Firm Pollution, Streams, and Biodiversity*

Seung Min Kim Sean Reese†

February 2, 2026

Abstract

While investors increasingly view firm impacts on nature as financially material, quantifying corporate biodiversity footprints remains challenging due to the complex nature of biodiversity. We introduce an ecologically grounded, data-driven method for estimating firm-level biodiversity footprints. We showcase our method by applying a spatial difference-in-difference design to a data set combining 30m-resolution forest biodiversity data, water pollution data by U.S. facilities, and hydrological model results. We find that toxic water releases have an adverse effect on the downstream forest ecosystem, reducing the tree density and its species diversity. A standard deviation increase in toxic effluent is associated with a 7.5% decline in tree population in the immediate downstream area, with aggregate water pollution leading to a total loss of 18 million trees over the seven-year sample period. A facility's biodiversity footprint is driven more by its siting in heavily forested areas than by variations in effluent volume.

*We thank Douglas Almond, Caroline Flammer, Shahid Naeem, and Viorel Popescu for helpful comments and guidance that greatly improved the study.

†Kim: Columbia University, 420 West 118th Street, New York, NY 10027 (email:sk5316@columbia.edu) . Reese: Columbia University, 420 West 118th Street, New York, NY 10027 (email:sean.reese@columbia.edu).

1 Introduction

Throughout the 21st century, vertebrates have been driven to extinction at rates up to 100 times larger than their historic levels and in 2019, the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services' (IPBES) Global Assessment revealed that at least one million animal and plant species (roughly a ninth of all known species) are at risk of extinction (IPBES, 2019; Ceballos et al., 2015). Concerns about the impact of biological degradation on investment are growing, with 48% of companies in a 2025 survey of a large asset owner indicating that they believe these losses are financially material (Gjerde et al., 2025). This has intensified investor scrutiny of portfolio firms' ecological impact, exemplified by the recent launch of the Taskforce on Nature-related Financial Disclosures (TNFD), whose participating organizations cover $22.4 trillion in assets under management. Simultaneously, firms increasingly fund biodiversity conservation efforts, through means such as blended finance and biodiversity finance (Flammer et al., 2024, 2025).

Central to these developments is the recognition that companies are susceptible to both physical and transition risks stemming from biodiversity loss. Physical risk arises from a company's reliance on nature within its supply chains and the vulnerabilities stemming from ecosystem degradation. Transition risk, by contrast, results from shifts in regulation, consumer preferences, or investor sentiment regarding biodiversity that may adversely affect the firm. Despite increasing investor demand for clarity, characterizing these risks remains difficult, primarily due to biodiversity's multifaceted nature and a lack of standardized data or measurements (Gjerde et al., 2025). In this data-constrained landscape, researchers have proposed various proxies for transition risk, each with its own strengths and weaknesses. Garel et al. (2024) and Coqueret et al. (2025) rely on the "Corporate Biodiversity Footprint" from the Iceberg DataLab, based on proprietary process-based models that couple firm pollution to a specific ecological index ("Mean Species Abundance," MSA). In contrast, Akbari et al. (2025) introduce the "Corporate Biodiversity Exposure" metric, which calculates the number of protected areas within a 30km buffer of a company's production sites, while Giglio et al. (2025) employ textual analysis of 10-K filings to quantify references to biodiversity in corporate disclosures.

Building on the existing suite of metrics, we present an ecologically grounded metric of a firm's biodiversity impact that aligns with the established indicators used by ecologists to directly characterize biodiversity such as species richness, species evenness, or the Shannon-Wiener diversity index (Shannon, 1948; Pielou, 1966). By integrating 30m-resolution forest biodiversity data with toxic effluent records and hydrological models, we analyze the spatially heterogeneous effects of industrial discharge on downstream environments. Our spatial difference-in-differences method accounts for forest location-specific and polluting facility-specific factors that drive biodiversity losses by using pixels immediately upstream of the

facilities as the control group, and by including a rich set of fixed effects and covariates.

We motivate our focus on water pollution and tree biodiversity by briefly surveying the relevant ecological literature. Contrary to the literature's strong focus on carbon and air pollution, we highlight that water pollution has an acute impact on organisms at local scales. Unlike air pollution, which diffuses rapidly across space, chemical stressors introduced via industrial effluent remain localized within the specific hydrological networks that facilities discharge into. These pollutants exert significant chronic pressure on vegetation through toxicological stress (affecting both the trees directly and their symbiotic microbial communities) and the general degradation of the abiotic environment. We also note that while tree diversity does not encompass the full spectrum of local biodiversity, it serves as a vital proxy for one. As "ecosystem engineers" and the dominant primary producers in forested systems, their health dictates the viability of the broader habitat. Ecological research demonstrates that the diversity of higher trophic levels, particularly herbaceous insects, is fundamentally constrained by the diversity of the plant community (Bernays and Chapman, 1994; Tilman et al., 1997; O'Brien et al., 2017). Hence, changes in the plant community will reverberate throughout the broader ecosystem.

Our empirical analysis reveals that surface water pollution leads to significant deterioration in the balance between tree species population, as measured by an ecological index of "species evenness" which captures the relative abundance of different species of trees. The results also suggest that toxic effluent significantly reduce the number of trees per area at the polluting facilities' immediate downstream, with a standard deviation increase in pollution being associated with 7.5% drop in population per area. To the best of our knowledge, this is the first reduced-form study to evaluate the impact of water pollution on observed biodiversity outcomes.

Additional results point to important nuances on the role of pollution and, in general, the complexity of ecosystem responses to firm activity. Importantly, we find that the reductions in tree population are *not* coupled with a loss in the embodied carbon of the forest–a result explained by increases in average tree *size*. The biodiversity effects of pollution are also spatially heterogeneous, with most of the population impact being focused on the first two kilometers downstream of the facility, versus the species evenness effect extending further downstream. We also find that the effect of water pollution depends on the chemical nature of the pollutants: the biodiversity impact is driven by non-nitrogenous pollutants.

We use the reduced-form coefficient estimates to compute a data-driven measure of plant-level biodiversity footprint. We find that industrial water pollution throughout our 7-year sample period has resulted in a loss of 18 million trees. The plant-level biodiversity footprint is strongly driven by the *placement* of plants in heavily forested areas, and to a lesser extent by variations in the volume of toxic effluent. The biodiversity impact is dominated by the

manufacturing sector, notably by facilities producing paper, chemicals, wood products, food, and petroleum and coal products.

Our high resolution data-based approach to estimating firm biodiversity impact differs from existing metrics of biodiversity footprint in that it is entirely data-driven; the estimates do not rely on process-based models that map pollution to a single biodiversity metric (e.g., "Mean Species Abundance" for Garel et al., 2024; Coqueret et al., 2025).[1] This reduced-form approach is even more valuable considering the multifaceted nature of biodiversity–unlike climate pollution (with its uniformly applicable metric of CO_2 equivalent emissions), an equivalent "unit of nature" does not exist (Wauchope et al., 2024; Gjerde et al., 2025). A comprehensive analysis of firm biodiversity footprint thus requires estimations over multiple outcome variables–a task greatly facilitated by gridded ecological data, which allows the researcher to calculate numerous well-established measures of biodiversity (e.g., Shannon's Diversity Index, see Section 2.1).

Further, the gridded ecological data, which we demonstrate in this paper, are public and developed independently from the polluting firms. This distinguishes the approach from text and self disclosure-based measures (e.g., Giglio et al., 2025). The metric also does not require the realization of transition risk into actual conservation policies for its measurement, contrary to Akbari et al. (2025), which captures firms' exposure to ecologically protected area policies, posterior to their imposition. Importantly, we emphasize that the data's use case is not confined to *adverse* impacts on biodiversity, but also could be extended to *beneficial* impacts, such as through firm-financed biodiversity credit projects. The high-resolution gridded data can also be used to calculate the *physical risk* from biodiversity loss, which is of more interest to firms than transition risk (Gjerde et al., 2025).

More broadly, this study also adds to the broader climate and sustainable finance literature by focusing on a relatively understudied aspect of firms' environmental footprint: water pollution. While most studies in this field (e.g., Hsu et al., 2023; Bolton and Kacperczyk, 2023) have been focused on air pollution and carbon pollution, our results suggest that for biodiversity, releases of toxic pollutants into surface water can have detrimental effects.

The paper also contributes to the limited empirical reduced-form literature on the economics of water quality policy (Keiser and Shapiro, 2019b; Paudel and Crago, 2021). Contrary to air quality policies, there has been limited studies that estimate the full spectrum of benefits that water quality improvements can confer (Keiser et al., 2019). Our study fills this gap by presenting the first reduced-form estimate of water quality's biodiversity impact. A prior "contingent valuation" study, which elicit stated preference from surveys, has found sizable

[1] The difference between our reduced-form approach and the life cycle assessment (LCA) model-based estimates from Garel et al. (2024); Coqueret et al. (2025) resembles the difference between reduced-form estimates and integrated assessment models in water pollution (Keiser et al., 2019) and in climate policy (Nordhaus, 1993; Pindyck, 2019).

non-market value of such ecological impacts of water quality (Vossler et al., 2023). Our empirical strategy comparing upstream to downstream pixels builds on similar methods introduced by Keiser and Shapiro (2019a) and Keiser (2019), which were focused on hedonic estimates of water quality policy's benefits.

Lastly, we add to the ecological literature on the impact of water pollution on neighboring ecosystems (Hamdhani et al., 2020). Relatively few studies have directly examined how industrial effluent affects terrestrial tree diversity, with most prior work instead focusing on aquatic ecosystems. Existing studies find greater primary producer (i.e., plants, algae, and bacteria) biomass and decreased fish biodiversity with larger effluent releases (Murdock et al., 2004; Gücker et al., 2006; Birge et al., 1989; Gafny et al., 2000). We extend this literature by analyzing the effects of industrial effluent on terrestrial forest ecosystems. Moreover, by leveraging spatially extensive data covering the contiguous United States, our study moves beyond site-specific case studies to assess whether localized ecological responses generalize across diverse forest types, hydrological contexts, and industrial settings.

The remainder of the paper is structured as follows. Section 2 discusses the fundamentals of biodiversity, its relationship with water pollution, and the recent advances in high-resolution biodiversity data. Section 3 presents the data and empirical strategy used in assessing the effect of pollution on forest biodiversity. Section 4 discusses the empirical results; Section 5 concludes.

2 Background

2.1 Biodiversity and water pollution

2.1.1 Biodiversity: Definition and measurement

According to the Dasgupta Review (Dasgupta, 2021), "biodiversity" encompasses the variety of life in all its forms and at all levels including: genes (the genetic material within individuals), species (populations capable of producing fertile offspring), and ecosystems (communities of organisms and their abiotic environment). While this definition is broad, our analysis focuses specifically on the diversity between species (interspecific diversity) and the diversity among individuals of the same species (intraspecific diversity).

Interspecific diversity consists of two fundamental dimensions: species richness (the number of unique species in an area) and species evenness (the relative abundances of those species). These dimensions can characterize biodiversity in fundamentally different ways: species richness captures changes in community membership through colonizations and extinctions and is therefore a central focus of conservation biology. However, richness alone provides limited information about the biomass, primary production, or competitive interactions within

an ecosystem. For example, an ecosystem may exhibit high species richness even when it is dominated by a single species. In such cases, high richness can mask underlying ecological degradation when the ecological community's functions becomes disproportionately reliant on a small subset of species.

By contrast, species evenness reflects the extent to which individual abundances are dominated by a few species. Changes in evenness within groups of organisms that perform similar ecological roles (i.e. functional groups) reflect shifts in dominance and competitive balance among species and can occur even in the absence of local species loss. This makes evenness particularly informative about compositional change and ecological stress. However, evenness alone is insufficient as a singular metric of biodiversity because it is conditional on species richness: a community composed of a single species is perfectly even by definition, yet lacks functional redundancy and trophic complexity. Evenness therefore conveys information about balance among existing species, but not about the breadth of ecological roles present.

As such, the ecological literature has long recognized the importance of considering both richness and evenness in characterizing biodiversity. Several indexes synthesize these two dimensions into a single measure, but we focus on the two most widely used metrics: the Shannon–Wiener Diversity Index H and Simpson's Diversity Index D (Shannon, 1948; Simpson, 1949). These are calculated as follows:

$$\text{Shannon-Wiener Diversity Index: } H = -\sum_{i=1}^{S} p_i \log(p_i) \tag{1}$$

$$\text{Simpson's Diversity Index: } D = 1 - \frac{\sum_{i=1}^{S} n_i(n_i - 1)}{N(N-1)} \tag{2}$$

where S denotes species richness (the total number of unique species), n_i represents the number of individuals belonging to species i, and $N = \sum n_i$ is the total number of individuals across all species. The term $p_i = \frac{n_i}{N}$ represents the relative size of abundance (i.e., population) of species i within the community. for what Although both indexes combine information on richness and evenness, they differ fundamentally in how they weight species abundances (species population). The Shannon–Wiener index places relatively greater weight on rare species, making it sensitive to the gain or loss of rare species. By contrast, Simpson's index places greater weight on common species and is therefore more responsive to shifts in dominance among the most abundant species.[2]

[2]Although the two indices are frequently highly correlated in empirical applications, divergence between them can nonetheless provide diagnostic insight into patterns of community restructuring. An increase in Shannon–Wiener diversity without a corresponding increase in Simpson's diversity indicates that gains in biodiversity are primarily driven by the colonization or proliferation of rare species. Conversely, a decline in Simpson's diversity accompanied by relatively smaller changes in the Shannon–Wiener index indicates increasing dominance by one

Lastly, the Shannon-Wiener index can be normalized to yield Pielou's Evenness Index (Pielou, 1966):

$$\text{Pielou's Evenness Index } J = \frac{H}{\log(S)} \tag{3}$$

Pielou's index isolates the evenness component of diversity by removing the influence of species richness, allowing changes in relative abundances to be examined directly. Unlike Shannon–Wiener and Simpson's indices, which jointly reflect richness and evenness, Pielou's index provides a complementary measure of compositional balance that is particularly informative when species richness remains stable.

By employing this suite of complementary metrics (species richness, Shannon–Wiener diversity, Simpson's diversity, and Pielou's evenness) we ensure that our analysis captures multiple relevant dimensions of biodiversity. Evaluating these outcomes in tandem allows us to better characterize the mechanisms underpinning the structural evolution of the community, revealing whether ecological shifts are occurring through changes in species occupancy or through shifts in competitive dominance that alter the community's internal balance.

2.1.2 Importance of forest biodiversity

This study focuses on the diversity of trees in forested ecosystems. While tree diversity is admittedly distinct from total biodiversity, it serves as a foundational proxy for the health of forested ecosystems broadly. As the dominant primary producers and structural engineers of the forest, trees exert strong bottom-up control over the broader community. This relationship is formalized in the classical ecological literature on the habitat heterogeneity hypothesis, which posits that structurally complex vegetation provides a wider array of niches (ranging from canopy nesting sites to specific leaf-litter compositions), thereby supporting a more diverse community of invertebrates, birds, and mammals (MacArthur and MacArthur, 1961; MacArthur, 1972; Bazzaz, 1975; Tews et al., 2004). Moreover, because many herbivorous insects exhibit high host-plant specificity, the diversity of insect communities is frequently constrained by plant community diversity (Strong et al., 1984; Bernays and Chapman, 1994; Siemann et al., 1998; O'Brien et al., 2017). Since these invertebrates often form the basis of the food web, a collapse or simplification of tree diversity can create a bottleneck that limits diversity among other organisms. Thus, while tree biodiversity metrics do not capture every dimension of biodiversity, they represent the necessary structural precondition for a diverse forest ecosystem.

Additionally, tree diversity governs the stability of forest ecosystem services–the economic benefits conferred by forests (Heal, 2000). In low-diversity systems, particularly monocultures, the ecosystem lacks functional redundancy; the loss of a dominant species can lead to a collapse

or a few common species, consistent with growing competitive imbalance.

in ecosystem functioning (Solan et al., 2004; Bunker et al., 2005). Monocultures may maximize the production of a specific timber product, but they carry risk regarding disease, pests, and climate stress. Higher diversity mitigates these risks through "compensatory growth," ensuring that the aggregate biomass and service provision remain stable even if individual species decline. Additionally, mixed-species forests have been shown to sequester carbon more reliably and support broader biodiversity among other organisms (e.g., birds and soil microorganisms) than their monoculture counterparts (Gamfeldt et al., 2013; Jucker et al., 2014; Liang et al., 2016).

2.1.3 Surface water pollution and forest biodiversity

Industrial effluent introduces a suite of chemical stressors such as heavy metals, water-soluble organic compounds, and inorganic effluent constituents such as salts and acids into the hydrological systems that support forest ecosystems. Unlike agricultural runoff, which primarily affects forests through nutrient enrichment in nitrogen or phosphorus, industrial effluent influences forest biodiversity largely through toxicological stress and degradation of the abiotic environment.

At the individual tree level, these pollutants impose severe physiological strain on trees through multiple pathways. One major pathway is the inhibition of photosynthesis. Heavy metals such as cadmium and copper can induce chlorosis (leaf yellowing) which impairs a plant's ability to capture solar energy and hence its ability to accumulate biomass (Marschner, 2011; John et al., 2009; Ghori et al., 2019). Additionally, exposure to organic industrial pollutants can elevate the production of reactive oxygen species, leading to oxidative damage of cellular membranes and proteins and forcing plants to divert metabolic resources from growth toward detoxification and repair processes (Schutzendubel and Polle, 2002). A second pathway operates through disruption of water uptake. Elevated concentrations of industrial salts and metals can damage fine root hairs and alter soil osmotic potential, impairing a tree's capacity to absorb water (Barceló and Poschenrieder, 1990; Munns and Tester, 2008). A third pathway is "crown dieback." Chronic exposure to industrial contaminants has been associated with the progressive mortality of distal twigs (a tree's outermost branch sections), branches, and fine roots, particularly in forests exposed to sustained metal deposition (Gawel et al., 1996).

Mortality resulting from these stressors is rarely uniform. Instead, it disproportionately affects individuals with less stored energy and species with limited physiological tolerance, effectively pruning forest stands of less resilient taxa. Because TRI-listed substances (such as lead, mercury, and chromium) exhibit heterogeneous toxicity across plant species, chronic exposure acts as a physiological filter. Over time, this selective pressure reduces species richness as sensitive, often rare, taxa are eliminated while more tolerant species persist.

Beyond their direct effects on trees, industrial pollutants also exert substantial indirect

impacts through soil microbial communities. Forest biodiversity depends critically on belowground symbiotic relationships, particularly those involving mycorrhizal fungi (Van Der Heijden et al., 1998). Industrial contaminants can alter soil pH and salinity or directly kill mycorrhizal fungi, disrupting nutrient acquisition and further compounding stress on forest vegetation (Giller et al., 1998; Schutzendubel and Polle, 2002).

2.2 Biodiversity data: A brief survey and application to biodiversity finance

As alluded to earlier, remarkable progress has been made in the availability of ecological data capturing biodiversity at fine spatial scales (Burgess et al., 2024). The biodiversity data sets differ in their generation process, temporal coverage, and the specific aspects of biodiversity covered.

The four panels of Table 1 correspond to the broad categories into which such biodiversity data fall. The first type, exemplified by the Biodiversity Intactness Index (BII, Newbold et al., 2016), provide model-based predictions of local biodiversity at fine spatial resolution. Such models are fitted on ecological data (e.g., abundance of specific taxon) from numerous peer-reviewed studies, with anthropogenic sources of biodiversity pressure (e.g., land-use change) as predictor variables. The models are then extrapolated to the entire globe.

To the extent that they are cross-sectional in nature, the first type of data sets are very useful in capturing between-firm variation in biodiversity impact. Nevertheless, this precise nature of the data limits their use for reduced-form estimations of firm impact, absent plausibly exogenous cross-sectional variations in firm activity. More importantly, considerable attention to the assumptions implicit in the data is in order. For example, both BII and Mean Species Abundance (MSA) data are based on models that take only abundance, not richness, as the outcome variable. The two models also do not distinguish species based on their conservation status, weighing critically endangered and non-threatened species equally. Using such data to capture biodiversity may not be appropriate when the firm activity is expected to disproportionately affect vulnerable species, or primarily affect species richness and evenness, while not altering the population size.

The second type of data (panel B, Table 1) obtain gridded estimates of biodiversity by directly relying on ecological surveys. For example, eBird Status and Trends data sets rely on citizen science submissions of bird observations to estimate species-specific trends in population and in their range (Fink et al., 2020, 2023). They overcome well-known challenges in using citizen-science data (Johnston et al., 2023) by focusing on research-grade observations and by explicitly taking into account observer efforts in modeling species abundance–resulting in 3km to 27km grid-level estimates of bird species' population and their trends. Similarly, TreeMap (Riley et al., 2022) provides 30m pixel-level estimate of tree species composition and other forest

Table 1. Examples of high-resolution biodiversity data

Name	Spatial Coverage and resolution	Original data source for biodiversity	Temporal coverage and Resolution	Measures abundance/richness?	Weigh species according to conservation status?	Reference
A. Model-based/predicted global biodiversity products						
Biodiversity Intactness Index (BII)	Global, 1km	Global site-level surveys	2015, cross-sectional	Yes/No	No	Scholes and Biggs (2005); Newbold et al. (2016)
Mean Species Abundance (MSA)	Global, 300m	Global site-level surveys	2015, cross-sectional	Yes/No	No	Schipper et al. (2020)
Species threat abatement and restoration (STAR)	Global, 5km	IUCN Red List range	2015, cross-sectional	Yes/Yes, indirectly†	Yes	Mair et al. (2021)
B. Ecological survey-based gridded biodiversity products						
eBird Trends	Global, 27km	Citizen science (eBird research quality observations)	2012-2021 trend, 2013-2022 trend	Yes/Yes*	Yes**	Fink et al. (2023)
eBird Status	Global, 3km	Citizen science (eBird research quality observations)	2021-, annual	Yes/Yes*	Yes**	Fink et al. (2020)
TreeMap (used in this paper)	Continental U.S., 30m	U.S. Forest Service's Forest Inventory and Analysis surveys	2016, 2020, 2022, annual	Yes/Yes	Yes**	Riley et al. (2022)
C. Ecological surveys at fixed sites						

Continued on next page

Table 1 – continued from previous page

Name	Spatial Coverage and resolution	Original data source for biodiversity	Temporal coverage and Resolution	Measures abundance/richness?	Weigh species according to conservation status?	Reference
North American Breeding Bird Survey	North America, ~ 5,000 routes	Bird identification by trained volunteers along vehicle routes	1966-, annual	Yes/Yes*	Yes**	Sauer et al. (2017)
ORSANCO Fish Population	Ohio River, 18 sites	Fish population and biotic conditions surveys	1957-, annual	Yes/Yes*	Yes**	Ohio River Valley Water Sanitation Commission (ORSANCO)
D. Raw species occurrence data/expert range estimates						
IUCN Red List Range	Global, Polygons	Expert consensus	Range map updated upon regular review	Yes/Yes, indirectly†	Yes	IUCN (2024)
Global Biodiversity Information Facility (GBIF)	Global, Points (cleaning and bias correction required)	Numerous reports of species occurrence (peer-reviewed & citizen science)	Reports added upon registration	No/Yes*	Yes**	GBIF (2025)

Note: Table presents a survey of biodiversity data that can be used to assess firm impact on local biodiversity. The table is not exhaustive, and we exclude products with spatial resolution insufficient for local biodiversity impact analysis (coarser than 0.25°, approximately 27km at equator, or products only available at the country level). † = Does not provide abundance or richness metrics, but provide species-specific data (which can be aggregated to calculate richness), and capture abundance by species-specific risk of extinction. *=Provides species-specific abundance, can be aggregated to obtain species richness and other biodiversity metrics (e.g., Shannon's Diversity). **=Provides species-specific data, can be weighted with rarity. IUCN = International Union for Conservation of Nature, ORSANCO = Ohio River Valley Water Sanitation Commission.

characteristics (e.g., biomass per area) using decades-long ecological surveys from the United States Forest Service.

Unlike model-based predictions, such survey-based gridded products provide a more direct measurement of biodiversity. This is because their species-level data can be used to compute well-known metrics of species biodiversity, such as the Shannon-Wiener Index, Simpson's Index, and Gini coefficients, and can also be used to assess heterogeneous impacts of firm activity across different species. Nevertheless, they are often limited in their temporal resolution (e.g., generated every 2-4 years for TreeMap) and in their taxa coverage. Such data sets are hence useful when variations in the measured taxon's biodiversity sufficiently capture the effect of firm activity in question.

In a similar vein, the third type of data (panel C) provide spatially constrained yet temporally extensive snapshots of local biodiversity, with many data sets spanning several decades. A careful empirical strategy that compares within-site, intertemporal variation in exposure to firm activities (e.g., firm pollution) to that of outcomes observed in the fixed sites could reveal the biodiversity impact of firm behavior.

Lastly, raw data on species presence and expert-based species-specific range estimates (panel D) provide essential data on presence of different species across the globe. Carefully combined with data sets on the biotic and abiotic conditions of the reported locations, the species presence data can be used to construct species distribution models (SDMs) which capture the suitability of locations to individual species (Araújo et al., 2019). However, species presence data require considerable data cleaning (Zizka et al., 2019), and the data by definition do not include non-presence; the absence of records cannot be interpreted as true absence of the species. The number of observations *per se* also cannot be treated as an unbiased estimate of abundance, as evidenced in the "data voids" in the tropics (Meyer et al., 2015).

This paper showcases how the second type of biodiversity data (panel B) can be used as outcome variable to assess the impact of toxic water pollution on local forest biodiversity. We leave the application of other forms of ecological data in Table 1 to future studies.

3 Data and Empirical Strategy

3.1 Data

3.1.1 Gridded forest biodiversity data

Forest biodiversity data were derived from TreeMap 2016 (Riley et al., 2021) and TreeMap 2022 (Houtman et al., 2025), which apply the imputation methods of Riley et al. (2022) to generate a 30-m resolution dataset for the United States. Using a random forest algorithm, TreeMap 2016

and 2022 impute US Forest Service Forest Inventory and Analysis (FIA, Bechtold and Patterson, 2005) plots to 30m grid cells based on environmental similarity. Consequently, each pixel is assigned a specific FIA plot, inheriting detailed tree-level attributes (e.g., species, abundance, diameter, height, vital status) and plot-level topographic variables (e.g., slope, aspect, elevation).

We use tree-level species data to compute different indexes of biodiversity, such as richness, the Shannon-Wiener Diversity Index, and Pielou's Evenness Index (Section 2.2). We also refer to other variables obtained from FIA plot data, such as aboveground biomass density and trees per acre.

3.1.2 Toxic releases inventory and ecotoxicity

We obtain data on facility-level toxic pollutant releases from the U.S. Environmental Protection Agency's (EPA) Toxic Release Inventory (TRI). Contrary to prior studies which focused on human health outcomes (Xu and Kim, 2022; Hsu et al., 2023) and hence focused on air pollution (or took the sum of all different forms of emission), we follow the ecological literature (Section 2.1) to focus on surfacewater releases. In accordance with forest biodiversity data, we compute for each plant and chemical the total mass of pollutants released into streamwater over the period of 2017 to 2022. While we focus on all 274 chemicals emitted over this period, in order to make the hydrological model runs more tractable (see below), we limit our sample to 1,037 facilities whose emissions cover at least 90% of emissions for all chemicals.[3]

In aggregating the annual TRI facility-chemical releases to the facility-year level, we note that different chemicals can have vastly different effects on the freshwater ecosystem. In particular, biologically active, water-soluble nitrogen emitted in forms of nitrate compounds, ammonia, and nitric acid can cause eutrophication, all the while having a fertilizing effect on neighboring plants if released in moderation (Odum et al., 1979; Camargo and Alonso, 2006). Further, non-nitrogenous molecules possess different toxicity, with some elements (e.g., mercury) having orders of magnitude larger effect per mass on aquatic organisms and plants than more commonly found components (e.g., zinc).

We thus separate nitrogenous chemicals from non-nitrogenous chemicals, and weight the latter using the ecotoxicological characteristic factor (CF) for individual chemicals, obtained from EPA's Tool for the Reduction and Assessment of Chemical and Other Environmental Impacts (TRACI v2.1, Bare, 2011). The CF, which is outcome-and emission medium-specific, captures the impact of a unit mass emission of individual chemicals to a specific medium. We use CF estimates for "freshwater" medium and "ecotoxicity" outcome. For nitrogenous

[3]To be more precise, we start by choosing top 10% facilities for each chemicals. As many facilities have zero emissions for many chemicals, this leads to covering 100% of emissions for 190 chemicals. 72 chemicals have between 90% to 100% of their emissions covered. For the remaining 12 chemicals, we include additional facilities until we reach the 90% mark.

emissions (from nitrate, ammonia, and nitric acid), we use molecular weight of the compounds to obtain the emissions that belong to nitrogen molecules[4].

3.1.3 Hydrological model data

For each of the 1,037 TRI facilities releasing toxic water pollutants, we identify the upstream and downstream pixels by using the EPA's National Hydrography Dataset Plus (NHDPlus, U.S. Environmental Protection Agency and U.S. Geological Survey, 2012). NHDPlus divides the rivers of the U.S. into short "flowlines" and categorizes them into different hierarchies (i.e., either belonging to the main river or its tributary).

We linked facilities to the stream network by following the NHDPlus digital elevation model from the facility coordinate (obtained from TRI) to the nearest downslope flowline ("initial flowline"). We then retrieve from NHDPlus the flowlines 150 km upstream and downstream of the initial flowline. NHDPlus contains additional data on the hydrologic characteristics (e.g., stream order, length, area, velocity) of the flowlines. The flowlines were spatially unioned into facility-level network geometries and buffered by 500 m. We defined the treatment group as pixels located within the downstream buffer and the control group as pixels located within the upstream buffer.

Figure 1 demonstrates the usage of NHDPlus data to identify treatment and control pixels, for a TRI facility in Cordova, Illinois. The grey dot represents the facility in question, with the neighboring grey line representing the initial flowline. The pixels that are within the buffer of the red (downstream) flowlines are assigned "treatment," while those in the blue buffer (upstream) are "control" pixels.

Note that the above mentioned method assumes that the TRI facility coordinates are good approximations of the outfall of the effluent (i.e., the exact points at which the effluent is released into the streams). While individually verifying the point of release is challenging due to the inconsistent availability of precise outfall coordinates in the EPA's National Pollutant Discharge Elimination System (NPDES), a verifying exercise of a subset of facilities suggest that the spatial offsets between the facility coordinates and their outfall were minimal.

3.2 Empirical strategy

We combine data sets on toxic surfacewater releases, forest biodiversity, and hydrological model results to construct a repeated-observation data set of total water pollution, the exposure

[4]In practice, nitrate (NO_3^-), ammonia (NH_3), and nitric acid (HNO_3) can be emitted in combination with other molecules–nitrate, for example, in the form of sodium nitrate ($NaNO_3$). EPA mandates the plants to report the share of emission that belongs to nitrate only in such cases (EPA, 2000). This allows us to reverse engineer the moles of NO_3^-, NH_3, and HNO_3 from their releases data, using their molecular weight (62.01g/mol, 17.03g/mol, and 63.01g/mol). This represents the moles of nitrogen molecule released.

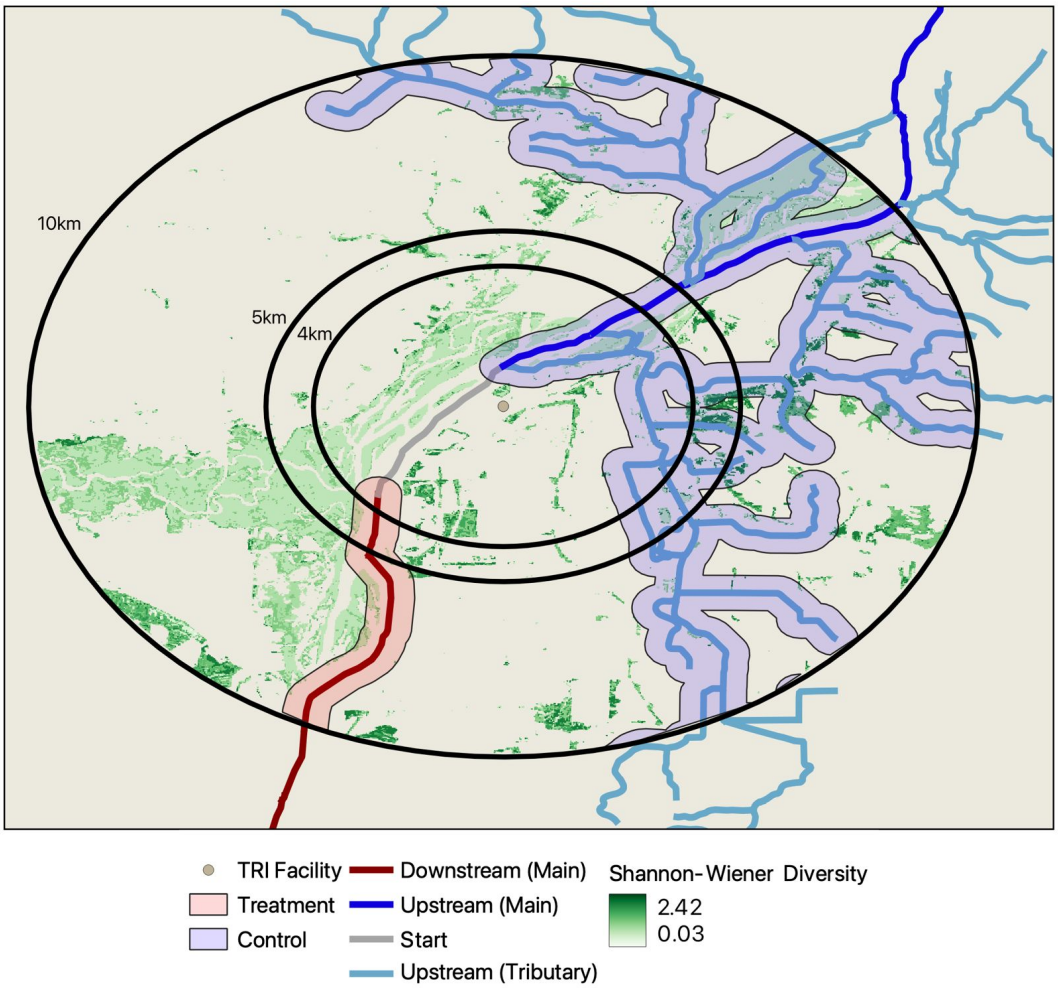


Figure 1. Forest biodiversity and streamflow around a polluting facility in Cordova, IL

Notes: Figure illustrates the processes underlying the combination of hydrological model data, water pollution data, and forest biodiversity data. The dot in grey, and flowline adjacent to it, represent a chemical manufacturing facility in Cordova, Illinois, and the closest NHDPlus (U.S. Environmental Protection Agency and U.S. Geological Survey, 2012) flowline downslope from the facility, respectively. 30m forested pixels from TreeMap (Riley et al., 2021) shown in green, with darker colors representing higher species diversity (calculated from equation (1)). Red and blue flowlines represent downstream and upstream flowlines, with their 500m buffers shown. The pixels within the 500m buffer of the donwstream lines are assigned "treatment" status, while those in the buffer of upstream lines are "control" pixels. Assignment is limited to pixels within 10km radius of the facility (shown in circle).

to such pollution, and shifts in forest biodiversity. We employ spatial difference-in-difference methods in the spirit of Keiser and Shapiro (2019a) and Keiser (2019) and assign upstream and downstream TreeMap pixels as "control" and "treatment" units, where the intensity of treatment is defined by the degree of water pollution. We capture the full spatial heterogeneity in pollutant's effects by dividing different pixels into distance bins.

Our reduced-form equation of interest is:

$$\Delta Y_{kf} = \sum_{k \neq 0} \beta_k \log(\text{WaterPollution}_f) \mathbb{1}[\text{Bin}_{if} = k] + X'_{kf} \lambda_{kf} + \alpha_f + \delta_k + \varepsilon_{if} \tag{4}$$

where ΔY_{kf} is the change in biodiversity outcome, between 2022 and 2016, for pixels in distance bin k for TRI facility f. WaterPollution_f is the amount of (either nitrogenous or non-nitrogenous) surfacewater emission for facility f, between 2017 and 2022. The reference distance bin, denoted as $k = 0$, is the bin 0 to 1km *upstream* of the facility. We include pixels up to 10km upstream, $k = -1, -2, \ldots, -9$, and up to 10km downstream, $k = 1, \ldots, 10$. The β_k estimates for $k > 0$ hence capture the marginal effect of water pollution on the *changes* in biodiversity outcome in pixels k kilometers downstream from the facility, relative to those immediately upstream of the facility. β_k for $k < 0$ in turn captures the effect on upstream pixels, akin to pre-trend terms in event-study methods. The difference-in-differences nature of model (4) controls for facility-bin-specific characteristics that define their baseline *levels* of biodiversity.

We employ a rich set of fixed effects and covariates to account for various confounders. The facility fixed effect α_f capture facility-specific shocks between 2016 and 2022, while bin fixed effects δ_k control the fixed effect of being downstream or upstream from a facility (e.g., improvements in biodiversity outcomes from increased water flow).

While α_f and δ_k account for facility-specific and bin-specific shocks, there is still a chance that the WaterPollution_f variable captures the industrial activity occurring at facility f, and that such activities have spatially heterogeneous effects on biodiversity across bins k. We thus compute $\text{NonWaterPollution}_f$, defined as the difference between 2017-2022 total toxic chemical releases and surfacewater pollution. We then include the interaction term $\text{NonWaterPollution}_f \mathbb{1}[\text{Bin}_{if} = k], k \neq 0$ in the vector of covariates X_{kf}. The interaction term controls for the spatially heterogeneous effect of fluctuations in plant activity on local biodiversity.

The error term ε_{if} is two-way clustered at the facility and state level to allow for arbitrary within-facility and within-state spatial correlation in residuals, while we test the robustness of our results to Conley standard errors. We weight the regression by the number of forested pixels in each distance bin, to make sure that the regression results are reflecting the entire population of forested pixels in the sample.

We focus on variations in multiple outcomes ΔY_{kf}, all obtained from TreeMap, in order to capture the different aspects of forest biodiversity. In particular, we use distance bin-averaged tree per acre, tree stocking (normalized density of trees), aboveground biomass per acre, tree diameter–all of which capture the overall health of the forest ecosystem at different distance bins. We also use tree-level data to compute the different species biodiversity indexes (Section 2.1) at the 30m pixel level, and take the average of such indexes for each distance bin. The bin-level

averaging reduces the noise from individual pixel-level changes in biodiversity metrics.

4 Results

4.1 Firm pollution and biodiversity

Figure 2 presents the spatial difference-in-differences coefficient estimates of model (4) for different metrics of forest biodiversity. Across all subplots, the treatment variable is the log of toxicity-weighted non-nitrogenous surface water effluent (henceforth effluent) released from the facility.

Figure 2a plots live tree stocking, defined as the percentage of potential growing space occupied by live trees (with 100% representing that tree density has reached its maximum biological potential in the forested pixel), as the outcome variable. The coefficient estimates for upstream pixels fail to find significant effects of effluent on upstream pixels, hence supporting the empirical strategy, while the coefficient estimates for downstream pixels suggest a sharp reduction in tree stocking up to 4 km downstream from the polluting facility, with the impact diminishing with distance. The coefficient estimate for 1 km (-0.43) suggests that a standard deviation increase in effluent is associated with a 1.40% drop (-0.43×3.26) in tree stocking in the immediately downstream pixels. This is equivalent to 0.146 of the baseline standard deviation (9.57%).

Figure 2b plots the number of trees per acre as the outcome variable and confirms the results shown in Figure 2a. A single SD increase in pollution leads to 7.5% decrease (-0.023×3.26) in tree population at the pixels 1 km downstream. The effects diminish and become statistically insignificant in pixels further down from the facility. Both Figure 2a and 2b results point to toxic water releases from TRI facilities adversely affecting the forests downstream.

Surprisingly, the results from Figure 2c suggest that such losses in tree populations are not associated with significant changes in per-area biomass. Figure 2d suggests a mechanism behind such results: population reductions are offset by *increases* in the size of remaining trees. This could be a consequence of competitive release for light or nutrients and, in water-constrained regions, with increased water availability from effluent. Growth on the intensive margin is no longer observed (and even reverses) once reductions along the extensive margin disappear at 3 km, further supporting our proposed mechanism of competitive release.

While Figure 2 focuses on biodiversity metrics of the overall health of the forest ecosystem, Figure 3 focuses on the aforementioned aspects of species biodiversity (evenness and richness, Section 2.1). Figure 3a suggests a muted effect on species evenness (equation (3)) for immediately downstream pixels, while further downstream (3–8 km) pixels experience a reduction in evenness. A single SD increase in effluent reduces the evenness index 3–4 km downstream

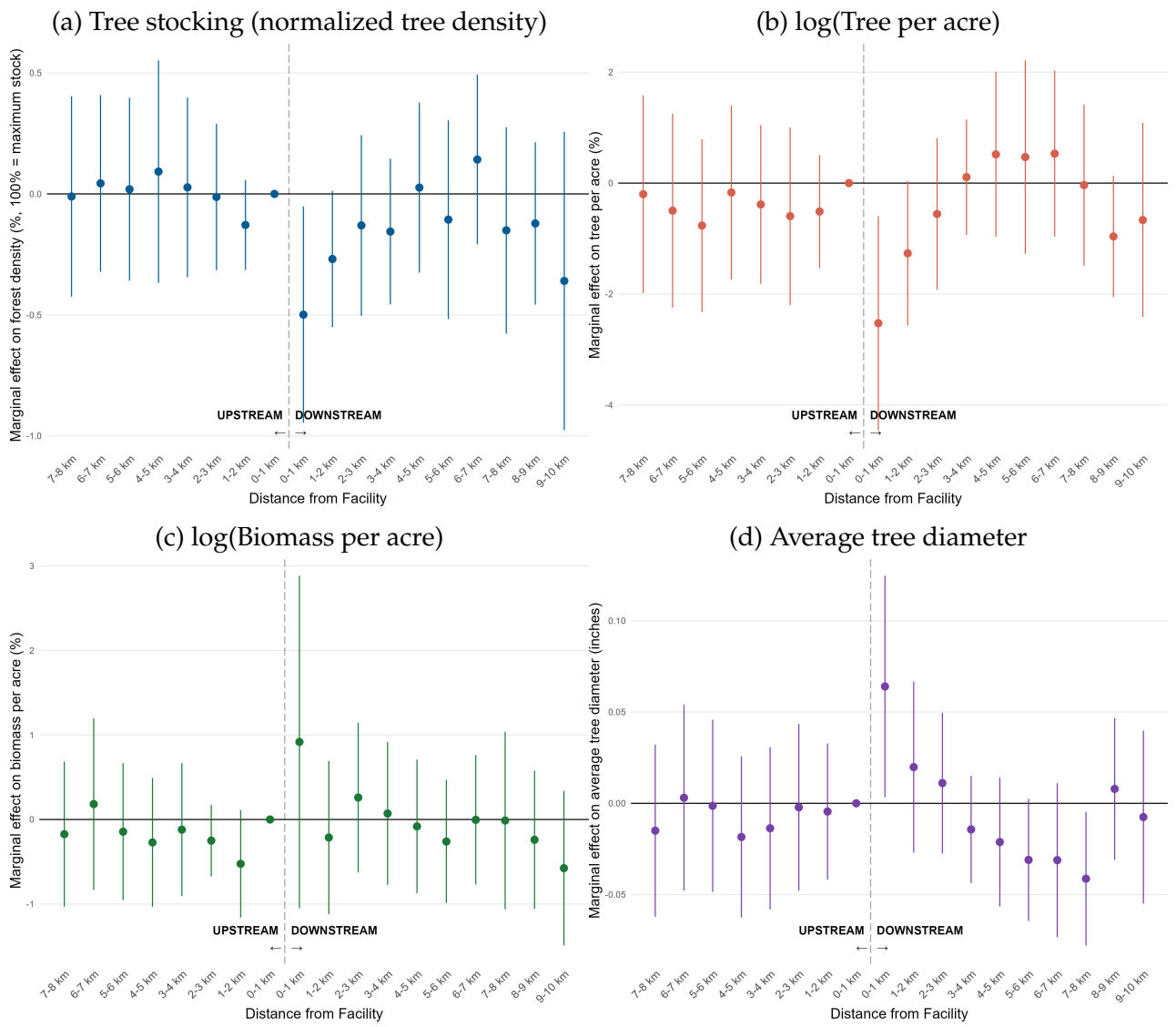


Figure 2. Water pollution and forest biodiversity: Forest ecosystem health

Notes: Figure presents the point estimates and 95% confidence intervals of spatial difference-in-difference coefficients (equation (4)) for different forest biodiversity metrics. The reference group is the change in the metric for the pixel 0 to 1km upstream of the facility. (a) takes alive tree stocking, a normalized metric of tree density, as the outcome variable. (b) and (c) use tree per acre and biomass per acre, both log-transformed, as the outcome. (d) is the same but for average tree diameter, in inches. Estimates and standard errors multiplied by 100 for (b) and (c) for legibility.

by 0.006 (-0.001829×3.26), which is 7.5% of the baseline SD of the index. Declines in evenness imply increasingly unequal species distributions, consistent with differential selection that disproportionately affects non-dominant species. In contrast, species richness remains largely unchanged across downstream distances (Figure 3b), indicating that pollution reshapes relative abundances without reducing the total number of species present.

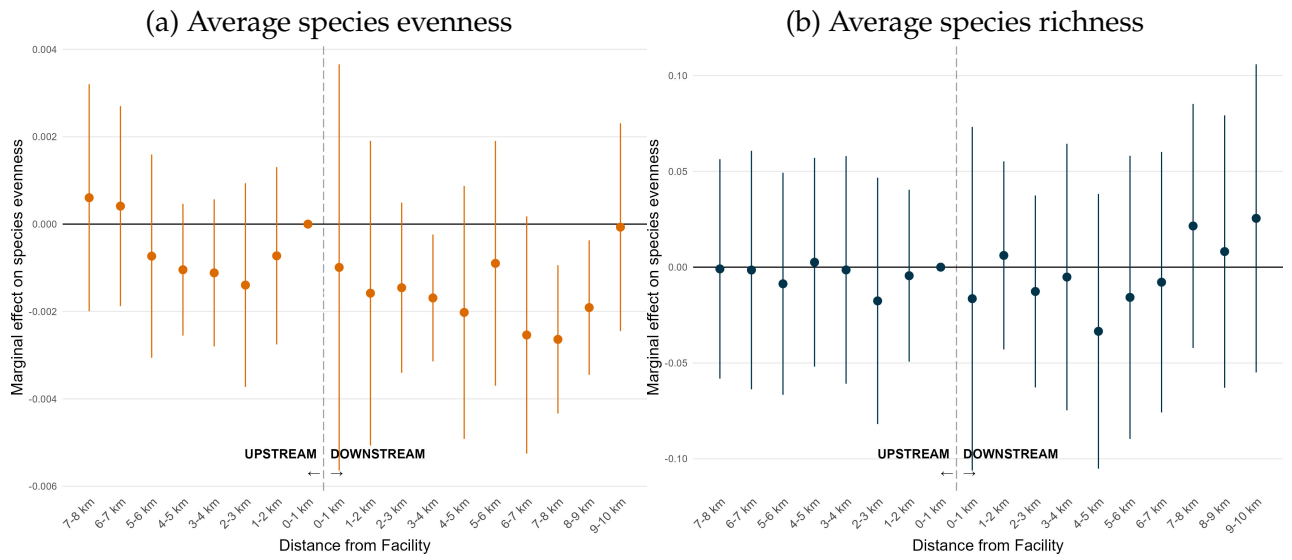


Figure 3. Water pollution and forest biodiversity: Species diversity

Notes: Figure presents the point estimates and 95% confidence intervals of spatial difference-in-difference coefficients (equation (4)) for different forest species diversity metrics. The reference group is the change in the metric for the pixel 0 to 1km upstream of the facility. (a) and (b) focus on the species evenness index (equation (3)) and species richness (number of unique species in each pixel) as outcome.

Taken together, the results suggest strong spatial heterogeneity in the effect of toxic water pollutants on forest biodiversity. The immediate vicinity of the effluent is characterized by countervailing effects on tree population and individual tree sizes (Figure 2), potentially due to relaxing competition or the dual role of effluent: providing essential water (especially in water-constrained regions) while simultaneously containing harmful pollutants (Hamdhani et al., 2020). The primarily adverse impact observed in the ensuing forest stands (3 to 8km, Figures 2d and 3a) suggests that after the first 2 kilometers, the impact of toxic pollutants overwhelm the effect of added water. In general, the impacts converge to zero in the pixels furthest from the facility (i.e., 10km downstream).

While Figures 2 and 3 focus on non-nitrogenous pollutants, Appendix Figures A1 and A2 explore the impact of nitrogenous pollution. Contrary to non-nitrogenous pollutants, we do not see a significant deterioration in tree population near the facility–with the point estimates for tree stocking, biomass, and tree diameter instead pointing to an imprecise yet positive effect of nitrogenous emissions.

4.2 Facility-level biodiversity footprint

We now use the coefficient estimates from the spatial difference-in-difference estimates to provide a data-driven estimate of firm biodiversity footprint. To this end, for each TRI facility

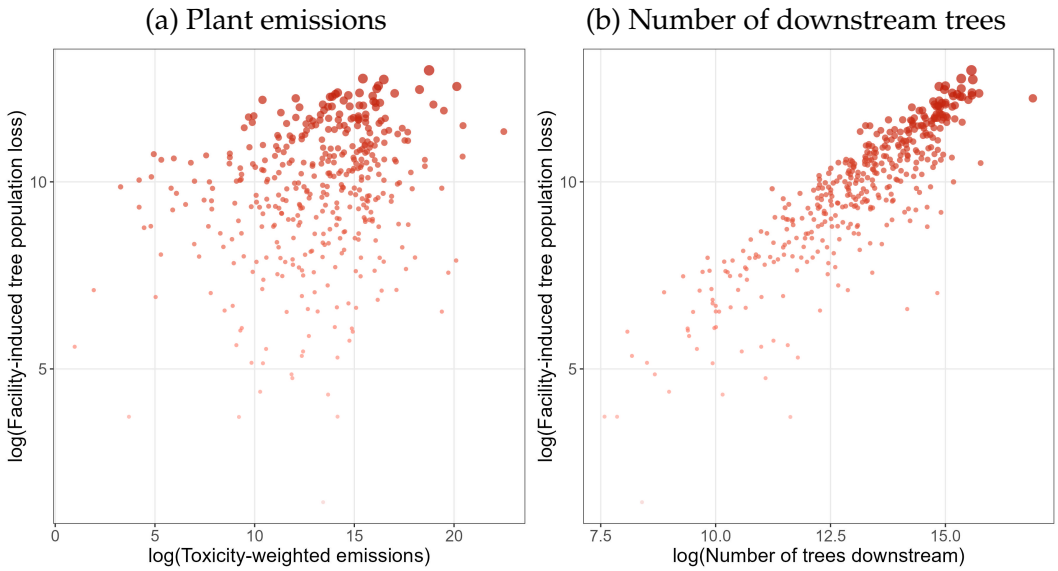


Figure 4. Pollution versus exposure: Drivers of facility-level biodiversity footprint

Notes: Figure assesses the drivers of TRI facility-level biodiversity footprint. The footprint, on the y axis of both Figures 4a and 4b, is calculated as the sum of trees lost at different distance bins downstream from the facility. Each point represents a TRI facility, with darker and larger points representing more negative biodiversity impacts. Figure 4a places the pollution intensity (ecotoxicity-weighted emissions) on the x axis. Figure 4b shares the same axis but uses the number of trees exposed to such pollution (number of trees downstream) in the x axis.

emitting non-nitrogenous pollutant, we compute the number of trees at different downstream distance-bins. We then take the coefficient estimates on the log of tree per acre (Figure 2b) in downstream distance bins, and take the sum of the coefficients' product with the downstream distance bin-level tree count. The resulting facility-level estimate represents the number of trees that have been lost from the plant's non-nitrogenous, industrial pollution.[5]

The results from this calculation suggest that between 2016 and 2022, a total of 18.1 million riparian trees have been lost to industrial pollution from TRI facilities. Such negative effect of pollution comes against the backdrop of an overall *increase* in the downstream tree population, which grew from 1,606 million to 1,616 million trees. Absent industrial pollution, the increase in tree population would have been more than double its observed rate.

To the extent that the facility-level biodiversity footprint is a function of both pollution intensity (total amount of toxic emissions) and exposure to such pollution (number of trees downstream), Figure 4 assesses which of these two aspects drive the footprint estimates. The results suggest that the biodiversity footprints are primarily driven by their placement in tree-rich areas (Figure 4b) and less so by their industrial pollution (Figure 4a). Such importance

[5]As before, we limit the back-of-the-envelope calculation to up to 10km downstream from a facility. The resulting figure will very likely be an underestimate of the total impact of industrial pollution on local biodiversity.

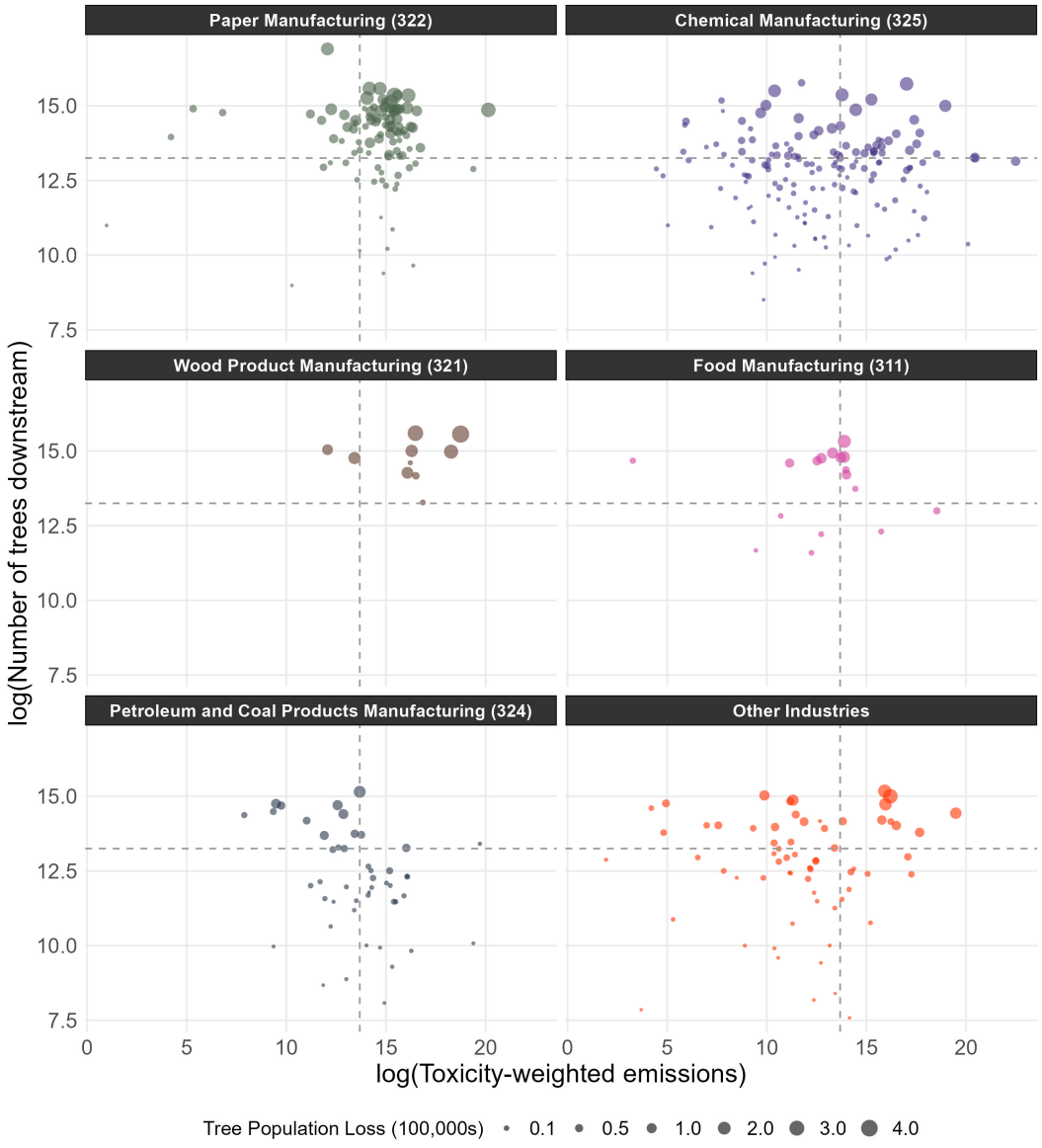


Figure 5. Components of facility-level biodiversity footprint, across different industries

Notes: Figure assesses the distribution of the two components of biodiversity footprint (pollution intensity and the number of trees downstream), across different industry sectors. Each dot represents a facility, with larger and darker dots representing larger biodiversity footprint. The median value of "pollution intensity" and "number of trees downstream," across all facilities, in dashed lines. First three digits of the NAICS code in parentheses.

of proximity to population-dense areas is analogous to air pollution's negative externality on human health outcomes (e.g., Muller and Mendelsohn, 2009).

Aggregating the facility-level biodiversity footprint to different industries using the associated NAICS code, we find that 40% of the tree losses (7.2 million trees) is driven by the paper manufacturing industry. This is followed by 26%, 10%, 6%, and 6% from the chemical

manufacturing, wood product manufacturing, food manufacturing, and petroleum and coal products manufacturing industries. The decomposition reveals once again the importance of facility location in driving biodiversity impact, as most wood- and paper manufacturing-related facilities are placed in the proximity of densely forested areas.

Figure 5 further illustrates how the industries differ in the way through which they reach such biodiversity impact. The figure plots for facilities the number of trees on the y axis and the emission intensity on the x axis. The dashed lines represent the median value of the component. Inspecting the distribution of facilities (points) along the y axis, we find that paper manufacturing and wood product manufacturing facilities are consistently *above* the median, suggesting that they are generally larger in the number of trees downstream. In contrast, facilities in the chemical manufacturing are almost equally distributed above and below the median, while most petroleum and coal products manufacturing facilities are lower in downstream tree population.

Moving onto the x axis, we find that most paper manufacturing and wood manufacturing facilities are also above median in their pollution intensity. Petroleum and coal products manufacturing facilities, despite their lower number of downstream tree population, have large pollution activities. In contrast, many facilities in chemical manufacturing and food manufacturing are lower than median in their emission intensities.

5 Conclusion

We apply spatial difference-in-difference methods to a data set combining 30m-resolution forest biodiversity data, surface water pollution records, and hydrological models. The results suggest that toxic water pollution, in particular the non-nitogenous toxicants, have a significant and negative impact on the forest ecosystem downstream–both in terms of their population size and in their species diversity. We also provide additional results suggesting the spatial heterogeneity of the impacts, and point to the limits of focusing on simple biomass measures to capture forest ecosystem health.

There are several refinements to our empirical approach that we plan to pursue in future work. First, one limitation of the current specification is that the straight-line distance from a TRI facility (Figure 1), which we use to define distance bins, does not correspond exactly to distance along the stream network. To address this, we plan to compute downstream distances from the TRI facility using flowline lengths from the NHDPlus dataset. Second, we will explore an alternative specification that constructs counterfactuals at the level of individual stream segments using their associated drainage basins. Rather than comparing treated pixels to areas upstream of the TRI facility (which may be geographically distant) we will instead use areas upstream of each flowline's drainage basin. This approach both increases the effective sample

size by incorporating additional forested pixels and allows for more spatially granular fixed effects. We could also significantly improve the precision of our estimates by using the facility-streamline, instead of facility, as the unit of clustering. Finally, we plan to substantiate the role of water in increasing tree diameter, by incorporating information on facility-level discharge volumes (from NPDES). Combined with NHDPlus stream-level data on the annual streamflow, this will allow us to classify streams into categories commonly used in the ecological literature: effluent-dependent (>90% effluent during baseflow), effluent-dominated (>50% effluent), and effluent-impacted (<50% effluent). This classification will allow us to explicitly test our proposed mechanism of water subsidization underlying the observed increases in tree diameters near TRI facilities.

Further, we also plan to translate our reduced-form estimates into a firm-level biodiversity footprint, using the parent company identifiers in the TRI facility data. Aggregating impacts across facilities operated by the same parent firm will yield a reduced form-based measure of firm biodiversity transition risk. Through a range of sensitivity analyses, we will analyze whether ignoring the chemical properties of the pollutants emitted, or not taking into account the between firm-heterogeneity in the number of neighboring forested pixels, leads to significant biases in the firm biodiversity footprint.

Lastly, we note that the methods already developed and discussed in this paper can be readily applied to other contexts where firms' impacts on biodiversity are of interest–such as in biodiveristy finance projects. We leave the application of other fine-scale biodiversity gridded data (Table 1) for future work.

References

Akbari, Amir, Lilian Ng, Man Duy (Marty) Pham, and Jing Yu, "The real effects of protecting biodiversity," *Review of Finance*, 10 2025, p. rfaf055.

Araújo, Miguel B, Robert P Anderson, A Márcia Barbosa, Colin M Beale, Carsten F Dormann, Regan Early, Raquel A Garcia, Antoine Guisan, Luigi Maiorano, Babak Naimi et al., "Standards for distribution models in biodiversity assessments," *Science advances*, 2019, *5* (1), eaat4858.

Barceló, JUAN and Ch Poschenrieder, "Plant water relations as affected by heavy metal stress: a review," *Journal of plant nutrition*, 1990, *13* (1), 1–37.

Bare, Jane C., "TRACI 2.0: the tool for the reduction and assessment of chemical and other environmental impacts 2.0," *Clean Technologies and Environmental Policy*, 2011, *13* (5), 687–696.

Bazzaz, FA, "Plant species diversity in old-field successional ecosystems in southern Illinois," *Ecology*, 1975, *56* (2), 485–488.

Bechtold, William A and Paul L Patterson, *The enhanced forest inventory and analysis program–national sampling design and estimation procedures* number 80, USDA Forest Service, Southern Research Station, 2005.

Bernays, Elizabeth A and RE Chapman, *Host-plant selection by phytophagous insects*, Springer, 1994.

Birge, Wesley J, Jeffrey A Black, Terry M Short, and Albert G Westerman, "A comparative ecological and toxicological investigation of a secondary wastewater treatment plant effluent and its receiving stream," *Environmental toxicology and chemistry*, 1989, *8* (5), 437–450.

Bolton, Patrick and Marcin Kacperczyk, "Global pricing of carbon-transition risk," *The Journal of Finance*, 2023, *78* (6), 3677–3754.

Bunker, Daniel E, Fabrice DeClerck, Jason C Bradford, Robert K Colwell, Ivette Perfecto, Oliver L Phillips, Mahesh Sankaran, and Shahid Naeem, "Species loss and aboveground carbon storage in a tropical forest," *Science*, 2005, *310* (5750), 1029–1031.

Burgess, Neil D, Natasha Ali, Jacob Bedford, Nina Bhola, Sharon Brooks, Alena Cierna, Roberto Correa, Matthew Harris, Ayesha Hargey, Jonathan Hughes et al., "Global metrics for terrestrial biodiversity," *Annual Review of Environment and Resources*, 2024, *49* (1), 673–709.

Camargo, Julio A and Álvaro Alonso, "Ecological and toxicological effects of inorganic nitrogen pollution in aquatic ecosystems: a global assessment," *Environment international*, 2006, *32* (6), 831–849.

Ceballos, Gerardo, Paul R Ehrlich, Anthony D Barnosky, Andrés García, Robert M Pringle, and Todd M Palmer, "Accelerated modern human–induced species losses: Entering the sixth mass extinction," *Science advances*, 2015, *1* (5), e1400253.

Coqueret, Guillaume, Thomas Giroux, and Olivier David Zerbib, "The biodiversity premium," *Ecological Economics*, 2025, *228*, 108435.

Dasgupta, Partha, *The economics of biodiversity: the Dasgupta review.*, Hm Treasury, 2021.

EPA, "Toxics Release Inventory: List of Toxic Chemicals Within the Water Dissociable Nitrate Compounds Category and Guidance for Reporting," Technical Report EPA 745-R-00-006, U.S. Environmental Protection Agency, Office of Environmental Information, Washington, DC December 2000.

Fink, Daniel, Alison Johnston, Matt Strimas-Mackey, Tom Auer, Wesley M Hochachka, Shawn Ligocki, Lauren Oldham Jaromczyk, Orin Robinson, Chris Wood, Steve Kelling et al., "A double machine learning trend model for citizen science data," *Methods in Ecology and Evolution*, 2023, *14* (9), 2435–2448.

_ , Tom Auer, Alison Johnston, Viviana Ruiz-Gutierrez, Wesley M Hochachka, and Steve Kelling, "Modeling avian full annual cycle distribution and population trends with citizen science data," *Ecological Applications*, 2020, *30* (3), e02056.

Flammer, Caroline, Thomas Giroux, and Geoffrey Heal, "Blended finance," Technical Report, National Bureau of Economic Research 2024.

_ , _ , and Geoffrey M Heal, "Biodiversity finance," *Journal of Financial Economics*, 2025, *164*, 103987.

Gafny, Sarig, Menachem Goren, and Avital Gasith, "Habitat condition and fish assemblage structure in a coastal mediterranean stream (Yarqon, Israel) receiving domestic effluent," *Hydrobiologia*, 2000, *422* (0), 319–330.

Gamfeldt, Lars, Tord Snäll, Robert Bagchi, Micael Jonsson, Lena Gustafsson, Petter Kjellander, María C Ruiz-Jaen, Mats Fröberg, Johan Stendahl, Christopher D Philipson et al., "Higher levels of multiple ecosystem services are found in forests with more tree species," *Nature communications*, 2013, *4* (1), 1340.

Garel, Alexandre, Arthur Romec, Zacharias Sautner, and Alexander F Wagner, "Do investors care about biodiversity?," *Review of Finance*, 2024, *28* (4), 1151–1186.

Gawel, James E, Beth A Ahner, Andrew J Friedland, and François MM Morel, "Role for heavy metals in forest decline indicated by phytochelatin measurements," *Nature*, 1996, *381* (6577), 64–65.

GBIF, "GBIF Data Processing and Quality Pipeline," 2025. Accessed on [Date].

Ghori, N-H, T Ghori, MQ Hayat, SR Imadi, A Gul, V Altay, and M Ozturk, "Heavy metal stress and responses in plants," *International journal of environmental science and technology*, 2019, *16* (3), 1807–1828.

Giglio, Stefano, Theresa Kuchler, Johannes Stroebel, and Xuxia Wang, "Biodiversity Risk," *Review of Finance*, 2025. Forthcoming/Working Paper version 2025.

Giller, Ken E, Ernst Witter, and Steve P Mcgrath, "Toxicity of heavy metals to microorganisms and microbial processes in agricultural soils: a review," *Soil biology and biochemistry*, 1998, *30* (10-11), 1389–1414.

Gjerde, Snorre, Zacharias Sautner, Alexander F Wagner, and Alexis Wegerich, "Corporate nature risk perceptions," *Review of Finance*, 2025, p. rfaf050.

Gücker, Björn, Mario Brauns, and Martin T Pusch, "Effects of wastewater treatment plant discharge on ecosystem structure and function of lowland streams," *Journal of the North American benthological society*, 2006, *25* (2), 313–329.

Hamdhani, Hamdhani, Drew E Eppehimer, and Michael T Bogan, "Release of treated effluent into streams: A global review of ecological impacts with a consideration of its potential use for environmental flows," *Freshwater Biology*, 2020, *65* (9), 1657–1670.

Heal, Geoffrey, "Valuing ecosystem services," *Ecosystems*, 2000, pp. 24–30.

Heijden, Marcel GA Van Der, John N Klironomos, Margot Ursic, Peter Moutoglis, Ruth Streitwolf-Engel, Thomas Boller, Andres Wiemken, and Ian R Sanders, "Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity," *Nature*, 1998, *396* (6706), 69–72.

Houtman, Rachel M, Lila ST Leatherman, Scott N Zimmer, Ian W Housman, Abhinav Shrestha, John D Shaw, and Karin L Riley, "TreeMap 2022 CONUS: A tree-level model of the forests of the conterminous United States circa 2022," 2025.

Hsu, Po-Hsuan, Kai Li, and Chi-Yang Tsou, "The pollution premium," *The Journal of Finance*, 2023, *78* (3), 1343–1392.

IPBES, *Global Assessment Report on Biodiversity and Ecosystem Services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services*, Bonn, Germany: IPBES secretariat, 2019.

IUCN, *Mapping Standards and Data Quality for IUCN Red List Spatial Data (Version 1.19)* International Union for Conservation of Nature 2024. Prepared by the IUCN Red List Unit.

John, Riffat, Parvaiz Ahmad, K Gadgil, and Satyawati Sharma, "Heavy metal toxicity: Effect on plant growth, biochemical parameters and metal accumulation by Brassica juncea L.," 2009.

Johnston, Alison, Eleni Matechou, and Emily B Dennis, "Outstanding challenges and future directions for biodiversity monitoring using citizen science data," *Methods in Ecology and Evolution*, 2023, *14* (1), 103–116.

Jucker, Tommaso, Olivier Bouriaud, Daniel Avacaritei, and David A Coomes, "Stabilizing effects of diversity on aboveground wood production in forest ecosystems: linking patterns and processes," *Ecology Letters*, 2014, *17* (12), 1560–1569.

Keiser, David A, "The missing benefits of clean water and the role of mismeasured pollution," *Journal of the Association of Environmental and Resource Economists*, 2019, *6* (4), 669–707.

__ and Joseph S Shapiro, "Consequences of the Clean Water Act and the demand for water quality," *The Quarterly Journal of Economics*, 2019, *134* (1), 349–396.

__ and __, "US water pollution regulation over the past half century: burning waters to crystal springs?," *Journal of Economic Perspectives*, 2019, *33* (4), 51–75.

__, Catherine L Kling, and Joseph S Shapiro, "The low but uncertain measured benefits of water quality policy," *Proceedings of the National Academy of Sciences*, 2019, *116* (12), 5262–5269.

Liang, Jingjing, Thomas W Crowther, Nicolas Picard, Susan Wiser, Mo Zhou, Giorgio Alberti, Ernst-Detlef Schulze, A David McGuire, Fabio Bozzato, Hans Pretzsch et al., "Positive biodiversity-productivity relationship predominant in global forests," *Science*, 2016, *354* (6309), aaf8957.

MacArthur, Robert H, *Geographical ecology: patterns in the distribution of species*, Princeton University Press, 1972.

__ **and John W MacArthur**, "On bird species diversity," *Ecology*, 1961, *42* (3), 594–598.

Mair, Louise, Leon A Bennun, Thomas M Brooks, Stuart HM Butchart, Friederike C Bolam, Neil D Burgess, Jonathan MM Ekstrom, EJ Milner-Gulland, Michael Hoffmann, Keping Ma et al., "A metric for spatially explicit contributions to science-based species targets," *Nature Ecology & Evolution*, 2021, *5* (6), 836–844.

Marschner, Horst, *Marschner's mineral nutrition of higher plants*, Academic press, 2011.

Meyer, Carsten, Holger Kreft, Robert Guralnick, and Walter Jetz, "Global priorities for an effective information basis of biodiversity distributions," *Nature communications*, 2015, *6* (1), 1–8.

Muller, Nicholas Z and Robert Mendelsohn, "Efficient pollution regulation: getting the prices right," *American Economic Review*, 2009, *99* (5), 1714–1739.

Munns, Rana and Mark Tester, "Mechanisms of salinity tolerance," *Annu. Rev. Plant Biol.*, 2008, *59* (1), 651–681.

Murdock, Justin, Daniel Roelke, and Frances Gelwick, "Interactions between flow, periphyton, and nutrients in a heavily impacted urban stream: implications for stream restoration effectiveness," *Ecological Engineering*, 2004, *22* (3), 197–207.

Newbold, Tim, Lawrence N Hudson, Andrew P Arnell, Sara Contu, Adriana De Palma, Simon Ferrier, Samantha LL Hill, Andrew J Hoskins, Igor Lysenko, Helen RP Phillips et al., "Has land use pushed terrestrial biodiversity beyond the planetary boundary? A global assessment," *Science*, 2016, *353* (6296), 288–291.

Nordhaus, William D, "Optimal greenhouse-gas reductions and tax policy in the" DICE" model," *The American Economic Review*, 1993, *83* (2), 313–317.

O'Brien, Michael J, Matteo Brezzi, Andreas Schuldt, Jia-Yong Zhang, Keping Ma, Bernhard Schmid, and Pascal A Niklaus, "Tree diversity drives diversity of arthropod herbivores, but successional stage mediates detritivores," *Ecology and Evolution*, 2017, *7* (21), 8753–8760.

Odum, Eugene P, John T Finn, and Eldon H Franz, "Perturbation theory and the subsidy-stress gradient," *BioScience*, 1979, *29* (6), 349–352.

Ohio River Valley Water Sanitation Commission (ORSANCO), "Ohio River Fish Population Database," Data set, retrieved from agency website 2025. Data retrieved from the ORSANCO Biological Programs website.

Paudel, Jayash and Christine L Crago, "Environmental externalities from agriculture: evidence from water quality in the United States," *American Journal of Agricultural Economics*, 2021, *103* (1), 185–210.

Pielou, Evelyn C, "The measurement of diversity in different types of biological collections," *Journal of theoretical biology*, 1966, *13*, 131–144.

Pindyck, Robert S, "The social cost of carbon revisited," *Journal of Environmental Economics and Management*, 2019, *94*, 140–160.

Riley, Karin L, Isaac C Grenfell, John D Shaw, and Mark A Finney, "TreeMap 2016 dataset generates CONUS-wide maps of forest characteristics including live basal area, aboveground carbon, and number of trees per acre," *Journal of Forestry*, 2022, *120* (6), 607–632.

_ , _ , Mark A Finney, and John D Shaw, "TreeMap 2016: A tree-level model of the forests of the conterminous United States circa 2016," 2021.

Sauer, John R, Keith L Pardieck, David J Ziolkowski Jr, Adam C Smith, Marie-Anne R Hudson, Vicente Rodriguez, Humberto Berlanga, Daniel K Niven, and William A Link, "The first 50 years of the North American breeding bird survey," *The Condor: Ornithological Applications*, 2017, *119* (3), 576–593.

Schipper, Aafke M, Jelle P Hilbers, Johan R Meijer, Laura H Antão, Ana Benítez-López, Melinda MJ de Jonge, Luuk H Leemans, Eddy Scheper, Rob Alkemade, Jonathan C Doelman et al., "Projecting terrestrial biodiversity intactness with GLOBIO 4," *Global Change Biology*, 2020, *26* (2), 760–771.

Scholes, Robert J and Reinette Biggs, "A biodiversity intactness index," *Nature*, 2005, *434* (7029), 45–49.

Schutzendubel, Andres and Andrea Polle, "Plant responses to abiotic stresses: heavy metal-induced oxidative stress and protection by mycorrhization," *Journal of experimental botany*, 2002, *53* (372), 1351–1365.

Shannon, Claude E, "A mathematical theory of communication," *The Bell system technical journal*, 1948, *27* (3), 379–423.

Siemann, Evan, David Tilman, John Haarstad, and Mark Ritchie, "Experimental tests of the dependence of arthropod diversity on plant diversity," *The American Naturalist*, 1998, *152* (5), 738–750.

Simpson, Edward H, "Measurement of diversity," *nature*, 1949, *163* (4148), 688–688.

Solan, Martin, Bradley J Cardinale, Amy L Downing, Katharina AM Engelhardt, Jennifer L Ruesink, and Diane S Srivastava, "Extinction and ecosystem function in the marine benthos," *Science*, 2004, *306* (5699), 1177–1180.

Strong, Donald R, John H Lawton, and Sir R Southwood, "Insects on plants. Community patterns and mechanisms.," 1984.

Tews, Jörg, Ulrich Brose, Volker Grimm, Katja Tielbörger, Matthias C Wichmann, Monika Schwager, and Florian Jeltsch, "Animal species diversity driven by habitat heterogeneity/diversity: the importance of keystone structures," *Journal of biogeography*, 2004, *31* (1), 79–92.

Tilman, David, Johannes Knops, David Wedin, Peter Reich, Mark Ritchie, and Evan Siemann, "The influence of functional diversity and composition on ecosystem processes," *Science*, 1997, *277* (5330), 1300–1302.

U.S. Environmental Protection Agency and U.S. Geological Survey, "NHDPlus Version 2.1: National Hydrography Dataset Plus," https://www.epa.gov/waterdata/nhdplus-national-hydrography-dataset-plus 2012. Accessed: 2025-12-18.

Vossler, Christian A, Christine L Dolph, Jacques C Finlay, David A Keiser, Catherine L Kling, and Daniel J Phaneuf, "Valuing improvements in the ecological integrity of local and regional waters using the biological condition gradient," *Proceedings of the National Academy of Sciences*, 2023, *120* (18), e2120251119.

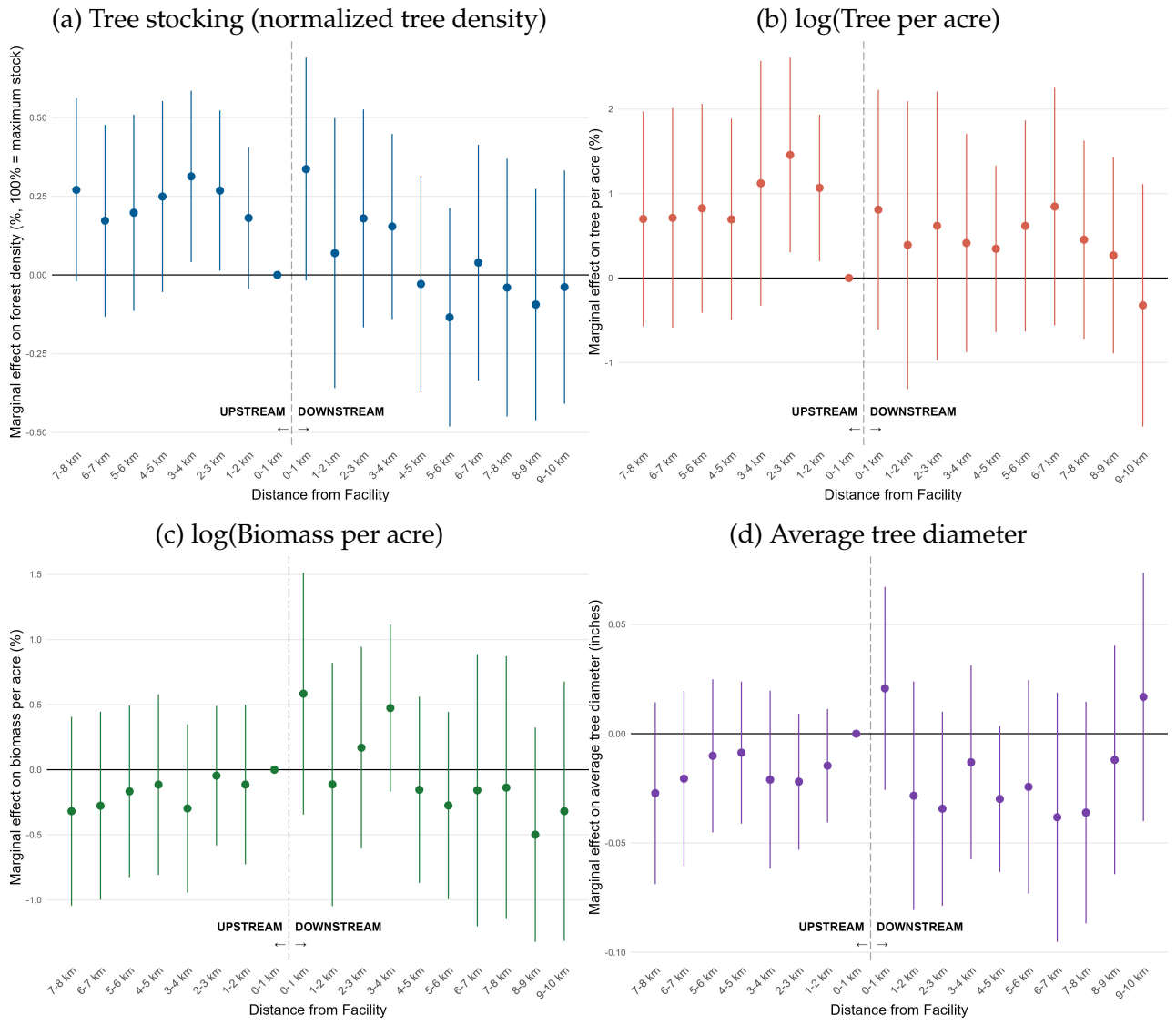
Wauchope, Hannah S, Sophus OSE Zu Ermgassen, Julia PG Jones, Harrison Carter, EJ Milner-Gulland et al., "What is a unit of nature? Measurement challenges in the emerging biodiversity credit market," *Proceedings of the Royal Society B: Biological Sciences*, 2024, *291* (2036).

Xu, Qiping and Taehyun Kim, "Financial constraints and corporate environmental policies," *The Review of Financial Studies*, 2022, *35* (2), 576–635.

Zizka, Alexander, Daniele Silvestro, Tobias Andermann, Josué Azevedo, Camila Duarte Ritter, Daniel Edler, Harith Farooq, Andrei Herdean, María Ariza, Ruud Scharn et al., "CoordinateCleaner: Standardized cleaning of occurrence records from biological collection databases," *Methods in Ecology and Evolution*, 2019, *10* (5), 744–751.

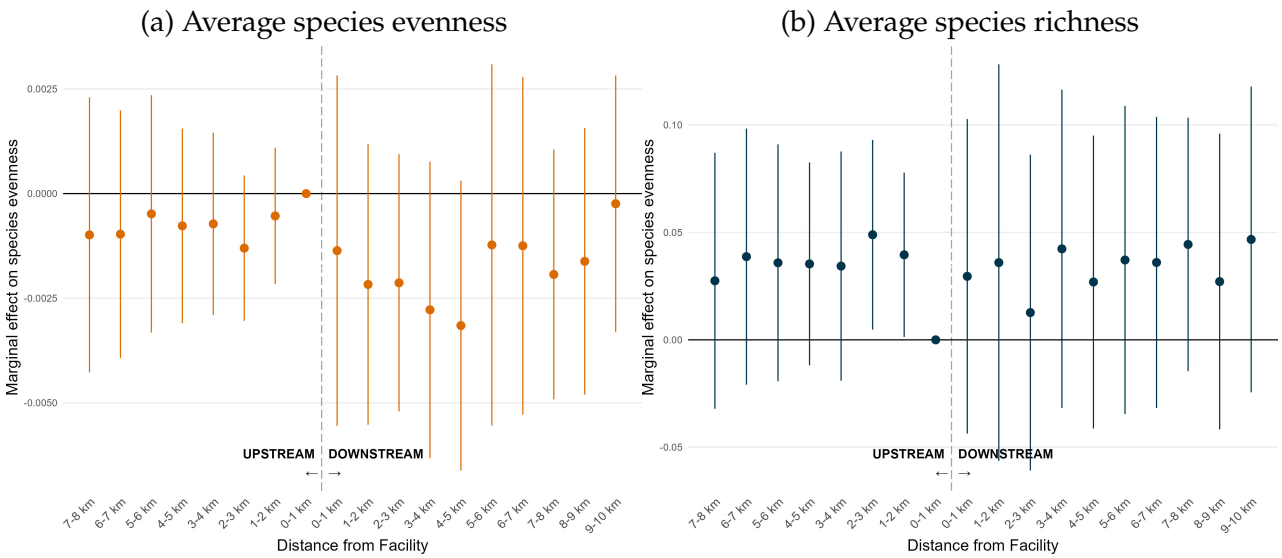
Appendix

A Additional Figures



Appendix Figure A1. Nitrogenous water pollution and forest biodiversity: Tree characteristics

Notes: Figure repeats Figure 2, but for nitrogenous water pollution, defined as the weighted sum of nitrate (NO_3^-), ammonia (NH_3), and nitric acid (HNO_3) releases, where the weight is the inverse of molecular mass of NO_3^-, NH_3, and HNO_3.



Appendix Figure A2. Water pollution and forest biodiversity: Species diversity

Notes: Figure repeats Figure 3, but for nitrogenous water pollution.