A New Dataset of Leaf Optical Traits to Include Biophysical Parameters in Addition to Spectral and Biochemical Assessment

Reisha D. Peters[a]* and Scott D. Noble[b]

[a]Chemical and Biological Engineering, reisha.peters@usask.ca [b]Mechanical Engineering, scott.noble@usask.ca

[a,b]College of Engineering, University of Saskatchewan, 57 Campus Drive, Saskatoon, Saskatchewan, S7N 5A9, Canada

*corresponding author

Abstract

To enable future improvement on current leaf optical property models, more data incorporating a larger range of measured properties is needed. To this end, a dataset was collected to associate spectral measurements (ultraviolet, visible, and near infrared) with biochemical and biophysical properties of leaves. The leaves represented in this dataset were selected to provide representation of agricultural species and of leaves with a wide variety of color (pigment) expression, surface characteristics, and age. Data collected for 290 leaf samples studied in this project included multiple spectral measurement geometries and ranges, biochemical assessment of chlorophyll a and b, carotenoids, and anthocyanins, and biophysical assessment of leaf thickness and surface characteristics that has not previously been a focus in other leaf datasets. The methods and results associated with this dataset are described in this work.

Keywords: Leaf Optical Properties, Spectral Measurements, Biophysical Measurements, BRDF, Spectral Reflectance

1 Introduction

Using spectral measurements to assess leaf biochemical traits has been a developing field for many years with applications in remote sensing and precision agriculture. To date, several models have been developed to estimate leaf biochemical properties (perhaps the most prominent

being the PROSPECT model and its derivative models (Jacquemoud and Baret, 1990)). These models can provide valuable non-destructive assessment in a high-throughput format. However, the datasets available for calibration do not necessarily provide an accurate or complete representation of the species being assessed.

A number of leaf datasets have been collected previously and are available through online resources (Hosgood et al., 1995; Jacquemoud et al., 2003) but advancements to data collection and leaf optical modelling has led to a need for expanded species representation and characterization of additional biophysical properties (Baranoski, 2006; Bousquet et al., 2005; Peters et al., 2020). Current leaf datasets are biased towards tree species and have limited representation of non-green leaves (Hosgood et al., 1995; Jacquemoud et al., 2003). Although chlorophyll a and b are collected independently in these datasets, the a to b ratio was reported to be relatively stable resulting in difficulty separating their spectral effects. These datasets also do not contain measured values for anthocyanin but have used measured indices to approximate anthocyanin content for modelling purposes (Féret et al., 2017). Obtaining greater variability in all pigment ratios allows for discrimination of each pigment's effects. By collecting samples with a variety of ages and colors, the chance of independent pigment variation should increase.

There are also few measurements associated with biophysical properties such as leaf thickness, surface roughness, or the presence of trichomes or surface waxes that can be linked to spectral datasets. The surface property characteristics are of particular importance when considering field applications as non-nadir illumination and bidirectional reflectance measurements are common as opposed to the hemispherical measurements that are often collected in laboratory settings. Non-nadir illumination introduces errors associated with leaf anisotropy and specular and diffuse reflectance properties (Bousquet et al., 2005; Comar et al., 2014). Hemispherical measurements

can mitigate these errors, but the bidirectional measurements of non-Lambertian surfaces are highly dependent on illumination and reflectance geometry as well as surface property characteristics (Jay et al., 2016; Woolley, 1971). The discrepancy between field and laboratory measurement orientations has often been ignored while monitoring crops as models developed with hemispherical measurements are commonly employed. However, data collection and monitoring can be improved by correlating hemispherical measurements to bidirectional measurements given the collection of both measurement geometries.

The dataset described in this work aims to provide sufficient data for relating hemispherical and bidirectional spectral measurements by collecting both measurement orientations for each sample. The traditional biochemical features that are collected in leaf spectral datasets are also collected in this dataset allowing for integration or combination of this dataset with those previously collected. Expanded assessment of biophysical properties is a focus in this new dataset as these features have not been previously studied in combination with both spectral and biochemical measurements. A comparison of the data collected in the ANGERS (Jacquemoud et al., 2003) and LOPEX (Hosgood et al., 1995) datasets and the data collected in the new LOTUS (Leaf Optical Traits collected at the University of Saskatchewan) dataset collected in this work is presented in Table 1. The data summarized in this work will help advance leaf optical modelling and calibration and will reduce errors associated hemispherical assumptions in field settings.

Table 1: Comparison of data collected in ANGERS, LOPEX, and LOTUS dataset.

	ANGERS	LOPEX	LOTUS
Number of samples	276	331	290
Spectral data			
Hemispherical	400 – 2500 nm	400 – 2500 nm	200 – 2500 nm
Angular			400 – 1650 nm
Polarized			400 – 1650 nm
Biophysical			
Water content	✓	✓	✓
Dry matter	✓	✓	✓
Leaf thickness		✓	✓
Cell size			✓
Cell shape			✓
Surface roughness			✓
Biochemical			
Chlorophyll a	✓	✓	✓
Chlorophyll b	✓	✓	✓
Carotenoids	✓	✓	✓
Pheophytin a			✓
Pheophytin b			✓
Anthocyanin			✓
Protein		✓	
Lignin		✓	
Cellulose		✓	
Starch		✓	
Additional data			
Leaf images			✓
Microscope images			✓
Elemental Concentrations		✓	

2 Methods

2.1 Samples and Collection

A total of 256 leaf samples were analyzed in this study. Among these samples, 53 different species were collected and 34 leaves were analyzed on both the abaxial and adaxial sides (resulting in 290 sets of spectral and surface measurements and 256 sets of biochemical measurements). Samples were collected from agricultural field trials, local outdoor gardens and

trees, and multiple greenhouses on the University of Saskatchewan campus. All samples were collected in Saskatoon, Saskatchewan, Canada. Up to 10 samples were collected at one time and were stored in individual plastic bags on ice in a dark cooler (Martin et al., 2018). All measurements (spectral and biophysical) were completed within 3 hours of collection. Table 2 summarizes the number of samples and each species in the dataset.

Table 2: Sample species and number collected in the LOTUS dataset.

Species	# of samples	Species	# of samples
Abelmoschus esculentus (L.) Moench	3	*Malus* sp. Mill.	6
Acer negundo L.	3	*Nicotiana tabacum* L.	4
Amelanchier alnifolia (Nutt.) Nutt.	5	*Ocimum basilicum* L.	2
Begonia sp. L.	1	*Parthenocissus quinquefolia* (L.)[†]	4
Beta vulgaris L.	2	*Penstemon barbatus* (Cav.) Roth	1
Brassica napus L.	87	*Petunia* sp. Juss.	1
Brassica oleracea L.	2	*Phaseolus vulgaris* L.	36
Capsicum annuum L.	11	*Prunus* sp. L.	13
Capsicum baccatum L.	1	*Quercus* sp. L.	10
Catharanthus sp. (L.) G.Don	1	*Rubus idaeus* L.[‡]	2
Celosia sp. L.	1	*Rumex acetosa* L.	1
Cornus sp. L.	3	*Salvia* sp. L.	1
Dianthus barbatus L.	1	*Sambucus racemosa* L.	1
Eucalyptus regnans F.Muell.	1	*Solanum lycopersicum* L.	8
Fragaria x ananassa Duchesne	4	*Solanum tuberosum* L.	5
Fraxinus pennsylvanica Marshall	10	*Spiraea* sp. L.	1
Geranium sp. L.	2	*Symphoricarpos* sp. Duhamel 1755	1
Glycine max (L.) Merr.	3	*Taraxacum* sp. F. H. Wigg.	1
Gynura sp. Cass.	1	*Tradescantia* sp. L.	6
Helianthus annuus L.	4	*Tropaeolum* sp. L.	2
Helianthus tuberosus L.	6	*Typha* sp. L.	1
Hosta sp. Tratt.[§]	1	*Ulmus americana* L.	8
Hypoestes phyllostachya Baker, 1887	2	*Verbena* sp. L.	1
Lactuca sativa L.	8	*Vitis vinifera* L.	2
Lathyrus odoratus L.	1	*Zea mays* L.	6
Ligularia sibirica (L.) Cass.	1	*Zinnia* sp. L.	1

[§]conserved name, not Jacq. (syn of *Cornutia* in Lamiaceae) nor Vell. Ex Pfeiff (Primulaceae)
[†]*Parthenocissus quinquefolia* (L.) Planch.
[‡]*Rubus idaeus* L. 1753 not Blanco 1837 nor Vell. 1829 nor Pursh 1814 nor Thunb. 1784

Leaves were assessed by first removing major veins and imaging the leaf on a flat surface to determine total leaf area. A preliminary mass of the leaf (m_{leaf1}) was measured and leaf thickness (t) was acquired using a micrometer. The leaves were returned to the cooler if there was to be any substantial length of time between measurements. Leaves were mounted in a leaf holder to

maintain a consistent assessment location and a series of spectral and microscopic measurements were then taken. The final leaf-level assessment was to record the leaf mass again at the end of all processing (m_{leaf2}), then remove two leaf disks of know area ($a_{disc} = 1.54\ cm^2$), and record the leaf mass without the two disks (m_{leaf3}). One disk was analyzed immediately (or frozen at -20°C for less than 24 hours) to determine biochemical attributes and one disk was frozen at -80°C to allow for future biochemical testing if desired (such as flavanol quantification). The leaf was then dried at 65°C for at least 72 hours and weighed to determine the dry leaf mass (m_{leaf4}).

2.2 Spectral Measurements

Spectral measurements were collected using two instruments. A Cary 5000 spectrophotometer (Agilent Technologies, Mississauga, ON, Can.) with a diffuse reflectance accessory was used to measure diffuse hemispherical reflectance and transmittance for each sample. These measurements were corrected using a well-characterized Spectralon reflectance standard (Labsphere, North Sutton, NH, USA). Illumination was set at nadir and diffuse reflectance was acquired. Due to the size and shape of the leaves, the sample viewing aperture on the sample holder was smaller than the reflectance port and spectrophotometer beam. Corrections for the dimensions of the aperture were performed using the measured transmittance through an empty sample holder (Noble and Crowe, 2007). Data were collected between 200 and 2500 nm at 1-nm intervals.

The second instrument (a goniospectropolarimeter, or GoSPo) used two spectrometers and an apparatus built in-house to assess bidirectional (biconical (Schaepman-Strub et al., 2006)) reflectance and transmittance of samples. The full description and calibration of this instrument is described by Peters et al. (2020). Bidirectional reflectance and transmittance measurements were obtained at 5° intervals in the zenith plane measured from nadir illumination (0°). These

measurements were later used to approximate hemispherical reflectance and transmittance for comparison with the Cary 5000 assessments. A well-characterized, 100 μm-thick sheet of polytetrafluoroethylene (PTFE) that was mounted in a sample holder was used as the calibration target in GoSPo.

Additional spectral measurements were obtained in GoSPo using illumination and detection at Brewster's angle (approximately 55° for leaves) with a rotating linear polarizer between the sample and sensor. These measurements were obtained for every 5° of rotation of the polarizer between 0 and 90° to identify the specular and diffuse reflectance from the leaf surface and were corrected using the total light as measured through the polarizing filter. All spectral measurements in GoSPo were obtained between 350 and 1650 nm at variable intervals less than 2 nm. The polarized spectra obtained at Brewster's angle contain an artifact associated with a thin film that was present on the polarizing filter. Although this artifact resulted in unexpected spectral variation, the averaged reflectance factors over a wavelength range at each polarizer orientation have been useful for correlation with biophysical leaf surface properties.

2.3 Biophysical Measurements

Biophysical measurements included moisture content, thickness, and several characterizations of leaf surface properties. The biophysical measurement symbols and descriptions are summarized in Table 3.

Table 3: Biophysical properties measured in LOTUS.

Parameter	Description	Unit
C_w	Per-area water content	g/cm^2 or cm
C_m	Per-area dry matter content	g/cm^2
t	Leaf thickness	mm
Cell size	Average size of epidermal cell	μm^2
MU	Degree of undulation of epidermal cell margin	unitless
CCAR	Ratio of epidermal cell width to height	unitless
SA	Average angle of incidence with nadir illumination	(°)

The per-area leaf water content (C_w) was determined two different ways by combining all four weight assessments of the leaf during processing. The first method used the leaf mass after the biochemical sample areas were removed (m_{leaf3}) and its dry mass (m_{leaf4}) as shown in Equation 1.

$$C_{w-low} = \frac{m_{leaf3} - m_{leaf4}}{a_{leaf} - 2 \times a_{disc}}, \quad (1)$$

where $C_{w\text{-}low}$ is the low estimate for water content, m_{leaf3} is the wet leaf mass after hole punching, m_{leaf4} is the dry leaf mass, and a_{leaf} is leaf area measured before hole punching.

The second method used to determine the water content involved a correction for the mass lost during the time required for spectral measurement in the spectrophotometer and GoSPo (Equation 2).

$$C_{w-high} = \frac{m_{leaf3} \times \frac{m_{leaf1}}{m_{leaf2}} - m_{leaf4}}{a_{leaf} - 2 \times a_{disc}}, \quad (2)$$

where $C_{w\text{-}high}$ is the high estimate for water content, m_{leaf1} is the mass of the leaf before scanning and m_{leaf2} is the mass of the leaf after scanning.

The actual water content in the leaf at any point during the spectral measurements would lie between the minimum value $C_{w\text{-}low}$ and the maximum value $C_{w\text{-}high}$.

Leaf mass per area (C_m) was determined using the dry mass of the leaf and the leaf area minus the area removed during hole punching as shown in Equation 3.

$$C_m = \frac{m_{leaf4}}{a_{leaf} - 2 \times a_{holepunch}}, \quad (3)$$

The leaf thickness was measured with a micrometer on an area of the leaf that did not present prominent veins. This metric was recorded in millimeters with precision to the nearest 0.01 mm.

Leaf surface assessment was done by visual observation as well as microscopic observation at 100 times and 500 times magnification. Images of the whole leaf were taken by placing the leaf flat (using small dark colored weights when necessary) on a light-colored surface and using a digital camera mounted at a consistent height. The surface was calibrated and mapped for area-based calculations that were later performed on the images in MATLAB using a simple color segmentation to discriminate leaf area (a_{leaf}).

For both levels of microscopic evaluation, a polarizing metallurgical microscope was used and images were taken using an 18MP digital camera (ME580TA-PZ-2L-18M3, AmScope). The images taken at 100 times magnification were collected with and without a polarizing filter to assist in classifying surface waxes. At 500 times magnification, the depth of field of the images was not sufficient to capture the entire leaf surface in focus. To obtain a single image that was in focus for each sample, an extended depth of focus technique was used (Peters, 2022). A series of 10 to 50 images were taken for each sample while moving the stage height 0.001 mm between each image. These images were then processed using a maximum gradient identification method that selected the most focused image for each pixel and stitched all focused pixels together. The 3D surface plot of the sample was then obtained by mapping the microscope stand heights of each pixel's chosen image.

Several metrics were obtained from the microscope images including the cell size, cell shape, and the cell cap aspect ratio. Based on the combined observations of the leaf and microscope images, each sample was classified as either waxy (glossy), glaucous, hairy (high trichome density), or glabrous. Glossy leaves appeared shiny and the microscope images had high variability between polarized and unpolarized images. Glaucous leaves had a blueish or whitish hue and the surface had a mottled appearance in microscope images. Hairy leaves had multiple hairs visible at 100x magnification. Glabrous leaves were classified in the absence of the aforementioned attributes. Average angle of incidence was also determined from the leaf surfaces.

2.4 Biochemical Measurements

Two 1.54 cm^2 leaf disks were removed from each sample. One disk was analyzed within 24 hours and one was frozen at -80°C. The chlorophyll a and b, carotenoid, pheophytin a and b, and anthocyanin content were all analyzed from a single leaf disk using a combination of protocols developed by Lichtenthaler (1987) and Gitelson et al. (2009). Multiple extractions were performed in series to ensure total chlorophyll and carotenoid extraction (Dunn et al., 2004). The leaf disk was first macerated by pestle and mortar in 1.65 mL of 60% acetone. The contents were then transferred to a micro centrifuge tube and centrifuged for 30 seconds or until a pellet formed. The supernatant was transferred to a 10 mL volumetric flask. The pestle and mortar was rinsed with 1.65 mL of 80% acetone which was then transferred to the micro centrifuge tube and used to suspend the pellet. The micro centrifuge tube was agitated and centrifuged again for 30 seconds. The supernatant was pooled in the 10 mL flask. The pestle and mortar was rinsed a second time with 1.65 mL of 100% acetone which was transferred and used to suspend the pellet. The micro centrifuge tube was again agitated and then centrifuged for 30 seconds and the

supernatant was pooled in the 10 mL flask. If at this time the pellet was not white or colorless, a fourth extraction was performed using 80% acetone. After complete extraction, the 10 mL flask was filled with 80% acetone. This solution was transferred to a cuvette and scanned in a Cary 5G, dual-beam spectrophotometer from 300 nm to 800 nm with 80% acetone in the reference cuvette. The solution was then mixed in equal volumetric parts with 0.1% HCl in 80% acetone and scanned a second time. The biochemical measurements and symbols are summarized in Table 4. The measurement pheophytin a and b in this dataset was primarily performed to facilitate anthocyanin measurement.

Table 4: Biochemical properties measured in LOTUS.

Parameter	Description	Unit
C_a	Chlorophyll a	$\mu g/cm^2$
C_b	Chlorophyll b	$\mu g/cm^2$
C_x	Carotenoids	$\mu g/cm^2$
C_{ph-a}	Pheophytin A	$\mu g/cm^2$
C_{ph-b}	Pheophytin B	$\mu g/cm^2$
C_{anth}	Anthocyanin	$\mu g/cm^2$

The scan taken before acidification with HCl was used to determine the chlorophyll a, chlorophyll b, and carotenoid concentration in the leaf. Using the equations outlined in Lichtenthaler (1987) for 80% acetone, Equations 4, 5, and 6 were used to determine chlorophyll a (C_a), chlorophyll b (C_b), and carotenoids (C_x) concentrations, respectively.

$$C_a = \frac{12.25 \times A_{663} - 2.79 \times A_{647}}{a_{holepunch}} \times v_{sol}, \quad (4)$$

$$C_b = \frac{21.50 \times A_{647} - 5.10 \times A_{663}}{a_{holepunch}} \times v_{sol}, \quad (5)$$

$$C_x = \frac{1000 \times A_{470} - 1.82 \times C_a - 85.02 \times C_b}{198 \times a_{holepunch}} \times v_{sol}, \quad (6)$$

where A_{xxx} is the absorbance at the specified wavelength and v_{sol} is the total volume of the extraction solution (10 mL for non-acidified).

After acidification, the pheophytin a and b and anthocyanin concentrations were determined. For pheopythin a and b ($C_{ph\text{-}a}$ and $C_{ph\text{-}b}$, respectively), Equations 7 and 8 were used based on those outlined in Lichtenthaler (1987). Equation 9 was then used to determine the anthocyanin concentration and is based on the work described in Gitelson et al. (2009).

$$C_{ph-a} = \frac{22.42 \times A_{665} - 6.81 \times A_{653}}{a_{holepunch}} \times v_{sol}, \quad (7)$$

$$C_{ph-b} = \frac{40.17 \times A_{653} - 18.58 \times A_{665}}{a_{holepunch}} \times v_{sol}, \quad (8)$$

$$C_{anth} = \left(A_{530} - \frac{MEC_{ph-a} \times C_{ph-a}}{MW_{ph-a}} - \frac{MEC_{ph-b} \times C_{ph-b}}{MW_{ph-b}}\right) \times \frac{MW_{anth} \times v_{sol}}{MEC_{anth} \times a_{holepunch}}, \quad (9)$$

where the *MEC* is the molar extinction coefficient and *MW* is the molecular weight. The values and sources for these numbers are summarized in Table 5. Equation 9 accounts for the absorption at 530 nm corresponding to the pheophytin a and b concentrations to accurately estimate the anthocyanin concentration.

Table 5: Summary of parameters in Equation 9 for anthocyanin concentration determination.

Parameter	Value	Source
$MEC_{ph\text{-}a}$	8.17 $mM^{-1} \cdot cm^{-1}$	(Gitelson et al., 2009)
$MEC_{ph\text{-}b}$	6.35 $mM^{-1} \cdot cm^{-1}$	(Gitelson et al., 2009)
MEC_{anth}	30 $mM^{-1} \cdot cm^{-1}$	(Gitelson et al., 2009)
$MW_{ph\text{-}a}$	871.2 $g \cdot mol^{-1}$	(PubChem, 2020a)
$MW_{ph\text{-}b}$	885.2 $g \cdot mol^{-1}$	(PubChem, 2020b)
MW_{anth}	207.2472 $g \cdot mol^{-1}$	(Khoo et al., 2017)

Two samples in the dataset (both *Beta vulgaris* L.) were betalain-containing species. For these samples, the betalain content was not estimated and the anthocyanin concentration was recorded as zero. Betalain estimation may be possible with the additional frozen samples.

3 Results

The results of each section of data collection are summarized here. Averages and trends among the data will be discussed. The full dataset is available at https://www.frdr-dfdr.ca/ DOI: 10.20383/103.0606.

3.1 Spectral Measurements

A wide variety of species and colors of leaves resulted in large variations in the visible spectrum of the hemispherical measurements. A select number of the hemispherical reflectance and transmittance spectra that capture the range of data collected in this dataset are shown in Figure 1. The corresponding 5-letter codes for each sample are provided (as labelled in the published dataset). The biochemical, biophysical, species, and observational data of the associated samples can be found in the full published dataset. For samples with both adaxial and abaxial measurements, the same 5-letter code was used with an additional letter “b” suffixed to the code to indicate the abaxial measurement.

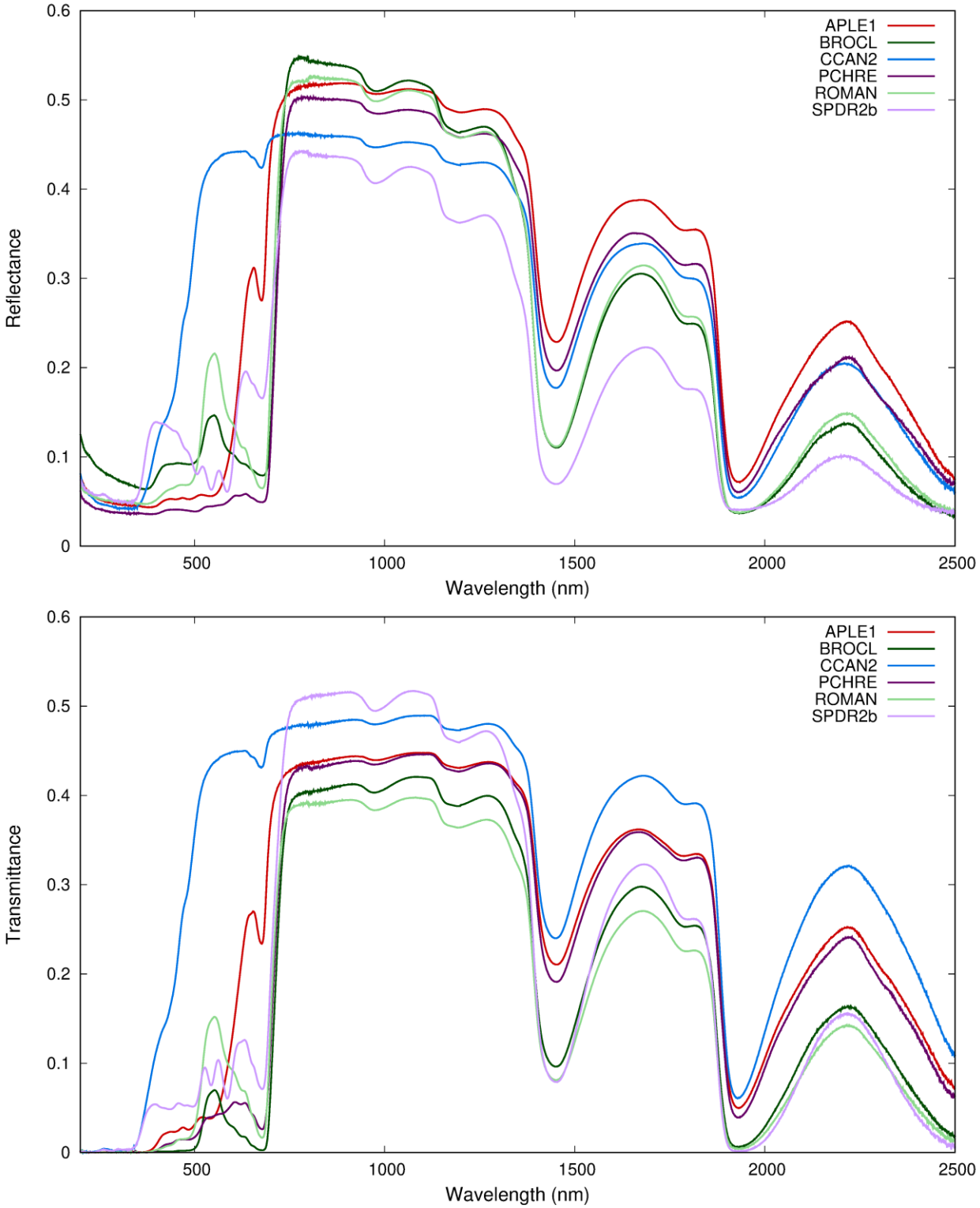

Figure 1: Hemispherical reflectance and transmittance spectra for select samples that cover a wide range of species and color presentations in the leaves.

In Figure 1, two green leaves (BROCL and ROMAN) are shown with high and low levels of chlorophyll concentration. The CCAN2 sample represents a sample with very low pigment concentrations and is a leaf that appeared very light yellow. The APLE1 leaf had very low chlorophyll content but high anthocyanin while the PCHRE leaf was high in both chlorophyll and anthocyanin. Finally, the SPDR2b leaf had a moderate amount of anthocyanin and chlorophyll. This last leaf displayed a unique feature that is very uncommon in leaf spectral measurements. The reflectance and transmittance just above 500 nm shows two distinct peaks that do not typically associate with any of the pigments measured in this study. This is likely due to accumulation of protoporphyrin IX and was more prevalent on the abaxial side of the leaves in which it was observed (Karcs et al., 2014). The biochemical and biophysical properties of the samples in this figure are summarized in Table 6.

Table 6: Biochemical and biophysical properties of the leaves associated with the spectra in Figure 1.

Parameter	APLE1	BROCL	CCAN2	PCHRE	ROMAN	SPDR2b
$C_a(\mu g/cm^2)$	1.75	43.05	1.27	34.31	19.92	29.63
$C_b\ (\mu g/cm^2)$	1.88	18.03	2.04	16.41	7.94	11.35
$C_x\ (\mu g/cm^2)$	3.27	10.55	0.41	9.82	6.83	7.53
$C_{ph-a}\ (\mu g/cm^2)$	2.14	42.52	1.02	35.19	20.32	29.79
$C_{ph-b}\ (\mu g/cm^2)$	1.65	13.85	1.35	14.36	7.02	9.17
$C_{anth}\ (\mu g/cm^2)$	7.34	1.16	0.50	7.60	0.95	3.25
$C_{w-low}(g/cm^2)$	0.0060	0.0150	0.0088	0.0100	0.0175	0.0324
$C_{w-high}(g/cm^2)$	0.0072	0.0171	0.0092	0.0112	0.0200	0.0337
$C_m(g/cm^2)$	0.0041	0.0028	0.0018	0.0059	0.0025	0.0019
$t\ (mm)$	0.17	0.24	0.13	0.19	0.27	0.44
Cell size (μm^2)	968	761	1120	621	2212	14706
MU	0.906	0.872	0.934	0.929	0.796	0.975
CCAR	7.39	15.92	17.68	7.43	11.19	19.98
SA (°)	13.66	6.16	7.98	13.02	17.17	6.82

Other spectral measurements included bidirectional angular reflectance and transmittance. In Figure 2, the reflectance factors at 5 degree intervals around the zenith plane are shown for a single sample with nadir illumination. These reflectance factors can be combined through hemispherical integration to approximate hemispherical reflectance (or transmittance) using a process described in Peters et al. (2020). These angular reflectance factors will help relate hemispherical reflectance and transmittance to bidirectional measurements that are more easily obtained in the field.

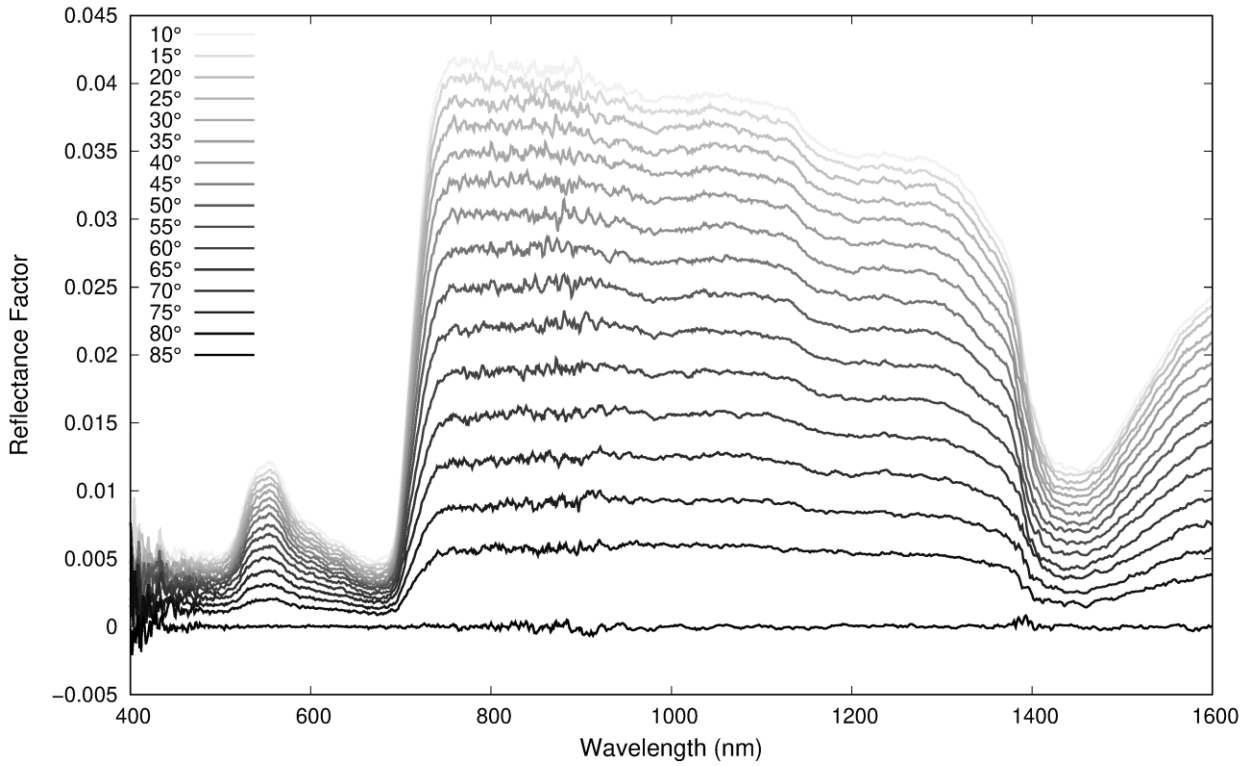

Figure 2: Bidirectional reflectance factor measurements for AY004 at 5 degree intervals around the zenith plane. Zero degrees represents illumination which is perpendicular to the leaf surface.

A second presentation of the bidirectional reflectance factor measurements is shown in Figure 3 for two leaf samples. These polar plots highlight the spatial variation of reflectance and transmittance factors that are affected by both the biophysical and biochemical properties of the

leaves. An optically smooth leaf with high specular reflectance results in a more oblong reflectance factor distribution whereas an optically rough surface results in a round reflectance factor distribution that is more comparable to Lambertian reflectance. The shape of the transmittance factor distribution is dependent on both the surface properties as well as the internal biophysical properties (not measured in this dataset).

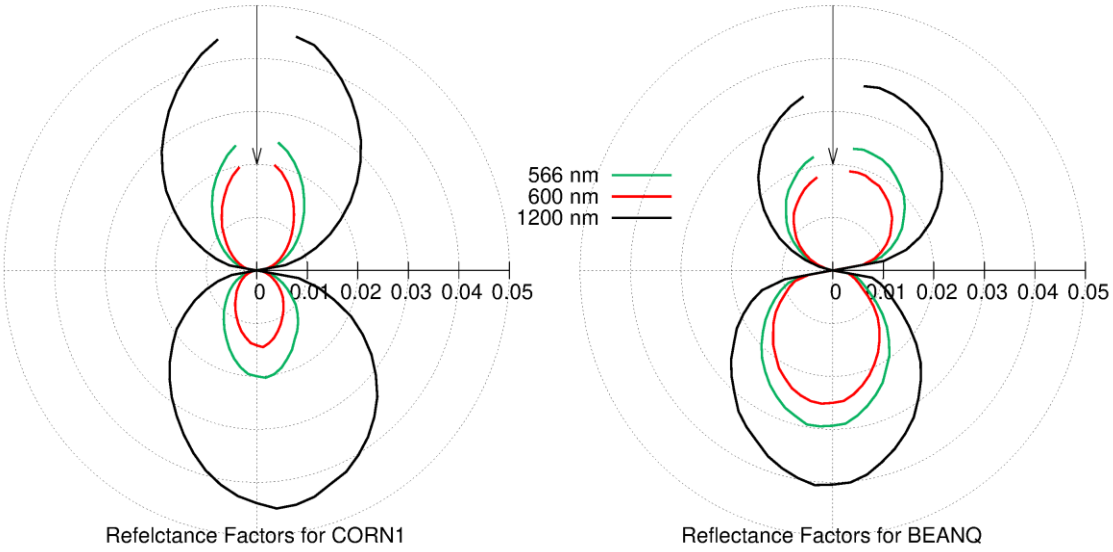

Figure 3: Polar plots at three wavelengths measured around the zenith plane for a glossy or shiny leaf (CORN1) and a leaf with a rough surface (BEANQ). Nadir illumination is represented by the arrow.

Spectral measurements at Brewster's angle (55°) with variation in polarizer orientation are summarized in Figure 4 for a waxy (glossy) leaf. With high wax loading, a smoother, shinier surface is observed on the leaf resulting in high variation in reflectance factor as the polarizer is rotated. Shown in Figure 4 are reflectance factors when the polarizer is rotated at 5 degree intervals with light source and sensor each at 55° from nadir.

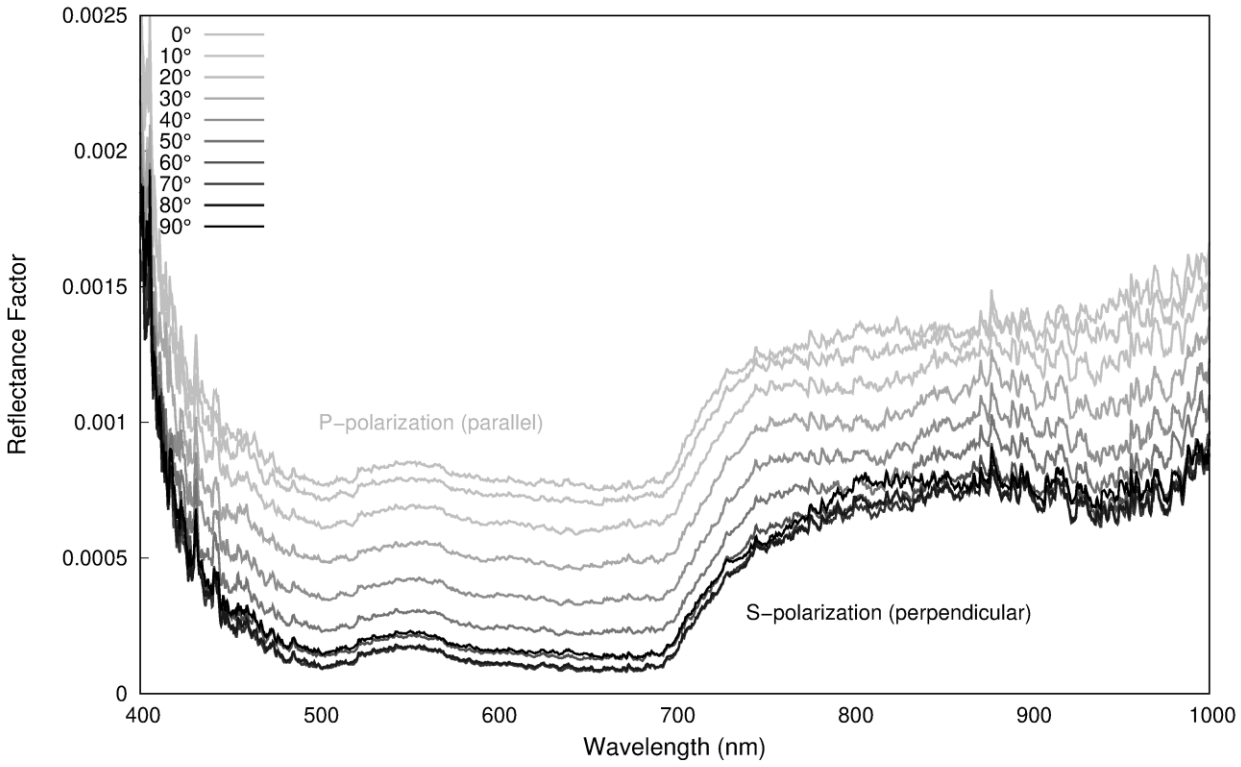

Figure 4: Reflectance factors measured at Brewster's angle (55°) with a rotating polarizer between the sample and sensor (for AELM1).

3.2 Biophysical Measurements

The distribution of the leaf dry matter content is shown in Figure 5a. The majority of samples had dry matter content between 0.002 and 0.005 g/cm^2. Figure 5b displays the distribution of leaf thickness in the dataset. Most leaf samples were between 0.2 and 0.3 mm thick. There was one extreme outlier with a thickness of 1.5 mm. This sample was from *Typha latifolia* L. (CATTL - a cattail plant).

Due to variations in leaf water loss over time, both the maximum and minimum calculated water content are shown in Figure 5c. The minimum estimated value for water content was between 60% and 96% of the maximum value for the samples in this dataset but the correlation between minimum and maximum values was very high. In general, leaves that had high water

concentrations had greater differences between the minimum and maximum values (higher water content loss).

In these samples, there is a slight correlation between the leaf thickness and water content (R^2 of 0.44 and 0.43 for maximum and minimum values, respectively), but no observable correlation between leaf thickness and dry matter content (R^2 of 0.03).

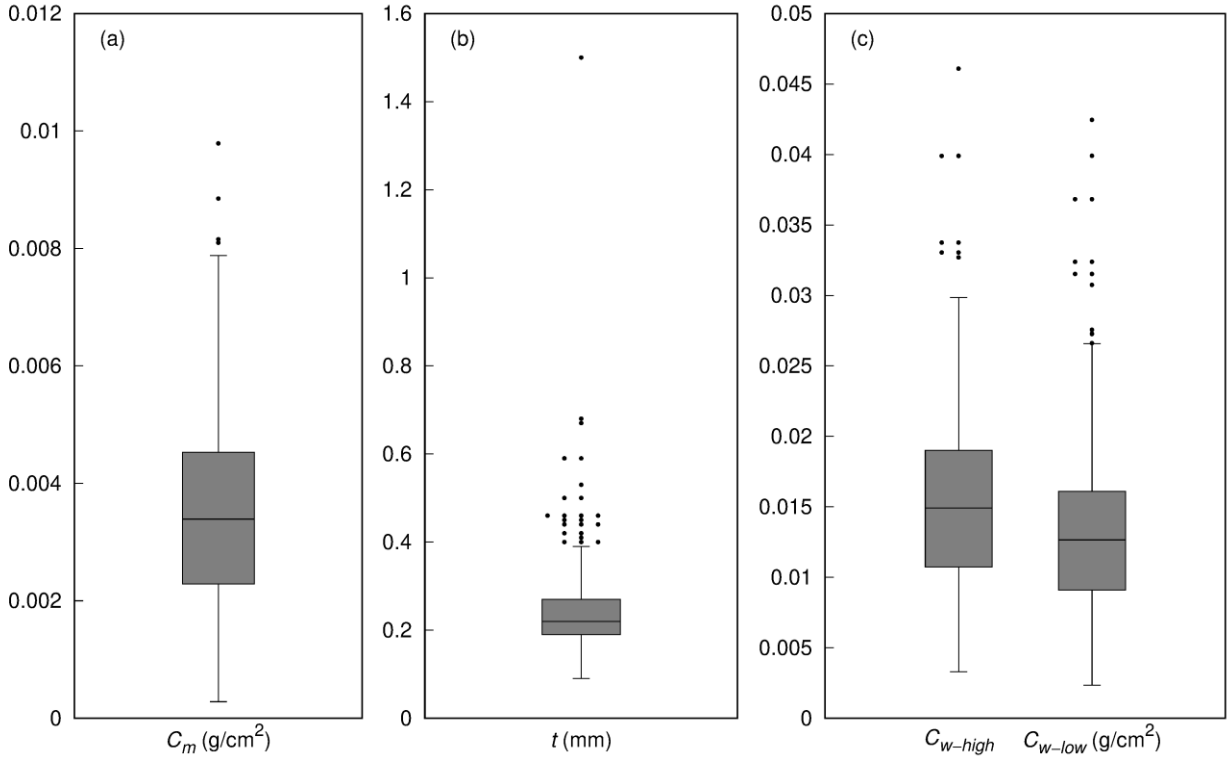

Figure 5: Sample distributions for (a) dry matter, (b) leaf thickness, and (c) water in the dataset.

Other biophysical assessment included images of the full leaves and microscopic imaging at 100 and 500 times magnification. In Figure 6, an example of 3 leaves with different surface phenotypes summarizes some of the differences at all three levels of observation within the dataset. At 100 times magnification, both the polarized and unpolarized images are shown to highlight the observable differences using this technique. The 500 times magnification images

305 are the results of the stitched extended depth of focus technique. The height map of the surface at
306 500 times magnification is also shown.

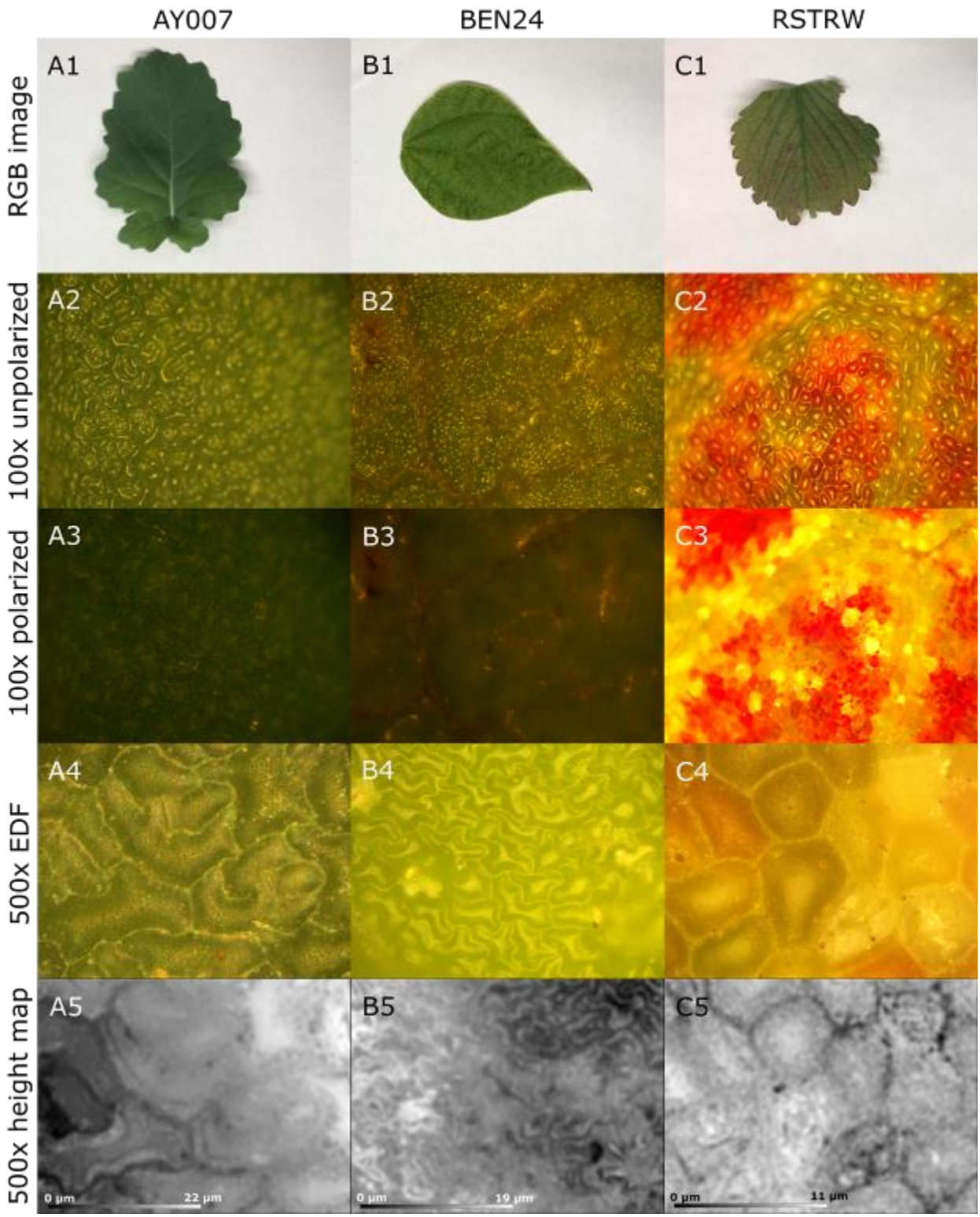

Figure 6: Imaging results for three samples: AY007 (A1-A5), BEN24 (B1-B5) and RSTRW (C1-C5). Each sample has a photo of the sample (A1, B1, C1), 100 times magnification without

polarizer (A2, B2, C2) and with polarizer (A3, B3, C3), stitched 500 times magnification (A4, B4, C4) and a relative height map at 500 times magnification (A5, B5, C5).

3.3 Biochemical Measurements

The distribution of measured biochemical attributes is summarized in Figure 7. Chlorophyll a and b are presented separately and show the expected higher estimation of a than b. A few samples had very low chlorophyll a concentrations; these leaves had very low pigment concentration in general and were mostly light yellow or white in color with a couple leaves that were bright red. Also in line with expectation, the carotenoid concentration is typically lower than both chlorophylls. The pheophytin a and b concentrations follow a similar trend to chlorophyll a and b and have high correlating values with these pigments. The anthocyanin concentration is distributed over a large range with many samples showing very little anthocyanin expression (less than 5 $\mu g/cm^2$). The samples with high anthocyanin concentrations were those that were visibly red or purple. There was very little correlation between anthocyanin content and chlorophyll content. Some dark red leaf samples were high in both anthocyanin and chlorophyll while some leaves with very low chlorophyll had high anthocyanin expression. This range in anthocyanin to chlorophyll concentration ratio will aid in modeling of these pigments. A low correlation between pigment concentration (or a high level of concentration independence) and a large range of pigment concentrations is important when differentiating between the effects of each model input.

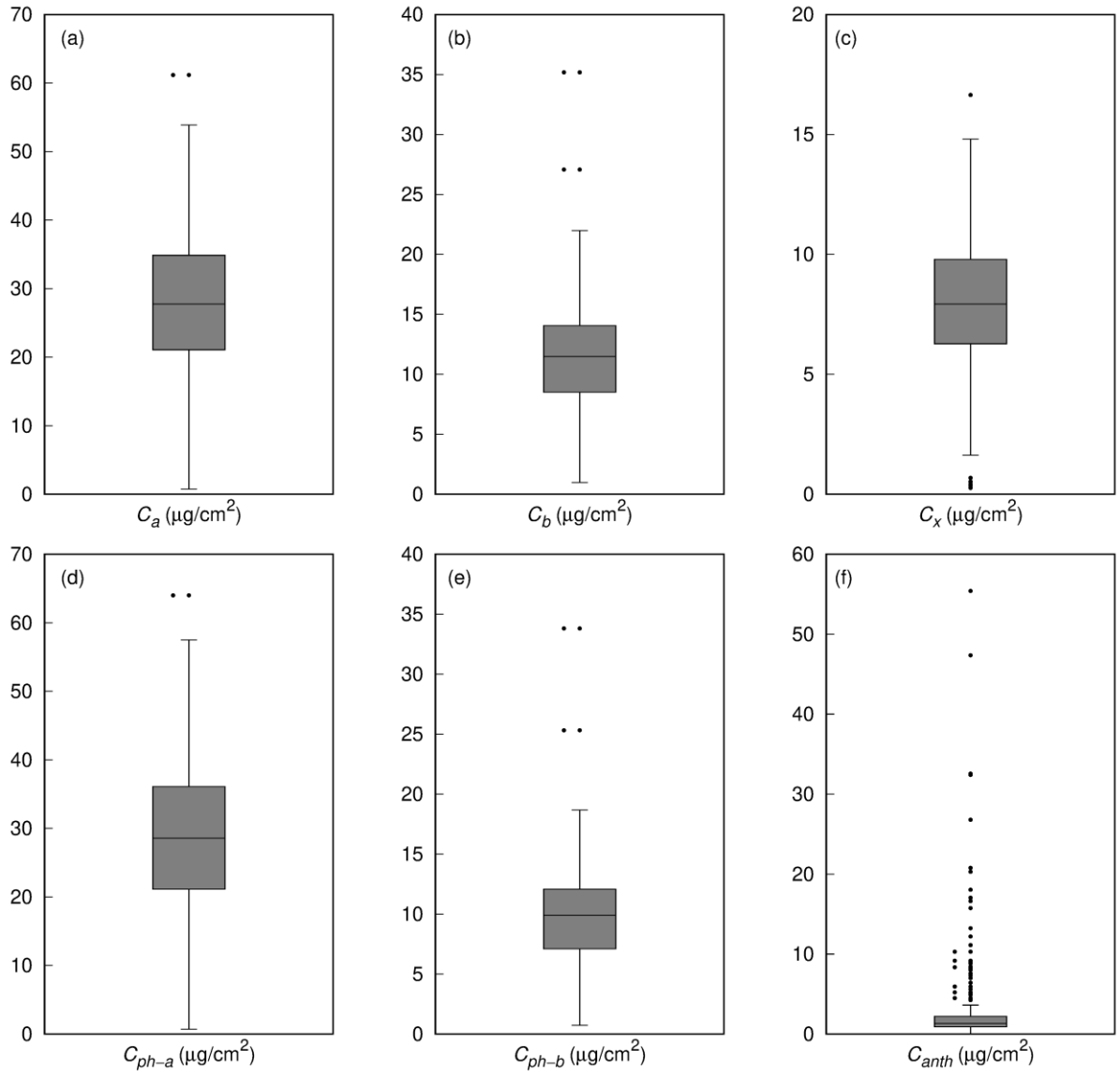

Figure 7: Distributions for biochemical concentrations in the dataset for (a) chlorophyll a, (b) chlorophyll b, (c) carotenoids, (d) pheophytin a, (e) pheophytin b, (f) anthocyanins.

Correlations between chlorophyll a and b concentrations were high as has been observed in previous studies (in the ANGERS dataset, the ratio of chlorophyll a to chlorophyll b was assumed to be approximately 3 (Féret et al., 2017)) but some leaves showed variation in this ratio. This was particularly true at very low and high overall chlorophyll concentrations as shown in Figure 8a. These samples that vary from the typical ratio assumptions may be useful for future

modelling when differentiating between these two pigments is of interest. It should be noted that pigment extractions were performed with acetone in LOTUS and with 95% ethanol in ANGERS (Feret et al., 2008). This pigment solvent difference could account for a higher overall chlorophyll b to a ratio as acetone extracts chlorophyll b more effectively (Sumanta et al., 2014). Figure 8b shows the correlation between carotenoids and the overall chlorophyll concentration in each sample. Similar to the chlorophyll a and b comparison, the correlation is high between these pigments, but deviations are more prominent at very high and low concentrations of chlorophyll. Samples with very low chlorophyll but significant carotenoid expression are important when considering model calibration if separation of these pigments is required. Carotenoids are also more effectively extracted with acetone than ethanol (Sumanta et al., 2014) but the increased expression potentially attributed to solvent type is not seen in Figure 8b.

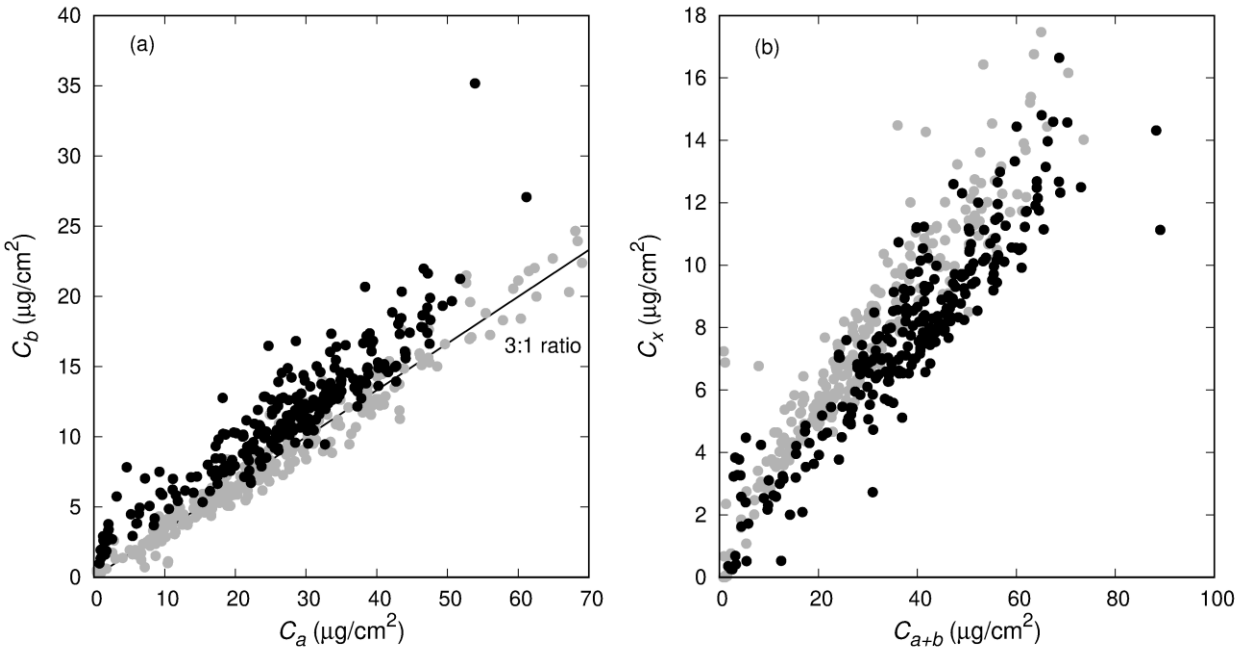

Figure 8: Correlations between chlorophyll a and b and carotenoids in ANGERS (gray) and LOTUS (black).

4 Conclusions

The LOTUS dataset was collected to help improve the application of leaf optical property modelling towards application in precision agriculture. A series of spectral, biophysical, and biochemical measurements provide a more complete assessment of features that can affect light reflectance and transmittance. A dataset that includes biophysical measurements in addition to spectral and biochemical has not been previously published, but has the potential to greatly improve leaf spectral modelling – especially in a field setting. This dataset includes a variety of agricultural species in addition to other samples with the goal of improving model calibration for applications in agriculture. Leaves at various stages in development with a large range of color expression improve upon previously collected datasets to better represent the full life cycle of a leaf. Measurement of anthocyanin concentration follows advancements in current leaf models (PROSPECT-D). Additional leaf disks were saved in this experiment for possible collection of future pigments of interest.

Acknowledgements

This work was supported in part by the Plant Phenotyping and Imaging Research Centre funded through the Canada First Research Excellence Fund (CFREF) and in part by the Natural Sciences and Engineering Research Council of Canada (NSERC).

References

Baranoski, G. V., 2006. Modeling the interaction of infrared radiation (750 to 2500 nm) with bifacial and unifacial plant leaves. Remote Sens. Environ. 100(3), 335-347.

Bousquet, L., Lachérade, S., Jacquemoud, S., Moya, I., 2005. Leaf BRDF measurements and model for specular and diffuse components differentiation. Remote Sens. Environ. 98(2-3), 201-211.

Dunn, J. L., Turnbull, J. D., Robinson, S. A., 2004. Comparison of solvent regimes for the extraction of photosynthetic pigments from leaves of higher plants. Funct. Plant Biol. 31(2), 195-202.

Comar, A., Baret, F., Obein, G., Simonot, L., Meneveaux, D., Viénot, F., De Solan, B., 2014. ACT: A leaf BRDF model taking into account the azimuthal anisotropy of monocotyledonous leaf surface. Remote Sens. Environ. 143, 112-121.

Feret, J. B., François, C., Asner, G. P., Gitelson, A. A., Martin, R. E., Bidel, L. P., Ustin, S. L., le Maire, G., Jacquemoud, S., 2008. PROSPECT-4 and 5: Advances in the leaf optical properties model separating photosynthetic pigments. Remote Sens. Environ. 112(6), 3030-3043.

Féret, J. B., Gitelson, A. A., Noble, S. D., Jacquemoud, S., 2017. PROSPECT-D: Towards modeling leaf optical properties through a complete lifecycle. Remote Sens. Environ. 193, 204-215.

Gitelson, A. A., Chivkunova, O. B., Merzlyak, M. N., 2009. Nondestructive estimation of anthocyanins and chlorophylls in anthocyanic leaves. Am. J. Bot. 96(10), 1861-1868.

Hosgood, B., Jacquemoud, S., Andreoli, G., Verdebout, J., Pedrini, G., Schmuck, G., 1995. Leaf optical properties experiment 93 (LOPEX93). Report EUR, 16095.

Jacquemoud, S., Baret, F., 1990. PROSPECT: A model of leaf optical properties spectra. Remote Sens. Environ. 34(2), 75-91.

Jacquemound, S., Bidel, L., Francois, C., Pavan, G., 2003. ANGERS Leaf Optical Properties Database (2003). Data Set. Available online: http://ecosis.org (accessed on 5 June 2018).

Jay, S., Bendoula, R., Hadoux, X., Féret, J. B., Gorretta, N., 2016. A physically-based model for retrieving foliar biochemistry and leaf orientation using close-range imaging spectroscopy. Remote Sens. Environ. 177, 220-236.

Karcz, D., Boroń, B., Matwijczuk, A., Furso, J., Staroń, J., Ratuszna, A., Fiedor, L., 2014. Lessons from chlorophylls: modifications of porphyrinoids towards optimized solar energy conversion. Molecules. 19(10), 15938-15954.

Khoo, H. E., Azlan, A., Tang, S. T., Lim, S. M., 2017. Anthocyanidins and anthocyanins: colored pigments as food, pharmaceutical ingredients, and the potential health benefits. Food Nutr. Res. 61(1), 1361779.

Lichtenthaler, H. K., 1987. Chlorophylls and carotenoids: pigments of photosynthetic biomembranes. Methods Enzymol. 148, 350-382.

Martin, R. E., Chadwick, K. D., Brodrick, P. G., Carranza-Jimenez, L., Vaughn, N. R., Asner, G. P., 2018. An approach for foliar trait retrieval from airborne imaging spectroscopy of tropical forests. Remote Sens. 10(2), 199.

National Center for Biotechnology Information. PubChem Database. Pheophytin a, CID=135398712, https://pubchem.ncbi.nlm.nih.gov/compound/Pheophytin-a (accessed on Apr. 24, 2020).

National Center for Biotechnology Information. PubChem Database. Pheophytin b, CID=135407446, https://pubchem.ncbi.nlm.nih.gov/compound/Pheophytin-b (accessed on Apr. 24, 2020).

Noble, S. D., Crowe, T. G., 2007. Sample holder and methodology for measuring the reflectance and transmittance of narrow-leaf samples. Appl. Opt. 36(22), 4968-4976.

Peters, R. D., 2022. Microscope EDF [Source code]. https://gitlab.com/usask-speclab/microscope-edf (accessed on July 7, 2022).

Peters, R. D., Hagey, S. R., Noble, S. D., 2020. GoSPo: a goniospectropolarimeter to assess reflectance, transmittance, and surface polarization as related to leaf optical properties. J. Appl. Remote Sens. 14(4), 047505.

Schaepman-Strub, G., Schaepman, M. E., Painter, T. H., Dangel, S., Martonchik, J. V., 2006. Reflectance quantities in optical remote sensing—Definitions and case studies. Remote Sens. Environ. 103(1), 27-42.

Sumanta, N., Haque, C. I., Nishika, J., Suprakash, R., 2014. Spectrophotometric analysis of chlorophylls and carotenoids from commonly grown fern species by using various extracting solvents. Res. J. Chem. 2231, 606X.

Woolley, J. T., 1971. Reflectance and transmittance of light by leaves. Plant Physiol. 47(5), 656-662.