ROYAL SOCIETY OPEN SCIENCE

royalsocietypublishing.org/journal/rsos

Research

Cite this article: Bean H, Field D, McNamara ME. 2026 Quantitative characterization of melanin-based feather colour using reflectance spectrophotometry. *R. Soc. Open Sci.* **13**: 251668.
https://doi.org/10.1098/rsos.251668

Received: 2 September 2025
Accepted: 13 November 2025

Subject Category:
Earth and environmental science

Subject Areas:
palaeontology

Keywords:
melanin, feathers, fossil colour, reflectance spectrophotometry, LDA, spectral analysis

Author for correspondence:
Hollie Bean
e-mail: hbean@ucc.ie

Supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.8153363.

THE ROYAL SOCIETY PUBLISHING

Quantitative characterization of melanin-based feather colour using reflectance spectrophotometry

Hollie Bean[1,2], Daniel Field[3,4] and Maria E. McNamara[1,2]

[1]School of Biological, Earth and Environmental Sciences, and [2]Sustainability Institute, University College Cork, Cork, County Cork, Republic of Ireland
[3]Department of Earth Sciences, and [4]Museum of Zoology, University of Cambridge, Cambridge, UK

HB, 0000-0002-0636-3262; DF, 0000-0002-1786-0352; MEMN, 0000-0003-0968-4624

Colour produced via melanin pigments is widespread in extant bird feathers and has been inferred in many fossil birds, feathered dinosaurs and pterosaurs. Current models of fossil feather colour are based on the geometry of feather melanosomes in extant birds, but to date, feather colour has been assigned on a purely visual basis. Here, we assess statistical support for visually assigned categories of melanin-based colour. We performed quantitative analysis of spectral reflectance data from black, brown, grey and rufous feathers, using both raw reflectance spectra and spectrally derived parameters. Only rufous feathers show a systematic division into distinct light and dark subcategories, whereas brown, grey and black categories are better treated as discrete groups. Analysis of spectral data demonstrates that colour categories are distinct. Predictive models for feather colour using linear discriminant analysis yield predicted accuracies of over 80% using parameter data and over 90% using raw spectral data. Results confirm the validity of these visually assigned colour categories. Future studies of feather colour should use spectral analysis to test whether visually assigned colour categories are statistically supported. This will provide a robust foundation for future efforts to investigate the evolution of feather colour in the fossil record.

1. Introduction

Integumentary coloration is a striking attribute of birds, exhibiting more variation in this group than in any other terrestrial vertebrate clade [1]. The colours and patterns of

feathers relate to a wide range of signalling functions, including camouflage and intraspecific signalling [2]. Colour expression in feathers results from the interaction of light with pigments, structural features or both [3]. Pigments preferentially absorb specific wavelengths of light and reflect non-absorbed wavelengths that contribute to visible colour [4]. Common feather pigments include melanins, carotenoids, pterins, porphyrins and psittacofulvins [4]. Melanins are of particular interest because they are the most common pigment in the feathers of extant birds [1] and are present in representatives of every extant avian order. Furthermore, melanins are the only pigments identified in fossil feathers to date [5]. As such, fossil evidence for melanin has the potential to provide unique insights into the evolution of colour and its functions in feathered animals [6]. This, however, requires a comprehensive understanding of both the biological basis of melanin-based coloration and how this is impacted by fossilization.

In the feathers of extant birds, melanins produce black, brown, grey and rufous pigmentary colours [4]. Melanin can also co-occur with other pigments and/or quasi-ordered nanostructures that give rise to non-iridescent structural colour and iridescent colour in feathers [4,7]. In these cases, melanin serves as a backing pigment and occurs primarily as a layer of melanosomes (melanin-containing organelles) below the colour-producing region of the feather cortex [8]. In this study, we focus on pigmentary colours produced primarily by melanins (black, brown, grey and rufous hues).

Visible feather colour has been linked to various attributes of melanosomes, notably melanosome size, geometry, internal structure and packing arrangement [9–11]. The link between melanosome geometry and feather colour has direct applications to the fossil record, as feathers are preserved in Cenozoic crown birds [10,12], both avialan and non-avialan dinosaurs [9,13–15] and pterosaurs [16–18]. Since the first report of melanosomes in fossil feathers [19], numerous studies have used preserved melanosome geometries in fossils to infer original melanin-based feather colour [9,10,12–15,18,20–23]. This process has required the construction of large datasets of melanosome geometry for feathers from extant birds [9,10,18], in which measurements of melanosome length and width derived from scanning electron microscopy images are used to calculate parameters such as aspect ratio and skew for feathers assigned to different colour categories [10,14]. Geometric data extracted from fossil melanosomes are then used to assign original feather colour based on multivariate analysis of data from extant birds [9,10,13,21].

To date, colour reconstructions based on melanosome geometry data have applied broad colour categories, including black, brown, grey, iridescent [9], non-iridescent structural colour [10] and a 'penguin-type' category [12]. The predictive accuracy of colour inference models is relatively low: the most recent reported accuracy is 61.8% [10] based on quadratic discriminant analysis. Black and grey colour categories are consistently predicted with the lowest accuracy of all categories [9,10], with predictions of black colour up to 55.6% accurate, grey colours 63.5% and brown colours up to 73% accurate [9].

The accuracy of colour predictions for fossil feathers is dependent on various factors, especially the quality of the extant bird feather colour dataset used for comparative analysis. Thus far, such datasets have assigned feathers to colour categories on a purely visual basis, with no quantitative or statistical support for colour assignments. However, this approach may be problematic as human colour perception is subjective and may be influenced by environmental factors [24]. As such, a fundamental step in improving the predictive power of inference models for melanin-based feather coloration is the development of a quantitative basis for colour assignments, but this is yet to be attempted. This may be at least partially because of known difficulties with the quantitative assessment of melanin-based colours, chiefly the lack of distinct peaks in reflectance spectra [25]. The near-uniform low reflectance of black, brown and grey feathers yields characteristically flat spectral shapes, and rufous feathers show a gradual increase in reflectance towards longer wavelengths, yielding non-peaked spectral curves with a gentle incline. Reflectance spectrophotometry has been widely used in analyses of feather colour [1,26–32], and parameters extracted from spectral data, such as saturation and maximum brightness, may be used to describe spectra quantitatively [4,30].

Here, we test the hypothesis that visual assignment of feather colour can be statistically supported. We characterized feather colour using reflectance spectrophotometry and a suite of spectral parameters derived from visible reflectance spectra. We then investigated whether visually assigned colour categories can be discriminated by significant differences in specific spectral parameters. Our results confirm that certain visually defined colour categories are statistically valid and that others can be subdivided, yielding greater colour specificity. We tested the accuracy of colour category prediction based on both raw reflectance spectra and spectral parameters. Such quantitative characterization of feather colour can contribute to a better understanding of the extent to which feather colour is linked

to other important feather characters and will assist with future efforts to investigate the evolution of feather colour in the fossil record.

2. Material and methods

2.1. Material

Our dataset comprises 226 spectral measurements of feather colour from 140 species of extant birds that represent all 44 extant avian orders (electronic supplementary material, table S1). Selected species exhibit at least one region ≥ 10 × 10 mm (herein termed a 'colour patch') of uniform melanin-based coloration in their plumage. Our dataset was initially constructed to capture within-order diversity using a semi-proportional dataset, with additional samples subsequently added to increase the sample size of under-represented colour categories. Shed feathers from the common ostrich (*Struthio camelus*) and lesser rhea (*Rhea pennata*) were provided by Fota Wildlife Park, Cork, Ireland. Feathers from all other taxa are from the collections of the University of Cambridge Museum of Zoology (Cambridge, UK). Where a single specimen showed plumage regions with different melanin-based colours (either in multiple body regions or within single feathers), each region was analysed.

We removed two data points from our dataset prior to analysis. Colour patch no. 21 was removed as it exhibits a metallic lustre, indicating a structural component. Colour patch no. 170 was removed owing to an unexplained mismatch between visual colour classification and the measured spectra. The total dataset used for subsequent analysis therefore comprises 224 colour patches from 139 extant bird species.

2.2. Definition of colour categories

Given that feathers were assigned to colour categories based on *in situ* plumage, spectral measurements were also performed *in situ* where possible. This assignment of colour was performed visually, without any equipment, initially by both H.B. and M.E.McN. to validate colour assignment, then by H.B. solely. This study assumes that the colour discrimination abilities of the authors are broadly representative of the general population. This has not, however, been tested experimentally and therefore does not imply that the results of this study can be applied automatically to other studies using the same colour categories.

Previous studies of melanin-based colours in feathers have consistently used black, brown, grey and orange/rufous colour categories [9,10,25]. One previous study, however, divided the colour categories brown, grey and orange into light and dark subcategories [25]. Given this lack of consensus on appropriate colour categories for feathers, we tested the validity of both undivided and subdivided categories for brown, grey and rufous colours.

2.3. Reflectance spectrophotometry

Reflectance spectrophotometry is the most widely used method for quantifying visible colour [4]. Reflectance spectra were collected using an Ocean Optics USB2000+ fibre optic spectrophotometer (range 200–1100 nm) with a tungsten halogen visible light source, a 400 µm diameter bifurcated reflection probe and OceanView spectral software (v. 2.0). A probe holder was used to maintain probe position perpendicular to the feather surface during colour measurements. A white standard spectrum for reference calibration and dark background spectrum were taken prior to collecting data from feathers. For each colour patch analysed, 3–4 reflectance spectra were measured from small (7 mm diameter) circular regions of the patch. Each measurement represents an average of 10 scans with a boxcar width of four. Spectral measurements were collected from feathers *in situ* on the bird specimen to preserve the optical properties of the plumage [33], excluding *S. camelus* and *R. pennata*. Feathers from these two taxa were unavailable *in situ* so several feathers were instead superimposed prior to analysis to minimize any contribution from the backing substrate in the analysed region.

Our dataset comprises reflectance spectra spanning the visible light spectrum (380–700 nm). Avian colour vision extends into the UV-A region (315–380 nm [4]), but as the aim of the present study was to assess colour classification by human vision (which does not extend into the UV range), we do not consider the limitation of our data to the human visible spectrum to adversely impact our results.

Reflectance spectra were processed and analysed in R (v. 4.3.1) [34]. Raw reflectance spectra were cut to the visible light spectrum (380–700 nm). Spectral parameters were extracted using functions from the *pavo* package (v. 2.9.0) [35]. Raw spectra were loess-smoothed with a span of 0.35 to remove noise using the procspec() function. Each spectrum was characterized using seven parameters: total brightness, maximum brightness, saturation, slope and red, green and blue (RGB) values. Total brightness was calculated as the sum of the reflectance values for each wavelength of light in the spectrum [4]. Maximum brightness (intensity) is the maximum value for percentage reflectance in the spectrum [4]. Saturation (chroma) was calculated as the difference between the maximum and minimum reflectance values, with this value then divided by the mean reflectance [4]. Slope is the gradient of the linear trendline fitted to the data following linear regression of reflectance against wavelength [25]. Red, green and blue are the RGB values for each spectrum, calculated using the spec2rgb() function [35].

2.4. Statistical analysis of spectral parameters

All statistical tests and data analyses were performed in R (v. 4.3.1) [34]. Data for each parameter were tested for normality using the Shapiro-Wilk test. Differences in the median values of each spectral parameter among colour categories were tested for significance using the non-parametric Kruskal-Wallis test, and the *post hoc* pairwise Wilcoxon test was used to test for significant differences between medians for each pair of colour categories for each parameter.

2.5. Multivariate analysis of spectral data

We used principal component analysis (PCA) as an exploratory multivariate analysis to visualize the extent of differences among colour categories without prior grouping specifications. We performed PCA separately on raw spectral data and spectral parameter data. We then performed linear discriminant analysis (LDA) to visualize differences among colour categories when data groupings were known, assessing the possibility of predictive potential by the extent of overlap in plots. Data were centred and scaled prior to PCA and LDA.

2.6. Colour prediction models

We used *k*-fold cross-validation to assess the predictive performance of LDA. Our datasets were split into separate training (80%) and testing (20%) partitions, with the proportion of each colour category retained in both partitions. LDA was performed on the training partition, with the testing partition used to assess the accuracy of colour category prediction by each model. The proportion of correct versus incorrect colour category predictions was used to evaluate the predictive power of each model. This procedure was repeated 100 times ($k = 100$) to allow for subsequent significance testing between models via *t*-tests. Predictive models were produced using: (i) spectral parameter data, (ii) log-transformed parameter data, (iii) raw spectral data, (iv) log-transformed raw spectral data, and (v) smoothed spectral data to assess differences in performance. Models were also run for each discrete combination of spectral parameters to assess which combination held the most predictive power. We measured mean predicted accuracy for each model and sensitivity and specificity for each colour category within each model. Sensitivity is a measure of the true-positive rate; that is, the ability of the model to correctly predict the colour of a feather. Specificity is a measure of the true-negative rate; that is, the ability of the model to correctly predict when a feather is not a particular colour. Finally, *t*-tests were used to assess the significance of differences between pairs of models.

3. Results

3.1. Definition of colour categories

Visual inspection of the feathers yielded distinct black, brown, grey and rufous categories. In addition, rufous feathers were readily visually separable into light and dark rufous subcategories. This

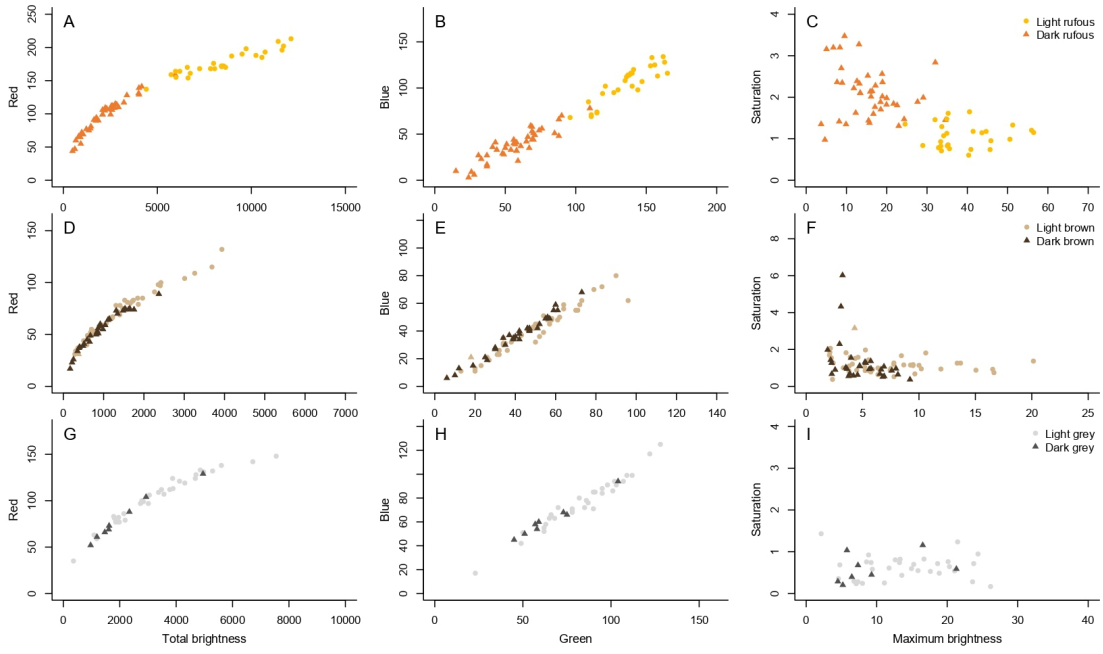


Figure 1. Scatterplots for various spectral parameters using light and dark subsets of each feather colour category. (A–C) Feathers coded visually as light ($n = 27$) and dark ($n = 42$) rufous. (D–F) Feathers coded visually as light ($n = 45$) and dark ($n = 32$) brown. (G–I) Feathers coded visually as light ($n = 32$) and dark ($n = 8$) grey. For each parameter, values refer to arbitrary units. Additional scatterplots for light and dark subsets of the colour categories are provided in the electronic supplementary material, figure S1. The scatterplots in this figure and in the electronic supplementary material, figure S1 show all spectral parameters used.

is supported by the minimal overlap of the data for light and dark rufous feathers for all spectral parameters (figure 1; electronic supplementary material, figure S1). For brown and grey feathers, however, there is a more gradational transition between light and dark hues in hand specimens and substantial overlap of the spectral parameter data for the light and dark subcategories of brown and grey (figure 1; electronic supplementary material, figure S1). This more gradational transition may be because of differences in melanin concentration within individual feathers [36]. As a result, we decided to apply a conservative approach whereby brown and grey colour categories were not subdivided. Based on these observations and data, our dataset comprises 38 colour patches categorized as black, 77 categorized as brown, 40 as grey, 42 as dark rufous and 27 as light rufous (electronic supplementary material, table S1).

3.2. Reflectance spectrophotometry

Spectra for black feathers are flat, with low reflectance values across the visible spectrum (figure 2). Spectra for brown feathers also show low reflectance across the spectrum, with reflectance values increasing slightly towards longer wavelengths. Spectra for grey feathers are similar in shape to those for brown feathers, but the former exhibit higher reflectance values across the visible range. Both dark rufous and especially light rufous feathers show a progressive increase in reflectance values towards longer wavelengths (figure 2). Full raw spectral data are reported in the electronic supplementary material, table S2.

3.3. Statistical analysis of spectral parameters

Shapiro-Wilk test results confirm that the data for each spectral parameter are non-normal ($p < 0.05$ for all parameters; electronic supplementary material, table S3). Kruskal-Wallis test results show that for each spectral parameter analysed, there are significant differences among colour categories for all parameters ($p < 0.001$; electronic supplementary material, table S3). Wilcoxon *post hoc* test results show significant differences between most pairs of colour categories (figure 3; electronic supplementary material, table S3). All pairs of colour categories differ significantly in values for total brightness, slope and green ($p < 0.05$; figure 3A,B; electronic supplementary material, table S3). The only pairs of colour

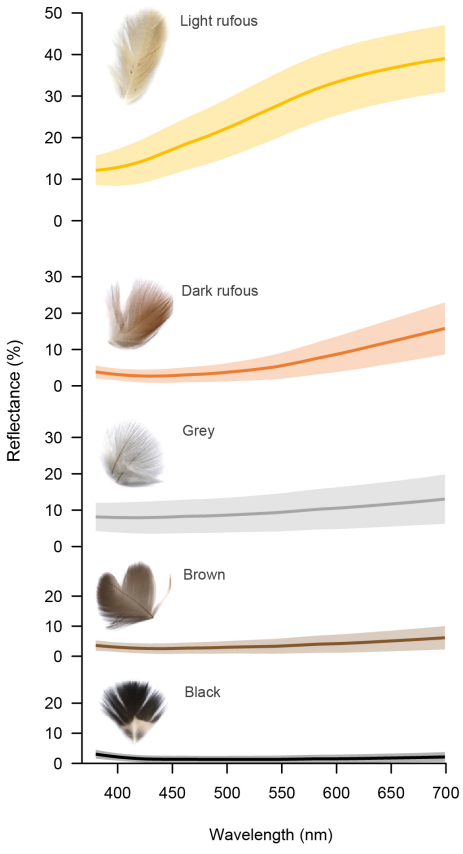


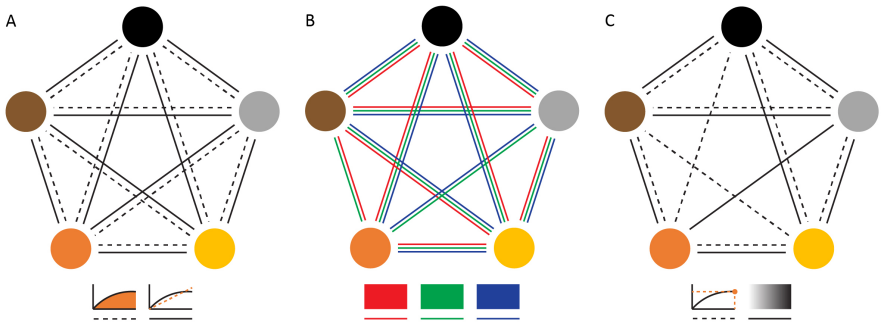
Figure 2. Mean reflectance spectra for feathers in each visually assigned colour category. Shaded regions denote standard deviation for the data in each category. Colour categories are light rufous ($n = 27$), dark rufous ($n = 42$), grey ($n = 40$), brown ($n = 77$) and black ($n = 38$). The inset associated with each plot shows a photograph of a representative feather from that colour category. Light rufous feather: *Eurypyga helias* (UMZC 15/Eur/1/a/3), dark rufous feather: *Tadorna variegata* (UMZC 12/Ana/13/d/2), grey feather: *Phalacrocorax punctatus* (UMZC 10/Phal/3/u/1), brown feather: *Nestor meridionalis* (UMZC 18/Psi/46/a/1), black feather: *Colaptes cafer* (UMZC 26/Pic/8/b/1).

categories for which the statistical tests yield a non-significant ($p > 0.05$) result are as follows (see also figure 3B,C; electronic supplementary material, table S3): grey and dark rufous feathers do not differ significantly in maximum brightness ($p = 0.13$) and red ($p = 0.73$) values; black and dark rufous feathers do not differ significantly for saturation ($p = 0.428$), nor do brown and light rufous feathers ($p = 0.594$); finally, brown and dark rufous feathers do not differ significantly for blue values ($p = 0.883$).

3.4. Statistical analysis of spectral parameters versus raw spectral data

PCA plots of spectral parameter data (figure 4A) and raw spectral data (figure 4B) provide a visual indication of the differences among colour categories without the input of predefined groupings. The PCA for spectral parameters (figure 4A) indicates that principal component (PC) 1 and PC2 collectively represent 95% of the variation within the dataset, with PC1 primarily explaining variation in red, green, total brightness and maximum brightness values, and PC2 primarily explaining variation in saturation. Although colour categories overlap, light rufous feathers plot separately, and dark rufous feathers show minimal overlap with brown and grey feathers. The PCA for the raw reflectance data indicates that PC1 and PC2 collectively represent 99% of the variation in the dataset (figure 4B). This plot shows better separation among colour categories, with dark and light rufous feathers separable from other colours and minimal overlap between black and other feathers.

LDA was used to assess the impact of assigning feather samples to predefined colour categories. The LDA for spectral parameter data indicates that linear discriminant (LD) 1 and LD2 represent 89% of the variation within the dataset (figure 4C). Colour categories overlap less extensively than in the PCA (figure 4A). Separation of light rufous feathers from feathers in all other colour categories is driven primarily by differences in slope, total brightness and maximum brightness (figure 4C). Separation of dark rufous feathers from black, brown and grey feathers is primarily the result of



Schematic	Spectral parameter	Description
	Total brightness	Total reflected light, measured as area under the spectrum
	Slope	The gradient of the linear regression for reflectance versus wavelength
	Red	Red RGB value
	Green	Green RGB value
	Blue	Blue RGB value
	Maximum brightness (intensity)	The value of maximum percentage reflectance
	Saturation (chroma)	$\frac{\text{(Maximum reflectance - Minimum reflectance)}}{\text{Mean reflectance}}$

Figure 3. Pentagrams illustrating the results of pairwise Wilcoxon tests. Coloured circles represent visually assigned feather colour categories (clockwise from top: black, grey, light rufous, dark rufous, brown). Connecting lines denote significant differences between categories. (A) Solid lines, slope; dashed lines, total brightness. (B) Solid lines, RGB values, with data for R (red), G (green) and B (blue) represented by the corresponding colours (red, green, blue, respectively). (C) Solid lines, saturation; dashed lines, maximum brightness. Schematics below the pentagrams represent spectral parameters analysed, from left to right: (A) total brightness, slope; (B) red, green, blue; (C) maximum brightness, saturation. All *p*-values can be viewed in the electronic supplementary material, table S3.

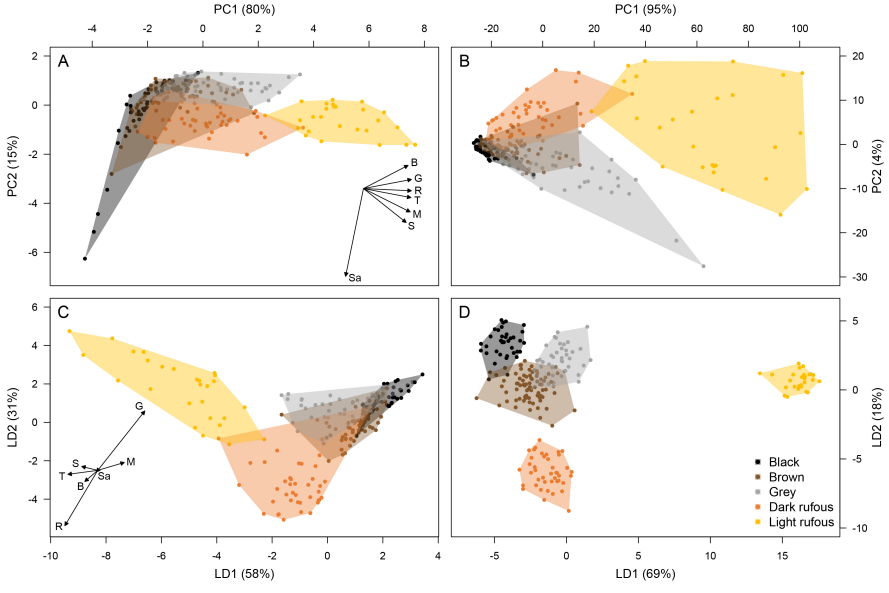


Figure 4. Multivariate analyses of spectral data. (A) Principal component analysis (PCA) plot using spectral parameter data. (B) PCA using raw spectral data. (C) Linear discriminant analysis (LDA) plot using spectral parameter data. (D) LDA using raw spectral data. (A) and (C) are accompanied by a biplot that shows the relative contribution of each spectral parameter to variation in the dataset. Biplots are not shown for (B) and (D) owing to the large number of variables included (each wavelength datapoint). Abbreviations: B, blue; G, green; M, maximum brightness; R, red; S, slope; Sa, saturation; T, total brightness.

differences in red and green values. The LDA plot for raw spectral data shows the best separation among colour categories (figure 4D), with 87% of the variation captured by LD1 and LD2. Light and dark rufous categories occupy distinct regions of the plot. Black and grey categories plot separately with only minimal overlap with brown feathers.

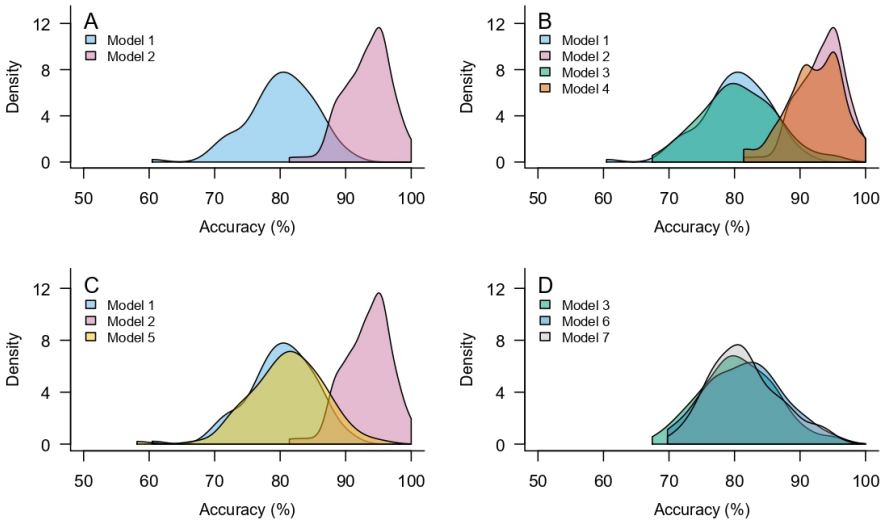


Figure 5. Density plots of percentage accuracy for 100 folds of *k*-fold cross-validation of LDA models. Model 1 uses spectral parameter data. Model 2 uses raw spectral data. Model 3 uses log-transformed parameter data. Model 4 uses log-transformed raw spectral data. Model 5 uses smoothed spectral data. Model 6 uses log-transformed total brightness, maximum brightness, red and green parameters. Model 7 uses log-transformed total brightness, saturation, red and green parameters. (A) Direct comparison of the accuracy of models 1 and 2. (B) Direct comparison of the accuracy of models 1-4. (C) Direct comparison of the accuracy of models 1, 2 and 5. (D) Direct comparison of the accuracy of models 3, 6 and 7.

Table 1. Results of *k*-fold cross-validation of LDA models. (Values represent mean percentage accuracy for each model.)

dataset	percentage accuracy (%)
raw spectra	93.3
log-transformed raw spectra	92.3
spectral parameters	80.2
log-transformed spectral parameters	80.6
smoothed spectra	81.1
log-transformed total brightness, maximum brightness, red and green	81.6
log-transformed total brightness, saturation, red and green	81.5

3.5. Colour prediction using linear discriminant analysis

Mean accuracies of colour prediction for full spectral and parameter data vary from 80.2 to 93.3% (table 1; figure 5). *k*-fold cross-validation of the colour prediction model using spectral parameter data produces a mean predicted accuracy of 80.2% (table 1; figure 5A). Log-transforming the data prior to analysis produces a near-identical accuracy of 80.6% (table 1; figure 5B). Predicted accuracy is notably higher for raw spectral data (93.3%; table 1; figure 5A). Log-transformation of raw spectral data yields a similarly high accuracy for colour prediction (92.3%; table 1; figure 5B). Smoothed spectral data produce an accuracy of 81.1% (figure 5C). *t*-tests show that differences in mean accuracy between raw spectral and log-transformed raw spectral data are non-significant. There is also no significant difference between untreated parameter data and log-transformed parameter data (electronic supplementary material, table S5).

Stepwise variable testing using log-transformed spectral parameter data yields predicted accuracies ranging from 46.5 to 82.8%. Two optimal parameter combinations reduce the number of independent spectral parameters while maintaining high levels of predicted accuracy: total brightness, maximum brightness, red and green (81.6%, table 1; figure 5D); and total brightness, saturation, red and green (82.5%, table 1; figure 5D). *t*-tests reveal no significant differences in predictive accuracy when using the log-transformed parameter data and either of the above sets of parameter combinations (electronic supplementary material, table S5).

Sensitivity (true-positive rate) and specificity (true-negative rate) describe model performance for the prediction of individual colour categories. Spectral parameter data produce values greater than

70% for both sensitivity and specificity for all colour categories, and values for most categories exceed 85% (electronic supplementary material, table S4). The lowest value is for black sensitivity (70.7%). Log-transformed parameter data show lower sensitivity and specificity values for the black and the two rufous categories, although grey sensitivity shows slightly higher values than untransformed parameter data. The highest sensitivity and specificity values are for raw spectral data (electronic supplementary material, table S4), whereby the values for both metrics for all colour categories are 99.8%. Log-transformed spectral data also yield high values, although there is notably lower sensitivity for dark rufous (84.2%). Using smoothed spectral data, sensitivity values are lower for black (72.2%) and grey feathers (74.9%; electronic supplementary material, table S4), but for all other categories, values are greater than 80%. Compared to the log-transformed parameter data, all sensitivity and specificity values are higher for each of the optimal parameter combinations described above (i.e. total brightness, maximum brightness, red and green; total brightness, saturation, red and green; electronic supplementary material, table S4).

4. Discussion

This study presents a statistically robust multivariate analysis of melanin-based feather colour by investigating the legitimacy of visually assigned colour categories using both raw reflectance spectra and their derived spectral parameters. Subjectively assigned colour categories are confirmed to differ significantly, yielding discrete colour categories compatible with subsequent predictive modelling. Our results also support discrete subcategories for light and dark rufous colours, whereby light rufous feathers appear more yellow in colour and dark rufous feathers more orange/reddish. These subcategories for rufous feathers have not been used in published models of fossil feather colour [9,10]. Inclusion of these subcategories will therefore add nuance to future inferences of fossil feather colour.

We find that feather colour can be predicted to a high level of accuracy using spectral and spectrally derived parameter data, confirming the validity of black, brown, grey, dark rufous and light rufous colour categories for the analysis of melanin-based colour in feathers. Using this framework, spectra or spectral characteristics relating to particular colour categories may subsequently be linked to other feather characters and potentially other aspects of avian biology, overcoming an aspect of error from unsupported colour category assignments. Strong statistical support for the colour categories defined using our dataset suggests that such categories are likely to also apply to datasets beyond those derived from birds, and as such, have potential to inform investigations in other animal groups where melanin-based colour is observed. Our results do not, however, automatically imply that the colour categories defined in our study can be used in future studies without statistical analysis of spectral data, nor does our study imply that the colour categories used in previous studies [9,10,25] are statistically robust. This partly relates to differences in the taxonomic composition of our dataset and those of other studies. We cannot exclude the possibility that feathers in certain taxa may be more difficult to categorize reliably than others. More importantly, however, the extent to which melanin-based feather colours can be consistently assigned has not been tested using a large human population or repeat assignment tests in individuals. Future studies may help constrain this potential source of variability.

The overlap of data for different colour categories in PCA and LDA plots (figure 4) is not unexpected, owing to strong similarities in spectral shape (and thus spectral parameters such as slope and maximum brightness; electronic supplementary material, table S1) among certain colours, particularly black, brown and grey (figure 2). Conversely, consistent separation of the data for light rufous feathers from those for all other feather colours in PCA and LDA plots is because of strong differences in spectral shape across the visible spectrum (figure 2), and hence also differences in parameters such as maximum brightness and slope (electronic supplementary material, table S1). Notably, despite extensive overlap in PCA and LDA space (figure 4A,C) for certain pairs of colour categories (black and brown, black and grey, and brown and grey), differences in spectral parameters are statistically significant for each pair (figure 3; electronic supplementary material, table S3). Overall, slope, green and total brightness differ significantly for all pairs of colour categories (figure 3; electronic supplementary material, table S3). Slope has been previously identified as a good predictor of human melanin-based colour categorization [25].

Colour prediction is more accurate using raw spectral data rather than spectral parameters (table 1; figure 5; electronic supplementary material, table S4). We conclude, therefore, that raw spectral data represent the optimal tool for future studies aiming to link feather colour to melanin content

(including melanosome attributes). These data will be essential for assessments of the accuracy of predicted reflectance spectra for fossil feathers with specific melanin attributes and assignment to a statistically valid, visually defined colour category. A similar model for predicting reflectance spectra from melanosome geometry has been attempted recently for fossil mammaliaforms [37]. Previous models for predicting fossil feather colour have included various melanosome morphological attributes [9,10,14], and differences in melanin monomer [5,25] and trace element [38,39] chemistry have also been linked to variation in feather colour. Future research should use statistically supported and quantitative colour categories to more confidently relate colour to specific aspects of melanosome geometry and chemistry. The availability of published datasets for which melanosome geometry and melanin chemistry data are labelled by colour category [9,10,14,25] already provides the opportunity to integrate colour category validation into future modelling efforts.

Colour prediction models based on spectral parameters have lower percentage accuracy than models using raw spectral data (table 1). This is not unexpected, as summary statistics will naturally result in loss of information compared to the data from which they derive. The relatively high accuracy of the models using spectral parameters, however, indicates that where full spectral data are unavailable or certain aspects of colour are to be investigated, a combination of parameters may be an acceptable alternative. This may allow the acquisition of data from high-quality photographs, for example, where parameters such as total brightness, maximum brightness, saturation and RGB values may be extracted from coloured areas. The two optimal combinations (total brightness, maximum brightness, red and green; total brightness, saturation, red and green) identified show no significant difference in accuracy compared to when all parameters are used. This is surprising as slope is absent from both parameter combinations, yet all pairs of colour categories differ significantly in slope (figure 3; electronic supplementary material, table S3). The two optimal parameter datasets produce slightly higher accuracies (table 1) than for smoothed raw spectral data (figure 5C), but the reason for this is unclear.

Our dataset has the potential to be expanded in several ways. The colour categories used in this study do not include certain categories used in previous studies of melanin-based fossil feather colour, such as iridescent and non-iridescent structural colour [9,10]. Reflectance spectra for these feathers differ strongly from those for feathers with non-structural melanin-based colour [1,40,41]. The categories of iridescent and non-iridescent structural colour encompass an extensive variety of hues [1], and in the case of iridescence, various optical effects [4], where the reflectance spectrum varies depending on the angle of the spectrophotometer probe to the feather surface. This may reduce the accuracy of colour prediction models. It is also important to note that feather coloration is not optimized for perception by human (trichromat) colour vision, and avian (tetrachromat) colour vision includes a UV component [4]. Future studies of melanin-based feather colour could potentially include the UV-A component of the spectrum (315–700 nm) and could examine the impact of this on the accuracy of colour prediction.

To summarize, this study provides a statistically robust method for quantitative analysis of melanin-based colours in feathers. By applying this approach to assess the validity of certain visually assigned colour categories, we confirm which categories are statistically discernible. Given that melanin-based colours are highly evolutionarily conserved [6], our findings may also be applicable to other tissues in birds and other animal groups that are coloured primarily by melanins, including skin, scales and hair. Our results have clear implications for studies of fossil feather colour: spectral datasets for feathers should divide rufous colours into light and dark subcategories and should provide statistical validation for visually assigned colour categories. These protocols will help constrain the accuracy of models for interpreting fossil feather colour and may have broader applications for colour reconstructions in the fossil record.

Ethics. This work did not require ethical approval from a human subject or animal welfare committee.

Data accessibility. All spectral and statistical data are available in the supplementary data [42].

Declaration of AI use. We have not used AI-assisted technologies in creating this article.

Authors' contributions. H.B.: conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, writing—original draft, writing—review and editing; D.F.: resources, writing—review and editing; M.E.McN.: conceptualization, funding acquisition, supervision, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. This study was funded by Research Ireland Postgraduate Scholarship GOIPG/2022/1231 awarded to H.B. and by ERC Consolidator Grant H2020-ERC-COG-101003293-PALAEOCHEM awarded to M.E.McN. D.F. was funded by UKRI Research and Innovation Grant MR/X015130/1.

Acknowledgements. We thank Mathew Lowe (Museum of Zoology, University of Cambridge) for valuable assistance with bird skin collections, Sean McKeown and Declan O'Donovan (Fota Wildlife Park, Cork) for supplying ostrich and rhea feathers, Kathleen O'Sullivan (University College Cork) for helpful PCA discussion and Naomi O'Reilly and Fay Clohessy (University College Cork) for research support. We thank Evan Saitta and one other anonymous reviewer for their helpful feedback on the manuscript.

References

1. Stoddard MC, Prum RO. 2011 How colorful are birds? Evolution of the avian plumage color gamut. *Behav. Ecol*. **22**, 1042–1052. (doi:10.1093/beheco/arr088)
2. Hill GE, McGraw KJ (eds). 2006 *Bird coloration (volume 2): function and evolution*. Cambridge, MA: Harvard University Press. (doi:10.2307/j.ctv22jnr8k)
3. Shawkey MD, D'Alba L. 2017 Interactions between colour-producing mechanisms and their effects on the integumentary colour palette. *Phil. Trans. R. Soc. B* **372**, 20160536. (doi:10.1098/rstb.2016.0536)
4. Hill GE, McGraw KJ (eds). 2006 *Bird coloration (volume 1): mechanisms and measurements*. Cambridge, MA: Harvard University Press. (doi:10.4159/9780674273788)
5. Slater TS, Ito S, Wakamatsu K, Zhang F, Sjövall P, Jarenmark M, Lindgren J, McNamara ME. 2023 Taphonomic experiments reveal authentic molecular signals for fossil melanins and verify preservation of phaeomelanin in fossils. *Nat. Commun*. **14**, 5651. (doi:10.1038/s41467-023-40570-w)
6. McNamara ME, Rossi V, Slater TS, Rogers CS, Ducrest AL, Dubey S, Roulin A. 2021 Decoding the evolution of melanin in vertebrates. *Trends Ecol. Evol.* **36**, 430–443. (doi:10.1016/j.tree.2020.12.012)
7. Price-Waldman RM, Ali JR, Shultz AJ, Hogan BG, Stoddard MC. 2025 Hidden white and black feather layers enhance plumage coloration in tanagers and other songbirds. *Sci. Adv*. **11**, 5857. (doi:10.1126/sciadv.adw5857)
8. Shawkey MD, Hill GE. 2006 Significance of a basal melanin layer to production of non-iridescent structural plumage color: evidence from an amelanotic Steller's jay (*Cyanocitta stelleri*). *J. Exp. Biol*. **209**, 1245–1250. (doi:10.1242/jeb.02115)
9. Li Q *et al*. 2012 Reconstruction of *Microraptor* and the evolution of iridescent plumage. *Science* **335**, 1215–1219. (doi:10.1126/science.1213780)
10. Babarović F, Puttick MN, Zaher M, Learmonth E, Gallimore EJ, Smithwick FM, Mayr G, Vinther J. 2019 Characterization of melanosomes involved in the production of non-iridescent structural feather colours and their detection in the fossil record. *J. R. Soc. Interface* **16**, 20180921. (doi:10.1098/rsif.2018.0921)
11. Nordén KK *et al*. 2019 Melanosome diversity and convergence in the evolution of iridescent avian feathers - implications for paleocolor reconstruction. *Evolution* **73**, 15–27. (doi:10.1111/evo.13641)
12. Clarke JA, Ksepka DT, Salas-Gismondi R, Altamirano AJ, Shawkey MD, D'Alba L, Vinther J, DeVries TJ, Baby P. 2010 Fossil evidence for evolution of the shape and color of penguin feathers. *Science* **330**, 954–957. (doi:10.1126/science.1193604)
13. Carney RM, Vinther J, Shawkey MD, D'Alba L, Ackermann J. 2012 New evidence on the colour and nature of the isolated *Archaeopteryx* feather. *Nat. Commun*. **3**, 637. (doi:10.1038/ncomms1642)
14. Li Q, Gao KQ, Vinther J, Shawkey MD, Clarke JA, D'Alba L, Meng Q, Briggs DEG, Prum RO. 2010 Plumage color patterns of an extinct dinosaur. *Science* **327**, 1369–1372. (doi:10.1126/science.1186290)
15. Zhang F, Kearns SL, Orr PJ, Benton MJ, Zhou Z, Johnson D, Xu X, Wang X. 2010 Fossilized melanosomes and the colour of Cretaceous dinosaurs and birds. *Nature* **463**, 1075–1078. (doi:10.1038/nature08740)
16. Pinheiro FL, Prado G, Ito S, Simon JD, Wakamatsu K, Anelli LE, Andrade JA, Glass K. 2019 Chemical characterization of pterosaur melanin challenges color inferences in extinct animals. *Sci. Rep*. **9**, 15947. (doi:10.1038/s41598-019-52318-y)
17. Yang Z, Jiang B, McNamara ME, Kearns SL, Pittman M, Kaye TG, Orr PJ, Xu X, Benton MJ. 2019 Pterosaur integumentary structures with complex feather-like branching. *Nat. Ecol. Evol*. **3**, 24–30. (doi:10.1038/s41559-018-0728-7)
18. Cincotta A *et al*. 2022 Pterosaur melanosomes support signalling functions for early feathers. *Nature* **604**, 684–688. (doi:10.1038/s41586-022-04622-3)
19. Vinther J, Briggs DEG, Prum RO, Saranathan V. 2008 The colour of fossil feathers. *Biol. Lett*. **4**, 522–525. (doi:10.1098/rsbl.2008.0302)
20. Li Q, Clarke JA, Gao KQ, Zhou CF, Meng Q, Li D, D'Alba L, Shawkey MD. 2014 Melanosome evolution indicates a key physiological shift within feathered dinosaurs. *Nature* **507**, 350–353. (doi:10.1038/nature12973)
21. Peteya JA, Clarke JA, Li Q, Gao KQ, Shawkey MD. 2017 The plumage and colouration of an enantiornithine bird from the Early Cretaceous of China. Palaeontology **60**, 55–71. (doi:10.1111/pala.12270)
22. Smithwick FM, Nicholls R, Cuthill IC, Vinther J. 2017 Countershading and stripes in the theropod dinosaur *Sinosauropteryx* reveal heterogeneous habitats in the Early Cretaceous Jehol biota. *Curr. Biol*. **27**, 3337–3343.(doi:10.1016/j.cub.2017.09.032)

23. Pan Y, Li Z, Wang M, Zhao T, Wang X, Zheng X. 2022 Unambiguous evidence of brilliant iridescent feather color from hollow melanosomes in an Early Cretaceous bird. *Natl Sci. Rev*. **9**, 227. (doi:10.1093/nsr/nwab227)
24. Emery KJ, Webster MA. 2019 Individual differences and their implications for color perception. *Curr. Opin. Behav. Sci*. **30**, 28–33. (doi:10.1016/j.cobeha.2019.05.002)
25. Galván I, Wakamatsu K. 2016 Color measurement of the animal integument predicts the content of specific melanin forms. *RSC Adv*. **6**, 79135–79142. (doi:10.1039/c6ra17463a)
26. Cuthill IC, Bennett ATD, Partridge JC, Maier EJ. 1999 Plumage reflectance and the objective assessment of avian sexual dichromatism. *Am. Nat*. **153**, 183–200. (doi:10.1086/303160)
27. Eliason CM, Bitton PP, Shawkey MD. 2013 How hollow melanosomes affect iridescent colour production in birds. *Proc. R. Soc. B* **280**, 20131505. (doi:10.1098/rspb.2013.1505)
28. Eliason CM, Shawkey MD, Clarke JA. 2016 Evolutionary shifts in the melanin-based color system of birds. *Evolution* **70**, 445–455. (doi:10.1111/evo.12855)
29. Freyer P, Wilts BD, Stavenga DG. 2021 Cortex thickness is key for the colors of iridescent starling feather barbules with a single, organized melanosome layer. *Front. Ecol. Evol*. **9**, 746254. (doi:10.3389/fevo.2021.746254)
30. Nordén KK, Eliason CM, Stoddard MC. 2021 Evolution of brilliant iridescent feather nanostructures. *eLife* **10**, 71179. (doi:10.7554/elife.71179)
31. Rogalla S, Patil A, Dhinojwala A, Shawkey MD, D'Alba L. 2021 Enhanced photothermal absorption in iridescent feathers. *J. R. Soc. Interface* **18**, 20210252. (doi:10.1098/rsif.2021.0252)
32. Jeon DJ, Ji S, Lee E, Kang J, Kim J, D'Alba L, Manceau M, Shawkey MD, Yeo JS. 2023 How keratin cortex thickness affects iridescent feather colours. *R. Soc. Open Sci*. **10**, 220786. (doi:10.1098/rsos.220786)
33. McCoy DE, Feo T, Harvey TA, Prum RO. 2018 Structural absorption by barbule microstructures of super black bird of paradise feathers. *Nat. Commun*. **9**, 1. (doi:10.1038/s41467-017-02088-w)
34. R Core Team. 2023 *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. See https://www.R-project.org.
35. Maia R, Gruson H, Endler JA, White TE. 2019 pavo2: new tools for the spectral and spatial analysis of colour in R. *Methods Ecol. Evol*. **10**, 1097–1107. (doi:10.1111/2041-210x.13174)
36. Field DJ, D'Alba L, Vinther J, Webb SM, Gearty W, Shawkey MD. 2013 Melanin concentration gradients in modern and fossil feathers. *PLoS ONE* **8**, e59451. (doi:10.1371/journal.pone.0059451)
37. Li R, D'Alba L, Debruyn G, Dobson JL, Zhou CF, Clarke JA, Vinther J, Li Q, Shawkey MD. 2025 Mesozoic mammaliaforms illuminate the origins of pelage coloration. *Science* **387**, 1193–1198. (doi:10.1126/science.ads9734)
38. Manning PL *et al*. 2013 Synchrotron-based chemical imaging reveals plumage patterns in a 150 million year old early bird. *J. Anal. At. Spectrom*. **28**, 1024–1030. (doi:10.1039/c3ja50077b)
39. Edwards NP *et al*. 2016 Elemental characterisation of melanin in feathers via synchrotron X-ray imaging and absorption spectroscopy. *Sci. Rep*. **6**, 34002. (doi:10.1038/srep34002)
40. McGraw KJ. 2004 Multiple UV reflectance peaks in the iridescent neck feathers of pigeons. *Naturwissenschaften* **91**, 125–129. (doi:10.1007/s00114-003-0498-0)
41. Freyer P, Wilts BD, Stavenga DG. 2018 Reflections on iridescent neck and breast feathers of the peacock, *Pavo cristatus*. *Interface Focus* **9**, 20180043. (doi:10.1098/rsfs.2018.0043)
42. Bean H, Field D, McNamara ME. 2025 Supplementary material from: Quantitative characterisation of melanin-based feather colour using reflectance spectrophotometry. FigShare (doi:10.6084/m9.figshare.c.8153363)