



2021 G2F Maize GxE Project Collaborator's Meeting North American Plant Phenotyping Network (NAPPN) Conference

Date: Tuesday, February 16th, 2021

Time: 8:30AM to 11:50AM

Connection Information: Forthcoming

Agenda

8:30 – 8:50AM - *Welcome and G2F maize GXE project update* – Natalia de Leon, University of Wisconsin, Madison

8:50 – 9:10AM – *Predicting the grain yield using temporal vegetation indices and plant heights via machine learning regressions in maize* – Alper Adak, Texas A&M University

9:10 – 9:30AM – *UAV-based estimation of core traits in Genomes to Field initiative* - Seth A. Tolley, Purdue University

9:30 – 9:50AM – *Detecting genetic and local environmental effects using high-throughput aerial imager* – Nicolas Morales, Cornell University

9:50 – 10:00AM - **Break**

10:00 – 10:20AM – *From Genomes to Fields: Exploring genotype-by-environment interactions and environment specific prediction in maize hybrids*- Anna Rogers, North Carolina State University

10:20 – 10:40AM – *Life isn't always trade-offs: The potential benefits and costs of resistance to northern corn leaf blight* - Danilo E. Moreta, Cornell University

10:40 – 11:00AM – *Structural root architecture modeling by digital twin from X-ray computed tomography* - Monica Herrero Huerta, Purdue University

11:00 – 11:20AM – *Analyzing maize leaf angles and modeling leaf curvature* - Zhongjie Ji and Erik Amezquita, Michigan State University

11:20 – 11:50AM - *2020 G2F Maize GXE drone project* - Edgar Spalding and Guosheng Wu

11:50AM - Adjourn

Title: Predicting the Grain Yield Using Temporal Vegetation Indices and Plant Heights via Machine Learning Regressions in Maize

Alper Adak¹, Seth C Murray^{1*}, Sofija Božinović², Regan Lindsey¹, Shakirah Nakasagga¹, Sumantra Chatterjee¹, Scott Wilde¹

¹ Department of Soil and Crop Sciences, Texas A&M University, College Station, TX

² Maize Research Institute, Zemun Polje, Belgrade, Serbia

ABSTRACT:

Field based high throughput phenotyping (HTP) provides high quality images for plant breeding nurseries where many genetic materials (eg. hybrids and inbreds) are repeatedly screened throughout the growth season at a low cost. In this study, the same set of 100 advanced hybrids were planted as optimal (IHOT trial) and delayed planting (DHOT trial) in randomized complete block designs with two replications, each containing two consecutive row plots for each maize hybrid in College Station, Texas. An unoccupied aerial vehicle (UAVs, aka drones) was flown twelve times over these two trials, each containing 400 row plots. Plot based grain yield, days to anthesis (DTA) and silking (DTS) data were collected for each trial to be predicted by 15 different vegetative indices (VIs) and canopy height measurements (CHMs) extracted from geometrically corrected aerial images (tif file and orthomosaics). BGI (Blue Green Pigment Index), BI (Brightness Index), blue index, EXG (excessive green), EXGR (excess green minus excess red index), GLI (green leaf index), Green index, MGVR (modified green red vegetation Index), MPRI (modified photochemical reflectance index), NGBDI (normalized green red difference index), NGRDI (normalized green red difference index), Red index, RGBVI (red green blue vegetation index), VARI (visible atmospherically resistant index), VDVI (visible-band difference vegetation index) and CHMs were generated for twelve time points for each trial. Temporal breeding values of 15 VIs and CHMs estimated from twelve time points for each maize hybrid were used as predictors in cross validation to predict grain yield and flowering times (DTA and DTS). Results showed that machine learning based regressions predicted the all dependent variables better in both trials since they have higher Rsquared and lower root mean squared error values. When machine learning regressions are compared, ridge regression was found to perform best in predicting the grain yield in both trials with between 0.81 and 0.84 Rsquared values while lasso and elastic net regressions were found to perform best in prediction the DTA and DTS with between 0.89 to 0.95 depending on the trials. Variable importance of predictors in the prediction models were found to have descending tendency towards the terminal stages of growth.



Title: UAV-based Estimation of Core Traits in Genomes to Field Initiative

Seth Tolley¹; Monica Herrero Huerta²; Mitchell Tuinstra¹; Yang Yang²

¹ Department of Agronomy, Purdue University, West Lafayette, IN, United States of America

² Institute for Plant Sciences, Purdue University, West Lafayette, IN, United States of America

ABSTRACT:

The genomes to fields (G2F) initiative is a public-private collaboration designed to integrate genomics and phenomics to support growers, consumers, and society. Since the adoption of next-generation sequencing in genomics, phenomics has been a new bottleneck that limits greater genetic gain. As a historically costly, time-consuming, and labor-intensive endeavor, the advent of high-throughput phenotyping (HTP) methods improves upon the sustainability of plant phenomics. Specifically, field-based HTP platforms, such as unmanned aerial vehicles (UAV), provide an opportunity to measure traits throughout the growing season to better understand plant growth and development and model genetic, environmental, and genotype-by-environment interaction factors. Typical traits manually measured in the G2F protocol include core traits centered around plant morphology, ear morphology, agronomic traits, and yield-related traits. The objective of this work was to describe the performance of various sensors to estimate or predict stand count, vegetative indices, and plant height. A GRYFN UAV platform was flown weekly throughout the growing season over the G2F experiment in West Lafayette, IN from May 21st to September 21st in 2020. In total, 19 flights were captured with Red-Green-Blue (RGB), multispectral, and LiDAR sensors. Stand count was measured on a per plot basis using the RGB camera and Phenix software. Various indices were computed using the RGB and multispectral data. LiDAR estimated plant height from post-anthesis was correlated to ($r=0.93$) and more heritable than reference terminal plant height. The implementation of UAV-based sensors could uncover additional relationships between genomics and phenomics, genotype-by-environment interactions, and genotype-by-genotype interactions.

Title: Detecting genetic and local environmental effects using high-throughput aerial imagery

Morales Nicolas^{1*}, Mueller Lukas A.², Robbins Kelly R.¹

¹ Plant Breeding and Genetics, School of Integrative Plant Science, Cornell University, Ithaca, NY 14853

² Boyce Thompson Institute, Ithaca, NY 14853

ABSTRACT:

High-throughput aerial imagery from unoccupied aerial vehicles (UAVs) allows measurement of large numbers of genotypes in the field over the course of the growing season. Local environmental effects within a field experiment, such as heterogeneous soil conditions, elevation differences, and soil curvature variations, confound the genetic effects existing in the measured phenotypes. Linear mixed models can partition genetic effects from spatial and temporal environmental effects. In this work, spatial effects are captured using two-dimensional splines, whereas temporal effects are captured as permanent environment solutions of random regressions. From six aerial imaging events on a 2019 Genomes-to-Fields (G2F) maize experiment, consisting of 612 hybrids planted over 800 experimental plots, vegetation indices were computed for all plots across the growing season. A UAV mounted MicaSense multi-spectral camera captured images of the field and ImageBreed (<https://imagebreed.org>) extracted the plot-level high-throughput phenotypes. Genotyping-by-sequencing (GBS) data produced genomic relationship matrices (GRMs) for the tested hybrids. Genetic effects were then partitioned from spatial and temporal environmental effects using a multivariate two-dimensional spline (2Dspline) model in R Sommer and a random regression (RR) model in REMLF90, respectively. Both the true genetic and environmental effects are unknown to us; therefore, simulation of five classes of environmental effects, both constant and changing in time, were tested within the 2Dspline and RR models. The RR model was most robust to all simulated environmental effects. To elucidate the true origins of the computed local environment effects, soil electrical conductance and elevation was collected prior to the start of the field season. Modeling of the soil and available climatic variables, such as precipitation and growing degree days (GDD), indicate the computed RR model environmental effects can be most easily explained. These findings demonstrate that repeated measurements through time may better capture the true environmental effects than two-dimensional splines.



Title: From Genomes to Fields: Exploring genotype-by-environment interactions and environment specific prediction in maize hybrids

Anna R. Rogers and James B. Holland

USDA-ARS and North Carolina State University

ABSTRACT:

Plant breeding programs are often faced with challenges in making initial selections among breeding materials based on evaluation in a single environment, with the ultimate goal of creating new varieties that will later be planted across multiple, more diverse conditions. In some cases, genotypes that initially seemed very promising are observed to vary dramatically for important agronomic traits across diverse environments. Genotype-by-Environment interactions (GxE) underlie relative differences in performance across environments but are difficult to predict without understanding how genotypes respond to specific environmental covariates. Recent advances in genomics and prediction modeling have accelerated the ability to perform selections using genomic data, but little has been done to incorporate environmental data into such modeling. Including environmental variables in GxE analysis often results in issues with multicollinearity, caused by presence of large numbers of predictors that are often highly correlated, each of which only explains a small amount of variance. Development of methods to incorporate both genomic and environmental data into genomic prediction models should provide ability to predict environment specific performance of individual genotypes. Using publicly available data for 1,919 maize hybrids spread across multiple locations over three years in North America, we explore GxE modeling using a mixed models approach incorporating high density DNA marker data and weather covariates. Using these data, we gain a clearer insight of what GxE means in context of plant development and response to fluctuating environmental conditions, and explore the possibility of predicting hybrid phenotypes in previously untested environments.

Title: Dissecting the benefits and costs of resistance to northern leaf blight in maize hybrids

Moreta, Danilo E. *, Kolkman, Judith M., and Nelson, R.

Plant Breeding & Genetics Section, School of Integrative Plant Science, Cornell University

ABSTRACT:

Breeding for quantitative disease resistance is an effective and sustainable disease management strategy to protect crops from pathogen attack. Genetic resistance, however, can be associated with trade-offs such as yield costs. As part of the Genomes to Fields initiative, we evaluated a population of maize hybrids to measure potential yield effects and trade-offs associated with resistance to northern leaf blight (NLB). Grain yield of hybrids in an uninoculated trial was regressed against NLB resistance values of the artificially inoculated trial in 2014. No costs of NLB resistance were observed at the yield level in the hybrid population field-tested in 2014 as indicated by a positive regression slope ($r = 0.23^*$, $p\text{-value} \leq 0.05$, $R^2 = 0.05$) between the inoculated and uninoculated trial. A preliminary genome-wide association study was conducted to identify genetic variants associated with NLB resistance and agronomically important related traits using genotyping-by-sequencing and phenotypic data from the field season 2014. Significant SNP associations were identified for stalk lodging, root lodging, and grain moisture under NLB pressure, but not for NLB resistance. Some of these initial findings confirm previous reports where selection for NLB quantitative resistance also improved stalk lodging resistance without affecting grain yield potential when measured under NLB-free conditions. A naturally occurring infestation of anthracnose leaf blight (ALB) in the NLB inoculated trial in 2018 provided the opportunity to determine putative disease resistance trade-offs. An inverse relationship ($r = -0.45^{***}$, $p\text{-value} \leq 0.05$, $R^2 = 0.21$) was observed between NLB and ALB, suggesting a putative resistance-susceptibility trade-off. Future efforts will focus on analyzing overlapping hybrids across years (2014-2019) and on using aerial imaging as a cost-and-time effective and more accurate way of screening field trials in the context of NLB resistance and trade-offs.

Title: Structural root architecture modeling by digital twins from X-ray computed tomography

Herrero-Huerta Monica ^{1*}, Meline Valerian², Hoagland Chris¹, Magalhanes Augusto¹, Yang Yang¹

¹ Institute for Plant Sciences, College of Agriculture, Purdue University, West Lafayette, IN, USA

² Department of Botany and Plant Pathology, Purdue University, West Lafayette, IN, USA

ABSTRACT:

Breakthrough imaging technologies are a potential solution to challenge the plant phenotyping bottleneck regarding marker-assisted breeding and genetic mapping. X-Ray CT (computed tomography) technology is able to acquire the digital twin of root system architecture (RSA). However, advances in computational methods to digitally model and parameterize the shape and spatial disposition of root system networks as well as analyzing the growth over time are urgently required.

We develop a structural root architectural modeling based on 3D data from X-ray CT. This novel approach is optimized for high-throughput, repeatable, and robust. Significant root architectural traits as number, length, growth angle, elongation rate, diameter, branching map, and even volume of axial and lateral roots are accurately extracted. Our pipeline is divided in two major steps: (i) first, we compute a curve-skeleton based on a constrained Laplacian smoothing algorithm. This skeletal structure determines the correspondences of the ramifications over time; (ii) after that, the root system is robustly modeled by a cylindrical fitting. The experiment was carried out at the Controlled Environment Phenotyping Facility (CEPF) from Purdue University in West Lafayette (IN, USA). Three samples of tomato root at 2 different times and three samples of corn root at 3 different times were scanned. Regarding the first step, the skeleton is able to accurately match the root ramifications over time based on a visual inspection. From the second step, the volume from the cylinder model were compared against the root digital twin, reaching a coefficient of determination (R^2) of 0.91.

The results confirm the feasibility of the proposed methodology, providing scalability to a comprehensive analysis to high throughput root phenotyping.

Keywords: *phenotyping, 3D modeling, skeleton, root system architecture (RSA), X-ray CT (computed tomography), digital twin.*



Title: Analyzing maize leaf angles and modeling leaf curvature

Ji, Zhongjie^{1*}, Amézquita, Erik^{2*}, Chitwood, Dan^{2,3}, Thompson, Addie¹

¹ Department of Plant, Soil & Microbial Sciences, Michigan State University, East Lansing, MI 48824

² Department of Computational Math, Science and Engineering, Michigan State University, East Lansing, MI 48824

³ Department of Horticulture, Michigan State University, East Lansing, MI 48824

ABSTRACT:

Leaf angle characteristics and their optimization are important considerations for crops as they directly impact photosynthesis and enable increased planting density. Due to the complexity of the canopy, the full arch of a leaf is difficult to measure in the field, so maize breeders and researchers have focused primarily on the angle of incidence between the base of the leaf and the stem as opposed to the entire leaf. In the summer of 2019, after flowering, we used a polar coordinate method to measure one leaf above and below the ear leaf in the Genomes2Fields hybrid field. For each leaf, we measured the angle at the stem, as well as the length and angle to both the apex and tip of the leaves. These five values for three spatial points provide a more comprehensive description than previous measurements for stem angle only. A dual polynomial-based model of leaf curve based on these points gives a better understanding of leaf angle as it occurs in the field. These descriptors will enable further modeling of leaf angles and canopy architecture to describe potential for light capture, and along with leaf number, could predict LAI and canopy cover. The angle descriptors not only reflect the physiology of the leaves and canopy but also reveal the genetic variance of leaf angles, and several associated SNPs have been discovered that influence leaf angle and curvature.