Capstone: A Custom *Carex* Color Palette

Noah Giebink

5/8/2021

ABSTRACT

Color is a conspicuous and variable trait across plants, yet it is challenging to glean taxonomic insight from color descriptions when the language used to describe color traits is inconsistent. The objective of the present study is to encourage a controlled color vocabulary by building a custom color palette tool which helps authors choose standardized color descriptions for plant specimens. Focusing on the genus *Carex*, I integrated data from author descriptions in Flora of North America Vol. 23 with corresponding ground truth color measurements to map real-life quantitative color values with the language used to describe them. Color descriptions fell within a small set of frequent color classes: brown, green, red, white, yellow-green, and yellow-brown. However, the striking overlap in the measured color values for samples given each of these names by authors further necessitates a controlled color vocabulary. Given color values and the color class for a subset of samples, I compared manual and machine learning approaches for predicting color class boundaries and present a proposed color palette for each.

INTRODUCTION

Plant color traits are of widespread research interest due to their significance to the ecology and evolution of plants, from mediating pollinator interactions to thermoregulation (Rosas-Guerrero et al., 2014; Dick et al., 2011). Though strides have been made to encourage controlled vocabularies for plants which would drastically increase our ability to leverage computational analyses of taxonomic descriptions (Endara et al., 2017), the inconsistency and subjectivity of color description hinders progress.

Questions For the present work, I addressed three major questions which are key to developing a controlled color vocabulary:

- *Q1. Can we create a custom color palette which uses both existing color description language and real values from* Carex *spp. sample measurements?*
- *Q2. Is a manual approach best for creating a palette or is it possible to automate the process via machine learning?*
- *Q3. How do existing author descriptions of colors compare with the visual appearance of measured colors?*

A key step to answering *Q1* is finding the appropriate color space representation for our sample measurements in order to distinguish colors in a way that is perceptually meaningful to users. While sRGB is a commonplace quantitative representation for colors, it is more suited to interpretation by machines than for capturing perceptual differences between colors. CIE Lab space, referred to as L*a*b* is specifically useful for encoding small differences between colors across axes that are relevant to human color perception. Axes a* and b* are based on the opponent process theory of color vision, where red opposes green and yellow opposes blue (Hurvich & Jameson, 1957). Centered at zero, negative values of a* are more green and positive values are red; negative values of b* are more blue and positive values are more yellow. color values in L*a*b* space are independent of lightness, encoded separately in L*. This marks a critical contrast with sRGB space in which color value cannot be disentangled from lightness. This is practically challenging for my purpose since similar colors may be distant in color space merely due to difference in brightness, not class membership.

I investigated the efficacy of representing real sample colors in L*a*b* space compared to sRGB space, for use in both manual and automated approaches to developing a custom *Carex* color palette. Finally, I used the color class boundaries predicted by these approaches to visualize incongruities in the existing, uncontrolled vocabulary for color descriptions.

METHODS

Data Preparation

Data Sources I joined two data sources: species level color descriptions from the Flora of North America Vol. 23 (henceforth “author descriptions”) (FNA; www.eflora.org.), and species to subspecies level ground truth, sRGB channel color measurements of specimen images (henceforth color measurements“). Specimen images for a total of 456 species were curated, mostly from online data sharing platforms (e.g. SeiNet, Intermountain Regional Herbarium Network). Specimens were filtered for quality such that all were high resolution, lacked preventable discoloration, and had a color control reference. For species whose specimens failed to meet these standards, original images were captured the herbaria of the Canadian Museum of Nature (CAN), Agriculture and Agri-Food Canada (DAO), or the Marie-Victorin Herbarium (MT). sRGB values were taken from each specimen over several pixels using ImageJ (ver. 1.50i; Schindelin et al., 2012). Author descriptions were specific to tissues including leaf, male scale, perigynium, female scale, and cataphyll. For both author descriptions and color measurements, I discarded observations from male and female scales because these were the average of the margin, blade, and axis and were thus not”real" colors shown on the specimen. I used the joined data set with both author descriptions and color measurements to manually sort colors into classes and further partitioned the data into train and test sets to train support vector machines for an automated approach. Table 1 includes sample sizes of each data set used.

Table 1: Sample size for each data set.

Data Set	Observations
Color Measurements	2883
Author Descriptions	956
Train	631
Train Down Sampled	215
Test	161

Color Classes I mined author description texts to identify the color classes present (R package ‘tm’; Feinerer et al., 2008). I first found the most frequent single-word color terms, not including stop words, used in author descriptions. I selected the most frequent terms (classes) and manually grouped the remaining terms with these classes as synonyms for their respective class, according to my subjective judgement of equivalence (Table 2). I validated these color classes and synonym groupings by subjecting them to review by three *Carex* experts.

I converted each author description string into a single color class (henceforth “author label”) based on the color class terms and synonyms present in the string. This wass a two-step process applied to each class and its synonyms: first, I filtered observations containing any one of the synonyms in a color class; second, I filtered the result of the first step to contain only the observations unique to each color class so that no observation had more than one color class. The result was a single color class label for each observation.

Color Space Transformation I engineered features in L*a*b* space from sRGB input values using the convertColor function from the built-in grDevices package (R Core Team, 2021).

Manual Approach - Thresholds I relied on natural breaks in the distribution of samples along a* and b* axes to set thresholds which divide colors. I continued to manually separate the data this way into sequentially smaller bins until they resembled the color classes brown, green, red, white, yellow-brown, and yellow-green. I visualize the full sequence of divisions in Figures 1-4. I color-code figures 1-4 and 6-8 with the actual values of samples which represent that respective color class or distribution of colors. I found each representative "centroid" by taking the mean a* - b* value of a set of samples, then calculating the nearest neighbor to the mean, an actual measurement. A limitation of this method is that it can only divide samples (create "thresholds") along vertical and horizontal lines: intercepts along the a* and b* axes, and rectangles formed via the combination of these thresholds.

Support Vector Machine Approach To overcome the limited division lines limitation of the manual thresholds approach, and because the author label data–while noisy–appear to be linearly or nearly linearly separable (Figure 6, bottom right), I trained two SVM models to predict color class (Meyer et al., 2021), one with the default radial kernel and the other linear. To overcome the class imbalance, I down sampled the data so that all classes had the same n, equal to the smallest class (Caret R package; Kuhn, 2020). Finally, I increased the regularization hyperparameter C from default 1 to 0.5 to penalize the model less for misclassifying noisy points. Finally, I calculated class-wise precision, recall, and f1 scores for both models and selected the model with the greatest mean f1 for the ultimate color palette.

Table 2: Class-wise sample sizes for developing thresholds ('Full Labeled Set') and SVM models ('Train Down Sampled (no white)'). SVM models and thresholds performance were assessed with a common test set (Test), except class 'white' was excluded from SVM model assessment.

set	brown	green	red	white	yellow-brown	yellow-green
Full Labeled Set	431	133	65	12	97	54
Train	344	106	52	9	77	43
Train Down Sampled (no white)	43	43	43	NA	43	43
Test	87	27	13	3	20	11

Final Color Palette I split the full data set (n = 2883; see Table 1) along color class thresholds defined via the manual thresholds method and predicted color class from a* and b* features using the best performing SVM. I further separated each resulting color class into light, medium, and dark using k-means on the L* variable, with k=3. To choose a reasonable number of color samples for users to select from and visualize, I sampled five values in each lightness level of each color class predicted by each approach. I down sampled using a custom algorithm which calculates the mean distance to k nearest neighbors for each point ("mean_dist"; default k = 3 used here), and samples the points with probability equal to the normalized mean_dist such that points with lower mean_dist (i.e. nearer neighbors) are less likely to be sampled. This allowed me to select a sample of colors to present in the palette which should be less redundant perceptually than a simple random sample.

RESULTS and DISCUSSION

Q1: Custom Color Palette The Red and Green axes of sRGB space contained most of the variance in color values (Figure 5). This may be due to the fact that natural vegetative tissue on plant specimens, including *Carex* spp., is unlikely to be blue. Therefore, I visually compared the differences in representation between R-G and a*-b* spaces (Figure 6). Representing color samples in L*a*b* space led to a greater spread in values and improved the linear separability of the color classes. Transforming sRGB measurements into perceptually-oriented L*a*b* space made it feasible to create a custom color palette from existing color language and ground-truth *Carex* color measurements.

Q2: Manual vs. Automated Approach The linear SVM had the best performance against the author labels in the test set (multiclass mean f1 = 0.4757042), followed by the radial SVM (multiclass mean f1 = 0.4573148), and the manual thresholds approach (mean multiclass f1 = 0.3998657 factoring out performance on 'white' class for equal comparison; else mean multiclass f1 = 0.388777). Detailed multiclass precision, recall, and f1 scores are summarized in tables 3 and 4.

Table 3: Class-wise precision, recall, and f1 scores for Linear SVM

	precision	recall	f1
brown	0.8709677	0.6206897	0.7248322
green	0.5500000	0.4074074	0.4680851
red	0.4333333	1.0000000	0.6046512
yellow-brown	0.2272727	0.2500000	0.2380952
yellow-green	0.2500000	0.5454545	0.3428571

Table 4: Class-wise precision, recall, and f1 scores for thresholds

	precision	recall	f1
brown	0.7261905	0.7011494	0.7134503
green	0.4285714	0.3333333	0.3750000
red	0.3500000	0.5384615	0.4242424
white	0.3333333	0.3333333	0.3333333
yellow-brown	0.2857143	0.1904762	0.2285714
yellow-green	0.2000000	0.3636364	0.2580645

While the SVM had greater performance classifying authors' labels (possibly because the manual approach was limited to horizontal and vertical lines in color space), the manual thresholds approach appears qualitatively at least as effective (Figure 9). This qualitative improvement may be due to the fact that the color language used by authors is inconsistent to begin with. So, while the SVM stays truer to author descriptions, my manual thresholds approach may be a better representation of true colors. However, the SVM provides the benefit of automation, speed, and easier reproducibility. Thus, I present two feasible methods for designing a sample-based color palette for a controlled vocabulary which can be employed based on the needs and goals of the curator.

Q3: Author Descriptions vs. Appearance Notably, the thresholds method did not rely on learning the author labels; matches between author labels and classes via the thresholds approach appeared when author labels agreed with my color designations. This justified using color class predictions from the thresholds approach to visually inspect the consistency of author descriptions, measured by the frequency that each author label appeared in the "correct" color class (Figure 8). Qualitatively, there is a great deal of overlap in color classes according to author labels in both R-G and a*b* space (Figure 7). Qualitatively, brown, green, and white contained majority "correct" author labels, whereas for red, yellow-brown, and yellow-green, the majority author label was "incorrect." While some variability of author labels is expected when pit against a color classification scheme that is not perfectly congruent with author labels (a model perfectly accurate at predicting author labels would show the author label to be correct each time), some of the deviance from predicted color class is certainly attributable to the striking overlap in the measured values of samples in each author label (see figure 7). Regardless, the variety in color descriptions given to visually similar colors necessitates a controlled color vocabulary.

Future Work The next stage will be to integrate the color palette produced here with web ontogeny software. Then, we will perform usability experiments and test whether the ability to make selections from a color palette containing a subset of real sample values (as shown here) improves consistency of color descriptions.

CODE AND REPRODUCIBILITY

All code and data necessary to reproduce this work can be found in this GitHub Repository.

LITERATURE CITED

Dick, C. A., Buenrostro, J., Butler, T., Carlson, M. L., Kliebenstein, D. J., & Whittall, J. B. (2011). Arctic Mustard Flower Color Polymorphism Controlled by Petal-Specific Downregulation at the Threshold of the Anthocyanin Biosynthetic Pathway. PLOS ONE, 6(4), e18230. https://doi.org/10.1371/journal.pone.0018230

Endara, L., Cole, H. A., Burleigh, J. G., Nagalingum, N. S., Macklin, J. A., Liu, J., Ranade, S., & Cui, H. (2017). Building the "Plant Glossary"—A controlled botanical vocabulary using terms extracted from the Floras of North America and China. Taxon, 66(4), 953–966. https://doi.org/10.12705/664.9

Feinerer, I. & K. Hornik (2020). tm: Text Mining Package. R package version 0.7-8. https://CRAN.R-project.org/package=tm

Hurvich, L. M., & Jameson, D. (1957). AN OPPONENT-PROCESS THEORY OF COLOR VISION. Psychological Review, 64(6), 21.

Kuhn, M. (2020). caret: Classification and Regression Training. R package version 6.0-86. https://CRAN.R-project.org/package=caret

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., and F. Leisch (2021). e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien. R package version 1.7-6. https://CRAN.R-project.org/package=e1071

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rosas-Guerrero, V., Aguilar, R., Martén-Rodríguez, S., Ashworth, L., Lopezaraiza-Mikel, M., Bastida, J. M., & Quesada, M. (2014). A quantitative review of pollination syndromes: Do floral traits predict effective pollinators? Ecology Letters, 17(3), 388–400. https://doi.org/10.1111/ele.12224

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., & Cardona, A. (2012). Fiji: An open-source platform for biological-image analysis. Nature Methods, 9(7), 676–682. https://doi.org/10.1038/nmeth.2019

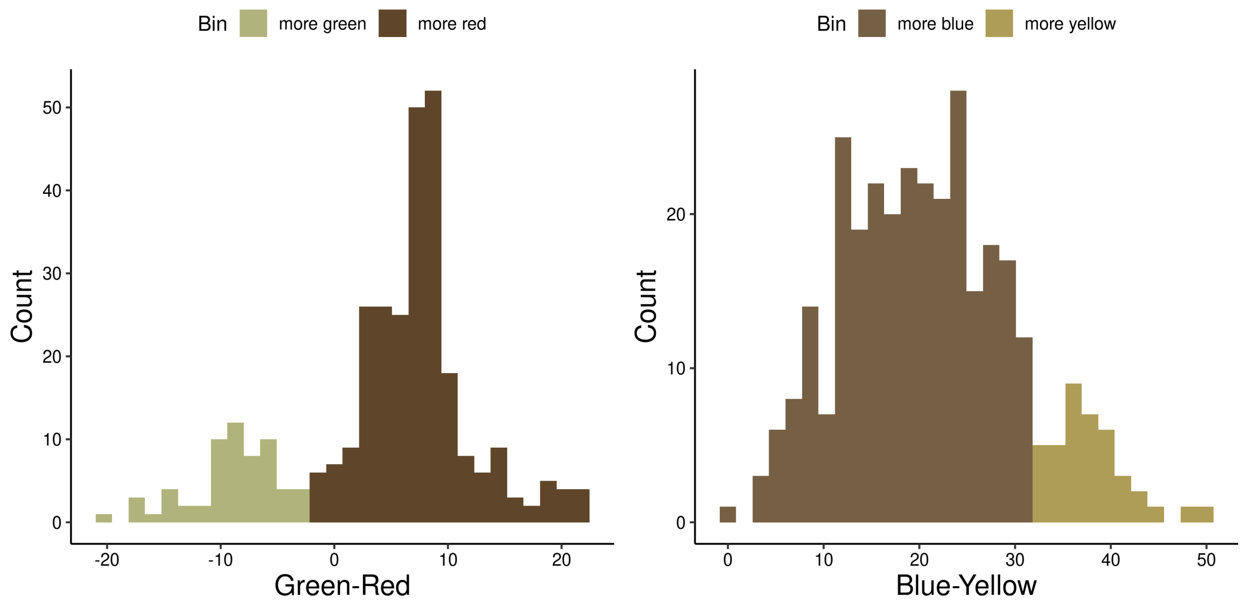

Figure 1: From left: Green-Red and Blue-Yellow divisions in the whole data set along natural breaks in distributions of color samples along a* (green-red) and b* (blue-yellow) axes of L*a*b*. Portions of the distributions are colored according to the sRGB values of centroid samples for corresponding to that portion of the distribution (see methods).

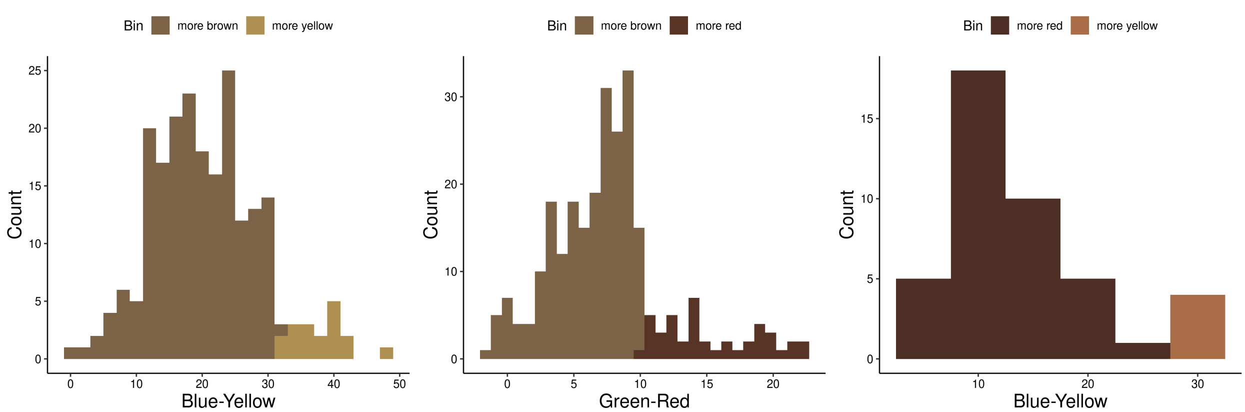

Figure 2: From left: brown-yellow-brown, brown-red, and red-brown (brown appears in two regions of L*a*b* space) divisions among a* values greater than -2 (after splitting on natural break in Figure 1, left). Divisions appear along natural breaks in distributions of color samples along a* (green-red) and b* (blue-yellow) axes of L*a*b* space. Portions of the distributions are colored according to the sRGB values of centroid samples for corresponding to that portion of the distribution (see methods).

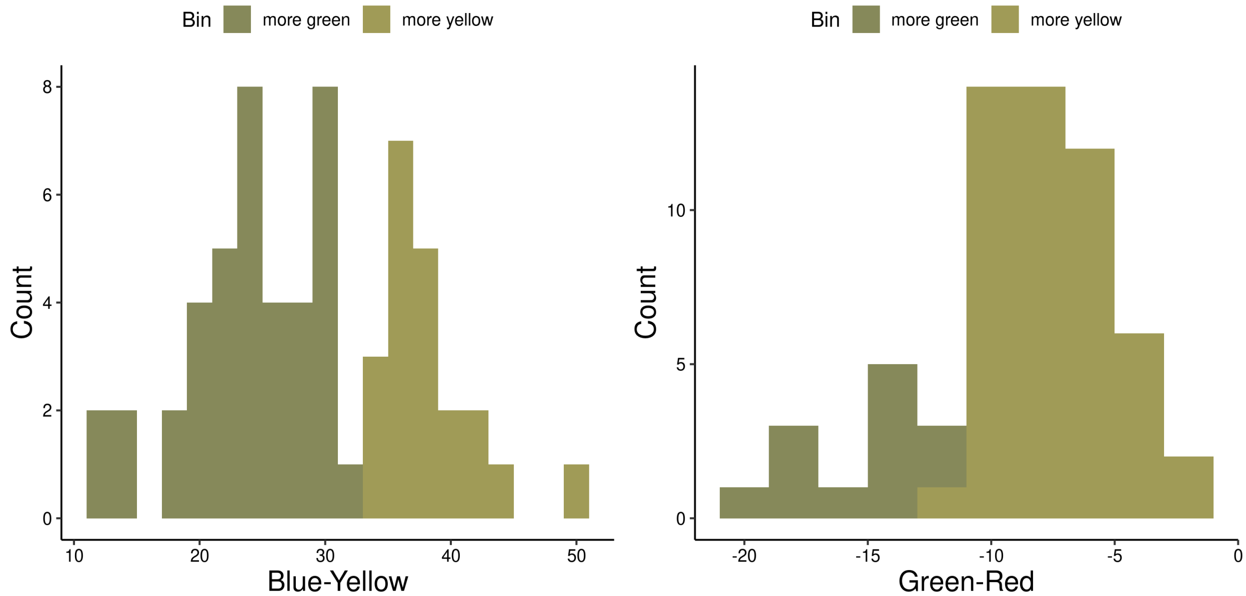

Figure 3: From left: green-yellow-green and yellow-green-green (green appears in two regions of L*a*b* space) divions among a* values less than -2 (after splitting on natural break in Figure 1, left). Divisions appear along natural breaks in distributions of color samples along a* (green-red) and b* (blue-yellow) axes of L*a*b* space. Portions of the distributions are colored according to the sRGB values of centroid samples for corresponding to that portion of the distribution (see methods).

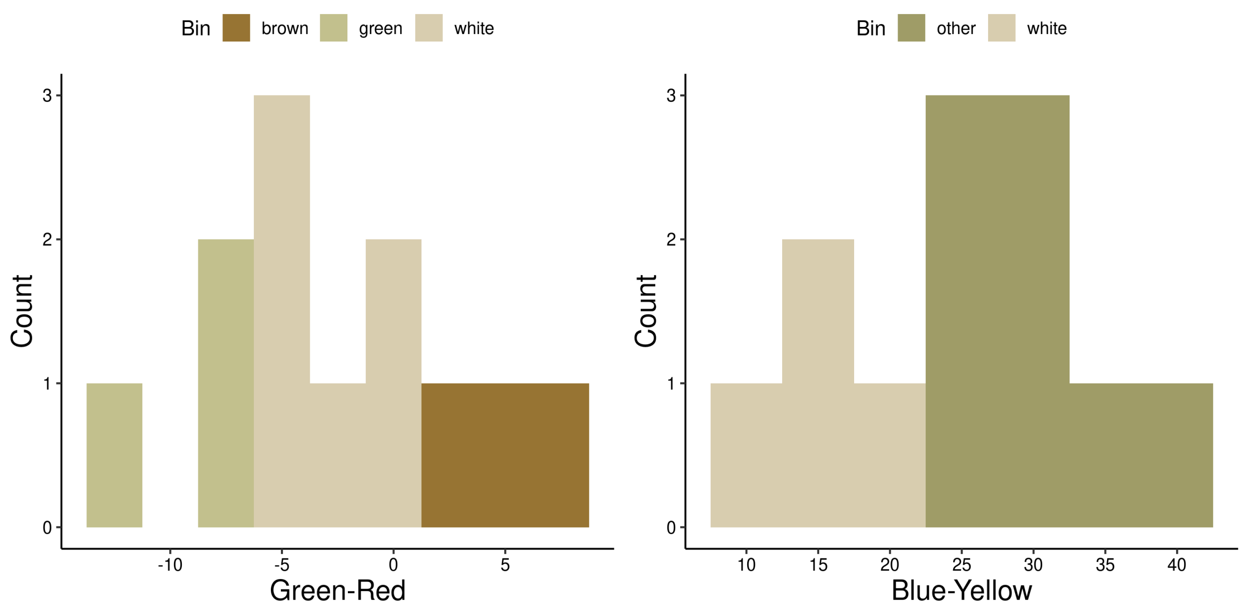

Figure 4: From left: green-white-brown and white-green divions among samples with author label "white." Divisions appear along natural breaks in distributions of color samples along a* (green-red) and b* (blue-yellow) axes of L*a*b* space. Portions of the distributions are colored according to the sRGB values of centroid samples for corresponding to that portion of the distribution (see methods).

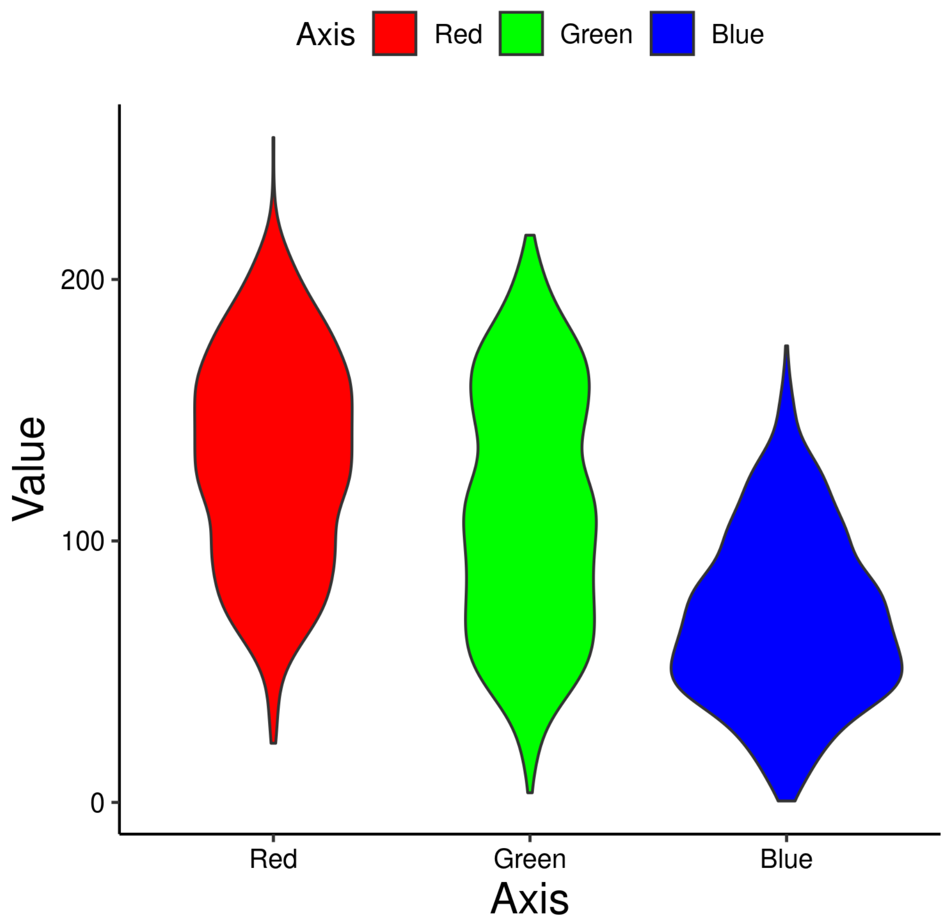

Figure 5: Red and Green axes in sRGB space contain most of the variance in color values.

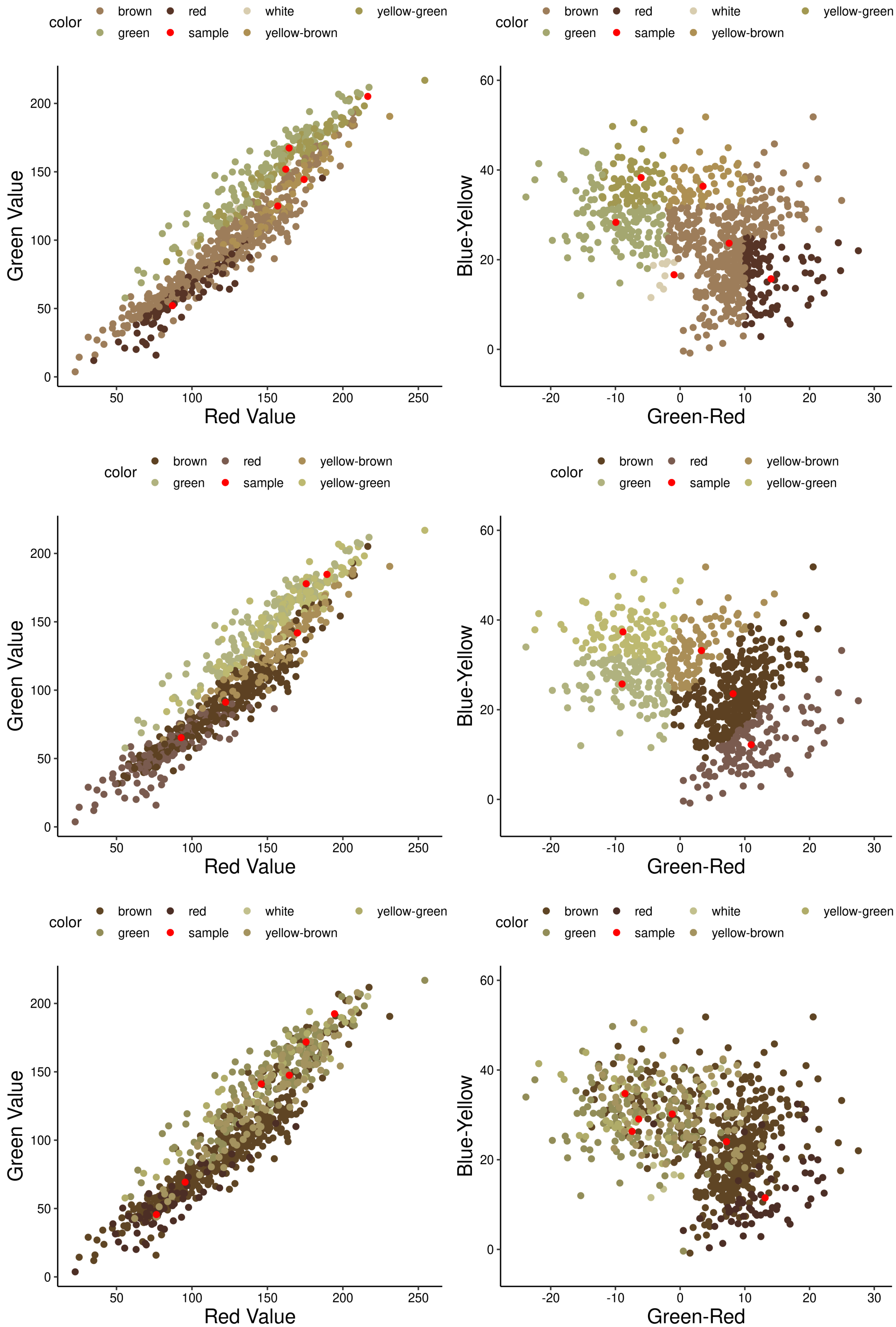

Figure 6: Points are colored according to the sRGB values of the centroid point for their class (see methods). The centroid point used for each class is indicated in red.

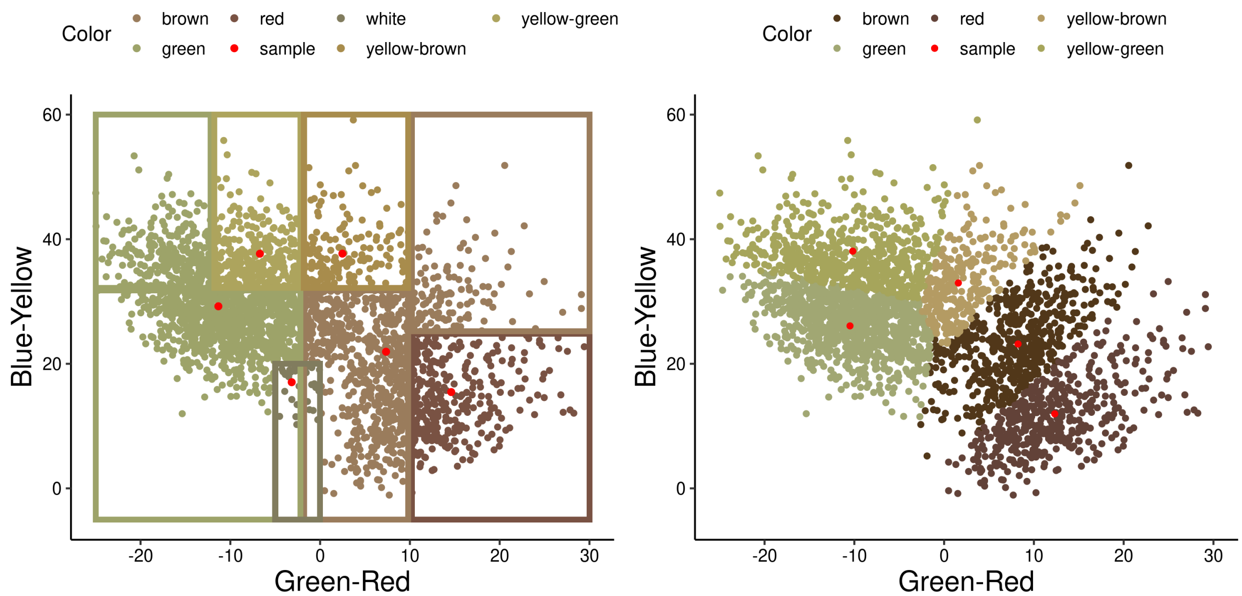

Figure 7: Color spaces reflecting spatial distribution of each class, predicted on the full data set by the thresholds approach (left) and a linear SVM (right) (n = 2883; see Table 1). Thresholds are shown as rectangles, colored according to the class they separate. Points are colored according to the sRGB values of the centroid point for their class (see methods). The centroid point used for each class is indicated in red.

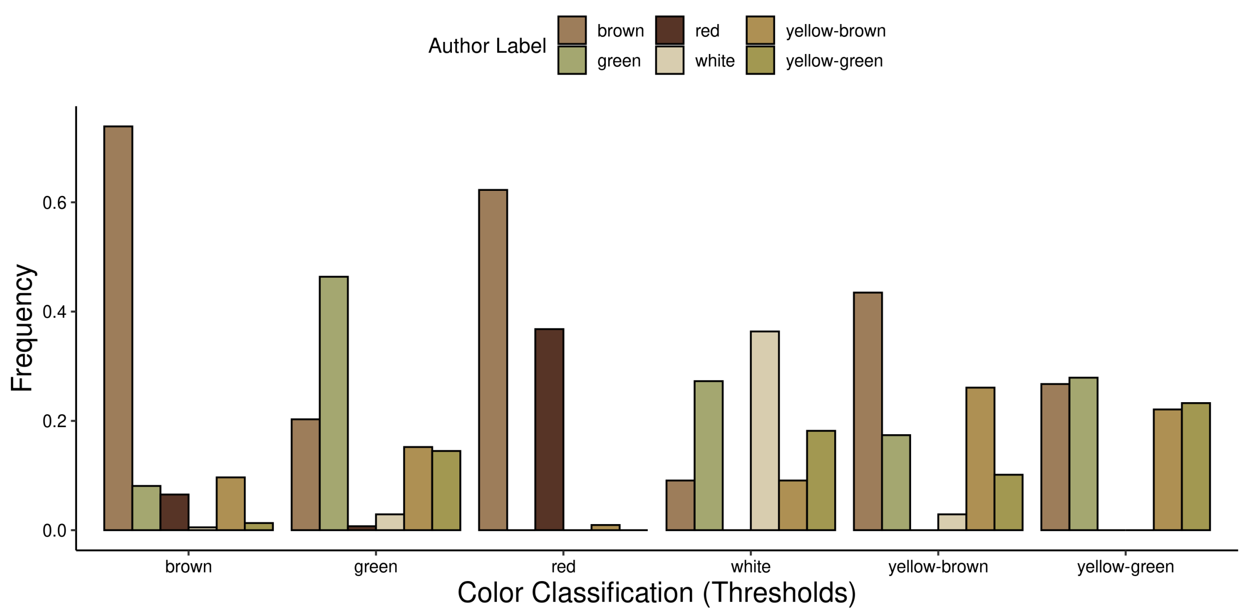

Figure 8: Frequency of author labels given to each color class, as classified via the thresholds approach.

	Brown	Green	Red	White	Yellow-brown	Yellow-green
1	Carex picta Perigynium_709	Carex cephaloidea Perigynium_173	Carex flava Cataphyll_273	Carex garberi Perigynium_341	Carex engelmannii Leaf_298	Carex tetanica Perigynium_913
2	Carex ursina Perigynium_962	Carex bicknellii Leaf_115	Carex pelocarpa Cataphyll_2_11	Carex elliottii Leaf_291	Carex californica Perigynium_158	Carex athrostachya Perigynium_77
3	Carex ramenskii Perigynium_755	Carex leptalea Leaf_508	Carex roanensis Cataphyll_658	Carex paniculata Leaf_675	Carex densa Perigynium_256	Carex adusta Perigynium_13
4	Carex tenuiflora Perigynium_911	Carex luzulina Leaf_538	Carex sartwelliana Cataphyll_2_4	Carex pluriflora Perigynium_723	Carex prairea Perigynium_734	Carex filifolia Leaf_310
5	Carex crawei Perigynium_213	Carex luzulifolia Leaf_536	Carex glacialis Cataphyll_297	Carex buxbaumii Perigynium_153	Carex muriculata Perigynium_607	Carex polymorpha Perigynium_728
6	Carex willdenowii Cataphyll_842	Carex viridula Leaf_990	Carex buxbaumii Cataphyll_132	Carex mackenziei Perigynium_544	Carex microdonta Perigynium_581	Carex collinsii Perigynium_190
7	Carex disperma Cataphyll_230	Carex tetrastachya Leaf_917	Carex gynandra Cataphyll_323	Carex garberi Perigynium_340	Carex sheldonii Perigynium_833	Carex laxiculmus Perigynium_488
8	Carex rossii Perigynium_777	Carex lemmonii Perigynium_495	Carex communis Cataphyll_164	Carex lucorum Perigynium_526	Carex typhina Perigynium_956	Carex venusta Leaf_974
9	Carex geophila Cataphyll_289	Carex bigelowii Perigynium_121	Carex deflexa Cataphyll_220	Carex nelsonii Perigynium_616	Carex macloviana Perigynium_546	Carex muskingumensis Perigynium_609
10	Carex conoidea Cataphyll_179	Carex egglestonii Leaf_286	Carex lenticularis Cataphyll_423	Carex paupercula Perigynium_685	Carex sterilis Perigynium_862	Carex feta Leaf_309
11	Carex leptalea Cataphyll_432	Carex eleusinoides Leaf_289	Carex fissa Cataphyll_270	Carex pedunculata Leaf_690	Carex elynoides Cataphyll_252	Carex lachenalii Leaf_467
12	Carex mariposana Cataphyll_471	Carex vestita Leaf_979	Carex oligosperma Cataphyll_552	Carex subspathacea Perigynium_891	Carex atrofusca Cataphyll_76	Carex gracillima Perigynium_369
13	Carex davyi Cataphyll_211	Carex acidicola Leaf_7	Carex arctata Cataphyll_56	Carex spectabilis Perigynium_851	Carex turgescens Perigynium_954	Carex triangularis Perigynium_936
14	Carex glaucodea Cataphyll_301	Carex glaucodea Leaf_357	Carex ramenskii Cataphyll_644	Carex arkansana Perigynium_70	Carex illota Perigynium_436	Carex marina Leaf_554
15	Carex calcifugens Cataphyll_133	Carex exilis Leaf_302	Carex ramenskii Cataphyll_643	Carex bushii Perigynium_151	Carex raynoldsii Perigynium_2_16	Carex scoparia Perigynium_816

	Brown	Green	Red	Yellow-brown	Yellow-green
1	Carex elliottii Leaf_291	Carex willdenowii Perigynium_1001	Carex oligosperma Cataphyll_551	Carex chordorrhiza Perigynium_186	Carex scoparia Leaf_817
2	Carex concinnoides Perigynium_202	Carex muriculata Leaf_606	Carex aquatilis Cataphyll_51	Carex chordorrhiza Perigynium_185	Carex fuliginosa Leaf_339
3	Carex misera Perigynium_590	Carex hassei Leaf_390	Carex gholsonii Cataphyll_294	Carex aquatilis Perigynium_59	Carex capillaris Perigynium_163
4	Carex bromoides Cataphyll_117	Carex neurophora Leaf_618	Carex pelocarpa Cataphyll_2_11	Carex rostrata Perigynium_779	Carex striatula Perigynium_875
5	Carex abrupta Cataphyll_4	Carex densa Leaf_256	Carex membranacea Cataphyll_2_2	Carex laxiculmis Cataphyll_414	Carex fraserianus Perigynium_337
6	Carex senta Perigynium_826	Carex silicea Leaf_843	Carex elliottii Cataphyll_250	Carex bromoides Cataphyll_119	Carex leptonervia Perigynium_511
7	Carex micropoda Cataphyll_493	Carex paeninculae Leaf_2_39	Carex atlantica Cataphyll_71	Carex brunnescens Cataphyll_124	Carex tenuiflora Leaf_914
8	Carex cephalophora Cataphyll_151	Carex sheldonii Leaf_834	Carex bigelowii Cataphyll_103	Carex rostrata Perigynium_780	Carex novae Leaf_632
9	Carex muriculata Cataphyll_514	Carex microchaeta Perigynium_579	Carex scirpoidea Cataphyll_686	Carex phaeocephala Cataphyll_601	Carex crus Perigynium_227
10	Carex bebbii Cataphyll_94	Carex umbellata Leaf_2_24	Carex retrorsa Cataphyll_652	Carex chapmanii Perigynium_180	Carex sheldonii Perigynium_833
11	Carex filifolia Cataphyll_266	Carex micropoda Perigynium_585	Carex lupulina Cataphyll_450	Carex aggregata Cataphyll_17	Carex cumberlandensis Perigynium_232
12	Carex chihuahuensis Cataphyll_158	Carex deflexa Leaf_253	Carex luzulifolia Perigynium_536	Carex turgescens Perigynium_953	Carex stenoptila Leaf_862
13	Carex tumulicola Cataphyll_807	Carex umbellata Leaf_2_23	Carex acidicola Cataphyll_8	Carex tenax Perigynium_905	Carex eleusinoides Leaf_289
14	Carex gracilior Cataphyll_312	Carex bigelowii Leaf_119	Carex tenera Cataphyll_768	Carex vexans Cataphyll_830	Carex digitalis Perigynium_262
15	Carex oreocharis Cataphyll_556	Carex lemmonii Leaf_494	Carex communis Cataphyll_165	Carex sychnocephala Perigynium_900	Carex viridula Leaf_987

Figure 9: Final color palettes produced by the thresholds approach (top) and a linear SVM (bottom).