Mechanisms of Multiple Global Change Factor Effects on Soil Ecosystems

Inaugural dissertation

to obtain the academic degree
Doctor of Philosophy in Natural Sciences (Ph.D. in Natural Sciences)
submitted to the

Department of Biology, Chemistry, Pharmacy

of Freie Universität Berlin

by

Mohan BI

Berlin, 2024

This work was carrired out between 2021 and 2024 under the supervision of Prof. Dr. Matthias C. Rillig.
Institute of Biology
Freie Universität Berlin, Germany.

Supervisor: Prof. Dr. Matthias C. Rillig

1^{st} reviewer: Prof. Dr. Matthias C. Rillig

2^{nd} reviewer: Prof. Dr. Masahiro Ryo

Date of defense: 26.11.2024

Declaration of Authorship

I declare that I alone am responsible for the content of my doctoral dissertation and that I have only used sources or references cited in the dissertation.

Mohan Bi

Foreword

This dissertation is a cumulative work of manuscripts, either published or submitted, selected from my publication list:

I Bi, M., Li, H., Meidl, P., Zhu, Y., Ryo, M., Rillig, M.C., 2024. Number and dissimilarity of global change factors influences soil properties and functions. Nature Communications 15(1). https://doi.org/10.1038/s41467-024-52511-2. (Corresponding to chapter 2)

II Cruz, J., Lammel, D., Kim, S.W., Bi, M., Rillig, M.C., 2024. COVID-19 pandemic-related drugs and microplastic from mask fibers jointly affect soil functions and processes. Enivronmental Science and Pollution Research 31, 50630-50641. https://doi.org/10.1007/s11356-024-34587-x. (Corresponding to chapter 3)

III Reprinted with permission from [Rillig, M.C., Ågerstrand, M., Bi, M., Gould, K.A., Sauerland, U., 2023. Risks and benefits of large language models for the environment. Environmental Science & Technology57(9). https://doi.org/10.1021/acs.est.3c01106]. Copyright [2023] American Chemical Society. (Corresponding to chapter 4)

IV Rillig, M.C., Mansour, I., Hempel, S., Bi, M., König-Ries, B., Kasirzadeh, A., 2024.

How widespread use of generative AI for images and video can affect the environment and the science of ecology. Ecology Latters 27(3).
https://doi.org/10.1111/ele.14397. (Corresponding to chapter 4)

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my supervisor, Prof. Dr. Matthias Rillig. His invaluable ideas and contributions shaped this project, and he provided a wonderfully open and friendly working environment that allowed me to grow as a researcher.

I am also profoundly thankful to Prof. Dr. Masahiro Ryo, who offered timely and insightful guidance, particularly during the publication process. His encouragement pushed me to think deeply about statistical analysis, for which I am immensely grateful.

Special thanks to Dr. Anika Lehmann, whose interesting personality and spirit have been a constant source of motivation, encouraging me to pursue a career in academia.

I would like to thank my colleagues, Huiying Li, Yanjie Zhu, and Peter Meidl, for their significant assistance with experiments. I am also deeply grateful to all other lab members, technicians, and staff who supported me in both my work and daily life. Special thanks to Dr. Shinwoong Kim, Dr. Milica Lakovic, Dr. Milos Bielcik, Dr. Baile Xu, Dr. Tingting Zhao, Dr. Han Yan, Sabine Buchert, Anja Wulf, Hongyu Chen, and Yaqi Xu.

A heartfelt thanks to my parents for their unwavering support throughout my studies abroad; their encouragement has been a constant source of strength.

Lastly, I would like to acknowledge the financial support provided by the CSC scholarship, which made this project possible.

Contents

List of Figures

List of Tables

Summary

This dissertation explores the mechanisms underpinning the joint effects of multiple global change factors (GCFs) on soil properties and functions incorporating the rapid development of machine learning methods and AI tools.

In chapter 2, we addresses the complex interactions between multiple GCFs and their joint effects on soil properties and functions. By developing a factor pool of 12 GCFs and calculating dissimilarity indices for factor combinations, the study shows that higher factor number and more dissimilar factors result in more pronounced deviations in soil properties and functions from null model predictions. These deviations often manifest as synergistic interactions, particularly in critical soil functions such as decomposition rates and enzyme activities. It highlights that not only the number of GCFs but also their dissimilarity plays a key role in driving soil responses, offering a new perspective for future research on the interactions of multiple global change factors.

In chapter 3, in addition to the mechanistic insights, the dissertation introduces a practical null model analysis workflow for evaluating interactions among multiple stressors in soil ecology. The workflow not only facilitates the identification of interactions among factors but also efficiently generates null model predictions for a large number of randomly selected factor combinations. By incorporating it with other modeling frameworks, this flexible model workflow can be adapted to various hypothesis testing in GCF studies. Two case studies demonstrate the utility of this approach, offering a robust framework for future research on impacts of multiple GCFs on soil ecosystems.

In chapter 4, the dissertation delves into the rapid development and increasing integration of artificial intelligence (AI), specifically generative models, into environmental science and ecology. This chapter explores the dual-edged nature of these

technologies. On the one hand, LLMs and generative AI offer significant advantages, such as streamlining research workflows, enhancing environmental communication, and broadening public engagement with ecological issues. On the other hand, we also emphasizes the potential risks associated with the unregulated use of AI in environmental sciences, such as the spread of misinformation and biased outputs. Moreover, the potential for AI-generated scientific data to be fabricated or manipulated raises concerns about the integrity of research findings in the field of ecology.

Overall, this dissertation contributes insights into two critical areas of environmental research: the conbined effects of multiple GCFs on soil ecosystems, and the transformative potential of AI technologies in ecological research and communication. By advancing our understanding of how these forces interact with and shape the environment, it provides important frameworks for addressing the multifaceted challenges posed by human activities and technological advancements.

Zusammenfassung

Diese Dissertation untersucht die Mechanismen, die den kombinierten Effekten mehrerer globaler Umweltveränderungsfaktoren (GCFs) auf Bodenbeschaffenheit und funktionen zugrunde liegen und berücksichtigt dabei die rasante Entwicklung von maschinellen Lernmethoden und KI-Tools.

In Kapitel 2 werden die komplexen Wechselwirkungen zwischen mehreren GCFs und deren gemeinsame Auswirkungen auf Bodenbeschaffenheit und -funktionen behandelt. Durch die Entwicklung eines Faktorpools von 12 GCFs und die Berechnung von Ähnlichkeitsindizes für Faktorkombinationen zeigt die Studie, dass eine höhere Anzahl und größere Unterschiedlichkeit der Faktoren zu stärkeren Abweichungen von den Vorhersagen des Nullmodells führen. Diese Abweichungen manifestieren sich häufig als synergetische Wechselwirkungen, insbesondere bei kritischen Bodenfunktionen wie Zersetzungsraten und Enzymaktivitäten. Die Studie hebt hervor, dass nicht nur die Anzahl der GCFs, sondern auch ihre Unterschiedlichkeit eine Schlüsselrolle bei der Beeinflussung von Bodenreaktionen spielt und bietet somit eine neue Perspektive für zukünftige Forschungen zu den Interaktionen mehrerer globaler Umweltveränderungsfaktoren.

In Kapitel 3 stellt die Dissertation zusätzlich zu den mechanistischen Erkenntnissen einen praktischen Workflow für die Analyse von Nullmodellen zur Bewertung von Wechselwirkungen zwischen mehreren Stressoren in der Bodenkunde vor. Der Workflow erleichtert nicht nur die Identifizierung von Wechselwirkungen zwischen den Faktoren, sondern generiert auch effizient Nullmodellvorhersagen für eine große Anzahl zufällig ausgewählter Faktorkombinationen. Durch die Integration in andere Modellierungsrahmen kann dieser flexible Workflow an verschiedene Hypothesentests in GCF-Studien angepasst werden. Zwei Fallstudien demonstrieren die Nützlichkeit dieses Ansatzes und bieten ein solides Rahmenwerk für zukünftige Forschungen zu den Auswirkungen mehrerer GCFs auf Ökosysteme.

In Kapitel 4 befasst sich die Dissertation mit der rasanten Entwicklung und zunehmenden Integration von Künstlicher Intelligenz (KI), insbesondere generativer Modelle, in die Umweltwissenschaften und Ökologie. Dieses Kapitel untersucht die zwiespältige Natur dieser Technologien. Einerseits bieten große Sprachmodelle (LLMs) und generative KI erhebliche Vorteile, wie die Optimierung von Forschungsabläufen, die Verbesserung der Umweltkommunikation und die Erweiterung des öffentlichen Engagements für ökologische Themen. Andererseits werden auch die potenziellen Risiken einer unregulierten Nutzung von KI in den Umweltwissenschaften betont, wie die Verbreitung von Fehlinformationen und voreingenommene Ergebnisse. Zudem wirft die Möglichkeit, dass KI-generierte wissenschaftliche Daten fabriziert oder manipuliert werden könnten, Bedenken hinsichtlich der Integrität von Forschungsergebnissen im Bereich der Ökologie auf.

Insgesamt liefert diese Dissertation Einblicke in zwei zentrale Bereiche der Umweltforschung: die kombinierten Effekte mehrerer GCFs auf Bodensysteme und das transformative Potenzial von KI-Technologien in der ökologischen Forschung und Kommunikation. Durch die Erweiterung unseres Verständnisses darüber, wie diese Kräfte auf die Umwelt wirken und sie formen, bietet sie wichtige Rahmenwerke zur Bewältigung der vielfältigen Herausforderungen, die durch menschliche Aktivitäten und technologische Fortschritte entstehen.

Chapter 1

General Introduction

Global change factors (GCFs) induced by human activities, such as climate change, pollution, and land-use changes, significantly impact the Earth's ecosystems, particularly soil ecosystems(Richardson et al.; 2023; Díaz et al.; 2019). These factors alter soil physicochemical properties, microbial communities, and ecosystem functions, with salinity, pesticides, and drought being notable examples of stressors that negatively affect soil health(Haj-Amor et al.; 2022; Tudi et al.; 2021; Schimel; 2018). While individual GCFs have been extensively studied, there is a growing recognition of the need to explore their combined effects, as multiple stressors often co-occur in real-world ecosystems(Zhou et al.; 2020). Yet, fewer than 2% of studies have addressed the concurrent impacts of three or more factors, largely due to the combinatorial explosion problem that makes traditional factorial designs impractical for high-dimensional ecological research(Rillig et al.; 2019). Innovative experimental designs that randomly select factors from predefined pools are emerging to tackle these challenges and enhance our understanding of complex GCF interactions(Rillig et al.; 2019; Speißer et al.; 2022; Yang et al.; 2022).

The profound and often unpredictable effects of multiple GCFs on soil ecosystems are further complicated by factor interactions. These interactions can be synergistic or antagonistic, meaning the combined effect of multiple factors can differ substantially from the sum of individual effects(Crain et al.; 2008; Dieleman; 2012; Holmstrup et al.; 2010). Understanding these complex dynamics is crucial for developing effective soil and ecosystem management strategies, particularly as ecosystem

biodiversity and functions face increasing threats from anthropogenic stressors. Furthermore, the classification of GCFs based on their traits and mechanisms is helping to uncover patterns and predictions related to how multiple stressors impact ecosystem processes(Rillig et al.; 2021; Orr et al.; 2022; Simmons et al.; 2021). However, despite progress, significant gaps remain in identifying net interactions and the mechanisms driving the joint effects of multiple co-acting GCFs .

In parallel, the rapid advancement of artificial intelligence (AI) is opening new frontiers in environmental sciences, particularly with the rise of generative models such as large language models and generative AI(Future of Life Institute.; 2023). These models, which can generate long text and transform textual prompts into detailed images, audio or videos, present both opportunities and challenges for ecological research. AI has the potential to streamline scientific workflows and enhance research communication, yet it also raises ethical concerns, such as the risk of generating misleading content with false authority(Birhane et al.; 2023; Samuelson; 2023). Moreover, these tools could revolutionize how ecological phenomena are visualized, communicated, and analyzed, particularly in studies involving complex GCF interactions. However, as AI becomes increasingly integrated into environmental science, a critical evaluation of its implications is essential.

This dissertation aims to explore the joint effects of multiple global change factors on soil properties and ecosystem functions, leveraging both advanced experimental designs and machine learning methods to address the complexities of high-dimensional factor interactions. Through a microcosm experiment, we investigated the role of factor dissimilarity and the number of co-acting GCFs in driving soil responses. By developing a Null model workflow, we provided a practical method to estimate null model predictions for GCF studies with complex experimental design. Additionally, this work assessed the potential of AI to enhance ecological research, providing new insights into the intersection of technology and environmental science.

Bibliography

Birhane, A., Kasirzadeh, A., Leslie, D. and Wachter, S. (2023). Science in the age of large language models, *Nature Reviews Physics* **5**(5): 277–280.

Crain, C. M. et al. (2008). Interactive and cumulative effects of multiple human stressors in marine systems, *Ecology Letters* **11**(12): 1304–1315.

Díaz, S. et al. (2019). Pervasive human-driven decline of life on Earth points to the need for transformative change, *Science* **366**(6471): eaax3100.

Dieleman, W. I. J. (2012). Simple additive effects are rare: A quantitative review of plant biomass and soil process responses to combined manipulations of CO_2 and temperature, *Global Change Biology* **18**(9): 2681–2693.

Future of Life Institute. (2023). Pause giant ai experiments: an open letter.
URL: *https://futureoflife.org/open-letter/pause-giant-ai-experiments/*

Haj-Amor, Z. et al. (2022). Soil salinity and its associated effects on soil microorganisms, greenhouse gas emissions, crop yield, biodiversity and desertification: A review, *Science of The Total Environment* **843**: 156946.

Holmstrup, M. et al. (2010). Interactions between effects of environmental chemicals and natural stressors: A review, *Science of The Total Environment* **408**(18): 3746–3762.

Orr, J. A. et al. (2022). Similarity of anthropogenic stressors is multifaceted and scale dependent, *Natural Sciences* **2**(1): e20210076.

Richardson, K. et al. (2023). Earth beyond six of nine planetary boundaries, *Science Advances* **9**(37): eadh2458.

Rillig, M. C. et al. (2019). The role of multiple global change factors in driving soil functions and microbial biodiversity, *Science* **366**(6467): 886–890.

Rillig, M. C. et al. (2021). Classifying human influences on terrestrial ecosystems, *Global Change Biology* **27**(11): 2273–2278.

Samuelson, P. (2023). Generative AI meets copyright, *Science* **381**(6654): 158–161.

Schimel, J. P. (2018). Life in Dry Soils: Effects of Drought on Soil Microbial Communities and Processes, *Annual Review of Ecology, Evolution, and Systematics* **49**(1): 409–432.

Simmons, B. I. et al. (2021). Refocusing multiple stressor research around the targets and scales of ecological impacts, *Nature Ecology & Evolution* **5**(11): 1478–1489.

Speißer, B. et al. (2022). Number of simultaneously acting global change factors affects composition, diversity and productivity of grassland plant communities, *Nature Communications* **13**(1): 7811.

Tudi, M. et al. (2021). Agriculture Development, Pesticide Application and Its Impact on the Environment, *International Journal of Environmental Research and Public Health* **18**(3): 1112.

Yang, G. et al. (2022). Multiple anthropogenic pressures eliminate the effects of soil microbial diversity on ecosystem functions in experimental microcosms, *Nature Communications* **13**(1): 4260.

Zhou, Z., Wang, C. and Luo, Y. (2020). Meta-analysis of the impacts of global change factors on soil microbial diversity and functionality, *Nature Communications* **11**(1): 3072.

Chapter 2

Factor number and dissimilarity drive effects of multiple global change factors on soil properties and functions

2.1 Abstract

Soil biota and functions are impacted by various anthropogenic stressors, including climate change, chemical pollution or microplastics. These stressors do not occur in isolation, and soil properties and functions appear to be directionally driven by the number of global change factors (GCFs) acting simultaneously. Building on this insight, we here hypothesize that co-acting GCFs with more diverse effect mechanisms, or higher dissimilarity, have greater impacts on soil properties and functions. We created a factor pool of 12 GCFs and calculated dissimilarity indices of randomly-chosen co-acting factors at high-GCF levels (2, 5 and 8 factors) based on the measured responses of soil properties and functions to the single factors. Results show that not only was the number of factors important, but factor dissimilarity was also key for predicting joint GCF effects. By analyzing deviations of soil properties and functions from three null model predictions, we demonstrate that higher factor dissimilarity and a larger number of factors could drive larger deviations from null models and trigger more frequent occurrence of synergistic factor net interactions on soil functions (decomposition rate, cellulase and β-glucosidase activity), which

provides new mechanistic insights for understanding high-dimensional effects of factors. Our work highlights the importance of considering factor similarity in future research on interacting GCFs.

2.2 Introduction

Global change factors (GCFs) induced by human activities have a significant impact on soil physicochemical properties, process rates and microbial communities in diverse terrestrial ecosystems(Rillig et al.; 2019). The effect of individual GCFs on soil properties and functions have been the focus of many prior studies. For example, salinity reduces the availability of soil nutrients and has negative impacts on soil microbial activities(Haj-Amor et al.; 2022), pesticides can pose adverse effects on non-target soil organisms(Tudi et al.; 2021), and drought affects soil processes by directly stressing soil organisms and indirectly by hindering substrate (Schimel; 2018). The multitude of GCFs collectively gives rise to concerns about soil ecosystem health.

Only a few studies have addressed the effects at a high-dimensional factor level with concurrent effects of a larger number of factors. A systematic mapping showed that fewer than 2% of experimental studies have explored the combined effects of three or more factors in the context of soils(Rillig et al.; 2019). One of the main obstacles for studying joint effects of multiple factors at a time is the combinatorial explosion problem(Katzir et al.; 2019). For traditional factorial designs, when the number of factors increases, the number of possible factor combinations will increase rapidly, meaning that such designs including a large number of factors are not feasible in ecology. To overcome this experimental challenge, recent studies investigating the effects of multiple factors followed an approach involving randomly selecting factors from a predefined factor pool; such a design avoids factor combination problems without losing generalizability(Rillig et al.; 2019; Speißer et al.; 2022; Yang et al.; 2022). Using this experimental approach addressed a general feature of multiple GCFs — the number of co-acting factors — which has been shown to directionally drive the effects of co-acting multiple factors on plants(Zandalinas et al.; 2021), soil ecosystems(Rillig et al.; 2019) and the plant community(Speißer et al.;

2022; Komatsu et al.; 2019). Another study also indicated that the increasing number of factors diminished the functions of soil microbial diversity(Yang et al.; 2022). However, in general, our knowledge about mechanisms underpinning the effects of multiple co-acting GCFs is still limited.

To gain more insights into this high-dimensionality problem, studies have highlighted the importance of ordering and classifying GCFs from trait-based perspectives(Rillig et al.; 2021a; Orr et al.; 2022b; Simmons et al.; 2021). In previous studies, factors are usually grouped by their sources instead of considering their effect mechanisms and ecological-scale dependency(Simmons et al.; 2021). Recently, an *a priori* factor classification system has been introduced, using inherent traits (physical, chemical and biological agents) and theoretical effects (effect mechanism, targets and key properties) of 30 different anthropogenic factors(Rillig et al.; 2021a). Building more comprehensive factor classification systems may enable extracting new features from factor traits to predict general patterns of multiple GCF effects. In this context, factor dissimilarity is a plausible feature that can be generated from multiple available factor traits and may capture patterns of high-dimensional effects(Orr et al.; 2022b). However, the role of factor dissimilarity in driving the effects of multiple GCFs has never been investigated experimentally.

Another important feature of the multiple co-acting GCF effects is the nature of factor interactions. Across marine and terrestrial ecosystems, many studies found that when two or more factors are present, the combined effects often differ from what is expected based on the single factor effects(Crain et al.; 2008; Dieleman; 2012; Holmstrup et al.; 2010). The interactions between factors are defined as antagonistic when combined effects are less than expected, while synergistic interactions cause combined effects larger than expected effects. Although understanding interactions among factors is crucial for prioritizing ecosystem stressor management, when more than three factors are acting simultaneously, testing every pairwise factor interaction or high-order interaction is extremely challenging unless every factor combination has been separately replicated(Smith et al.; 2024). In this case, revealing the overall net interactive effects of multiple GCFs is a more practical solution and can also indicate potential interactions among multiple GCFs. Nevertheless, there is still a lack of established methods for identifying the net interactions for multiple co-acting GCFs

and insufficient knowledge about the potential mechanisms underpinning such effects.

We here aim to investigate the joint effects of multiple GCFs on soil properties and functions, examining the effects of number of factors and dissimilarity of GCF combinations. We present a microcosm experiment (Fig. 1) with 2, 5 and 8 factor levels to test the following hypotheses:

- (i) Factor dissimilarity can help predict soil biological and ecological responses to multiple co-acting GCFs in addition to the number of factors;
- (ii) a larger dissimilarity among factors or larger number of factors will cause greater deviation of joint effects on soil properties and functions from expected effects;
- (iii) Factor dissimilarity or number of factors may drive the emergence of factor interactions (synergistic or antagonistic).

2.3 Methods

2.3.1 Experimental design

The experiment was set up with a GCF pool that includes 12 factors: salinity, drought, microplastic, fungicide, herbicide, antibiotic, insecticide, surfactant, nitrogen deposition, heavy metal pollution, perfluoroalkyl and polyfluoroalkyl substances (PFAS) and lithium. The selected factors were chosen from the most frequently occurring anthropogenic factors in soil ecosystems subject to intense human influence(Riedo et al.; 2021; Zhou et al.; 2020), and differ in intrinsic features (physical, chemical etc.) and effect mechanisms (mode of action, effect targets etc.) in affecting soil properties and functions(Rillig et al.; 2021a; Orr et al.; 2022b; Schäfer and Piggott; 2018). On the basis of previous experimental designs1, three multi-factor levels (2, 5 and 8 factors, 50 replicates) were created by a random factor-selection method(Rillig et al.; 2019; Huang et al.; 2018; Tilman et al.; 1996). To achieve this, first, complete sets of factor combinations for each factor level were generated (e.g., for the 5 factor level, there were in total 792 different factor combinations for choosing 5 factors from a 12 factor pool). Then we randomly selected 50 factor combinations from all the

possible combinations at each factor level without replacement to avoid selecting repeated factor combinations. Furthermore, we set up single factor treatments with 8 replicates for each factor and the control group including 20 replicates. Finally, to test for the effects of organic solvents (dimethyl sulfoxide (DMSO) and acetone) used to apply chemical GCFs (fungicide and herbicide) on soil properties and functions, we included 10 additional replicates (water control) that received the same rate of water instead of organic solvent in the experiment. Collectively, we had $(50 \times 3) + (12 \times 8) + 20 + 10 = 276$ units in our experiment (FIGURE 2.1).

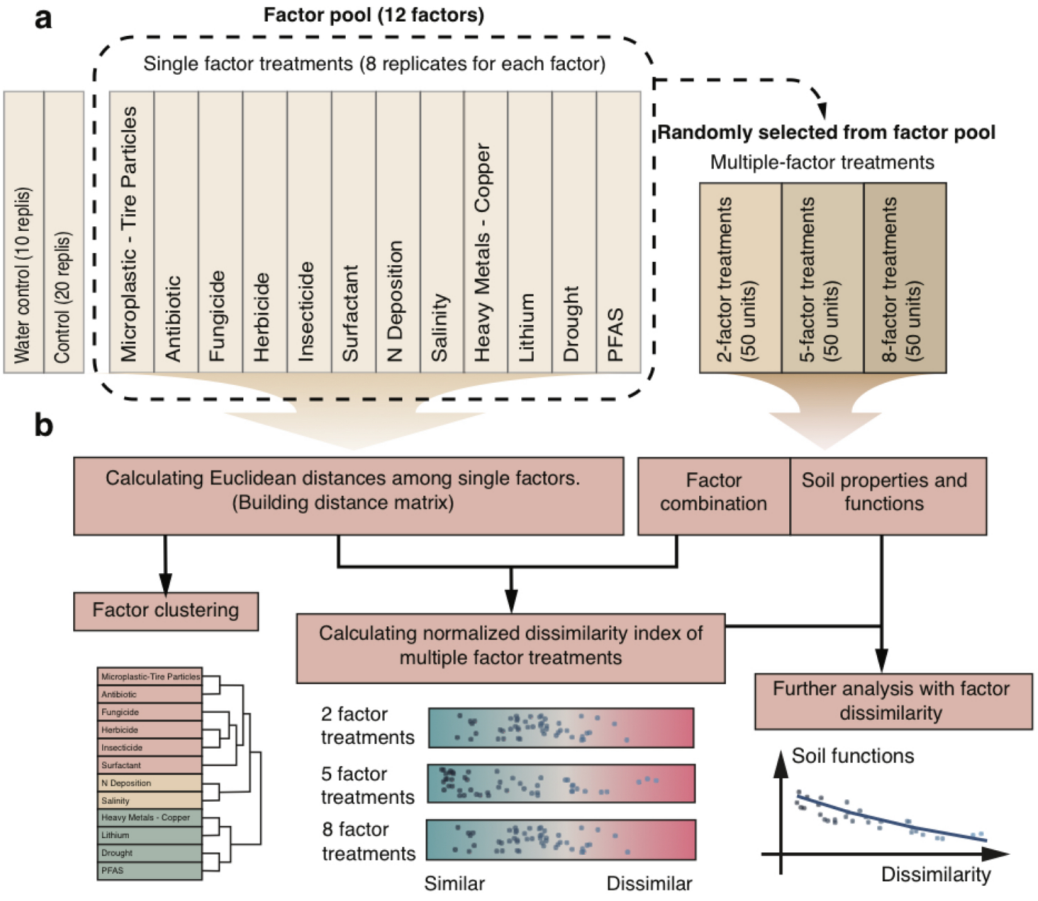

FIGURE 2.1: **Experimental design and analysis workflow.**

(a) The design of the multiple factor experiment. There were 20 replicates for the control; 10 replicates for water control (without adding organic solvents); 8 replicates for single GCFs; 50 replicates for each factor level in the multiple factor group; the total number of experimental units = 20+10+812+350=276. Factor combinations in 2, 5 and 8 levels are randomly selected from the full combinations by drawing 2, 5 and 8 factors from the 12 factors pool without repetition. (b) Analysis workflow. The normalized dissimilarity index for each multi-factor treatment is calculated based on the Euclidean distances among single factors and the randomly selected factor combinations.

2.3.2 Soil preparation and incubation system

The soil used in the experiment was collected in February 2022 from a local grassland at an experimental site of Freie Universität Berlin (52° 28′ N, 13°18′ E, Berlin, Germany) with a sandy loamy texture. Before the start of the experiment, the soil was air dried and passed through a 2 mm sieve to remove large stones and big grass roots. To prepare the "loading soil" for the factor implementation, one eighth of the air-dried sieved soil was sterilized under 121°C for 20 mins. Loading soil was used to more effectively mix small amounts of chemicals into the experimental units; it was sterilized to avoid large local effects on soil microbes.

The experimental unit was a 50 mL mini-bioreactor (Product Nr: 431720, Corning®, USA) with a vented film, which allows gas exchange but prevents microbial contamination (FIGURE 2.2). Inside the bioreactor we added 40.0 g (dry weight, d.w.) soil with the respective GCF treatments.

2.3.3 Implementation of GCFs and harvest

This study focuses on 12 frequently occurring GCFs in soil ecosystems that differ in their nature (physical, chemical, biological) and effect mechanisms: salinity, drought, microplastic, fungicide, herbicide, antibiotics, insecticide, surfactant, nitrogen deposition, heavy metals, PFAS and lithium. Detailed information about the selected GCFs is presented in the supplementary information.

To homogeneously mix GCFs with the testing soil, we added chemical factors to 5.0 g (d.w.) loading soil first then mixed the loading soil with the other 35.0 g (d.w.) soil. Concentrated solutions were prepared for every chemical factor except for salinity and microplastic. Most of the chemicals we used were dissolved in distilled water, except for fungicide (carbendazim, dissolved in DMSO) and herbicide (Diflufenican, dissolved in acetone). According to the designed factor combination for each treatment, 100 μL solution (water, DMSO or acetone) carrying appropriate chemical dose for 40.0 g (d.w.) soil was added to 5.0 g (d.w.) loading soil inside a 150 mL cup. To standardize the amount of solvents we added into every treatment, treatments which had fewer factors, for instance, single factor treatments and control treatments, additionally received solvents (water, DMSO or acetone) to the

same amount of solvents added in the eight factor treatments. To further test the effects of the solvents (DMSO and acetone) on soil properties and functions, another 10 control treatments only received the same amount of distilled water. The effects of organic solvents on soil properties and functions are shown in Appendix FIGURE 1. For the experimental units that include microplastic or salinity treatments, 40.0 mg tire particles (1-2 mm diameter) or 200.0 mg sodium chloride were added into the 150 mL cup accordingly. Then an additional 35.0 g of air-dried soil was added to every 150 mL cup. After covering with a cap, all the soil treatments were mixed for 30 mins with a shaking machine (Product Nr: 541-21009-00, Reax2, Heidolph Instrument GmbH & Co. KG, Schwabach, Germany) at a speed of 80 rpm to achieve a homogeneous distribution. After the mixing process, the soil-chemical mixture was transferred to the 50 mL mini bioreactors. For tracking the litter decomposition rate during the experiment, a sterilized tea bag was placed vertically in the center of the soil. Finally, distilled water was added to bring the soil water content to 60% of the soil water-holding capacity (30% of water-holding capacity for drought treatments).

All 50 mL mini-bioreactors were incubated at 25 °C in a dark environment for 42 days. As there was on average 0.5 g weight loss every week for each mini-bioreactor, we added 0.5 mL distilled water to each treatment every week to keep the water content constant. After 42 days, all units were harvested. Soil cores were taken from the bioreactors, the tea bags (see below) were removed and the soil of each treatment was homogeneously mixed by a spoon in a sterilized Petri dish for 2 min. 5.0 g fresh soil was collected and stored at 4 °C for enzymatic activity measurement, and the remaining soil was air dried at room temperature for soil property measurements. Tea bags from all units were collected and oven dried (60 °C) before measuring litter decomposition rate.

2.3.4 Soil response variables

The soil response variables we measured in this experiment are: litter decomposition rate, soil pH, water-stable aggregates and the activity of four extracellular soil enzymes. Here we present the methods of each response variable measurement.

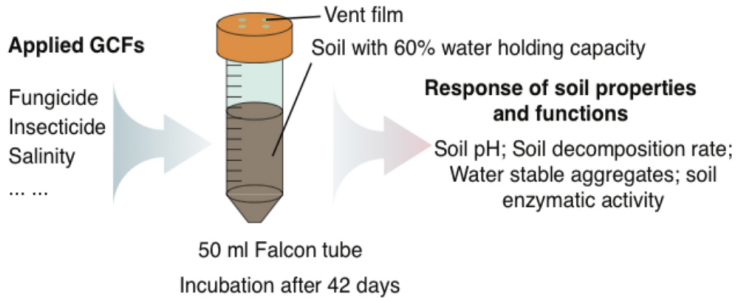

FIGURE 2.2: **Microcosm experimental system**

pH measurement

We collected 5 g air dried soil from each unit to a 50 mL centrifuge tube (Product Nr: 62.547.255, Sarstedt AG & Co., Germany) after harvest and mixed with 25 mL distilled water. The tubes were shaken vertically for 30 mins with a speed of 200 rpm and centrifuged at 3000 rpm for 5 mins. The pH of each unit was determined with a pH meter (Hanna Instrument, Smithfield, USA) by placing the pH probe in the supernatant for 2 mins.

Water-stable soil aggregate measurement

The proportion of water-stable soil aggregates (WSA) was measured following a modified previous protocol(Klute; 1986). 4.0 g air dried soil was weighed and recorded as the sample weight. Then it was placed in a small 250 μm sieve above a weighing boat, rewetted by distilled water for 5 mins and inserted into a sieving machine (Eijkelkamp, Netherlands) for 3 mins running. Afterwards the remaining matter was washed into a plastic plate and dried at 60 °C overnight, then weighed and recorded as dry matter. Then we rewetted the dry matter and gently crushed the aggregates with fingers and flushed it on a 250 μm sieve.The remaining material was dried at 60 °C overnight and it was recorded as coarse matter. The WSA is calculated by using the following formula:

$$WSA(\%) = \frac{dry\ matter - coarse\ matter}{sample\ weight - coarse\ matter} \times 100 \tag{2.1}$$

Decomposition rate

The decomposition rate was measured and calculated using a modified protocol(Keuskamp et al.; 2013). A sealed tiny bag (with 38 μm mesh size) containing 300.0 mg (d.w.) tea biomass was placed into the soil. We autoclaved the tea bags and dried them in the oven at 60°C to sterilize before using. After the harvest, we dried all tea bags at 60°C for two days and removed the soil particles attached to the tea bags using a brush. Then the proportional weight loss was calculated based on the tea biomass (d.w.) inside the bag before and after the incubation.

Enzyme acticities measurement

The measurements of N-acetyl-glucosaminidase (chitin degradation), cellulase (cellulose degradation), β-glucosidase (cellulose degradation) and phosphatase (organic phosphorus mineralization) activity followed a high throughput microplate protocol(Jackson et al.; 2013). The enzyme substrate solution was prepared using ρNP-β-glucopyranoside (Sigma no. N7006), ρNP-β-D-cellobioside (Sigma no. N5759), ρNP-β-N-acetylglucosaminide (Sigma no. N9376) and ρNPphosphate(Sigma no. 71768) accordingly. We collected 5 g fresh soil from each unit during the harvest and stored it in 50 mL centrifuge tubes (Product Nr: 62.547.255, Sarstedt AG & Co., Germany) at 4 °C. The enzyme activity measurement was conducted 2 days after the completion of the harvest. We mixed 8 mL 50 mM acetate buffer with the 5.0 g fresh soil to form a soil slurry, vortexed for 30 seconds and pipetted 150 μL slurry into four wells of each 96-well plate. The process was repeated in four plates because of four types of measured enzyme activity. After all wells were filled with soil slurry in a plate, 150 μL enzyme substrate solution was added to the four wells with soil slurry using a 8-channel pipettor and 150 μL acetate buffer was added to the following two wells as control. Then the plate was incubated at 20 °C in the dark for 2 hours (phosphatase and β-glucosidase) or 4 hours (β-D-cellobiosidase and β-N-acetylglucosaminidase). After the incubation, the plates were centrifuged at a speed of 3000 rpm for 5 mins. We used a 8-channel pipettor to transfer 100 μL of the upper suspension to a new 96-well plate, where 200 μL of 0.05 M NaOH solution was also added into each well. The final measurement was conducted by a microplate reader (BioRad, Benchmark

Plus, Japan) to measure the absorbance at a wavelength of 410 nm. The final enzymatic activity was calculated with the standard curve with the unit of ρNP (μmoles) per hour per gram of dry soil.

2.3.5 Effect size calculation and significance test of single and multiple factor groups

Data were analyzed with R Version 4.1.1(R Core Team; 2021). For single factor and multiple factor groups, the effect size and 95% confidence intervals (CIs) of each group were estimated with a nonparametric bootstrap method with 10,000 permutations(Efron and Tibshirani; 1986). Considering the multiple testing problem, the statistical significance of single and multiple factor effects was evaluated by using adjusted P-values based on the Benjamini-Hochberg method.

2.3.6 Calculating factor dissimilarity

We used the "vegan"(Oksanen J; 2018) R package to calculate the Euclidean distances between all pairwise factor combinations based on the corresponding standardized effect sizes of singly applied factors on the seven soil properties (including four soil enzyme activity, WSA, soil decomposition rate and soil pH). Clustering of single factors was conducted based on Euclidean distance by using hierarchical clustering analysis ("ggdendro"(de Vries and Ripley; 2024) and "dendextend" R packages were used)(FIGURE 2.3). Then, we used principal coordinate analysis to visualize the distances among factors, resulting in the PCoA1 and PCoA2 axes explaining 55.43% and 23.38% of the variation respectively (FIGURE 2.4).

For the multiple-factor treatments, we calculated a dissimilarity index (DI) for each unique factor combination by adding up the Euclidean distances between every two component factors in the multiple-factor treatments:

$$DI_i = \sum_{j \in N_i} d_j \tag{2.2}$$

, where (i = 1, 2, ..., 50) is the dissimilarity index of multiple-factor treatment's i_{th} combination in a specific level of number of factors (2, 5, and 8), d_j is the Euclidean

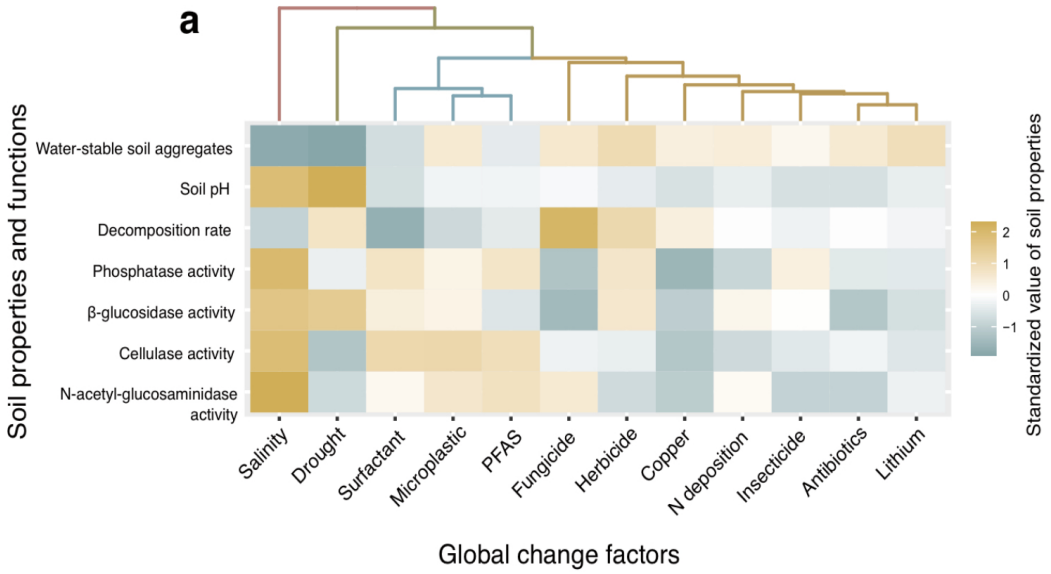

FIGURE 2.3: **Clustering factors by seven soil responses.**

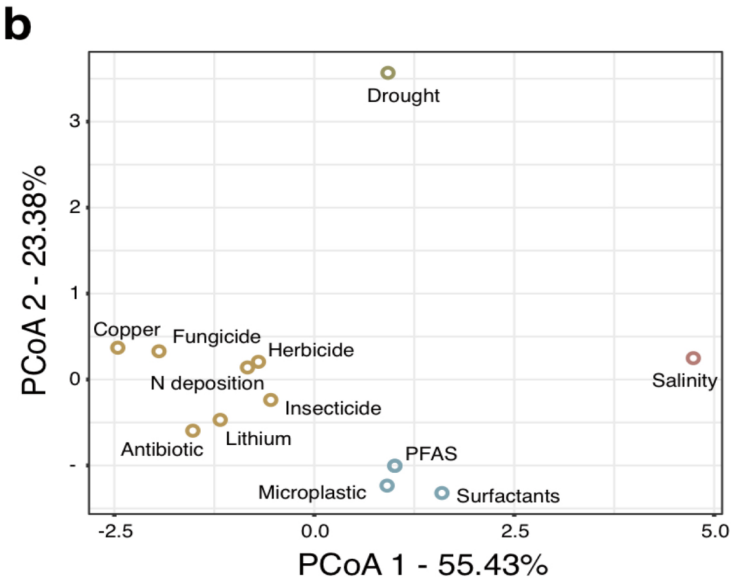

FIGURE 2.4: **Principal coordinates analysis (PCoA) of the Euclidean distances between single factors.**

distance between the j_{th} two factor pair estimated based on the single factor experiment, and N_i is the set of all unique factor pairs of the treatment i. To compare dissimilarity indices between different numbers of factor levels, we normalized the dissimilarity indices of each factor level to a range between 0 and 1 by using the "range" method of the preProcess function from the "caret" R package(Kuhn M; 2015). To do this, we subtract the minimum value from each dissimilarity index and divide it by the range of the dissimilarity indices of each number of factor level. The distributions of normalized factor dissimilarity indices in three factor levels and the reasons for choosing the normalization method are shown in Appendix FIGURE2.

2.3.7 Correlations between soil responses and factor dissimilarity within factor levels

To show the changing trend of soil property and function responses across the range of factor dissimilarity within factor levels, we applied Spearman correlation analyses to the normalized dissimilarity index and soil properties and functions in each factor level. Estimated P-value and coefficient are provided respectively for each correlation.

2.3.8 Predicting effects of multiple co-acting factors by null models

In ecological studies, null models are used for predicting the joint effect of multiple factors without considering interactions(Schäfer and Piggott; 2018). For commonly-used null models, the additive model assumes that the joint effect of multiple factors will be the sum of the effects of the single factors, indicating that the sensitivities of the target to factors are negatively correlated. The multiplicative model assumes that the effects of single factors are combined by proportional change, meaning that the factor sensitivities are non-correlated. In the dominative model, the factor with the largest absolute effect overrides other factors, implying the factor sensitivities are positively correlated. To make plausible predictions of multiple-factor effects on soil responses, we imposed three null model assumptions (i.e., additive, multiplicative and dominative assumption) for generating predictions for the multiple factor treatments instead of arbitrarily selecting one(Schäfer and Piggott; 2018). For each null

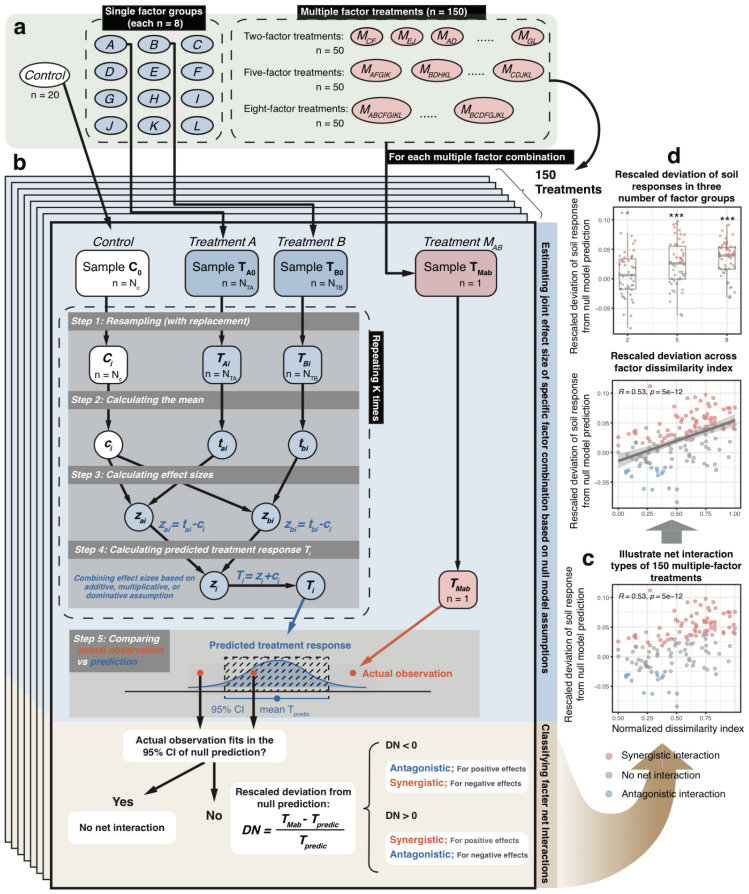

FIGURE 2.5: **Calculating soil response deviation from null model prediction and net interaction type classification for 150 multi-factor treatments.**

(a) Treatments in the experimental design. Single factor treatments are shown in blue ovals. Each multi-factor treatment is shown by a red oval. The subscript of a multi-factor treatment indicates the component factors. (b) Interaction type classification workflow for multi-factor treatments. The workflow includes two parts: (1) estimating the joint response distributions of component factors of multi-factor treatments; (2) identifying the net interaction type for multi-factor treatments. For illustration purposes, one two-factor treatment (includes factor A and B) is taken as an example. In Step 1, we resampled from each control, single factor A and B treatment with replacement to generate C_i, T_{Ai} and T_{Bi}. Then, in Step 2, mean values of each resampled treatment (c_i, t_ai and t_bi) are calculated. In step 3, absolute effect sizes from control (Z_{ai} and Z_{bi}) for A and B single factor treatments are calculated. In step 4, combined effect size of A and B (Z_i) are calculated depending on different null model assumptions (additive, multiplicative or dominative). Then the control mean is added to z_i to generate predicted joint response (T_i). Steps 1-4 are repeated K times to generate the distribution of the predicted joint response of factor A and B. Then in Step 5, we compared the actual joint response of factor A and B (T_{Mab}) to the predicted response distribution. If the actual observation fitted within the 95% confidence intervals (CIs) of prediction distribution, then it was regarded as no net interaction. If it did not fit, then we calculated the rescaled Deviation from Null model prediction (DN). Then we classified the net interaction type based on the rescaled DN. (c) Visualization of the rescaled DN and net interaction types of 150 multi-factor treatments. (d) Statistical analysis of rescaled DN of soil response across factor dissimilarity index and in three different number of factor groups.

model assumption, we applied the calculation methods from a previous study(Rillig et al.; 2019). For each number of factors level, the unique subset of factor combinations randomly chosen from the 12 factor pool is denoted as A_n (n = 2, 5, 8). For each multiple-factor combination $K_m \in A_n$ (e.g. K_1 = [Microplastic, Drought], K_2 = [Antibiotic, Fungicide] ..., K_{50} = [Salinity, PFAS]; K_1, K_2, ..., $K_{50} \in A_2$), K_m includes N component factors, denoted as (F_{m_1}, F_{m_2}, ..., F_{m_N}) (N = 2 for A_2, N = 5 for A_5, N = 8 for A_8). ES_{m_i} is the mean of estimated effect size of the factor F_{m_i} observed from the single factor treatment. In additive assumption, the predicted effect size of factor combination K_m:

$$P_{additive_m} = \sum_{i=1}^{N} ES_{m_i} \tag{2.3}$$

Considering each set of A_n has 50 elements (K_m), we applied a bootstrapping method (with 1,000 iterations; see FIGURE 2.5) for each K_m. Each K_m has 1,000 iterated effect size predictions, in total 50,000 effect size predictions were made for all treatments for each number of factor level, which should be sufficient for generating reliable estimates. Afterwards, the mean value and 95% CI were calculated form the distribution of each factor combination. The same bootstrapping procedures were used in multiplicative and dominative assuptions. For multiplicaitve addumption, based on a previous method(Schäfer and Piggott; 2018), the predicted effect size of factor combination K_m is shown as:

$$P_{multiplicative_m} = CT \prod_{i=1}^{N} \left(1 + \frac{ES_{m_i}}{CT}\right) - CT \tag{2.4}$$

CT is the estimated response of the control group. For the dominative null models, the predicted effect size is:

$$P_{dominative_m} = ES_{m_i}(|ES_{m_i}| = max(|ES_{m_1}|, |ES_{m_2}|, ..., |ES_{m_N}|)) \tag{2.5}$$

2.3.9 Hierarchical modeling framework for hypothesis testing

To disentangle the contribution of possible drivers (number of factor effect and factor dissimilarity effect) on the variability of soil properties and functions in response to multiple GCFs, a hierarchical modeling framework was implemented (FIGURE 2.6, Appendix TABLE 1).

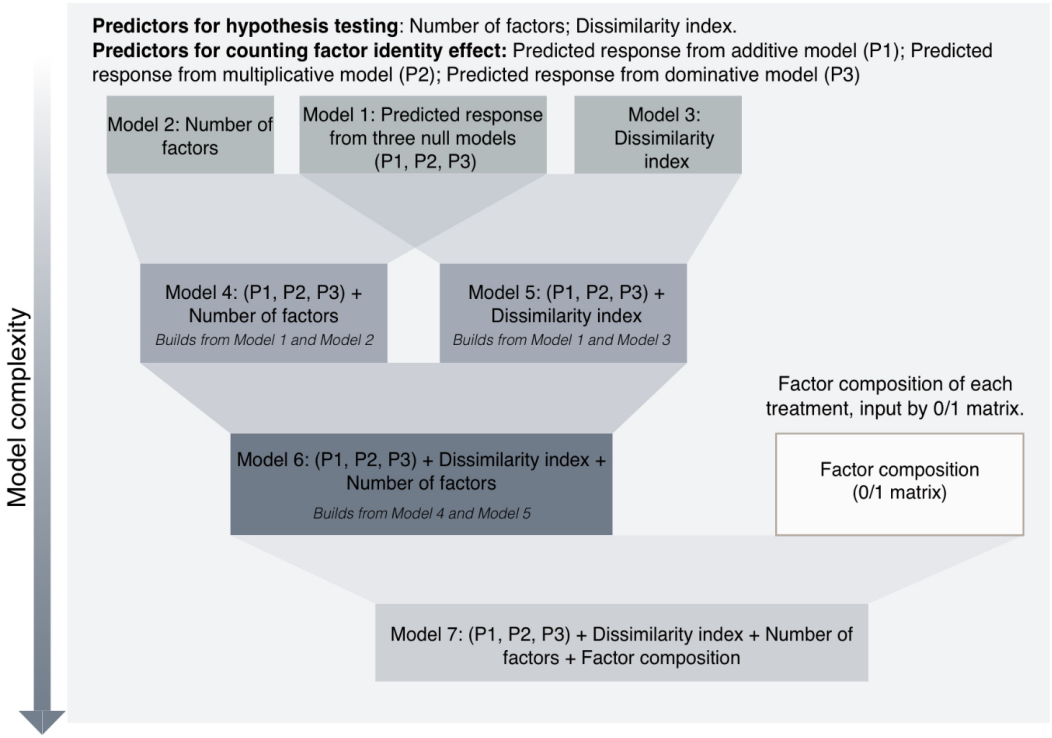

FIGURE 2.6: **Structure of hierarchical modeling framework for hypothesis testing.**

Structure of hierarchical modeling framework. Different compositions of predictors are included in Model 1 to 7. Model complexity increases when more predictors are included.

TABLE 2.1: **Model formulas for Hierarchical modeling analysis.**

Model	**Model formula**
Model 1	Response variables ∼ P1 + P2 + P3
Model 2	Response variables ∼ Number of factors
Model 3	Response variables ∼ Dissimilarity index
Model 4	Response variables ∼ P1 + P2 + P3 + Number of factors
Model 5	Response variables ∼ P1 + P2 + P3 + Dissimilarity index
Model 6	Response variables ∼ P1 + P2 + P3 + Number of factors + Dissimilarity index
Model 7	Response variables ∼ P1 + P2 + P3 + Number of factors + Dissimilarity index + salinity + drought + microplastic + fungicide + herbicide + antibiotics + insecticide +surfactant + nitrogen deposition + heavy metals + PFAS + lithium

To generate robust results, we compared the modeling results generated by both machine learning and generalized linear model (GLM). In the modeling, data from all the treatments except for the controls were used. In both algorithms, to separate the factor identity effects, the null model predictions (from additive, multiplicative and dominative models) were first included as predictors in the baseline model (Model 1), which is regarded as the soil response variability explained by the contributions of factor identity. Number of factors was solely included as the predictor in Model 2, and in Model 3 factor dissimilarity indices were included instead. Then, on the basis of the baseline model, each soil response was modeled by adding the number of factors or factor dissimilarity indices as an additional predictor in Model 4 and Model 5 respectively. Furthermore, in Model 6, both factor dissimilarity indices and number of factors were added on the basis of the baseline model. Lastly, factor composition (i.e., a binary matrix coding the features for each treatment, where 1 or 0 represent the presence or absence of each stressor.) was included as the last predictor for the final model (Model 7). The formula describing each model is shown in Appendix TABLE 2.1. In Model 7, due to different model algorithms, factor composition has a different meaning. For the Random Forest algorithm(Breiman; 2001), the factor composition stands for all the information from the experimental design (also includes the information of other predictors, e.g., number of factors), and theoretically it can provide the best model fits. Thus, in the hierarchical modeling framework, the factor composition is only being added at the end to show the variability that can be explained by the experimental treatments (the randomly-drawn factors). In the GLM, including factor composition does not stand for the factor identity effects and also does not have a specific statistical meaning in this case. But for comparison to the Random Forest model, we still provide the modeling results of Model 7.

We evaluated the variability of soil responses explained by all seven models with model R^2 values (%). To evaluate the contribution of each model predictor, for the GLM, we compared models by their AIC (Akaike information criterion) values based on the ANOVA tests (Supplementary Table 7 and 8) and evaluated the increase in model R-squared values. For the Random Forest models, the contribution of each model predictor was evaluated by the increase in the model R-squared

values. To address the statistical inference of predictor contributions in the Random Forest models, we used a permutation-based random forest model approach with 1,000 permutations to calculate the relative importance of each predictor. Adjusted P-values of relative importance for each model predictor are shown in Appendix TABLE 2.

2.4 Results

2.4.1 Effects of individual factors on soil functions and properties

The 12 single factors, our factor pool, produced a variety of responses on soil properties and functions, including positive, neutral and negative trends (FIGURE 2.7 and FIGURE 2.8). However, according to the significance tests based on adjusted P values (n = 8), none of the 12 single factors had significant effects on soil decomposition rate and four soil enzyme activity. Salinity and drought caused soil pH to increase ($P = 0.019$ and < 0.001, respectively, TABLE 2.2), while decreasing the proportion of water-stable soil aggregates ($P = 0.021$ and 0.042, respectively, TABLE 2.2).

TABLE 2.2: **Significance test for the effects of single GCFs on soil responses.**

Factor	N-acetyl-glucosaminidase	Cellulase activity	β-glucosidase activity	Phosphatase activity	Decomposition rate	Soil pH	Water stable soil aggregate
PFAS	0.833	0.603	0.879	0.819	0.768	0.168	0.181
Copper	0.833	0.603	0.755	0.819	0.813	0.588	0.736
Lithium	0.862	0.603	0.879	0.819	0.768	0.574	0.972
N deposition	0.8362	0.603	0.755	0.819	0.813	0.168	0.816
Antibiotic	0.833	0.603	0.755	0.819	0.813	0.739	0.706
Insecticide	0.833	0.603	0.755	0.819	0.768	0.776	0.335
Surfactant	0.862	0.603	0.755	0.819	0.433	0.805	0.245
Fungicide	0.833	0.603	0.755	0.819	0.071	0.168	0.816
Herbicide	0.833	0.603	0.755	0.819	0.768	0.168	0.971
Microplastic	0.833	0.603	0.755	0.819	0.572	0.168	0.749
Salinity	0.833	0.603	0.755	0.609	0.271	**0.001***	**0.021**
Drought	0.862	0.603	0.755	0.819	0.768	<**0.001***	**0.042**

Adjusted P values based on Benjamini-Hochberg method obtained from two sided t-tests between each treatment group and controls. (Significant differences with 0.01 < P 0.05 are shown in bold, and P < 0.01 are marked by * additionally.)

2.4.2 Effects of multiple co-acting GCFs on soil functions and properties

The simultaneous effects of multiple factors on soil functions and properties changed directionally with an increase in the number of factors. When multiple factors were applied, water-stable soil aggregates ($P < 0.001$ for all of 2, 5 and 8 factor groups) decreased, while soil pH ($P < 0.001$ for all of 2, 5 and 8 factor groups) increased compared to control. Soil decomposition rate decreased only in the eight-factor group (FIGURE 2.7, TABLE 2.3). Activity of β-glucosidase increased in all three factor groups ($P = 0.015$ for 2 factor group, $P < 0.001$ for 5 and 8 factor groups). Phosphatase activity did not change in any factor group, while activity of N-acetyl-glucosaminidase and cellulase increased in 5, 8 factors groups and 8 factors group,

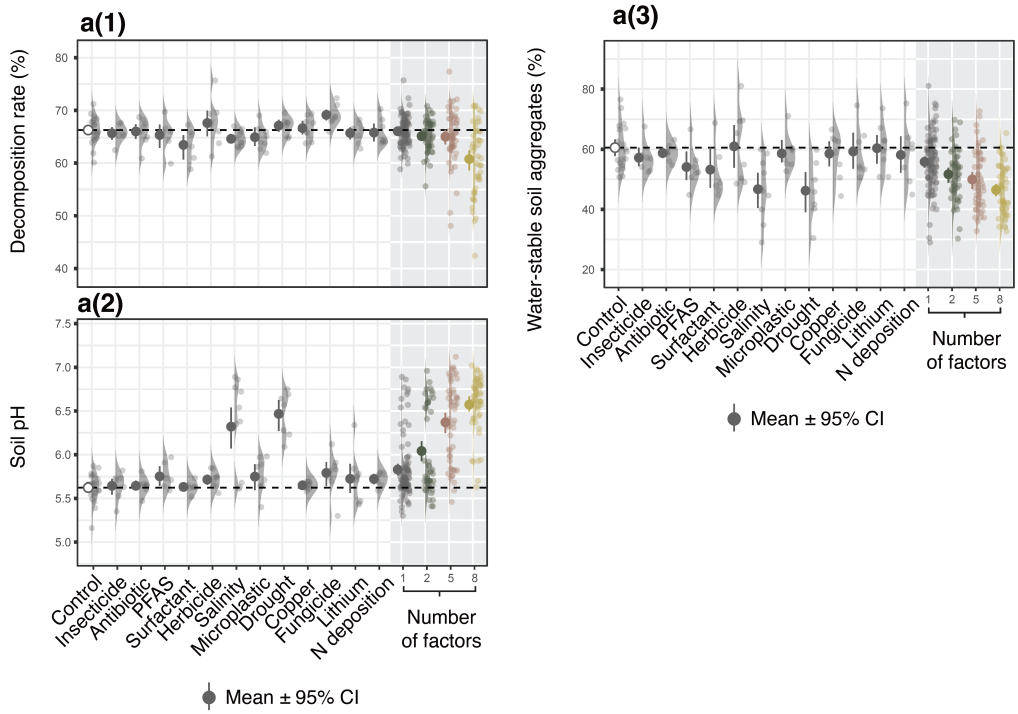

FIGURE 2.7: **Effect sizes of soil property responses to single and multiple factor groups**

For each soil property, effect sizes of single factors (n =8) and multiple factor groups (2, 5 and 8 factors, 50 treatments included in each factor group) were estimated [a(1), a(2) and a(3)]

respectively (P = 0.017, 0.002 and 0.008, respectively) (FIGURE 2.8, TABLE 2.3).

TABLE 2.3: **Significance test for the effects of groups in different number of factors on soil responses.**

Number of factor group	N-acetyl-glucosaminidase	Cellulase activity	β-glucosidase activity	Phosphatase activity	Decomposition rate	Soil pH	Water stable soil aggregate
1 (n = 93)	0.843	0.578	0.472	0.933	0.769	**<0.001***	**0.012**
2 (n = 47)	0.068	0.575	**0.015**	0.869	0.271	**<0.001***	**<0.001***
5 (n = 50)	**0.017**	0.058	**<0.001***	0.191	0.271	**<0.001***	**<0.001***
8 (n = 50)	**0.002***	**0.008***	**<0.001***	0.191	**<0.001***	**<0.001***	**<0.001***

Adjusted P values based on Benjamini-Hochberg method obtained from two sided t-tests between each treatment group and controls. (Significant differences with 0.01 < P 0.05 are shown in bold, and P < 0.01 are marked by * additionally.)

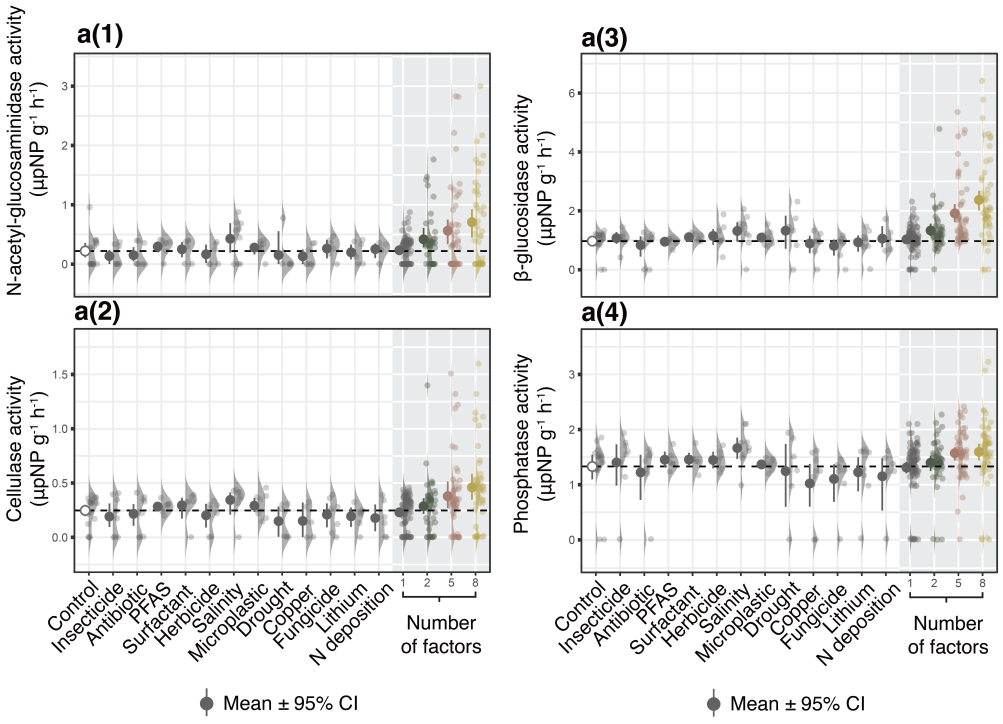

FIGURE 2.8: **Effect sizes of soil enzymatic acticity responses to single and multiple factor groups**

For each soil enzymatic activity, effect sizes of single factors (n =8) and multiple factor groups (2, 5 and 8 factors,50 treatments included in each factor group) were estimated [a(1), a(2), a(3) and a(4)].

2.4.3 Correlations of soil property and function responses to co-acting GCFs with factor dissimilarity

We used Spearman correlation analyzes to show the changing trend of soil property and function responses along factor dissimilarity index range within each factor level. Factor dissimilarity was positively associated with soil pH, but negatively associated with water-stable soil aggregates and soil decomposition rate (FIGURE 2.9). Soil enzymatic activity was positively correlated with the factor dissimilarity index (FIGURE 2.10). However, the correlation of soil responses with the factor dissimilarity index could be caused by the artefact that factor combinations with larger dissimilarity indices also have a higher chance of including the factors with extreme effect size. Therefore, only the correlation analysis by itself is insufficient for evaluating the real effect of factor dissimilarity.

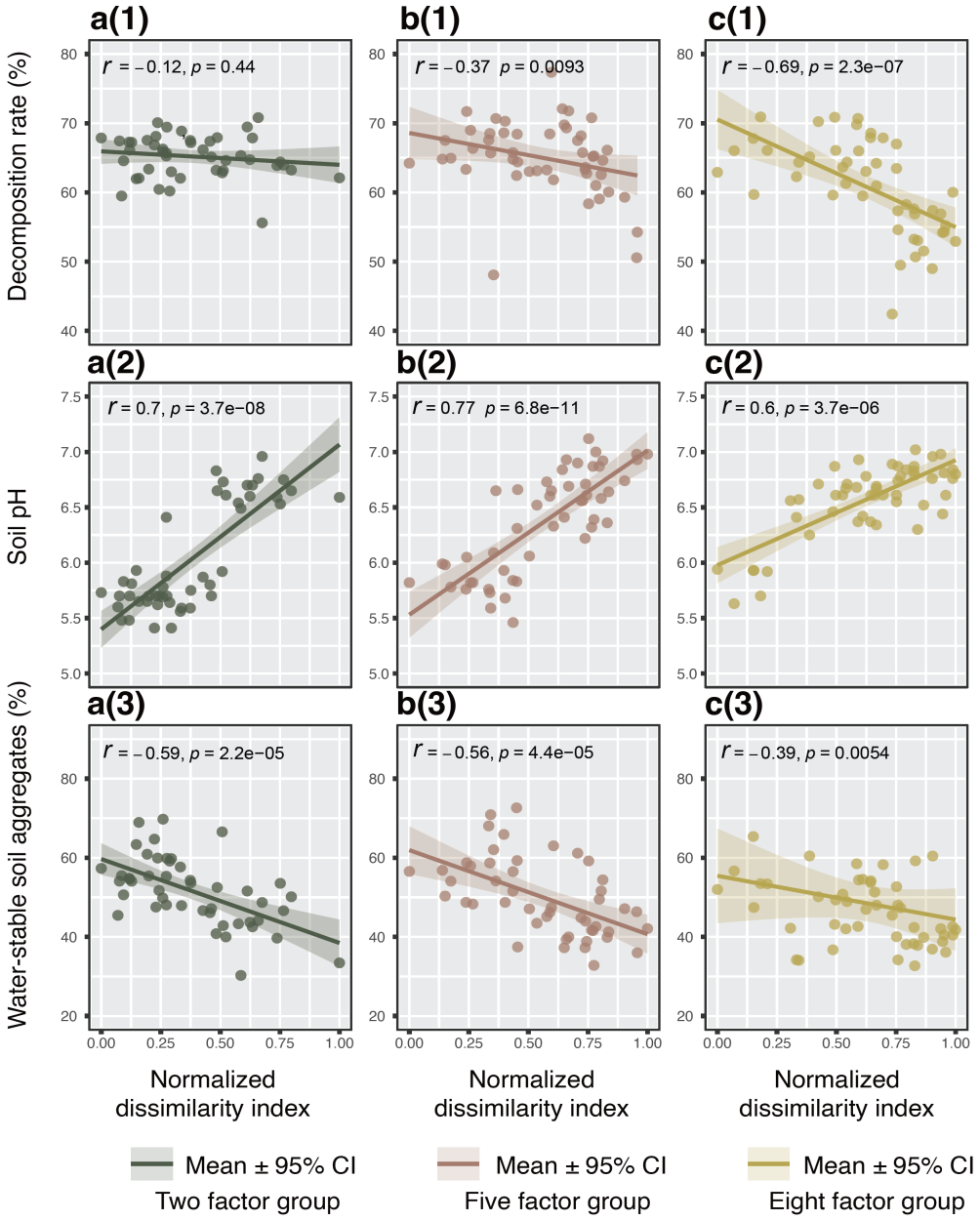

FIGURE 2.9: **Spearman correlations of soil property responses to the normalized factor dissimilarity index**

the correlations of soil property responses to the normalized factor dissimilarity index are shown in scatter plots [a(1) to c(1), a(2) to c(2) and a(3) to c(3)], Spearman correlation coefficients and significance of correlations are indicated by r and p, respectively.

2.4.4 Hypothesis testing by hierarchical modeling framework

To test the possible drivers (number of factors and factor dissimilarity) of the variability of soil responses to simultaneously acting multiple factors and to separate the factor identity contribution, a hierarchical modeling framework was implemented

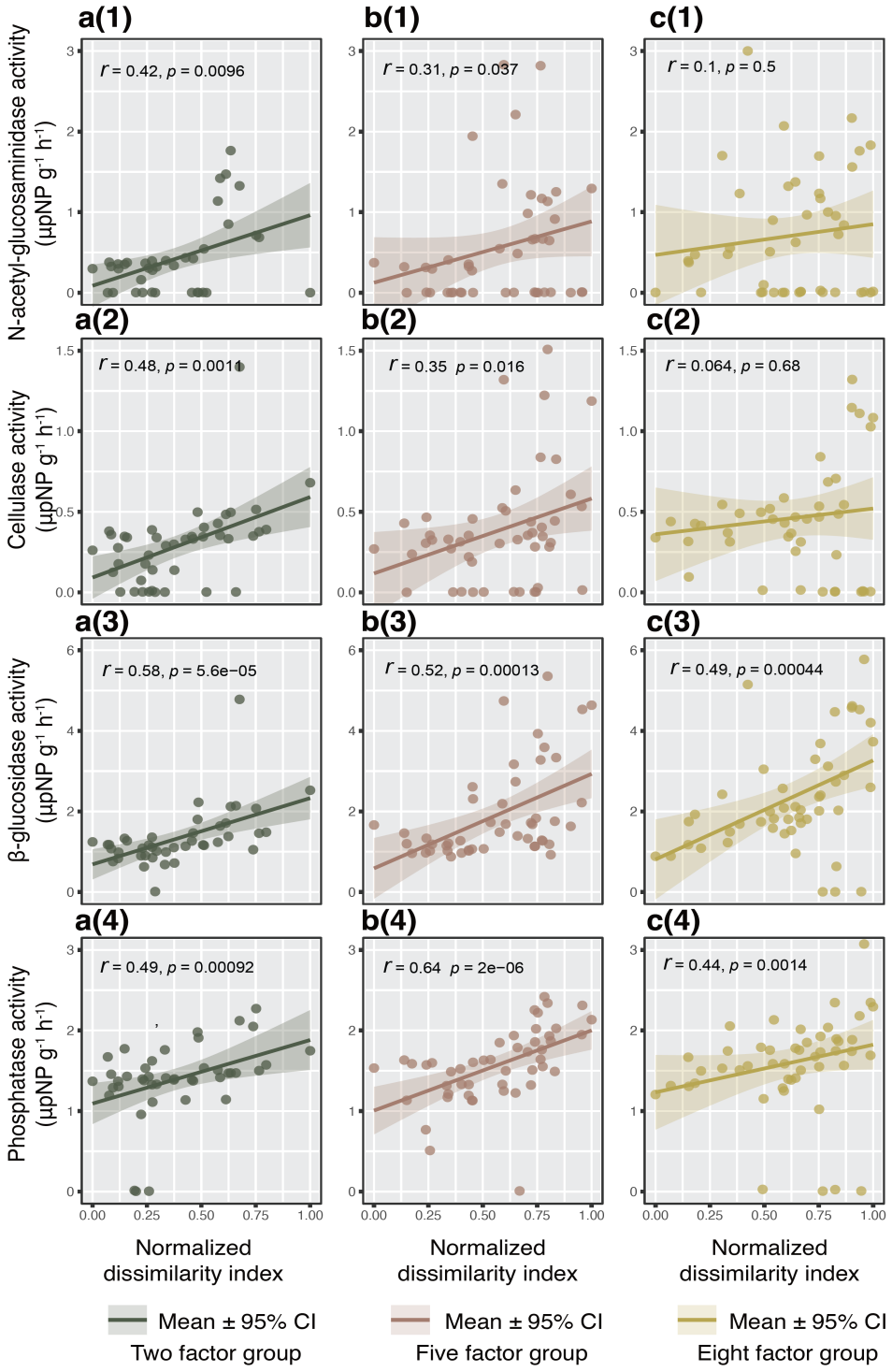

FIGURE 2.10: **Spearman correlations of soil enzymatic activity responses to the normalized factor dissimilarity index**

the correlations of soil enzymatic activity responses to the normalized factor dissimilarity index are shown in scatter plots [a(1) to c(1), a(2) to c(2), a(3) to c(3) and a(4) to c(4)], Spearman correlation coefficients and significance of correlations are indicated by r and p, respectively.

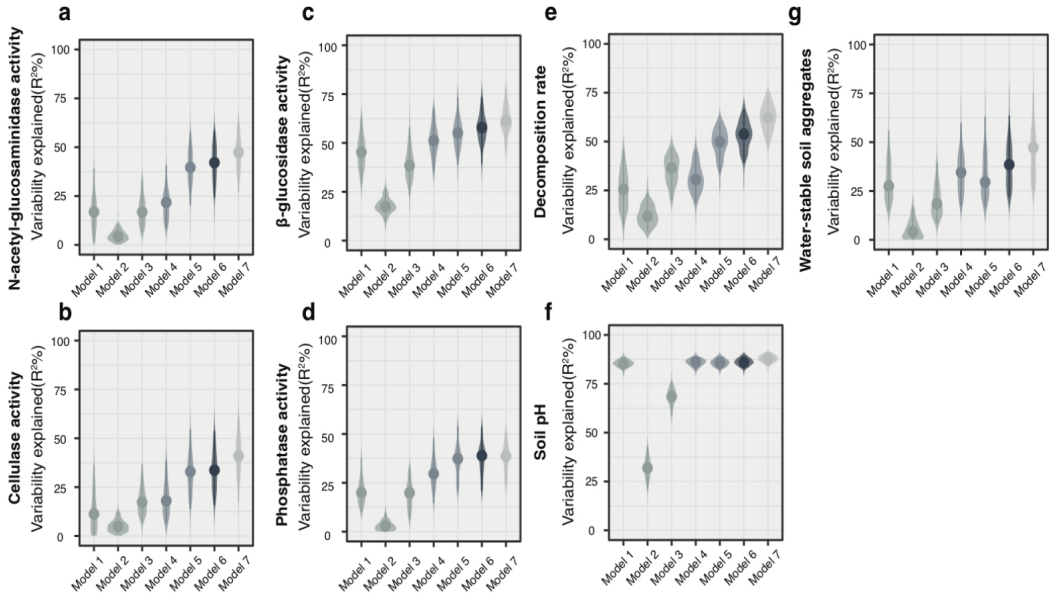

FIGURE 2.11: **Explained variance of soil properties and functions by seven models from the hierarchical modeling framework.**

based on machine learning and generalized linear model (GLM) algorithms (FIGURE 2.6). To separate the contribution of factor identity, we first built the baseline model by predicting the soil functions and properties using the three null model predictions (i.e. predicted responses calculated by the responses of single factor treatments based on additive, multiplicative and dominative null model algorithms). Then, we tested the effect of factor dissimilarity and number of factors by adding additional predictors on the basis of the baseline model. The contribution of each predictor was evaluated by the increment of an R-squared value for Random Forest (RF) models and by comparing the changes of model AIC for the GLM.

The hierarchical modeling results based on the random forest algorithm showed that adding the number of factors improved the model R^2 for soil decomposition rate and three types of soil enzymatic activity (cellulase, β-glucosidase and phosphatase) (FIGURE 2.11, Appendix TABLE 3 and 4). Adding the dissimilarity indices further improved the model R^2 largely for soil decomposition rate and four types of soil enzymatic activity (FIGURE 2.11, Appendix TABLE 3 and 4). The permutation-based random forest approach also indicates that the importance of number of factors is significant for predicting soil decomposition rate and three types of soil enzymatic activity (cellulase, β-glucosidase and phosphatase), and factor dissimilarity index is

significant for predicting all the soil responses except for water-stable soil aggregates (Appendix TABLE 2). The hierarchical modeling based on the GLM algorithm also showed similar results as the RF models (TABLE 2.4, 2.5, 2.6, 2.7, 2.8, 2.9 and 2.10). Collectively, both machine learning and GLM algorithms indicate that the number of factors and factor dissimilarity are important predictors for the variability of soil responses to multiple GCFs.

TABLE 2.4: **Contributions of model predictors on soil decomposition rate based on general linear models.**

Model(Reference model)	Df	AIC	logLik	P value	Adjusted R^2
Model 1	5	-433.8994	221.9497	-	0.024
Model 2	3	-443.0140	224.507	-	0.0001
Model 3	3	-457.2668	231.6334	-	0.1734
Model 4 (Model 1)	6	-439.3643	225.6822	**0.007***	0.0847
Model 5 (Model 1)	6	-457.5266	234.7633	<**0.001***	0.1911(+)
Model 6 (Model 4)	7	-456.9633	235.4817	<**0.001***	0.1933(+)
Model 6 (Model 5)	7	-456.9633	235.4817	0.241	0.1933
Model 7 (Model 6)	18	-479.9469	257.9734	<**0.001***	0.3557(+)

TABLE 2.5: **Contributions of model predictors on soil pH based on general linear models.**

Model(Reference model)	Df	AIC	logLik	P value	Adjusted R^2
Model 1	5	-35.84877	22.92439	-	0.8098
Model 2	3	172.84632	-83.42316	-	0.2028
Model 3	3	53.81937	-23.90968	-	0.6453
Model 4 (Model 1)	6	-35.71041	23.8552	0.181	0.8109
Model 5 (Model 1)	6	-39.45372	25.72686	**0.020**	0.8156(+)
Model 6 (Model 4)	7	-37.62369	25.81185	0.053	0.8145
Model 6 (Model 5)	7	-37.62369	25.81185	0.687	0.8145
Model 7 (Model 6)	18	-57.58191	46.79096	<**0.001***	0.8488(+)

TABLE 2.6: **Contributions of model predictors on water-stable soil aggregate based on general linear models.**

Model(Reference model)	Df	AIC	logLik	P value	Adjusted R^2
Model 1	5	-221.9365	115.9683	-	0.1267
Model 2	3	-204.6608	105.3304	-	0.1975
Model 3	3	-224.0199	115.0099	-	0.1274
Model 4 (Model 1)	6	-227.3576	119.6788	**0.007***	0.1448(+)
Model 5 (Model 1)	6	-221.55372	116.7769	0.212	0.1302
Model 6 (Model 4)	7	-225.4000	119.7	0.841	0.1582
Model 6 (Model 5)	7	-225.4000	119.7	0.687	0.1582
Model 7 (Model 6)	18	-222.4485	129.2243	0.096	0.1979

TABLE 2.7: **Contributions of model predictors on N-acetyl-glucosaminidase activity based on general linear models.**

Model(Reference model)	Df	AIC	logLik	P value	Adjusted R^2
Model 1	5	218.6840	-104.342	-	-0.0217
Model 2	3	288.4251	-141.2125	-	0.0216
Model 3	3	282.2736	-138.1368	-	0.0617
Model 4 (Model 1)	6	213.7671	-100.8836	**0.010**	0.0329(+)
Model 5 (Model 1)	6	198.7713	-93.38566	$<$**0.001***	0.1594(+)
Model 6 (Model 4)	7	198.9904	-92.4952	$<$**0.001***	0.1650(+)
Model 6 (Model 5)	7	198.9904	-92.4952	0.196	0.1650
Model 7 (Model 6)	18	193.1288	-78.5644	**0.011**	0.2778(+)

TABLE 2.8: **Contributions of model predictors on cellulase activity based on general linear models.**

Model(Reference model)	Df	AIC	logLik	P value	Adjusted R^2
Model 1	5	71.33657	-30.6682	-	-0.0227
Model 2	3	84.18846	-39.09423	-	0.0301
Model 3	3	74.15474	-34.07737	-	0.0941
Model 4 (Model 1)	6	68.12162	-28.06081	**0.027**	0.0230(+)
Model 5 (Model 1)	6	59.88618	-23.94309	$<$**0.001***	0.1075(+)
Model 6 (Model 4)	7	61.69514	-23.84757	**0.005***	0.0989(+)
Model 6 (Model 5)	7	61.69514	-23.84757	0.674	0.0989
Model 7 (Model 6)	18	61.16782	-12.58391	0.054	0.1929(+)

TABLE 2.9: **Contributions of model predictors on β-glucosidase activity based on general linear models.**

Model(Reference model)	Df	AIC	logLik	P value	Adjusted R^2
Model 1	5	385.9900	-187.995	-	0.2690
Model 2	3	460.6354	-227.3177	-	0.1113
Model 3	3	429.0648	-211.5324	-	0.2831
Model 4 (Model 1)	6	382.6148	-185.3074	**0.023**	0.2918(+)
Model 5 (Model 1)	6	371.9674	-179.9837	<**0.001***	0.3447(+)
Model 6 (Model 4)	7	371.3913	-178.6956	<**0.001***	0.3520(+)
Model 6 (Model 5)	7	371.3913	-178.6956	0.117	0.3520
Model 7 (Model 6)	18	356.5627	-160.2814	<**0.001***	0.4594

TABLE 2.10: **Contributions of model predictors on phosphatase activity based on general linear models.**

Model(Reference model)	Df	AIC	logLik	P value	Adjusted R^2
Model 1	5	226.4944	-108.2472	-	0.0415
Model 2	3	228.0747	-111.0373	-	0.0564
Model 3	3	207.4480	-100.724	-	0.1467
Model 4 (Model 1)	6	223.7225	-105.8612	**0.032**	0.0656(+)
Model 5 (Model 1)	6	210.2900	-99.14499	<**0.001***	0.1472(+)
Model 6 (Model 4)	7	212.2295	-99.11475	<**0.001***	0.1415(+)
Model 6 (Model 5)	7	212.2295	-99.11475	0.81	0.1415
Model 7 (Model 6)	18	214.7683	-89.38413	0.086	0.1843

2.4.5 Emrgence of factor net interactions in multiple-factor treatments

We developed a methodology to assess the emergence of GCF interactions based on the deviation of soil responses from null model predictions (see methods). Based on this approach, we identified the net interaction type of 150 multiple-factor treatments for each soil response. At the two-factor level, net interaction represents pairwise interaction. When the number of component factors is more than two, net interaction represents the overall effect of all pairwise interactions and higher-order interactions among factors. For soil decomposition rate, soil cellulase and β-glucosidase activity, based on three null model predictions, more synergistic net interactions emerged when the factor dissimilarity index increased (FIGURE 2.12 and FIGURE

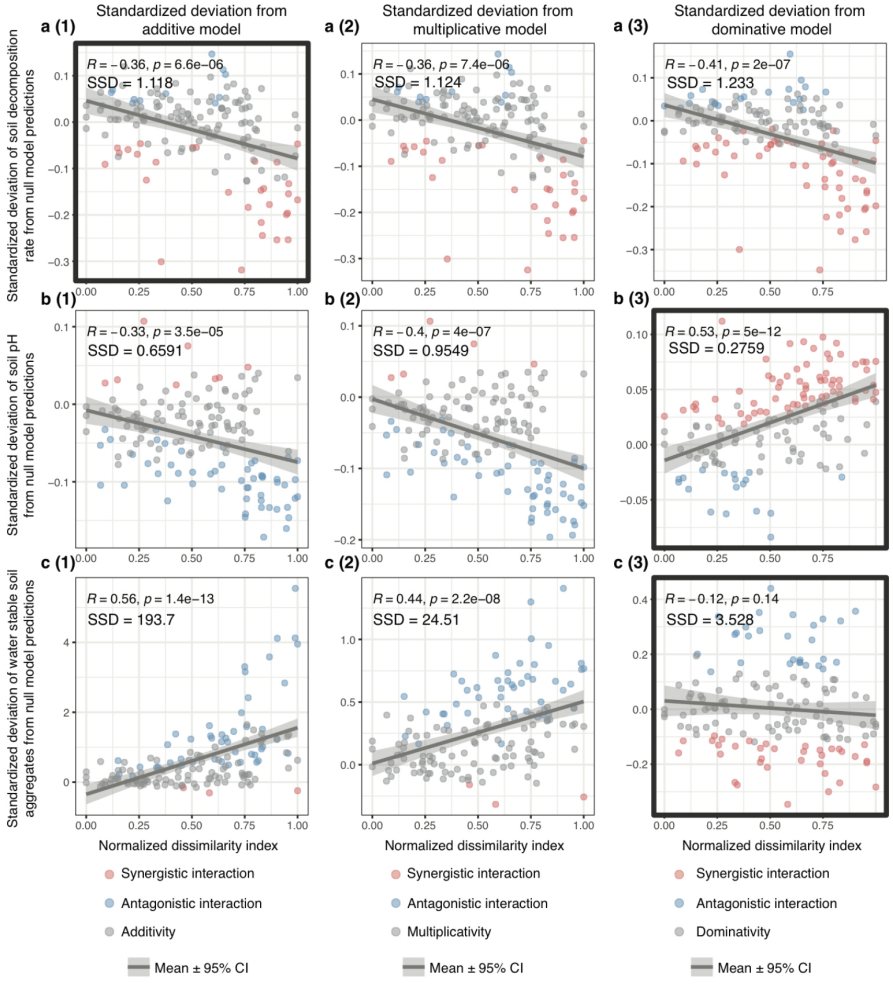

FIGURE 2.12: **Correlation between rescaled deviation of soil responses (decomposition rate, soil pH and water-stable soil aggregates) from three null model predictions and normalized dissimilarity index.**

[a (1) to c (3)] Scatter plots show the standardized deviation of soil responses from null model predictions for multiple-factor treatments. Net interaction type of each multiple-factor treatment is marked as different colored points (antagonistic, blue; synergistic, red; no interaction, gray). The betterst-fitting null model for each soil response has been selected based on the smallest model sum of squared deviation (SSD), and it is indicated by the bold frame [a (1), b (3) and c (3)]. The correlations between standardized deviation of soil responses from null model predictions and normalized dissimilarity index are shown in linear correlation with 95% confidence intervals. Spearman correlation coefficients and significance of correlations are indicated by R and p, respectively.

2.13). Across three number of factor levels, no obvious change of the emergence of factor net interactions was observed (Appendix FIGURE 2.14 and 2.15).

To evaluate the drivers of factor interactions, we assessed the standardized deviations of soil responses from three null model predictions across the dissimilarity range and three factor levels, respectively. From the three null models, the model

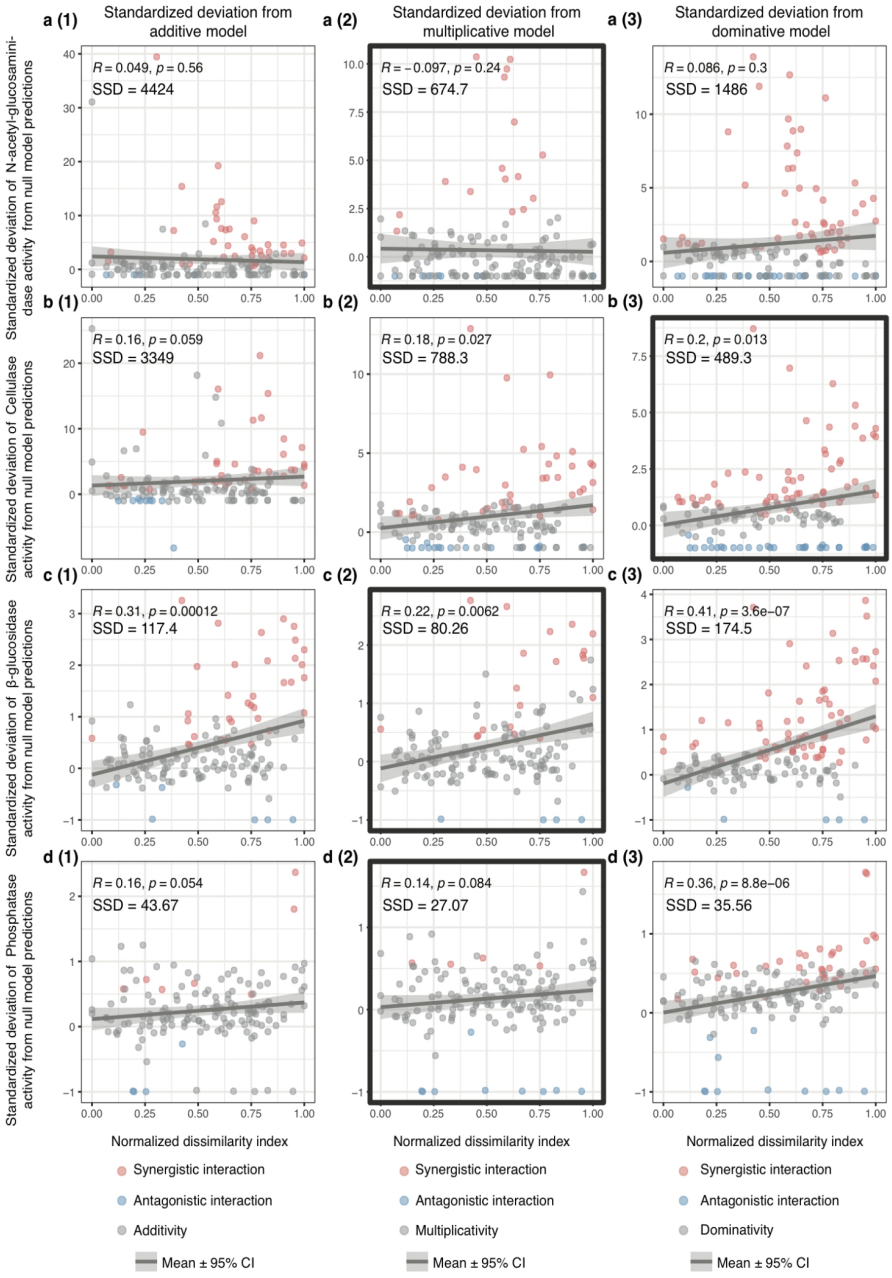

FIGURE 2.13: **Correlation between rescaled deviation of soil enzymatic activity from three null model predictions and normalized dissimilarity index.**

[a(1) to d(3)] Scatter plots show the standardized deviation of soil enzymatic activity from null model predictions for multiple-factor treatments. Net interaction type of each multiple-factor treatment is marked as different colored points (antagonistic, blue; synergistic, red; no interaction, gray). The betterst-fitting null model for each soil enzymatic activity has been selected based on the smallest model sum of squared deviation (SSD), and is indicated by the bold frame [a (2), b (3), c (2) and d (2)]. The correlations between standardized deviation of soil enzymatic activity from null model predictions and normalized dissimilarity index are shown in linear correlation with 95% confidence intervals. Spearman correlation coefficients and significance of correlations are indicated by R and p, respectively.

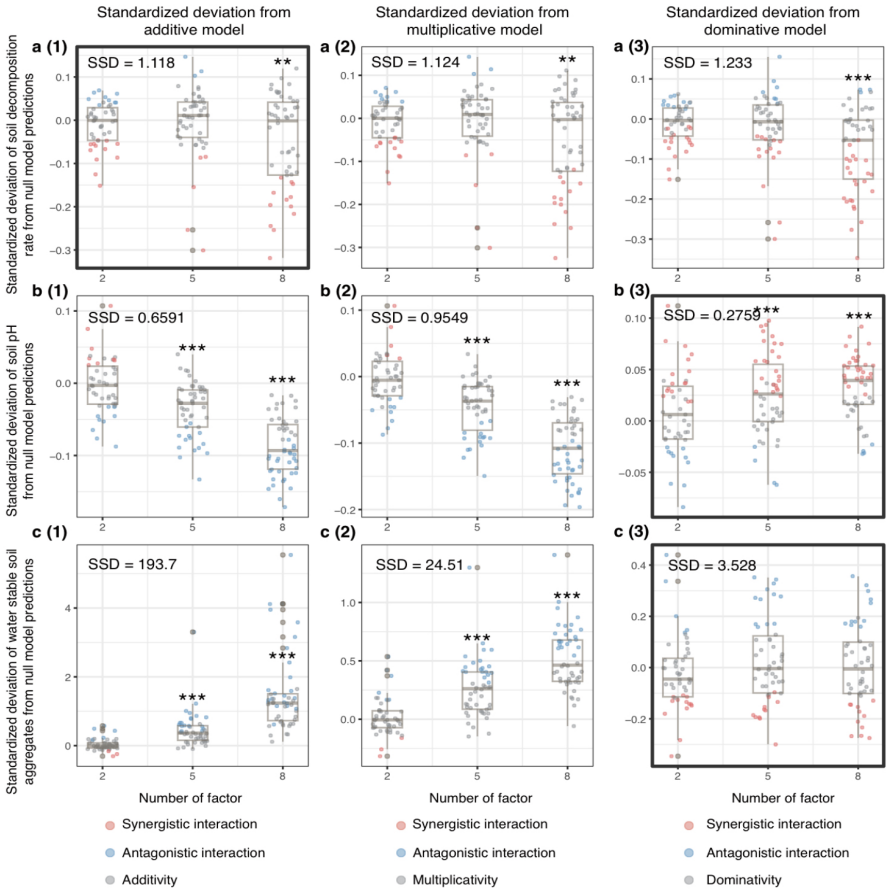

FIGURE 2.14: **Rescaled deviation of soil responses (soil decomposition rate, soil pH and water-stable soil aggregates) from three null model predictions with different numbers of factors.**

[a (1) to c (3)] Scatter points represent the standardized deviation of soil responses from null model predictions for multiple-factor treatments. Net interaction type of each multiple-factor treatment is marked as different colored points (antagonistic, blue; synergistic, red; no interaction, gray). The best-fitting null model for each soil response has been selected based on the smallest model sum of squared deviation (SSD), and it is indicated by the bold frame [a (1), b (3) and c (3)]. The overall deviances of soil responses from null models in three number of factor groups (factor = 2, 5 and 8) are shown by boxplot. The significant deviations from zero were evaluated by two sided t-tests (* 0.01 < P 0.05; ** 0.001 < P 0.01; *** P 0.001).

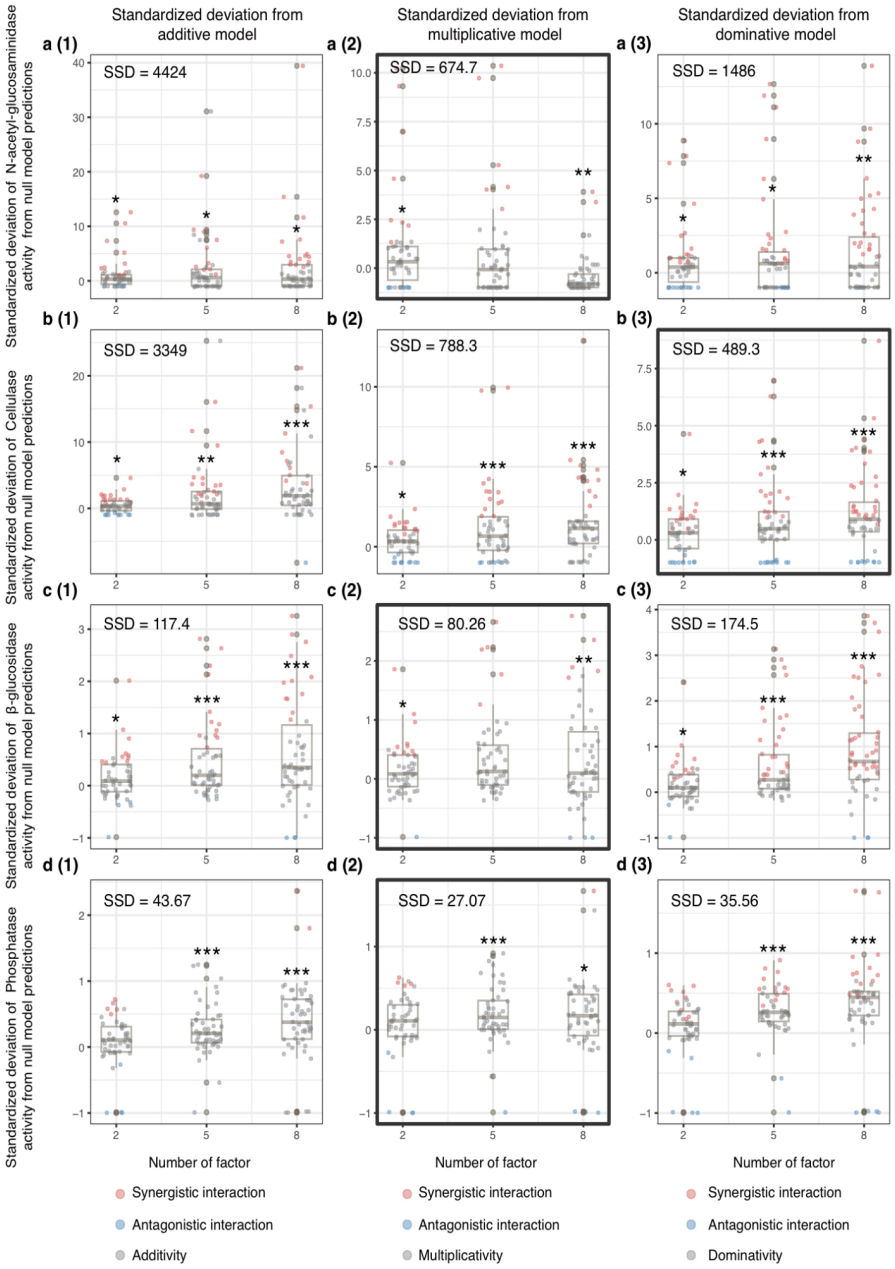

FIGURE 2.15: **Rescaled deviation of soil enzymatic activity from three null model predictions for different numbers of factors.**

[a (1) to c (3)] Scatter points represent the standardized deviation of soil enzymatic activity from null model predictions for multiple-factor treatments. Net interaction type of each multiple-factor treatment is marked as different colored points (antagonistic, blue; synergistic, red; no interaction, gray). The best-fitting null model for each soil enzymatic activity has been selected based on the smallest model sum of squared deviation (SSD), and it is indicated by the bold frame [a (2), b (3), c (2) and d (2)]. The overall deviances of soil enzymatic activity from null models in three number of factor groups (factor = 2, 5 and 8) are shown by boxplot. The significant deviations from zero were evaluated by two sided t-tests (*0.01 < P 0.05; ** 0.001 < P 0.01; *** P 0.001)

with the smallest sum of squared deviations (SSD) is used for estimating the deviation for each soil response. For soil decomposition rate, the additive model has the smallest SSD, while the dominative models have the smallest SSD for soil pH and water-stable soil aggregates. For soil enzymatic activity, the multiplicative models have the smallest SSD, except for cellulase activity (dominative model) (Appendix TABLE 5). The standardized deviation of soil decomposition rate from additive model predictions is correlated with the factor dissimilarity index, with synergistic interactions becoming more frequent with higher dissimilarity (FIGURE 2.12). The standardized deviation of cellulase and β-glucosidase activity from the null models with the smallest SSD also show correlations with factor dissimilarity index, with more frequent synergistic interactions appearing with higher dissimilarity (FIGURE 2.13). The soil decomposition rate responses to eight-factor treatments deviate significantly from the additive model (Appendix FIGURE 2.14), and also the soil enzymatic activity responses to higher number of factor treatments show significant deviation from the null models with the smallest SSD(Appendix FIGURE 2.15). These results suggest that the joint GCF effects on soil properties and functions deviate from null model predictions, and the number of factors and factor dissimilarity may drive the occurrence of more synergistic factor interactions.

2.5 Discussion

By assessing soil ecological responses to a large set of factor combinations (150 different factor combinations) at three different factor levels (2, 5 and 8), our study suggests that, in addition to the number of factors, factor dissimilarity also drives the effects of multiple GCFs. Our study supports previous findings that the number of co-acting factors affects soil responses to GCFs(Rillig et al.; 2019; Yang et al.; 2022). As hypothesized, (i) the effects of factor dissimilarity played an important role in predicting the variability of soil responses to multiple GCFs; (ii) a larger dissimilarity among factors or larger number of factors will cause greater deviation of joint effects on soil properties and fucntions from null models; (iii) co-acting factors with higher dissimilarity tend to have more synergistic interactions. This provides a new mechanistic perspective for predicting the joint effects of multiple GCFs on

soil properties and functions and highlights the importance of systematically understanding the properties and mode of action of single GCFs. Our findings also open new opportunities towards improving management approaches; management should prioritize local GCFs not just in terms of the most severe factor(s), but should also take into account factor dissimilarity building on known factor traits (FIGURE 2.16).

2.5.1 Separating factor identity effects in multiple GCFs studies

A simultaneous manipulation of a large number of factors (usually more than six) is needed in multiple factor research to test general rules for multiple factor effects. However, the contribution of the factor identity effect is an important point that cannot be ignored, otherwise this may result in contradictory results. The concept of "species identity effect" has been first raised in biodiversity studies for separating it from the diversity effect(Loreau; 1998; Loreau and Hector; 2001). Similarly, in multiple factor studies, factors following the design we use here are randomly chosen from a factor pool. When the number of selected factors increases, there is also an increasing probability of including single factors with extremely strong effects in high-level factor combinations. This higher chance of including extreme factors in the higher number of factors group results in a "GCF-number effect" but this effect is not due to the number of factor effects, but an increased rate of selecting extreme single factors. When testing the effect of factor dissimilarity, similarly, because the factor dissimilarity index is calculated based on the single factor effects, a GCF with more extreme effect on soil properties and functions will also have larger effect 'distances' to other GCFs. Therefore, a factor combination that includes the extreme factor is likely to have stronger combined effect and at the same time has a relatively larger dissimilarity index. In this case, the correlations between factor dissimilarity indices and soil properties and functions could be only caused by the higher selecting rate of extreme factors, and are insufficient to support the effect of factor dissimilarity. Thus, an appropriate statistical method is needed for disentangling the effect of factor number and dissimilarity from the factor identity effects.

Although only a few studies on multiple GCFs have assessed factor identity effects(Komatsu et al.; 2019; Zandalinas et al.; 2021), attempsts have been made in

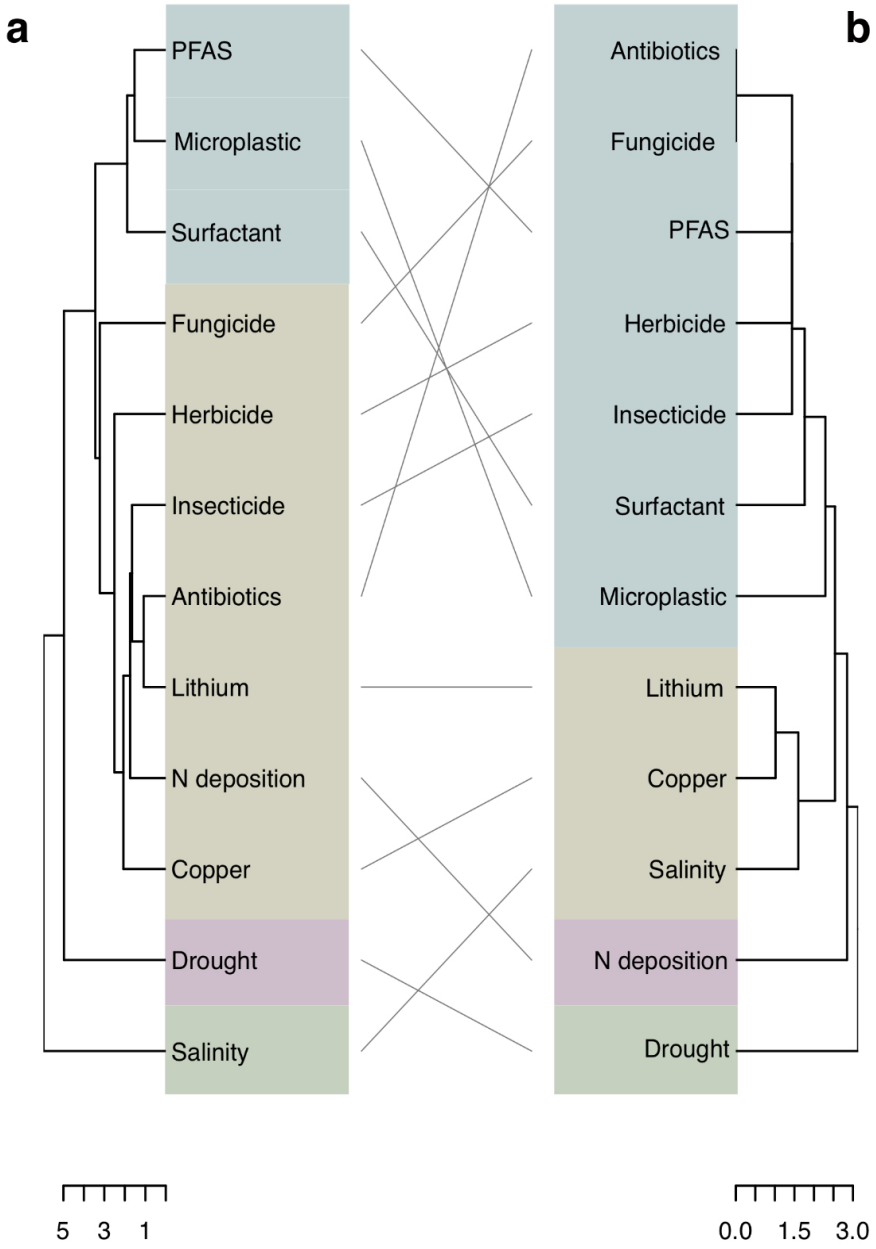

FIGURE 2.16: **A comparison of dendrograms of clustered 12 GCFs**

Dendrograms of 12 GCFs clustered based on their effects on soil properties and functions (measured in this experiment) (a) and based on a priori trait-based classification from expert opinions (b). Cophenetic correlation between two trees was calculated based on a bootstrapping method developed by a previous study(Rillig et al.; 2021a). The Cophenetic correlation test calculates the correlation between two cophenetic distance matrices of the two trees. The value of the cophenetic correlation coefficient can range between -1 to 1. With near 0 values meaning that the two trees are not statistically similar. After applying 1,000 times of permutation, the results show that the two dendrograms have structures more similar to each other than expected by chance (cophenetic correlation coefficient: mean = 0.2869 [95%CI: 0.073-0.475] > 0), indicating a priori trait-based ordering of factors having similar traits affecting soil properties and functions. It may provide advantages for future GCF management evaluating factor dissimilarity based on the priori trait-based factor classification systems.

some studies to use both parametric and nonparametric methods(Rillig et al.; 2019; Speißer et al.; 2022). Inspired by identifying contributions of species identities, previous work has employed a hierarchical diversity-interaction modeling framework based on linear mixed-effects models to assess the contribution of GCF identities by ANOVA tests(Speißer et al.; 2022). But the limitation of traditional statistical models is that they lack the power to capture unknown nonlinear patterns of the hypothesis in higher dimensionality(Ryo and Rillig; 2017). Another way to identify the individual factor contribution is comparing observed effects with effects predicted by null models. Null models assume that there are no interactions among factors(Schäfer and Piggott; 2018), thus the predictions of null models can be viewed as the combinations of single factor effects. To better deal with potential nonlinear relationships, a previous study explored the potential to combine machine learning algorithms with null model predictions to address the effects of factor number and to differentiate the factor identity contribution1. In our study, since null model prediction results have been integrated as the baseline models in the hierarchical modeling framework, the increase of the model predictability from baseline models can be interpreted as the contribution of factor number effect or factor dissimilarity effect other than the factor identity effect. Additionally, we further analyzed the change of standardized deviation of factor joint effect from the best-fitting null model across the dissimilarity index range, and indicated the direction of deviations (the types of the emergent interactions). In this way, the effect of factor dissimilarity can be depicted by both the explained variability increased from the baseline model and the changes of standardized deviation of soil responses from null models , but not directly from the correlation of factor dissimilarity with soil responses. For example, in our study, soil pH is strongly associated with factor dissimilarity, but the correlation is mostly caused by the contribution of extreme factors, because the baseline model can explain 85.4% of the response variability and adding other predictors hardly increase model R squared (FIGURE 2.9 and TABLE 2.5). And also, water-stable soil aggregation is negatively correlated with the dissimilarity index. However, from the standardized deviation analysis of WSA from the dominative null model prediction, we find the WSA responses are mostly subjected to the dominative model and there is no correlation between factor dissimilarity and standardized deviation from the

dominative model (FIGURE 2.12 and Appendix TABLE 5). As we saw contradictory results by analyzing data with or without considering the factor identity effect, to avoid misinterpreting results, we here suggest separating factor identity effects from other effects driven by multiple GCFs. Our study also provides a practical method to evaluate the contribution of factor identity effects and it should be used as the baseline for testing other hypotheses in future multiple GCFs studies based on randomly-drawn factors.

2.5.2 The role of the factor dissimilarity in driving the emergence of GCF interactive effects on soil properties and functions

Our study has found there is more emergence of synergistic factor interaction on soil decomposition rate, soil cellulase and β-glucosidase enzymatic activity when factor dissimilarity increases. By analyzing the deviation of factor joint effects from the best-fitting null model predictions across the dissimilarity index range, we found that factor dissimilarity drives the interactions of multiple GCFs towards a more synergistic direction (FIGURE 2.12 and FIGURE 2.15). Our finding indicates that factor dissimilarity underpins the interactive mechanisms of GCFs on soil properties and functions.

The role of factor dissimilarity in driving the interactive effects of multiple GCFs could be due to three distinct mechanisms. Firstly, factors that differ in their physicochemical nature may be more likely to have direct interactions compared to factors with the same physicochemical nature. These direct factor interactions are only related to the physical or chemical properties of the factor itself, without considering how they affect soil organisms and processes(Rillig et al.; 2021b). For example, drought can interact with other chemical factors as a concentration amplifier(De Vries et al.; 2020), and surfactants can increase the solubility or movement of organic pollutants(Dollinger et al.; 2018). Factor direct interactions usually amplify the intensity of single factors. Thus, the emergence of synergistic effects of high-dissimilarity multiple factors may be derived from the occurrences of direct factor interactions.

Secondly, factor dissimilarity may play a role in affecting species adaptation to

a multiple-GCF environment. Performance trade-offs are common when populations are exposed to multiple-factor environments(Langley et al.; 2022; Orr et al.; 2022a). Based on the pareto optimality theory, species cannot optimize adaptation to multiple factors at the same time(Shoval et al.; 2012; Tikhonov et al.; 2020). When factors are more dissimilar, these trade-offs would be larger, leading to a lower overall adaptation performance to the multiple-factor environments. By contrast, when the effects of factors are similar, the adaptation strategy of a population to one factor could also allow its adaptation to another factor by applying the same genetic or metabolic responses (e.g., cross-protection)(Bubliy and Loeschcke; 2005; MacMillan et al.; 2009).

Thirdly, factor dissimilarity could reshape the co-tolerance space of species to multiple factors. In the co-tolerance theory framework, the resistance of a community to multiple-stressor environments is affected by the relatedness of species' tolerances to different stressors(D. Vinebrooke et al.; 2004). If species tolerances to stressors are positively correlated, the overall species loss will be less than if the tolerances are unrelated, but if species tolerances to stressors are negatively correlated, more species (and the functions they drive) will be lost in multi-stressors environments. When multiple factors are quite different, their mode of action or target species range might also be different. From the perspective of the community, species tolerances to factors would be more negatively correlated. In this scenario, more species from the community are likely to be lost when more dissimilar GCFs are applied simultaneously. Biodiversity loss likely leads to reduced ecosystem functions (e.g., litter decomposition rate) based on biodiversity insurance theory(Loreau et al.; 2021); this may explain why there are more synergistically negative effects on soil decomposition rate when factor dissimilarity indices are higher (FIGURE 2.12 and 2.13).

2.6 Conclusion

Our study investigated the effects of factor dissimilarity on soil properties and ecological functions, suggesting that factor dissimilarity can drive more frequent occurrence of synergistic effects on several soil functions. Future work should address the

effects of factor dissimilarity at different levels of the ecological hierarchy (organism, population and community levels), or the temporal variation in effects of factor dissimilarity. Incorporating the effect of factor dissimilarity in future multiple GCFs studies will be helpful for estimating effects of co-acting GCFs and may be useful in informing protocols for ecosystem management and restoration.

Bibliography

Breiman, L. (2001). Random forests, *Machine Learning* **45**(1): 5–32.

Bubliy, O. A. and Loeschcke, V. (2005). Correlated responses to selection for stress resistance and longevity in a laboratory population of *Drosophila melanogaster*, *Journal of Evolutionary Biology* **18**(4): 789–803.

Crain, C. M. et al. (2008). Interactive and cumulative effects of multiple human stressors in marine systems, *Ecology Letters* **11**(12): 1304–1315.

D. Vinebrooke, R. et al. (2004). Impacts of multiple stressors on biodiversity and ecosystem functioning: The schimellifedrysoils2018role of species co-tolerance, *Oikos* **104**(3): 451–457.

de Vries, A. and Ripley, B. D. (2024). ggdendro: Create dendrograms and tree diagrams using 'ggplot2'. R package version 0.2.0.
URL: *https://andrie.github.io/ggdendro/*

De Vries, F. T. et al. (2020). Harnessing rhizosphere microbiomes for drought-resilient crop production, *Science* **368**(6488): 270–274.

Dieleman, W. I. J. (2012). Simple additive effects are rare: A quantitative review of plant biomass and soil process responses to combined manipulations of CO_2 and temperature, *Global Change Biology* **18**(9): 2681–2693.

Dollinger, J. et al. (2018). Effect of surfactant application practices on the vertical transport potential of hydrophobic pesticides in agrosystems, *Chemosphere* **209**: 78–87.

Efron, B. and Tibshirani, R. (1986). Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy, *Statistical Science* **1**(1).

Haj-Amor, Z. et al. (2022). Soil salinity and its associated effects on soil microorganisms, greenhouse gas emissions, crop yield, biodiversity and desertification: A review, *Science of The Total Environment* **843**: 156946.

Holmstrup, M. et al. (2010). Interactions between effects of environmental chemicals and natural stressors: A review, *Science of The Total Environment* **408**(18): 3746–3762.

Huang, Y. et al. (2018). Impacts of species richness on productivity in a large-scale subtropical forest experiment, *Science* **362**(6410): 80–83.

Jackson, C. R. et al. (2013). Determination of Microbial Extracellular Enzyme Activity in Waters, Soils, and Sediments using High Throughput Microplate Assays, *Journal of Visualized Experiments* (80): 50399.

Katzir, I. et al. (2019). Prediction of ultra-high-order antibiotic combinations based on pairwise interactions, *PLOS Computational Biology* **15**(1): e1006774.

Keuskamp, J. A. et al. (2013). Tea Bag Index: A novel approach to collect uniform decomposition data across ecosystems, *Methods in Ecology and Evolution* **4**(11): 1070–1075.

Klute, A. (1986). *Methods of Soil Analysis. Part 1 : Physical and Mineralogical Methods*, 2nd ed edn, American Society of Agronomy : Soil Science Society of America Madison, Wis., Madison, Wis.

Komatsu, K. J. et al. (2019). Global change effects on plant communities are magnified by time and the number of global change factors imposed, *Proceedings of the National Academy of Sciences* **116**(36): 17867–17873.

Kuhn M (2015). Classification and regression training, CRAN Repository.
URL: *https://cran.r-project.org/web/packages/caret/index.html*

Langley, J. A. et al. (2022). Do trade-offs govern plant species' responses to different global change treatments?, *Ecology* **103**(6): e3626.

Loreau, M. (1998). Separating Sampling and Other Effects in Biodiversity Experiments, *Oikos* **82**(3): 600.

Loreau, M. and Hector, A. (2001). Partitioning selection and complementarity in biodiversity experiments, *Nature* **412**(6842): 72–76.

Loreau, M. et al. (2021). Biodiversity as insurance: From concept to measurement and application, *Biological Reviews* **96**(5): 2333–2354.

MacMillan, H. A. et al. (2009). The effects of selection for cold tolerance on cross-tolerance to other environmental stressors in *Drosophila melanogaster*, *Insect Science* **16**(3): 263–276.

Oksanen J (2018). vegan: Community Ecology Package, R Foundation for Statistical Computing.

Orr, J. A. et al. (2022a). Rapid evolution generates synergism between multiple stressors: Linking theory and an evolution experiment, *Global Change Biology* **28**(5): 1740–1752.

Orr, J. A. et al. (2022b). Similarity of anthropogenic stressors is multifaceted and scale dependent, *Natural Sciences* **2**(1): e20210076.

R Core Team (2021). R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing.

Riedo, J. et al. (2021). Widespread Occurrence of Pesticides in Organically Managed Agricultural Soils—the Ghost of a Conventional Agricultural Past?, *Environmental Science & Technology* **55**(5): 2919–2928.

Rillig, M. C. et al. (2019). The role of multiple global change factors in driving soil functions and microbial biodiversity, *Science* **366**(6467): 886–890.

Rillig, M. C. et al. (2021a). Classifying human influences on terrestrial ecosystems, *Global Change Biology* **27**(11): 2273–2278.

Rillig, M. C. et al. (2021b). Mechanisms underpinning nonadditivity of global change factor effects in the plant–soil system, *New Phytologist* **232**(4): 1535–1539.

Ryo, M. and Rillig, M. C. (2017). Statistically reinforced machine learning for nonlinear patterns and variable interactions, *Ecosphere* **8**(11): e01976.

Schäfer, R. B. and Piggott, J. J. (2018). Advancing understanding and prediction in multiple stressor research through a mechanistic basis for null models, *Global Change Biology* **24**(5): 1817–1826.

Schimel, J. P. (2018). Life in Dry Soils: Effects of Drought on Soil Microbial Communities and Processes, *Annual Review of Ecology, Evolution, and Systematics* **49**(1): 409–432.

Shoval, O. et al. (2012). Evolutionary Trade-Offs, Pareto Optimality, and the Geometry of Phenotype Space, *Science* **336**(6085): 1157–1160.

Simmons, B. I. et al. (2021). Refocusing multiple stressor research around the targets and scales of ecological impacts, *Nature Ecology & Evolution* **5**(11): 1478–1489.

Smith, T. P. et al. (2024). High-throughput characterization of bacterial responses to complex mixtures of chemical pollutants, *Nature Microbiology* **9**(4): 938–948.

Speißer, B. et al. (2022). Number of simultaneously acting global change factors affects composition, diversity and productivity of grassland plant communities, *Nature Communications* **13**(1): 7811.

Tikhonov, M. et al. (2020). A model for the interplay between plastic tradeoffs and evolution in changing environments, *Proceedings of the National Academy of Sciences* **117**(16): 8934–8940.

Tilman, D. et al. (1996). Productivity and sustainability influenced by biodiversity in grassland ecosystems, *Nature* **379**(6567): 718–720.

Tudi, M. et al. (2021). Agriculture Development, Pesticide Application and Its Impact on the Environment, *International Journal of Environmental Research and Public Health* **18**(3): 1112.

Yang, G. et al. (2022). Multiple anthropogenic pressures eliminate the effects of soil microbial diversity on ecosystem functions in experimental microcosms, *Nature Communications* **13**(1): 4260.

Zandalinas, S. I. et al. (2021). The impact of multifactorial stress combination on plant growth and survival, *New Phytologist* **230**(3): 1034–1048.

Zhou, Z. et al. (2020). Meta-analysis of the impacts of global change factors on soil microbial diversity and functionality, *Nature Communications* **11**(1): 3072.

Chapter 3

A null model workflow developed for multiple GCF studies

3.1 Abstract

Human activities are driving significant changes in Earth's ecosystems, with soil functioning increasingly threatened by multiple anthropogenic factors. While many studies have focused on the effects of individual factors, recent research underscores the importance of evaluating the combined effects of multiple factors. Null model approaches, widely used in toxicology and ecotoxicology to assess the combined impacts of multiple stressors, have been recently adopted in soil ecology studies for evaluating factor interactions. However, there are still obstacles for applying null model approaches in GCF studies when more factors are considered simultaneously. Here we present a practical null model analysis workflow that generates predictions based on three null model assumptions for specified factor combinations. This workflow not only identifies net interactions within replicated factor combinations but also efficiently generates predictions for a large number of randomly selected factor combinations. This flexible modeling workflow can also be used with other modeling frameworks for hypotheses testing in multiple GCF studies, such as evaluating the contribution of selection effect and factor identity. Here, we detail the workflow and demonstrate its utility through two case studies involving global change factor research.

3.2 Introduction

Human activities are inducing significant alterations to the Earth's ecosystems in various ways(Richardson et al.; 2023). Soil ecosystem biodiversity and functionality are increasingly threatened by a range of human-induced factors(Díaz et al.; 2019), including climate change, pollutants, microplastics, nitrogen deposition etc(Rillig et al.; 2021). Beyond examining the isolated effects of each factor on soil ecosystem functioning, recent studies highlight the importance of understanding how these factors interact(Rillig et al.; 2019; Speißer et al.; 2022). The combined impacts of multiple co-acting factors on soil ecosystems are often profound and unpredictable due to complex factor interactions(Rillig et al.; 2019; Larsen et al.; 2011). Developing a practical method for quantifying the factor interactions is crucial for studying the combined effects of multiple factors on soil ecosystems.

The investigation of multiple stressors is well-developed in fields such as toxicology and ecotoxicology, where studies commonly assess the combined effects of multiple chemical stressors on individual organisms or populations using null models(Schäfer et al.; 2023; Schäfer and Piggott; 2018). In soil ecology, studies on multiple factors have similarly integrated null models to evaluate factor interactions and test hypotheses(Rillig et al.; 2019). However, as the number of factors increases, it becomes impractical for controlled experiments to test every possible combination with sufficient replication(Katzir et al.; 2019). Thus, many studies employ a random selection experimental design, in which factors are randomly chosen from a predefined pool, to address the "curse of dimensionality"(Speißer et al.; 2022; Rillig et al.; 2019). Despite the advantages of this approach, the random selection of factors also poses challenges for applying null model approaches for further understanding the underlying mechanisms of multiple factor interactions.

To address these challenges, we developed a null model analysis workflow capable of generating model predictions based on three null model assumptions for specified factor combinations. This workflow not only identifies factor interactions within replicated factor combinations but also efficiently generates predictions for a large number of randomly selected factor combinations using one-hot encoded factor identity features (where the presence or absence of factor is represented by "1"

and "0", respectively). Here we provide a step-by-step introduction to the workflow and present two case studies demonstrating its application in multiple global change factor (GCF) studies. The first case study investigated the combined effects of multiple COVID-19 pharmaceutical drugs and microplastics on soil microbial activity and enzymatic activity. The second case study evaluated the effects of 250 different combinations of GCF from a pool of 14 factor on soil pH, decomposition rate and water-stable soil aggregates, across eight levels of factor inclusion (0, 1, 5, 6, 7, 8, 9 and 10 factors).

3.3 Null model analysis workflow for multiple GCF studies

We created a null model analysis workflow that can be easily applied to multiple GCF studies with various data structure. The whole analysis workflow was performed with R (4.1.1) (R Core Team; 2021).

3.3.1 Step 1. Calculating effect sizes for individual factors

.

To make this workflow easily applied to any kind of dataset from multiple factor studies, we make this workflow compatible to a well-known package "dabestr"(Ho et al.; 2019). "Dabestr" is a R package that use Bootstrap-coupled methods to generate estimates of effect sizes for different treatments. First we calculate the single factor effect sizes using the function dabest() and mean_diff(). By using 'dabestr' package, the effect sizes of all single factors can be estimated by a bootstrapping method with defult 500 permutations. The output of mean_diff() function will be used in the following steps for generating null model predictions of any specified combination of single factors.

```
Response_ES<-data%>%
  dabest(Treatment, Response,
         idx = c("Control","factor_A","factor_B","factor_C","factor_D","factor_E","factor_F","factor_G","factor_H"),
         paired = FALSE)

Response_meandiff<-mean_diff(Response_ES,reps = 1000)
```

3.3.2 Step 2. Calculating null model prediction of specific factor combinations by NullModel function

The "NullModel" function was developed for generating bootstrapped estimates of three null model predictions (additive, multiplicative and dominative null models) based on specified factor combinations. The rationale of the three null model calculation is provided in Chapter 2 Methods.

```
NullModel = function(object, selected_factors=vector(), n_perm = 500){

  output = list()
  CT = object[["summary"]][["mean"]][[1]]
  input=object[["data"]]
  population_CT= input[input[,as.character(object[["x"]][[2]])] == object[["result"]][["control_group"]][1], object[["result"]][["variable"]][1]]
  size_CT=length(population_CT)

  for (type in c("additive","multiplicative", "dominative")) {
    bs = numeric(0)
    for (id in c(1:n_perm)) {
    each_effect = numeric(0)
    k_CT = mean(sample(population_CT, size_CT, replace = T))
      for (treatment in selected_factors) {
      population_TR = input[input[,as.character(object[["x"]][[2]])] == object[["result"]]
      [["test_group"]][which(object[["result"]][["test_group"]]==treatment)], object[["result"]][["variable"]][1]]
      size_TR = length(population_TR)
      k_TR = mean(sample(population_TR,size_TR,replace = T))

      # ES estimate depending on the type of null hypothesis
      if(type == "additive")        each_effect = append(each_effect,(k_TR - k_CT))
      if(type == "multiplicative")  each_effect = append(each_effect, (k_TR-k_CT)/k_CT)
      if(type == "dominative")      each_effect = append(each_effect,(k_TR - k_CT))
      }

    if(type == "additive") {
      joint_effect = sum(each_effect)
      pre_response = joint_effect+CT
    }

    if(type=="multiplicative"){
      z = 1
      for(m in c(1:length(selected_factors))) {
      z = z * (1 + each_effect[m])
      joint_effect = (z - 1)*k_CT
      }
      pre_response =joint_effect+CT
    }

    if(type=="dominative")  {
      joint_effect = each_effect[which(max(abs(each_effect))==abs(each_effect))]
      pre_response =joint_effect+CT
    }

    bs = append(bs, pre_response)
    }
  output[[type]] = bs
  }
  return(output)
}
```

In "NullModel" function, input object is the output of "mean_diff" function of "dabestr" package. The specified factor combination is required by the input parameter: `selected_factors = c()`. The times of permutation can be defined by: `n_per = 500`.

```
Null_modle = NullModel(Response_meandiff, selected_factors = c("factor_A","factor_C","factor_G"))
```

3.3.3 Step 3. Significance testing between null model predictions and actual data

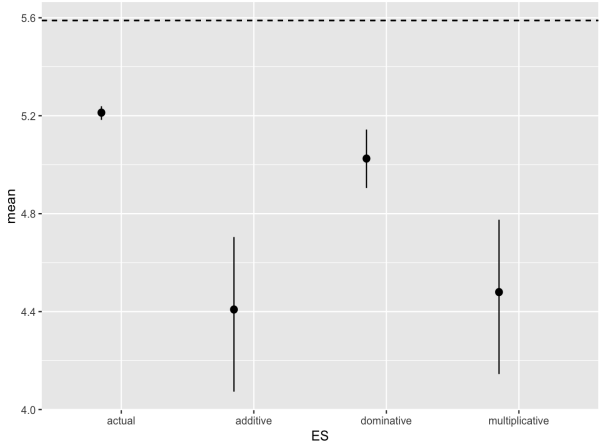

FIGURE 3.1: **Significant testing between three null model prediction and actual data**

The "NullModel_summary" function was developed for calculating the 95% confidence interval of each null model prediction and applying significance testing between actual multi-factor treatment data and null model predictions.

```
NullModel_summary = function(null_data, actual_data = vector()){
  output = list()

  if(length(actual_data)<=2){
    output[["actual"]] = c(mean(actual_data),mean(actual_data),mean(actual_data),1)
    for (type in c("additive","multiplicative", "dominative")) {
    bs = rep(mean(actual_data),length(null_data[[type]])) - null_data[[type]]
    p = length(which(bs>0))/length(bs)
    p = min(p, 1-p)
    output[[type]] = c(quantile(null_data[[type]], .025), mean(null_data[[type]]), quantile(null_data[[type]], .975), p)
    }
  }

  if(length(actual_data)>=3){
    #re-sampling of actual data
    size_actul=length(actual_data)
    bs_actual = numeric(0)
    for (i in c(1:length(null_data[[1]]))) {
      k_actual = mean(sample(actual_data, size_actul, replace = T))
      bs_actual = append(bs_actual, k_actual)
    }

    output[["actual"]] = c(quantile(bs_actual, .025), mean(bs_actual), quantile(bs_actual, .975), 1)

    for (type in c("additive","multiplicative", "dominative")) {
      bs = bs_actual - null_data[[type]]
      p = length(which(bs>0))/length(bs)
      p = min(p, 1-p)
      output[[type]] = c(quantile(null_data[[type]], .025), mean(null_data[[type]]), quantile(null_data[[type]], .975), p)
```

```
    }
  }

  output = t(data.frame(output))
  colnames(output) = c("X2.5%", "mean", "X97.5%", "p_value")
  output = data.frame(ES = c("actual", "additive","multiplicative","dominative"), output)
  row.names(output) = c()

  return(output)
}
```

The significant difference between actual data and null model prediction indicates the null model failed to predict the results (There is factor net interaction).

```
Null_modle_summary = NullModel_summary(Null_modle, actual_data = actual_data)
#Ploting
p_null = ggplot()+
  coord_flip() +
  geom_estci(data=Null_modle_summary, aes(x = mean, y = ES, xmin=X2.5., xmax=X97.5.,
                                  xintercept=pH.plot.meandiff[["summary"]][["mean"]][1]), center.linecolour = "black",
             size=0.6, ci.linesize = 0.5, position=position_nudge(y = -0.15))

p_null
```

3.4 Application case 1

3.4.1 Experimental design and measurements

During the worldwide COVID-19 pandemic, an increasing amount of different pharmaceutical drugs have been consumed, as well as plastic face masks. Here we designed an experiment to investigate the potential impacts of combined COVID-19 drugs and microplastic(MP) generated by facial masks on soil microbial activity and enzymatic activity.

In this study, we chose three drugs commonly used during COVID-19 pandemic: remdesivir (antiviral), azithromycin (antibacterial) and ivermectin (antiparasitic) and microplastics from FFP2 face mask. We tested individual effects and combined effects of those three drugs on soil microbial activity and enzymatic activity with low and high concentration. The working concentrations of drugs with high and low concentration are shown in Appendix TABLE 6. The name of each treatment and number of replicates are shown in TABLE 3.1.

TABLE 3.1: **Experimental design.**

Treatment	Applied factor(s)	Replicates
Control	None	8
R_low	Remdesivir (low concentration)	8
R_high	Remdesivir (high concentration)	8
A_low	Azithromycin (low concentration)	8
A_high	Azithromycin (high concentration)	8
I_low	Ivermectin (low concentration)	8
I_high	Ivermectin (high concentration)	8
MP	Microplastic	8
AIMP_low	Azithromycin, Ivermectin (low concentration), Microplastic	8
AIMP_high	Azithromycin, Ivermectin (High concentration), Microplastic	8
RIMP_low	Remdesivir, Ivermectin (low concentration), Microplastic	8
RIMP_high	Remdesivir, Ivermectin (high concentration), Microplastic	8
ARMP_low	Azithromycin, Remdesivir (low concentration), Microplastic	8
ARMP_high	Azithromycin, Remdesivir (high concentration), Microplastic	8
RAIMP_low	Azithromycin, Remdesivir, Ivermectin (low concentration), Microplastic	8
RAIMP_high	Azithromycin, Remdesivir, Ivermectin (high concentration), Microplastic	8

Every experimental unites has been conducted in a microcosm system, by putting 40.0 g (d.w.) soil in a 50.0 mL falcon tube. The pharmaceutical drugs and 0.4% microplastics have been applied to the corresponding soil treatments at the beginning at one time. After 42 days incubation, 5.0 g fresh soil has been harvested from each treatment for further soil enzymatic activity (N-acetyl-glucosaminidase, cellulase, β-glucosidase and phosphatase) and microbial activity (FDA hydrolysis) measurements.

3.4.2 Statistical analysis

Data analysis and visualization were conducted using R (R Core Team; 2021). Effect sizes and 95% confidence intervals (CIs) for both single and multiple factor treatments were estimated employing a bootstrap method with 10,000 permutations. Visualizations depicting effect sizes and raw data distributions for each treatment group were created. A positive effect indicates that the measured response variable

was higher in the treatment group compared to the control, while a negative effect signifies the opposite.

To assess the interactive potential of multifactorial treatments on soil enzymatic activity and FDA hydrolysis, this study employed three null model assumptions — additive, multiplicative, and dominative — following the null model workflow we developed. Under the additive model, the joint effects of factors are estimated by summing the effect sizes of individual components. The dominative model assumes the joint effect equals the largest absolute effect size of any single factor. The multiplicative model estimates combined effects by multiplying the proportional changes induced by each single-factor treatment against a control. These models presuppose no interactions among factors, serving as baseline predictions. Interactions are identified through significant deviations of observed data from these null model predictions, indicating either synergistic or antagonistic interactions among the factors involved. Data visualization was conducted using the ggplot2 package (Wickham; 2016).

3.4.3 Results

Effects of drugs and microplastics on soil enzymatic and microbial activity

For evaluating the effects of COVID-19 drugs and microplastics from masks on soil microbial activity and enzymatic activity, We measured overall microbial activity by FDA hydrolysis and four excellular soil enzyme activity (Appendix FIGURE 3 and 4). Adding microplastic had negative effects on β-glucosidase, phosphatase activity and FDA hydrolysis activity. We found an general decrease in FDA activity for drug treatments applied singly and in combination. Low-concentration of pharmaceutical drugs had negative effects on N-acetyl-glucosaminidase and cellulase activity, while high-concentration of drugs showed positive effects (Appendix FIGURE 3 and 4).

Net interactions between drugs and microplastics on soil enzymatic and microbial activity

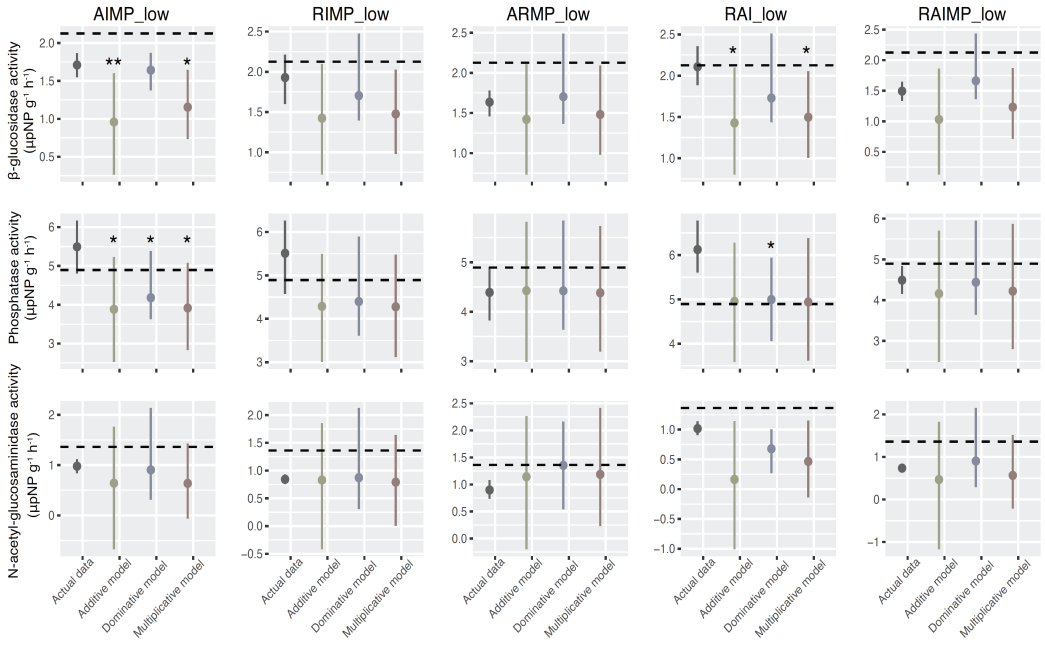

FIGURE 3.2: **The null model predictions for soil enzymatic response to different combinations of low concentration pharmaceutical drugs and mask MP.**

For each soil enzymatic activity response, the figure shows the predictions of three null models (additive, multiplicative and dominative) for each low-concentration pharmaceutical drups combination (with or without MP. The 95% Confidential intervals are generated by a bootstrapping method with 1,000 time permutations. The significant difference of actual data from null model predictions were evaluated by two sided t-tests (*0.01 < P 0.05; **0.001 < P 0.01;***P 0.001). (remdesivir, R; azithromycin, A; ivermectin, I; mask microplastic, MP)

To investigate the potential net interactions among COVID-19 drugs and mask microplastics, we used the null model workflow to compare the actual soil responses to different combinations of drugs and microplastics to the null model predictions generated from the same factor combination based on the actual soil responses to factors applied individually (FIGURE 3.2, 3.3, 3.4 and 3.5).

For low-concentration pharmaceutical drugs, most of the soil enzymatic activity responses fit to at least one model from the three null model predictions (additive, multiplicative and dominative model), except for phosphatase activity to AIMP_low treatment and FDA hydrolysis activity to ARMP_low treatment. The actual response

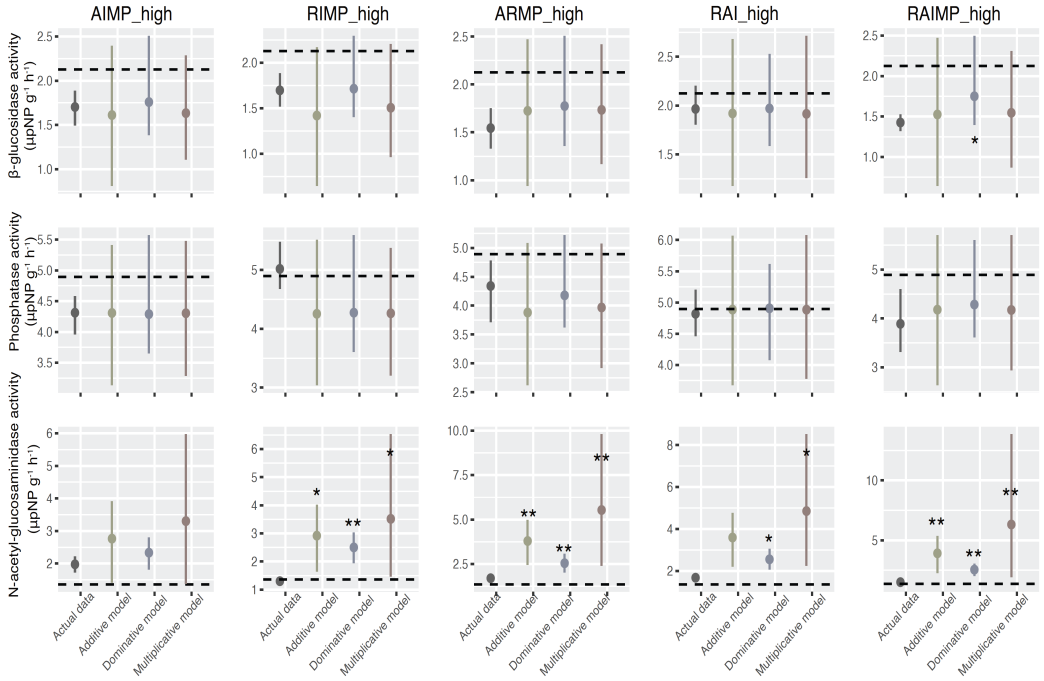

FIGURE 3.3: **The null model predictions for soil enzymatic response to different combinations of high concentration pharmaceutical drugs and mask MP.**

For each soil enzymatic activity response, the figure shows the predictions of three null models (additive, multiplicative and dominative) for each high-concentration pharmaceutical drups combination (with or without MP. The 95% Confidential intervals are generated by a bootstrapping method with 1,000 time permutations. The significant difference of actual data from null model predictions were evaluated by two sided t-tests (*0.01 < P 0.05; **0.001 < P 0.01;***P 0.001). (remdesivir, R; azithromycin, A; ivermectin, I; mask microplastic, MP)

of soil phosphatase activity to co-occurence of low concentration azithromycin, ivermectin and microplastics is positive, while the null model predictions are all negative (FIGURE 3.2). The actual joint effect of low concentration azithromycin, remdesivir and microplastics on soil FDA hydrolysis activity is smaller than the negative effects predicted by three null models (FIGURE 3.4).

For the combined effects of high-concentration drugs on soil microbial and enzymatic activity, most of the null models predictions are significantly different from the responses of N-acetyl-glucosaminidase and cellulase activity and FDA hydrolysis activity. The predicted effect sizes of combined factors are smaller than the actual responses, indicating that there are potential antagonistic interactions among the high-concentration drug and microplastics (FIGURE 3.3 and 3.5).

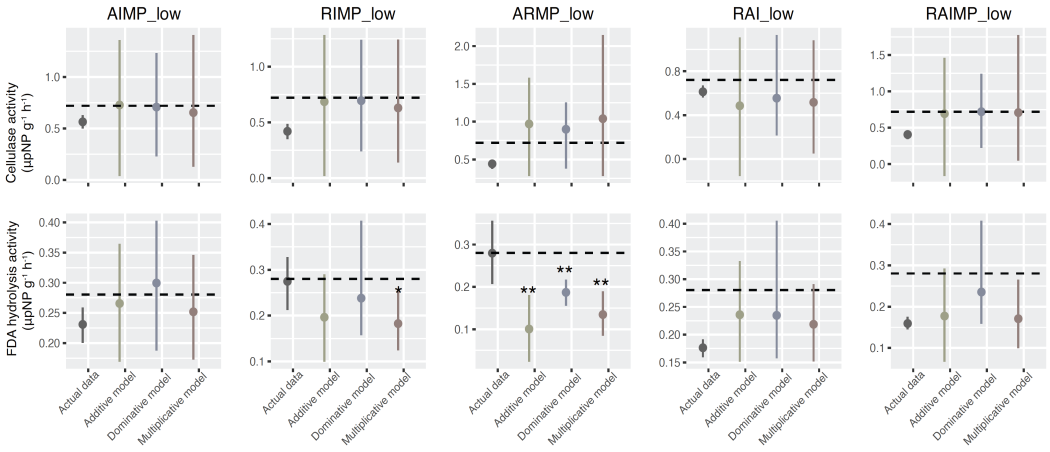

FIGURE 3.4: **The null model predictions for soil enzymatic response to different combinations of low concentration pharmaceutical drugs and mask MP.**

For each soil enzymatic activity response, the figure shows the predictions of three null models (additive, multiplicative and dominative) for each low-concentration pharmaceutical drups combination (with or without MP. The 95% Confidential intervals are generated by a bootstrapping method with 1,000 time permutations. The significant difference of actual data from null model predictions were evaluated by two sided t-tests (*0.01 < P 0.05; **0.001 < P 0.01;***P 0.001). (remdesivir, R; azithromycin, A; ivermectin, I; mask microplastic, MP)

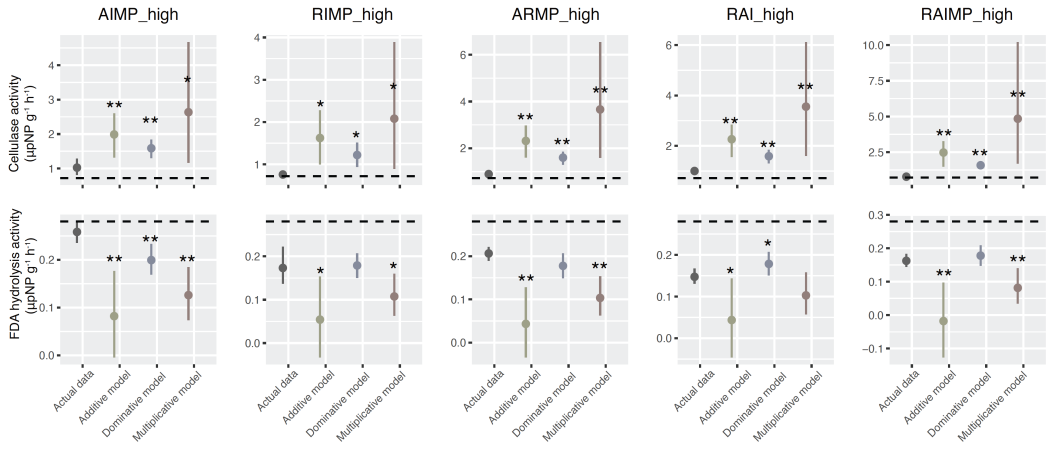

FIGURE 3.5: **The null model predictions for soil enzymatic response to different combinations of high concentration pharmaceutical drugs and mask MP.**

For each soil enzymatic activity response, the figure shows the predictions of three null models (additive, multiplicative and dominative) for each high-concentration pharmaceutical drups combination (with or without MP. The 95% Confidential intervals are generated by a bootstrapping method with 1,000 time permutations. The significant difference of actual data from null model predictions were evaluated by two sided t-tests (*0.01 < P 0.05; **0.001 < P 0.01;***P 0.001). (remdesivir, R; azithromycin, A; ivermectin, I; mask microplastic, MP)

3.5 Application case 2

3.5.1 Experimental design and soil measurements

When soil ecosystems are exposed to an increasing number of anthropogenic stressors, the potential high-order interactions among these factors remains largely unknown. To investigate the effect of such interactions, we designed an experiment using a pool of 14 global change factors (GCFs): tire particles, antibiotic, fungicide, herbicide, insecticide, surfactants, nitrogen deposition, salinity, copper, lithium, drought, PFAS, microfibers and soil compaction. The working concentration of each factor is provided in the TABLE 3.2. To create factor combinations that could have potential high-order interactions, we choose five multi-factor levels (6, 7, 8, 9 and 10 factors; 50 replicates each factor level). Similar to a previous experimental design Rillig et al. (2019), factors were randomly selected for each combination from the GCF pool. Additionaly, single-factor treatments (with six replicates per factor) and a control group (10 replicates) were included. In total, the experiment consisted of 344 experimental units.

a: Water holding capacity

TABLE 3.2: **Working concentration and applications of 14 factors.**

Factors	**Abbreviations**	**Applied factor**	**Concentration**
Microplastics	M	Tire particles	1 $g\,kg^{-1}$
Antibiotic	A	oxytetracycline	3.00 $mg\,kg^{-1}$
Fungicide	FU	carbendazim	6.00 $mg\,kg^{-1}$
Herbicide	H	diflufenican	1.00 $mg\,kg^{-1}$
Insecticide	I	imidacloprid	0.05 $mg\,kg^{-1}$
Surfactant	SU	sodium dodecylbenzenesulfonate	16.0 $mg\,kg^{-1}$
N deposition	N	NH_4NO_3	439.6 $mg\,kg^{-1}$
Salinity	S	NaCl	4.00 $g\,kg^{-1}$
Heavy Metals	C	$CuSO_4 * 5H_2O$	100 $mg\,Cu\,kg^{-1}$
Lithium	L	LiCl	11.40 $mg\,kg^{-1}$
Drought	D	water content	30% WHC^a
PFAS	P	PFOA perfluorooctanoic acid	1.00 $mg\,kg^{-1}$
Microfiber	MF	polyester fibers	1 $g\,kg^{-1}$
Soil compaction	CP	soil density	1.7 $g\,cm^{-3}$

The soil used for this experiment was collected from a local grassland at an experimental site of Free university Berlin. The soil was a sandy loam texture. prior to the experiment, the soil was air-dried and sieved through a 2 mm mesh to remove large stones and coarse grass roots. Each experimental unit consisted of a 50 mL mini-bioreactor equipped with a vented film to allow gas exchange while preventing microbial contamination. Inside each bioreactor, 40.0g (dry weight) of soil was added. Loading soil (5.0 g) was used to ensure even distribution of small amounts of chemicals in the experimental units, with sterilization performed to prevent localized (FIGURE 3.6 and 3.7).

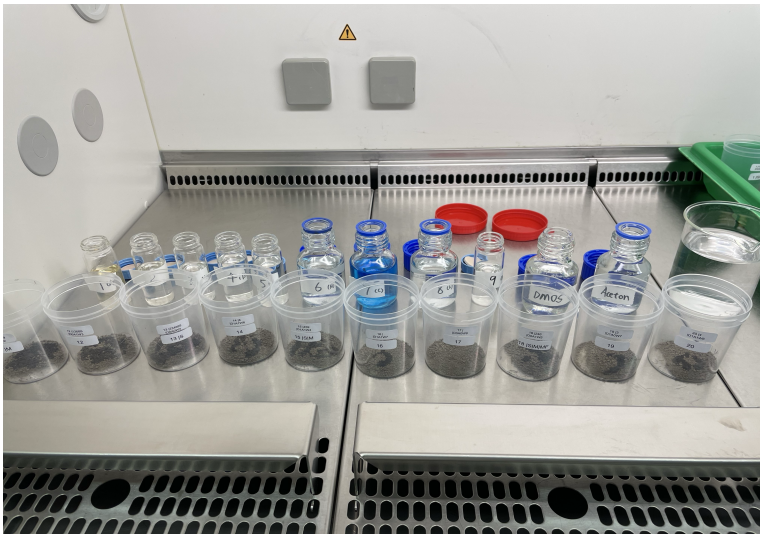

FIGURE 3.6: **Adding multiple GCF treatments to the loading soils.**

The soil response variables measured in this experiment included litter decomposition rate, soil pH, and water-stable aggregates (WSA). The litter decomposition rate was assessed using the tea bag index method(Keuskamp et al.; 2013). In brief, a sealed mesh bag (38 μm mesh size) containing 300.0 mg (dry weight) of tea biomass was placed into the soil. The proportional weight loss of the tea biomass before and after incubation was used to calculate the decomposition rate. Soil pH was determined by mixing 5.0 g of air-dried soil with 25 mL of distilled water in a 50 mL centrifuge tube, followed by measurement with a pH meter (Hanna Instruments, Smithfield, USA). The proportion of WSA was measured using a modified protocol based on a previously established method(Klute; 1986).

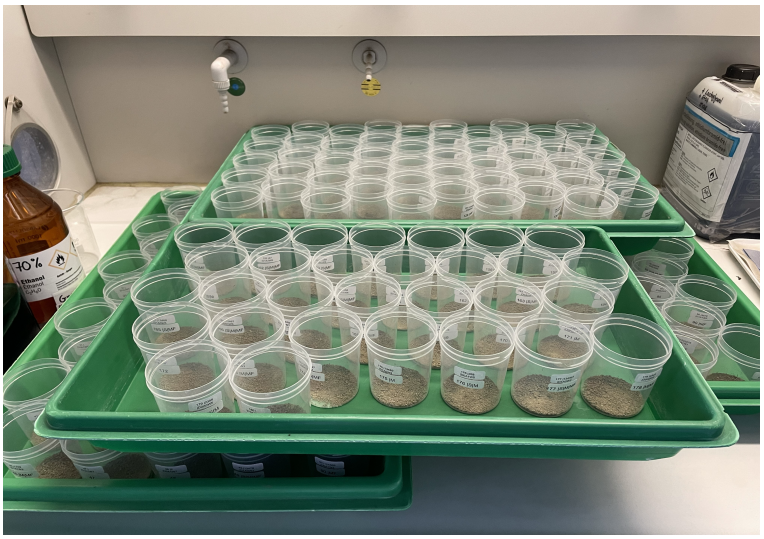

FIGURE 3.7: **Loading soils for the treatments of multiple factor experiment**

3.5.2 Statistical analysis

Data analysis and visualization were performed using R(R Core Team; 2021). Effect sizes and 95% confidence intervals (CIs) for single and multi-factor treatments were estimated using a bootstrap approach with 10,000 permutations. Visualizations, including effect sizes and raw data distributions for each treatment group, were generated to facilitate interpretation of the results.

To investigate the potential interactions among factors, we used the Null model workflow to identify the factor net interaction for 250 factor combinations. First, we calculated the null model predictions for each factor combination depending on the component single factor effects. Secondly, we compared the actual effect of a certain multi-factor treatment to the Null model-predicted effect. Actual effects that fit into the 95% of the null model confidence intervals (generated by 1,000 bootstrapping permutations) were classified as no interaction, others were categorized as antagonistic or synergistic net interaction. Then we calculated the rescaled deviations of actual effects from the null model predictions and assessed the overall deviation of all multi-factor treatments from three null model assumptions, indicating by the sum of squared deviation (SSD). Rescaled deviation of actual effects from collective null model predictions were also calculated for comparing the methods. Collective null model predictions were generated by non-specified factor combinations. For each

factor level, the null model predictions were generated by the single factor effects of randomly chosen n-factors (n = 6, 7, 8, 9 or 10).

3.5.3 Results

3.5.3.1 Effects of single and multiple factors on soil decomposition rate, soil pH and WSA

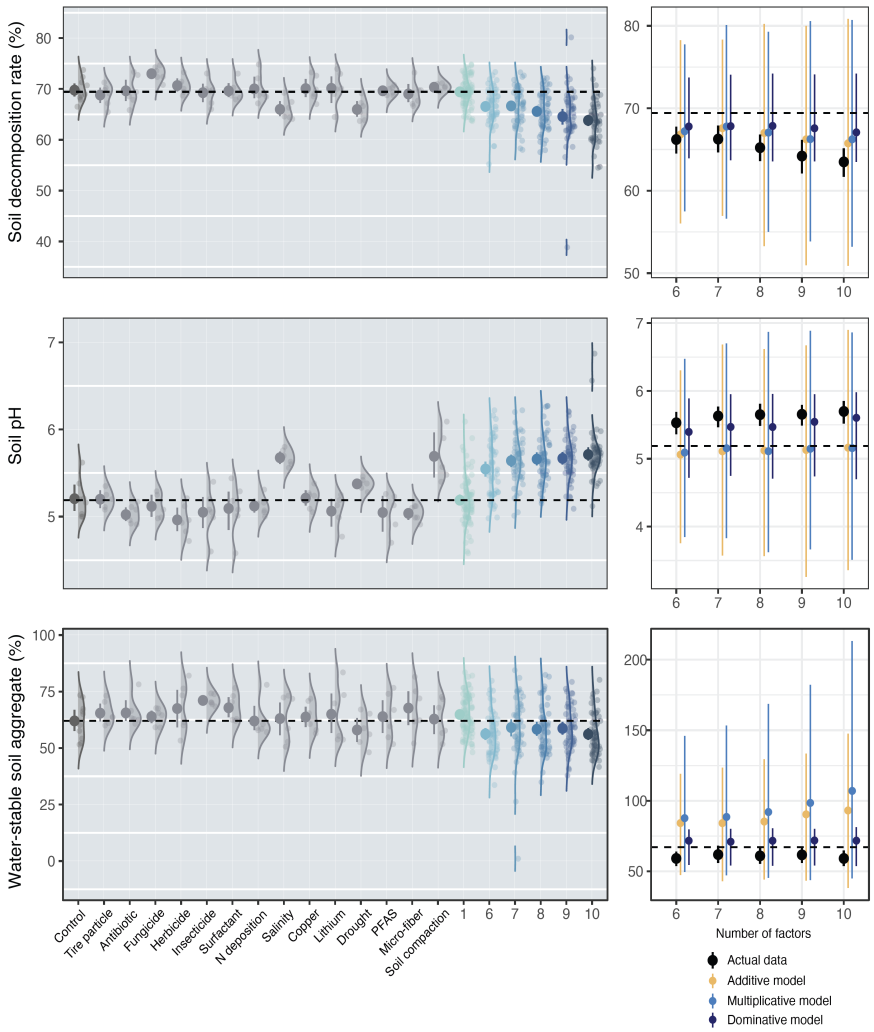

FIGURE 3.8: **Response of soil decomposition rate, soil pH and WSA to multiple GCF treatments applied singly or simultaneously.**

On the left panel, for each soil function and property, effect sizes of single factor treatments and multiple factor treatments (6, 7, 8, 9, 10 factors, 50 treatments included in each factor group) were estimated by bootstrapping method with 1,000 permutations. On the right panel, single-factor effect sizes were used to predict multi-factor effects based on three different assumptions (additive, multiplicative and dominative) in each number of factor group.

The 14 different GCFs have different effects on soil decomposition rate, soil pH and water-stable soil aggregate. Fungicide has positive effects on soil decomposition rate. Salinity increased soil pH. Single GCF does not have significant effect on water stable soil aggregates (TABLE 3.3). For the multiple factor treatments, soil decomposition decreased while soil pH increased when there are multiple GCFs applied to the soils (FIGURE 3.8).

TABLE 3.3: **Significance test for the effects of single GCFs on soil responses.**

Factor	Decompo-	Soil pH	Water stable soil
PFAS	0.9029	0.8981	0.9657
Copper	0.9029	0.8981	0.9657
Lithium	0.9029	0.8981	0.9657
N deposition	0.9029	0.8981	0.9657
Antibiotic	0.9029	0.1654	0.9657
Insecticide	0.9029	0.8981	0.0961
Surfactant	0.9029	0.8981	0.9657
Fungicide	**0.0475**	0.8981	0.9657
Herbicide	0.9029	0.1536	0.9657
Microplastic	0.9029	0.8981	0.9657
Salinity	0.0643	**0.006***	0.9657
Drought	0.0643	0.8981	0.9657
Microfibers	0.9029	0.1675	0.9657
Soil compaction	0.9029	0.1675	0.9657

Adjusted P values based on Benjamini-Hochberg method obtained from two sided t-tests between each treatment group and controls. (Significant differences with 0.01 < P 0.05 are shown in bold, and P < 0.01 are marked by * additionally.)

TABLE 3.4: **Significance test for the effects of multiple GCFs on soil responses.**

Factor level	Decompo-	Soil pH	Water stable soil
6	**0.0016***	**0.0073***	0.2742
7	**0.0016***	**0.0015***	0.3831
8	**0.0003***	**0.0015***	0.3831
9	**0.0001***	**0.0015***	0.3831
10	**<0.0001***	**0.0009***	0.2742

Adjusted P values based on Benjamini-Hochberg method obtained from two sided t-tests between each treatment group and controls. (Significant differences with 0.01 < P 0.05 are shown in bold, and P < 0.01 are marked by * additionally.)

For the multiple factor treatments, soil decomposition rate decreased significantly for all 6, 7, 8, 9 and 10 factor groups. Soil pH increased significantly for all 6, 7, 8, 9 and 10 factor groups. However, multiple factor treatment do not have significant effects on water-stable soil aggregation (FIGURE 3.8 and TABLE 3.4).

3.5.3.2 Deviations of multi-factor treatment effects from the null model predictions

By analysing the rescaled deviations of soil decomposition rate from the null model predictions for multi-factor treatments, we found that no interactions emerged from the 5 factor levels (FIGURE 3.9). The sum of squared deviation (SSD) from the specified null model predictions are smaller than the SSD from the collective null model predictions.

The rescaled deviations of soil pH responses from the specified null model predictions indicate that there is emergence of net synergistic interactions among co-acting multi-factors. The SSD from the specified null model predictions are larger than the SSD from the collective null model predictions, indicating factor interacting effects might be underestimated by using collective null model prediction (FIGURE

3.10).

The rescaled deviations of water-stable soil aggregate from null model predictions show that there are also potential synergistic interactions among co-acting factors. The SSD from the specified null model predictions are not very different from

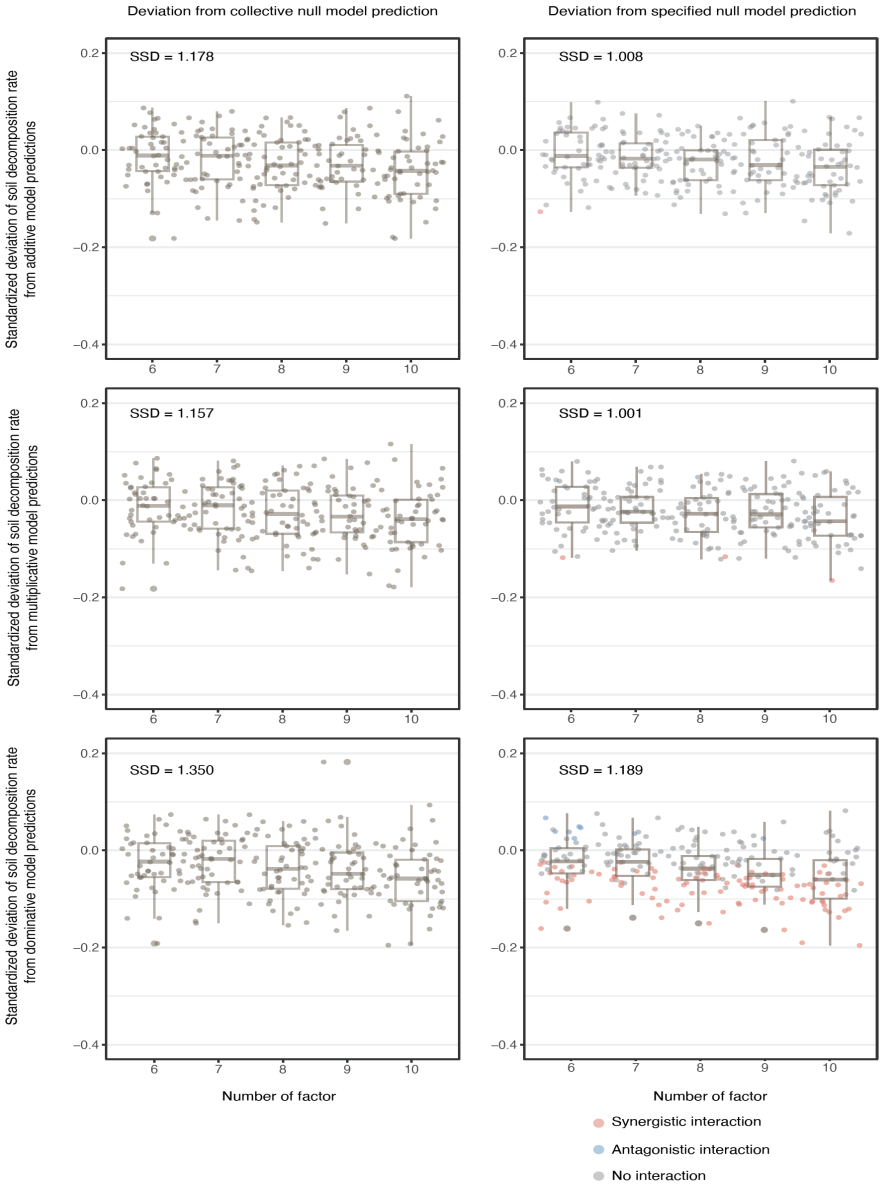

FIGURE 3.9: **Rescaled deviation of soil decomposition rate from three null model predictions.**

Scatter points represent the standardized deviation of soil decomposition rate responses from collective null model predictions (on the left panel) and from specified null model predictions (on the right panel) for multiple-factor treatments. For the deviations form specifed null model prediction, the net interaction type of each multiple-factor treatment is marked as different colored points (antagonistic, blue; synergistic, red; no interaction, gray). The smallest model sum of squared deviation (SSD) is shown for each figure.

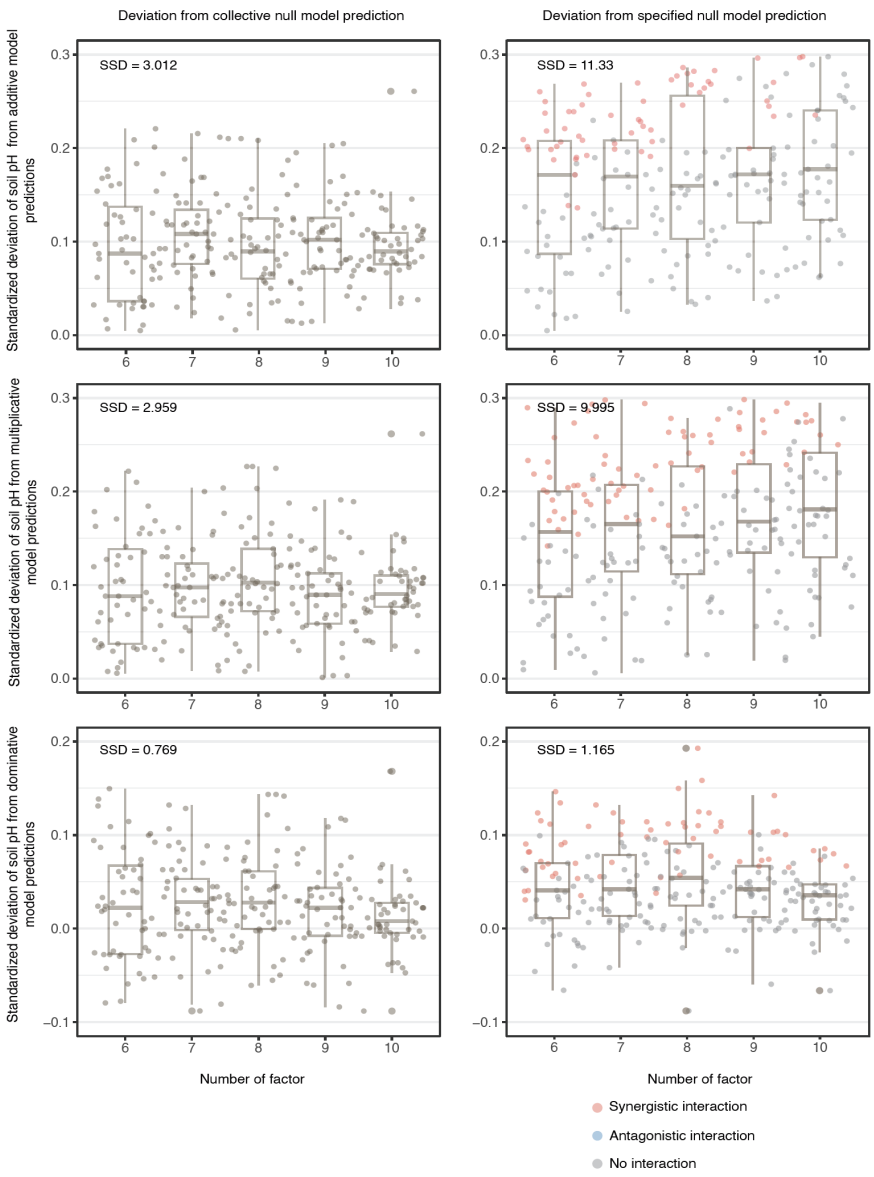

FIGURE 3.10: **Rescaled deviation of soil pH from three null model predictions.**

Scatter points represent the standardized deviation of soil pH responses from collective null model predictions (on the left panel) and from specified null model predictions (on the right panel) for multiple-factor treatments. For the deviations form specifed null model prediction, the net interaction type of each multiple-factor treatment is marked as different colored points (antagonistic, blue; synergistic, red; no interaction, gray). The smallest model sum of squared deviation (SSD) is shown for each figure.

the SSD from the collective null model predictions (FIGURE 3.11).

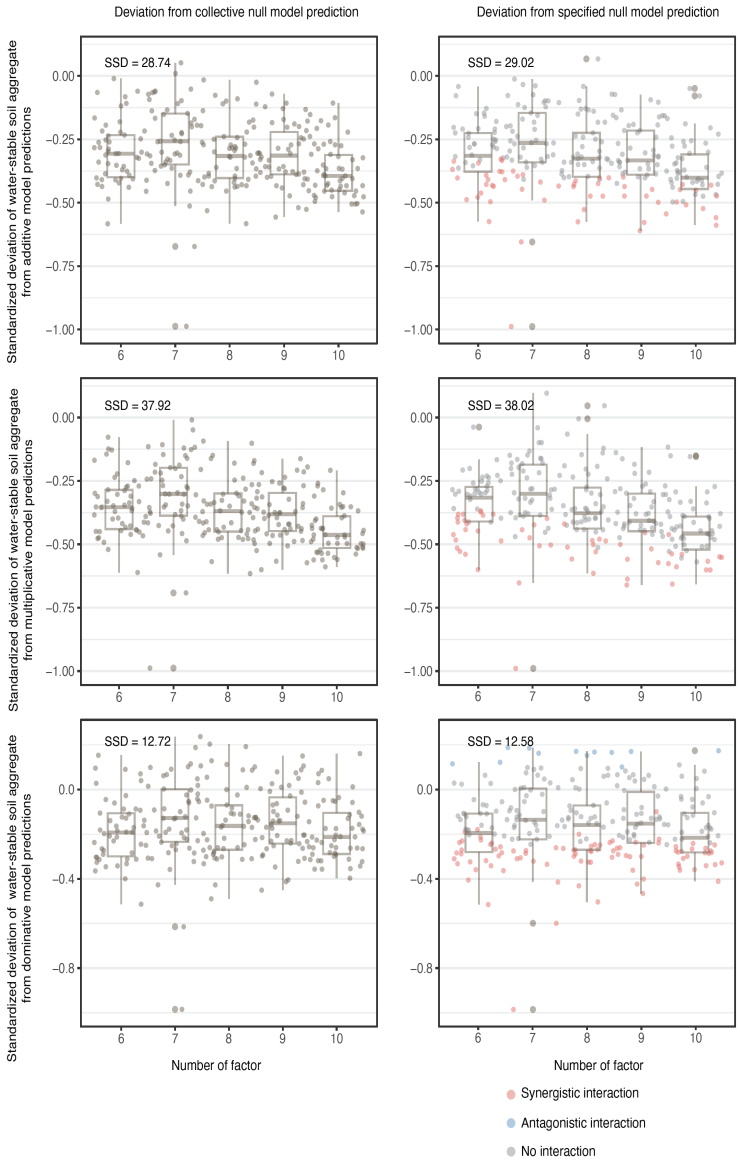

FIGURE 3.11: **Rescaled deviation of water-stable soil aggregate from three null model predictions.**

Scatter points represent the standardized deviation of WSA responses from collective null model predictions (on the left panel) and from specified null model predictions (on the right panel) for multiple-factor treatments. For the deviations form specifed null model prediction, the net interaction type of each multiple-factor treatment is marked as different colored points (antagonistic, blue; synergistic, red; no interaction, gray). The smallest model sum of squared deviation (SSD) is shown for each figure.

3.6 Discussion

The Null model workflow provided here can be applied to easily estimate the combined effects of factors, based on three different null model assumptions (additive, multiplicative and dominative). For replicated factor combinations, this workflow

is very practical for generating null model predictions by specifying co-acting factors, and it can also indicate the statistical inference for factor net interactions. For randomly-drawn factor designs in multiple factor studies, although no statistical inference for factor interaction can be drawn from each multi-factor treatment, the strength of this approach is that it can be used for batch processing a large number of different factor combinations. By analysing the deviation of observed effects from the null model predictions, it can reveal patterns in how co-acting factor interactions contribute to the variations of ecosystem functioning. Compared to the collective null model prediction, the benefit of estimating the null model predictions by specified factor identity is especially important when there are extreme single factor effects.

Although this workflow has advantages in estimating combined factor effects and indicating interactions among multiple factors, it does not fully address the common limitations associated with the use of null models in ecological studies. Previous studies have emphasized that studies incorporating null models often focusing on identifying statistically significant interactions rather than providing mechanistic insights (Schäfer and Piggott; 2018; Orr et al.; 2020). Further research should focus on selecting appropriate null models that are underpinned by a mechanistic understanding of single and multiple factor effects.

3.7 Conclusion

In conclusion, this study presents a practical and flexible null model analysis workflow designed to evaluate the combined effects of multiple global change factors on soil ecosystems. By using three distinct null model assumptions (additive, multiplicative and dominative), the workflow efficiently generates predictions for both replicated and randomly selected factors. This approach proves particularly useful in assessing the contribution of factor interactions to ecosystem variability, providing critical insights into the role of co-acting factors in shaping soil functions. The ability to batch-process large numbers of factor combinations and incorporate specific factor identities further enhances the utility of this method, particularly in cases

where extreme single-factor effects dominate. Overall, the workflow offers a robust framework for future GCF studies, facilitating the exploration of high-order interactions and advancing our understanding of complex ecological responses to anthropogenic stressors.

Bibliography

Díaz, S. et al. (2019). Pervasive human-driven decline of life on Earth points to the need for transformative change, *Science* **366**(6471): eaax3100.

Ho, J. et al. (2019). Moving beyond p values: Everyday data analysis with estimation plots.

Katzir, I. et al. (2019). Prediction of ultra-high-order antibiotic combinations based on pairwise interactions, *PLOS Computational Biology* **15**(1): e1006774.

Keuskamp, J. A. et al. (2013). Tea Bag Index: A novel approach to collect uniform decomposition data across ecosystems, *Methods in Ecology and Evolution* **4**(11): 1070–1075.

Klute, A. (1986). *Methods of Soil Analysis. Part 1 : Physical and Mineralogical Methods*, 2nd ed edn, American Society of Agronomy : Soil Science Society of America Madison, Wis., Madison, Wis.

Larsen, K. S. et al. (2011). Reduced N cycling in response to elevated CO2, warming, and drought in a Danish heathland: Synthesizing results of the CLIMAITE project after two years of treatments: EFFECTS OF CLIMATE CHANGE ON N CYCLING, *Global Change Biology* **17**(5): 1884–1899.

Orr, J. A. et al. (2020). Towards a unified study of multiple stressors: Divisions and common goals across research disciplines, *Proceedings of the Royal Society B: Biological Sciences* **287**(1926): 20200421.

R Core Team (2021). R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing.

Richardson, K. et al. (2023). Earth beyond six of nine planetary boundaries, *Science Advances* **9**(37): eadh2458.

Rillig, M. C. et al. (2019). The role of multiple global change factors in driving soil functions and microbial biodiversity, *Science* **366**(6467): 886–890.

Rillig, M. C. et al. (2021). Classifying human influences on terrestrial ecosystems, *Global Change Biology* **27**(11): 2273–2278.

Schäfer, R. B. and Piggott, J. J. (2018). Advancing understanding and prediction in multiple stressor research through a mechanistic basis for null models, *Global Change Biology* **24**(5): 1817–1826.

Schäfer, R. B. et al. (2023). Chemical Mixtures and Multiple Stressors: Same but Different?, *Environmental Toxicology and Chemistry* **42**(9): 1915–1936.

Speißer, B. et al. (2022). Number of simultaneously acting global change factors affects composition, diversity and productivity of grassland plant communities, *Nature Communications* **13**(1): 7811.

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*, Springer-Verlag New York.
URL: *https://ggplot2.tidyverse.org*

Chapter 4

How artificial intelligence models affect the environment and the science of ecology

4.1 Abstract

The rapid development of artificial intelligence (AI), particularly generative models like large language models (LLMs) and generative AI, is transforming many fields, including environmental sciences and ecology. Since the release of OpenAI's ChatGPT in November 2022, the integration of these technologies into research, communication, and education has grown exponentially. While LLMs and generative AI present opportunities for streamlining research workflows, improving environmental communication, and enhancing public engagement, they also pose substantial risks. These include potential misinformation, biased outputs, energy consumption, and even the fabrication of scientific data. This chapter examines the implications of LLMs and generative AI on environmental sciences, focusing on the potential environmental impacts, the spread of misinformation, and the benefits of enhanced communication. It also discusses how AI tools could revolutionize ecological research and education while highlighting the urgent need for policies to regulate AI use and mitigate risks, advocating for responsible AI integration to safeguard the integrity of environmental science.

4.2 Introduction

These are turbulent times in the field of artificial intelligence (AI), with a flurry of new AI tools for productivity, research and creation becoming available to the general public on a daily basis, in part accelerating since the release of the chatbot ChatGPT by OpenAI in November 2022. Amid the tech enthusiasm, there are also voices that warn about the potential dangers of the unchecked development and deployment of advanced generative AI models. In March 2023, for example, there was an Open Letter (the motivation behind which is still subject to debate) calling for a temporary pause on the development of models more powerful than GPT-4 (Future of Life Institute.; 2023).

Beyond the potential negative socio-political implications of generative models(Dai et al.; 2021; Weidinger et al.; 2022) and legal concerns such as copyright infringement(Samuelson; 2023), the implications for environmental sciences and ecology are significant. Generative AI involves the use of machine learning approaches to generate new content (e.g., text, images, audio, or video) based on characteristics of training data and user input. The impacts and challenges of the use of large language models on scientific practice(Birhane et al.; 2023), and on environmental sciences and ecology has already been explored(Rillig et al.; 2023), where advantages include aspects of streamlining environmental scientists' workflows or enhancing teaching materials, and challenges related to producing fraudulent texts with an air of simulated authority, among other points. However, an emerging trend that is currently unfolding, and that requires in-depth assessment, is the rise and refinement of text-to-image generative models and their multimodal counterparts. These generative models can transform textual prompts into detailed images or videos(Zhang et al.; 2023a), virtually indistinguishable from actual photos or other original work, including images of environmental content. Such capabilities, accessible without programming expertise, could affect all visual-related aspects of ecological research and environmental advocacy. Such AI models are widely available (and often for free), and include apps such as Dall-E-2, Stable Diffusion, Midjourney, Leondardo.ai, Sora and many others.

Given the prevalent belief that 'a picture is worth a thousand words'— highlighting the ability of images to transcend linguistic barriers and convey emotions swiftly — they raise concerns when produced by advanced models. What then are the implications and opportunities of text-to-image generative models or their multimodal counterparts for environmental and ecological research? We would like to emphasize that the opportunities and risks we discuss below would more generally also apply to other fields of science. However, for clarity, our specific focus will be on the environment and the science of ecology.

4.3 Risks and Benefits of Large Language Models for the Environment

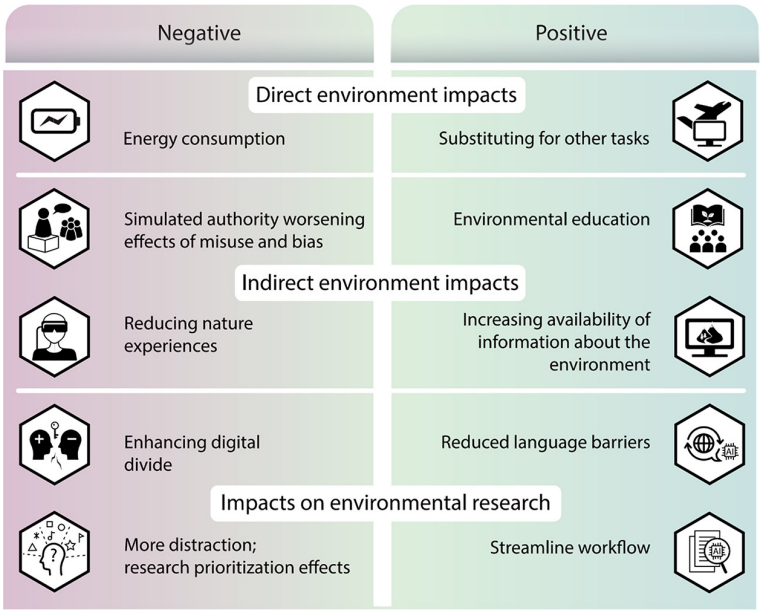

FIGURE 4.1: **Large language models come with risks and opportunities for the environment.**

Increased use of large language models could affect the environment positively or negatively, with possible direct and indirect effects on the environment and on the way environmental research is conducted.

Large language models (LLMs) are artificial intelligence (AI) models with complex architecture and a large number of parameters that have been trained on very large amounts of text (billions of words). These models, arising from the field of natural language processing, can generate natural, human-like writing and have

been designed to react to user input, enabling conversations and customized output according to prompts. The release of the chatbot ChatGPT by OpenAI (OpenAI; 2023) in late 2022 has rapidly spread this technology to a wide range of users (GPT means generative pretrained transformer and denotes the type of large language model used in this chatbot). Other companies are offering their own apps featuring different kinds of LLMs, and this technology is also being rapidly integrated into existing apps and online tools. Because LLMs will likely become extremely common, the potentially transformative nature of these models has already sparked a lively debate about their use. This discussion focuses on academic integrity and the future of research and teaching, the meaning of authorship, potential consequences for the general workforce, and unresolved copyright issues (Dis et al.; 2023). However, the debate has so far largely missed the potential implications of current and future LLM tools for the environment (Bender et al.; 2021b). We see the possibility of direct and indirect environmental impacts and effects, and opportunities and risks for researchers in the environmental sciences (FIGURE 4.1).

The first point to consider is the positive or negative direct environmental impact (Figure 1). Other potentially transformative technological innovations, such as the metaverse (Rillig et al.; 2022), likely will have direct consequences on the environment via increased energy use and thus resource consumption and production of carbon dioxide. This clearly is also a concern for LLMs (Strubell et al.; 2019), with both the training of LLMs and inference having large energy demands, prompting an early call for algorithmic efficiency (Bender et al.; 2021b). The carbon footprint will depend on the energy use and the carbon intensity of the energy source being used. In addition to carbon dioxide emissions, the computing facilities may also have other environmental impacts such as water use and soil pollution or sealing, which could have broader implications for environmental quality. Conversely, can the use of text-based chats in the future partially replace video conferences or travel to in-person meetings that might consume more resources by comparison? It is unclear if this will be the case, given that human verification and expertise will likely remain indispensable (Dis et al.; 2023).

There are likely also indirect consequences of increased LLM use, which are potentially more important (FIGURE 4.1). The first issue is the level of artificial expertise. LLM output comes with a certain degree of simulated authority, given the extensive amount of information with which LLMs have been trained and the polished language in which output is written. Therefore, such output can easily be confused with expert opinions, even though LLMs will continue to have limited ability to judge the reliability and relevance of information, in part because LLMs have not achieved natural language understanding. Thus, false output is created, as anybody who has played with these apps on topics of their own expertise will have noticed. More worrying is the potential to inadvertently or purposefully introduce bias at three points: the training data (the input to the model), the algorithm (how sources are used), and the form of output (e.g., disclaimers, statements of uncertainty, and references). At each of these points, special interest groups and networks could exploit the ability of LLMs to generate text with unprecedented efficiency, thus offering misinformation under the guise of "artificial intelligence", and flooding public spaces with it. We think this is the biggest concern of the more widespread use of LLMs for environmentally relevant topics. But even without ill intent, existing biases on complex environmental topics, including environmental racism (Bender et al.; 2021b), climate change and global environmental change, biodiversity loss, or pollution, could be perpetuated and multiplied by the training data the LLM uses. On the contrary, writing informative content about environmental issues by actors interested in environmental education could be made more efficient through LLMs. For example, materials for use in environmental education could be more easily adapted for different target groups, such as different ages or educational levels.

Differential access to LLM-based apps could worsen or improve digital divide effects within and among societies. The availability of these tools could further favor people with already excellent access to environmental information. On the positive side, LLMs could broaden the participation of people in the environmental debate, especially because the LLM services are being offered in various languages.

It is also possible that increasingly relying on technology-guided interactions

could contribute to reduced nature experiences (Soga and Gaston; 2016), thus reducing value attributed to the environment. More screen time spent on fascinating tools could translate to less time spent in nature, with potential consequences for how people will appreciate biodiversity and ecosystems. Conversely, the public may also benefit from unprecedented, up-to-date, accessible, and personalized information and educational opportunities on environmental topics. This increasingly available "information at your fingertips" could lead to greater curiosity about environmental topics, thus enhancing environmental literacy.

There are certainly many benefits of LLMs for research in the environmental sciences. The efficient use of these tools will likely streamline the workflow of environmental scientists and potentially improve the quality of writing, thus freeing up researchers' time for other tasks, i.e., designing and analyzing experiments and developing innovative ideas. By offering a tool to sharpen their English language scientific writing, LLMs might increase representation of researchers from non-English speaking countries in the environmental sciences. This all could potentially accelerate scientific progress on important environmental topics. Conversely, these tools might prove to be a distraction, and researcher time may increasingly be consumed by dealing with adverse effects of LLM use, for example, in university education and training (e.g., misuse and cheating) (Dis et al.; 2023). If LLMs are increasingly used to summarize information and to support funding decisions in the future, this could render human expertise and insight less important, potentially leading to adverse effects on research prioritization in the environmental sciences.

Clearly, there are benefits and risks for the environment in the increased use of LLMs, and it is important to start this discussion early to avert possible harm. To harness the opportunities, we ought to protect such powerful tools from the undue influence of individual groups. To ensure that LLMs are constructed to yield unbiased information, governments, intergovernmental bodies, and large organizations (such as OECD) need to devise policy tools that could include oversight committees, guidelines, or regulations that lead to disclosure of data sources (for training of models) and funding. Another important goal is to increase literacy in the use of these tools, and this is where environmental scientists can act now, by including LLMs in their university courses and by clarifying in their laboratories how LLMs

can and should be used in the environmental sciences.

4.4 How widespread use of generative AI for images and video can affect the environment and the science of ecology

4.4.1 Opportunities

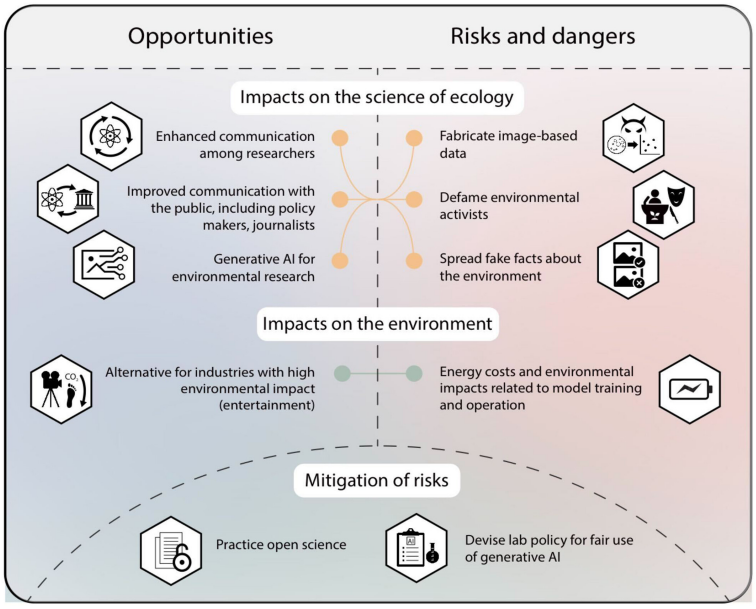

FIGURE 4.2: **Opportunities and risks associated with the use of generative AI for images and video in the field of ecology/environmental science, and effects on the environment itself, as well as possible routes for risk mitigation that are immediately available to researchers and their labs.**

Exploring the opportunities for environmental and ecological research, many are related to enhancing our ability to communicate (FIGURE 4.2). First, ecologists use images and videos to communicate with other scientists. For example, they develop conceptual models that come to life in the form of visual representations. These depictions are very important, because they help other researchers understand complex relationships. Perhaps in the near future, the ecologist's workflow could be streamlined by creating such visuals using text prompts, saving time (and money) for other work. This way, ecologists could communicate their thoughts and ideas about ecological topics more effectively. Certainly, all of biology deals with complex topics, but in ecology visuals could for example be created for future scenarios

of climate change, to depict concepts in theory-heavy fields such as population or community ecology, or to help explain experimental designs using tailor-made diagrams.

Second, ecologists also use images for communicating with the public, for example, policy makers or journalists, and in environmental education (Schäfer; 2023). Environmental scientists could likely use AI-generated images (or videos) to help them better explain the often complex environmental issues (such as pollution) and intricate ecological relationships, such as in biodiversity research. This could be done more cheaply, and more effectively, thus potentially reaching more people with visually appealing content. One powerful application could be the creation of visual what-if scenarios, illustrating risks for the environment (Luccioni et al.; 2021). Imagine creating an image or video of your city in the year 2050 given the presence of certain global change factors, including climate change, invasive species, pollution, land use change and other drivers. Perhaps reporting on ecological research findings overall could benefit for exactly the same reasons, including reporting done by journalists or institutional press offices. Given the command line-based image creation process, images and movies can be easily adapted to different target groups with different backgrounds or different scopes. However, we should also be aware of and prevent potential negative impacts on scientific communication professionals (illustrators, artists, storytellers); in the near future, generative AI is unlikely to surpass their specific skill and professional input, which is essential for excellent communication.

Direct benefits for the environment could occur by replacing energy-intensive operations with AI-generated video. For example, traditional filmmaking has a high carbon footprint and could be partially supplemented or replaced with AI-generated content (careful carbon budgets would need to verify this assertion). This could make certain industries, such as entertainment, more environmentally sustainable.

Could generative AI for images also be useful in the scientific process of ecology? There are several opportunities that can be explored (TABLE 4.1): we identify image data augmentation and gap filling (Kebaili et al.; 2023), biodiversity monitoring and enhanced citizen science as potential areas where AI can be employed to enhance research outcomes.

TABLE 4.1: **Ideas for how generative AI for images can be useful in ecological research; three examples are discussed: data augmentation and gap filling, biodiversity monitoring and enhanced citizen science.**

Research opportunity	Explanation	Example
Data augmentation and gap filling	In environmental research, it is common to encounter data gaps due to inaccessible regions, seasonal changes, or instrument failures. Generative AI models can potentially predict and fill these gaps, providing a more complete dataset for researchers	In remote sensing, where consistent satellite imagery is important, AI could generate images for days where data might be missing due to cloud cover or technical glitches. This continuous data stream can be invaluable for real-time monitoring of environmental changes
Biodiversity monitoring	Generative AI can assist in creating a library of images representing various species in multiple poses (for animals) or perspectives (for plants and macrofungi) or environments	This can help in training more robust identification models, vital for biodiversity monitoring. It might help practitioners with species identification
Enhanced citizen science	Engaging the public in data collection has seen a rise with citizen science initiatives. Generative AI can assist by providing participants with visual aids, helping them correctly identify and report findings	Explanatory materials can be of key importance in citizen science, and illustrations could be created more easily with generative AI

4.4.2 Risks and dangers

There are direct environmental risks that are similar to those of text-based AI, in that AI-generated image and video will consume computing-related resources, leading to effects on energy and resource consumption and carbon emissions in addition to those of text-based models (Bender et al.; 2021a; Bird et al.; 2023; Rillig et al.; 2023), since they also make use of large language models as part of their model architecture. Energy needs arise during model training, during inference (that is, queries made to the model), and also for server infrastructure, and corresponding carbon emissions depend on the carbon intensity of the energy source. One estimate per query is 0.5 g CO_2 for a large language model (Vanderbauwhede; 2024). Luccioni et al. (Luccioni et al.; 2022) estimate the carbon cost of BLOOM (a 176-billion parameter language model) to be 50.5 t CO_2 over its life cycle.

However, we contend that indirect effects on the environment and the science of ecology are likely going to be of greatest concern (FIGURE 4.2). First and foremost, we see risks associated with spreading misinformation via deep fakes (Zhang et al.; 2023b). Imagine central figures, environmental activists, representatives of environmental NGOs, politicians or ecology experts, for example, being depicted in compromising settings that cast doubt on their integrity, defaming them; thus undermining

the validity of the message they send. In addition, prominent communicators of science, or simply well-recognized people in public life, could be instrumentalized in spreading fake facts about the environment, and it will be very difficult to get these under control once they have been spread via social media. For example, imagine a politician, well-respected TV personalities and conservationists, or the president of an ecological society in a deepfaked video declaring that all problems of environmental pollution and climate change are well under control or that there is no biodiversity crisis to worry about. The damage to the credibility of the science of ecology caused by bad actors (governments, organizations, companies, individuals) generating such materials would be immense.

A threat to the integrity of ecology from within looms large as well. Image or video generation opens a door to rapidly and easily fabricating data; this is an already existing problem in science (Bik et al.; 2016) but could be exacerbated immensely by the availability of advanced text-to-image tools (Gu et al.; 2022). Perhaps images can also be manipulated to a point that it will be extremely difficult (or impossible) to detect. This will affect all areas of environmental science in which visual evidence plays a role: think of gel images, microscopic evidence, and really anything that can be used as supporting visual evidence in a scientific study on an environmental or ecological topic. Such deepfaked visual scientific data produced by bad actors would not only be a nuisance but might also undermine some of the trust in environmental science and ecology overall.

4.4.3 Risk mitigation

How could we effectively counteract some of these effects, for example protect reporting on ecological topics from fabricated, false information? How could we safeguard the integrity of our science? We will likely have an arms race between image generating AI tools and those that can detect their products; but improved detection will only solve part of the problem, because once images are posted on social media, effects will be very difficult to control. Banning these AI-tools will not be a realistic option as they are already widely adopted, and it is unclear how effective regulations would be when certain governments step in to regulate these products. We here propose local solutions, at the lab or institutional level, that can be rapidly

implemented; however, we stress that in the long-term regulatory action on both the national and international level is needed.

To counter deepfaked visual information, we envision transitioning to radically open science as one potential option to safeguard ecological research and reporting. In the long term, multi-platform science communication in ecology has the potential to build public engagement and trust (Pavelle and Wilkinson; 2020) and might be less prone to fabricated visuals. Since such social media open science activities go beyond the open science policies already implemented in funding agencies (e.g. the European Research Council), these would need to be included in future funding schemes, providing adequate support to ecologists. Transparent reporting of the responsible use of AI-generated content in scientific work including detailed (and ideally machine-actionable) provenance information must be part of such an open science approach (Wahle et al.; 2023). This would also safeguard gap-filling techniques (Table 1) from misuse via fabricated data.

Responsible use of generative AI also includes taking steps to protect others' careers with connection to environmental science, such as those of science communication experts. This includes carefully weighing the options when working on communication-related tasks, for example by asking how visualization could profit from professional input.

Another important step that can be taken by institutions or individual research groups is the development of policies on the fair use of generative AI (see here (Rillig; 2023) for an example). Such policies provide a starting point for larger-scale agreements in environmental science and they offer an opportunity to train scientists, early career researchers and more experienced colleagues, in these issues and involve them in the discussion during this critical time of transformation.

Bibliography

Bender, E. M. et al. (2021a). *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, FAccT '21, Association for Computing Machinery, New York, NY, USA, p. 610–623.

URL: *https://doi.org/10.1145/3442188.3445922*

Bender, E. M. et al. (2021b). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?, *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, ACM, Virtual Event Canada, pp. 610–623.

Bik, E. M., Casadevall, A. and Fang, F. C. (2016). The Prevalence of Inappropriate Image Duplication in Biomedical Research Publications, *mBio* **7**(3): e00809–16.

Bird, C., Ungless, E. and Kasirzadeh, A. (2023). Typology of risks of generative text-to-image models, *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*, AIES '23, Association for Computing Machinery, New York, NY, USA, p. 396–410.
URL: *https://doi.org/10.1145/3600211.3604722*

Birhane, A., Kasirzadeh, A., Leslie, D. and Wachter, S. (2023). Science in the age of large language models, *Nature Reviews Physics* **5**(5): 277–280.

Dai, J., Fazelpour, S. and Lipton, Z. (eds) (2021). *Proceedings of the 2021 Aaai/Acm Conference on Ai, Ethics, and Society*.

Dis, V. et al. (2023). ChatGPT: Five priorities for research, *Nature* **614**(7947): 224–226.

Future of Life Institute. (2023). Pause giant ai experiments: an open letter.
URL: *https://futureoflife.org/open-letter/pause-giant-ai-experiments/*

Gu, J. et al. (2022). AI-enabled image fraud in scientific publications, *Patterns* **3**(7): 100511.

Kebaili, A., Lapuyade-Lahorgue, J. and Ruan, S. (2023). Deep Learning Approaches for Data Augmentation in Medical Imaging: A Review, *Journal of Imaging* **9**(4): 81.

Luccioni, A. S., Viguier, S. and Ligozat, A.-L. (2022). Estimating the carbon footprint of bloom, a 176b parameter language model.
URL: *https://arxiv.org/abs/2211.02001*

Luccioni, A. et al. (2021). Using Artificial Intelligence to Visualize the Impacts of Climate Change, *IEEE Computer Graphics and Applications* **41**(1): 8–14.

OpenAI (2023). Chatgpt: Optimizing lnaguage models for dialogue, openai.
URL: *https://openai.com/index/chatgpt/*

Pavelle, S. and Wilkinson, C. (2020). Into the Digital Wild: Utilizing Twitter, Instagram, YouTube, and Facebook for Effective Science and Environmental Communication, *Frontiers in Communication* **5**: 575122.

Rillig, M. (2023). Lab policy about the use of generative AI tools.

Rillig, M. C. et al. (2022). Opportunities and Risks of the "Metaverse" For Biodiversity and the Environment, *Environmental Science & Technology* **56**(8): 4721–4723.

Rillig, M. C. et al. (2023). Risks and Benefits of Large Language Models for the Environment, *Environmental Science & Technology* **57**(9): 3464–3466.

Samuelson, P. (2023). Generative AI meets copyright, *Science* **381**(6654): 158–161.

Schäfer, M. S. (2023). The Notorious GPT: Science communication in the age of artificial intelligence.

Soga, M. and Gaston, K. J. (2016). Extinction of experience: The loss of human–nature interactions, *Frontiers in Ecology and the Environment* **14**(2): 94–101.

Strubell, E., Ganesh, A. and McCallum, A. (2019). Energy and Policy Considerations for Deep Learning in NLP, *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, Association for Computational Linguistics, Florence, Italy, pp. 3645–3650.

Vanderbauwhede, W. (2024). The climate cost of the ai revolution. ripe labs.
URL: *https://labs.ripe.net/author/wim-vanderbauwhede/the-climate-cost-of-the-ai-revolution/*

Wahle, J. P. et al. (2023). Ai usage cards: Responsibly reporting ai-generated content, *2023 ACM/IEEE Joint Conference on Digital Libraries (JCDL)*, pp. 282–284.

Weidinger et al. (2022). Taxonomy of risks posed by language models, *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, Association for Computing Machinery, p. 214–229.

Zhang, C., Zhang, C., Zhang, M. and Kweon, I. S. (2023a). Text-to-image Diffusion Models in Generative AI: A Survey.

Zhang, C., Zhang, C., Zhang, M. and Kweon, I. S. (2023b). Text-to-image diffusion models in generative ai: A survey.

URL: *https://arxiv.org/abs/2303.07909*

Chapter 5

General discussion

We are living in a turbulent era where an increasing number of anthropogenic stressors are threatening Earth's ecosystems, while technologies like artificial intelligence are rapidly advancing. Global environmental change involves the simultaneous action of multiple factors, introducing substantial uncertainty in predicting their combined effects. Soil ecosystems, which are subject to a range of these interacting factors, are particularly vulnerable; however, the outcomes of these factors acting in concert are not well understood. In this PhD thesis, I investigated the potential drivers of multiple global change factor influencing soil properties and functions, focusing on the impact of factor number and dissimilarity. Additionally, I explored how machine learning-based methods can be harnessed to analyze complex datasets from multiple GCFs studies and also examined the impacts of development of AI tools on the environment and the field of ecological science.

Chapter 2 underscores the significance of both the number of factors and the dissimilarity among global change factors in shaping soil responses. By analyzing a wide range of factor combinations, we demonstrate that higher factor dissimilarity tends to result in more synergistic interactions, leading to greater deviations from null model predictions and amplifying the effects on key soil functions, such as soil decomposition rate, water-stable soil aggregate and enzymatic activities. Our findings suggest that factor dissimilarity plays a crucial role in driving these interactions. This highlights the importance of considering not only the number of co-acting factors but also their distinct properties and modes of actions in studies in multiple GCFs.

Additionally, the study introduces practical approaches for disentangling the

effects of factor number, dissimilarity, and factor identity contribution, helping to avoid misinterpretation of results in future research. By accounting for factor dissimilarity, researchers can better predict how multiple factors interact to impact soil properties and ecosystem functions. We also discussed three mechanisms that may explain how factor dissimilarity drives the impacts of multiple GCFs on soil ecosystems: (i) direct interactions of physicochemically dissimilar factors: Factors with distinct physicochemical properties are more likely to interact directly, amplifying the intensity of individual factors, as seen with drought amplifying chemical factors or surfactants enhancing pollutant mobility; (ii) impact on species adaptation: Factor dissimilarity increases performance trade-offs, as species cannot simultaneously optimize adaptation to multiple distinct stressors, resulting in lower overall adaptation performance compared to environments with more similar stressors; (iii) reshaping co-tolerance spaces: Dissimilar factors can negatively affect the co-tolerance of species to multiple stressors, leading to higher species loss and diminished ecosystem functions when factors are highly dissimilar. Moving forward, incorporating factor dissimilarity into multiple GCF studies could enhance our understanding of complex ecological interactions and support more informed decision-making for ecosystem restoration and management.

In chapter 3, we introduce a practical null model analysis workflow for evaluating the combined effects of multiple global change factors on soil ecosytems. By employing three distinct null model assumptions - additive, multiplicative and dominative - the workflow efficiently generates predictions for both replicated and randomly selected factor combinations. This approach is particularly effective in evaluating the contribution of factor interaction to soil response variability, offering insights into the influence of co-acting factors on soil functions. Its ability to batch-process a large number of factor combinations and include specific factor identities further increases its applicability, especially in scenarios where extreme single-factor effects are prevalent. However, this appraach also does not fully overcome the inherent limitations of null models in ecological research. The chapter emphasizes the need for future studies to focus on selecting null models grounded in mechanistic

understanding. Despite these limitations, the workflow provides a robust framework for advancing GCF studies, facilitating the exploration of multi-factor interactions and deepening our understanding of ecological responses to anthropogenic stressors.

Chapter 4 explores the rapid advancements in artificial intelligence (AI), specifically focusing on generative models like large language models (LLMs) and their impact on environmental sciences and ecology. AI technologies have been increasingly integrated into research, communication, and education. While LLMs offer significant opportunities, such as streamlining research workflows, enhancing environmental communication, and increasing public engagement, they also pose considerable risks. These risks include the potential spread of misinformation, biased outputs, high energy consumption and even the fabrication of scientific data.

To harness the benefits while mitigating the risks, chapter 4 underscores the need for early and responsible use of AI in environmental science. It advocates for the development of policies and guidelines to ensure that LLMs provide unbiased information, including oversight by governments and organizations.

Chapter 6

Conclusion

This dissertation addresses the complex challenges posed by the increasing number of anthropogenic stressors and rapid technological advancements, particularly in the context of global environmental change and soil ecosystems. The simultaneous action of multiple global change factors (GCFs) introduces considerable uncertainty in predicting their combined effects, with soil ecosystems being especially vulnerable. By focusing on the impact of factor number and dissimilarity, as well as leveraging machine learning methods, this research provides new insights into the mechanisms driving soil responses to GCFs and examines the broader implications of artificial intelligence (AI) in ecological science. Overall, it advances our understanding of how multiple GCFs interact to affect soil ecosystems and underscores the importance of incorporating factor dissimilarity into future studies. It also provides a robust framework for evaluating complex ecological interactions and emphasizes the need for responsible integration of AI technologies in environmental research. By addressing both the ecological and technological dimensions of global change, this work contributes to more informed decision-making in ecosystem management and offers a path forward for sustainable ecological practices in an era of rapid change.

List of publications and contributions

I Bi, M., Li, H., Meidl, P., Zhu, Y., Ryo, M., Rillig, M.C., 2024. Number and dissimilarity of global change factors influences soil properties and functions. Nature Communications 15(1). https://doi.org/10.1038/s41467-024-52511-2.
Own contribution: I participated in experimental design, conducting the experiment; I led the data analysis pipeline design, carried out the data analysis and co-led the writing of the manuscript.

II Cruz, J., Lammel, D., Kim, S.W., **Bi, M.**, Rillig, M.C., 2024. COVID-19 pandemic-related drugs and microplastic from mask fibers jointly affect soil functions and processes. Environmental Science and Pollution Research 31, 50630-50641. https://doi.org/10.1007/s11356-024-34587-x.
Own contribution: I participated in the data analyis and helped with revision of the manuscript.

III Rillig, M.C., Ågerstrand, M., **Bi, M.**, Gould, K.A., Sauerland, U., 2023. Risks and benefits of large language models for the environment. Environmental Science & Technology 57(9). DOI: 10.1021/acs.est.3c01106
Own contribution: I designed the figure for this paper and also helped with framing the paper, helped with writing the paper, and revised the final version.

IV Rillig, M.C., Mansour, I., Hempel, S., **Bi, M.**, König-Ries, B., Kasirzadeh, A., 2024. How widespread use of generative AI for images and video can affect the environment and the science of ecology. Ecology Latters 27(3). DOI: 10.1111/ele.14397

Own contribution: I contributed to the writing of the paper, specifically by contributing concepts and ideas. I designed the figure and revised the final version of the manuscript.

Appendix for chapter 2

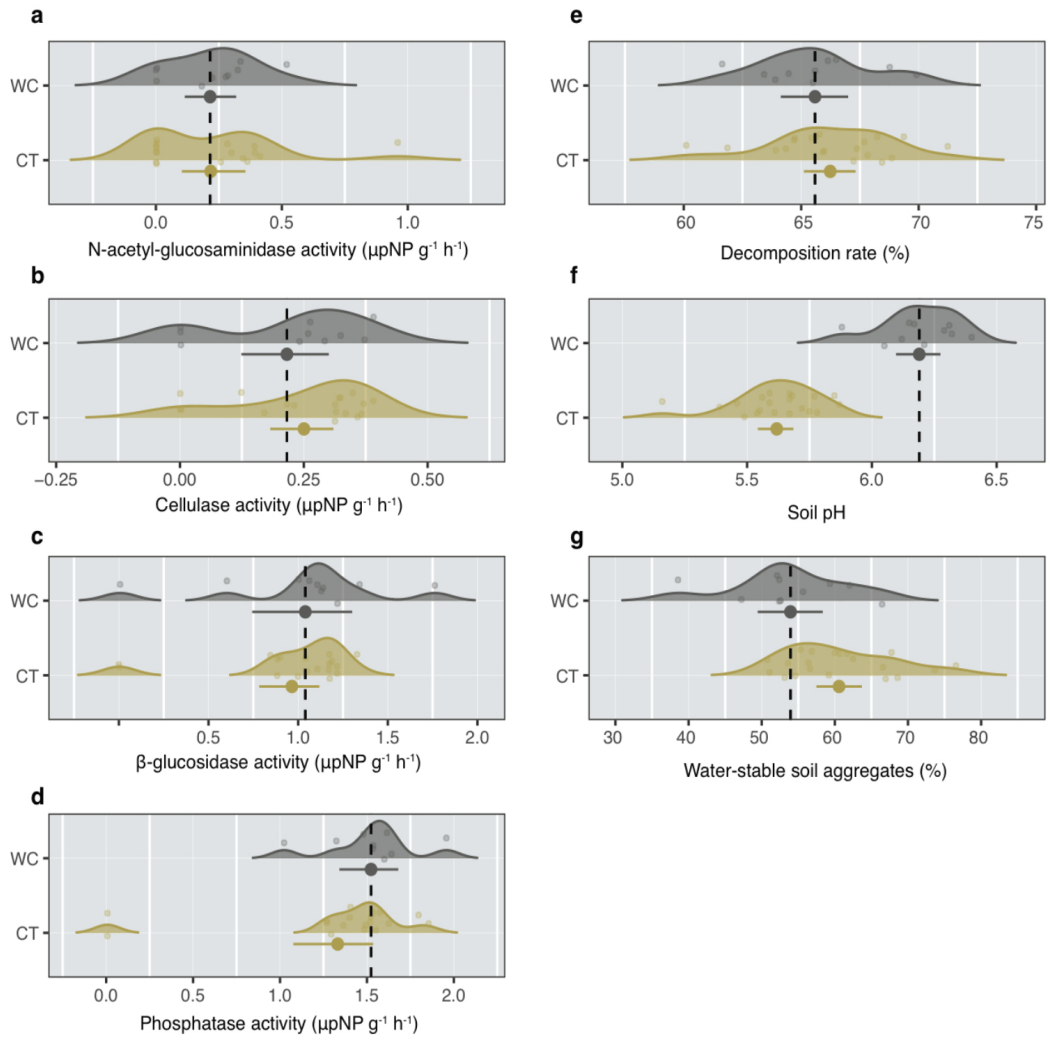

FIGURE 1: **Effects of organic solvents on soil properties and functions.**

Illustrated are effect sizes of organic solvents (DMSO and acetone), that are used in experimental treatments for dissolving chemical factors (herbicide and fungicide), on soil properties and functions. Tested groups are Control Treatment (CT, n=20, applied by corresponding solvents including organic solvents) and Water Control

Treatment (WC, n=10, organic solvents were replaced by the same amount of water). CT and WC treatments are incubated together with other experimental treatments in the same condition, and the soil properties and functions of those treatments are measured also at the same time with other treatments.

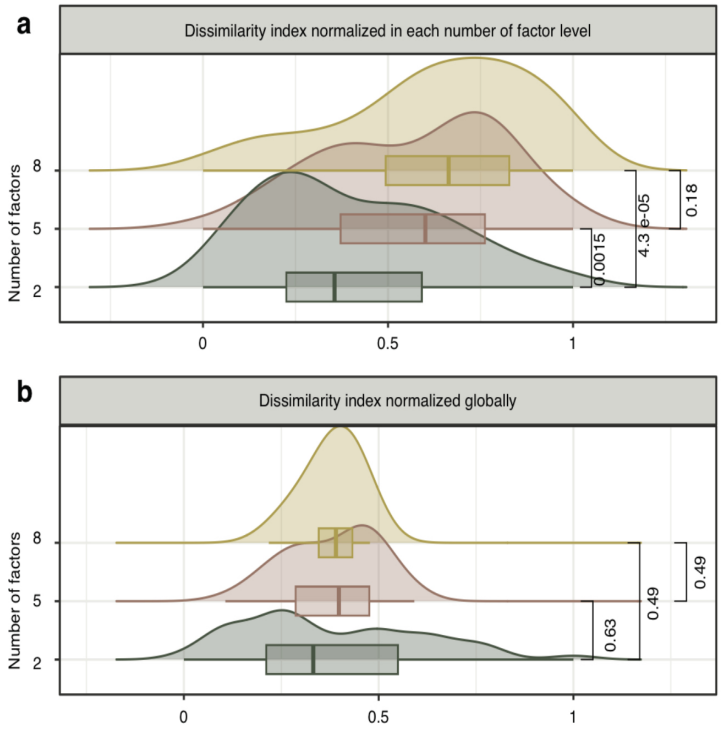

FIGURE 2: **Density distribution of factor dissimilarity indices of multiple-factor treatments in different number of factor groups normalized separately for each number of factor level (a) and normalized globally (b).**

P values obtained from pairwise t-tests are shown between every number of factor groups. In panel (a) and (b) we compared the difference of normalizing the dissimilarity index in each factor level and normalizing it globally. When dissimilarity indices are normalized in each number of factor level, their distributions are significantly different from each other ($P < 0.05$), which means the dissimilarity indices co-vary with factor levels. We think this bias partially comes from the random selection method (for example we may select more 2-factor combinations with lower dissimilarity by chance) and partially comes from the intrinsic unevenness of distances among 12 single factors. Even though we can not fully disentangle the dissimilarity

TABLE 1: **Testing hypothesis by comparing models.**

Model(s)	**Tested hypothesis**
Model 1	Variability of soil response to multiple GCFs explained by three null model predictions, which is regarded as the factor identity effect.
Model 2	Variability of soil response to multiple GCFs explained by number of factor effect.
Model 3	Variability of soil response to multiple GCFs explained by factor dissimilarity indices.
Model 4→ Model 1	Variability of soil response to multiple GCFs explained by number of factor effect on the basis of including factor identity effect.
Model 5→ Model 1	Variability of soil response to multiple GCFs explained by factor dissimilarity effect on the basis of including factor identity effect.
Model 6→ Model 4	Variability of soil response to multiple GCFs explained by factor dissimilarity effect on the basis of including factor identity effect and number of factor effect.
Model 6→ Model 5	Variability of soil response to multiple GCFs explained by number of factor effect on the basis of including factor identity effect and factor dissimilarity effect.
Model 7→ Model 6	Variability of soil response to multiple GCFs explained by factor composition information on the basis of including all hypothetical predictors and factor identity effect.

indices from co-varying with factor level, our further approaches enable us to disentangle the effects of factor dissimilarity through hierarchical modeling methods. These methods enable us to assess the unique contribution of factor dissimilarity to the model predictability by comparing to a model taking into account the number of factor effects. However, when dissimilarity indices are normalized globally, even though their mean values are not different from each other (P > 0,05), their distributions are quite different for different factor groups, for example, the range for 2-factor group is 1 but the range for 8-factor group is 0.259. These widely varying ranges will interfere with our further analysis for assessing the effect of factor dissimilarity. Due to the aforementioned reasons, we opted to utilize dissimilarity indices normalized within each factor level for our analysis.

TABLE 2: **Significance tests for importance measures of predictors of random forest models.**

Predictors	P1	P2	P3	Number of factor	Dissimilarity index
Random forest model for N-acetyl-glucosaminidase activity					
Relative importance	0.00339	-0.00042	0.02923	0.00311	0.11128
Adjusted P-values	0.214	0.499	**0.008***	0.214	<**0.001***
Random forest model for cellulase activity					
Relative importance	-0.00062	0.00033	-0.00076	0.00387	0.01403
Adjusted P-values	0.701	0.460	0.701	**0.048**	<**0.001***
Random forest model for β-glucosidase activity					
Relative importance	0.07323	0.02102	0.09116	0.02690	0.37618
Adjusted P-values	**0.002***	**0.016**	<**0.001***	**0.014**	<**0.001***
Random forest model for phosphatase activity					
Relative importance	0.01084	0.01777	0.01073	0.00675	0.04817
Adjusted P-values	**0.020**	**0.003***	**0.023**	**0.045**	<**0.001***
Random forest model for soil decomposition rate					
Relative importance	0.00005	0.00004	0.00003	0.00029	0.00102
Adjusted P-values	**0.042**	0.070	0.077	<**0.001***	<**0.001***
Random forest model for soil pH					
Relative importance	0.07558	0.03572	0.19785	0.00028	0.00264
Adjusted P-values	<**0.001***	<**0.001***	<**0.001***	0.242	<**0.005***
Random forest model for water-stable soil aggregates					
Relative importance	0.00037	0.00121	0.00425	0.00025	0.00003
Adjusted P-values	0.023	<**0.001***	<**0.001***	0.061	0.459

Adjusted P values for relative importance of model predictors calculated from a permutation-based random forest model approach. (P1 indicates the predicted response from the additive model; P2 indicates the predicted response from the multiplicative model; P3 indicates the predicted response from the dominative model; Significant differences with 0.01 < P 0.05 are shown in bold, and P 0.01 are marked by * additionally.)

TABLE 3: **Mean value and 95% confidence interval of R^2 (%) explained by seven random forest models of soil decomposition rate, soil pH and water-stable soil aggregation.**

Models	Soil decomposition rate			Soil pH			Water stable soil aggregate		
	CI.2.5%	Mean	CI.97.5%	CI.2.5%	Mean	CI.97.5%	CI.2.5%	Mean	CI.97.5%
Model 1	10.68	25.86	43.89	81.19	85.41	88.62	15.26	27.55	55.91
Model 2	3.70	11.78	23.47	21.20	31.96	42.85	<0.01	4.17	16.78
Model 3	20.00	36.48	47.00	58.51	68.51	76.34	8.44	18.51	39.24
Model 4	16.77	29.70	43.40	82.01	86.17	89.08	17.43	34.53	59.97
Model 5	33.86	49.94	60.27	81.98	85.82	89.04	16.81	29.59	63.09
Model 6	39.41	54.12	66.40	82.37	86.05	89.63	21.91	38.47	63.60
Model 7	39.41	62.07	73.79	84.11	87.81	90.75	24.91	47.25	75.41

TABLE 4: **Mean value and 95% confidence interval of R^2 (%) explained by seven random forest models of four soil enzymatic activities.**

Models	N-acetyl-glucosaminidase activity			Cellulase activity		
	CI.2.5%	Mean	CI.97.5%	CI.2.5%	Mean	CI.97.5%
Model 1	1.38	16.82	39.15	0.16	11.21	36.41
Model 2	0.82	4.38	10.58	0.91	4.89	11.85
Model 3	6.53	16.79	33.30	7.27	17.26	36.50
Model 4	8.50	21.73	41.00	6.40	18.01	39.10
Model 5	20.21	39.67	57.51	13.82	32.94	54.70
Model 6	24.21	42.10	58.43	15.94	33.65	53.70
Model 7	26.96	47.30	62.42	17.84	40.93	61.55
Models	β-glucosidase activity			Phosphatase activity		
	CI.2.5%	Mean	CI.97.5%	CI.2.5%	Mean	CI.97.5%
Model 1	29.15	45.22	62.70	10.08	19.93	37.69
Model 2	10.10	17.43	26.65	0.07	2.78	9.46
Model 3	23.17	38.46	58.08	6.55	19.83	32.83
Model 4	35.29	51.19	67.57	14.62	29.63	47.91
Model 5	39.37	55.12	73.15	20.98	37.35	53.81
Model 6	39.37	57.82	73.57	20.92	38.90	53.79
Model 7	44.83	61.03	77.76	22.33	38.68	53.81

TABLE 5: **Statistical assessment of rescaled multi-factor treatment response deviations from three null model predictions.**

	Null models					
	Additive		Multiplicative		Dominative	
Factors	Mean	SSD	Mean	SSD	Mean	SSD
Soil decomposition rate	-0.0199	1.118	-0.0206	1.124	-0.0349	1.233
Soil pH	-0.0431	0.6591	-0.0540	0.9549	0.0218	0.2759
WSA	0.6598	193.7	0.2713	24.51	0.0029	3.538
N-acetyl-glucosaminidase activity	1.844	4424	0.3381	674.7	1.195	1486
Cellulase activity	2.047	3349	1.025	788.3	0.8159	489.3
β-glucosidase activity	0.4295	117.4	0.2837	80.26	0.5941	174.5
Phosphatase activity	0.2506	43.67	0.1389	27.07	0.2477	35.56

Appendix for chapter 3

TABLE 6: **Basic information of the pharmaceutical compounds used.**

Compound	Manufacturer	Class	Working concentration (mg/kg)	
			Low	High
Remdesivir	Cayman Chemical	Antiviral	0.011	1.1
Azithromycin	Sigma-Aldrich	Macrolide antibiotic	0.014	1.4
Ivermectin	Sigma-Aldrich	Antiparasitic	0.05	5

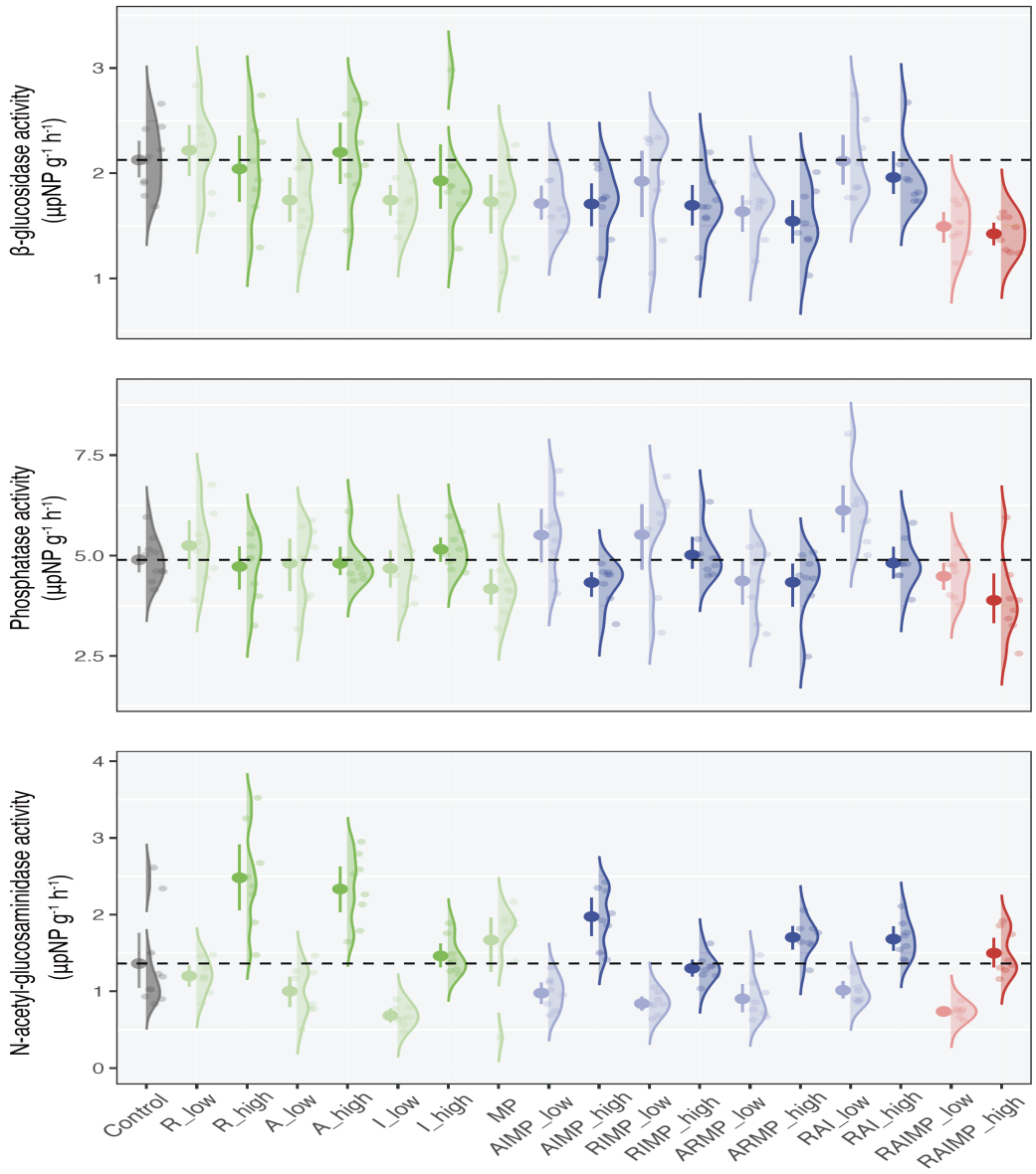

FIGURE 3: **Response of soil enzymatic activities to pharmaceutical drugs and MP treatments applied singly or simultaneously.**

For each soil enzymatic activity, effect sizes of single factor treatments (low/high-concentration of pharmaceutical drugs and MP) and multiple factor treatments (combinations of 2 and 3 pharmaceutical drugs with and without MP) were estimated by bootstrapping method with 1,000 permutations. (remdesivir, R; azithromycin, A; ivermectin, I; mask microplastic, MP)

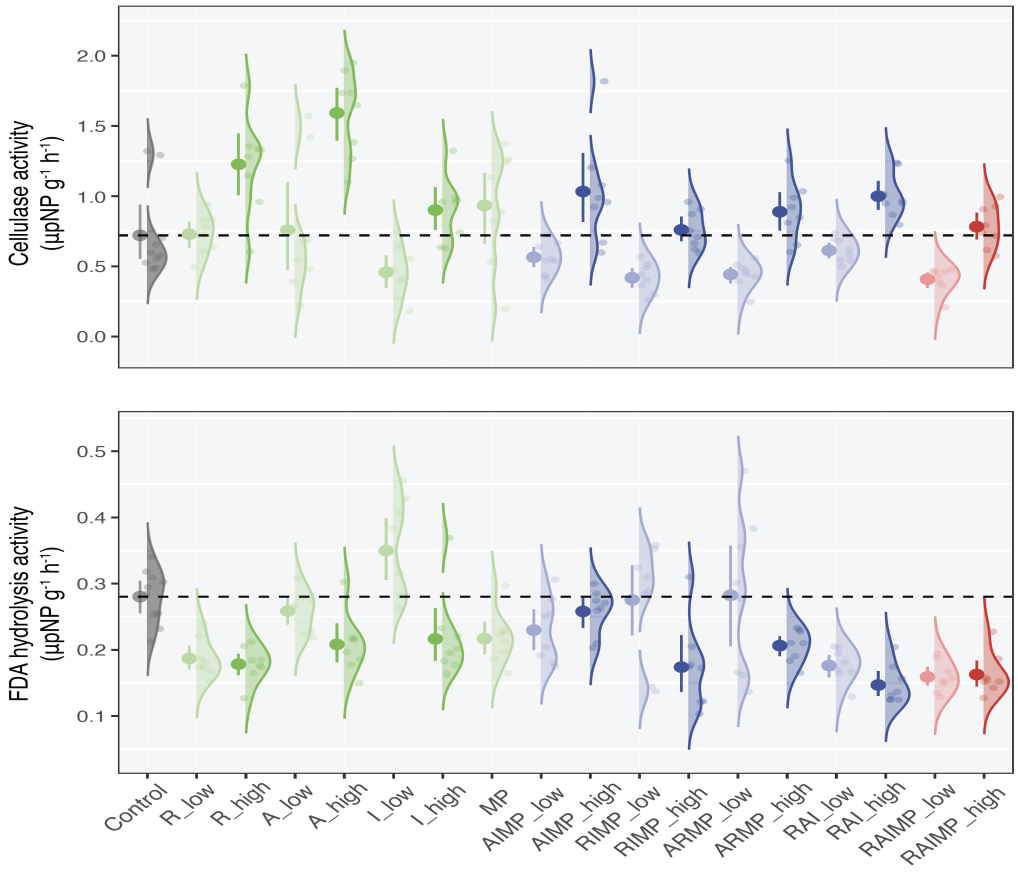

FIGURE 4: **Response of cellulase activity and FDA activity to pharmaceutical drugs and MP treatments applied singly or simultaneously.**

For each soil cellulase and FDA hydrolysis activity, effect sizes of single factor treatments (low/high-concentration of pharmaceutical drugs and MP) and multiple factor treatments (combinations of 2 and 3 pharmaceutical drugs with and without MP) were estimated by bootstrapping method with 1,000 permutations. (remdesivir, R; azithromycin, A; ivermectin, I; mask microplastic, MP)