Recommended terminology for aerobiological studies: automatic and real-time monitoring methods

Aerobiologia
https://doi.org/10.1007/s10453-025-09879-5

BRIEF REPORT

Recommended terminology for aerobiological studies: automatic and real-time monitoring methods

Fiona Tummon · Katharina Bastl · Maximilian Bastl · Paul J. Beggs · Jordina Belmonte · Nicolas Bruffaerts · Athanasios Charalampopoulos · Bernard Clot · Antonella Cristofori · Benoît Crouzy · Athanasios Damialis · Sophie Erb · Carmen Galan · Elias Graf · Jörg Haus · Kilian Koch · Gian Lieberherr · Maria Lbadaoui-Darvas · Marie-Pierre Meurville · Erny Niederberger · Jose Oteros · Rosa Pérez-Badia · Helena Ribeiro · Andreas Schwendimann · Branko Sikoparija · Carsten Skjoth · Matt Smith · Mikhail Sofiev · Olga Sozinova · Laura Sukiene · Konstantina Vasilatou · Willem W. Verstraeten · Yanick Zeder · Franziska Zemmer

Received: 28 April 2025 / Accepted: 12 August 2025

Keywords Automatic monitoring · Pollen · Fungal spores · Terminology

1 Introduction

Since Galán et al. (2017) published important definitions for terms to be used in aerobiological studies, techniques have evolved regarding automatic and real-time monitoring of pollen and fungal spores, with a range of different technologies having been established over the past decade. These developments have been accompanied by the emergence of a new set of specific and common terminologies related not only to the instrumental hardware but also particularly to the machine learning algorithms that are used to classify particles. The aim of this article is to update the recommended aerobiological terminology, including definitions related to automatic and

F. Tummon · B. Clot · B. Crouzy · S. Erb · G. Lieberherr · M. Lbadaoui-Darvas · M.-P. Meurville
Federal Office of Meteorology and Climatology MeteoSwiss, Payerne, Switzerland
e-mail: Fiona.tummon@meteoswiss.ch

K. Bastl · M. Bastl
Department of Otorhinolaryngology, Medical University of Vienna, Vienna, Austria

P. J. Beggs
School of Natural Sciences, Faculty of Science and Engineering, Macquarie University, Sydney, Australia

J. Belmonte
Department of Animal Biology, Plant Biology and Ecology, Universitat Autònoma de Barcelona, Barcelona, Spain

J. Belmonte
Institute of Environmental Sciences and Technology, ICTA-UAB, Universitat Autònoma de Barcelona, Barcelona, Spain

N. Bruffaerts
Sciensano, Brussels, Belgium

A. Charalampopoulos · A. Damialis
Department of Ecology, School of Biology, Aristotle University of Thessaloniki, Thessaloniki, Greece

A. Cristofori (✉)
Research and Innovation Centre, Fondazione Edmund Mach, San Michele all'Adige, Italy
e-mail: antonella.cristofori@fmach.it

A. Cristofori
NBFC, National Biodiversity Future Center, Palermo, Italy

C. Galan · J. Oteros
Department of Botany, Ecology and Plant Physiology, University of Córdoba, Córdoba, Spain

E. Graf · K. Koch · E. Niederberger · A. Schwendimann · Y. Zeder
Swisens AG, Emmen, Switzerland

Published online: 16 October 2025

real-time monitoring and to new applications using real-time data, such as numerical forecast models. This is particularly important for the standardisation of procedures, the homogenisation of time series, using data from both manual and automatic methods, as well as simply for ensuring the community has a common language to work with.

2 Terms

Airflow cytometry: technology that enables rapid analysis and quantification of the physical and chemical characteristics of single airborne particles as they pass through a detection chamber and interact with a beam of light and/or other detectors. For ***bioaerosol***, typical measurement techniques include, but are not limited to, ***elastic light scattering***, ***induced fluorescence***, ***polarisation***, and ***digital holography***.

Automatic bioaerosol measurement: unsupervised measurement that is taken by an instrument with no human intervention required to sample, identify, and quantify ***bioaerosol*** concentrations.

Bioaerosol (or primary biological aerosol particles (PBAP)): solid airborne particles derived from biological organisms, including microorganisms such as bacteria, viruses, or algae, as well as pollen, fungal spores, and fragments of biological material such as cell fragments, plant debris, fungal hyphae, plant trichomes, and animal dander (Després et al., 2012).

Calibration: a two-step operation that is carried out under specific conditions. In the first step, the relation between an instrument under test and a measurement standard, including estimations of measurement uncertainty, is established. In the second step, this information is used to post-process the measured data to align with the standard. A calibration may be expressed by a statement, calibration function, calibration diagram, calibration curve, or calibration table (adapted from BIPM, 2025). For automatic bioaerosol monitors, a range of different methods exist for calibration purposes (see Tummon et al., 2024).

Classification algorithm (or classifier; to be used instead of identification algorithm, model, or recognition algorithm): a procedure that is applied to raw data, measured by an instrument, to classify particles. Classification algorithms may or may not use machine learning principles. Typically, a classifier provides a class label and an associated confidence level (e.g. a confidence of 95% that a given particle is a birch pollen grain).

Classification confidence threshold: a percentage value that defines the minimum classification confidence level required for an ***event*** classification to be counted towards the ***particle number concentration*** of a specific class. This allows the selection of

J. Haus
Helmut Hund GmbH, Wetzlar, Germany

R. Pérez-Badia
Institute of Environmental Sciences, University of Castilla-La Mancha, Toledo, Spain

H. Ribeiro
Department of Geosciences, Environment and Spatial Planning, University of Porto, Porto, Portugal

H. Ribeiro
Earth Sciences Institute, Pole of the Faculty of Sciences, University of Porto, Porto, Portugal

B. Sikoparija
BioSense Institute - Research Institute for Information Technologies in Biosystems, University of Novi Sad, Novi Sad, Serbia

C. Skjoth
Department of Environmental Science, Aarhus University, Roskilde, Denmark

M. Smith
School of Science and the Environment, University of Worcester, Worcester, UK

M. Sofiev
Finnish Meteorological Institute, Helsinki, Finland

O. Sozinova
Faculty of Science and Technology, University of Latvia, Riga, Latvia

L. Sukiene
Siauliai Academy, Vilnius University, Siauliai, Lithuania

K. Vasilatou
Department of Chemistry and Biology, Federal Institute of Metrology METAS, Bern, Switzerland

W. W. Verstraeten
Royal Meteorological Institute of Belgium, Ukkel, Belgium

F. Zemmer
Independent Researcher, San Michele all'Adige, Italy

only those ***events*** with a high probability of accurate classification.

Cleaned dataset: a digital dataset that has been manually or automatically processed to remove unwanted ***events*** from a ***training dataset*** or a ***test dataset***. This process may include handling missing data, removing outliers, and filtering environmental noise or agglomerates from a dataset of individual particle ***events***.

Continuous sampling: measurements taken uninterruptedly, with a known volume of air being sucked through the instrument continuously.

Counting efficiency: ratio of the ***particle number concentration*** determined by an instrument under test to the ***particle number concentration*** determined by a reference instrument for the same test aerosol. A counting efficiency of one indicates the instrument detects all target particles measured by a reference instrument, while a low counting efficiency indicates undercounting due to particle losses and other possible sensor/instrument limitations.

Data assimilation: techniques that combine observational data and a numerical model to provide a more accurate estimate of a system's state than what is obtained using just the measurements or the model alone. In recent years, data assimilation has been extended to cover various other aspects of model improvement: more accurate external forcing parameters (e.g. emission of atmospheric tracers) or refinement of internal model parameters. Data assimilation adjusts the internal or external variables of the corresponding model, so that the model results better align with observational data.

Data augmentation: a method to create new data from pre-existing data to increase the size and diversity of a dataset. Data augmentation can improve the performance and robustness of machine learning algorithms.

Data flag: a value assigned to each data point that qualifies the measurement, i.e. indicates whether the measurement is valid, uncertain, invalid, missing, etc. The flags can also identify the reason for missing/abnormal data.

Data labelling: the process by which human experts assign a class (e.g. a genus determination such as *Corylus*) to an ***event***. In aerobiology, this is typically carried out on the images obtained with ***digital microscopy***.

Data fusion: the process of combining data from multiple sources to create a more accurate representation of the state of a system. This could involve merging data from different instruments, models, or combinations thereof. The goal is to enhance the numerical model's predictions or the accuracy of the measurements by integrating various information to reduce uncertainty and improve overall accuracy.

Data levels: description of the degree of processing applied to data obtained from monitoring instruments. The data directly obtained from the instrument (raw data) are referred to as ***level-0.*** For bioaerosols, the ***level-1*** data refer to the concentration values of classified particles.

Digital holography: a means of creating images of particles. A coherent light beam interacts with the particle, causing interference patterns on the camera sensor. An algorithm is used to reconstruct an image of the particle using this interference pattern, providing information about the particle morphology.

Digital microscopy: a variation of optical microscopy using the same optics but in combination with a digital camera that acts as a detector.

Effective detection flow rate: volumetric flow rate through the sensing zone, which might be optically and/or aerodynamically limited. ***Calibration*** of the effective detection flow rate is necessary for the evaluation of the ***counting efficiency*** (adapted from ISO 21501-1). Note: Whenever the effective flow rate cannot be directly measured, the ***sampling flow rate*** should be used for the calculation of the ***counting efficiency***.

Elastic light scattering: the process by which light is forced to deviate from a straight trajectory by particles present in the air. In the case of elastic scattering, the scattered light has the same wavelength/frequency as the incident light.

Event: the measurement of a signal (e.g. light scattering, microscopic image, holographic image, fluorescence spectrum) associated with the detection of a single particle.

False negative: a particle wrongly identified as not being a member of a particular class (e.g. an *Alnus* pollen grain that is not identified as *Alnus*).

False positive: a particle wrongly identified as being a member of a particular class (e.g. an *Alternaria* spore misclassified as a *Betula* pollen grain or a water droplet misclassified as a Poaceae pollen grain).

Fluorescence intensity: the intensity of light emitted by a particle upon excitation at a defined wavelength, measured at a specific wavelength or across a waveband, typically in arbitrary units.

Fluorescence lifetime: the average time a fluorescent molecule remains in its excited state after absorbing light before returning to the ground state by emitting fluorescence. It is typically measured in nanoseconds. The fluorescence lifetime depends on the molecule, its immediate environment, and the excitation and emission wavelengths.

Fluorescence spectrum: the emission characteristics of a molecule that fluoresces (fluorophore). Typically, this is provided as a graphical representation of the emitted fluorescence light intensity as a function of wavelength.

F-score: the F-score provides an overall value regarding the performance of a classification algorithm. It takes into account both the precision and recall and thus serves to optimise both values.

Hyperparameter: a configurable variable that is set prior to training a ***classification algorithm***. It governs the learning process and algorithm structure (e.g. the learning rate or the number of layers in a neural network). In contrast, parameters are internal algorithm values (such as weights or biases) that are learned from the training data during the optimisation process.

Image stack: a number of microscopic images taken along the viewing direction of a microscope, differing only in their focusing plane. A dedicated stacking software is required to extract all sharp regions from all images of the stack, resulting in a high-resolution image with an extended depth of field.

Impaction (or aerosol impaction): the process by which particles of a given aerodynamic size are separated by inertia from an air stream, and collide and adhere to an impaction surface.

Induced fluorescence: the natural emission of light resulting from the relaxation of excited electron states after the absorption of light by certain molecules. Such molecules can be found inside or on the surface of airborne particles.

Inference: the application of a ***classification algorithm*** to measurement data which generates a prediction of a class for each particle.

Intermittent sampling: measurements that are taken at intervals, i.e. non-continuously. For example, sampling one minute every five minutes.

Level-0 data (or raw data): data and metadata as provided by the instrument at its ***native time resolution*** with no transformation. Data are in physical units (e.g. length, electric current) either directly provided by the instrument or converted from basic units (e.g. mV) to physical units.

Level-1 data (or concentration data): data to which the ***classification algorithm*** and any ***post-processing*** have been applied to obtain concentration values and that is averaged to a higher temporal resolution (e.g. hourly) than the instrument's ***native time resolution***. Typically, this level of data has been quality checked; thus, ***data flags*** are applied.

Measurement efficiency: the ratio of the ***particle number concentration*** per class determined by the monitoring instrument to that measured by a reference instrument. This includes both the ***counting efficiency*** and the recognition rate of the ***classification algorithm***.

Measurement system: the hardware of a monitoring instrument together with any software, including ***classification algorithms*** and ***pre-*** or ***post-processing*** steps required to provide values of airborne ***bioaerosol*** concentrations.

Measurement uncertainty: a non-negative range of values that characterises the dispersion associated with a measurement result. It indicates how well a measurement result represents the real value of what is being measured.

Native time resolution: the temporal resolution inherent to an instrument without any averaging applied.

Near-real time (NRT): data that are automatically processed with a slight delay, often due to the measurement principle or network transmission. These data have not passed through any manual quality control or review. In aerobiology, the generally accepted delay is between 3 and 12 h. Note that there are large differences in this definition between disciplines. For example, in computing, near-real time refers to delays ranging from seconds to a few minutes.

Offline numerical model: numerical model that simulates the emission, transport, and deposition of atmospheric constituents, including ***bioaerosol***, using as input meteorological data from another model that has been run separately. This is computationally

cheaper than ***online models***, making it possible to run many simulations (e.g. testing pollen emission scenarios) or for emergency responses. However, numerical errors introduced by offline approaches increase with grid resolution (Baklanov et al., 2017) and consequently may be of importance for high-resolution models.

Online numerical model: numerical model that simulates the emission, transport, and deposition of atmospheric constituents, including ***bioaerosol***, at the same time as meteorological processes as part of a single unified numerical system. This provides a more consistent description of the interactions between processes such as advection, cloud microphysics, and radiative transfer (e.g. Baklanov et al., 2017). Such models are considerably more complex than ***offline models*** and require higher computational resources to run.

Particle number concentration: number of ***bioaerosol*** particles per unit volume of air (e.g. pollen grains or fungal spores per m^3 air). Depending on the standard (varies between countries), the concentration can be provided for ambient atmospheric pressure during the observation or normalised to a standard sea-level pressure.

Particle size resolution: the ability of an instrument to differentiate between particles of different sizes. High size resolution means the instrument can differentiate between particles of similar sizes.

Post-processing: an analytical procedure applied to data after the application of a ***classification algorithm*** (e.g. filtering data using the confidence of a predicted taxon) or after a numerical forecast model has been run.

Precision: the proportion of classifications that are true (defined as the number of ***true positives*** divided by the sum of the number of ***true positives*** and ***false positives***). It indicates how often a ***classification algorithm*** predicts the correct class.

Pre-processing: an analytical procedure applied to data prior to the application of a ***classification algorithm*** or before a numerical forecast model is run. This may include removing unwanted data (e.g. filtering out particles using physical characteristics such as shape or size, or flagged ***events***), handling missing values, or ***data augmentation***.

Real time: automatically processed and delivered data with no perceptible delay. These data have not passed through any manual inspection or review. In aerobiology, real-time data are provided within 3 h of the measurement (adapted from Laj et al., 2024). Note that there are large differences in this definition between disciplines. For example, in computing, real time refers to delays measured in milliseconds.

Recall: the proportion of classifications that are correctly made (defined as the number of ***true positives*** divided by the sum of the number of ***true positives*** and ***false negatives***). It indicates how well the ***classification algorithm*** identifies all objects of a particular class.

Relative fluorescence intensity: a normalised representation of ***fluorescence intensities*** to ensure comparability between different instruments. For automatic ***bioaerosol*** instruments, it is typically calculated by dividing the fluorescence intensity of each emission wavelength by the sum of all emission intensities for a particular excitation wavelength. The result is a dimensionless value ranging from 0 to 1 for each emission wavelength.

Re-classification: the application of a new ***classification algorithm*** to raw data to obtain a new concentration time series. This is typically carried out when a ***classification algorithm*** has been improved and is important for ensuring homogeneous time series.

Sample: a set of particles extracted from ambient air. For impactor-based instruments with subsequent optical microscopy, this is a physical collection on some kind of impaction surface over a defined time interval. For flow cytometry-based instruments, the sample is a number of particles registered in the detection zone over a defined time interval.

Sampling flow rate: the volume of air per unit of time that a monitoring instrument samples through the inlet (typically per minute). Note that the sampling flow rate can be larger than the ***effective detection flow rate***.

Scaling factor (or calibration factor): classifier- and/or instrument-specific ratio between the reference measurement value and the instrument under test, both converted to ***particle number concentration***.

Scan: for instruments based on ***digital microscopy***, a process applied to ensure the detection of all particles in a ***sample*** when the detector's field of view is smaller than the total sampling area. At high particle concentrations, the full sampling area may not be scanned.

Segmentation algorithm: for instruments based on ***digital microscopy***, an algorithm is required to isolate

all target particles before they can be fed into the ***classification algorithm***. This algorithm must be trained to find particles of a specific shape or size.

Test dataset: an independent dataset (i.e. unseen during the training) used to provide an unbiased evaluation of a ***classification algorithm***'s performance. The primary purpose of this dataset is to allow a fair assessment of how the ***classification algorithm*** is likely to perform when it encounters new data in a live operational environment.

Time series homogenisation: methods that are used to combine datasets from various monitoring instruments to create a new, combined long-term dataset. Typically, the old dataset is adapted to the new one. A period of parallel measurements is required to ensure the homogenisation method can be correctly applied.

Traceability: property of a measurement result whereby the result can be related to a reference through a documented unbroken chain of ***calibrations***, each contributing to the measurement uncertainty. Typically, this includes a national or regional reference instrument which in turn is compared to an international standard. There may be numerous steps in the calibration chain, and a high degree of confidence and accuracy is required along each of these steps.

Training dataset: a set of data used to train a ***classification algorithm***, typically containing data only from a known particle type.

True negative: a particle correctly identified as not being of a class of interest (e.g. a dust particle identified as not being pollen).

True positive: a particle correctly identified as being of a particular class (e.g. an *Olea* pollen grain identified as *Olea*).

Validation dataset: a digital dataset used to provide an evaluation of a ***classification algorithm's*** fit to the training data. It is likewise used to adjust or optimise the algorithm's ***hyperparameters*** before it is finally evaluated with the ***test dataset***.

3 Conclusion

This effort brings together a large group of aerobiologists to define terminology to be used for automatic pollen and fungal spore monitoring. It sets out a number of definitions that are aimed at aiding the community to better understand and communicate around the theme of automatic real-time measurements as well as their uses. As the field evolves, the terminology will also change in accordance and need to be updated. This will particularly be the case when a European standard related to automatic pollen and fungal spore monitoring is released.

Acknowledgements The authors acknowledge the work of the EUMETNET AutoPollen Programme, to which this manuscript contributes. This paper is also a contribution to the work of the Horizon Europe SYLVA project.

Author contribution FT wrote the main manuscript. All authors contributed to and reviewed the text.

Funding Open access funding provided by Fondazione Edmund Mach - Istituto Agrario di San Michele all'Adige within the CRUI-CARE Agreement.

Data availability No datasets were generated or analysed during the current study.

Declarations

Competing interests The authors have no competing interests to declare that are relevant to the content of this article; however, Carmen Galan and Matt Smith are editors-in-chief of Aerobiologia, Athanasios Damialis, Helena Ribeiro, Branko Sikoparija, and Fiona Tummon are associate editors of Aerobiologia, and Mikhail Sofiev and Olga Sozinova are on the Aerobiologia editorial board. The peer review of this article was independently handled by another member of the journal's editorial board.

References

Baklanov, A., Brunner, D., Carmichael, G., Flemming, J., Freitas, S., Gauss, M., Hov, Ø., Mathur, R., Schlünzen, K. H., Seigneur, C., & Vogel, B. (2017). Key issues for seamless integrated chemistry-meteorology modeling. *Bulletin*

Of The American Meteorological Society, 98(11), 2285–2292. https://doi.org/10.1175/BAMS-D-15-00166.1

BIPM (Bureau International des Poids et Mesures). https://jcgm.bipm.org/vim/en/2.39.html (accessed 25 March 2025).

Després, V. R., Huffman, J. A., Burrows, S. M., Hoose, C., Safatov, A. S., Buryak, G. A., Fröhlich-Nowoisky, J., Elbert, W., Andreae, M. O., Pöschl, U., & Jaenicke, R. (2012). Primary biological aerosol particles in the atmosphere: A review. *Tellus Series B-Chemistry Physics and Meteorology, 64*(1), Article 15598. https://doi.org/10.3402/tellusb.v64i0.15598

Galán, C., Ariatti, A., Bonini, M., Clot, B., Crouzy, B., Dahl, A., Fernandez-González, D., Frenguelli, G., Gehrig, R., Isard, S., Levetin, E., Li, D. W., Mandrioli, P., Rogers, C. A., Thibaudon, M., Sauliene, I., Skjoth, C., Smith, M., & Sofiev, M. (2017). Recommended terminology for aerobiological studies. *Aerobiologia, 33*(3), 293–295. https://doi.org/10.1007/s10453-017-9496-0

ISO 21501-1. (2009). Determination of particle size distribution: Single particle light interaction methods. Part 1: Light scattering aerosol spectrometer.

Laj, P., Myhre, C. L., Riffault, V., Amiridis, V., Fuchs, H., Eleftheriadis, K., et al. (2024). Aerosol, clouds and trace gases research infrastructure (ACTRIS): The European research infrastructure supporting atmospheric science. *Bulletin of the American Meteorological Society, 105*(7), E1098–E1136. https://doi.org/10.1175/BAMS-D-23-0064.1

Tummon, F., Bruffaerts, B., Celenk, S., Choël, M., Clot, B., Galán, C., Gilge, S., Mokin, V., O'Connor, D., Rodinkova, V., Sauliene, I., Sikoparija, B., Sofiev, M., Sozinova, O., Tesendic, D., & Vasilatou, V. (2024). Towards standardisation of automatic pollen and fungal spore monitoring: Best practices and guidelines. *Aerobiologia*. https://doi.org/10.1007/s10453-022-09755-6