Supplementary Materials: Meteorological and Air Quality Effects on Bioaerosol Detection Using WIBS-NEO and IBAC-2 in Dublin City

Emma Markey [1], Jerry Hourihane Clancy [1,*], Moisés Martínez-Bracero [1,2], José María Maya-Manzano [3,4], Raúl Pecero-Casimiro [3], Eoin Joseph McGillicuddy [4], Gavin Sewell [4], Roland Sarda-Estève [5], Andrés M. Vélez-Pereira [6] and David J. O'Connor [1]

1. Methods

1.0.1. Statistical Testing and Validation

Multiple Linear Regression (MLR) was used to examine relationships between bioaerosol concentrations and WIBS/IBAC data, meteorological variables, and pollutant levels.

To address multicollinearity among the predictors, the Variance Inflation Factor (VIF) was calculated for each variable. Variables with a VIF greater than 5 were excluded from the final model to ensure the reliability of the regression coefficients. The MLRs were validated using cross-validation techniques, and the results were compared with observed bioaerosol concentrations to assess algorithm performance.

To quantify out-of-sample performance and reduce the risk of overfitting, we evaluated each WIBS-Hirst relationship using leave-one-out cross-validation (LOOCV). LOOCV was applied to each bioaerosol identified by the Hirst sampler using the WIBS-NEO fluorescent particle metrics as predictors. Model performance was summarised using the cross-validated mean absolute error (MAE), root mean square error (RMSE), and mean bias error (MBE), together with the correlation between observed and cross-validated predicted values. Where model selection was required (e.g., choice of fluorescence threshold or particle size bands), the final configuration reported in the Results was chosen based on cross-validated performance to ensure comparability across bioaerosol types.

2. Results

2.1. Multivariate Regression Analysis

Table S1. MLR models of analysed bioaerosols, showing predictors used, R^2 and adjusted R^2 values, along with results from tests for multicollinearity (VIF) and autocorrelation (DW test).

Target	WIBS Min Particle Size (μm)	Predictors	R^2	Max VIF	DW Stat	DW p
Pollen Total	9	ABC1, BC, B, Rh, Cloud, Day_L	0.718	2.98	2.06	0.92
Urticaceae	9	AC, BC, Day_L, soil, Cloud	0.696	2.76	2.12	0.92
Poaceae	9	AB, BC, Day_L, evap, Cloud	0.475	4.14	1.94	0.634
Cladosporium	2	ABC1, AC	0.49	1.03	1.33	0.014
Basidiospores	2	ABC1, BigBioIBAC, AC, Max_temp	0.667	1.25	2.25	0.6
Ascospores	2	ABC1, ABC2, AC, PM_{10}	0.294	1.76	1.9	0.51
Alternaria	9	BC, NO, wdsp, soil, Cloud	0.869	1.75	2.32	0.586
Spore Total	2	ABC1, BigIBAC, B, BC, PM_{10}	0.632	1.64	1.51	0.054

Multiple Linear Regression (MLR) algorithims were developed to better analyse and approximate pollen and fungal spore concentrations based on meteorological parameters, pollution levels, and WIBS FAPs (Table S1). This analysis ascertained the influence of multiple parameters on the observed pollen concentrations (R^2 = 0.72), with cloud cover, daylight, and WIBS BC-type particles (>9 μm) identified as the most significant predictors.

This aligns with the known relationships between these variables and pollen release and dispersal.

The calculated fungal spore concentration, based upon the same combination of meteorological parameter, pollution values, and real-time instrumentation outputs, possessed an R^2 of 0.63 for combined spores indicating a strong positive correlation. Among fungal spore parameters, the most commonly significant parameters were the first mode of the WIBS ABC-type particles, and WIBS AC-type particles, with each spore type possessing other different, unique significant parameters, indicating that a broad "Fungal Spores" MLR output is less appropriate. These results do however indicate that accurate models and forecasts could potentially be developed. This would require a longer monitoring campaign and a sufficiently large pool of valid WIBS data combined with meteorological and pollution datasets [1].

When looking at individual species of fungal spores, the one parameter that appeared as being a highly significant predictor was the first mode of ABC particles. When looking at the MLR analysis of basidiospore concentrations throughout the campaign, the ABC first mode had a significance value of <0.0001, indicating that it is an extremely significant predictor of basidiospore concentrations. It also had a standard error of 0.01, suggesting that the result is highly precise and that the sample is likely to be a reliable representation of the broader population.

The results above (Table S1) are also encouraging due to the number of parameters that were removed during the analysis, via the stepwise selection. Recommendations for the future are to carry out another WIBS monitoring campaign, to further investigate the potential of these MLR results, which have not been replicated in other studies [2–5].

The WIBS-NEO data, particularly the BC-type particles, proved to be a valuable input for the predictive models. These particles, which showed a strong correlation with Urticaceae pollen and *Alternaria* spores, provided a proxy for bioaerosol concentrations that could be used to improve the accuracy of pollen and fungal spore forecasts. However, the presence of other interfering particles, such as combustion-related aerosols, highlights the need for further refinement of the WIBS-NEO's detection algorithms. Leave-one-out cross-validation (LOOCV) was used to assess predictive performance of the Table S1 multiple linear regression models. LOOCV indicated strong generalisation for Alternaria (R^2 LOOCV = 0.796), moderate generalisation for Urticaceae (0.353) and total spores (0.264), and weaker generalisation for total pollen (0.178) and Poaceae (0.176). For *Cladosporium* and ascospores, LOOCV R^2 was negative, indicating limited predictive skill for unseen days under the current linear model specification.

2.2. Additional correlation analysis between meteorological parameters/air quality with bioaerosol concentrations, and with IBAC particles

Table S2. Spearman's rank correlation coefficients between daily IBAC particle classes and meteorological/air-quality parameters

Parameter	Small Particles	Large Particles	Small Fluorescent	Large Fluorescent
Tmed	-0.10	-0.15	-0.26	-0.22
Tmax	-0.24	-0.28	-0.23	-0.17
Tmin	0.03	-0.03	-0.20	-0.20
Gmin	-0.04	-0.10	-0.38*	-0.35*
Rain	-0.11	-0.00	-0.24	-0.13
Pres	-0.05	-0.17	0.16	-0.08
Wind_s	0.34*	0.36*	0.01	0.06
Rh	0.12	0.15	-0.17	-0.04
Wind_d	-0.15	-0.08	-0.21	-0.11
Day_L	-0.04	0.08	0.08	0.25
G_rad	-0.15	0.03	-0.08	0.17
Pe	-0.12	0.04	-0.16	0.10
Evap	-0.08	0.10	-0.15	0.13
$PM_{2.5}$	0.39*	0.33*	0.66**	0.53**
PM_{10}	0.58**	0.61**	0.74**	0.79**
NOx	-0.17	-0.17	0.13	0.16

* p<0.05, ** p<0.01 (two-sided Spearman).

Table S3. Spearman's rank correlation coefficients between meteorological/air-quality parameters and daily spore and pollen concentrations

Parameter	PollenTot	Urticaceae	Poaceae	*Cladosporium*	Basidiospores	Ascospores	*Alternaria*	Others	Total
Tmed	0.07	0.06	0.09	0.02	0.17	-0.01	0.17	-0.04	-0.06
Tmax	0.27	0.28	0.36*	0.29	0.40*	-0.12	0.34*	0.13	0.21
Tmin	-0.09	-0.12	-0.13	-0.17	-0.02	-0.02	-0.01	-0.13	-0.25
Gmin	-0.19	-0.20	-0.13	-0.34*	-0.06	0.12	-0.23	-0.08	-0.33*
Rain	-0.55**	-0.54**	-0.47**	-0.63**	-0.57**	0.39*	-0.53**	-0.12	-0.44**
Pres	0.02	0.05	0.03	0.16	0.01	-0.46**	0.23	0.00	-0.03
Wind_s	0.03	-0.05	0.14	-0.22	-0.05	-0.03	-0.16	0.13	-0.25
Rh	-0.11	-0.16	0.08	-0.33*	-0.08	-0.01	-0.35*	-0.20	-0.48**
Wind_d	-0.06	-0.02	0.00	-0.19	-0.08	0.08	-0.38*	-0.21	-0.52**
Day_L	0.20	0.19	0.24	0.25	0.13	0.10	0.52**	0.17	0.35*
G_rad	0.20	0.20	0.29	0.19	0.14	0.25	0.52**	0.16	0.35*
Pe	0.18	0.15	0.29	0.11	0.23	0.14	0.47**	0.02	0.16
Evap	0.19	0.15	0.33*	0.12	0.18	0.21	0.48**	0.03	0.20
$PM_{2.5}$	0.43**	0.45**	0.39*	0.35*	0.16	-0.26	0.35*	0.15	0.23
PM_{10}	0.51**	0.54**	0.45**	0.43**	0.24	-0.19	0.47**	-0.02	0.31
NOx	0.05	0.03	0.06	0.08	0.07	-0.22	0.03	0.30	0.07

* p<0.05, ** p<0.01 (two-sided Spearman).

2.3. Additional bioaerosol concentration information

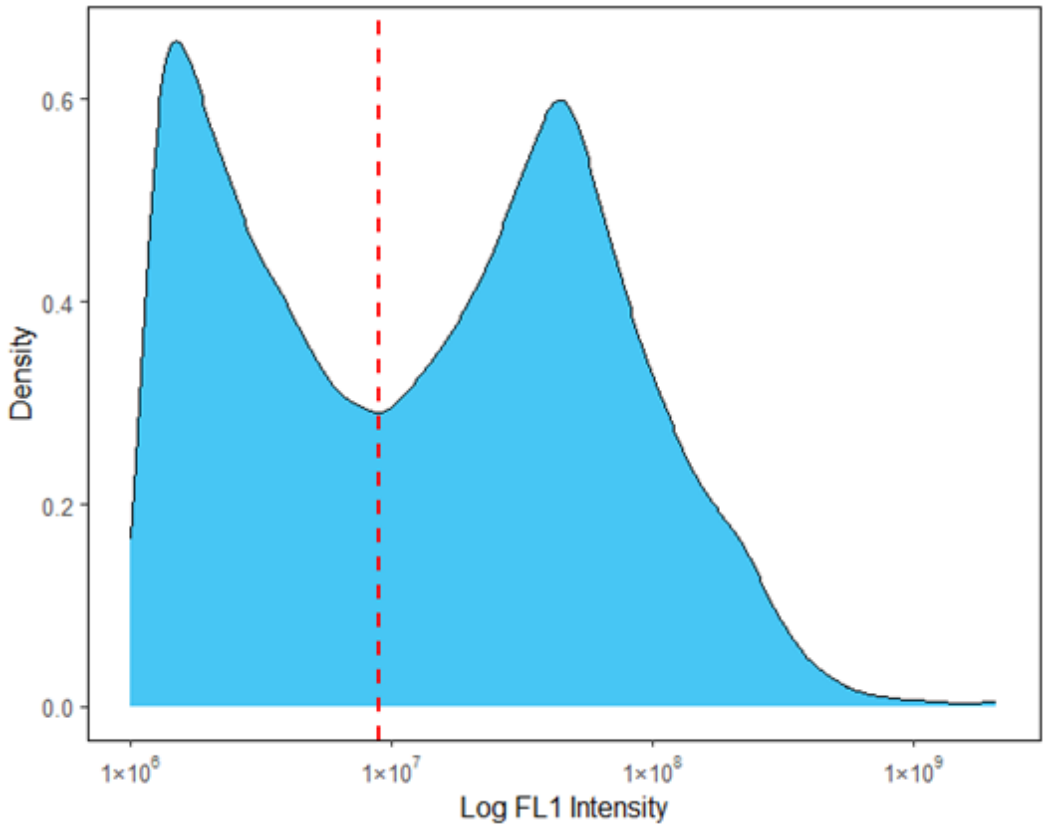


Figure S1. : Bimodal Distribution of the ABC WIBS Channel.

3. Discussion

The development of MLR algorithms marked a significant advancement, achieving strong ability to accurately describe the relationships between bioaerosol release and transport, with ambient meteorological conditions. This was reflected in the achieved accuracy of both fungal spores (R^2 = 0.63) and pollen (R^2 = 0.72) [1,6]. Key predictors included WIBS-NEO fluorescence signals (e.g., BC-type particles), meteorological variables (e.g., soil temperature), and pollution metrics (e.g., PM_{10}) [7,8], highlighting the utility of integrating real-time sensor data with environmental datasets

3.1. Multiple Linear Regression (MLR) Analysis

The MLR models developed in this study provide a framework to disentangle the combined effects of meteorological conditions, air quality, and real-time fluorescence data on bioaerosol concentrations. For pollen prediction, the models achieved strong performance (R^2 = 0.82), with BC-type fluorescent particles (>8 μm) detected by the WIBS-NEO emerging as the strongest predictor. These particles, which align closely with Urticaceae pollen trends, act as potential indicators of grass and herbaceous pollen in urban settings. Soil temperature and easterly wind direction also played significant roles that reflected the localized release and transport of pollen from green spaces such as Phoenix Park. Notably, higher nitrogen oxide (NOx) levels were associated with reduced pollen detection which has been seen previously to be linked with pollution-driven fragmentation of pollen grains into sub-micron particles undetectable by traditional methods [9]. For fungal spores, the models (R^2 = 0.78) highlighted distinct environmental drivers. *Cladosporium*, the dominant spore type, was strongly associated with coarse particulate matter (PM_{10}), suggesting these spores become attached to, and are transported by, dust particles during dry and windy

conditions. In contrast, ascospores peaked during low air pressure time periods, consistent with their passive dispersal from maritime air masses. Grass minimum temperature and rainfall exerted suppressive effects on overall spore concentrations, underscoring how cool, wet weather generally inhibits spore release.

Urban pollutants introduced notable challenges to the interpretation of the data. Combustion-related particles, such as black carbon and PAHs, skewed fluorescence signals, particularly for smaller particles (<2 μm). This interference, evident in the strong correlation between NOx and fluorescent AB-type particles, mirrors findings from polluted regions like Tokyo and the Yangtze River Delta [5,10]. To mitigate these effects, wind speed thresholds were incorporated into the models, improving accuracy by accounting for sensor limitations during turbulent conditions. Mean absolute errors were between 12% and 18%,with several key limitations remaining. For example, rooftop sensor placement underestimates ground-level pollen and PM_{10} concentrations which is a known issue in urban air quality studies. Future work will be aimed at assessing similar fluorescence based instruments that posses wider excitation and emission bands, such as the Swisens Poleno Jupiter, as the addition of wider ranges and additional bands has shown promise in recent studies [2,11]. Expanding data collection to rural sites would further enhance the model's ability to generalise, particularly for spores like *Alternaria* that straddle both urban and agricultural sources.

Future efforts should focus on long-term monitoring to validate MLR robustness [2,3,5] and address urban-specific challenges in bioaerosol discrimination [9,12].

These observations emphasize the need for algorithm refinements to mitigate anthropogenic interferences, particularly from combustion-related particles correlated with NO_X [4,13,14].

3.2. Limitations

This study is constrained by the short campaign period (41 days total) and modest sample size, which limits representativeness across meteorological regimes and increases sensitivity to influential days. Although LOOCV indicated strong out-of-sample generalisation for *Alternaria* and moderate generalisation for Urticaceae and total spores, results were mixed, and performance dropped for some taxa, most notably for *Cladosporium* and ascospores, indicating limited predictive skill for unseen days. In addition, some spore models showed residual autocorrelation, implying that the independence assumption is not fully met. This can overstate in-sample fit and reduce performance on unseen days when conditions persist from one day to the next. More broadly, WIBS fluorescence types should not be treated as a direct representative of biology in a mixed urban environment, as some of the signal can come from non-biological fluorescent particles, so the relationships we see may reflect a mixture of biological variability and interference. The analysis also uses a deliberately simple linear approach, which may miss non-linear behaviour, interactions, and threshold effects (for example under high humidity), and it does not explicitly account for lagged effects or seasonal structure. Taken together, these points support future work using longer time series, stricter validation, and a wider set of predictors while checking for collinearity.

4. Conclusion

Building on these findings, the MLR algorithims developed in this study quantified the relative influence of key predictors: BC-type particles (>8 μm) emerged as the strongest indicator of pollen concentrations, while PM_{10} and grass minimum temperature drove fungal spore variability. These models strongly indicated that combustion-related aerosols (NOx, $PM_{2.5}$) contributed significantly to fluorescence interference, particularly for smaller

particles (<2 μm), explaining the 20% residual variance. Notably, wind speed thresholds (>16.6 km/h) improved model accuracy by 12% highlighting the need for wind speed-limited sampling protocols in future urban deployments, along with further investigation into high-wind conditions. The MLR analysis also highlighted the WIBS's dual sensitivity to biological and anthropogenic aerosols, supporting its potential as a hybrid air quality monitor. For instance, correlations between AB-type particles and NOx mirrored trends observed in various other studies, where similar fluorescence was attributed to soot and PAHs [5,10]. Future iterations could integrate machine learning classifiers to mitigate the impact of these overlaps. Such advancements, coupled with height-adjusted PM calibrations, would enhance the WIBS's utility in urban settings, bridging gaps between bioaerosol detection and comprehensive air quality assessment.

References

1. Filho, D.B.F.; Júnior, J.A.S.; Rocha, E.C. What is R2 all about? *Leviathan (São Paulo)* **2011**, pp. 60–68. Number: 3, https://doi.org/10.11606/issn.2237-4485.lev.2011.132282.
2. Markey, E.; Hourihane Clancy, J.; Martínez-Bracero, M.; Neeson, F.; Sarda-Estève, R.; Baisnée, D.; McGillicuddy, E.J.; Sewell, G.; O'Connor, D.J. A Modified Spectroscopic Approach for the Real-Time Detection of Pollen and Fungal Spores at a Semi-Urban Site Using the WIBS-4+, Part I. *Sensors* **2022**, *22*, 8747. Number: 22 Publisher: Multidisciplinary Digital Publishing Institute, https://doi.org/10.3390/s22228747.
3. O'Connor, D.J.; Healy, D.A.; Hellebust, S.; Buters, J.T.M.; Sodeau, J.R. Using the WIBS-4 (Waveband Integrated Bioaerosol Sensor) Technique for the On-Line Detection of Pollen Grains. *Aerosol Science and Technology* **2014**, *48*, 341–349. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/02786826.2013.872768, https://doi.org/10.1080/02786826.2013.872768.
4. Toprak, E.; Schnaiter, M. Fluorescent biological aerosol particles measured with the Waveband Integrated Bioaerosol Sensor WIBS-4: laboratory tests combined with a one year field study. *Atmospheric Chemistry and Physics* **2013**, *13*, 225–243. Publisher: Copernicus GmbH, https://doi.org/10.5194/acp-13-225-2013.
5. Yu, X.; Wang, Z.; Zhang, M.; Kuhn, U.; Xie, Z.; Cheng, Y.; Pöschl, U.; Su, H. Ambient measurement of fluorescent aerosol particles with a WIBS in the Yangtze River Delta of China: potential impacts of combustion-related aerosol particles. *Atmospheric Chemistry and Physics* **2016**, *16*, 11337–11348. Publisher: Copernicus GmbH, https://doi.org/10.5194/acp-16-11337-2016.
6. Gross, J.; Ligges, U. nortest: Tests for Normality, 2015.
7. Clancy, J.H.; Markey, E.; Martínez-Bracero, M.; Maya-Manzano, J.M.; McGillicuddy, E.J.; Sewell, G.; Sarda-Estève, R.; Baisnée, D.; Vélez-Pereira, A.M.; Davis, G.; et al. Comparative Analysis of Real-Time Fluorescence-Based Spectroscopic Instruments: Bioaerosol Detection in the Urban Environment of Dublin City, Ireland. *Atmosphere* **2025**, *16*, 275. Number: 3 Publisher: Multidisciplinary Digital Publishing Institute, https://doi.org/10.3390/atmos16030275.
8. Agency (EPA), E.P. Reports & SAFER-Data. Publisher: Environmental Protection Agency (EPA).
9. Mampage, C.B.A.; Hughes, D.D.; Jones, L.M.; Metwali, N.; Thorne, P.S.; Stone, E.A. Characterization of sub-pollen particles in size-resolved atmospheric aerosol using chemical tracers. *Atmospheric Environment: X* **2022**, *15*, 100177. https://doi.org/10.1016/j.aeaoa.2022.100177.
10. Miyakawa, T.; Kanaya, Y.; Taketani, F.; Tabaru, M.; Sugimoto, N.; Ozawa, Y.; Takegawa, N. Ground-based measurement of fluorescent aerosol particles in Tokyo in the spring of 2013: Potential impacts of nonbiological materials on autofluorescence measurements of airborne particles. *Journal of Geophysical Research: Atmospheres* **2015**, *120*, 1171–1185. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/2014JD022189, https://doi.org/10.1002/2014JD022189.
11. Markey, E.; Hourihane Clancy, J.; Martínez-Bracero, M.; Sarda-Estève, R.; Baisnée, D.; McGillicuddy, E.J.; Sewell, G.; Skjøth, C.A.; O'Connor, D.J. Spectroscopic detection of bioaerosols with the wibs-4+: Anthropogenic and meteorological impacts. *Science of The Total Environment* **2024**, *943*, 173649. https://doi.org/10.1016/j.scitotenv.2024.173649.
12. Markey, E.; Clancy, J.H.; Martínez-Bracero, M.; Maya-Manzano, J.M.; Smith, M.; Skjøth, C.; Dowding, P.; Sarda-Estève, R.; Baisnée, D.; Donnelly, A.; et al. A comprehensive aerobiological study of the airborne pollen in the Irish environment. *Aerobiologia* **2022**, *38*, 343–366. https://doi.org/10.1007/s10453-022-09751-w.
13. Gabey, A.M.; Stanley, W.R.; Gallagher, M.W.; Kaye, P.H. The fluorescence properties of aerosol larger than 0.8 μm in urban and tropical rainforest locations. *Atmospheric Chemistry and Physics* **2011**, *11*, 5491–5504. Publisher: Copernicus GmbH, https://doi.org/10.5194/acp-11-5491-2011.
14. Twohy, C.H.; McMeeking, G.R.; DeMott, P.J.; McCluskey, C.S.; Hill, T.C.J.; Burrows, S.M.; Kulkarni, G.R.; Tanarhte, M.; Kafle, D.N.; Toohey, D.W. Abundance of fluorescent biological aerosol particles at temperatures conducive to the formation of mixed-phase and cirrus clouds. *Atmospheric Chemistry and Physics* **2016**, *16*, 8205–8225. Publisher: Copernicus GmbH, https://doi.org/10.5194/acp-16-8205-2016.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.