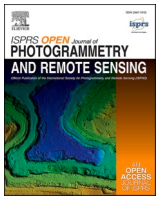
ISPRS Open Journal of Photogrammetry and Remote Sensing 12 (2024) 100063

Contents lists available at ScienceDirect

ISPRS Open Journal of Photogrammetry and Remote Sensing

journal homepage: www.journals.elsevier.com/isprs-open-journal-of-photogrammetry-and-remote-sensing



Natural color dispersion of corbicular pollen limits color-based classification

Parzival Borlinghaus [a,*], Frederic Tausch [b], Richard Odemer [c]

[a] Karlsruhe Institute of Technology (KIT), Institute for Operations Research, Karlsruhe, Germany
[b] apic.ai GmbH, Karlsruhe, Germany
[c] Julius Kühn-Institut (JKI) - Federal Research Centre for Cultivated Plants, Institute for Bee Protection, Braunschweig, Germany

ARTICLE INFO

Keywords:
Bee botany
Chromatic assessment
Pollen diversity
Pollination
Precision beekeeping
Pollenyzer

ABSTRACT

Various methods have been developed to assign pollen to its botanical origin. They range from technically complex approaches to the less precise but sophisticated chromatic assessment, in which the pollen colors are used for identification. However, a common challenge lies in the similarity of colors of pollen from different plant species. The advent of camera-based bee monitoring systems has sparked renewed interest in classifying pollen based on color and offers potential advances for honey bee biomonitoring. Despite the promise of improved sensor accuracy, a critical examination of whether color diversity within a single species may be the primary limiting factor has been lacking. Our comprehensive analysis, which includes over 85,000 corbicular pollen from 30 major pollen species, shows that the average color variation within each species is distinguishable to a human observer, similar to the difference between two dissimilar colors. From today's perspective, the considerable color variation within a single pollen source makes the use of color alone to classify pollen impractical. When picking a single pollen color from the entire dataset, we report a correct pollen type classification rate of 67 %. The accuracy was highly dependent on the type and ranged from 0 % for rare types with common colors to 99 % for distinct colors. The large color dispersion within species highlights the need for complementary methods to improve the accuracy and reliability of color-based pollen identification in biomonitoring applications.

1. Introduction

Pollen collection by honey bees (*Apis mellifera* L.) is of multiple ecological and agricultural importance (Ollerton et al., 2011). At its core, pollination is a vital ecosystem service with immense economic value that supports agricultural productivity and biodiversity conservation efforts worldwide. The importance of this service is particularly pronounced given agriculture's increasing dependence on pollinators, a trend underscored by the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (2016). Recent findings by Casas Restrepo et al. (2023) have highlighted the potential of pollen-based regional differentiation and offer new insights into the geographical classification of apicultural products. Consequently, the analysis of pollen is emerging not only as a means to delineate honey varieties, but also as a promising way to monitor biodiversity dynamics in different ecosystems (Milla et al., 2022), facilitating informed decision-making in land conservation and management.

In the field of pollen analysis, a variety of methods have been developed to recognize plant taxa within samples, each offering unique advantages and insights (Campos et al., 2021; Layek et al., 2022). DNA metabarcoding, for example, has revolutionized our understanding of floral diversity by revealing a broader range of plant taxa in the vicinity of beehives compared to traditional manual surveys (Milla et al., 2022). Similarly, studies utilizing amino acid composition analysis have successfully delineated pollen samples based on temporal and spatial variation, providing valuable temporal insights into ecosystem dynamics (Ares et al., 2022). In addition, the integration of spectroscopic techniques, including Fourier transform near-, mid- and Raman spectroscopy (ATR-FTIR), together with color analysis and microscopy has enabled precise discrimination of pollen origins (Bleha et al., 2021; Swiatly-Blaszkiewicz et al., 2021), demonstrating the versatility of modern analytical approaches in elucidating botanical origins.

Despite the methodological advancements, the practical application of pollen analysis often defaults to the technically less complex yet gold standard method of light microscopy (Campos et al., 2021; Louveaux et al., 1978; Von Der Ohe et al., 2004). This technique relies on the

* Corresponding author.
E-mail address: parzival.borlinghaus@kit.edu (P. Borlinghaus).

https://doi.org/10.1016/j.ophoto.2024.100063
Received 22 October 2023; Received in revised form 19 March 2024; Accepted 20 March 2024
Available online 16 April 2024

morphological examination of individual pollen grains, which are subsequently classified into pollen types utilizing established pollen atlases (Fægri et al., 2007). However, the limitations inherent in this method become evident when attempting to identify plant taxa at the species level due to morphological similarities among pollen grains. Nonetheless, the designation of pollen types provides valuable insights into taxonomic diversity, with naming conventions reflecting varying degrees of taxonomic resolution (De Klerk and Joosten, 2007). It is important to understand in the context of this work that one type of pollen can be produced by several species. Notably, the concept of floral constancy, wherein honey bees tend to collect pollen pellets of the same botanical origin during a single foraging flight (Knoll, 1956; Newstrom-Lloyd et al., 2017; Percival, 1947), has facilitated chromatic assessments of pollen pellets, offering a less labor-intensive yet economically viable means of assessing floral diversity (Brodschneider et al., 2021; Conti et al., 2016).

In previous studies, we have already utilized machine learning and computer vision to advance entomological research (Borlinghaus et al., 2022b, 2023a; 2023b). We have shown that results from tedious manual chromatic assessment can be reproduced equally well in a fraction of time using the Pollenyzer app, a photograph of a pollen sample and everyday objects for color calibration (Borlinghaus et al., 2023a). Although accurate color-based botanical classification remains problematic (Kirk, 2006), existing pollen color books provide valuable guidelines for classification that are supplemented by spatial and seasonal considerations (Hodges, 1952; Kirk, 2006). However, the inherent variability of pollen colors poses a significant challenge for standardized classification systems, especially the natural dispersion of pollen colors and its impact on classification accuracy have not yet been sufficiently investigated. In particular, the quantification of natural color dispersion requires the compilation of a data set containing large amounts of pollen of the same origin, as the limited data from hand-picked corbicular pollen collections prove to be insufficient (Hodges, 1952; Kirk, 2006).

To overcome this challenge, we combined images of pollen trap contents with standard palynological laboratory analyses. Our approach requires that the pooled results from the laboratory can be linked to the individual pollen colors derived from the image data, i.e. that the gap between the type information at the sample level and the color information at the pollen basket level is bridged. Therefore, we used Gaussian mixture models (GMMs) to assign the most likely pollen type from the laboratory results to the individual pollen baskets. After cleaning the data, we were able to assign 62,343 calibrated pollen colors from 86 samples to one of 30 pollen types, including 14 of the 31 major European types (Keller et al., 2005). Subsequently, 253 chromatic means and covariance matrices describing the color dispersion of the occurrence of each pollen type were determined from the data. Finally, the mean distance to the center of the distribution was calculated, averaged for each pollen type and interpreted to quantify the type-specific dispersion.

2. Material and methods

Fig. 1 illustrates the methodological approach of this study in 6 steps. In step (1), pollen pellets were collected using pollen traps, the contents of which we refer to as pollen samples. Each sample was digitized (2) before the pollen colors were calibrated and extracted from the image in step (4) using the Pollenyzer app (Borlinghaus et al., 2023a). In addition, the composition of the sample was palynologically analyzed in the laboratory (3). Both pieces of information were combined in a Gaussian mixture model (5) to determine the species for each pollen basket. Finally, in step (6), the type-specific color dispersion was determined by interpreting the covariance matrices estimated by the GMMs.

2.1. Data collection

In this study, four different data sets were provided with a total of 107 pollen samples, 86 of which were suitable for further analysis and are listed in Table 1. Each sample originated from a single full-sized honey bee colony (*A. mellifera*) and was collected over a 24-h period with an entrance pollen trap (Campos et al., 2021). For logistical reasons, the pollen collection for data set D was extended to a duration of 48 h.

For datasets A, B, and C, the pollen pellets were imaged in a specially designed dome light setup (Tausch et al., 2023). This setup was tailored for this particular application. To ensure accurate color measurements, the camera used was calibrated with a Calibr8 color chart, resulting in L*a*b* color values for the subsequent analyses. The L*a*b* color space expresses color in three values: L* for perceived brightness and a* and b* for the unique colors of human vision: red, green, blue, and yellow.

Dataset D was scanned with a commercial flatbed scanner and calibrated with the same Calibr8 color calibration card. Because scanning was done after palynological analysis, some pollen were missing from the images compared to their microscopic equivalents. Subsequently, the Pollenyzer software was used to detect the pollen on all images, extract the pollen colors and to obtain calibrated pollen colors in L*a*b* color space (Borlinghaus et al., 2023a).

Table 1
Dataset Overview. Samples were collected in Germany in 2021 and 2022.

name	N samples	Collection	Location	N colonies	Image acquisition
A	27	24h, Apr–Sep 2021	Wolfsburg, DE	4	Light dome
B	9	24h, May–Aug 2021	Mayen, DE	4	Light dome
C	40	48h, unknown	3 federal states, DE	10	Light dome
D	31	24h, Apr–Sep 2022	Wolfsburg, DE	2	Flatbed scanner

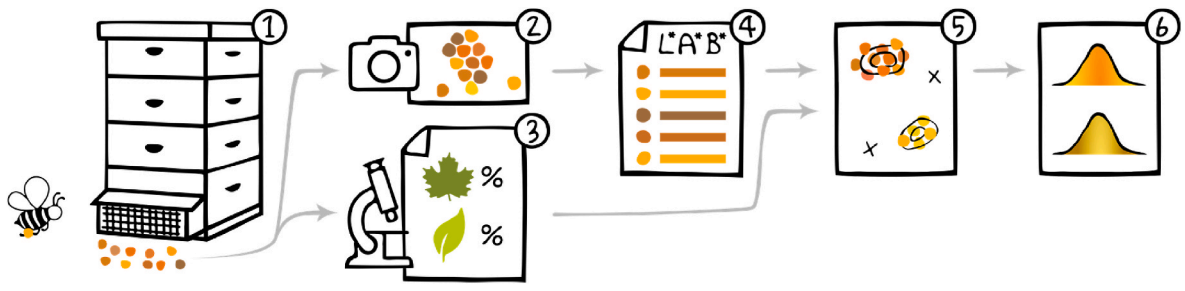


Fig. 1. Over periods of 24–48h, pollen samples were collected in pollen traps (1). The pollen samples were then digitized (2) and the pollen types identified (3). Using the Pollenyzer software, the calibrated pollen colors were extracted from the photos (4). Based on prior information (3) and pollen colors (4) a GMM was fitted (5) and the distributions' parameters reported (6).

2.2. Palynological analyses

All samples were palynologically analyzed by light microscopy and expert judgment at the Expert Center for Bees and Beekeeping, Mayen, Germany (for detailed methodological protocol, see data repository in Borlinghaus et al. (2024)).

In brief, each sample was subjected to morphological analysis, which included routine homogenization of the corbicular pollen. Approximately 100 individual pollen grains were then extracted from the mixture and subjected to expert classification to determine the relative pollen type abundance. It is important to point out that although experts can occasionally identify taxa to species level, this is generally not feasible. In such cases, only the pollen type representing a group of often closely related plant taxa is determined. In cases where pollen types are named after a species (e.g., *Phacelia*), current knowledge indicates that the pollen type (identified by its morphology) is produced exclusively by a single plant. A less precise identification is highlighted by an abbreviated type suffix (e.g. *Taraxacum* T.), indicating that some, but not all, taxa of the given rank share the same morphology. When pollen grains from a whole family cannot be distinguished morphologically, the type is designated by the family name (De Klerk and Joosten, 2007).

In total, the pollen colors from the Pollenyzer app and the proportion per pollen type determined in the laboratory were available for each of the 86 samples. Henceforth, the individual samples were labeled with data set and sample designation, e.g. sample D03.

2.3. Color dispersion model

A Gaussian Mixture Model was applied to find the missing link between pollen colors of individual pollen baskets and pollen type probabilities from laboratory analyses (step 5, Fig. 1).

2.3.1. Gaussian Mixture Model

When data consists of normally distributed components, a Gaussian Mixture Model (GMM) can be constructed by combining several (multivariate) Gaussian distributions, each of which has a certain weight related to the component's proportion of the overall data. The parameters of a GMM, the weights, mean vectors and covariance matrices, can be fitted by the Expectation-Maximization (EM) algorithm (Dempster et al., 1977). The EM algorithm ensures incremental improvements but no global optimum. The iterative procedure is stopped when convergence is reached. Similar to the well-known k-means algorithm (Lloyd, 1982), the EM algorithm requires an initial parametrization for the weights, means and covariances. When fitting a GMM, it is assumed that the underlying data is a mixture of well-separated Gaussian distributions and that the number of components is known (Reynolds, 2009).

2.3.2. Visualization of pollen colors in two dimensions

Pollen colors were handled as trichromatic L*a*b* color values. Only for visualization, the color dimensions were reduced to two using principal component analysis (PCA). The PCA was fitted once on all available 85,531 pollen colors. To make the visualizations easy to compare, the same transformations (and axis limits) were used for all dimensionally reduced L*a*b* pollen colors, both here and in the data repository (Borlinghaus et al., 2024).

2.3.3. Cluster annotation

To make the fitting of the GMM more robust, the initial chromatic mean values were provided by annotation. Fig. 2 exemplifies how the cluster mean values were annotated with the corresponding pollen type. On the left, all colors extracted from sample D17 are shown as a two-dimensional scatter plot. Contour lines provide information about the distribution's density and reveal three distinct clusters. In the middle, the result of the laboratory analysis is shown, which lists three pollen types that account for more than 5% of the total. Through prior knowledge, literature and logical combination, the individual clusters were assigned to the three major pollen types: (1) *Phacelia* is known for its distinct dark purple color (Kirk, 2006). (2) According the laboratory analysis, pollen type *Rosaceae* is much more common than *Centaurea cyanus* and must therefore produce gray pollen. (3) The remaining pollen type, *C. cyanus*, was identified to species level under the microscope and could therefore be verified as light brown pollen using a pollen atlas. Alternatively, it could simply be assumed to match the remaining cluster. The annotation was simplified by the fact that often only a few, large clusters were found. The laboratory results showed that on average 2.15 pollen types made up an abundance of more than 10% per sample, which is corroborated by literature (Ares et al., 2022; Bilisk Tosunoglu et al., 2008).

2.3.4. Model assumptions

It is plausible to assume that the pollen colors of the same botanical origin follow a normal distribution around a chromatic mean. Since different pollen types may be present, it is reasonable to assume that the data within a sample is distributed like a mixture of three-dimensional Gaussian distributions.

The number of components were known as they were taken from the palynological analysis. Pollen types with a share of less than 5 % were not considered and added to the share of outliers. Further adjustments to the number of components were made based on visual inspection considering the ease of cluster separation. Samples had to be excluded if the modes of the color distribution could not be easily assigned to pollen types, i.e. the underlying mixtures were not well-separated. For the remaining samples, the GMM assumption of separability was deemed satisfied.

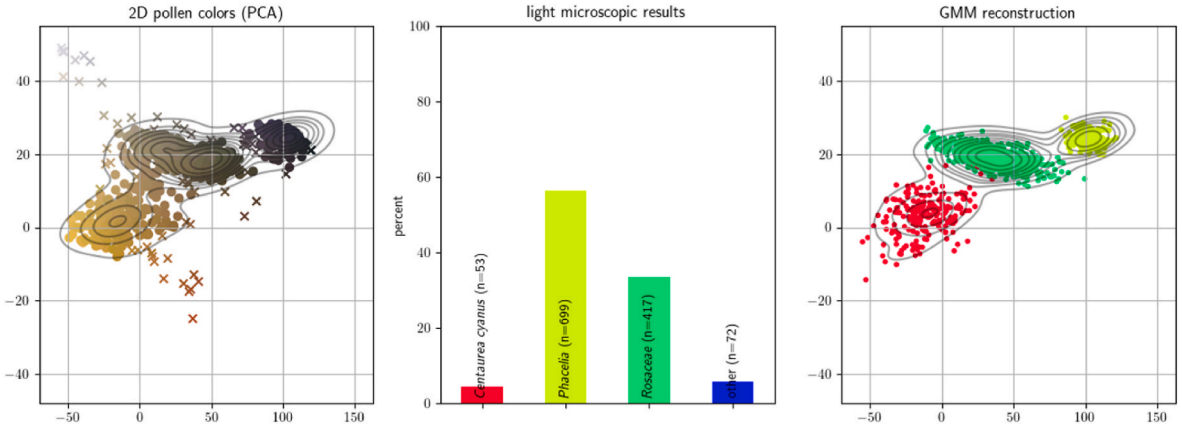


Fig. 2. Exemplary pollen color scatter plot of sample D17 (left), corresponding palynological assessment (middle) and fitted GMM model (right). Outliers are plotted as ×.

Considering that a single sample contains only very few abundant species, it is unlikely that two of them have the same pollen type. It is therefore plausible to assume that in a single sample each pollen type comes from a single floral origin. If this assumption is violated, two cases can occur. Either both species produce the same colors or different ones. The first case can be ignored as it does not change the interpretation of the results and in the second case the annotation of the pollen types is ambiguous and the sample can be discarded.

2.3.5. Steps to determine the natural pollen color distribution

Annotation step: The samples were annotated as described above. Samples that could not be accurately annotated were discarded.

EM-initialization: The EM algorithm was used to determine the chromatic mean and the color dispersion for each annotated pollen type of each sample. Gaussian distributions were initialized with the corresponding annotated cluster centroid. As a starting point for the (full) 3x3 covariance matrix, all pollen in a sample were temporarily assigned to the nearest (annotated) cluster centroid and the covariance matrix was calculated for each such cluster. The weights of the Gaussian distributions corresponded to the known proportion of each pollen type.

EM-algorithm: In each iteration, the mean vectors and covariances were updated. Only the weight vector was treated as known prior information and was kept constant.

Outlier removal: Since GMMs are considered sensitive to outliers, the EM algorithm was run repeatedly and 20% of the outliers were removed in each of five runs. The number of outliers was calculated beforehand by multiplying the proportion of unidentified pollen types and non-annotated pollen types by the number of pollen in the sample. In each outlier-removal step, those pollen colors were removed that had the lowest probability of originating from any of the fitted Gaussian

Table 2
Chromatic mean values and the expected deviation therefrom based on the fitted GMM parameters. The latter is measured as the CIE76 ΔE color difference. IQR values are omitted for pollen types with a single observation. Humans perceive color dissimilarties greater than 5 as different colors.

Pollen type	Color dispersion (median (IQR))	Chromatic means
Acer	6.33 (5.38-8.59)	
Achillea T.	7.27 (6.83-7.58)	
Aesculus	13.95 (13.60-15.08)	
Anemone T.	8.95 (8.57-9.34)	
Artemisia T.	8.69 (7.16-10.21)	
Asparagus	8.44	
Balsaminaceae	8.87 (8.25-8.92)	
Brassica T.	5.07 (4.40-8.39)	+20
Buddleja	7.76	
Castanea sativa	8.91	
Centaruea cyanus	10.14 (8.58-11.22)	
Chenopodium	9.21 (8.24-9.85)	
Cornus T.	7.17 (6.32-9.81)	
Filipendula	11.30 (11.11-12.09)	
Hedera	12.24 (12.06-12.67)	
Hydrangeaceae	10.92 (10.61-11.08)	
Parthenocissus	10.57 (9.50-10.99)	
Phacelia	13.60 (6.20-16.46)	+12
Plantaginaceae	9.95 (8.39-14.13)	+5
Poaceae	13.53	
Potentilla	12.38 (11.00-13.75)	
Prunus T., *Pyrus* T., *Rubus* T.	10.66 (9.65-13.03)	+13
Ranunculaceae	12.48 (11.78-13.18)	
Rosaceae	15.02 (10.82-16.93)	+9
Salix	15.76 (12.37-18.19)	+7
Sinapis T.	7.41 (6.25-8.00)	
Taraxacum T.	8.57 (6.87-9.67)	+9
Trifolium pratense	14.42 (13.82-15.45)	
Trifolium repens	14.13 (12.82-14.30)	
Vicia faba	10.38 (9.84-10.70)	

distributions.

Fitting: After removing the outliers, a final fitting was performed for each sample. The estimated mean vectors and covariance matrices of each pollen type's distribution were filed. Fig. 2 shows the fitted Gaussian distributions on the right. The number of plotted elements per component corresponds to the proportion of the pollen type in the palynological assessment multiplied by the number of pollen pellets in the sample.

Preparation of results: For most pollen types, several Gaussian distributions have been fitted, i.e. several mean vectors and covariance matrices were available as they occurred in multiple samples. Each of these distributions belongs to the same pollen type, but not necessarily to the same floral origin. To make the covariance accessible to interpretation, the mean Euclidean distance to the mode of the distribution was calculated. The result, a scalar, is henceforth referred to as the standard deviation of the distribution. Such a Euclidean distance in L*a*b* color space is usually denoted as CIE76 ΔE and is a measure for color differences (Robertson, 1977). Unlike the RGB color space, the L*a*b* color space seeks to reflect perceived color differences thereby making ΔE values comparable and expressive throughout the color space. For each pollen type, the median and inter-quartile range of the standard deviation of the color distributions were reported.

We used Mokrzycki and Tatol's (2011) ranking to classify and interpret our own results: Human observers do not perceive color differences with CIE76 ΔE values from 0 to 1; differences with values from 1 to 2 are detectable only by experienced observers; values from 2 to 3.5 are also perceptible by inexperienced observers; differences between 3.5 and 5 are perceived as significant color differences; values above 5 are interpreted as completely different colors.

3. Results

On 86 images of pollen samples, a total of 253 pollen types were annotated, 30 of which were different. For each sample, a GMM was fitted and the covariance was converted into standard deviations that quantify the expected color dispersion from the chromatic mean. Table 2 shows the median and inter-quartile range of the standard deviation of the colors for each pollen type, as well as the color means for each occurrence. *Brassica* T. had the lowest color dispersion (5.07) whereas *Salix* had the highest (15.76).

The data repository contains comprehensive results for individual samples, which would go beyond the scope of the presentation in this article (Borlinghaus et al., 2024). The fitted GMM parameters in JSON format alongside visualizations similar to Fig. 2 are available in the 'pollen samples' folder. Also scatter plots of the fitted Gaussian distributions for each pollen type are provided in the 'pollen types' folder (Fig. 3).

Taking the assigned pollen type predictions of the GMM models as ground truth, a pollen classification can be performed that simulates a real-world use case ignoring the known class probabilities from the laboratory analysis. Therefore, for all 62,343 available pollen colors that were not previously labeled as outliers, all 30 pollen types were assumed to be equally likely. Classification was performed by selecting the most likely color distribution (out of 253) and assigning their corresponding pollen type as the resulting class.

The aim of achieving few false positives (precision) while simultaneously minimizing false negatives (recall) is typically mutually exclusive, as demonstrated by the precision/recall curve in Fig. 4. The curve aggregates 30 one-versus-rest classifiers through non-weighted averaging.

The axes of the confusion matrix in Fig. 5 were arranged by color, which results in a more compact representation of confusion. As shown on the vertical axis, class sizes range from 22 (*Asparagus*) to 8,328 (*Prunus* T., *Pyrus* T., *Rubus* T.). The diagonal values represent the one-versus-rest classification accuracy which ranges from 0 % (*Anemone*

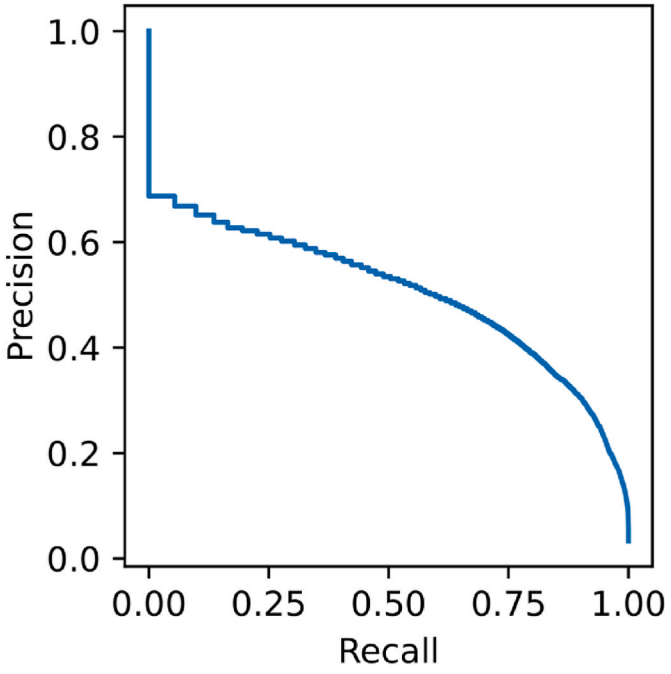


Fig. 4. Precision/recall curve of a color-based pollen classifier. The curve was calculated by averaging over 30 class-wise one-versus-rest precision/recall curves.

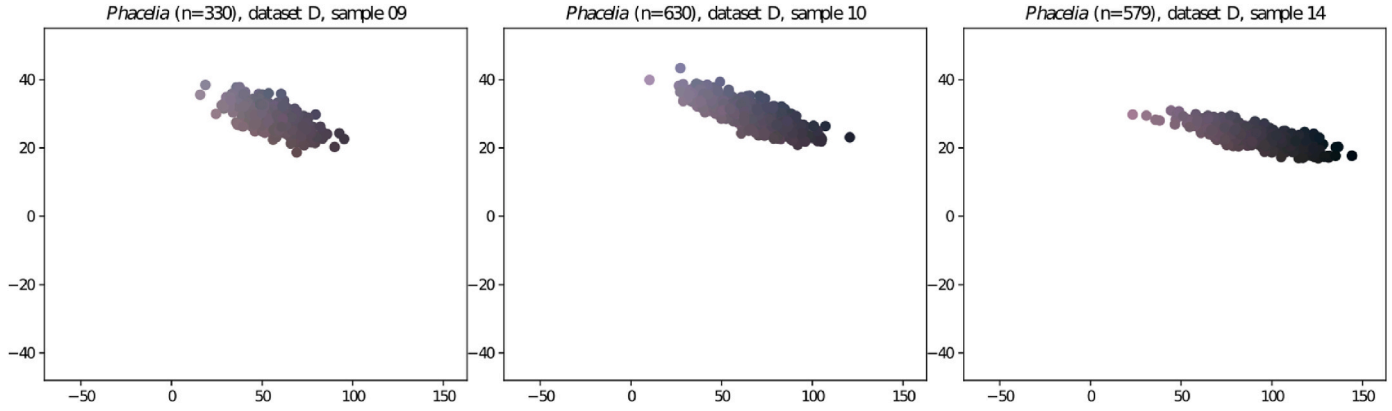


Fig. 3. Colour dispersion for 3 of 23 exemplary Phacelia observations (D09, D10 and D14). The supplementary contains all 253 available scatter plots for all pollen types.

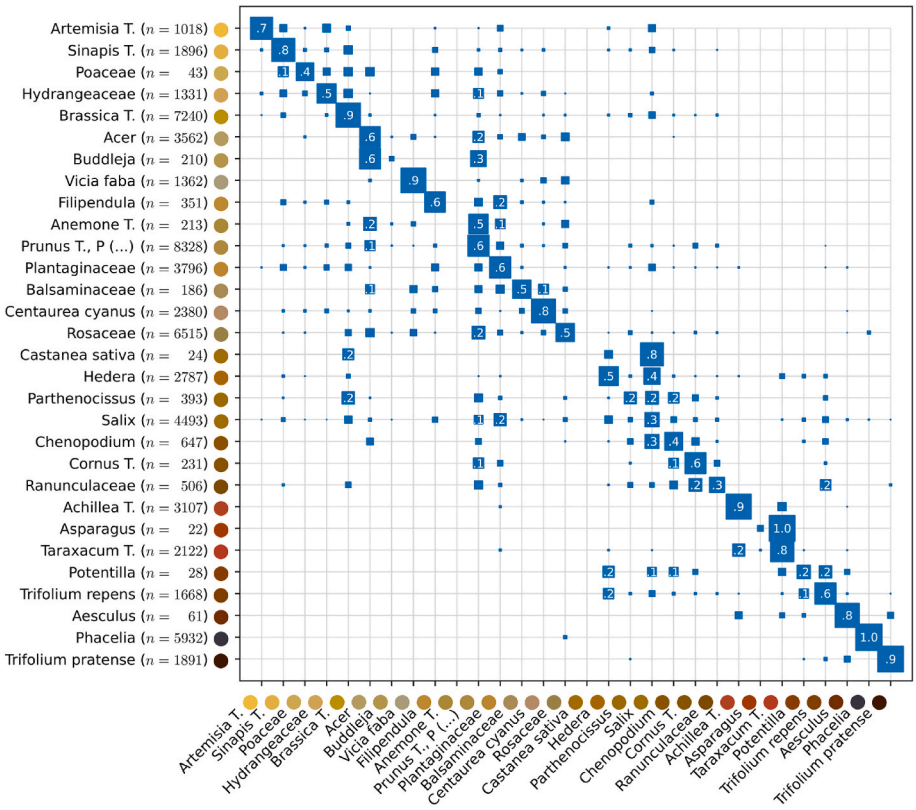


Fig. 5. Confusion matrix for pollen type classification. Similar colors are likely to be confused. Types are ordered by color, class sizes are given in brackets on the y-axis. Values greater than 0.1 are annotated and rounded to one decimal place. The true classes are shown on the vertical axis, the predicted classes are plotted horizontally. The values are row-wise normalized. Note how small class sizes led to unstable GMM that made larger, similarly colored classes more likely than true classes (e.g. castanea sativa).

T., *Castanea sativa*) to 99 % (*Phacelia*). Considering all data points, the correct classification rate is 67 % (n = 62,343, classes = 30).

4. Discussion

The availability of food resources in a landscape is discussed as an important factor in the health of honey bees and other pollinators. Quantifying this availability, and thus biodiversity, may be useful for designing specific insect pollinator support programs. Currently, there is a lack of methods that can automatically detect available food resources in relevant landscapes. Identifying plant species by pollen color could facilitate the evaluation of such support programs and may improve food resources for important insect pollinators. However, it is not known if (automatic) chromatic assessment of pollen is sufficient for more than an estimate of species diversity, as identification has proven to be even more difficult than previously known.

In our study, we collected and analyzed 86 pollen samples from beehives in four German states. Using light microscopy, we identified the pollen types and cataloged the results in a database, which was supplemented by calibrated pollen colors. The subsequent adaptation of Gaussian mixture models enabled us to interpret the distribution parameters. We were able to show that the expected color deviation from the chromatic mean was greater than five for each species and pollen type in the data set. Thereby, a value of five marks the limit above which humans perceive colors as completely different (Mokrzycki and Tatol, 2011)). It highlights the complex nature of pollen colors, which exhibit a multitude of nuances even within a single species. Consequently, the effort to categorize pollen colors under a single color proves insufficient, as the natural dispersion already spans several colors on average.

The measured colors of the pollen species that was both highly abundant and had the lowest dispersion was *Brassica* T. However, the mean standard deviation was still 5.07, which leads us to conclude that the classification of individual pollen pellets cannot be successful without additional resources, e.g. in the form of extensive expertise. The finding that plant species produce (or obtain through typical collection practices and storage) pollen colors of an unexpectedly high color spectrum limits the use of color books for species identification when using pollen traps as collection method. The reason for this is not primarily that pollen of different origins looks the same, but that pollen of the same origin does not look the same. This result is based on 85,531 corbicular pollen in 86 samples containing 14 of the 31 major European pollen types (Keller et al., 2005), distributed over several sites and a two-year sampling period in Germany.

While we are aware of the potential limitations of our GMM model, confirmation using samples that allow clear color assignment supports our conclusions (e.g., A09, A11, A12, A20, A21, A23, C14, C15, C16, C20, C21, C30, C35, C36, C37, D08, D18, D30). By additionally reporting the median and IQR of all color deviations of a pollen type, outliers were largely excluded from the interpretation, which corroborates our conclusion. As mentioned above, pollen is known to change color, for example due to drying (Campos et al., 2021). We therefore argue that color dispersion would be lower if the colors were determined immediately, ideally already in the field (Kirk, 1994). While this is true in principle, the collection period of 24–48 h and storage in the freezer make the results of this work more applicable from a practical point of view since this is common practice (Campos et al., 2008, 2021).

The confusion matrix showed instabilities when estimating a GMM with small class sizes. Against intuition, two types had a correct classification rate of zero. Given the identity of training and test data, one would expect that these two normal distributions would over fit to their corresponding pollen colors. Due to the few data points, the data did not allow a robust mean estimation leading to slightly higher matching

probabilities with wrong, similar looking types. On the contrary, the confusion matrix also showed that some pollen types can be easily distinguished. For example, *Brassica* T, *Vicia faba* and *Phacelia* produce distinct yellow, gray and dark blue colors respectively. In general, the accuracy of classification always depends on the number of classes and similarity. In another data set with many dark pollen colors, *Phacelia* could be much more difficult to recognize. However, for Central Europe, especially Germany, we consider the explanatory power of the confusion matrix to be acceptable due to the diversity of collection sites and duration. The same caution applies to the generalization of the overall correct classification rate of 67 %, which is also closely related to our data set. It would be desirable to validate the classifier on a test split at the sample level. However, this would have required that both sets contained the same species, as the species determine the pollen color. Since we only had pollen type class labels, this approach deemed unfeasible. Convinced that within-sample splitting does not test the generality of the results, we did not perform a test split. We claim that this is permissible, as the classifier is not used in an inferential but a purely descriptive way and is therefore inverted in its typical application.

In contrast to *Brassica* T., which we described as a relatively homogeneous pollen type with little dispersion, the dispersion in *Hydrangeaceae*, for example, was strikingly large. This raises the suspicion that the pollen type in samples D09, D10, and D12 was produced by several species, contrary to what we assumed. A similar observation was made for the pollen type *Plantaginaceae* in samples D19, D20, D23, and D24.

When comparing the photographed pollen with the results of our melissopalynological laboratory analysis, it became clear that a laboratory analysis is far from an ideal method. On the one hand, only a limited number of pollen grains are examined under the microscope, which inevitably leads to different proportions per pollen type depending on the subsample size. On the other hand, humans themselves also appear to be a source of error. For example, in the selection of pollen grains examined and the level of rigor. Obvious discrepancies between laboratory findings and actual pollen type occurrence even led to the exclusion of samples in our study.

Recent technological advances in pollen research could provide a more reliable identification. Dunker et al. (2021) combined multispectral imaging flow cytometry and deep learning to rapidly and accurately identify species, quantify pollen grains, and extract features of recent pollen. In addition, metabarcoding methods could also be considered to improve the reliability, cost-effectiveness, and processing speed of pollen analyses (Campos et al., 2021).

The integration of visible/near infrared (Vis/NIR) and hyperspectral imaging (HSI) technologies would be a promising way to improve automated pollen classification. These modalities provide detailed spectral signatures that could enable robust classification models through advanced machine learning algorithms (Reddy et al., 2022). This approach would promote the use of accessible tools to increase the reach of pollen analysis and avoid complex laboratory setups. These advances could significantly improve the standardization of pollen data to reduce discrepancies, as shown by the observed variations in our samples.

For example, we found a notable discrepancy in sample C07. The laboratory tests showed that 22.67 % *Phacelia* pollen was present in this sample. In fact, only 8.18% corbicular *Phacelia* pollen was visible in the photo, which could be reliably identified due to its distinct color (see Table 1). Laboratory tests carried out according to the common standard (Louveaux et al., 1978; Von Der Ohe et al., 2004) can therefore sometimes deviate considerably from reality. With an estimated 10,700 pollen grains analyzed manually based on their morphology, errors are to be expected and the question is raised whether sampling of pollen grains is sufficiently standardized to obtain comparable data.

In our previous studies (Borlinghaus et al., 2023a), we achieved an overall precision of 98.77% for automatic pollen detection, indicating low error rates. However, it is important to point out a limitation we identified: The algorithm occasionally failed to detect dark-colored pollen, mistaking it for shadows in the image (Borlinghaus et al., 2023a). This systematic overlooking of darker shades could lead to a bias in our analyses. Further limitations result from the use of GMMs. The method assumes that pollen colors of the same botanical origin follow a normal distribution around a chromatic mean and that the dataset as a whole is distributed like a mixture of three-dimensional Gaussian distributions. Such assumptions may not hold true in all cases and could lead to inaccuracies in color determination. Furthermore, the annotation step relies on prior knowledge, literature, and logical combination to assign pollen clusters to specific types. This process introduces subjectivity and potential errors, particularly in cases where clusters are not clearly separable or when multiple species produce similar colors. This must be considered when interpreting the results.

While we argue that today individual pollen cannot be reliably classified by color because of color dispersion within species, this might be possible with a sample of many pollen of the same origin. A larger number of pollen would allow estimation of both a robust chromatic mean and the covariance matrix. Because the structure of the covariance matrix varies, it could be used to identify pollen type, which has not yet been addressed. We therefore propose to test this hypothesis in future work. Clearly, the knowledge of the flowering periods and the location allows great restrictions of the possible candidates and can simplify the classification. Even without a database on flowering periods, similar results could be achieved by taking into account the probability of co-occurrence of individual types.

5. Conclusion

Our study, using real-life pollen samples collected from pollen traps, revealed a strikingly wide variation in the color dispersion among pollen pellets presumed to originate from the same botanical source. This variance is of such magnitude that to the human eye, the average color dispersion appears equivalent to the range between distinctly dissimilar colors within each of the analyzed cases. Given this substantial variability, our findings strongly caution against relying solely on color as a reliable criterion for the classification of individual pollen pellets. In light of our findings, it is evident that there is a critical need for further research and exploration in the field of pollen classification. Technologically more complex, but also more accurate methods than color determination, exist. For example combined multispectral imaging flow cytometry and deep learning, as well as molecular techniques, may enable more robust replication and higher sample throughput needed to compensate for the observed shortcomings. Ultimately, this would finally allow to assess pollen diversity in order to create a European pollen map, as suggested in previous research.

CRediT authorship contribution statement

Parzival Borlinghaus: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Visualization, Writing – original draft. **Frederic Tausch:** Conceptualization, Data curation, Writing – review & editing. **Richard Odemer:** Data curation, Resources, Writing – review & editing.

Declaration of competing interest

The authors 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

Parzival Borlinghaus receives a scholarship from the German Federal Environmental Foundation (DBU). The DBU was not involved in the study design, the collection, analysis, and interpretation of data, the

writing of the report, or the decision to submit the article for publication.

This work was supported by funds of the Federal Ministry of Food and Agriculture (BMEL) based on a decision of the Parliament of the Federal Republic of Germany via the Federal Office for Agriculture and Food (BLE) under the innovation support program OCELI (Grant Number 281C307E199).

This work was additionally supported by funds from the Federal Ministry of Food and Agriculture (BMEL) based on a decision of the Parliament of the Federal Republic of Germany via the Federal Office for Agriculture and Food (BLE) under the innovation support program VIBEE. Furthermore, this work was supported by funds of the German Government's Special Purpose Fund held at Landwirtschaftliche Rentenbank (Grant Number 28RZ3IP092).

We would like to express our particular thanks to the staff of the Expert Centre for Bees and Beekeeping in Mayen for the pollen collection and analysis.

References

Ares, A.M., Toribio, L., Tapia, J.A., González-Porto, A.V., Higes, M., Martín-Hernández, R., Bernal, J., 2022. Differentiation of bee pollen samples according to the apiary of origin and harvesting period based on their amino acid content. Food Biosci. 50, 102092 https://doi.org/10.1016/j.fbio.2022.102092.

Bilisk Tosunoglu, A., Cakmak, I., Saatcioglu, G., Bicakci, A., Malyer, H., 2008. Spectrum of pollen collected by honeybees in bursa lowland area in high season. Bee. Sci. 8, 143–148.

Bleha, R., Shevtsova, T.V., Živčáková, M., Korbářová, A., Ježková, M., Saloň, I., Brindza, J., Synytsya, A., 2021. Spectroscopic discrimination of bee pollen by composition, color, and botanical origin. Foods 10, 1682. https://doi.org/10.3390/foods10081682.

Borlinghaus, P., Jung, J., Odemer, R., 2023a. Introducing pollenyzer: an app for automatic determination of Colour diversity for corbicular pollen loads. Smart. Agric. Technol., 100263 https://doi.org/10.1016/j.atech.2023.100263.

Borlinghaus, Parzival, Odemer, Richard, Tausch, Frederic, 2024. Color dispersion of corbicular pollen loads - Supplementary Material. figshare. Dataset.

Borlinghaus, P., Odemer, R., Tausch, F., Schmidt, K., Grothe, O., 2022. Honey bee counter evaluation – introducing a novel protocol for measuring daily loss accuracy. Comput. Electron. Agric. 197, 106957 https://doi.org/10.1016/j.compag.2022.106957.

Borlinghaus, P., Tausch, F., Rettenberger, L., 2023b. A purely visual Re-id approach for bumblebees (bombus terrestris). Smart. Agric. Technol. 3, 100135 https://doi.org/10.1016/j.atech.2022.100135.

Brodschneider, R., Kalcher-Sommersguter, E., Kuchling, S., Dietemann, V., Gray, A., Božič, J., Briedis, A., Carreck, N.L., Chlebo, R., Crailsheim, K., Coffey, M.F., Dahle, B., González-Porto, A.V., Filipi, J., de Graaf, D.C., Hatjina, F., Ioannidis, P., Ion, N., Jørgensen, A.S., Kristiansen, P., Lecocq, A., Odoux, J.-F., Özkirim, A., Peterson, M., Podrižnik, B., Rašić, S., Retschnig, G., Schiesser, A., Tosi, S., Vejsnæs, F., Williams, G., van der Steen, J.J.M., 2021. CSI pollen: diversity of honey bee collected pollen studied by citizen scientists. Insects 12, 987. https://doi.org/10.3390/insects12110987.

Campos, M.G., Anjos, O., Chica, M., Campoy, P., Nozkova, J., Almaraz-Abarca, N., Barreto, L.M.R.C., Nordi, J.C., Estevinho, L.M., Pascoal, A., Paula, V.B., Chopina, A., Dias, L.G., Tešić, Ž.L.J., Mosić, M.D., Kostić, A.Ž., Pešić, M.B., Milojković-Opsenica, D.M., Sickel, W., Ankenbrand, M.J., Grimmer, G., Steffan-Dewenter, I., Keller, A., Förster, F., Tananaki, C.H., Liolios, V., Kanelis, D., Rodopoulou, M.-A., Thrasyvoulou, A., Paulo, L., Kast, C., Lucchetti, M.A., Glauser, G., Lokutova, O., De Almeida-Muradian, L.B., Szczęsna, T., Carreck, N.L., 2021. Standard methods for pollen research. J. Apicult. Res. 60, 1–109. https://doi.org/10.1080/00218839.2021.1948240.

Campos, M.G., Bogdanov, S., De Almeida-Muradian, L.B., Szczesna, T., Mancebo, Y., Frigerio, C., Ferreira, F., 2008. Pollen composition and standardisation of analytical methods. J. Apicult. Res. 47, 154–161. https://doi.org/10.1080/00218839.2008.11101443.

Casas Restrepo, L.C., Gutierrez Alabat, I.E., Salamanca Grosso, G., de Assis Ribeiro dos Santos, F., 2023. Markers for the spatial and temporal differentiation of bee pollen harvested by Apis mellifera L. in the Eastern Andes of Colombia. J. Apicult. Res. 62, 556–569. https://doi.org/10.1080/00218839.2021.1916188.

Conti, I., Medrzycki, P., Grillenzoni, F.V., Corvucci, F., Tosi, S., Malagnini, V., Spinella, M., Mariotti, M.G., 2016. Floral diversity of pollen collected by honey bees (Apis mellifera L.) – validation of the chromatic assessment method. J. Apicult. Sci. 60, 209–220. https://doi.org/10.1515/jas-2016-0028.

De Klerk, P., Joosten, H., 2007. The difference between pollen types and plant taxa: a plea for clarity and scientific freedom. Quat. Sci. J. 56, 162–171. https://doi.org/10.3285/eg.56.3.02.

Dempster, A.P., Laird, N.M., Rubin, D.B., 1977. Maximum likelihood from incomplete data via the EM algorithm. J. Roy. Stat. Soc. B 39, 1–38.

Dunker, S., Motivans, E., Rakosy, D., Boho, D., Mäder, P., Hornick, T., Knight, T.M., 2021. Pollen analysis using multispectral imaging flow cytometry and deep learning. New Phytol. 229, 593–606. https://doi.org/10.1111/nph.16882.

Fægri, K., Iversen, J., Kaland, P.E., Krzywinski, K., 2007. In: Textbook of Pollen Analysis, IV ed. Blackburn Press, Caldwell, NJ.

Hodges, D., 1952. The Pollen Loads of the Honeybee. A Guide O Their Identification by Colour and Form. Bee Research Association, London.

Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES), 2016. Assessment Report on Pollinators, Pollination and Food Production. Zenodo. https://doi.org/10.5281/zenodo.3402857.

Keller, I., Fluri, P., Imdorf, A., 2005. Pollen nutrition and colony development in honey bees: part 1. Bee World 86, 3–10. https://doi.org/10.1080/0005772X.2005.11099641.

Kirk, W.D.J., 2006. In: A Colour Guide to Pollen Loads of the Honey Bee, second ed. International Bee Research Association, Cardiff.

Kirk, W.D.J., 1994. Recording the colours of pollen loads. Bee World 75, 169–180. https://doi.org/10.1080/0005772X.1994.11099224.

Knoll, F., 1956. Die Biologie der Blüte, Verständliche Wissenschaft. Springer Berlin Heidelberg, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-86220-5.

Layek, U., Das, N., Kundu, A., Karmakar, P., 2022. Methods employed in the determining nectar and pollen sources for bees: a review of the global scenario. Ann. Entomol. Soc. Am. 115, 417–426. https://doi.org/10.1093/aesa/saac013.

Lloyd, S., 1982. Least squares quantization in PCM. IEEE Trans. Inf. Theor. 28, 129–137. https://doi.org/10.1109/TIT.1982.1056489.

Louveaux, J., Maurizio, A., Vorwohl, G., 1978. Methods of melissopalynology. Bee World 59, 139–157. https://doi.org/10.1080/0005772X.1978.11097714.

Milla, L., Schmidt-Lebuhn, A., Bovill, J., Encinas-Viso, F., 2022. Monitoring of honey bee floral resources with pollen DNA metabarcoding as a complementary tool to vegetation surveys. Ecol. Solut. Evid. 3 https://doi.org/10.1002/2688-8319.12120.

Mokrzycki, W.S., Tatol, M., 2011. Colour difference Delta E - a survey. MG&V 20, 383–411.

Newstrom-Lloyd, L., Raine, I., Li, X., 2017. The power of pollen profiles for planting trees for bees. Trees. Bees 12.

Ollerton, J., Winfree, R., Tarrant, S., 2011. How many flowering plants are pollinated by animals? Oikos 120, 321–326. https://doi.org/10.1111/j.1600-0706.2010.18644.x.

Percival, M., 1947. Pollen collection by Apis mellifera. New Phytol. 46, 142–165. https://doi.org/10.1111/j.1469-8137.1947.tb05076.x.

Reddy, P., Guthridge, K.M., Panozzo, J., Ludlow, E.J., Spangenberg, G.C., Rochfort, S.J., 2022. Near-infrared hyperspectral imaging pipelines for pasture seed quality evaluation: an overview. Sensors 22, 1981. https://doi.org/10.3390/s22051981.

Reynolds, D., 2009. Gaussian mixture models. In: Li, S.Z., Jain, A. (Eds.), Encyclopedia of Biometrics. Springer, US, Boston, MA, pp. 659–663. https://doi.org/10.1007/978-0-387-73003-5_196.

Robertson, A.R., 1977. The CIE 1976 color-difference formulae. Color Res. Appl. 2, 7–11. https://doi.org/10.1002/j.1520-6378.1977.tb00104.x.

Swiatly-Blaszkiewicz, A., Pietkiewicz, D., Matysiak, J., Czech-Szczapa, B., Cichocka, K., Kupcewicz, B., 2021. Rapid and accurate approach for honeybee pollen analysis using ED-XRF and FTIR spectroscopy. Molecules 26, 6024. https://doi.org/10.3390/molecules26196024.

Tausch, F., Wagner, J., Klaus, S., 2023. Pollinators as data collectors: estimating floral diversity with bees and computer vision. In: Presented at the Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 643–650.

Von Der Ohe, W., Persano Oddo, L., Piana, M.L., Morlot, M., Martin, P., 2004. Harmonized methods of melissopalynology. Apidologie 35, S18–S25. https://doi.org/10.1051/apido:2004050.