Journal of Environmental Management 393 (2025) 127033

Contents lists available at ScienceDirect

Journal of Environmental Management

journal homepage: www.elsevier.com/locate/jenvman

Research article

Real-time field measurements of bioaerosols in the agricultural environment: Concentrations, components and environmental impacts

Zhuo Chen [a,*], Ian Crawford [a], Emily Matthews [a], Michael Flynn [a], Thomas Bannan [a], Laura Cardenas [b], Jon S. West [c], Hugh Coe [a], David Topping [a], Martin Gallagher [a]

[a] Department of Earth and Environmental Sciences, The University of Manchester, Manchester, M13 9PL, UK
[b] Rothamsted Research, North Wyke, Okehampton, EX20 2SB, UK
[c] Rothamsted Research, Harpenden, AL5 2JQ, UK

ARTICLE INFO

Keywords:
Bioaerosols
Agriculture
Machine learning
Detection

ABSTRACT

In agricultural production, bioaerosols inevitably pose health hazards to animals and workers. Currently, there is a lack of research on real-time bioaerosol concentration monitoring at agricultural sites. We conducted a real-time airborne bioaerosol measurement study using the Multiparameter Bioaerosol Spectrometer (MBS) and applied a Uniform Manifold Approximation and Projection (UMAP) approach to classify bioaerosol emissions from the North Wyke Farm Platform between April and May. *Penicillium* and *Cladosporium* were the most dominant fungi. Another machine learning approach, Generalized Additive Model (GAM) was also constructed to explore the relationship with meteorological data and selected trace gases. It was found that animal houses and agricultural fields were the main sources of bioaerosols, and significant dispersion was observed downwind of these point sources. Two main bioaerosol types were *Cladosporium* and *Penicillium*, which accounted for 29.8 % and 24.1 % of the total, respectively. *Cladosporium* had an average concentration of 3.79 L^{-1} in the animal house direction, which is 2.19 L^{-1} higher and about 2.37 times that in the farmland direction (1.60 L^{-1}). For *Penicillium*, the average concentration was 2.44 L^{-1} in the animal house direction, 0.93 L^{-1} higher and 1.61 times that in the farmland direction (1.52 L^{-1}). And both bioaerosols are more active at temperatures above 15 °C and relative humidity above 80 %.

These results may provide recommendations for detection and identification of bioaerosol composition and emission patterns in the agricultural environments, and emission profiles associated with animal farms to provide better understanding for agricultural regional planning and public health perspectives.

1. Introduction

As defined by Whitby et al. (2022) bioaerosols are "suspensions of airborne particulate matter of biological origin, representing a distinct category of aerosols". They can be defined as suspended airborne particles that are emitted by living organisms (Crawford et al., 2023; GSJ et al., 2023). They contain diverse biological particles with a wide range of sizes such as pollen (10–100 μm) (Bennett and Willis, 2001), bacteria (0.2–10 μm) (Katz et al., 2003), fungal spores (2–50 μm) (Patel et al., 2018), viruses (<0.2 μm) (Grgacic and Anderson, 2006). Due to the small aerodynamic diameter of some airborne bioaerosols, they can easily enter the body through the respiratory system and be deposited in the lungs (GSJ et al., 2023; Sajjad et al., 2023). Prolonged exposure can lead to respiratory diseases such as asthma and even severe organ damage (Bennett and Willis, 2001; Maya-Manzano et al., 2021; Sauvageat et al., 2020). Bioaerosols can also act as cloud condensation and ice nuclei, accelerating the rate of cloud and precipitation formation (Tang et al., 2022; Uetake et al., 2019). Cloud formation and rainfall, in turn, can further influence regional climate, as well as local bioaerosol release contributing to the so-called bio-precipitation cycle (Bigg et al., 2014).

Meanwhile, human agriculture activities, and natural agricultural bioaerosol emissions also have an impact on local and regional air quality and ecosystems (Sabban and van Hout, 2011). In the livestock industry, processes such as feed preparation, animal activity and manure management all lead to increased concentrations of bioaerosols, especially fungi, and bacteria, in the farming area (Kumar et al., 2024). Subsequent dispersion into the atmosphere through ventilation equipment poses a major challenge to the diversity and concentration of

* Corresponding author.
E-mail address: zhuo.chen-2@manchester.ac.uk (Z. Chen).

https://doi.org/10.1016/j.jenvman.2025.127033
Received 16 June 2025; Received in revised form 4 August 2025; Accepted 16 August 2025
Available online 20 August 2025

bioaerosols in the surrounding area (Kumar et al., 2021, 2024). Elevated concentrations of bacteria, fungi, and endotoxins were reportedly detected downwind of intensive farms (Gladding et al., 2020; Ko et al., 2008). For example, zoonotic pathogens and antibiotic-resistant bacteria (e.g. *Escherichia coli* and *Staphylococcus aureus*) can be detected in the atmosphere within 150–200 m of pig and poultry farms (Gibbs et al., 2006; Kumar et al., 2024). However, for the time being, it remains uncertain whether farm-sourced bioaerosols pose a health risk to the public. Radon et al. (2007) noted a corresponding increase in asthma prevalence around the animal house as exposure to animal feeding operations. However, Smit et al. (2014) showed a statistically negative correlation between PM concentrations around the farm and public health outcomes. The UK Environment Agency reported in 2008 that chicken coops have the highest emission rates for total microorganisms, fungi and endotoxins. Fungi and bacteria were detected above background values at distances of 200 and 400 m from the barns (Scaife et al., 2008). Also, for livestock barns, the more frequent the animal activity, the more bioaerosols are emitted (Scaife et al., 2008).

Based on the serious challenges posed by bioaerosols to the agricultural sector, rapid monitoring and characterisation of bioaerosol particles is of great importance. Based on the physical and chemical characteristics of bioaerosols, such as size, fluorescence spectra, absorption spectra, and shape, a series of single particle on-line detection system have recently been developed. The advantage of such systems is that they enable detection of relatively low concentrations of bioaerosol particles within generally larger ambient concentrations of other atmospheric aerosol particles (Pan et al., 2022). Currently, the key technology being applied for online detection of airborne bioaerosols are based on single particle ultraviolet light induced fluorescence (UV-LIF) spectrometry (Gabbarini et al., 2019; Huffman et al., 2020). The Multi-Parameter Bioaerosol Spectrometer (MBS) employed in this study is such an typical instrument based on this technology (Foot et al., 2008).

The interpretation of data generated by real-time monitoring instruments involves two aspects: first, the use of machine learning classifiers to categorize data into distinguishable taxonomic classes, and second, the application of statistical models to quantify and visually represent the relationship between bioaerosols and meteorological factors. In this manuscript, we employ Uniform Manifold Approximation and Projection (UMAP) for dimensional reduction to classify data from the MBS (Crawford et al., 2023; McInnes et al., 2018). UMAP has been demonstrated to exhibit robust classification performance (Crawford et al., 2023; Zhao et al., 2024). By reducing dimensionality, it preserves both global and local relationships within the data, making the classification process more concise and accurate (Mpaka and von der Heyden, 2024). After obtaining the classified data, we introduce the Generalized Additive Model (GAM), proposed by Hastie and Tibshirani, to construct models for feature parameters (whether linear or nonlinear) (Hastie, 2017; Yan et al., 2024). Traditional regression models, such as linear regression, are not particularly effective in capturing nonlinear relationships between meteorological parameters and aerosol concentrations (Cheng et al., 2021). While deep learning models such as neural network models may achieve more accurate results, they require significantly greater computational resources and time (Du et al., 2022). In comparison to these two types of statistical models, GAM outperforms regression models in capturing nonlinear relationships while providing relatively accurate and interpretable results with less computational time (Lundberg and Lee, 2017). Consequently, GAM is also widely applied in atmospheric pollutant monitoring and impact assessment (Cheng et al., 2021; Qi et al., 2021; Ramsey et al., 2014).

Kumar et al. (2024) noted that although much research has been done, little is known about the bioaerosol species, exposure levels, and their environmental impacts within feedlots and dispersal areas. To elucidate these knowledge gaps, here we conducted a focussed ambient sampling experiment using an MBS to monitor concentrations of bioaerosols in real-time dispersed downwind of an animal farm, the North Wyke Farm Platform (NWFP). The aims of the experiment were to evaluate a) the bioaerosol species detected in the agricultural area; b) the identification of their sources; c) the quantification of the relationship between detected bioaerosol, meteorological parameters and select trace gases using a machine learning approach. Through these objectives, the quantification of real-time bioaerosol species and the observation of their impacts on the environment can deepen the understanding of the composition of bioaerosols in the agricultural sector and provide a basis for future decision-making on scientific management and land use, as well as provide a reference for the risk assessment of these agricultural facilities.

2. Methods

2.1. Sampling location

The North Wyke Farm Platform (NWFP) is situated at the North Wyke grassland site of Rothamsted Research (location: 50°46′10″ N, 3°54′05″ W), 20 km north of Dartmoor National Park, which is the largest upland area in south-west England (Orr et al., 2011). NWFP comprises three farming systems in "farmlets", each consisting of five component catchments totalling approximately 21 ha per farmlet. These farmlets are used to monitor soil, livestock and silage performance (Hawkins et al., 2023a). Fig. 1 shows a summary map of NWPF farm station and the associated land management regions (Hawkins et al., 2023b).

The asterisk indicates the location of the MBS instrument on the 'Top Burrows' field, which is elevated compared with the surrounding fields. To the north-west of the instrument lies several livestock sheds which were housing cattle and sheep during the measurement period. The surrounding fields are a mixture of arable and pastoral land.

2.2. Sampling methods

2.2.1. Modulair sensor for atmospheric pollutants

A QuantAQ sensor (Modulair) was used to continuously measure CO, NO_2, O_3, and size resolved particulate matter (PM) concentrations. Also included in this sensor system is a measure of internal temperature and relative humidity, and a Global System for Mobile Communications unit for sending logged data to an external database. The sensor system was coupled with a Davis® Sonic Anemometer to provide in-situ wind direction and wind speed measurements. Prior to deployment at North Wyke, the sensor underwent calibration and verification with scientific grade reference instrument at the NERC Air Quality Supersite in Manchester (Diez et al., 2022, 2024). The sensor was mounted at a height of 2 m within 200 m of the MBS detector. Data were processed using the QUANT AQ dashboard and exported for further analysis.

2.2.2. The multiparameter bioaerosol spectrometer

The Multiparameter Bioaerosol Spectrometer (MBS) is a biofluorescence spectrometer developed by the University of Hertfordshire. By analysing the autofluorescence spectra, size, and morphological parameters of single particles it is possible to detect and classify bioaerosol particles in real time. Previous research has demonstrated that the outputs of real-time measurement instruments with similar technology are in good agreement with offline fungal tracers (Gosselin et al., 2016). In this study, we calibrate the size and fluorescence performance of MBS by using standard and fluorescent doped PSLs following the approach described previously for similar UVLIF spectrometers (Wideband Integrated Bioaerosol Sensors, WIBS), as described by Robinson et al. (2013) and Crawford et al. (2015). The MBS detection principle is briefly described below. An optically filtered xenon flash lamp (280 nm) is used to excite individual aerosol particles in the sensing area through deep ultraviolet (UV) excitation; the resultant autofluorescence is detected via a grating spectrometer and multichannel photodetector over 8 bands in the 315–640 nm range. This configuration covers the emission ranges

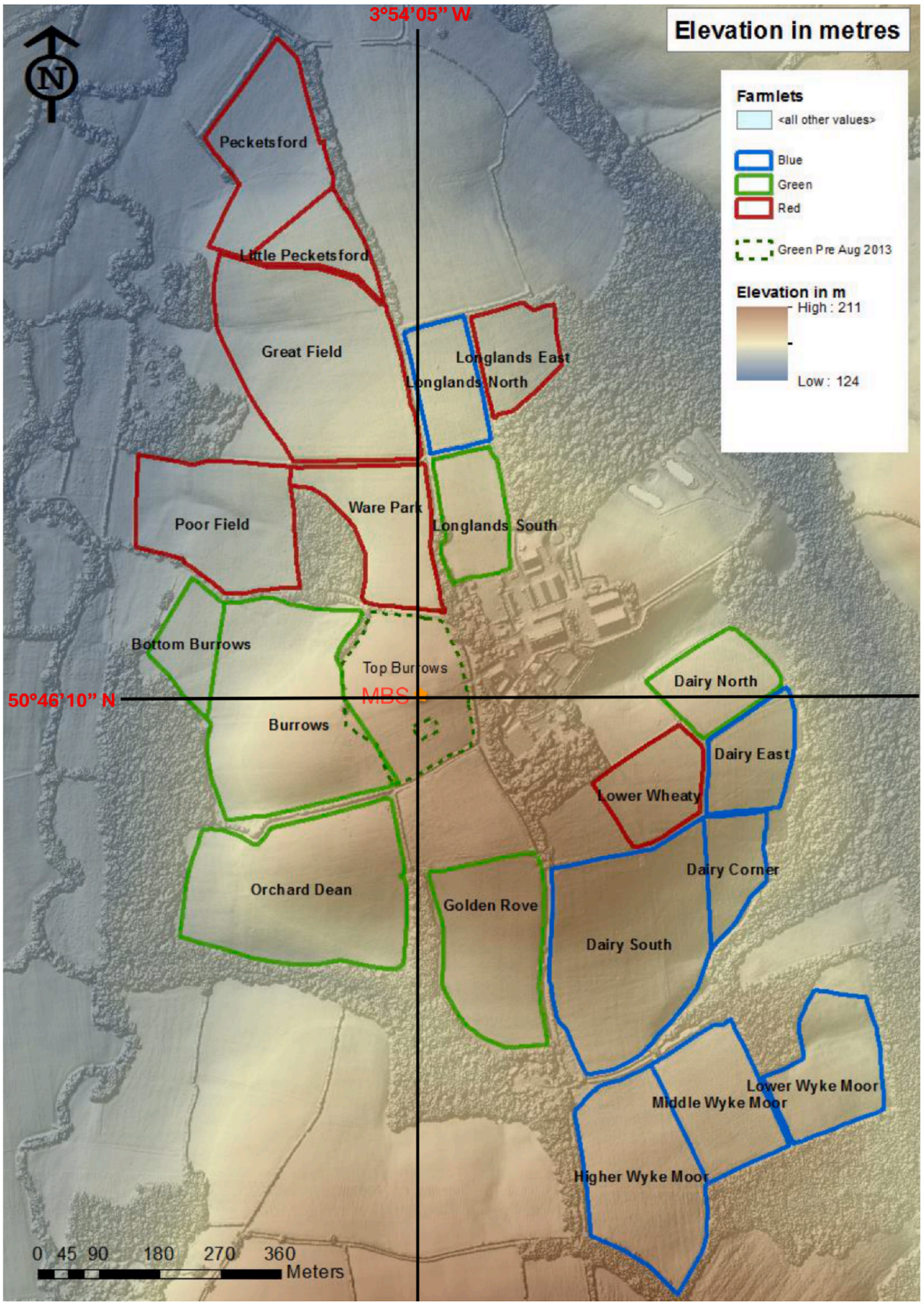

Fig. 1. The North Wyke Farm Platform fields. Farmlets and land use and management are shown with: High sugar grass areas outlined in red, white clover mix area outlined in blue, permanent pasture areas are outlined in green. The purple boxed sections highlight the cattle and sheep. The location of the bioaerosol instrument measurements (MBS) is indicated by the asterisk (orange).

of several key biofluorophores commonly found in bioaerosol of interest at 280 nm excitation which are now listed (Könemann et al., 2019).

- Tyrosine: 310 ± 20 nm;
- Tryptophan: 365 ± 40 nm;
- Riboflavin: 520 ± 30 nm;
- Chlorophyll *b*: 640 ± 10 nm.

Ambient air is drawn into the instrument at a total flow rate of 1.8 L min^{-1} through an inlet tube. 1.5 L min^{-1} of the airflow is filtered and used as bleed flow and sheath flow. The bleed flow is used to keep the inner optical chamber clean, and the sheath flow is used to constrain the remaining sample flow (0.3 L min^{-1}). The constrained sample flow provides a single file of collimated aerosol for the detection system. A single particle carried in the sample flow is first detected and sized by a low power laser beam (12 mW, 635 nm) in the 0.5–15 μm diameter range via Mie scattering. A second high power pulsed laser (250 mW, 637 nm) will be triggered when particles in this size range are detected. The high-power laser illuminates the particles with sufficient intensity to detect the morphology of particles via a dual 512-pixel complementary metal-oxide-semiconductor (CMOS) image sensor array. The array can collect the scattered light from the illuminated particle and construct two cross-sectional chords from the 2D profile of the scattering pattern. Particle morphological parameters are automatically generated from the CMOS distributions which are now described (Crawford et al., 2020).

- Peakwidth: An estimate of the average width of the array peak.
- Peakmean: The ratio of the peak to mean parameters.
- Mirror: A measure of scattering symmetry between top and bottom half of each array.
- AsymLR: A measure of the scattering symmetry between the left and right arrays.
- AsymLRinv: Same with AsymLR, however the right hand array is inverted.

10 μs after this initial detection, the xenon flashlamp is triggered which illuminates the particle with a UV pulse at 280 nm for approximately 1 μs. Any autofluorescence generated by this UV excitation is focused with two hemispherical mirrors onto a grating spectrometer to segregate the fluorescence by wavelength and this spectral signal is recorded by an 8-channel photomultiplier tube (PMT). The maximum strobe rate of the xenon flashlamp is approximately 125 Hz, limiting the maximum acquisition rate. However, this upper limit is rarely approached in practice so is still sufficient for expected ambient concentrations (Crawford et al., 2023).

The fluorescence threshold is defined by the forced trigger (FT) cycle, where the xenon flashlamp is strobed at 10 Hz in the absence of particles during a "pump off" phase to determine the optical system background noise levels. The MBS is run in "forced trigger" mode for 10 s at the start of each new data file, where a new file is stared every 30,000 particles. The average of the FT spectral data +7 standard deviations is used to define the threshold for each channel (referred to as the 7σ threshold), which is then subtracted from the acquired data in post processing, and any negative result are set to zero. The value of 7σ retains most of the biofluorescence while rejecting non-biological interfering particles that are normally weakly fluorescent (Crawford et al., 2023; Savage et al., 2017). Additionally, the particle must also satisfy the 7σ threshold in at least two channels to be classified as fluorescent to remove false positives arising from spurious reflections and noise (Crawford et al., 2023; Könemann et al., 2019).

2.3. Data preparation

A brief summary of the meteorological, trace gas and bioaerosol data collected in the experiment is shown in Table 1.

Table 1
Summary of meteorological parameters and trace gas concentration data during sampling period.

	Units	Mean	Min	Max
Bioaerosol	Numbers/L	44	0	2050
Relative Humidity (RH)	%	69.8	33.5	87.5
Temperature (T)	°C	10.4	3.4	18.8
Wind Direction (WD)	°	/	0	360
Wind Speed (WS)	m/s	3.6	0	17.0
CO	ppm	191.25	188.51	442.05
NO	ppb	1.96	1.32	29.16
NO_2	ppb	7.81	1.34	29.01
O_3	ppb	39.72	14.55	53.72

Meteorological and trace gas data were recorded at 1-min intervals, single particle bioaerosol counts were averaged over 5-min periods. We extracted the meteorological and trace gas data corresponding to the bioaerosol sampling time and merged them with the bioaerosol data. The merging method was to map the meteorological data corresponding to the time points of sampling of bioaerosols. Subsequent calculation of hourly averages for the data was done by vector averaging for wind direction. The OpenAir package in R was used to produce distribution polar plots of the various aerosol and trace gas parameters (Carslaw and Ropkins, 2012).

2.4. Bioaerosol classification method

The Uniform Manifold Approximation and Projection for dimension reduction (UMAP) was used to classify particulate matter captured by MBS into more broadly representative BioPM species (McInnes et al., 2018). UMAP downscales and compresses the spectral data and the CMOS morphological provided by MBS into 2D space and optimises the results using the training data, constructing a transformed space with a high degree of spatial separation for each of the classification results. Since UMAP achieves optimal performance when all input features are on a similar scale, each parameter was normalized based on its expected maximum value. The MBS training data used for UMAP were provided by experiments previously conducted at the ChAMBRe simulation chamber facility as described in detail by Crawford et al. (2023). The same analysis structures and procedures were adopted as described by Crawford et al. (2023) to discriminate bioaerosol classes. The basic fluorescence criteria for acceptance of MBS single particle data as potential biofluorescent aerosols for subsequent classification required fluorescence threshold intensity to 1) exceed a the 7σ threshold and 2) exhibit the fluorescence in at least two of the MBS fluorescence wavebands.

2.5. Generalized Additive Model (GAM)

GAM extends the framework of generalized linear models (GLMs) by replacing linear predictors with additive smooth functions (Hastie, 2017; Herman and Hastie, 1990). The fundamental formulation follows:

$$g[E(y_i)] = \beta_0 + f_1(X_1) + f_2(X_2) + f_3(X_3) + \cdots + f_n(X_n) \quad (1)$$

where $g()$ is the link function; $E(y_i)$ denotes the conditional expectation of the response variable; β_0 is the intercept; X_n is the explanatory variables; $f_n()$ is the smooth function.

$f_n()$ in GAM is a core tool for defining nonlinear smooth terms for independent variables, allowing the model to automatically learn nonlinear relationships through a data-driven approach.

In this experiment, we used the python-based pyGAM package for model construction (Servén and Brummitt, 2018). Created a single-factor GAM formula based on Equation (1), which is Equation (2):

$$g[Y_i] = \beta_0 + S(x_1) + S(x_2) + S(x_3) + \cdots + S(x_n) \quad (2)$$

Where Y_i is the Processed bioaerosol data; $S()$ is the spline term, which has a penalty on their 2nd derivative, which encourages the functions to be smoother (Servén and Brummitt, 2018). x_n represents meteorological and trace gas data been collected. The data itself needs to pass the variance inflation factor (VIF) test to eliminate multicollinearity before importing it where the VIF value > 5 would indicate multicollinearity. The VIF test is done by SPSS (Landau and Everitt, 2003).

2.6. Model development

2.6.1. Box-Cox transform

It was found that the data distribution of the raw bioaerosol data was right-skewed due to low atmospheric bioaerosol concentrations which was limited by the instrumental sample volume. We therefore applied the Box-Cox transformation to the data (Loaiza et al., 2023; Sakia, 1992). To explain the superiority of this data transformation, we illustrate it by introducing the Akaike Information Criterion (AIC) where a smaller AIC value indicates a better model performance (Akaike, 1987). After testing, the Box-Cox transformed data contributed to smaller AIC values (−11237 and −10296). Quantile-Quantile plots (QQ-plots) of the data distribution before and after the Box-Cox transformation, consistent with the Gamma distribution, are provided in the appendix. This is consistent with the general trend observed in environmental pollution data, which often follow Poisson or Gamma distributions (Cheng et al., 2021; Yan et al., 2024).

2.6.2. Model training and evaluation

To ensure robust performance evaluation, the dataset was partitioned using a 5-fold cross-validation scheme (Anguita et al., 2012; dos Santos et al., 2021). 80 % of the observations (4-fold) were used for training and 20 % (1-fold) for testing in each iteration. The random state was set to 42.

For the model training, based on the API pyGAM application manual, we selected hyperparameters for each spline term, which include "*n_splines*" and "*lam*". Among these two, "*n_splines*" indicates the number of splines to use for the feature function and "*lam*" indicates the strength of smoothing penalty (Servén and Brummitt, 2018). The choice of these two parameters is based on a hyperparameter optimization package called Optuna, which allows users to search the best hyperparameter match in the set time period (Akiba et al., 2019).

The evaluation of the model is based on two indicators, mean square error (MSE) and coefficient of determination (R^2). For MSE, the equation can be written as:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(f(x_i) - y_i)^2 \quad (3)$$

The lower the MSE value the better model performance (Chai and Draxler, 2014). The model selection is based on the higher R^2 value and lower MSE value.

3. Results and discussion

3.1. UMAP classification results

The classification results based on the UMAP downscaled two-dimensional map, are shown in Fig. 2. We then defined the boundary of the transform space for each classification as twice the mean of the standard deviations of the x, y components, written as $2\,\overline{\sigma_{x,y}}$ in Fig. 2 (right). The training data for the model refer to Crawford et al. (2023), and we use the same framework here. Based on the labels of the ChAMBRe training data, the classification results were categorized into the following categories: "Nettle", "Bacteria", "*Penicillium*", "*Cladosporium*", "Unclassified" and "Non-bacteria bioaerosols" where the "Unclassified" and "Non-bacteria bioaerosols" were bioaerosol species that are not precisely defined or potentially strongly fluorescing interferents.

The classification of the fungal kingdom is clear. Both *Cladosporium* and *Penicillium* are in their own transformed space. However, there was a high degree of convergence in the classification of bacteria. Some were distinguished as overlapping non-bacteria bioaerosols and bacterial particles. These conflated particles may be potential fungal particles. Unclassified particles were linked in three taxonomic spaces, especially between fungal spaces. A small number of particles located on the right side of the two-dimensional space during the sampling period were classified as nettle; the low counts observed is consistent with the start of the nettle pollen season in southwest England.

Fig. 3 compares the particle fluorescence response and morphological parameters between the ambient sample and training data. In the laboratory samples used as training data, bacteria rarely displayed fluorescence in the fifth to eighth fluorescence channels, therefore any particles initially classified as bacteria with fluorescence detected in channels 5–8 are rejected from the bacteria classification in post processing based on the laboratory results of Crawford et al. (2023). After rejecting this fraction of particles, further examination of the remaining subset revealed that some particles within the bacterial classification

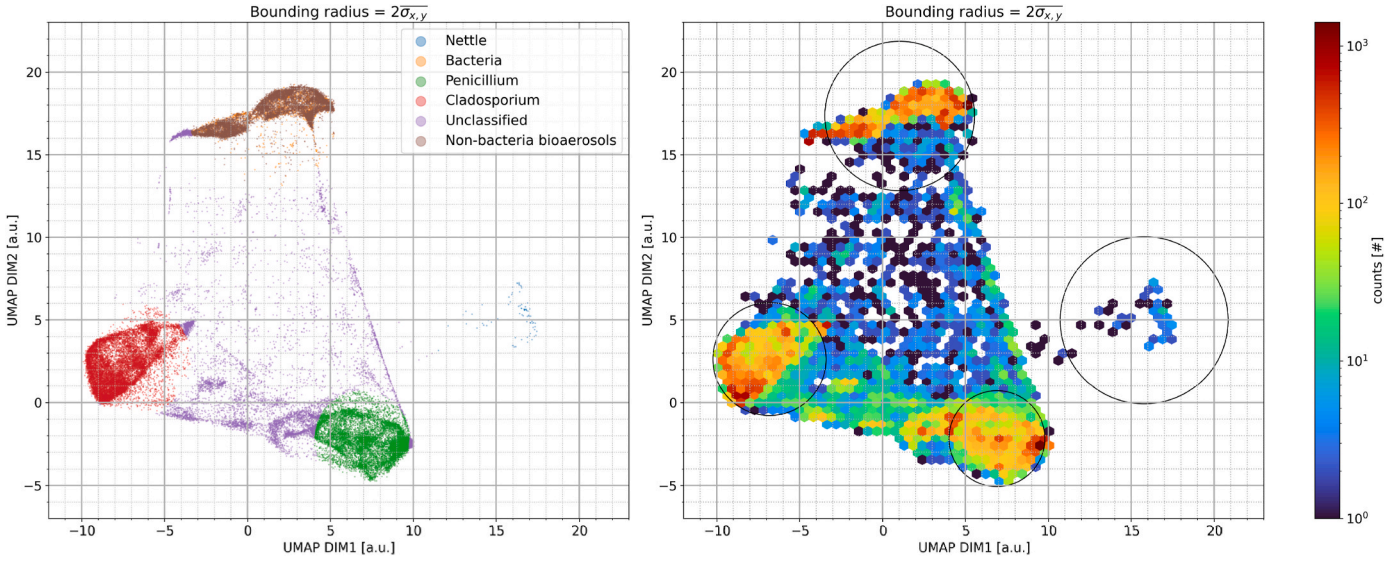

Fig. 2. Left: Devon data classification dimension reduce into 2D space. Right: 2D density scatter plot of the Devon data in the transformed space. The radius of the classification boundary for each category was defined as twice the mean standard deviation of the x and y components of each category 2 $\overline{\sigma_{x,y}}$.

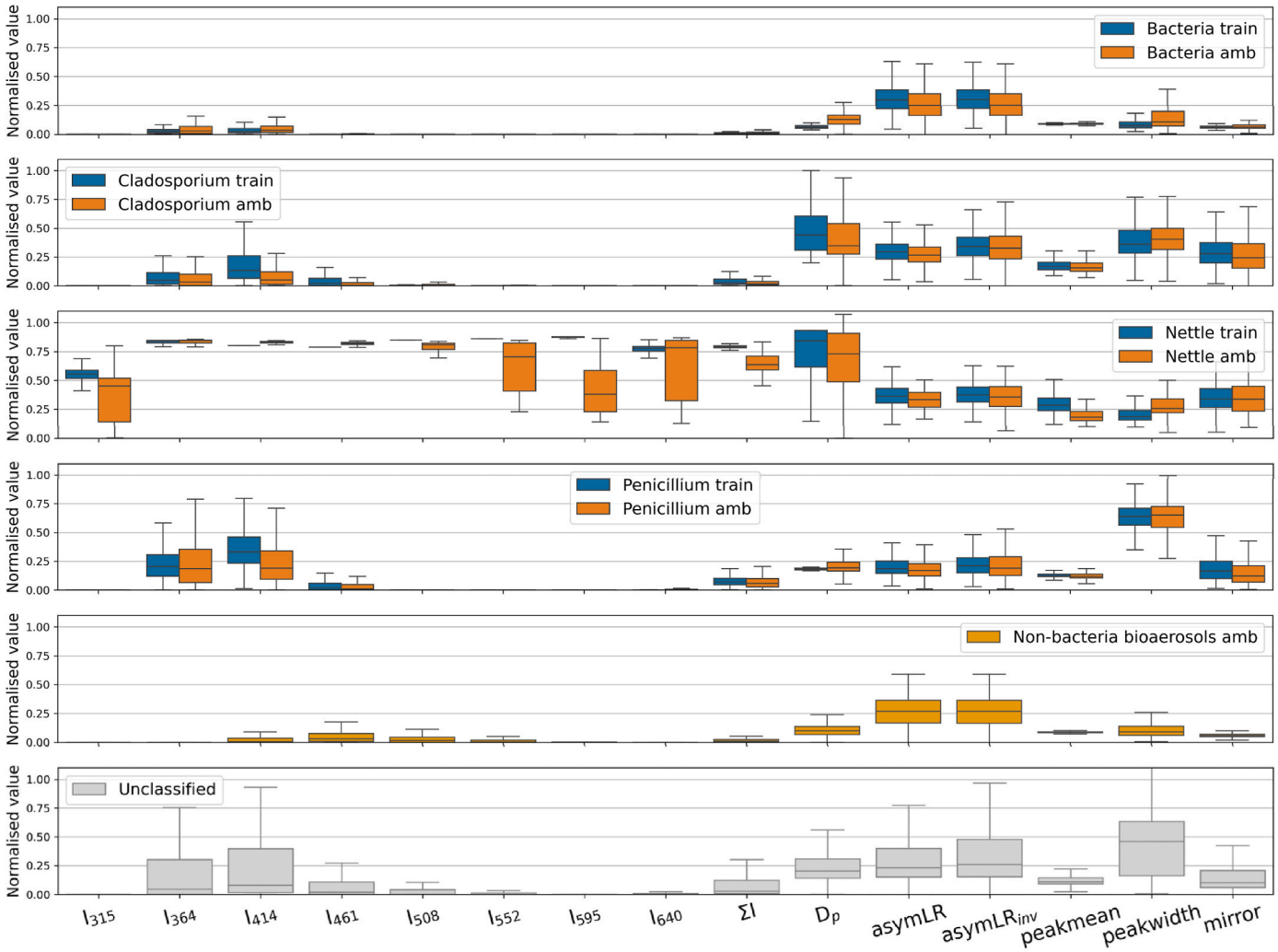

Fig. 3. Compared training data and environmental data, presented as box and whisker figures. Whiskers represent 5 %/95 %. Ith: fluorescent intensity at wavelength x [nm]; ∑I: sum of Ith; Dp: particle diameter; AsymLR: symmetry between left and right CMOS arrays; AsymLRinv: as AsymLR, but with the right array inverted; Peakmean: ratio of peak to mean CMOS array values; Peakwidth: estimate of the mean width of the array, defined as mid-point between mean and peak values; Mirror: measure of symmetry between top and bottom half of each array.

showed a fluorescent response in channel 4; based on procedure described in Crawford et al. (2023), the intensity of fluorescence in channel 4 (461 nm) should not be greater than that of channel 3 (414 nm) for bacterial particles. An additional filter is applied to the bacterial subset, where any particles where the intensity of the 4th channel exceeds that of the 3rd channel are also rejected from the bacterial classification.

Bacterial particles rejected by post processing were relabelled as "Non-bacteria bioaerosols". It maybe that these "Non-bacteria bioaerosols" may arise due to optical alignment drift, differences in the emission characteristics of the detected spores, or as a result of atmospheric or other processing. Generally, the rejected bacterial particles share characteristic similarities to *Penicillium* spores, and are likely misclassified *Penicillium* or fungal spores with similar characteristics.

According to the UMAP classification results, the most dominant class was *Cladosporium,* which accounted for 29.8 % of the total detected aerosols, followed by *Penicillium* with 24.1 % of the total detected aerosols. The fungal spores represented by *Cladosporium* and *Penicillium* together accounted for 54 % of the total. Bacteria is the next most abundant at 7.5 % contribution to the fluorescent population, while the lowest was Nettle pollen, at 0.1 %. The remaining "Unclassified" results accounted for 19.1 % and the "Non-bacteria bioaerosols" class accounted for 19.3 %. Potential fungal spores were included in the "Non-bacteria bioaerosols" class, which implies that the actual fraction of fungal spores is greater than 54 %. These two detected fungi are also noted in the UK Environment Agency report as the main types of fungi present on farms (Scaife et al., 2008).

3.2. *Temporal and spatial distribution of bioaerosol number concentrations*

Fig. 4 shows the time evolution of the classified data and meteorological parameters from April 13th to 29th. According to Fig. 4a, prior to 21st April, three bioaerosol classes originated mainly in the direction of agricultural land. Significant increases in concentration between April 24th and April 27th were seen followed by a gradual decline, maintaining elevated residual concentrations. The animal house source episodes clearly overlap with significant concentration enhancements. Among the fungal spores, the *Cladosporium* and *Penicillium* classes have similar variations. Both fungal spore types showed several peaks in concentration between April 20th and 27th, with large fluctuations during April 27th. Similarly, there were more peaks and higher concentrations in the direction of the animal house than in the direction of the farmland. *Cladosporium* had an average concentration of 3.79 L^{-1} in the animal house direction and 1.60 L^{-1} in the farmland direction. For *Penicillium*, the average concentration was 2.44 L^{-1} in the animal house

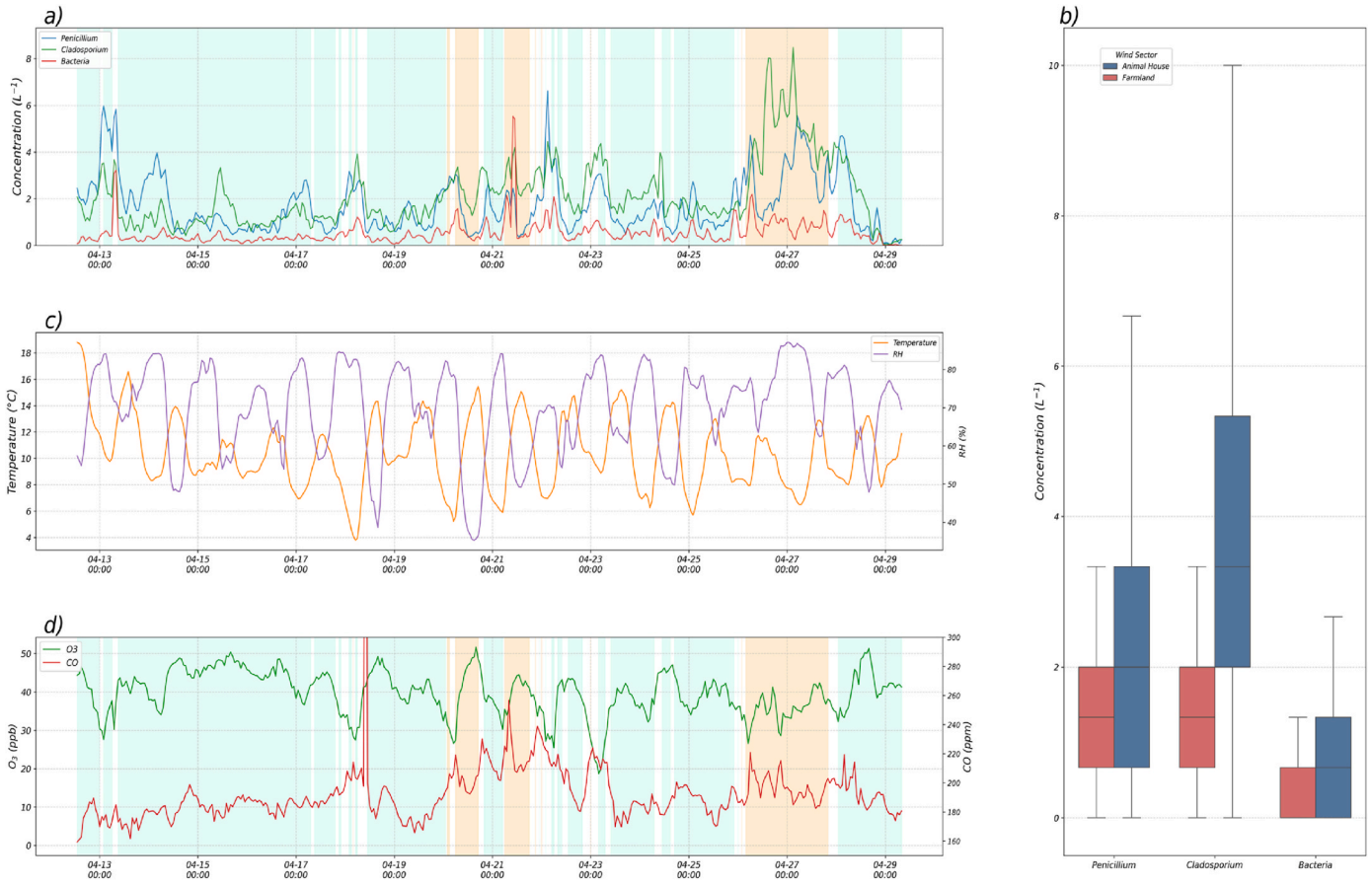

Fig. 4. Illustrates temporal variations in atmospheric concentrations of distinct microbial categories recorded between April 13 and 29, 2024. Plot a): hourly rolling averages of classified fungi; shaded areas indicate farmland sources (southwest) and animal house sources (northeast), respectively; b): box and whisker plot of three classified bioaerosols concentration comparation between the direction from the farmland and animal house. Whiskers denote the 5th and 95th percentiles; c): the time series of temperature versus relative humidity; d): the time series of CO (ppm) and O_3(ppb); shaded areas indicate farmland sources (southwest) and animal house sources (northeast), respectively.

direction and 1.52 L^{-1} in the farmland direction. This concentration range is similar to the bioaerosol concentrations observed on agricultural farms by Tarigan et al. (2017). This finding suggests that the release of bioaerosols is influenced by animal and human activities within the animal house area. Additionally, the diurnal variation curve of *Cladosporium* concentrations shown in Supplementary Fig. S5 does not align with the expected release timings of these spores, which typically exhibit high emissions between 11 a.m. and 2 p.m. (Stępalska and Wołek, 2009). Instead, an increase in emissions is observed between 4 a.m. and 10 a.m., coinciding with peak worker/animal activity. Meanwhile, following April 27th, significant livestock behavioral events were recorded at the station. Bacterial concentrations were observed to be relatively low with initially little variability. Similarly, April 21st the animal house direction showed a larger peak. Further comparing the concentration values of the two sources, as shown in the box plot in Fig. 4b, all three classified bioaerosol classes exhibit higher median values in the animal house sector. Previous studies have reported that 90 % of emitted particulates associated with animal husbandry are in the form of bioaerosols, the main sources of which include feed, litter, and excreta, with bacterial components including *Salmonella*, *Staphylococcus, E.coil etc.* (Gohel et al., 2024). It is reasonable to assume that these feeding-related products and biowaste are likely to have produced the main bacterial species detected downwind of the animal house by the MBS. Changes in meteorological parameters show periodicity. Preliminary analysis shows that the bioaerosol concentration increases during the time periods when the temperature has increased. Fig. 4d shows the curves for ozone and CO, with a significant peak in CO on 18 April, which could be instrumental calibration or artificial disturbance.

The spatial distribution of the concentrations of the main bioaerosol classes are shown in Fig. 5. The polar plots for the two main bioaerosol species, *Cladosporium* and *Penicillium*, both show clear areas of high concentrations in the northeast, while Bacteria, although less concentrated overall relative to the other two fungal species, were also a strong source in the northeast. However, the *Cladosporium* and *Penicillium* classes also exhibit sources to the southwest with a much stronger *Cladosporium* source appearing at higher wind speed, in particular, bioaerosol emissions are strongest at wind speeds of 5–15 m/s. Overall, all three bioaerosols shared a common point source of emissions in the direction of the northeastern animal house, and all were detected diffusing downwind. However, the transport distance of the emissions cannot be determined.

3.3. Single-factor modeling of bioaerosols and environmental variables

3.3.1. Spearman analysis

Fig. 6 shows the Spearman's correlation coefficient analysis results, which can determine the possible correlation between selected bioaerosol classes, in this case fungal spores, meteorological parameters and trace gases were the focus of the analysis. Between the meteorological parameters, relative humidity (RH) was positively correlated with fungal spores, and temperature (T), wind speed (WS), and wind direction (WD) were all negatively correlated. The statistical significance of all meteorological parameters were less than 0.001. Among the range of trace gases measured, *CO*, *NO* and *NO_2* showed positive

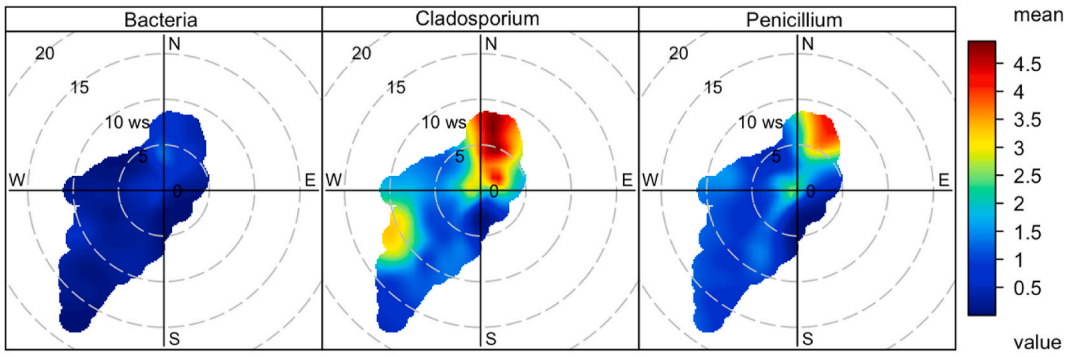

Fig. 5. Polar map for Bacteria, Cladosporium and Penicillium concentrations. The numbers on the circles in the figure represent wind speeds in m/s. The legend on the right represents the bioaerosol concentration per 5 min (L^{-1}).

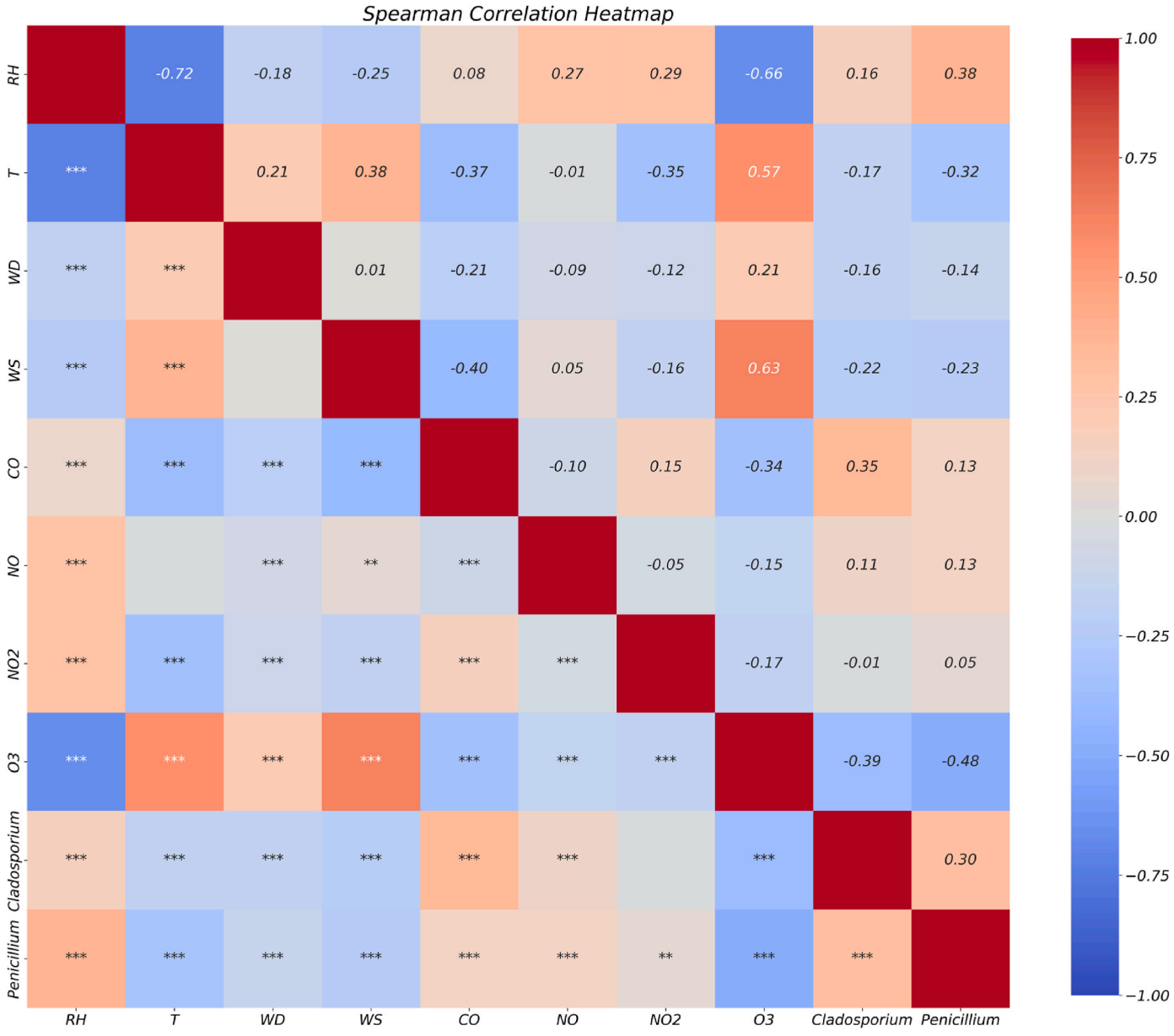

Fig. 6. *Heat map of Spearman's correlation between meteorological factors, trace gases and fungal spore concentrations. (*** is $p < 0.001$, ** is $p < 0.01$, * is $p < 0.05$).*

Spearman correlation coefficients. However, the correlation for NO_2 was not significant. O_3 showed a significant negative Spearman's coefficient.

Among the trace gases, the nitrogen dioxide, NO_2 P value was larger than 0.05, but could not pass the significance test. The nitric oxide, *NO*, Spearman value was relatively close to 0. After comprehensive analysis it was decided to that these three variables were not significant in this particular data set. Prior to introducing the remaining variables to the GAM model, it was necessary to test for the presence of multicollinearity between the individual predictor parameters. Multicollinearity affects the fitting performance of the model and can be judged by the VIF value. Through testing, *RH, T, WD, WS, CO and* O_3 VIF values were all less than 5, indicating they can be used in the GAM regression model.

3.3.2. GAM and SHAP analysis

When using UMAP classified fungal spore concentrations as response variables, the explanatory variables include *RH, T (°C), WD, WS, CO (ppm) and* O_3 *(ppb)*. In the single factor model, the R^2 value for *Penicillium* and *Cladosporium* on the test set were 0.36 and 0.34, and MSE were 0.005 and 0.006. The partial dependence plot (PDP), was then used which shows the marginal effect of the different features on the model predictive outcomes (Friedman, 2001). The result for two different fungal spore species acts with each response parameter in the model is shown in Fig. 7.

For *Penicillium*, Relative Humidity (RH) (Edof = 17.1, $P < 0.001$), trends show an increase in fungal concentrations at high humidity. Temperature (Edof = 20.7, $P < 0.001$) showed two peaks, which were between 6 °C and 10 °C and greater than 15 °C. Overall, *Penicillium* concentrations increased with increasing temperature. Source location, based on wind direction (*WD*) and wind speed (*WS*) (Edof = 16, $P < 0.001$; Edof = 18.7), matched the results shown in the polar map. Most of the *Penicillium* particles were found to be concentrated within wind directions 0–90° indicating they were significantly influenced by the proximity of animal housing, and with increased wind speeds, showing a significant positive response, particularly for wind speeds greater than 10 m/s. Carbon monoxide (*CO*) (Edof = 9, $P < 0.001$) and Ozone (O_3) (Edof = 25.8, $P < 0.001$) concentrations showed a non-linear response, where the positive effects of *CO* and O_3 on fungal spore concentrations diminished with increasing concentration.

For *Cladosporium*, Relative Humidity (RH) (Edof = 23.2, $P < 0.001$) showed a similar trend to *Penicillium*. However, concentrations of *Cladosporium* spores were more readily observed when RH was less than 50 %. Temperature (Edof = 12.6, $P < 0.001$) curves, in contrast, showed that *Cladosporium* was more inactive at lower temperatures (<8 °C) relative to *Penicillium*, whereas the positive effect was more pronounced at greater than 16 °C. Wind direction and wind speed (Edof = 11.5, $P < 0.001$; Edof = 7.1, $P < 0.001$) showed the source location for these particles was mainly from the animal house area. Meanwhile, the highest contribution to the particle concentrations were observed in the WS range 15–17 m/s. Carbon monoxide (Edof = 10.2, $P < 0.001$) and Ozone (Edof = 3.7, $P < 0.001$) have similar pattern as *Penicillium,* especially for O_3 which will be discussed in more detail below.

The overall model predictions are significant, However, the R^2 metrics are not high due to the size of the database, but the model results reproduce the observed patterns well. The results are consistent with both *Penicillium* and *Cladosporium* spores being produced in humid conditions and released passively, which means that higher wind speed and lower RH would favour their release after a period of high humidity. Typically, these conditions occur during a night and day cycle.

The specific contribution of each parameter to the variation of the GAM model can be visualised by the SHapley Additive exPlanations (SHAP) values (Lundberg and Lee, 2017). The magnitude of the effect of each parameter can be quantified by the absolute SHAP value. Fig. 8 displays the ranked impacts of each parameter on two fungal spore concentrations.

Our analysis shows that for both fungal spores, O_3 is the most important model parameter associated with changes in *Cladosporium* and *Penicillium* concentrations. We used the model to capture the fact that at a lower ozone concentration window (35–50 ppb), increases in ozone concentration can still negatively affect fungal spore release. This concentration window corresponds to common atmospheric ozone concentrations in many non-urban ecosystems, e.g. average ozone concentrations around 35 ppb in rural central England (Derwent et al., 2010). In Supplementary Fig. S5, the diurnal variation curves of the two bioaerosols exhibit high emissions during the early morning and evening hours, with relatively low emissions at midday and in the afternoon. In contrast, the diurnal variation of ozone shows the opposite pattern. Additionally, it has been observed that in different wind sectors, particularly in the direction of agricultural fields, there is a lag of approximately 1–1.5 h following bioaerosol emissions. This conclusion can be extended to a more environmentally representative range of concentrations, suggesting that higher atmospheric ozone concentrations may act as a proxy indicator for fungal spore release mechanisms. Despite the lack of exploration of the possible mechanisms of direct or ozone proxy related emission mechanisms in natural bioaerosol environments, these results suggest a novel detailed approach for future environmental studies. Previous studies on ozone related bioaerosol responses have been limited. In a controlled laboratory environment, Korzun et al. (2008) explored the impact of ozone on *Cladosporium* spp. Their findings revealed that exposure to very high ozone concentrations ranging from 11 to 12 ppm significantly compromised the survival of *Cladosporium* spp. The extent of conidial viability reduction was directly proportional to the duration of ozone exposure. Wen et al. (2020) compared the total nitrogen (TN) content in a suspension of fungal spores after ozone inactivation and found that the TN content in the suspension increased after 10 min of ozone application, demonstrating that ozone disrupts cellular integrity. In outdoor urban environments, Yang et al. (2024) noted that during high ozone episodes (HO episodes, ozone concentration: 102.3 ± 66.2 μg m^{-3}), a significant negative correlation was found between ozone concentrations and total airborne microbe (TAM) concentrations and that most of the bioaerosols were in the form of fine particles (<2.1 μm). Mechanistically, ozone destroys microbial DNA as well as cellular structure, and research has shown that the structure of atmospheric Gram-negative bacterial communities can be regulated by atmospheric ozone concentrations (Wang et al., 2020; Xu et al., 2017). Based on this, some studies have applied ozone as a method of suppressing fungal hazards in livestock houses. A study of dairy farms in the province of Giza found that fumigation at 80 ppm ozone for 10 min or 20 and 40 ppm for 20 min significantly inhibited the growth of fungi and bacteria and was more economical and efficient technique than traditional antimicrobials (Hassan et al., 2017). Despite the consistency between the modelled and laboratory analyses of ozone impacts, it is still important to consider that there may be some 'noise' in the correlations, as concentrations of background airborne spores arriving from more distant windward sources may be released from 1 h to several days earlier than the sample time under different conditions (Hirst et al., 1967). Further assessment of the effect of ozone on fungal spore release and transport will requires additional measurements to identify sources within the animal house.

CO ranked fourth in the *Penicillium* model and third in the *Cladosporium* model, and it is worth noting that *CO* has a negative effect on *Penicillium* concentration and more positive feedback on *Cladosporium* concentration. Interpretation of this phenomenon requires caution and may be due to synergistic effects of carbon monoxide with other trace gases or as a potential CO_2 transformation pathway affecting fungal activity, or simply an artifact of effects such as wind speed. It has been observed that under CO_2 enrichment conditions, the involvement of ozone affects the enzymatic activity of fungal spores, which could also be one of the underlying causes (Chung et al., 2006). However, we still cannot confirm with this data set whether *CO* will influence the concentration of fungal spores in the atmosphere.

To further investigate the links between environmental variables and bioaerosol concentrations, Fig. 9 shows the orthogonal distance

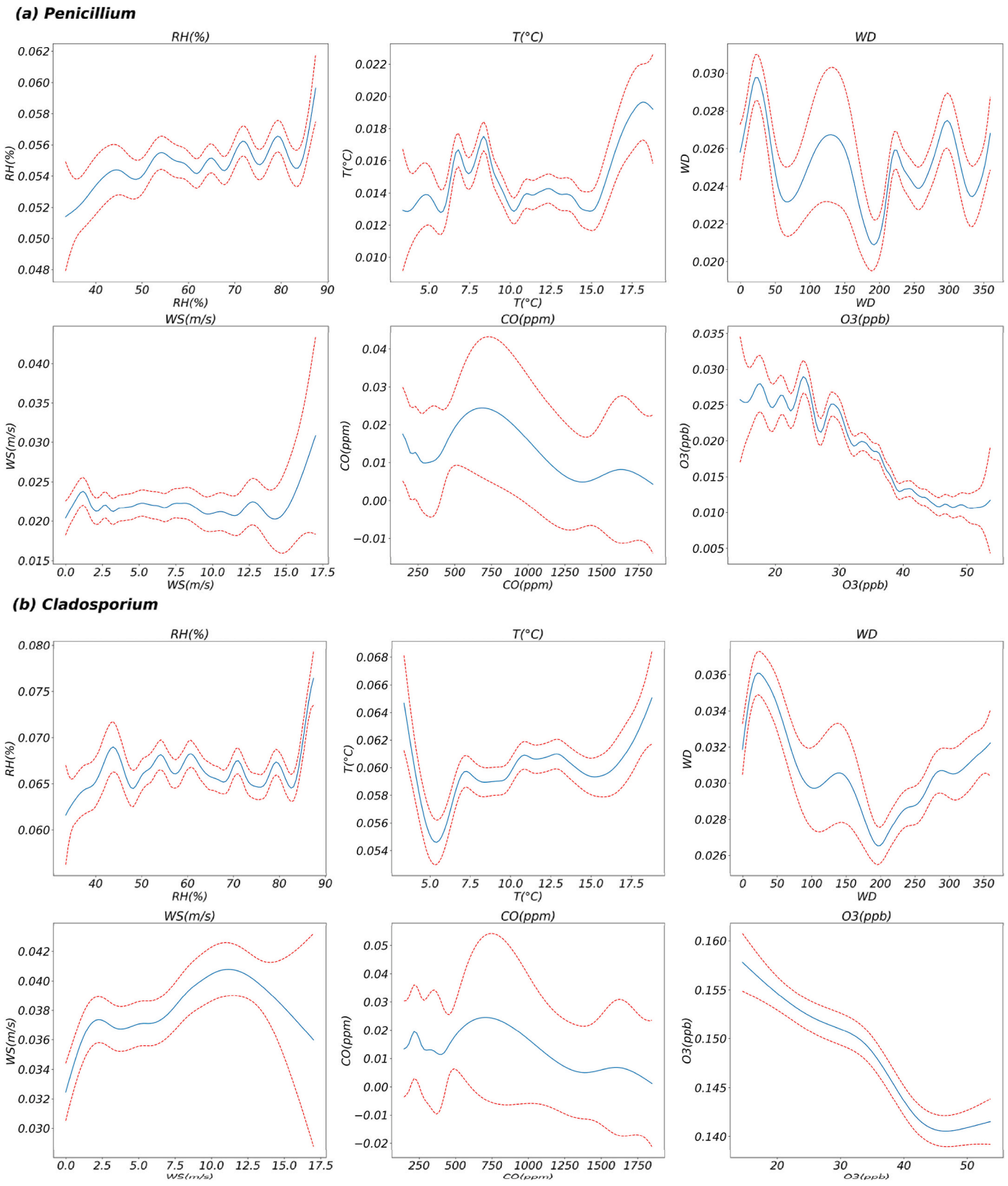

Fig. 7. Partial dependence plots (PDP) of two fungi species and environmental variables (including meteorological factors and trace gas data). The y-axis represents the degree to which bioaerosols respond to environmental variables. The x-axis denotates measured values of each factor. The blue line represents the response curve, and the red dotted line represents the 95 % confidence interval.

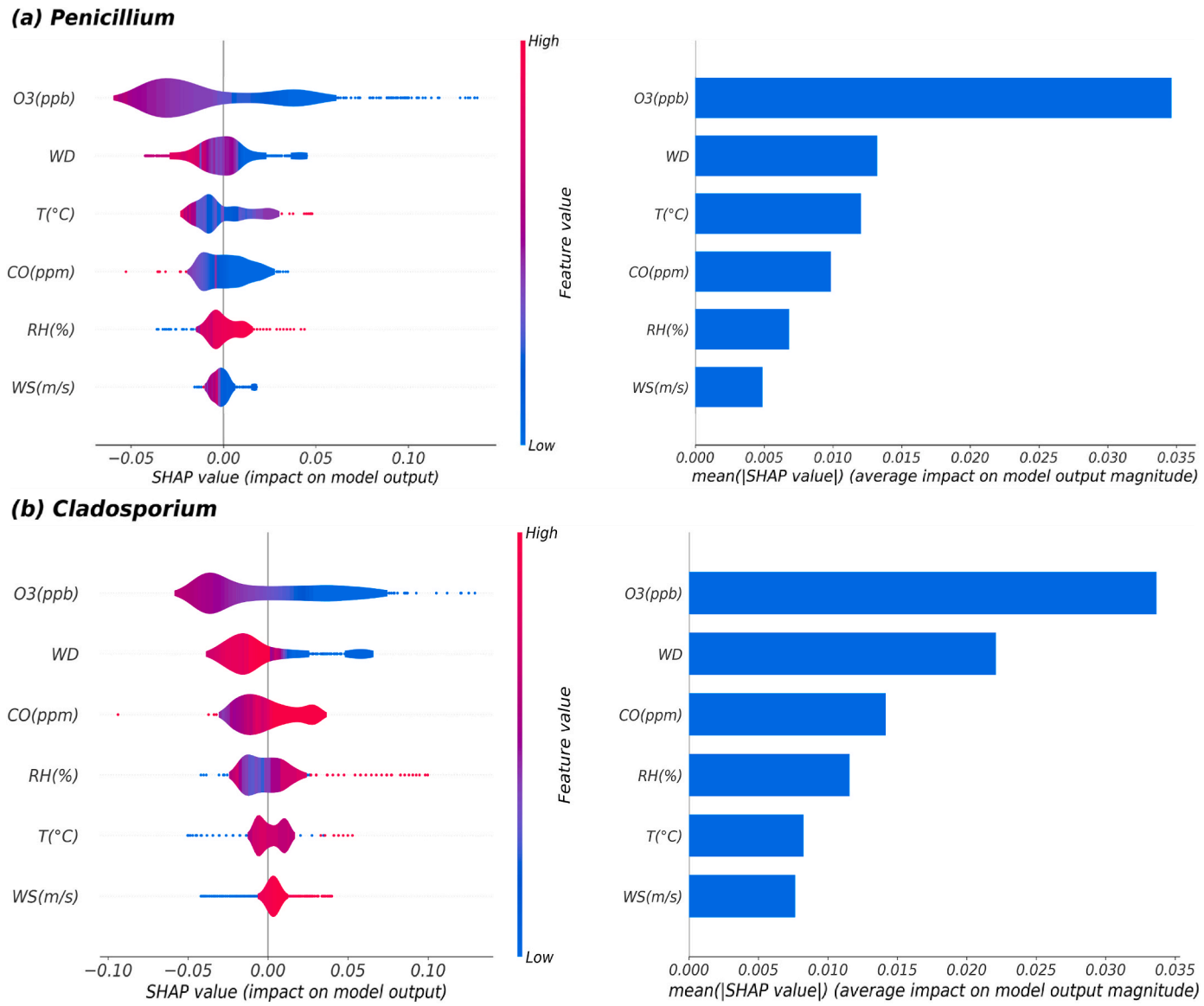

Fig. 8. Shap analysis of influential factors.

regression (ODR) analysis between the two fungal spores, from the direction of the animal house and the farmland, respectively, and the environmental parameters.

For *Penicillium* and *Cladosporium*, the deviation of the fit between the data points and meteorological data in the direction of the animal house was small, whereas the deviation of the fit in the direction of the farmland was large. Notably, for *Cladosporium*, relative humidity had a positive effect on the concentration, especially at RH of 75 % or more, which was particularly noticeable in the direction of the animal house. The release pattern of *Cladosporium* is confirmed by most studies to be inhibited at high relative humidity (Almeida et al., 2018; Grinn-Gofroń et al., 2019). The opposite trend in the present study may be due to 1) Higher relative temperatures (and therefore lower RH) in the animal house compared to the outdoor area, and greater release of *Cladosporium* at higher temperatures; 2) Insect activities in the animal house, as well as feed, straw mattresses, and faeces, may also be potential sources (Breitenbach and Simon-Nobbe, 2002; Nicoletti et al., 2024).

According to Tang et al. (2015), alternating cycles of high (97 %) and low relative humidity, where the low RH is below a critical threshold between 11 % and 43 %, can effectively inhibit the growth of Cladosporium. Therefore, introducing controlled fluctuations in relative humidity—such as through dehumidification and improved ventilation—may serve as a practical approach to suppress fungal proliferation in indoor environments like animal housing.

3.4. Implications for bioaerosol impacts and management

This study is the first to conduct real-time monitoring of bioaerosol concentrations within a single animal housing and the surrounding farmland area in the UK, generating a unique dataset on bioaerosol emission concentrations caused by animal activity behaviour. Machine learning methods were employed to classify the monitored bioaerosols, yielding reliable high time resolution results, which saved time and reduced labour costs as compared to offline methods. Additionally, the study captured the impact of various environmental variables on bioaerosol concentration changes under natural conditions.

Regarding the two primary bioaerosols, *Cladosporium* and *Penicillium*, their small particle size facilitates deposition in the upper respiratory tract of farm workers and livestock, potentially causing a range of allergic reactions and pulmonary diseases (Bamotra et al., 2025; Calderón-Ezquerro et al., 2025). In statistical studies in the UK and Europe, *Cladosporium* has been shown to cause a high allergy risk at airborne

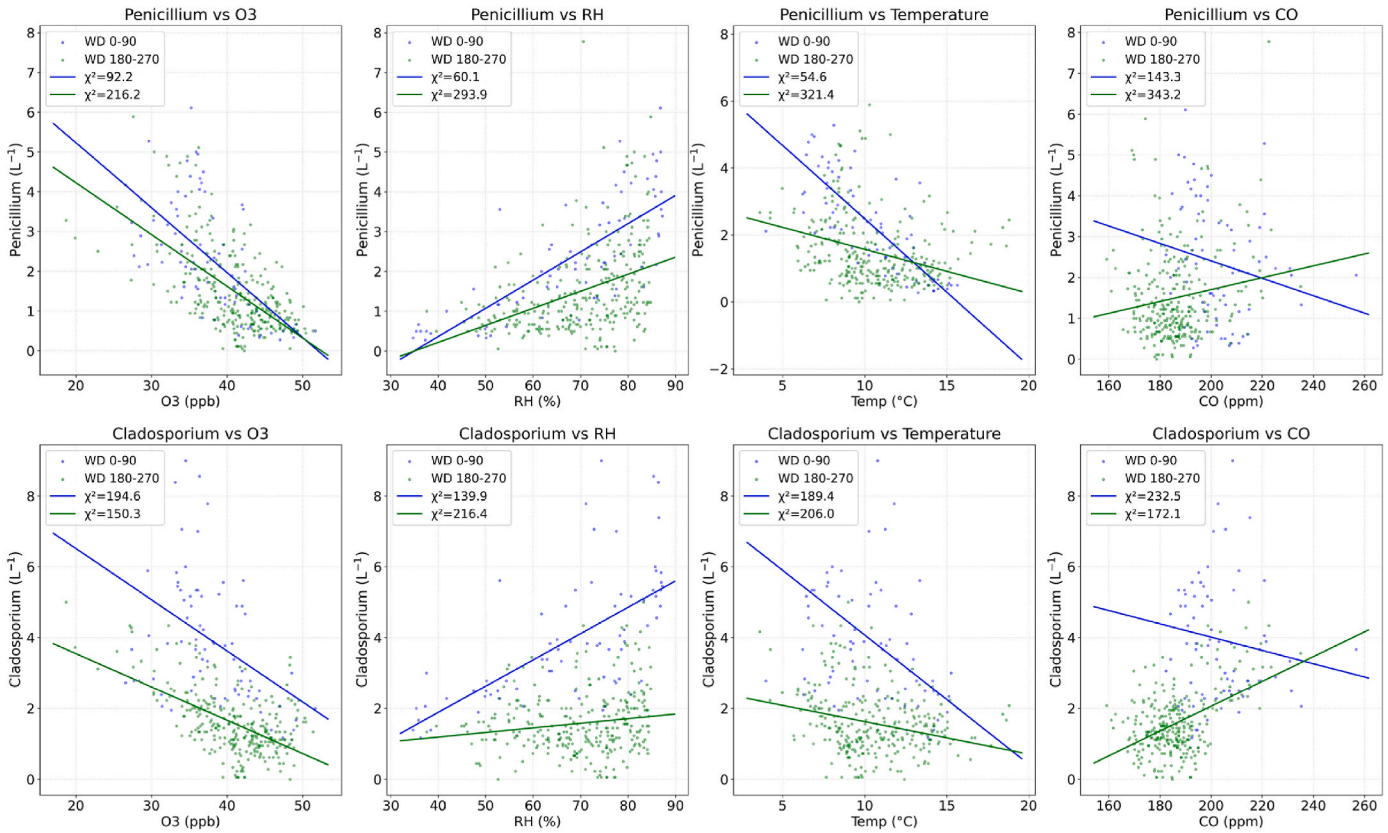

Fig. 9. Orthogonal Distance Regression (ODR) plots of Penicillium and Cladosporium concentrations (L^{-1}) versus environmental parameters. Each fungal species was divided into two source directions, animal house (WD 0–90) shown in blue and farmland (WD 180–270) shown in green. The χ^2 value was calculated for each fitted curve.

concentrations above 3000 spores per cubic metre (Breitenbach and Simon-Nobbe, 2002; Sadyś et al., 2016). Allergic reactions caused by *Cladosporium* are predominantly nasal congestion, which can lead to sinusitis and upper respiratory tract infections (Bamotra et al., 2025; Weryszko-Chmielewska et al., 2018). *Penicillium*, on the other hand, has been found to cause epidermal infections of the skin and nails, especially in immunocompromised individuals (Egbuta et al., 2017). Furthermore, when exposed to certain temperatures and humidity, filamentous fungi such as *Penicillium* can produce mycotoxins that are not only harmful to human health, but can also lead to hormonal disruption, immunosuppression and even death of livestock in animal houses (Brown et al., 2012; Egbuta et al., 2017; Richard, 2007).

The control of bioaerosols from an environmental management point of view is therefore crucial, especially in indoor closed environments such as animal houses. Filtration is a widely used method to capture and remove airborne bioaerosols through collision by varying the filter material, pore size and air flow rate (Liu et al., 2017). Combining the appropriate growth temperature and relative humidity thresholds for the target bioaerosol and suppressing the emission of the corresponding bioaerosol by changing the temperature and humidity. Ozone as a strong oxidising agent has a strong effect on reducing bacteria, fungi, spores etc. production, but at the same time it can cause irritation to the respiratory tract of humans or animals (Song et al., 2022). The current use of ozone for disinfection must be in a controlled environment because the effective disinfection concentration exceeds the recommended safety threshold (Lu et al., 2025). However, this study found that at lower ozone concentrations there may also be negative effects on certain bioaerosol emissions, providing insights for the future development of low-concentration ozone inactivation technologies.

3.5. Limitations

In this experiment, based on the work by Crawford et al. (2023), the post processing used to classify between airborne fungi and bacteria was further refined by setting the bacterial filter threshold for the fluorescence channels 5–8 and restricting modal fluorescence reduced the conflation pathway and showed good classification results. The 2D classification space presented by the UMAP classifier may however not contain all hotspot counts when using twice the standard deviation as the classification boundary and will be examined in future studies. Meanwhile, other frequently occurring allergy-causing fungi such as *Aspergillus* and *Basidiomycetes* were not included due to the lack of spectral and morphological training data. This dataset can be re-investigated with updated classifiers following further laboratory experiments.

Secondly, due to the limited duration of the database, interpretation of some variables using GAM may result in high complexity and uncertainty in interpretation. Temperature and RH did not change significantly during the sampling period, and there were no significant weather events such as precipitation, which may have enhanced or mitigated certain bioaerosol emissions. Therefore, the model may need to be retrained under a wider range of weather conditions to fully capture the T, RH responses. At the same time, there is still space to improve the fitting accuracy of the model due to the low concentration of bioaerosols in the natural environment and the limitations of the current real-time measurement instruments (An et al., 2024; Šantl-Temkiv et al., 2020).

Thirdly, a potential improvement of this study would be to continue cross-season sampling with supporting offline reference measurements (e.g., Hirst sport trap) to better reflect the seasonal universality of bioaerosols. With seasonal changes, changes in relative humidity and

temperature can also affect the release of bioaerosols and their concentration in the air. For indoor animal housing environments, the tendency to close doors/windows to keep animals warm as temperatures drop in autumn and winter, this behavior may lead to an accumulation of bioaerosol concentrations indoors, causing health impacts for workers and livestock. (Anderson et al., 2016).

4. Conclusions

In agricultural environments—particularly during the planning and operation of livestock housing systems—it is essential to implement effective protective measures to mitigate farm workers and nearby inhabitants' exposure to bioaerosols. A key step toward this goal is the identification of bioaerosol sources and a better understanding of their emission patterns and airborne concentrations. The main findings are summarized as follows.

(1) Bioaerosol particles at the NWFP station were detected mainly from the Northeast and south-west sectors, and high-intensity point sources were clearly identified. The north-eastern sector was in the direction of animal houses while the south-western sector consisted mainly of farmlands. And the peak changes in bioaerosol concentrations monitored in real time coincide with the timing of animal behaviour recorded at the site. *Cladosporium* and *Penicillium* animal house sources mean concentration were 2.37 times and 1.61 times farmland sources, respectively.
(2) Bioaerosols appear to be generated from mainly local source emissions compared to other wind sectors. The temporal trend showed a peak on the second day of measurement and several smaller peaks at intervals in the following days, but these dissipated quickly showing an overall decreasing trend. This is likely due to specific agricultural activities and or animal activities. At the same time, the dispersion can be seen downwind of the point source of the animal house, but the dispersion distance is uncertain.
(3) For bioaerosols, our model is able to reproduce the observation that bioaerosol releases are low at low ozone concentration levels within this study environment. This finding is informative for improving future real-time bioaerosol emission modelling applications. However, there is still a lack of monitoring of the response of bioaerosols to real-time variations with respect to relevant meteorological drivers and atmospheric trace gases within different ecosystems. It may also suggest the possibility of using ambient ozone concentrations to inform future laboratory studies for investigating potential mechanisms on the effect on bioaerosols in this concentration window. However, it also needs to be considered an uncertainty that there may be a negative correlation effect due to a reduction in turbulence, an increase in atmospheric concentrations of mixed ozone and a decrease in emissions of source fungal spores. This uncertainty is of interest to inform further research into these parameters.

Overall, from a bioaerosol regulatory perspective, it is important to strengthen monitoring around and downwind of agricultural facilities. From a pollution control perspective, the possibility of applying temperature-humidity regulation mechanism, and ozone sterilisation within safe concentration thresholds for livestock houses cleaning could be discussed. Future research will focus on 1) improvement of bioaerosol classification techniques; 2) increasing routine real-time monitoring to improve sampling periods to observe the effects of a wider range of meteorological drivers on bioaerosols; 3) Study of general patterns and relationships between bioaerosol releases and target trace gases in conjunction with laboratory controlled variable experiments. These works can further help to accurately classify bioaerosols and improve the existing instrument technology to gain more in-depth insights into the general laws of bioaerosol emissions.

CRediT authorship contribution statement

Zhuo Chen: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Formal analysis, Data curation. **Ian Crawford:** Writing – review & editing, Visualization, Validation, Software, Methodology, Formal analysis. **Emily Matthews:** Writing – review & editing, Resources, Methodology, Funding acquisition, Data curation, Conceptualization. **Michael Flynn:** Resources, Project administration, Investigation, Conceptualization. **Thomas Bannan:** Methodology, Investigation, Funding acquisition, Data curation. **Laura Cardenas:** Writing – review & editing, Resources. **Jon S. West:** Writing – review & editing, Supervision, Resources, Project administration, Investigation, Funding acquisition. **Hugh Coe:** Supervision, Resources, Project administration, Methodology, Funding acquisition, Conceptualization. **David Topping:** Writing – review & editing, Supervision, Software, Resources, Funding acquisition. **Martin Gallagher:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

There is no conflict of interest.

Zhuo Chen reports financial support was provided by EPSRC Centre for Doctoral Training in Aerosol Science. Jon West reports financial support was provided by Biotechnology and Biological Sciences Research Council. Laura Cardenas reports financial support was provided by Biotechnology and Biological Sciences Research Council. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

ZC acknowledges the EPSRC Centre for Doctoral Training in Aerosol Science for funding, grant reference EP/S023593/1; LC and JSW were supported by the Biotechnology and Biological Sciences Research Council Institute Strategic Programmes: Resilient Farming Futures (BB/X010961/1; BBS/E/RH/230004A) and Growing Health (BBS/E/RH/230003A; BBS/E/RH/230003C), and the North Wyke Farm Platform (BBS/E/RH/23NB0008).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jenvman.2025.127033.

Data availability

Data will be made available on request.

References

Akaike, H., 1987. Factor analysis and AIC. Psychometrika 52, 317–332.
Akiba, T., Sano, S., Yanase, T., Ohta, T., Koyama, M., 2019. Optuna: a next-generation hyperparameter optimization framework. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pp. 2623–2631.
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.
An, T., Liang, Z., Chen, Z., Li, G., 2024. Recent progress in online detection methods of bioaerosols. Fundamental Research 4, 442–454.
Anderson, B.D., Ma, M., Xia, Y., Wang, T., Shu, B., Lednicky, J.A., Ma, M.-J., Lu, J., Gray, G.C., 2016. Bioaerosol sampling in modern agriculture: a novel approach for emerging pathogen surveillance? The Journal of infectious diseases 214, 537–545.
Anguita, D., Ghelardoni, L., Ghio, A., Oneto, L., Ridella, S., 2012. The'K'In K-fold Cross Validation. ESANN, pp. 441–446.
Bamotra, S., Yadav, S., Tandon, A., 2025. Air Pollution from Dumping Sites in India: Implications on Environmental Health. Water, Air, & Soil Pollution 236, 1–58.

Bennett, K.D., Willis, K.J., 2001. Pollen. Tracking Environmental Change Using Lake Sediments: Terrestrial, Algal, and Siliceous Indicators, pp. 5–32.
Bigg, E., Soubeyrand, S., Morris, C., 2014. Rainfall feedback via persistent effects on bioaerosols. Atmos. Chem. Phys. Discuss. 14, 25503–25532.
Breitenbach, M., Simon-Nobbe, B., 2002. The allergens of cladosporium. Fungal allergy and pathogenicity 81, 48–72.
Brown, D.W., Butchko, R.A., Baker, S.E., Proctor, R.H., 2012. Phylogenomic and functional domain analysis of polyketide synthases in Fusarium. Fungal Biology 116, 318–331.
Calderón-Ezquerro, M.C., Brunner-Mendoza, C., Guerrero-Guerra, C., Sanchez-Flores, A, Salinas-Peralta, I., Toriello, C., Ponce-de León, A., Ortega-Rosas CI, C.I., 2025. Genetic sequencing of the airborne fungal spectrum and air quality at a public hospital in Mexico City. PLOS Global Public Health 5, e0004784.
Carslaw, D.C., Ropkins, K., 2012. Openair—An R package for air quality data analysis. Environ. Model. Software 27, 52–61.
Chai, T., Draxler, R.R., 2014. Root mean square error (RMSE) or mean absolute error (MAE). Geosci. Model Dev. Discuss. (GMDD) 7, 1525–1534.
Cheng, B., Ma, Y., Feng, F., Zhang, Y., Shen, J., Wang, H., Guo, Y., Cheng, Y., 2021. Influence of weather and air pollution on concentration change of PM2. 5 using a generalized additive model and gradient boosting machine. Atmospheric environment 255, 118437.
Chung, H., Zak, D.R., Lilleskov, E.A., 2006. Fungal community composition and metabolism under elevated CO 2 and O 3. Oecologia 147, 143–154.
Crawford, I., Bower, K., Topping, D., Di Piazza, S., Massabò, D., Vernocchi, V., Gallagher, M., 2023. Towards a UK airborne bioaerosol climatology: real-time monitoring strategies for high time resolution bioaerosol classification and quantification. Atmosphere 14, 1214.
Crawford, I., Ruske, S., Topping, D., Gallagher, M., 2015. Evaluation of hierarchical agglomerative cluster analysis methods for discrimination of primary biological aerosol. Atmos. Meas. Tech. 8, 4979–4991.
Crawford, I., Topping, D., Gallagher, M., Forde, E., Lloyd, J.R., Foot, V., Stopford, C., Kaye, P., 2020. Detection of airborne biological particles in indoor air using a real-time advanced morphological parameter uv-lif spectrometer and gradient boosting ensemble decision tree classifiers. Atmosphere 11, 1039.
Derwent, R.G., Witham, C.S., Utembe, S.R., Jenkin, M.E., Passant, N.R., 2010. Ozone in central England: the impact of 20 years of precursor emission controls in Europe. Environ. Sci. Pol. 13, 195–204.
Diez, S., Lacy, S., Coe, H., Urquiza, J., Priestman, M., Flynn, M., Marsden, N., Martin, N. A., Gillott, S., Bannan, T., 2024. Long-term evaluation of commercial air quality sensors: an overview from the QUANT (quantification of utility of atmospheric network technologies) study. Atmos. Meas. Tech. 17, 3809–3827.
Diez, S., Lacy, S.E., Bannan, T.J., Flynn, M., Gardiner, T., Harrison, D., Marsden, N., Martin, N.A., Read, K., Edwards, P.M., 2022. Air Pollution Measurement Errors: Is Your Data Fit for Purpose? Atmospheric Measurement Techniques, vol. 15, pp. 4091–4105.
dos Santos, M.P., Heinemann, A.B., Stone, L.F., da Matta, D.H., de Castro, J.R., dos Santos, A.B., 2021. Nitrogen determination in irrigated rice using spectral reflectance. Agron. J. 113, 5087–5101.
Du, J., Qiao, F., Lu, P., Yu, L., 2022. Forecasting ground-level ozone concentration levels using machine learning. Resour. Conserv. Recycl. 184, 106380.
Egbuta, M.A., Mwanza, M., Babalola, O.O., 2017. Health risks associated with exposure to filamentous fungi. International journal of environmental research and public health 14, 719.
Foot, V.E., Kaye, P.H., Stanley, W.R., Barrington, S.J., Gallagher, M., Gabey, A., 2008. Low-Cost real-time Multiparameter bio-aerosol Sensors, Optically Based Biological and Chemical Detection for Defence IV. SPIE, pp. 78–89.
Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. Ann. Stat. 1189–1232.
Gabbarini, V., Rossi, R., Ciparisse, J., Puleio, A., Malizia, A., Gaudio, P., 2019. An UltraViolet laser-induced fluorescence (UV-LIF) system to detect, identify and measure the concentration of biological agents in the environment: a preliminary study. J. Instrum. 14, C07009.
Gibbs, S.G., Green, C.F., Tarwater, P.M., Mota, L.C., Mena, K.D., Scarpino, P.V., 2006. Isolation of antibiotic-resistant bacteria from the air plume downwind of a swine confined or concentrated animal feeding operation. Environmental Health Perspectives 114, 1032–1037.
Gladding, T., Rolph, C.A., Gwyther, C., Kinnersley, R., Walsh, K., Tyrrel, S., 2020. Concentration and composition of bioaerosol emissions from intensive farms: pig and poultry livestock. J. Environ. Manag. 272, 111052.
Gohel, R., Siabbweka, M., Singh, R., Thanki, A.A., Jadeja, U., 2024. Sampling, Detection, and Health Impacts of Bioaerosols Emitted from Livestock Facilities, Bioaerosols Emission from Anthropogenic Sources. Elsevier, pp. 141–161.
Gosselin, M.I., Rathnayake, C.M., Crawford, I., Pöhlker, C., Fröhlich-Nowoisky, J., Schmer, B., Després, V.R., Engling, G., Gallagher, M., Stone, E., 2016. Fluorescent bioaerosol particle, molecular tracer, and fungal spore concentrations during dry and rainy periods in a semi-arid forest. Atmos. Chem. Phys. 16, 15165–15184.
Grgacic, E.V., Anderson, D.A., 2006. Virus-like particles: passport to immune recognition. Methods 40, 60–65.
Grinn-Gofroń, A., Nowosad, J., Bosiacka, B., Camacho, I., Pashley, C., Belmonte, J., De Linares, C., Ianovici, N., Manzano, J.M.M., Sadyś, M., 2019. Airborne alternaria and cladosporium fungal spores in Europe: forecasting possibilities and relationships with meteorological parameters. Sci. Total Environ. 653, 938–946.
Gsj, S., Ramakodi, M.P., Tvbps, R., 2023. Review of bioaerosols from different sources and their health impacts. Environ. Monit. Assess. 195, 1321.
Hassan, A., Howayda, M., Hanan, K.M., 2017. Antimicrobial potential of ozone on fungal and bacterial contamination of animal feed that caused diseases in some buffalo farms. In: 1st International Conference, Animal Health Research Institute. ARC, Egypt, pp. 9–13.
Hastie, T.J., 2017. Generalized Additive Models. Statistical Models in S, pp. 249–307.
Hawkins, J., Beaumont, D., Sint, H., Harris, P., 2023a. The North Wyke Farm Platform: Field Survey Data.
Hawkins, J., Griffith, B., Sint, H., Harris, P., 2023b. The North Wyke Farm Platform: Design, Establishment and Development.
Herman, A.A., Hastie, T.J., 1990. An analysis of gestational age, neonatal size and neonatal death using nonparametric logistic regression. J. Clin. Epidemiol. 43, 1179–1190.
Hirst, J., Stedman, O., Hurst, G., 1967. Long-distance spore transport: vertical sections of spore clouds over the sea. Microbiology 48, 357–377.
Huffman, J.A., Perring, A.E., Savage, N.J., Clot, B., Crouzy, B., Tummon, F., Shoshanim, O., Damit, B., Schneider, J., Sivaprakasam, V., 2020. Real-time sensing of bioaerosols: review and current perspectives. Aerosol. Sci. Technol. 54, 465–495.
Katz, A., Alimova, A., Xu, M., Rudolph, E., Shah, M.K., Savage, H.E., Rosen, R.B., McCormick, S.A., Alfano, R.R., 2003. Bacteria size determination by elastic light scattering. IEEE J. Sel. Top. Quant. Electron. 9, 277–287.
Ko, G., Simmons Iii, O.D., Likirdopulos, C.A., Worley-Davis, L., Williams, M., Sobsey, M. D., 2008. Investigation of bioaerosols released from swine farms using conventional and alternative waste treatment and management technologies. Environmental science & technology 42, 8849–8857.
Könemann, T., Savage, N., Klimach, T., Walter, D., Fröhlich-Nowoisky, J., Su, H., Pöschl, U., Huffman, J.A., Pöhlker, C., 2019. Spectral intensity bioaerosol sensor (SIBS): an instrument for spectrally resolved fluorescence detection of single particles in real time. Atmos. Meas. Tech. 12, 1337–1363.
Korzun, W., Hall, J., Sauer, R., 2008. The Effect of Ozone on Common Environmental Fungi, vol. 21. American Society for Clinical Laboratory Science, pp. 107–111.
Kumar, P., Kausar, M.A., Singh, A., Singh, R., 2021. Biological contaminants in the indoor air environment and their impacts on human health. Air Qual. Atmos. Health 14, 1723–1736.
Kumar, P., Tiwari, S., Uguz, S., Li, Z., Gonzalez, J., Wei, L., Samuel, R.S., Zhang, Y., Yang, X., 2024. Bioaerosols downwind from animal feeding operations: a comprehensive review. J. Hazard Mater., 135825
Landau, S., Everitt, B.S., 2003. A Handbook of Statistical Analyses Using SPSS. Chapman and Hall/CRC.
Liu, G., Xiao, M., Zhang, X., Gal, C., Chen, X., Liu, L., Pan, S., Wu, J., Tang, L., Clements-Croome, D., 2017. A review of air filtration technologies for sustainable and healthy building ventilation. Sustain. Cities Soc. 32, 375–396.
Loaiza, J.G., Rangel-Peraza, J.G., Monjardín-Armenta, S.A., Bustos-Terrones, Y.A., Bandala, E.R., Sanhouse-García, A.J., Rentería-Guevara, S.A., 2023. Surface water quality assessment through remote sensing based on the box–cox transformation and linear regression. Water 15, 2606.
Lu, K., Zhang, J., Li, Z., Li, Y., 2025. Bioaerosols in various working and living environments and their control measure: a review. Curr. Pollut. Rep. 11, 24.
Lundberg, S.M., Lee, S.-I., 2017. A unified approach to interpreting model predictions. Adv. Neural Inf. Process. Syst. 30.
Maya-Manzano, J.M., Smith, M., Markey, E., Hourihane Clancy, J., Sodeau, J., O′ Connor, D.J., 2021. Recent developments in monitoring and modelling airborne pollen, a review. Grana 60, 1–19.
McInnes, L., Healy, J., Melville, J., 2018. Umap: uniform manifold approximation and projection for dimension reduction. arXiv preprint arXiv:1802.03426.
Mpaka, Y.W., von der Heyden, B.P., 2024. Enhanced classification of pyrite generations based on mineral chemistry using uniform manifold approximation and projection (UMAP). J. Afr. Earth Sci. 218, 105363.
Nicoletti, R., Russo, E., Becchimanzi, A., 2024. Cladosporium—insect relationships. Journal of Fungi 10, 78.
Orr, R., Griffith, B., Rose, S., Hatch, D., Hawkins, J., Murray, P., 2011. Designing and creating the north wyke farm platform. Abstracts Catchment Science 2011. Dublin, 14-16 September 2011, 35.
Pan, Y.-L., Aptowicz, K., Arnold, J., Cheng, S., Kalume, A., Piedra, P., Wang, C., Santarpia, J., Videen, G., 2022. Review of elastic light scattering from single aerosol particles and application in bioaerosol detection. J. Quant. Spectrosc. Radiat. Transf. 279, 108067.
Patel, T.Y., Buttner, M., Rivas, D., Cross, C., Bazylinski, D.A., Seggev, J., 2018. Variation in airborne fungal spore concentrations among five monitoring locations in a desert urban environment. Environ. Monit. Assess. 190, 1–10.
Qi, X., Mei, G., Cuomo, S., Liu, C., Xu, N., 2021. Data analysis and mining of the correlations between meteorological conditions and air quality: a case study in beijing. Internet of Things 14, 100127.
Radon, K., Schulze, A., Ehrenstein, V., Van Strien, R.T., Praml, G., Nowak, D., 2007. Environmental exposure to confined animal feeding operations and respiratory health of neighboring residents. Epidemiology 18, 300–308.
Ramsey, N.R., Klein, P.M., Moore III, B., 2014. The impact of meteorological parameters on urban air quality. Atmos. Environ. 86, 58–67.
Richard, J.L., 2007. Some major mycotoxins and their mycotoxicoses—An overview. International journal of food microbiology 119, 3–10.
Robinson, N.H., Allan, J., Huffman, J., Kaye, P.H., Foot, V., Gallagher, M., 2013. Cluster analysis of WIBS single-particle bioaerosol data. Atmos. Meas. Tech. 6, 337–347.
Sabban, L., van Hout, R., 2011. Measurements of pollen grain dispersal in still air and stationary, near homogeneous, isotropic turbulence. J. Aerosol Sci. 42, 867–882.
Sadyś, M., Kennedy, R., West, J.S., 2016. Potential impact of climate change on fungal distributions: analysis of 2 years of contrasting weather in the UK. Aerobiologia 32, 127–137.

Sajjad, B., Hussain, S., Rasool, K., Hassan, M., Almomani, F., 2023. Comprehensive insights into advances in ambient bioaerosols sampling, analysis and factors influencing bioaerosols composition. Environmental Pollution, 122473.
Sakia, R.M., 1992. The box-cox transformation technique: a review. J. R. Stat. Soc. - Ser. D Statistician 41, 169–178.
Šantl-Temkiv, T., Sikoparija, B., Maki, T., Carotenuto, F., Amato, P., Yao, M., Morris, C. E., Schnell, R., Jaenicke, R., Pöhlker, C., 2020. Bioaerosol field measurements: challenges and perspectives in outdoor studies. Aerosol. Sci. Technol. 54, 520–546.
Sauvageat, E., Zeder, Y., Auderset, K., Calpini, B., Clot, B., Crouzy, B., Konzelmann, T., Lieberherr, G., Tummon, F., Vasilatou, K., 2020. Real-time pollen monitoring using digital holography. Atmos. Meas. Tech. 13, 1539–1550.
Savage, N.J., Krentz, C.E., Könemann, T., Han, T.T., Mainelis, G., Pöhlker, C., Huffman, J. A., 2017. Systematic characterization and fluorescence threshold strategies for the wideband integrated bioaerosol sensor (WIBS) using size-resolved biological and interfering particles. Atmos. Meas. Tech. 10, 4279–4302.
Scaife, H., Crook, B., Jordinson, G., 2008. PPC Bioaerosols (Dust and Particulates) Potentially Emanating from Intensive Agriculture and Potential Effects on Human Health. Environment Agency.
Servén, D., Brummitt, C., 2018. Pygam: Generalized Additive Models in Python. Zenodo.
Smit, L.A., Hooiveld, M., van der Sman-de Beer, F., Opstal-van Winden, A.W., Beekhuizen, J., Wouters, I.M., Yzermans, C.J., Heederik, D., 2014. Air pollution from livestock farms, and asthma, allergic rhinitis and COPD among neighbouring residents. Occup. Environ. Med. 71, 134–140.
Song, L., Zhou, J., Wang, C., Meng, G., Li, Y., Jarin, M., Wu, Z., Xie, X., 2022. Airborne pathogenic microorganisms and air cleaning technology development: a review. J. Hazard Mater. 424, 127429.
Stępalska, D., Wołek, J., 2009. Intradiurnal periodicity of fungal spore concentrations (alternaria, botrytis, cladosporium, didymella, ganoderma) in cracow, Poland. Aerobiologia 25, 333–340.
Tang, K., Sánchez-Parra, B., Yordanova, P., Wehking, J., Backes, A.T., Pickersgill, D.A., Maier, S., Sciare, J., Pöschl, U., Weber, B., 2022. Bioaerosols and atmospheric ice nuclei in a mediterranean dryland: community changes related to rainfall. Biogeosciences 19, 71–91.
Tang, W., Kuehn, T., Simcik, M.F., 2015. Effects of temperature, humidity and air flow on fungal growth rate on loaded ventilation filters. J. Occup. Environ. Hyg. 12, 525–537.
Tarigan, Y.G., Chen, R.-Y., Lin, H.-C., Jung, C.-Y., Kallawicha, K., Chang, T.-P., Hung, P.-C., Chen, C.-Y., Chao, H.J., 2017. Fungal bioaerosol exposure and its effects on the health of mushroom and vegetable farm workers in Taiwan. Aerosol Air Qual. Res. 17, 2064–2075.
Uetake, J., Tobo, Y., Uji, Y., Hill, T.C., DeMott, P.J., Kreidenweis, S.M., Misumi, R., 2019. Seasonal changes of airborne bacterial communities over Tokyo and influence of local meteorology. Front. Microbiol. 10, 1572.
Wang, B., Li, Y., Xie, Z., Du, S., Zeng, X., Hou, J., Ma, T., 2020. Characteristics of microbial activity in atmospheric aerosols and its relationship to chemical composition of PM2. 5 in Xi'an, China. J. Aerosol Sci. 146, 105572.
Wen, G., Liang, Z., Xu, X., Cao, R., Wan, Q., Ji, G., Lin, W., Wang, J., Yang, J., Huang, T., 2020. Inactivation of fungal spores in water using ozone: kinetics, influencing factors and mechanisms. Water Res. 185, 116218.
Weryszko-Chmielewska, E., Kasprzyk, I., Nowak, M., Sulborska, A., Kaczmarek, J., Szymanska, A., Haratym, W., Gilski, M., Jedryczka, M., 2018. Health hazards related to conidia of Cladosporium—biological air pollutants in Poland, central Europe. Journal of environmental sciences 65, 271–281.
Whitby, C., Ferguson, R.M., Colbeck, I., Dumbrell, A.J., Nasir, Z.A., Marczylo, E., Kinnersley, R., Douglas, P., Drew, G., Bhui, K., 2022. Compendium of analytical methods for sampling, characterization and quantification of bioaerosols. Adv. Ecol. Res. 101–229. Elsevier.
Xu, C., Wei, M., Chen, J., Wang, X., Zhu, C., Li, J., Zheng, L., Sui, G., Li, W., Wang, W., 2017. Bacterial characterization in ambient submicron particles during severe haze episodes at Ji'nan, China. Sci. Total Environ. 580, 188–196.
Yan, Y., Ren, P., Meng, Q., 2024. Quantitative evaluation of the synergistic effects of multiple meteorological parameters on air pollutants based on generalized additive models. Urban Clim. 55, 101965.
Yang, Y., Yang, L., Hu, X., Shen, Z., 2024. Characteristics of bioaerosols under high-ozone periods, haze episodes, dust storms, and normal days in Xi'an, China. Particuology 90, 140–148.
Zhao, Y., Hou, Y., Wang, F., 2024. Ecological risk and pollution assessment of heavy metals in farmland soil profile with consideration of atmosphere deposition in central China. Toxics 12, 45.