**The critical role of tree species and human disturbance in determining the macrofungal diversity in Europe**

**Appendix S2: Tables and Figures**



**Figure S2.1** The cumulative number of fungal species observed in eight European countries (i.e., Norway, Sweden, Finland, Denmark, Netherlands, Germany, France, and Spain) over time. “Fungi original” represents all fungal species in the original database; “Macrofungi original” represents species from seven orders (i.e., *Agaricales*, *Boletales*, *Cantharellales*, *Hymenochaetales*, *Polyporales*, *Russulales*, and *Thelephorales*) in the original database; “Macrofungi in use” represents the macrofungal species used in this study.



**Figure S2.2** The taxon distribution of species. The proportion of species number in (a) all orders in the phylum Basidiomycota in original database and (b) orders of macrofungi (i.e. *Agaricales*, *Boletales*, *Cantharellales*, *Hymenochaetales*, *Polyporales*, *Russulales*, and *Thelephorales*) used in model; (c) the proportion of species that are shared by different numbers of countries in our model; (d) the number of species for seven orders in the eight countries in our model.

**Table S2.1** The number of records for all fungal species in original database as well as for macrofungal species in seven orders (i.e., *Agaricales*, *Boletales*, *Cantharellales*, *Hymenochaetales*, *Polyporales*, *Russulales*, and *Thelephorales*) before and after filtering

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Country | Number of records for all fungi species in original database | Number of records for macrofungal species in seven orders | | | | | |
| Original records number | After removing species with taxonomic issues | After removing records with spatial issue | | Final use in the model | |
| After removing species with positional uncertainty larger than 5 km | After removing species with occurrence less than 30 grids | Saprotrophic macrofungi | Ectomycorrhizal macrofungi |
| Denmark | 586,110 | 474,854 | 467,001 | 466,982 | 460,304 | 99,570 | 54,818 |
| Finland | 33,356 | 29,284 | 28,543 | 27,240 | 26,709 | 8,170 | 7,203 |
| France | 133,349 | 62,772 | 60,758 | 60,758 | 55,692 | 15,559 | 11,176 |
| Germany | 198,818 | 99,814 | 95,628 | 95,628 | 92,271 | 32,122 | 18,399 |
| Netherlands | 507,136 | 243,385 | 231,138 | 231,138 | 229,520 | 46,575 | 18,319 |
| Norway | 426,647 | 275,231 | 272,399 | 272,394 | 270,391 | 57,755 | 61,387 |
| Spain | 137,579 | 45,613 | 43,039 | 42,199 | 36,870 | 13,222 | 9,942 |
| Sweden | 1,422,039 | 1,323,201 | 1,314,662 | 1,314,635 | 1,307,631 | 223,522 | 201,239 |

**Table S2.2** The number of macrofungal species from seven orders (i.e., *Agaricales*, *Boletales*, *Cantharellales*, *Hymenochaetales*, *Polyporales*, *Russulales*, and *Thelephorales*) before and after filtering

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Country | Number of macrofungal species in seven orders | | | | | |
| Original species number | After removing species with taxonomic issues | After removing species with spatial issue | | Final use in the model | |
| After removing species with positional uncertainty larger than 5 km | After removing species with occurrence less than 30 grids | Saprotrophic macrofungi | Ectomycorrhizal macrofungi |
| Denmark | 2703 | 2697 | 2696 | 1721 | 952 | 581 |
| Finland | 1426 | 1426 | 1399 | 1184 | 616 | 416 |
| France | 3040 | 3008 | 3008 | 1734 | 917 | 627 |
| Germany | 2999 | 2986 | 2984 | 1829 | 990 | 637 |
| Netherlands | 1631 | 1614 | 1614 | 1273 | 735 | 395 |
| Norway | 2314 | 2293 | 2293 | 1752 | 898 | 647 |
| Spain | 2802 | 2802 | 2787 | 1647 | 872 | 564 |
| Sweden | 2983 | 2952 | 2952 | 1957 | 1002 | 731 |

*### This R code replicates the prepossessing of GBIF data in the paper entitled:*

*# "The critical role of tree species and human disturbance in determining the macrofungal diversity in Europe"*

*# by H. Yu et al. Submitted to Global Ecology and Biogeography*

*#1. Download fungal occurrence data (csv format) for eight European countries:*

*#Finland: GBIF.org (7th December 2018) GBIF Occurrence Download https://doi.org/10.15468/dl.cbs3ys*

*#France: GBIF.org (16 January 2020) GBIF Occurrence Download https://doi.org/10.15468/dl.log0qg*

*#Spain: GBIF.org (17 January 2020) GBIF Occurrence Download https://doi.org/10.15468/dl.9pjyeg*

*#Germany: GBIF.org (17 January 2020) GBIF Occurrence Download https://doi.org/10.15468/dl.wbuq93*

*#Norway: GBIF.org (23rd November 2018) GBIF Occurrence Download https://doi.org/10.15468/dl.vjjigy*

*#Netherlands: GBIF.org (16 January 2020) GBIF Occurrence Download https://doi.org/10.15468/dl.fnarp9*

*#Denmark: GBIF.org (28th November 2018) GBIF Occurrence Download https://doi.org/10.15468/dl.w773at*

*#Sweden: GBIF.org (7th December 2018) GBIF Occurrence Download https://doi.org/10.15468/dl.wisfph*

*#2. Preprocess of macrofungi data*

*#2.1 Choose macrofungi: seven orders under Agaricomycetes class*

FL <- **read.csv**("./Finland.csv",sep=",")

FR <- **read.csv**("./France.csv",sep=",")

ES <- **read.csv**("./Spain.csv",sep=",")

DE <- **read.csv**("./Germany.csv",sep=",")

NO <- **read.csv**("./Norway.csv",sep=",")

NL <- **read.csv**("./Netherlands.csv",sep=",")

DM <- **read.csv**("./Denmark.csv",sep=",")

SW <- **read.csv**("./Sweden.csv",sep=",")

keep=c("gbifID","datasetKey","kingdom","phylum","class","order","family","genus","species","taxonRank","scientificName","countryCode","decimalLatitude","decimalLongitude","coordinateUncertaintyInMeters","year","basisOfRecord","issue")

FL <- FL[,keep]

FR <- FR[,keep]

ES <- ES[,keep]

DE <- DE[,keep]

NO <- NO[,keep]

NL <- NL[,keep]

DM <- DM[,keep]

SW <- SW[,keep]

all <- **rbind**(FL,FR,ES,DE,NO,NL,DM,SW)

order\_macro <- c("Agaricales", "Boletales", "Cantharellales", "Hymenochaetales", "Polyporales", "Russulales","Thelephorales")

all\_macro <- all[**which**(all$order=="Agaricales"),]

for (i in 2:7) {

all\_macro0 <- all[**which**(all$order==order\_macro[i]),]

all\_macro <- **rbind**(all\_macro,all\_macro0)

}

*#2.2 Data culling: remove records with taxonomic issue, without species name, and with coordinate uncertainty larger than 5km*

all\_macro <- all\_macro[**which**(all\_macro$species!=""),]

all\_macro <- all\_macro[!**grepl**("TAXON\_MATCH\_FUZZY",all\_macro$issue),]

all\_macro0 <- all\_macro[**which**(all\_macro$coordinateUncertaintyInMeters>5000),]

**for** (i **in** 1:l**ength**(all\_macro0$gbifID)) {

all\_macro <- all\_macro[**which**(all\_macro$gbifID!=all\_macro0[i,1]),]

}

**write.csv**(all\_macro,"./all\_macro.csv")

*#2.3 Remove species with species number (grid cell) less than 30*

*#Project coordinate of 'all\_macro' occurrence point into 'ETRS\_1989\_LAEA' using 'Project' tools in ArcGIS 10.8.1; and add the transferred coordinate into all\_macro.shp in ArcGIS*

**library**(raster)

**library**(sf)

P <- **sf::read\_sf**("./all\_macro.shp")

eu <- **sf::read\_sf**("./5kmgrid.shp",fid\_column\_name="ID")

J <- **st\_join**(P,eu,left=TRUE)

J <- **as.data.frame**(J)[,c(1:25)]

J <- **cbind**(J,a=1)

L <- **aggregate**(J$a,by=list(J$species,J$ID),FUN=sum)

L <- **cbind**(L,a=1)

L <- **aggregate**(L$a,by=list(L$Group.1),FUN=sum)

L <- L[**which**(L$x>=30),]

species30 <- L$Group.1

macro\_use <- J[**which**(J$species==species30[1]),]

**for** (i **in** 2:**length**(species30)) {

macro\_use0 <- J[**which**(J$species==species30[i]),]

macro\_use <- **rbind**(macro\_use,macro\_use0)

}

*#2.4 Remove duplicated species records within 5 km grid*

A <- macro\_use[**which**(macro\_use$species==L[1,1]),]

A <- A[!**duplicated**(A$ID),]

**for** (i **in** 2:**length**(L$Group.1)) {

A0 <- macro\_use[**which**(macro\_use$species==L[i,1]),]

A0 <- A0[!**duplicated**(A0$ID),]

A <- **rbind**(A,A0)

}

macro\_use=A[!**is.na**(A$ID),]

*#3. Assign nutritional mode information to species occurrence data*

*#checklist of nutritional mode at species-level(Lspecies.csv): Andrew, et al.(2018), SÃ¡nchez-GarcÃ­a et al.(2020) and Nguyen et al.(2016),*

*#genus-level(Lgenera.csv): Rinaldi, Comandini, & Kuyper(2008) and Tedersoo & Smith(2013) (more details can be found in paper).*

macro\_use <- **cbind**(macro\_use,stype=0,gtype=0)

**for** (i **in** 1:**length**(Lspecies$species)) {

macro\_use[**which**(macro\_use$species==Lspecies[i,1]),26]=Lspecies[i,2]

}

**for** (i **in** 1:**length**(Lgenera$species)) {

macro\_use[**which**(macro\_use$species==Lgenera[i,1]),27]=Lgenera[i,2]

}

#4. Manually double check the nutritional mode of each species