

1 The role of long-distance dispersal for community 2 dynamics of fungi

3 Jussi Mäkinen^{1,*}, Janne Mäyrä¹, Yalda Fatahi², Mikhail Sofiev², Otso Ovaskainen³, Nerea
4 Abrego³, and Veera Norros¹

5 ¹Finnish Environment Institute (Syke), Latokartanonkaari 11, 00790 Helsinki, Finland

6 *Corresponding author. Email: jussi.a.makinen@syke.fi

7 ²Finnish Meteorological Institute, Erik Palménin aukio 1, 00560 Helsinki, Finland

8 ³Department of Biological and Environmental Science, University of Jyväskylä, Surfontie 9, 40500
9 Jyväskylä, Finland

10 ABSTRACT

Realized dispersal ability of wind-dispersing fungi in natural conditions remains poorly understood. Here, we study how climate, habitat type and long-distance dispersal jointly structure fungal communities and how their effects vary between species that have different morphological and life history traits. We integrate global fungal (spores and other organismal material) occurrence data collected in 2018-2021 with atmospheric dispersion projections that are used to define sample connectivity and to assess land cover in the spore source areas. Connectivity and land cover similarity increase community similarity, but the effects are small compared to the effect of climatic sorting. In contrast, connectivity is the main factor controlling the occurrence pattern of relatively common species and higher for plant pathogens than for all other species. We present the framework as a proof of concept -type of work to highlight the progress and future potential in studying dispersal of microbial organisms at the global scale.

Keywords: metacommunity, fungi, atmospheric dispersion modeling, airborne eDNA, metabarcoding, joint species distribution modeling

1 INTRODUCTION

Understanding global patterns of diversity of fungal species has lagged behind the respective understanding of animals and plants (Green et al., 2004). Similarly to other groups of microbial species, most fungi could be expected to be cosmopolitan and to have high dispersal abilities (Finlay, 2002). However, according to recent findings, fungal communities are spatially structured latitudinally and regionally (Tedersoo et al., 2014; Davison et al., 2015; Vetrovsky et al., 2019) with remarkable levels of endemism (Talbot et al., 2014). On a global scale, the patterns of fungal diversity and their community structure are driven by climate, especially the annual mean temperature (Vetrovsky et al., 2019; Abrego et al., 2024b) challenging the idea of fungi and their spores as homogeneously distributed propagules that reflect mainly stochastic processes (Vellend, 2010). Together, previous findings indicate a relatively high importance of climatic sorting and dispersal limitation compared to random fluctuations in the structure of the fungal community (Pellitier et al., 2025). Still, direct links between dispersal vectors, such as wind patterns, and fungal community structure remain poorly understood (Pellitier et al., 2025). Thus, a comprehensive quantification of the impacts of environmental sorting, dispersal limitation, and random factors for the community structure of fungi is needed.

For most fungal species, wind is the main dispersal vector at large spatial scales (Chaudhary et al., 2022). However, species differ in terms of their physical properties and therefore their expected airborne dispersal ability (Norros et al., 2014; Wang et al., 2021). Although fungal dispersal traits and dispersal syndromes are still incompletely understood (Cuprewich et al., 2025), it is clear that spore size is one of the key traits influencing dispersal ability due to its strong effect on the velocity of deposition (Hussein et al., 2013). Smaller spores have longer potential dispersal distances (Norros et al., 2014), while allowing a bet-hedging strategy through a larger number of spores (Norros et al., 2023). There is a large and systematic variation in spore size across the fungal kingdom (Calhim et al., 2018), with a general tendency to larger spores in symbiotic vs. free-living species (Aguilar-Trigueros et al., 2023). At the same time, dispersal ability cannot be equated with spore size, as dispersal can be dramatically affected by other traits such as viability and release conditions that may also covary with spore size (Norros et al., 2015; Lagomarsino Oneto et al., 2020; Golan et al., 2023).

The effect of wind on spore dispersal has previously been studied in controlled experiments (Norros et al., 2012) and in long- and short-term observational studies (Pellitier et al., 2025; Naranjo-Orrico et al., 2025). These previous studies show that species differ in terms of their abilities for passive dispersal (Norros et al., 2014, 2012) and that dispersal limitation through wind conditions may structure the composition of the fungal community (Pellitier et al., 2025; Naranjo-Orrico et al., 2025). However, important knowledge gaps remain. Importantly, the importance of daily to seasonal dispersal events for the momentary community structure of fungi in the air is unknown. In addition, it is unclear how increasing the source area of spores increases their species richness and whether this species-area relationship varies between scales and bioclimatic regions (Veresoglou and Johnson, 2023), as has been shown, for example, for aquatic species (He et al., 2024; Soininen and Hillebrand, 2007). The effects of short-term and inter-annual variability of meteorological parameters have been shown to be significant for other types of primary biogenic particles, such as pollen (Siljamo et al., 2008; Skjoth et al., 2008; Sofiev, 2017, 2019; Veriankaitė et al., 2009).

The main objectives of our study are to quantify the importance of airborne dispersal for the taxonomic community composition and occurrence of fungi. We associate community composition and species distributions with climate and wind conditions in a short time frame, i.e., a few days, to assess the importance of realised dispersal limitation. We take advantage of novel data streams of airborne DNA-based spore data, predictions of atmospheric dispersion models, and Earth observation data to associate the species occurrence and community composition of fungal spores with the characteristics of the spore source areas. We characterise the similarity of sample source areas in two procedures: 1) through spatial overlap between sample footprints (potential source areas) and 2) through similarity in forest, agriculture, urban and sea coverage of sample footprints. We assume that forests, urban and agricultural areas have different fungal communities, which is supported by the findings of Verchot (2010); Delelegn et al. (2018); Abrego et al. (2020).

We predict that dispersal source areas quantified by connectivity (procedure 1) or land cover (procedure 2) are an important factor for community composition and species occurrence patterns. We break these predictions into the following more specific hypotheses. We hypothesise that increasing the sum of the land area in the source area increases the richness of airborne species (H1.1), indicating a generally higher niche variety in larger areas and corresponding to a classic species-area relationship (Peay et al., 2007, 2016). We predict that the coverage of the forest, agricultural and urban areas modify the composition of the airborne fungal community composition (H1.2) and that samples that are connected through air dispersion are more similar than samples that are not connected (H1.3). The effect of connectivity depends on the temporal window in the definition of connectivity (H1.4). Due to the mechanistic dispersal information, we predict that connectivity is a better predictor of community dissimilarity than geographical distance (H1.5). We predict that connectivity and land cover of the source areas affect the probability of the presence of individual species (H2.1), but their importance varies in relation to the species' trophic guild and spore size. We predict that due to the greater importance of dispersal limitation for species with large spores and resource specialists such as wood saprotrophs (Norros et al., 2012), these groups are more sensitive to connectivity and land

cover characteristics in the source area than other species groups (H2.2).

2 MATERIALS AND METHODS

Spore, trait and climate data extraction

The spore data set, sampling methods, and bioinformatic workflow have been described in Ovaskainen et al. (2024b). Here we summarise the methods of the study. Ovaskainen et al. (2024b) collected airborne fungal particles from 47 globally distributed sites using cyclone air samplers over standardised 24-hour periods. The fungal DNA material is assumed to be from spores, although it could also represent other types of fungal dispersal propagules or fungal material. The samples captured particles larger than 1 μ m and included field blanks and negative controls to monitor contamination. From air samples, DNA was extracted, the fungal ITS2 region was amplified with PCR, and the resulting amplicons were sequenced on an Illumina MiSeq platform. Most of the negative controls did not produce fungal readings and the tests showed that low-level cross-contamination did not affect the main patterns in the data.

Bioinformatic processing involved trimming, denoising, and chimera checking of reads, followed by alignment against the UNITE fungal reference database (Abarenkov et al., 2024) supplemented with synthetic spike-in controls. After quality filtering, the data set consisted of 65912 ASV $\times$ 2768 sample matrix of read abundance. The authors developed a taxonomically guided clustering approach to assign ASVs to species-level operational taxonomic units (OTUs) (Furneaux et al., 2025). First, ASVs were probabilistically assigned to taxa using Protax-fungi (Abarenkov et al., 2018). The unassigned ASVs were matched with the closest cluster with optimised similarity thresholds. As the last step, the unassigned ASVs were clustered using de novo single-linkage clustering. The final data set contained 27954 species-level fungal OTUs in 2768 samples (see Fig. 1: panel A).

The data of the trophic guild and spore size were assembled in Aguilar-Trigueros et al. (2023). The spore size data for sexual and asexual cycles originate from species-level taxonomic descriptions in Mycobank (Robert et al., 2013a). The trophic guild data have been assembled from publicly available databases (see the full list in Aguilar-Trigueros et al. (2023)). We used species-level estimates of spore size and trophic guild assembled by Abrego et al. (2024b). Linking spore volume data to molecularly identified taxa is challenging because some organisms were not identified to the species level and existing spore volume databases are incomplete. When available, species-level spore volume estimates were used. If these were missing, the authors substituted genus-level averages, then family-level averages if needed. When no estimate was available at any of these taxonomic levels, the spore volume for that taxon was treated as missing data.

Trophic guild assignments were limited to the most common guilds, with some categories grouped. OTUs reliably identified to the species-level were matched directly to an existing trophic guild database. When only genus- or family-level identifications were available, all guilds associated with species in that genus or family were assigned to the OTU. Consequently, some OTUs were linked to multiple possible trophic guilds and these assignments were treated as provisional, based on information from related taxa, rather than direct species-level evidence.

As the mean annual temperature, we included the climatic mean over 1979–2018 from ERA5 hourly data on the single level dataset (Hersbach et al., 2020) (available at the Copernicus Climate Data Store <https://cds.climate.copernicus.eu/datasets/reanalysis-era5-land-timeseries>). The species occurrence and trait and the climate dataset were sourced from Abrego et al. (2024a). See code and data availability for data sources.

Atmospheric dispersion modelling

The SILAM (System for Integrated modelLling of Atmospheric coMposition, <https://silam.fmi.fi> (Sofiev et al., 2015)) atmospheric composition model developed by the Finnish Meteorological Institute (FMI) computes dispersion and inverse dispersion and can be used to assess atmospheric footprints as part of its capabilities (Verstraeten et al., 2025; Sofiev et al., 2023, 2024).

A footprint in atmospheric modelling describes the geographical and temporal source areas for emission concentrations observed at a receptor (or vice-versa), showing which source areas contribute most to what is measured at a given location and time (Veriankaitė et al., 2009). High footprint values indicate regions where emissions would have the strongest impact on the receptor. An atmospheric footprint is produced using inverse (adjoint) dispersion modelling. Instead of simulating how emissions spread forward in time from sources to receptors, SILAM can run the transport equations backward in time from a receptor (e.g., a measurement site, population location or observation period) (Sofiev et al., 2006, 2015). The resulting footprint field quantifies the sensitivity of a receptor concentration to upwind emissions as a function of space, time, and altitude. Mathematically, it represents a transfer coefficient that links the emissions in each grid cell to the observed concentration at the receptor. In practice, SILAM footprints have been used to trace the long-range transport of aerosols (Meinander et al., 2020), pollen (Veriankaitė et al., 2009; Sofiev et al., 2022), and pollutants (Sofiev, 2000) back to their source regions, helping to identify the source areas that affect a receptor site during days of atmospheric transport.

Here, we applied SILAM to compute footprints in a 36 hour window preceding the day-long spore sampling. We used only the lowest layer (5 metres) of the SILAM calculations as that best corresponds to the sampling height of the spore samples. As a reference particle diameter in SILAM, we used 18 μm , which on average corresponds to fungal spores. The footprints have a spatial resolution of 0.5°. We defined connectivity between two samples as the spatial overlap between their footprints. The spatial overlap was defined as the cosine similarity, where footprint concentration was considered as statistical probability distributions. Then

$$\text{cosine}_{sim}(u, v) = \frac{\sum_{i=1}^n u_i v_i}{\sqrt{\sum_{i=1}^n u_i^2} \sqrt{\sum_{i=1}^n v_i^2}}$$

where u and v are footprint values and i is an index for spatial coordinate (here a cell in a global grid). Similarity is defined as the overlap of the footprint densities normalised with the total mass of the footprint. A high proportion of the footprint mass originates from relatively short distances. An example of a footprint and the area from which 100% and the top 99, 95, and 75% of the concentration originate are shown in Fig. 1.

We tested the impact of different time windows by setting footprint similarity to zero if samples had not been sampled within a certain time window. We tested four different time windows, +/- day, month, three months, or a year. The time window is computed over the preceding and subsequent days, months, or years. The number of samples that are connected to any other sample depends on the width of the temporal window (Fig. 2: panels A and B and Fig. S3 and S4).

We transformed the similarity index into a distance metric by applying the Floyd-Warshall algorithm. Similarity is not a valid distance metric as it does not fulfil the triangle inequality: a distance between two samples can decrease if it is calculated through a third sample. Floyd-Warshall algorithm computes the shortest path between two samples using all other samples as potential nodes along the path (Floyd, 1962). Thus, the distances correspond to the distances in a non-linear network. Non-linearity is caused by the non-stationary wind conditions during the sampling period. We tested only the month-long temporal window for defining the distance matrix of the sample connectivity. The distance in relation to footprint overlap was then used as a distance matrix of a random effect in the species distribution model to account for the sample covariance that is due to connectivity.

Land cover data

We characterised land cover using the Global Dynamic Land Cover of 2019 in 100 metre resolution from the Copernicus Global Land Service (information, 2020). These data provide a detailed view of land cover at three classification levels and achieves 80% precision at class level 1 on each continent. Because these data do not extend north of 80°N or south of 60°S, we used ocean cover data from `rnaturalearth` R-package (Massicotte and South, 2026) to classify these areas into either Open sea or Unknown classes. For each spore sample, we calculated two types of land cover data. First, we weighted land cover with the footprint to derive the relative contribution of different land cover types as source areas, i.e. the relative particle concentration that would be predicted to originate from forest, agriculture, urban, and water areas. Second, we calculated the unweighted proportion of these land cover classes within a 500 km radius circle around the sampling site. We used 500 km as a crude prior estimate of the source area. However, we acknowledge that the 500 km radius clearly underestimates the average reach of footprints from focal sampling sites. See Fig. S1 for a comparison of footprint thresholds in relation to the distribution of their spatial radius.

Statistical modelling of community turnover

We estimated the effect of mean annual temperature, seasonality, sample connectivity, land cover around and at the sampling site, and sequencing depth on the community structure by associating sample-level pair-wise community turnover with the difference in explanatory variables and sample connectivity. We conducted this only for sample-pairs from different sites and thus the model estimates do not represent the stochasticity in within-site community turnover.

We measured community turnover with Jaccard dissimilarity index and the difference in explanatory variables with Euclidean distance (see Fig. S2 for distribution of sample-pair community turnover). We measured seasonality with periodic functions $\sin(2\frac{\pi d}{365})$ and $\cos(2\frac{\pi d}{365})$ where d is the Julian day. The sin function captures the seasonality in the Northern Hemisphere whereas the cos-function captures the seasonality in the Southern Hemisphere. They are both weighted by latitude to capture the latitudinally increasing effect of seasonality (Abrego et al., 2024b). We computed the difference in land cover as the Euclidean distance in relation to two metrics: coverage of land cover classes within 500 km radius (non-weighted) or proportions of footprint weighted land cover classes under the footprint (footprint weighted). The distance is measured in four dimensional space whereas the difference for other variables is measured in one dimension for each variable separately. With this approach, we could capture the divergence jointly in all land cover dimensions. Lastly, the Jaccard dissimilarity index and explanatory variables were scaled by first adding one and taking the logarithm of the sum, and second by standardising the resulting variables.

Land cover at the sampling site reflects whether two sites are in the same land cover category (forest, agriculture, urban, water, or other land cover type) and was treated as a binary variable. We transformed similarity measures of footprint overlap by taking the logarithm of the non-zero values and standardizing them. The effects of the binary overlapping/non-overlapping variable and the continuous variable for the strength of the overlap are accounted for in the statistical model. We accounted for the statistical dependence between sample pairs by including two sample-level random effects representing the samples that were used for sample pairs, following the approach of Row et al. (2017). This relaxes the dependence between sample-pairs and returns valid credibility intervals for the model estimates.

For computing community turnover, we discarded two samples that did not have a calculated footprint ($n=2$). In addition, samples with fewer than 10,000 sequence reads were excluded, and variation in sequencing depth was handled statistically rather than by rarefaction by including pairwise difference in sequencing depth as an explanatory variable in the model. This resulted in 2716 samples that contained in total 27935 species-level OTUs.

We did a model comparison on two levels: weighting scheme of land cover (non-weighted and footprint weighted) and community turnover of a specific species group. To assess how community turnover varies between species groups, we fitted the linear model separately for all species (1), plant pathogens only (2) or wood saprotrophs only (3). To test the explanatory power of connectivity, we fitted models in which connectivity was replaced with a spatial distance between sample sites. We applied lme4 R-package (v. 1.1-38) for fitting linear mixed effect models.

Finally, we measured the importance of the variable with variance partitioning. We computed the linear predictor with the fitted model while assigning original values for each variable at the time and fixing other variables to zero for continuous variables or to a certain class for random effects. We computed the variance of the linear predictors of each variable and divided each variance by the sum of the variances. This resulted as proportions with which each variable contributed to the full linear predictor (Schulz et al., 2025).

Joint species distribution model

To estimate the effects of connectivity and land cover on single species distributions, we fitted a joint species distribution model (JSDM) (Ovaskainen et al., 2017). We explained the presence-absence status of all species jointly with mean annual temperature, seasonality, land cover around and at the sampling site, and log-transformed sequencing depth. To model the presence-absences, we fitted a probit model. We included connectivity in the JSDM as the distance matrix of a random effect. We included the first- and second order polynomial effects of mean annual temperature and the first order effect of other explanatory variables in the model. For land cover, we used land cover at the sample site (categorical) and either non-weighted or footprint weighted land cover.

To achieve enough statistical power in the JSDM, we kept species that had been observed in at least 50 samples and excluded samples that had no footprint ($n=2$). Similarly to community turnover, samples with fewer than 10,000 reads were excluded and variation in sequencing depth was handled statistically rather than by rarefaction by including sequencing depth as an explanatory variable in the model. This resulted in 2716 samples that contained in total 485 species.

We accounted for phylogenetic relatedness in the species-level covariate effect estimates by estimating the magnitude of interspecific covariance in the covariate effects that is structured by the phylogenetic tree. We assumed equal branch lengths at the levels of phylum, class, order, family, genus, and species in the phylogenetic tree. We fitted the models with the Hmsc r-package (versions 3.1-2) (Tikhonov et al., 2020) in the R software (version 4.5.2; R Core Team (2021)) and applied the parallelized HPC-extension of Hmsc (Rahman et al., 2024). Hmsc allows a fully Bayesian model inference and a comprehensive uncertainty assessment through posterior samples. We conducted posterior sampling with five chains and per chain we drew 100,000 samples of which we discarded the first 50000 and kept one per 50 remaining samples. This resulted in 1000 samples per chain and 5000 posterior samples in total. We checked the convergence of the model for each parameter with potential scale reduction factors (Gelman and Rubin, 1992).

3 RESULTS

Species-area relationship

We found that species richness did not depend on the concentration of air particles that originate from land areas (footprint weighted model; Fig.3: panel A). When conditioning species-area relationship on the latitudinal zones ($< 30^\circ$ N/S, $30-60^\circ$ N/S and $> 60^\circ$ N/S) and growing season (all year when $< 30^\circ$ N/S, days 275-60 when $> 30^\circ$ S and $> 30^\circ$ N), the footprint-weighted species-area relationship appeared in low latitudes ($< 30^\circ$ N/S), but not in mid to high-latitudes in either of the seasons. The non-weighted species-area relationship was positive and consistent across the latitudinal zones in the growing season, though slightly stronger in the low latitudes (Fig.3: panels B, C).

Community turnover

The samples connected through footprints were on average more similar than non-connected samples. Community turnover was primarily explained by temperature followed by land cover, seasonality, and connectivity (Fig. 4: panels A and C). The effect of connectivity on community turnover varied in relation to the temporal window of connectivity (Fig. 4: panels B). Although the effect was the largest in a short temporal window (+/- day), the importance of connectivity in variance partitioning increased towards longer time windows (Fig. 4: panel D). The discrepancy between the effect size and the importance of the variable was due to the low levels of connectivity between the samples when connectivity was defined in strict temporal terms. Although connectivity increased similarity in community composition, its overall effect on global turnover was minor due to the low frequency of connectivity. There was no difference between the land cover variables that reflect the footprint source area or the area within a circle around the sampling sites (Fig. S5).

Models with footprint informed connectivity predicted the rate of community turnover more accurately than models where connectivity was defined with spatial distance (Fig. 4: panel E).

Connectivity had a relatively smaller effect on plant pathogen turnover than all species turnover, while land cover explained relatively less of the wood saprotroph turnover compared to all species and plant pathogens (Fig. 4: panel A).

Species distribution modelling

The occurrence patterns of 485 most prevalent species (≥ 50 presence records) were mainly explained by a connectivity related sample-level random effect followed by mean annual temperature and either site-level random effect in the footprint weighted model and land cover of the surrounding area in the non-weighted model (Fig. 5: panel A). In the non-weighted model, land cover soaked up variation especially from the site-level random effect compared with the footprint weighted model.

We tested how well models captured the expected association between land cover and spore occurrence: In the footprint weighted model, no species responded positively and 125 negatively to water areas, while the response was uncertain for 360 species. In the non-weighted model, 18 species responded positively and 160 negatively to water areas, whereas the response was uncertain for 307 (Fig. 5: panel B).

In both footprint-weighted and non-weighted models, the importance of land cover was higher for species with large spores (Fig. 5: panel C). The importance of connectivity did not vary in relation to the spore size (Fig. S6). Regarding trophic guilds, we found that land cover was more important for plant pathogens than other species groups, but only when land cover was weighted by footprints (Fig. S7). In contrast, we found a smaller importance of connectivity (random effect) for wood saprotrophs than for other species (Fig. S8). We found satisfying levels of model convergence for all model parameters (Fig. S9).

4 DISCUSSION

We developed and tested a framework that relates fungal spore observations with their expected source areas and quantifies the importance of environmental and seasonal sorting and dispersal limitation for community composition. The framework was enabled by a novel data stream from atmospheric dispersion models to approximate the connectivity of spore samples and source areas of fungal spores. We found that the occurrence of fungal spores depended on the connectivity between the samples and the land cover in source area. However, the importance of connectivity depended on the suitability of the dispersion events (time window for connectivity). In addition, we found that the species' trophic guild and spore size contributed to the importance of connectivity in spore community structuring (see Table 1 for hypothesis and results). Next, we break down our findings and discuss their ecological and practical implications.

Community turnover was primarily explained by temperature that refuted large scale community turnover between biomes and climatic regions (Abrego et al., 2024b; Tedersoo et al., 2014; Vetrovsky et al., 2019) and debunked the prediction that fungal spores are cosmopolitan and homogeneously distributed throughout space (Finlay, 2002). However, community turnover responded to connectivity and land cover in the particle source area reflecting short-term dispersion events. Although connectivity shaped community composition, sample connections were relatively rare in the dataset and thus the overall importance of connectivity stayed low. The connectivity measure could capture dispersal events and explain community turnover across spatial scales of a few 100 to 1000 km that correspond to the geographical extent of footprints. Accounting for a temporal window of a day to three months provided the highest effect sizes of connectivity (Fig. 4: panel B). This indicates that seasonal turnover (found in Abrego et al. (2024b)) in community composition constrains the time window that can be used to define the community composition of the source areas. In the joint species distribution model, connectivity explained on average half of the variation. This strong impact reflects that the occurrence of a species is controlled by atmospheric connectivity between samples and that the prevalence of species at a site over time depends on the consistency of the atmospheric source area.

From a predictive perspective, connectivity is a better predictor of community turnover than spatial proximity indicating that spores and fungal particles do not distribute evenly across space but along wind corridors, as found previously by Pellitier et al. (2025). This finding was not consistent across temporal windows, as in the shortest and strictest time window (+/- day)

the model with spatial distance provided a higher predictive performance (Fig. 4: panel D). This indicates that such a short time window may not capture the temporal turnover of fungal communities and connectivity dynamics.

The effect of connectivity and land cover on community turnover and species occurrence reflected the ecological differences in relation to trophic guilds and spore size. The community turnover of plant pathogens was less sensitive to sample connectivity than all species on average, whereas the occurrence pattern of wood saprotrophs was slightly less controlled by connectivity than other species on average. These findings could reflect the fact that plant pathogens and wood saprotrophs depend more on dispersal over short geographical distances than all species on average. Another possible explanation is that the communities of plant pathogens are more homogeneous over space and their community structure depends more on random processes. We did find a slightly lower average turnover rate for plant pathogens and wood saprotrophs than for all species (Fig. S2). The importance of land cover in explaining species occurrence patterns increased along with spore size, reflecting the expectation that larger spores are on average more dispersal limited (Norros et al., 2014) and their community assembly thus more strongly structured by dispersal.

The novel data stream of atmospheric dispersion allowed us also to study a fundamental ecological rule, the species-area relationship, universally and in relation to latitudinal regions and seasonality (Fig. 3). We found a positive relationship only in low latitudes ($< 30^\circ$ N/S) reflecting spatial turnover and high beta-diversity. In higher latitudes and especially outside of the growing season, a lack of species-area relationship reflects homogeneous communities and lower beta-diversity. These results align with expectations in Veresoglou and Johnson (2023) and have been previously shown to be a general rule for other species groups (Soininen et al., 2007).

We suspect that non-weighted land cover within a fixed radius can be confounded by other bioclimatic factors that impact species prevalence (Fig. 1). This is reflected by a consistently positive species-area relationship when using non-weighted land cover and by positive response of a few species to water areas in joint species distribution models. Our results indicate that atmospheric dispersion predictions (i.e. footprints) can mitigate the confounding impact of omitted bioclimatic factors on the analysis of metacommunity processes.

Despite the ecologically intuitive results, we have still applied atmospheric dispersion predictions in a solely statistical framework without a mechanistic understanding of spore deposition under natural conditions. For example, spore observations were collected in 1 m height under canopy layer. It is unclear how large a proportion of the spores that disperse above the canopy layer can descend to the ground level. We expect that integrating knowledge from more fine-grained mechanistic models and experiments on spore dispersion could alleviate these concerns. This would be an essential piece of information to further increase the accuracy of the effect of connectivity at distances of hundreds to thousands of kilometres.

Our model framework and findings provide a novel understanding of how long-distance air dispersal constrains the community structure of fungal spores. This study presents a proof-of-concept on the relation of short-term atmospheric dispersion to momentary spore dispersal. In particular, our findings highlight the spatial and temporal scale of dispersal and its importance in comparison to environmental filtering and seasonality. Since air and wind are the main dispersal means for most fungi, quantifying the importance and scale of dispersal for community structuring represents a step change in understanding their metacommunity processes. Especially given the pressure of climate and land cover change on ecosystems, accurate predictions of how, where, and when species can disperse to find suitable habitats are valuable for environmental mitigation strategies.

REFERENCES

- Abarenkov, K., Nilsson, R. H., Larsson, K. H., Taylor, A. F. S., May, T. W., Froslev, T. G., Pawlowska, J., Lindahl, B., Poldmaa, K., Truong, C., Vu, D., Hosoya, T., Niskanen, T., Piirmann, T., Ivanov, F., Zirk, A., Peterson, M., Cheeke, T. E., Ishigami, Y., Jansson, A. T., Jeppesen, T. S., Kristiansson, E., Mikryukov, V., Miller, J. T., Oono, R., Ossandon, F. J., Pauperio, J., Saar, I., Schigel, D., Suija, A., Tedersoo, L., and Koljalg, U. (2024). The unite database for molecular identification and taxonomic communication of fungi and other eukaryotes: sequences, taxa and classifications reconsidered. *Nucleic Acids Res*, 52(D1):D791–D797.
- Abarenkov, K., Somervuo, P., Nilsson, R. H., Kirk, P. M., Huotari, T., Abrego, N., and Ovaskainen, O. (2018). Protax-fungi: a web-based tool for probabilistic taxonomic placement of fungal internal transcribed spacer sequences. *New Phytol*, 220(2):517–525.
- Abrego, N., Crosier, B., Somervuo, P., Ivanova, N., Abrahamyan, A., Abdi, A., Hamalainen, K., Junninen, K., Maunula, M., Purhonen, J., and Ovaskainen, O. (2020). Fungal communities decline with urbanization-more in air than in soil. *ISME J*, 14(11):2806–2815.
- Abrego, N., Furneaux, B., Hardwick, B., Somervuo, P., Palorinne, I., Aguilar, C., Andrew, N., Babiy, U., Bao, T., Bazzano, G., Bondarchuk, S., Bonebrake, T., Brennan, G., Bret-Harte, S., Bässler, C., Cagnolo, L., Cameron, E., Chapurlat, E., Creer, S., D’Acqui, L., de Vere, N., Desprez-Loustau, M.-L., Dongmo, M., Dyrholm Jacobsen, I., Fisher, B., Flores de Jesus, M., Gilbert, G., Griffith, G., Gritsuk, A., Gross, A., Grudd, H., Halme, P., Hanna, R., Hansen, J., Hansen, L. H., Hegbe, A.,

Hill, S., Hogg, I., Hultman, J., Hyde, K., Hynson, N., Ivanova, N., Karisto, P., Kerdraon, D., Knorre, A., Krisai-Greilhuber, I., Kurhinen, J., Kuzmina, M., Lecomte, N., Lecomte, E., Loaiza, V., Lundin, E., Meier, A., Mešić, A., Miettinen, O., Monkhouse, N., Mortimer, P., Müller, J., Nilsson, H., Nonti, P. Y., Nordén, J., Nordén, B., Norros, V., Paz, C., Pellikka, P., Pereira, D., Petch, G., Pitkanen, J.-M., Popa, F., Potter, C., Purhonen, J., Pätsi, S., Rafiq, A., Raharinjanahary, D., Rakos, N., Rathnayaka, A., Raundrup, K., Rebriev, Y., Rikkinen, J., Rogers, H., Rogovsky, A., Rozhkov, Y., Runnel, K., Saarto, A., Savchenko, A., Schlegel, M., Schmidt, N. M., Seibold, S., Skjøth, C., Stengel, E., Sutyryna, S., Syvänperä, I., Tedersoo, L., Timm, J., Tipton, L., Toju, H., Uscka-Perzanowska, M., van der Bank, M., van der Bank, H., Vandenbrink, B., Ventura, S., Vignisson, S., Wang, X., Weisser, W., Wijesinghe, S., Wright, J., Yang, Y., Yorou, N., Young, A., Yu, D., Zakharov, E., Hebert, P., Roslin, T., and Ovaskainen, O. (2024a). Data and scripts for: Airborne dna reveals predictable spatial and seasonal dynamics of fungi.

Abrego, N., Furneaux, B., Hardwick, B., Somervuo, P., Palorinne, I., Aguilar-Trigueros, C. A., Andrew, N. R., Babiy, U. V., Bao, T., Bazzano, G., Bondarchuk, S. N., Bonebrake, T. C., Brennan, G. L., Bret-Harte, S., Bassler, C., Cagnolo, L., Cameron, E. K., Chapurlat, E., Creer, S., D'Acqui, L. P., de Vere, N., Desprez-Loustau, M. L., Dongmo, M. A. K., Jacobsen, I. B. D., Fisher, B. L., Flores de Jesus, M., Gilbert, G. S., Griffith, G. W., Gritsuk, A. A., Gross, A., Grudd, H., Halme, P., Hanna, R., Hansen, J., Hansen, L. H., Hegbe, A., Hill, S., Hogg, I. D., Hultman, J., Hyde, K. D., Hynson, N. A., Ivanova, N., Karisto, P., Kerdraon, D., Knorre, A., Krisai-Greilhuber, I., Kurhinen, J., Kuzmina, M., Lecomte, N., Lecomte, E., Loaiza, V., Lundin, E., Meire, A., Mesic, A., Miettinen, O., Monkhouse, N., Mortimer, P., Muller, J., Nilsson, R. H., Nonti, P. Y. C., Norden, J., Norden, B., Norros, V., Paz, C., Pellikka, P., Pereira, D., Petch, G., Pitkanen, J. M., Popa, F., Potter, C., Purhonen, J., Patsi, S., Rafiq, A., Raharinjanahary, D., Rakos, N., Rathnayaka, A. R., Raundrup, K., Rebriev, Y. A., Rikkinen, J., Rogers, H. M. K., Rogovsky, A., Rozhkov, Y., Runnel, K., Saarto, A., Savchenko, A., Schlegel, M., Schmidt, N. M., Seibold, S., Skjøth, C., Stengel, E., Sutyryna, S. V., Syvanpera, I., Tedersoo, L., Timm, J., Tipton, L., Toju, H., Uscka-Perzanowska, M., van der Bank, M., van der Bank, F. H., Vandenbrink, B., et al. (2024b). Airborne dna reveals predictable spatial and seasonal dynamics of fungi. *Nature*, 631(8022):835–842.

Aguilar-Trigueros, C. A., Krah, F. S., Cornwell, W. K., Zanne, A. E., Abrego, N., Anderson, I. C., Andrew, C. J., Baldrian, P., Bassler, C., Bissett, A., Chaudhary, V. B., Chen, B., Chen, Y., Delgado-Baquerizo, M., Deveau, C., Egidi, E., Flores-Moreno, H., Golan, J., Heilmann-Clausen, J., Hempel, S., Hu, Y., Kausarud, H., Kivlin, S. N., Kohout, P., Lammell, D. R., Maestre, F. T., Pringle, A., Purhonen, J., Singh, B. K., Veresoglou, S. D., Vetrovsky, T., Zhang, H., Rillig, M. C., and Powell, J. R. (2023). Symbiotic status alters fungal eco-evolutionary offspring trajectories. *Ecol Lett*, 26(9):1523–1534.

aguilart (2023). *aguilart/symbiotic-status-and-fungal-spore-size: v1.0.2*.

Calhim, S., Halme, P., Petersen, J. H., Laessoe, T., Bassler, C., and Heilmann-Clausen, J. (2018). Fungal spore diversity reflects substrate-specific deposition challenges. *Sci Rep*, 8(1):5356.

Chaudhary, V. B., Aguilar-Trigueros, C. A., Mansour, I., and Rillig, M. C. (2022). Fungal dispersal across spatial scales. *Annual Review of Ecology, Evolution, and Systematics*, 53(1):69–85.

Cuprewich, S. A., Barbour, K. M., Afkhami, M. E., Lynn, K. M. T., Romero-Olivares, A. L., Aguilar-Trigueros, C., Chaverri, P., Egan, C. P., Norros, V., Peay, K., Ramos, R. J., Stephens, R., Ward, L., and Chaudhary, V. B. (2025). One hundred unanswered questions on the dispersal ecology of fungi. *ISME J*, 20(1).

Davison, J., Moora, M., Öpik, M., Adholeya, A., Ainsaar, L., Bâ, A., Burla, S., Diedhiou, A., Hiiesalu, I., and Jairus, T. (2015). Global assessment of arbuscular mycorrhizal fungus diversity reveals very low endemism. *Science*, 349(6251):970–973.

Delelegn, Y. T., Puraong, W., Sanden, H., Yitaferu, B., Godbold, D. L., and Wubet, T. (2018). Transition of ethiopian highland forests to agriculture-dominated landscapes shifts the soil microbial community composition. *BMC Ecol*, 18(1):58.

Finlay, B. J. (2002). Global dispersal of free-living microbial eukaryote species. *science*, 296(5570):1061–1063.

Floyd, R. W. (1962). Algorithm 97: shortest path. *Communications of the ACM*, 5(6):345–345.

Furneaux, B., Anslan, S., Somervuo, P., Hultman, J., Abrego, N., Roslin, T., and Ovaskainen, O. (2025). Optimotu: Taxonomically aware otu clustering with optimized thresholds and a bioinformatics workflow for metabarcoding data. *arXiv preprint arXiv:2502.10350*.

Gelman, A. and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical science*, 7(4):457–472.

Golan, J. J., Lagomarsino Oneto, D., Ding, S., Kessenich, R., Sandler, M., Rush, T. A., Levitis, D., Gevens, A., Seminara, A., and Pringle, A. (2023). Differences in spore size and atmospheric survival shape stark contrasts in the dispersal dynamics of two closely related fungal pathogens. *Fungal Ecology*, 66.

Green, J. L., Holmes, A. J., Westoby, M., Oliver, I., Briscoe, D., Dangerfield, M., Gillings, M., and Beattie, A. J. (2004). Spatial scaling of microbial eukaryote diversity. *Nature*, 432(7018):747–750.

He, S., Wang, B., Chen, K., Li, N., and Soininen, J. (2024). Species–environment sorting explains latitudinal patterns in spatiotemporal -diversity for freshwater macroinvertebrates. *Ecography*, 2024(9).

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., and

- Schepers, D. (2020). The era5 global reanalysis. *Quarterly journal of the royal meteorological society*, 146(730):1999–2049.
- Hussein, T., Norros, V., Hakala, J., Petäjä, T., Aalto, P. P., Rannik, , Vesala, T., and Ovaskainen, O. (2013). Species traits and inertial deposition of fungal spores. *Journal of Aerosol Science*, 61:81–98.
- information, E. U. C. L. M. S. (2020).
- Lagomarsino Oneto, D., Golan, J., Mazzino, A., Pringle, A., and Seminara, A. (2020). Timing of fungal spore release dictates survival during atmospheric transport. *Proc Natl Acad Sci U S A*, 117(10):5134–5143.
- Massicotte, P. and South, A. (2026). *rnaturalearth: World Map Data from Natural Earth*. R package version 1.2.0.9000.
- Meinander, O., Kontu, A., Kouznetsov, R., and Sofiev, M. (2020). Snow samples combined with long-range transport modeling to reveal the origin and temporal variability of black carbon in seasonal snow in sodankylä (67°n). *Frontiers in Earth Science*, 8.
- Naranjo-Orrico, D., Ovaskainen, O., Furneaux, B., Purhonen, J., Arancibia, P. A., Burg, S., Moser, N., Niku, J., Tikhonov, G., Zakharov, E., Monkhouse, N., and Abrego, N. (2025). Wind is a primary driver of fungal dispersal across a mainland-island system. *Mol Ecol*, page e17675.
- Norros, V., Halme, P., Norberg, A., and Ovaskainen, O. (2023). Spore production monitoring reveals contrasting seasonal strategies and a trade-off between spore size and number in wood-inhabiting fungi. *Functional Ecology*, 37(3):551–563.
- Norros, V., Karhu, E., Norden, J., Vahatalo, A. V., and Ovaskainen, O. (2015). Spore sensitivity to sunlight and freezing can restrict dispersal in wood-decay fungi. *Ecol Evol*, 5(16):3312–26.
- Norros, V., Penttilä, R., Suominen, M., and Ovaskainen, O. (2012). Dispersal may limit the occurrence of specialist wood decay fungi already at small spatial scales. *Oikos*, 121(6):961–974.
- Norros, V., Rannik, U., Hussein, T., Petaja, T., Vesala, T., and Ovaskainen, O. (2014). Do small spores disperse further than large spores? *Ecology*, 95(6):1612–21.
- Ovaskainen, O., Abrego, N., Furneaux, B., Hardwick, B., Somervuo, P., Palorinne, I., Andrew, N., Babiy, U., Bao, T., Bazzano, G., Bondarchuk, S., Bonebrake, T., Brennan, G., Bret-Harte, S., Bässler, C., Cagnolo, L., Cameron, E., Chapurlat, E., Creer, S., D’Acqui, L., de Vere, N., Desprez-Loustau, M.-L., Dongmo, M., Dyrholm Jacobsen, I., Fisher, B., Flores de Jesus, M., Gilbert, G., Griffith, G., Gritsuk, A., Gross, A., Grudd, H., Halme, P., Hanna, R., Hansen, J., Hansen, L. H., Hegbe, A., Hill, S., Hogg, I., Hultman, J., Hyde, K., Hynson, N., Ivanova, N., Karisto, P., Kerdraon, D., Knorre, A., Krisai-Greilhuber, I., Kurhinen, J., Kuzmina, M., Lecomte, N., Lecomte, E., Loaiza, V., Lundin, E., Meier, A., Mešić, A., Miettinen, O., Monkhouse, N., Mortimer, P., Müller, J., Nilsson, H., Nonti, P. Y., Nordén, J., Nordén, B., Paz, C., Pellikka, P., Pereira, D., Petch, G., Pitkänen, J.-M., Popa, F., Potter, C., Purhonen, J., Patsi, S., Rafiq, A., Raharinjanahary, D., Rakos, N., Rathnayaka, A., Raundrup, K., Rebriev, Y., Rikkinen, J., Rogers, H., Rogovsky, A., Rozhkov, Y., Runnel, K., Saarto, A., Savchenko, A., Schlegel, M., Schmidt, N. M., Seibold, S., Skjøth, C., Stengel, E., Sutyryna, S., Syvänperä, I., Tedersoo, L., Timm, J., Tipton, L., Toju, H., Uscka-Perzanowska, M., van der Bank, M., van der Bank, H., Vandenbrink, B., Ventura, S., Vignissin, S., Wang, X., Weisser, W., Wijesinghe, S., Wright, J., Yang, Y., Yorou, N., Young, A., Yu, D., Zakharov, E., Hebert, P., and Roslin, T. (2024a). Data from: Global spore sampling project: A global, standardized dataset of airborne fungal dna.
- Ovaskainen, O., Abrego, N., Furneaux, B., Hardwick, B., Somervuo, P., Palorinne, I., Andrew, N. R., Babiy, U. V., Bao, T., Bazzano, G., Bondarchuk, S. N., Bonebrake, T. C., Brennan, G. L., Bret-Harte, S., Bassler, C., Cagnolo, L., Cameron, E. K., Chapurlat, E., Creer, S., D’Acqui, L. P., de Vere, N., Desprez-Loustau, M. L., Dongmo, M. A. K., Dyrholm Jacobsen, I. B., Fisher, B. L., Flores de Jesus, M., Gilbert, G. S., Griffith, G. W., Gritsuk, A. A., Gross, A., Grudd, H., Halme, P., Hanna, R., Hansen, J., Hansen, L. H., Hegbe, A., Hill, S., Hogg, I. D., Hultman, J., Hyde, K. D., Hynson, N. A., Ivanova, N., Karisto, P., Kerdraon, D., Knorre, A., Krisai-Greilhuber, I., Kurhinen, J., Kuzmina, M., Lecomte, N., Lecomte, E., Loaiza, V., Lundin, E., Meire, A., Mesic, A., Miettinen, O., Monkhouse, N., Mortimer, P., Muller, J., Nilsson, R. H., Nonti, P. Y. C., Norden, J., Norden, B., Paz, C., Pellikka, P., Pereira, D., Petch, G., Pitkanen, J. M., Popa, F., Potter, C., Purhonen, J., Patsi, S., Rafiq, A., Raharinjanahary, D., Rakos, N., Rathnayaka, A. R., Raundrup, K., Rebriev, Y. A., Rikkinen, J., Rogers, H. M. K., Rogovsky, A., Rozhkov, Y., Runnel, K., Saarto, A., Savchenko, A., Schlegel, M., Schmidt, N. M., Seibold, S., Skjoth, C., Stengel, E., Sutyryna, S. V., Syvanpera, I., Tedersoo, L., Timm, J., Tipton, L., Toju, H., Uscka-Perzanowska, M., van der Bank, M., Herman van der Bank, F., Vandenbrink, B., Ventura, S., et al. (2024b). Global spore sampling project: A global, standardized dataset of airborne fungal dna. *Sci Data*, 11(1):561.
- Ovaskainen, O., Tikhonov, G., Norberg, A., Guillaume Blanchet, F., Duan, L., Dunson, D., Roslin, T., and Abrego, N. (2017). How to make more out of community data? a conceptual framework and its implementation as models and software. *Ecol Lett*, 20(5):561–576.
- Peay, K. G., Bruns, T. D., Kennedy, P. G., Bergemann, S. E., and Garbelotto, M. (2007). A strong species-area relationship for eukaryotic soil microbes: island size matters for ectomycorrhizal fungi. *Ecol Lett*, 10(6):470–80.
- Peay, K. G., Kennedy, P. G., and Talbot, J. M. (2016). Dimensions of biodiversity in the earth mycobiome. *Nat Rev Microbiol*, 14(7):434–47.
- Pellitier, P. T., Kling, M. M., Qin, C., Van Nuland, M. E., Zhu, K., and Peay, K. G. (2025). Wind patterns influence the dispersal

and assembly of north american soil fungal communities. *Ecol Lett*, 28(5):e70130.

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

Rahman, A. U., Tikhonov, G., Oksanen, J., Rossi, T., and Ovaskainen, O. (2024). Accelerating joint species distribution modelling with hmsc-hpc by gpu porting. *PLoS Comput Biol*, 20(9):e1011914.

Robert, V., Vu, D., Amor, A. B. H., Van De Wiele, N., Brouwer, C., Jabas, B., Szoke, S., Dridi, A., Triki, M., and Daoud, S. b. (2013a). Mycobank gearing up for new horizons. *IMA fungus*, 4(2):371–379.

Robert, V., Vu, D., Amor, A. B. H., Van De Wiele, N., Brouwer, C., Jabas, B., Szoke, S., Dridi, A., Triki, M., and Daoud, S. b. (2013b). Mycobank gearing up for new horizons. *IMA fungus*, 4(2):371–379.

Row, J. R., Knick, S. T., Oyler-McCance, S. J., Lougheed, S. C., and Fedy, B. C. (2017). Developing approaches for linear mixed modeling in landscape genetics through landscape-directed dispersal simulations. *Ecol Evol*, 7(11):3751–3761.

Schulz, T., Saastamoinen, M., and Vanhatalo, J. (2025). Model-based variance partitioning for statistical ecology. *Ecological Monographs*, 95(1).

Siljamo, P., Sofiev, M., Severova, E., Ranta, H., Kukkonen, J., Polevova, S., Kubin, E., and Minin, A. (2008). Sources, impact and exchange of early-spring birch pollen in the moscow region and finland. *Aerobiologia*, 24(4):211–230.

Skjoth, C. A., Sommer, J., Brandt, J., Hvidberg, M., Geels, C., Hansen, K. M., Hertel, O., Frohn, L. M., and Christensen, J. H. (2008). Copenhagen—a significant source of birch (betula) pollen? *Int J Biometeorol*, 52(6):453–62.

Sofiev, M. (2000). A model for the evaluation of long-term airborne pollution transport at regional and continental scales. *Atmospheric Environment*, 34(15):2481–2493.

Sofiev, M. (2017). On impact of transport conditions on variability of the seasonal pollen index. *Aerobiologia (Bologna)*, 33(1):167–179.

Sofiev, M. (2019). On possibilities of assimilation of near-real-time pollen data by atmospheric composition models. *Aerobiologia*, 35(3):523–531.

Sofiev, M., Buters, J., Tummon, F., Fatahi, Y., Sozinova, O., Adams-Groom, B., Bergmann, K. C., Dahl, A., Gehrig, R., Gilge, S., Seliger, A. K., Kouznetsov, R., Lieberherr, G., O'Connor, D., Oteros, J., Palamarchuk, J., Ribeiro, H., Werchan, B., Werchan, M., and Clot, B. (2023). Designing an automatic pollen monitoring network for direct usage of observations to reconstruct the concentration fields. *Sci Total Environ*, 900:165800.

Sofiev, M., Palamarchuk, J., Kouznetsov, R., Abramidze, T., Adams-Groom, B., Antunes, C. M., Arino, A. H., Bastl, M., Belmonte, J., Berger, U. E., Bonini, M., Bruffaerts, N., Buters, J., Carinanos, P., Celenk, S., Ceriotti, V., Charalampopoulos, A., Clewlow, Y., Clot, B., Dahl, A., Damialis, A., De Linares, C., De Weger, L. A., Dirr, L., Ekeboom, A., Fatahi, Y., Fernandez Gonzalez, M., Fernandez Gonzalez, D., Fernandez-Rodriguez, S., Galan, C., Gedda, B., Gehrig, R., Geller Bernstein, C., Gonzalez Roldan, N., Grewling, L., Hajkova, L., Hanninen, R., Hentges, F., Jantunen, J., Kadantsev, E., Kasprzyk, I., Kloster, M., Kluska, K., Koenders, M., Laffersova, J., Leru, P. M., Lipiec, A., Louna-Korteniemi, M., Magyar, D., Majkowska-Wojciechowska, B., Makela, M., Mitrovic, M., Myszkowska, D., Oliver, G., Ostensson, P., Perez-Badia, R., Piotrowska-Weryszko, K., Prank, M., Przedpelska-Wasowicz, E. M., Patsi, S., Rajo, F. J. R., Ramfjord, H., Rapiejko, J., Rodinkova, V., Rojo, J., Ruiz-Valenzuela, L., Rybnicek, O., Saarto, A., Sauliene, I., Seliger, A. K., Severova, E., Shalaboda, V., Sikoparija, B., Siljamo, P., Soares, J., Sozinova, O., Stangel, A., Stjepanovic, B., Teinmaa, E., Tyuryakov, S., Trigo, M. M., Uppstu, A., Vill, M., Vira, J., Visez, N., Vitikainen, T., Vokou, D., Weryszko-Chmielewska, E., and Karppinen, A. (2024). European pollen reanalysis, 1980–2022, for alder, birch, and olive. *Sci Data*, 11(1):1082.

Sofiev, M., Siljamo, P., Valkama, I., Ilvonen, M., and Kukkonen, J. (2006). A dispersion modelling system silam and its evaluation against etex data. *Atmospheric Environment*, 40(4):674–685.

Sofiev, M., Sofieva, S., Palamarchuk, J., Sauliene, I., Kadantsev, E., Atanasova, N., Fatahi, Y., Kouznetsov, R., Kuula, J., Noreikaite, A., Peltonen, M., Pihlajamaki, T., Saarto, A., Svirskaitė, J., Toiviainen, L., Tyuryakov, S., Sukiene, L., Asmi, E., Bamford, D., Hyvarinen, A. P., and Karppinen, A. (2022). Bioaerosols in the atmosphere at two sites in northern europe in spring 2021: Outline of an experimental campaign. *Environ Res*, 214(Pt 2):113798.

Sofiev, M., Vira, J., Kouznetsov, R., Prank, M., Soares, J., and Genikhovich, E. (2015). Construction of the silam eulerian atmospheric dispersion model based on the advection algorithm of michael galperin. *Geoscientific Model Development*, 8(11):3497–3522.

Soininen, J. and Hillebrand, H. (2007). Disentangling distance decay of similarity from richness gradients: response to baselga (2007). *Ecography*, 30(6):842–844.

Soininen, J., Lennon, J. J., and Hillebrand, H. (2007). A multivariate analysis of beta diversity across organisms and environments. *Ecology*, 88(11):2830–8.

Talbot, J. M., Bruns, T. D., Taylor, J. W., Smith, D. P., Branco, S., Glassman, S. I., Erlandson, S., Vilgalys, R., Liao, H. L., Smith, M. E., and Peay, K. G. (2014). Endemism and functional convergence across the north american soil mycobiome. *Proc Natl Acad Sci U S A*, 111(17):6341–6.

- 489 Tedersoo, L., Bahram, M., Põlme, S., Kõljalg, U., Yorou, N. S., Wijesundera, R., Ruiz, L. V., Vasco-Palacios, A. M., Thu, P. Q.,
490 and Suija, A. (2014). Global diversity and geography of soil fungi. *science*, 346(6213):1256688.
- 491 Tikhonov, G., Opedal, O. H., Abrego, N., Lehtikainen, A., de Jonge, M. M. J., Oksanen, J., and Ovaskainen, O. (2020). Joint
492 species distribution modelling with the r-package hmsc. *Methods Ecol Evol*, 11(3):442–447.
- 493 Vellend, M. (2010). Conceptual synthesis in community ecology. *Q Rev Biol*, 85(2):183–206.
- 494 Verchot, L. V. (2010). *Impacts of forest conversion to agriculture on microbial communities and microbial function*, pages
495 45–63. Springer.
- 496 Veresoglou, S. D. and Johnson, D. (2023). Species-area relationships in microbial-mediated mutualisms. *Trends Microbiol*,
497 31(11):1111–1117.
- 498 Veriankaitė, L., Siljamo, P., Sofiev, M., Šaulienė, I., and Kukkonen, J. (2009). Modelling analysis of source regions of
499 long-range transported birch pollen that influences allergenic seasons in lithuania. *Aerobiologia*, 26(1):47–62.
- 500 Verstraeten, W. W., Kouznetsov, R., Bruffaerts, N., Sofiev, M., and Delcloo, A. W. (2025). Analyzing the airborne birch and
501 grass pollen monitoring network of belgium. *Aerobiologia*, 41(2):211–227.
- 502 Vetrovsky, T., Kohout, P., Kopecky, M., Machac, A., Man, M., Bahnmann, B. D., Brabcova, V., Choi, J., Meszarosova, L.,
503 Human, Z. R., Lepinay, C., Llado, S., Lopez-Mondejar, R., Martinovic, T., Masinova, T., Morais, D., Navratilova, D.,
504 Odriozola, I., Stursova, M., Svec, K., Tlaskal, V., Urbanova, M., Wan, J., Zifcakova, L., Howe, A., Ladau, J., Peay, K. G.,
505 Storch, D., Wild, J., and Baldrian, P. (2019). A meta-analysis of global fungal distribution reveals climate-driven patterns.
506 *Nat Commun*, 10(1):5142.
- 507 Wang, M., Kriticos, D. J., Ota, N., Brooks, A., and Paine, D. (2021). A general trait-based modelling framework for revealing
508 patterns of airborne fungal dispersal threats to agriculture and native flora. *New Phytol*, 232(3):1506–1518.

CODE AND DATA AVAILABILITY

Data and code are shared with the reviewers as a repository that will be published in Zenodo after the manuscript has been accepted.

Land cover data was derived from the Copernicus Global Land Service: Land Cover 100m: collection 3: epoch: 2019 (Marcel Buchhorn, Bruno Smets, Luc Bertels, Bert De Roo, Myroslava Lesiv, Nandin-Erdene Tsendbazar, Martin Herold, & Steffen Fritz. (2020). Copernicus Global Land Service: Land Cover 100m: collection 3: epoch 2019: Globe (V3.0.1) [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.3939050>)

For a requested citation format: this publication has been prepared using European Union's Copernicus Land Monitoring Service information; <https://doi.org/10.2909/c6377c6e-76cc-4d03-8330-628a03693042>.

For data above 80°N and below 60°S, land-ocean classification was done with dissolved land cover data from rnatuarearth R-package (Massicotte and South, 2026)

Species, trait and environmental data: was derived from the data repository (Abrego et al., 2024a) (Data and scripts for: Airborne DNA reveals predictable spatial and seasonal dynamics of fungi. <https://doi.org/10.5281/zenodo.10896659>). The dataset consists of spore occurrence data from Ovaskainen et al. (2024a) (Data from: Global Spore Sampling Project: A global, standardized dataset of airborne fungal DNA. <https://doi.org/10.5281/zenodo.11125610>), climate data from the Copernicus Climate Change Service Climate Data Store; 'ERA5 hourly data on single levels dataset' and 'sis biodiversity era5 global dataset' (Hersbach et al., 2020) ('ERA5 hourly data on single levels dataset' and 'sis biodiversity era5 global dataset'. <https://cds.climate.copernicus.eu/datasets/reanalysis-era5-land-timeseries>) and spore size and trophic guild data extracted by aguilar (2023) (aguilar/Symbiotic-status-and-fungal-spore-size: v1.0.2. <https://doi.org/10.5281/zenodo.7954364>). Spore size data has been originally assembled from Mycobank (mycobank.org) (Robert et al., 2013b).

Table 1. Summary of hypothesis, parametric tests and results.

Hypothesis #	Hypothesis explanation	Parameter test	Result
H1.1	Land area in spore source area increases species richness in spore sample (species-area relationship).	Correlation between log-transformed species richness and log-transformed land area.	Positive species-area relationship in low latitudes but not in high latitudes.
H1.2	Community composition in spore samples that are connected through air dispersion are more similar than samples which are not connected and the strength of connectivity increases community similarity.	Effect size estimates for the connectivity status (0/1) and the strength of connectivity in a linear mixed effect model explaining the Jaccard turnover index of the community composition.	Relatively large negative effect of connectivity on turnover.
H1.3	Land cover characteristics in the spore source area of sampling sites modify spore community composition.	Effect size estimate for the change in land cover in a linear model explaining the Jaccard turnover index of the community composition.	Land cover has a relatively small effect on turnover.
H1.4	The effect of connectivity depends on the definition of connectivity in relation to temporal window.	Comparison of models for community turnover. Models differ in relation to the definition of connectivity.	Connectivity has the largest effect on turnover when temporal window is smaller than a year but the importance of connectivity in variance partitioning increases along with longer time window.
H1.5	Connectivity is a better predictor than spatial distance for predicting community dissimilarity.	Comparison between linear mixed effect models that explain Jaccard turnover index of the community composition with connectivity (binary and continuous variables) or spatial distance.	Footprint model has a smaller root mean squared error of model predictions in four-fold cross-validation than a spatial model has.
H2.1	Land cover characteristics in the spore source area and connectivity affect the presence probability of single species.	Effect size estimate and variance partitioning of land cover variables and connectivity related random effect in a joint species distribution model.	Land cover explains on average less than 10% whereas connectivity related random effect explains 50% of the occurrence patterns.
H2.2	Habitat specialists such as wood saprotrophs and species with large spores are more sensitive to land cover characteristics in the source area and connectivity than species on average.	Correlation between the importance of land cover variables and connectivity in variance partitioning and the species trophic guild and spore size.	Land cover explains larger part of variation for plant pathogens than other species and species that have large spores. Connectivity explains smaller part of variation for wood saprotrophs than other species.

FIGURES

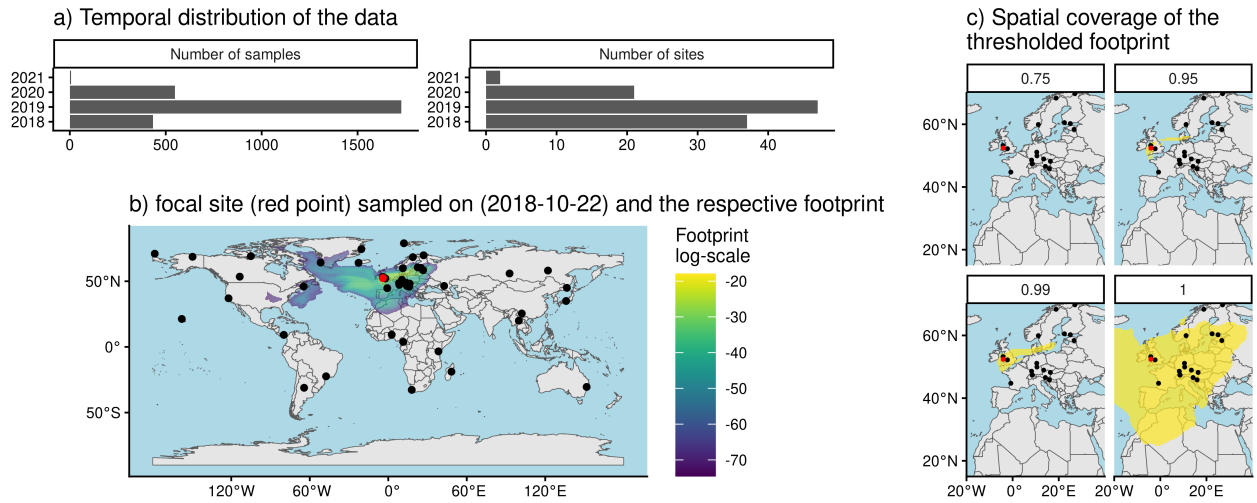


Figure 1. Panel A shows the number samples and the number of sampled sites per year. Panel B shows the geographical locations of the sampling sites and a footprint related to one spore sample. Footprint concentration is shown on a log scale. Panel C shows the spatial distribution of the discretized footprint, where the footprint is set to cover the top 75%, 95%, 99% or 100% of the total footprint mass.

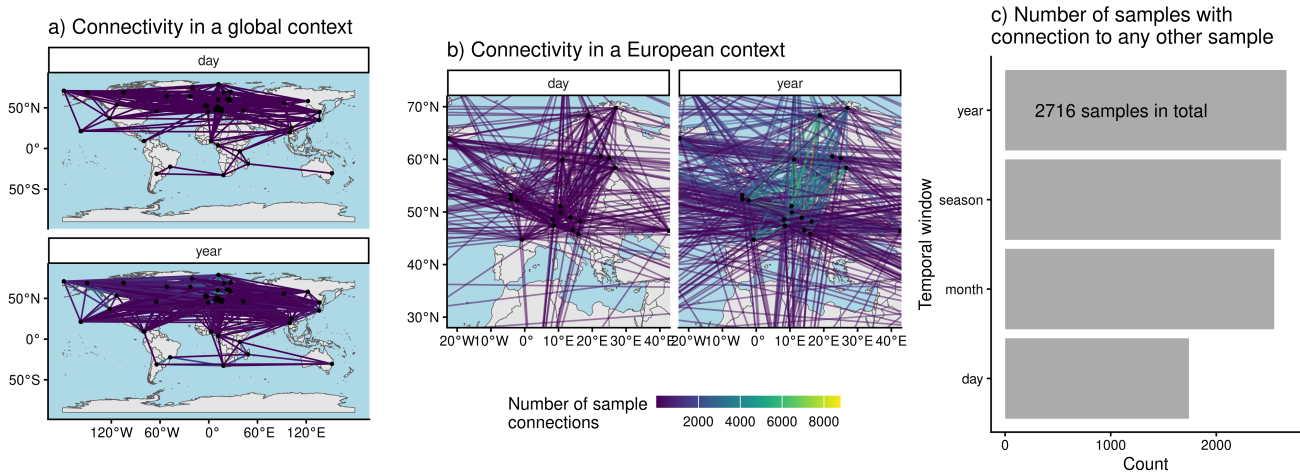


Figure 2. Panels A and B show the number of sample level connections per each site-pair in a global and European extent. Panel C shows the number of samples which are connected to any other sample given different temporal windows. Samples from the same site are by default considered as non-connected.

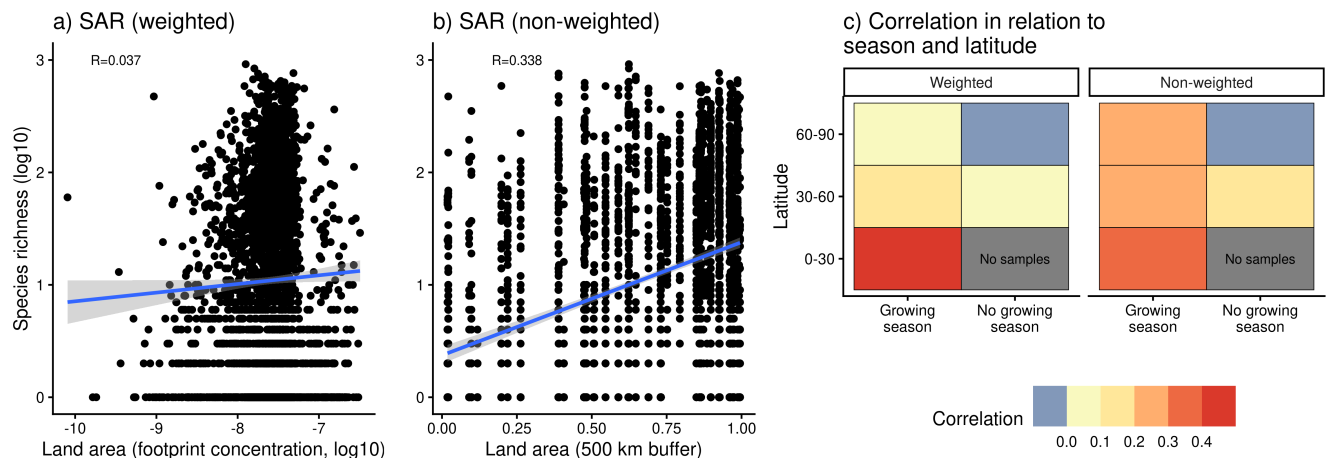


Figure 3. Panels A and B show the species-area relationship (SAR), where log-transformed species richness of fungal spores is associated with log-transformed footprint concentration that originates from land area (panel A; weighted model) and proportion of land cover within a 500 km radius around sampling site (B; non-weighted model). Panel C shows the respective species-area relationship for scenarios, where the sampling sites are sub-sampled in relation to latitude and growing season.

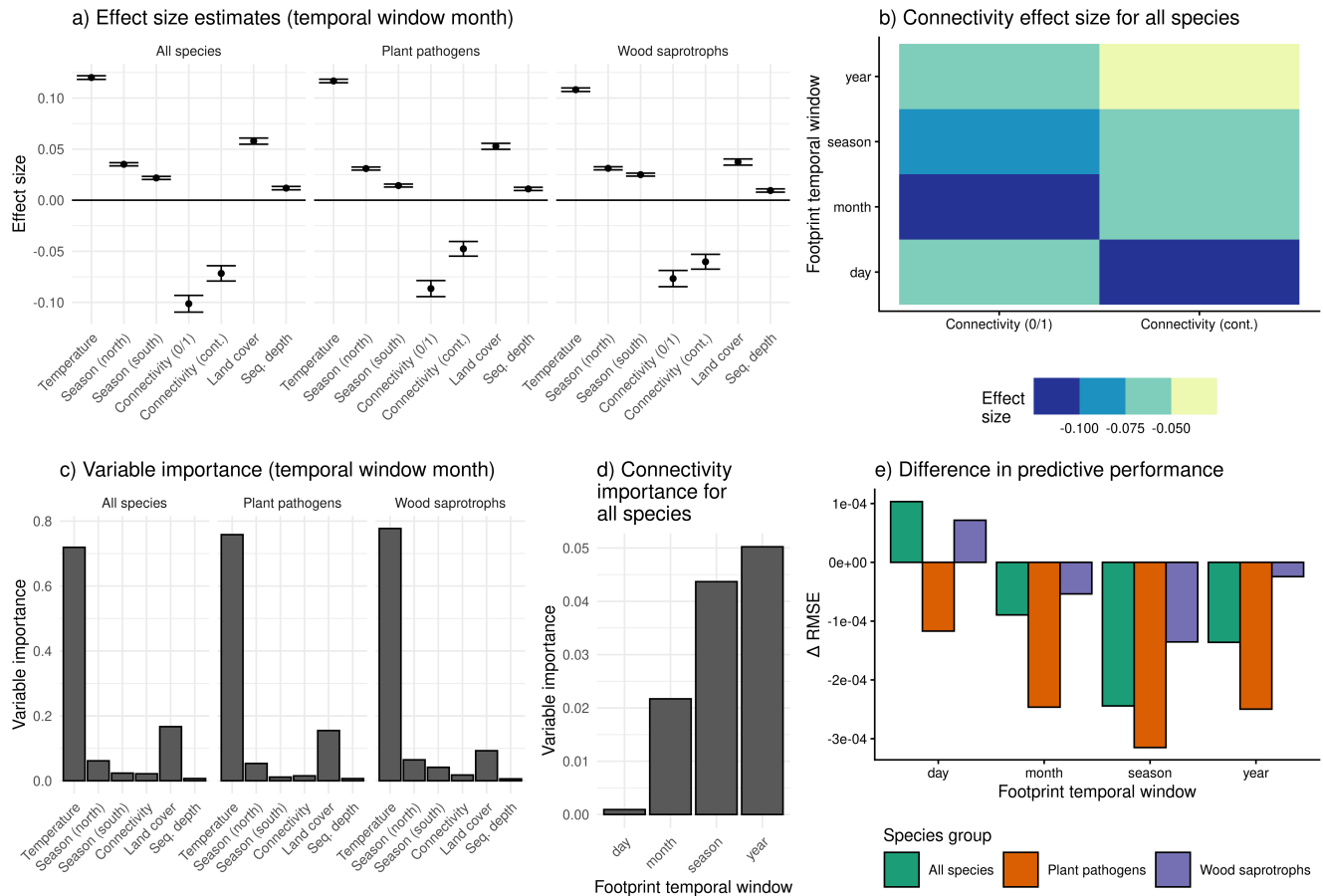


Figure 4. Covariate effect sizes in a linear model using connectivity among other explanatory variables explaining community turnover between sample pairs. Panels A summarizes the effect sizes (dots) and the associated 95 % credibility intervals (error bars) that are estimated with the linear models and Jaccard index of community turnover as the response variable. The effect sizes are shown for all study species and for plant pathogens and wood saprotrophs. Panel B summarizes the estimated effect sizes for the connectivity variables given different temporal constraints of the connectivity. Panel C shows the variable importance of the covariates for different species groups. Note that the importance of binary and continuous connectivity effects are summed together. Panel D shows the importance of connectivity for all species given different temporal constraints of the connectivity. Panel E summarizes the difference in the predictive performance between models for which connectivity is defined through footprint overlap or geographical distance. Negative values denote a smaller root mean squared error (higher accuracy) for the footprint informed model.

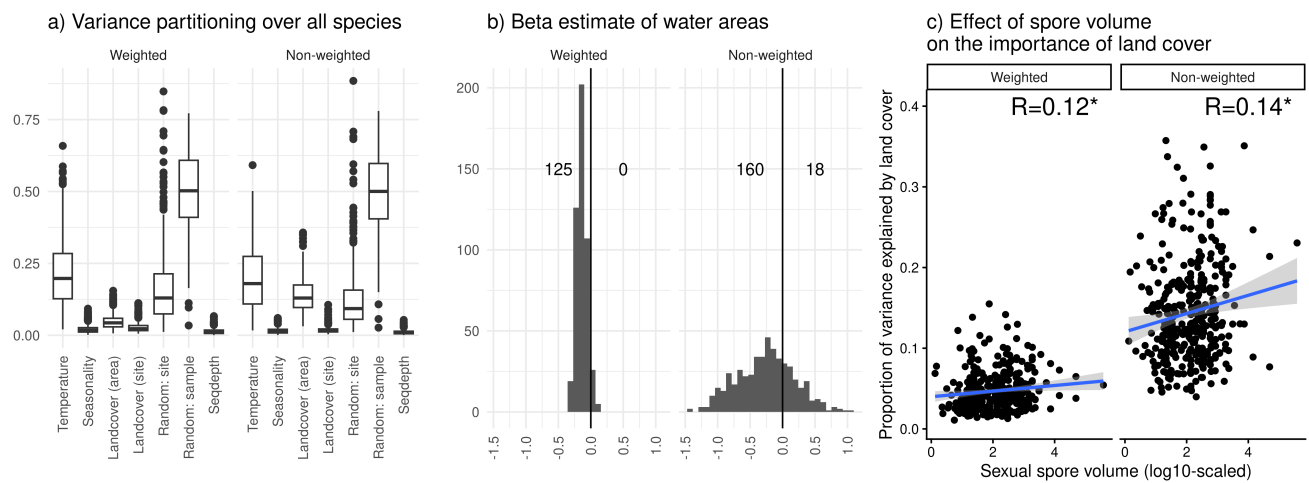


Figure 5. Results from the joint species distribution models, in which land cover from the surrounding area has been either weighted or not weighted with a footprint. In panel A, boxplots show the importance of different variables in the model across all species. In Panel B, histograms show the distribution of the species-specific responses to the coverage of water areas. Panel C shows the association between sexual spore size (x-axis) and the relative importance of land cover (area) in explaining species occurrence pattern. Blue line denotes a linear association and the grey shading denotes a 95 % credibility interval of the association. The asterisk in the top right corner denotes whether the linear association is not zero with the 95% credibility.

Supporting information: *The role of long-distance dispersal for community dynamics of fungi*

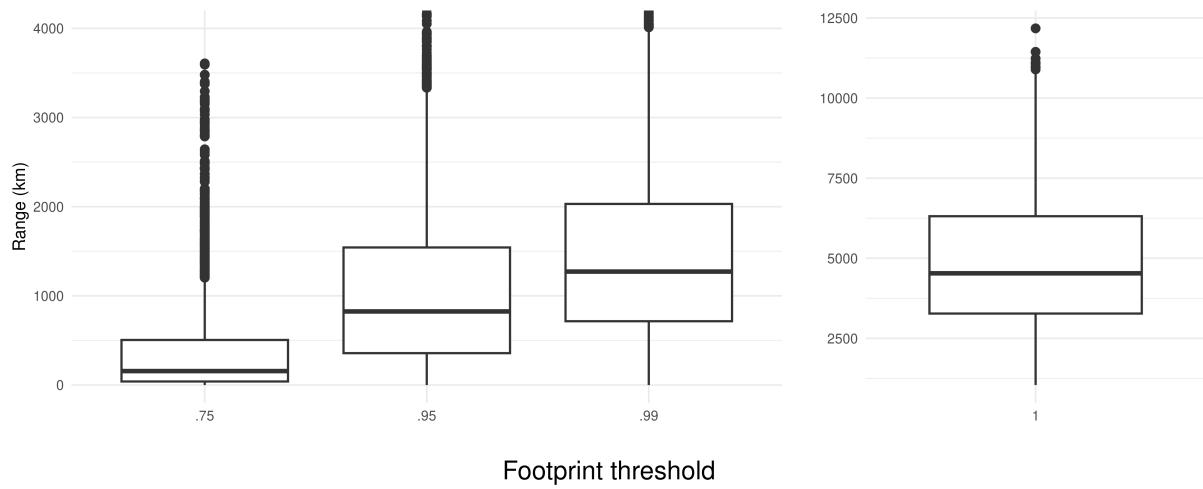


Figure 1: Max. distance from sampling site to footprint edge.

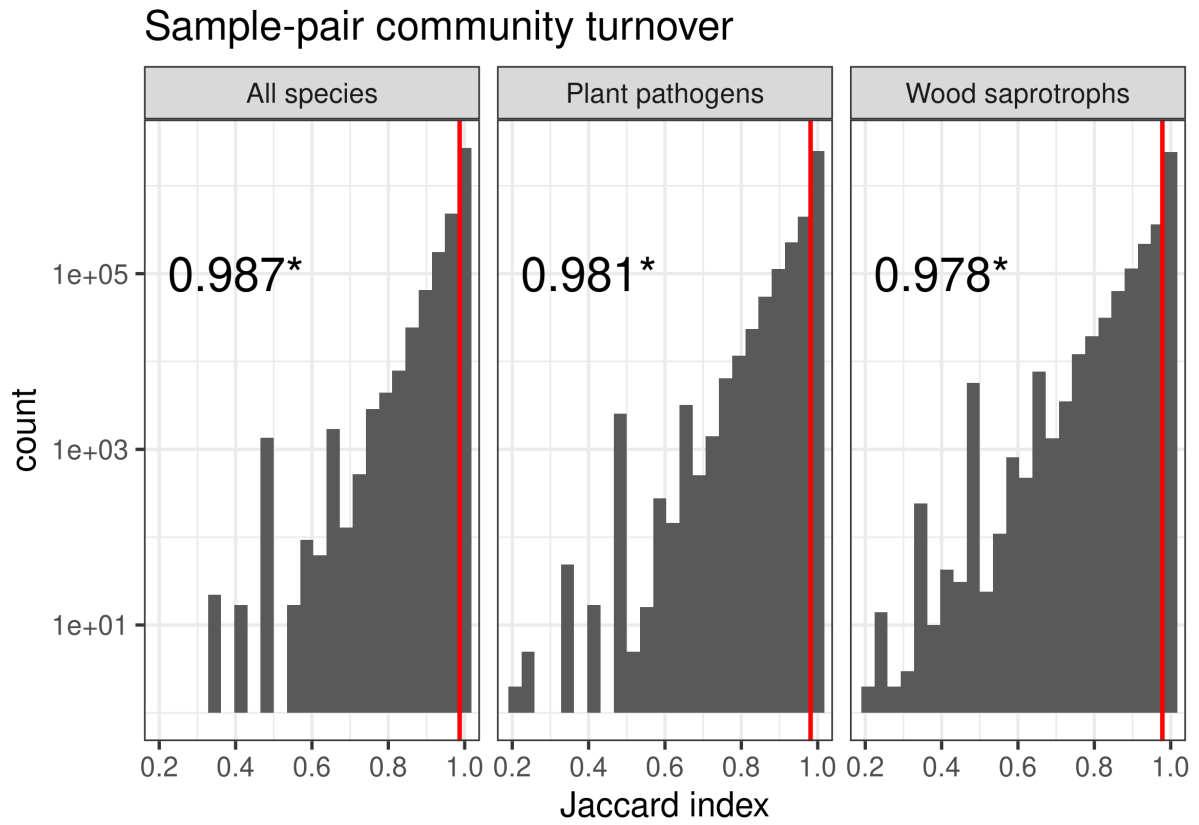


Figure 2: Empirical distribution of sample-to-sample community turnover. Asterisks indicate whether distributions have statistically significantly different population mean values ($p < 0.05$).

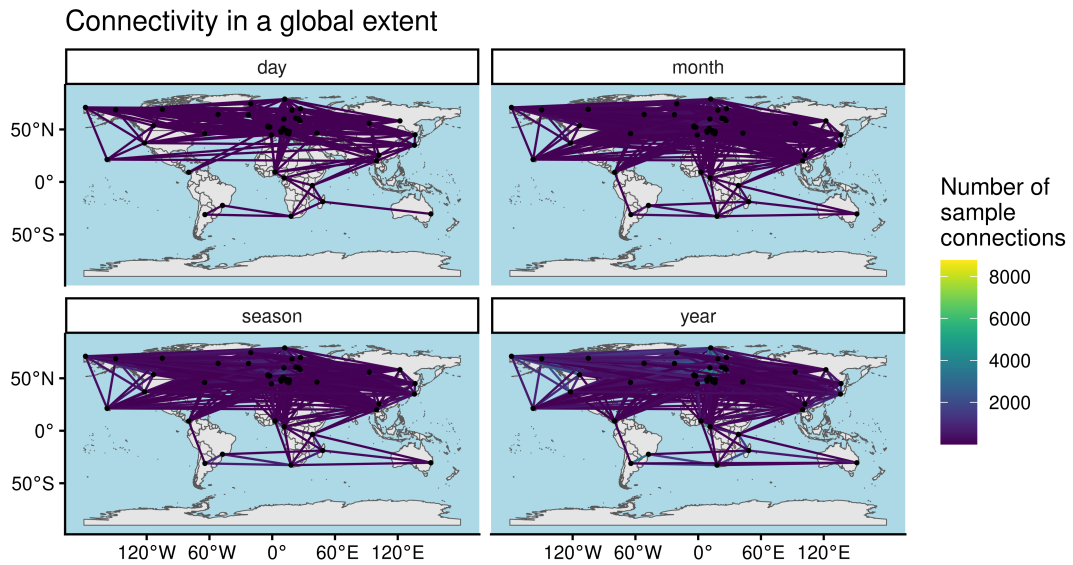


Figure 3: Connectivity in a global extent given different temporal constraints of the connectivity.

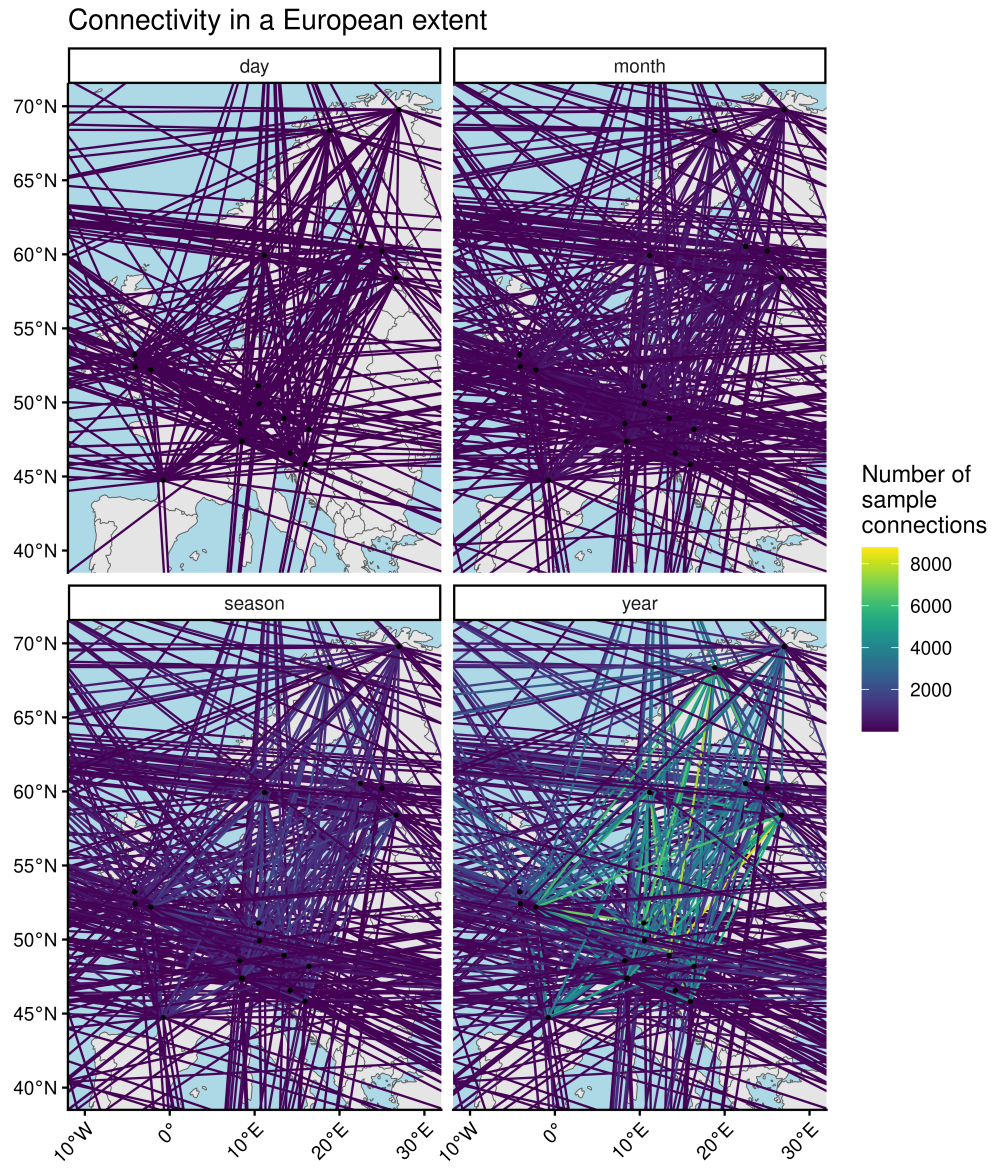


Figure 4: Connectivity in a European extent given different temporal constraints of the connectivity.

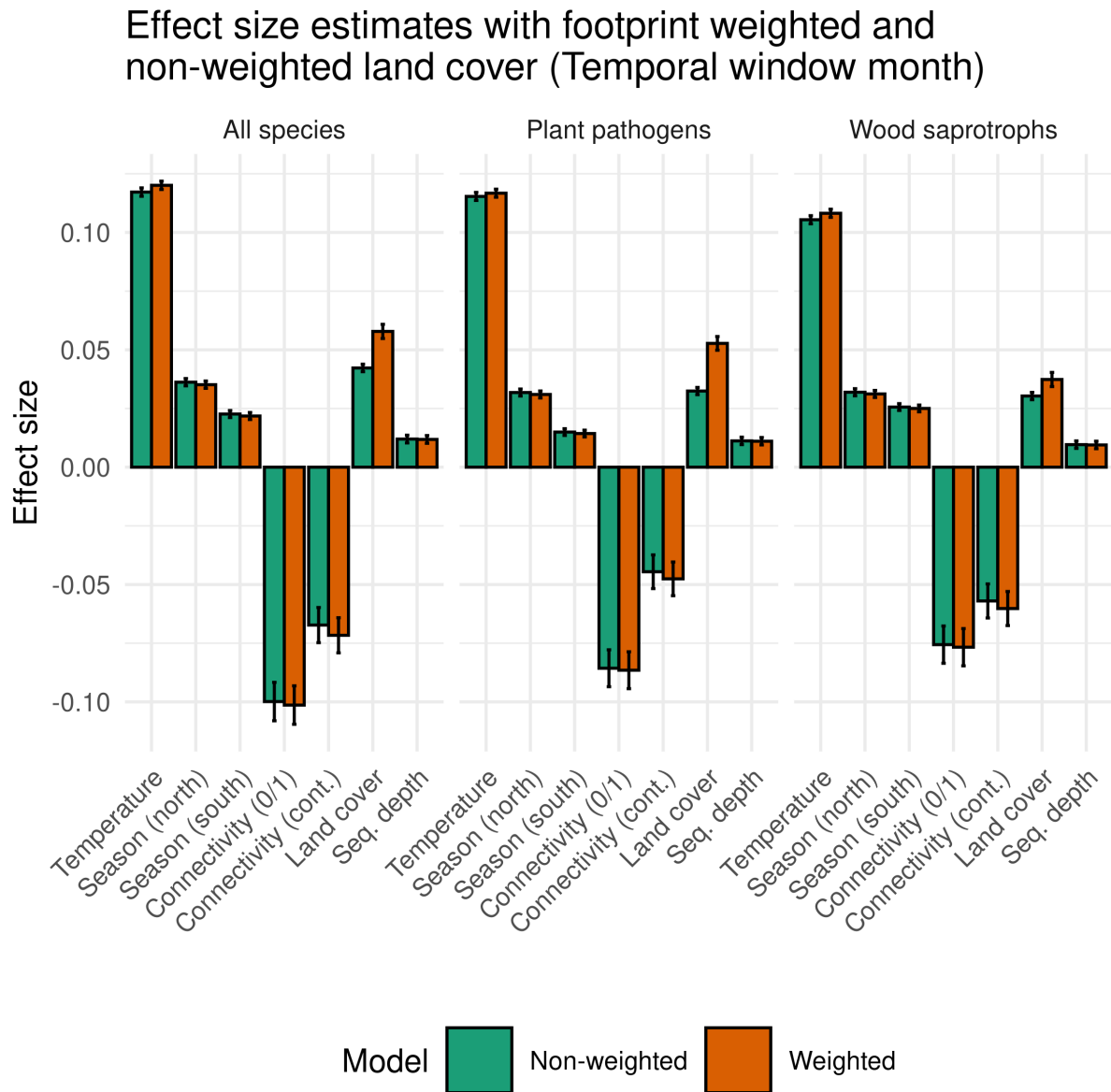


Figure 5: Effect sizes (bars) and the associated 95 % credibility intervals (error bars) for two linear models, in which land cover from the surrounding area has been weighted or not weighted with a footprint.

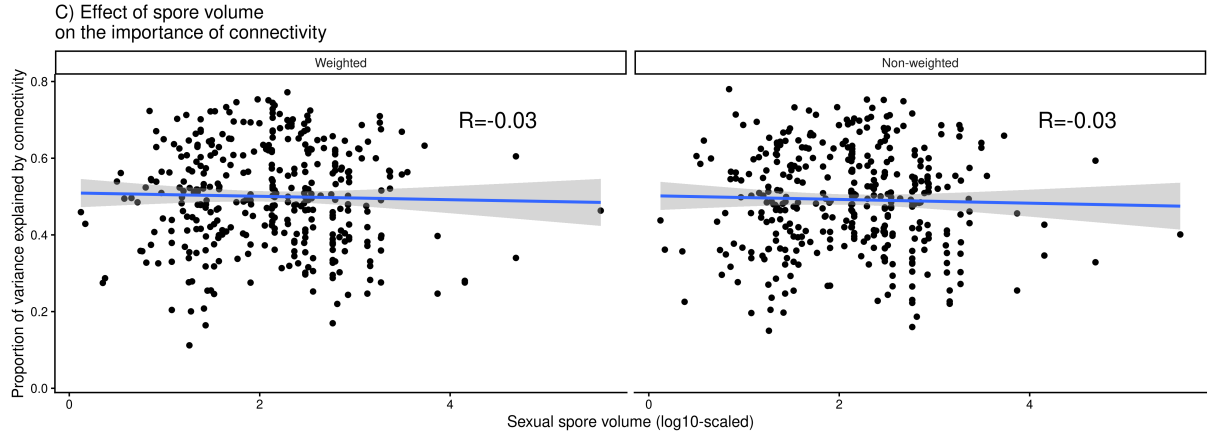


Figure 6: Association between sexual spore size (x-axis) and the relative importance of connectivity (random: sample) in explaining species occurrence pattern. Blue line denotes a linear association and the grey shading denotes a 95 % credibility interval of the association. The asterisk in the top right corner denotes whether the linear association is not zero with the 95% credibility.

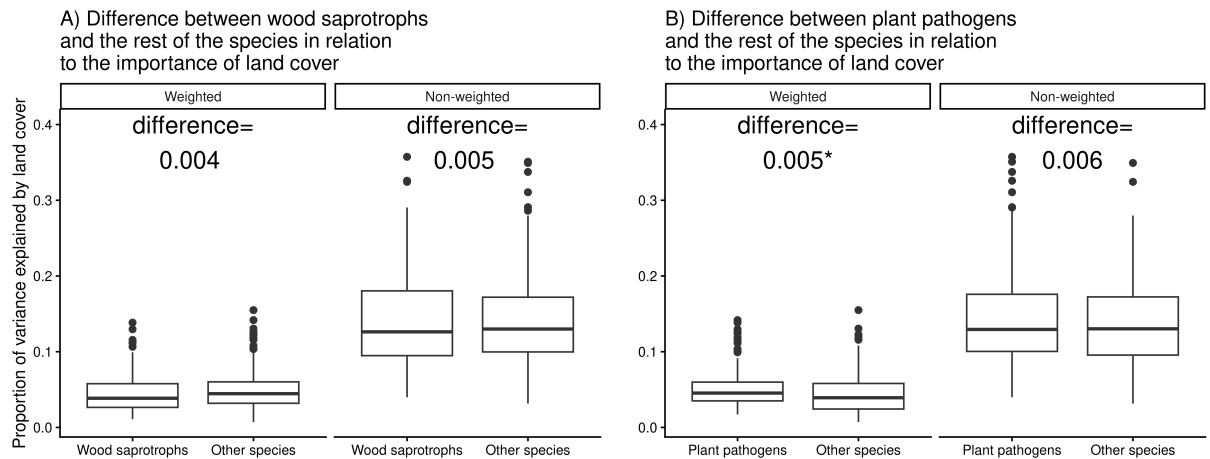


Figure 7: Association between trophic guild status (wood saprotrophs in panel A and plant pathogens in panel B) and the relative importance of land cover (area) in explaining species occurrence pattern. An asterisk denotes that the group means are statistically different with 95% credibility.

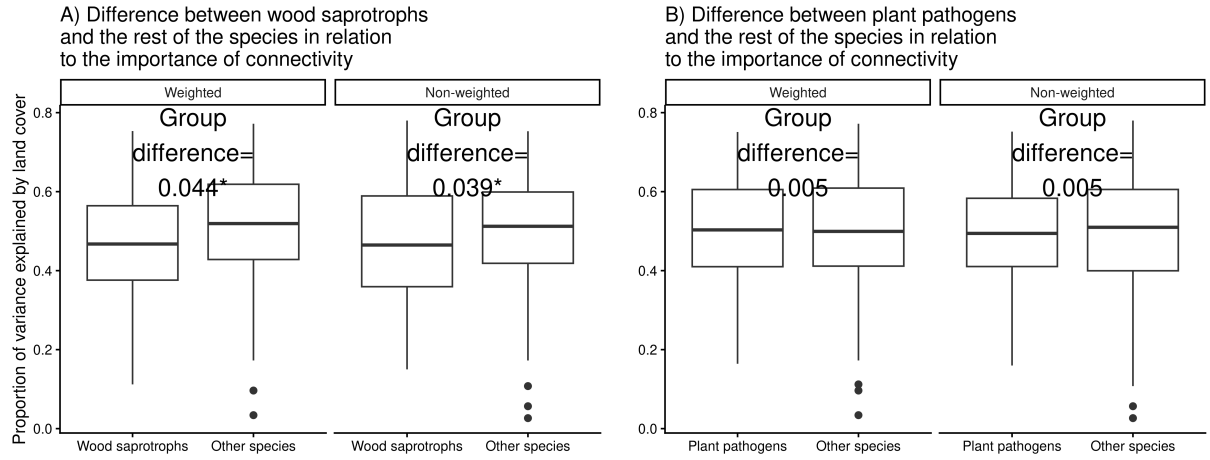


Figure 8: Association between trophic guild status (wood saprotrophs in panel A and plant pathogens in panel B) and the relative importance of connectivity related random effect (random: sample) in explaining species occurrence pattern. An asterisk denotes that the group means are statistically different with 95% credibility.

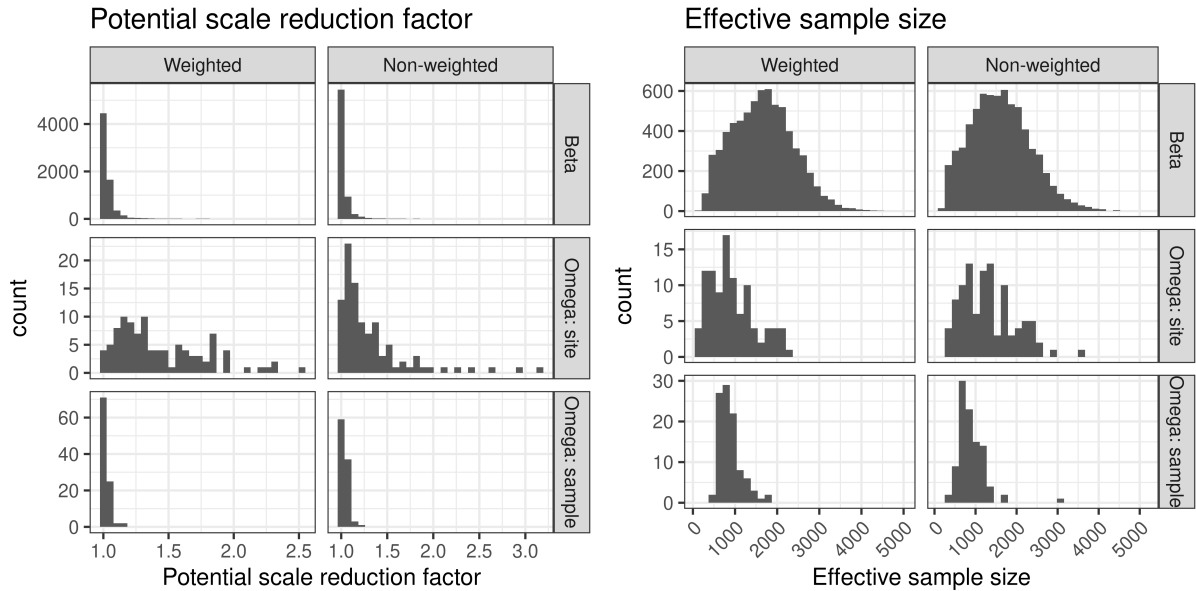


Figure 9: Potential scale reduction factor and effective sample size for Beta and Omega parameters in Hmsc.