

# **Airborne *Alternaria* and *Cladosporium* fungal spores in Europe: forecasting possibilities and relationships with meteorological parameters**

Agnieszka Grinn-Gofroń<sup>a</sup>, Jakub Nowosad<sup>b</sup>, Beata Bosiacka<sup>a</sup>, Irene Camacho<sup>c</sup>, Catherine Pashley<sup>d</sup>, Jordina Belmonte<sup>e,f</sup>, Concepción De Linares<sup>e,f</sup>, Nicoleta Ianovici<sup>g</sup>, Jose María Maya Manzano<sup>h</sup>, Magdalena Sadyś<sup>i</sup>, Carsten Skjoth<sup>i</sup>, Victoria Rodinkova<sup>j</sup>, Rafael Tormo-Molina<sup>k</sup>, Despoina Vokou<sup>l</sup>, Santiago Fernández-Rodríguez<sup>k</sup>, Athanasios Damialis<sup>l,m\*</sup>

<sup>a</sup>Department of Plant Taxonomy and Phytogeography, Faculty of Biology, University of Szczecin, Szczecin, Poland, agofr@univ.szczecin.pl

<sup>b</sup>Space Informatics Lab, University of Cincinnati, 219 Braunstein Hall, Cincinnati, OH 45221, USA, nowosad.jakub@gmail.com

<sup>c</sup>Madeira University, Faculty of Life Sciences, Campus Universitário da Penteada, 9000-390 Funchal, Portugal, camire@uma.pt

<sup>d</sup>Institute for Lung Health, Department of Infection, Immunity and Inflammation, University of Leicester, Leicester LE1 7RH, UK, chp5@leicester.ac.uk

<sup>e</sup>Unidad de Botánica, Facultad de Ciencias, Universidad Autónoma de Barcelona, Barcelona, Spain, Jordina.Belmonte@uab.cat

<sup>f</sup>Botany Unit, Dept. Of Animal Biology, Plant Biology and Ecology, Universitat Autònoma de Barcelona, Bellaterra (Cerdanyola del Vallès), Spain

<sup>g</sup>West University of Timisoara, Department of Biology, Faculty of Chemistry-Biology-Geography, Romania, nicole\_ianovici@yahoo.com

<sup>h</sup>University of Extremadura, Department of Plant Biology, Ecology and Earth Sciences, Faculty of Science, Avda Elvas s/n, 06071 Badajoz, Spain, jmmaya@unex.es

<sup>i</sup>University of Worcester, National Pollen and Aerobiology Research Unit, Henwick Grove, Worcester WR2 6AJ, United Kingdom, magdalenasadys@wp.pl

<sup>j</sup>National Pirogov Memorial Medical University, Vinnytsia, Ukraine, vikarodi@gmail.com

<sup>k</sup>Department of Construction, School of Technology, University of Extremadura, Avda. de la Universidad s/n, Cáceres, Spain, santiferro@unex.es

<sup>l</sup>Department of Ecology, School of Biology, Aristotle University of Thessaloniki, Thessaloniki, GR-54124, Greece, vokou@bio.auth.gr

<sup>m</sup>Chair and Institute of Environmental Medicine, UNIKA-T, Technical University of Munich and Helmholtz Zentrum München, Germany - German Research Center for Environmental Health, Neusaesser Str. 47, DE-86156, Augsburg, Germany, thanos.damialis@tum.de

\*Corresponding author: Dr. Athanasios Damialis, Chair and Institute of Environmental Medicine, UNIKA-T, Technical University of Munich, Neusaesser Str. 47, 86156 Augsburg, Germany, E-mail: thanos.damialis@tum.de

## **Abstract**

Airborne fungal spores are prevalent components of bioaerosols with a large impact on ecology, economy and health. Their major socioeconomic effects could be reduced by accurate and timely prediction of airborne spore concentrations. The main aim of this study was to create and evaluate models of *Alternaria* and *Cladosporium* spore concentrations based on data on a continental scale. Additional goals included assessment of the level of

generalization of the models in space and description of the main meteorological factors influencing fungal spore concentrations.

Aerobiological monitoring was carried out at 18 sites in six countries across Europe over 3 to 21 years depending on site. Quantile random forest modelling was used to predict spore concentrations values. Generalization of the *Alternaria* and *Cladosporium* models was tested using (i) one model for all the sites, (ii) models for groups of sites, and (iii) models for individual sites.

The study revealed the possibility of reliable prediction of fungal spore levels using gridded meteorological data. The classification models also showed the capacity for providing larger scale predictions of fungal spore concentrations. Regression models were distinctly less accurate than classification models due to several factors, including measurement errors and distinct day-to-day changes of concentrations. Temperature and vapour pressure proved to be the most important variables in the regression and classification models of *Alternaria* and *Cladosporium* spore concentrations.

Accurate and operational daily-scale predictive models of bioaerosol abundances contribute to the assessment and evaluation of relevant exposure and consequently more timely and efficient management of phytopathogenic and of human allergic diseases.

## **Keywords**

Advanced statistical models; aerobiology; bioaerosols; biometeorology; continental scale; molds

## **Highlights**

- No operational forecasting model for allergenic fungal spore exposure exists in Europe
- Potential exposure in Europe was assessed and predicted for 2 major allergenic fungi
- Random forest modelling was applied to >7,000 daily time series
- Air temperature and vapour pressure were the most significant variables
- Classification models showed higher capacity for large-scale spore predictions

## 1. Introduction

Fungal spores are one of the most prevalent components of bioaerosols, found across a wide range of biogeographic regions over long time periods each year. The primary source of spore emissions are the substrates on which fungi grow, such as plants, soil and decaying organic matter. This means the majority of fungal spores in the air originate from farms, forest stands and decomposing plant material (Bowers et al., 2013). They will remain airborne for variable amounts of time and will be transported over distances ranging from a few centimeters to hundreds of kilometers (Ansari et al., 2015; Heald and Spracklen, 2009; Stockmarr et al., 2007).

*Alternaria* and *Cladosporium* are ubiquitous asexually reproducing fungal genera that produce spores, known as conidia, which are readily airborne. Both genera contain plant pathogens (Carlile et al., 2007; Chaerani and Voorrips, 2006; Lee et al., 1997; Nowicki et al., 2012; Thomma et al., 2005), many of which also produce phytotoxic metabolites that affect mammalian cells (De Lucca, 2007; Friesen et al., 2008; Mamgain et al., 2013). Fungal spores are also important aeroallergens, causing adverse health effects (Krouse et al., 2002).

Aerobiological surveys reported *Alternaria* as one of the most prevalent airborne fungal types and an important aeroallergen (Budd, 1986; Mitakakis et al., 2001). Also, *Cladosporium* is frequently reported as the most abundant aeroallergen and the second most

allergenic fungal type worldwide (Tariq et al., 1996). Allergy to fungi from both genera has been responsible for hospital admissions due to severe asthma attacks in sensitized individuals, particularly among children (Bush and Prochnau, 2004; Dales et al., 2000). Due to their ubiquitous nature (Damialis et al., 2017; Jędryczka, 2014) exposure to these aeroallergens is literally inevitable during their dispersion season.

It is important to assess the potential risks of the atmospheric presence of fungal spores, and more research is needed to evaluate this (Beggs, 2004; Crameri et al., 2014). There is limited data on long-term trends in airborne fungal spore abundance, in part due to a lack of representative fungal spore data sets longer than a decade across the globe. Airborne fungal spores as allergens have received comparatively less attention than pollen, and associated public health consequences are likely to have been underestimated (Damialis et al. 2015).

Sporulation and dispersion of fungi are influenced by several meteorological factors, including air temperature, relative humidity, precipitation, atmospheric turbulence, wind speed, and UVB radiation (Al-Subai, 2002; Carlile et al., 2007; Cecchi et al., 2010; Straatsma et al., 2001). However, the exact influence of climatic variability on fungal ecology at a larger scale is still not understood. There are indications that changing climate may lead to alterations in phenology (Corden et al., 2003; Gange et al., 2007; Kauserud et al., 2010) and dynamics of fungal communities (Gange et al., 2011).

This complexity of biological and ecological processes represent one of the key problems of modelling biological systems. There are inconsistencies regarding what drives and controls the distribution of fungal bioaerosols both at a local and regional scale. Multiple meteorological factors may alter the spatiotemporal distribution of *Alternaria* and *Cladosporium* spores (Corden et al., 2003; Damialis and Gioulekas, 2006; De Linares et al., 2010; Escudero et al., 2011; Iglesias et al., 2007; Recio et al., 2012; Sindt et al., 2016; Skjøth

et al., 2016). The relationships between fungal development and environmental factors, including major climatic variables, are often the only component used for disease forecasting systems (van Maanen and Xu, 2003).

Atmospheric dispersion models have been used to describe the spatiotemporal dispersal of fungal pathogens (Burie et al., 2012; Oteros et al., 2015; Stockmarr et al., 2007; van Leuken et al., 2016). Additionally, descriptive, predictive or conceptual modelling of airborne fungal spores concentration is another promising tool, albeit challenging (Grinn-Gofroń and Bosiacka, 2015; Grinn-Gofroń et al., 2018; Iglesias et al., 2007; Jędryczka et al., 2015). At a minimum, high-resolution data on meteorological and geographic variables should be included in such models. It is important to include and analyse all associated factors, so that the resulting models can accurately describe the complex environmental interdependencies.

Given the widespread, ecological, economical and health impact of spores from *Alternaria* and *Cladosporium*, the main goal of the present study was to build models that generalize beyond the observed data and are capable of estimating the spatiotemporal distribution and concentration of fungal spores on a broad scale and to address the factors that influence the fungal spatiotemporal patterns. Created models could help to answer several questions - (i) is it possible to make a reliable prediction of fungal spore concentrations using only one model for all sites, (ii) is there a difference in quality between one model for all the sites, models for the groups of sites, and models for individual sites, (iii) what are the main meteorological factors influencing fungal spore concentrations, and (iv) do these factors differ between created models (analysed sites)?

## **2. Materials and Methods**

### **2.1. Aerobiological data**

The taxa selected for this study were those of *Alternaria* and *Cladosporium*. These comprise some of the most well studied fungal taxa worldwide due to their allergenic and phytopathogenic properties. In the current research, the spore abundance ranges greatly from >15% to over 96% of the total annual spore index per site (Table 2). This variation alone challenges for the elaboration of a universal forecasting model in Europe. But even if we had added the rest of the fungal diversity per site, this would have not most probably provided any additional insight to either ecological/biological processes or health impacts, as responses of different fungi are highly individualistic and their sensitivity (or not) to environmental stress (i.e. climate change) would not be reflected reliably (e.g. Damialis et al. 2015). Also, in most stations in the world, mainly these two types alone are counted. So as to obtain similar data in such a large-scale spatial study design and consequently comparable results, we did not extend to the investigation of additional fungal taxa.

Fungal spores were collected in the frame of long-term aerobiological monitoring in 18 sites from six countries across Europe and for a time span of 1987-2015 (Fig. 1, Table 1). *Alternaria* spore concentrations were measured at all sites and *Cladosporium* spore concentrations at 15 out of the 18 (Table 1).

In brief, for this research, all data providers have been inquired regarding major changes in methodological procedures and no significant alteration has been reported. Moreover, in all participating countries, microscopic identification of airborne fungal spores has been conducted by experts with long-standing experience in such techniques. Therefore, in all stations, the same method of collection and analysis was used. Standard sampling, processing and analysis techniques were followed, according to the recommendations of the European Aerobiology Society (Frenguelli, 2003) and the British Aerobiology Federation (1995). At each location, a Hirst-type volumetric spore trap was used to sample airborne fungal spores (Hirst, 1952), which is considered the gold-standard device for sampling

airborne particles of biological origin (Galán et al., 2014). Samples were collected and analysed weekly, applying standard methods for sample processing and microscopic identification (e.g. British Aerobiology Federation, 1995; Galán et al., 2014; Grant Smith, 1984). Final measurements referred to daily resolution and were expressed as concentrations of fungal spores per cubic meter of air on a given date.

## 2.2. Meteorological data

Eight meteorological factors were included as co-factors in the data analysis, namely maximum temperature, minimum temperature, average temperature, vapour pressure, sum of precipitation, potential evaporation from a free water surface, potential evapotranspiration from a crop canopy, and total global radiation. These were acquired from the AGRI4CAST Interpolated Meteorological Database (Baruth et al., 2007).

Based on maximum ( $T_{max}$ ) and minimum ( $T_{min}$ ) temperature, an additional ninth parameter was included, growing degree days (GDD). Cumulative GDD is an indicator measuring a heat accumulation, and as a proxy, can represent plant and fungal development.

GDD value is calculated as follows:

$$GDD = \frac{T_{max} + T_{min}}{2} - T_{base}$$

Values of GDDs were accumulated starting from January 1. GDD do not accumulate when the daily mean temperature ( $(T_{max} + T_{min})/2$ ) is lower or equal to the base temperature. Value of  $T_{base}$  was set to 5. No relevant previous information on base temperatures for fungal spores existed, so we assumed that these would be similar to later-flowering plant species (being abundant mainly during May-August) and which could be found across a variety of latitudes and climates in Europe; therefore, we set the base temperature at 5, in a similar manner to studies on grass pollen (Emberlin 1993; Frenguelli et al., 1989).

All the the abovementioned factors were considered either based on their availability, or because of the focus of the current study, or based on previous literature on the topic. For



instance, relative humidity data are not available via AGRI4CAST as these cannot be corrected for altitude differences among sites within the target climatic grid cell. Regarding other factors, like wind vectors (speed, persistence and direction), these were excluded from this analysis, as this would make more sense in a smaller-scale temporal data processing, when the intermittent nature of wind would be possible to take into account. Finally, there have been previously published reports of particular meteorological factors in site-specific studies, included in the current analysis, proven to be the most decisive for obtaining accurate and reliable forecasts (e.g. Damialis and Gioulekas, 2006).

### **2.3. Models**

Partial autocorrelation function (Durbin, 1960) was applied independently to *Alternaria* and *Cladosporium* daily counts to check for the temporal autocorrelation of spore data. This summarizes the relationship between an observation  $x_t$  and observations with lagged time steps (days) removing the impact of the values at all shorter lags.

Redundancy among the lagged values of meteorological predictors was explored using principal coordinates analysis (PCA) (Jolliffe, 1986). PCA transforms a number of correlated variables into a smaller set of uncorrelated variables. The role of PCA in this study was to reduce the dimension of the data, reduce computational time of models building, minimize spurious effects of single lags and ease interpretation of the final models.

Two main types of modeling techniques were used, regression and classification. The Quantile Random Forest method was used to create regression models. This is a generalization of Random Forests that infer the full conditional distribution of a response variable (Li et al., 2011). We decided not to predict mean value, but rather a median value, because of a non-symmetrical distribution of fungal spore values. Daily spore concentrations of *Alternaria* or *Cladosporium* were used as dependent variables (model output) in regression models. Breiman's Random Forest (Breiman, 2001) was used in classification models. So as

to eliminate high levels of statistical noise, daily spore concentrations were divided into two levels, low and high, according to the thresholds of *Alternaria* and *Cladosporium* allergens to evoke allergic symptoms (Gravesen, 1979). *Alternaria* values lower than 100 spores were considered as ‘low’ and beyond that threshold they were characterized as ‘high’. For *Cladosporium*, this threshold was 3,000 spores. As airborne fungal spore measurements could still exhibit a huge disparity in the frequencies of the observed classes, we adopted an optimizing probability threshold technique (Kuhn and Johnson, 2013; Nowosad, 2016). In this approach, alternative cutoffs for the predicted probabilities were determined using resampling. Sensitivity (Sens), specificity (Spec), positive predictive value (Ppv), and negative predictive value (Npv) were calculated for 20 different threshold values. For each model, optimal threshold value was established minimizing the distance between Sens, Spec, Ppv, Npv and the value of 1. This value indicates the best possible performance.

Attempting to also reduce the spatial variability among monitoring sites across Europe, analyzed sites were divided into three groups based on the annual temporal changes of fungal spore concentrations and on data availability (Table 1), as follows: northeastern Spain (Barcelona, Bellaterra, Girona, Lleida, Manresa, Roquetes-Tortosa, Tarragona, Vielha), western Spain (Don Benito, Plasencia, Zafra) and non-Spanish sites (Derby, Leicester, Szczecin, Thessaloniki, Timisoara Vinnytsia, Worcester). The grouping was based on the seasonality and multi-modality of airborne spore concentrations, i.e. the first group exhibited the highest seasonality and normality of data, with longer seasons and fewer outliers. The third group presented a bi- or multi- modal yearly pattern, whilst the second group included those sites with only *Alternaria* spore measurements.

The modeling was performed in three ways: 1) per site, 2) per group of sites as mentioned above, and 3) an integrated model for all sites combined. This resulted in 22 combinations for *Alternaria* (18 sites, 3 groups of sites, and 1 whole dataset) and 18

combinations for *Cladosporium* (15 sites, 2 groups of sites, and 1 whole dataset) generating a total of 40 regression and 40 classification models.

## **2.4. Validation metrics**

The accuracy of all models was assessed using a repeated k-fold cross-validation (Kuhn and Johnson, 2013). Regression models were evaluated using the Symmetric Mean Absolute Percentage Error (SMAPE). Thus, accuracy measure was based on relative errors instead of standard metrics like Root Mean Square Error (RMSE) or Mean Absolute Error (MAE), which was particularly important in our study because of high among-site data variability. Classification modelling results were characterized using balanced accuracy. This metric is calculated as sensitivity + specificity/2. It is a better measurement in the case of imbalanced datasets as it gives the same weight for correctly predicted cases with low concentration and correctly predicted cases with high concentration. Influence of the predictors was determined using a scaled permutation importance (mean decrease in accuracy) (Breiman, 2001).

## **3. Results**

### **3.1. Spatiotemporal variability of spore concentrations**

Fungal spore concentrations differed between *Alternaria* and *Cladosporium*, between sites, and between years of measurements (Table 2, Figs. 2, 3). The annual sum of *Alternaria* daily concentrations varied between 907 (Vielha) and 67,166 (Lleida) and had a mean value of 13,448. In addition to Lleida having the highest average annual sums of *Alternaria*, the values here were also the most variable between years. Other sites with substantial changes between years were Derby, Thessaloniki, Vielha, and Girona. The annual sum of the daily concentrations of *Cladosporium* varied between 24,637 (Thessaloniki) and 1,500,699 (Derby) and had a mean value of 405,237. Annual spore concentrations in Lleida were also the most changeable for *Cladosporium*. Overall, the order of average values for all sites was similar for

*Alternaria* and *Cladosporium* with two exceptions, Worcester and Thessaloniki. The annual sum of the daily concentrations of *Alternaria* in Worcester were relatively small and invariable, whilst *Cladosporium* values were high and changeable. In Thessaloniki *Alternaria* levels were moderate and varied substantially, whilst *Cladosporium* annual values were small and more consistent.

For *Alternaria*, sites could be split into two groups based on the time course of the season: (i) sites with period of high concentration and period of lower concentrations (such as Barcelona, Lleida, etc.), (ii) sites with one period of high concentrations and period with absence of *Alternaria* spores (such as Szczecin, Thessaloniki and Vielha) (Fig. 2). Time course of *Cladosporium* spore concentrations is more heterogeneous, with probably two main groups: (i) Spanish sites with period of high concentration and period of moderate concentrations, (ii) sites with one period of high concentration and period with low concentrations or absence of spore concentrations (such as Szczecin, Derby, Vinnytsia). Additionally, a third group, consisting of Don Benito, Plasencia, and Zafra, was separated due to missing data of *Cladosporium* (Fig. 3).

### **3.2. Predictor variables**

Based on the spatiotemporal analysis, lagged daily values of nine meteorological parameters between 1-15 days were created. A principal component analysis (PCA) for each of the parameters was run and the results of the PCA gave an insight into the variability of predictors. In most of the cases, the two first components explained the majority of variations. The cumulative value was between 0.85 for radiation and 0.94 for average temperature in these variables. The first component expressed the value of a given parameter (large values of loadings for all of the lags), whilst the second parameter expressed the temporal changes of a given parameter (the largest, positive value for the first lag, and the lowest, negative value for

the last lag). In case of cumulated GDD, we used only the first component, which explained 0.95 of variation. Precipitation values were the most changeable, and the first two components had a cumulative variance of 0.22. Therefore based on the autocorrelation plot, we used raw values of precipitation for lags between 1 and 4 days.

The final group of predictors consisted of 19 variables - the first principal component of cumulated GDD, one to four day lags of precipitation, and the first two principal components of the rest of the meteorological parameters (Table 3).

### **3.3. Regression models**

#### **3.3.1. Performance of the models**

The final 40 regression models (22 for *Alternaria* and 18 for *Cladosporium*) were built and results compared using Symmetric Mean Absolute Percentage Error (SMAPE) (Table 4). These values ranged between 0.56 and 0.90 for *Alternaria* (average value of 0.69), and 0.53 and 0.73 for *Cladosporium* (average value of 0.61). For all of the sites, the models gave a value of 0.76 for *Alternaria* and a value of 0.73 for *Cladosporium*. Those values were slightly higher than values for separate groups and distinctly higher than values for most of the individual sites. Similarly, most of the models for the site groups performed worse than the models for the individual sites. *Alternaria* model had a SMAPE value of 0.73 for the first group (average for the models for the individual sites in this group was 0.67), 0.75 for the second group (average of 0.74), and 0.69 for the third group (average of 0.61). These differences were higher in *Cladosporium* models with SMAPE of 0.67 for the first group (average of 0.60) and 0.71 for the second group (average of 0.59). Only in four *Alternaria* models, Derby, Worcester, and Vielha, values of SMAPE were higher than values for models of groups of sites.

#### **3.3.2. Variable importance**

The same set of variables seemed to influence the models of both taxa (Fig. 4). The most important predictors were the first principal coordinates of vapor pressure and temperatures (minimum, average, maximum and cumulated GDD). They were followed by the first principal coordinates of evapotranspiration, evaporation, and radiation. The second principal coordinates of those parameters had small importance. The smallest values were observed for lagged values of precipitation.

### **3.4. Classification models**

#### **3.4.1. Performance of the models**

The final classification models were compared using the balanced accuracy metric (Table 4). Average value of balanced accuracy were 0.78 for *Alternaria* and 0.73 for *Cladosporium*. These values varied among sites between 0.50 and 0.90 for *Alternaria* and between 0.50 and 0.99 for *Cladosporium*. Models for all the sites gave similar results of 0.80 and 0.78. The *Alternaria* model gave a balanced accuracy value of 0.77 for the first group of sites compared to the average of individual site's models of 0.75, a value of 0.85 for the second group of sites comparing to the average of individual site's models of 0.81, and a value of 0.80 for the third group of sites compared to the average of individual site's models of 0.71. Values of a balanced accuracy for the models of groups of sites were also higher for *Cladosporium*. The first group had a value of 0.68 (average for individual sites was 0.66), and the second group had a value of 0.85 (average for individual sites was 0.80). About 44% (15 of 34) models for individual sites gave worse values of balanced accuracy than models for groups of sites.

#### **3.4.2. Variable importance**

Values of variable importance for classification models were more diverse than for regression models. *Alternaria* and *Cladosporium* classification models were influenced mostly by the same predictors (Fig. 4). Temperature (cumulated GDD) was the most

important variable, followed by the first and second principal coordinate of vapor pressure. The rest of the predictors showed moderate to low importance. Similarly to regression models, predictors with values of precipitation had the smallest importance.

#### 4. Discussion

The present study revealed that wide-scale, accurate, operational modeling of fungal spore abundances is feasible with one universal model, which answers the first research question. Of course, there are restrictions because of annual and spatial variability, which result in varying performance of the obtained models. Based on the model used, these individualistic responses (also based on the fungal taxon examined) can be decreased to an extent. We found that specific meteorological factors significantly contributed to the forecasting power, with air temperature playing the leading role. Consequently, the current research highlights the possibility and need for developing universal predictive models of airborne fungal spore abundances, something currently lacking and therefore making this study novel.

In the current research, an integrative approach of the variables affecting the distribution of these bioaerosols was adopted, on a local or a regional scale, providing a deeper comprehension of the dynamics of these fungal taxa. The survey revealed distinct relationships between spore concentrations and sites and years of sampling. The overall spore frequency and the annual sum of the daily concentrations of *Cladosporium* was higher than *Alternaria*'s, a finding also reported previously in different regions of the world. In Cartagena (Spain), *Cladosporium* represented 62.2% of the total spore count and *Alternaria* only 5.3%, however, *Alternaria* was still the second most abundant fungal type (Elvira-Rendueles et al., 2013). Likewise, in Bursa (Turkey), *Cladosporium* represented 88.1% of the total spore count followed by *Alternaria* at 4.9% (Ataygul et al., 2007). A similar trend was found in

Thessaloniki (Greece) by Gioulekas et al. (2004) and in Madrid (Spain) by Sabariego et al. (2007).

Two types of models, regression and classification, were built as a part of this study. The goal of the first one was to predict values of the fungal spore concentration in the studied sites, while the classification models were created to predict high levels of the fungal spore concentrations. Classification models were more accurate than regression models. This is due to a number of factors. Concentration values of *Alternaria* and *Cladosporium* were classified into two groups prior to the modelling. This procedure generalises the fungal spore values and, therefore, changes in values. As a result, it gives more general trends instead of showing the local trends and day-to-day differences. Moreover, the obtained values of concentration are an estimation of the whole population and are prone to random and systematic (bias) errors (Comtois et al., 1999; Oteros et al., 2013). The benefit of using a classification method is that it reduces the influence of methodological differences between sites, such as the relative position of samplers and the heights of buildings, or instrumental and human errors.

The life cycles of many fungal pathogens are strongly determined by weather. The airborne spore concentrations are affected by biological factors (reproduction and survival), weather parameters, land use, resource availability and competition (Boddy et al., 2014). Dispersal and circulation are highly influenced by wind and rainfall, while germination and infection rates are often dependent upon liquid water on the plant surface (sometimes high relative humidity) and species-specific optimal temperature ranges. In this study the most important meteorological variables were the same for both *Alternaria* and *Cladosporium*. This shows that despite the differences in the values of spore concentration, both taxa are mostly affected by the same meteorological factors, such as temperature (minimum, average, maximum and cumulated GDD) and vapour pressure, which determines the water content in the air (related to air humidity).



In many aerobiological studies, temperature and relative humidity of the air are meteorological parameters that significantly influence concentrations of *Cladosporium* and *Alternaria* spores, with temperature being positively associated and relative humidity negatively associated (e.g. Grinn-Gofroń and Strzelczak 2009; O'Connor et al. 2014; Sadyś et al. 2016; Ianovici 2016; Almeida et al. 2018)

Other variable - global radiation includes both the direct solar radiation and the diffuse radiation resulting from reflected or scattered sunlight, and can be considered as a function of temperature (Meza and Varas 2000). One of the solar radiation components, UVB radiation, is reported to affect the survival of airborne fungal spores during movement over long distances through the atmosphere (Al-Subai 2002). This predictor showed moderate importance in our models.

Rain has been often cited as being one of the most influential factors in reducing airborne pollen (e.g. Damialis et al., 2005). However, precipitation was identified as the least important factor in our models. There are two main reasons behind the small effect of precipitation on the fungal spore models. Firstly, variation of daily precipitation values does not closely correlate with the values and levels of *Alternaria* and *Cladosporium* concentrations, both of which are regarded as dry weather spores. For example, it is possible to have a rainless day and a high fungal spore concentration (the middle of the season) and to have a rainless day without airborne fungal spores (off-season). Secondly, precipitation is characterized by high variability, which could not be captured in a daily timescale (sum of precipitation). To better understand the effects of rainfall on fungal spore diversity and abundance, and on circulation patterns, finer resolution data (hourly scale), with different statistical techniques (i.e. artificial intelligence models) would be required.

## **5. Conclusions**

- Classification models were more accurate than regressions for *Alternaria* and *Cladosporium* fungal spores.
- Regression models gave better results for individual sites compared to grouped sites, resulting potentially from strong effects from local meteorological conditions.
- Classification models gave better results in grouped sites rather than for individual sites, thus, displaying the capacity for accurately providing larger scale predictions of fungal spore concentrations (compared to the more localized regression models).
- Temperature (in the form of minimal, average, maximum temperature, and cumulated GDD) and vapour pressure were the most important variables in models of *Alternaria* and *Cladosporium*, while radiation and daily sum of precipitation had a smaller impact on the models.

## References:

- Almeida, E., Caeiro, E., Todo-Bom, A., Ferro, R., Dionísio, A., Duarte, A., Gazarini, L., 2018. The influence of meteorological parameters on *Alternaria* and *Cladosporium* fungal spore concentrations in Beja (Southern Portugal): preliminary results. *Aerobiologia* 34, 219–226
- Al-Subai, A.A.T., (2002). Air-borne fungi at Doha, Qatar. *Aerobiologia* 18, 175–183.
- Ansari, T.U., Valsan, A.E., Ojha, N., Ravikrishna, R., Narasimhan, B., Gunthe, S.S., 2015. Model simulations of fungal spore distribution over the Indian region. *Atmos. Environ.* 122, 552–560.
- Ataygul, E., Celenk, S., Canitez, Y., Bicakci, A., Malyer, H., Sapan, N., 2007. Allergenic Fungal Spore Concentrations in the Atmosphere of Bursa, Turkey. *J. Biol. Environ. Sci.* 1, 73–79.

- Baruth, B., Genovese, G., Leo, O., 2007. CGMS Version 9.2 - User Manual and Technical Documentation. Office for Official Publications of the European Communities, Luxembourg.
- Beggs, P.J., 2004. Impacts of climate change on aeroallergens. Past and future. Clin. Exp. Allergy 34, 1507–1073.
- Berman, D. 2011. Climate change and aeroallergens in South Africa. Curr. Allergy Clin. Immunol. 24, 65–71.
- Boddy, L., Büntgen, U., Egli S., Gange, A.C., Heegaad, E., Kirk, P.M. et al., 2014. Climate variation effects on fungal fruiting. Fungal Ecol. 10, 20–33.
- Bowers, R.M., Clements, N., Emerson, J.B.E., Wiedinmyer, C., Hannigan, M. P., Fierer, N., 2013. Seasonal Variability in Bacterial and Fungal Diversity of the Near-Surface Atmosphere. Environ. Sci. Technol. 47, 12097–12106.
- Breiman, L. 2001. Random Forests. University of California, accessed on 30<sup>th</sup> May 2018, <https://www.stat.berkeley.edu/~breiman/randomforest2001.pdf>
- British Aerobiology Federation, 1995. Airborne pollens and spores. A guide to trapping and counting: National Pollen and Hayfever Bureau, Rotherham, United Kingdom.
- Budd T.W., 1986. Allergens of *Alternaria*. Grana 25, 147–154
- Burie, J.B., Calonnec, A., Langlais, M., Mammeri, Y., 2012. Modelling the spread of a pathogen over a spatially heterogeneous growing crop. Paper presented at the 2012 IEEE 4th International Symposium on Plant Growth Modeling, Simulation, Visualization and Applications (PMA), Shanghai, China.
- Bush, R. K., Prochnau, J.J., 2004. *Alternaria*-induced asthma. J. Allergy Clin. Immunol. 113, 227–234.
- Carlile, M.J., Watkinson, S.C., Gooday, G.W., 2007. The fungi, 2nd edition: Elsevier, UK, The Netherlands, USA.

- Cecchi, L., D'Amato, G., Ayres, J.G., Galan, C., Forastiere, F., Forsberg, B., Gerritsen, J., Nunes, C., Behrendt, H., Akdis, C., Dahl, R., Annesi-Maesano, I., 2010. Projections of the effects of climate change on allergic asthma: the contribution of aerobiology. *Allergy* 65, 1073–1081.
- Chaerani, R., Voorrips, R. E., 2006. Tomato early blight (*Alternaria solani*): the pathogen, genetics, and breeding for resistance. *J. Gen. Plant Pathol.* 72, 335–347.
- Comtois, P., Alcazar, P., Néron, D., 1999. Pollen counts statistics and its relevance to precision. *Aerobiologia* 15, 19–28.
- Corden, J.M., Millington, W.M, Mullins, J., 2003. Long-term trends and regional variation in the aeroallergen *Alternaria* in Cardiff and Derby UK – are differences in climate and cereal production having an effect? *Aerobiologia* 19, 191–199.
- Crameri, R., Garbani, M., Rhyner, C., Huitema, C. 2014. Fungi: The neglected allergenic sources. *Allergy* 69,176–185.
- Dales, R. E., Cakmak, S., Burnett, R. T., Judek, S., Coates, F., Brook, J. R., 2000. Influence of ambient fungal spores on emergency visits for asthma to a regional children's hospital. *Am. J. Respir. Critic. Care Med.* 162, 2087–2090.
- Damialis, A., Gioulekas, D., 2006. Airborne allergenic fungal spores and meteorological factors in Greece: Forecasting possibilities. *Grana* 45,122–129.
- Damialis, A., Gioulekas, D., Lazopoulou, C., Balafoutis, C., Vokou, D., 2005. Transport of airborne pollen into the city of Thessaloniki: the effects of wind direction, speed and persistence. *Int. J. Biometeorol.* 49, 139-145.
- Damialis, A., Kaimakamis, E., Konoglou, M., Akritidis, I., Traidl-Hoffmann, C., Gioulekas, D., 2017. Estimating the abundance of airborne pollen and fungal spores at variable elevations using an aircraft: how high can they fly? *Scientific Reports.* 7: 44535, Doi:10.1038/srep44535

- Damialis, A., Mohammad, A.B., Halley, J.M., Gange, A.C., 2015. Fungi in a changing world: growth rates may be elevated, but spore production will decrease in future climates. *Int. J. Biometeorol.* 59, 1157–1167.
- De Linares, C., Belmonte, J., Canela, M., Díaz de la Guardia, C., Alba-Sanchez, F., Sabariego, S., et al., 2010. Dispersal patterns of *Alternaria* conidia in Spain. *Agric. Forest Meteorol.* 150, 1491–1500.
- De Lucca, A.J., 2007. Harmful fungi in both Agriculture and Medicine. *Rev. Iberoam. Micol.* 24, 3–13.
- Durbin, J., 1960. The Fitting of Time-Series Models. *Revue de l'Institut International de Statistique / Review of the International Statistical Institute.* 28, 233–244.
- Elvira-Rendueles, B., Moreno, J., Garcia-Sanchez, A., Vergara, N., Martinez-Garcia, M. et al., 2013. Air-spore in Cartagena, Spain: Viable and non-viable sampling methods. *Ann. Agric. Environ. Med.* 20, 664–671.
- Emberling, J., 1993. Variations in pollen concentrations – the need for accurate monitoring. *Clin. Exp. Allergy* 23, 45.
- Escudero, O., Seijo, M.C., Fernández-González M., Iglesias I., 2011. Effects of meteorological factors on the levels of *Alternaria* spores on a potato crop. *Int. J. Biometeorol.* 55, 243–252.
- Frenguelli, G., 2003. Basic microscopy, calculating the field of view, scanning of slides, sources of error. *Postępy Dermatologii i Alergologii* 2, 227–229.
- Frenguelli, G., Bricchi, E., Romano, B., Mincigrucci, G., Spieksma, F.T.M., 1989. A predictive study on the beginning of the pollen season for Gramineae and *Olea europaea* L. *Aerobiologia* 5, 64–70.
- Friesen, T.L., Faris, J.D., Solomon, P.S., Oliver, R.P., 2008. Host-specific toxins: effectors of necrotrophic pathogenicity. *Cell Microbiol.* 10, 1421–1428.

- Galán, C., Smith, M., Thibaudon, M., Frenguelli, G., Oteros, J., Gehrig, R., Berger, U., Clot, B., Brandao, R., 2014. Pollen monitoring: minimum requirements and reproducibility of analysis. *Aerobiologia* 30, 385-395.
- Gange, A.C., Gange, E.G., Mohammad, A.B., Boddy L., 2011. Host shifts in fungi caused by climate change? *Fungal Ecol.* 4, 154–190.
- Gange, A.C., Gange, E.G., Sparks, T.H., Boddy L., 2007. Rapid and recent changes in fungal fruiting patterns. *Science* 316, 71.
- Gioulekas D., Damialis A., Mpalafoutis C., Papakosta D., Giouleka P., and Patakas D., 2004. Allergenic fungal spore records (15 years) and relationship with meteorological parameters in Thessaloniki, Greece. *Allergy Clin. Immunol. Int.* 16, 52-59.
- Grant Smith, E., 1984. Sampling and identifying allergenic pollens and molds. An Illustrated Manual for Physicians and Lab Technicians: Blewstone Press, San Antonio, Texas.
- Gravesen, S., 1979. Fungi as a cause of allergic disease. *Allergy* 34, 135–154.
- Grinn-Gofroń, A., Bosiacka, B., 2015. Effects of meteorological factors on the composition of selected fungal spores in the air. *Aerobiologia* 31, 63–72.
- Grinn-Gofroń, A., Bosiacka, B., Bednarz, A., Wolski, T., 2018. A comparative study of hourly and daily relationships between selected meteorological parameters and airborne fungal spore composition. *Aerobiologia* 34, 45–54.
- Grinn-Gofroń, A., Strzelczak, A., 2009. Hourly predictive ANN and MRT models of *Alternaria* and *Cladosporium* spore concentrations in Szczecin (Poland). *Int. J. Biometeorol.* 53, 555-562.
- Heald, C. L., Spracklen, D.V., 2009. Atmospheric budget of primary biological aerosol particles from fungal spores. *Geophys. Res. Lett.* 36.
- Helfer, S., 2014. Rust fungi and global change. *New Phytologist*. 201, 770–780.
- Hirst, J., 1952. An Automatic Volumetric Spore Trap. *Ann. Appl. Biol.* 39, 257–265.

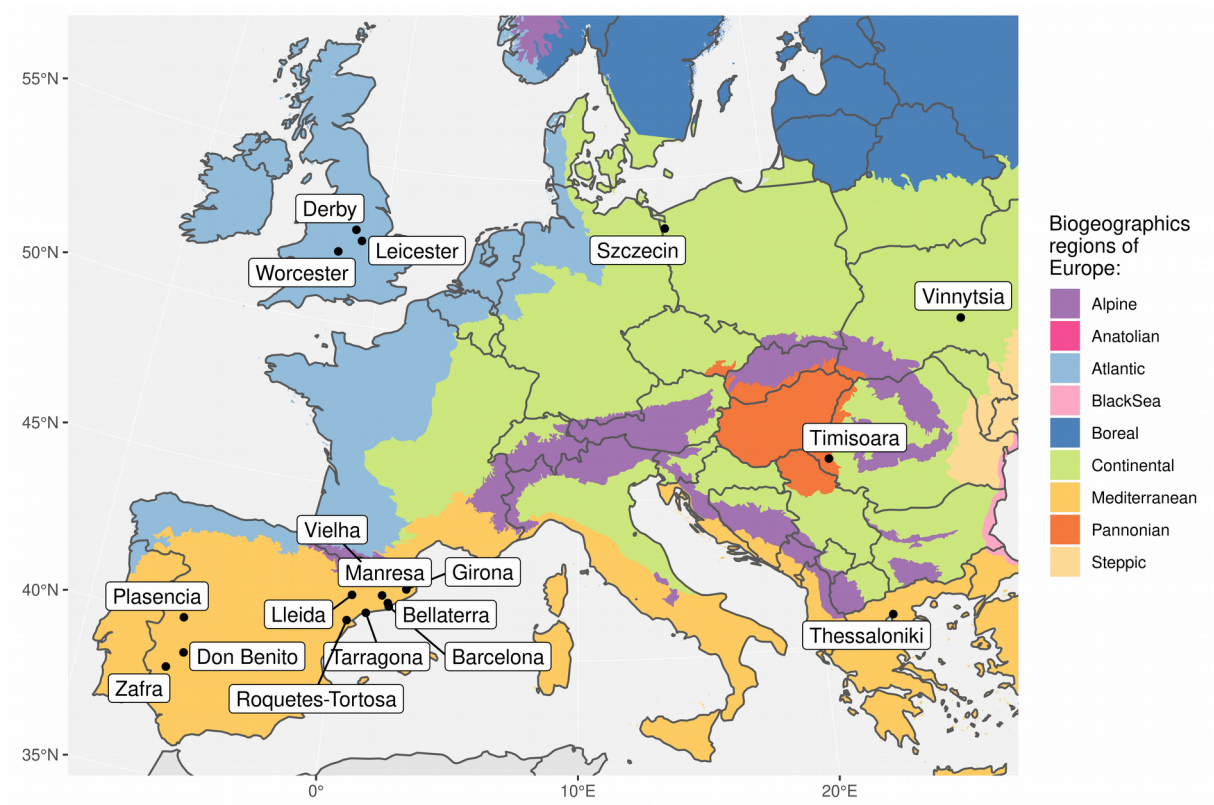
- Ianovici, N., 2016. Atmospheric concentrations of selected allergenic fungal spores in relation to some meteorological factors, in Timisoara (Romania). *Aerobiologia* 32, 139-156.
- Iglesias, I., Rodríguez-Rajo, F. J., Méndez, J., 2007. Evaluation of the different *Alternaria* prediction models on a potato crop in A Limia (NW of Spain). *Aerobiologia* 23, 27-34.
- Jedryczka, M., 2014. Aeromycology: studies of fungi in aeroplankton. *Folia Biologica et Oecologica* 10, 18–26.
- Jedryczka, M., Strzelczak, A., Grinn-Gofron, A., Nowak, M., Wolski, T., Siwulski, M., et al., 2015. Advanced statistical models commonly applied in aerobiology cannot accurately predict the exposure of people to *Ganoderma* spore-related allergies. *Agric. Forest Meteorol.* 201, 209–217.
- Jolliffe, I.T., 1986. Principal component analysis. Second Edition (2002). Springer. 1-518.
- Kauserud, H., Heegaard, E., Semenov, M.A., Boddy, L., Halvorsen, R., Stige, L.C., Sparks, T.H., Gange, A.C., Stenseth, N.C., 2010b. Climate change and spring-fruited fungi. *Proceedings of the Royal Society B. Biological Sciences.* 277, 1169–1177.
- Krouse, J.H., Chadwick, S.J., Gordon, B.R., Derebery, M.J., 2002. Allergy and Immunology. An Otolaryngic Approach, Lippincott Williams & Wilkins, USA.
- Kuhn, M., Johnson, K., 2013. Applied Predictive Modeling. Springer, New York, pp, 1–595.
- Lee, K., Young, K., Kang, H., Ahn, K., Min, K., Cha, B., 1997. Cucumber scab caused by *Cladosporium cucumerinum* in Korea. *Korean J. Plant Pathol.* 13, 288-294.
- Li, R., L. Peng, C. Road., 2011. Quantile Regression for Left-Truncated Semicompeting Risks Data. *Biometrics.* 7, 701–710.
- Mamgain, A., Roychowdhury, R., Tah, J., 2013. *Alternaria* pathogenicity and its strategic controls. *Res. J. Biol.* 1, 1-9.
- Meza, F., & Varas, E. 2000. Estimation of mean monthly solar global radiation as a function of temperature. *Agric. Forest Meteorol.* 100, 231-241.

- Mitakakis, T.Z., Barnes, Ch., Tovey R., 2001. Spore germination increases allergen release from *Alternaria*. *J. Allergy Clin. Immunol.* 107, 338-381.
- Nowicki, M., Nowakowska, M., Niezgoda, A., Kozik, E., 2012. *Alternaria* black spot of crucifers: symptoms, importance of disease, and perspectives of resistance breeding. *Veget. Crops Res. Bull.* 76, 5-19.
- Nowosad, J., 2016. Spatiotemporal models for predicting high pollen concentration level of *Corylus*, *Alnus*, and *Betula*. *Int. J. Biometeorol.* 60, 843–855.
- O'Connor, D. J., Sadyś, M., Skjøth, C. A., Healy, D. A., Kennedy, R., & Sodeau, J. R., 2014. Atmospheric concentrations of *Alternaria*, *Cladosporium*, *Ganoderma* and *Didymella* spores monitored in Cork (Ireland) and Worcester (England) during the summer of 2010. *Aerobiologia* 30, 397–411.
- Oteros, J., Galán, C., Alcázar, P., Domínguez-Vilches E., 2013. Quality control in bio-monitoring networks, Spanish Aerobiology Network. *Sci. Total Environ.* 443, 559–565.
- Oteros, J., García-Mozo, H., Alcázar, P., Belmonte, J., Bermejo, D., Boi, M., et al., 2015. A new method for determining the sources of airborne particles. *J. Environ. Manage.* 155, 212-218.
- Recio, M., del Mar Trigo, M., Docampo, S., Melgar, M., García-Sánchez, J., Bootello, L., et al., 2012. Analysis of the predicting variables for daily and weekly fluctuations of two airborne fungal spores: *Alternaria* and *Cladosporium*. *Int. J. Biometeorol.* 56, 983-991.
- Sabariago, S., Díez, A., Gutiérrez M., 2007. Monitoring of airborne fungi in Madrid (Spain). *Acta Bot. Croat.* 66, 117–126.
- Sadyś, M., Adams-Groom, B., Herbert, R.J., Kennedy, R., 2016. Comparisons of fungal spore distributions using air sampling at Worcester, England (2006–2010). *Aerobiologia* 32 (4), 619–634.

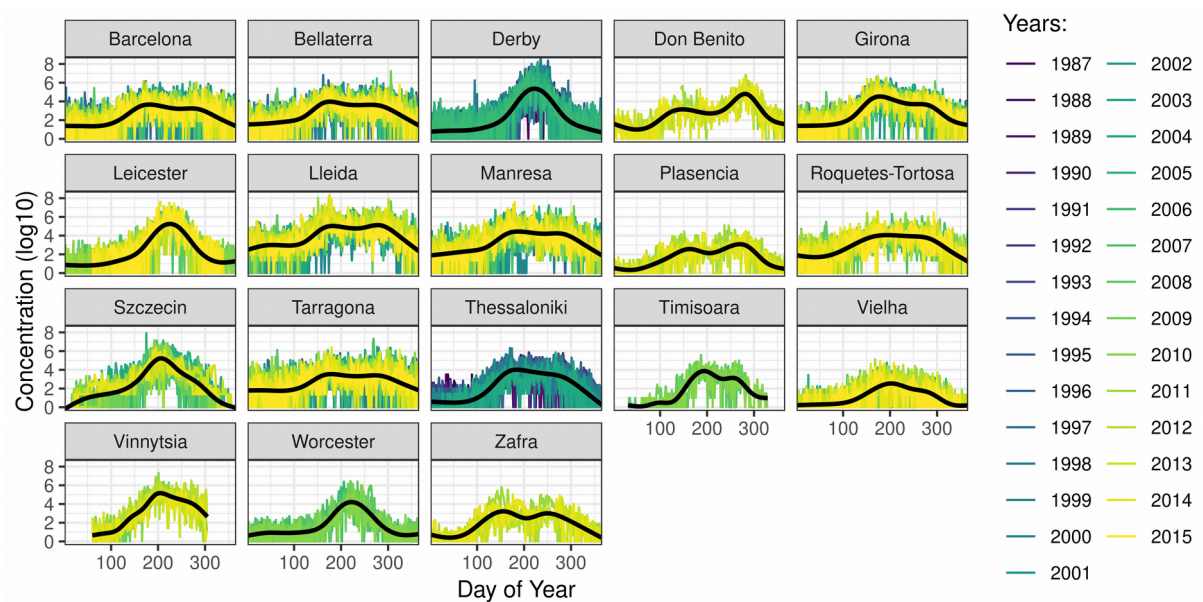


- Sindt, C., Besancenot, J-P., Thibaudon, M., 2016. Airborne *Cladosporium* fungal spores and climate change in France. *Aerobiologia* 32, 53–68.
- Skjøth, C.A., Sommer, J., Frederiksen, L., Gosewinkel Karlson, U., 2012. Crop harvest in Denmark and Central Europe contributes to the local load of airborne *Alternaria* spore concentrations in Copenhagen. *Atmos. Chem. Phys.* 12, 11107–11123.
- Stockmarr, A., Andreasen, V., Ostergård, H., 2007. Dispersal distances for airborne spores based on deposition rates and stochastic modeling. *Phytopathol.* 97, 1325–1330.
- Straatsma, G., Ayer, F., Egli, S., 2001. Species richness, abundance, and phenology of fungal fruit bodies over 21 years in a Swiss forest plot. *Mycol. Res.* 105, 515–523.
- Tariq, S.M., Matthews, S.M., Stevens, M., Hakim, E.A., 1996. Sensitization to *Alternaria* and *Cladosporium* by the age of 4 years. *Clin. Exp. Allergy.* 26, 794–798.
- Thomma, B.P., van Esse, H.P., Crous, P.W., de Wit, P.J., 2005. *Cladosporium fulvum* (syn. *Passalora fulva*), a highly specialized plant pathogen as a model for functional studies on plant pathogenic Mycosphaerellaceae. *Mol. Plant. Pathol.* 6, 379–393.
- Van Leuken, J.P.G., Swart, A.N., Havelaar, A.H., VanPul, A., VanderHoek, W., Heederik, D., 2016. Atmospheric dispersion modelling of bioaerosols that are pathogenic to humans and livestock – A review to inform risk assessment studies. *Microb. Risk Anal.* 1, 19–39.
- van Maanen, A., Xu, X-M., 2003. Modelling plant disease epidemics. *Eur. J. Plant Pathol.* 109, 669–682.

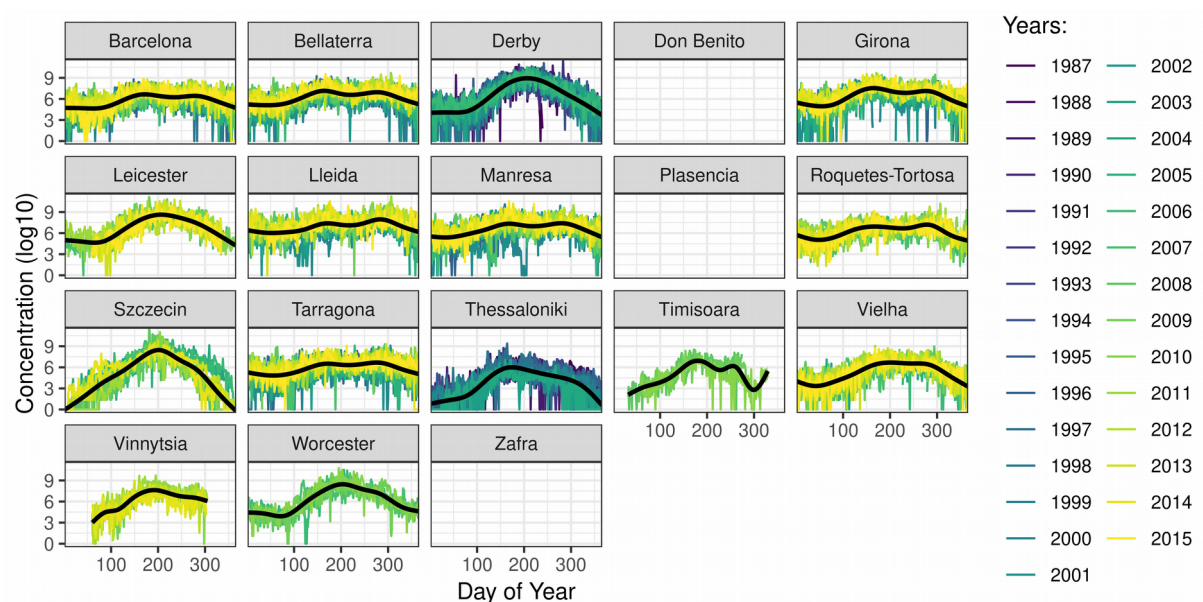
## Figures:



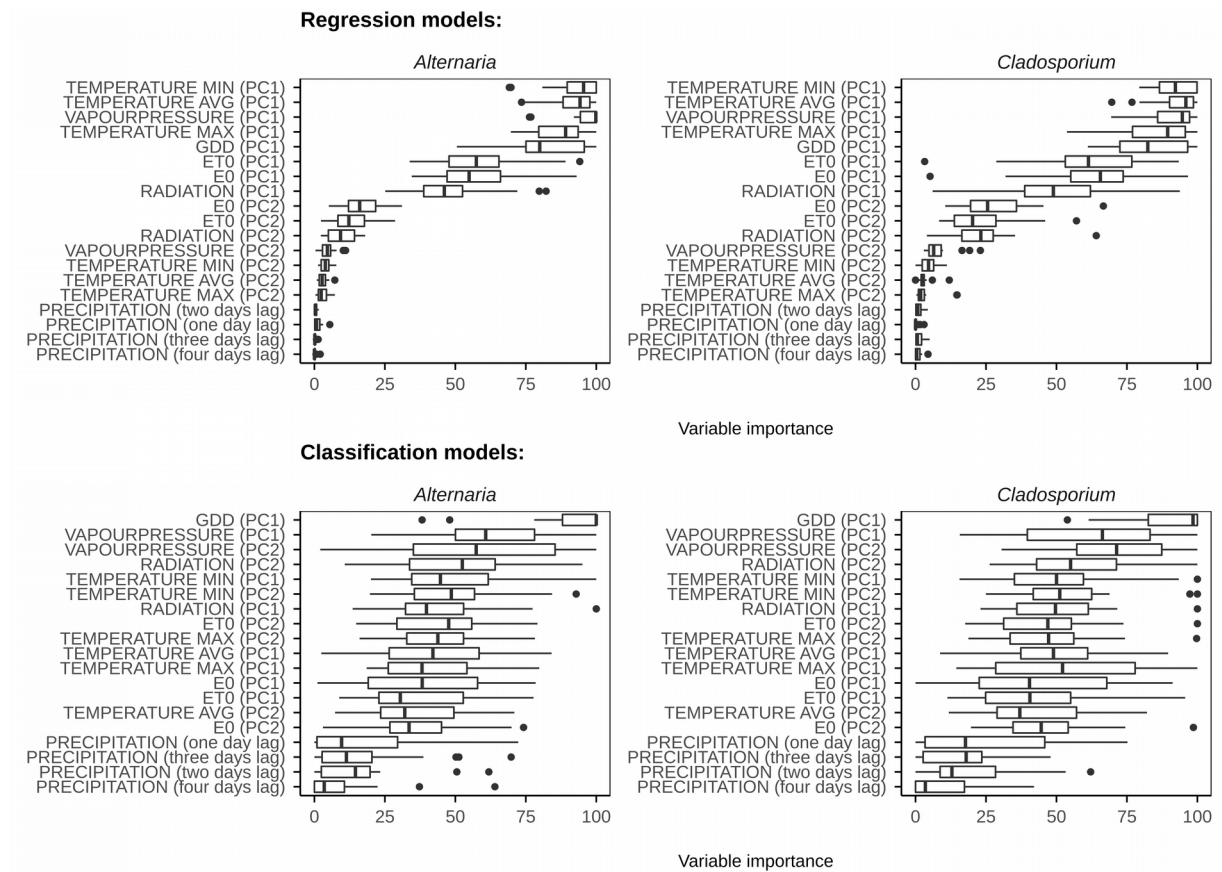
**Figure 1.** Location of measurement points against the background of biogeographical conditions in Europe (according to European Environmental Agency: [www.eea.eu.int](http://www.eea.eu.int) modified)



**Figure 2.** Daily concentrations of airborne *Alternaria* spores by day of year for all of the analyzed sites on a logarithmic scale ( $\log(1+x)$ ). Black lines are cubic spline smoothers for each site, whereas colored lines represent the actual data. Variable coloration of Figures represents different timespan of data, with darker colors indicating lack of more recent spore data.



**Figure 3.** Daily concentrations of airborne *Cladosporium* spores by day of year for all of the analyzed sites on a logarithmic scale ( $\log(1+x)$ ). Black lines are cubic spline smoothers for each site, whereas colored lines represent the actual data. Variable coloration of Figures represents different timespan of data, with darker colors indicating lack of more recent spore data.



**Figure 4.** Variable importance of each input variable for *Alternaria* and *Cladosporium* - regression models (top) and classification models (bottom). The variables are showed by the mean value of variable importance for all of the taxa in descending order

**Table 1**

List of monitoring sites, their exact locations, the years of sampling and the fungal genera examined

*Alt: Alternaria, Clad: Cladosporium*

Site	Country	Longitude	Latitude	Start	End	Missing years	Genera examined
Barcelona	Spain	2.164722	41.39361	1995	2015		<i>Alt &amp; Clad</i>
Bellaterra	Spain	2.107778	41.50056	1995	2015		<i>Alt &amp; Clad</i>
Derby	UK	-1.496360	52.93862	1987	2006		<i>Alt &amp; Clad</i>
Don Benito	Spain	-5.833333	38.96667	2011	2014		<i>Alt</i>
Girona	Spain	2.823056	41.98417	1996	2015	1997, 1998	<i>Alt &amp; Clad</i>
Leicester	UK	-1.122710	52.62316	2007	2015		<i>Alt &amp; Clad</i>
Lleida	Spain	0.595556	41.62806	1996	2015	1997, 1998	<i>Alt &amp; Clad</i>
Manresa	Spain	1.839722	41.72000	1996	2015	1998, 1999	<i>Alt &amp; Clad</i>
Plasencia	Spain	-6.086855	40.04897	2011	2014		<i>Alt</i>
Roquetes-	Spain	0.493056	40.82028	2006	2015		<i>Alt &amp; Clad</i>
Tortosa							
Szczecin	Poland	14.541667	53.44167	2004	2014		<i>Alt &amp; Clad</i>
Tarragona	Spain	1.243611	41.12000	1996	2015	1998, 1999	<i>Alt &amp; Clad</i>
Thessaloniki	Greece	22.950000	40.61667	1987	2004	2001, 2002	<i>Alt &amp; Clad</i>
Timisoara	Romania	21.231239	45.74722	2008	2010		<i>Alt &amp; Clad</i>
Vielha	Spain	0.797222	42.70222	2004	2015		<i>Alt &amp; Clad</i>
Vinnytsia	Ukraine	28.472074	49.23473	2009	2014		<i>Alt &amp; Clad</i>
Worcester	UK	-2.233333	52.18333	2006	2010		<i>Alt &amp; Clad</i>
Zafra	Spain	-6.416667	38.41667	2011	2014		<i>Alt</i>

**Table 2**

Summary of fungal spore concentration data per site. Minimum, maximum and average yearly spore concentration is given for the examined genera and their contribution % to the total annual fungal spore index.

Site	<i>Alternaria</i>				<i>Cladosporium</i>			
	min.	max.	mean	% of all fungal taxa	min.	max.	mean	% of all fungal taxa
Barcelona	4,967	12,925	9,310	1.8	84,636	339,962	204,770	38.9
Bellaterra	6,272	16,430	11,622	1.1	152,177	478,794	311,729	29.4
Derby	6,805	52,007	24,918	0.7	509,554	1,500,699	950,719	25.3
Don Benito*	8,789	14,878	11,774	25.2				
Girona	5,032	17,696	12,873	0.7	165,634	539,613	361,858	19.9
Leicester	16,522	31,831	21,519	0.7	523,894	1,088,832	852,812	26.4
Lleida	3,282	67,166	35,380	2.0	48,636	1,194,049	579,855	32.1
Manresa	7,571	37,859	20,981	1.6	144,782	660,514	473,642	35.6
Plasencia*	2,701	5,413	3,610	10.6				
Roquetes-Tortosa	10,086	18,080	13,983	0.9	214,236	445,494	323,300	20.0
Szczecin	10,333	31,694	17,553	12.3	276,520	813,364	498,877	72.0
Tarragona	5,580	13,986	10,149	1.1	73,741	323,176	221,793	23.1
Thessaloniki	2,962	18,925	10,383	11.2	24,637	138,678	74,279	70.5
Timisoara	4,714	5,944	5,176	4.5	68,372	175,750	121,541	92.2
Vielha	907	3,956	1,979	0.2	63,672	272,798	162,606	16.4
Vinnytsia	10,712	23,976	17,55	4.5	128,95	494,932	279,87	66.9

			6		1		3	
Worcester	6,022	9,279	7,776	0.7	437,171	902,136	660,895	62.5
Zafra*	3,535	9,383	5,516	14.7				
All sites	907	67,166	13,448		24,637	1,500,699	405,237	



**Table 3**

The final list of predictor variables

Predictor name	Description
GDD_PC1	First principal coordinate of cumulated GDD
PRECIPITATION_lag1	Average daily precipitation lagged by one day
PRECIPITATION_lag2	Average daily precipitation lagged by two days
PRECIPITATION_lag3	Average daily precipitation lagged by three days
PRECIPITATION_lag4	Average daily precipitation lagged by four days
E0_PC1	First principal coordinate of potential evaporation from a free water surface
E0_PC2	Second principal coordinate of potential evaporation from a free water surface
ET0_PC1	First principal coordinate of potential evapotranspiration from a crop canopy
ET0_PC2	Second principal coordinate of potential evapotranspiration from a crop canopy
RADIATION_PC1	First principal coordinate of average total global radiation
RADIATION_PC2	Second principal coordinate of average total global radiation
TEMPERATURE_AVG_PC1	First principal coordinate of average temperature
TEMPERATURE_AVG_PC2	Second principal coordinate of average temperature
TEMPERATURE_MAX_PC1	First principal coordinate of average maximum temperature
TEMPERATURE_MAX_PC2	Second principal coordinate of average maximum temperature
TEMPERATURE_MIN_PC1	First principal coordinate of average minimum temperature
TEMPERATURE_MIN_PC2	Second principal coordinate of average minimum temperature
VAPOURPRESSURE_PC1	First principal coordinate of average vapour pressure
VAPOURPRESSURE_PC2	Second principal coordinate of average vapour pressure

**Table 4.**

Performance of the regression and classification models. Regression models are described by the SMAPE (less is better), classification models are described by the balanced accuracy (more is better)

Group	Regression Models		Classification Models	
	<i>Alternaria</i>	<i>Cladosporium</i>	<i>Alternaria</i>	<i>Cladosporium</i>
Barcelona	0.72	0.59	0.72	0.50
Bellaterra	0.65	0.57	0.76	0.55
Girona	0.62	0.53	0.81	0.80
Lleida	0.63	0.63	0.84	0.75
Manresa	0.68	0.66	0.78	0.76
Roquetes-Tortosa	0.61	0.56	0.78	0.61
Vielha	0.90	0.63	0.50	0.50
FIRST GROUP	0.73	0.67	0.77	0.68
Derby	0.82	0.55	0.89	0.86
Leicester	0.73	0.54	0.90	0.81
Szczecin	0.69	0.60	0.88	0.83
Vinnytsia	0.60	0.56	0.85	0.75
Thessaloniki	0.65	0.69	0.80	0.99
Worcester	0.75	0.53	0.84	0.81
Timisoara	0.66	0.61	0.66	0.75
SECOND GROUP	0.75	0.71	0.85	0.83
Don Benito	0.56		0.88	
Plasencia	0.60		0.50	
Zafra	0.66		0.77	
THIRD GROUP	0.69		0.80	
ALL	0.76	0.73	0.80	0.78