# PREDICTION OF LEAF RELATIVE WATER CONTENT USING PATTERNS OF HYPERSPECTRAL INTENSITY

by

Mark Gee

## A Thesis

Submitted to the Faculty of Purdue University In Partial Fulfillment of the Requirements for the degree of

Master of Science in Agricultural and Biological Engineering



School of Agricultural and Biological Engineering West Lafayette, Indiana May 2022

# THE PURDUE UNIVERSITY GRADUATE SCHOOL STATEMENT OF COMMITTEE APPROVAL

# Dr. Jian Jin, Chair

School of Agricultural and Biological Engineering

## **Dr. Mitch Tuinstra**

School of Agronomy

## **Dr. David Umulis**

School of Agricultural and Biological Engineering

# Approved by:

Dr. Nathan S. Mosier

# ACKNOWLEDGMENTS

Thank you to Dr. Jian Jin for your mentorship and support of my research. Thank you to Dr. Mitch Tuinstra for your guidance and pushing me to design stronger experiments. Thank you to all the lab members who pitched in with good cheer and made it possible to complete the work. There are too many to name, but Zhihang, Yikai, Thirwat, Tanzeel, Libo, and the Tuinstra Tigers deserve special recognition. To Donald, Glen, and Margaret – you pitched in when I needed it most and for that I am grateful.

# TABLE OF CONTENTS

LIST (	OF T	ABLES	. 6
LIST (	OF F	IGURES	. 7
ABST	RAC	Т	. 9
1. IN	NTRO	DDUCTION	10
2. L	ITER	ATURE REVIEW	11
2.1	Mai	ize Water Status and Mechanisms of Drought Adaptation	11
2.	1.1	Agronomic Impact of Drought	11
2.	.1.2	Maize Water Status	11
2.	.1.3	Drought Adaptation Mechanisms	13
2.	.1.4	Measuring Plant Water Content	16
2.2	Bre	eding for Drought Tolerance	17
2.	.2.1	Direct Selection	17
2.	.2.2	Selection for Secondary Traits	18
2.	.2.3	Genomic Selection	20
2.3	Hyp	perspectral Phenotyping of Leaf Water Content	21
2.	.3.1	Early Work: Hyperspectral Indices to Predict Plant Water Content	22
2.	.3.2	Current Standard: Hyperspectral Data and Partial Least Squares Regression	23
2.	.3.3	Next Steps: Utilization of High Resolution Imagery to Improve Prediction	25
3. M	IETH	IODS	28
3.1	Exp	perimental Design	28
3.2	Har	ndheld Hyperspectral Device and Plant Sampling	28
3.3	Mea	asurement of Leaf Relative Water Content	29
3.4	Ima	ge Processing and Segmentation	30
3.5	Mo	del Development, Variable Selection, and Accuracy Comparison	33
3.6	Spa	tial Pattern Feature Development	35
3.	.6.1	Mean Spectrum	35
3.	.6.2	Multiple Locations on the Leaf	35
3.	.6.3	Tissue Types	36
3.	.6.4	Cross Sectional Pattern	37

3	.6.5 Contours	. 38
3	.6.6 Texture	. 39
4. F	RESULTS	. 41
4.1	Experimental Verification	. 41
4.2	Mean Spectrum Model	. 43
4.3	Multiple Locations on the Leaf	. 44
4.4	Tissue Type	. 49
4.5	Cross Sectional Pattern	. 51
4.6	Contours	. 54
4.7	Texture	. 55
5. I	DISCUSSION AND CONCLUSIONS	. 58
REFE	ERENCES	. 60

# LIST OF TABLES

Table 1. Model performance for multiple locations on the leaf. The p-value is the result of test for difference relative to the R2Prediction for the mean spectrum model	the t-
Table 2. Model Performance for various tissue types	51
Table 3. Model performance for WI features. The p-value is the result of the t-test for different relative to the R2Prediction for the mean spectrum model	erence 53
Table 4. Model Performance with GLCM Texture Features.	56

# LIST OF FIGURES

Figure 1. Symptoms of common nutrient deficiencies (Shandong)
Figure 2. Handheld hyperspectral imaging of maize leaves. (a) LeafSpec in use. (b) Mean spectrum obtained from a leaf scan. (c) Full resolution NDVI heatmap obtained from a leaf scan
Figure 3. Leaf with markings showing the location of tissue samples for relative water content.29
Figure 4. Reconstructed RGB image from a hyperspectral leaf scan after calibration and segmentation
Figure 5. Hyperspectral scan at full resolution (top) and after resizing (bottom)
Figure 6. Heatmap of NDVI intensity for a leaf with high nitrogen (left) and low nitrogen (right) fertilizer treatment. Color scale is the same for both images
Figure 7. Heatmap of WI intensity for a leaf with $RWC = 0.91$ (left) and $RWC = 0.71$ (right). Color scale is the same for both images
Figure 8. Plot of measured RWC vs predicted RWC. Red line is the least squares best fit. $R^2Prediction = 0.691$
Figure 9. Segmentation regions for midrib and green sections
Figure 10. Regions used to construct vertical and horizontal cross sections
Figure 11. Contour map from a representative leaf
Figure 12. Water treatment groups and soil moisture content measured at sampling
Figure 13. Plot of soil volumetric water content vs leaf relative water content at the time of measurement
Figure 14. Plot of soil volumetric water content vs leaf relative water content at the time of measurement. Each sub-plot is colored by the experimental factor in the sub-plot title.Figure 15. Plot of soil volumetric water content vs leaf relative water content at the time of measurement. 42
Figure 14. Plot of soil volumetric water content vs leaf relative water content at the time of measurement. Each sub-plot is colored by the experimental factor in the sub-plot title
Figure 15. Variable Importance in Projection plot for the mean spectrum PLSR model
Figure 16. Heatmap of R <sup>2</sup> Prediction at each pixel location for resized leaf scans
Figure 19. Heatmaps of R2Prediction at each pixel location over resizing iterations. The plot shows the maximum R2Prediction for an individual pixel at each resizing. The mean spectrum model is represented by resizing eight
Figure 20. Plot of R2Prediction for every pixel combination relative to the R2Prediction for the best individual pixel and whole leaf mean spectrum model

## ABSTRACT

Drought is the leading cause of crop loss globally. Breeding for drought tolerance is difficult due to the polygenetic nature of the trait and low heritability of yield under drought. Plant relative water content is a secondary trait that may advance drought breeding programs.

The LeafSpec, a newly developed hyperspectral leaf scanner, was used to test the hypothesis that distribution of hyperspectral information across the leaf can be used to improve prediction of leaf relative water content. Data was collected across two experiments from five different maize genotypes representing temperate and tropical hybrids with varying levels of drought tolerance and inbreds with varying stomatal densities. The hyperspectral intensity averaged across the entire leaf was used to predict relative water content with an R<sup>2</sup>Prediction of 0.7989. Model performance was tested using additional predictors that quantify:

- Spectral information from multiple regions in the leaf (e.g. base, middle, tip).
- Spectral information from regions segmented by tissue type.
- The distribution of hyperspectral intensity in a cross section parallel to the midrib or in a cross-section perpendicular to the midrib.
- A contour pattern of hyperspectral intensity from the outside edge of the leaf to the midrib.
- Texture features extracted from each wavelength.

The mean spectrum model outperformed previously reported results, potentially due to the elimination of sources of noise and higher quality data produced by the LeafSpec. None of the models with expanded feature sets outperformed the mean spectrum model at a statistically significant level. The hyperspectral signal from the green tissue a third of the way from the base of the leaf and half way between the midrib and edge was the most correlated with relative water content. Models without midrib and vein tissue signals had increased performance. Distribution of the Water Index visually showed improved ability to discriminate leaf RWC as compared to individual wavelengths but this did not translate to improved model performance.

For future work, more data should be collected to improve model robustness and hyperspectral imaging should include SWIR wavelengths that have previously been found useful for predicting relative water content. Exploring indices composed from current spectral bands may lead to improved prediction performance.

## **1. INTRODUCTION**

Drought is the leading cause of crop loss globally. Breeding for improved drought tolerance is difficult due to the polygenetic nature of the trait and low heritability of yield under drought conditions. To address this, breeders have turned to breeding for secondary traits such as leaf rolling under water stress or anthesis-silking interval which are more heritable under drought stress.

Genomic breeding has opened up a new frontier for drought tolerance breeding since genomic markers are fully heritable and can be consistently identified regardless of environmental conditions. However, identification of relevant markers requires study of large populations and phenotyping methods that are scalable, repeatable, and cost effective. For this reason, image based techniques have emerged as the phenotyping tool of choice.

Plant relative water content is a secondary trait that has the potential to be useful in drought breeding programs. While previously measured manually, hyperspectral imaging has proven to be reasonably effective method for predicting plant relative water content across diverse genetics. The state-of-the-art method for predicting plant relative water content from hyperspectral data is to collect hyperspectral images, segment out plant pixels, average the hyperspectral intensities across all plant pixels, and to use this mean spectrum to build a partial least squares regression model.

Work on classification of leaves based on nitrogen treatment has shown that analyzing the distribution of hyperspectral data across the leaf can improve predictive performance over the mean hyperspectral data. In this work, we use data from the LeafSpec, a newly developed hyperspectral leaf scanner, to test the hypothesis that distribution of hyperspectral information across the leaf can be used to improve prediction of leaf relative water content.

## 2. LITERATURE REVIEW

## 2.1 Maize Water Status and Mechanisms of Drought Adaptation

#### 2.1.1 Agronomic Impact of Drought

In the modern era, maize has become the largest grain crop and almost outweighs wheat and rice combined in terms of tons produced per year (*Agricultural Output - Crop Production - OECD Data*, 2021). Maize is grown all over the globe and plays a significant role for both industrialized food systems and consumption oriented small-holder farmers.

The effect of drought on maize and the people who grow it has been a major factor for as long as maize has been cultivated. As far back as the 12th century, drought induced failures in maize crops forced the Anasazi Native Americans to abandon their homeland and search for wetter climates (Benson et al., n.d.) As US maize yields have risen, the sensitivity to drought has increased because the improved performance requires access to adequate water. Some estimates state that drought sensitivity in US corn yields have risen 55% from 1999 to 2018 (Lobell et al., 2020). Currently, drought is the leading cause of crop losses from natural disaster and the damage is only expected to increase with rising global temperatures (FAO, 2018).

## 2.1.2 Maize Water Status

Water status maintenance is critical to healthy development at every maize growth stage. The root is the first thing to emerge from a germinating maize seed and immediately grows downward in search of water to support shoot growth. Maize leaves are oriented so that rainfall is redirected towards the stalk where it flows downwards directly to the root zone.

On an agronomic scale, growing a maize crop requires a significant quantity of water. In central Nebraska, seasonal water use for maize is 22-25 in/yr as compared to 21-23 in/yr for soybeans and 19-22/yr in for sorghum (Kranz, 2015). However, when compared based on grain produced per use of water consumed, maize is reasonably efficient. In a three year study of rainfed crops in Nebraska, maize had a water use efficiency of 5.6 bu/in, sorghum was 6.37 bu/in and soybean was 2.43 bu/in (Rees, 2015). Daily water demand is highly variable and is influenced by weather conditions. Overall, water demand is low during crop establishment, picks up around V4

as the canopy closes, and reaches a maximum just before the reproductive stages of growth with a slow tapering during the grain filling stages (Kranz, 2015).

In well-watered plants, water content is highest in the leaf sheaths, followed by the stem and roots, with water content being the lowest in the green leaf tissue (Sun et al., 2015). Along the leaf, water content is generally observed to decrease slightly from the base towards the middle with a more drastic decline towards the leaf tip (Pick et al., 2011). These studies did not separate tissue types when measuring water content, so the observed trend may be due to the declining quantities of midrib and xylem tissues closer to the leaf tip. More work is needed to understand if there is a water gradient across the leaf in the green leaf tissue.

Within each plant, transpiration creates a gradient of water potential extending upward from the roots to the elongating portion of the leaf. Experiments with V5 maize plants grown in controlled environments quantified the water potential profile at -0.07 to -0.01 MPa in the soil, -0.15 MPa in the mature root tissue, -and -0.13 to -0.17 MPa across the leaf tissue (Tang & Boyer, 2002). When water in the soil is limited, the water potential in the plant becomes approximately ten times greater (-1.5 MPa) and is strongest at the leaf tip. This results in a stronger driving force to extract water from dry soil. Some of the water is diverted from the transpiration path by a growth induced water potential that extends radially between xylem veins in the elongation portion of the leaf. This potential between two veins ranged from -0.44 MPa during the day and fell found to -0.38MPa during the night. Comparatively, transpiration moves water under a gradient of -0.17 MPa m<sup>-1</sup> and growth creates a gradient of -1000 MPa m<sup>-1</sup> between vascular bundles.

The flow of water in leaves can be analyzed using radioactive isotopes. As water evaporates, there is an enrichment of heavy isotopes due to the molecular kinetics of evaporation (Merlivat & Coantic, 1975). Several models have been developed to describe this phenomena in leaves. The most famous is the Craig-Gordon but this has been improved upon by models such as the one developed by Farquar-Gan (Craig & Gordon, 1965; Gan et al., 2003). Studies of <sup>18</sup>O enrichment in plant leaves show the evaporation leads to isotope enrichment along the transpiration path from the base of the leaf to the tip and radially from the midrib to the leaf edge (Gan et al., 2003). This enrichment is humidity dependent as it is driven by the evaporation of water along the leaf tissue (Gan et al., 2003). The authors used the Faquar-Gan model to quantify flow regimes using the Peclet number which is a dimensionless number defined as the rate of advective flow to the rate of diffusive flow. In that study, the longitudinal variability in <sup>18</sup>O was best modeled by a Peclet

number as high as 10^7, indicating a flow dominated by advection rather than diffusion. Peclet numbers describing the radial gradient in <sup>18</sup>O were on the range of 0-2.5 indicating a diffusive flow.

## 2.1.3 Drought Adaptation Mechanisms

Under a shortage of water, maize plants have several adaptive behaviors to maintain a healthy water content in plant tissues for as long as possible.

Leaf rolling is the most obvious response to water stress. Maize leaves are normally flat and elongated to expose the maximum surface area to the sun for photosynthesis. Under water stress, maize leaves begin to fold at the midrib and curl along its axis to create a corkscrew shape. This is caused by a difference in elastic contraction between the top and bottom surfaces of the leaf as moisture content declines (Hay et al., 2000). Rolled leaves have lower rates of transpiration and are exposed to less photosynthetic damage from sunlight (O'Toole et al., 1979; Smith et al., 1997). While leaf rolling may help plants maintain internal hydration, it does not seem to contribute to productivity and may be more of a desperate attempt at survival. Genotypes which exhibit higher levels of leaf rolling tend to have reduced yield under drought conditions (Allah, 2009; Effendi et al., 2019). Thus, other mechanisms of maintaining water status are more advantageous from an agronomic and breeding perspective.

Stomata are critical to water stress adaptation. Stomatal opening and closing is a rapid response that is able to help plants cope with short-term environmental changes such as daily variations in water stress. Short-term opening and closing can be regulated passively by water turgidity but is more commonly managed via an active stress response signaled through calcium, ABA, or another signal molecule (Le et al., 2011). However, it has been hypothesized that oversensitive stomatal response with early closure can lead to reduced adaptation to short term dehydration and lower agronomic performance under normal conditions (Benešová et al., 2012). Changes in stomatal density are observed under longer periods of drought stress. Studies that transgenically altered stomatal density found high correlations between stomatal density and both water use efficiency and drought survival (Liu et al., 2015). Under severe drought stress, xylem embolism will occur in the midrib, but this is not observed in leaf veins due to stomatal closure (Cochard, 2002). Maize plants that have been exposed to a week or more of drought conditions do not increase their stomatal density once the water stress has ended while drought hardy plants like

sorghum are capable of recovering normal stomatal density even after a fourteen day drought (Glover, 1959).

Roots are the foundation for maintaining plant water status. Under water stress, maize plants will stop shoot growth at a deficit of -1.00 MPa but will maintain root growth all the way until a deficit of -1.4 MPa is reached (Westgate & Boyer, 1985). An ideotype of "steep, deep, and cheap" was proposed as an optimal strategy for accessing deep soil moisture during water limited conditions (Lynch, 2013). Often, root density is higher than necessary for optimal for water uptake, potentially for increased extraction of nutrients (Robertson et al., 1993). When water becomes limiting, plants can reduce rooting density at shallower depths and divert more resources to growing towards deeper soil moisture. In a greenhouse drought study, plants with deep, low biomass root systems were found to be more efficient in terms of transpiration efficiency (van Oosterom et al., 2016). There are many other mechanisms of root adaptation to drought stress. Aeration in root cortical aerenchyma have been associated with higher levels of shoot biomass under drought stress, deeper rooting patterns, and higher leaf water content (Zhu et al., 2010). The authors of the study attribute these correlations to the reduced metabolic cost of soil exploration caused by higher air volume in roots. Roots are an under explored aspect of plant biology and more work needs to be done to understand the mechanisms of drought adaptation.

Leaf growth is reduced under water stress. The rate of leaf elongation is consistent for wellwatered plants, but decreases exponentially with the water potential of the soil as stress is applied (Acevedo et al., 1971). In well-watered plants, a strong water potential gradient of -1000 MPa m<sup>-1</sup> exists between the vascular bundles of the elongation region causing leaves to expand along the midrib and widen axially (Tang & Boyer, 2002). Under water stress, this gradient is eliminated and does not return until sufficient soil moisture is present (Tang & Boyer, 2002). Leaf extension has been observed to be linearly related to the water potential in the leaf tissue (Acevedo et al., 1971). Mildly stressed plants resume normal leaf elongation seconds after rewatering, supporting the idea that the xylem acts as a hydraulic unit bringing about immediate changes in leaf water potential as the soil water potential changes (Acevedo et al., 1971). This adaptation sacrifices new growth, but conserves water for already matured leaf tissues.

Osmotic and metabolic adjustments are made to cope with the altered biochemistry that occurs under water limiting conditions. Studies of xylem sap exudates show significant differences in signaling hormones and the proteome under drought conditions. (Alvarez et al., 2008). Levels

of signaling hormones BAP and ABA were significantly higher 7 to 12 days after the initiation of water stress. ABA is a common long-distance stress signal between roots and leaves, so the increase was expected. BAP is implicated in delaying leaf senescence and promoting the accumulation of proline as an osmotic adjustment (McDavid et al., 1973; Thomas et al., 1992). Amino acid levels circulating in xylem sap also tended to increase, notably proline which has been implicated as an osmotic adjustment in response to salt stress (Thomas et al., 1992). The components for building lignin in cell walls were found to increase, indicating that lignin biosynthesis had declined in water stressed plants, possibly as a side effect of reduced growth overall. Expression increased for peroxidases responsible for crosslinking cell wall components. More crosslinking increases the rigidity of cell walls and could be a mechanism by which plants reduce growth and conserve water. The method used to extract xylem sap involved cutting off the top portion of the plant which almost certainly affects hormone signaling. All expression levels were compared to a set of control plants, but it is possible that there was an interaction effect between the water stress treatment and plant wounding that affected relative levels of hormone changes. Comparison of protein levels found differential expression in cell wall metabolism such as peroxidase, xyloglucan endotransglycosylase, polygalacturonase inhibitor and pectin methylesterase, and plant defense mechanisms such as thaumatin-like pathogenesis-related protein, zeatin-like protein, cupin family protein, putative germin A, class IV chitinase and b-1,3-glucanase.

Once the environmental conditions exceed the capacity of the plant's adaptive defenses, water stress develops from the bottom up when there is low soil moisture or damage to the vascular system. Sun et al. (2015) characterized the internal redistribution of water within maize plants experiencing drought stress. Maize seedlings were grown for two weeks, at which point half of the seedlings were no longer watered while the other half continued to receive regular watering. Six plants were harvested from each water treatment group and destructively measured at the points where the soil moisture in the drying pots reached 75%, 55%, 35%, 25%, 20%, 15%, and 10% of the field water capacity. They found that drying first occurred in the green tissue and sheath of lower leaves and moved upwards to the next leaf as moisture became more limited. The moisture content of the upper leaves was maintained above 75% if the soil moisture was above 15% field capacity. Once soil moisture dropped to 15% or below, the upper leaves dried, followed by drying of the stem and roots as extreme stress continued. At 20% field capacity, inward shrinkage of cell walls was observed along with declining volume of bulliform cells, declining number of

chloroplasts, deformation of chloroplasts, and a decline of photosynthesis to almost zero. They found that growth could be recovered within hours by plants that had experienced water stress at 15% field capacity but if water stress of 10% field capacity was experienced, then a week or more of recovery time at normal soil moisture was needed for plants to reinitiate growth.

## 2.1.4 Measuring Plant Water Content

To understand the physiology of maize response to water stress, we must know the water content profile throughout the plant tissues. Physical measurements of relative water content have been the gold standard for assessing the water content of plant tissues. The general protocol is to cut a sample of tissue from the plant, record the fresh weight, submerge the tissue in water for several hours until turgidity is reached, blot the sample dry, record the turgid weight, dry the sample in an oven until equilibrium, and record the dry weight. The RWC value is calculated according to the formula below.

#### Equation 1. Relative Water Content Calculation

RWC = (FW - DW)/(TW - DW) \* 100%

This is a widely used technique and significant work has gone into understanding what parameters affect measurement quality. As far back as 1962, Barrs and Weatherly conducted an assessment of sources of experimental error associated with the relative turgidity method (Barrs & Weatherley, 1962). They identified three main sources of error: changes in sample dry weight due to continued photosynthesis or respiration, growth of the leaf disk resulting in absorption of water beyond full turgidity, and irregularities in water holding capacity due to the cut edges of the disks. To prevent the accumulation or loss of dry matter, they suggest maintaining the external lighting at the compensation point. In their experiments with castor leaves, they found 500 to 800lux to be the appropriate light intensity to balance rates of photosynthesis and respiration. Upon submersion, there are clearly two phases of water absorption, the first phase characterized by a rapid uptake lasting approximately 4 hours and the second phase characterized by a slower linear increase that continues indefinitely. The general consensus is that the first phase of water absorption is due to achievement of turgidity and the second phase is caused by continued growth of the leaf disk (Barrs

& Weatherley, 1962; Čatský, 1959; Weatherley, 1950). Introduction of metabolic inhibitors or elimination of oxygen were effective at eliminating phase two water absorption, but these methods are impractical (Barrs & Weatherley, 1962). They recommend ending rehydration after 4hr to eliminate the increases caused by phase two absorption. This recommendation has been confirmed in a wide range of crops and the 4hr recommendation was specifically verified for maize (Bliss et al., 1957; Čatský, 1960; Namken & Lemon, 1960). Barrs and Weatherley (1950) describe a species dependent effect of irregularities in water holding capacity due to the cut edge of the sample. In our work with maize, we have found this effect to be important. Using a sharp razor blade to cut tissues produces samples that absorb water more quickly and absorb a higher total volume of water when compared to samples taken with a hole punch. This could be due to damage in vascular structure caused by the crushing action of the hole punch.

Manual measurement of plant water status is the foundation of physiological drought research. In the current literature, this method is applied with widely varying submersion times, lighting conditions, etc. Researchers to rely on the consistency approach, in that if they treat all of their samples consistently then the RWC measurements are still comparable within their experiment. However, more consistent, efficient, cost-effective, and scalable methods are needed for breeding applications.

### 2.2 Breeding for Drought Tolerance

Improving drought tolerance has been an objective of breeding programs for a very long time. While unintentional selection for drought tolerance has probably been occurring since the domestication of maize, modern commercial programs began to adopt this as an important secondary trait in the 1930s after a series of years with severe drought (Janick, 2010). Drought tolerance has continued to be important as increasing plant densities put more pressure on soil moisture resources and drought events increase in severity and frequency (Barbosa et al., 2021; Janick, 2010).

### 2.2.1 Direct Selection

The most direct breeding approach is to select for yield under drought stress conditions. Recurrent selection of a segregating population under multi-location drought-stress experiments is a natural extension of current breeding programs that focus on improving yield in ideal conditions. For this approach to work, the genes for drought tolerance must be present in the narrow pool chosen for multi-location trials, they must be heritable, and progress from optimal environments must carry over to stressful environments (Maazou et al., 2016).

All of these assumptions are limited in some way. Drought tolerance genes are less abundant in elite germplasm and many efforts have been made to introgress improved alleles via wide crosses to non-elite germplasm (Barbosa et al., 2021; Meseka et al., 2013; Moeinizade et al., 2021). Heritability of yield under drought conditions is low. In a study of 3509 inbred lines across a range of moisture environments, yield was found to have a broad sense heritability of 0.6 under well watered conditions, but only 0.4 under stressful environments (Bolaños & Edmeades, 1996). Another study with testcrosses of 238 recombinant inbreds from a B73xMo17 population evaluated in multilocation trials with managed drought and non-drought conditions found that well watered yield heritability was 0.60 but only 0.37 in drought conditions (Ziyomo & Bernardo, 2013). Performance in optimal conditions is linked to, but does not guarantee performance in water limiting environments. If selection under optimal conditions was sufficient, then there would be no need to focus on drought trials. Ziyomo found only a 0.60 correlation between yield across watering treatments in a controlled breeding program (Ziyomo & Bernardo, 2013). A study of 108 genotypes grown in six locations with a range of rainfall conditions found that the repeatability of the grain yield trials ranged from 39 to 80% (Eze et al., 2020). Water stress was just one of many factors that caused the observed genotype by environment interaction, but nevertheless this study shows the limited carryover from optimal conditions to a drought environment.

Due to these limitations, results from directly breeding for yield under drought conditions are variable, with some authors reporting 12.6% gain in yield per selection cycle while others found very little genetic gain over three generations of selection (Chapman & Edmeades, 1999; Magorokosho & Tongoona, 2003)

## 2.2.2 Selection for Secondary Traits

To overcome the inconsistency of selecting for yield under drought, breeders have turned to selection for secondary traits that are more heritable under stress and are strongly correlated with yield. One of the earliest commercial breeding programs, Pioneer Hibred, has had a long term strategy of selecting for secondary traits such as anthesis-silking interval, ears per 100 plants, and delayed senescence to improve the efficiency of their stress breeding (Janick, 2010). Numerous studies have found gains in efficiency when selecting for secondary traits in addition to yield under drought stress (Bänziger & Lafitte, 1997; Ziyomo & Bernardo, 2013).

Selection for anthesis-silking interval (ASI) is one of the most successful examples. Drought stress during flowering causes a delay in silk growth relative to pollen shed. This can result in poor seed set or even total crop loss in extreme cases. Ability to maintain silk growth in water limiting conditions directly addresses the pollination challenge and may also generally indicate ability to maintain normal plant function under water stress. The previously referenced study that quantified the low heritability of yield under drought stress found that ASI maintained heritability under stressed conditions and had a 0.9 correlation with drought stressed grain yield (Bolaños & Edmeades, 1996). They found that pollination related secondary traits were much more correlated with drought yield than water status secondary traits. Magorokosho found ASI to be less correlated with drought yields (r = 0.40), but it was still one of the most useful and heritabile traits in their study (Magorokosho & Tongoona, 2003). Many other studies find ASI to be a reliable predictor of drought yield and is often the highest performing secondary trait for selection (Bänziger & Lafitte, 1997; Messmer et al., 2009; Ziyomo & Bernardo, 2013). As pollination becomes robust to drought stress, the importance of other secondary traits may increase.

Senescence traits (also called stay green) have been widely correlated with drought yields and yields in non-stressed conditions (Bolaños & Edmeades, 1996; Cairns et al., 2012; Ziyomo & Bernardo, 2013). Maintaining photosynthetic capacity by delaying senescence under stress increases yield because assimilate flux is a major determinant of kernel number and higher photosynthetic capacity later in the season increases grain fill (Schussler & Westgate, 1995).

Failure to develop an ear is a common drought stress response, so selection for consistency of ear setting under stress effectively increases yield stability (Wang et al., 2019). Bolanos found that ears per plant had a heritability of 0.77 under drought stress and had a 0.9 correlation with drought yield while yield in drought conditions had a heritability of only 0.4 (Bolaños & Edmeades, 1996).

Visual assessment of leaf rolling has been widely applied across breeding programs and automated methods have been developed to quantify this trait from drone imagery (Baret et al., 2018). However, leaf rolling is less correlated with yield under drought than other secondary traits

(Bruce et al., 2002). This may be due to the extreme stress required to induce leaf rolling having surpassed the relevant levels of water stress for assessing variation in drought yield.

#### 2.2.3 Genomic Selection

Genomic selection has distinct advantages for improving drought tolerance. Once qualitative trait loci (QTL) are identified, they can reliably be identified in a segregating population and are more consistent and much cheaper to measure than yield under stress. Speed of breeding is a critical advantage of genomic selection with three selection cycles per year possible in a greenhouse or winter nursery setup as compared to the two years needed for a cycle of testcross phenotypic selection (Ziyomo & Bernardo, 2013).

Many QTL have been identified for yield under drought stress and genomic methods of breeding are widely applied (Bankole et al., 2017; Bruce et al., 2002; Messmer et al., 2009; Ribaut & Ragot, 2007). A comparison of the breeding efficiency for selection on various traits relative to selection on drought yield alone (efficiency = 1.00) found that selection on the secondary trait ASI had a relative efficiency of 1.04, an index of drought yield and ASI had a relative efficiency of 1.13 and that genome wide selection for drought tolerance had a relative efficiency of 1.24 (Ziyomo & Bernardo, 2013). This experiment was conducted with testcrosses of 238 recombinant inbreds from a B73xMo17 population evaluated in multilocation trials with managed drought and non-drought conditions.

As the cost of sequencing has declined, breeders are shifting from targeted marker assisted selection and moving to more general whole genome selection (Bhat et al., 2016). Genomic breeding approaches are better suited to improving drought tolerance than marker assisted selection because the broad genome coverage is better able to identify the important components of complex quantitative traits. Cerrudo et al found that genomic selection outperforms marker assisted selection for grain yield and secondary traits in a study with 169 doubled haploid lines derived from the cross between CML495 and LPSC7F64 and 190 testcrosses (tester CML494) were evaluated in a total of 11 treatment-by-population combinations under WW and DS conditions (Cerrudo et al., 2018).

Several challenges for genomic breeding include the interaction of QTL with the environment, instability in yield QTL, sensitivity to different genetic contexts, and complexity of genetic interactions for drought stability (Messmer et al., 2009). The QTL controlling yield are

different in well watered vs water stressed environments indicating the continued need for selection under both conditions. Additionally, QTL identified in inbreds are not always valuable for predicting performance of hybrids and models often require training data from testcross populations to be useful (Cerrudo et al., 2018; Zhang et al., 2021). Work to identify QTL for salinity tolerance in rice using a fluorescence based imaging phenotype found significant QTL, but the QTL were not significant on all days of imaging (Campbell et al., 2015). The authors concluded that repeated phenotyping over time would be required to more reliably identify QTL for complex traits that would require non-destructive techniques such as imaging.

Genotypic selection requires a large training population to be genotyped and phenotyped. This information is used to build a model that identifies the important genomic components and future generations are genotyped to identify recombinant individuals with high genome estimated breeding values. Phenotypic verification of the genomic predictions can be carried out in tandem with genomic selection or at the end of several cycles. Because of the large training population required, the phenotyping methods must be cost effective, scalable, and consistent. Traditional phenotypes of yield under drought stress and manually measured secondary traits are not able to scale to the size needed for genomic prediction to reach it's full potential (Bhat et al., 2016). For this reason, much work has been put into developing automated phenotyping methods that can be deployed at scale.

## 2.3 Hyperspectral Phenotyping of Leaf Water Content

The ability to predict leaf water content at scale would provide breeders with a better tool to understand genetic adaptation to drought stress over a much more physiologically relevant range than current methods. Non-destructive imaging techniques are emerging as the dominant phenotyping method for large breeding programs pursing genomic selection (Bhat et al., 2016). These methods are repeatable due to their mechanical nature, high-throughput because each image only requires seconds to acquire, and cost effective at scale because the marginal cost of each image is almost zero after the initial equipment investment. Additionally, imaging data can be used to assess a wide range of traits at once using a single measurement.

#### 2.3.1 Early Work: Hyperspectral Indices to Predict Plant Water Content

Hyperspectral imaging is the most promising method to assess maize water content at scale. This method emerged from the NASA Landsat project, which is the "longest continuous spacebased record of Earth's land in existence" with the first satellite launched in 1972 (Landsat Science -, n.d.). The multispectral imagery from this project provided the impetus for developing spectral indices to predict plant characteristics, the most famous of which is the Normalized Difference Vegetative Index (NDVI) which is used to identify regions with vegetative cover and has since been correlated with nitrogen content (Edalat et al., 2019; Rouse et al., 1974). The prediction of plant characteristics from leaf reflectance relies on variations in light absorption due to excitation of molecular bonds, primarily C-H, N-H, and O-H bonds at wavelengths in the visible (400-700 nm), near-infrared (NIR; 700-1,100 nm), and short-wave infrared (SWIR; 1,100-2,400 nm) regions (Cotrozzi et al., 2020). In 1987, a water prediction index was developed using the Landsat imaging wavelengths in the range of 760-900 nm that was not significantly different from whole plant nor leaf water content in Agave deserti (Raymond Hunt et al., 1987). Over the subsequent decades, this index was refined and validated over a wider range of crops (PEÑUELAS et al., 1993). In 1996, a Normalized Difference Water Index (NDWI) was proposed that was the normalized difference of intensities at 860nm and 1240nm. This work broadened the spectral range used for predicting water content into the short-wave infrared region (SWIR). Since its inception, the index has been validated in many crops and refined for specific applications (Chai et al., 2021). Other indices have been proposed including the Global Vegetation Moisture Index which uses NIR and SWIR wavelengths in a mathematical combination distinct from the traditional NDVI form (Ceccato et al., 2002). Generally, these indices were applied to satellite imagery to improve analysis of land cover and agricultural water usage (Cheng et al., 2008; Sims & Gamon, 2003).

These efforts utilized small numbers of wavelengths in their final indices due to sensor limitations that prevented collection of full spectral data and insufficient computational power and methods to process spectral data if it could be obtained. As sensor technology improved, it became easier and more cost effective to collect images with full spectrum data. A method of using full spectral data for prediction of sample composition emerged from the field of chemometrics and began to be used more widely in the early 2000s (Wold et al., 2001). Partial Least Squares (PLS) was proposed as a statistical method to take the very large number of correlated predictors, reduce it to a set of key components, and use the key components to predict the response variable. This

method is similar to Principal Component Analysis (PCA) in that both methods reduce high dimensionality data into a set of critical components. However, the method of constructing these components is different. PCA maximizes the variance in predictor variables captured in each component while PLS maximizes the covariance between predictor and response captured. PLS is robust to having highly correlated predictor variables and is a preferred method for using spectral data to predict sample properties. Modern machine learning techniques are also capable of handling highly correlated predictor data, but are susceptible to overtraining and require very large datasets (>10,000 samples) to develop robust models.

### 2.3.2 Current Standard: Hyperspectral Data and Partial Least Squares Regression

As technology advanced, more work started to focus on using spectral data to predict relative water content in individual maize plants and leaves. Research done at the University of Nebraska Lincoln has been at the forefront of this effort. Early studies conducted in the greenhouse sought to correlate leaf hyperspectral reflectance with leaf water content using partial least squares regression (Ge et al., 2016). Spectral data from 500 – 1750 nm was collected using a Headwall scanner on 40 plants each from B73 and Fast Flowering Mini Maize-A. Scans were averaged across all plant pixels to create a representative spectrum used to predict leaf water content through PLSR modeling. The authors reported prediction  $R^2$  of 0.92 and 0.81 when modeling each genotype independently and 0.87 when modeling them both together. They repeated the experiment with 60 plants of B73 and were able to achieve a relative water content prediction  $R^2$ of 0.93 (Pandey et al., 2017). This approach was applied to the maize diversity panel consisting of 282 distinct genotypes in the field under nitrogen sufficient and nitrogen limiting conditions as well as the greenhouse under nitrogen limiting conditions (Ge et al., 2019). Leaf 2, 3, and 4 were sampled from a representative plant out of each plot. Hyperspectral data was collected at the tip, middle, and base of each leaf using a benchtop spectroradiometer with a range of 350-2500nm. The nine spectral readings were averaged to create a representative spectrum for the plant and compared to the mean leaf water content of the three sampled leaves. Data was divided into a training and test set according to a 60%/40% split. When modeled using both partial least squares regression and support vector machine regression techniques the prediction R<sup>2</sup> was 0.70 (Ge et al., 2019). The decline in accuracy across experiments indicates the challenge of building models that are robust to genotypic variation and noise generated by environmental variation.

Since the publication of these studies, others have replicated the results but failed to substantially improve the prediction accuracy across diverse genetics. A field study with six cultivars was able to predict canopy water content with an R<sup>2</sup> of 0.70 using a predetermined hyperspectral index (Elsayed & Darwish, 2017). A greenhouse study with several hundred plants across multiple experiments was able to achieve plant water content prediction R<sup>2</sup> of 0.76 and found that PLSR outperformed all other analysis methods (Mertens et al., 2021). A greenhouse study with 85 seedling plants used a variety of modeling methods including principal component analysis (PCA) and kullback-leibler divergence (KLD) to predict leaf water content from hyperspectral data (Gao et al., 2019). The authors were only able to report prediction  $R^2$  of 0.123 for PCA and 0.590 for KLD. When applied in a breeding context, hyperspectral modeling techniques were able to discriminate among drought sensitive and tolerant genotypes with a R<sup>2</sup> of 0.65 in the training environment which fell to an  $R^2$  of 0.36 when the model was applied to data from a different environment (Ryckewaert et al., 2021). A study comparing support vector regression (SVR), partial least square regression (PLSR), and deep convolutional neural network for regression (CNN-R), and decision fusions of these methods found that all methods produced similar results with a prediction R2 of about 0.85 (Zhou et al., 2021). Most of the work using advanced machine learning techniques is focused on using RGB imagery to predict different classes of water stress due to the massive computational burden of processing hyperspectral imagery (Chandel et al., 2020; Zhuang et al., 2017).

Other sensing technologies have been used to accurately predict leaf water content. Noncontact resonant ultrasound spectroscopy data paired with convolutional neural network and random forest modeling was able to predict leaf water content with  $R^2$  of 0.85 and 0.71 respectively (Fariñas et al., 2019). Singh et al. used terahertz imaging to map the 3D distribution of water in agave leaves (Singh et al., 2020). The disadvantage of these approaches is that they often require large, stationary equipment and are only suitable for indoor analysis of potted plants that can be moved into the imaging station.

## 2.3.3 Next Steps: Utilization of High Resolution Imagery to Improve Prediction

The maize leaf is complex. with multidimensional gradients of water content, gene expression, photosynthesis and tissue structure. Spatial patterns of color on leaves have been linked to a variety of plant conditions. Figure 1 shows several common nutrient deficiencies and the color patterns they cause (Shandong, n.d.). Pick et al. conducted a systems analysis of the maize leaf developmental gradients (Pick et al., 2011). They divided leaves into ten sections and analyzed each section to assess the distribution of metabolic and physiological properties finding that moisture content decreased linearly from the base of the leaf to the tip. Water exists in a gradient across the leaf and it is possible that water stress creates a unique pattern.

One of the main limitations of current hyperspectral phenotyping of maize water content is the loss of distribution information. Current approaches average spectral reflectance across the entire plant and loose any information





across the entire plant and loose any information that may have been contained in the leaf pattern.

Work in other crops shows that a distribution of hyperspectral signal exists across plant leaves and that this pattern may be useful for predicting leaf water properties. Higa et al. generated a PLSR water content prediction model and applied it to the individual pixels of high resolution hyperspectral images of golden pothos leaves (Higa et al., 2013). They found a distinct pattern in the predicted water content across the leaf with the edges and tip having less predicted moisture than the center and base. This result shows that there is variation in the spectral signature across the leaf, but conclusions about the actual water distribution can not be made since there was no experimental validation of the predicted pattern. They reported a 3% increase in model accuracy

when utilizing selected regions on the leaf to predict whole leaf water content, but did not test to ensure that this difference was statistically significant. Kim et al. used histogram analysis to try and differentiate between stressed and unstressed leaves (Kim et al., 2015). Leaf stress levels were generated by detaching leaves from a boxwood plant and allowing them to desiccate for various lengths of time. Spectral differences were attributed to differences in water content, but could have arisen from a wide range of stress responses caused by detachment. Nevertheless, the authors took high resolution hyperspectral images of the leaves and generated a histogram based on the pixel intensities for each wavelength and a variety of calculated indices. By comparing the Bhattacharyya statistics from histograms of two leaves, they were able to differentiate between the leaves when the difference in water content was at least 20%. This sensitivity is low, but indicates that higher level resolution data can be used to assess leaf characteristics. In their study of salinity on lettuce leaves, Lara et al. built a hyperspectral model to classify the effect of salinity treatment (Lara et al., 2016). They applied the model to the individual pixels from the leaf hyperspectral images and found higher levels of predicted salinity effects at the leaf edges. Texture features from RGB imagery were used classify young maize plants among three watering treatments with a 81% accuracy (Zhuang et al., 2017). Murphey et al. studied the effect of tissue types on water content distribution and hyperspectral prediction accuracy in lettuce leaves (Murphy et al., 2019). They measured the water content of whole leaves, green tissue, and stem and build correlations based on intensity values of the moisture stress index, normalized difference water index, and selected SWIR bands for each tissue type. They found that the relationship between hyperspectral signal and water content was highest for the green tissue and that hyperspectral signals from one leaf component were much less accurate at predicting water content of a different component. Leaf patterns are not used in current maize hyperspectral prediction models partly because previous work has not had access to a sensor capable of capturing a high resolution scan of the entire corn leaf.

The recently developed LeafSpec is a sensor with the capability to capture hyperspectral data from whole maize leaves with high enough resolution to observe potential patterns related to water stress signals (Wang et al., 2020). The LeafSpec is a push broom scanner capable of capturing high resolution hyperspectral images from 450 to 950 nm. Imaging noise from ambient lighting and leaf angle is reduced because the scanner is clamped around the leaf slid from the base to the leaf tip. In an experimental validation reported in the literature, the LeafSpec was used to

scan the top collared leaves of 59 hybrid B73xMo17 maize plants at the V7-V8 stage under a lattice experimental design with a high/low water treatment and high/low nitrogen treatment (Wang et al., 2020). The mean spectral data from each leaf scan was used to predict the relative water content of the leaf using PLSR and a prediction  $R^2$  of 0.771 was obtained. This result was consistent with the results reported in the literature where spectral data was averaged across the entire plant.

The high resolution capabilities of the LeafSpec have been used to improve nitrogen phenotyping. Ma et al. analyzed the distribution of NDVI from the base to tip of leaves and found that the pattern of NDVI was different for leaves from high and low nitrogen treatments. Using this pattern, they were able to develop an algorithm with improved sensitivity for discriminating between high and low nitrogen treatments (Ma et al., 2020).We hope to improve the prediction accuracy for leaf relative water content with the development of algorithms capable of utilizing the distribution of hyperspectral information across the leaf.



Figure 2. Handheld hyperspectral imaging of maize leaves. (a) LeafSpec in use. (b) Mean spectrum obtained from a leaf scan. (c) Full resolution NDVI heatmap obtained from a leaf scan.

## 3. METHODS

## **3.1** Experimental Design

Two experiments were conducted in the Purdue Lilly Greenhouse (40°25'19.7"N, 86°55'7.8"W) to generate the data for this research. The temperature in the greenhouse was maintained at 23–29 °C with supplemental lighting for 12 h a day. Plants were grown in three liter pots filled with a 67% / 33% blend of Sun Gro Horticulture MetroMix 510 and Greens Grade Turface Profile. To ensure proper establishment, three seeds were sown per pot and thinned to one plant per pot one week after emergence. Two weeks after emergence, plants began to receive fertilizer treatments of 0.42g of 6-24-24 per week along with an additional 0.38g per week of 46-0-0 fertilizer for the high nitrogen treatments. All plants were fully watered until one week before the sampling date. During that week, water was restricted to maintain designated soil moisture targets as measured with a HydroSense II Handheld Soil Moisture Sensor.

The first experiment was planted in February, 2021 and consisted of 72 plants grown in a fully interleaved design with three genotypes (Hybrids B73 X Mo17, P1105AM, and DTMA-205 a cross of tropical and temperate germplasm), two nitrogen treatments (50mL solution per week of 200 ppm and 25 ppm nitrogen fertilizer), and four watering treatments based on soil moisture content (<5% Threshold for leaf rolling, 6-8% Stress observed without leaf rolling, 10-12% No stress observed, >30% Fully watered 2hr before sampling). There were three replicates for each GxNxW treatment.

The second experiment was planted in July, 2021 and consisted of 120 plants grown under the same experimental design with the addition of two inbred genotypes (Hybrids B73 X Mo17, P1105AM, and DTMA-205. Inbreds B73 and B97) The hybrids selected include tropical and temperate germplasm with varying levels of drought adaptation. The inbreds include the reference genotype B73 and B97, which has an altered stomatal density phenotype.

## 3.2 Handheld Hyperspectral Device and Plant Sampling

For each experiment, plants were grown in three blocks with full within block randomization. Experiment one was not randomized until one week before imaging, but experiment two was randomized at the time of sowing. Plants were sampled over three successive days at the V7-V8 stage, with one experimental block sampled per day.

For each measurement day, all measurements were conducted within the four hour period centered on solar noon. For each plant, the top collared leaf was scanned using the LeafSpec developed by the Purdue Phenotyping lab group (Wang et al., 2020). The LeafSpec has a spectral range of 450 to 950nm with a spectral resolution of 0.74nm. It is a push broom imager with 224 pixels per line with the number of samples determined by an encoder wheel that rotates as the leaf slides through the device. A typical maize leaf will trigger approximately 500 imaging samples. In real time, a sample image is sent to a smartphone application for verification before the data is stored and the plant is destructively sampled for relative water content.

## 3.3 Measurement of Leaf Relative Water Content

For each leaf, six tissue samples were collected along the leaf with even spacing from base to tip as shown in Figure 3. Leaf with markings showing the location of tissue samples for relative water content.. This was done to enable future work on the distribution of water content along the leaf.



Figure 3. Leaf with markings showing the location of tissue samples for relative water content.

For each sample, two cuts were made perpendicular to the midrib, spaced 1 cm apart. The segment of leaf is then ripped away from the plant and weighed to obtain leaf fresh weight (FW). Then the sample was immediately hydrated for six hours until it was completely turgid under normal room lighting and temperature. Leaf samples were blotted dry on a paper towel then

weighed to obtain the fully turgid weight (TW). Finally, the leaf sample was fully dried with the dry oven (140 °C) for at least 36 h to obtain the dry weight (DW). An analytical balance was used for weighing to ensure accurate measurements with the leaf segments. To obtain the whole leaf RWC needed for this work, the six sample weights were summed into a whole leaf fresh weight, turgid weight, and dry weight then RWC was computed according to Equation 1.

## 3.4 Image Processing and Segmentation

After collection, the raw hyperspectral images were calibrated relative to the signal from the empty device as a reference according to Equation 2. This normalization helps correct for the non-uniform lighting generated by the LEDs inside the LeafSpec.

## Equation 2. Hyperspectral Normalization

$$I_{Calibrated} = \frac{I_{Raw}}{I_{White Reference}}$$

After calibration, plant tissue was segmented from background using an intensity threshold in the 450nm and 680nm bands, which worked due to the occlusion of blue light by the plant tissue. This method was an improvement over the more traditional NDVI threshold segmentation because the difference in nitrogen treatments made setting an NDVI threshold ineffective. Segmentation results were manually verified and segmentation thresholds tuned until satisfactory quality was achieved. Figure 4. Reconstructed RGB image from a hyperspectral leaf scan after calibration and segmentation. shows an RGB image of a leaf reconstructed from a hyperspectral scan after calibration and segmentation.



Figure 4. Reconstructed RGB image from a hyperspectral leaf scan after calibration and segmentation.

Due to the difference in leaf sizes, scan lengths ranged from 300 samples for leaves from inbred plants with low nitrogen fertilizer to over 600 samples for hybrid leaves with high nitrogen fertilizer. To accommodate this difference, scans were rescaled to a common size of 110 by 220 pixels using the MATLAB imresize() function. Figure 5. Hyperspectral scan at full resolution (top) and after resizing (bottom). shows a hyperspectral leaf scan at full resolution and after resizing. Note that the veins and other leaf features are still visible in the resized image.



Figure 5. Hyperspectral scan at full resolution (top) and after resizing (bottom).

Spectral indices have been commonly used to enhance the performance of prediction models relative to predictions from intensities based on individual wavelengths. NDVI was calculated according to Equation 3 and a heatmap of NDVI for a high nitrogen and low nitrogen leaf is shown in Figure 6. Heatmap of NDVI intensity for a leaf with high nitrogen (left) and low nitrogen (right) fertilizer treatment. Color scale is the same for both images.

Most spectral indices related to water require hyperspectral data from SWIR wavelengths, which the LeafSpec does not currently capture. In unpublished work, a member of our lab developed a water index (WI) utilizing the wavelengths the LeafSpec can currently capture. This

index is defined in Equation 4 and a heatmap of intensity values is shown in Figure 7. Heatmap of WI intensity for a leaf with RWC=0.91 (left) and RWC=0.71 (right). Color scale is the same for both images

There is a noticeable difference between the intensity of the high water and low water leaf. However, the midrib shows the lowest intensity even though it is the wettest part of the leaf indicating that the WI may not be applicable to all tissue types. Both the NDVI and WI images were included as additional wavelengths in all analyses.

Equation 3. Formula to calculate NDVI.

$$NDVI = \frac{I_{800nm} - I_{650nm}}{I_{800nm} + I_{650nm}}$$



Figure 6. Heatmap of NDVI intensity for a leaf with high nitrogen (left) and low nitrogen (right) fertilizer treatment. Color scale is the same for both images.

Equation 4. Formula to calculate the Water Index for LeafSpec.

$$WI = \frac{(I_{552nm})^2 x I_{722nm} + I_{722nm}}{I_{552nm} x I_{680nm}}$$



Figure 7. Heatmap of WI intensity for a leaf with RWC = 0.91 (left) and RWC = 0.71 (right). Color scale is the same for both images.

## 3.5 Model Development, Variable Selection, and Accuracy Comparison

For a given set of features, PLSR modeling was used to predict relative water content of the leaves. PLSR was chosen as the modeling method because it was designed to deal with highly correlated predictors as was observed in hyperspectral data and it was less susceptible to overtraining than other machine learning algorithms (Wold et al., 2001). PLSR models were implemented in MATLAB using the Eigenvector Toolbox (Eigenvector, n.d.). The number of components for each model was determined using the built-in recommendation, which is based on a score calculated from the reduction in RMSECV provided by each component. The PLS model generates a Variable Importance in Projection score for each predictor variable, which was used to investigate the predictors and conduct variable selection using the algorithm built into the toolbox.

Machine learning methods such as PLSR are susceptible to overtraining, especially when the number of predictor features is greater than the number of datapoints. To validate each model and assess overtraining, the data was split into a training set (2/3) and test set (1/3). The training set was used to calibrate the PLSR model using a 10-fold Venetian blind cross validation procedure. The calibrated model was then applied to the test data to generate a set of predicted RWC. The plot of the measured RWC vs the predicted RWC was fit with a line to create a R<sup>2</sup>Prediction. If the predicted RWC values perfectly match the measured values then R<sup>2</sup>Prediction = 1. Figure 8. Plot of measured RWC vs predicted RWC. Red line is the least squares best fit. R<sup>2</sup>Prediction = 0.691 shows an example plot of measured RWC vs predicted RWC for a PLSR model using the mean spectrum from each leaf as the set of predictor features. If a model was overtrained, it would have a high R<sup>2</sup>Cross Validation but a low R<sup>2</sup>Prediction.

To determine if two models differ in their accuracy, they are compared based on their  $R^2P$ rediction values. However, the arbitrary division in the training and test datasets introduces the potential for bias, so the modeling exercise was repeated for a chosen number of times with randomly selected training and test datasets to produce a sample of  $R^2P$ rediction values. The two samples for each model can be compared using standard hypothesis testing to determine if the samples came from populations with different means. For this work, fifteen data splits were tested for each model comparison and the  $R^2P$ rediction samples were compared using the standard t-test procedure.



Figure 8. Plot of measured RWC vs predicted RWC. Red line is the least squares best fit.  $R^{2}Prediction = 0.691$ 

#### **3.6** Spatial Pattern Feature Development

To develop features that capture the spatial pattern of water, we first sought to visually understand how the spatial pattern of hyperspectral intensity varied across leaves with different water contents. Heatmaps of intensity at each wavelength were generated for each leaf and then sorted based on the measured RWC from each leaf. The images were visually studied for patterns similar to the nutrient stress symptoms shown in Figure 1 and those reported in similar work with nitrogen (Ma et al., 2020). While no patterns were immediately obvious from the single wavelength images nor the NDVI or WI heatmaps, this helped generate several hypotheses for further investigation. For each hypothesis, an algorithm was developed to calculate relevant features from the hyperspectral scans, a prediction model was built using the features, and the prediction accuracy was compared to the prediction based on the mean spectrum model using the hypothesis test defined in the model development section.

## 3.6.1 Mean Spectrum

The state-of-the-art method for predicting leaf RWC from hyperspectral data is to average hyperspectral data across all plant pixels, build a PLSR model, and assess model performance using a training/test split on the data (Ge et al., 2019). This approach was implemented to validate the data collected in this work and to serve as a baseline for comparison. Spectral data was preprocessed using a Transmission to Absorbance log(1/T) operation followed by autoscaling. Four latent variables were included and model performance was assessed as the mean performance across 50 iterations.

### **3.6.2** Multiple Locations on the Leaf

Hypothesis: Utilizing spectral information from multiple regions in the leaf (base, middle, tip) will have improved predictive performance relative to the spectral information averaged over the entire leaf.

To better understand which leaf regions were the most correlated with whole leaf RWC, predictive PLSR models were built using the spectral information at each pixel location on the hyperspectral scan. Since all scans were resized to the same dimensions, each pixel refers to the same relative position on the leaf (e.g. When the scans have been resized to 110x200, pixel 55x100

represents a sample half way down the midrib and half way in the cross-sectional dimension for all leaves.). Five latent variables were included in the models and spectral data was preprocessed using a Transmission to Absorbance log(1/T) operation followed by autoscaling. Model performance was assessed using the mean performance after five modeling iterations due to limitations on computational power and a heatmap was generated using the mean R<sup>2</sup>Prediction for each pixel. This process was repeated using images downscaled by a factor of two until the images became a single pixel, which is the same as the mean spectrum.

To combine spectral information from multiple locations, the spectral data was appended and the larger predictor set used in a PLSR model to predict RWC as described previously. Every combination of pixels was investigated from leaf scans resized to 7x13 over three modeling iterations due to computational limitations. The number of latent variables included in each model was six. The R<sup>2</sup>Prediction performance was compared for each combination relative to the best individual pixel and the mean spectrum model. Heatmaps of the best pixel locations were used to choose locations for verification with more modeling iterations.

Due to the large number of combinations, some were expected to outperform the mean spectral model simply due to random variability, especially because the number of modeling iterations was low to conserve computational resources. Based on the heatmap of successful individual pixels and combination pixels, three regions (base, middle, tip) were defined for further verification. Models for each region and every combination of the regions were built and compared to the performance of the mean spectrum model.

## 3.6.3 Tissue Types

Hypothesis: Utilizing spectral information from regions segmented by tissue type will improve prediction relative to the mean spectral signal from that region.

Midrib and green regions were segmented from the scans using predefined segmentation boundaries shown in Figure 9. Hyperspectral data from each region was averaged and used to create a PLSR model that was compared against the model for the whole leaf mean spectrum.



Figure 9. Segmentation regions for midrib and green sections.

The effect of different tissue types was investigated further within the green region. Veins were segmented from green tissue using the WI heatmap due to enhanced contrast between tissue types. The segment was binarized using the MATLAB function imbinarize() with the adaptive thresholding set to a sensitivity of 0.65. Holes in the binary image were filled using imfill() to obtain the vein mask. The green tissue mask was obtained by eroding the inverse of the vein mask with a linear structuring element four pixels in size. Segmentation results were visually inspected for every image. Results were usually satisfactory (Figure 22. Segmentation of veins and green tissue (green). B. Mean spectrum for vein tissue for every leaf. C. Mean spectrum for green tissue for every leaf.), but some images were washed out and quality segmentation was not possible.

## 3.6.4 Cross Sectional Pattern

Hypothesis: The distribution of hyperspectral intensity in a cross section parallel to the midrib or in a cross section perpendicular to the midrib varies based on water content in a way that can be used to predict relative water content.

Work done to predict nitrogen content from hyperspectral leaf scans demonstrated that evaluating the distribution of NDVI values from the base to the tip of the leaf using SVM modeling can improve classification between high and low nitrogen treatments relative to the performance of mean NDVI (Ma et al., 2020). For this work, cross sections perpendicular to the midrib were constructed by taking a section centered at the middle of the leaf covering 20% of the leaf length. The section was averaged along the midrib direction to create a representative perpendicular cross-section. The cross section parallel to the midrib was constructed by taking a section a quarter of the way down the leaf width covering 20% of the leaf width. The section was averaged in the direction perpendicular to the midrib to create a representative parallel cross section. The two sections are shown in Figure 10. The cross-sections were smoothed with a Gaussian filter with a standard deviation of 8 then appended to the mean spectrum. Models were built using PLSR and SVM and compared to the performance of the mean spectrum PLSR model.



Figure 10. Regions used to construct vertical and horizontal cross sections.

For every wavelength, the cross sections were plotted and colored based on three RWC groups (high, middle, low). These plots were visually inspected to look for differences in pattern beyond mean intensity that could be used to predict RWC. The only wavelength with clear differences were the WI cross sections perpendicular to the midrib. From the plot of WI cross section, several features were calculated: mean WI, max WI, min WI, value of WI at middle of the distribution, slope of WI distribution, coefficients of second order polynomial fit to WI distribution. These features were appended to the mean spectrum and PLSR model performance was compared to using the mean spectrum alone.

## 3.6.5 Contours

Hypothesis: A contour pattern of hyperspectral intensity exists from the outside edge of the leaf to the inside that can be used to predict relative water content.

Nutrient stress patterns often create a color gradient from the edge of the leaf inward towards the center and a contour map could quantify that gradient in a way that is informative. To build the contour map for each leaf, the binary segmentation mask was sequentially eroded using MATLAB's imerode() function with a disk structuring element with a ten pixel diameter. Four erosion iterations were conducted to create five contour levels (Figure 11. Contour map from a representative leaf.). The contour map was applied to the images at each wavelength and the regionprops() function was used to obtain the mean intensity for each contour level. At every wavelength, the contour distribution for each leaf was plotted colored based on the leaf RWC. These plots were visually inspected for trends that could be used to predict RWC. Additionally, for every wavelength contour distribution was appended to the mean spectral data and used to build a PLSR model to predict RWC. Performance was compared to a model based on the mean spectral data alone to determine if the contour distribution improved prediction quality.



Figure 11. Contour map from a representative leaf.

## 3.6.6 Texture

Hypothesis: Differences in water content create a difference in visual texture that can be used to predict leaf relative water content.

There are many ways to analyze image texture. Two widely used algorithms, gray-level co-occurrence matrix (GLCM) and local binary pattern (LBP), have been implemented in MATLAB. GLCM calculates the frequency that pairs of pixels with specific intensities exist in specific spatial relationships, organizes the frequencies in a matrix, and extracts four statistics from the matrix – contrast which measures local variations in the matrix, correlation which measures

the probability of occurrence for specific pixel pairs, energy which is the sum of squared elements in the matrix, and homogeneity which quantifies the distribution of elements in the matrix (*Texture Analysis Using the Gray-Level Co-Occurrence Matrix* (*GLCM*) - *MATLAB* & *Simulink*, n.d.). LBP applies an operation to every pixel in an image that quantifies the relation to neighboring pixels then returns a histogram of the values (Rosebrock, 2015). Both algorithms were applied to the middle leaf segment shown in Figure 20. To determine if the traits improved predictive value, they were appended to the mean spectrum data and used to build a PLSR model to predict RWC. Performance was compared to a model based on the mean spectral data alone to determine if the texture data improved prediction quality.

## 4. **RESULTS**

## 4.1 Experimental Verification

Differing levels of water stress were successfully maintained in the experiment based on soil moisture content monitoring. Figure 12 shows the soil moisture content boxplots for each water treatment at the time of sampling. A clear gradient of soil moisture was created across the treatment groups. The mean for each treatment group falls within the desired range, but the quartiles do not. This is acceptable given the limitations of hand watering and variability in daily transpiration rates. The purpose of the watering treatments was to create leaves with variation in leaf water content. Since the leaf RWC was measured directly, the absolute value of the soil moisture content is not essential.



Figure 12. Water treatment groups and soil moisture content measured at sampling.

A plot of leaf RWC and soil moisture content at the time of measurement is shown in Figure 13 with a replotting of the data colored by several experimental factors shown in Figure 14. Plants in pots with high soil moisture had leaf moisture contents in the range of 90-100%. One data point exceeded 100% RWC, but this outlier is due to experimental error. Moisture content of the leaves did not drop unless the soil moisture content decreased below 10%, at which point leaf RWC decreased rapidly. At soil moisture contents less than 5%, leaf rolling began and plants looked severely water stressed. Because of the challenge of maintaining soil moisture in this narrow range with hand watering and unpredictable transpiration rates influenced by daily sunlight intensity, the data is highly skewed towards plants with unstressed moisture contents. The differentiation between nitrogen treatments was significant, with a clear difference in leaf color and plant size. Plants in the low nitrogen treatment group rarely experienced enough water stress to exhibit reduced leaf RWC, so they were not included in the rest of analysis. Leaf moisture content was evenly distributed across genotype and measurement day.



Figure 13. Plot of soil volumetric water content vs leaf relative water content at the time of measurement.



Figure 14. Plot of soil volumetric water content vs leaf relative water content at the time of measurement. Each sub-plot is colored by the experimental factor in the sub-plot title.

Of the 96 plants that received the high nitrogen fertilizer treatment, ten had scans with poor quality leaving a total of 86 samples for model development.

#### 4.2 Mean Spectrum Model

The mean spectrum model was able to predict leaf RWC with an R2Prediction of 0.7989 across diverse genetics and experiments. Model performance metrics after 50 modeling iterations are shown in Table 1. Slight overtraining is present as expected, but model performance is robust when applied to unseen data in the test set. The VIP score plot is shown in Figure 15. Variable Importance in Projection plot for the mean spectrum PLSR model.. The most noticeable peaks are

in the range of 500-550nm and the red-edge of 650-700nm with several other regions contributing to the model.



Figure 15. Variable Importance in Projection plot for the mean spectrum PLSR model.

## 4.3 Multiple Locations on the Leaf

Regions on the leaf have higher predictive performance than others. Figure 16. Heatmap of R<sup>2</sup>Prediction at each pixel location for resized leaf scans. shows a heatmap of R2Predicted values for models built using the spectrum from individual pixel locations in resized leaf scans. The green leaf tissue in the middle of the leaf was most correlated with whole leaf RWC while the midrib and leaf tip show less correlation. It is not surprising that the midrib was not as correlated given that the physical RWC measurements were made on green tissue samples. Figure 17. Heatmaps of R2Prediction at each pixel location over resizing iterations. The plot shows the maximum R2Prediction for an individual pixel at each resizing. The mean spectrum model is represented by

resizing eight. shows a plot of the maximum R2Prediction for an individual pixel over several resizing iterations. In the high resolution heatmaps, individual pixels beat the performance of the mean spectrum model. However, the performance across the entire leaf improves as pixel resolution decreases. This could be due to reduction in noise created by leaf texture variation. Interestingly, model performance peaked at resizing iteration four (shown in Figure 16. Heatmap of R<sup>2</sup>Prediction at each pixel location for resized leaf scans.) and then converged to the performance obtained by the mean spectrum model. This indicates that there may be a balance between averaging hyperspectral data to reduce noise and maintaining a high enough resolution to stay within the most informative region of the leaf.



Figure 16. Heatmap of R<sup>2</sup>Prediction at each pixel location for resized leaf scans.



Figure 17. Heatmaps of R2Prediction at each pixel location over resizing iterations. The plot shows the maximum R2Prediction for an individual pixel at each resizing. The mean spectrum model is represented by resizing eight.

Even though the spectral data from tip of the leaf was less correlated with RWC than spectral data from the middle of the leaf, it may combine useful information that is able to improve prediction when added to the spectral data from another region. Figure 18 shows the plot of R2Prediction for every pixel combination relative to the R2Prediction for the best individual pixel and whole leaf mean spectrum model. Some combinations did show improved performance, but this is expected due to the large number of tests and variance in the data. However, the pairs that did have improved performance consistently came from two regions as shown in Figure 19. Pixels from best combination are grouped in two regions. A heatmap of each region is shown.. While the two regions overlap, one is centered towards the base of the leaf and the other is centered towards the middle of the leaf. None of the best combinations included pixels from the leaf tip.



Figure 18. Plot of R2Prediction for every pixel combination relative to the R2Prediction for the best individual pixel and whole leaf mean spectrum model.

Three regions (base, middle, tip shown in Figure 20) were chosen for further analysis based on the heatmaps in Figure 17 and Figure 19. Spectral data was averaged in each region and model performance was calculated using the spectral data from each region individually and for every combination of regions (shown in Table 1). Model accuracies were compared for a difference relative to the mean spectrum model using a t-test on the R2Prediction value from 50 iterations. None of the regions nor combinations outperformed the mean spectrum model at a significant level. All of the individual regions showed worse performance than the mean spectrum model. Only the combination of all three regions was able to outperform the mean spectrum level, but the gain was very small and not statistically significant. This indicates that the previously observed pixels and combinations that outperformed were the result of statistical noise from testing large numbers of samples.



Figure 19. Pixels from best combination are grouped in two regions. A heatmap of each region is shown.



Figure 20. Regions selected for further analysis (base, middle, tip) superimposed on a representative leaf.

Table 1. Model performance for multiple locations on the leaf. The p-value is the result	of the t-
test for difference relative to the R2Prediction for the mean spectrum model.	

Model	<b>R<sup>2</sup>Calibration</b>	<b>R<sup>2</sup>Cross Validation</b>	<b>R</b> <sup>2</sup> <b>Prediction</b>	p-value
Mean Spectrum	0.8302	0.7699	0.7989	
Base	0.8040	0.7407	0.7638	0.1051
Middle	0.7305	0.6176	0.6182	8.50e-14
Tip	0.7311	0.6188	0.6205	9.41e-13
Base + Middle	0.8697	0.7993	0.7993	0.0403
Base + Tip	0.8123	0.7061	0.7115	2.26e-05
Middle + Tip	0.8343	0.7370	0.7736	0.0789
Base + Middle + Tip	0.8785	0.7859	0.8053	0.1855

## 4.4 Tissue Type

The midrib and green regions of the leaf have clearly different spectrum (Figure 21). The LeafSpec transmits light from LEDs, through the leaf, to the imaging sensor, so the thicker midrib has lower intensity than the green leaf. The spectrum from 700-800nm is distinctly different between the tissue types which could be due to the presence of chlorophyll in the green tissue.



Figure 21. Mean spectrum for midrib and green region of leaves.

The highest resolution heatmap in Figure 17 shows a clear pattern of veins and green tissue, implying that one of the tissue types is better for predicting RWC. Figure 22 shows the spectrum of vein and green tissue for each leaf after segmentation. While the segmentation results appear to be good, there is not a noticeable difference in the spectrum among the tissue types from the same region.



Figure 22. Segmentation of veins and green tissue. A. Representative region of leaf with segmentation masks for veins (blue) and green tissue (green). B. Mean spectrum for vein tissue for every leaf. C. Mean spectrum for green tissue for every leaf.

The model results for the different tissue types are shown in Table 2.. No model outperformed the whole leaf mean spectrum model though the models from the green region were not significantly different. This provides further evidence that the whole leaf spectrum reduces noise levels and provides more consistent prediction accuracy. Models that included midrib spectral data had a decrease in performance indicating that spectral data from the midrib over 450-950nm is not as useful for predicting leaf RWC. This is expected given that the physical RWC measurements were made from the green tissue.

Model	<b>R<sup>2</sup>Calibration</b>	<b>R<sup>2</sup>Cross Validation</b>	<b>R</b> <sup>2</sup> <b>Prediction</b>	p-value
Mean Spectrum	0.8302	0.7699	0.7989	
Midrib	0.7394	0.6226	0.6708	9.92e-12
Green Region	0.8225	0.7650	0.7854	0.4407
Midrib + Green Tissue	0.8299	0.7518	0.7635	0.1508
Veins from Green	0.7845	0.6994	0.7404	0.0011
Region				
Green Tissue from	0.8145	0.7562	0.7563	0.1381
Green Region				
Veins+Green Tissue	0.8265	0.7397	0.7766	0.0403
Veins + Green Tissue	0.8292	0.7447	0.7726	*0.2253
compared to Green				
Region Model				

Table 2. Model Performance for various tissue types.

## 4.5 Cross Sectional Pattern

Model performance was not improved by including hyperspectral distribution data in addition to the mean spectrum. Figures 23 and 24 show the model performance for a PLSR model with distribution data appended at every wavelength. None of the models with higher R2Prediction performed at a level that was statistically different than the mean spectrum model. For the perpendicular cross sections, models with data from 750nm and above seemed to perform better than the models at smaller wavelengths, but the trend was not statistically significant. SVM regression had been successfully applied to NDVI distribution data for nitrogen prediction (Ma et al., 2020). When applied in this work for prediction of RWC, no SVM model outperformed the mean spectrum PLSR model.



Figure 23. Model performance for a PLSR model with mean spectral data and the perpendicular distribution for an individual wavelength. The line is the performance of a model with only the mean spectral data.



Figure 24. Model performance for a PLSR model with mean spectral data and the parallel distribution data for an individual wavelength. The line is the performance of a model with only the mean spectral data.

Visual inspection of the cross section plots of individual wavelengths did not show clear patterns that could be used to separate leaves based on RWC (Figure 25). However, the cross-section plot of WI perpendicular to the midrib shown in **Error! Reference source not found.** had clear differences among RWC levels. Individual WI features were able to predict about 30% of the variance in water content. Table 3 shows the performance of PLSR model when the calculated WI features are appended to the mean spectrum. None of the models outperformed the mean spectrum alone.

Table 3. Model performance for WI features. The p-value is the result of the t-test for difference relative to the R2Prediction for the mean spectrum model.

Model	<b>R<sup>2</sup>Calibration</b>	<b>R<sup>2</sup>Cross Validation</b>	<b>R</b> <sup>2</sup> <b>Prediction</b>	p-value
Mean Spectrum (MS)	0.8302	0.7699	0.7989	
MS + WI Max	0.8318	0.7729	0.7921	0.7970
MS + WI Min	0.8379	0.7801	0.7802	0.9343
MS + WI Mean	0.8286	0.7658	0.8028	0.8274
MS + WI Middle	0.8375	0.7762	0.7738	0.8578
MS + WI Slope	0.8433	0.7854	0.7845	0.8190
MS + WI Poly Coeff	0.8359	0.7687	0.7871	0.5321
MS + WI All Traits	0.8460	0.7756	0.7639	0.3608



Figure 25. Vertical Cross Section of Water Index. Each line represents a leaf and has been colored based on RWC. Blue is greater than 0.87 RWC, green is less than 0.87 RWC and greater than 0.77 RWC, and red is less than 0.77 RWC.

### 4.6 Contours

Visual inspection of the contour distributions at each wavelength showed differences in mean intensity but did not reveal any patterns that could be used to distinguish between leaves based on RWC. This can be seen in the contour plot for the WI shown in Figure 26, where each line is from a different leaf and is colored based on three RWC classes. The mean value for each leaf is different and can be used to identify RWC class, but the shape of the curve is the same across the different RWC levels.



Figure 26. Contour distribution for water index intensity values. One is the outermost contour and five is the innermost contour. Each line represents a leaf and has been colored based on RWC. Blue is greater than 0.87 RWC, green is less than 0.87 RWC and greater than 0.77 RWC, and red is less than 0.77 RWC.

To verify that the contours did not contain additional predictive value beyond the mean intensity, the contour distribution for every wavelength was appended to the mean spectral data and used to build a PLSR model to predict RWC. Performance was compared to a model based on

the mean spectral data alone. Figure 27. Model performance for a PLSR model with mean spectral data and the contour data for an individual wavelength. The horizontal line is the performance of a model with only the mean spectral data. plots the R2Prediction value for each model at the wavelength for the contour distribution.. None of the models with additional contour data differed from the mean spectrum model at a statistically significant level.



Figure 27. Model performance for a PLSR model with mean spectral data and the contour data for an individual wavelength. The horizontal line is the performance of a model with only the mean spectral data.

## 4.7 Texture

Performance for models based on each GLCM texture trait are shown in Table 4. The individual traits show some predictive value, but there is severe overtraining. When combined with

the mean spectrum data, models that include the texture data outperform during calibration but suffer performance drops when used to predict the RWC of new samples. This indicates that the GLCM texture data is inducing overtraining and does not improve the true predictive capability of the RWC model.

Model	<b>R</b> <sup>2</sup> Calibration	<b>R<sup>2</sup>Cross Validation</b>	<b>R<sup>2</sup>Prediction</b>	p-value
Mean Spectrum	0.8302	0.7699	0.7989	
Contrast	0.3133	0.1038	0.1147	1.06e-64
Correlation	0.2391	0.0530	0.0852	5.98e-75
Energy	0.8625	0.0654	0.1154	1.46e-60
Homogeneity	0.5828	0.2189	0.3213	6.08e-39
Mean Spectrum &	0.8602	0.7183	0.7273	0.0015
Contrast				
Mean Spectrum &	0.8729	0.7170	0.7262	0.0011
Correlation				
Mean Spectrum &	0.9449	0.5642	0.6418	8.16e-11
Energy				
Mean Spectrum &	0.8716	0.7081	0.7495	0.0011
Homogeneity				

Table 4. Model Performance with GLCM Texture Features.

The LBP algorithm produces 59 features for each greyscale image segment analyzed using the default MATLAB parameters. For every wavelength, the LBP features were appended to the mean spectrum and model performance from the combined predictor set was compared to the performance of the mean spectrum alone (Figure). None of the models that performed better than the mean spectrum model did so at a statistically significant level.



Figure 28. Model performance for a PLSR model with mean spectral data and the LBP data for an individual wavelength. The horizontal line is the performance of a model with only the mean spectral data.

## 5. DISCUSSION AND CONCLUSIONS

Prediction of plant relative water content using hyperspectral data has proven to be a robust and effective method. In this work, hyperspectral leaf scans were collected using the LeafSpec, a recently developed hyperspectral imaging device. Data were collected across two experiments from five different maize genotypes representing temperate and tropical hybrids with varying levels of drought tolerance and inbreds with varying stomatal densities. The hyperspectral intensity averaged across the entire leaf was used to predict relative water content with an R<sup>2</sup>Prediction of 0.7989. This result is comparable to previous reports in the literature studying a narrow range of genotypes and much better than performance for models applied to the maize diversity panel. The improved quality of prediction from LeafSpec data may be due to removal of noise from leaf angle and external lighting conditions during imaging.

Distribution of hyperspectral information has been used to improve classification of leaves based on nitrogen treatment relative to classification based on mean hyperspectral data (Ma et al., 2020). We explored the hypothesis that prediction of leaf relative water content could be improved over the mean hyperspectral PLSR models by developing features that quantify:

- Spectral information from multiple regions in the leaf (e.g. base, middle, tip).
- Spectral information from regions segmented by tissue type.
- The distribution of hyperspectral intensity in a cross section parallel to the midrib or in a cross-section perpendicular to the midrib.
- A contour pattern of hyperspectral intensity from the outside edge of the leaf to the midrib.
- Texture features extracted from each wavelength.

PLSR predictive performance with the new features was compared to the mean hyperspectral model over multiple testing iterations. None of the models with expanded feature sets outperformed the mean spectrum model at a statistically significant level.

In exploring various regions of the leaf, it was found that the hyperspectral signal from the green tissue a third of the way from the base of the leaf and half way between the midrib and edge was the most correlated with relative water content. This location may be biased by the RWC sampling method used during the experiment, but confirms the standard procedure for spectrometer sampling (Xiong et al., 2015).

Midrib and vein tissue models performed worse at a statistically significant level indicating that there may be some value to segmenting them out from the green tissue. However, the green tissue model with midrib and vein data removed did not perform better than the mean spectrum model at a statistically significant level indicating that the extra segmentation may not be worth the effort. Averaging the hyperspectral data across the entire leaf creates a spectrum very similar to the spectrum from green leaf tissue alone. Cross-sections of hyperspectral intensity from individual wavelengths did not improve model performance and patterns that separate leaves by RWC could not be visually discerned. The Water Index cross-section perpendicular to the midrib visibly separated leaves by RWC and a set of features were developed to quantify the pattern. None of the features improved model performance relative to the mean spectral model. Contour patterns did not improve model performance. Visual inspection of the contour plots revealed that leaves with different RWC had similarly shaped distributions of hyperspectral intensity that simply differed in mean intensity level. Thus, the mean hyperspectral data should contain equivalent predictive information as the hyperspectral contours. Models that included the texture features had improved calibration performance, but reduced predictive performance. This indicates that the addition of texture features induces overtraining and does not improve the model's ability to predict RWC.

For future work, more data should be collected to improve model robustness. At least several thousand datapoints are probably needed to build a robust model. Hyperspectral data generates large numbers of predictors which makes machine learning methods susceptible to overtraining. Generating leaves with low RWC is a challenge due to the difficulty of properly dosing water with unpredictable rates of transpiration. In future work, the experimental design should be biased toward lower soil moisture contents to ensure a more even spread of leaf RWC.

Distribution of hyperspectral data may still hold opportunity for improving RWC prediction. Previously reported spectral indices for leaf water content frequently include SWIR wavelengths. The current version of the LeafSpec is not capable of capturing SWIR, but a high resolution leaf scan in these wavelengths may reveal patterns of water distribution. Additionally, exploring more indices composed from current spectral bands may lead to an interesting finding. The Water Index was a first step and visually showed improved ability to discriminate leaf RWC as compared to individual wavelengths.

## REFERENCES

- Acevedo, E., Hsiao, T. C., & Henderson, D. W. (1971). Immediate and Subsequent Growth Responses of Maize Leaves to Changes in Water Status. *Plant Physiology*, 48(5), 631–636. https://doi.org/10.1104/pp.48.5.631
- Agricultural output—Crop production—OECD Data. (2021, February 7). TheOECD. http://data.oecd.org/agroutput/crop-production.htm
- Allah, A. A. (2009). Genetic studies on leaf rolling and some root traits under drought conditions in rice (Oryza sativa L.). *African Journal of Biotechnology*, 8(22), Article 22. https://doi.org/10.4314/ajb.v8i22.66127
- Alvarez, S., Marsh, E. L., Schroeder, S. G., & Schachtman, D. P. (2008). Metabolomic and proteomic changes in the xylem sap of maize under drought. *Plant, Cell & Environment*, 31(3), 325–340. https://doi.org/10.1111/j.1365-3040.2007.01770.x
- Bankole, F., Menkir, A., Olaoye, G., Crossa, J., Hearne, S., Unachukwu, N., & Gedil, M. (2017). Genetic Gains in Yield and Yield Related Traits under Drought Stress and Favorable Environments in a Maize Population Improved Using Marker Assisted Recurrent Selection. *Frontiers in Plant Science*, 8, 808. https://doi.org/10.3389/fpls.2017.00808
- Bänziger, M., & Lafitte, H. R. (1997). Efficiency of Secondary Traits for Improving Maize for Low-Nitrogen Target Environments. Crop Science, 37(4), cropsci1997.0011183X003700040013x. https://doi.org/10.2135/cropsci1997.0011183X003700040013x
- Barbosa, P. A. M., Fritsche-Neto, R., Andrade, M. C., Petroli, C. D., Burgueño, J., Galli, G., Willcox, M. C., Sonder, K., Vidal-Martínez, V. A., Sifuentes-Ibarra, E., & Molnar, T. L. (2021). Introgression of Maize Diversity for Drought Tolerance: Subtropical Maize Landraces as Source of New Positive Variants. *Frontiers in Plant Science*, *12*, 2023. https://doi.org/10.3389/fpls.2021.691211
- Baret, F., Madec, S., Irfan, K., Lopez, J., Comar, A., Hemmerlé, M., Dutartre, D., Praud, S., & Tixier, M. H. (2018). Leaf-rolling in maize crops: From leaf scoring to canopy-level measurements for phenotyping. *Journal of Experimental Botany*, 69(10), 2705–2716. https://doi.org/10.1093/jxb/ery071
- Barrs, H. D., & Weatherley, P. E. (1962). A Re-Examination of the Relative Turgidity Technique for Estimating Water Deficits in Leaves. *Australian Journal of Biological Sciences*, 15(3), 413–428. https://doi.org/10.1071/bi9620413

- Benešová, M., Holá, D., Fischer, L., Jedelský, P. L., Hnilička, F., Wilhelmová, N., Rothová, O., Kočová, M., Procházková, D., Honnerová, J., Fridrichová, L., & Hniličková, H. (2012). The Physiology and Proteomics of Drought Tolerance in Maize: Early Stomatal Closure as a Cause of Lower Tolerance to Short-Term Dehydration? *PLOS ONE*, *7*(6), e38017. https://doi.org/10.1371/journal.pone.0038017
- Benson, L., Petersen, K., & Stein, J. (n.d.). ANASAZI (PRE-COLUMBIAN NATIVE-AMERICAN) MIGRATIONS DURING THE MIDDLE-12TH AND LATE-13TH CENTURIES – WERE THEY DROUGHT INDUCED? 28.
- Bhat, J. A., Ali, S., Salgotra, R. K., Mir, Z. A., Dutta, S., Jadon, V., Tyagi, A., Mushtaq, M., Jain, N., Singh, P. K., Singh, G. P., & Prabhu, K. V. (2016). Genomic Selection in the Era of Next Generation Sequencing for Complex Traits in Plant Breeding. *Frontiers in Genetics*, 7, 221. https://doi.org/10.3389/fgene.2016.00221
- Bliss, L. C., Kramer, P. J., & Wolf, F. A. (1957). Drought Resistance 1• n Tobacco. *Tobacco Science*, 4.
- Bolaños, J., & Edmeades, G. O. (1996). The importance of the anthesis-silking interval in breeding for drought tolerance in tropical maize. *Field Crops Research*, 48(1), 65–80. https://doi.org/10.1016/0378-4290(96)00036-6
- Bruce, W. B., Edmeades, G. O., & Barker, T. C. (2002). Molecular and physiological approaches to maize improvement for drought tolerance. *Journal of Experimental Botany*, *53*(366), 13–25. https://doi.org/10.1093/jexbot/53.366.13
- Cairns, J. E., Sanchez, C., Vargas, M., Ordoñez, R., & Araus, J. L. (2012). Dissecting Maize Productivity: Ideotypes Associated with Grain Yield under Drought Stress and Wellwatered Conditions. *Journal of Integrative Plant Biology*, 54(12), 1007–1020. https://doi.org/10.1111/j.1744-7909.2012.01156.x
- Campbell, M. T., Knecht, A. C., Berger, B., Brien, C. J., Wang, D., & Walia, H. (2015). Integrating Image-Based Phenomics and Association Analysis to Dissect the Genetic Architecture of Temporal Salinity Responses in Rice. *Plant Physiology*, *168*(4), 1476–1489. https://doi.org/10.1104/pp.15.00450
- Čatský, J. (1959). The role played by growth in the determination of water deficit in plants. *Biologia Plantarum*, 1(4), 277. https://doi.org/10.1007/BF02920847
- Čatský, J. (1960). Determination of water deficit in disks cut out from leaf blades. *Biologia Plantarum*, 2(1), 76–78. https://doi.org/10.1007/BF02920701
- Ceccato, P., Flasse, S., & Grégoire, J.-M. (2002). Designing a spectral index to estimate vegetation water content from remote sensing data: Part 2. Validation and applications. *Remote Sensing of Environment*, 82(2), 198–207. https://doi.org/10.1016/S0034-4257(02)00036-6

- Cerrudo, D., Cao, S., Yuan, Y., Martinez, C., Suarez, E. A., Babu, R., Zhang, X., & Trachsel, S. (2018). Genomic Selection Outperforms Marker Assisted Selection for Grain Yield and Physiological Traits in a Maize Doubled Haploid Population Across Water Treatments. *Frontiers in Plant Science*, 9, 366. https://doi.org/10.3389/fpls.2018.00366
- Chai, L., Jiang, H., Crow, W. T., Liu, S., Zhao, S., Liu, J., & Yang, S. (2021). Estimating Corn Canopy Water Content From Normalized Difference Water Index (NDWI): An Optimized NDWI-Based Scheme and Its Feasibility for Retrieving Corn VWC. *IEEE Transactions on Geoscience and Remote Sensing*, 59(10), 8168–8181. https://doi.org/10.1109/TGRS.2020.3041039
- Chandel, N. S., Chakraborty, S. K., Rajwade, Y. A., Dubey, K., Tiwari, M. K., & Jat, D. (2020). Identifying crop water stress using deep learning models. *Neural Computing and Applications*. https://doi.org/10.1007/s00521-020-05325-4
- Chapman, S. C., & Edmeades, G. (1999). Selection Improves Drought Tolerance in Tropical Maize Populations: II. Direct and Correlated Responses among Secondary Traits. *Crop Science - CROP SCI*, 39. https://doi.org/10.2135/cropsci1999.3951315x
- Cheng, Y.-B., Ustin, S. L., Riaño, D., & Vanderbilt, V. C. (2008). Water content estimation from hyperspectral images and MODIS indexes in Southeastern Arizona. *Remote Sensing of Environment*, 112(2), 363–374. https://doi.org/10.1016/j.rse.2007.01.023
- Cochard, H. (2002). Xylem embolism and drought-induced stomatal closure in maize. *Planta*, 215(3), 466–471. https://doi.org/10.1007/s00425-002-0766-9
- Cotrozzi, L., Peron, R., Tuinstra, M. R., Mickelbart, M. V., & Couture, J. J. (2020). Spectral Phenotyping of Physiological and Anatomical Leaf Traits Related with Maize Water Status. *Plant Physiology*, *184*(3), 1363–1377. https://doi.org/10.1104/pp.20.00577
- Craig, H., & Gordon, L. I. (1965). *Deuterium and Oxygen 18 Variations in the Ocean and the Marine Atmosphere*. Laboratorio di geologia nucleare.
- Edalat, M., Naderi, R., & Egan, T. P. (2019). Corn nitrogen management using NDVI and SPAD sensor-based data under conventional vs. Reduced tillage systems. *Journal of Plant Nutrition*, 42(18), 2310–2322. https://doi.org/10.1080/01904167.2019.1648686
- Effendi, R., Priyanto, S. B., Aqil, M., & Azrai, M. (2019). Drought adaptation level of maize genotypes based on leaf rolling, temperature, relative moisture content, and grain yield parameters. *IOP Conference Series: Earth and Environmental Science*, 270(1), 012016. https://doi.org/10.1088/1755-1315/270/1/012016
- Eigenvector. (n.d.). Chemometrics—Data Analysis Software—PLS\_Toolbox. *Eigenvector*. Retrieved December 6, 2021, from https://eigenvector.com/software/pls-toolbox/
- Elsayed, S., & Darwish, W. (2017). Hyperspectral remote sensing to assess the water status, biomass, and yield of maize cultivars under salinity and water stress. *Bragantia*, 76, 62–72. https://doi.org/10.1590/1678-4499.018

- Eze, C. E., Akinwale, R. O., Michel, S., & Bürstmayr, H. (2020). Grain yield and stability of tropical maize hybrids developed from elite cultivars in contrasting environments under a rainforest agro-ecology. *Euphytica*, 216(6), 89. https://doi.org/10.1007/s10681-020-02620-y
- FAO. (2018). *The impact of disasters and crises on agriculture and food security 2017*. Food and Agriculture Organization of the United Nations.
- Fariñas, M. D., Jimenez-Carretero, D., Sancho-Knapik, D., Peguero-Pina, J. J., Gil-Pelegrín, E., & Gómez Álvarez-Arenas, T. (2019). Instantaneous and non-destructive relative water content estimation from deep learning applied to resonant ultrasonic spectra of plant leaves. *Plant Methods*, 15(1), 128. https://doi.org/10.1186/s13007-019-0511-z
- Gan, K. S., Wong, S. C., Yong, J. W. H., & Farquhar, G. D. (2003). Evaluation of models of leaf water 18O enrichment using measurements of spatial patterns of vein xylem water, leaf water and dry matter in maize leaves. *Plant, Cell & Environment*, 26(9), 1479–1495. https://doi.org/10.1046/j.1365-3040.2003.01070.x
- Gao, Y., Qiu, J., Miao, Y., Qiu, R., Li, H., & Zhang, M. (2019). Prediction of Leaf Water Content in Maize Seedlings Based on Hyperspectral Information. *IFAC-PapersOnLine*, 52(30), 263–269. https://doi.org/10.1016/j.ifacol.2019.12.532
- Ge, Y., Atefi, A., Zhang, H., Miao, C., Ramamurthy, R. K., Sigmon, B., Yang, J., & Schnable, J. C. (2019). High-throughput analysis of leaf physiological and chemical traits with VIS– NIR–SWIR spectroscopy: A case study with a maize diversity panel. *Plant Methods*, 15. https://doi.org/10.1186/s13007-019-0450-8
- Ge, Y., Bai, G., Stoerger, V., & Schnable, J. C. (2016). Temporal dynamics of maize plant growth, water use, and leaf water content using automated high throughput RGB and hyperspectral imaging. *Computers and Electronics in Agriculture*, 127, 625–632. https://doi.org/10.1016/j.compag.2016.07.028
- Glover, J. (1959). The apparent behaviour of maize and sorghum stomata during and after drought. *The Journal of Agricultural Science*, 53(3), 412–416. https://doi.org/10.1017/S0021859600020839
- Hay, J. O., Moulia, B., Lane, B., Freeling, M., & Silk, W. K. (2000). Biomechanical analysis of the Rolled (RLD) leaf phenotype of maize. *American Journal of Botany*, 87(5), 625–633. https://doi.org/10.2307/2656848
- Higa, S., Kobori, H., & Tsuchikawa, S. (2013). Mapping of Leaf Water Content Using Near-Infrared Hyperspectral Imaging. *Applied Spectroscopy*, 67(11), 1302–1307. https://doi.org/10.1366/13-07028
- Janick, J. (2010). Plant Breeding Reviews, Part 2: Long-term Selection: Crops, Animals, and Bacteria. John Wiley & Sons.

- Kim, D. M., Zhang, H., Zhou, H., Du, T., Wu, Q., Mockler, T. C., & Berezin, M. Y. (2015). Highly sensitive image-derived indices of water-stressed plants using hyperspectral imaging in SWIR and histogram analysis. *Scientific Reports*, 5(1), 15919. https://doi.org/10.1038/srep15919
- Kranz, B. (2015, April 8). Understanding Crop Water Use. https://water.unl.edu/documents/Section%20I.pdf
- Landsat Science -. (n.d.). Retrieved December 3, 2021, from https://landsat.gsfc.nasa.gov/
- Lara, M. A., Diezma, B., Lleó, L., Roger, J. M., Garrido, Y., Gil, M. I., & Ruiz Altisent, M. (2016). Hyperspectral imaging to evaluate the effect of irrigation water salinity in lettuce. *Applied Sciences*, 6(12), 412. https://doi.org/10.3390/app6120412
- Le, A., S, T., Je, O., & Kk, T. (2011). Stomatal Responses to Drought Stress and Air Humidity. In *Abiotic Stress in Plants—Mechanisms and Adaptations*. IntechOpen. https://doi.org/10.5772/24661
- Liu, Y., Qin, L., Han, L., Xiang, Y., & Zhao, D. (2015). Overexpression of maize SDD1 (ZmSDD1) improves drought resistance in Zea mays L. by reducing stomatal density. *Plant Cell, Tissue and Organ Culture (PCTOC)*, 122(1), 147–159. https://doi.org/10.1007/s11240-015-0757-8
- Lobell, D. B., Deines, J. M., & Tommaso, S. D. (2020). Changes in the drought sensitivity of US maize yields. *Nature Food*, 1(11), 729–735. https://doi.org/10.1038/s43016-020-00165-w
- Lynch, J. P. (2013). Steep, cheap and deep: An ideotype to optimize water and N acquisition by maize root systems. *Annals of Botany*, *112*(2), 347–357. https://doi.org/10.1093/aob/mcs293
- Ma, D., Wang, L., Zhang, L., Song, Z., U. Rehman, T., & Jin, J. (2020). Stress Distribution Analysis on Hyperspectral Corn Leaf Images for Improved Phenotyping Quality. *Sensors*, 20(13), 3659. https://doi.org/10.3390/s20133659
- Maazou, A.-R. S., Tu, J., Qiu, J., & Liu, Z. (2016). Breeding for Drought Tolerance in Maize (Zea mays L.). American Journal of Plant Sciences, 7(14), 1858–1870. https://doi.org/10.4236/ajps.2016.714172
- Magorokosho, C., & Tongoona, P. (2003). Selection for drought tolerance in two tropical maize populations. *African Crop Science Journal*, 11(3), 151–161. https://doi.org/10.4314/acsj.v11i3.27566
- McDAVID, C. R., Sagar, G. R., & Marshall, C. (1973). The Effect of Root Pruning and 6-Benzyl-Aminopurine on the Chlorophyll Content, 14co2 Fixation and the Shoot/Root Ratio in Seedlings of Pisvm Sativum L. New Phytologist, 72(3), 465–470. https://doi.org/10.1111/j.1469-8137.1973.tb04395.x

- Merlivat, L., & Coantic, M. (1975). Study of mass transfer at the air-water interface by an isotopic method. *Journal of Geophysical Research* (1896-1977), 80(24), 3455–3464. https://doi.org/10.1029/JC080i024p03455
- Mertens, S., Verbraeken, L., Sprenger, H., Demuynck, K., Maleux, K., Cannoot, B., De Block, J., Maere, S., Nelissen, H., Bonaventure, G., Crafts-Brandner, S. J., Vogel, J. T., Bruce, W., Inzé, D., & Wuyts, N. (2021). Proximal Hyperspectral Imaging Detects Diurnal and Drought-Induced Changes in Maize Physiology. *Frontiers in Plant Science*, 12, 240. https://doi.org/10.3389/fpls.2021.640914
- Meseka, S., Fakorede, M., Ajala, S., Badu-Apraku, B., & Menkir, A. (2013). Introgression of Alleles from Maize Landraces to Improve Drought Tolerance in an Adapted Germplasm. *Journal of Crop Improvement*, 27. https://doi.org/10.1080/15427528.2012.729259
- Messmer, R., Fracheboud, Y., Bänziger, M., Vargas, M., Stamp, P., & Ribaut, J.-M. (2009). Drought stress and tropical maize: QTL-by-environment interactions and stability of QTLs across environments for yield components and secondary traits. *Theoretical and Applied Genetics*, 119(5), 913–930. https://doi.org/10.1007/s00122-009-1099-x
- Moeinizade, S., Han, Y., Pham, H., Hu, G., & Wang, L. (2021). A look-ahead Monte Carlo simulation method for improving parental selection in trait introgression. *Scientific Reports*, *11*(1), 3918. https://doi.org/10.1038/s41598-021-83634-x
- Murphy, R. J., Whelan, B., Chlingaryan, A., & Sukkarieh, S. (2019). Quantifying leaf-scale variations in water absorption in lettuce from hyperspectral imagery: A laboratory study with implications for measuring leaf water content in the context of precision agriculture. *Precision Agriculture*, 20(4), 767–787. https://doi.org/10.1007/s11119-018-9610-5
- Namken, L. N., & Lemon, E. R. (1960). Field Studies of Internal Moisture Relations of the Corn Plant1. *Agronomy Journal*, 52(11), 643–646. https://doi.org/10.2134/agronj1960.00021962005200110010x
- Oosterom, E., Yang, Z., Zhang, F., Deifel, K., Cooper, M., Messina, C., & Hammer, G. (2016). Hybrid variation for root system efficiency in maize: Potential links to drought adaptation. *Functional Plant Biology*, *43*. https://doi.org/10.1071/FP15308
- O'Toole, J. C., Cruz, R. T., & Singh, T. N. (1979). Leaf rolling and transpiration. *Plant Science Letters*, *16*(1), 111–114. https://doi.org/10.1016/0304-4211(79)90015-4
- Pandey, P., Ge, Y., Stoerger, V., & Schnable, J. C. (2017). High Throughput In vivo Analysis of Plant Leaf Chemical Properties Using Hyperspectral Imaging. *Frontiers in Plant Science*, 8. https://doi.org/10.3389/fpls.2017.01348
- PEÑUELAS, J., FILELLA, I., BIEL, C., SERRANO, L., & SAVÉ, R. (1993). The reflectance at the 950–970 nm region as an indicator of plant water status. *International Journal of Remote Sensing*, 14(10), 1887–1905. https://doi.org/10.1080/01431169308954010

- Pick, T. R., Bräutigam, A., Schlüter, U., Denton, A. K., Colmsee, C., Scholz, U., Fahnenstich, H., Pieruschka, R., Rascher, U., Sonnewald, U., & Weber, A. P. M. (2011). Systems Analysis of a Maize Leaf Developmental Gradient Redefines the Current C4 Model and Provides Candidates for Regulation. *The Plant Cell*, 23(12), 4208–4220. https://doi.org/10.1105/tpc.111.090324
- Raymond Hunt, E., Rock, B. N., & Nobel, P. S. (1987). Measurement of leaf relative water content by infrared reflectance. *Remote Sensing of Environment*, 22(3), 429–435. https://doi.org/10.1016/0034-4257(87)90094-0
- Rees, J. (2015, September 17). Crop Water Use Comparison of Rainfed Corn, Sorghum, and Soybean from 2009 to 2011. CropWatch. https://cropwatch.unl.edu/crop-water-usecomparison-rainfed-corn-sorghum-and-soybean-2009-2011
- Ribaut, J.-M., & Ragot, M. (2007). Marker-assisted selection to improve drought adaptation in maize: The backcross approach, perspectives, limitations, and alternatives. *Journal of Experimental Botany*, 58(2), 351–360. https://doi.org/10.1093/jxb/erl214
- Robertson, M. J., Fukai, S., Ludlow, M. M., & Hammer, G. L. (1993). Water extraction by grain sorghum in a sub-humid environment. I. Analysis of the water extraction pattern. *Field Crops Research*, 33(1), 81–97. https://doi.org/10.1016/0378-4290(93)90095-5
- Rosebrock, R. (2015, December 7). Local Binary Patterns with Python & OpenCV. *PyImageSearch*. https://www.pyimagesearch.com/2015/12/07/local-binary-patterns-with-python-opencv/
- Rouse, J. W., Haas, R. H., Schell, J. A., & Deering, D. W. (1974, January 1). *Monitoring vegetation* systems in the Great Plains with ERTS. https://ntrs.nasa.gov/citations/19740022614
- Ryckewaert, M., Gorretta, N., Henriot, F., Gobrecht, A., Héran, D., Moura, D., Bendoula, R., & Roger, J.-M. (2021). Potential of high-spectral resolution for field phenotyping in plant breeding: Application to maize under water stress. *Computers and Electronics in Agriculture*, 189, 106385. https://doi.org/10.1016/j.compag.2021.106385
- Schussler, J. R., & Westgate, M. E. (1995). Assimilate Flux Determines Kernel Set at Low Water Potential in Maize. *Crop Science*, 35(4), cropsci1995.0011183X003500040026x. https://doi.org/10.2135/cropsci1995.0011183X003500040026x
- Shandong. (n.d.). Nutrient Deficiency Disease Leaf State of Corn (Maize). Shandong Lvfeng Fertilizer Co.,Ltd. Retrieved December 5, 2021, from http://www.acefertilizer.com/info/nutrient-deficiency-disease-leaf-state-of-corn-2453004.html
- Sims, D. A., & Gamon, J. A. (2003). Estimation of vegetation water content and photosynthetic tissue area from spectral reflectance: A comparison of indices based on liquid water and chlorophyll absorption features. *Remote Sensing of Environment*, 84(4), 526–537. https://doi.org/10.1016/S0034-4257(02)00151-7

- Singh, A. K., Pérez-López, A. V., Simpson, J., & Castro-Camus, E. (2020). Three-dimensional water mapping of succulent Agave victoriae-reginae leaves by terahertz imaging. *Scientific Reports*, 10(1), 1404. https://doi.org/10.1038/s41598-020-58277-z
- Smith, W. K., Vogelmann, T. C., DeLucia, E. H., Bell, D. T., & Shepherd, K. A. (1997). Leaf Form and Photosynthesis. *BioScience*, 47(11), 785–793. https://doi.org/10.2307/1313100
- Sun, Y., Wang, H., Sheng, H., Liu, X., Yao, Y., & Gong, C. (2015). Variations in internal water distribution and leaf anatomical structure in maize under persistently reduced soil water content and growth recovery after re-watering. *Acta Physiologiae Plantarum*, 37(12), 263. https://doi.org/10.1007/s11738-015-2006-5
- Tang, A., & Boyer, J. S. (2002). Growth-induced water potentials and the growth of maize leaves.JournalofExperimentalBotany,53(368),489–503.https://doi.org/10.1093/jexbot/53.368.489
- Texture Analysis Using the Gray-Level Co-Occurrence Matrix (GLCM)—MATLAB & Simulink.(n.d.).RetrievedDecember8,2021,fromhttps://www.mathworks.com/help/images/texture-analysis-using-the-gray-level-co-<br/>occurrence-matrix-glcm.htmlfromfrom
- Thomas, J. C., McElwain, E. F., & Bohnert, H. J. (1992). Convergent Induction of Osmotic Stress-Responses: Abscisic Acid, Cytokinin, and the Effects of NaCl. *Plant Physiology*, 100(1), 416–423. https://doi.org/10.1104/pp.100.1.416
- Wang, B., Liu, C., Zhang, D., He, C., Zhang, J., & Li, Z. (2019). Effects of maize organ-specific drought stress response on yields from transcriptome analysis. *BMC Plant Biology*, 19(1), 335. https://doi.org/10.1186/s12870-019-1941-5
- Wang, L., Jin, J., Song, Z., Wang, J., Zhang, L., Rehman, T. U., Ma, D., Carpenter, N. R., & Tuinstra, M. R. (2020). LeafSpec: An accurate and portable hyperspectral corn leaf imager. *Computers and Electronics in Agriculture*, 169, 105209. https://doi.org/10.1016/j.compag.2019.105209
- Weatherley, P. E. (1950). Studies in the Water Relations of the Cotton Plant. *New Phytologist*, 49(1), 81–97. https://doi.org/10.1111/j.1469-8137.1950.tb05146.x
- Westgate, M. E., & Boyer, J. S. (1985). Osmotic adjustment and the inhibition of leaf, root, stem and silk growth at low water potentials in maize. *Planta*, *164*(4), 540–549. https://doi.org/10.1007/BF00395973
- Wold, S., Sjöström, M., & Eriksson, L. (2001). PLS-regression: A basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, 58(2), 109–130. https://doi.org/10.1016/S0169-7439(01)00155-1
- Xiong, D., Chen, J., Yu, T., Gao, W., Ling, X., Li, Y., Peng, S., & Huang, J. (2015). SPAD-based leaf nitrogen estimation is impacted by environmental factors and crop leaf characteristics. *Scientific Reports*, 5(1), 13389. https://doi.org/10.1038/srep13389

- Zhang, A., Pérez-Rodríguez, P., San Vicente, F., Palacios-Rojas, N., Dhliwayo, T., Liu, Y., Cui,
  Z., Guan, Y., Wang, H., Zheng, H., Olsen, M., Prasanna, B. M., Ruan, Y., Crossa, J., &
  Zhang, X. (2021). Genomic prediction of the performance of hybrids and the combining
  abilities for line by tester trials in maize. *The Crop Journal*.
  https://doi.org/10.1016/j.cj.2021.04.007
- Zhou, L., Zhang, C., Taha, M. F., Qiu, Z., & He, Y. (2021). Determination of Leaf Water Content with a Portable NIRS System Based on Deep Learning and Information Fusion Analysis. *Transactions of the ASABE*, 64(1), 127–135. https://doi.org/10.13031/trans.13989
- Zhu, J., Brown, K. M., & Lynch, J. P. (2010). Root cortical aerenchyma improves the drought tolerance of maize (Zea mays L.). *Plant, Cell & Environment*, 33(5), 740–749. https://doi.org/10.1111/j.1365-3040.2009.02099.x
- Zhuang, S., Wang, P., Jiang, B., Li, M., & Gong, Z. (2017). Early detection of water stress in maize based on digital images. *Computers and Electronics in Agriculture*, 140, 461–468. https://doi.org/10.1016/j.compag.2017.06.022
- Ziyomo, C., & Bernardo, R. (2013). Drought tolerance in maize: Indirect selection through secondary traits versus genomewide selection. *Crop Science*, 53(4), 1269–1275. https://doi.org/10.2135/cropsci2012.11.0651