

Use of Crop Growth Models with Whole-Genome Prediction: Application to a Maize Multienvironment Trial

Mark Cooper,* Frank Technow, Carlos Messina, Carla Gho, and L. Radu Totir

ABSTRACT

High throughput genotyping, phenotyping, and envirotyping applied within plant breeding multi-environment trials (METs) provide the data foundations for selection and tackling genotype \times environment interactions (GEIs) through whole-genome prediction (WGP). Crop growth models (CGM) can be used to enable predictions for yield and other traits for different genotypes and environments within a MET if genetic variation for the influential traits and their responses to environmental variation can be incorporated into the CGM framework. Furthermore, such CGMs can be integrated with WGP to enable whole-genome prediction with crop growth models (CGM-WGP) through use of computational methods such as approximate Bayesian computation. We previously used simulated data sets to demonstrate proof of concept for application of the CGM-WGP methodology to plant breeding METs. Here the CGM-WGP methodology is applied to an empirical maize (*Zea mays* L.) drought MET data set to evaluate the steps involved in reduction to practice. Positive prediction accuracy was achieved for hybrid grain yield in two drought environments for a sample of doubled haploids (DHs) from a cross. This was achieved by including genetic variation for five component traits into the CGM to enable the CGM-WGP methodology. The five component traits were a priori considered to be important for yield variation among the maize hybrids in the two target drought environments included in the MET. Here, we discuss lessons learned while applying the CGM-WGP methodology to the empirical data set. We also identify areas for further research to improve prediction accuracy and to advance the CGM-WGP for a broader range of situations relevant to plant breeding.

M. Cooper, and C. Messina, DuPont Pioneer, 7250 NW 62nd Avenue, Johnston, IA 50131; F. Technow, DuPont Pioneer, 596779 County Road 59N, Woodstock, Ontario N4S 7W1, Canada; C. Gho, DuPont Pioneer, Semillas Pioneer Chile Ltda, Santa Filomena 1609–Buin, PO Box 267, Chile; and L.R.Totir, DuPont Pioneer, 8305 NW 62nd Avenue, Johnston, IA 50131. Received 20 Aug. 2015. Accepted 30 Nov. 2015. *Corresponding author (mark.cooper@pioneer.com).

Abbreviations: ABC, approximate Bayesian computation; AM, area of the largest leaf; BLUP, best linear unbiased prediction; CGM-WGP, crop growth models with whole-genome prediction; CGM, crop growth model; DH, doubled-haploid; FS, flowering stress; GBLUP, genomic best linear unbiased prediction; GEI, genotype \times environment interaction; MEB, minimum ear biomass; MET, multienvironment trial; QTL, quantitative trait loci; SNP, single-nucleotide polymorphism; TLN, total leaf number; TPE, target population of environments; TUS, timing of pollen shed; VPDB, vapor pressure deficit breakpoint; WGP, whole-genome prediction.

THE UBIQUITOUS NATURE and importance of GEIs for plant breeding have been widely documented and studied; conferences have been held, papers written, and books published. The optimistic view is that through understanding the causal genetic and environmental components of trait GEIs, breeding strategies can be designed to exploit repeatable components of the GEIs to improve realized genetic gains achieved from selection. With genotypes developed from such breeding strategies agronomists can then identify positive combinations of genotypes and crop management (Cooper and Hammer, 1996). Widespread adoption of the improved genotypes by farmers together with the appropriate management practices can enable superior industry scale productivity (e.g., Gaffney et al., 2015). Some examples of positive repeatable components of GEIs have been reported. However, for the majority of cases in plant breeding the relative magnitudes of

Published in Crop Sci. 56:2141–2156 (2016).
doi: 10.2135/cropsci2015.08.0512

© Crop Science Society of America
5585 Guilford Rd., Madison, WI 53711 USA
This is an open access article distributed under the CC BY-NC-ND license
(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

sources of genotypic and GEIs have been documented, and the GEIs are regularly treated as a source of uncertainty that reduce the rates of genetic gain achieved by breeding. In a few cases, the genetic and environmental contributions to specific GEIs have been described and in some cases these descriptions lend themselves to prediction of GEIs (Chapman et al., 2002, 2003; Cooper et al., 2002; Chenu et al., 2009; Messina et al., 2011, 2015). Nevertheless, to date, there has been limited use of such information to improve prediction in plant breeding. Typically the incidence of GEIs and a lack of their understanding results in a reduction in the predictability of trait phenotypes and a reduction in the realized rates of genetic gain achieved from selection. These negative experiences with GEIs have shaped many of the components of the plant breeding programs that are operating today, for example, extensive use of METs.

Successful design of METs requires a strategy for sampling the important environmental conditions that are expected to occur within the target geography of the breeding program. The concept of a target population of environments (TPE) for a breeding program can be defined as the set of key types of environmental conditions together with their frequencies of occurrence that are expected to occur within the target geography of the breeding program. For example, for dryland maize in the western region of the US Corn Belt, water limitations can occur and different drought environment types can be expected as components of the TPE. Designing METs that adequately represent a TPE requires characterization of the frequent repeatable environment types within the TPE—referred to here as envirotyping—and a strategy for sampling and creating environmental conditions at different locations to represent these environment types.

A common breeding strategy is to avoid some of the GEIs and place restrictions on the geographical scope of the TPE for any one breeding program. Then, within the defined geographical target, ensure that there is adequate sampling of the range of environmental conditions that are expected to be encountered across years, locations, and crop management practices adopted by farmers. Breeding programs that are so designed are the backbone of the long-term rates of genetic gain for yield and other complex traits that have been reported for crops in US agriculture (Smith et al., 2014). To do better than what has been achieved to date, plant breeders will require quantitative methods that identify positive components of GEIs and improve the predictability of genotype performance for key repeatable environmental components of the TPE. Further, the prediction methodology must operate at the scale of the breeding program and provide timely results to support selection decisions.

The development of marker and statistical technologies to map the genetic architecture of quantitative traits

has enabled the study of GEIs at the level of quantitative trait loci (QTL). For commercial maize breeding programs, multiple mapping studies of quantitative traits in the elite breeding populations have been conducted. Collectively, these studies have revealed that the genetic architecture of the important traits of maize involves many QTL, with multiple functional alleles, each with small effects (e.g., Melchinger et al., 2004; Boer et al., 2007; van Eeuwijk et al., 2010; Cooper et al., 2014a). Following multiple empirical studies, marker-assisted selection for specific QTL has not had the widespread impact on genetic improvement of quantitative traits that was originally anticipated. More recently the development of WGP methodology has opened up encouraging opportunities for working with contributions from multiple small QTL (Meuwissen et al., 2001; Bernardo, 2008; Heffner et al., 2009; Cooper et al., 2014a). In the case of WGP, the genetic contributions of all regions of the genome are assessed within a statistical model and the cumulative effects of their contributions across the whole genome is used to predict the breeding value or the genetic value of individual genotypes. The WGP methodology is implemented by applying a statistical algorithm to fit a suitable quantitative genetic model to data obtained from an appropriately chosen training data set. The individuals comprising the training data set are genotyped with numbers of markers across the genome consistent with the expected linkage disequilibrium within the reference population, and they are phenotyped for the relevant traits. The statistical model applied to the training data set associates the polymorphic markers with the phenotypic variation for the traits. The selected model is then applied to make predictions for suitable application data sets. In a typical application, the individuals in the application data sets are genotyped with markers across the whole genome, but they have not been phenotyped for the target traits in the target environments. The statistical model developed in the training data set is combined with the marker fingerprints of the individuals in the application data sets to predict the unobserved trait phenotypes of the targeted individuals for the target environments. Simulation and empirical studies have been conducted to evaluate the accuracy of the predictions in suitable application data sets.

While the models used to enable WGP are founded on quantitative genetic theory they are statistical representations of a complex biological system. The statistical models do not explicitly take into consideration much of the biology that contributes to GEIs (Hammer et al., 2006). The methodology for WGP relies heavily on the creation of suitable training data sets that span the highly combinatorial inference space (new genetic combinations in the environments of the TPE) for the intended applications. Fitting the models to the marginal effects across the environments of the training data set to predict marginal performance in the application set is the most common approach; however,

the presence of important GEIs in the application TPE presents some challenges. Extensions to the quantitative genetic models that rely on the identification of suitable environmental covariates have been investigated (Boer et al., 2007; Heslot et al., 2014). Defining suitable environmental breakouts within the training and application data sets based on environmental criteria associated with important target components of the GEIs has been applied to generate separate predictions for each breakout target with some success (Cooper et al., 2014b). However, these methods do not explicitly take into consideration the dynamic nature of the biology underlying the GEIs. Further, they are resource intensive to apply. Recently, Technow et al. (2015) proposed a novel methodology for WGP that can take into consideration the biology underpinning the GEIs. This was achieved through integration of an appropriate CGM within the model parameter estimation step using approximate Bayesian computation (ABC; Marjoram et al., 2014). They referred to this methodology as CGM-WGP. Here we review the salient aspects of developments in three components, managed environments, crop growth models, and extended prediction algorithms, which can be combined to provide such a quantitative prediction methodology. We then demonstrate an application of CGM-WGP methodology to a real maize drought data set as a first application of CGM-WGP to an empirical MET.

Managed Environments

Plant breeders routinely conduct METs for multiple stages of breeding programs—sampling multiple locations and years—to test the genotypes under consideration for the different stages. The data sets obtained from these METs can be used to construct suitable training data sets for WGP. A limitation of this approach is that it is possible to miss many of the important environmental conditions of the TPE that are not sufficiently sampled in the METs. Where possible, the samples of environments obtained from the TPE can be augmented by conducting additional designed managed environment experiments to ensure the genotypes are tested in key environmental conditions that are likely to be missed or underrepresented in the METs. This approach has been successfully applied to develop drought tolerant maize hybrids for the US Corn Belt (Cooper et al., 2014b; Gaffney et al., 2015).

Crop Growth Models

Similar to plant breeders, agronomists routinely conduct METs. The treatment combinations and objectives of these agronomy trials differ from the plant breeding METs. Typically, the agronomists are seeking appropriate agronomic management strategies (combinations of planting date, plant population, fertilizer rates, irrigation strategies, etc.) for the elite genotypes developed from the breeding programs. Crop scientists have developed CGMs

to augment and extend the results of these agronomic experiments (Keating et al., 2003; Holzworth et al., 2014). The motivation for such applications of CGMs is to extend the results of the empirical studies to the broader range of conditions that can be expected by farmers in the TPE. Recently, extensions of these CGMs have been proposed to incorporate functional relationships between the environmental variables and traits that are influential on the genetic variation for yield and agronomic performance in elite reference populations of interest to plant breeders (Chapman et al., 2002, 2003; Cooper et al., 2002; Hammer et al., 2006; Chenu et al., 2009; Messina et al., 2011, 2015). These developments of CGMs provide tools for extending the study of GEIs to the level of the prediction problems of the plant breeder seeking to create novel improved genotypes from within the breeding program.

Extended Prediction Algorithms

We seek a practical method of applying WGP that takes into consideration the consequences of GEIs for genotype performance within the environments and management practices of the TPE. If this can be achieved, the breeder could consider predicted performance of genotypes for key environment types of the TPE and move beyond predictions based solely on marginal performance across environments. Technow et al. (2015) used simulation to demonstrate how CGMs could be applied with WGP for such purposes. This extension of WGP was achieved within a high performance computing environment by application of an approximate Bayesian computation (ABC; Marjoram et al., 2014) algorithm that used a suitable CGM in place of the likelihood function to obtain the WGP outcomes; we identify the WGP methodology with the CGM embedded within the prediction algorithm as CGM-WGP. The simulation results reported by Technow et al. (2015) were encouraging. It is necessary to consider how to scale the proposed methods to the empirical data sets generated by breeding programs. This paper complements the proof-of-concept paper of Technow et al. (2015) by applying the CGM-WGP methodology to an empirical maize data set and focuses on key practical considerations that are involved in the reduction-to-practice of the CGM-WGP methodology.

Application of Crop Growth Models with Whole-Genome Prediction to a Maize Multienvironment Trial

Breeding for improved grain yield of maize hybrids for the western region of the US Corn Belt represents a complex breeding challenge. The maize agriculture practices of the region create a complex TPE where GEIs for yield are common; maize is grown in a mixture of water-availability environments that range from fully irrigated through multiple limited-irrigation strategies to

many different dryland farming practices. Improvements in WGP that take into consideration these GEIs would provide breeders and agronomists with new opportunities to accelerate the codevelopment of improved maize hybrids and management strategies that deliver sustainable improvements in on-farm productivity. Maize breeding programs designed with the objective of developing improved hybrids for the range of water environments of the western region have used METs that combine both on-farm and managed-environment testing (Cooper et al., 2014b). The managed-environment experiments are designed to represent a range of the on-farm environmental conditions that can occur within the western region TPE. Complex GEIs for yield are a common feature of these managed-environment experiments.

The objective of this paper is to report the application of the CGM-WGP methodology to prediction of grain yield variation within a maize cross using a representative MET data set that was obtained from the early-stage testing conducted by a commercial maize breeding program. The focus here is on the practical steps required to achieve a successful implementation of the CGM-WGP method to an empirical data set. This represents the first step in moving the application of the CGM-WGP methodology from the proof-of-concept phase, which was based on simulated data (Technow et al., 2015), to assess the ability of the CGM-WGP methodology to provide predictions for an experimental data set of the type generated by a commercial maize breeding program. Further research and steps required to broaden the initial application reported here are discussed.

MATERIALS AND METHODS

Data Set

The MET data set selected to evaluate the empirical implementation of the CGM-WGP method was based on grain yield results obtained for a single biparental cross of maize entries evaluated as testcross hybrids in two drought managed environments (treatments) at a single location.

The inbred parents of the biparental cross were selected to contrast for their grain yield breeding value under drought: one parent previously characterized to have high breeding value and the other low breeding value. The parents were crossed to produce the F_1 generation and the F_1 was self-pollinated to produce the F_2 generation. The biparental cross was represented by 106 DH lines derived from a random sample of individuals from the F_2 generation. The 106 DH lines were genotyped with a total of 86 single-nucleotide polymorphism (SNP) markers distributed across the 10 chromosomes. The SNPs were previously identified to be polymorphic between the two parents. The 106 DH lines were crossed with an inbred tester line to generate testcross hybrid seed. The tester line was selected from the complementary heterotic group and was considered to have high breeding value for grain yield under drought. All grain yield data were generated on the testcross hybrid seed for the 106 DH lines.

The 106 DH lines were evaluated for grain yield in experimental plots in two drought environments. The two drought environments were generated by creating two drought treatments in two experiments conducted in adjacent fields at the Pioneer Viluco station in Chile in the 2011–2012 season. Quantity and timing of irrigation was used to generate the different drought treatments. Irrigation was managed through a drip tape system installed in the experimental plots at planting. The experimental plots were each two rows, 4.5 m long with 0.75-m spacing between rows. The drip tape was inserted into the soil at planting beside each row in each plot of the experiment. The drought treatments were implemented by regulating the amount of irrigation water that was supplied to the plots through the drip tape system installed within the experiment. The supply of water was managed differently between the two drought experiments to generate two different levels of water supply, thus two different drought treatments. The irrigation schedule was managed to coincide the timing of the maximum water deficit with the flowering period of the 106 DH lines. A characterization of the temporal patterns of water deficit that was achieved in the two drought treatments is shown in Fig. 1. The water deficit was measured by the ratio of the availability of water (supply) to the developing crop relative to the quantity of water required by the crop canopy (demand); calculation of supply/demand ratio is described below.

The supply/demand ratio was calculated from separate estimates for each component. The supply term was calculated from the aggregated supply of water from all of the soil layers occupied by roots, that is, the amount of water available in the soil layer and a coefficient that represents the combined conductivity of water through the soil and roots following Robertson et al. (1993). Irrigation and precipitation were inputs to the model that determined the amount of water available to the plant. The demand term was estimated using dry matter growth, the transpiration efficiency coefficient (9 kPA), and the vapor pressure deficit (VPD).

A supply/demand ratio of 1.0 indicates water supply meets canopy demand and there is no water deficit for growth. As water supply drops below the level that is required to meet the canopy demand the supply/demand ratio decreases below 1.0, and a supply/demand ratio of 0 would occur when there was no available water to meet the demand of the crop canopy. In both drought environments, the irrigation quantity and timing was managed to impose a water deficit that coincided with flowering time. Thus, the environments were characterized as flowering stress (FS) environments and distinguished as FS1 (Flowering Stress Environment 1) and FS2 (Flowering Stress Environment 2). For FS1 the water deficit commenced earlier and reached lower levels of the supply/demand ratio, indicating more intense drought than for FS2. Following the flowering period, the water deficit was relieved by restarting the irrigation schedule.

The experimental design for both environments was based on two replicates. The 106 DH lines were evaluated in a row-column configuration together with a number of other DH lines and a set of commercial hybrid checks. For the objectives of this paper, these additional DH lines and the commercial hybrid checks will not be considered further other than to recognize that they were part of the data set from which the 106 DH lines were obtained. The grain yield data were obtained using a two-plot combine harvest system that measured the

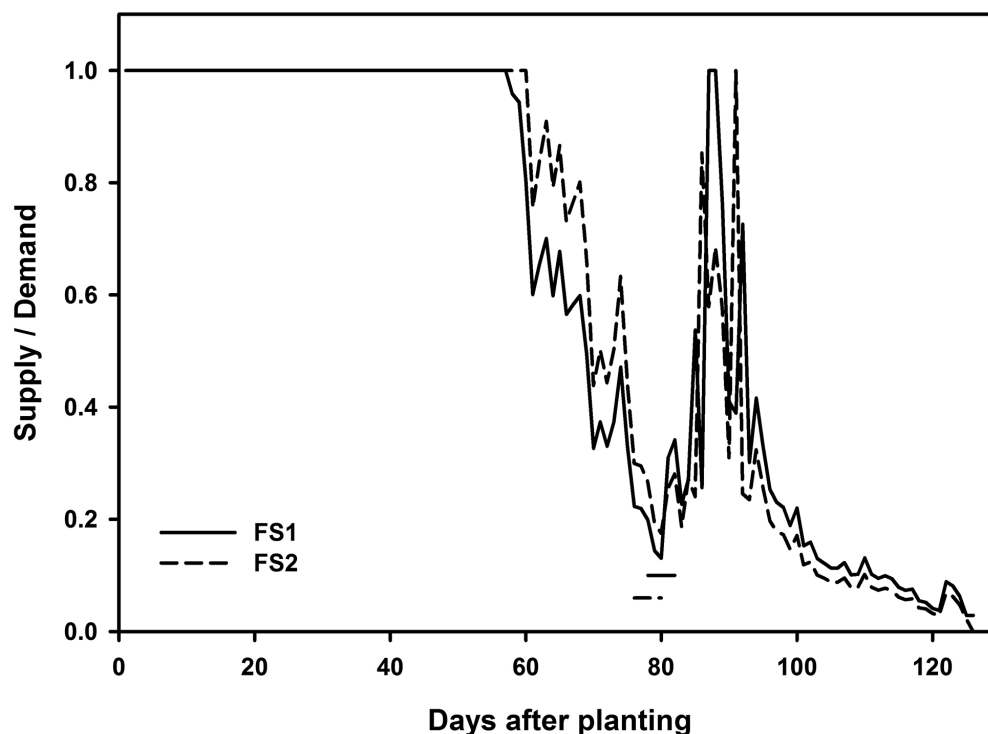


Fig. 1. Temporal pattern, measured as a daily time step from planting, of modeled water supply/demand ratio for drought environments FS1 (Flowering Stress 1) and FS2 (Flowering Stress 2) conducted at the Pioneer Viluco station in the 2011–2012 season. The supply/demand ratio indicates the balance of water supply from the soil to meet the demand from the canopy, given the atmospheric conditions. A supply/demand ratio of 1.0 indicates supply meets demand and no stress. Conversely, a supply/demand ratio below 1.0 indicates that supply does not meet demand and that there is a level of water deficit for the day. The window of the flowering time period for the set of 106 doubled-haploid entries measured as the time of pollen shed is indicated for both environments by the horizontal bars at ~80 d after planting.

weight of grain obtained from the plot and the grain moisture content. The grain harvest weight per plot was adjusted to grain yield per unit area at 15% moisture content. The grain yield data were analyzed using a mixed model that included terms for the row and column position of the plots and the spatial correlation of the estimated plot residuals. The 106 DH lines were considered to represent a random sample of the possible DHs that could have been obtained from the biparental cross. Accordingly, the genotypic term for the trait variation among the 106 DH lines was treated as random and best linear unbiased predictions (BLUPs) were obtained for grain yield of each of the 106 DH lines for both of the drought treatments.

Crop Growth Model

The CGM used in this study was based on the mechanistic model developed by Muchow et al. (1990) and integrated within the ABC algorithm by Technow et al. (2015). Briefly, the CGM uses concepts of resource use, resource use efficiency, and resource allocation to grain to simulate grain yield. Figure 2 provides a schematic of the key relationships that operated within the CGM used for this study. Light interception is modeled based on leaf appearance rate, the area of the largest leaf (AM), total leaf number (TLN), planting density, and a coefficient of light extinction within the crop canopy. Simulation of daily increase in total mass results from the product of light interception and radiation use efficiency on a given day. Grain yield is simulated from the daily increase in harvest index starting 3 d after the end of the flag leaf expansion and ending at physiological maturity (Muchow et al., 1990). Irrigation, precipitation, temperature,

and solar radiation are environmental variables that are input to the model. The vapor pressure deficit is calculated from temperature data following Messina et al. (2015).

Since the motivation of this study was to demonstrate the CGM-WGP methodology for a maize population evaluated under drought stress conditions, the model was modified to simulate soil water balance, transpiration, and growth response to water deficit (Fig. 2). The soil water balance was modeled using a multilayer approach as described by Ritchie (1998). The components of the soil water balance, infiltration, runoff and evaporation were simulated as described by Muchow and Sinclair (1991). Evaporation was modeled using a two-stage model. Transpiration was modeled based on mass growth and a transpiration efficiency coefficient equal to 9 Pa (Tanner and Sinclair, 1983). The limited transpiration trait was implemented as described by Messina et al. (2015) with the difference that in this study the transpiration response to vapor pressure deficit above a vapor pressure deficit breakpoint (VPDB) was modeled as a continuous linear function rather than a constant maximum value. Root water uptake was simulated using first order kinetics (Robertson et al., 1993) with the exponent of the function describing root occupancy and hydraulic conductivity; this parameter was set to 0.08 (Dardanelli et al., 1997). The sum of the potential water uptake across soil layers determined the soil water supply, while the potential transpiration calculation based on light limited growth determined the water demand term. The ratio of these two components—supply and demand—define a stress index that was used to affect mass growth and leaf expansion.

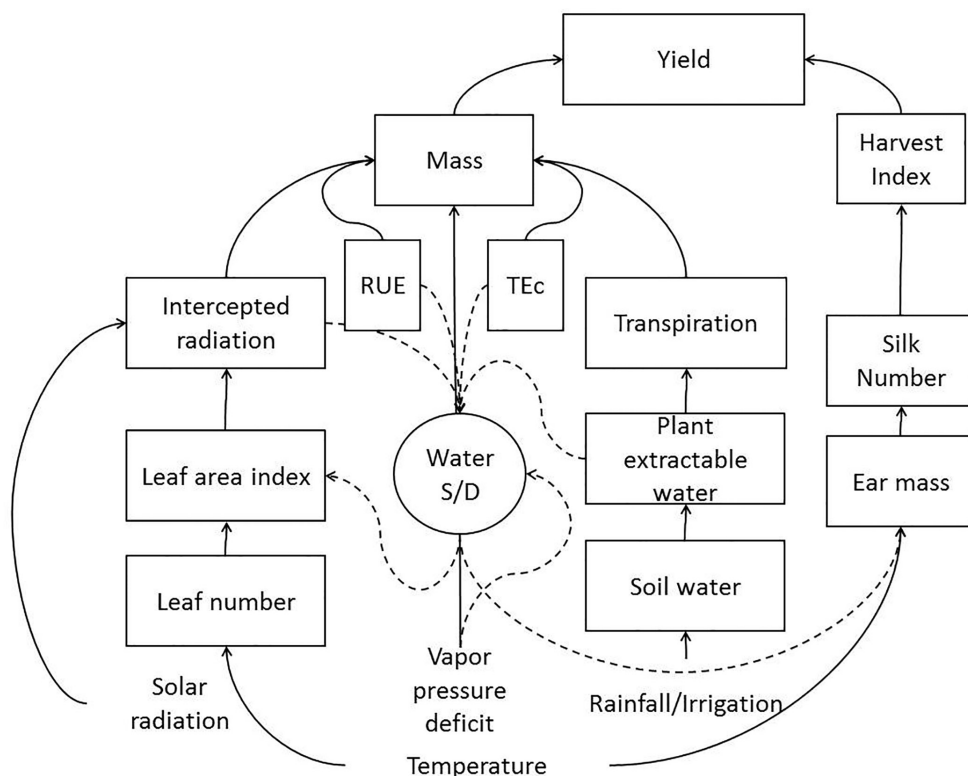


Fig. 2. Schematic representation of the structure of the crop growth model (CGM) based on environmental resource capture, resource use efficiency and partitioning to determine the ultimate trait grain yield. The diagram identifies the key environmental inputs relevant for the current study, solar radiation, and water; the main plant functional components involved in water and radiation resource capture; the interactions among the functional components; and the environmental inputs and temperature conditions to determine the water supply/demand ratio. In combination with the physiological processes that determine resource capture, resource use efficiency, quantified as solar radiation use efficiency (RUE), and transpiration efficiency (TEc) determine plant mass. Physiological processes of mass partitioning, that are dependent on the supply/demand ratio and temperature, then determine ear mass and silk number and ultimately grain yield through the harvest index function.

Because the objective of the model was to simulate maize yield subject to water deficit at flowering time (Fig. 1), and the harvest index approach was inadequate to simulate maize grain yields in these types of stress environments (Sinclair et al., 1990), the model was modified to incorporate elements important to describe the dynamics of silk emergence and ear growth (Fig. 2), processes that are sensitive to water deficit (Edmeades et al., 1993; Westgate and Boyer, 1985). The attainable harvest index was modeled as a function of a potential harvest index, which corresponds to that attained in the absence of water deficit; a potential number of silks that results from the maximum number of rows and rings of kernels in the ear; the exerted number of silks 3 d after silking; and the potential increase in kernel weight when the source exceeds the sink capacity. Borrás et al. (2004) demonstrated that kernel weight can increase about 20% under these conditions. The number of exerted silks was modeled using a negative exponential function as proposed by Cooper et al. (2014a). The parameter trait minimum ear biomass (MEB) corresponds to the threshold in ear mass growth below which silks do not emerge from the husk enclosing the ear. The potential number of silks defines the yield potential. The exponent of the function defines the rate of silk appearance per unit ear growth, which was modeled using an exponential function of thermal time and a stress factor directly proportional to the supply/demand ratio. To account for the plant-to-plant variation

in flowering time, growth and development of three ears were implemented. The weighted average of the emerged silks for these three ears was used to determine the final attainable harvest index. The onset of ear growth was set at vegetative stage fifteen. Yields were simulated using a daily increment in harvest index, which was updated from the potential harvest index (Muchow et al., 1990) to that determined at flowering time based on the effects of water deficit on ear growth and silk emergence.

Approximate Bayesian Computation

Five traits were identified as key components of the CGM for investigation within the ABC framework: TLN, AM, VPDB, MEB, and the cumulative thermal units from completion of canopy development as measured by the completion of flag leaf expansion and the timing of pollen shed (TUS). Together, the component traits TLN and AM influence canopy size, which influences soil water balance in water-limited environments. The component trait VPDB influences transpiration rate of the canopy and can also influence soil water balance. The component trait MEB influences reproductive resiliency and ultimately kernel set when water limitations coincide with the flowering period. The component trait TUS allowed for a source of genetic variation for flowering time other than that associated with the variation for TLN.

Following the methodology introduced by Technow et al. (2015) the five component traits were treated as latent variables for prediction by the CGM-WGP methodology. The latent value for each trait for each DH entry (y_{TRAIT_i}) was modeled as a linear function of trait specific marker effects:

$$y_{\text{TLN}_i} = \mu_{\text{TLN}} + \mathbf{z}_i \mathbf{u}_{\text{TLN}}$$

$$y_{\text{AM}_i} = \mu_{\text{AM}} + \mathbf{z}_i \mathbf{u}_{\text{AM}}$$

$$y_{\text{VPDB}_i} = \mu_{\text{VPDB}} + \mathbf{z}_i \mathbf{u}_{\text{VPDB}}$$

$$y_{\text{MEB}_i} = \mu_{\text{MEB}} + \mathbf{z}_i \mathbf{u}_{\text{MEB}}$$

$$y_{\text{TUS}_i} = \mu_{\text{TUS}} + \mathbf{z}_i \mathbf{u}_{\text{TUS}}$$

where \mathbf{z}_i is the vector of the observed biallelic SNP markers of DH entry i ; μ_{TLN} , μ_{AM} , μ_{VPDB} , μ_{MEB} , and μ_{TUS} are the intercepts for the five traits; and \mathbf{u}_{TLN} , \mathbf{u}_{AM} , \mathbf{u}_{VPDB} , \mathbf{u}_{MEB} , and \mathbf{u}_{TUS} are the vectors of marker effects for the five traits. For brevity, we will use $\boldsymbol{\theta}$ to denote the joint parameter vector [μ_{TLN} , ..., μ_{TUS} , \mathbf{u}_{TLN} , ..., \mathbf{u}_{TUS}].

Approximate Bayesian computation is a novel technique developed in population genetics that allows analysis of otherwise intractable models (Marjoram et al., 2014). The ABC rejection sampling algorithm has three basic steps. In step one a candidate parameter vector is drawn from the prior distribution of the parameter vector $\boldsymbol{\theta}$ of the component trait (treated as latent variables) specific marker effects. Then, in step two, new data are simulated from the biological model (CGM) by using the candidate parameter vector as the input to the CGM. Finally, in step three, the distance between the CGM simulated data (e.g., yield) and the observed data (e.g., experimental yield BLUPs) for the training data set is calculated, and if this distance is below a preset tolerance threshold, the set of values in the candidate parameter vector $\boldsymbol{\theta}$ of trait-specific marker effects is accepted as an independent sample from the posterior distribution of $\boldsymbol{\theta}$. This process is repeated until a sufficient number of samples is obtained. These accepted samples then approximate the posterior distribution of the parameter vector; the lower the tolerance threshold, the better the approximation.

Defining Prior Information for Crop Growth Model Component Traits

Independent normal distribution priors were used for the five traits for all components of $\boldsymbol{\theta}$. The prior for the intercepts μ_{TRAIT} was $N(m_{\text{TRAIT}}, \sigma_{\mu_{\text{TRAIT}}}^2)$, where m_{TRAIT} is the prior mean and $\sigma_{\mu_{\text{TRAIT}}}^2$ the prior variance, which quantifies uncertainty in the intercept. The prior for the marker effects $\mathbf{u}_{\text{TRAIT}}$ was $N(0, \sigma_{u_{\text{TRAIT}}}^2)$, where the variance parameter $\sigma_{u_{\text{TRAIT}}}^2$ controls the shrinkage of the marker effects toward 0. This prior corresponds to the BayesC prior (Habier et al., 2011). The m_{TRAIT} , $\sigma_{\mu_{\text{TRAIT}}}^2$, and $\sigma_{u_{\text{TRAIT}}}^2$ values for the five traits are given in Table 1.

Different sources of information were used to obtain the prior values for the five component traits. For TLN, AM, and MEB, a subset of 38 of the 106 DH lines was evaluated for

Table 1. Prior parameter values for the five component traits that were identified to influence grain yield in the two drought environments and treated as sources of genetic variation within the crop growth model (CGM) used within the CGM-whole genome prediction methodology.

	Trait				
	Total leaf number	Area of largest leaf	Vapor pressure deficit breakpoint	Minimum ear biomass for silk exertion	Timing of pollen shed
		cm ²	kPa	g	°C
m_{TRAIT}	19.37	842.90	1.90	0.76	40.00
$\sigma_{\mu_{\text{TRAIT}}}^2$	0.01	2.00	0.1	0.001	5.00
$\sigma_{u_{\text{TRAIT}}}^2$	0.016 ²	7.19	0.022 ²	0.005 ²	0.176

TLN, AM, and MEB in an experiment conducted in Iowa in 2011. The data for these three traits were obtained using the same testcross seed source used to obtain the grain yield data. As for grain yield, the trait measurements were analyzed using a mixed model, and BLUPs were obtained for the DH lines. For the component traits TLN, AM, and MEB, m_{TRAIT} was then computed as the average of the trait BLUPs for the subset of 38 DHs included in the 2011 Iowa experiment. Also for these three traits, $\sigma_{\mu_{\text{TRAIT}}}^2$ was computed as $\text{var}(\text{TRAIT})/M$, where $\text{var}(\text{TRAIT})$ is the observed variance of the trait BLUPs in the Iowa experiment and M the number of markers.

For the component traits VPDB and TUS, no direct measurements were made on the DH entries. All information used to define the prior parameters was based on published information for maize. For the VPDB trait, the results reported by Gholipour et al. (2013) were used. For TUS, the prior parameters were determined based on a combination of published information indicating a TUS interval of 3 d (Muchow et al., 1990) and field observations indicating synchronous termination of leaf expansion and commencement of shedding for drought tolerant hybrids.

The ABC algorithm was implemented as described by Technow et al. (2015). The simulation model operator Model ($y_{ik}^* | \boldsymbol{\theta}$) comprised the CGM $F(\cdot)_{ik}$ as the deterministic component, indexed for DH entry i in environment k , and a Gaussian noise variable distributed as $N(0, \sigma_{\epsilon}^2)$ as the stochastic component. The value of σ_{ϵ}^2 was set equal to 5% of the observed variance of the grain yield BLUPs. The tolerance level was tuned to an acceptance rate of approximately 1×10^{-6} . The number of posterior samples drawn was 400. The CGM-WGP algorithm was implemented as a C routine integrated with the R software environment (R Core Development Team, 2014).

Whole-Genome Prediction with Crop Growth Models Estimation, Prediction, and Testing Procedure

The CGM-WGP models were fitted and parameter estimates obtained using either data from the FS1 or FS2 environment. A random set of 50 DH entries was used as the training set, referred to from hereon as the estimation set. The remaining 56 DH entries were then used to test the model performance

and are referred to as the test set. The environment from which the data were sampled to fit the CGM-WGP model will be referred to as the estimation environment. The other environment will be referred to as the new environment. For the purposes of this paper, the other environment is new in the sense that no data from that environment were used to select the CGM-WGP model or estimate the parameters of θ . The selected CGM-WGP model was then tested in both the estimation environment and the new environment; for example, the model was selected based on the sample of 50 DH entries in the FS1 environment—in this case the FS1 environment is the estimation environment and the FS2 environment is the new environment—and then was tested on the remaining 56 DH entries in the FS1 estimation environment and the FS2 new environment. Once the CGM-WGP model was selected, the parameter estimates were used to predict yield of the DH entries for both the estimation and test sets of DH entries in both the FS1 and FS2 environments. Predictions for the same environment as the estimation environment will be referred to as observed environment predictions (e.g., predictions for FS1 with models fitted with FS1 data). Predictions for an environment from which no data were used in fitting the model will be referred to as new environment predictions (e.g., predictions for FS2 with models fitted with FS1 data). This process was replicated 20 times for the FS1 and FS2 environments. As a point estimate for the predicted grain yield of a DH entry i in a specific environment k , we used the mean of the posterior predictive distribution for the DH entry in question. The posterior predictive distribution was obtained by evaluating the CGM $F(\cdot)_{ik}$ over the accepted θ samples using the weather, soil, irrigation, and management data for that environment.

Prediction accuracy for the CGM-WGP was computed as the Pearson product moment correlation between predicted and observed performance of the DH entries in the environment for which the prediction was made. As a performance benchmark, genomic best linear unbiased prediction (GBLUP; Meuwissen et al., 2001) was also applied to all data sets.

RESULTS

The irrigation schedule applied to the two experiments resulted in two temporal patterns of water deficit over the course of the experiments (Fig. 1). The period of maximum water deficit coincided with the timing of flowering for the DH entries in both experiments. The differential irrigation management between the two experiments generated the managed drought environments FS1 and FS2.

Analysis of variance for the grain yield data indicated that there was significant genotypic variation among the DH entries (Fig. 3). There was also significant genotypic variation for the timing of flowering, as measured by growing degree day units from planting to pollen shed. However, there was no linear or nonlinear association between growing degree day units from planting to pollen shed and grain yield in the two environments ($P > 0.05$; data not shown). Therefore, variation for flowering time was not considered to have had a direct effect on grain yield for the progeny of the chosen cross in the two

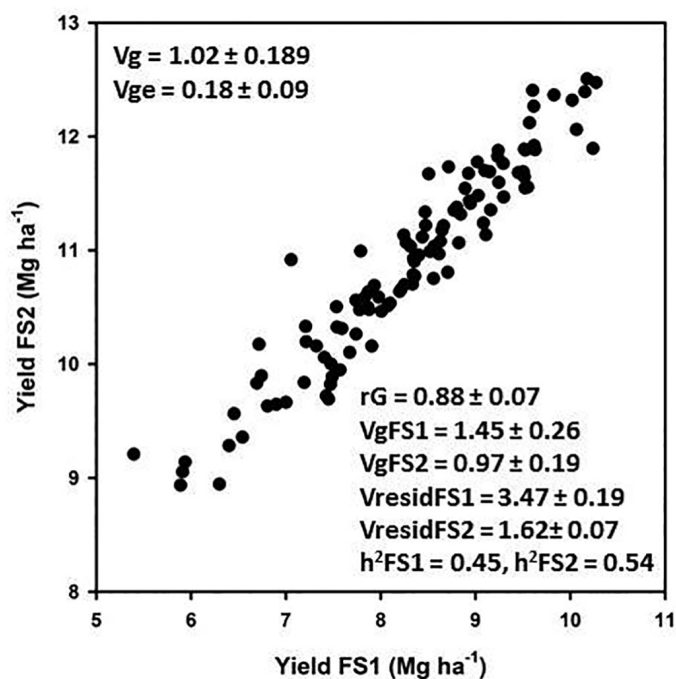


Fig. 3. Grain yield best linear unbiased predictors for the complete set of 106 doubled-haploid entries evaluated in the drought environments FS1 (Flowering Stress 1) and FS2 (Flowering Stress 2) conducted at the Pioneer Viluco station in the 2011–2012 season. Summary results from analysis of variance of grain yield are provided. V_g , genetic component variance for yield across FS1 and FS2; V_{ge} , genotype \times environment interaction component of variance for yield between FS1 and FS2; r_G , genetic correlation for yield between FS1 and FS2; V_{gFS1} , genotypic variance component for yield within FS1; V_{gFS2} , genotypic variance component for yield within FS2; $V_{residFS1}$, residual variance component within FS1; $V_{residFS2}$, residual variance component within FS2; h^2_{FS1} and h^2_{FS2} , estimates of line mean heritability for yield within FS1 and FS2, respectively.

drought environments. Consequently, the major component of grain yield variation for the DH entries was considered to be associated with trait variation other than timing of flowering: TLN, AMAX, VPDB, and MEB. While strong direct effects of the genetic variation for flowering time on yield were not detected, the possibility of indirect conditional effects of flowering time on yield was not excluded. Thus, genetic variation for the trait TUS was included in the CGM to allow for the possibility that grain yield variation could still be associated with flowering time effects conditional on the genetic variation for the other four traits.

The GEI for grain yield between the two environments was significant. The GEI component of variance was smaller than the genotypic component of variance (Fig. 3). The genetic correlation for grain yield between the two environments, while <1.0 , was estimated to be relatively high (0.88). A component of the detected GEI for grain yield was associated with heterogeneity in the magnitude of genotypic variance between the two environments (Fig. 3); the genotypic variance component for

grain yield within FS1 was approximately 1.5 times that of FS2. A scatter plot comparing the grain yield BLUPs between the two environments indicated that there were some rank differences for grain yield of individual DH entries between FS1 and FS2, but overall, there was reasonable agreement in the relative yield of the DH entries between the two environments (Fig. 3).

Thus, for the chosen MET data set there was a level of repeatability of grain yield performance of the DH entries between environments FS1 and FS2, together with a significant level of GEI for grain yield that, in part, was due to lack of genetic correlation and heterogeneity of genotypic variance. Therefore, the drought environments FS1 and FS2 both revealed genotypic variation for grain yield among the DH entries with a combination of similar and different genetic components of yield under drought revealed between FS1 and FS2. These grain yield results are similar to previous drought MET experiments conducted at the same stage of a maize breeding program (Cooper et al., 2014b).

Given these grain yield results, the chosen MET data set was considered suitable to be used as a case study for a first empirical evaluation of the CGM-WGP methodology. The relatively limited levels of change in rank of the DH entries between the two environments is expected to improve the chances of successful prediction between the environments for the GBLUP methodology in comparison with other situations where greater levels of rank change occur between the environments. Exploration of a wider range of GEI scenarios than the single example shown in Fig. 3 is discussed further below. Here, we focus on the requirements for the successful implementation of the CGM-WGP method for an empirical MET data set generated as part of a maize breeding program.

The initial set of environmental inputs that were used to run the CGM component of the CGM-WGP resulted in poor agreement between the grain yield predictions and observed results within the FS2 environment. There was good agreement for the FS1 environment. The predicted yield values obtained from the CGM were consistently lower than the observed yields. This resulted in a re-evaluation of the environmental inputs for the two environments. The initial assumption, based on available soil maps, was that the soil depth for the adjacent fields was the same and the different yield levels would be explained by the different irrigation schedules used for the two environments. Further investigation of the characterization of the soils for the two adjacent fields, based on soil cores, revealed that there was a significant ($P < 0.001$) difference in soil depth of ~0.2 m between the adjacent fields; the soil depth of the FS2 environment was deeper than originally assumed. Once this adjustment was made to the inputs for the CGM the predicted yields for the FS2 environment aligned with the observed yields. This experience is

Table 2. Average prediction accuracy obtained for the crop growth model-whole genome prediction (CGM-WGP) and genomic best linear unbiased prediction methods for grain yield of doubled-haploid (DH) entries evaluated as testcross hybrids in two drought environments (Flowering Stress 1 [FS1] and Flowering Stress 2 [FS2]), averaged over 20 replications. For each replication the 106 DH entries belonged to either the estimation set (50 DH entries) or test set (56 DH entries). For each implementation the two environments were defined as either the estimation environment or the prediction environment.

Estimation environment	Prediction environment	Estimation DH entries		Test DH entries	
		CGM-WGP	GBLUP	CGM-WGP	GBLUP
FS1	FS1	0.82	0.78	0.23	0.24
	FS2	0.53	0.51	0.21	0.21
FS2	FS1	0.50	0.53	0.22	0.23
	FS2	0.77	0.82	0.38	0.41

provided as an example of some of the additional detailed envirotyping requirements associated with applying the CGM-WGP in practice. While this may be seen as an additional cost, it also demonstrates that the CGM-WGP is responsive to the environmental inputs for the CGM component, a requirement to accommodate the effects of GEI.

Average prediction accuracy was positive for all CGM-WGP scenarios considered (Table 2). This result demonstrates that the CGM and the five component traits—TLN, AM, MEB, VPDB, and TUS—provided a relevant framework to define models that capture genetic variation for yield in the form of the approximate posterior distributions of the parameters of θ obtained by applying the ABC algorithm.

The highest average prediction accuracy for CGM-WGP was achieved for the scenarios where the estimation and prediction environments were the same (Table 2). This was the case for both the estimation and test sets of DH entries. This result is expected, since predictions within an environment do not have to accommodate the effects of any GEI that occur between different environments.

The prediction accuracy was consistently higher when the CGM-WGP was applied to the estimation set of DH entries in comparison to the application of the CGM-WGP to the test set of DH entries (Table 2). Thus, there was loss of model adequacy for purposes of prediction when the selected parameters of θ were applied to new DH entries sampled from the same reference population. This loss of predictive skill occurred whether the estimation and prediction environments were the same (i.e., FS1 to FS1 and FS2 to FS2) or different (i.e., FS1 to FS2 and FS2 to FS1).

The average prediction accuracy for grain yield achieved by CGM-WGP was similar to that for GBLUP for all scenarios considered (Table 2). While average prediction accuracy was similar, there were differences in prediction accuracy between CGM-WGP and GBLUP for

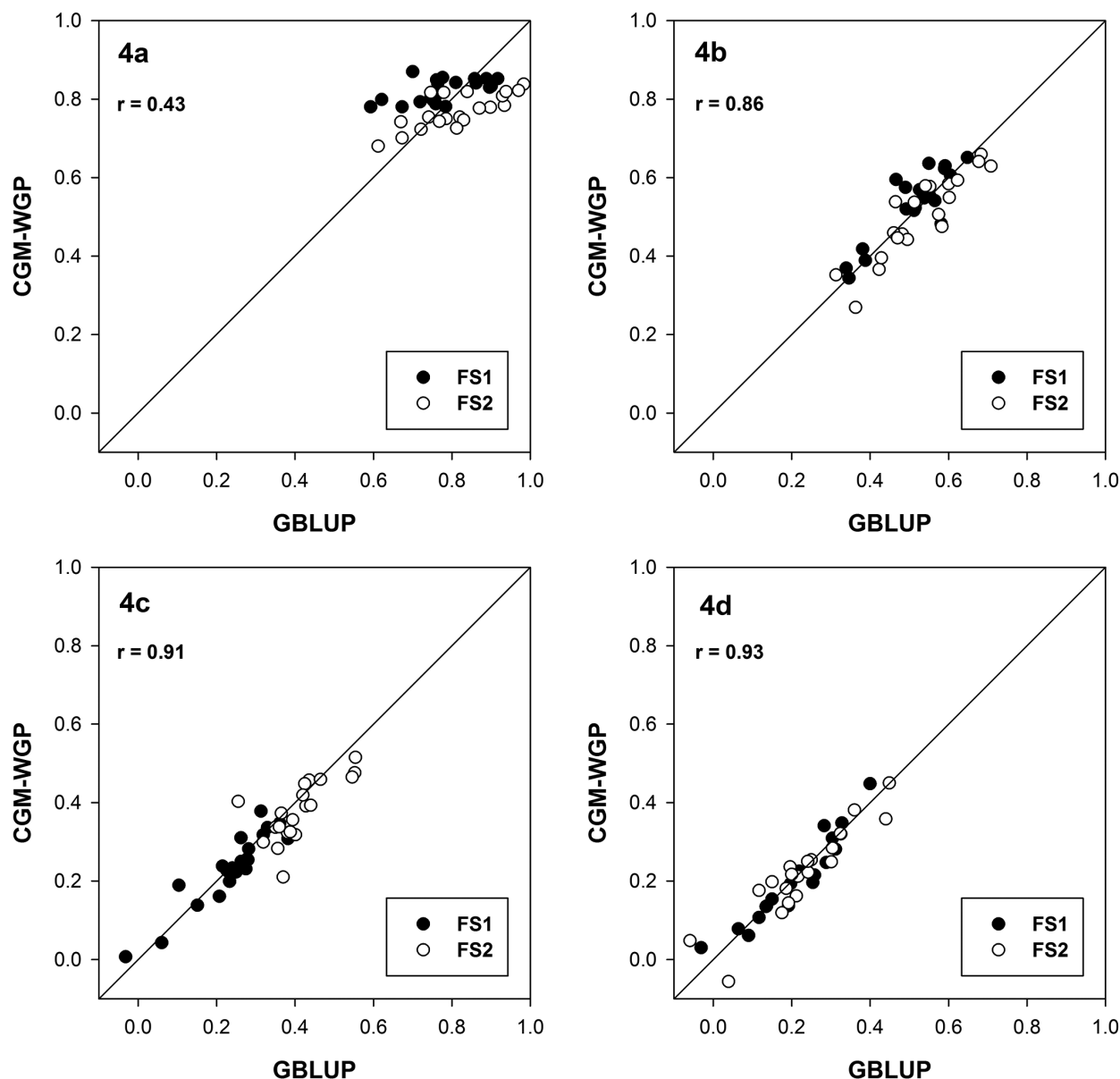


Fig. 4. Comparison of prediction accuracy obtained for testcross grain yield between crop growth model–whole genome prediction and genomic best linear unbiased prediction for 20 replications for a single maize cross evaluated in drought environments FS1 and FS2: (a) estimation entries in observed environments; (b) estimation entries in new environments; (c) test entries in observed environments; (d) test entries in new environments. Legend identifies the environment where genetic model parameters were estimated. The 1:1 line is included for reference.

individual replications (Fig. 4). Prediction accuracy for the individual replications was considered for four scenarios: (i) prediction of the estimation entries in the observed environments (Fig. 4a), (ii) prediction of the estimation entries in the new environments (Fig. 4b), (iii) prediction of the test entries in the observed environments (Fig. 4c), and (iv) prediction of the test entries in the new environments (Fig. 4d). Across all four scenarios, for some replications, the prediction accuracy was the similar for CGM-WGP and GBLUP, indicated by points close to the 1:1 line, and for other replications, the prediction accuracy differed and the points deviated from the 1:1 line; there were 15 cases where the prediction accuracy differed by >0.1 (Fig.

4). These differences indicate that in some cases different genetic models for yield were selected by the CGM-WGP and GBLUP methods when they were applied to the same estimation data sets. A consequence of the selection of different genetic models by the two prediction methods was that yield predictions for different DH entries changed with the prediction method and the ranking of the individual DH entries could also change based on the predictions (Fig. 5). Thus, depending on the selection intensities applied by the breeder different sets of DH entries could be selected for the different cases represented by the 20 replications even though the average prediction accuracies were similar for both prediction methods.

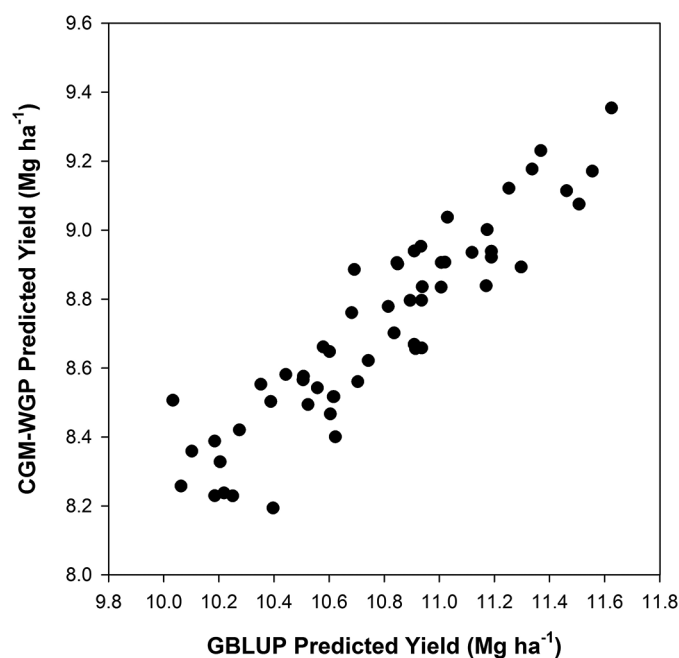


Fig. 5. Comparison of grain yield predictions based on genomic best linear unbiased prediction and crop growth model-whole genome prediction for one replication. Model selection was based on grain yield of 50 doubled-haploid (DH) entries in the FS2 (Flowering Stress 2) environment and predictions were for 56 DH entries in the FS1 (Flowering Stress 1) environment.

Inspection of the distribution of prediction accuracy for the replicates indicates that there are important differences in the prediction accuracy obtained among the four scenarios (Fig. 4). As expected, the prediction accuracy was consistently higher for the scenario where the predictions were made for the estimation entries in the observed environments (Fig. 4a). However, for this scenario there was only a weak correlation ($r = 0.43$) between the CGM-WGP and GBLUP prediction accuracy across the replicates. In contrast, for the other three scenarios, where prediction accuracy was estimated for the estimation entries in new environments (Fig. 4b) or the test entries in observed (Fig. 4c) and new (Fig. 4d) environments, the prediction accuracy was generally lower, and there was a stronger correlation ($r = 0.86$ – 0.93) between the CGM-WGP and GBLUP prediction accuracy across the replicates. The different distributions of the estimates of prediction accuracy for both CGM-WGP and GBLUP across the replicates for the four scenarios (Fig. 4) may reveal different features of the underlying genetic models selected by the CGM-WGP and GBLUP methods. For the scenario where predictions are made for the estimation entries in the observed environments (Fig. 4a), the lower correlation and the different distribution of prediction accuracy of the replicates between CGM-WGP and GBLUP suggests a potential difference between the CGM-WGP and GBLUP methods. For this scenario, the prediction accuracy ranged from 0.59 to 0.98 for GBLUP

and from 0.68 to 0.87 for CGM-WGP (Fig. 4a). Thus, the prediction accuracy for the CGM-WGP was restricted to a narrower range than for GBLUP. The prediction accuracy for GBLUP could drop to lower values than for CGM-WGP but also could achieve higher values.

DISCUSSION

Following the simulation study designed to demonstrate proof of concept for the CGM-WGP methodology (Technow et al., 2015), the present study reports the first application of the CGM-WGP to an empirical MET data set. Unlike in the simulation scenarios used for the proof-of-concept study, for the empirical data set considered here, the true genetic and physiological bases of the observed grain yield variation in environments FS1 and FS2 are unknown. Thus, we have no measure of the upper limits of prediction accuracy that can be achieved. Therefore, in this reduction to practice study, we focused on describing the important steps required to implement the CGM-WGP to an empirical data set, the average prediction accuracy that was achieved by CGM-WGP for the empirical example, the variation for prediction accuracy among different replicates of the CGM-WGP that resulted from different sample sets of the DH entries allocated to the estimation and test sets of entries, and we also used the prediction accuracy achieved by GBLUP as a benchmark. The key results from the application to the empirical data set are first summarized and some differences between the empirical study reported here and the simulation study by Technow et al. (2015) are discussed. Then, based on the combined results from the proof-of-concept study (Technow et al., 2015) and the empirical study reported here, we identify and discuss areas for investigation to further develop the CGM-WGP methodology for application in breeding.

Positive prediction accuracy was consistently achieved for grain yield by applying the CGM-WGP methodology to the selected maize cross in the two drought environments. For this first empirical implementation the average prediction accuracy of the CGM-WGP was similar to that achieved by GBLUP. However, there were differences between the CGM-WGP and GBLUP methods for specific cases that justify further investigation, discussed further below.

Beyond the necessary computational infrastructure (Technow et al., 2015), key components that were required to implement the CGM-WGP methodology to the empirical data set included definition of a suitable CGM (e.g., Fig. 2), identification of a suitable set of traits to model genetic variation for yield in the reference population, and definition of suitable parameters for the priors of the traits. Further, in the case of the drought empirical data set considered here, additional requirements included high-quality characterization of the soil, weather, and management conditions in each environment to enable the running of the CGM; high-quality management of irrigation

to ensure appropriate coincidence of the peak water deficit with the flowering window of the DH entries (Fig. 1); and high-quality phenotyping of all traits. Based on experience gained from multiple years of drought phenotyping of elite maize hybrids at the Viluco station, we had access to the necessary information required to apply the CGM-WGP methodology to the chosen MET data set. The reference genetic population chosen for this empirical study represents a case where a sample of DH lines was created from a cross between two parents that were selected on the basis of contrasting breeding values for drought performance. This situation is representative of many other crosses that are generated in the course of conducting a maize breeding program designed with the objective of developing new hybrids for the western region of the US Corn Belt. Therefore, the positive prediction accuracy results that were obtained from the application of the CGM-WGP to this cross for the two drought environments are encouraging for broader application of the methodology to other similar types of breeding crosses.

The prediction accuracy advantage of CGM-WGP over GBLUP that was reported for the proof-of-concept example by Technow et al. (2015) was not realized for the empirical example considered in the present study. In the case of the proof-of-concept study, the simulation scenario was designed to create strong levels of nonlinear physiological epistasis and crossover GEI that impacted grain yield. In the case of the empirical example reported here, the levels of physiological epistasis among the studied traits and their influence on yield is unknown but expected to be less than in the simulation study, and the level of GEI was quantified to be less than for the simulation study. Another key difference is that for the case of the simulation study, the CGM was assumed to capture all relevant physiological processes, environmental parameters, and their interactions that determine yield in the studied environments. This situation was not assumed for the empirical study. Thus, the simulated data set studied by Technow et al. (2015) is likely to represent the best-case scenario for the strengths of the CGM-WGP, whereas the empirical data set is expected to be suitable for successful application of the GBLUP method. Thus, to achieve an average prediction accuracy for application of the CGM-WGP to the empirical data set that was similar to that achieved for GBLUP is encouraging.

While the average prediction accuracy was similar between CGM-WGP and GBLUP (Table 2), there were important differences in the prediction accuracy distributions across replicates (Fig. 4). One possible reason for the different distributions could be that the properties of the two methods resulted in selection of different genetic models for the target trait grain yield. For the case of the CGM-WGP method, the prediction accuracy that can be achieved is conditioned by the component physiological traits that are implemented as sources of variation within

the CGM and the specification of the trait priors. In cases where important component traits that are influential on the expressed yield variation are not included as sources of genetic variation within the CGM, we can anticipate this would limit the upper levels of prediction accuracy that could be achieved by the CGM-WGP method. Equally, if traits important for explaining the yield variation are included in the CGM, this can be anticipated to reduce the likelihood of selecting inadequate genetic models that result in low prediction accuracy. No such restrictions to trait contributions to the selected genetic model are imposed with the GBLUP methodology. This potential difference between the CGM-WGP and GBLUP methods requires further investigation. One approach to investigate the hypothesized difference would be to conduct a series of simulation experiments that allow relationships to be present within the CGM used to generate the simulated data sets that are to be analyzed that are then removed from the version of the CGM for application of the CGM-WGP.

A key consideration for successful application of the CGM-WGP methodology is availability of a suitable CGM that can be used to incorporate the physiological effects and associated environmental responses of genetic variation for traits that influence genetic variation for yield. Ultimately, the CGM-WGP methodology can only be as good as the CGM used to represent the genetic variation for the ultimate trait yield. Other limitations will be imposed by the quality of the data sets and the data inputs for the CGM. To emphasize this point, we discussed in some detail important experience obtained while performing the characterization of the soil inputs for the CGM. The soil input data for the CGM provided a real example, not revealed by simulation studies, where incorrect specification of important environmental inputs or physiological processes compromised the application of the CGM-WGP methodology. Further, another important consideration for successful application is the decision on which traits should be used within the CGM to represent the genetic variation for yield. In the case of the empirical study reported here, the selection of the five component traits—TLN, AM, MEB, VPDB, and TUS—and the incorporation of their physiological effects within the CGM (Fig. 2) enabled positive prediction accuracy for grain yield in the drought environments. Other trait choices could have been made, and the results of the present study suggest other traits should be investigated to improve the prediction accuracy of the CGM-WGP. Also, it should be understood that to broaden the application of the CGM-WGP methodology to a wider range of environments and different GEI scenarios, additional traits will have to be included into the CGM for application of CGM-WGP.

As a first approximation, the CGM-WGP methodology can be viewed as a special case of a broader neural network approach for extending additive WGP methods to deal with the effects of nonadditivity in the trait genetic

model. Both can be represented as a network diagram or a graph (e.g., Fig. 2). In the neural network framework, the intermediate nodes would correspond to the component physiological traits. In the case of the CGM-WGP method, the mapping from SNPs to intermediate nodes to the output trait yield is defined by the CGM instead of arbitrarily derived from the data at hand as in the case of the neural network. An important difference between the neural network and the CGM is that there are functions in the biological network that the neural network method could suppress, while this is not possible in a CGM. In an unsupervised training exercise, the neural network may assign parameter values to a particular relation that suppresses the biological function altogether, for example, the value of radiation use efficiency could be zero. The CGM supervises the network grounded in biological principles and provides constraints that could be absent if the model behaved as a neural network. The fact that the CGM-WGP correctly predicts the mean and the range of yield in the application environment, while parameterized using data from the estimation environment, and was sensitive to the initial misspecification of soil parameters shows that the physiological and environmental information is being used by the CGM in the CGM-WGP methodology. This example demonstrates that the CGM does not behave entirely as an unsupervised neural network.

There are opportunities to improve the CGM-WGP methodology over the first implementation considered here. Areas for further research include the following:

- With the current CGM-WGP implementation applied in the present study, conduct a survey of additional crosses that have undergone comparable drought testing to that reported here. This will build a broader germplasm diversity perspective of the prediction accuracy that can be expected from the current implementation.
- Broaden the range of environmental conditions included in the MET to sample additional levels of GEI. In particular, focus on GEI between stress and favorable environmental conditions. As discussed above, the relatively low levels of GEI contributing to changes in rank of the DH entries for yield in the chosen MET were expected to have assisted successful application of the GBLUP method. We hypothesize that as the levels of GEI increase and there is an increase in the occurrence of changes in ranks of the entries there will be an increasing frequency of cases where the CGM-WGP methodology will have superior prediction accuracy over GBLUP. This advantage of the CGM-WGP methodology over GBLUP was observed in the simulation study reported by Technow et al. (2015).

- Extend the ABC algorithm to enable the use of information from multiple environments within the estimation step.
- Extend the ABC algorithm to enable the use of phenotypes available on multiple traits in the estimation step.
- Extend the ABC algorithm to allow both multiple-environment and multitrait data to be explicitly used within the estimation step. The broader range of environments will enable improved parameterizations of the functions describing trait norms of reactions embedded in the CGM. We hypothesize that the accuracy of CGM-WGP will improve with availability of algorithms to leverage the expanded datasets.
- Investigate the scope for improving prediction accuracy through alternative design of estimation data sets.
- Refine the version of the CGM used here to include additional trait variation that is known to exist in elite maize crosses that was not considered in the present study.
- Extend the genetic component of the CGM to incorporate a polygenic component to capture genetic signal that was not captured by the embedded equations for traits in the CGM. The imperfect correlation between the yield predicted by CGM-WGP and GBLUP (Fig. 5) indicates that there is an opportunity to combine noncommon components of the genetic models identified by both approaches to further increase prediction accuracy (Cooper et al., 2005, Messina et al., 2011).
- Following consideration of the outcomes of the research areas described above, compare prediction accuracy achieved by CGM-WGP to other WGP methods for a wide range of situations relevant to breeding programs.

As with the initial stages of any new methodology to be used in plant breeding, many limitations of the current implementation of CGM-WGP can be identified. However, the combined results obtained from the proof-of-concept study (Technow et al., 2015) and the reduction to practice for the empirical study reported here are encouraging. The CGM-WGP methodology opens up new possibilities for tackling GEI in plant breeding. Further, we have identified key areas of research that will enable enhancements to the current implementation and broaden its utility to a wider range of applications in breeding programs than those that have been considered to date.

Whole-genome prediction is now being used routinely as an integrated component of commercial maize breeding programs (Cooper et al., 2014a). Combining a CGM with WGP into an integrated prediction method (CGM-WGP) has interesting implications for future opportunities to parsimoniously combine crop physiology and plant breeding to tackle the genetic improvement of complex traits such as grain yield. There is much literature advocating how crop physiology could be used to improve the efficiency of plant breeding. While many of the arguments are appealing, there are however few convincing examples where crop physiology has identified a novel path to crop genetic improvement for grain yield that would not have been achieved otherwise by direct selection on grain yield in the target environments. Critical review supports the view that the sustained long-term genetic improvements of grain yield by plant breeding have largely been driven by plant breeders playing a numbers game. Plant and crop physiology research has helped to explain how trait variation contributed to these long term yield gains and has identified putative sources of trait genetic variation that could be important for further genetic improvement of yield. Many trait physiology studies have identified screens for the identified traits. Physiologists have advocated the use of such screens by breeders as part of the selection processes used in breeding programs. However, while many of these screens for physiological traits have been studied, few, if any, have been scaled and routinely adopted by breeders as integrated components of their breeding programs. The CGM-WGP methodology provides an opportunity to overcome some of the limitations that are associated with developing and applying additional physiological trait screening within a breeding program context. An important feature of the CGM-WGP methodology is that the additional traits that are chosen to vary as putative sources of genetic variation for grain yield within the framework of the CGM (e.g., TLN, AM, MEB, VPDB, and TUS in this study) are treated as latent variables in the CGM-WGP application to predict grain yield variation. Thus, in the implementation reported here, the physiological traits are not directly measured on any entries in any environments of the MET. The only trait that is directly measured in the MET is grain yield. The potential conditional influences of genetic variation for the traits on the grain yield variation in the MET are assessed through the formal quantitative relationships among the traits encoded in the equations of the CGM. The plausible estimated genetic models for the traits, and thus their estimated contributions to grain yield in the MET, out of the many potential genetic models, are identified by applying the ABC algorithm within the estimation step to sort the samples of genetic models drawn from the priors and to build the approximate posterior distribution of parameter vector θ . In addition to the direct application

for prediction, as discussed here, the parameter estimates retained in the posterior joint parameter vector θ can also be investigated to determine how the different traits are predicted to contribute to the grain yield variation within the environments of the MET. Thus, through focusing physiological research on developing an appropriate CGM and building an understanding of which traits vary for different sets of germplasm, the CGM-WGP methodology provides a path for integrating our physiological understanding of trait variation in elite germplasm into the core selection steps of a breeding program without the cost and complexity associated with measuring all of the relevant traits on all of the entries within all of the environments sampled in METs. This potential application of the CGM-WGP methodology, in combination with the areas identified for further research, open up new opportunities for a highly productive multidisciplinary approach to the genetic improvement of crops.

CONCLUSIONS

The results of the empirical example reported here demonstrate that the CGM-WGP methodology introduced by Technow et al. (2015) can be applied to a breeding program MET to generate useful predictions for yield. For the CGM-WGP implementation used here, the average prediction accuracy was similar to that achieved applying GBLUP methodology, although the predictions differed, and consequently, in some cases, different DH entries would have been selected by both methods. Further research is necessary to determine the scope to refine the CGM-WGP methodology to improve prediction accuracy and realize the opportunities for greater prediction accuracy that were identified in the proof-of-concept study by Technow et al. (2015). A requirement for implementation of the CGM-WGP methodology is availability of a suitable CGM to provide the biological framework for modeling trait genetic variation that contributes to the observed genetic variation for yield or other traits of interest. This requirement is different than for other WGP methods. However, most other WGP methods have limited ability to explicitly deal with nonadditive genetic effects such as physiological epistasis and GEI that are often associated with complex traits such as yield. Therefore, the cost and effort involved in developing an appropriate CGM for specific breeding program objectives has to be weighed against the importance of GEI and other sources of non-additivity. For grain yield of maize in the US Corn Belt, there are many situations where having additional capacity to deal with GEI for grain yield and improve prediction accuracy would justify such investment.

Another significant motivation for investing in the development of CGM-WGP is to open up new opportunities for integrating the large body of knowledge from plant and crop physiology research within plant breeding

methodology (Jackson et al., 1996; Hammer et al., 2006). Much has been advocated and discussed on this topic with limited practical progress to date. Organizing our efforts to extend physiological and agronomic research beyond the traditional outcomes of trait phenotyping and envirotyping would enable the integration of the trait–environmental results into an advanced CGM. The CGM would provide the link between trait and yield variation across the environmental conditions relevant to breeding programs (Messina et al., 2015). Combined with the availability of the CGM–WGP methodology, this provides a practical multidisciplinary path for the integration of physiology, agronomy, and breeding that has long been sought.

References

- Bernardo, R. 2008. Molecular markers and selection for complex traits in plants: Learning from the last 20 years. *Crop Sci.* 48:1649–1664. doi:10.2135/cropsci2008.03.0131
- Boer, M.P., D. Wright, L. Feng, D.W. Podlich, L. Luo, M. Cooper, and F.A. van Eeuwijk. 2007. A mixed-model quantitative trait loci (QTL) analysis for multiple-environment trial data using environmental covariables for QTL-by-environment interactions, with an example in maize. *Genetics* 177:1801–1813. doi:10.1534/genetics.107.071068
- Borras, L., G.A. Slafer, and M.E. Otegui. 2004. Seed dry weight response to source-sink manipulations in wheat, maize and soybean: A quantitative reappraisal. *Field Crops Res.* 86:131–146. doi:10.1016/j.fcr.2003.08.002
- Chapman, S.C., M. Cooper, and G.L. Hammer. 2002. Using crop simulation to generate genotype by environment interaction effects for sorghum in water-limited environments. *Aust. J. Agric. Res.* 53:379–389. doi:10.1071/AR01070
- Chapman, S., M. Cooper, D. Podlich, and G. Hammer. 2003. Evaluating plant breeding strategies by simulating gene action and dryland environment effects. *Agron. J.* 95:99–113. doi:10.2134/agronj2003.0099
- Chenu, K., S.C. Chapman, F. Tardieu, G. McLean, C. Welcker, and G.L. Hammer. 2009. Simulating the yield impacts of organ-level quantitative trait loci associated with drought response in maize: A “gene-to-phenotype” modeling approach. *Genetics* 183:1507–1523. doi:10.1534/genetics.109.105429
- Cooper, M., S.C. Chapman, D.W. Podlich, and G.L. Hammer. 2002. The GP problem: Quantifying gene-to-phenotype relationships. *In Silico Biol.* 2:151–164.
- Cooper, M., C. Gho, R. Leafgren, T. Tang, and C. Messina. 2014b. Breeding drought-tolerant maize hybrids for the US corn-belt: Discovery to product. *J. Exp. Bot.* 65:6191–6204. doi:10.1093/jxb/eru064
- Cooper, M., and G.L. Hammer, editors. 1996. *Plant adaptation and crop improvement*. CAB International, Wallingford, UK.
- Cooper, M., C.D. Messina, D. Podlich, L.R. Totir, A. Baumgarten, N.J. Hausmann, D. Wright, and G. Graham. 2014a. Predicting the future of plant breeding: Complementing empirical evaluation with genetic prediction. *Crop Pasture Sci.* 65:311–336. doi:10.1071/CP14007
- Cooper, M., D.W. Podlich, and O.S. Smith. 2005. Gene-to-phenotype models and complex trait genetics. *Aust. J. Agric. Res.* 56:895–918. doi:10.1071/AR05154
- Dardanelli, J.L., O.A. Bachmeier, R. Sereno, and R. Gil. 1997. Rooting depth and soil water extraction patterns of different crops in a silty loam Haplustoll. *Field Crops Res.* 54:29–38. doi:10.1016/S0378-4290(97)00017-8
- Edmeades, G.O., J. Bolaños, M. Hernández, and S. Bello. 1993. Causes for silk delay in a lowland tropical maize population. *Crop Sci.* 33:1029–1035. doi:10.2135/cropsci1993.0011183X003300050031x
- Gaffney, J., J. Schussler, C. Löffler, W. Cai, S. Paszkiewicz, C. Messina, J. Groetke, J. Keaschall, and M. Cooper. 2015. Industry-scale evaluation of maize hybrids selected for increased yield in drought-stress conditions of the US corn belt. *Crop Sci.* 55:1608–1618. doi:10.2135/cropsci2014.09.0654
- Gholipour, M., S. Choudhary, T.R. Sinclair, C.D. Messina, and M. Cooper. 2013. Transpiration response of maize hybrids to atmospheric vapor pressure deficit. *J. Agron. Crop Sci.* 199:155–160. doi:10.1111/jac.12010
- Habier, D., R. Fernando, K. Kizilkaya, and D. Garrick. 2011. Extension of the Bayesian alphabet for genomic selection. *BMC Bioinformatics* 12:186. doi:10.1186/1471-2105-12-186
- Hammer, G., M. Cooper, F. Tardieu, S. Welch, B. Walsh, F. van Eeuwijk, S. Chapman, and D. Podlich. 2006. Models for navigating biological complexity in breeding improved crop plants. *Trends Plant Sci.* 11:587–593. doi:10.1016/j.tplants.2006.10.006
- Heffner, E.L., M.E. Sorrells, and J.L. Jannink. 2009. Genomic selection for crop improvement. *Crop Sci.* 49:1–12. doi:10.2135/cropsci2008.08.0512
- Heslot, N., D. Akdemir, M.E. Sorrells, and J.L. Jannink. 2014. Integrating environmental covariates and crop modeling into the genomic selection framework to predict genotype by environment interactions. *Theor. Appl. Genet.* 127:463–480. doi:10.1007/s00122-013-2231-5
- Holzworth, D.P., N.I. Huth, P.G. deVoil, E.J. Zurcher, N.I. Herrmann, G. McLean, et al. 2014. APSIM— evolution towards a new generation of agricultural systems simulation. *Environ. Model. Software* 62:327–350. doi:10.1016/j.envsoft.2014.07.009
- Jackson, P., M. Robertson, M. Cooper, and G. Hammer. 1996. The role of physiological understanding in plant breeding; from a breeding perspective. *Field Crops Res.* 49:11–37. doi:10.1016/S0378-4290(96)01012-X
- Keating, B.A., P.S. Carberry, G.L. Hammer, M.E. Probert, M.J. Robertson, D. Holzworth, et al. 2003. An overview of APSIM, a model designed for farming systems simulation. *Eur. J. Agron.* 18:267–288. doi:10.1016/S1161-0301(02)00108-9
- Marjoram, P., A. Zubair, and S.V. Nuzhdin. 2014. Post-GWAS: Where next? More samples, more SNPs or more biology? *Heredity* 112:79–88. doi:10.1038/hdy.2013.52
- Melchinger, A.E., H.F. Utz, and C.C. Schön. 2004. QTL analyses of complex traits with cross validation, bootstrapping and other biometric methods. *Euphytica* 137:1–11. doi:10.1023/B:EUPH.0000040498.48379.68
- Messina, C.D., D. Podlich, Z. Dong, M. Samples, and M. Cooper. 2011. Yield-Trait performance landscapes: From theory to application in breeding maize for drought tolerance. *J. Exp. Bot.* 62:855–868. doi:10.1093/jxb/erq329
- Messina, C.D., T.R. Sinclair, G.L. Hammer, D. Curran, J. Thompson, Z. Oler, C. Gho, and M. Cooper. 2015. Limited-transpiration trait may increase maize drought tolerance in the US corn belt. *Agron. J.* 107:1978–1986. doi:10.2134/agronj15.0016
- Meuwissen, T.H.E., B.J. Hayes, and M.E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157:1819–1829.
- Muchow, R.C., and T.R. Sinclair. 1991. Water deficit effects on maize yields modeled under current and “greenhouse” climates.

- Agron. J. 83:1052–1059. doi:10.2134/agronj1991.00021962008300060023x
- Muchow, R.C., T.R. Sinclair, and J.M. Bennett. 1990. Temperature and solar radiation effects on potential maize yield across locations. *Agron. J.* 82:338–343. doi:10.2134/agronj1990.00021962008200020033x
- R Core Development Team. 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Ritchie, J.T. 1998. Soil water balance and plant water stress. In: G.Y. Tsuji, G. Hoogenboom, and P.K. Thornton, editors, Understanding options for agricultural production. Kluwer Academic Publishers, Dordrecht, The Netherlands. p. 41–54.
- Robertson, M.J., S. Fukai, M.M. Ludlow, and G.L. Hammer. 1993. Water extraction by grain sorghum in a sub-humid environment. I. Analysis of the water extraction pattern. *Field Crops Res.* 33:81–97. doi:10.1016/0378-4290(93)90095-5
- Sinclair, T.R., J.M. Bennett, and R.C. Muchow. 1990. Relative sensitivity of grain yield and biomass accumulation to drought in field-grown maize. *Crop Sci.* 30:690–693. doi:10.2135/cropsci1990.0011183X003000030043x
- Smith, J.S.C., M. Cooper, and J. Gogerty. C. Löffler, D. Borcherting, and K. Wright. 2014. Maize (*Zea mays L.*). In: J.S.C. Smith, B. Carver, B.W. Diers, and J.E. Specht, editors, Genetic gains of major US field crops. ASA, CSSA, and SSSA, Madison, WI.
- Tanner, C.B., and T.R. Sinclair. 1983. Efficient water use in crop production: Research or re-search? In: H.M. Taylor and W.R. Jordan, editors, Limitations to efficient water use in crop production. ASA, CSSA, and SSSA, Madison, WI. p. 1–27.
- Technow, F., C.D. Messina, L.R. Totir, and M. Cooper. 2015. Integrating crop growth models with whole genome prediction through approximate Bayesian computation. *PLoS ONE* 10:E0130855. doi:10.1371/journal.pone.0130855
- van Eeuwijk, F.A., M. Boer, L.R. Totir, M. Bink, D. Wright, C.R. Winkler, D. Podlich, K. Boldman, A. Baumgarten, M. Smalley, M. Arbelbide, C.J.F. ter Braak, and M. Cooper. 2010. Mixed model approaches for the identification of QTLs within a maize hybrid breeding program. *Theor. Appl. Genet.* 120:429–440. doi:10.1007/s00122-009-1205-0
- Westgate, M.E., and J.S. Boyer. 1985. Osmotic adjustment and the inhibition of leaf, root, stem and silk growth at low water potentials in maize. *Planta* 164:540–549. doi:10.1007/BF00395973