1 **Airborne *Alternaria* and *Cladosporium* fungal spores in**

2 **Europe: forecasting possibilities and relationships with**

3 **meteorological parameters**

4

5 Agnieszka Grinn-Gofroń[a], Jakub Nowosad[b], Beata Bosiacka[a], Irene Camacho[c], Catherine

6 Pashley[d], Jordina Belmonte[e,f], Concepción De Linares[e,f], Nicoleta Ianovici[g], Jose María Maya

7 Manzano[h], Magdalena Sadyś[i], Carsten Skjoth[i], Victoria Rodinkova[j], Rafael Tormo-Molina[k],

8 Despoina Vokou[l], Santiago Fernández-Rodríguez[k], Athanasios Damialis[l,m*]

9

10 [a]Department of Plant Taxomony and Phytogeography, Faculty of Biology, University of

11 Szczecin, Szczecin, Poland, agofr@univ.szczecin.pl

12 [b]Space Informatics Lab, University of Cincinnati, 219 Braunstein Hall, Cincinnati, OH 45221,

13 USA, nowosad.jakub@gmail.com

14 [c]Madeira University, Faculty of Life Sciences, Campus Universitário da Penteada, 9000-390

15 Funchal, Portugal, camire@uma.pt

16 [d]Institute for Lung Health, Department of Infection, Immunity and Inflammation, University

17 of Leicester, Leicester LE1 7RH, UK, chp5@leicester.ac.uk

18 [e]Unidad de Botánica, Facultad de Ciencias, Universidad Autónoma de Barcelona, Barcelona,

19 Spain, Jordina.Belmonte@uab.cat

20 [f]Botany Unit, Dept. Of Animal Biology, Plant Biology and Ecology, Universitat Autònoma de

21 Barcelona, Bellaterra (Cerdanyola del Vallès), Spain

22 [g]West University of Timisoara, Department of Biology, Faculty of Chemistry-Biology-

23 Geography, Romania, nicole_ianovici@yahoo.com

24  hUniversity of Extremadura, Department of Plant Biology, Ecology and Earth Sciences,

25  Faculty of Science, Avda Elvas s/n, 06071 Badajoz, Spain, jmmaya@unex.es

26  iUniversity of Worcester, National Pollen and Aerobiology Research Unit, Henwick Grove,

27  Worcester WR2 6AJ, United Kingdom, magdalenasadys@wp.pl

28  jNational Pirogov Memorial Medical University, Vinnytysia, Ukraine, vikarodi@gmail.com

29  kDepartment of Construction, School of Technology, University of Extremadura, Avda. de la

30  Universidad s/n, Cáceres, Spain, santiferro@unex.es

31  lDepartment of Ecology, School of Biology, Aristotle University of Thessaloniki,

32  Thessaloniki, GR-54124, Greece, vokou@bio.auth.gr

33  mChair and Institute of Environmental Medicine, UNIKA-T, Technical University of Munich

34  and Helmholtz Zentrum München, Germany - German Research Center for Environmental

35  Health, Neusaesser Str. 47, DE-86156, Augsburg, Germany, thanos.damialis@tum.de

36

37  *Corresponding author: Dr. Athanasios Damialis, Chair and Institute of Environmental

38  Medicine, UNIKA-T, Technical University of Munich, Neusaesser Str. 47, 86156 Augsburg,

39  Germany, E-mail: thanos.damialis@tum.de

40

41  **Abstract**

42  Airborne fungal spores are prevalent components of bioaerosols with a large impact on

43  ecology, economy and health. Their major socioeconomic effects could be reduced by

44  accurate and timely prediction of airborne spore concentrations. The main aim of this study

45  was to create and evaluate models of *Alternaria* and *Cladosporium* spore concentrations

46  based on data on a continental scale. Additional goals included assessment of the level of

47  generalization of the models in space and description of the main meteorological factors

48  influencing fungal spore concentrations.

49  Aerobiological monitoring was carried out at 18 sites in six countries across Europe over 3 to

50  21 years depending on site. Quantile random forest modelling was used to predict spore

51  concentrations values. Generalization of the *Alternaria* and *Cladosporium* models was tested

52  using (i) one model for all the sites, (ii) models for groups of sites, and (iii) models for

53  individual sites.

54  The study revealed the possibility of reliable prediction of fungal spore levels using gridded

55  meteorological data. The classification models also showed the capacity for providing larger

56  scale predictions of fungal spore concentrations. Regression models were distinctly less

57  accurate than classification models due to several factors, including measurement errors and

58  distinct day-to-day changes of concentrations. Temperature and vapour pressure proved to be

59  the most important variables in the regression and classification models of *Alternaria* and

60  *Cladosporium* spore concentrations.

61  Accurate and operational daily-scale predictive models of bioaerosol abundances contribute to

62  the assessment and evaluation of relevant exposure and consequently more timely and

63  efficient management of phytopathogenic and of human allergic diseases.

64

65  **Keywords**

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

67  molds

68

69  **Highlights**

3

70 • No operational forecasting model for allergenic fungal spore exposure exists in Europe

71 • Potential exposure in Europe was assessed and predicted for 2 major allergenic fungi

72 • Random forest modelling was applied to >7,000 daily time series

73 • Air temperature and vapour pressure were the most significant variables

74 • Classification models showed higher capacity for large-scale spore predictions

75

76

77 **1. Introduction**

78    Fungal spores are one of the most prevalent components of bioaerosols, found across a

79 wide range of biogeographic regions over long time periods each year. The primary source of

80 spore emissions are the substrates on which fungi grow, such as plants, soil and decaying

81 organic matter. This means the majority of fungal spores in the air originate from farms, forest

82 stands and decomposing plant material (Bowers et al., 2013). They will remain airborne for

83 variable amounts of time and will be transported over distances ranging from a few

84 centimeters to hundreds of kilometers (Ansari et al., 2015; Heald and Spracklen, 2009;

85 Stockmarr et al., 2007).

86    *Alternaria* and *Cladosporium* are ubiquitous asexually reproducing fungal genera that

87 produce spores, known as conidia, which are readily airborne. Both genera contain plant

88 pathogens (Carlile et al., 2007; Chaerani and Voorrips, 2006; Lee et al., 1997; Nowicki et al.,

89 2012; Thomma et al., 2005), many of which also produce phytotoxic metabolites that affect

90 mammalian cells (De Lucca, 2007; Friesen et al., 2008; Mamgain et al., 2013).  Fungal spores

91 are also important aeroallergens, causing adverse health effects (Krouse et al., 2002).

92    Aerobiological surveys reported *Alternaria* as one of the most prevalent airborne

93 fungal types and an important aeroallergen (Budd, 1986; Mitakakis et al., 2001). Also,

94 *Cladosporium* is frequently reported as the most abundant aeroallergen and the second most

4

95    allergenic fungal type worldwide (Tariq et al., 1996). Allergy to fungi from both genera has

96    been responsible for hospital admissions due to severe asthma attacks in sensitized

97    individuals, particularly among children (Bush and Prochnau, 2004; Dales et al., 2000). Due

98    to their ubiquitous nature (Damialis et al., 2017; Jędryczka, 2014) exposure to these

99    aeroallergens is literally inevitable during their dispersion season.

100    It is important to assess the potential risks of the atmospheric presence of fungal

101    spores, and more research is needed to evaluate this (Beggs, 2004; Crameri et al., 2014).

102    There is limited data on long-term trends in airborne fungal spore abundance, in part due to a

103    lack of representative fungal spore data sets longer than a decade across the globe. Airborne

104    fungal spores as allergens have received comparatively less attention than pollen, and

105    associated public health consequences are likely to have been underestimated (Damialis et al.

106    2015).

107    Sporulation and dispersion of fungi are influenced by several meteorological factors,

108    including air temperature, relative humidity, precipitation, atmospheric turbulence, wind

109    speed,  and UVB radiation (Al-Subai, 2002; Carlile et al., 2007; Cecchi et al., 2010; Straatsma

110    et al., 2001). However, the exact influence of climatic variability on fungal ecology at a larger

111    scale is still not understood. There are indications that changing climate may lead to

112    alterations in phenology (Corden et al., 2003; Gange et al., 2007; Kauserud et al., 2010) and

113    dynamics of fungal communities (Gange et al., 2011).

114    This complexity of biological and ecological processes represent one of the key

115    problems of modelling biological systems. There are inconsistencies regarding what drives

116    and controls the distribution of fungal bioaerosols both at a local and regional scale. Multiple

117    meteorological factors may alter the spatiotemporal distribution of *Alternaria* and

118    *Cladosporium* spores (Corden et al., 2003; Damialis and Gioulekas, 2006; De Linares et al.,

119    2010; Escudero et al., 2011; Iglesias et al., 2007; Recio et al., 2012; Sindt et al., 2016; Skjøth

5

120  et al., 2016). The relationships between fungal development and environmental factors,

121  including major climatic variables, are often the only component used for disease forecasting

122  systems (van Maanen and Xu, 2003).

123  Atmospheric dispersion models have been used to describe the spatiotemporal

124  dispersal of fungal pathogens (Burie et al., 2012; Oteros et al., 2015; Stockmarr et al., 2007;

125  van Leuken et al., 2016). Additionally, descriptive, predictive or conceptual modelling of

126  airborne fungal spores concentration is another promising tool, albeit challenging (Grinn-

127  Gofroń and Bosiacka, 2015; Grinn-Gofroń et al., 2018; Iglesias et al., 2007; Jędryczka et al.,

128  2015). At a minimum, high-resolution data on meteorological and geographic variables should

129  be included in such models. It is important to include and analyse all associated factors, so

130  that the resulting models can accurately describe the complex environmental inter-

131  dependencies.

132  Given the widespread, ecological, economical and health impact of spores from

133  *Alternaria* and *Cladosporium*, the main goal of the present study was to build models that

134  generalize beyond the observed data and are capable of estimating the spatiotemporal

135  distribution and concentration of fungal spores on a broad scale and to address the factors that

136  influence the fungal spatiotemporal patterns. Created models could help to answer several

137  questions - (i) is it possible to make a reliable prediction of fungal spore concentrations using

138  only one model for all sites, (ii) is there a difference in quality between one model for all the

139  sites, models for the groups of sites, and models for individual sites, (iii) what are the main

140  meteorological factors influencing fungal spore concentrations, and (iv) do these factors differ

141  between created models (analysed sites)?

142  **2. Materials and Methods**

143  **2.1. Aerobiological data**

144 The taxa selected for this study were those of *Alternaria* and *Cladosporium*. These comprise
145 some of the most well studied fungal taxa worldwide due to their allergenic and
146 phyopathogenic properties. In the current research, the spore abundance ranges greatly from
147 >15% to over 96% of the total annual spore index per site (Table 2). This variation alone
148 challenges for the elaboration of a universal forecasting model in Europe. But even if we had
149 added the rest of the fungal diversity per site, this would have not most probably provided any
150 additional insight to either ecological/biological processes or health impacts, as responses of
151 different fungi are highly individualistic and their sensitivity (or not) to environmental stress
152 (i.e. climate change) would not be reflected reliably (e.g. Damialis et al. 2015). Also, in most
153 stations in the world, mainly these two types alone are counted. So as to obtain similar data in
154 such a large-scale spatial study design and consequently comparable results, we did not
155 extend to the investigation of additional fungal taxa.

156    Fungal spores were collected in the frame of long-term aerobiological monitoring in
157 18 sites from six countries across Europe and for a time span of 1987-2015 (Fig. 1, Table 1).
158 *Alternaria* spore concentrations were measured at all sites and *Cladosporium* spore
159 concentrations at 15 out of the 18 (Table 1).

160    In brief, for this research, all data providers have been inquired regarding major
161 changes in methodological procedures and no significant alteration has been reported.
162 Moreover, in all participating countries, microscopic identification of airborne fungal spores
163 has been conducted by experts with long-standing experience in such techniques. Therefore,
164 in all stations, the same method of collection and analysis was used. Standard sampling,
165 processing and analysis techniques were followed, according to the recommendations of the
166 European Aerobiology Society (Frenguelli, 2003) and the British Aerobiology Federation
167 (1995). At each location, a Hirst-type volumetric spore trap was used to sample airborne
168 fungal spores (Hirst, 1952), which is considered the gold-standard device for sampling

7

169 airborne particles of biological origin (Galán et al., 2014). Samples were collected and

170 analysed weekly, applying standard methods for sample processing and microscopic

171 identification (e.g. British Aerobiology Federation, 1995; Galán et al., 2014; Grant Smith,

172 1984). Final measurements referred to daily resolution and were expressed as concentrations

173 of fungal spores per cubic meter of air on a given date.

174 **2.2. Meteorological data**

175 Eight meteorological factors were included as co-factors in the data analysis, namely

176 maximum temperature, minimum temperature, average temperature, vapour pressure, sum of

177 precipitation, potential evaporation from a free water surface, potential evapotranspiration

178 from a crop canopy, and total global radiation. These were acquired from the AGRI4CAST

179 Interpolated Meteorological Database (Baruth et al., 2007).

180 Based on maximum ($T_{max}$) and minimum ($T_{min}$) temperature, an additional ninth

181 parameter was included, growing degree days (GDD). Cumulative GDD is an indicator

182 measuring a heat accumulation, and as a proxy, can represent plant and fungal development.

183 GDD value is calculated as follows:

184 $$GDD = \frac{T_{max} + T_{min}}{2} - T_{base}$$

185 Values of GDDs were accumulated starting from January 1. GDD do not accumulate when the

186 daily mean temperature (Tmax + Tmin/2) is lower or equal to the base temperature. Value of

187 $T_{base}$ was set to 5. No relevant previous information on base temperatures for fungal spores

188 existed, so we assumed that these would be similar to later-flowering plant species (being

189 abundant mainly during May-August) and which could be found across a variety of latitudes

190 and climates in Europe; therefore, we set the base temperature at 5, in a similar manner to

191 studies on grass pollen (Emberlin 1993; Frenguelli et al., 1989).

192    All the the abovementioned factors were considered either based on their availability,

193    or because of the focus of the current study, or based on previous literature on the topic. For

194    instance, relative humidity data are not available via AGRI4CAST as these cannot be

195    corrected for altitude differences among sites within the target climatic grid cell. Regarding

196    other factors, like wind vectors (speed, persistence and direction), these were excluded from

197    this analysis, as this would make more sense in a smaller-scale temporal data processing,

198    when the intermittent nature of wind would be possible to take into account. Finally, there

199    have been previously published reports of particular meteorological factors in site-specific

200    studies, included in the current analysis, proven to be the most decisive for obtaining accurate

201    and reliable forecasts (e.g. Damialis and Gioulekas, 2006).

202    **2.3. Models**

203    Partial autocorrelation function (Durbin, 1960) was applied independently to

204    *Alternaria* and *Cladosporium* daily counts to check for the temporal autocorrelation of spore

205    data. This summarizes the relationship between an observation $x_t$ and observations with

206    lagged time steps (days) removing the impact of the values at all shorter lags.

207    Redundancy among the lagged values of meteorological predictors was explored using

208    principal coordinates analysis (PCA) (Jolliffe, 1986). PCA transforms a number of correlated

209    variables into a smaller set of uncorrelated variables. The role of PCA in this study was to

210    reduce the dimension of the data, reduce computational time of models building, minimize

211    spurious effects of single lags and ease interpretation of the final models.

212    Two main types of modeling techniques were used, regression and classification. The

213    Quantile Random Forest method was used to create regression models. This is a

214    generalization of Random Forests that infer the full conditional distribution of a response

215    variable (Li et al., 2011). We decided not to predict mean value, but rather a median value,

216    because of a non-symmetrical distribution of fungal spore values. Daily spore concentrations

9

217  of *Alternaria* or *Cladosporium* were used as dependent variables (model output) in regression

218  models. Breiman's Random Forest (Breiman, 2001) was used in classification models. So as

219  to eliminate high levels of statistical noise, daily spore concentrations were divided into two

220  levels, low and high, according to the thresholds of *Alternaria* and *Cladosporium* allergens to

221  evoke allergic symptoms (Gravesen, 1979). *Alternaria* values lower than 100 spores were

222  considered as 'low' and beyond that threshold they were characterized as 'high'. For

223  *Cladosporium*, this threshold was 3,000 spores. As airborne fungal spore measurements could

224  still exhibit a huge disparity in the frequencies of the observed classes, we adopted an

225  optimizing probability threshold technique (Kuhn and Johnson, 2013; Nowosad, 2016). In this

226  approach, alternative cutoffs for the predicted probabilities were determined using

227  resampling. Sensitivity (Sens), specificity (Spec), positive predictive value (Ppv), and

228  negative predictive value (Npv) were calculated for 20 different threshold values. For each

229  model, optimal threshold value was established minimizing the distance between Sens, Spec,

230  Ppv, Npv and the value of 1. This value indicates the best possible performance.

231      Attempting to also reduce the spatial variability among monitoring sites across

232  Europe, analyzed sites were divided into three groups based on the annual temporal changes

233  of fungal spore concentrations and on data availability (Table 1), as follows: northeastern

234  Spain (Barcelona, Bellaterra, Girona, Lleida, Manresa, Roquetes-Tortosa, Tarragona, Vielha),

235  western Spain (Don Benito, Plasencia, Zafra) and non-Spanish sites (Derby, Leicester,

236  Szczecin, Thessaloniki, Timisoara Vinnytsia, Worcester). The grouping was based on the

237  seasonality and multi-modality of airborne spore concentrations, i.e. the first group exhibited

238  the highest seasonality and normality of data, with longer seasons and fewer outliers. The

239  third group presented a bi- or multi- modal yearly pattern, whilst the second group included

240  those sites with only *Alternaria* spore measurements.

10

241      The modeling was performed in three ways: 1) per site, 2) per group of sites as

242 mentioned above, and 3) an integrated model for all sites combined. This resulted in 22

243 combinations for *Alternaria* (18 sites, 3 groups of sites, and 1 whole dataset) and 18

244 combinations for *Cladosporium* (15 sites, 2 groups of sites, and 1 whole dataset) generating a

245 total of 40 regression and 40 classification models.

246 **2.4. Validation metrics**

247      The accuracy of all models was assessed using a repeated k-fold cross-validation

248 (Kuhn and Johnson, 2013). Regression models were evaluated using the Symmetric Mean

249 Absolute Percentage Error (SMAPE). Thus, accuracy measure was based on relative errors

250 instead of standard metrics like Root Mean Square Error (RMSE) or Mean Absolute Error

251 (MAE), which was particularly important in our study because of high among-site data

252 variability. Classification modelling results were characterized using balanced accuracy. This

253 metric is calculated as sensitivity + specifity/2. It is a better measurement in the case of

254 imbalanced datasets as it gives the same weight for correctly predicted cases with low

255 concentration and correctly predicted cases with high concentration. Influence of the

256 predictors was determined using a scaled permutation importance (mean decrease in

257 accuracy) (Breiman, 2001).

258 **3. Results**

259 **3.1. Spatiotemporal variability of spore concentrations**

260      Fungal spore concentrations differed between *Alternaria* and *Cladosporium*, between

261 sites, and between years of measurements (Table 2, Figs. 2, 3). The annual sum of *Alternaria*

262 daily concentrations varied between 907 (Vielha) and 67,166 (Lleida) and had a mean value

263 of 13,448. In addition to Lleida having the highest average annual sums of *Alternaria*, the

264 values here were also the most variable between years. Other sites with substantial changes

11

265    between years were Derby, Thessaloniki, Vielha, and Girona. The annual sum of the daily

266    concentrations of *Cladosporium* varied between 24,637 (Thessaloniki) and 1,500,699 (Derby)

267    and had a mean value of 405,237. Annual spore concentrations in Lleida were also the most

268    changeable for *Cladosporium*. Overall, the order of average values for all sites was similar for

269    *Alternaria* and *Cladosporium* with two exceptions, Worcester and Thessaloniki. The annual

270    sum of the daily concentrations of *Alternaria* in Worcester were relatively small and

271    invariable, whilst *Cladosporium* values were high and changeable. In Thessaloniki *Alternaria*

272    levels were moderate and varied substantially, whilst *Cladosporium* annual values were small

273    and more consistent.

274       For *Alternaria,* sites could be split into two groups based on the time course of the

275    season: (i) sites with period of high concentration and period of lower concentrations (such as

276    Barcelona, Lleida, etc.), (ii) sites with one period of high concentrations and period with

277    absence of *Alternaria* spores (such as Szczecin, Thessaloniki and Vielha) (Fig. 2). Time

278    course of *Cladosporium* spore concentrations is more heterogeneous, with probably two main

279    groups: (i) Spanish sites with period of high concentration and period of moderate

280    concentrations, (ii) sites with one period of high concentration and period with low

281    concentrations or absence of spore concentrations (such as Szczecin, Derby, Vinnytsia).

282    Additionally, a third group, consisting of Don Benito, Plasencia, and Zafra, was separated due

283    to missing data of *Cladosporium* (Fig. 3).

284

285    **3.2. Predictor variables**

286    Based on the spatiotemporal analysis, lagged daily values of nine meteorological parameters

287    between 1-15 days were created. A principal component analysis (PCA) for each of the

288    parameters was run and the results of the PCA gave an insight into the variability of

289    predictors. In most of the cases, the two first components explained the majority of variations.

290  The cumulative value was between 0.85 for radiation and 0.94 for average temperature in

291  these variables. The first component expressed the value of a given parameter (large values of

292  loadings for all of the lags), whilst the second parameter expressed the temporal changes of a

293  given parameter (the largest, positive value for the first lag, and the lowest, negative value for

294  the last lag). In case of cumulated GDD, we used only the first component, which explained

295  0.95 of variation. Precipitation values were the most changeable, and the first two components

296  had a cumulative variance of 0.22. Therefore based on the autocorrelation plot, we used raw

297  values of precipitation for lags between 1 and 4 days.

298  The final group of predictors consisted of 19 variables - the first principal component

299  of cumulated GDD, one to four day lags of precipitation, and the first two principal

300  components of the rest of the meteorological parameters (Table 3).

301  **3.3. Regression models**

302  **3.3.1. Performance of the models**

303  The final 40 regression models (22 for *Alternaria* and 18 for *Cladosporium*) were built

304  and results compared using Symmetric Mean Absolute Percentage Error (SMAPE) (Table 4).

305  These values ranged between 0.56 and 0.90 for *Alternaria* (average value of 0.69), and 0.53

306  and 0.73 for *Cladosporium* (average value of 0.61). For all of the sites, the models gave a

307  value of 0.76 for *Alternaria* and a value of 0.73 for *Cladosporium*. Those values were slightly

308  higher than values for separate groups and distinctly higher than values for most of the

309  individual sites. Similarly, most of the models for the site groups performed worse than the

310  models for the individual sites. *Alternaria* model had a SMAPE value of 0.73 for the first

311  group (average for the models for the individual sites is this group was 0.67), 0.75 for the

312  second group (average of 0.74), and 0.69 for the third group (average of 0.61). These

313  differences were higher in *Cladosporium* models with SMAPE of 0.67 for the first group

13

314    (average of 0.60) and 0.71 for the second group (average of 0.59). Only in four *Alternaria*

315    models, Derby, Worcester, and Vielha, values of SMAPE were higher than values for models

316    of groups of sites.

317    **3.3.2. Variable importance**

318    The same set of variables seemed to influence the models of both taxa (Fig. 4). The

319    most important predictors were the first principal coordinates of vapor pressure and

320    temperatures (minimum, average, maximum and cumulated GDD). They were followed by

321    the first principal coordinates of evapotranspiration, evaporation, and radiation. The second

322    principal coordinates of those parameters had small importance. The smallest values were

323    observed for lagged values of precipitation.

324    **3.4. Classification models**

325    **3.4.1. Performance of the models**

326    The final classification models were compared using the balanced accuracy metric

327    (Table 4). Average value of balanced accuracy were 0.78 for *Alternaria* and 0.73 for

328    *Cladosporium*. These values varied among sites between 0.50 and 0.90 for *Alternaria* and

329    between 0.50 and 0.99 for *Cladosporium*. Models for all the sites gave similar results of 0.80

330    and 0.78. The *Alternaria* model gave a balanced accuracy value of 0.77 for the first group of

331    sites compared to the average of individual site's models of 0.75, a value of 0.85 for the

332    second group of sites comparing to the average of individual site's models of 0.81, and a value

333    of 0.80 for the third group of sites compared to the average of individual site's models of 0.71.

334    Values of a balanced accuracy for the models of groups of sites were also higher for

335    *Cladosporium*. The first group had a value of 0.68 (average for individual sites was 0.66), and

336    the second group had a value of 0.85 (average for individual sites was 0.80). About 44% (15

14

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

360 deeper comprehension of the dynamics of these fungal taxa. The survey revealed distinct

361 relationships between spore concentrations and sites and years of sampling. The overall spore

362 frequency and the annual sum of the daily concentrations of *Cladosporium* was higher than

363 *Alternaria*'s, a finding also reported previously in different regions of the world. In Cartagena

364 (Spain), *Cladosporium* represented 62.2% of the total spore count and *Alternaria* only 5.3%,

365 however, *Alternaria* was still the second most abundant fungal type (Elvira-Rendueles et al.,

366 2013). Likewise, in Bursa (Turkey), *Cladosporium* represented 88.1% of the total spore count

367 followed by *Alternaria* at 4.9% (Ataygul et al., 2007). A similar trend was found in

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

369 (2007).

370　　　　Two types of models, regression and classification, were built as a part of this study.

371 The goal of the first one was to predict values of the fungal spore concentration in the studied

372 sites, while the classification models were created to predict high levels of the fungal spore

373 concentrations. Classification models were more accurate than regression models. This is due

374 to a number of factors. Concentration values of *Alternaria* and *Cladosporium* were classified

375 into two groups prior to the modelling. This procedure generalises the fungal spore values

376 and, therefore, changes in values. As a result, it gives more general trends instead of showing

377 the local trends and day-to-day differences. Moreover, the obtained values of concentration

378 are an estimation of the whole population and are prone to random and systematic (bias)

379 errors (Comtois et al., 1999; Oteros et al., 2013). The benefit of using a classification method

380 is that it reduces the influence of methodological differences between sites, such as the

381 relative position of samplers and the heights of buildings, or instrumental and human errors.

382　　　　The life cycles of many fungal pathogens are strongly determined by weather. The

383 airborne spore concentrations are affected by biological factors (reproduction and survival),

384 weather parameters, land use, resource availability and competition (Boddy et al., 2014).

16

385 Dispersal and circulation are highly influenced by wind and rainfall, while germination and

386 infection rates are often dependent upon liquid water on the plant surface (sometimes high

387 relative humidity) and species-specific optimal temperature ranges. In this study the most

388 important meteorological variables were the same for both *Alternaria* and *Cladosporium.* This

389 shows that despite the differences in the values of spore concentration, both taxa are mostly

390 affected by the same meteorological factors, such as temperature (minimum, average,

391 maximum and cumulated GDD) and vapour pressure, which determines the water content in

392 the air (related to air humidity).

393    In many aerobiological studies, temperature and relative humidity of the air are

394 meteorological parameters that significantly influence concentrations of *Cladosporium* and

395 *Alternaria* spores, with temperature being positively associated and relative humidity

396 negatively associated (e.g. Grinn-Gofroń and Strzelczak 2009; O'Connor et al. 2014; Sadyś et

397 al. 2016; Ianovici 2016; Almeida et al. 2018)

398    Other variable - global radiation includes both the direct solar radiation and the diffuse

399 radiation resulting from reflected or scattered sunlight, and can be considered as a function of

400 temperature (Meza and Varas 2000). One of the solar radiation components, UVB radiation,

401 is reported to affect the survival of airborne fungal spores during movement over long

402 distances through the atmosphere (Al-Subai 2002). This predictor showed moderate

403 importance in our models.

404    Rain has been often cited as being one of the most influential factors in reducing

405 airborne pollen (e.g. Damialis et al., 2005). However, precipitation was identified as the least

406 important factor in our models. There are two main reasons behind the small effect of

407 precipitation on the fungal spore models. Firstly, variation of daily precipitation values does

408 not closely correlate with the values and levels of *Alternaria* and *Cladosporium*

409 concentrations, both of which are regarded as dry weather spores. For example, it is possible

410　to have a rainless day and a high fungal spore concentration (the middle of the season) and to

411　have a rainless day without airborne fungal spores (off-season). Secondly, precipitation is

412　characterized by high variability, which could not be captured in a daily timescale (sum of

413　precipitation). To better understand the effects of rainfall on fungal spore diversity and

414　abundance, and on circulation patterns, finer resolution data (hourly scale), with different

415　statistical techniques (i.e. artificial intelligence models) would be required.

416　**5. Conclusions**

417　• Classification models were more accurate than regressions for *Alternaria* and
418　　*Cladosporium* fungal spores.

419　• Regression models gave better results for individual sites compared to grouped sites,
420　　resulting potentially from strong effects from local meteorological conditions.

421　• Classification models gave better results in grouped sites rather than for individual
422　　sites, thus, displaying the capacity for accurately providing larger scale predictions of
423　　fungal spore concentrations (compared to the more localized regression models).

424　• Temperature (in the form of minimal, average, maximum temperature, and cumulated
425　　GDD) and vapour pressure were the most important variables in models of *Alternaria*
426　　and *Cladosporium*, while radiation and daily sum of precipitation had a smaller impact
427　　on the models.

428

429　**References:**

430　Almeida, E., Caeiro, E., Todo-Bom, A., Ferro, R., Dionísio, A., Duarte, A., Gazarini, L.,
431　　2018. The influence of meteorological parameters on *Alternaria* and *Cladosporium* fungal
432　　spore concentrations in Beja (Southern Portugal): preliminary results. Aerobiologia 34,
433　　219–226

18

434  Al-Subai, A.A.T., (2002). Air-borne fungi at Doha, Qatar. Aerobiologia 18, 175–183.

435  Ansari, T.U., Valsan, A.E., Ojha, N., Ravikrishna, R., Narasimhan, B., Gunthe, S.S., 2015.

436      Model simulations of fungal spore distribution over the Indian region. Atmos. Environ.

437      122, 552–560.

438  Ataygul, E., Celenk, S., Canitez, Y., Bicakci, A., Malyer, H., Sapan, N., 2007. Allergenic

439      Fungal Spore Concentrations in the Atmosphere of Bursa, Turkey. J. Biol. Environ. Sci. 1,

440      73–79.

441  Baruth, B., Genovese, G., Leo, O., 2007. CGMS Version 9.2 - User Manual and Technical

442      Documentation. Office for Official Publications of the European Communities,

443      Luxembourg.

444  Beggs, P.J., 2004. Impacts of climate change on aeroallergens. Past and future. Clin. Exp.

445      Allergy 34, 1507–1073.

446  Berman, D. 2011. Climate change and aeroallergens in South Africa. Curr. Allergy Clin.

447      Immunol. 24, 65–71.

448  Boddy, L., Büntgen, U., Egli S., Gange, A.C., Heegaad, E., Kirk, P.M. et al., 2014. Climate

449      variation effects on fungal fruiting. Fungal Ecol. 10, 20–33.

450  Bowers, R.M., Clements, N., Emerson, J.B.E., Wiedinmyer, C., Hannigan, M. P., Fierer, N.,

451      2013. Seasonal Variability in Bacterial and Fungal Diversity of the Near-Surface

452      Atmosphere. Environ. Sci. Technol. 47, 12097–12106.

453  Breiman, L. 2001. Random Forests. University of California, accessed on 30th May 2018,

454      https://www.stat.berkeley.edu/~breiman/randomforest2001.pdf

455  British Aerobiology Federation, 1995. Airborne pollens and spores. A guide to trapping and

456      counting: National Pollen and Hayfever Bureau, Rotherham, United Kingdom.

457  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.

20

482  Damialis, A., Gioulekas, D., 2006. Airborne allergenic fungal spores and meteorological

483      factors in Greece: Forecasting possibilities. Grana 45,122–129.

484  Damialis, A., Gioulekas, D., Lazopoulou, C., Balafoutis, C., Vokou, D., 2005. Transport of

485      airborne pollen into the city of Thessaloniki: the effects of wind direction, speed and

486      persistence. Int. J. Biometeorol. 49, 139-145.

487  Damialis, A., Kaimakamis, E., Konoglou, M., Akritidis,I., Traidl-Hoffmann, C., Gioulekas,

488      D., 2017. Estimating the abundance of airborne pollen and fungal spores at variable

489      elevations using an aircraft: how high can they fly? Scientific Reports. 7: 44535,

490      Doi:10.1038/srep44535

491  Damialis, A., Mohammad, A.B., Halley, J.M., Gange, A.C., 2015. Fungi in a changing world:

492      growth rates may be elevated, but spore production will decrease in future climates. Int. J.

493      Biometeorol. 59, 1157–1167.

494  De Linares, C., Belmonte, J., Canela, M., Díaz de la Guardia, C., Alba-Sanchez, F., Sabariego,

495      S., et al., 2010. Dispersal patterns of *Alternaria* conidia in Spain. Agric. Forest Meteorol.

496      150, 1491–1500.

497  De Lucca, A.J., 2007. Harmful fungi in both Agriculture and Medicine. Rev. Iberoam. Micol.

498      24, 3–13.

499  Durbin, J., 1960. The Fitting of Time-Series Models. Revue de l'Institut International de

500      Statistique / Review of the International Statistical Institute. 28, 233–244.

501  Elvira-Rendueles, B., Moreno, J., Garcia-Sanchez, A., Vergara, N., Martinez-Garcia, M. et al.,

502      2013. Air-spore in Cartagena, Spain: Viable and non-viable sampling methods. Ann. Agric.

503      Environ. Med. 20, 664–671.

504  Emberling, J., 1993. Variations in pollen concentrations – the need for accurate monitoring.

505      Clin. Exp. Allergy 23, 45.

**Formatted:** English (United States)

21

506 Escudero, O., Seijo, M.C., Fernández-González M., Iglesias I., 2011. Effects of

507 meteorological factors on the levels of *Alternaria* spores on a potato crop. Int. J.

508 Biometeorol. 55, 243–252.

509 Frenguelli, G., 2003. Basic microscopy, calculating the field of view, scanning of slides,

510 sources of error. Postępy Dermatologii i Alergologii 2, 227–229.

511 Frenguelli, G., Bricchi, E., Romano, B., Mincigrucci, G., Spieksma, F.T.M., 1989. A predictive

512 study on the beginning of the pollen season for Gramineae and *Olea europaea* L.

513 Aerobiologia 5,64–70.

514 Friesen, T.L., Faris, J.D., Solomon, P.S., Oliver, R.P., 2008. Host-specific toxins: effectors of

515 necrotrophic pathogenicity. Cell Microbiol. 10, 1421–1428.

516 Galán, C., Smith, M., Thibaudon, M., Frenguelli, G., Oteros, J., Gehrig, R., Berger, U., Clot,

517 B., Brandao, R., 2014. Pollen monitoring: minimum requirements and reproducibility of

518 analysis. Aerobiologia 30, 385-395.

519 Gange, A.C., Gange, E.G., Mohammad, A.B., Boddy L., 2011. Host shifts in fungi caused by

520 climate change? Fungal Ecol. 4, 154–190.

521 Gange, A.C., Gange, E.G., Sparks, T.H., Boddy L., 2007. Rapid and recent changes in fungal

522 fruiting patterns. Science 316, 71.

523 Gioulekas D., Damialis A., Mpalafoutis C., Papakosta D., Giouleka P., and Patakas D., 2004.

524 Allergenic fungal spore records (15 years) and relationship with meteorological parameters

525 in Thessaloniki, Greece. Allergy Clin. Immunol. Int. 16, 52-59.

526 Grant Smith, E., 1984. Sampling and identifying allergenic pollens and molds. An Illustrated

527 Manual for Physicians and Lab Technicians: Blewstone Press, San Antonio, Texas.

528 Gravesen, S., 1979. Fungi as a cause of allergic disease. Allergy 34, 135–154.

529 Grinn-Gofroń, A., Bosiacka, B., 2015. Effects of meteorological factors on the composition of

530 selected fungal spores in the air. Aerobiologia 31, 63–72.

22

531   Grinn-Gofroń, A., Bosiacka, B., Bednarz, A., Wolski, T., 2018. A comparative study of hourly
532       and daily relationships between selected meteorological parameters and airborne fungal
533       spore composition. Aerobiologia 34, 45–54.

534   Grinn-Gofroń, A., Strzelczak, A., 2009. Hourly predictive ANN and MRT models of
535       *Alternaria* and *Cladosporium* spore concentrations in Szczecin (Poland). Int. J.
536       Biometeorol. 53, 555-562.

537   Heald, C. L., Spracklen, D.V., 2009. Atmospheric budget of primary biological aerosol
538       particles from fungal spores. Geophys. Res. Lett. 36.

539   Helfer, S., 2014. Rust fungi and global change. New Phytologist. 201, 770–780.

540   Hirst, J., 1952. An Automatic Volumetric Spore Trap. Ann. Appl. Biol. 39, 257–265.

541   Ianovici, N., 2016. Atmospheric concentrations of selected allergenic fungal spores in relation
542       to some meteorological factors, in Timisoara (Romania). Aerobiologia 32, 139-156.

543   Iglesias, I., Rodríguez-Rajo, F. J., Méndez, J., 2007. Evaluation of the different *Alternaria*
544       prediction models on a potato crop in A Limia (NW of Spain). Aerobiologia 23, 27-34.

545   Jedryczka, M., 2014. Aeromycology: studies of fungi in aeroplankton. Folia Biologica et
546       Oecologica 10, 18–26.

547   Jedryczka, M., Strzelczak, A., Grinn-Gofron, A., Nowak, M., Wolski, T., Siwulski, M., et al.,
548       2015. Advanced statistical models commonly applied in aerobiology cannot accurately
549       predict the exposure of people to *Ganoderma* spore-related allergies. Agric. Forest
550       Meteorol. 201, 209–217.

551   Jolliffe, I.T., 1986.Principal component analysis. Second Edition (2002). Springer. 1-518.

552   Kauserud, H., Heegaard, E., Semenov, M.A., Boddy, L., Halvorsen, R., Stige, L.C., Sparks,
553       T.H., Gange, A.C., Stenseth, N.C., 2010b. Climate change and spring-fruiting fungi.
554       Proceedings of the Royal Society B. Biological Sciences. 277, 1169−1177.

23

555   Krouse, J.H., Chadwick, S.J., Gordon, B.R., Derebery, M.J., 2002. Allergy and Immunology.

556       An Otolaryngic Approach, Lippincott Williams & Wilkins, USA.

557   Kuhn, M., Johnson, K., 2013. Applied Predictive Modeling. Springer, New York, pp, 1–595.

558   Lee, K., Young, K., Kang, H., Ahn, K., Min, K., Cha, B., 1997. Cucumber scab caused by

559       *Cladosporium cucumerinum* in Korea. Korean J. Plant Pathol. 13, 288-294.

560   Li, R., L. Peng, C. Road., 2011. Quantile Regression for Left-Truncated Semicompeting Risks

561       Data. Biometrics. 7, 701–710.

562   Mamgain, A., Roychowdhury, R., Tah, J., 2013. *Alternaria* pathogenicity and its strategic

563       controls. Res. J. Biol. 1, 1-9.

564   Meza, F., & Varas, E. 2000. Estimation of mean monthly solar global radiation as a function

565       of temperature. Agric. Forest Meteorol. 100, 231-241.

566   Mitakakis, T.Z., Barnes, Ch., Tovey R., 2001. Spore germination increases allergen release

567       from *Alternaria*. J. Allergy Clin. Immunol. 107, 338-381.

568   Nowicki, M., Nowakowska, M., Niezgoda, A., Kozik, E., 2012. *Alternaria* black spot of

569       crucifers: symptoms, importance of disease, and perspectives of resistance breeding. Veget.

570       Crops Res. Bull. 76, 5-19.

571   Nowosad, J., 2016. Spatiotemporal models for predicting high pollen concentration level of

572       *Corylus*, *Alnus,* and *Betula*. Int. J. Biometeorol. 60, 843–855.

573   O'Connor, D. J., Sadyś, M., Skjøth, C. A., Healy, D. A., Kennedy, R., & Sodeau, J. R., 2014.

574       Atmospheric concentrations of *Alternaria, Cladosporium, Ganoderma* and *Didymella*

575       spores monitored in Cork (Ireland) and Worcester (England) during the summer of 2010.

576       Aerobiologia 30, 397–411.

577   Oteros, J., Galán, C., Alcázar, P., Domínguez-Vilches E., 2013. Quality control in bio-

578       monitoring networks, Spanish Aerobiology Network. Sci. Total Environ. 443, 559–565.

579    Oteros, J., García-Mozo, H., Alcázar, P., Belmonte, J., Bermejo, D., Boi, M., et al., 2015. A

580       new method for determining the sources of airborne particles. J. Environ. Manage. 155,

581       212-218.

582    Recio, M., del Mar Trigo, M., Docampo, S., Melgar, M., García-Sánchez, J., Bootello, L., et

583       al., 2012. Analysis of the predicting variables for daily and weekly fluctuations of two

584       airborne fungal spores: *Alternaria* and *Cladosporium.* Int. J. Biometeorol. 56, 983-991.

585    Sabariego, S., Díez, A., Gutiérrez M., 2007. Monitoring of airborne fungi in Madrid (Spain).

586       Acta Bot. Croat. 66, 117–126.

587    Sadyś, M., Adams-Groom, B., Herbert, R.J., Kennedy, R., 2016. Comparisons of fungal spore

588       distributions using air sampling at Worcester, England (2006–2010). Aerobiologia 32 (4),

589        619–634.

590    Sindt, C., Besancenot, J-P., Thibaudon, M., 2016. Airborne *Cladosporium* fungal spores and

591       climate change in France. Aerobiologia 32, 53–68.

592    Skjøth, C.A., Sommer, J., Frederiksen, L., Gosewinkel Karlson, U., 2012. Crop harvest in

593       Denmark and Central Europe contributes to the local load of airborne *Alternaria* spore

594       concentrations in Copenhagen. Atmos. Chem. Phys. 12, 11107–11123.

595    Stockmarr, A., Andreasen, V., Ostergård, H., 2007. Dispersal distances for airborne spores

596       based on deposition rates and stochastic modeling. Phytopathol. 97, 1325–1330.

597    Straatsma, G., Ayer, F., Egli, S., 2001. Species richness, abundance, and phenology of fungal

598       fruit bodies over 21 years in a Swiss forest plot. Mycol. Res. 105, 515–523.

599    Tariq, S.M., Matthews, S.M., Stevens, M., Hakim, E.A., 1996. Sensitization to *Alternaria* and

600       *Cladosporium* by the age of 4 years. Clin. Exp. Allergy. 26, 794–798.

601    Thomma, B.P., van Esse, H.P., Crous, P.W., de Wit, P.J., 2005. *Cladosporium fulvum* (syn.

602       *Passalora fulva*), a highly specialized plant pathogen as a model for functional studies on

603       plant pathogenic Mycosphaerellaceae. Mol. Plant. Pathol. 6, 379–393.

604 Van Leuken, J.P.G, Swart, A.N., Havelaar, A.H., VanPul, A., VanderHoek, W., Heederik, D.,

605      2016. Atmospheric dispersion modelling of bioaerosols that are pathogenic to humans and

606      livestock – A review to inform risk assessment studies. Microb. Risk Anal. 1, 19–39.

607 van Maanen, A., Xu, X-M., 2003. Modelling plant disease epidemics. Eur. J. Plant Pathol.

608      109, 669–682.

609

610

611

**Figure captions:**


**Figure 1.** Location of measurement points against the background of biogeographical conditions in Europe (according to Europaean Environmental Agency: 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