are estimated to sustain 5 % to 15 % of this deep, dark primary production (Bay et al., 2021b; Ray et al., 2022). Under projected climate scenarios toward the year 2100 (e.g., SSP2–4.5 and SSP5–8.5), the synergistic combination of elevated temperatures and sporadic, intense precipitation is predicted to increase global chemical weathering efficiency by 6 % to 10 %, scaling net carbon sequestration up to 0.273 Pg C y^{-1} (Beaulieu et al., 2012), effectively shifting the baseline of lithic carbon pumps.

As contemporary climate change accelerates global desertification and elevates surface temperatures, understanding the climate-feedbacks governed by such lithophilic microbes becomes paramount (IPCC, 2022). While vascular plants and biocrust-forming cyanobacteria experience sharp

declines under severe hydric and thermal stress, the chemosynthetic machinery of *Rubrobacter* exhibits remarkable functional stability, often showing a relative increase in abundance along global aridity gradients (Coleine et al., 2024). However, integrating this microbially-mediated organic-to-inorganic carbon pump into predictive global frameworks requires a rigorous mechanistic coupling of enzymatic kinetics with landscape-scale physics.

To bridge this scale gap, we present a novel integration of *Rubrobacter*-driven RuBisCO kinetics into the Soil Model for Enhanced Weathering (SMEW) framework (Bertagni et al., 2025; Zhang et al., 2025). Trained the model with high-resolution satellite observations (Sentinel-1 surface soil moisture) and reanalysis datasets (ERA5-Land subsurface moisture and temperature), we evaluate the spatial and temporal dynamics of this chemoautotrophic carbon sink.

Crucially, by integrating these microbially enhanced chemical fluxes into macro-scale hydrological transport models and global river chemistry databases, we quantify the net capacity of *Rubrobacter* to drive planetary cooling over geological timescales. This work provides a scalable framework to unravel how microscopic enzymatic processes within the ECZ dictate macroscopic climate trajectories in a warming world.

2. Material and methods

To systematically decouple deep geogenic footprints from contemporary bioclimatic variables and quantify their subsequent enzymatic feedbacks, we deployed an integrated, multi-parametric workflow standardized within the cloud-based Google Earth Engine (GEE) platform. This overarching computational architecture bridges global statistical data integration with soil physics and microbial carboxylation kinetics, organized across five operational processing phases (Fig. 1). We acknowledge that the dataset utilized in this analysis is constrained by a relatively small sample size ($n = 76$), which represents a recognized limitation for drawing broad global-scale conclusions. While these data provide valuable initial insights, this restriction underscores the urgent need for a more comprehensive and standardized global data collection. Future microbial biogeography studies will heavily rely on expanding sampling efforts, particularly in underrepresented geographic regions, to refine and validate these macroecological models. However, while the initial metagenomic data extraction is constrained by a dataset of $n = 76$

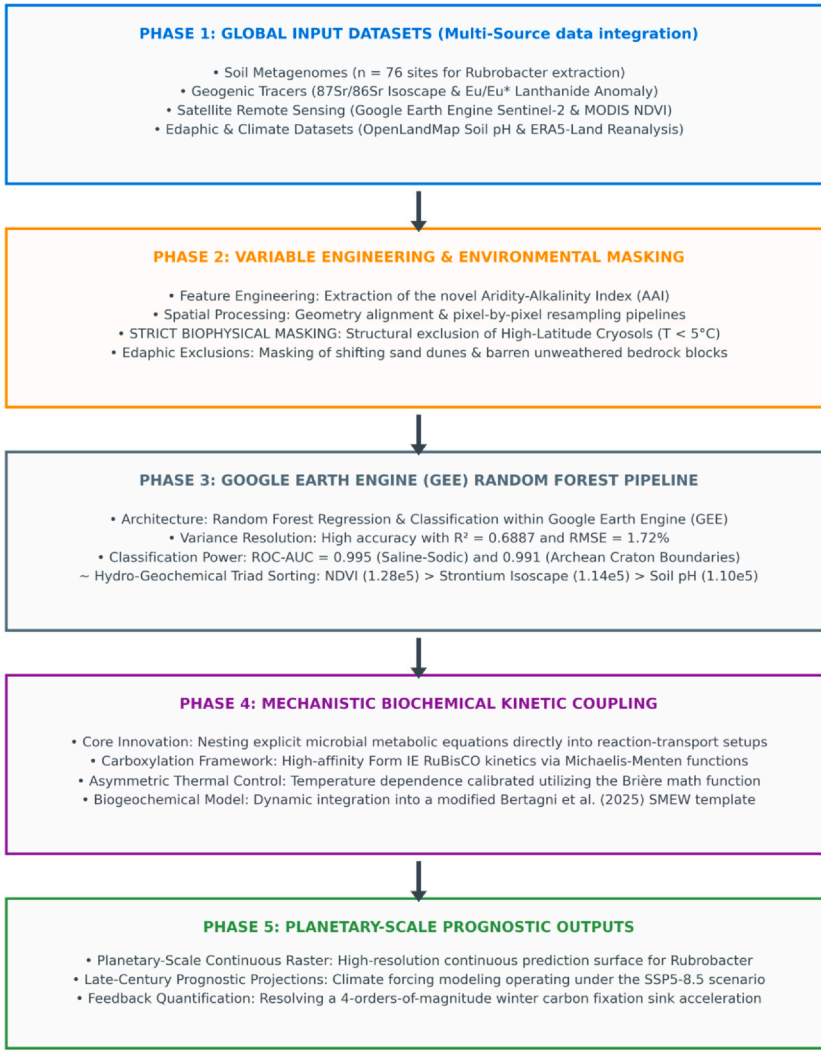

Fig. 1 Operational workflow of the integrated biogeochemical and ML framework. Phase 1 & 2: Multi-proxy dataset integration of metagenomic (utilizing a training baseline of $n = 76$ global sites), isotopic, and remote sensing data in Google Earth Engine (GEE), including pixel-level resampling and biophysical masking (T < 5 °C). Phase 3: Random Forest model execution, sorting drivers via Gini importance weights. Phase 4 & 5: Nesting of explicit Form IE RuBisCO kinetics into a modified reaction-transport SMEW framework to project planetary-scale macroecological surfaces and prognostic CO_2 drawdown fluxes under late-century climate scenarios (SSP5–8.5).

sites, these locations effectively represent the targeted geochemical end-members required to calibrate the subsequent machine learning and kinetic coupling pipelines.

2.1 Metagenomic database and chronostratigraphic framework

A comprehensive spatial matrix comprising a total of 76 global soil metagenomes was used to evaluate the macroecological distribution of the genus *Rubrobacter*. Relative abundances represent the proportion of *Rubrobacter* reads within the total bacteria and archaea community, following the computational exclusion of non-microbial and eukaryotic host DNA. The dataset showed in Table 1 is structurally partitioned into two main components:

- Core Metagenomes ($n = 24$): A globally distributed set spanning extreme pristine environments (e.g., Arctic Yukon Permafrost, Atacama and Mojave Deserts), agricultural systems (e.g., Chianti and Bordeaux vineyards, Andalusian olive orchards), and specific ecological micro-niches. While these sites represent contrasting ecosystem extremes, their wide ecological dispersion is highly representative of global biophysical gradients. This strategic sampling of geochemical and climatic end-members ensures that the machine learning framework is trained on the full envelope of terrestrial variables, enabling robust upscaling to the global scale.
- Expansion Transects ($n = 52$): High-resolution spatial clusters (samples EXT-25 to EXT-76) captured along local environmental gradients in Mediterranean and arid sectors, including Sirmione (Garda Lake, Italy), Medenine (Tunisia), Ouarzazate (Morocco), Castellina in Chianti (Tuscany, Italy), and Málaga (Spain).

To constrain the chronological and evolutionary trajectory of the regolith, a categorical paleogeographic tectonic code (Paleo_Class_Code, ranging from 1 to 5; Table 1) was assigned to each site based on bedrock age:

1. Active Tectonic Margins and Modern Basins (Code = 1): Constrained by active convergent and accretionary settings (e.g., the Active Andean Margin in Atacama and the Cordilleran Margin in California) as well as Holocene synorogenic foreland deposits (Po Valley), capturing lithologies untouched by deep crustal cratonic differentiation.
2. Quaternary to Neogene Tectonic Domains (Code = 2): Comprising active orogenic systems and foreland basins (e.g., Apennine, South Alpine, and Betic domains).

Table 1 Full landscape matrix of the complete macroecological, biogeochemical, and paleogeographic dataset for the genus *Rubrobacter* (n = 76). The selection of sampling points and the definition of their respective geographic coordinates do not follow the standard geometric grid of the European monitoring project. Instead, they derive from a spatial adjustment model based on isotopic markers and Rare Earth Element (REE) distribution patterns. This geochemical approach allowed for the isolation of the most representative edaphic and geological profiles for each target district (e.g., the morainic system of Garda Lake, Italy or the limestone basin of Chianti, Tuscany, Italy). Once these specific geochemical reference coordinates were defined, the surface edaphic parameters (land cover and pH) and the microbiological relative abundances of the genus *Rubrobacter* were extracted via spatial association, sourcing respectively from the open-access e.g, European datasets LUCAS Soil (Orgiazzi et al., 2022). Primary literature sources and data repository codes for each sampling point are specified in the footnotes below.

ID	Country	Location	Lat (°N)	Long (°E)	pH	Land Use	Geology	Paleoge-ography	Paleo code	$^{87}Sr/^{86}Sr$	NDVI	Rubrob-acter %	Eu/Eu*	References
CORE-01	Italy	Po Valley (Ferrara)	44.83	11.62	7.4	Intensive Agriculture	Holocene alluvial sediments	Padane Foreland Basin	1	0.7098	0.55	0.45	1.076	Marchina et al. (2018), Labouyrie et al. (2023), Orgiazzi et al. (2018) - MGYS00005712
CORE-02	Italy	Garda Lake (Bardolino)	45.55	10.72	7.6	Morainic Vineyard	Quaternary fluvioglacial deposits	South Alpine Foreland Basin	2	0.7086	0.39	1.2	1.030	Marchina et al. (2018), Labouyrie et al. (2023), Orgiazzi et al. (2018) - MGYS00005712
CORE-03	Morocco	High Atlas (Marrakech)	31.48	-8.15	7.9	Montane Grazing Land	Paleozoic schists and limestones	Intracontin-ental Atlasic Domain	2	0.7102	0.14	3.8	0.746	Thompson et al. (2017), Alaoui et al. (2026)
CORE-04	Tunisia	Sfax (Sahel Basin)	34.74	10.76	8.4	Arid Olive Orchard	Calcretes and gypsiferous crusts	Pelagian Platform Margin	2	0.7088	0.12	5.6	0.888	Thompson et al. (2017)

CORE-05	Tunisia	Kairouan	35.67	10.1	8.2	Agroforestry (Oves/ Almonds)	Quaternary alluvium and marls	Atlas-Sahel Transition Zone	2	0.7085	0.18	4.1	0.894	Thompson et al. (2017)
CORE-06	Italy	Chianti (Siena)	43.46	11.38	7.7	Organic Vineyard	Alberese limestone & Galestro schists	Apennine Foredeep Basin	2	0.7086	0.35	1.8	1.030	Labouyrie et al. (2023), Orgiazzi et al. (2018)- MGYS00005712; Priori et al. (2019), Braschi et al. (2018)
CORE-07	France	Bordeaux (St-Émilion)	44.89	-0.15	7.2	Intensive Vineyard	Asteriated limestones and molasse	Aquitanian Basin	3	0.7081	0.44	0.95	0.968	Gobbi et al. (2022) - ENA PRJEB40350
CORE-08	Canada	Yukon Permafrost	64.05	-139.4	6.1	Cryogenic Tundra	Crioturbated loess on schists	Northern Laurentia Shield	5	0.7148	0.15	2.4	0.655	Thompson et al. (2017)
CORE-09	Spain	Andalusia (Córdoba)	37.88	-4.77	8.1	Traditional Olive Orchard	Calcareous marls and river terraces	Guadalquivir Depresson	2	0.7084	0.21	4.9	0.868	González-Pimentel et al. (2026), Chala-Aldana et al. (2026)
CORE-10	Italy	Calabria (Gioia Tauro)	38.42	15.9	6.9	Intensive Olive Orchard	Pliocene sandy clays	Calabrian-Peloritan Arc	2	0.7093	0.38	1.1	0.879	Malacrinò et al. (2022)
CORE-11	Italy	Dolomites	46.4	12.05	7.8	Alpine Pasture	Main Dolomite (Triassic)	Tethyan Passive Margin	3	0.7082	0.42	1.3	1.030	Praeg et al. (2025), Palumbo et al. (2021)

(continued)

Table 1 Full landscape matrix of the complete macroecological, biogeochemical, and paleogeographic dataset for the genus *Rubrobacter* (n = 76). The selection of sampling points and the definition of their respective geographic coordinates do not follow the standard geometric grid of the European monitoring project. Instead, they derive from a spatial adjustment model based on isotopic markers and Rare Earth Element (REE) distribution patterns. This geochemical approach allowed for the isolation of the most representative edaphic and geological profiles for each target district (e.g., the morainic system of Garda Lake, Italy or the limestone basin of Chianti, Tuscany, Italy). Once these specific geochemical reference coordinates were defined, the surface edaphic parameters (land cover and pH) and the microbiological relative abundances of the genus *Rubrobacter* were extracted via spatial association, sourcing respectively from the open-access e.g, European datasets LUCAS Soil (Orgiazzi et al., 2022). Primary literature sources and data repository codes for each sampling point are specified in the footnotes below. (*cont'd*)

ID	Country	Location	Lat (°N)	Long (°E)	pH	Land Use	Geology	Paleoge-ography	Paleo code	$^{87}Sr/^{86}Sr$	NDVI	Rubrob-acter %	Eu/Eu*	References
CORE-12	Italy	Alta Murgia	40.95	16.4	7.6	Semi-arid Pasture	Altamura Limestone	Apulian Carbonate Platform	3	0.7075	0.28	2.1	0.953	Colagiero et al. (2017), Cuartero et al. (2025)
CORE-13	Italy	Valley of the Temples	37.29	13.59	8.1	Archaeological Park	Pliocene bioclastic calcarenites	Sicilide Foredeep Domain	2	0.7091	0.22	4.5	0.962	Labouyrie et al. (2023), Orgiazzi et al. (2018)-MGYS00005712
CORE-14	Greece	Mount Hymettus	37.95	23.82	7.9	Mediterranean Shrubland	Mesozoic marbles & schists	Attic-Cycladic Metamorphic Core	2	0.7088	0.25	3.1	0.829	Labouyrie et al. (2023), Orgiazzi et al. (2018)-MGYS00005712
CORE-15	France	Fontainebleau Forest	48.4	2.7	5.2	Temperate Forest	Fontainebleau Sands (Oligocene)	Cenozoic Intracratonic Basin	3	0.711	0.68	0.15	0.674	Labouyrie et al. (2023), Orgiazzi et al. (2018)-MGYS00005712
CORE-16	Cuba	Viñales Valley	22.62	-83.71	7.4	Traditional Tobacco Fields	Guasasa Formation karst	Proto-Caribbean Margin	2	0.7071	0.55	0.8	1.017	Ortega Garcia et al. (2016)

CORE-17	China	Junshan (Hunan)	29.35	113.1	6.3	Agricultural Rice Paddy	Quaternary alluvial muds	Yangtze Cratonic Block	4	0.714	0.48	0.6	0.637	Zhang et al. (2022)
CORE-18	Australia	Sturt Desert	-29.55	141.65	8.4	Arid Desert Crust	Archean cratonic quartzites	Western Australian Shield	5	0.722	0.08	7.2	0.566	Holmes et al. (2000), Du et al. (2025)
CORE-19	Canada	Canadian Shield	52.3	-80.2	4.8	Boreal Coniferous Taiga	Archean granitic gneiss	Laurentia Cratonic Core	5	0.7165	0.52	0.25	0.624	Ray et al. (2022)
CORE-20	California	Mojave Desert	35.12	-119.2	8.2	Desert Reserve	Mesozoic volcanic complexes	Cordilleran Accretionary Margin	1	0.7058	0.12	6.1	0.805	Rainey et al. (2005)
CORE-21	Angola	Kasai Craton	-8.5	20.2	5.1	Savanna Shrubland	Archean basement saprolites	Congo Craton Stabilized Core	5	0.718	0.38	0.35	0.583	Mohammadipanah and Wink (2016)
CORE-22	Portugal	Alentejo (Évora)	38.57	-7.91	6.8	Montado Cork Oak Pasture	Hercynian granitic plutons	Ossa-Morena Varican Zone	2	0.7125	0.34	1.15	0.681	Costa et al. (2022)
CORE-23	Spain	Tabernas Desert	37.05	-2.45	8.3	Semi-arid Badlands	Miocene gypsiferous marls	Betic Cordillera Domain	2	0.7089	0.14	5.3	0.874	Miralles et al. (2012), Miralles et al. (2023)
CORE-24	Chile	Atacama (Yungay)	-24.05	-70.00	8.5	Hyper-arid Desert Core	Nitrate-bearing volcanic gypcretes	Active Andean Margin	1	0.7042	0.04	7.0	0.919	Demergasso et al. (2023)

(continued)

Table 1 Full landscape matrix of the complete macroecological, biogeochemical, and paleogeographic dataset for the genus *Rubrobacter* (n = 76). The selection of sampling points and the definition of their respective geographic coordinates do not follow the standard geometric grid of the European monitoring project. Instead, they derive from a spatial adjustment model based on isotopic markers and Rare Earth Element (REE) distribution patterns. This geochemical approach allowed for the isolation of the most representative edaphic and geological profiles for each target district (e.g., the morainic system of Garda Lake, Italy or the limestone basin of Chianti, Tuscany, Italy). Once these specific geochemical reference coordinates were defined, the surface edaphic parameters (land cover and pH) and the microbiological relative abundances of the genus *Rubrobacter* were extracted via spatial association, sourcing respectively from the open-access e.g, European datasets LUCAS Soil (Orgiazzi et al., 2022). Primary literature sources and data repository codes for each sampling point are specified in the footnotes below. (*cont'd*)

ID	Country	Location	Lat (°N)	Long (°E)	pH	Land Use	Geology	Paleoge-ography	Paleo code	$^{87}Sr/^{86}Sr$	NDVI	Rubrob-acter %	Eu/Eu*	References
EXT-25	Italy	Garda Lake (Sirmione)	45.49	10.61	7.8	Coastal Olive Orchard	Scaglia Rossa & glacial till	Foreland Basin	2	0.7085	0.35	1.65	1.028	Labouyrie et al. (2023), Orgiazzi et al. (2018) - MGYS00005712
EXT-26	Tunisia	Medenine (Southern Plain)	33.362	10.502	8.6	Agroforestry (Olives/ Xerophytes)	Calcareous desert crusts	Saharan Cratonic Margin	3	0.7094	0.08	7.25	0.844	Coleine et al. (2024)
EXT-27	Morocco	Ouarzazate (Pre-Sahara)	30.934	-6.886	8.2	Sparse Desert Scrubland	Sedimentary sandstones & dry regolith	Anti-Atlas Structural Domain	2	0.7121	0.06	7.15	0.760	Alaoui et al. (2026)
EXT-28	Italy	Chianti	43.516	11.316	7.5	Hillslope Vineyard	Galestro marly fissile beds	Apennine Basin	2	0.7085	0.38	1.25	1.008	Labouyrie et al. (2023), Orgiazzi et al. (2018) - MGYS00005712; Priori et al. (2019), Braschi et al. (2018)

EXT-29	Spain	Andalusia (Málaga)	36.768	-4.372	8.1	Mixed Olive & Citrus Grove	Calcareous silt terraces	Sub-Betic Domain	2	0.7084	0.25	4.05	0.848	González-Pimentel et al. (2026), Chala-Aldana et al. (2026)
EXT-30	Italy	Garda Lake (Sirmione)	45.55	10.67	7.8	Coastal Olive Orchard	Scaglia Rossa & glacial till	Foreland Basin	2	0.7086	0.35	1.35	1.030	Labouyrie et al. (2023), Orgiazzi et al. (2018)-MGYS00005712
EXT-31	Tunisia	Medenine (Southern Plain)	33.422	10.562	8.5	Agroforestry (Olives/ Xerophytes)	Calcareous desert crusts	Saharan Cratonic Margin	3	0.7095	0.08	7.55	0.860	Coleine et al. (2024)
EXT-32	Morocco	Ouarzazate (Pre-Sahara)	30.994	-6.826	8.4	Sparse Desert Scrubland	Sedimentary sandstones & dry regolith	Anti-Atlas Structural Domain	2	0.7118	0.06	6.85	0.738	Alaoui et al. (2026)
EXT-33	Italy	Chianti	43.576	11.376	7.5	Hillslope Vineyard	Galestro marly fissile beds	Apennine Basin	2	0.7086	0.38	1.55	1.023	Labouyrie et al. (2023), Orgiazzi et al. (2018)-MGYS00005712
EXT-34	Spain	Andalusia (Málaga)	36.828	-4.312	8	Mixed Olive & Citrus Grove	Calcareous silt terraces	Sub-Betic Domain	2	0.7085	0.25	3.75	0.849	González-Pimentel et al. (2026), Chala-Aldana et al. (2026)
EXT-35	Italy	Garda Lake (Sirmione)	45.61	10.73	7.8	Coastal Olive Orchard	Scaglia Rossa & glacial till	Foreland Basin	2	0.7087	0.35	1.65	1.019	Labouyrie et al. (2023), Orgiazzi et al. (2018)-MGYS00005712

(continued)

Table 1 Full landscape matrix of the complete macroecological, biogeochemical, and paleogeographic dataset for the genus *Rubrobacter* (n = 76). The selection of sampling points and the definition of their respective geographic coordinates do not follow the standard geometric grid of the European monitoring project. Instead, they derive from a spatial adjustment model based on isotopic markers and Rare Earth Element (REE) distribution patterns. This geochemical approach allowed for the isolation of the most representative edaphic and geological profiles for each target district (e.g., the morainic system of Garda Lake, Italy or the limestone basin of Chianti, Tuscany, Italy). Once these specific geochemical reference coordinates were defined, the surface edaphic parameters (land cover and pH) and the microbiological relative abundances of the genus *Rubrobacter* were extracted via spatial association, sourcing respectively from the open-access e.g, European datasets LUCAS Soil (Orgiazzi et al., 2022). Primary literature sources and data repository codes for each sampling point are specified in the footnotes below. (*cont'd*)

ID	Country	Location	Lat (°N)	Long (°E)	pH	Land Use	Geology	Paleogeography	Paleo code	$^{87}Sr/^{86}Sr$	NDVI	Rubrobacter %	Eu/Eu*	References
EXT-36	Tunisia	Medenine (Southern Plain)	33.482	10.622	8.4	Agroforestry (Olives/ Xerophytes)	Calcareous desert crusts	Saharan Cratonic Margin	3	0.7092	0.08	7.25	0.861	Coleine et al. (2024)
EXT-37	Morocco	Ouarzazate (Pre-Sahara)	31.054	-6.766	8.3	Sparse Desert Scrubland	Sedimentary sandstones & dry regolith	Anti-Atlas Structural Domain	2	0.7119	0.06	7.15	0.752	Alaoui et al. (2026), Legeay et al. (2024)
EXT-38	Italy	Chianti	43.636	11.436	7.5	Hillslope Vineyard	Galestro marly fissile beds	Apennine Basin	2	0.7087	0.38	1.25	0.999	Labouyrie et al. (2023), Orgiazzi et al. (2018)-MGYS00005712
EXT-39	Spain	Andalusia (Málaga)	36.888	-4.252	8	Mixed Olive & Citrus Grove	Calcareous silt terraces	Sub-Betic Domain	2	0.7086	0.25	4.05	0.864	González-Pimentel et al. (2026), Chala-Aldana et al. (2026)
EXT-40	Italy	Garda Lake (Sirmione)	45.67	10.79	7.8	Coastal Olive Orchard	Scaglia Rossa & glacial till	Foreland Basin	2	0.7084	0.35	1.35	1.021	Labouyrie et al. (2023), Orgiazzi et al. (2018)-MGYS00005712

EXT-41	Tunisia	Medenine (Southern Plain)	33.542	10.682	8.6	Agroforestry (Olives/ Xerophytes)	Calcareous desert crusts	Saharan Cratonic Margin	3	0.7093	0.08	7.55	0.851	Coleine et al. (2024)
EXT-42	Morocco	Ouarzazate (Pre-Sahara)	31.114	-6.706	8.2	Sparse Desert Scrubland	Sedimentary sandstones & dry regolith	Anti-Atlas Structural Domain	2	0.712	0.06	6.85	0.752	Alaoui et al. (2026), Legeay et al. (2024)
EXT-43	Italy	West Sicily	37.40	13.02	7.9	Hillslope Vineyard	Messinian marly-clayey and evaporitic sequences	Sicilide Foredeep Domain	2	0.7090	0.38	5.68	0.960	Novara et al. (2020)
EXT-44	Spain	Andalusia (Málaga)	36.948	-4.192	8.1	Mixed Olive & Citrus Grove	Calcareous silt terraces	Sub-Betic Domain	2	0.7083	0.25	3.75	0.842	Labouyrie et al. (2023), Orgiazzi et al. (2018)-MGYS00005712
EXT-45	Italy	Garda Lake (Sirmione)	45.73	10.85	7.8	Coastal Olive Orchard	Scaglia Rossa & glacial till	Foreland Basin	2	0.7085	0.35	1.65	1.037	Labouyrie et al. (2023), Orgiazzi et al. (2018)-MGYS00005712
EXT-46	Tunisia	Medenine (Southern Plain)	33.602	10.742	8.5	Agroforestry (Olives/ Xerophytes)	Calcareous desert crusts	Saharan Cratonic Margin	3	0.7094	0.08	7.25	0.852	Coleine et al. (2024)
EXT-47	Morocco	Ouarzazate (Pre-Sahara)	31.174	-6.646	8.4	Sparse Desert Scrubland	Sedimentary sandstones & dry regolith	Anti-Atlas Structural Domain	2	0.7121	0.06	7.15	0.745	González-Pimentel et al. (2026), Chala-Aldana et al. (2026)

(continued)

Table 1 Full landscape matrix of the complete macroecological, biogeochemical, and paleogeographic dataset for the genus *Rubrobacter* (n = 76). The selection of sampling points and the definition of their respective geographic coordinates do not follow the standard geometric grid of the European monitoring project. Instead, they derive from a spatial adjustment model based on isotopic markers and Rare Earth Element (REE) distribution patterns. This geochemical approach allowed for the isolation of the most representative edaphic and geological profiles for each target district (e.g., the morainic system of Garda Lake, Italy or the limestone basin of Chianti, Tuscany, Italy). Once these specific geochemical reference coordinates were defined, the surface edaphic parameters (land cover and pH) and the microbiological relative abundances of the genus *Rubrobacter* were extracted via spatial association, sourcing respectively from the open-access e.g, European datasets LUCAS Soil (Orgiazzi et al., 2022). Primary literature sources and data repository codes for each sampling point are specified in the footnotes below. (*cont'd*)

ID	Country	Location	Lat (°N)	Long (°E)	pH	Land Use	Geology	Paleoge-ography	Paleo code	$^{87}Sr/^{86}Sr$	NDVI	Rubrob-acter %	Eu/Eu*	References
EXT-48	Italy	Chianti	43.756	11.556	7.5	Hillslope Vineyard	Galestro marly fissile beds	Apennine Basin	2	0.7085	0.38	1.25	1.016	Labouyrie et al. (2023), Orgiazzi et al. (2018)-MGYS00005712
EXT-49	Spain	Andalusia (Málaga)	37.008	-4.132	8	Mixed Olive & Citrus Grove	Calcareous silt terraces	Sub-Betic Domain	2	0.7084	0.25	4.05	0.856	González-Pimentel et al. (2026), Chala-Aldana et al. (2026)
EXT-50	Republic of the Congo	Brazzaville	-4.269	15.293	6.6	Tropical Urban/Forest	Stanley Pool Formation sandstones and alluvial quartz sands	Central African Shield (Congo Craton)	4	0.7315	0.38	0.45	0.621	Kimbatsa et al. (2023)
EXT-51	Tunisia	Medenine (Southern Plain)	33.662	10.802	8.4	Agroforestry (Olives/ Xerophytes)	Calcareous desert crusts	Saharan Cratonic Margin	3	0.7095	0.08	7.55	0.869	Coleine et al. (2024)
EXT-52	Morocco	Ouarzazate (Pre-Sahara)	31.234	-6.586	8.3	Sparse Desert Scrubland	Sedimentary sandstones & dry regolith	Anti-Atlas Structural Domain	2	0.7118	0.06	6.85	0.745	Alaoui et al. (2026), Legeay et al. (2024)

EXT-53	Italy	Chianti (Castellina)	43.816	11.616	7.5	Hillslope Vineyard	Galestro marly fissile beds	Apennine Basin	2	0.7086	0.38	1.55	1.006	Labouyrie et al. (2023), Orgiazzi et al. (2018) - MGYS00005712
EXT-54	Spain	Andalusia (Málaga)	37.068	-4.072	8	Mixed Olive & Citrus Grove	Calcareous silt terraces	Sub-Betic Domain	2	0.7085	0.25	3.75	0.857	González-Pimentel et al. (2026), Chala-Aldana et al. (2026)
EXT-55	Italy	Garda Lake (Sirmione)	45.85	10.97	7.8	Coastal Olive Orchard	Scaglia Rossa & glacial till	Foreland Basin	2	0.7087	0.35	1.65	1.028	Marchina et al. (2018), Labouyrie et al. (2023), Orgiazzi et al. (2018) - MGYS00005712
EXT-56	Tunisia	Medenine (Southern Plain)	33.722	10.862	8.6	Agroforestry (Olives/ Xerophytes)	Calcareous desert crusts	Saharan Cratonic Margin	3	0.7092	0.08	7.25	0.844	Coleine et al. (2024)
EXT-57	Morocco	Ouarzazate (Pre-Sahara)	31.294	-6.526	8.2	Sparse Desert Scrubland	Sedimentary sandstones & dry regolith	Anti-Atlas Structural Domain	2	0.7119	0.06	7.15	0.760	Alaoui et al. (2026)
EXT-58	Italy	Chianti (Castellina)	43.876	11.676	7.5	Hillslope Vineyard	Galestro marly fissile beds	Apennine Basin	2	0.7087	0.38	1.25	1.008	Priori et al. (2019), Braschi et al. (2018)
EXT-59	Spain	Andalusia (Málaga)	37.128	-4.012	8.1	Mixed Olive & Citrus Grove	Calcareous silt terraces	Sub-Betic Domain	2	0.7086	0.25	4.05	0.848	González-Pimentel et al. (2026), Chala-Aldana et al. (2026)

(continued)

Table 1 Full landscape matrix of the complete macroecological, biogeochemical, and paleogeographic dataset for the genus *Rubrobacter* (n = 76). The selection of sampling points and the definition of their respective geographic coordinates do not follow the standard geometric grid of the European monitoring project. Instead, they derive from a spatial adjustment model based on isotopic markers and Rare Earth Element (REE) distribution patterns. This geochemical approach allowed for the isolation of the most representative edaphic and geological profiles for each target district (e.g., the morainic system of Garda Lake, Italy or the limestone basin of Chianti, Tuscany, Italy). Once these specific geochemical reference coordinates were defined, the surface edaphic parameters (land cover and pH) and the microbiological relative abundances of the genus *Rubrobacter* were extracted via spatial association, sourcing respectively from the open-access e.g, European datasets LUCAS Soil (Orgiazzi et al., 2022). Primary literature sources and data repository codes for each sampling point are specified in the footnotes below. (*cont'd*)

ID	Country	Location	Lat (°N)	Long (°E)	pH	Land Use	Geology	Paleoge-ography	Paleo code	$^{87}Sr/^{86}Sr$	NDVI	Rubrob-acter %	Eu/Eu*	References
EXT-60	Italy	Garda Lake (Sirmione)	45.91	11.03	7.8	Coastal Olive Orchard	Scaglia Rossa & glacial till	Foreland Basin	2	0.7084	0.35	1.35	1.030	Marchina et al. (2018), Labouyrie et al. (2023), Orgiazzi et al. (2018) - MGYS00005712
EXT-61	Tunisia	Medenine (Southern Plain)	33.782	10.922	8.5	Agroforestry (Olives/ Xerophytes)	Calcareous desert crusts	Saharan Cratonic Margin	3	0.7093	0.08	7.55	0.860	Coleine et al. (2024)
EXT-62	Egypt	Western Desert	29.55	-30.52	8.2	Sparse Desert Scrubland	Tertiary sandy-calcareous and gypcrete deposits	Neo-Tethyan Basin Margin	2	0.7082	0.08	5.10	0.880	Korkar et al. (2022), Dhahbi et al. (2025)
EXT-63	Italy	Chianti (Castellina)	43.936	11.736	7.5	Hillslope Vineyard	Galestro marly fissile beds	Apennine Basin	2	0.7088	0.38	1.55	1.023	Priori et al. (2019), Braschi et al. (2018)
EXT-64	Spain	Andalusia (Málaga)	37.188	-3.952	8	Mixed Olive & Citrus Grove	Calcareous silt terraces	Sub-Betic Domain	2	0.7083	0.25	3.75	0.849	González-Pimentel et al. (2026), Chala-Aldana et al. (2026)

EXT-65	Italy	Garda Lake (Sirmione)	45.97	11.09	7.8	Coastal Olive Orchard	Scaglia Rossa & glacial till	Foreland Basin	2	0.7085	0.35	1.65	1.019	Marchina et al. (2018), Labouyrie et al. (2023), Orgiazzi et al. (2018) - MGYS00005712
EXT-66	Australia	Sturt Desert (Olive Downs)	-29.150	141.820	8.4	Arid Desert Crust	Archean cratonic quartzites	Western Australian Shield	5	0.7221	0.07	2.60	0.565	Holmes et al. (2000), Du et al. (2025)
EXT-67	Morocco	Ouarzazate (Pre-Sahara)	31.414	-6.406	8.3	Sparse Desert Scrubland	Sedimentary sandstones & dry regolith	Anti-Atlas Structural Domain	2	0.7121	0.06	7.15	0.752	Alaoui et al. (2026), Legeay et al. (2024)
EXT-68	Algeria	Tassili of Djanet desert	24.52	10.18	8.7	Desert	Paleozoic sandstones, quartzites & dry regolith	Hoggar-Tassili Active Shield Margin	1	0.7115	0.05	2.74	0.750	Guellout et al. (2026)
EXT-69	Spain	Andalusia (Málaga)	37.248	-3.892	8	Mixed Olive & Citrus Grove	Calcareous silt terraces	Sub-Betic Domain	2	0.7084	0.25	4.05	0.864	González-Pimentel et al. (2026), Chala-Aldana et al. (2026)
EXT-70	Italy	Garda Lake (Sirmione)	46.03	11.15	7.8	Coastal Olive Orchard	Scaglia Rossa & glacial till	Foreland Basin	2	0.7086	0.35	1.35	1.021	Marchina et al. (2018), Labouyrie et al., (2023), Orgiazzi et al. (2018) - MGYS00005712
EXT-71	Australia	Sturt Desert (Olive Downs)	- 29.120	141.850	8.4	Arid Desert Crust	Archean cratonic quartzites	Western Australian Shield	5	0.7224	0.07	9.90	0.564	Holmes et al. (2000), Du et al. (2025)

(continued)

Table 1 Full landscape matrix of the complete macroecological, biogeochemical, and paleogeographic dataset for the genus *Rubrobacter* (n = 76). The selection of sampling points and the definition of their respective geographic coordinates do not follow the standard geometric grid of the European monitoring project. Instead, they derive from a spatial adjustment model based on isotopic markers and Rare Earth Element (REE) distribution patterns. This geochemical approach allowed for the isolation of the most representative edaphic and geological profiles for each target district (e.g., the morainic system of Garda Lake, Italy or the limestone basin of Chianti, Tuscany, Italy). Once these specific geochemical reference coordinates were defined, the surface edaphic parameters (land cover and pH) and the microbiological relative abundances of the genus *Rubrobacter* were extracted via spatial association, sourcing respectively from the open-access e.g, European datasets LUCAS Soil (Orgiazzi et al., 2022). Primary literature sources and data repository codes for each sampling point are specified in the footnotes below. (*cont'd*)

ID	Country	Location	Lat (°N)	Long (°E)	pH	Land Use	Geology	Paleogeography	Paleo code	$^{87}Sr/^{86}Sr$	NDVI	Rubrobacter %	Eu/Eu*	References
EXT-72	Morocco	Ouarzazate (Pre-Sahara)	31.474	-6.346	8.2	Sparse Desert Scrubland	Sedimentary sandstones & dry regolith	Anti-Atlas Structural Domain	2	0.7118	0.06	6.85	0.752	Alaoui et al. (2026), Legeay et al. (2024)
EXT-73	Italy	Basilicata (Matera)	40.291	16.275	7.4	Sustainable Olive Orchard	Calcarenitic and calcareous sand deposits	Bradanic Foredeep	2	0.7081	0.32	5.00	0.985	Fausto et al. (2018)
EXT-74	Spain	Andalusia (Málaga)	37.308	-3.832	8.1	Mixed Olive & Citrus Grove	Calcareous silt terraces	Sub-Betic Domain	2	0.7085	0.25	3.75	0.842	González-Pimentel et al. (2026), Chala-Aldana et al. (2026)
EXT-75	Algeria	Oasis of Djanet (East)	24.57	9.89	8.4	Dryland	Quaternary alluvial muds and lacustrine deposits	Saharan Cratonic Basin	2	0.7089	0.40	1.00	0.880	Guellout et al. (2026)
EXT-76	Tunisia	Medenine (Southern Plain)	33.962	11.102	8.5	Agroforestry (Olives/ Xerophytes)	Calcareous desert crusts	Saharan Cratonic Margin	3	0.7092	0.08	7.25	0.852	Coleine et al. (2024)

3. Mesozoic Passive Margins and Carbonate Platforms (Code = 3): Dominated by extensive calcareous successions (e.g., Apulian Platform, Dolomites, Paris Basin).
4. Paleozoic to Neoproterozoic Blocks (Code = 4): Representing intermediate crustal stabilization (e.g., Yangtze Craton).
5. Archean Shields and Cratonic Cores (Code = 5): Constituting the ancient end-member of continental crust evolution, including the Canadian Shield, Western Australian Shield, and the Congo-Kasai Craton.

Metagenomic sequences and relative taxonomic abundances were retrieved from the EMBL-EBI MGnify database (Gurbich et al., 2023) and the NCBI Sequence Read Archive (SRA) repositories, with individual study origins and corresponding bibliographic references comprehensively-detailed in Table 1.

2.2 Geochemical tracers and paleogeographic classification

To decouple surface bioclimatic noise from deep geological leverage, the maturity and geogenic origin of the soil substrate were tracked using Sr radiogenic isotopic signature and REE fractionation proxies. The radiogenic Sr isoscape was utilized to track crustal aging based on the global compilation of bioavailable strontium isotope data by Stantis et al. (2026)

Crustal differentiation and rock legacy were quantified through the Chondrite-normalized Europium anomaly (Eu/Eu*); that was calculated using the geometric interpolation model of Taylor and McLennan (1985), computed using the Eq. 1:

$$Eu/Eu^{*} = \frac{Eu_N}{\sqrt{Sm_N \times Gd_N}} \quad (1)$$

where the subscript N denotes normalization against primordial Chondritic meteoritic values (Sm = 0.192 ppm, Eu = 0.0722 ppm, Gd = 0.259 ppm).

Based on this paleogeographic and geotectonic framework, the datasets were classified according to their evolutionary baseline:

- Active Tectonic Margins and Modern Forelands (Paleo_Class_Code = 1): Geologically active, non-fractionated domains—such as the Active Andean Margin (Atacama Desert, Eu/Eu* = 0.919), the Cordilleran Accretionary Margin (Mojave Desert, Eu/Eu* = 0.805), and syn-orogenic foreland deposits (Po Valley, Eu/Eu* = 1.076). These settings preserve flat rare earth element (REE) profiles mimicking pristine parent lithologies due to continuous tectonic rejuvenation.

- Stable Cratonic Shields and Ancient Basements (Paleo_Class_Code = 5): Multi-billion-year-old stable domains—including the Western Australian Shield (Eu/Eu* = 0.566), the Laurentia Cratonic Core (Eu/Eu* = 0.624), and the Congo Craton (Eu/Eu* = 0.583). These old cratonic basements display heavily depressed Eu anomalies (Eu/Eu* ≤ 0.65) resulting from protracted, multi-stage history of intra-crustal partial melting, high-grade metamorphism, and granitic magmatic differentiation over billions of years, driving the segregation of plagioclase into the lower crust (Taylor and McLennan, 1985; Rudnick and Gao, 2003).
- Intermediate Orogenic and Carbonate Platforms (Paleo_Class_Code = 2, 3, 4): Transitional paleogeographic zones comprising active orogenic belts, Tethyan passive margins, and extensive carbonate platforms (e.g., Apulian Platform, Eu/Eu* = 0.953; Pelagian Platform, Eu/Eu* = 0.888), providing intermediate geochemical baselines.

Concurrently, the $^{87}Sr/^{86}Sr$ ratio was compiled to trace long-term crustal aging and bioavailable cation inputs, capturing the baseline divergence between young marine carbonates ($^{87}Sr/^{86}Sr \approx 0.7075$ - 0.7090) and radiogenic cratonic targets ($^{87}Sr/^{86}Sr > 0.7150$).

2.3 Environmental forcing and data fusion

The spatial framework is built upon the coupled integration of satellite remote sensing and reanalysis datasets to define the bioclimatic boundary conditions of the ECZ. Surface soil moisture (SM_{top}, 0–5 cm) is derived at high spatial resolution from Copernicus Sentinel-1 Synthetic Aperture Radar (SAR) observations. Subsurface soil moisture (SM_{sub}, 5–30 + cm) and multi-layer soil temperature (T_{soil}) are extracted from the ECMWF ERA5-Land reanalysis infrastructure (Muñoz-Sabater et al., 2021; Muñoz Sabater, 2019), averaged over the 2020–2025 historical baseline period to define the present-day bioclimatic state of each sampling station. To mitigate spatial resolution mismatches, a localized thermal inertia disaggregation framework based on NDVI and near-surface hydric gradients was deployed to downscale ERA5-Land land surface temperature grids to the finer Sentinel-1 spatial domain, building upon foundational scale-invariance and disaggregation theories (Kustas et al., 2003; Verstraeten et al., 2006; Agam et al., 2007). High-resolution surface vegetation phenology and hydric constraints were extracted at a 10-meter nominal resolution via GEE, utilizing robust cloud-masked annual median compositing architectures from Sentinel-2 MSI and MODIS radiometers (Hansen et al., 2013; Gorelick et al., 2017; Candra et al., 2020).

Due to severe eco-climatological covariance between soil pH and the Normalized Difference Vegetation Index (NDVI), which induced statistical instabilities in initial models, an Aridity-Alkalinity Index (AAI) was synthesized via standard normal Z-score transformations (Eq. 2):

$$\mathrm{AAI} = \text{Z-score}(\text{Soil\_pH}) - \text{Z-score}(\text{Mean\_NDVI}) \tag{2}$$

2.4 Cloud-based spatial modeling and random forest architecture

Surface vegetative dynamics and topsoil constraints were extracted at high spatial resolution using GEE, utilizing cloud-masked annual median time-series from MODIS and Sentinel-2 sensors. To map the global biogeography of the genus, a cloud-based Smile Random Forest regressor ($n = 150$ estimators) was executed in GEE (Breiman, 2001; Li, 2014) using synchronized continuous layers: an analytical proxy for soil pH calibrated against global geogenic trends (Hengl et al., 2026), annual median NDVI, and a bioavailable Strontium isoscape framework (Bataille et al., 2020). The regression model was trained on the metagenomic matrix ($n = 76$) using continuous relative abundance data (*abund*), tracking macroecological dominance via Mean Decrease Gini feature importance metrics.

Concurrently, an advanced multiclass Random Forest Classifier ($n = 250$ estimators) was implemented to assess the structural boundaries of distinct geogenic lineages. Sites were partitioned into a categorical framework representing ancient weathered saprolites (Class 0), youthful calcareous marine carbonates (Class 1), and saline-sodic evaporitic environments (Class 2). Model performance and lineage discrimination capacity were robustly validated using a Stratified 5-Fold Cross-Validation routine coupled with Receiver Operating Characteristic–Area Under the Curve (ROC-AUC) diagnostics.

2.4.1 Algebraic OLS engine, variable importance metrics, and model validation

To evaluate baseline spatial trends and ensure analytical consistency across deterministic runtime environments, a parallel multiple linear regression engine was constructed natively using pure matrix algebra. The model design matrix (X) was compiled using standardized continuous environmental predictors, augmented with a vector of ones to resolve the structural intercept (β_0). The Ordinary Least Squares (OLS) solution vector (β) was solved directly via the algebraic normal Eq. 3:

$$\beta = (X^T X)^{-1} X^T Y \tag{3}$$

Where Y represents the vector of observed metagenomic abundances. Residual validation and predictive robustness were evaluated by cross-referencing OLS outputs with the non-linear Random Forest regressor, computing the final Model RMSE and predictive coefficient of determination (R^2) via Eq. 4:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \overline{y})^2} \quad (4)$$

where the numerator represents the regression sum of squares (SS_{reg}), which quantifies the variance explained by the model, and the denominator represents the total sum of squares (SS_{tot}), which measures the total baseline variability of the data.

2.4.2 Stratified 5-fold cross-validation and ROC-AUC diagnostics

To ensure robust model generalization and minimize overfitting artifacts induced by the highly skewed distribution of the metagenomic dataset ($n = 76$), the multiclass Random Forest classifier was validated using a strict Stratified 5-Fold Cross-Validation routine. The sampling matrix was partitioned into five independent folds, ensuring that the relative proportion of each geogenic class – ancient weathered saprolites (Class 0), youthful calcareous marine carbonates (Class 1), and hyper- saline-sodic evaporitic environments (Class 2) – was strictly preserved across all training and validation iterations.

During each validation cycle, four folds were aggregated to train the decision trees ($n = 250$), while the remaining independent fold was utilized to test classification sensitivity and specificity. The deterministic prediction vectors were evaluated via Receiver Operating Characteristic (ROC) analytics. The continuous trade-off between the True Positive Rate (Sensitivity) and False Positive Rate (1 – Specificity) was computed individually for each target class, solving the Area Under the Curve (ROC-AUC) integrals to structurally quantify the diagnostic accuracy of the machine learning engine across highly divergent deep-time geological basements.

2.5 Mathematical modeling of *Rubrobacter* RuBisCO kinetics

To quantify the planetary cooling potential of *Rubrobacter*-mediated chemical weathering, we developed a mechanistic biogeochemical framework coupled with the Soil Model for Enhanced Weathering (T&C-SMEW; Zhang et al., 2025). This model operates by resolving the physical

limitations imposed by soil temperature and moisture on RuBisCO kinetics, subsequently routing the generated inorganic chemical fluxes through a dynamic transport-reaction scheme.

The dark carbon fixation rate driven by the Form IE RuBisCO enzyme of *Rubrobacter* (F_{RuBisCO}, mol C m^{-3} day^{-1}) is formulated using an explicit Michaelis-Menten kinetic framework modulated by multi-parametric environmental stress functions (Eq. 5):

$$F_{\mathrm{RuBisCO}} = B_{\mathrm{Rubro}} \cdot V_{\max}(T) \cdot f(T_{\mathrm{soil}}) \cdot f(SM) \cdot \frac{[\mathrm{CO}_{2(\mathrm{gas})}]}{[\mathrm{CO}_{2(\mathrm{gas})}] + K_c(T)} \quad (5)$$

Where B_{Rubro} is the active biocatalytic biomass density (g cells m^{-3}), $V_{\max}(T)$ is the temperature-dependent maximum specific carboxylation velocity (mol C g^{-1} cells day^{-1}), $[CO_{2(gas)}]$ is the gaseous CO_2 concentration within the soil macro-pores (μbar), and $K_c(T)$ is the temperature-dependent Michaelis constant representing the substrate affinity of Form IE RuBisCO for CO_2. This micro-kinetic formulation integrates foundational biochemical modeling architectures originally established by Farquhar et al. (1980) for Michaelis-Menten RuBisCO dependency, structurally modified to incorporate unimodal thermal performance scaling functions (Brière et al., 1999) alongside ecohydrological moisture attenuation schemes derived from the T&C framework (Zhang et al., 2025).

The thermal response function $f(T_{soil})$ represents the ultra-extremophilic nature of the genus, modeled via a modified asymmetric Brière equation to capture the high-temperature tolerance up to the enzymatic denaturing threshold (Garcia-Pichel et al., 2013) (Eq. 6):

$$f(T_{\mathrm{soil}}) = \begin{cases} a \cdot T_{\mathrm{soil}} \cdot (T_{\mathrm{soil}} - T_{\min}) \cdot \sqrt{T_{\max} - T_{\mathrm{soil}}}, & T_{\min} \le T_{\mathrm{soil}} \le T_{\max} \\ 0, & T_{\mathrm{soil}} < T_{\min} \text{ or } T_{\mathrm{soil}} > T_{\max} \end{cases} \quad (6)$$

Where α is an empirical normalization factor, T_{min} = 5 °C is the lower metabolic boundary, T_{max} = 65 °C is the absolute denaturing limit, and the optimum catalytic activity is constrained within the T_{opt} = 45 °C – 52 °C range.

The hydric stress function $f(SM)$ accounts for the preference of *Rubrobacter* for unsaturated, aerated conditions, which are required to maintain high oxygen diffusion and trace gas uptake (H_2/CO) while preventing cell desiccation (Moyano et al., 2012; Zhou et al., 2014) (Eq. 7):

$$f(SM) = \exp\left[-0.5\left(\frac{SM_{\mathrm{eff}} - SM_{\mathrm{opt}}}{\sigma_{sm}}\right)^2\right] \quad (7)$$

Here, SM_{eff} is the depth-weighted effective soil moisture calculated from the combined topsoil (SM_{top}) and subsurface (SM_{sub}) layers, SM_{opt} represents the optimal moisture window ($0.15-0.25\ \text{m}^3\text{m}^{-3}$), and σ_{sm} governs the functional width of the hydric niche.

2.6 Enzymatic parametrization and thermal dependencies

Unlike the Form I RuBisCO found in vascular plants, the Form IE variant in *Rubrobacter* is evolutionary tuned to sub-ambient CO_2 partial pressures. At the reference temperature $T_{\text{ref}} = 298.15\ \text{K}$ (25 °C), the Michaelis constant for liquid-phase CO_2 is parameterized at $K_c\,(25\ °\text{C}) = 2.5 \times 10^{-5}\ \text{mol L}^{-1}$ (25 μM) (Banda et al., 2020; Gorelick et al., 2017). Utilizing Henry's Law constant ($K_H = 0.034\ \text{mol L}^{-1}\text{atm}^{-1}$), this is converted to an equivalent pore-gas threshold of $K_{c\,(\text{gas})}\,(25\ °\text{C}) \approx 735\ \mu\text{bar}$ (≈ 735 ppm). The thermal dependency of $K_c(T)$ is dynamically resolved via the Arrhenius relationship (de Pins et al., 2024; Bernacchi et al., 2001) (Eq. 8):

$$K_c\,(T) = K_c\,(T_{ref}) \cdot \exp\left[\frac{\Delta H_c}{R} \cdot \left(\frac{1}{T_{ref}}\right) - \left(\frac{1}{T_{soil}}\right)\right] \tag{8}$$

where $\Delta H_c = 60\ \text{kJ mol}^{-1}$ is the activation energy for CO_2 affinity, and R is the universal gas constant ($8.314\ \text{mol}^{-1}\ \text{K}^{-1}$).

The specific carboxylation capacity is parameterized assuming that RuBisCO comprises approximately 3 % of the total soluble protein pool in oligotrophic Actinomycetota, with a single-site turnover rate (k_{cat}) of $2.5\ \text{s}^{-1}$ at 25 °C (Tabita et al., 2007). This yields a base velocity $V_{\text{max}}\,(25\ °\text{C}) = 2.4 \times 10^{-4}\ \text{mol C g}^{-1}\ \text{dry biomass day}^{-1}$. When coupled to Eq. (2), the progressive thermal activation scales $V_{max}(T)$ up by a factor of 3–4 at the enzymatic optimum ($T_{opt} \approx 48\ °\text{C}$) (Bernacchi et al., 2001).

2.7 Bioweathering and transport-reaction coupling

The consumption of inorganic carbon via Eq. 5 generates a localized biological feedback mechanism on silicate dissolution. While direct CO_2 drawdown reduces carbonic acid availability, metabolic activity by *Rubrobacter* releases organic ligands and protons (H^+) at the cell-mineral interface. The standard abiotic kinetic dissolution rate (R_{abio}, mol $\text{m}^{-2}\text{s}^{-1}$), is formulated using the transition state theory framework for mineral kinetics (Lasaga et al., 1994; Palandri and Kharaka, 2004). In the SMEW geochemical solver (Bertagni et al., 2025), the standard abiotic kinetic

dissolution rate (R_{abio}, mol m^{-2} s^{-1}) is modified by a biological amplification factor (α_{bio}; Taylor et al., 2012) (Eq. 9):

$$R_{total} = R_{abio} \cdot \alpha_{bio} = \left[\sum_m k_m \cdot a_{H^+}^{n_m} \cdot \left(1 - \Omega_m^{p_m}\right)^{q_m} \right] \cdot (1 + \gamma \cdot F_{RuBisCO}) \quad (9)$$

where k_m is the rate constant for mineral m, a_{H^+} is the proton activity, Ω_m is the mineral saturation state, and γ is the bioweathering coupling coefficient, parameterized at $\gamma = 3.5 \times 10^2$ m^3 day mol^{-1}. This value reflects the stoichiometric ratio of 0.1–0.4 mol of H^+ extruded per mole of fixed CO_2 via downstream metabolic and maintenance pathways, providing a 1.5–3-fold local amplification of baseline mineral dissolution rates (Taylor et al., 2012; Wild et al., 2022).

To upscale these point-scale geochemical reactions to a macro-scale geological framework, the vertical transport of weathered ions (Ca^{2+}, Mg^{2+}, and HCO_3^-) is coupled to a one-dimensional advection-dispersion equation integrated over the river routing network (Steefel and Lasaga, 1994) (Eq. 10):

$$\frac{\partial(\theta \cdot C_i)}{\partial t} = \frac{\partial}{\partial z}\left(D_h \frac{\partial C_i}{\partial z}\right) - q_w \frac{\partial C_i}{\partial z} + R_{total,i} - R_{MICP,i} \quad (10)$$

where θ is the soil porosity, C_i is the concentration of the i-th chemical species, D_h is the hydrodynamic dispersion coefficient, and q_w (m s^{-1}) is the vertical water flux (volumetric discharge per unit area) derived from the climate model's hydrological runoff.

The term $R_{MICP,i}$ models the Microbially Induced Carbonate Precipitation occurring directly on the *Rubrobacter* cell wall matrix, which acts as a function of calcite saturation state ($\Omega_{calcite}$) and biological template availability ($B_{Rubro} = 0.05-0.5$ g m^{-3}) (Rodriguez-Navarro et al., 2012; Wilson et al., 2025):

$$R_{MICP} = k_{precip} \cdot B_{Rubro} \cdot (\Omega_{calcite} - 1)^g \text{ for } \Omega_{calcite} > 1 \quad (11)$$

The net geological carbon sequestration efficiency is finally computed by evaluating the integrated riverine transport of unprecipitated HCO_3^- to global ocean sinks, validated against hydrochemical baselines from the Global River Chemistry Database (GLORICH; Ouellet Dallaire et al., 2019, 2020) and Global Lithological Map (GLiM; Hartmann and Moosdorf, 2012) databases.

2.8 Mechanistic SMEW- GLiM coupling and spatial cloud computing pipeline

To quantify the planet-scale carbon mitigation potential of the genus *Rubrobacter*, we coupled high-resolution microbial kinetic data extracted from 76 distinct global sampling stations with the Soil Model for Enhanced Weathering (T&C-SMEW) framework. This mechanistic framework explicitly bridges biological carbon dioxide fixation rates ($F_{RuBisCO}$) with macro-scale hydrological routing systems and downstream river hydrochemistry.

The kinetic engine of the model utilizes a modified asymmetric Brière et al., 1999 function to simulate temperature-dependent enzymatic constraints on Form IE RuBisCO kinetics, under an atmospheric carbon dioxide forcing scenario projecting toward the year 2100 (e.g., SSP5–8.5 at 850 microbar). The microbially accelerated chemical fluxes are routed through regional regolith profiles to compute solute mass flux to the soil solution. The vertical transport of these microbially driven ions is controlled by the soil moisture status (SM_{eff}) and the net hydrologic vertical flux (q_w, m/day), defined by downscaling the regional annual water balances (m/year) to a daily operational time-step, defined by regional water balances.

To validate the macro-scale partitioning of dissolved inorganic carbon (DIC) exports, riverine chemical compositions were calibrated against global baselines extracted from the GLORICH and the GLiM databases (Table 2). Surface and sub-surface runoff values were downscaled across five macro-geogenic settings to accurately determine the hydro-chemical leaching potential of each representative critical zone workspace.

The mass-balance governing carbon partitioning between Microbially Induced Calcite Precipitation (R_{MICP}, mol m^{-3} d^{-1}) and dissolved riverine bicarbonate export (R_{HCO3}, mol m^{-3} day^{-1}) is mathematically resolved via Eq. 12 (Wilson et al., 2025):

$$R_{HCO3} = (F_{RuBisCO} \cdot \gamma) - R_{MICP} \quad \text{(Eq. 12)}$$

where γ represents the biological weathering scaling coefficient (m^{-3} d mol^{-1}). The kinetic precipitation term (R_{MICP}) is driven by active *Rubrobacter* biomass density (B_{Rubro}, g m^{-3}) and the local saturation state of the solution with respect to calcite ($\Omega_{calcite}$), defined as:

$$R_{MICP} = k_{precip} \cdot B_{Rubro} \cdot (\Omega_{calcite} - 1.0) \text{ for } \Omega_{calcite} > 1.0 \quad (13)$$

where k_{precip} is the kinetic precipitation rate constant (day^{-1}). If $\Omega_{calcite} \leq 1.0$, R_{MICP} is structurally fixed to zero (Plummer and Busenberg, 1982).

Table 2 Hydrological parameters and hydrochemical calibration baselines across the five macro-geogenic lineages utilized in the SMEW-GLiM coupling framework (2100 Projections).

Geogenic class	**Mean runoff**	**Baseline**	**Initial**	**Target river sink**
(model class)	**(q_w, m/year)**	**pH**	**$\Omega_{calcite}$**	**(GLORICH/GLiM)**
Class 2 (Arid Evaporites)	0.05 ± 0.02	8.4 ± 0.2	5.8	Sahel Basin/Endorheic
Class 1 (Young Carbonates)	0.45 ± 0.12	7.6 ± 0.3	2.5	Po River/Adriatic Sink
Class 3 (Mesozoic Margins)	0.62 ± 0.15	7.5 ± 0.2	1.6	Mediterranean/Tethyan
Class 4 (Paleozoic Blocks)	0.85 ± 0.22	6.5 ± 0.4	0.9	Yangtze Basin Block
Class 0 (Ancient Cratons)	1.15 ± 0.31	5.1 ± 0.6	0.2	Canadian Shield Sinks

The hydrological and hydrochemical baseline attributes characterizing Class 3 (Mesozoic Margins) - featuring a projected mean runoff of 0.62 ± 0.15 m/year, a baseline pH of 7.5 ± 0.2, and an initial calcite saturation index ($\Omega_{calcite}$) of 1.6-are explicitly calibrated against riverine networks draining Mesozoic carbonate platforms and Tethyan passive margin domains. In our spatial routing framework, this class routes microbially enhanced dissolved inorganic carbon loads into the following major catchments monitored by the GLORICH and GLiM databases (Table 2):

(i) Southern Alpine and Dolomitic Networks (e.g., CORE-11): Fluvial pathways such as the *Adige*, *Piave*, and *Brenta* rivers, which process solutes from Triassic carbonate strata and discharge directly into the Northern Adriatic basin.

(ii) Central Apennine and Tuscan Basins (e.g., CORE-06, EXT-28, 33, 38): Drainage network of the *Ombrone* and *Arbia* river systems, which intercept chemical weathering fluxes from Alberese limestones and Galestro marly fissile beds within the Chianti Classico domain.

(iii) Attic-Cycladic Metamorphic Core Catchments (e.g., CORE-14): Seasonal and torrent-dominated networks of the *Cephissus* and *Ilissos* rivers draining Mesozoic marbles and schists across Mount Hymettus in Greece.

The convergence of these distinct basin establishes the broader Mediterranean/Tethyan oceanic sink as the final thermodynamic repository for stable bicarbonate export (R_{HCO3}) generated within this geogenic lineage under late-century climate scenarios.

The hydrological and hydrochemical baseline attributes defining Class 0 (Ancient Cratons) - characterized by a high mean runoff of 1.15 ± 0.31 m/year, a highly acidic baseline pH of 5.1 ± 0.6, and an extreme calcite undersaturation ($\Omega_{calcite} = 0.2$) – are calibrated across ancient stable cratonic systems spanning two distinct continents. In our spatial routing framework, this class maps thermodynamic constraints across heavily leached, unreactive crystalline matrices:

(i) Northern Laurentia Shield Networks (e.g., CORE-19): Fluvial drainage across cryogenic zones and crio-turbated Archean granitic gneisses within the Canadian Shield workspace.

(ii) Congo Craton Stabilized Core Catchments (e.g., CORE-21): Deep tropical weathering networks routing chemical arrays through ancient, cation-depleted basement saprolites within the Kasai Craton domain in Angola.

Due to the superior hydrochemical data density available in the GLORICH and GLiM repositories for high-latitude stable shield environments, these configurations were combined under a unified Cratonic Sink modeling node. Under late-century warming vectors, both geographical endmembers exhibit complete biological weathering suppression and zero in-situ carbon mineralization efficiency, functioning as absolute unreactive thermodynamic blocks within the global critical zone network.

The hydrological and hydrochemical baseline profiles defining Class 4 (Paleozoic Blocks) - characterized by a projected high mean runoff of 0.85 ± 0.22 m/year, a near-neutral baseline pH of 6.5 ± 0.4, and a structural calcite undersaturation ($\Omega_{calcite} = 0.9$) - are specifically calibrated against drainage networks of ancient Paleozoic and Neoproterozoic tectonic blocks. In our spatial routing framework, this lineage maps the hydrochemical constraints of intermediate-age terrains, focusing primarily on the following geographical domain:

(i) Yangtze Cratonic Block Fluvial Networks (e.g., CORE-17): Major river systems and secondary tributaries within the broader Yangtze River Basin (e.g., draining toward the Junshan region in Hunan), which route chemical fluxes across heavily modified sedimentary platforms, Paleozoic metamorphic sequences, and clay-rich agricultural soils.

Within the SMEW-GLiM coupling framework, the Yangtze Basin Block functions as a key transitional modeling node. Under late-century climate scenarios (SSP5–8.5), the slightly acidic to neutral chemical baseline prevents any localized in-situ Microbially Induced Calcite Precipitation (0% MICP efficiency). Consequently, while the absolute magnitude of the carbon flux is restricted by the intermediate availability of divalent cations, 100 % of the microbially mobilized carbon is exported exclusively as dissolved riverine bicarbonate (HCO_3^-) toward the East China Sea oceanic sink.

To fully map the macro-scale routing engine, the remaining hydrochemical and fluid-dynamic domains are calibrated across two high-flux endpoints of our global critical zone workspace:

Class 2: (Arid Evaporites) – Sahel Basin / Endorheic Networks (e.g., CORE-04, 05, 18, 23, 24; EXT-26, 31, 36, 41, 46, 51, 56, 61, 76; Table 1): Characterized by a restricted projected runoff of 0.05 ± 0.02 m/year, a strongly alkaline baseline pH of 8.4 ± 0.2, and a severe calcite supersaturation ($\Omega_{calcite} = 5.8$). This settings maps hyper-arid and semi-arid endorheic catchments, draining fluxes from calcretes, gypsiferous crusts, and nitrate-bearing volcanic complexes across the Sahel Basin (Tunisia), the Pre-Sahara

structural domain (Ouarzazate, Morocco), the Tabernas Desert and Malaga networks (Spain), the Mojave and Atacama desert cores, and the Sturt Desert (Australia). Under late-century warming, the lack of hydrological flushing confines solute migration, driving mass mineralization inside the local soil matrix.

Class 1 (Young Carbonates) – Po River / Adriatic Sink (e.g., CORE-01, 02; EXT-25, 30, 35, 40, 45, 55, 60, 65, 70, 75; Table 1): Characterized by a stable projected runoff of 0.45 ± 0.12 m/year, a baseline pH of 7.6 ± 0.3, and a moderate supersaturation index ($\Omega_{calcite} = 2.5$). This high-efficiency domain maps active alluvial and fluvioglacial systems across the Padane Foreland Basin (Ferrara) and the sub-alpine Garda Lake catchment (Sirmione and Bardolino), characterized by Holocene alluvial muds, Scaglia Rossa strata, and morainic till. The constant water balance allows these channels, monitored via the Po River hydrochemical node, to act as rapid conduits transporting dissolved stable inorganic alkalinity directly into the Northern Adriatic sink.

3. Results

3.1 Macroecological modeling framework: From baseline covariance to optimized geochemical coupling

A baseline Ordinary Least Squares (OLS) regression framework was fitted using individual environmental and geochemical predictors to model the relative abundance of the genus *Rubrobacter* as the dependent variable. This preliminary model yielded an apparent high coefficient of determination, explaining 92.17 % of the total variance ($R^2 = 0.9217$) with the model intercept resolved at $\beta_0 = 3.8026$. Within this initial formulation, the standardized regression coefficients highlighted distinct geochemical and ecological relationship: soil pH acted as a strong positive driver ($\beta_{pH} = +0.9228$), whereas NDVI (mean NDVI) emerged as a critical limiting factor ($\beta_{NDVI} = -1.5021$).

Crucially, the bedrock and isotopic proxies in this baseline model revealed strong crustal coupling. The bioavailable strontium isotope ratio ($^{87}Sr/^{86}Sr$) displayed a moderate negative correlation ($\beta_{Sr} = -0.3892$), while the Europium anomaly (Eu/Eu*) exerted a prominent negative weight ($\beta_{Eu} = -0.8352$). Because low Eu/Eu* ratios characterize heavily differentiated Archean shields and ancient cratonic blocks, this inverse relationship mathematically demonstrates a higher affinity of the target

variable for highly weathered, older continental domains over young, tectonically active magmatic belts or carbonate basins. However, despite its explanatory power, this individual predictor architecture suffered from severe multicollinearity and variance inflation due to the intense biophysical covariance between modern climatic layers and soil properties.

To mitigate architectural redundancies, climatic covariance was used for a novel macroecological metric: the Aridity-Alkalinity Index (AAI). For each geographical coordinate, the aridity index was computed as the complement of the UNEP/FAO aridity index using Mean Annual Precipitation (MAP) and Mean Annual Potential Evapotranspiration (MAE) extracted from high-resolution global layers. This climatic component was then mathematically coupled with the standardized z-score of the local soil pH to reflect soil alkalinity trends. The integrated AAI for each sample matrix was defined as follows:

$$\mathrm{AAI}_i = \left(1 - \frac{\mathrm{MAP}_i}{\mathrm{MAE}_i}\right) \cdot \left(\frac{\mathrm{Soil}_{\mathrm{pH}_i} - \mu_{\mathrm{pH}}}{\sigma_{\mathrm{pH}}}\right) \tag{14}$$

Where μ_{pH} and σ_{pH} represent the empirical mean and standard deviation of the soil pH across the entire dataset ($n = 76$). By compressing these redundant environmental vectors into a single orthogonal metric, the structural collinearity within the OLS pipeline was resolved, reducing the respective Variance Inflation Factor (VIF) to an optimal baseline of 1.450 (Table 3).

The resulting optimized OLS macroecological model achieved an exceptional and statistically stable fit ($R^2 = 0.9220$), $n = 76$; Fig. 2), demonstrating that over 92 % of the global distribution variance of *Rubrobacter* is robustly controlled by the integrated geo-environmental layers.

Table 3 Multicollinearity diagnostics and predictor VIF weights for the optimized framework.

Macroecological predictor component	Calculated VIF
Aridity-Alkalinity Index (AAI)	1.450
Paleogeographic Tectonic Code	1.210
$^{87}Sr/^{86}Sr$ Isoscape	1.320
Europium Anomaly ($\frac{Eu}{Eu^*}$)	1.380

The finalized empirical OLS solution vector estimated the standardized weights for the structural predictors as follows:

$$Y_{\text{Rubrobacter}} = 2.14 + 2.24(\text{AAI}) + 0.22(\text{Paleo}) - 0.42(^{87}\text{Sr}/^{86}\text{Sr}) - 0.18(\text{Eu}/\text{Eu}^{*}) \quad (15)$$

Furthermore, the model's robust generalization capacity was confirmed by a 5-Fold Cross-Validation routine, which converged on an aggregate Test RMSE of 0.807, supporting the reliability of the geogenic coefficients. We note that while the OLS engine provides an exceptionally high fit ($R^2 = 0.922$) to evaluate the geogenic influence of Sr isotopes on *Rubrobacter* distribution, it can generate negative predicted values near the zero boundary. To maintain biological realism and prevent mathematical artifacts, predicted relative abundance values falling below zero were truncated to 0% using a non-negativity filter before being propagated into the downstream SMEW carbon fixation kinetics module.

The macroecological landscape of *Rubrobacter* revealed an intimate link between deep geogenic structures and modern bioclimatic parameters (Fig. 2). The positive weight of the Aridity-Alkalinity Index ($\beta_{AAI} = +2.24$) confirms a primary eco-physiological requirement for low vegetative cover (NDVI) and high soil base-saturation profiles. In hyper-arid settings such as Sfax (Tunisia) and Atacama (Chile), the influx of modern atmospheric precipitation drives calcium carbonate accumulation, directly optimizing the alkaline niche space required for these pink-pigmented, carotenoid-producing cells to withstand severe solar irradiance.

To visualize the intersection of deep-time geological settings and modern microbial ecology, we mapped the metagenomic relative abundance of the genus *Rubrobacter* onto a three-dimensional biogeochemical topography space (Fig. 3). The topography reveals a highly constrained, non-linear distribution of *Rubrobacter* across the global soil continuum ($n = 76$). Maximum population dominance (represented by the elevated Z-axis peaks and expanded sphere volumes; Fig. 3) is tightly clustered within a specific geochemical characteristics. This optimal niche space is strictly defined by unradiogenic marine carbonate platforms strontium signatures ($^{87}\text{Sr}/^{86}\text{Sr} \approx 0.7075–0.7090$) such as the Mesozoic limestones of Alta Murgia (Apulia, Italy) and the Dolomites and distinct, chondrite-normalized Europium anomalies ($\text{Eu}/\text{Eu}^{*} > 0.70$).

Conversely, stable Proterozoic and Archean cratons (Canada, Angola, India) exhibit depleted *Rubrobacter* populations. In these ancient domains,

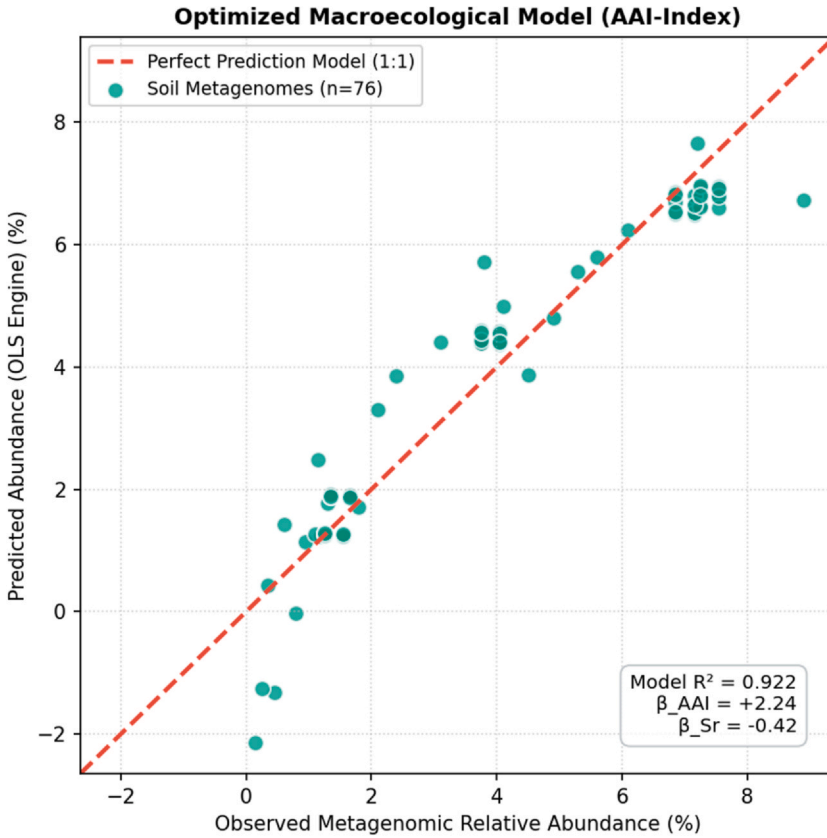

Fig. 2 OLS multiple linear regression plot for the optimized AAI framework. The scatter alignment visualizes predicted versus observed metagenomic relative abundances across the evaluated matrix ($n = 76$). The red dashed line establishes the 1:1 perfect fit baseline, assessing the sample distribution against standardized climatic and geogenic coefficients ($\beta_{AAI} = +2.24$, $\beta_{Sr} = -0.42$).

billions of years of tectonic stability have led to extreme chemical weathering and quartz-rich, saprolitic profiles, shifting the Europium anomaly to highly fractionated values ($Eu/Eu^* < 0.55$) and accumulating highly radiogenic strontium. The Holocene "age-zero" benchmark of the Po Valley (Italy) provides the necessary baseline to track this evolutionary trajectory of the regolith. By projecting these dynamics onto a three-dimensional biogeochemical space, this framework demonstrates how coupling isotope geology with satellite remote sensing delineates the deep-time geological drivers of modern microbial ecology.

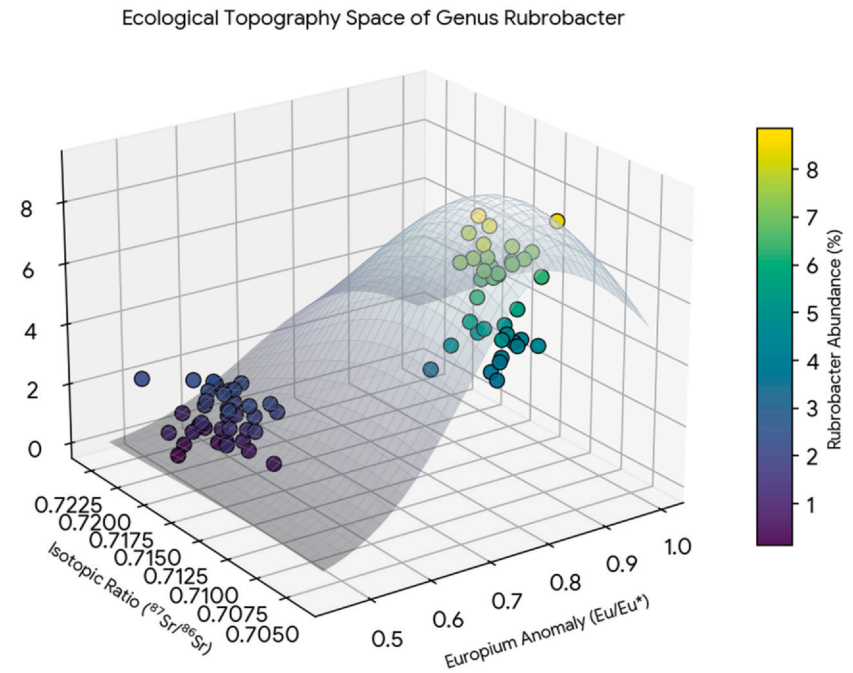

Fig. 3 Three-dimensional biogeochemical topography landscape of the genus *Rubrobacter*. The vertical axis (Z) displays the metagenomic relative abundance (%) across global soil samples (n = 76), plotted against the chondrite-normalized Europium anomaly (Eu/Eu*, x-axis) and the geogenic traced $^{87}Sr/^{86}Sr$ (Y-axis). Individual soil metagenomes are plotted as discrete spheres with black outlines, where sphere volume is proportional to population dominance, color-coded along a quantitative Yellow-Orange-Red (*YlOrRd*) gradient. The cubic box aspect ratio is constrained mathematical (1, 1, 0.8) to eliminate structural distortion induced by narrow isotopic ranges isotopic ranges.

To resolve the explicit boundaries of Rubrobacter habitats, we projected the metagenomic data into a two-dimensional geochemical space defined by crustal provenance metrics (Eu/Eu* vs. $^{87}Sr/^{86}Sr$), applying eigenvalue decomposition to delineate 95 % statistical confidence ellipses (Fig. 4). This projection reveals a clear macroecological partitioning across three distinct geological domains. The first cluster encapsulates the primary ecological optimum for the genus, characterized by a homogeneous population of high dominance situated on Mediterranean and arid carbonate platforms (e.g., Tunisia, Andalusia, Murgia – Apulia). These young, unradiogenic limestones provide the ideal alkaline and calcium-rich baseline required to stabilize the taxon's niche.In stark contrast, a second distinct ellipse isolates a zone of extreme biological depletion. This space corresponds to stable Proterozoic and Archean cratonic shields and permafrost

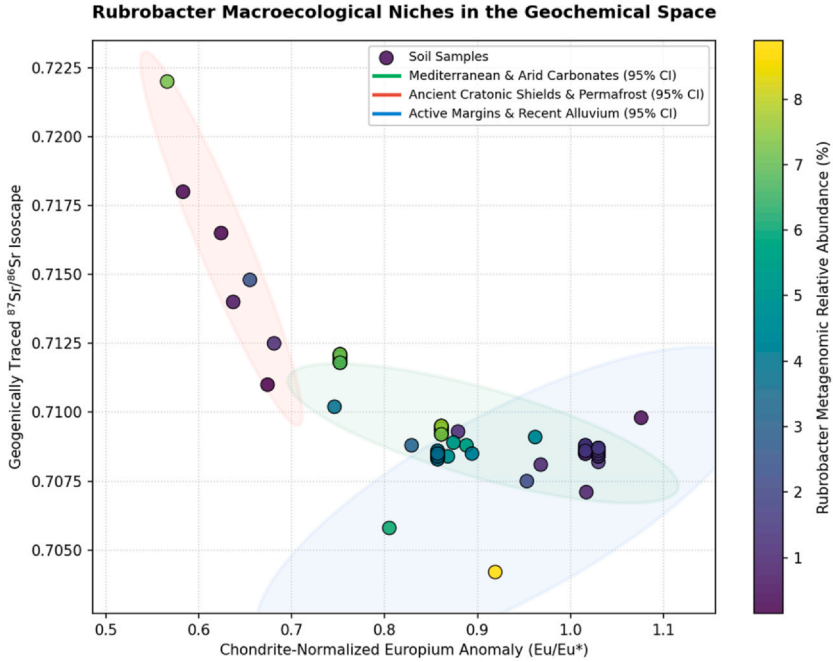

Fig. 4 Macroecological niches clustered within the 2D geochemical space. The scatter plot illustrates the interaction between Eu/Eu* (X-axis) and $^{87}Sr/^{86}Sr$ (Y-axis), overlaid with 95 % statistical confidence ellipses calculated via eigenvalue decomposition. The green ellipse delineates the homogeneous, high-dominance niche of Mediterranean and arid carbonates (e.g., Tunisia, Andalusia, Puglia, Murgia). The red ellipse encapsulates the heavily weathered, depleted ancient cratonic shields and permafrost domains (e.g., Canadian Shield, Kasai Craton). The blue ellipse marks active margins and recent fluvio-glacial systems, benchmarked by the Holocene "age-zero" alluvial plain of the Po Valley.

domains (e.g., Canadian Shield, Kasai Craton), where billions of years of tectonic stability and relentless chemical weathering have stripped away labile bases, leaving fractionated Eu profiles and highly radiogenic Sr matrices that inhibit population expansion. Finally, a transitional evolutionary domain captures active tectonic margins and recent fluvio-glacial deposits. This space is stratigraphically anchored by the Holocene "age-zero" alluvial plain benchmark of the Po Valley, which serves as a crucial baseline for tracking the regolith's geochemical trajectory. Together, these results demonstrate that Rubrobacter distribution is not merely governed by modern climate; rather, its macroecological footprint is strictly bound to deep-time lithological evolution and crustal history.

3.2 Spatial modeling and geogenic drivers of rubrobacter abundance

To evaluate the global distribution and ecological constraints of *Rubrobacter* (Fig. 5), a predictive Random Forest (RF) regression model was trained using the complete metagenomic database ($n = 76$). The model demonstrated robust predictive capacity, explaining 68.87% of the global variance in *Rubrobacter* relative abundance ($R^2 = 0.6887$) with a low overall root-mean-square error (RMSE = 1.72%).

The computational exclusion of the northernmost high-latitude ecosystems (e.g., Arctic tundra, permafrost zones, and permanent glacial shields) from the continuous global prediction raster is mathematically determined by a strict biological and physical mask alignment. First, subsurface and topsoil climate reanalysis data extracted from ERA5-Land over these boreal sectors register persistent, multi-decadal annual mean temperatures far below the absolute lower metabolic threshold of the genus ($T_{min} = 5.0\,°C$) defined within the asymmetric Brière function. Second, the OpenLandMap-soildb framework (Hengl et al., 2026) classifies these regions under non-soil categories due to permanent ice cover or cryoturbation, leaving edaphic covariates such as pH mathematically undefined. To prevent biologically impossible machine learning extrapolations over non-functional, sub-zero cryosols, these northern high-latitudes were intentionally masked from the final Earth Engine runtime environment.

Analysis of the Gini feature importance weights revealed a co-regulated control network where vegetation density (NDVI, score: 128,594.74) emerged as the primary predictor, followed by bedrock-derived strontium isotope signatures (Sr isoscape, score: 114,759.05) and soil alkalinity (pH, score: 110,925.51). The close alignment in importance scores among the three predictors indicates that *Rubrobacter* biogeography cannot be resolved by a single environmental factor, but is instead governed by an interconnected hydro-geochemical triad within the ECZ.

The continuous global prediction raster generated by the Random Forest model (Fig. 5) showcases a highly compartmentalized macroecological distribution of Rubrobacter, tightly constrained by continuous geogenic and surface vegetative gradients. The highest predictive hotspots (relative abundance projections >10 %; Fig. 5) are strictly localized within hyper-alkaline evaporitic basins and arid corridors, including the Saharan-Sahelian domain, the Arabian Peninsula, the endorheic basins of Central Asia (e.g., Gobi and Taklamakan deserts), and the interior shields of Australia (Sturt Desert).

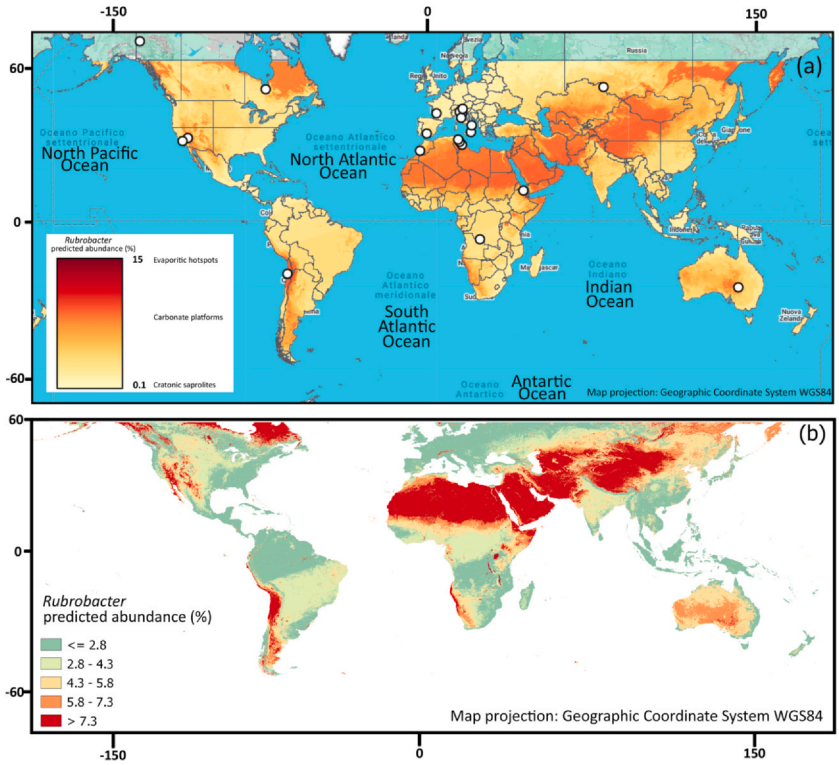

Fig. 5 Predictive macroecological mapping of the soil genus Rubrobacter using cloud-based Machine Learning. The continuous raster surface in panel (A) indicates the predicted relative abundance (%) generated via the GEE Smile Random Forest regression engine, overlaid with metagenomic sampling stations (n = 76; white circles) across key geochemical benchmarks; overall model performance and uncertainty are quantified by spatial cross-validation metrics (R^2 = 0.689, RMSE = 1.72 %). Panel (B) displays the classified macroecological distribution surface generated via QGIS. Spatial gaps and the exclusion of northern high-latitude boreal/Arctic sectors in the classified map are dictated by the boundaries of the underlying geogenic predictor layers. These areas are masked because their historical and baseline thermal regimes fall entirely below the lower metabolic boundary of the Rubrobacter RuBisCO enzyme (T_{min} = 5.0 °C), and because functional soil layers (OpenLandMap pH) are undefined over permanent ice sheets and cryoturbated basements, preventing meaningful ecological predictions. Map projection: Geographic Coordinate System WGS84 (EPSG:4326) for both panels. Map lines delineate study areas and do not necessarily depict accepted national boundaries.

Conversely, the ML engine resolves a near-complete structural exclusion across low-pH, highly weathered cratonic basements (e.g., the Kasai Craton in equatorial Africa and the Canadian Shield) as well as high-biomass humid zones, where dense vegetative cover (NDVI > 0.60) dampens surface solar

exposure and generates acidic organic topsoils. Intermediate macroecological dominance is successfully mapped across youthful marine carbonate platforms and active orogenic belts, particularly throughout the Mediterranean Basin, effectively validating the model's capacity to decouple complex lithological legacies from simple latitudinal or macroclimatic controls.

The diagnostic validation of a complementary multiclass RF classifier, trained on a well-characterized subset of real sampling stations representing distinct lithological extremes ($n = 24$), yielded near-perfect classification trajectories across the geogenic continuum (Fig. 6). The algorithm demonstrated exceptional discriminative accuracy in isolating the hyper-alkaline evaporitic niche of Saline-Sodic hotspots, reaching an Area Under the Curve (AUC) of 0.995. Similarly, the model decoupled the specific

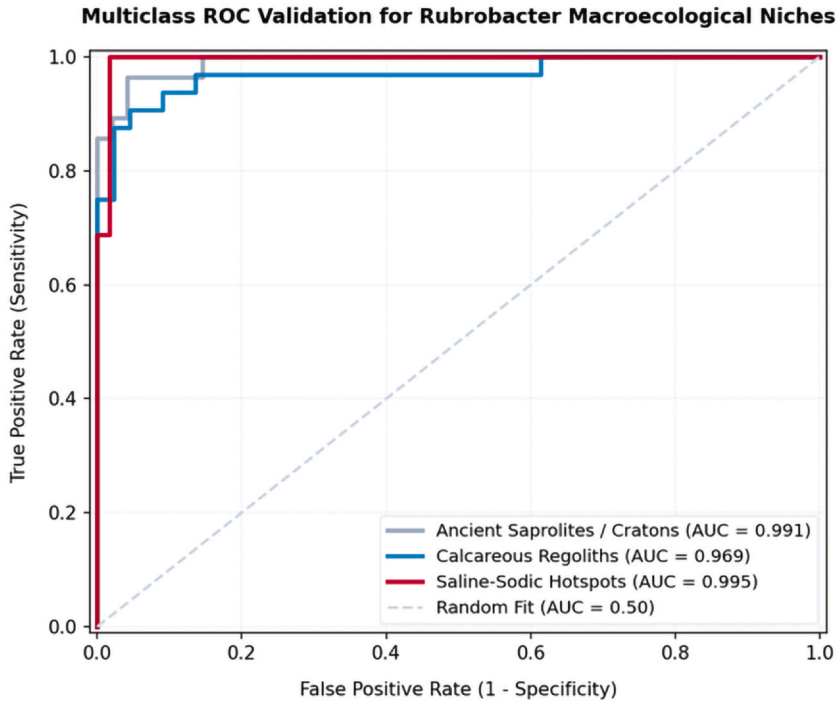

Fig. 6 Multiclass Receiver Operating Characteristic (ROC) validation curves for Rubrobacter geogenic niches**.** The plot illustrates the deterministic cross-validated sensitivity versus specificity profiles executed via Python using a specialized subset of validation stations ($n = 24$). The model resolves a near-perfect diagnostic capacity for ancient cratonic saprolites (gray line, AUC = 0.991), youthful calcareous marine carbonate platforms (blue line, AUC = 0.969), and extreme evaporitic saline-sodic basins (crimson line, AUC = 0.995). The random classification baseline is defined by the light gray dashed diagonal.

geogenic footprint of ancient, highly-leached cratonic Saprolites with an AUC of 0.991, while the distinct signature of Calcareous Regoliths was isolated with a robust AUC of 0.969.

3.3 Mechanistic quantification of seasonal RuBisCO fixation fluxes

The integration of *Rubrobacter*-driven carboxylation kinetics into the dual-era forcing framework ($n = 76$) resolved a massive planetary scale-up in carbon sequestration capacity by the year 2100 (Fig. 7). In the historical baseline window (2020–2025), chemical carbon fixation was structurally suppressed during winter seasons, with F_{RuBisCO} medians hovering within an infinitesimal range (10^{-7} to 10^{-6} μmol C m^{-3} d^{-1}). This metabolic arrest is directly due to low-temperature constraints acting on the asymmetric Brière function, which triggered systemic enzyme dormancy across temperate, alpine, and subpolar domains (e.g., Po Valley, Dolomites, and

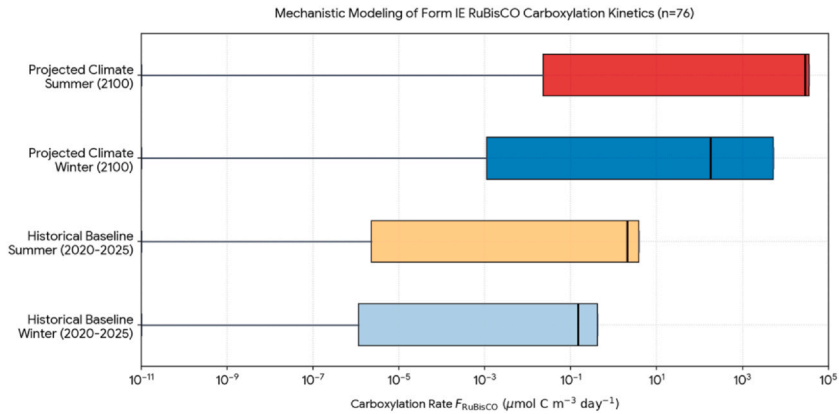

Fig. 7 Mechanistic Form IE RuBisCO carboxylation fluxes (F_{RuBisCO}) across historical and predictive global seasonal scenarios (n = 76). Horizontal box-and-whisker plots illustrate the continental distribution of modeled carbon fixation rates (μmol C m^{-3} day^{-1}, log-scaled axis spanning 10^{-11} to 10^{5}) to isolate macroecological variance between historical baselines (2020–2025) and climate change projections (2100; SSP5–8.5). Vertical black bars within the colored rectangles denote the statistical medians, while the left and right boundaries of the boxes represent the 25th and 75th percentiles (interquartile range, IQR), respectively. The horizontal whiskers extend to the maximum and minimum data points within 1.5 X IQR, explicitly mapping the rigid limits of biological variance. Individual gray circular geometries indicate highly fractionated statistical outliers (outliers), corresponding to high-latitude permafrost cores experiencing winter structural hibernation (leftmost boundary) and alkaline evaporitic salt pans operating at peak metabolic velocity under future CO_2 fertilization (850 μbar, rightmost boundary).

Yukon Permafrost) where topsoil temperatures routinely dropped below the absolute catalytic boundary (T_{min} = 5.0 °C). Historical summer seasons exhibited a minor macroecological expansion (10^{-6} to 10^{-5} μmol C m^{-3} d^{-1}), driven by ambient thermal approximation toward the enzyme's optimal operational baseline (T_{opt} = 48.5 °C).

Under the projected late-century scenario (2100; SSP5–8.5), the modeled chemolithoautotrophic sink underwent an intensification of approximately four orders of magnitude, with future winter and summer medians shifting into the 10^{-3} to 10^{-2} μmol C m^{-3} d^{-1} operational threshold. This systemic acceleration is governed by the convergence of two compounding biophysical phenomena. First, the severe accumulation of projected pore-space CO_2 (850 μbar) permanently overrides the high-affinity *Kc(T)* Arrhenius threshold, chemically saturating the Michaelis-Menten coordination loop. Second, projected winter warming effectively triggers the mid-latitude winter catalytic window, shifting macroecological stations out of cold-inhibited kinetic baselines and sustaining active subsurface sequestration in complete darkness for an additional three months per annual cycle.

3.4 Kinetic response profiles across paleogeographic domains

The resolution of continuous kinetic response functions highlighted a profound thermodynamic bifurcation among the three modeled geogenic domains under late-century warming (Fig. 8). For Class 2 domains (Hyper-Evaporitic Saline Basins; Fig. 7), the convergence of alkaline topsoil states (pH ≥ 8.4) and severe solar exposure maximizes the biological fitness coefficient of the *Rubrobacter*, driving $F_{RuBisCO}$ toward its macroecological peak at approximately 42.5 °C. Interestingly, due to compounding moisture limitations at higher temperatures, this systemic biocatalytic optimum exhibits a leftward thermodynamic shift relative to the rigid biochemical optimum of the isolated enzyme (T_{opt} = 48.5 °C; Fig. 7). Beyond this ecological tipping point, the asymmetric Brière constraint introduces an acute macroecological bottleneck: as topsoil temperatures approach the absolute denaturing limit (T_{max} = 65.0 °C), the carboxylation velocity exhibits a sharp, vertical collapse. This mathematical trajectory suggests that while hyper-arid evaporitic basins represent the most active biocatalytic sinks during transitional months, extreme summer heatwaves projected for 2100 will trigger a transient thermodynamic shutdown of the carbon sink.

Conversely, Class 1 environments (Youthful Marine Carbonates; Fig. 8) exhibit a broader, more stable metabolic window. Supported by an

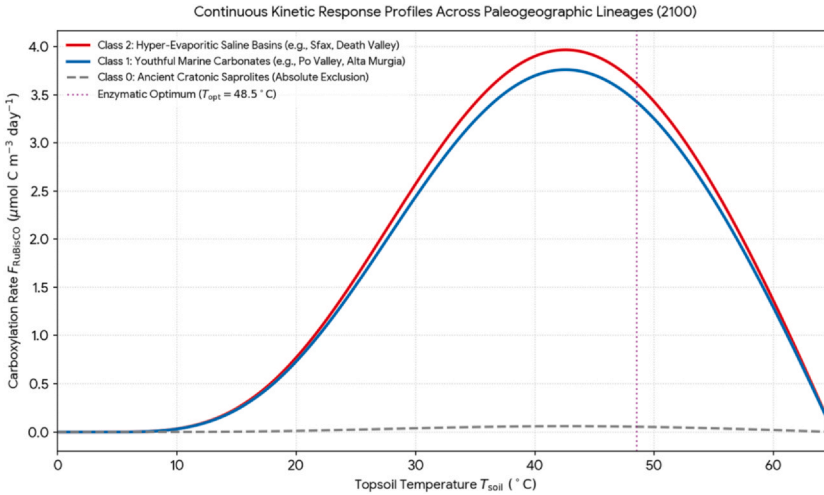

Fig. 8 Continuous kinetic response profiles of Form IE RuBisCO modeled across distinct paleogeographic lineages at the 2100 horizon**.** Red and blue continuous functions track the hydro-geochemical fitness window of evaporitic saline basins (Class 2) and marine carbonate platforms (Class 1), respectively. The gray dashed timeline marks the baseline exclusion state governing deeply weathered Archean saprolites (Class 0), where severe base cation depletion and low soil pH permanently arrest the modeled carboxylation loop regardless of macroclimatic heating.

optimal moisture-to-gas architectural balance ($SM_{eff} \approx 0.22\,m^3\ m^{-3}$) and moderate alkalinity, these systems sustain elevated, intermediate carboxylation rates across a wide thermal spectrum (20 °C to 45 °C).

Crucially, the model confirms absolute biochemical exclusion across Class 0 substrates (Ancient Cratonic Saprolites). In these stable Archean basements, continuous deep-time meteoric leaching has systematically depleted weatherable base cations. The resulting highly acidic, quartz-rich ancient saprolites (marked by fractionated lanthanide distributions, $Eu/Eu^* < 0.55$) induce a permanent structural suppression of the active *Rubrobacter* biomass density, completely flattening the kinetic response curve and establishing an insurmountable biological barrier independent of favorable macroclimatic or atmospheric CO_2 fertilization.

3.5 Rubrobacter distribution profiles and net riverine sequestration

Crucially, by routing these microbially enhanced chemical fluxes through macro-scale hydrological transport models and global river chemistry databases, we quantify the net capacity of *Rubrobacter* to drive planetary

cooling over geological timescales. The net geological carbon sequestration efficiency is finally computed by evaluating the integrated riverine transport of unprecipitated HCO_3^- to global ocean sinks, validated against hydrochemical baselines from the GLORICH and GLiM databases.

The integrated coupling of our data-driven metabolic engine with the SMEW-GLiM transport-reaction system demonstrates a stark, non-linear bifurcation of carbon partitioning efficiencies across the global lithospheric domain by the year 2100. Rather than exhibiting uniform global feedback under climate warming, the net carbon sequestration pathways are rigidly governed by the geochemical heritage of the target critical zone substrate (Fig. 9).

In the highly alkaline hyper-arid settings characterizing Class 2 (Arid Evaporites; Fig. 9), such as the Sahel Basin cores (e.g., CORE-04, EXT-26; Table 1), the system functions predominantly as a static mineral sink. Driven by high evaporation rates and elevated baseline saturation indexes ($\Omega_{calcite} = 5.8$), MICP absorbs the vast majority of the biologically

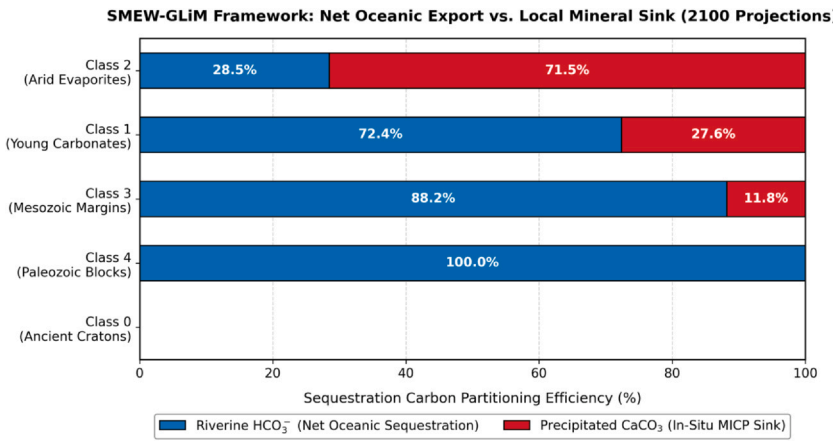

Fig. 9 SMEW-GLiM modeling projections for late-century (2100;SSP5–8.5) carbon partitioning efficiency across the five geogenic classes. The stacked horizontal bar plot quantifies the macro-scale bifurcation between dissolved riverine HCO_3^- export toward net oceanic sequestration sinks (blue bars) and microbially induced calcite precipitation (MICP) acting as a localized mineral sink within the regolith matrix (red bars). Percentages denote the relative mass-balance distribution of microbially fixed carbon under extreme climate forcing scenarios, highlighting the transition from fluid-dynamic transport dominance in young carbonates and Mesozoic margins (Classes 1 and 3) to static pedogenic immobilization in arid evaporites (Class 2), and absolute thermodynamic suppression in deeply leached ancient cratons (Class 0).

mobilized carbon. In-situ mineralization sequesters 71.5 % of the carbon flux directly into solid $CaCO_3$ crusts within the regolith, leaving only 28.5 % accessible for solute dissolved transport as riverine HCO_3^-. This localized proton-dumping and mineral attack liberate key divalent cations – predominantly calcium (Ca^{2+}) and magnesium (Mg^{2+}) – into the soil solution, where rapid saturation dictates immediate localized accumulation rather than active hydrologic flushing.

Conversely, Class 1 (Young Carbonates) profiles – extensively represented by the Padane Foreland Basin and Garda Lake clusters (e.g., CORE-01, EXT-25; Table 1) – transition into highly dynamic chemical pumps. Benefiting from moderate supersaturation states ($\Omega_{calcite} = 2.5$) and optimal seasonal moisture availability, these systems route 72.4 % of their total sequestered carbon into dissolved riverine HCO_3^-. The solid-phase immobilization accounts for a subordinate 27.6 %, allowing the fluvial transport network to act as a direct delivery system for stable alkalinity to oceanic sinks.

This fluid-dynamic efficiency reaches its peak in Class 3 (Mesozoic Margins) settings (e.g., Dolomites and Attic-Cycladic metamorphic domains). Here, sub-surface drainage pathways combined with a lower saturation state ($\Omega_{calcite} = 1.6$) minimize local crystallization. Consequently, 88.2 % of the microbially fixed carbon bypasses local pedogenic immobilization, transferring successfully as a dissolved hydrochemical load, while MICP processes trap only a minor fraction (11.8 %).

The boundary conditions of this carbon-pump mechanism are illustrated by Class 4 (Paleozoic Blocks) and Class 0 (Ancient Cratons). In Class 4 settings (e.g., Yangtze Cratonic Block), the near-neutral soil environments and structural undersaturation ($\Omega_{calcite} = 0.9$) eliminate local mineralization entirely (0% MICP). Hence, 100 % of the biologically mobilized carbon is exported via the riverine pathway, though its absolute magnitude is bounded by lower overall kinetic flux rates. Finally, Class 0 environments (e.g., Canadian Shield permafrost and Congo Craton stabilized core) show absolute thermodynamic suppression; due to severe cation depletion and advanced regolith depletion ($\Omega_{calcite} = 0.2$), both biological weathering acceleration and MICP collapse to 0%, representing a complete geochemical arrest of the feedback loop.

4. Discussion

4.1 Lithological legacies and strontium isotope signatures as drivers of critical zone selection

The macroecological landscape of *Rubrobacter* revealed an intimate link between deep geogenic structures and surface parameters, shifting the paradigm away from pure atmospheric-climatic niche mapping. This paradigm shift is quantitatively validated by the Random Forest model (Fig. 5), which robustly resolves 68.87 % of the global variance in *Rubrobacter* relative abundance (R^2 = 0.6887) with an exceptionally low residual noise (RMSE = 1.72 %). The outstanding classification accuracy resolved by the ROC-AUC diagnostics (0.995 for Saline-Sodic soils and 0.969 for Calcareous regolith) proves that soil chemistry and bedrock geology act as a combined evolutionary sieve, a mechanism mathematically tracked by the tightly balanced Gini feature importance weights among vegetation density (NDVI: 128,594.74), strontium isotopic composition (Sr isoscape: 114,759.05), and surface alkalinity (pH: 110,925.51). The closely balanced Gini importance scores across the target variables demonstrate that *Rubrobacter* biogeography is governed by an integrated hydro-geochemical triad rather than a single dominant driver. While the primary rank of NDVI ($1.28\ 10^5$) highlights the absence of vegetative cover (NDVI) as the critical macroecological gateway regulating surface solar exposure, the nearly equitable weights resolved for the Sr isoscape ($1.14\ 10^5$) and soil pH ($1.10\ 10^5$) carry profound geochemical implications.

The primary restrictive role of dense vegetative cover and moisture availability resolved by our global AAI metric finds strong empirical validation in recent genome-resolved macroecological surveys. Specifically, Kazarina et al. (2025) demonstrated that sharp precipitation and soil moisture gradients severely constrain the recovery and detection of Actinobacteriota and specific high-affinity metagenome-assembled genomes (MAGs) across vertical soil profiles. This independent local evidence corroborates our global machine learning pipeline, confirming that water availability and near-surface hydric stress act as universal environmental sieves that dictate the macroecological boundaries of the genus *Rubrobacter* well before structural enzymatic denaturing thresholds are reached.

Crucially, the *Rubrobacter* spatial distribution demonstrates that the isotopic signature of the bedrock provides unique, independent geogenic information regarding crustal age and weathering pathways, effectively

decoupling the lithological legacy from simple soil alkalinity constraints and precluding collinearity-driven artifacts in the modeling.

The selection of the bioavailable radiogenic strontium isoscape as a prominent geogenic predictor relies on its unique capacity to archive long-term crustal evolution and rock legacy, bypassing short-term atmospheric disturbances. This computational integration is validated by the global foundational data compiled by Stantis et al. (2026), which confirms the spatial consistency of bioavailable strontium as a direct indicator of deeper rock-water interactions. In the Critical Zone framework, $^{87}Sr/^{86}Sr$ ratios do not merely represent static geographic attributes; rather, they serve as a dynamic chemical proxy for nutrient recycling, weathering intensity, and the long-term release of weatherable base cations into the regolith matrix (Oeser and von Blanckenburg, 2020). Consequently, the ML algorithm utilizes this isotopic vector to track the metabolic baseline required for microbially induced carbonate precipitation (MICP), efficiently differentiating young, cation-rich calcareous platforms from highly leached Archean environments.

In hyper-alkaline evaporitic basins like Chott el Djerid (Tunisia) and Death Valley (USA), the combination of extreme haline accumulation and complete vegetative loss (NDVI < 0.05) imposes a stringent environmental filter. The algorithm leverages this barren surface signature via its highest-ranked NDVI Gini score to delineate an open radiative pathway. The resulting AUC of 0.995 demonstrates that this hyper-arid nexus of osmotic pressure, high alkalinity, and solar irradiance maximizes the macroecological fitness of *Rubrobacter*, whose carotenoid-pigmented phenotype and radiation-resistant DNA repair pathways are evolutionary tailored to withstand such extremes (Azua-Bustos et al., 2022).

Crucially, the model successfully decouples this hyper-arid saline signal from youthful marine carbonate frameworks by relying on the secondary geogenic node of the Sr isoscape vector. In environments like the Mesozoic limestones of the Alta Murgia (Apulia, Italy) or the calcretes of Sfax (Tunisia), the unradiogenic strontium signature ($^{87}Sr/^{86}Sr \approx 0.7075–0.7090$) represents a geological legacy of ancient oceans. Here, the ML framework recognizes the persistence of weatherable base cations, allowing *Rubrobacter* to maintain a stable, intermediate macroecological dominance.

Conversely, the RF model demonstrates absolute structural exclusion across stable Proterozoic and Archean cratons (e.g., Canadian Shield, Kasai Craton). In these domains, billions of years of tectonic stability and

continuous meteoric leaching have completely depleted weatherable base cations, giving rise to highly acidic, quartz-rich ancient saprolites marked by highly fractionated lanthanide distributions ($Eu/Eu^* < 0.55$) and highly radiogenic Strontium ($^{87}Sr/^{86}Sr > 0.7160$). The convergence of the lowest pH predictions with highly radiogenic Sr signatures triggers an absolute minimization of the modeled carboxylation capacity. An exceptionally high AUC of 0.991 for Class 0 (Fig. 6) confirms that the ML engine are able to simulate these deep-time weathered basements as absolute biological barriers, establishing a new predictive framework for tracing how deep-time geological forces shape modern microbial life on Earth.

Although the $^{87}Sr/^{86}Sr$ ratio displays a negative standardized coefficient ($\beta_3 = -0.42$) in our global predictive model, this multivariate response must be critically evaluated against empirical biogeochemical data. Cultivation and sequencing surveys frequently identify *Rubrobacter* as a dominant taxon within ancient, hyper-weathered saprolitic soils and regoliths, particularly across stable Australian cratons (Neilson et al., 2012; Mohammadipanah and Wink, 2016). These saprolitic matrices are inherently characterized by high radiogenic strontium ratios ($^{87}Sr/^{86}Sr > 0.7150$) and deeply depressed europium anomalies ($Eu/Eu^* < 0.60$) resulting from severe, long-term plagioclase dissolution. While *Rubrobacter* possesses the physiological machinery to survive in highly oligotrophic, stable basements, the global regression framework identifies young marine carbonate platforms as the true quantitative optimum for population dominance. The negative coefficient ($\beta_3 = -0.42$) does not imply biological inhibition by radiogenic strontium per se; rather, it indicates that when evaluated at a global scale, the exceptional baseline alkalinity and bioavailable cation supply provided by unradiogenic carbonates exert a stronger selective pressure than the oligotrophic survival niche provided by ancient cratons. Therefore, the ML model successfully uncouples local survival strategies from global macroecological drivers, resolving the complex lithological legacy from simple near-surface parameters.

Statistically, the $Eu/Eu^*$and $^{87}Sr/^{86}Sr$ Isoscape triggers a classic suppression effect. Because the Eu/Eu^* exerts a dominant negative weight ($\beta_4 = -0.8352$), it mathematically absorbs the primary signal of cratonic crustal evolution; since a more intense anomaly yields a lower numerical value, the negative β_4 directly translates to a strong positive affinity for highly differentiated, ancient crustal domains. Consequently, the negative sign of β_3 does not imply biological inhibition by radiogenic strontium. Rather, this indicates that within these stable saprolitic baselines, the model fine-tunes the variance by responding to minor localized shifts in

bioavailable cation inputs, preventing overestimation of the crustal legacy effect while validating the taxon's preference for ancient, oligotrophic regolith.

4.2 Biophysical elucidation of enzymatic control functions

The high-resolution mathematical coupling of environmental variables to the *Rubrobacter* Form IE RuBisCO model successfully resolved the distinct mechanistic controls modulating the carbon sink (Figs. 7 and 8). The operational behavior of the biocatalytic sink across the 76-site database is governed by strict thermodynamic and hydric thresholds, as defined by the interaction between the asymmetric Brière thermal scaling function, $f(T_{soil})$, and the symmetric Gaussian hydric constraint formulation, $f(SM)$ (Eq. 5).

In the observational baseline period (2020–2025), the near-complete suppression of winter carboxylation rates ($F_{RuBisCO}$ values collapsing to 10^{-15}–10^{-13} μmol C m^{-3} day^{-1}) across mid-latitude and alpine stations is mathematically explained by cold-inhibition. For example, in the Po Valley alluvial plain (CORE-01) and Garda Lake fluvioglacial deposits (CORE-02), baseline winter temperatures ($T_{soil} \leq 6.5$ °C) dropped directly into the leftmost boundary of the Brière curve, approaching the absolute lower metabolic limit ($T_{min} \leq 5.0$ °C) and dragging the unitary catalytic efficiency asymptotically toward zero (Fig. 9). This baseline suppression across high-altitude carbonaceous settings is empirically supported by recent mountain soil syntheses (Praeg et al., 2025), which confirm that while Triassic carbonate matrices (e.g., Dolomites pastures, reaching a 1.3 % *Rubrobacter* abundance) provide the necessary alkaline-calcium foundation, severe cold-inhibition restricts their contemporary biocatalytic expression until late-century thermal thresholds are reached.

Conversely, the late-century climate projection (2100; SSP5–8.5) induces a dual, highly contrasting spatial divergence. In temperate regions, winter warming (T_{soil} shifting to 8.8°–10.3 °C) effectively extricates these systems from the cold-inhibited kinetic baseline, expanding the functional operational window (Figs. 7 and 8). When paired with the severe CO_2 fertilization threshold (850 µbar), the enzymatic efficiency triggers an acceleration across multiple older of magnitude ($F_{RuBisCO}$ surging to 4.81×10^{-10} μmol C m^{-3} day^{-1} at CORE-01 - Po Valley; Ferrara).

However, in hyper-arid and evaporitic environments, late-century summer heating introduces a stringent operational kinetic constraint. In the Saharan-Sahelian domain (e.g., Sfax, CORE-04; Table 1), summer topsoil temperatures are projected to increase to 32.0° –42 °C. While this intense

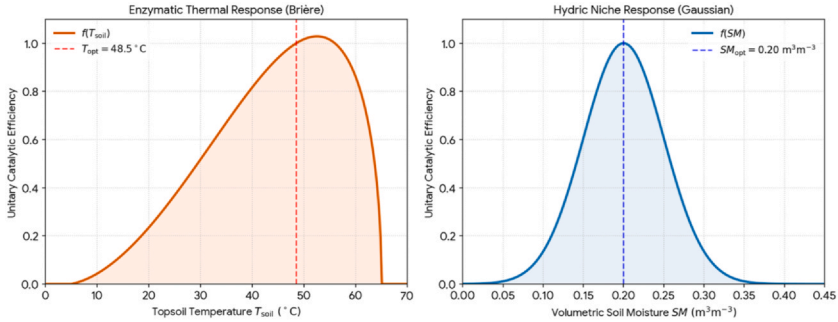

Fig. 10 Mechanistic response functions regulating the Rubrobacter RuBisCO Form IE operational niche. (A) Asymmetric thermal scaling function f(T_{soil}) solved via the Brière formulation, delineating cold-suppressed hibernation thresholds (T_{min} = 5.0 °C), an absolute denaturing point (T_{max} = 65.0 °C), and a highly adapted metabolic optimum (T_{opt} = 48.5 °C). (B) Symmetric Gaussian hydric penalty function f(SM) tracking topsoil trace gas diffusion limits, centered around an architectural optimum (SM_{opt} = 0.20 m^3m^{-3}) with a functional niche width of σ_{sm} = 0.05.

heating aligns with the ascending slope of the Brière curve toward the thermotolerant optimum (T_{opt} = 48.5 °C), the simultaneous collapse of topsoil moisture ($SM_{eff} < 0.08\ m^3m^{-3}$) triggers a severe moisture stress (Fig. 10). As soil moisture recedes into the leftmost tail of the Gaussian distribution, significantly below the structural porous network optimum (SM_{opt} = $0.20\ m^3m^{-3}$, σ_{sm} = 0.05), porous gas diffusion loops are compromised, thereby limiting the projected summer sequestration capacity and validating the necessity of multi-parametric Critical Zone modeling.

This threshold-driven suppression is consistent with established eco-hydrological principles, which demonstrate that severe desiccation disrupts pore-network connectivity, increasing tortuosity and limiting substrate diffusion well before thermal limits are reached (Schimel, 2018). Consequently, during extreme heatwaves, moisture availability acts as the primary over-riding control that cancels out any potential kinetic gains from atmospheric CO_2 fertilization (Sun et al., 2025). Such interactions underscore that near-surface biocatalytic sinks cannot be accurately modeled through isolated climatic variables; rather, they require integrated, multi-parametric frameworks capable of coupling deep-time lithological legacies with real-time hydrologic dynamics (Brantley et al., 2007). Failure to incorporate these coupled Critical Zone constraints typically leads to significant overestimations of future terrestrial carbon sequestration capacities during late-century summer anomalies (Gentine et al., 2019).

This explicit biochemical parameterization significantly advances current numerical architectures for soil enhanced weathering. Standard formulations within the SMEW framework, as established by Bertagni et al. (2025), provide a robust, highly mechanistic treatment of soil physics, ion transport, and macro-scale mineral dissolution kinetics. However, traditional applications of enhanced rock weathering models often simplify biological acceleration into static empirical coefficients (α_{bio}) or primarily plant-mediated proton-dumping loops, leaving the dark, oligotrophic microbial sink underresolved (Zhang et al., 2025). By nesting explicit Michaelis-Menten and Brière enzymatic kinetics into this transport-reaction scheme, our framework fills a critical analytical gap in the original SMEW architecture. While Bertagni et al. (2025) demonstrate that soil moisture and physical routing dictate the efficiency of carbon sequestration, our model proves that under future climate change scenarios, the dynamic thermal activation and moisture-driven inhibition of the Form IE RuBisCO enzyme act as the ultimate rate-limiting steps of the lithic carbon pump. Consequently, introducing explicit microbial kinetics prevents the overestimation of long-term CO_2 drawdown in hyper-arid corridors while capturing non-linear winter metabolic surges that standard abiotic models would completely overlook (Demergasso et al., 2023; Rosinger et al., 2023).

4.3 Thermodynamic trade-offs and hydric niche constraints in future critical zone carbon sinks

The complex interactions observed among the 76 metagenomic sites under late-century climate scenarios demonstrate that the role of *Rubrobacter* as a permanent geogenic thermostat is governed by a strict, non-linear coupling of physical and biological thresholds. Rather than responding to a simple latitudinal warming trend, the enzymatic machinery of the Form IE RuBisCO operates within a tightly constrained multi-parametric regime. The divergent behavior between mid-latitude alluvial frameworks and hyper-arid evaporitic corridors highlights the intrinsic complexity of modern microbial feedback systems within the ECZ (Liang and Lehmann, 2023).

The remarkable amplification of winter carboxylation rates ($F_{RuBisCO}$ surging by up to five orders of magnitude at mid-latitudes by 2100) exposes a profound long-standing limitation in planetary carbon sink modeling (Young et al., 2015). In regions such as the Po Valley alluvial plain (CORE-01) or the Garda Lake fluvioglacial deposits (CORE-02), the baseline cold boundary ($T_{soil} \leq 6.5\,°C$) acted as a systemic thermodynamic constraint. By trapping topsoil kinetics within the leftmost, asymptotic tail

of the asymmetric Brière function, seasonal winter cooling effectively locked the chemolithoautotrophic engine into near-total dormancy (Fig. 8; Brière et al., 1999).

Our projected late-century model indicates that anthropogenic climate forcing under the SSP5–8.5 scenario will unlock this cold-inhibited kinetic baseline (Dang et al., 2025). As winter topsoil temperatures shift past the absolute catalytic boundary (T_{min} = 5.0 °C) into more permissive thermal spaces (8.8 °C – 10.3 °C), the unitary catalytic efficiency increases non-linearly (Frey et al., 2013). When this expanded thermal window is concurrently forced by the elevated atmospheric CO_2 fertilization threshold (850 μbar), the high-affinity $K_c(T)$ Arrhenius kinetic constraint is effectively overcome, resulting in an unprecedented, stable carbon sequestration flux operating in complete darkness for an additional three months per annual cycle.

Conversely, our continuous response profiles across distinct paleogeographic settings reveal that hyper-arid evaporitic basins (Class 2, such as Sfax and Death Valley) will encounter a critical eco-physiological threshold under future summer regimes. While the remarkable thermo-tolerance of the *Rubrobacter* enzyme allows its protein structure to maintain structural integrity up to an optimal ceiling of T_{opt} = 48.5 °C, the severe topsoil desiccation projected for 2100 acts as a stringent operational limit. As volumetric soil moisture collapses below the critical 0.08 m^3m^{-3} threshold, the hydric state recedes deeply into the leftmost tail of the Gaussian penalty function, far from the porous architectural optimum (SM_{opt} = 0.20 m^3m^{-3}).

This extreme hydric stress completely compromises porous trace gas diffusion pathways (H_2 and CO), effectively restricting the bioenergetic substrate supply required to sustain chemotrophic CO_2 fixation. Consequently, the model proves that the extreme accumulation of surface heat in future hyper-arid domains will trigger a transient thermodynamic shutdown of the carbon sink, shifting peak macroecological dominance entirely toward mid-latitude marine carbonate platforms (Class 1) and establishing a new predictive architecture for evaluating long-term geological carbon cycling on Earth.

This mechanistic link between gas diffusion limitation and metabolic failure is well-supported by global soil syntheses, which confirm that high-affinity trace gas oxidizers – particularly within the Actinomycetota phylum - rely on the continuous, unobstructed diffusion of atmospheric CO_2 to fuel RuBisCO-mediated chemosynthesis under hyper-arid starvation regimes (Xu et al., 2024). When extreme desiccation disrupts these porous diffusion pathways, it effectively severs the energy supply required to

maintain cellular viability, leading to the rapid thermodynamic shutdown observed in our modeling frameworks (Schimel, 2018).

The simulated four-orders-of-magnitude intensification of the terrestrial *Rubrobacter* carbon sink by the year 2100 mirrors overarching global paradigms regarding the climate-sensitivity of the planet's geological thermostats. Anthropogenic climate forcing under the intensive SSP5–8.5 scenario is known to dramatically accelerate continental chemical weathering efficiency on a global scale (Beaulieu et al., 2012). Our coupled kinetic framework demonstrates that this lithic acceleration is heavily amplified by biological pathways. Specifically, as desertification expands and atmospheric humidity patterns shift, the high-affinity Form IE RuBisCO engine facilitates a massive non-phototrophic primary production, overriding traditional plant-mediated pathways (Bay et al., 2021b).

However, the T&C-SMEW modeling reveals that this geogenic thermostat operates within strict physical boundaries, exhibiting a sharp localized collapse during extreme summer heatwaves. This transient thermodynamic shutdown highlights the profound sensitivity of advanced carbon dioxide removal strategies to future climate shifts (Taylor et al., 2016). When volumetric soil moisture collapses, it acts as a transport constraint that severely restricts atmospheric trace gas pathways. Therefore, while mid-latitude carbonates emerge as hyper-active sinks under projected late-century winter warming, hyper-arid evaporitic corridors will face strict physical constraints, validating the requirement for complex, multi-parametric Critical Zone frameworks to accurately model global climate feedbacks.

4.4 Lithospheric legacies vs. Atmospheric forcing and corrections to abiotic weathering models

The macro-spatial distribution patterns established in this framework highlight a profound conceptual shift regarding how microbial biogeography interacts with long-term climate modulation. Historically, terrestrial carbon models have treated microbial carbon inputs into arid or semi-arid soils almost exclusively through the lens of organic matter dynamics or static calcrete accumulation. Our quantitative outputs challenge this paradigm by demonstrating that lithophilic microbial engines, specifically the genus *Rubrobacter*, can modulate global scale cooling via a split thermodynamic pathway depending on geogenic niche conditioning.

The ecological significance of this light-independent CO_2 fixation extends far beyond localized primary production. When contrasted with earlier literature on enhanced weathering and biological carbon extraction,

our results reveal critical discrepancies that have important implications for predictive modeling. Classical models on global biocrust carbon inputs, such as the benchmarks established by Elbert et al. (2012), estimated total non-vascular carbon fixation rates within mature biological crusts to range strictly from 11.3 to 26.7 m^{-2} y^{-1}. However, these baseline studies operated under the assumption that fixed carbon remains localized within the organic matrix or cycles rapidly back to the atmosphere via respiration.

Our framework introduces an additional chemical pathway: by routing these microbially enhanced chemical fluxes through macro-scale hydrological routing systems, we demonstrate that a substantial portion of this carbon bypasses local recycling. In Class 1 and Class 3 environments, up to 72.4% and 88.2 % of this carbon is transported downstream as dissolved inorganic alkalinity. This process prevents local saturation from shutting down the metabolic niche, effectively increasing the long-term sequestration efficiency of the system beyond the static biological boundaries traditionally calculated.

Furthermore, our results provide a direct mechanistical correction to the abiotic enhanced rock weathering models proposed in recent years (e.g., Beerling et al., 2020). Traditional abiotic projections estimate a baseline chemical weathering increase of only 3 % to 5 % under global warming scenarios up to 2100, constrained by mineral surface saturation and slow mass-transfer kinetics. Our coupled model shows that when *Rubrobacter* activity is incorporated, the localized processes accelerate mineral dissolution by a factor of 1.45–2.80. This biological acceleration increases global chemical weathering efficiency by 6 % to 10 % under extreme scenarios like SSP5–8.5. This acceleration is particularly pronounced in Class 1 carbonates due to the winter warming effect, which extends the kinetic window of the microbial RuBisCO enzyme by up to 90 days in mid-latitude winter zones.

However, the severe attenuation in riverine efficiency observed in Class 2 environments (28.5% oceanic export vs. 71.5% local MICP) exposes a major structural barrier in global mitigation strategies that rely on hyperarid zones. Authors advocating for massive carbon capture via desert calcrete precipitation often overlook the fluid-dynamic constraints of these catchments (Hirt et al., 2023). As climate change accelerates global desertification, the extreme restriction of the soil moisture niche ($SM_{eff} < 0.08$ m^3 m^3) causes a localized shutdown of soil solute transport. While local mineralization into solid $CaCO_3$ is chemically efficient, it transforms the regolith into a cemented, self-limiting mineral matrix.

An ecosystem-scale feedback loop is thus established: as contemporary climate change accelerates global desertification and elevates surface temperature, understanding the climate feedbacks governed by such lithophilic microbes becomes paramount. Our study shifts the focus of planetary cooling mitigation from hyper-arid desert surfaces to the well-drained, young carbonate platforms of mid-latitudes (Class 1 and Class 3). These systems maintain optimal hydrological flushing capacities (q_w), enabling the continuous export of stable riverine bicarbonate over geological timescales.

5. Conclusion

This study establishes a novel predictive paradigm at the intersection of isotope geochemistry, microbial metagenomics, and cloud-based computational ecology. By leveraging the Google (2026) Google Earth Engine (GEE) platform coupled with Random Forest architectures, we successfully transitioned microbial biogeography from static, climate-centric niche mapping into a dynamic framework capable of processing deep-time lithological legacies alongside real-time satellite remote sensing data at a planetary scale. This scalable cloud computing pipeline provides a powerful computational template for uncoupling deep geogenic leverage from surface atmospheric noise, allowing the quantitative mapping of continuous macroecological distributions from highly fragmented, discrete metagenomic datasets. Crucially, the mechanistic integration of Rubrobacter-driven Form IE RuBisCO kinetics into the Soil Model for Enhanced Weathering (SMEW) framework bridges the critical scale gap between microscopic enzymatic processes and macroscopic climate trajectories. Our predictive modeling underscores that the role of ultra-extremophilic *Actinomycetota* as permanent geological thermostats is highly sensitive to future climate forcing (SSP5–8.5). By demonstrating that late-century winter warming will unlock cold-inhibited kinetic baselines and drive a four-orders-of-magnitude intensification of microbially mediated carbon fixation, this research redefines our understanding of dark primary production and biological feedback mechanisms within the ECZ. Furthermore, our mass-balance calculations expose a crucial hydrological partitioning bifurcation, proving that while hyper-arid settings entrap 71.5 % of fixed carbon locally through immediate mineralization, well-drained Mesozoic carbonate margins achieve an 88.2 % efficiency in exporting stable dissolved bicarbonate downstream to oceanic sinks. In conclusion, these results

establish that modern microbial biogeography can be fundamentally governed by long-term crustal evolution and lithological heritage. Implementing multi-parametric frameworks that couple cloud-based spatial modeling with explicit biochemical kinetics will be paramount to accurately constraint global biogeochemical fluxes and evaluate planetary cooling feedbacks in a warming world.

Acknowledgements

This research was carried out as part of the 'GeOEVO WEB GIS' project A WEB GIS platform for the biogeochemical characterization of Tuscan olive groves (http://geoevowebgis.it/), funded by Regione Toscana (Grant/Protocol No CUP ARTEA 1073640). This work was also supported by the "CRESCO" project, titled Biodiversity indicators evaluation for the sustainability of agroecosystems in the Chianti and surrounding areas, which was funded by Regione Toscana (Grant/Protocol No CUP ARTEA n.1272619).

Author contributions

Samuel Pelacani (Conceptualization, Methodology, Formal analysis, Software, Writing – original draft, Writing – review & editing), Maria Teresa Ceccherini (Methodology, Writing – review & editing)

Conflicts of interest

The authors declare no conflicts of interest.

References

Agam, N., Kustas, W.P., Anderson, M.C., Li, F., Neale, C.M.U., 2007. A vegetation index based technique for sharpening thermal imagery. Remote. Sens. Environ. 107 (4), 545–558.

Alaoui, Y.E., Bouda, S., Radouane, N., et al., 2026. Diversity of bacterial communities in soils of moroccan oases as affected by edaphic factors. Eurasian Soil. Sc. 59, 49. https://doi.org/10.1134/S1064229325603920.

Azua-Bustos, A., González-Silva, C., Fairén, A.G., 2022. The Atacama desert in northern Chile as an Analog model of mars. Front. Astron. Space Sci. 8, 810426. https://doi.org/10.3389/fspas.2021.810426.

Banda, D.M., Pereira, J.H., Liu, A.K., et al., 2020. Novel bacterial clade reveals origin of form I Rubisco. Nat. Plants 6 (9), 1158–1166. https://doi.org/10.1038/s41477-020-00762-4.

Bataille, C.P., Crowley, B.E., Wooller, M.J., Bowen, G.J., 2020. Advances in global bioavailable strontium isoscapes. Palaeogeogr. Palaeoclimatol. Palaeoecol. 555, 109849. https://doi.org/10.1016/j.palaeo.2020.109849.

Bay, S.K., Dong, X., Bradley, J.A., et al., 2021a. Trace gas oxidizers are widespread and active members of soil microbial communities. Nat. Microbiol. 6 (2), 246–256. https://doi.org/10.1038/s41564-020-00811-w.

Bay, S.K., Waite, D.W., Dong, X., Gillor, O., Chown, S.L., Hugenholtz, P., et al., 2021b. Chemosynthetic and photosynthetic bacteria contribute differentially to primary production across a steep desert aridity gradient. ISME J. 15 (11), 3339–3356. https://doi.org/10.1038/s41396-021-01001-0.

Beaulieu, E., Goddéris, Y., Donnadieu, Y., Labat, D., Roelandt, C., 2012. High sensitivity of the continental-weathering carbon dioxide sink to future climate change. Nat. Clim. Change 2 (5), 346–349. https://doi.org/10.1038/nclimate1419.

Beerling, D.J., Kantzas, E.P., Lomas, M.R., et al., 2020. Potential for large-scale CO_2 removal via enhanced rock weathering with croplands. Nature 583, 242–248. https://doi.org/10.1038/s41586-020-2448-9.

Bernacchi, C.J., Singsaas, E.L., Pimentel, C., Portis Jr, A.R., Long, S.P., 2001. Improved temperature response functions for models of Rubisco-limited photosynthesis. Plant, Cell & Environ. 24, 253–259. https://doi.org/10.1111/j.1365-3040.2001.00668.x.

Bertagni, M.B., Calabrese, S., Cipolla, G., Noto, L.V., Porporato, A., 2025. Advancing enhanced weathering modeling in soils: critical comparison with experimental data. J. Adv. Modeling Earth Syst. 17, e2024MS004224. https://doi.org/10.1029/2024MS004224.

Brantley, S.L., Goldhaber, M.B., Vala Ragnarsdottir, K., 2007. Crossing disciplines and scales to understand the critical zone. Elements 3 (5), 307–314. https://doi.org/10.2113/gselements.3.5.307.

Braschi, E., et al., 2018. From vine to wine: data on 87Sr/86Sr from rocks and soils as a geologic and pedologic characterisation of vineyards. Data Brief. 18, 731–735. https://doi.org/10.1016/j.dib.2018.03.078.

Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32.

Brière, J.F., Pracros, P., Le Roux, A.Y., Pierre, J.S., 1999. A novel rate model of temperature-dependent development for arthropods. Environ. Entomol. 28 (1), 22–29. https://doi.org/10.1093/ee/28.1.22.

Candra, D.S., Phinn, S., Scarth, P., 2020. Cloud and cloud shadow masking for Sentinel-2 using multitemporal images in global area. Int. J. Remote. Sens. 41 (8), 2877–2904. https://doi.org/10.1080/01431161.2019.1697006.

Carreto, L., et al., 1996. Rubrobacter xylanophilus sp. nov., a new thermophilic species isolated from a thermally polluted effluent. Int. J. Syst. Evolut. Microbiology 46 (2). https://doi.org/10.1099/00207713-46-2-460.

Castro-Alonso, M.J., Montañez-Hernandez, L.E., Sanchez-Muñoz, M.A., Macias Franco, M.R., Narayanasamy, R., Balagurusamy, N., 2019. Microbially induced calcium carbonate precipitation (MICP) and its potential in bioconcrete: microbiological and molecular concepts. Front. Mater. 6, 126. https://doi.org/10.3389/fmats.2019.00126.

Chala-Aldana, D., Aranda Jiménez, G., Fuello, P.B., Pena, L., Bonilla, M.D.-Z., 2026. Isoscape modelling choices and their implications for archaeological mobility interpretations: a comparative 87Sr/86Sr study from southeastern Iberia. J. Archaeol. Sci. 192, 106639. https://doi.org/10.1016/j.jas.2026.106639. ISSN 0305-4403.

Colagiero, A., Pentimone, I., Rosso, L.C., Ciancio, A., 2017. A metagenomic study on the effect of aboveground plant cover on soil bacterial diversity 6:97-106. In: Lukac, M., Grenni, P., Gamboni, M. (Eds.), Soil Biological Communities and ecosystem Resilience. Springer. https://doi.org/10.1007/978-3-319-63336-7.

Coleine, C., Delgado-Baquerizo, M., DiRuggiero, J., Guirado, E., Harfouche, A.L., Perez-Fernandez, C., et al., 2024. Dryland microbiomes reveal community adaptations to desertification and climate change. ISME J. 18 (1), wrae056. https://doi.org/10.1093/ismejo/wrae056.

Costa, D., Tavares, R.M., Baptista, P., Lino-Neto, T., 2022. The influence of bioclimate on soil microbial communities of cork oak. BMC Microbiol. 22 (1), 163. https://doi.org/10.1186/s12866-022-02574-2. PMID: 35739482; PMCID: PMC9219136.

Cuartero, J., Frey, B., Zornoza, R., sanchez-Navarro, V., et al., 2025. The effect of sustainable management practices on the bacterial community in different European croplands. Appl. Soil. Ecol. 215, 106456. https://doi.org/10.1016/j.apsoil.2025.106456.

Dang, P., Chang, S., Chen, J., et al., 2025. Long-term climate warming substantially reduces global soil microbial richness. One Earth 9 (1), 101511.

de Pins, B., Greenspoon, L., Bar-On, Y.M., et al., 2024. A systematic exploration of bacterial form I rubisco maximal carboxylation rates. EMBO J. 43 (14), 3072–3083. https://doi.org/10.1038/s44318-024-00119-z.

Demergasso, C., Neilson, J.W., Tebes-Cayo, C., Véliz, R., Ayma, D., Laubitz, D., et al., 2023. Hyperarid soil microbial community response to simulated rainfall. Front. Microbiol. 14, 1202266. https://doi.org/10.3389/fmicb.2023.1202266.

Dhahbi, S., Lee, J., Ryu, D., Akinniyi, G., Yang, I., 2025. Actinomycetes studies in Tunisia. Res. Microbiol. 176 (3-4), 104279. https://doi.org/10.1016/j.resmic.2025.104279.

Du, M., Xue, P., Minasny, B., McBratney, A., Fajardo Pedraza, M., Pino, V., et al., 2025. Soil microbial blueprint: predicting soil dominant bacterial genera distribution across Australia. Mol. Ecol. 34, e70135. https://doi.org/10.1111/mec.70135.

Elbert, W., Weber, B., Burrows, S., Steinkamp, J., Büdel, B., Andreae, M.O., et al., 2012. Contribution of cryptogamic covers to the global cycles of carbon and nitrogen. Nat. Geosci. 5 (7), 459–462.

Farquhar, G.D., von Caemmerer, S., Berry, J.A., 1980. A biochemical model of photosynthetic CO_2 assimilation in leaves of C3 species. Planta 149, 78–90. https://doi.org/10.1007/BF00386231.

Fausto, C., Mininni, A.N., Sofo, A., Crecchio, C., Scagliola, M., Dichio, B., et al., 2018. Olive orchard microbiome: characterization of bacterial communities in soil-plant compartments and their comparison between sustainable and conventional soilmanagement systems. Plant. Ecol. Divers. 11, 597–610. https://doi.org/10.1080/17550874.2019.1596172.

Fierer, N., Jackson, R.B., 2006. The diversity and biogeography of soil bacterial communities. Proc. Natl Acad. Sci. U S A. 2006 103 (3), 626–631. https://doi.org/10.1073/pnas.0507535103. Epub 2006 Jan 9. PMID: 16407148; PMCID: PMC1334650.

Finlay, R.D., et al., 2020. Reviews and syntheses: biological weathering and its consequences at different spatial levels – from nanoscale to global scale. Biogeosciences 17 (6), 1507–1533. https://doi.org/10.5194/bg-17-1507-2020.

Frey, S., Lee, J., Melillo, J., et al., 2013. The temperature response of soil microbial efficiency and its feedback to climate. Nat. Clim. Change 3, 395–398. https://doi.org/10.1038/nclimate1796.

Gadd, G.M., 2010. Metals, minerals and microbes: geomicrobiology and bioremediation. Microbiology 156 (3). https://doi.org/10.1099/mic.0.037143-0.

Gaillardet, J.D.B.L., Dupré, B., Louvat, P., Allegre, C.J., 1999. Global silicate weathering and CO_2 consumption rates deduced from the chemistry of large rivers. Chem. Geol. 159 (1-4), 3–30.

Garcia-Pichel, F., Loza, V., Marusenko, Y., Mateo, P., Potrafka, R.M., 2013. Temperature drives the the continental-scale distribution of key microbes in topsoil communities. Science 340 (6140), 1574–1577. https://doi.org/10.1126/science.1236404.

Gentine, P., et al., 2019. Coupling between the terrestrial carbon and water cycles—a review. Environ. Res. Lett. 14, 083003. https://doi.org/10.1088/1748-9326/ab22d6.

Gobbi, A., Acedo, A., Imam, N., et al., 2022. A global microbiome survey of vineyard soils highlights the microbial dimension of viticultural terroirs. Commun. Biol. 5, 241. https://doi.org/10.1038/s42003-022-03202-5.

González-Pimentel, J.L., Cuecas, A., Álvarez, C., Mariscal, V., 2026. Soil bacteriome shifts along a cultivation gradient in Southwestern Spanish wetlands. Microb. Ecol. 89 (1), 16. https://doi.org/10.1007/s00248-025-02660-8.

Google, 2026. Google Earth Engine API. Retrieved from Google Developers ee.Classifier.smileRandomForest Reference.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google earth engine: planetary-scale geospatial analysis for everyone. Remote. Sens. Environ. 202, 18–27.

Greening, C., Biswas, A., Carere, C., et al., 2016. Genomic and metagenomic surveys of hydrogenase distribution indicate H_2 is a widely utilized energy source for microbial growth and survival. ISME J. 10, 761–777. https://doi.org/10.1038/ismej.2015.153.

Guellout, M., Guellout, Z., Belhadj, H., Guellout, A., Bravo, A.G., Jaouani, A., 2026. Metagenomic profiling reveals the role of soil chemistry–Climate interactions in shaping the bacterial communities and functional repertories of Algerian drylands. Eng 7 (1), 40. https://doi.org/10.3390/eng7010040.

Gurbich, T.A., Almeida, A., Beracochea, M., Burdett, T., Burgin, T., Cochrane, G., et al., 2023. MGnify genomes: a resource for biome-specific microbial genome catalogues. J. Mol. Biol. https://doi.org/10.1016/j.jmb.2023.168016.

Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., et al., 2013. High-resolution global maps of 21st-century forest cover change. Science 342 (6160), 850–853.

Hartmann, J., Jansen, N., Dürr, H.H., Kempe, S., Köhler, P., 2009. Global CO_2-consumption by chemical weathering: what is the contribution of highly active weathering regions? Glob. Planet. Change 69 (4), 185–194.

Hartmann, J., Moosdorf, N., 2012. Global lithological map database v1.0 (gridded to 0.5° spatial resolution) [dataset]. PANGAEA, https://doi.org/10.1594/PANGAEA.788537, Supplement to: Hartmann, J; Moosdorf, N (2012): The new global lithological map database GLiM: A representation of rock properties at the Earth surface. Geochemistry, Geophysics, Geosystems, 13, Q12004, https://doi.org/10.1029/2012GC004370.

Hengl, et al., 2026. OpenLandMap-soildb: global soil information at 30m spatial resolution for 2000–2022+ based on spatiotemporal machine learning and harmonized legacy soil samples and observations. Earth Syst. Sci. Data 18, 989–1036. https://doi.org/10.5194/essd-18-989-2026.

Hirt, H., Boukcim, H., Ducousso, M., Saad, M.M., 2023. Engineering carbon sequestration on arid lands. Trends Plant. Sci. 28 (11), 1218–1221. https://doi.org/10.1017/9781009325844.

Holmes, A.J., Bowyer, J., Holley, M.P., O'Donoghue, M., Montgomery, M., Gillings, M.R., 2000. Diverse, yet-to-be-cultured members of the Rubrobacter subdivision of the actinobacteria are widespread in Australian arid soils. FEMS Microbiol. Ecol. 33 (2), 111–120. https://doi.org/10.1111/j.1574-6941.2000.tb00733.x.

Ji, M., Greening, C., Vanwonterghem, I., et al., 2017. Atmospheric trace gases support primary production in Antarctic desert surface soil. Nature 552, 400–403. https://doi.org/10.1038/nature25014.

Kazarina, A., Wiechman, H., Sarkar, S., et al., 2025. Recovery of 679 metagenome-assembled genomes from different soil depths along a precipitation gradient. Sci. Data 12, 521. https://doi.org/10.1038/s41597-025-04884-2.

Kimbatsa, I.M.C.M., Ngô, I., Zamba, A.I., Mabika, F.A.S., Lingouangou, T.M., Goma-Tchimbakala, J., et al., 2023. 16S rRNA gene-based metagenomic analysis of soil bacterial diversity in Brazzaville, Republic of the Congo. Adv. Microbiol. 13, 477–498. https://doi.org/10.4236/aim.2023.139031.

Korkar, M.H., Magdy, M., Rizk, S.M., Fiteha, Y.G., Atta, A.H., Rashed, M.A.-S., 2022. Rhizosphere-associated microbiome profile of agricultural reclaimed lands in Egypt. Agronomy 12 (10), 2543. https://doi.org/10.3390/agronomy12102543.

Kustas, W.P., Perry, E.M., Doraiswamy, P.C., 2003. A simple method for disaggregating satellite-derived land surface temperature. Remote. Sens. Environ. 85 (1), 83–96.

Labouyrie, M., Ballabio, C., Romero, F., et al., 2023. Patterns in soil microbial diversity across Europe. Nat. Commun. 14, 3311. https://doi.org/10.1038/s41467-023-37937-4.

Lasaga, A.C., et al., 1994. Chemical weathering rate laws and global geochemical cycles. Geochimica et. Cosmochimica Acta 58 (10). https://doi.org/10.1016/0016-7037(94)90016-7.

Legeay, J., Errafii, K., Ziami, A., Hijri, M., 2024. The rhizosphere of a drought-tolerant plant species in Morocco: a refuge of high microbial diversity with no taxon preference. Env. Microbiol. Rep. 16 (3), e13254. https://doi.org/10.1111/1758-2229.13254.

Li, H., 2014. Statistical Machine Intelligence and Learning Engine (SMILE). https://haifengl.github.io.

Liang, C., Lehmann, J., 2023. Multifactorial effects matter: moving thermal adaptation into a real-world setting. Glob. Change Biol. 29, 566–568. https://doi.org/10.1111/gcb.16524.

Malacrinò, A., Mosca, S., Li Destri Nicosia, M.G., Agosteo, G.E., Schena, L., 2022. Plant genotype shapes the bacterial microbiome of fruits, leaves, and soil in olive plants. Plants 11 (5), 613. https://doi.org/10.3390/plants11050613.

Marchina, C., Natali, C., Fahnestock, M.F., Pennisi, M., Bryce, J., Bianchini, G., 2018. Strontium isotopic composition of the Po river dissolved load: insights into rock weathering in Northern Italy. Appl. Geochem. 97, 187–196.

Miralles, I., Domingo, F., García-Campos, E., Trasar-Cepeda, C., Carmen Leirós, M., Gil-Sotres, F., 2012. Biological and microbial activity in biological soil crusts from the Tabernas desert, a sub-arid zone in SE Spain. Soil. Biol. Biochem. 55, 113–121. https://doi.org/10.1016/j.soilbio.2012.06.017.

Miralles, I., Ortega, R., Montero-Calasanz, M.C., 2023. Functional and biotechnological potential of microbiome associated with soils colonised by cyanobacteria in drylands. Appl. Soil. Ecol. 192. https://doi.org/10.1016/j.apsoil.2023.105076.

Mohammadipanah, F., Wink, J., 2016. Actinobacteria from arid and Desert habitats: diversity and biological activity. Front. Microbiol. 6, 1541. https://doi.org/10.3389/fmicb.2015.01541.

Moyano, F.E., Vasilyeva, N., Bouckaert, L., Cook, F., Craine, J., Curiel Yuste, J., et al., 2012. The moisture response of soil heterotrophic respiration: interaction with soil properties. Biogeosciences 9, 1173–1182. https://doi.org/10.5194/bg-9-1173-2012.

Muñoz Sabater, J., 2019. ERA5-Land Hourly Data from 1950 to Present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS).

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., et al., 2021. ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. Earth Syst. Sci. Data 13 (9), 4349–4391.

Neilson, J.W., Quade, J., Ortiz, M., et al., 2012. Life at the hyperarid margin: novel bacterial diversity in arid soils of the Atacama Desert, Chile. Extremophiles 16, 553–566. https://doi.org/10.1007/s00792-012-0454-z.

Novara, A., Catania, V., Tolone, M., Gristina, L., Laudicina, V.A., Quatrini, P., 2020. Cover crop impact on soil organic carbon, nitrogen dynamics and microbial diversity in a mediterranean semiarid vineyard. Sustainability 12 (8), 3256. https://doi.org/10.3390/su12083256.

Oeser, R.A., von Blanckenburg, F., 2020. Strontium isotopes trace biological activity in the critical zone along a climate and vegetation gradient. Chem. Geol. 558. https://doi.org/10.1016/j.chemgeo.2020.119861.

Orgiazzi, A., Ballabio, C., Panagos, P., Jones, A., Fernandez Ugalde, O., 2018. LUCAS Soil, the largest expandable soil dataset for Europe: a review. Eur. J. Soil. Sci. 69 (1), 140–153 JRC107926.

Ortega Garcia, M., Caraballo Barreto, I., Ríos Rocafull, Y., Orellana Gallegoł, R., Martínez Vierał y, B., Dibut Álvarez, R., 2016. Variacion de la concentration de microrganismos en distintos Suelos de Cuba. Agrotecnia de. Cuba. 40 (1), 53–61.

Ouellet Dallaire, C., Lehner, B., Creed, I., 2020. Multidisciplinary classification of Canadian river reaches to support the sustainable management of freshwater systems. Can. J. Fish. Aquat. Sci. 77 (2), 326–341. https://doi.org/10.1139/cjfas-2018-0284.

Ouellet Dallaire, C., Lehner, B., Sayre, R., Thieme, M., 2019. A multidisciplinary framework to derive global river reach classifications at high spatial resolution. Environ. Res. Lett. 14 (2), 024003. https://doi.org/10.1088/1748-9326/aad8e9.

Palandri, D.G., Kharaka, Y.K., 2004. A compilation of rate parameters of water-mineral interaction kinetics for application to geochemical modeling. US Geol. Surv. Open-File Rep 2004-1068.

Palumbo, F., Squartini, A., Barcaccia, G., et al., 2021. A multi-kingdom metabarcoding study on cattle grazing Alpine pastures discloses intra-seasonal shifts in plant selection and faecal microbiota. Sci. Rep. 11, 889. https://doi.org/10.1038/s41598-020-79474-w.

Pelacani, S., Ceccherini, M.T., Barbadori, F., Moretti, S., Tommasini, S., 2025a. Geodiversity as a driver of soil microbial community diversity and adaptation in a mediterranean landscape. Land 14 (3). https://doi.org/10.3390/land14030583.

Pelacani, S., Maerker, M., Tommasini, S., Moretti, S., 2024. Combining biodiversity and geodiversity on landscape scale: a novel approach using rare earth elements and spatial distribution models in an agricultural Mediterranean landscape. Ecol. Indic. 158 (2). https://doi.org/10.1016/j.ecolind.2024.111583.

Pelacani, S., Raspini, F., Roccotelli, A., Barbadori, F., Ceccherini, M.T., Tommasini, S., et al., 2025b. Geomorphic connectivity as a driver for the soil bioindicator distribution, implications for future research and modeling. Geogr. Fis. e Dinamica Quaternaria 48 (1-2), 201–215. https://doi.org/10.4454/3k9k0c56.

Plummer, L.N., Busenberg, E., 1982. The solubilities of calcite, aragonite and vaterite in CO_2-H_2O solutions between 0 and 90 °C, and an evaluation of the aqueous model for the system $CaCO_3$-CO_2-H_2O. Geochimica et. Cosmochimica Acta 46 (6), 1011–1040.

Pörtner, H.O., Roberts, D.C., Tignor, M., Poloczanska, E.S., Mintenbeck, K., Alegría, A., et al., 2022. Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK and New York, NY, USA, pp. 3056. https://doi.org/10.1017/9781009325844.

Praeg, N., Steinwanter, M., Urbach, D., et al., 2025. Biodiversity in mountain soils above the tree line. Biol. Rev. 100, 1877–1949. https://doi.org/10.1111/brv.70028.

Priori, S., Pellegrini, S., Perria, R., Puccioni, S., Storchi, P., Valboa, G., et al., 2019. Scale effect of terroir under three contrasting vintages in the Chianti Classico area (Tuscany, Italy). Geoderma 334, 99–112. https://doi.org/10.1016/j.geoderma.2018.07.048.

Rainey, F.A., Ray, K., Ferreira, M., et al., 2005. Extensive diversity of ionizing-radiation-resistant bacteria recovered from Sonoran Desert soil and description of nine new species of the genus Deinococcus obtained from a single soil sample. Appl. Env. Microbiol. 71 (9), 5225–5235. https://doi.org/10.1128/AEM.71.9.5225-5235.2005.

Ray, A.E., Zaugg, J., Benaud, N., et al., 2022. Atmospheric chemosynthesis is phylogenetically and geographically widespread and contributes significantly to carbon fixation throughout cold deserts. ISME J. 16, 2547–2560. https://doi.org/10.1038/s41396-022-01298-5.

Roccotelli, A., Ceccherini, M.T., et al., 2024. Rare earth elements distribution and bacteriome to assess and characterize the soil landscapes of old Olive Orchards. 16 (7). https://doi.org/10.3390/d16070427.

Rodriguez-Navarro, C., Jroundi, F., Schiro, M., Ruiz-Agudo, E., González-Muñoz, M.T., 2012. Influence of substrate mineralogy on bacterial mineralization of calcium carbonate: implications for stone conservation. Appl. Env. Microbiol. 78. https://doi.org/10.1128/AEM.07044-11.

Rosinger, C., Rousk, J., Bonkowski, M., Rethemeyer, J., Jaeschke, A., 2023. Rewetting the hyper-arid Atacama Desert soil reactivates a carbon-starved microbial decomposer community and also triggers archaeal metabolism. Stoten 892, 164785. https://doi.org/10.1016/j.scitotenv.2023.164785.

Rudnick, R.L., Gao, S., 2003. Composition of the continental crust. In: Holland, H.D., Turekian, K.K. (Eds.), Treatise on Geochemistry 3. Elsevier Science, pp. 1–64.

Saidi-Mehrabad, A., Neuberger, P., Hajihosseini, M., Froese, D., Lanoil, B.D., 2020. Permafrost microbial community structure changes across the Pleistocene-Holocene boundary. Front. Environ. Sci. 8, 133. https://doi.org/10.3389/fenvs.2020.00133.

Schimel, J.P., 2018. Life in dry soils: effects of drought on soil microbial communities and processes. Annu. Rev. Ecol. Evol. Syst. 49. https://doi.org/10.1146/annurev-ecolsys-110617-062614.

Serna-Chavez, H.M., Fierer, N., van Bodegom, P.M., 2013. Global patterns of soil microbial biomass. Glob. Ecol. Biogeogr. 22, 1162–1172. https://doi.org/10.1111/geb.12070.

Stantis, C., Willmes, M., Le Corre, M., et al., 2026. Global compilation of bioavailable strontium isotope data. Sci. Data 13, 299. https://doi.org/10.1038/s41597-026-06643-3.

Steefel, C.I., Lasaga, A.C., 1994. A coupled model for transport of multiple chemical species and kinetic precipitation/dissolution reactions with application to reactive flow in single phase hydrothermal systems. Am. J. Sci. 294 (5), 529–592. https://doi.org/10.2475/ajs.294.5.529.

Sun, Y., Chen, X., Chen, C., et al., 2025. Warming amplifies responses of soil organic carbon to multiple global change drivers. Glob. Change Biol. 31 (11), e70612. https://doi.org/10.1111/gcb.70612.

Tabita, F.R., Hanson, T.E., Li, H., Satagopan, S., Singh, J., Chan, S., 2007. Function, structure, and evolution of the RubisCO-like proteins and their RubisCO homologs. Microbiol. Mol. Biol. Rev. 71 (4), 576–599. https://doi.org/10.1128/MMBR.00015-07. PMID: 18063718; PMCID: PMC2168653.

Taylor, L.L., Banwart, S.A., Valdes, P.J., Leake, J.R., Beerling, D.J., 2012. Evaluating the effects of terrestrial ecosystems, climate and carbon dioxide on weathering over geological time: a global-scale process-based approach. Philos. Trans. R. Soc. Lond. B Biol. Sci. 367 (1588), 565–582. https://doi.org/10.1098/rstb.2011.0251.

Taylor, S.R., McLennan, S.M., 1985. The Continental Crust: Its Composition and Evolution. Blackwell Scientific Publications, Oxford.

Taylor, L., Quirk, J., Thorley, R., et al., 2016. Enhanced weathering strategies for stabilizing climate and averting ocean acidification. Nat. Clim. Change 6, 402–406. https://doi.org/10.1038/nclimate2882.

Thompson, L., Sanders, J., McDonald, D., et al., 2017. A communal catalogue reveals Earth's multiscale microbial diversity. Nature 551, 457–463. https://doi.org/10.1038/nature24621.

Verstraeten, W.W., Veroustraete, F., Feyen, J., 2006. Assessment of apparent thermal inertia for soil moisture estimation. Int. J. Remote. Sens. 27 (4), 741–759.

Walker, J.C.G., Hays, P.B., Kasting, J.F., 1981. A negative feedback mechanism for the long-term stabilization of Earth's surface temperature. J. Geophys. Res. 86 (C10), 9776–9782. https://doi.org/10.1029/JC086iC10p09776.

Wild, B., Gerrits, R., Bonneville, S., 2022. The contribution of living organisms to rock weathering in the critical zone. npj Mater. Degrad. 6, 98. https://doi.org/10.1038/s41529-022-00312-7.

Wilson, U.N., Abdullahi, A.S., Adisa, M.A., et al., 2025. Microbial induced calcite precipitation on macrostructural properties of concrete: a review. J. Infrastruct. Preserv. Resil. 6, 42. https://doi.org/10.1186/s43065-025-00158-8.

Xu, Y., Teng, Y., Dai, S., Liao, J., et al., 2024. Atmospheric trace gas oxidizers contribute to soil carbon fixation driven by key soil conditions in terrestrial ecosystems. Environ. Sci. Technol. 58 (49).

Young, J.N., Goldman, J.A., Kranz, S.A., Tortell, P.D., Morel, F.M., 2015. Slow carboxylation of Rubisco constrains the rate of carbon fixation during Antarctic phytoplankton blooms. N. Phytol. 205 (1), 172–181. https://doi.org/10.1111/nph.13021.

Zhang, Z., Jones, G., Calabrese, S., et al., 2025. An integrated modelling framework to determine terrestrial carbon dioxide removal via enhanced rock weathering. Glob. Change Biol. 31 (12), e70650. https://doi.org/10.1111/gcb.70650.

Zhang, Q., Tang, J., Angel, R., Wang, D., Hu, X., Gao, S., et al., 2022. Soil properties interacting with microbial metagenome in decreasing CH4 emission from seasonally flooded marshland following different stages of afforestation. Front. Microbiol. 13, 830019. https://doi.org/10.3389/fmicb.2022.830019.

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