# Applications of remote sensing in agriculture via unmanned aerial systems and satellites

by

Sebastian Varela

B.S., University of the Republic, 2009

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Agronomy
College of Agriculture

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2018

# Abstract

The adoption of Remote Sensing (RS) in agriculture have been mainly utilized to inference about biological processes in a scalable manner over space and time. In this context, this work first explores two non-traditional approaches for rapid derivation of plant performance under field conditions. Both approaches focus on plant metrics extraction exploiting high spatial resolution from Unmanned Aerial Systems (UAS). Second, we investigate the spatial-temporal dynamics of corn (Zea mays L.) phenology and yield in the corn belt region utilizing high temporal resolution from satellite. To evaluate the impact of the adoption of RS for deriving plant/crop performance the following objectives were established: i) investigate the implementation of digital aerial photogrammetry to derive plant metrics (plant height and biomass) in corn; ii) implement and test a methodogy for detecting and counting corn plants via very high spatial resolution imagery in the context of precision agriculture; iii) derive key phenological metrics of corn via high temporal resolution satellite imagery and identify links between the derived metrics and yield trends over the last 14 years for corn within the corn belt region. For the first objective, main findings indicate that digital aerial photogrammetry can be utilized to derive plant height and assist in plant biomass estimation. Results also suggest that plant biomass predictability significantly increases when integrating the aerial plant height estimate and ground stem diameter. For the second objective, the workflow implemented demostrates adequate performance to detect and count corn plants in the image. Its robustness highly dependends on the spatial resolution of the image, limitations and future research paths are further discussed. Lastly, for the third objective, outcomes evidenced that for a long-term perspective (14 years), a lengthened reproductive stage significantly correlates with high yield for corn. When considering a shorter-term period (last 4 years) mainly characterized by optimal

growth conditions, early season green-up rate and late season senescence rate positively describe yield trend in the region. The significance of the variables changed according to the time-span considered. It is noticed that when optimal growth conditions are met, modern-hybrids can capitalize by increasing yield, due to primarily a faster (green-up) rate before flowering and on senescence rate better describes yield under these conditions.

The entire research project investigates opportunities and needs for integrating remote sensing into the agronomic-based inference process.

# Applications of remote sensing in agriculture via unmanned aerial systems and satellites

by

Sebastian Varela

B.S., University of the Republic, 2009

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Agronomy
College of Agriculture

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2018

Approved by:

Major Professor
Dr. Ignacio A. Ciampitti

# Abstract

The adoption of Remote Sensing (RS) in agriculture have been mainly utilized to inference about biological processes in a scalable manner over space and time. In this context, this work first explores two non-traditional approaches for rapid derivation of plant performance under field conditions. Both approaches focus on plant metrics extraction via high spatial resolution from Unmanned Aerial Systems (UAS). Second, we investigate the spatial-temporal dynamics of corn (Zea mays L.) phenology and yield in the corn belt region utilizing high temporal resolution from satellite. To evaluate the impact of the adoption of RS for deriving plant/crop performance the following objectives were established: i) investigate the implementation of digital aerial photogrammetry to derive plant metrics (plant height and biomass) in corn; ii) implement and test a methodogy for detecting and counting corn plants via very high spatial resolution imagery in the context of precision agriculture; iii) derive key phenological metrics of corn via high temporal resolution satellite imagery and identify links between the derived metrics and yield trends over the last 14 years for corn within the corn belt region. For the first objective, main findings indicate that digital aerial photogrammetry can be utilized to derive plant height and assist in plant biomass estimation. Results also suggest that plant biomass predictability significantly increases when integrating the aerial plant height estimate and ground stem diameter. For the second objective, the workflow implemented demostrates adequate performance to detect and count corn plants in the image. Its robustness highly dependends on the spatial resolution of the image, limitations and future research paths are further discussed. Lastly, for the third objective, outcomes indicate that lengthened vegetative and reproductive stages, green-up and senescence rate metrics describe yield increase between 2003 and 2017. Both the spatial and temporal components of the model were significant to

describe yield trend. Moreover, when including the temporal component, the model receives lower penalization as an indicator of superior fit on describing yield trend in the region. Overall, the outcomes indicate that in the last 14 years, a significate trend in both space and time on lengthened seasons, faster green-up and senescence rates significantly describe USDA NASS increase on yield in the region.

The entire research project investigates opportunities and needs for integrating remote sensing into the agronomic-based inference process.

# Table of Contents

# List of Figures

# List of Tables

# Acknowledgements

To my advisor, Ignacio for giving me the chance to be a PhD student in his program at K-State. I truly appreciate the freedom and guidance I had from him to put my research ideas into practice.

To my committee members Dr. Vara Prasad, Dr. Juan Du, Dr. Terry Griffin, and Dr. Allison Ferguson, for their commitment and guidance through all the process.

Thanks to Gerard Kluitenberg for keeping me up to date with all my program and for helping and supporting me through my studies. Thanks to Dr. Douglas Goodin, the Outside Chair on my committee.

Special appreciation to Dr. William Hsu and MSc Pruthvidhar Reddy Dhodda for the enriching interaction and learning I had during our work.

Thanks to Ana Julia Azevedo, Olser Ortez, Damaris Handsel, Santiago Tamagno, Guillermo Balboa and Javier Fernandez, Graduate Student at KSU Crops Production Team, for our daily interaction and all their help and feedback on my project.

I would also like to thank KSU Crops Production Team Visiting Scholars 2015, 2016 and 2017, for their valuable help during exhausting field campaigns across the state of Kansas in this last 4 years. Special thanks to Mario Secchi and Matias Aseguinolaza for their tireless help in the field during 2017.

To the Fulbright Program and PrecisionHawk for supporting this research. Thanks to ANII (National Research and Innovation Agency) in Uruguay for allowing me to pursue my PhD program at K-State. I want to acknowledge the Agronomy Department for all the support during my stay at Kansas State University.

# Dedication

To my family, who support me to pursue this journey. To my mother and father, Maria Teresa and Gustavo, and my sister Cecilia whose encourage me to follow my dreams. To my wife Lorena, for these 4 years of unconditional support and mutual effort.

# Preface

Remote sensing is the acquisition of information about an object without making physical contact with the object. The carrier of information in remote sensing is electromagnetic radiation, which travels in vacuum at the speed of light in the form of waves of different lengths. Remote sensing can significantly contribute to providing inference about biological processes in agriculture with the advantage of being suitable for gathering information either over small or large areas with high spatial and temporal resolutions. The overall dissertation objective was to investigate and report applications of remote sensing with significant contribution on crop monitoring at different scales.

The primary objectives for each chapter are as follows:

1. to examine the relationship between plant height data collected from UAS at critical developmental stages and the final biomass estimation of corn hybrids under different fertilizer nitrogen (N) management and planting densities (Chapter 1);
2. to develop a reliable, timely, and unbiased method for identifying and counting corn plants based on ultra-high resolution imagery acquired from UAS to automatically scout fields using real field conditions (Chapter 2);
3. to derive key phenological metrics for corn over the US corn belt via high temporal resolution MODIS vegetation index, benchmark these phenology metrics against ground-truth and CPCRs data at the ASD level, and identify the links between the phenological metrics and yield trends over the last 14 years (Chapter 3).

# Chapter 1 - Spatio-temporal evaluation of plant height via Unmanned Aerial Systems.

Varela, S., Assefa, Y; Prasad, P. V. V., Peralta, N. R., Griffin, T, W., Sharda, A., Ferguson, A., Ciampitti, I, A. (2017). Spatio-temporal evaluation of plant height via Unmanned Aerial Systems. *J. of Applied Remote Sensing* 11(3). doi:10.1117/1.JRS.11.036013

## ABSTRACT

Detailed spatial and temporal data on plant growth are critical to guide crop management. Conventional methods to determine field plant traits are intensive, time-consuming, expensive, and limited to small areas. The objective of this study was to examine the integration of data collected via unmanned aerial systems (UAS) at critical corn (Zea mays L.) developmental stages for plant height and its relation to plant biomass. The main steps followed in this research were (1) workflow development for an ultrahigh resolution crop surface model (CSM) with the goal of determining plant height (CSM-estimated plant height) using data gathered from the UAS missions; (2) validation of CSM-estimated plant height with ground-truthing plant height (measured plant height); and (3) final estimation of plant biomass via integration of CSM-estimated plant height with ground-truthing stem diameter data. Results indicated a correlation between CSM-estimated plant height and ground-truthing plant height data at two weeks prior to flowering and at flowering stage, but high predictability at the later growth stage. Log–log analysis on the temporal data confirmed that these relationships are stable, presenting equal slopes for both crop stages evaluated. Concluding, data collected from low-altitude and with a low-cost sensor could be useful in estimating plant height.

**Keywords:** unmanned aerial systems; structure from motion; corn; imagery.

# INTRODUCTION

Use of unmanned aerial systems (UAS) to evaluate crop growth, development, and performance is a promising new area of agricultural research (Hunt et al., 2005; Lee et al., 2010; Peña et al., 2013; Primicerio et al., 2012). Because piloted aircraft and satellite imagery are either prohibitively expensive or not easily available to the required spatio-temporal resolution, the use of UAS has been presented as an alternative (Herwitz et al., 2004).

The flexibility of UAS to conduct low-altitude flight and facilitate high-resolution imagery has proven useful for site-specific weed management (Torres-Sanchez et al., 2013); to evaluate crop nutrient requirement (Hunt et al., 2005), soil water status (Ryo et al., 2007), and crop water stress (Zarco-Tejada et al., 2011); and to monitor vegetation growth (Berni et al., 2009). Plant height is one of the major indicators of plant growth and development. Plant height is positively correlated with plant grain yield (Law et al., 1978; Shrestha et al., 2002; Yin et al., 2011) plant biomass, and soil nitrogen (N) supply (Yin et al., 2013; Gul et al., 2015). Most cereals attain maximum plant height and yield potential at the onset of the reproductive stage (Mourtzinis et al., 2013), with approximately half of biomass and N accumulated relative to maturity (Freeman et al., 2007; Ciampitti and Vyn, 2012). Therefore, early-season estimation of yield potential in cereals can be generated when the plant attained its maximum height (at flowering) or right before this point (1 or 2 weeks before flowering). Specifically for corn (Zea mays L.), plant height is needed for biomass estimation via stem volume calculation (measured via the cylindrical formula based on plant height and stem diameter both determined at comparable phenological stages). Previous researchers documented a high degree of correlation between ground-truthing based stem volume calculation and plant biomass at flowering in corn

(Miles, 1993; Vega et al., 2000; Borras et al., 2003; Maddonni and Otegui, 2004; Pagano and Maddonni, 2007; D'Andrea et al., 2008; Ciampitti et al., 2012a).

Application and process involved in plant height measurement conducted using UAS platforms were discussed by few researchers (Anthony et al., 2014; Grenzdörffer, 2014; Bendig et al., 2015). The process involves (i) collecting aerial data imagery from a camera mounted onboard in UAS, (ii) generating ultrahigh resolution crop surface models (CSMs), and (iii) determining plant height from the CSM (Bendig et al., 2013), herein, defined as CSM-estimated plant height. However, studies validating CSM-estimated plant height via ground-truthing measurements to predict field crop yields are scarce in the scientific literature. Early- or even mid-season crop production forecasts assist producers to make informed decisions regarding crop and nutrient management, yield estimation, marketing, storage, and transportation (Hammer et al., 2001; Raun et al., 2005; Kantanantha et al., 2010; Franzen et al., 2014). Various models have been used to make such predictions but current applications of most of these models are only for large-scale (regional- or state-level) production systems. As crop management progresses from large-scale uniform management to site-specific using precision agriculture technologies, evaluation of within-field variation and more accurate yield forecasts should be pursued. Following this rationale, plant height relates not only to plant growth during the vegetative stages, but this plant trait can also be used to improve the relationship between active optical sensor readings and yield estimates (Raun et al., 2005; Franzen et al., 2014). Therefore, accurate and rapid plant height prediction could facilitate and improve yield forecast in corn. The overall objective of this study was to examine the relationship between plant height data collected from UAS at critical developmental stages and the final biomass estimation of corn hybrids of different maturity groups under different N management and planting densities.

# MATERIALS AND METHODS

## Study area and dataset

During the 2015 growing season, four corn experiments were established in 1.2 hectares at Ashland Bottoms Farm, Manhattan, Kansas (39.13°N, −96.6°E, 314 m above sea level) (Fig. 1.1). The nitrogen experiment (NE) was implemented in 9.1 m × 10.6 m plots, with five N fertilization levels using urea ammonium-nitrate ranging from 0 to 200 kgNha−1, in 50 kgNha−1 intervals. The plant density experiment (PE) and plant density gap experiment (PGE) were conducted in 6.1 m × 10.6 m plots, whereas the hybrid experiment (HE) was planted in 6.1 m × 12.2 m plots. All experiments were evaluated in randomized complete block design with five replications. Across all studies, row spacing was 0.76 m. Target plant density was 8.4 plantsm−2 in NE and HE, and a range between 4.4 and 10.4 plantsm−2 for the PE and PGE. Corn hybrid used in NE, PE, and PGE was DK61-88 (Dekalb®, Monsanto) 111 days commercial relative maturity (CRM). For the HE, corn hybrids evaluated were DK61-88, DK63-55 (113 CRM), DK64-69 (114 CRM), and P1105 and P1151 (111 CRM; Dupont Pioneer®). All four corn trials were used as a base line to generate spatio-temporal variability of plant height, biomass, and yield to evaluate UAS and structure from motion under different plant height scenarios.

## Platform, sensor and ground-truthing

An UAS platform (S800, DJI, Shenzhen, China) was used to collect aerial imagery. This platform includes the Wookong-Monboard autopilot system and GPS v2 unit (S800, DJI, Shenzhen, China). Flight missions and parameter settings were assigned using UgCS ground station software (SPH,2013). The platform sensor included in each flight was Alpha ILCE A5100 RGB Sony (Tokyo, Japan), mounted with a Sony SELP1650 PZ 16-50 mm lens (sensor resolution is 6000 × 4000 pixels). Aperture and exposure time were adjusted manually prior to

each flight mission considering the ground speed of the UAS and light conditions at the time of flights. In both flights, camera setting was performed using manual exposure control; shutter speed was set to 1⁄4000 s, aperture to f5, and ISO to 640 and 16 mm focal length configuration. Two UAS missions were performed (17 and 29 July). Highly visible yellow and black (1 m×1 m) cross-centered plastic ground targets were used as ground control points (GCPs). In this project, 14 yellow and black cross-centered plastic (1 m×1 m) GCPs were used as main sources for imagery geolocation. The GCPs were distributed on the borders and internal alleys of the experiments following the (Gomez-Candon et al., 2014) recommendations. The average distance between GCPs was 42 m in both missions. Two critical corn growth stages were identified as target candidates for UAS missions: (1) the late vegetative herein termed as pre-flowering and (2) onset of reproductive or flowering stage (Ritchie et al., 1996). These UAS mission timings were relevant because of the importance of the aforementioned corn growth stages to determine if plant height estimates can populate crop yield forecasting models. The goal of this step was to overlay CSMs from UAS with ground-truthing data then check goodness-of-fit of CSMs to capture spatio-temporal change of plant height at both stages. The ground-based data collection was divided into destructive biomass sampling and nondestructive in situ plant height measurements. GCPs and plant samples were georeferenced by implementing a Global Navigation Satellite System-Real-Time Kinematic survey for spatial and temporal monitoring. The data layer containing the geolocated plant positions was overlaid with the orthomosaic and CSMs using ArcMap (ArcGIS v10.3, Environmental System Research Institute Inc.) (ESRI, 2014). Absolute plant height, field ground-truthing, was measured via a centimeter resolution wooden ruler. Field sampling procedures define absolute plant height as the vertical distance between the base of stem and the top region of the plant where leaves reach maximum height

without any external intervention (n = 331 plants measured 2 weeks before flowering and n = 331 plants determined at flowering). Stem diameter (n = 331 measured) was determined at the base of the plant following the procedure described by Ciampitti et al.20 The field measurement performed 2 weeks before flowering was separated by 5 days from the UAS mission; thus plant height was adjusted to the date of the UAS mission using the observed plant height change rate computed between flowering and 2-weeks prior. This adjustment did not modify the proportion of variation accounted for the aerial imagery but significantly reduced the bias in the final observed plant height values, with lower plant height values for the ground-truth data (adjusted by 5 days within the period of height growth). For biomass determination, each individual plant was cut at the stem base and fresh weight was collected in situ. Both stem diameter and plant biomass were measured only at flowering time.

## Data Processing Workflow: Crop Surface Model, Orthomosaic Generation, and Plant Height

The UAS missions were conducted at 65-m altitude to achieve a ground sampling distance, expressed as the distance between the centers of two consecutive pixels measured on the ground, of 0.015 m. An overlapping and side lapping of 80% was employed in accordance with Photoscan manual recommendations for successful CSM reconstructions [Fig. 1.2 (A)] (Agisoft, 2016). Ground speed setting of the UAS was 7 ms−1 obtaining one image per 1.8 s to achieve the expected overlapping on the track of the UAS. A total of 265 images were collected per mission. The GCPs and UAS imagery data set were integrated and processed for true color [red, green, and blue (RGB)] orthomosaic and CSMs for plant height. A workflow for CSMs reconstruction was implemented using Photoscan [Fig. 1.2 (A)]. Processing steps included: feature matching, solving camera intrinsic, and extrinsic orientation parameters, reconstructing

of the dense point cloud (DPC), and texture mapping. Parameter setting for imagery alignment presented the following characteristics: low accuracy and referenced pair preselection, tie and key points limited to 0 and 40000. The Photoscan imagery alignment algorithm detects points in the source images, which are stable under changing viewpoints and lighting conditions. Then, Photoscan software generates a descriptor for each point based on its local neighborhood. These descriptors are used to detect correspondences across the images. Later the software estimates the camera intrinsic and extrinsic orientation parameters using the internal bundle-adjustment algorithm to approximate accurate camera locations. The distance between all GCPs was comparable and located along the image data set to minimize horizontal and vertical geometrical error. The DPC was reconstructed by Photoscan by implementing the height-field algorithm based on pairwise depth map computation. Moreover, the quality value for the DPC reconstruction was set to medium for optimizing the computation time and data set size following Photoscan manual recommendations. The DPC reconstruction achieved 2550 and 2765 points⁄m2, respectively, for each mission timing. A spatial interpolation procedure, via inverse distance weighting (IDW), was applied to the DPC to generate the CSM. Orthomosaic and CSM native Photoscan spatial resolutions were 1.0 and 2.0 cm⁄pixel for data sets from both missions.  The absolute plant height estimation was solved as the difference between the CSM and a digital terrain model (DTM) of bare ground surface [Fig. 1.2 (B)]. The DTM was reconstructed from the flowering RGB orthomosaic (captured 2-weeks prior) and CSM data sets.26 The first step includes the ground class segmentation in the RGB orthomosaic [Fig. 1.2 (B)]. A support vector machine (SVM) classification was implemented in ENVI (Exelis, 2010) to solve the ground class segmentation [Fig. 1.2 (B)]. The training data set included 4000 pixels and iterated in “vegetation,” “bare soil,” and “shadow” classes. In the iteration phase, a linear

discriminant was explored with unsatisfactory results (over all accuracy = 0.55). Thus, a nonlinear classification approach was implemented on the decision surface hyperplane and a radial kernel function was utilized for discrimination between classes. The gamma in kernel function was set to 0.25, the inverse of the number of computed attributes, (Oyewole et al., 2015; Hsu et al., 2003) and the penalty parameter was set to 95 (Hsu et al., 2003). The overall accuracy of the nonlinear SVM classification on the three classes was 0.79. The "bare soil" raster was exported from the CSM with bare soil areas into ArcMap and the DTM solved by using the overlapping CSM vertical and horizontal determined from the bare soil over the bare soil class data from segmented ground class regions. Ground class regions utilized in the IDW interpolation included the borders and alleys of the trials, considering an average distance of 12 m between adjacent alleys. The absolute plant height estimated data were obtained by a map algebra subtraction between the CSM and the DTM over 0.08-m cylinder radius length assigned to each plant center location [Fig. 2(b)]. Estimated plant height was assigned to upper mean quintile CSM pixels in the cylinder area assigned to each plant.

## Plant Height Validation and its Relationship with Stem Volume and Biomass

Plant height data extracted from UAS imagery analysis and collected from field measurements (ground-truth data) were linearly regressed using the GraphPad Prism software (Motuslky and Christopoulos, 2003). The proportion of variation accounted by the fitted model at each developmental stage was evaluated. In addition, a linear relationship between plant biomass and stem volume calculation was examined; whereas an exponential model was fitted for the plant biomass and plant height obtained via CSM. Both fits were performed using the GraphPad Prism software. For plant height validation, model fit was calculated by determination of the root mean square error (RMSE, measurement of estimated versus observed values).

Outlier detection was executed via the robust standard deviation of the residuals (Motulsky and Brown, 2006). An allometric evaluation was performed for plant height data extracted from UAS imagery and within-field measurements. Thus, reduced major axis was performed with the Standardized Major Axis Estimation and Testing Routines (SMATR) contributed package (Warton et al., 2012) to R development software (R Development Core Team, 2017). For the different phenological timings, slopes were tested to compare independent fit versus a shared fit for this parameter (if slopes are equal or not). Parameters were log10 transformed ($Y = \alpha X\beta \rightarrow \log Y = \log \alpha + \beta \log X$) prior to the analysis (Niklas, 2006) and normal distribution of residuals was verified.

# RESULTS AND DISCUSSION

## Crop Surface Model and Orthomosaic generation

Early process of CSM construction is presented in (Fig. 1.3). The UAS images taken at different sections were stitched together using GCPs as a reference (Fig. 1.3) in PhotoScan software. The importance and the number of GCPs necessary to ensure accuracy of UAS image construction have been previously discussed by other researchers (Tahar et al., 2012; Tahar, 2013; Prajwal et al., 2016). In terms of geometric quality, the accumulated horizontal and vertical error was 0.7 cm⁄pixel in orthomosaic and CSM from 2-weeks prior to flowering, and 0.5 cm⁄pixel for the flowering raster products. Furthermore, a woody table (dimensions = 0.8 m length × 0.4 m wide × 0.6 m height) was used for non-vegetation geometric evaluation. A total of six local GCPs were implemented along the top of the table to evaluate the vertical displacement between the original GCPs and the same one located in the CSM reconstruction. The vertical error in this case was 0.6 cm⁄pixel.

## Plant Height from Unmanned Aerial Systems Versus Ground-Truth Plant Trait

A strong positive correlation was obtained between CSM-estimated plant height and ground-truth data collected when corn plants were at flowering stage ($R^2$ = 0.79, RMSE = 0.09 m, n = 331, and mean = 1.84 m) [Fig. 1.4 (B)]. The correlation between CSM and ground-truth data measured two weeks prior to flowering ($R^2$ = 0.63, RMSE = 0.11 m, n = 331, and mean = 1.05 m) was relatively weaker (lower $R^2$) and with a slightly higher RMSE [Fig. 1.4 (A)]. The RMSE to mean plant height ratio prior to flowering was 14%, close to threefold higher compared to the ratio estimated at flowering time (5%). For the pre-flowering measurement, the lower proportion of the variation accounted for the CSM-estimated plant height was primarily due to

lack of uniform development within the corn canopy and plants emerging at different timing due to soil–weather factors (e.g., saturated soil areas, low residue with less temperature). At flowering, maximum plant height was attained (Ciampitti et al., 2016), corn canopy become more uniform with less heterogeneity (lower RMSE to mean plant height ratio) and better prediction power (higher $R^2$). A correlation obtained between measured and CSM-estimated plant height is consistent with previous findings for corn (Anthony et al., 2014), ( Grenzdörffer, 2014) (Geipel et al., 2014) barley (Hordeum vulgare L.), and rice (Oryza sativa L.) (Bendig et al., 2015) (Tilly et al., 2014). A significant correlation between plant height measurements at flowering stage support the conclusion drawn by (Geipel et al., 2014) that imagery taken at end of stem elongation is better correlated with ground-truth data. Few researchers have studied the corn growth stage that UAS imagery should be taken to improve plant height estimation (Geipel et al., 2014; Sharma et al., 2016; Shi et al., 2016). Other studies not using UAS imagery to evaluate the relationship between actual plant height and remotely sensed plant height (Sharma et al., 2016) also concluded that late vegetative stage sensor-based plant height measurements correlated with actual plant height. Additionally, plant height measurements at late stage of corn were found to correlate with grain yield (Warton et al., 2012).

It is worth noting that at both corn stages plant height was underestimated, similar to the findings presented by (Grenzdörffer, 2014) and (Shi et al., 2016). For understanding the stability of the plant height estimation, two evaluations were executed. The first one was done by comparing the linear regression slopes (for equality) of the estimated- and observed-plant height relationship [Figs. 1.4 (A) and 1.4 (B)] between the two growth stages evaluated to understand the stability of the estimation between dates and across plant height class (log–log transformation analysis) [Fig. 1.4 (C)]. Results showed similar slopes across classes for both mission timings. A

second evaluation was performed to understand the absolute and relative magnitude of the plant height estimation for both dates. Ground-truth plant height data were divided into three equal classes for each crop stage. Prior flowering ground-truth mean plant height data classes were low 1.22 m, medium 1.52 m, and high 1.71 m. For the 2 weeks before flowering timing, within the low plant height group, 22% of the data in this class were underestimated by the CSM-estimated plant height trait; while for the high plant height group, this analysis resulted in 24% of the plant height observations being underestimated. At flowering, only 10% of the data on plant height across all classes (low 1.77 m, medium 2.12 m, and high 2.24 m) were underestimated by the CSM-estimated plant height trait. Synthesizing, this analysis allowed us to conclude that there was a better prediction of plant height due to a lower underestimation at flowering, which was also related to lower plant heterogeneity within the corn canopy.

## Unmanned Aerial Systems Based Plant Height Relation with Biomass

Since ground-truth plant height was better estimated at flowering, the biomass data collected at the same growth stage were utilized to better understand the relationship between plant height and plant biomass. Plant biomass and CSM-estimated plant height exhibited a statistically significant correlation at flowering [Fig. 1.5 (A)]. However for the plant biomass trait, the proportion of the variation accounted by the CSM-estimated parameter alone was low ($R^2 = 0.31$, $n = 332$, and $P < 0.05$). Examination of Fig. 1.5 (A) shows substantial variation present in the data, and possibly nonlinear behavior at greater plant height values. Plant biomass estimation substantially improved ($R^2 = 0.79$, $n = 332$, and $P < 0.05$) when the stem diameter (determined at equal growth stage and for the same plants as the plant height trait) was considered as a part of the stem volume calculation [Fig. 1.5 (B)]. Allometric equations were previously utilized in corn to predict biomass with different levels of success depending on the

variation of the data (genotype by environment by management interaction) and the timing of the sampling (Mourtzinis et al., 2013; Pordesimo et al., 2004; Barten, 2013). More accurate biomass estimation performed via utilization of allometric models could be utilized as a tool to forecast corn yields. Last, improvements of biomass prediction for corn after flowering stage were documented when the apical ear shoot diameter (maximum diameter of the ear organ) was included in the stem volume calculation (Pagano and Maddonni, 2007). Thus, improvement in corn biomass prediction will be of a great challenge for the remote sensing discipline because the reproductive organs (ears) are placed at varying positions within the corn canopy. In a simplified approach, a combination of various data layers collected from multiple sensors [e.g., plant height, stand counts, normalized difference vegetation index (NDVI) (Vergara-Diaz et al., 2016)] in a spatio-temporal fashion might allow to adjust in real-time corn yield estimations based not only on plant size but also considering plant nutrient status and the complex interaction with the environment. For corn crop, a correlation between ground-truth plant height measured at late vegetative or early reproductive and plant biomass has been previously documented (Yin et al., 2011; Mourtzinis et al., 2013). For the current study, plant height was adequately predicted for both corn growth stages: 2-weeks before (with more variation detected) and at flowering. Similarly, a significant positive relationship between corn plant height measured late-vegetative using sensors, mounted on satellite or run manually, with biomass or yield was reported (Freeman et al., 2007; Bach, 1998). Not many UAS based results are available for corn (Geipel et al., 2014; Shi et al., 2016) but our results on the relationship between plant height measured by UAS platform with biomass is in agreement with previous research reported for other field crops (Bendig et al., 2015; Tilly et al., 2014). Last, the relationship documented in this study for corn crops between stem diameter and plant biomass is in line with findings previously presented by

(Mourtzinis et al., 2013). A nondestructive way of measuring stem diameter from images mounted on UAS and other ground-truth sensors remains as a critical research gap for improving plant biomass prediction and the potential for yield forecasting purposes. From a remote sensing standpoint, different vegetation indices and multi/hyperspectral sensors can be investigated to improve plant biomass prediction and yield forecast procedure.

# CONCLUSIONS

Spatial-temporal correlation between CSM-estimated versus ground-truthing plant height trait suggested that the CSM integration could assist in biomass estimation. Both dates evidence plant height underestimation but with higher departure for this trait for the pre-flowering stage. Imagery overlapping and plant height heterogeneity become critical factors in the plant height estimation process. At flowering stage, plant biomass and yield prediction could still be used for late management practices, such as nutrient fertilization and fungicide/insecticide protection. Nonetheless, accurate corn yield prediction at early growth stage (before flowering) remains a topic needing additional research. The evidence suggests that both plant traits such as stem diameter and/or nutrient content estimation should be targeted to increasing reliability of forecasting yield procedures. Future research should also look into the integration of UAS and spectral remote sensing data into ultrahigh spatial-resolution analysis for crop growth modeling.

# REFERENCES

Agisoft LLC, AgiSoft PhotoScan User Manual, Professional Edition v.1.2.6, Agisoft LLC, St. Petersburg, Russia (2016).

Anthony, D., Elbaum, S., Lorenz, A., Detweile, C. On Crop Height Estimation with UAVs. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2014), September 14-18, 2014, Chicago, IL, USA, pp .4805-4812.

Bach, H. (1998). Yield estimation of corn based on multi-temporal Landsat-TM data as input for an agrometeorological model," Pure Appl. Opt. 7, 809–825.

Barten, T.J. Evaluation and prediction of corn stover biomass and composition from commercially available corn hybrids, PhD Dissertation, Iowa State University, Ames, Iowa (2013).

Bendig, J., Bolten, A., Bareth, G. (2013). UAV-based imaging for multi-temporal, very high-resolution crop surface models to monitor crop growth variability. Photogramm. Fernerkundung Geoinf. 2013(6), 551–562.

Bendig, J., Yu, K., Aasen, H., Bolten, A., Bennertz, S., Broscheit, J., Gnyp, M.L., Bareth, G. (2015). Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. Int. J. Appl. Earth Obs. Geoinf. 39, 79–87.

Berni, J., Zarco-Tejada, P., Suarez, L., Fereres, E. (2009). Thermal and Narrowband Multispectral Remote Sensing for Vegetation Monitoring from an Unmanned Aerial Vehicle.

Geoscience and Remote Sensing, IEEE Transactions 47(3), 722-738.doi: 10.1109/TGRS.2008.2010457.

Borrás, L ., Westgate, M. E ., Otegui, M. E. (2003). Control of Kernel Weight and Kernel Water Relations by Post-flowering Source–sink Ratio in Maize. Annals of Botany, 91(7), 857-867.

Ciampitti, I.A., Vyn, T. J. (2012). Physiological perspectives of changes over time in maize yield dependency on nitrogen uptake and associated nitrogen efficiencies: A review. Field Crops Res. 133, 48-67.

Ciampitti, I. A., Zhang, H, Friedemann, P, Vyn, T. J. (2012a). Potential physiological frameworks for mid-season field phenotyping of final plant nitrogen uptake, nitrogen use efficiency, and grain yield in maize. Crop Science, 52(6), 2728-2742.

Ciampitti, I.A., Elmore, R.W., Lauer, J. Corn Growth and Development, Kansas State University Agricultural Experiment Station and Cooperative Extension Service, MF3305, Manhattan, Kansas (2016).

D'Andrea, K.E., Otegui, M.E., Cirilo, A.G. (2008). Kernel number determination differs among maize hybrids in response to nitrogen. Field Crops Res. 105, 228–239.

ESRI, The mapping platform for your organization. ArcGIS, 2014, http://www.arcgis.com/features/index.html (29 July 2016).

Exelis Visual Information Solutions, "ENVI 5.3.1," 2010, Exelis Visual Information Solutions, Boulder, Colorado, http://harrisgeospatial.com/ProductsandTechnology/Software/ENVI.aspx (4 August 2017).

Franzen, D., Sharma, L.K., Bu, H. Active Optical sensor algorithms for corn yield prediction and a corn side-dress Nitrogen rate aid. North Dakota State University, SF1176-5 (2014).

Freeman, K., Girma, K., Arnall, D., Mullen, R., Martin, K., Teal, R., & Raun, W. (2007). By-Plant Prediction of Corn Forage Biomass and Nitrogen Uptake at Various Growth Stages Using Remote Sensing and Plant Height. Agronomy Journal, 99(2), 530-536.

Geipel, J. Link, J., Claupein, W. (2014). Combined spectral and spatial modeling of corn yield based on aerial images and crop surface models acquired with an unmanned aircraft system. Remote Sens. 6, 10335–10355.

Gomez-Candon, D., De Castro, A.I., Lopez-Granados, F. (2014). Assessing the accuracy of mosaics from unmanned aerial vehicle (UAV) imagery for precision agriculture purposes in wheat. Precis. Agric. 15, 44-56.

Grenzdörffer, G. J. Crop height determination with UAS point clouds. in Int. Archives of the Photogrammetry, Remote Sensing and Spatial Information Science, Volume XL-1, ISPRS Technical Commission I Symp., Denver, Colorado (2014).

Gul, S., Khan, M. H., Khanday, B.A., Nabi , S. (2015). Effect of Sowing Methods and NPK Levels on Growth and Yield of Rainfed Maize (Zea mays L.). Scientifica, 6.

Hammer, G. L., Hansen, J. W., Phillips, J. G., Mjelde, J. W., Hill, H., Love, A., and Potgieter, A. (2001). Advances in application of climate prediction in agriculture. Agric. Syst. 70, 515–553.

Herwitz, S. R., Johnson, L. F., Dunagan, S. E., Higgins, R. G., Sullivan, D. V., Zheng, J. (2004). Imaging from an unmanned aerial vehicle: Agricultural surveillance and decision support. Computers and Electronics in Agriculture, 44, 49-61. doi:10.1016/j.compag.2004.02.006.

Hsu, C.-W., Chang, C.-C., and Lin, C.-J. (2003). A practical guide to support vector classification. Tech. rep., Department of Computer Science, National Taiwan University. (40) http://ntu.csie.org/~cjlin/papers/guide/guide.pdf (Visited 1 December 2016).

Hunt, E.R., Cavigelli, M., Daughtry, C.S.T. Precision Agric (2005) 6: 359. doi:10.1007/s11119-005-2324-5.

Kantanantha, N., Serban, N., Griffin, P. (2010). Yield and price forecasting for stochastic crop decision planning. J. Agric. Biol. Environ. Stat. 15, 362–380.

Law, C. N., Snape, J. W., Worland, A. J. The genetic relationship between height and yield in wheat. Heredity 40, 133–151. (1978).

Lee, W.S., Alchanatis, V., Yang, C., Hirafuji, M., Moshoue, D., Li, C. (2010). Sensing technologies for precision specialty crop production. Comput. Electron. Agric, 74, 2–33. doi:10.1016/j.compag.2010.08.005.

Maddonni, G.A., Otegui, M.E. (2004). Intra-specific competition in maize: early establishment of hierarchies among plants affects final kernel set. Field Crops Res. 85, 1–13.

Miles, C.A. “Divergent selection of sweet corn (Zea mays L. var. saccharata) under low and conventional nitrogen environments,” PhD Dissertation, Cornell University, Ithaca, New York (1993).

Mourtzinis, S., Arriaga, F., Balkcom, K., Ortiz, B. (2013). Corn Grain and Stover Yield Prediction at R1 Growth Stage. Agronomy Journal, 105(4), 1045-1050.

Motulsky, H., and Christopoulos., A. Fitting models to biological data using linear and nonlinear regression, in A Practical Guide to Curve Fitting 2nd ed., GraphPad Software Inc., San Diego, California (2003).

Motulsky, H., and Brown, R.E. (2006). Detecting outliers when fitting data with nonlinear regression - a new method based on robust nonlinear regression and the false discovery rate. BMC Bioinf. 7, 123 (2006).

Niklas, K.J. (2006). A phyletic perspective on the allometry of plant biomass-partitioning patterns and functionally equivalent organ-categories. New Phytol. 171, 27–40.

Oyewole, S., Olugbara, O., Adetiba, E., Nepal, T. Classification of Product Images in Different Color Models with Customized Kernel for Support Vector Machine. Artificial Intelligence, Modelling and Simulation (AIMS), 2015 3rd International Conference on, 153-158.

Pagano, E., Maddonni, G.A. (2007). Intra-specific competition in maize: early established hierarchies differ in plant growth and biomass partitioning to the ear around silking. Field Crops Res. 101, 306–320.

Peña, J.M., Torres-Sánchez, J., De Castro A.I., Kelly, M., López-Granados, F. (2013). Weed Mapping in Early-Season Maize Fields Using Object-Based Analysis of Unmanned Aerial Vehicle (UAV) Images. PLoS ONE 8(10). doi:10.1371/journal.pone.0077151.

Pordesimo, L., Edens, W.C., Sokhansanj, S. (2004). Distribution of above ground biomass in corn stover. Biomass Bioenergy 26, 337-343.

Prajwal, M., Rishab, J., Vaibhav, S., Karthik, K.S. (2016). Optimal number of ground control points for a UAV based corridor mapping. Int. J. Innovative Res. Sci. Eng. Technol. 5, 28-32.

Primicerio, J., Gennaro, S. F. D., Fiorillo, E., Genesio, L., Lugato, E., Matese, A. (2012). A flexible unmanned aerial vehicle for precision agriculture. Precision Agriculture. doi: 10.1007/s11119-012-9257-6.

R Development Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna (2009).

Raun, W.R., Solie, J.B., Stone, M.L., Martin, K.L., Freeman, K.W., Mullen, R.W., Zhang, H., Schepers, J.S., Johnson., G.V. (2005). Optical Sensor-Based Algorithm for Crop Nitrogen Fertilization, Communications in Soil Science and Plant Analysis, 36:19-20, 2759-2781, doi: 10.1080/00103620500303988.

Ritchie, S.W., Hanway, J.J., Thompson, H.E. How a corn plant develops. Special Report 48, Cooperative Extension Service Ames, Iowa State University, Ames, Iowa (1996).

Ryo, S., Noguchi, N., Ishii, K. (2007). Correction of low-altitude thermal images applied to estimating of soil water status," Biosyst. Eng. 96, 301-313. doi:10.1016/j.biosystemseng.2006.11.006.

Sharma, L.K., Bu, H., Franzen, D.W., Denton, A. (2016). Use of corn height measured with an acoustic sensor improves yield estimation with ground based active optical sensors. Comput. Electron. Agric, 124, 254-262.

Shi, Y., Thomasson, J.A., Murray, S.C., Pugh, N.A., Rooney, W.L., Shafian, S., Rajan, N., Rouze, G., Morgan, C.L., Neely, H.L. (2016). Unmanned Aerial Vehicles for High-Throughput Phenotyping and Agronomic Research. PLoS ONE, 11, e0159781.

Shrestha, D.S., Steward, B.L., Birrell, S.J., Kaspar, T.C. Plant height estimation using two sensing systems. In: Proceedings of the ASAE Annual International Meeting, 2002. CD-ROM, St. Joseph, MI.

SPH, UGCS Software for mission planning and execution for all types of unmanned vehicles. SPH Engineering, 2013, https://www.ugcs.com/en/page/products (26 November 2016).

Tahar, K.N., Ahmad, A., Akib, W.A.A.W.M., Mohd, W.M.N.W. (2012). Assessment on Ground Control Points in Unmanned Aerial System Image Processing for Slope Mapping Studies. Int. J. Sci. Eng. Res, 3, 1-10.

Tahar, K.N. An evaluation on different number of ground control points in unmanned aerial vehicle photogrammetric block. In Proceedings of the ISPRS, International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Istanbul, Turkey, 27–29 November 2013, Volume XL-2/W2.

Tilly, N., Hoffmeister, D., Cao, Q., Huang, S., Lenz-Wiedemann, V.I.S., Miao, Y., Bareth, G. (2014). Multitemporal crop surface models: Accurate plant height measurement and biomass estimation with terrestrial laser scanning in paddy rice. J. Appl. Remote Sens, 8, 083671–083671.

Torres-Sánchez, J., López-Granados, F., De Castro, A, I., Peña-Barragán, J.M. (2013). Configuration and Specifications of an Unmanned Aerial Vehicle (UAV) for Early Site Specific Weed Management. PLOS ONE 8(3). doi:10.1371/journal.pone.0058210.

Vega, C.R.C., Sadras, V.O., Andrade, F.H. & Uhart, S.A. (2000). Reproductive allometry in soybean, maize and sunflower. Annals of Botany. 85, 461–468.

Vergara-Díaz, O., Zaman-allah, M. A., Masuka, B., Hornero, A., Zarco-Tejada, P., Prasanna, B. M., et al. (2016). A novel remote sensing approach for prediction of maize yield under different conditions of nitrogen fertilization. Front. Plant Sci. 7:666. doi: 10.3389/fpls.2016.00666.

Warton, D. I., Duursma, R. A., Falster, D. S., Taskinen, S. (2012) smatr 3– an R package for estimation and inference about allometric lines. Methods Ecol. Evol. 3, 257–259.

Yin, X., Jaja, N., McClure, M.A., Hayes, R.M. (2011). Comparison of models in assessing relationship of corn yield with plant height measured during early- to mid-season. J. Agric. Sci, 3, 14–24.

Yin, X., McClure, M.A., Jaja, N., Tyler, D.D., Hayes, R.M. (2011a). In-season prediction of corn yield using plant height under major production systems. Agron. J, 103, 923-929.

Yin, X., McClure, M. (2013). Relationship of Corn Yield, Biomass, and Leaf Nitrogen with Normalized Difference Vegetation Index and Plant Height. Agronomy Journal, 105(4), 1005-1016.

Zarco-Tejada, P.J., González-Dugo, V., Berni, J.A.J. (2011). Fluorescence, temperature and narrow-band indices acquired from a UAV platform for water stress detection using a micro-hyperspectral imager and a thermal camera. Remote Sensing of Environment, 117(C), 322-337. doi:10.1016/j.rse.2011.10.007.

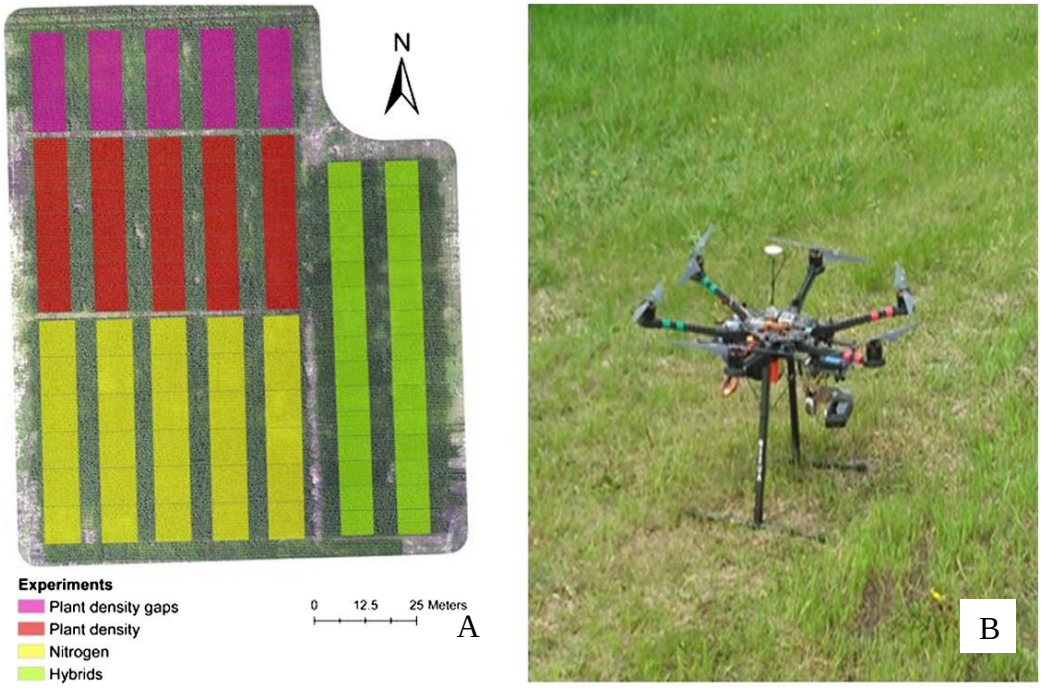


**Figure 1.1 (A) Study area with four corn experiments evaluating: (i) hybrids, (ii) fertilizer N rates, (iii) plant densities, and (iv) plant density gaps and (B) photo of the UAS S800 DJI hexacopter mounted with RGB sensor.**

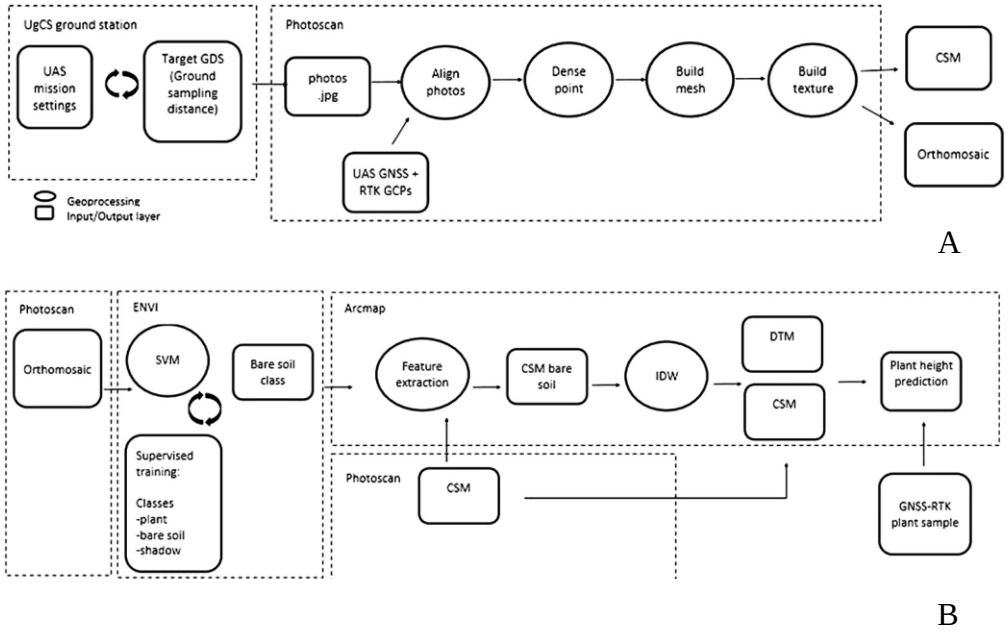


**Figure 1.2. (A) Workflow data integration between UAS and Photoscan and (B) Photoscan–ENVI– ArcMap data workflow.**

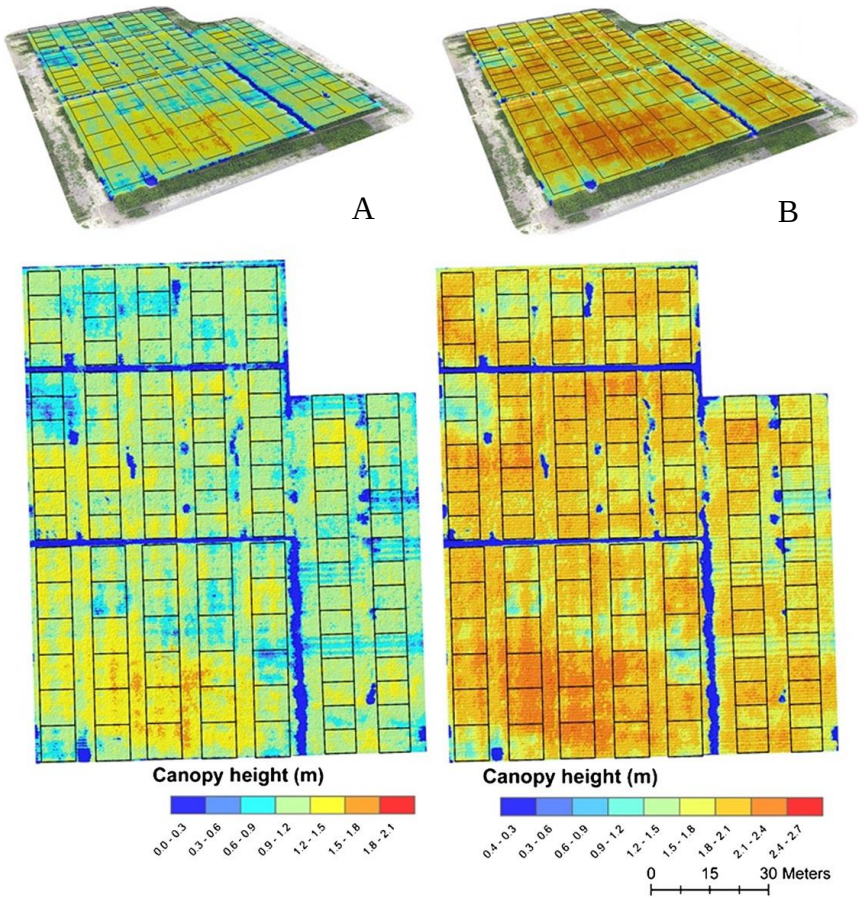


**Figure 1.3. CSMs for estimated absolute plant height on top of the corn canopy: (A) 2-weeks prior to flowering and (B) flowering time. Upper part: 3-D view and; lower part: 2-D perspectives for corn plant height. Note: The blue color represents ground and low vegetation, the yellow refers to short medium corn plants, and the brown and red colors represent taller plants within the corn canopy.**

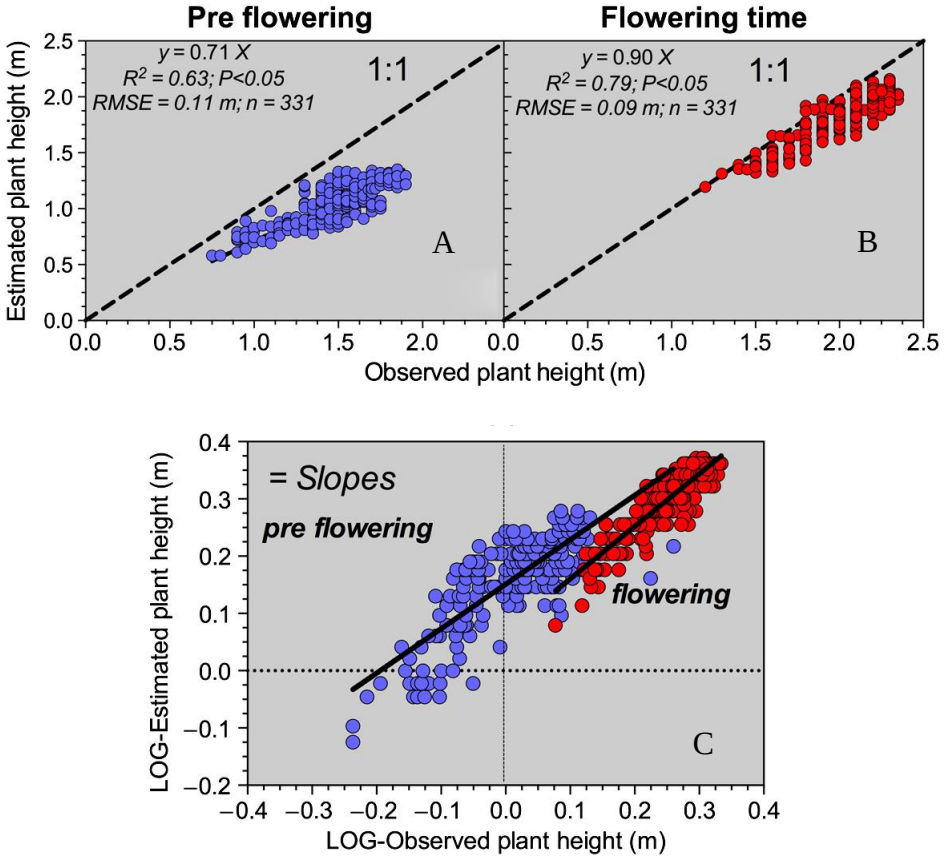


**Figure 1.4. Plant height estimation via UAS imagery collection (A) 2-weeks prior, (B) at flowering, and (C) log–log linear regression of estimated- to observed-plant height (determined from the ground base to the top of the canopy). RMSE, root mean square error.**

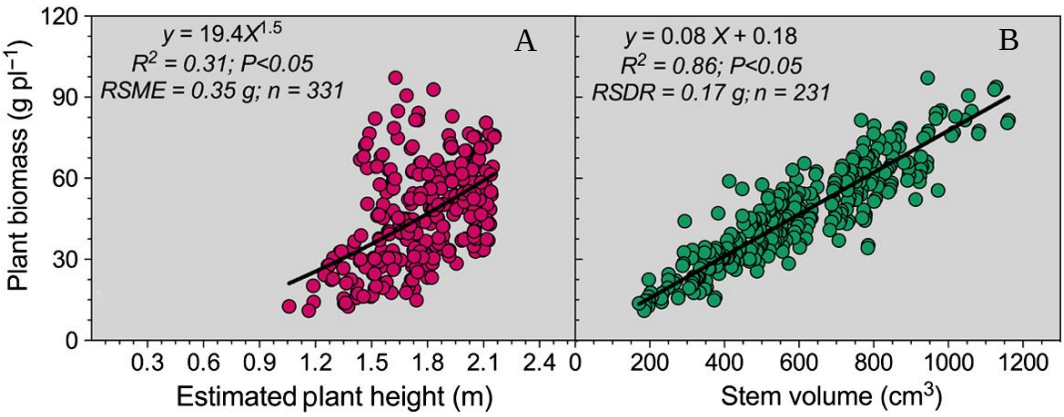


**Figure 1.5. Per-plant biomass (dry basis, expressed in g pl−1) versus (A) plant height trait estimated via CSM and (B) stem volume estimated via implementation of the volumetric cylinder equation (including plant height estimated via CSM, CSM-estimated plant height) all parameters determined at flowering.**

# Chapter 2 - Early-season stand count determination in corn via integration of imagery from Unmanned Aerial Systems and supervised learning techniques

Varela, S., Reddy Dhodda, P., Hsu., W. H., Prasad, P. V.V., Assefa, Y., Peralta, N, R., Terry, G., Ajay Sharda, A., Ferguson, A and Ciampitti I. A. *Remote Sensing.* **2018**, *10*, 343. doi: 10.3390/rs10020343

## ABSTRACT

Corn (Zea mays L.) is one of the most sensitive crops to planting pattern and early-season uniformity. The most common method to determine number of plants is by visual inspection on the ground but this field activity becomes time-consuming, labor-intensive, biased, and may lead to less profitable decisions by farmers. The objective of this study was to develop a reliable, timely, and unbiased method for counting corn plants based on ultra-high-resolution imagery acquired from unmanned aerial systems (UAS) to automatically scout fields and applied to real field conditions. A ground sampling distance of 2.4 mm was targeted to extract information at a plant-level basis. First, an excess greenness (ExG) index was used to individualized green pixels from the background, then rows and inter-row contours were identified and extracted. A scalable training procedure was implemented using geometric descriptors as inputs of the classifier. Second, a decision tree was implemented and tested using two training modes in each site to expose the workflow to different ground conditions at the time of the aerial data acquisition. Differences in performance were due to training modes and spatial resolutions in the two sites. For an object classification task, an overall accuracy of 0.96, based on the proportion of corrected assessment of corn and non-corn objects, was obtained for local (per-site) classification, and an accuracy of 0.93 was obtained for the combined training modes. For

successful model implementation, plants should have between two to three leaves when images are collected (avoiding overlapping between plants). Best workflow performance was reached at 2.4 mm resolution corresponding to 10 m of altitude (lower altitude); higher altitudes were gradually penalized. The latter was coincident with the larger number of detected green objects in the images and the effectiveness of geometry as descriptor for corn plant detection.
**Keywords:** unmanned aerial system; supervised learning; corn; farm management; precision agriculture.

# INTRODUCTION

Corn (Zea mays L.) is one of the most responsive grain crops to agronomic management practices including planting pattern and plant density (Lauer and Rankin, 2004; Ciampitti and Vyn, 2011). Corn has a limited capacity to compensate for missing plants within a row, consequently penalizing grain yield per unit land area at the end of the season (Weiding et al., 2004; Nielsen, 2001; Nafziger and Carter, 1991). One of the most frequent practices to determine the final number of emerged plants is by visual inspection on the ground (De Bruin and Pedersen, 2004). This is a labor-intensive, time-demanding, and cumbersome activity for farmers or researchers. Therefore, there is a need to find alternative and novel techniques to quantify plant stands. The novel process should also include quick data processing and data analyses so that the outcomes can help for efficient planning of operations (e.g., re-planting decisions) on the farm (Nielsen, 2003).

Recent advances in ground sensors and computer vision have provided new insights into plant counting via proximal sensing (Nakarmi and Tang, 2012; Nakarmi and Tang, 2014). The proximal sensing method can provide potential applications of automation and mechanization, which substantially reduces the cost of field scouting (Shi et al., 2013). Shertha et al. 2005 reported the use of the size and shape of corn plants to estimate plant density and row spacing via video frame sequencing, segmentation, and object classification using ground vehicles. In the same context, ground laser line-scanning was adopted to automatically locate stalk and interplant spacing (Shi et al., 2013). Ground vehicles are used to mount proximal sensors or cameras to document images and videos. However, the use of ground vehicles is limited to small areas and often dependent on good trafficable conditions to successfully implement a programmed task.

To overcome this, implementation and the use of remote sensing using aerial or satellite images and data is gaining importance. Thorp et al. 2008 reported the use of aerial hyperspectral data and principal components analysis (PCA) for estimating densities of plants in corn fields. From the same authors, the best performance for the proposed method was reported at the later-vegetative stage ($R^2 = 0.79$) using 6-m resolution imagery. Early-season estimation of plant densities was significantly limited due to the dominant soil background signal when using meter level resolution imagery (Thorp et al., 2008; Thorp et al., 2004). The use of small unmanned aerial systems (UAS) fills the gap of information between proximal ground sensing and meter spatial resolution platforms.

The UAS platforms deliver unprecedented ultra-high spatial resolution imagery and flexible revisit time and offer high versatility under adverse weather conditions (Torrez-Sanchez et al., 2015; Salami and Barrado, 2014). In this context, the use of UAS has been reported in agriculture for crop and weed detection (Peña et al., 2015; Lottes et al., 2016; Lopez-Granados, 2011; Torres-Sanchez et al., 2014). For weed management, detailed knowledge on the spatial distribution of crops and weeds can significantly reduce the impact of agrochemicals on the environment by using site-specific interventions (Lottes et al., 2016). Moreover, early detection of crop and weeds aligns with best practices to maximize the effectiveness of agrochemical applications and yield potential (Lottes et al., 2016). Perez-Ortiz et al., 2016 reported the use of a support vector machine (SVM) classifier, utilizing color intensity and geometrical information as input features for weed and crop mapping. The spatial resolution was critical in the performance of the classifier as also identified by Peña et al. 2015. In general, the implementation of UAS in agriculture has been focused on the extraction of information at the “canopy scale” for further biophysical and yield prediction (Geipel et al., 2014; Zhou et al., 2017). This approach has been

extensively reported via integration of UAS and sensors: RGB, multi-spectral, hyperspectral, and thermal imagery had been used to estimate biomass (Bendig et al., 2013), LAI (Bendig et al., 2013; Mathews et al., 2013; Yao et al., 2017; Haboudane et al., 2004; Pölönen et al., 2013; Potgieter et al., 2017), canopy height (Geipel et al., 2014; Bendig et al., 2013; De Souza et al., 2017;Iqbal et al., 2017), nitrogen (Pölönen et al., 2013; Caturegli et al., 2016; Clevers and Kooistra, 2012), chlorophyll (Clevers and Kooistra, 2012; Uto et al., 2013), and temperature (Gomez-Candon et al., 2016; Gonzalez-Dugo et al., 2013; Berni et al., 2009). Recently, Jin et al. 2017 estimated plant density in wheat from UAS observations using a RGB sensor, ultra-high-resolution imagery, and a support vector machine classifier.

Modern approaches on smart farming typically need detailed knowledge of the current status of crops in the fields. The earlier the yield-limiting factors are identified at the field level, the greater the chances to understand the causes and identify potential farming solutions (Lottes et al., 2016). However, most of the studies on plant density estimation have been implemented via utilization of RGB sensors and computer vision via ground vehicles (Nakarmi, 2012; Nakarmi and Tang, 2014; Guo et al., 2013; Montalvo et al., 2012). Scarce attention has been focused on counting and segmenting individual plants in real field conditions via UAS. Recently, (Gnädinger and Schmidhalter, 2017) implemented a digital counting procedure using a decorrelation stretch contrast enhancement in the RGB feature space domain via UAS. The developed method utilizes the color differences between young and old leaves to estimate plants of different age groups in the image, with an $R^2$ of 0.89 between ground-truthing and estimated plant count. However, the challenge of image thresholding techniques is that they may be prone to misclassification due to the similar spectral response of target and non-target vegetation in the image (Baxes, 1994; Savitzky and Golay, 1964). The current work aims to contribute to the

transition from passive and time-delayed workflows into more automatized, reactive, and integrated systems of managing information on monitoring crop performance on farmers' fields (Tokekar et al., 2016; Hale Group, 2014; Henry, 2015) by developing a tool for quantifying early-season stand counts for corn. Briefly, the present work has been implemented using ultra-high-resolution imagery for plant metric extraction and the workflow was developed by applying the following steps: (i) identify green and non-green regions, (ii) perform a row detection procedure, (iii) extract geometric descriptors of the green objects, and as a last step, (iv) train a decision tree classifier to retrieve information on count and location of the corn plants.

# MATERIALS AND METHODS

## Experimental Sites

Two fields were included to test the workflow under different field conditions such as crop residue, soil backscatter, and crop growth stages. Farmers' fields sites were located in the NE region of Kansas (KS), US (Fig. 2.1). Site 1 was located at Atchison County, KS (39°33′14.84″N, 95°33′46.07″W). Site 2 was located at Jefferson County, KS (39°3′23.60″N, 95°23′19.70″W). Both fields were managed in a soybean-corn rotation. The size of the field in site 1 was 18 hectares, managed under rainfed conditions. The size of the field in site 2 was 64 hectares, under irrigation. The plant density in both the fields was 7.5 plants m−2 and inter-row distance 0.75 m.

## Platform, sensor, and field data collection

A UAS octocopter platform (S1000, DJI, Shenzhen, China) was utilized to collect the aerial images and data. The platform included the A2 flight controller and Global Positioning System (GPS) units used to set up flight missions (S1000, DJI, Shenzhen, China). The flight parameters setting was controlled using UgCs ground station software (UgCs, Riga, Latvia). A PX4 Pixhawk autopilot (Meier et al., 2015) was installed in the same platform for full control of the intervalometer of the sensor via Mission Planner ground station, an open-source software developed by Michael Osborne (http://planner.ardupilot.com). Nine sample areas of 0.2 hectares were randomly selected and marked in each field prior to the growing season to account for varied spatial conditions from existing residues and non-corn objects. Flights used an automatic setting pattern of 4 parallel lines with a time-lapse of 4 s between images, targeting 25–30 images in each sample region of the field. Side lapping and overlapping were set to 20%,

targeting a consistent distribution of sample images in each sample area. The low overlapping requirement increases the efficiency of the flight and post-processing time compared to other data collection approaches (orthomosaic stitching) to analyze UAS data. The platform, camera orientation, and flight direction were set parallel to the direction of the rows. UAS flight autonomy was set for 15 min; 2 and 3 flights were needed to cover the nine sample areas for sites 1 and 2, respectively.

The platform sensor included was an Alpha ILCE A5100 RGB Sony camera (Tokyo, Japan), mounted with a Sony SELP1650 PZ 16-50 mm lens (sensor resolution is 6000 × 4000 pixels). The aperture and exposure time was adjusted manually prior to each mission considering the ground speed of the UAS and light conditions at the time of flights. In all flights, the camera settings used manual exposure control; shutter speed was set to 1/3000 s, aperture to f5, ISO to 400 and 16 mm focal length configuration. One flight in each site was performed between May and June with full sun and 2–3 m s−1 wind conditions (Table 1). On the date of the flights, sites 1 and 2 were at 2 and 2–3 visible leaves growth stage, respectively. Higher soil temperatures and adequate soil moisture conditions during the planting–emergence period in site 2 explained similar growth stages encountered in both locations on the date of the flights, despite a late planting date in site 2. The flight altitude was set to 10 m reaching a spatial resolution of 2.4 mm.

### Data processing

The following workflow including five steps (Fig. 2.2) was developed and implemented after the images were collected from the fields: (1) images were converted into excess greenness (ExG)–vegetation index, (2) row detection and contours were delineated, (3) geometric descriptors were built from contours, (4) classifier training, and (5) classifier testing.

Steps 1, 2, and 3 were implemented via OpenCV Python modules (Laganiere, 2014), steps 4 and 5 were implemented via Sklearn Python modules (Pedregosa et al., 2011). For each site, image data sets were randomly divided into training (60%) and testing (40%) data sets. The training data set was used to predict the value of a target class by learning the decision rules inferred from the geometric descriptors of that class (corn or non-corn objects). The trained decision rules were then evaluated in a new data set (testing) to evaluate the performance of the model exposed to an independent data set.

### Vegetation detection

In the training data set, the images were first utilized to classify vegetation and background pixels. The excess green index (ExG) (Meyer and Neto, 2008) helps stretch the contrasting intensity response between green and background pixels. In addition, a bilateral filter was applied to decrease the noise intensity of each channel while preserving the edges of the green objects (Baxes, 1994).

$$ExG = 2 \times Green - Red - Blue \quad (1)$$

A morphological operation was implemented to facilitate the isolation of green contours in the image by computing the corresponding intensity between contours and background. It includes both erosion and dilation transformations by utilizing a predefined kernel size to preserve the integrity of the green objects in the image (Baxes, 1994). An Otsu threshold procedure was adopted to transform the ExG grey scale into a binary image by using a discriminant criterion in the ExG scale. The method automatically finds an optimal threshold

value between both green and background classes (Otsu, 1979) by minimizing the intra-class variance as much as possible.

$$\sigma^2 w(t) = \omega 0(t)\sigma^2 0\ (t) + \omega 1(t)\sigma^2 1\ (t) \qquad (2)$$

ω0 and ω1 are the probabilities of the two classes and σ20 and σ21 are the variances of these two classes. "t" is the desired threshold that minimizes weighted sum of variances of these two classes. The binary transformation assigns a value of 1 to green pixels and 0 to background. Small non-target green contours, <400 connected pixels, are eliminated using a conditional rule.

### Row detection

First, canny edge detection was implemented to map structures with contrasting ExG intensity in the image. Edges are mostly related to the transitional regions between green objects and background pixels (Baxes, 1994). Hough transformation was adopted to define the orientation angle of the images (Hough, 1962). The solution to the angle rotation was solved by a voting process of all possible angles between the Hough lines and the reference horizontal axis of the image. The angle that received more votes was chosen as the solution for the entire image rotation.

The ExG intensity was projected to the vertical axis of the image. The Savitzky–Golay (Savitzky and Golay, 1964) filter was utilized to smooth local-maxima peaks to better target the candidate areas for rows location (Fig. 2.3). A relative threshold value defines the peaks that define the rows in the vertical projection of the image as follows: each peak must reach one third of the previous one, ExG intensity to be assumed as rows in the vertical axis. The selected peaks represent the rows of the crops in the image. In the same manner, the width of each row was equal to the width of the crest at the thresholding region. The process does not require external

user supervision (automated process) to define an optimal threshold to locate the rows, allowing massive scaling of this step.

$$I_A = \begin{cases} I_A & if\ I_i \geq I_{i-1}/3 \\ I_i & if\ I_i < I_{i-1}/3 \end{cases} \quad (3)$$

$I_A$ denotes the threshold that defines whether a peak is a crop row. $I_i$ is the sum of intensities of pixels in $i$th peak. The equation is from i = 2 to $n$ where $n$ is the number of peaks found.

### Feature descriptors

All contours were extracted from the row and inter-row areas of the image and labeled as corn and non-corn contours, respectively. This approach enables the scaling of the training as no manual tagging of classes is needed. The procedure assumes all contours in the inner region of the row belong to “corn class” and all inter-row contours belong to “non-corn class”. Each contour is characterized by a set of 10 geometric descriptors. This step explores the potential of different geometries to efficiently characterize corn and non-corn objects.

Geometric descriptors were evaluated using the feature importance procedure based on the mean decrease of impurity (Guyon and Elisseeff, 2003). Features decreasing the impurity have more importance in the selection, which accounts for potential collinearity between features by penalizing collinear features. According to the feature selection, aspect ratio, axis–diameter ratio, convex area, thinness, and solidity were found as significant contributors to characterize the two types of objects in the training data set.

### Classifier training

A decision tree (DT) classifier was implemented using the information of the geometric descriptors in each class as input features (Alaydin, 2004; Patel, 2012). A 6-fold, cross-validation

(CV) procedure was implemented by leaving one out cross validation (LOOCV). It was utilized as a first approach as to how the classifier may generalize to the new independent data set (testing), and to identify potential overfitting of the model (Alydin, 2004) (Table 2). The DT was trained to relate the descriptors to the labeled corn and non-corn objects. Due to unbalanced sizes between classes, decision nodes were differently weighed to prevent class overfitting in the classifier. The 6-fold CV was used as an intermediate checkpoint of the classifier performance evaluation. The goal of this step was to create a model that predicts the value of a target class by learning the decision rules inferred from the geometric descriptors of that class. A model-selection procedure was used to determine the DT structure by finding a non-dominated solution representing a trade-off between the classifier performance and the computational cost following (Patel, 2012; Breiman et al., 1984) recommendations. Bottom-up pruning of the tree was implemented via a cost-complexity curve (Patel, 2012) removing statistical non-significant nodes, preventing overfitting, and saving the computational cost of the classifier (Eastwood et al., 2012). The optimal structure that minimized computation time without penalizing the classifier performance had a tree depth of 10 levels and 20 sample leaves.

Ground-truthing was implemented via visual inspection of individual plants on the testing data by accounting for: matching, non-matching, and non-detected plants, differences between the labeling output of the classifier, and the visual inspection of the contours. To evaluate the scalability of the classifier, two training modes were considered: (a) local training and local site testing (LTLT) in each site, and (b) combined or joint training and local site testing (JTLT). The LTLT utilizes the site n training data set in training and evaluates the workflow using the site n testing data set. The JLTL utilizes the site n + m training data set in training and evaluates the workflow using the site n testing data set, and later the same evaluation in site m testing data set.

### Classifier performance evaluation

Precision: for x class is the number of true positives (Tp), the number of objects correctly labeled as belonging to the x class divided by the total number of (Tp) plus false positives (Fp) as elements labeled as belonging to the x class but actually were part of class y.

$$\text{Precision} = Tp/(Tp + Fp) \quad (4)$$

Recall: for x class is the number of (Tp) divided by the total number of objects that actually belong to the x class false negatives (Fn), including the (Tp).

$$\text{Recall} = Tp/(Tp + Fn) \quad (5)$$

Accuracy is a global evaluator of the classifier performance for n classes evaluated. The number of objects of n classes have been corrected classified (Tp) and true negatives (Tn) divided by all the objects have been classified.

$$\text{Accuracy} = (Tp + Tn)/(Tp + Tn + Fp + Fn \quad (6)$$

# RESULTS AND DISCUSSION

## Evaluation metrics: training modes

The classifier ability to discriminate classes was evaluated by elaborating receiver operating characteristic (ROC) and precision-recall curves (Powers, 2011; Fawcett, 2006). The performance of the classifier was accounted for at the plant-level basis, predicted object versus ground-truthing findings. A random selection of images was implemented in JLTL to account for a balanced training size and comparison between training modes (Fig. 2.4). The JLTL recall outperforms LTLT in site 1, 0.92 to 0.97. The LTLT better targeted the classification of corn plants (ground-truthing) by reducing "false negatives", non-corn class (ground-truthing) classified as corn objects. Contrarily, LTLT outperforms JTLT in site 2, recall decreases from 0.95 to 0.93, JTLT presents lower power to correctly classify ground-truth non-corn objects ''false negative" as non-corn objects. A higher number of ground-truth corn plants were misclassified as non-corn class objects. Precision slightly decreases when using JTLT, from 0.97 to 0.94 in site 1. A lower performance of the classifier on the "false positive" detection was documented due to a higher number of ground-truth non-corn objects classified as corn. Precision remained stable (0.96–0.97) in site 2 as an indication that "false positive" detection remained stable across training modes. Nevertheless, the overall accuracy followed a decreasing trend between sites when transitioning from LTLT to JTLT mode as noticed in the area under the curves (AUC) (Fig. 2.5). The LTLT reached an accuracy of 0.96 in both sites and decreased for JTLT to 0.92 for site 1 and 0.93 for site 2. The penalization was mainly due to a lower performance of the JTLT classifier on "false positives" detection, a slightly higher misclassification of ground-truth non-corn objects as corn. Outcomes of the LTLT are in

accordance with (Lottes et al., 2016) findings, reporting an accuracy of 0.96, recall of 0.99, and precision of 0.97 between crop and weed objects detection.

### Evaluation metrics: spatial resolution

A data downscaled resolution was simulated to evaluate the sensitivity of the workflow on plant detection by recreating degraded resolutions of increasing flight altitudes. The original resolution of 2.4 mm in site 1 was resized to 4.8, 9.6, and 19.2 (Fig. 2.6), simulating 20, 40, and 80 m flight altitudes, respectively. For downscaling the data, simple linear kernels were implemented: 2 × 2, 4 × 4, 8 × 8 mean values of the original pixels scale values into the resulting downscaled pixel. All workflow steps were fully re-implemented at each downscaled resolution. Manual tuning of the pixels row size was utilized to prevent losing and incorrectly accounting for row and inter-row green objects during the training of the downscaled data set. The classifier accuracy was consistently penalized when the spatial resolution was degraded. Original resolution reached the highest accuracy of 0.96, and decreased to 0.89, 0.85, and 0.68 for each 2.4, 4.8, 9.6, and 19.2 mm resolution, respectively (Fig. 2.6). The P/R curve was penalizing the downscaling following the same trend. Consequently, the overall performance of the classifier was penalized due to a lower sensitivity of geometry as efficient descriptor to differentiate corn and non-corn classes of objects.

It should be noticed that downscaled resolution penalizes the ExG binarization step, and consequently, the ability of the workflow to distinguish objects in the image. The departure between ground-truth objects and the classifier detection assists with a metric on the sensitivity of the workflow to detect green objects in the image. A total of 15 images were selected for this analysis. The departure from ground-truthing (Fig.2.7) represents the relative distance between the number of true detected objects and the ones reported by the classifier when analyzing the

images. When using the original 2.4-mm resolution, the penalization on the sensitivity to detect green objects remains very low (1.5%). When downscaled to 4.8, 9.6, and 19.2 mm, the penalization increases to 6%, 12%, and 42%, respectively. Downscaled spatial resolution increases the ground sampling distance (GSD), meaning that a larger area on the ground is sensed per pixel unit. Thus, it becomes critical for transitional (green objects borders-background) areas of the image for quality contours delineation in the image. An increasing number of double objects by unit of contour due to "mixed signals" was progressively found when transitioning from finer to downscaled data generating an underestimation of the total number of contours (green objects) (Fig. 2.7).

Current methods propose the use of ground vehicles or satellite data to estimate detailed information of plant status at the fields. The first one evidenced limitations by only being able to cover small areas and depending on good trafficable conditions. The second one does not provide the needed spatial resolution and the performance on this kind of task remains weak. The proposed workflow exploits synergistically the versatility of UAS platforms and a supervised learning procedure to identify crops and non-crops in the field enabling the differentiation between corn plants and weeds early in the season. In addition, the proposed workflow allows the identification and mapping of plants at a very early time in the season using real farm conditions and balancing the classifier performance between both corn and non-corn objects.

A few limitations from the tested method include: (i) late within-growing-season flights are prone to plant overlapping degrading the workflow performance and causing underestimation in the plant count; and (ii) plant density was not evaluated at field-scale since (1) the focus of this project remains in the evaluation of the classifier performance by itself corn plant identification, and (2) accurate field-scale plant density estimation needs accurate and precise ground scaling of

the individual imagery via RTK (Real Time Kinematic) or PPK (Post Processed Kinematic) global navigation satellite systems (GNSS). An opportunity for delivering large scale, more efficient, and faster models can be pursued by collecting UAS data using a sub-sampling imagery strategy and spatial analysis. The latter appears as a potential solution for saving computational costs of processing data and preventing a degraded resolution from the original imagery when building an orthomosaic via the stitching procedure.

The main contribution of this paper is related to the development of a procedure to detect corn plants to better guide early season operations for farmers. The foundation of the method relies on the combination of traditional imagery and a supervised learning procedure. The outcome of the workflow allows the digital counting of plants using a low-cost UAS and RGB camera contributing to quantify early-season data of crop performance at on-farm conditions.

Future work should (a) study the potential of spectral and texture descriptors for classes delineation, (b) explore the potential of including multiclass non-corn objects by reducing the internal variance of non-corn objects, and (c) investigate the penalization of high wind conditions in the geometric descriptors and classifier performance.

.

# CONCLUSIONS

In this work, a workflow to identify corn plants in real field conditions was implemented using vegetation detection, feature extraction, and classification using aerial images by exploiting geometric descriptors information. The developed workflow utilizes the spatial arrangement of crops to scale up the training of the classifier. The proposed approach was implemented and tested with imagery data collected via UAS at two farm fields to evaluate the upscaling robustness of the workflow and the potential applications on farm operations. Even though the combined sites' training (0.92 and 0.93) performed lower than local site training mode (0.96), the combined training mode is still robust for scaling up the processes and, most importantly, to save computational time when dealing with massive amounts of data in the post-processing steps. The original 2.4 mm resolution portrayed the best performance to detect corn objects. Downscaled spatial resolutions gradually negatively impacted the workflow performance at two levels: (i) evidencing a lower sensitivity to detect green contours in the image due to an increased mixed signal (soil background-green objects) that degraded the contours delineation and (ii) decreasing the power of the classifier itself due to a degraded power of the geometry as an effective descriptor to differentiate both classes of objects. Results suggest that the optimal growth for accurate estimation for a field setting of corn plants is between two and three leaves.

# REFERENCES


Alaydin, E. Introduction to Machine Learning; MIT Press: Cambridge, MA, USA, 2004.

Baxes, G.A. Digital Image Processing, Principles and Application; John Wiley & Sons: Hoboken, NJ, USA, 1994; ISBN 0-471-00949-0.

Bendig, J.; Willkomm, M.; Tilly, N.; Gnyp, M.L.; Bennertz, S.; Qiang, C.; Miao, Y.; Lenz Wiedemann, V.I.S.; Bareth, G. Very high resolution crop surface models (CSMs) from UAV-based stereo images for rice growth monitoring in Northeast China. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. 2013, 40, 45–50.

Berni, J.; Zarco-Tejada, P.; Suarez, L.; Fereres, E. Thermal and narrowband multispectral remote sensing for vegetation monitoring from an unmanned aerial vehicle. IEEE Trans. Geosci. Remote Sens. 2009, 47, 722–738.

Breiman, L.; Friedman, J.; Olshen, R.; Stone, C. Classification and Regression Trees; Wadsworth: Belmont, CA, USA, 1984.

Ciampitti, I.A.; Vyn, T.J. A comprehensive study of plant density consequences on nitrogen uptake dynamics of maize plants from vegetative to reproductive stages. Field Crop. Res. 2011, 121, 2–18.

Caturegli, L.; Corniglia, M.; Gaetani, M.; Grossi, N.; Magni, S.; Migliazzi, M.; Angelini, L.; Mazzoncini, M.; Silvestri, N.; Fontanelli, M.; et al. Unmanned Aerial Vehicle to Estimate Nitrogen Status of Turfgrasses. PLoS ONE 2016, 11, e0158268.

Clevers, J.G.P.W.; Kooistra, L. Using Hyperspectral Remote Sensing Data for Retrieving Canopy Chlorophyll and Nitrogen Content. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2012, 5, 574–583.

De Bruin, J.; Pedersen, P. Early Season Scouting; Extension and Outreach. IC-492:7; Iowa State University: Ames, IA, USA, 2004; pp. 33–34.

De Souza, C.H.W.; Lamparelli, R.A.C.; Rocha, J.V.; Magalhaes, P.S.G. Height estimation of sugarcane using an unmanned aerial system (UAS) based on structure from motion (SfM) point clouds. Int. J. Remote Sens. 2017, 38, 2218–2230.

Eastwood, M.; Gabrys, B. Generalised bottom-up pruning: A model level combination of decision trees. Expert Syst. Appl. 2012, 39, 9150–9158.

Fawcett, T. An introduction to ROC analysis. Pattern Recognit. Lett. 2006, 27, 861–874.

Gnädinger, F.; Schmidhalter, U. Digital counts of maize plants by unmanned aerial vehicles (UAVs). Remote Sens. 2017, 9, 544.

Geipel, J.; Link, J.; Claupein, W. Combined spectral and spatial modeling of corn yield based on aerial images and crop surface models acquired with an unmanned aircraft system. Remote Sens. 2014, 6, 10335–10355.

Gonzalez-Dugo, V.; Zarco-Tejada, P.; Nicolas, E.; Nortes, P.A.; Alarcon, J.J.; Intrigliolo, D.S.; Fereres, E. Using high resolution UAV thermal imagery to assess the variability in the water status of five fruit tree species within a commercial orchard. Precis. Agric. 2013, 14, 660–678.

Gomez-Candon, D.; Virlet, N.; Labbe, S.; Jolivot, A.; Regnard, J.L. Field phenotyping of water stress at tree scale by UAV-sensed imagery: New insights for thermal acquisition and calibration. Precis. Agric. 2016, 17, 786–800.

Guyon, I.; Elisseeff, A. An Introduction to Variable and Feature Selection. J. Mach. Learn. Res. 2003, 3, 1157–1182.

Haboudane, D.; Miller, J.R.; Pattey, E.; Zarco-Tejada, P.J.; Strachan, I.B. Hyperspectral

vegetation índices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. Remote Sens. Environ. 2004, 90, 337–352.

Hale Group. The Digital Transformation of Row Crop Agriculture, AgState Electronic Survey Findings. 2014. Available online: http://www.iowacorn.org/document/filelibrary/membership/agstate.AgState_Executive_Summary_0a58d2a59dbd3.pdf (accessed on 19 December 2017).

Hough, P.V.C. A Method and Means for Recognizing Complex Patterns. U.S. Patent 3,069,654, 18 December 1962.

Iqbal, F.; Lucieer, A.; Barry, K.; Wells, R. Poppy Crop Height and Capsule Volume Estimation from a Single UAS Flight. Remote Sens. 2017, 9, 647.

Lauer, J.G.; Rankin, M. Corn Response to Within Row Plant Spacing Variation. Agron. J. 2004, 96, 1464–1468.

López-Granados, F. Weed detection for site-specific weed management: Mapping and real-time approaches. Weed Res. 2011, 51, 1–11.

Lottes, P.; Hoferlin, M.; Sander, S.; Muter, M.; Schulze-Lammers, P.; Stachniss, C. An effective classification system for separating sugar beets and weeds for precision farming applications. In Proceedings of the IEEE International Conference on Robotics & Automation (ICRA), Stockholm, Sweden, 16–21 May 2016.

Nafziger, E.D.; Carter, P.R.; Graham, E.E. Response of corn to uneven emergence. Crop Sci. 1991, 31, 811–815.

Nakarmi, A.D.; Tang, L. Automatic inter-plant spacing sensing at early growth stages using a 3D vision sensor. Comput. Electron. Agric. 2012, 82, 23–31.

Nakarmi, A.D.; Tang, L. Within-row spacing sensing of maize plants using 3D computer vision. Biosyst. Eng. 2014, 125, 54–64.

Nielsen, R.L. Stand Establishment Variability in Corn; Publication AGRY-91-01; Purdue University: West Lafayette, IN, USA, 2001.

Nielsen, B. Estimating Yield and Dollar Returns from Corn Replanting; Purdue University Cooperative Extension Service: West Lafayette, IN, USA, 2003.

Mathews, A.J.; Jensen, J.L.R. Visualizing and Quantifying Vineyard Canopy LAI Using an Unmanned Aerial Vehicle (UAV) Collected High Density Structure from Motion Point Cloud. Remote Sens. 2013, 5, 2164–2183.

Peña, J.M.; Torres-Sanchez, J.; Serrano-Perez, A.; de Castro, A.I.; Lopez-Granados, F. Quantifying efficacy and limits of unmanned aerial vehicle UAV technology for weed seedling detection as affected by sensor resolution. Sensors 2015, 15, 5609–5626.

Perez-Ortiz, M.; Peña, J.M.; Gutierrez, P.A.; Torres-Sanchez, J.; Hervas-Martınez, C.; Lopez-Granados, F. Selecting patterns and features for between- and within-crop-row weed mapping using UAV imagery. Expert Syst. Appl. 2016, 47, 85–94.

Pölönen, I.; Saari, H.; Kaivosoja, J.; Honkavaara, E.; Pesonen, L. Hyperspectral imaging-based biomass and nitrogen content estimations from light-weight UAV. In Proceedings of the SPIE Remote Sensing, Dresden, Germany, 23–26 September 2013.

Potgieter, A.B.; George-Jaeggli, B.; Chapman, S.C.; Laws, K.; Suárez Cadavid, L.A.; Wixted, J.; Wason, J.; Eldridge, M.; Jordan, D.R.; Hammer, G.L. Multi-spectral imaging from an unmanned aerial vehicle enables the assessment of seasonal leaf area dynamics of sorghum breeding lines. Front. Plant Sci. 2017, 8.

Jin, X.; Liu, S.; Baret, F.; Hemerlé, M.; Comar, A. Estimates of plant density of wheat crops at

emergence from very low altitude UAV imagery. Remote Sens. Environ. 2017, 198, 105–114.

Guo, W.; Rage, U.K.; Ninomiya, S. Illumination invariant segmentation of vegetation for time series wheat images based on decision tree model. Comput. Electron. Agric. 2013, 96, 58–66.

Henry, M. Big Data and the Future of Farming; Australian Farm Institute Newsletter: Surry Hills, Australia, 2015; Volume 4.

Montalvo, M.; Pajares, G.; Guerrero, J.M.; Romeo, J.; Guijarro, M.; Ribeiro, A.; Ruz, J.J.; Cruz, J.M. Automatic detection of crop rows in maize fields with high weeds pressure. Expert Syst. Appl. 2012, 39, 11889–11897.

Meier, L.; Honegger, D.; Pollefeys, M. PX4: A node-based multithreaded open source robotics framework for deeply embedded platforms. In Proceedings of the IEEE International Conference on Robotics & Automation (ICRA), Seattle, WA, USA, 26–30 May 2015.

Laganiere, R. OpenCV 2 Computer Vision Application Programming Cookbook; Packt Publishing Ltd.: Birmingham, UK, 2014.

Meyer, G.E.; Neto, J.C. Verification of color vegetation indices for automated crop imaging applications. Comput. Electron. Agric. 2008, 63, 282–293.

Otsu, N. A threshold selection method from gray-level histogram. IEEE Trans. Syst. Man Cybern. 1979, 9, 62–66.

Patel, N.; Upadhyay, S. Study of various decision tree pruning methods with their empirical comparison in WEKA. Int. J. Comput. Appl. 2012, 60, 20–25.

Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V. Scikit-learn: Machine Learning in Python. J. Mach. Learn. Res. 2011, 12, 2825–2830.

Powers, D.M.W. Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation. J. Mach. Learn. Technol. 2011, 2, 37–63.

Salami, E.; Barrado, C.; Pastor, E. UAV flight experiments applied to the remote sensing of vegetated areas. Remote Sens. 2014, 6, 11051–11081.

Savitzky, A.; Golay, M.J.E. Smoothing and differentiation of data by simplified least squares procedures. Anal. Chem. 1964, 36, 1627–1639.

Shi, Y.; Wang, N.; Taylor, R.K. Automatic corn plant location and spacing measurement using laser line-scan technique. Precis. Agric. 2013, 14, 478–494.

Shrestha, D.S.; Steward, B.L. Size and Shape Analysis of Corn Plant Canopies for Plant Population and Spacing Sensing. Appl. Eng. Agric. 2005, 21, 295–303.

Thorp, K.R.; Steward, B.L.; Kaleita, A.L.; Batchelor, W.D. Using aerial hyperspectral remote sensing imagery to estimate corn plant stand density. Trans. ASABE 2008, 51, 311–320.

Thorp, K.R.; Tian, L.; Yao, H.; Tang, L. Narrow-band and derivative-based vegetation indices for hyperspectral data. Trans. ASAE 2004, 47, 291–299.

Tokekar, P.; Hook, J.V.; Mulla, D.; Isler, V. Sensor planning for a symbiotic UAV and UGV system for precision agriculture. IEEE Trans. Robot. 2016, 32, 5321–5326.

Torres-Sanchez, J.; Lopez-Granados, F.; Peña, J.M. An automatic object-based method for optimal thresholding in UAV images: Application for vegetation detection in herbaceous crops. Comput. Electron. Agric. 2015, 114, 43–52.

Torres-Sánchez, J.; Peña, J.M.; de Castro, A.I.; López-Granados, F. Multi-temporal mapping of the vegetation fraction in early-season wheat fields using images from UAV. Comput. Electron. Agric. 2014, 103, 104–113.

Uto, K.; Seki, H.; Saito, G.; Kosugi, Y. Characterization of Rice Paddies by a UAV-Mounted

Miniature Hyperspectral Sensor System. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2013, 6, 851–860.

Wiedong, L.; Tollenaar, M.; Stewart, G.; Deen, W. Impact of planter type, planting speed, and tillage on stand uniformity and yield of corn. Agron. J. 2004, 96, 1668–1672.

Yao, X.; Wang, N.; Liu, Y.; Cheng, T.; Tian, Y.; Chen, Q.; Zhu, Y. Estimation of Wheat LAI at Middle to High Levels Using Unmanned Aerial Vehicle Narrowband Multispectral Imagery. Remote Sens. 2017, 9, 1304.

Zhou, X.; Zheng, H.B.; Xu, X.Q.; He, J.Y.; Ge, X.K.; Yao, X.; Cheng, T.; Zhu, Y.; Cao, W.X.; Tian, Y.C. Predicting grain yield in rice using multi-temporal vegetation indices from UAV-based multispectral and digital imagery. ISPRS J. Photogramm. Remote Sens. 2017, 130, 246–255.

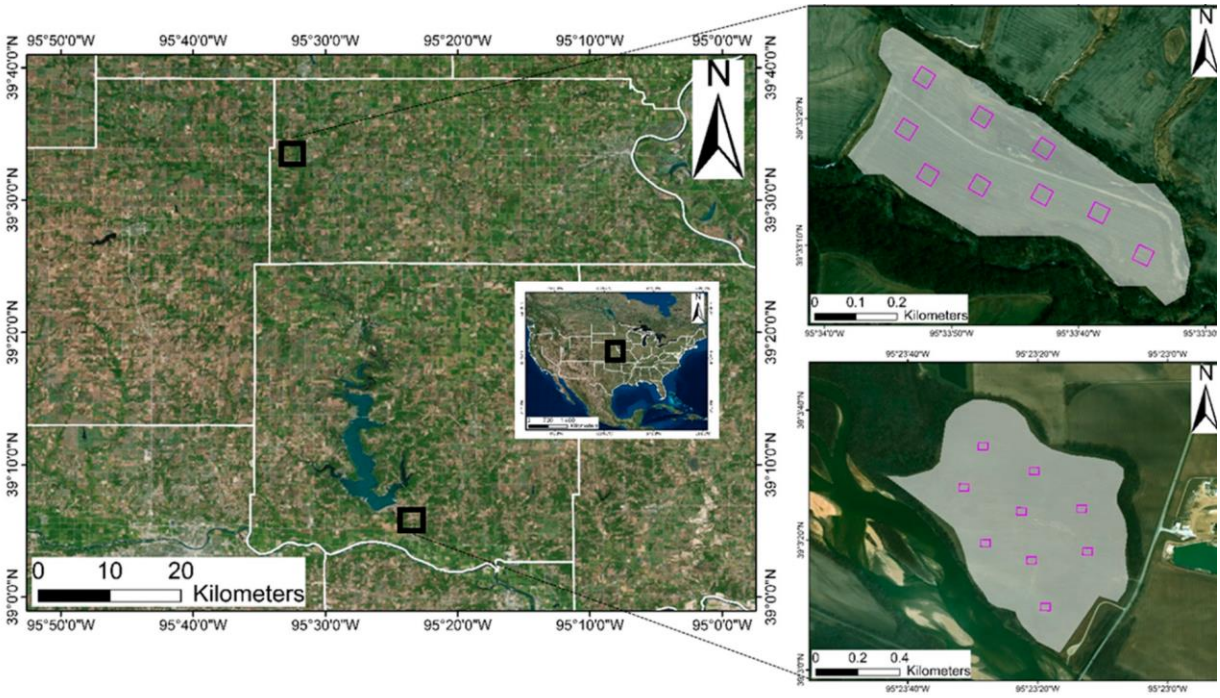


**Figure 2.1. Left: On-farm fields located in the northeast region of Kansas. Top-right: Site 1, Atchison, KS; bottom-right: Site 2, Jefferson, KS. Purple squares = field sampled areas.**

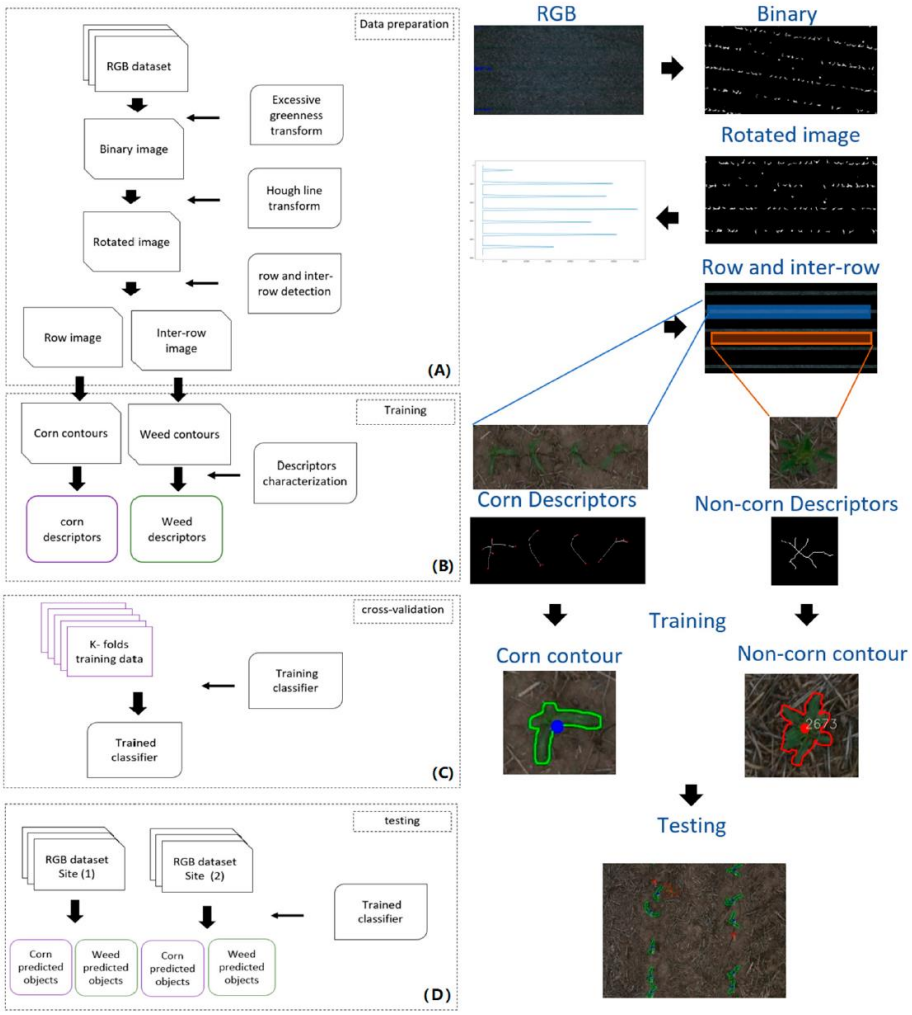


**Figure 2.2 Workflow for plant estimation via unmanned aerial systems (UAS). (A) data pre-processing, (B) training, (C) cross-validation, and (D) testing.**

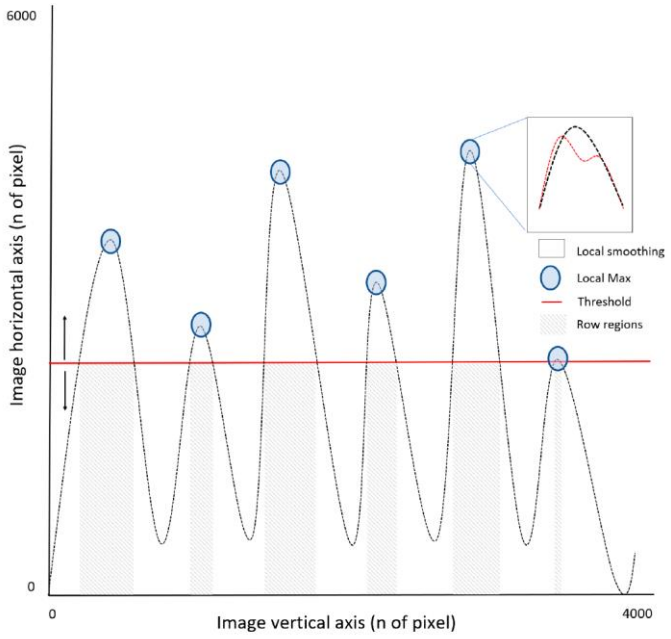


**Figure 2.3. Diagram of the Excess Greenness (ExG) index projection, local-maxima smoothing, and thresholding for rows location.**

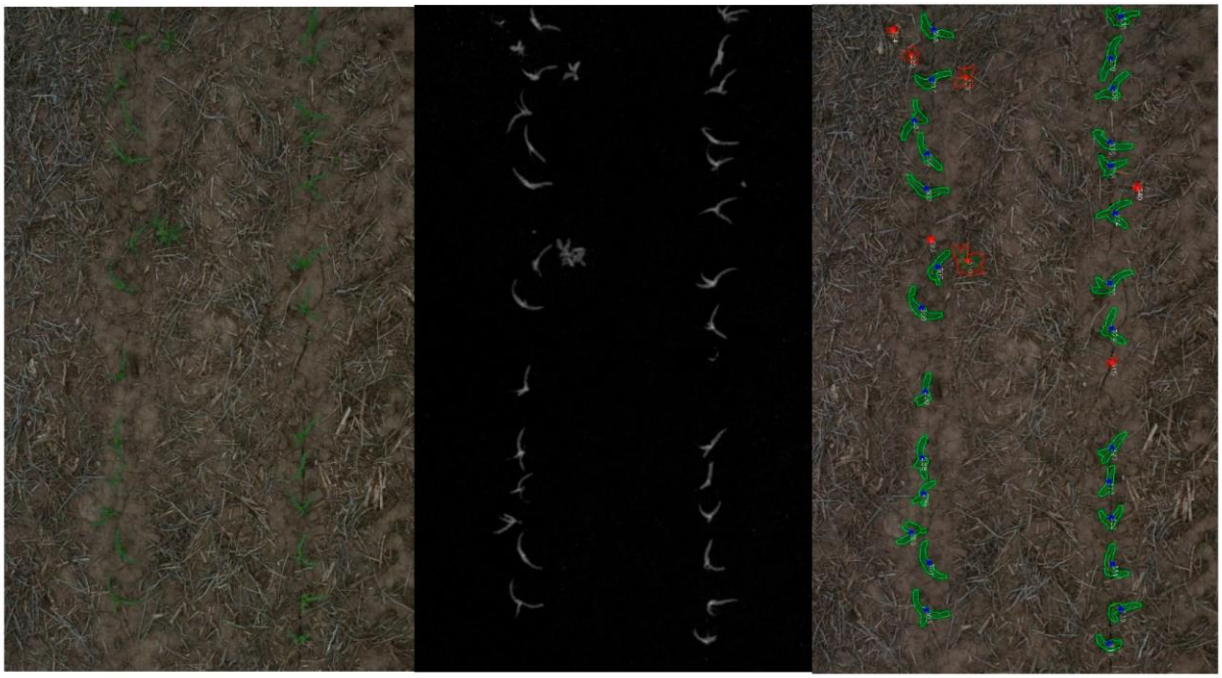

**Fig. 2.4. Left: RGB, center: ExG, right: classifier output on testing data in site 1, green contours: corn objects, red contours: non-corn objects**.

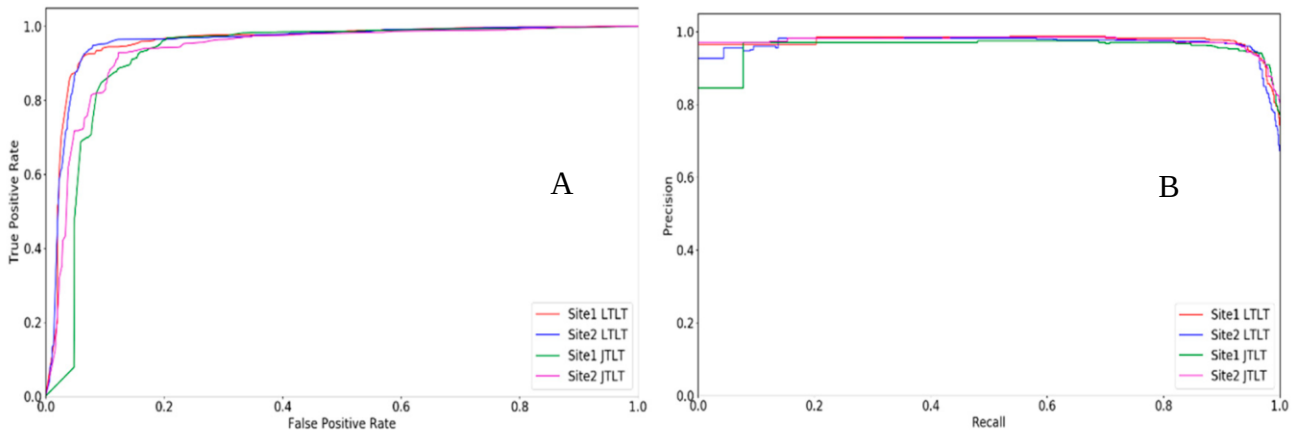


**Figure 2.5. Receiver operating characteristic (ROC) curves (A) and positive rate (PR) plots (B) based on testing data for each site.**

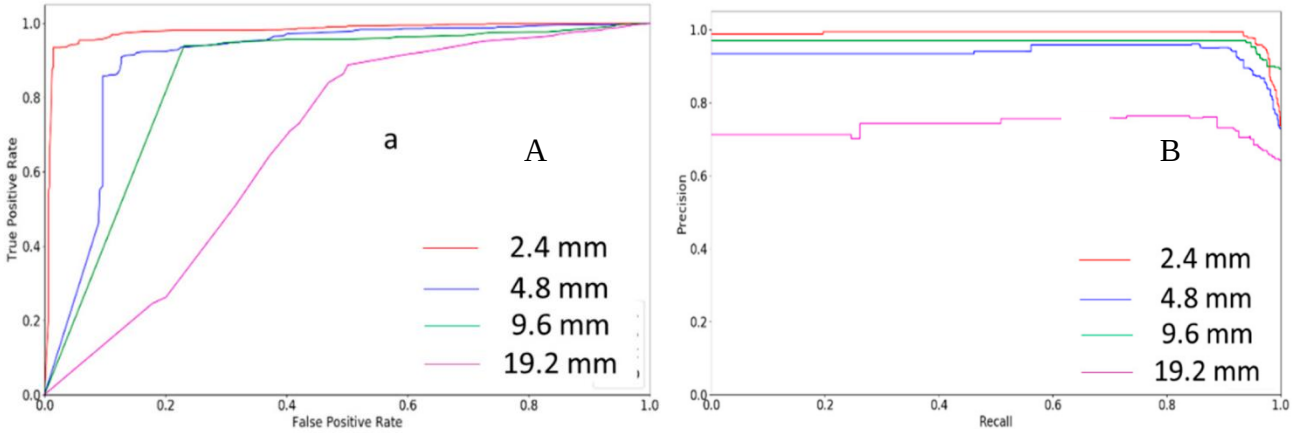


**Figure 2.6. ROC curves (A), and PR plots (B) of downscaled testing data set in testing resolutions.**

## Table 2.1. Information about sites and flights during the 2017 growing season.

Table 1. Information about sites and flights during the 2017 growing season.

| Fields | Previous Crop | Planting Date (DOY) | Growth Stage | Flight Day (DOY) | Flight Altitude (m) |
|---|---|---|---|---|---|
| Site 1 | Soybean | 116 | v2 | 135 | 10 |
| Site 2 | Soybean | 130 | v2–v3 | 153 | 10 |

## Table 2.2 Data sets used for training and testing of the classifier.

Table 2. Data sets used for training and testing of the classifier.

| | Site 1 | | Site 2 | |
|---|---|---|---|---|
| Data Set | Training | Testing | Training | Testing |
| Images | 94 | 75 | 87 | 75 |
| Contours | 17,608 | 15,378 | 16,855 | 15,246 |

# Chapter 3 - Monitoring phenological footprint of corn in the Mid-western USA via MODIS time-series


## ABSTRACT

Monitoring crop phenology is a fundamental proxy to understand crop adaptation over large regions. 250-m daily time-series of Moderate Resolution Imaging Spectroradiometer (MODIS) Normalized Difference Vegetation Index (NDVI) was utilized for screening relevant transition growth stages in corn (*Zea mays* L.) development. The spatial and temporal dynamic of phenology and yield was investigated via Hierarchical Bayesian Spatiotemporal Modeling (HBSTM) to link the timing of growth stages of the crop and yield in the United States (U.S.) corn belt between 2003 and 2017. The robustness of MODIS -based phenology metrics was evaluated at two levels: i) in-season ground truth (GT) via an extensive field survey in Kansas, and ii) with data obtained from the Crop Progress and Condition Report (CPCR) from the U.S. Department of Agriculture (USDA) National Agricultural Statistics Service (NASS) during the 2006-2017 time period. A threshold value of 20% and 35% of relative amplitude of the time-series profile reported the best agreement with GT and CPCR emergence and physiological maturity dates, respectively. A root-mean square error (RMSE) of 16 days for start of the season, 14 days for the vegetative-reproductive transition, and 17 days for the end of season on metrics extraction was reported. The length of vegetative and reproductive stage, green-up and senescence rates were positively related to the increase on yield. MODIS-based phenology metrics significantly describe the increase on yield over both, space and time in the region during the last 14 years.



**Keywords:** MODIS, phenology, corn, time-series, spatial-temporal.

# INTRODUCTION

Monitoring crop phenology is a key component of studies related to crop adaptation (Sahu and Bakar., 2012; Zheng et al., 2002; Tao et al., 2006), yield forecasting (Bolton et al 2013; Fang et al., 2011), crop mapping (Zhong et al., 2016) and growth modeling (Moulin et al., 1997; Fang et al., 2011). The spatio-temporal dynamics of phenology can reveal adaptation over extensive regions at reduced cost (Yusoff et al., 2017; Estel et al., 2015). In United States, the Department of Agriculture (USDA) generates weekly Crop Progress and Condition Reports (CPCRs) between March and November each year. The CPCRs are field surveys-based reports describing crop condition and progress of major crops in the country. For the CPCRs, the reported information is weighed by county crop acreage and summarized at the agriculture statistical district (ASD) and state level by the regional field offices of the National Agricultural Statistics Service (NASS). These reports had been used as input for governmental agencies, policy-makers, industry, and universities. Private stakeholders and insurance companies use these reports for early-season diagnosing of overall crop productivity, short-term price and future market negotiation. Past researchers have been using this information as to track long-term trends over the past several decades related to changes in planting date and phenology (Sacks and Kucharik, 2011; Kucharick, 2006; Shen and Liu, 2015).

Tracking the occurrence of key crop growth and development stages via remote sensing (RS) time-series had been previously reported. For example, using Advanced Very High Resolution Radiometer (AVHRR) and Moderate Resolution Imaging Spectroradiometer (MODIS) researchers developed operational land surface spatio-temporal data for global phenology monitoring (Moulin et al., 1997; Zhang et al 2003). However, monitoring annual crops implies significant challenge due to the high spatial and temporal dynamics of this type of

vegetation (Zheng et al., 2002). Wardlow et al. 2006 derived green-up onset dates for corn (*Zea mays* L.), soybean (*Glycine max* L.), and sorghum (*Sorghum bicolor* L.) in Kansas using 16-day MODIS Normalized Difference Vegetation Index (NDVI) data. The same authors compared MODIS-derived green-up onset dates with the USDA CPCRs data finding large inconsistencies across ASDs. Sakamoto et al. 2010 developed a two-step filtering (TSF) method to detect phenological stages of corn and soybean using time-series Wide Dynamic Range Vegetation Index (WDRVI) derived from MODIS, data from a 6-yr period. Their RS-based metric outcomes were consistent with ground-based observations for two irrigated and one rainfed sites in Nebraska. This crop phenology detection method, however, has not yet been examined on a large geographical scale and using a longer-term data, more than 10 years. On a broader scale, (Ren et al. 2017) reported the extraction of RS-based phenology parameters (start and end of season) at state level in the US corn belt region, but quantification of in-seasonal stages of the crop has been scarce or not yet reported. At ASD scale, it is still unclear whether there is a robust relationship between in-season RS-based metrics and ground truth-based estimates, especially for extensive field crops such as corn in the US corn belt region. Although CPCRs have their own uncertainties reflecting non-biased observations, they are the most widely accepted source of information and statistics relative to crop phenology in the US. Furthermore, the potential of daily-based satellite imagery for detecting short-term anomaly/change on canopies have not been fully exploited (Zhao et al., 2012). However, it is a challenge to generalize such algorithms to a broad-scale, agricultural landscape and for relatively long-term periods with varying weather conditions. Broad-scale identification and monitoring of phenological dynamics of crops over extensive time can help to better understand the drivers of adaptation and performance of modern hybrids over extensive regions.

Information reporting the timing and transition of corn phenology such as start of the season (SOS), vegetative-reproductive transition (VRT) and end of the season (EOS) and further derivation of each of the phenology phases is scarce. (Sacks and Kucharik, 2011) reported the relevance of seasonal length and crop phenology for crop adaptation in Mid-western US between 1981 and 2005. They reported that the period between corn planting to maturity was about 12 days longer around 2005 than it was in 1981. A larger driver of this change was a 14% increase in the number of GDD (Growing Degree Days) needed for corn to progress through the reproductive period, probably reflecting the adoption of longer season cultivars. New insights that unveil the connection between the timing of the season, development and yield in the crop are needed to better understand crops adaptation and management practices in different regions. The overall goals of this work were to: i) derive key phenological metrics for corn over the US corn belt via high temporal resolution vegetation index, ii) benchmark these phenology metrics against ground-truth and CPCRs data at the ASD level, and iii) identify the links between the phenological metrics and yield trends over the last 14 years in the corn belt region.

# MATERIALS AND METHODS

## Data Processing

The region of the study involved the US states of Kansas, Missouri, Iowa, and Indiana, characterized as a relevant corn producing region for the country (Fig. 3.1). The abovementioned US states have in combination an overall production of 335 million tonnes representing 41.4% of the total corn production at the country level according to USDA NASS (USDA, 2017).

NDVI is the most widely adopted vegetation index for crop monitoring (Wardlow et al., 2006). It denotes high sensitivity in the red and near-infrared regions of the electromagnetic spectrum, particularly relevant to detect vegetation onset, plant growth, fraction of photosynthetic active radiation (fPAR), and leaf area index (LAI) as bio-physical parameters.

$$NDVI = (NIR - Red)/(NIR + Red) \quad (1)$$

MODIS Terra Surface Reflectance Daily L2G Global 250-m SIN Grid V006 (MOD09GQ) product was utilized as primary input in the process. Imagery data for this study were acquired from the National Aeronautics and Space Administration (NASA) (https://e4ftl01.cr.usgs.gov/) from 2003 to 2017. Processing was first implemented via MODIStsp R package (Bussetto et al., 2016), including download of MODIS (h09-h11, v04-v5) extent tiles, extract, mosaic, project (Albert Equal Area Conic), generate and stack daily NDVI layers for further processing. For the 2003-2017 period, 12,460 MODIS scenes were included in the analysis. Corn cropland masks were generated from Crop Data Layer (CDL) produced by USDA NASS (Boryan et al., 2011). Complete CDL coverage for the states involved in the analysis were available since 2003. Multi-year USDA NASS CDL layers were included as key

inputs in our workflow to locate/extract pixels of corn fields (2003-2017) for the entire region. CDL corn pixels were sampled using area weight sample via "GridSample" (Tatem et al., 2017) contributed package to R [R Core Team, 2017] to prevent geographical bias when validating versus CPCRs. Point samples geolocation were imported back to "MODIStsp" R and used as centroids for the extraction of time-series profiles using only pixels fully contained within fields of corn. Raw time series contain pixels with unreliable values. Although MOD09GQ contains some data quality information, it lacks pixel quality information that MOD09GA (Vermote and Wolfe, 2015) has stored. The MOD09GA contains binary-coded cloud state, cloud shadow, sensor zenith angle, solar zenith angle information useful for contaminated pixel filtering. The MOD09GA quality data layers have 500m resolutions and are re-sampled to match the 250m reflectance data layers. The binary-coded quality information is then interpreted into various quality statuses for filtering purpose. The pixel validation status, cloud state and cloud shadow were used as the criteria for filtering cloud contaminated and low-quality pixels. A second step includes a data smoothing procedure via SG (Savitzky-Golatz) (Savitzky and Golay, 1964) to improve the signal to noise ratio but at the same time preserve in the signal critical inflexion regions related to phenological transitions in the season. SG filter is utilized as low pass filter to better characterize corn seasonal dynamic and for extraction of phenological metrics (Chen et al., 2004). The SG filter applies a moving window and fits a quadratic polynomial function to the raw time-series and estimates new values for the center point of each moving window. Smoothing of daily MODIS NDVI was implemented using "savgol" function in "Prospectr" R package (Stevens, 2013). The SG function was set to third order polynomial and two consecutive rolling fits were utilized using first a window of 30 days to reduce larger peaks followed by a second 16 days window accounting for small peaks as final smooth. This two-step procedure

reported the best performance to denoise NDVI time-series signal and preserve key temporal features (local minimum, inflection points and local maximum) of the profile during the growing season (further details in Fig. 3.2).

Final metrics extraction was implemented in R via self-designed code considering the following definitions for phenology metrics (Table 3.1). MODIS-based phenology parameters (SOS, VRT, EOS) were evaluated at two levels (Fig. 3.2). First, implementing an extensive field survey including with five revisited times of 290 fields across Kansas during the 2017 growing season. The timing of summer crop field operations and physiological growth stages vary across Kansas. The southeast region has earlier planting dates, ahead from the Northeast and Western regions. This is mainly due to soil temperatures warming earlier with more growing degree (GDU) accumulation from southern to northern parts of the state and with an increasing gradient on precipitation from West to East. Second, MODIS-based phenology was compared using a large collection of CPCRs between 2006-2017 period for the designated US states. The CPCRs of crop development stages reported as percentages of completed phenological phases in the weekly reports (emergence, silking/tasseling, dough, maturity) were used as benchmark for metric extraction for this larger region (Fig. 3.3, portrays pictures of corn stages).

A reduced time window (April-November) was utilized following (Ren et al., 2017) recommendations in order to enhance polynomial smoothing and metrics extraction. Decreasing this window reduces the percent of snow in a given pixel for that period reducing the noise signal in the time series profile. Farmers in the Northcentral US corn belt region typically start planting corn around mid-April and complete harvest by early November. Unfavored early season conditions (e.g., cold temperatures and drought) could negatively impact farm operations, delay crop development and negatively impact yield potential. Therefore, corn SOS, VRT and EOS

detection can be focused within the analytical window of early-April to end-October for the states located in the Northcentral US corn belt region (e.g., Kansas, Iowa, Missouri, and Indiana). The phenology metrics extracted from the field centroids within each county were averaged and assigned to the county scale and matched to year-based yield (USDA, NASS) for further comparison and modeling.

HBSTM (Hierarchical Bayesian Spatiotemporal Modeling) was implemented to account for the spatial and temporal structure between MODIS-based phenology metrics and yield at the county level in the region. Bayesian methods allows to reduce uncertainty in inference statements that arise from joint space-time modeling (Cressie and Wikle, 2011). They become popular because of their ability to combine information from different sources over long-length time series data (Bakar and Sahu, 2013). It is process-based inference and incorporates random spatial and temporal process in the outcome statements. In this paper, a Bayesian hierarchical spatio-temporal model was proposed to analyze the data.

spTimer contributed package (Bakar and Sahu, 2013) in R environment was utilized for HBSTM implementation. The package incorporates the derived phenology metrics as covariables of the explanatory (yield) that vary in both space and time. The Bayesian spatio-temporal models can be represented in a hierarchical structure, where, according to (Gelfand, 2012), we specify distributions for data, process and parameters in three stages:

First: [data| process, parameter]

Second: [process| parameter]

Third: [parameter]

In the second stage, the process can add different levels, for example in Gaussian process models (Cressie and Winkle, 2011), we have true underlying process in the first level and the spatio-temporal random effect in the second level of the hierarchy. In the third stage of the hierarchy we introduce the prior distribution of the hyperparameters.

We first explore an independent GP (Gaussian Process) approach which considers a hierarchical nugget effect together with an independent GP (Gaussian Process) model at each time point. The Gaussian process implies a spatio-temporal random effect that captures the space-time interactions (Cressie and Wikle, 2011). Overall, this model parallels the spatial random effect model in "spatial only" data analysis and naturally provides a very simple starting model where successive events in time are considered independent. GP can be specified as follows:

$$Zlt = Olt + \epsilon lt \qquad (2)$$

$$Olt = Xlt\beta + \eta lt \qquad (3)$$

Let $Z_{lt}(s_1, t)$ denote the observed point referenced data and $O_{lt}(s_1, t)$ be the true value corresponding to $Z_{lt}(si, t)$ at site si, i = 1, ..., *n* at time denoted by two indices *l* and *t*. Let $Zlt = (Z_{lt}(s_1, t), ...,Z_{lt}(s_n, t))'$ and $O_{lt} = (O_{lt}(s_1, t), ...,O_{lt}(s_n, t))'$. Let $N = nrT$ be the total number of observations to be modeled. Throughout, the notation $O_{lt} = (O_l(s_1, t), ..., O_l(s_n, t))'$ will be used to denote the so called nugget effect or the pure error term assumed to be independently normally distributed $N(0, \sigma^2\varepsilon I_n)$ where $\sigma 2\varepsilon$ is the unknown pure error variance and $I_n$ is the identity matrix of order *n*. The spatio-temporal random effects will be denoted by $\eta lt = (\eta l (s1, t), ..., \eta l (sn, t))'$ and these will be assumed to follow $N(0,\Sigma\eta)$ independently in time, where $\Sigma\eta = \sigma^2\eta S\eta$, $\sigma^2\eta$ is the

site invariant spatial variance and $S_\eta$ is the spatial correlation matrix obtained from the general Matern correlation function (Matern, 1996). We assume that there are *p* covariates, including the intercept, denoted by the n × p matrix $X_{lt}$. Some of these covariates may vary in space and time. The notation $\beta = (\beta 1, ..., \beta p)$ will be used to denote the p×1 vector of regression coefficients.

We further implement a hierarchical AR (Auto-Regressive) model for space-time with an explicit auto-regressive term where the underlying true spatio-temporal process is assumed in a hierarchical set-up that includes the overall nugget-effect (Sahu and Bakar, 2012). The main objective was to explicitly identify and valuate the "temporal" autoregressive component in the structure of the data and understand the temporal dependency in the data. The model can be specified as follows:

$$Zlt = Olt + \epsilon lt \quad (4)$$

$$Olt = \rho Olt_{-1} + Xlt\beta + \eta lt, \quad (5)$$

for all l and t; where, ρ denotes the unknown temporal correlation parameter assumed to be in the interval (−1, 1). Obviously, when $\rho = 0$, these models reduce to the GP models. We continue to assume the Gaussian distributions, introduced, for $\epsilon lt$ and $\eta lt$ for all values of l and t.

The significance of variables (phenological metrics) parametrization on yield response was evaluated via trace, autocorrelation, and density plots for further interpretation of the model convergence (Fig. 3.7). The quality of the GP and AR models' fit is evaluated via predictive model choice criteria (PMCC), (Gelfand and Ghosh, 1998) which is given by,

$$PMCC = \sum_{i=1}^{n}\sum_{l=1}^{r}\sum_{t=1}^{T}\left\{E(Z_l(\mathbf{s}_i,t)_{rep} - z_l(\mathbf{s}_i,t))^2 + Var(Z_l(\mathbf{s}_i,t)_{rep})\right\} \quad (6)$$

where, $Zl(s_i, t)rep$ denotes a future replicate of the data $zl(s_i, t)$. The first term in the PMCC assesses the goodness of fit and second term is a penalty term for model complexity. The PMCC, justified using a squared error loss function, is most suitable for comparing Bayesian hierarchical models that involve a first stage Gaussian model.

# RESULTS AND DISCUSSION

## Threshold selection and MODIS -derived metrics selection at regional scale

A first evaluation of the extraction of SOS metric utilizing full-year images sequence as input and varying thresholds ranging from 10 (0.1) to 50% (0.5) of the relative NDVI amplitude of each time-series profile, with a root mean-square error (RMSE) in the range of 23–53 days for the entire region (Fig. 3.4). However, the corresponding coefficient of determination ($R^2$) for the linear relationship between MODIS-derived SOS dates and CPCR dates was weak ($R^2 = 0.18$). The effect of snow in the winter signal generated biased, lack of SG smoothing fit, thus, low performance in the extraction of the phenological metrics. Using partial year images as input, RMSE values were in the range of 16-28 days. For SOS, the best performance was obtained using a threshold value of 20% for SOS detection, reporting a $R^2$ of 0.67, evidencing a RMSE 16 days and an average delay of 9 days between the expected and estimated values. The coefficient of determination increased with the threshold value (Fig. 3.5), however, MODIS-derived SOS values were displaced out from the 1:1 line in (Fig. 3.5).

For tasseling (VRT) detection, NASS CPCR weekly reports silking and/or tasseling depending on each regional NASS office. Thus, we assumed synchronized silking and tasseling when validating MODIS-derived metrics versus CPCRs. The detection of tasseling (VRT) is assumed to be the maximum observed NDVI value in the time-series profile and considered as the end of the green-up period. Tasseling reported a RMSE of 13.9 days with a delay bias of 10 days. When estimating EOS, the best outcomes were obtained when utilizing a threshold value of 35% of the relative amplitude of the time-series profile reporting a $R^2$ of 0.48, RMSE of 16.9 days and a delay of 4 days (Fig.4). Overall, SOS and EOS detection had reasonable agreement

with CPCR data, as documented by Wardlow et al., 2006 for the state of Kansas, and with Ren et al. 2017 at state level in the corn belt.

**MODIS-Derived EOS Metrics and ground-truth validation**

Ground truth validation was carried out between May-October in 2017. Overall, SOS, VRT and EOS metrics outperformed the regional validation. They reported lower RMSE, higher coefficient of determination and equivalent average delay. Overall, it is noted that lower RMSE and higher $R^2$ can be imputed to a different scale of control of geographical weight uncertainty. CPCRs summarize information at ASD scale, and certain geographic bias may occur, which differs from the direct ground-truth validation where field location and its corresponding time-series profile are directly matched and evaluated. Thus, this uncertainty is reduced when coupling local Kansas ground-truth data and time-series from known field locations. Differences on RMSE between Kansas and regional validation are mainly due to: 1) subjective interpretation and unbalanced area weight of weekly CPCRs. 2) Broader soil moisture and texture conditions at regional scale early in the season may affect the ratio between background and vegetation thus, the sensitivity of NDVI to consistently detect consecutive increases of NDVI assigned to plant emergence.

**Spatial-temporal patterns of the MODIS-based metrics and yield**

A threshold value of 20 and 35% of the relative amplitude of the time-series profile reported the most robust SOS and EOS metrics. The annual SOS, VRT, and EOS derivation were utilized to investigate the spatial patterns. From ground-truth and CPCR validation, annual mean over the 2003-2017 time period was further investigated in each county for all metrics.
The visual array for mean SOS and VRT follows an increasing pattern from the South to North region. Earlier SOS emergence in the SE region of Kansas, SE of Missouri and South of Illinois.

The mean EOS follows the same geographical pattern, but its distribution appears more compacted in the calendar. The length of vegetative, reproductive and full length is shorter in regions with high yield potential, Central and Northern of Iowa and Illinois, North East and South West regions of Kansas. Mean length of the season does not evidenced positive correlation with mean yield across the region.

The implementation of robust derivation of phenological metrics clearly depends on temporal and spatial resolution of the satellite. Other relevant aspect influencing the consistent estimation is the smoothing algorithm, we utilized Savitzy-Golatz since main priority was to preserve critical temporal features related to SOS, EOS and VRT. That later allows us to extract information and assign phenological metrics in both vegetative and reproductive periods. Using a reduced time window through the year significantly benefit the metrics extraction.

Threshold values used to detect SOS and EOS were also important (Zhao et al., 2004; Boryan et al., 2011; Tatem et al., 2017). For a large study area such as the Midwest US, we were searching for a threshold applicable to the entire region, rather than variable thresholds across the region. Compared to a 50% seasonal amplitude as the default threshold, a 20% value appeared to be better associated with CPCR 50% emerged dates at ASD level. We used daily MODIS time-series data to derive estimates and exploit the added value of very short-term temporal sensitivity to capture rapid changes on the canopy of the region. We also conducted additional analysis to examine how the use of different vegetation indices EVI and NDVI affects SOS detection. For each pair of NDVI-derived SOS and EVI-derived values, we conducted a correlation analysis and found high correlation $R^2$ =0.92. We further investigated the spatial and temporal relation between the derived satellite seasonal and phenological metrics and yield at two date frames. The length of the reproductive stage was significant retained in the model on describing yield trend

between 2003-2017. These results were in accordance with the one reported by (Sacks and Kucharik ,2011) for the region between 1981-2005. Most traditional modelling techniques (e.g. multiple linear regression) do not consider any spatio-temporal correlations, and hence are less appropriate to describe a data set that exhibits a spatial and temporal correlation structure, or random effects (Clark and Linzer 2015). In order to quantify the presence of spatial dependency in the data, we calculate the Moran's I statistic (Moran, 1950) through a permutation test. Moran's I statistic can be used only to test spatial autocorrelation at a single time point. Moran's I statistic was computed for all fourteen yield time points. All fourtheen time points indicated strong spatial autocorrelation for yield (Table 3.2). This motivates to include the spatial and temporal autocorrelations into the modelling hierarchy, which could capture and account for the spatial and temporal dependencies between the derived phenolical metrics and yield over the last 14 years in the region.

During HBSTM implementation, the best stability of variables parametrization was found at 7000 iterations, sampling autocorrelation was rapidly reduced. After Monte Carlo integration convergence of variables parameters were evaluated via density plots, trace and autocorrelation graphs (Fig. 3.7 and 3.8) [Hadfield, 2012; Martin et al., 2012). Results indicated significant parameter retention of length of the vegetative and reproductive stages, green-up and senescence rates (Table 3.3).

All four variables were significant (confidence interval of the coefficient outside 0 value) in the MCMC parametrization for both, GP HBSTM (Fig. 3.7) and AR HBSTM (Fig. 3.8). And, there were positively related to the increase on yield trend in the last 14 years. In GP HBSTM model, the spatial component was significant given by sigma2eta value (Table 3.3). This outcome confirms that the spatial structure between the phenological metrics and yield was

significant. But in addition, AR HBSTM confirmed that the spatial and temporal component (rho) were significant on describing yield. Moreover, when comparing the goodness of fit via PMCC, AR HBSTM outperforms GP HBSTM, meaning that the inclusion of the temporal component reduces the penalization (lower PMCC) of the model to describe yield. Overall, the outcomes indicate that a spatial but also a temporal trend that positively links the increase of the length of the vegetative, reproductive stages, faster green-up and senescence rates with yield is reported for the region in the last 14 years.

The significance of a higher green-up and senescence rate reported should be considered in the context of management practices. Higher plant population is increasingly utilized by farmers to maximize yield (Assefa et al., 2018). Higher plant density can lead to faster rate to reach full canopy cover/soil background prior to flowering and senescence speed rate before maturity. The lengthening outcomes are in accordance with last trends on yield improvement documented in the US corn belt region, an extended length of the vegetative and reproductive stages as one of the factors positively impacted yield gain (Sacks and Kucharik, 2011). The authors reported that a large driver of this change was a 14% increase in the number of GDD (Growing Degree Days) needed for corn to progress through the reproductive period, probably reflecting an adoption of longer season cultivars.

# CONCLUSION

We implemented a workflow using 250-m MODIS NDVI time-series to derive critical SOS, VRT, EOS temporal metrics. We compared our approach with both an extensive field campaign and with CPCR at 50% for each indicator evaluated. We apply a reduced window time to decrease winter signal noise ("denoising") and for facilitate the metric extraction procedure. We evaluate several threshold levels in order to find the best combination for each region of the time series profile. Best combination was threshold 20% and 35% or SOS and EOS, respectively. The outcomes unveil that lengthened duration of the vegetative and reproductive stages (full season), increase on green-up and senescence rates positively describe yield trend on both space and time for the region in the last 14 years.

# REFERENCES


Assefa, Y., Carter, P; Hinds, M., Bhalla, G., Schon, R., Jeschke, M., Paszkiewicz, S., Smith, S., Ciampitti, I.A. (2018). Analysis of long term study indicates both agronomic optimal plant density and increase maize yield per plant contributed to yield gain. Scientific Reports, 8, 4937, (2018).

Bakar K.S., Sahu, S. K (2013). spTimer: Spatio-Temporal Bayesian Modelling Using R. R package version 0.05, URL http://cran.r-project.org/web/packages/spTimer.

Banerjee, S., Carlin, B.P., Gelfand, A.E. (2004). Hierarchical modeling and analysis for spatial data. Chapman & Hall/CRC, Boca Raton.

Banerjee, S, Gelfand, A.E, Finley, A.O, Sang, H. (2008). Gaussian predictive process models for large spatial data sets. Journal of Royal Statistical Society B, 70, 825–848.

Bolton, D. K.; Friedl, M.A. Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics. Agric. For. Meteorol. 2013, 173, 74–84.

Boryan, C., Yang, Z., Mueller, R., Craig, M. Monitoring Us Agriculture: The US Department of Agriculture, National Agricultural Statistics Service, Cropland Data Layer Program. Geocarto Int. 2011, 26, 341–358.

Busetto, L., Ranghetti, L. MODIStsp: An R package for automatic preprocessing of MODIS Land Products time series, Computers & Geosciences. 2016, Volume 97, Pages 40-48.

Chen, J., Josson, P., Tamura, M., Gu, Z., Matsushita, B., and Eklundh, L.,"A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky - Golay filter", Remote Sensing of Environment, 2004, 91, (3-4), pp. 332-344.

Clark T.S., Linzer, D.A. 2015. Should I use fixed or random effects? Pol Sci Res Methods

3(2):399–408

Cressie, N.A.C. (1993). Statistics for spatial data. John Wiley & Sons, New York.

Cressie, N.A.C, Wikle, C.K. (2011). Statistics for spatio-temporal data. John Wiley & Sons, New York.

De Beurs, K.M.; Henebry, G.M. Spatio-temporal statistical methods for modelling land surface phenology. In Phenological Research; Springer: Dordrecht, The Netherlands, 2010; pp. 177–208.

Duvick, D.N. Possible genetic causes of increased variability in U.S. maize yields. In Variability in Grain Yields: Implications for Agricultural Research and Policy in Developing Countries; Anderson, J.R., Hazel, P.B.R., Eds.; Johns Hopkins University Press: Baltimore, MA, USA, 1989; pp. 147–156.

Estel, S., Kuemmerle, T., Alcántara, C., Levers, C., Prishchepov, A., Hostert, P. Mapping farmland abandonment and recultivation across Europe using MODIS NDVI time series. *Remote Sens. Environ.* 2015, *163*, 312–325.

Fang, H., Liang, S., Hoogenboom, G. Integration of MODIS Lai and Vegetation Index Products with the Csm–Ceres–Maize Model for Corn Yield Estimation. Int. J. Remote Sens. 2011, 32, 1039–1065.

Gelfand, A. E. 2012. Hierarchical modeling for spatial data problems. Spatial Statistics, 1, 30–39.

Gelfand A.E, Ghosh, S.K. 1998. Model Choice: A Minimum Posterior Predictive Loss Approach. Biometrika, 85, 1-11.

Gelman, A., Carlin J.B., Stern H.S., Rubin D.B. (2004). Bayesian Data Analysis. 2nd edition. Chapman & Hall/CRC, Boca Raton.

Hadfield, J (2012). MCMCglmm: MCMC Generalised Linear Mixed Models. R package version 2.17, URL http://cran.r-project.org/web/packages/MCMCglmm.

Jonsson, P., Eklundh, L. Seasonality Extraction by Function Fitting to Time-Series of Satellite Sensor Data. IEEE Trans. Geosci. Remote Sens. 2002, 40, 1824–1832.

Jönsson, P., Eklundh, L. Timesat—A Program for Analyzing Time-Series of Satellite Sensor Data. Comput. Geosci. 2004, 30, 833–845.

Kucharik., C.J. A Multidecadal Trend of Earlier Corn Planting in the Central USA. Agron. J. 2006, 98, 1544–1550.

Martin, A., Quinn K, Park, J (2012). MCMCpack: Markov chain Monte Carlo (MCMC) Package. R package version 1.2-4, URL http://cran.rproject.org/web/packages/MCMCpack.

Matern, B. (1986). Spatial Variation. Second edition. Springer-Verlag, Berlin.

Moulin, S.; Kergoat, L.; Viovy, N.; Dedieu, G. Global-Scale Assessment of Vegetation Phenology Using Noaa/Avhrr Satellite Measurements. J. Clim. 1997, 10, 1154–1170.

Moulin, S., Bondeau, A. and Delécolle, R. 1998. Combining agricultural crop models and satellite observations: from field to regional scale. *International Journal of Remote Sensing*, 19: 1021–1036.

Moran, P. 1950. A Test for the Serial Independence of Residuals. Biometrika, 37, 178-181

NASS, USDA. Agricultural Statistics Data Base. (https://quickstats.nass.usda.gov/api accessedon 8/23/2018).

R Development Core Team. R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna (2009).

Ren, J., Campbell, J.B., Shao, Y. Estimation of SOS and EOS for Midwestern US Corn and Soybean Crops. *Remote Sens.* 2017, *9*, 722.

Sacks, W., Kucharik, C. Crop management and phenology trends in the US corn belt: Impacts on yields, evapotranspiration and energy balance. *Agric. For. Meteorol.* 2011, *151*, 882–894.

Sahu, S.K., Bakar, K.S. (2012). "A Comparison of Bayesian Models for Daily Ozone Concentration Levels." Statistical Methodology, 9(1), 144–157.

Sakamoto, T., Wardlow, B.D., Gitelson, A.A., Verma, S.B., Suyker, A.E., Arkebauer, T.J. A Two-Step Filtering Approach for Detecting Maize and Soybean Phenology with Time-Series MODIS Data. Remote Sens. Environ. 2010, 114, 2146–2159.

Savitzky, A., and Golay, M.J.E., 1964. Smoothing and differentiation of data by simplified least squares procedures. Anal. Chem. 36, 1627-1639. Wentzell, P.D., and Brown, C.D., 2000. Signal processing in analytical chemistry. Encyclopedia of Analytical Chemistry, 9764-9800.

Shao, Y., Lunetta, R.S., Wheeler, B., Iiames, J.S., Campbell, J.B. An Evaluation of Time-Series Smoothing Algorithms for Land-Cover Classifications Using MODIS-NDVI Multi-Temporal Data. Remote Sens. Environ.2016, 174, 258–265.

Shen, Y., Liu, X. Phenological changes of corn and soybeans over US by Bayesian change-point model. Sustainability 2015, 7, 6781–6803.

Stevens, A., Ramirez-Lopez, L. 2013. An introduction to the prospectr package. https://cran.r-project.org/web/packages/prospectr/vignettes/prospectr-intro.pdf

Tao, F., Yokozawa, M., Xu, Y., Hayashi, Y., Zhang, Z. Climate Changes and Trends in Phenology and Yields of Field Crops in China, 1981–2000. Agric. For. Meteorol. 2006, 138, 82–92.

Tatem, A.J, Thomson, D.R, Stevens, F.R, Castro, M.C, Ruktanonchai, N.W. (2017) GridSample: an R package to generate household survey primary sampling units (PSUs) from gridded

population data. Int. J. Health Geogr. 16:25.

USDA-NASS. Crop Production 2017 Summary. January 2018.

Vermote, E., Wolfe, R. (2015). MOD09GA MODIS/Terra Surface Reflectance Daily L2G Global 1kmand 500m SIN Grid V006 [Data set]. NASA EOSDIS LP DAAC. doi: 10.5067/MODIS/MOD09GA.006

Wardlow, B.D., Kastens, J.H., Egbert, S.L. Using USDA Crop Progress Data for the Evaluation of Greenup Onset Date Calculated from MODIS 250-Meter Data. Photogramm. Eng. Remote Sens. 2006, 72, 1225–1234.

Yusoff, N.M., Muharam, F.M., Takeuchi, W., Darmawan, S., Razak, M.H.A. Phenology and classification of abandoned agricultural land based on ALOS-1 and 2 PALSAR multi-temporal measurements Int. J. Digit. Earth, 10 (2017), pp. 155-174.

Zhang, X., Friedl, M.A., Schaaf, C.B., Strahler, A.H., Hodges, J.C., Gao, F., Reed, B.C., Huete, A. Monitoring Vegetation Phenology Using MODIS. Remote Sens. Environ. 2003, 84, 471–475.

Zhang, X., Friedl, M.A., Schaaf, C.B., Strahler, A.H., 2004. Climate controls on vegetation phenological patterns in northern mid- and high latitudes inferred from MODIS data. Glob. Change Bio. 10 (7), 1133–1145.

Zhao, H., Yang, Z., Di, L., Pei, Z., 2012. Evaluation of temporal resolution effect in remote sensing-based crop phenology detection studies. Computer and Computing Technologies in Agriculture V the series IFIP Advances in Information and Communication Technology. 369, pp. 135–150.

Zheng, J., Ge, Q., Hao, Z., 2002. Impacts of climate warming on plants phenophases in China for the last 40 years. Chin. Sci. Bull. 47 (21), 1826–1831.

Zhong, L., L. Hu, L. Yu, P. Gong, and G. S. Biging. 2016. Automated Mapping of Soybean and Corn Using Phenology. *Isprs Journal of Photogrammetry & Remote Sensing* 119: 151–164. doi:10.1016/j.isprsjprs.2016.05.014.

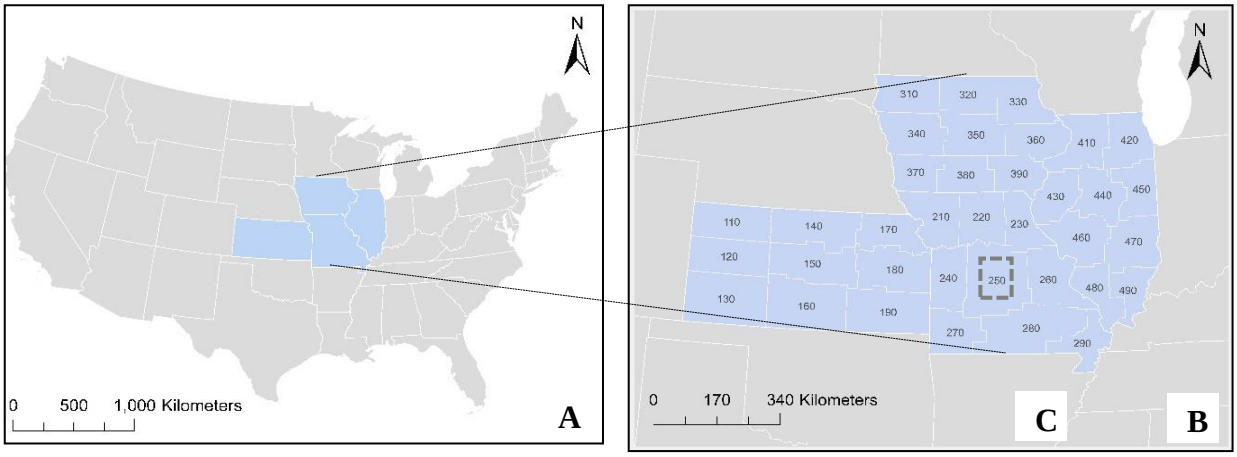


**Figure 3.1. (A): Study area: states of Iowa, Illinois, Missouri and Kansas. (B): Extent, location and (C): USDA-NASS ASDs (Agriculture Districts) code in the area of study.**

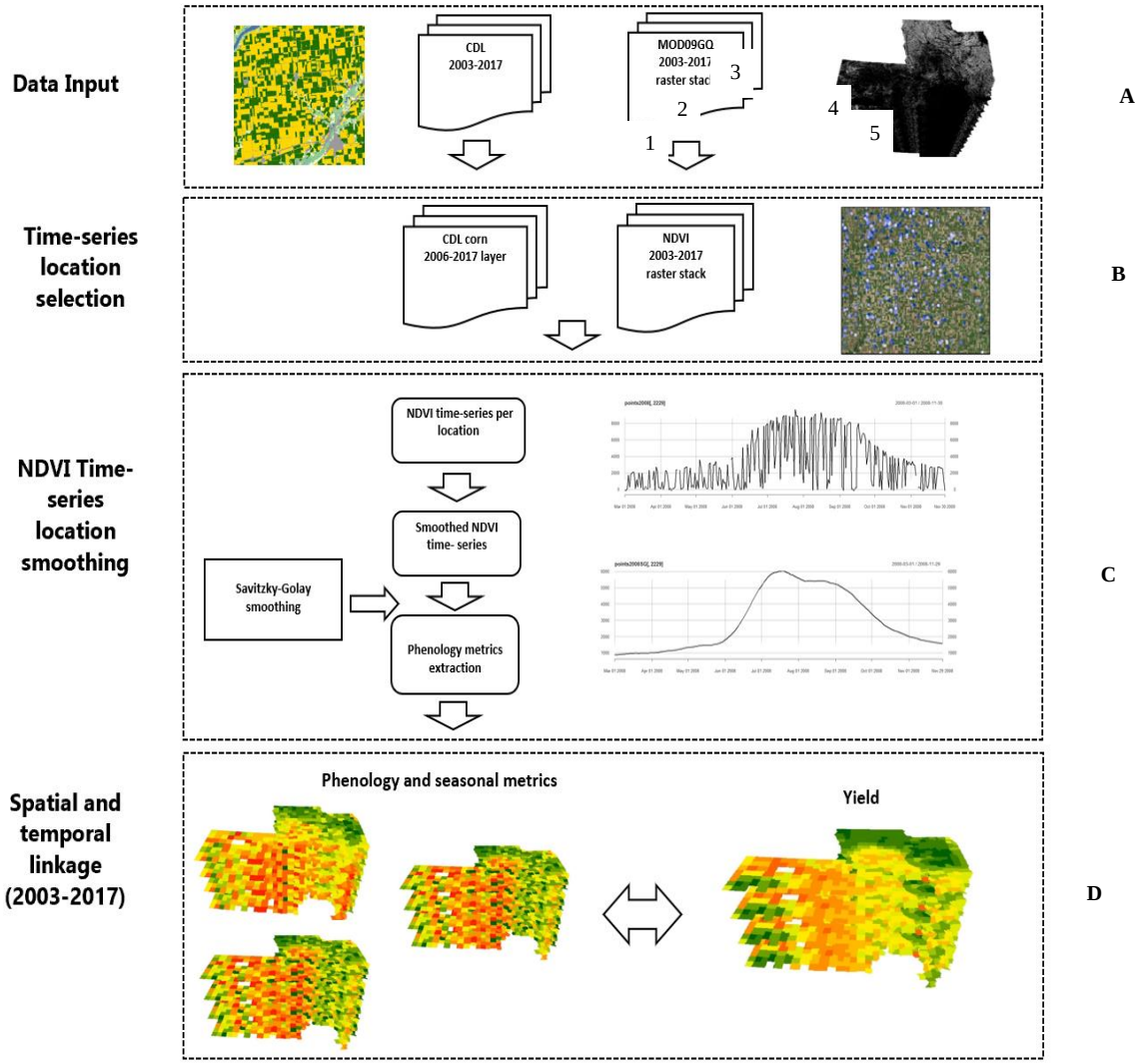


**Figure 3.2. Workflow process for phenological metrics extraction. (A) Data downloaded, MOD09GQ rasterstacks and CDL corn layer extraction. (B) Generation of MOD09 NDVI rasterstack, clipping of CDL corn and MOD09 NDVI pixels. (C) Extraction of NDVI time-series profiles, using selected geo-locations on MOD09 NDVI rasterstack. (D) Spatio-temporal model linkage. Down: (C.1) = start of season, (C.2) = green-up rate, (C.3) = tasseling (maximum NDVI value), (C.4) = browndown rate (senescence), (C.5) = end of season.**

**Figure 3.3. Phenology transition during 2017 field growing season, left panel: start of the season (SOS), center panel: tasseling (VRT), right panel: end of the season (EOS).**

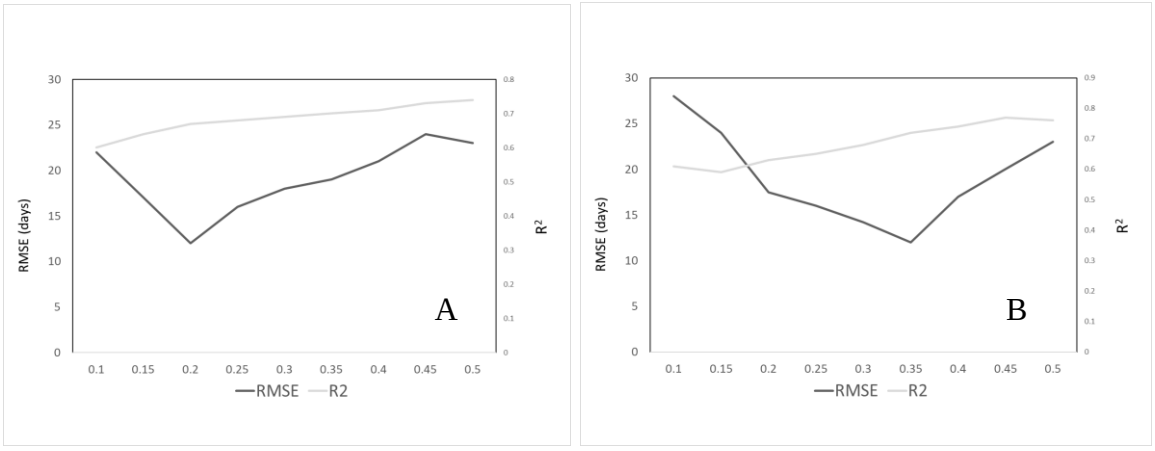


**Figure 3.4. Impact of threshold value on RMSE and coefficient of determination on: (A) start of season (SOS) and (B) end of season.**

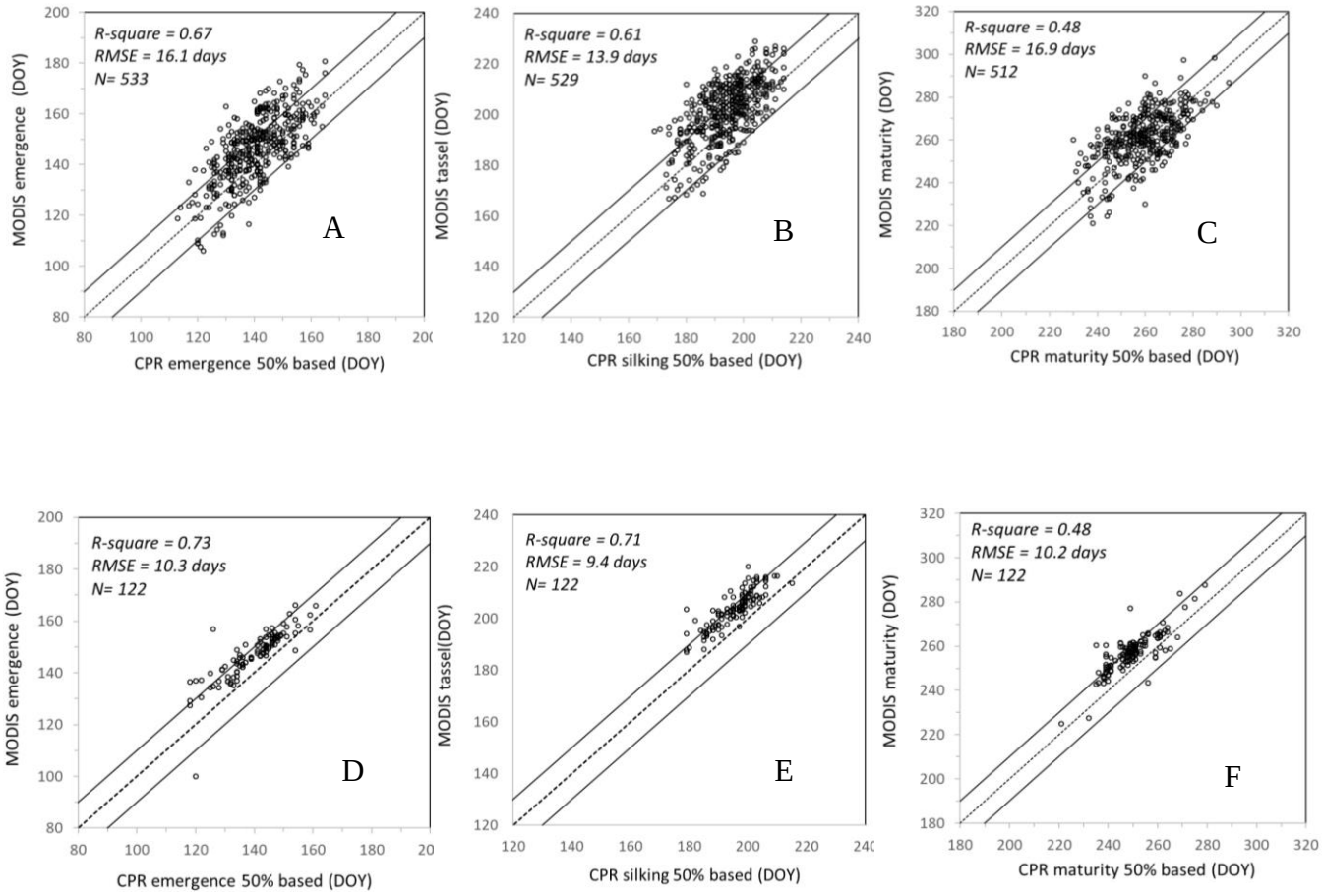


**Figure 3.5. Top panel: MODIS derived metrics versus CPCRs in the region of study: (A) start of the season (SOS), (B) vegetative-reproductive transition (VRT), (C) end of the season (EOS) across the states (Indiana, Kansas, Missouri, and Illinois) data between 2006 and 2017. Bottom panel: MODIS derived metrics versus ground-truth data from field surveys in the state of Kansas during the 2017 growing season: (D) SOS, (E) VRT, (F) EOS all relative to MODIS-derived phenology metrics.**

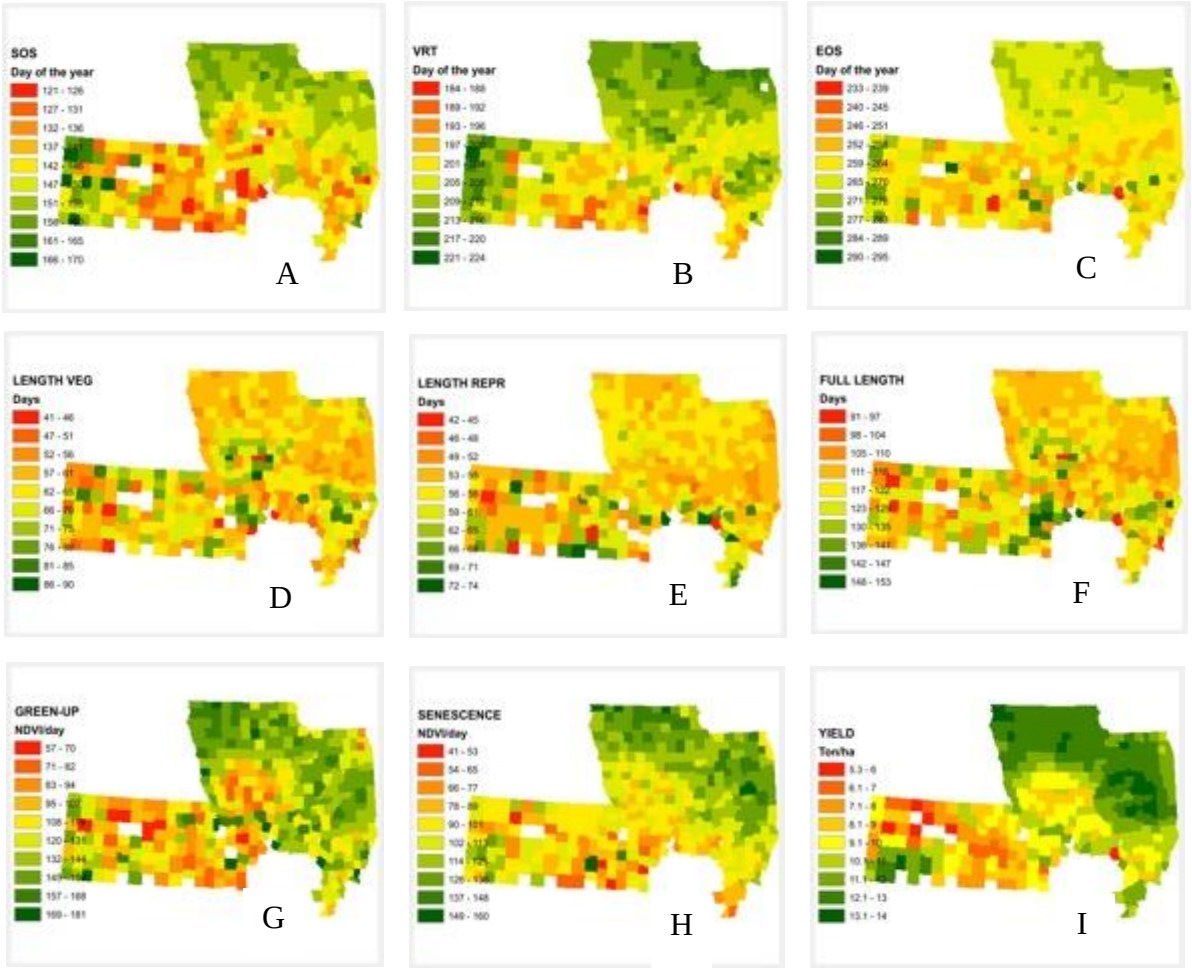


**Figure 3.6. Maps of derived phenology metrics means (2003-2017). (A) start of the season (SOS), (B) vegetative-reproductive transition (VRT), (C) end of the season (EOS), (D) Length of vegetative stage, (E) Length of reproductive stage and (F) Length of season, (G) green-up, (H) senescence, (I) yield.**

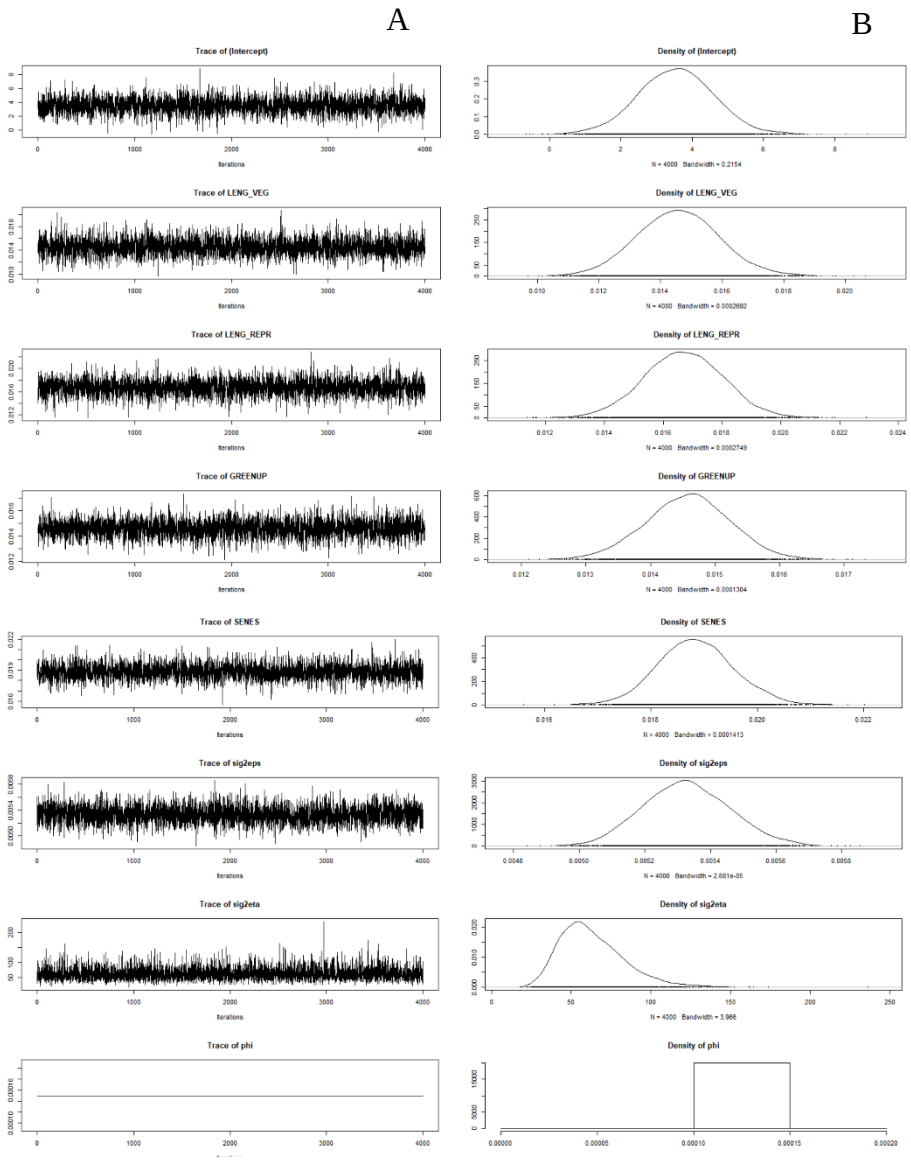


**Figure 3.7. (A) Left panel: Trace evaluation of variables in GP HBSTM parametrization and (B) Monte Carlo integration, variables parameter distribution, confident interval approximation for length of vegetative and reproductive stages, green up, senescence and parameters of the model (sig2eps, sig2eta, phi).**

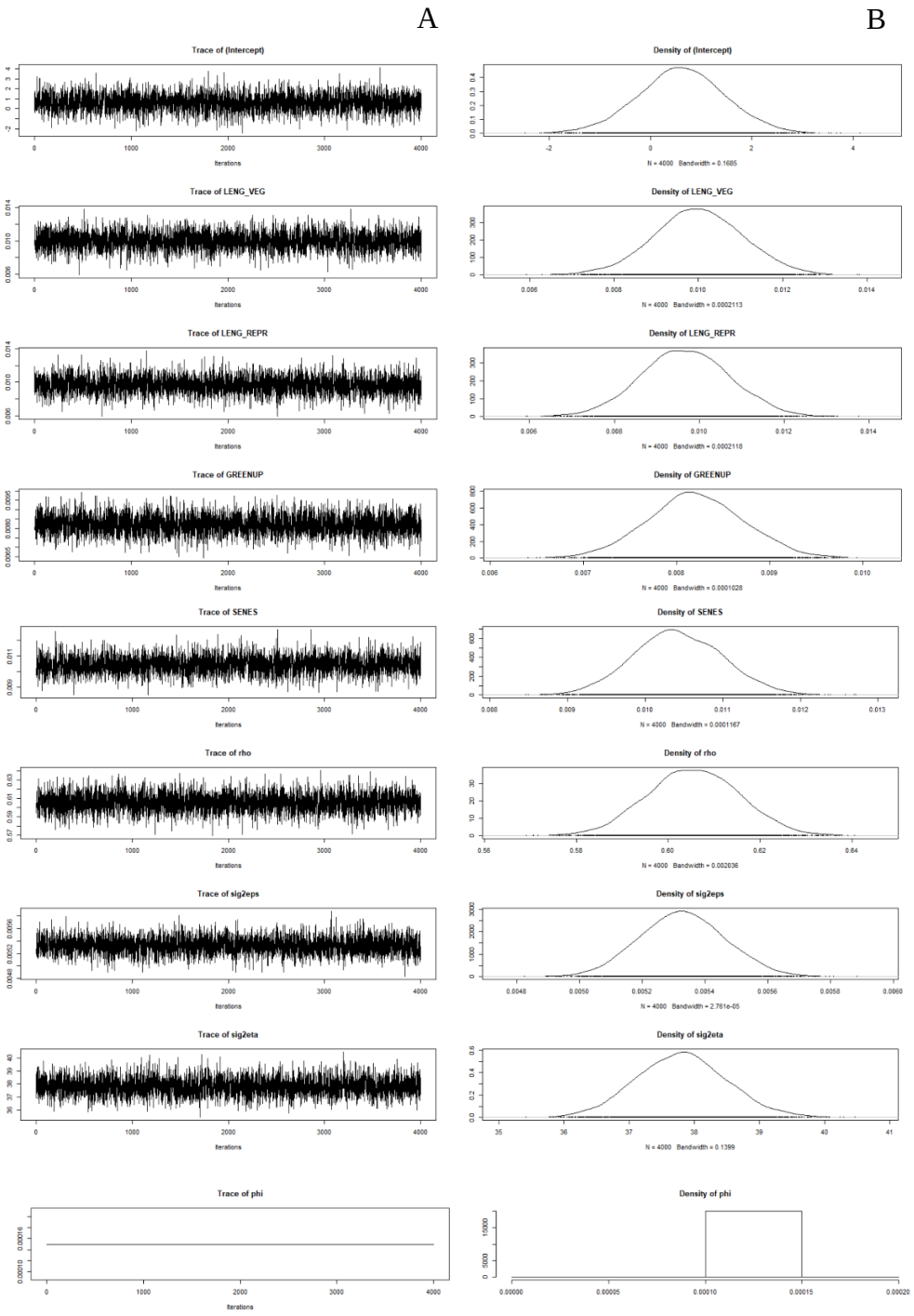


**Figure 3.8. a=Left panel: Trace evaluation of variables in AR HBSTM parametrization and b= Monte Carlo integration, variables parameter distribution, confident interval approximation for length of vegetative and reproductive stages, green up, senescence and parameters of the model (rho, sig2eps, sig2eta, phi).**

**Table 3.1. Phenology metrics defined and extracted from time-series profile description.**

| **Phenology metrics** | **Description** |
|---|---|
| Start of season (SOS) | Beginning of growing season, time which NDVI start to increase, 20% of time-series relative amplitude |
| Green-up rate | Calculated as the difference between the NDVI between 30-and-70 % of the time-series relative amplitude over time |
| Vegetative -Reproductive transition (VRT) value | Largest NDVI value for the fitted function. Approximation for tasseling in corn |
| Vegetative -Reproductive transition (VRT) time | Time during the growing season when the maximum NDVI peak occurs |
| End of season (EOS) | End of growing season, time which NDVI ends to decrease, 35% of time-series relative amplitude after maximum peak. |
| Vegetative length | Time from the start of the growing season to tasseling |
| Reproductive length | Time from tasseling to physiological maturity |
| Seasonal Length | Time from the start to the end of the growing season |
| Browndown rate | Calculated as the difference between the NDVI between 30-and-70% of the senescence process (decay phase of NDVI seasonal curve) over time |
| Amplitude | Difference between maximum and base NDVI level |

**Table 3.2. Spatial autocorrelation of USDA-NASS yield via Moran I test at the county level for the region.**

| Metric | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Moran I | 0.62 | 0.58 | 0.63 | 0.49 | 0.50 | 0.57 | 0.66 | 0.85 | 0.91 | 0.44 | 0.83 | 0.83 | 0.76 | 0.90 | 0.82 |

- Expectation = $-2.4e^{-3}$, Variance = $8.0e^{-4}$, p-value = $2.2e^{-16}$

**Table 3.3. Inference of the GP and AR models parameters (i.e. median and statistical significance from Markov chain Monte Carlo samples) in the region between 2003-2017.**

| Phenological metrics: | GP Model | AR Model |
|---|---|---|
| Length of Vegetative | 1.460E-02* | 1.00E-02* |
| Length of Reproductive | 1.670E-02* | 9.70E-03* |
| Green-up rate | 1.460E-02* | 8.20E-03* |
| Senescence rate | 1.880E-02* | 1.04E-02* |
| Other parameters in the models: | | |
| Intercept | 3.55E+00* | 6.16E-01 |
| Spatial white noise (*sig2eps*) | 5.00E-03* | 5.00E-03* |
| Spatial correlation (*sig2eta*) | 6.00E+01* | 3.78E+01* |
| Spatial decay (*phi*) | 1.00E-04* | 1.00E-04* |
| Temporal correlation (*rho*) | -- | 6.06E-01* |
| PMCC | 3.00E+05 | 1.49E+05 |

*****Statistically significant considering 95% Bayesian credible interval.

# Chapter 4 - General discussion

## Conclusions and implications for agriculture

A strong positive correlation was obtained between CSM-estimated plant height and ground-truth data collected when corn plants were at flowering stage ($R^2$ = 0.79, RMSE =0.09 m). The correlation between CSM and ground-truth data measured two weeks prior to flowering was penalized ($R^2$ = 0.63, RMSE = 0.11 m). For the pre-flowering measurement, the lower proportion of the variation accounted for the CSM-estimated plant height was primarily due to lack of uniform development within the corn canopy and plants emerging at different timing due to soil–weather factors in the season. At flowering, maximum plant height was attained, corn canopy become more uniform with less heterogeneity (lower RMSE to mean plant height ratio) and better prediction power (higher $R^2$). The correlation obtained between measured and CSM-estimated plant height is consistent with previous findings for corn and other crops. A significant correlation between plant height measurements at flowering stage support the conclusion that imagery taken at end of stem elongation is better correlated with ground-truth data. Plant height estimate reported a stable response through heights on both dates, slopes of the linear regression slopes between dates were no significant different. We also report that there was a better prediction of plant height at flowering due to a lower underestimation, which is also related to lower plant heterogeneity within the canopy. Spatial-temporal correlation between CSM-estimate and biomass suggested that digital photogrammetry can be implemented to assist plant biomass estimation. But plant biomass predictability is not strong; however, predictability significantly increases when joining CSM-derived plant height and stem diameter measured on the ground.

In Chapter 2, the developed workflow integrates traditional color thresholding and supervised learning to detect corn plant in the images with the objective of proposing a solution to move from a manual to an automatized and unbiased process. Real farm field conditions were utilized for collecting data, implementation and validation steps. Even tough combined sites training performed slightly weaker (0.93), the implementation of a prior-trained classifier in a new site evidenced potential with the benefit of reducing the computational cost in large datasets. Overall performance significantly dependens on the flight altitude/spatial resolution, best outcomes were reached in the lower flight altitude (10 m) imagery. Degraded spatial resolutions gradually penalizes the outcomes due to the lower sensitivity to detect green contours (mixed signal between vegetation and soil background) and by depleting the significance of "geometry" as a meaningful descriptor to differentiate between green contours classes.

In chapter 3, a scalable approach was proposed to first screens key phenological metrics of corn via satellite data, in a second step these metrics were utilized to understand the spatial and temporal dynamics of the crop phenology and yield over an extensive region for 14 years. Outcomes indicate that lengthened vegetative and reproductive stages, faster green-up and senescence rates positively describes yield. In addition, the positive linkage between satellite-based phenology metrics and yield trends were significant in space but also in time for the region.

## Contribution to science

The present research contributes to the knowledge on the adoption of remote sensing in the context of agriculture enabling scalable inference over biological processes that in general evidence site, space and time dependencies. The utilization of digital aerial photogrammetry is still novel in agriculture, and literature on the adoption of this approach for crop metrics derivation is scarce. The outcomes demonstrate that the technique can be a valid low-cost alternative to derive allometric plant metrics in the context of precision agriculture and plant breeding.

The second research proposes an alternative to shift from a traditional manual biased operation into an automatized procedure to derive early season crop performance. But, strictly in the arc of science, it attempts to leverage previously reported color threshold approaches highly dependent on local site conditions, prone to miss-classification and limited on scalability. In this context, the proposed method relies in a novel integration of the UAS, color thresholding, and supervised learning. The proposed workflow reported less local site-dependency and higher performance than previous works reported in the literature. The main contribution of this paper is related to the development of a procedure to detect corn plants to better guide early season operations for farmers. The outcome of the workflow allows the digital counting of plants using a low-cost UAS and RGB camera contributing to quantify early-season data of crop performance at on-farm conditions.

For the third research, even though vegetation phenology via satellite imagery had been widely studied globally, not much attention had been focused to better understand the links between the spatial and temporal dynamics of crops phenology and yield. To the best of our knowledge, this is the first attempt to understand these dynamics over an extensive region in the

last 14 years. The outcomes allow to identify the most significant phenology indicators describing yield trend. The results can help to better understand crop adaptation in a context of a large-scale interpretation. Identify the links between phenology and yield across latitudes, varied weather conditions, across regions with contrasting yield potentials enables better screening of crop adaptation and decision-making processes of farmers.

### Future research

New questions arose when addressing the objectives of each chapter. In the case of the adoption of digital photogrammetry, future work should account for the impact of early season Digital Terrain Models on plant height estimate. This will likely evidence the prior-site magnitude dependency of the technique that will help to better understand the trade-off between precision of the estimate and the cost of its operations. Another line of work should investigate the viability of digital photogrammetry on deriving "canopy" metrics linked to yield prediction in corn. In this context, new insights on the relation between volume estimates and actual biomass are needed. Lastly, but not less relevant is the study of potential integration between digital photogrammetry. multi and hyperspectral domains, but also the redundancy of the approaches should be further studied and identified. For the second research, future work should investigate the integration of geometry, spectral and texture descriptors for classes delineation on leveraging overall performance. To explore opportunities of integrating multiclass non-corn objects to reduce the internal high variance of non-corn objects class boosting overall performance. To investigate the penalization of changing environment conditions (wind, clouds) in the geometric descriptors and classifier performance requires further insights.

In the third study, further insights are needed to better understand the relevance of the length of stay-green post-flowering stage. This is not considered in the current design on metrics

extraction, since the  estimation includes the full length of the reproductive stage. However, it is a significant aspect from a physiology stand point, since this metric may represent a valuable contribution of selection in modern hybrids to leverage yield. In addition to that, there is a clear need to integrate satellite phenology estimates into predictive models, our outcomes evidenced this potential. There is a synergistically opportunity to integrate the high spatial resolution and large area coverage of satellites with the precision of local weather station estimates based on temperature to improve phenology predictability.