[redacted]

form a re-
PE and that the MET
ying and estimating G × E
decide upon a strategy on how to

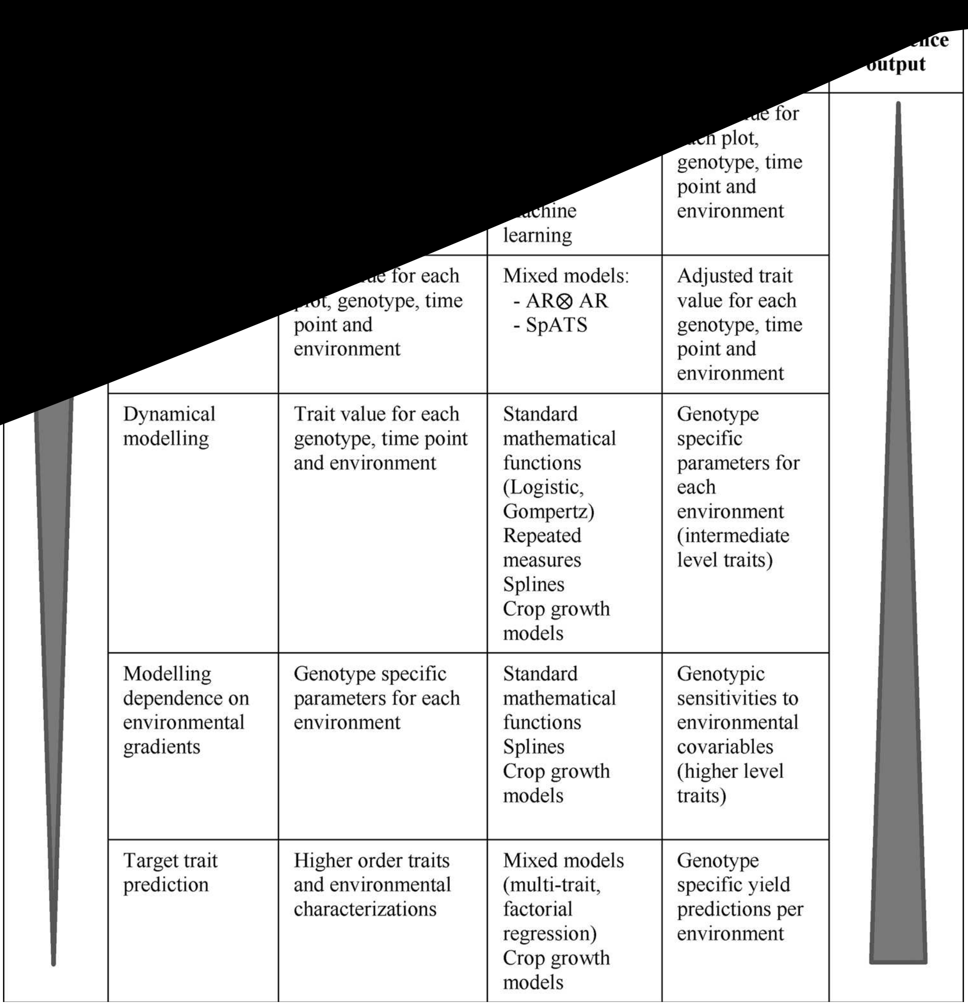
it was recognized that expecting METs to contain
sets of environmental conditions was not an optimal
strategy for developing adapted genotypes to defined biotic and abiotic stress conditions. Managed stress trials were introduced to complement classical location by year METs [8–10]. Furthermore, in many crops, physiology and physics driven measurement protocols were developed in attempts to predict the behaviour of genotypes under abiotic stress conditions [11]. Therefore, many phenotyping techniques were being used to help the breeding process attain genotypes with superior yields.

However, phenotyping as a separate discipline within plant biology is a new phenomenon. The rapid development of new phenotyping technologies has led to an enormous increase in the types and amounts of phenotypic data that can be monitored and registered. All levels of biological organisation have become accessible from within the cell up to the level of the ecosystem, and at spatial and temporal resolutions that were unthinkable before, with whole new types of processes being open to study [12,13]. New phenotyping devices occur that can be classified as: 1) high precision platforms, which are not necessarily high throughput, for omics, anatomical and functional traits at cell and organ level, covering short time scales [14,15]; 2) whole-plant field and controlled environment platforms for plant canopy traits [8,16,17]; 3) networks of field experiments (METs), with high throughput phenotyping techniques across environmental ranges [18,19]. The diversity and complexity of new phenotypic data types raises the questions of how to use this information efficiently to increase biological understanding of plant systems and of how to improve the accuracy of prediction for complex traits like yield. Should we copy the genomic prediction paradigm and simply include all phenotypic data in our prediction models without any feature selection? Or, will we do better by selecting specific features from our phenotyping information, *i.e.*, identify those traits that are relevant for better adaptation, as proposed by [12,20]?

In this paper, we will look at ways to incorporate new phenotyping data into predictive models for complex traits, where for simplicity of exposition we will focus on yield. We will describe various model classes, mainly statistical, but also crop growth models, that provide the means to evaluate and quantify the utility of new phenotyping information. We start with a presentation of a classification of phenotypic trait data according to the amount of modelling the data underwent. The next two sections, describe statistical models and crop growth models for predicting phenotypic traits from DNA profiles, environmental information, and additional phenotypic information, together with some illustrative examples. In these sections, attention will be given to some promising new G2P models and the role of modelling in relation to breeding strategies is considered. The last section presents some challenges ahead.

at the genetic
and genotypes precisely and
suitable G2P models that allow
genetic and environmental factors
and to predict the phenotype from genetic
inputs. For complex traits, G2P models will have to
explicitly the phenomenon of G × E [19,21,22]. Traditionally,
the target trait of the breeding process, yield, MET evaluations of breeding material play an important role, where the included germplasm can consist of segregating offspring populations as well as diversity panels. Some popular G2P models for MET data are the following. For genotype i in environment j (trial, year by location combination) a simple model is $y_{ij} = \mu_{ij} + \varepsilon_{ij}$, where stands for the expectation or mean of genotype i in environment j, while ε_{ij} represents a random residual. Roughly, in this model the expectation expresses the adaptation part of the phenotype of which we think we understand it well enough to predict it for given genotypes and environments, while the residual represents a combination of biological instability (to be modelled by genotype specific stability variances), not modelled phenotypic variation (polygenic variation), and experimental error. For the expectation, we try to find a model that makes the expectation a function of genotype specific sensitivities to environmental covariables, where, for notational convenience, we will ignore intercept terms depending on genotypes or environments: $\mu_{ij} = \mu(\beta_i; z_j)$, for a single environmental covariable z with a value z_j in environment j to which the genotype i has a sensitivity β_i. Genetic information can be inserted by making the genotypic sensitivities functions of marker profiles, $\boldsymbol{x_i}$, $\mu_{ij} = \mu(\beta(\boldsymbol{x}_i); z_j)$. When the phenotypic responses for genotypes to the environmental covariable are linear, *i.e.*, the reaction norm is linear, then $\mu_{ij} = \mu(\beta_i; z_j) = \beta_i z_j$. The formulation $\mu_{ij} = \mu(\boldsymbol{\beta}_i; z_j)$ allows the reaction norms also to be non-linear. In the latter case, the genotype can be characterized by multiple parameters as well, with $\boldsymbol{\beta}_i$ a vector of parameters determining the form of the reaction norm $\mu_{ij} = \mu(\boldsymbol{\beta}_i; z_j)$. For example, for a logistic dependence, $\mu_{ij} = \frac{\beta_{1i}}{1+\exp(-\beta_{2i}(z_j - z_0))}$, with β_{1i} the plateau for the reaction norm of genotype i, β_{2i} the growth rate, and z_0, or β_{3i}, the value for which the response reaches half of its maximum. More flexible formulations are possible by choosing a spline basis for the environmental covariable z, $s(z_j) = \sum_v b_{iv} h_v(z_j)$, with $h_v(z_j)$ one of the elements in a set of the B-spline basis functions relevant to environment j, and b_{iv} a genotype specific spline coefficient. Then, $\mu_{ij} = \mu(\beta_i; s(z_j)) = \mu(\Sigma_v b_{iv} h_v(z_j))$ [23,24].

In traditional phenotyping, most resources were spent on METs to estimate yield itself, the primary trait, but it was not uncommon either to try to estimate other traits than yield, secondary phenotypes, that could be used as genotype specific covariables in prediction models for yield. For example, $\mu_{ij}^f = \mu^f(\mu_i^s; z_j)$, with μ_{ij}^f the expectation for the focus trait (target trait, primary trait, highest order trait), $\mu^f(a; b)$ a function to generate the expectation for the focus trait with a genotypic input a and an environmental input b, μ_i^s is the genotype specific expectation for a secondary trait measured under controlled conditions and z_j is an environmental characterisation. This type of G2P model will also be suitable to incorporate new phenotyping information. However, the new secondary phenotyping information has higher spatial and temporal resolution than the more traditional secondary phenotyping information, it can come from many biological levels of organization, and the number of secondary phenotypes can be huge. We will give examples of such secondary phenotypes below. Therefore, it is not obvious which secondary phenotypes are useful for predicting yield and how the secondary phenotypes should enter existing G2P models. In statistical terms, within a high dimensional regression context, we have



	[illegible]	[illegible]	[illegible]	[illegible] output	
	[illegible]	[illegible]	[illegible] learning	[illegible] for [illegible] plot, genotype, time point and environment	
	[illegible]	[illegible] for each [illegible] genotype, time point and environment	Mixed models: - AR⊗ AR - SpATS	Adjusted trait value for each genotype, time point and environment	
	Dynamical modelling	Trait value for each genotype, time point and environment	Standard mathematical functions (Logistic, Gompertz) Repeated measures Splines Crop growth models	Genotype specific parameters for each environment (intermediate level traits)	
	Modelling dependence on environmental gradients	Genotype specific parameters for each environment	Standard mathematical functions Splines Crop growth models	Genotypic sensitivities to environmental covariables (higher level traits)	
	Target trait prediction	Higher order traits and environmental characterizations	Mixed models (multi-trait, factorial regression) Crop growth models	Genotype specific yield predictions per environment	

simultaneously a covariable subset selection problem and a functional form or transformation problem. Furthermore, our current G2P models may have to be modified to benefit from the large quantity of secondary phenotyping information.

To facilitate the development of a new G2P framework that can incorporate both large quantities of secondary phenotyping information as well as genomic information and environmental characterizations, we want to introduce a conceptual classification of new phenotyping traits. Advanced statistical and crop growth modelling methods will be required to generalize prediction and inference from the genotypes and environmental conditions included in METs to a wider set of genotypes representing the full collection of selection candidates and a wider set of environmental conditions representing the TPE. Table 1 shows how raw secondary phenotyping data are elaborated and converted into genotype specific covariables that enter G2P models for yield. This conversion consists of a number of discrete modelling steps that transform large numbers of basic and raw secondary phenotype data with low predictive power and utility into relatively few genotype specific parameters with high predictive power.

2.2. Feature extraction

Let us define a typical phenotypic observation or measurement as $y_{ijr}^{k}(t_{ijrm})$, for trait k ($k = 1. ..n_T$), genotype i ($i = 1. ..n_g$), in environment j ($j = 1. ..n_e$), for the r-th replicate ($r = 1. ..n_R$) in environment j, at time point t_{ijrm} ($m = 1... n_{M_{ijrk}}$). The observation $y_{ijr}^{k}(t_{ijrm})$ can be interpreted as a raw data point. Alternative modelling steps add value to the data and aggregate information from raw data points into model parameters. The raw data are null level traits, they are not aggregated over time or environmental gradients and no value addition by modelling has taken place. In a sequence of modelling efforts, the raw data are converted into parameters, higher level traits, that integrate data over time and environmental gradients to become predictors of complex traits like yield.

Modern phenotyping devices and techniques can produce large numbers of variables that each by themselves not necessarily relate in a one-to-one fashion to phenotypic traits of interest. So an important first step in phenotyping is to extract from large numbers of variables those features that are potentially useful for modelling biological and genetic processes and structures. Imaging has become an important secondary phenotyping technique that generates large amounts of information of which only a part will be relevant for phenotypic prediction. There are several methods to extract the features from images (e.g. Fig. 1). As images are usually quite noisy, pre-processing and feature-extraction methods are a key first step to produce phenotypic information with improved signal to noise ratios that at a later stage may prove to be

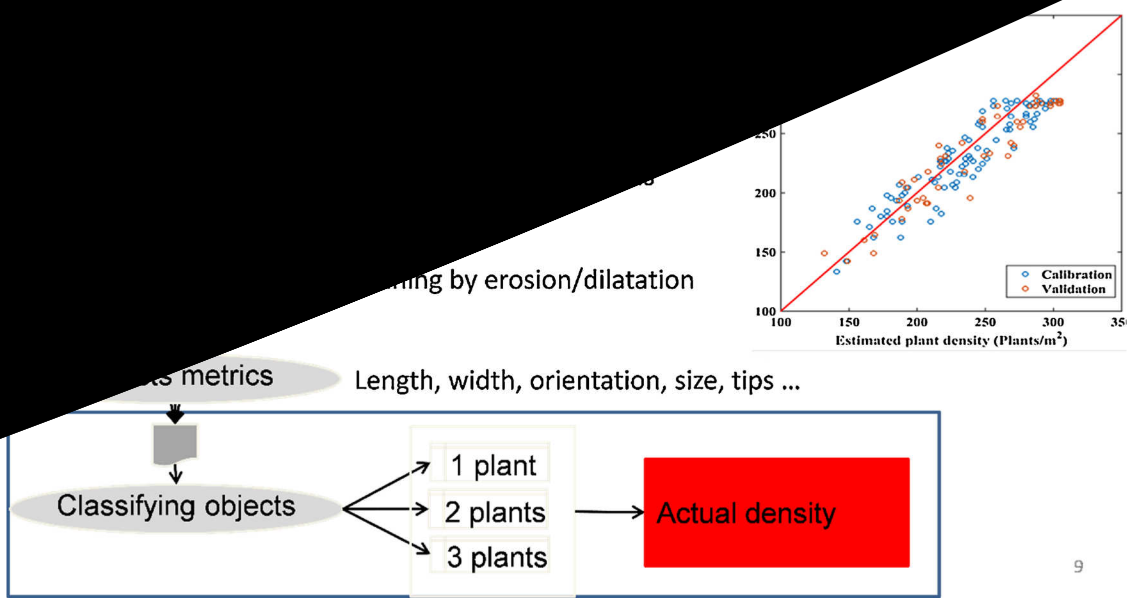


Fig. 1. Identification of young wheat plants to count plant density per unit area. After objects are isolated and sized from rows, the data are processed to predict intersecting objects and estimate their number (Liu et al. [26]).

useful for predicting a target trait [25,26].

Strategies to obtain trait information from images might involve segmentation [27,28], tracking procedures to deal with the problem of occlusions between plant organs [29] and using signal intensity at specific wavelengths (or indices derived from these intensities) as a proxy to plant traits. Common proxies using this approach are the normalized difference vegetation index (NDVI) to characterize biomass accumulation dynamics or canopy temperature to identify drought-tolerant genotypes [18,28,30,31]. An alternative is to identify morphological changes in sequential images (*e.g.* time to flowering) and segment specific organs (*e.g.* spikes) or organ surfaces (*e.g.* canopy drought stress [32]).

2.3. Correcting for experimental design factors and spatial variation

The extraction of features from the ample information generated by new phenotyping devices contains aspects of covariable subset selection, transformation and noise reduction. The result of this process is a set of secondary phenotypic traits for which it later on remains to be shown that they are useful for the prediction of primary traits. A first modelling step is now to estimate genotype and treatment means correcting for experimental design factors and spatial variation. Such an analysis is usually done per environment and per trait so that we can simplify the model to $y_{ir}(t_{im}) = \mu_i(t_{im}) + \varepsilon_{ir}(t_{im})$, with $\mu_i(t_{im})$ the genotype specific expectation for the trait in its dependence on time, while $\varepsilon_{ir}(t_{im})$ is a residual. When all genotypes are observed at the same times, we can write $y_{ir}(t) = \mu_i(t) + \varepsilon_{ir}(t)$, where we want to estimate $\mu_i(t)$. In principle, when correcting for design and spatial variation we should take into account the dependence in time between the observations on the same genotype and plant, but such an analysis is immediately non-trivial. Therefore, analyses are often performed per time point. Effectively we then fit the model $y_{ir} = \mu_i + \varepsilon_{ir}$ for each time point, which is equivalent to an analysis of variance, or a mixed model with autoregressive formulations for the dependence in row and column directions [33]. A recently proposed alternative is a mixed model with a two dimensional P-spline basis for spatial variation [34,35]. This spline approach has the advantage of avoiding the difficult model selection step. An example using the so-called SpATS model is given in (Fig. 2). The corresponding R-package SpATS is available on CRAN (https://cran.r-project.org/package=SpATS), and produces adjusted means (Best Linear Unbiased Predictions, BLUPs, or Best Linear Unbiased Estimates, BLUEs) as well as generalized heritabilities [37]. SpATS is flexible and user-friendly and performs comparably to more elaborate spatial models that require for each experiment a model selection process [34]. The absence of a model selection requirement allows a fully automatic implementation of this method for the analysis of field trial and platform experiments, especially convenient for the analysis of time series of phenotyping data obtained by the use of High Throughput Phenotyping (HTP) devices. As an example, Fig. 2 shows the Arabidopsis data coming from the Phenovator platform [38] for the light-use efficiency of PSII of 344 Arabidopsis accessions [39].

2.4. Dynamic modelling of spatially adjusted secondary phenotypes

Adjustment of genotypic and treatment means for experimental design and spatial variation was done per time point above. The next step is to model the genotypic means (or treatment means of any type) as adjusted for experimental design and spatial variation in relation to time, where these genotypic means are treated as if they were observations themselves, $y_i(t) = \mu_i(t) + \varepsilon_i(t)$, with $y_i(t)$ actually being the estimated genotypic mean of the previous modelling step (see Section 2.3), which we can write as $y_i(t) := \hat{\mu}_i$ at time *t*.

Models for $\mu_i(t)$ can be parametric, $\mu_i(t) = \mu(t; \boldsymbol{\theta}_i)$, with $\boldsymbol{\theta}_i$ a parameter vector for genotype *i* that describes the dependence of its dynamic phenotype $y_i(t)$ on time. Common functions to model trait dynamics are the logistic, Gompertz and exponential growth functions [40–43]. For a logistic relation to time the model would be $\mu_i(t) = \frac{\theta_{1i}}{1 + \exp(-\theta_{2i}(t - \theta_{3i}))}$ with θ_{1i} the upper asymptote, θ_{2i} the growth rate, and θ_{3i} the time at which the mid-point is reached [41,44].

Fig. 3 illustrates the dynamics of leaf length of two genotypes that

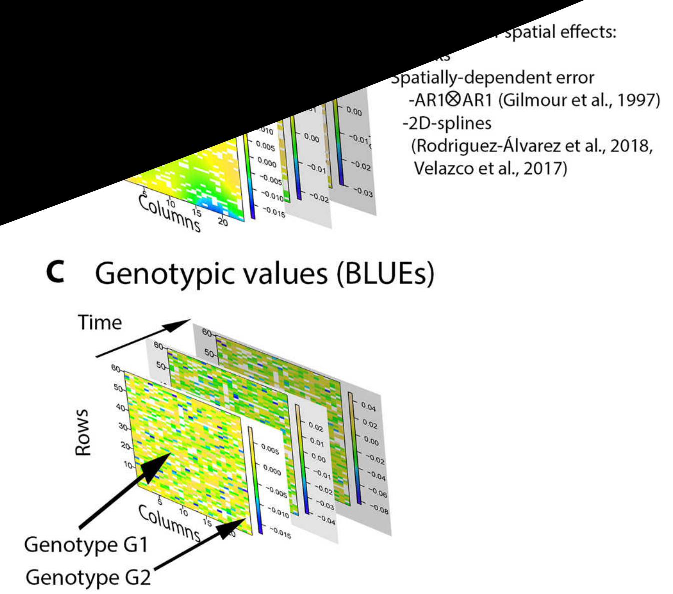


Fig. 2. Raw plot data (A) are adjusted for experimental design factors and spatial variation (B). The results are adjusted genotypic means (best linear unbiased estimators, BLUEs) or predictions (best linear unbiased predictors, BLUPs) per time point and environment (C).

have been phenotyped over time, with both genotypes grown under three levels of temperature and two levels of water stress. The leaf elongation dynamics are summarized by a logistic curve. Genotype G1 is sensitive to temperature (reflected in the three red curves running clearly apart from each other, Fig. 3A), but less sensitive to water stress (reflected in the two blue curves running close to each other). In contrast, genotype G2 is not sensitive to temperature, but quite sensitive to water stress. By fitting a logistic curve we can extract slope parameters (leaf elongation rate) that are at this stage both genotype and environment specific.

Parametric models for $\mu_i(t)$ may be too restrictive to capture the dynamics of traits like canopy temperature, leaf area or senescence. A more flexible alternative for modelling trait dynamics is offered by P-splines [23,24]. For example, P-splines were used to model haulm senescence in potato and identify QTLs (Quantitative Trait Locus) for the genotype-dependent slopes [45]. The use of splines can be extended to genomic prediction models for incorporating information from canopy temperature and NDVI over time [46].

In this context of dynamic models for phenotypes, we mention further the work on function-valued trait models, in which explicit mathematical functions describe trait dynamics during the growing season [47]. Genotype specific curve parameters as conditioned by the environment were modelled in relation to a genetic basis within a functional mapping framework by [42,48,49].

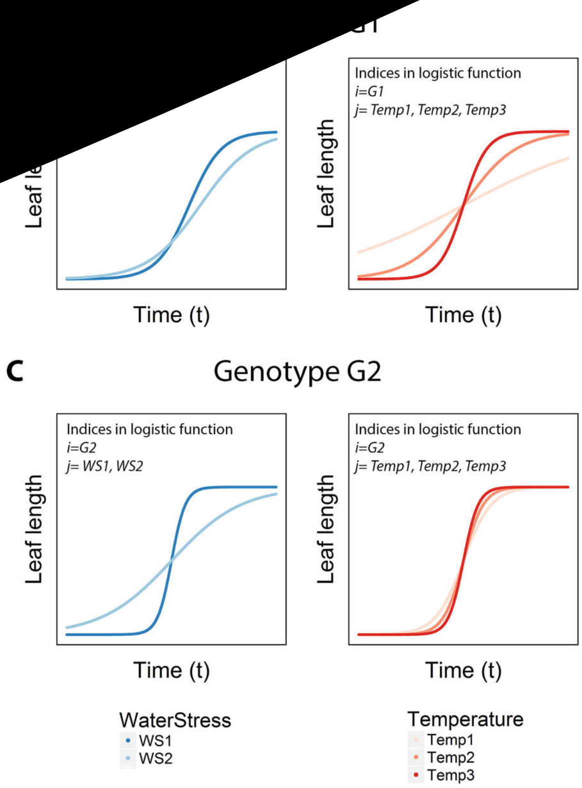


Fig. 3. Modelling the dynamics of phenotypic traits. Spatially adjusted genotypic means are modelled in their dependence on time for each environment. A logistic curve is fitted and genotype specific parameters describing the dynamics are extracted. (A) Model formulation for logistic curve. (B) Logistic curves as fitted for a genotype G1 on the dynamics of leaf length under two water stress conditions (WS1, WS2) and three temperature stress conditions (Temp1, Temp2, Temp3). (C) as (B) for a genotype G2. Genotype G1 in (B) shows a low sensitivity to water stress and a high sensitivity to temperature stress, whereas G2 in (C) shows a high sensitivity to water stress and a low sensitivity to temperature stress. (For interpretation of the references to colour in the text, the reader is referred to the web version of this article.)

Some phenotypic traits are repeatedly measured over time without the genotype specific expectation changing over time. This brings us in the classical repeated measurements situation, $y_i(t) = \mu_i + \varepsilon_i(t)$, where for the residual term a correlation structure should be defined that allows for deceasing correlation between time points with increasing separation [50,51]. For example, repeated measures on NDVI were analysed by a multi-trait model with a special structure for the autocorrelation between successive measurements [46].

The curve parameters, θ_i, represent traits of a higher level or order than the initial time dependent trait $y_i(t)$. They represent an integration over time of this lower level trait. New phenotyping technologies facilitate the characterization of growth and development during the growing season. High temporal resolution of measurements is achieved for secondary traits that are at a lower biological organization level than target or focus traits: leaf and stem size, biomass and disease scores [52,53]. Modelling the dynamics of such lower and intermediate level traits fits into a strategy to replace large amounts of data by limited numbers of genotype specific parameters. These parameters are often rates, like leaf elongation rate. The estimates for the time dynamics parameters can once again form the starting point of a new round of

[redacted]

write,
to a set of
The function $\mu(a; b)$
specific expectation from a
ironment specific input b where, for
will ignore intercept terms depending on
onments. In its simplest version, a linear regression
environmental characterization, $y_{ij} = \beta_i z_j + \varepsilon_{ij}$.

he genotype specific sensitivities of curve parameters, $\boldsymbol{\beta}_i$, represent again a higher level of integration of phenotypic information. When we consider the dependence of curve parameters on environmental conditions, we effectively have integrated across temporal and environmental gradients. Examples of such high level traits are the physiological parameters that predict yield across environments [53–55]. In Fig. 4A and B, they would correspond to the sensitivities of leaf elongation rate (a higher order trait) to temperature and water stress. The sensitivities to multiple environmental covariables can be incorporated in a higher level prediction model for yield.

High precision platform and to some extent whole plant platform measurements aim at the phenotyping of higher-order traits that are expected to show a reduced complexity for the control of G × E because, as they represent sensitivities, their nature already embodies the genotypic response across environmental conditions. The information on the genetic and environmental controls of higher order traits is expected to be transferable to field conditions and to help in predicting

a range of environ-
of G × E and calculation
–59].

higher order traits in prediction models for the

cture of G × E observed for the target trait in the TPE and its underlying environmental drivers have a large influence on which phenotypic traits need to be estimated at various biological levels for a successful prediction of the target trait across environmental gradients. When we have an a priori idea of a statistical or physiological prediction model, the structure of such a model can guide us in which secondary traits to concentrate on for measurements at a phenotyping platform. As an illustration, we consider a crop growth model (CGM) that predicts a target trait from small sets of genotype specific inputs and environmental characterizations [60,61]. The inputs for the CGM were at the genotypic side, the physiological parameters total leaf number (TLN, a low level trait), area of largest leaf (AM, a low level trait), solar radiation use efficiency (RUE, a higher order trait), and thermal units to physiological maturity (MTU, an intermediate level trait). For the environmental side of the model, the important inputs were daily average temperature (Temp) and solar radiation (Rad). The structure of the CGM was: $y_{ij}^f = \mu^{CGM}(\boldsymbol{y}_i^s; \boldsymbol{z}_j) + \varepsilon_{ij}$, with y_{ij}^f the focus trait, yield, for genotype i in environment j, $\boldsymbol{y}_i^s$ the input vector of secondary phenotypes (TLN, AM, RUE, MTU), $\boldsymbol{z}_j$ the input vector of relevant environmental characterizations (Temp, Rad), and $\mu^{CGM}(a; b)$ the CGM function that converts genotype specific inputs and environmental inputs in to predictions for the target trait.

In the original study [61], only meteorological data from the weather station Champaign (Illinois) were used to estimate predictions $\hat{y}_{ij}^f$, whereas for additional calculations we used the information of another 19 weather stations in Illinois. After the integration over time and the meteorological characterizations by the CGM, the results of our

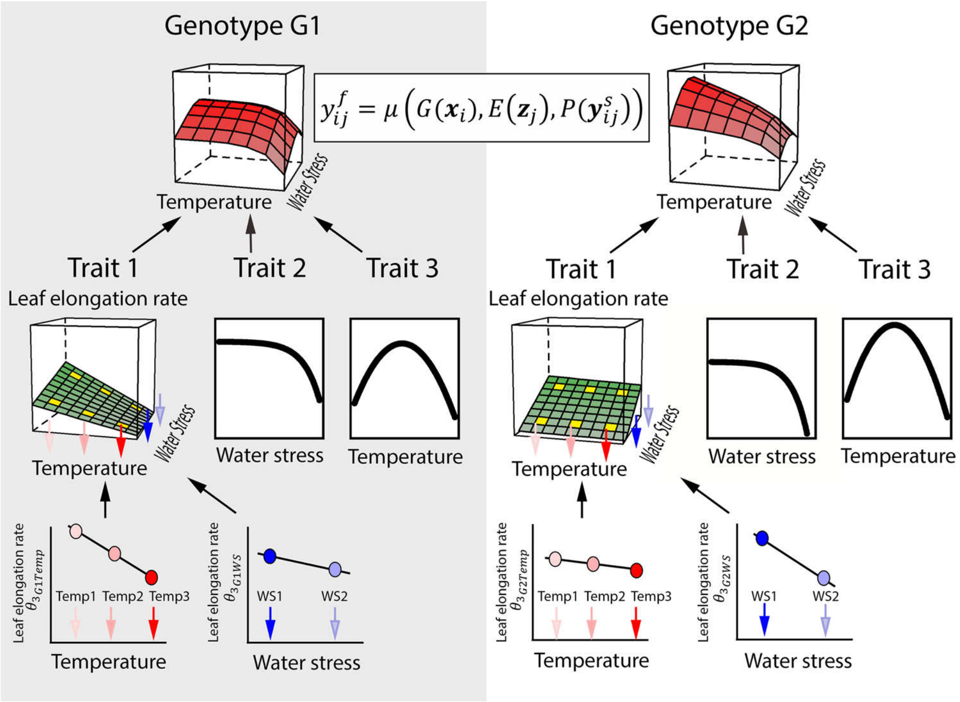


Fig. 4. Genotype specific parameters expressing dynamics of leaf length on time as extracted across different environmental conditions, see Fig. 3, are modelled in relation to water and temperature stress characterizations. For leaf elongation rate (trait 1), a two dimensional response surface is fitted. For other higher order traits (trait 2 and 3) one dimensional reaction norms are fitted as these traits depend in a simpler way on the environmental conditions. In the fitted response surface for leaf elongation rate the yellow points indicate conditions that were sampled in experiments. The higher order traits 1, 2 and 3 together determine a response surface for yield, the target trait. The prediction model for yield has trait 1, 2 and 3 as inputs as well as temperature and water stress characterizations. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

[redacted]

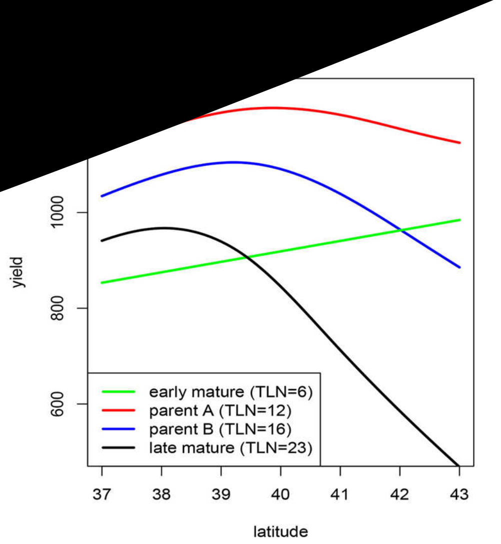


Fig. 5. A response surface and four response curves for yield in maize. (A) The response surface for a parent of a bi-parental cross obtained from crop growth simulations using inputs from 20 weather station locations in Illinois for season 2012, with a heritability of 0.85 per location. (B) For the longitude indicated by the dotted vertical red line in (A) four genotypes are compared for their dependence of yield on latitude; the two parents of a bi-parental cross, plus two offspring lines showing extreme values for the trait total leaf number (TLN). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

calculations show non-linear responses for the target trait, $\hat{y}_{ij}^f$, in relation to the environmental covariables latitude and longitude (Fig. 5A). The impression is raised that a suitable statistical model would be $y_{ij}^f = \mu^{Stat}\left(\boldsymbol{\beta}_i; \boldsymbol{z}_j\right) + \varepsilon_{ij}$, with $\boldsymbol{z}_j$ is latitude and longitude, and $\boldsymbol{\beta}_i$ the genotype specific curve parameters that describe the dependence on latitude and longitude. Both the CGM and statistical prediction model may produce acceptable accuracies (correlation between observed and predicted yield). The difference between both approaches may reside in the robustness of the predictions with respect to the specification of the training set of environmental conditions.

The simulated heritability in our example was 0.85 per location, as in the original study [61]. Fig. 5A shows the two-dimensional response surface for yield in 2012 for one of the parents of the simulated DH (Double Haploid) population [61], where a mixed model two-dimensional P-spline method was used to fit the spatial trend [62]. Fig. 5B shows the response curves for four genotypes as function of latitude, for predictions with the same longitude as Champaign. The (extremely) early mature genotype (TLN = 6) is performing best at the highest observed latitude, while the other three genotypes show optima at lower latitudes.

puts.
d possibilities for in-
ation. Within the framework
econdary phenotyping to prediction
sed in a number of ways. In later sections,
will be given of possible improvements by the
secondary phenotyping information into G2P models.
ation we envision for our modelling purposes is the following:
e have at our disposal a set of field experiments for a target trait that belong to a MET that is supposed to represent the TPE. The target trait is estimated only once, at harvest, and no dynamical information is available for the target trait. We have measurements with phenotyping tools in one or more trials of the MET as well as on possible phenotyping platforms. With respect to secondary phenotyping, we may want to compare different technologies as well as sampling and measurement schemes.

For statistical G2P models, and ignoring the time dependency of the focus or target trait, we can think of observations for genotype i in environment j,y_{ij}^f, and we want to predict yield from genetic and environmental information, as well as from secondary phenotyping information:

$$y_{ij}^f = \mu(G(\boldsymbol{x}_i), E(\boldsymbol{z}_j), P(\boldsymbol{y}_{ij}^s)) + \varepsilon_{ij}^f$$

in which $\mu(G(\boldsymbol{x}_i), E(\boldsymbol{z}_j), P(\boldsymbol{y}_{ij}^s))$ is a function that generates the expectation for the target phenotype on the basis of three arguments: 1) a genotype related function, *G(.)*, working on genotype specific genetic information, which can include marker information (SNPs, sequence profile), contained in the input vector $\boldsymbol{x}_i$; 2) an environment related function, *E(.)*, working on environment specific information contained in the vector $\boldsymbol{z}_j$ consisting of summaries for meteorological, soil and management variables; 3) a phenotype related function *P(.)* working on secondary phenotypic information, $\boldsymbol{y}_{ij}^s$, with $s \in S$, the set of secondary phenotypes, in which each phenotype can be genotype specific or, more often, can depend on a combination of genotype and environment (Figs. 3–5).

In many statistical-genetic G2P models, the functions *G(.)* and *E(.)* are simply identity functions and a dependence on higher order traits is absent, while the function $\mu(G(\boldsymbol{x}_i), E(\boldsymbol{z}_j), P(\boldsymbol{y}_{ij}^s))$ is a multiplication of genotypic and environmental information. For example, $\mu(G(\boldsymbol{x}_i), E(\boldsymbol{z}_j), P(\boldsymbol{y}_{ij}^s)) = \mu(G(\boldsymbol{x}_i), E(\boldsymbol{z}_j)) = \mu(x_i; z_j) = \theta x_i z_j$, with x_i a count for the number of minor alleles for a marker in a QTL or candidate gene (between 0 and 2), and z_j an environmental covariable like the average temperature in a critical developmental stage, and θ is the QTL allele substitution effect that scales with the environment [19,21,22,63–65]. Traditional G2P models need to be generalized to accommodate secondary phenotypic information. Various options will be discussed in the remainder of this paper.

Traditional statistical models for MET data do not contain dynamical aspects, developmental time is not explicitly present in such models. In contrast, CGMs have the dynamics of development as the core of their model formulation. One possible mathematical-statistical description of a CGM is as follows. Assume we consider yield as the focus trait, *f*, for genotype *i* in environment *j* at time T_j (harvest time in environment *j*, taken to be the same for all genotypes):

$$y_{ij}^f(T_j) = \int_0^{T_j} \mathfrak{F}\left[\left(\boldsymbol{y}^{s\in S_P}(\boldsymbol{x}_i)\right);\ \left(\boldsymbol{y}^{s\in S_I}(\boldsymbol{x}_i, \boldsymbol{z}(t_j))\right)\right] dt + \varepsilon_{ij}^f$$

In a crop growth model, yield depends on two sets of underlying or component traits: 1) physiological parameters that are functions of DNA variation, represented by a SNP vector, $\boldsymbol{x}_i$, that have no environmental and no time dependence, $(\boldsymbol{y}^{s\in S_P}(\boldsymbol{x}_i))$, with the secondary phenotypes, denoted by s, belonging to the set S_P set of physiological parameters; 2) intermediate yield

traits is de-

traction of such G2P

redict phenotypes for combi-

ironments from marker profiles and

ations.

tatistical models, for the estimation of parameters, a

can be defined on the discrepancy between observations

the primary trait for MET data representing the TPE and the predictions of the G2P model [75]. A loss function puts weights on the discrepancies $(y_{ij} - \hat{y}_{ij})$ during the estimation of model parameters. The choice of loss function determines the estimation and inference procedure including the model building. A well-known loss function is least squares, where parameters are estimated to minimize $(y_{ij} - \hat{y}_{ij})^2$. An alternative estimation and inference procedure is based on maximization of (residual) maximum likelihood [76]. Given a G2P model for a primary trait and an inference procedure we can investigate to which extent the insertion of secondary phenotyping information itself or features extracted from secondary information into the G2P model leads to a smaller residual variance or higher likelihood.

Another popular way of assessing the quality of secondary phenotyping information is by the correlation between observed and predicted primary trait values, *i.e.* corr $(y_{ij}; \hat{y}_{ij})$, where this correlation is often calculated as part of a cross validation process. Utility of phenotyping information can then be established by inspecting these correlations for different G2P models. This cross validation approach can be applied to both statistical and crop growth models.

For CGMs, genotype specific information on a number of physiological parameters may be required that can be difficult to obtain. Phenotyping platforms may then produce direct estimates for such parameters or approximations to such parameters. In such cases, the contribution of secondary phenotyping to genetic gain may be less straightforward to assess. In general, any correlation of secondary phenotypes to genotype specific inputs for G2P models will contribute to genetic gain, where higher level phenotypes will be more useful than lower level phenotypes because they are closer to the target trait, their phenotypic distance to the target trait is smaller [77]. Just like any correlation of secondary traits with G2P inputs will be useful, so will QTLs that are shared between secondary traits and G2P inputs.

In the remainder of this paper, with respect to G2P models to consider, we will focus on linear mixed models (LMMs) and crop growth models (CGMs), although we will also address Bayesian approaches. We will describe G2P models for a primary trait like yield and illustrate how these models help in assessing the usefulness of phenotyping strategies.

3. Statistical G2P models

3.1. A correlated response framework connecting measurements on plant platforms and field observations

Consider the simple case that a secondary trait is measured on a

use the platform es-

nvironmental sensitivity for

the TPE can be represented by a

g to a MET.

e platform and the field trait is a convenient

pply classical quantitative genetic theory on correlated

as a framework to assess the utility of phenotyping.

alistically, the single trait on the platform may be a higher order secondary trait, a function of platform traits like a genotype specific summary across multiple measuring times or environments, a principal component calculated from a set of platform traits, or a selection index calculated to maximize the selection response for a trait in the field/TPE. The single trait in the TPE can again be a function of a number of field traits, *i.e.* a summary statistic like a genotypic mean, a sensitivity to an environmental gradient, or a selection index calculated from the experiments in a MET.

Let us assume that the trait as measured on the platform will consist of an intercept, genetic part and an error part, $y_i^s = \mu_i^s + G_i^s + \varepsilon_i^s$, similarly so for the trait in the field/MET/TPE, $y_i^f = \mu_i^f + G_i^f + \varepsilon_i^f$. We assume that the TPE is known and the field experiments in the MET represent a random draw from the TPE, so experiments are exchangeable and G × E is a source of error variation, and no repeatable interactions can be identified. For the latter case, a regression approach is more suitable than a correlation approach.

We may want to select genotypes on superior performances in the platform response, anticipating that the correlated selection response for the field trait will be larger than when selecting directly on the field response itself. Using indirect selection response theory we can say that this approach makes sense when the genetic correlation between platform and field, $\rho^{(s,f)}$ is high, *i.e.*, the same genetic basis is involved, and the heritability on the platform, $h^{2,s} = \frac{V_{G^s}}{V_{y^s}} = \frac{V_{G^s}}{V_{G^s} + V_{\varepsilon^s}}$, is higher than the heritability in the field, $h^{2,f} = \frac{V_{G^f}}{V_{G^f} + V_{\varepsilon^f}}$. More precisely, when the condition is fulfilled that $\rho^{2,(s,f)} h^{2,s} > h^{2,f}$ it makes sense to select on the platform representation of the trait in place of the field trait itself. The heritability at the platform may be higher because the conditions at the platform can be better controlled or more replicates can be taken, hence the error variance, V_{ε^s}, is smaller than the error variance in the field, V_{ε^f}. Additionally, the genetic variance at the platform, V_{G^s}, can be increased by choosing environmental conditions that are more strongly discriminating between genotypes and it is larger than the genetic variance in the field, V_{G^f}. The genetic correlation between platform and field depends on the extent to which the conditions in the field induce the same genes or QTLs to be expressed as on the platform, where the sign of the QTL effects should coincide and the magnitude of the effects should be proportional between platform and field.

For a full evaluation of correlated responses *versus* direct responses, economic considerations for measurements on platforms and fields should be included as well [160]. In that case, a selection index with economic weights is recommended. For example, economic weighting coefficients were included for the combination of several primary traits for sugar cane (biomass yield, sugar and fibre content) [78,79]. The index can also consider traits (*e.g.* physiological measurements) that do not have a direct economic impact, but correlate with economically important traits. In sugarcane for example, it was found that across diverse genotypes, the secondary phenotype mid-season plant stomatal conductance was highly correlated with total biomass yield, and this trait could be used as a proxy during the earlier stages of selection when genotype numbers are high, and plots are small, such that biomass yield *per se* is not reliable [80]. (Although the secondary trait is here not

[redacted]

The size of the error variance is thus largely under control of the breeder. This is to a lesser extent the case for the value of the genetic variance and the correlation between platform and field. The magnitude of these genetic parameters will depend on the timing of measurement and the method of measurement on platform and field as well as on the environmental conditions up to the measurements [81].

3.2. Illustrative examples of a correlated response framework for high precision and whole-plant platforms to predict phenotypes in the field

Grain yield in maize depends on the ability of leaves and silks to maintain growth under fluctuating environmental conditions, especially under drought. Characterizing growth of leaves and silks under a range of environmental conditions in the field is a difficult task. Platforms have the advantages of facilitating more frequent and detailed measurements and also offer larger possibilities of controlling the environmental conditions than field experiments. For that reason, they are an interesting alternative to characterize relevant traits for drought adaptation in maize. For example, platform phenotypes have been used to estimate leaf elongation rate per unit of thermal time and the slope of leaf elongation to evaporative demand and soil water status [52]. QTLs of maximum leaf elongation rate on the platform co-located with QTLs of the anthesis-silking interval in well-watered fields, with alleles conferring high leaf elongation rate conferring a low anthesis-silking interval. The QTLs of the response of leaf elongation rate to water deficit at the platform co-located with QTLs of anthesis-silking interval in water deficit fields. For these QTLs, the allele conferring a larger anthesis-silking interval in the field (hence a reduced silk elongation rate) was also the one leading to a smaller leaf elongation rate on the platform. This suggested that common genetic mechanisms are shared between the growth of reproductive organs in the field and leaf growth on the platform. These conclusions were further supported by a metaQTL analysis in which QTLs for maximum leaf elongation rate on the platform co-located with QTLs involved in the growth of shoots, roots, but also reproductive organs in controlled conditions and fields [82,83]. Further, the sensitivity of leaf growth to soil water deficit at the phenotyping platform was related to the sensitivity of maize grain number to soil water deficit around flowering time in the field [84]. A more detailed characterization of the sensitivity of leaf growth to environmental conditions was demonstrated and a model to predict leaf area in the field, as a function of leaf length and width and their respective sensitivities to intercepted light and evaporative demand was proposed

[redacted]

external data set, a network of 15 field experiments. The model to predict leaf length and width resulted in an accurate prediction of individual leaf area in the whole field dataset ($R^2 = 0.62$). QTL allelic effects underlying leaf width and length were smaller in the platform than in the field, but they were clearly correlated between platform and field experiments. Thus, although there was G × E between platform and field experiments, the correlated QTL effects would still allow to use platform data to improve selection for leaf area in field conditions [57].

Another example of the use of indirect selection for field conditions based on traits measured in controlled conditions is the following: nine early vigour characteristics of wheat F2:4 plants grown in trays were evaluated to improve biomass production in the field for F2:6 plants [85]. Tray performance was a good predictor of field performance for leaf breadth and length, leaf area, and plant biomass. Genetic correlation between a trait measured on the trays and in the field was highest for early biomass ($r = 0.61$), suggesting that early biomass measured on the trays was a good predictor for early biomass in the field. Other traits measured on the trays that showed a moderate to high genetic correlation with field biomass were plant leaf area ($r = 0.59$), length of leaf 2 ($r = 0.43$), length of leaf 3 ($r = 0.41$) and mean leaf breadth ($r = 0.40$). The most promising for indirect selection were mean leaf breadth and breadth of leaf 2, with a relative gain from indirect selection that was 61 and 60% of the gain to be obtained from directly selecting for biomass in the field. Although plant leaf area on the tray showed a larger genetic correlation with field biomass than mean leaf breadth and breadth of leaf 2, it had a lower heritability ($H^2_{plant\ leaf\ area} = 0.30$, $H^2_{mean\ leaf\ breadth} = 0.82$, and $H^2_{breadth\ leaf\ 2} = 0.76$) reducing its potential to be used for indirect selection (relative gain from indirect selection for leaf area on the tray was 55%). This example illustrates the importance of jointly considering the magnitude of genetic correlations between traits measured on the platforms and those in the field and trait heritability to assess the potential of traits to aid selection for early biomass in the field. One aspect that would need to be further assessed in this study [85] is whether selection for early biomass is indeed correlated to yield at the end of the growing season. Only if the correlation between early biomass and yield is reasonably large, it would be advantageous to use early biomass measured in the trays as early selection trait. In the previous example, all phenotyping was done by hand, which is a time-consuming process. However, thanks to imaging technologies, the approach could be scaled up to a breeding programme. For example, the dynamics of early growth for individual plants grown in greenhouse trays could be characterized with multi-view images [86]. Examples presented above are summarized in Table 2.

3.3. A multi-trait prediction framework for yield using high throughput phenotyping information

HTP devices in the forms of sensors, drones, and unmanned aerial vehicles generate high dimensional secondary phenotypic data for experiments that are part of METs representing the TPE. The HTP information is used to approximate yield related traits and components

		cient	
[illegible]	[illegible]	the growth of shoots, roots, ve organs	[82,83]
[illegible]	[illegible]	Sensitivity of maize grain number to soil water deficit around flowering time	[84]
[illegible]	*t with genotype-specific ers*	Predict individual leaf area in a network of field trials.	[57]
[illegible]	*Correlation*	Allelic effects underlying leaf width and length	
xperiments, correlation of QTL effects allows the use of platform data for improving leaf area in field			
nd 3, leaf breadth	*Correlation and indirect selection*	Biomass	[85]
tween a platform and field trait needs to be considered together with the heritability for those traits to arrive at a valid assessment of ion on a platform for field conditions (early biomass)			

over the growing season. The HTP information is introduced in the form of additional traits alongside yield in multi-trait G2P models.

A multi-trait mixed model for genotype *i* and trait *k* can be defined as follows $y_i^k = \mu_i^k + G_i^k + \varepsilon_i^k$, with μ_i^k the fixed intercept for trait *k*, and G_i^k the random trait-specific genetic effect for genotype *i* and trait *k*, while ε_i^k is a residual. The random genetic effects G_i^k will have a multivariate normal matrix distribution with mean zero and variance-covariance matrix (VCOV) $\boldsymbol{\Sigma}$: $\{G_i^k\} \sim MVN(\mathbf{0}, \boldsymbol{\Sigma})$ [159]. The VCOV for the genetic effects, $\boldsymbol{\Sigma}$, has a special structure that can be interpreted as a product of a VCOV defined on the genotypes, $\boldsymbol{\Sigma}^G$, where pedigree and/or marker information determine correlations between genotypes, and another VCOV defining the genetic variances and correlations between traits, $\boldsymbol{\Sigma}^T$, where each trait has its unique genetic variance and each pair of traits has its unique genetic correlation. The above multi-trait model is customarily turned into a genomic prediction model by defining the matrix $\boldsymbol{\Sigma}^G$ in terms of similarities between marker profiles of genotypes [87].

The multi-trait model above can be generalized to multiple environments: $y_{ij}^k = \mu_{ij}^k + G_{ij}^k + \varepsilon_{ij}^k$ with intercept μ_{ij}^k and a random genetic effect G_{ij}^k and a residual ε_{ij}^k. The VCOV for the genetic effects has a multivariate normal distribution with zero mean and will have to represent the trait by environment variances and correlations. In the context of genomic prediction models for multi-environment data with G × E, VCOV structures for the environments have been described [88,89]. The VCOV for environments can be based on similarity in environmental characterization between environments, a generalization of the fixed factorial regression models, see next section, like in [89,90] and [91]. For the simultaneous modelling of VCOV structures for genotypes, environments and traits, see [92].

For the identification of multi-trait and multi-environment mixed models one can use log-likelihood ratio tests to compare different VCOV structures for G_i^k or G_{ij}^k (see [93]) where the differences relate to how to combine pedigree and marker information for the genotypes in $\boldsymbol{\Sigma}^G$ [94–96], which traits to include alongside yield in $\boldsymbol{\Sigma}^T$, and how to efficiently represent the relations between environments in a VCOV for environments, $\boldsymbol{\Sigma}^E$, and/or which environmental characterizations to use to calculate environmental correlations [89]. To test the utility of secondary phenotyping information, log-likelihood ratio tests can compare the fit of multi-trait models with differing sets of secondary traits. See also [97]. If these tests cannot be applied, because the VCOV models are not nested, then information criteria like AIC or BIC may be used [98,99]. Of course, models can also look at the predictive ability of models with different sets of secondary phenotypes.

The most common multi-trait prediction scenario considers yield and basic phenotypes measured simultaneously at the end of the growing season (*e.g.* [100,101]). A second (and less explored) scenario combines information from yield measured at harvest with low level phenotypes measured over multiple time points during the growing season, or with higher order traits that summarize the response of low level traits over time. Incorporating the genotype specific responses during the growing season into a prediction model provides a better insight in the traits underlying adaptation to particular growing conditions [8,72,102].

3.4. Illustrative examples of multi-trait prediction to incorporate high precision and whole-plant platforms

One example for multi-trait prediction is the use of aerial measurements of canopy temperature, and green and red NDVI as secondary phenotypes to increase accuracy for grain yield in wheat [30]. In [30], single trait predictions were compared with multi-trait predictions across environments differing for the level of drought and temperature stress. In this example, prediction accuracy for yield increased by 70%, averaged across environments. The benefit from modelling multiple traits simultaneously was similar for all environmental conditions.

An important issue to be considered when doing multi-trait prediction in multiple environments is to which extent those secondary phenotypes are related to phenology. Correcting for days to heading improved single-trait prediction accuracies across environments, suggesting that the G × E variance for grain yield corrected for days to heading is lower than that of uncorrected grain yield [30]. For both within- and across-environment prediction, correcting for days to heading reduced the genetic correlations between grain yield and the secondary traits, which in turn reduced the accuracy gained from including secondary trait data. A strategy to take full advantage of secondary trait data while avoiding indirect selection on a phenological trait may be to include data on the phenological trait in a multivariate prediction model alongside with any available secondary traits, and then use the multivariate BLUPs to calculate a selection index with yield and the phenological trait weighted appropriately [30].

[redacted]

on the that are in- ptions that are made residual.

value of a primary field trait (yield) j is: $y_{ij}^{f} = \mu_j^{f} + g_i^{f} + ge_{ij}^{f} + \varepsilon_{ij}^{f}$, with μ_j^{f} a intercept, g_i^{f} a fixed genetic main effect, ge_{ij}^{f} fixed and ε_{ij}^{f} a random residual, which can be structured in various ways. To start, for the G × E term, the purpose will be to identify genotypic covariables that can be combined with environmental covariables and leave a proportionality constant to be estimated, $ge_{ij}^{f} = \sum_{c\in C,d\in D}\theta_{cd}\,y(\boldsymbol{x}_i)_c^{s} z_{jd}^{e} + \delta_{ij}^{f}$, with $y(\boldsymbol{x}_i)_c^{s}$ a genotype specific predictor measured at the phenotyping platform with a genetic basis contained in $\boldsymbol{x}_i$, z_{jd}^{e} an environmental characterization, and θ_{cd} a scaling constant for the product $y(\boldsymbol{x}_i)_c^{s} z_{jd}^{e}$ that requires estimation. Lastly, δ_{ij}^{f} represents a lack of fit term. The full sets of genotypic and environmental covariables are denoted by *C* and *D*, respectively. Although the genotype specific platform traits can be related to their genetic basis by fitting a QTL or genomic prediction model this extra step of modelling is optional in that the field trait can be predicted directly from the platform traits at the phenotypic level as well as from their predictions from a genetic model. The use of a genetic model for the platform traits allows extending the selection set of genotypes by predicting platform trait values for non-phenotyped genotypes (at the platform). The genotypic and environmental covariables can include quadratic and cross product terms as well to allow the fitting of response surfaces. Another useful generalization is to allow the covariables to be expressed with respect to a spline basis.

The approximation of the field response G × E model term by products of platform genotypic covariables and field environmental characterization covariables can equally be applied to the total of genotype dependent terms as follows:
$g_i^{f} + ge_{ij}^{f} = \sum_{c\in C,d\in D}\theta_{cd}\,y(\boldsymbol{x}_i)_c^{s} z_{jd}^{e} + \delta_{ij}^{f}$.

For the fitting of predictive models of the above types, linear mixed models, penalized regressions and Bayesian approaches are required that allow for the modelling of the residual terms δ_{ij}^{f} and ε_{ij}^{f} with general variance-covariance structures allowing for kinship relationships between genotypes and heterogeneity of variance and correlation for environments. An important problem is the selection of predictive genotypic and environmental covariables in variable selection procedures. For penalized regression the choice of penalties requires attention, which translates to the choice of priors in the Bayesian context. The complicating issue is the simultaneous selection or penalization of information in the genotypic and environmental direction.

Selection of predictive environmental covariables becomes even more difficult - yet even more vital - when considering high-frequency data obtained from sensors. Weather stations and soil moisture probes report environmental conditions throughout the growing season, on a daily, hourly, or even more frequent basis. These characterizations potentially can be incorporated as environmental covariables after variable selection or as principal components after dimensionality reduction, but then the G × E interactions at any given time point of observation will be lost.

spent in different next step is to calculate variables like minimum, max- rainfall, radiation, and further vari- in each developmental stage. The assigned covariable in environment *j* for a genotype *i* growth stage depends on the beginning and end time of growth stage for that genotype in that environment. Therefore, the covariable should be indexed by genotype, *i*, and environment, *j*, as well as the covariable itself, *d*: z_{ijd}. G × E can be modelled in terms of such genotype specific environmental covariables as $ge_{ij} = \sum_{d\in D}\beta(\boldsymbol{x}_i)_d z_{ijd}$, with the genotypic sensitivities, $\beta(\boldsymbol{x}_i)_d$, being a (possible) function of a SNP profile. In maize, the progression of phenological stages closely follows thermal time, with a nearly constant leaf appearance rate [102]. Thermal time based on meristem temperature can be used to calculate leaf stages at the platform and in the field, provided that the leaf emission rates are the same. Leaf stages correspond to developmental stages of the ear [103] and can be used to define phenological periods. So, measuring leaf appearance on the platform allows to define the length of phenological stages in the field and to calculate the environmental conditions working at a particular genotype in a field experiment. In that way, environmental covariables can be calculated that are genotype and environment specific. With respect to those covariables, genotypic sensitivities are estimated. For the maize panel, three environmental covariables could be identified that explained a substantial part of the G × E [19]. These sensitivities could successfully be predicted from marker profiles in a genomic prediction. Therefore, the G × E in this example could be predicted from marker data and environmental covariables by using a factorial regression model. The platform served to estimate the length of genotype specific phenological stages and environmental characterizations.

3.7. Structural equations and network models

Structural equation models (SEMs) are an alternative class of statistical models describing the relations between primary traits and secondary traits [104]. In their simplest form, SEMs describe functional relations among traits in a single environment and at a single time point. The response for genotype *i* and trait *k* is then modelled as $y_i^{k} = \mu^{k} + \sum_{v\in pa(k)}\lambda_v^{k}\,y_i^{v} + \varepsilon_i^{k}$, where μ^{k} is a trait specific intercept, $pa(v)$ is the set of parents of trait *k* (*i.e.* the traits affecting trait *k*), and the λ_v^{k} 's are path coefficients, describing the strength of the relations. In the classical SEM literature [105], the functional relations are linear and the errors Normal, but generalizations are possible [106]. See *e.g.* [107] for a non-linear example in rice. SEMs are conceptually similar to factorial regression and crop growth models (in the sense that a primary trait is modelled in terms of component traits), but are more suitable for modelling additional levels of biological variation: the components can in turn depend on metabolites, methylation, proteins, gene expression, *etc.* The main advantage of structural models over regression models is the ability to predict the behaviour of the system after an intervention (*i.e.* selection decision), which mathematically is defined as a change in one or several of the structural equations. This property makes SEMs a tool for ideoptype design, helping breeders to define a selection strategy.

While SEMs rely on functional relations specified beforehand, methods for causal inference aim to learn relations between traits from observational data, which is of particular interest for traits that are only partially understood [108]. The earliest causal inference methods such as the PC-algorithm learn relations by estimating and comparing all the [redacted] etic cor- ons is an im- odels could make such gful way.

as G2P models

rop growth modelling framework aiming at prediction

The G2P models in Section 3 are all static, the time dimension proper to growth and development can only be incorporated in a limited and somewhat artificial way. Crop growth models (CGMs) present a class of G2P models that integrate genetic and environmental variables in a natural way over time [53,61,127,128]. This solves several problems. First, the need for dimensionality reduction or variable selection of environmental covariables is eliminated. Second, the interaction of the crop with the environment across time is automatically accounted for. This essentially eliminates the need to correct for flowering time, as phenology is incorporated into the model. Finally, it takes advantage of previously determined biological interactions and observations of plant growth and development to estimate the target (such as yield) from a set of higher order traits [129]. The higher order traits used to parameterize the CGM for the target trait may exhibit higher heritability and stability across environments than the target trait itself, and therefore selection on those secondary traits may be more efficient than selection on the target trait itself, as long as the physiological processes are properly modelled [129].

In Section 2.7, we described a CGM for a target trait f for genotype i in environment j as follows

$$y_{ij}^{f}(T_j) = \int_0^{T_j} \mathfrak{F}\left[(\boldsymbol{y}^{s\in S_P}(\boldsymbol{x}_i));\ (\boldsymbol{y}^{s\in S_I}(\boldsymbol{x}_i, \boldsymbol{z}(t_j)))\right]dt + \varepsilon_{ij}^{f}$$

The target trait, yield, depends on physiological parameters without environment and time dependence, $(\boldsymbol{y}^{s\in S_P}(\boldsymbol{x}_i))$, and intermediate traits with G × E that are environment and time dependent, $(\boldsymbol{y}^{s\in S_I}(\boldsymbol{x}_i, \boldsymbol{z}(t_j)))$. The time varying environmental covariables, $\boldsymbol{z}(t_j)$ determine the environment dependence of the intermediate traits. The secondary traits are functions of SNP variation *via* the argument $\boldsymbol{x}_i$, the SNP profile. The target trait is evaluated at the end of the growing season, for $t = T_j$. The dynamical function $\int_0^{T_j} \mathfrak{F}[[;]]dt$ integrates physiological and intermediate traits over time,

For most commonly used models, the time dimension is expressed at a daily time step, although hourly time steps have been recently proposed for sugar cane [130] and potato [131]. Therefore, most CGMs are not continuous time but discrete time models. Models combining processes occurring at different times scales and with facilities for feedback loops between physiological parameters and intermediate traits can generate reproducible emergent properties at plant level [12,55].

For the estimation of parameters in CGMs, Bayesian approaches are attractive [132,133]. Bayesian approaches have powerful capability to optimize multiple parameters in a nonlinear and complex model and to quantify the uncertainty in estimated parameters and predictions. Good

[redacted] re- nent traits which are other advantage of Bayesian expert knowledge of breeders and/ a as prior information.

be useful to evaluate the design of phenotyping he efficiency of resource allocation. Key design issues the number and type of environments to include in METs and e number and type of genotypes and traits that can be phenotyped. Other relevant design issues are how to choose the selection intensity on individual traits to arrive at the realisation of an ideotype, and how frequent and how precise phenotyping methods needs to be to increase prediction accuracy.

4.2. *Construction of a training set of environments*

When G × E is present in the TPE, the choice of which environments to use for phenotyping becomes crucial to obtain high prediction accuracies, as the environments used for training should represent well the environmental conditions in which future varieties will be grown [91]. The structure of G × E has a large influence on whether it is convenient to obtain field phenotyping information from METs (relying on the natural year-to-year variation to represent the range of environmental conditions relevant to the TPE), or whether it is necessary to use managed stress environments to ensure that all relevant levels of environmental variation are covered. The choice for one or the other strategy will depend on how well locations in an average year represent the whole range of environmental conditions and on the estimated year-to-year variation [8,136,137].

Crop growth models offer a valuable tool to characterize the G × E structure and to separate repeatable from non-repeatable G × E components. For example, the crop growth model APSIM was used to classify environments according to their water-deficit patterns [138] and to give an impression of the repeatability of a given water-deficit pattern at a given location, across years. The APSIM model has also been used to define the drought patterns of maize in Europe [139]. These results were used to develop a detailed QTL model with environment dependent QTL expression. Combining the drought patterns with temperature variation resulted into six different environmental scenarios that were introduced as a classifying factor in a multi-environment QTL model [19]. Such an approach allows to identify the most likely water-stress scenario in a particular location and the QTL-alleles that should be selected. The CGM output can also be used to develop statistical criteria to optimize METs to increase prediction accuracy for the target trait [140] or to characterize genotype-by-environment-by-management interactions, helping breeders, physiologists and agronomists to better understand the drivers of genotype adaptation [129,141].

4.3. *Construction of a training set of genotypes*

The parametrization of statistical and crop growth models requires intensive phenotyping efforts. In place of measuring expensive physiological traits on all genotypes of a selection set, one can try to estimate these parameters on a reduced training set of genotypes and then predict the parameters for the total of the selection set of genotypes by QTL or genomic prediction models. For the construction of a training set of genotypes that well represent the selection set, a diversity analysis based on molecular markers is useful. See [142–144] for examples of strategies to select genotypes for the training set. Alternatively, and relying only on phenotypic data, principal components analysis of morpho-physiological traits can be used to identify an appropriate [illegible] ent and [illegible] models (the [illegible] incorporate the 3-D [illegible] explicitly modelling plant ar- [illegible]

[illegible] GECROS has been used to evaluate the im- [illegible] photosynthesis on rice biomass production [151]. The [illegible] odel considers the relationships between traits at different [illegible] vels of organization over time, enabling to upscale from a basic process as photosynthesis to its effects at a whole crop level. A similar CGM aiming to scale up from photosynthesis to canopy level is the Diurnal Canopy Photosynthesis Simulator (DCaPS), implemented as a web-application [152]. The DCaPS package allows to assess likely canopy-level consequences of changes in photosynthetic properties, connecting photosynthesis with crop growth and development.

Virtual phenotyping using a functional-structural plant modelling approach can help to support the decision cycle of plant performance analysis by integrating different traits into a spatial-temporal whole plant simulation [146–150]. For example, an optimization procedure was applied to the functional–structural plant model MAppleT to evaluate which parameter combinations would allow to identify the trait combinations leading to the optimal phenotypes for the target trait [153].

4.5. *Evaluating the impact of phenotyping schedule on prediction accuracy*

The approach to characterize G × E using CGM [138,154,155] can be extended by combining the APSIM model with knowledge from quantitative genetics, simulating an explicit genetic basis for the APSIM parameters that are segregating in the population [156]. Such an approach would simulate trait dynamics across environments, characterizing G × E patterns and the change of trait correlations over time. For example, a wheat diversity panel segregating for 12 parameters of APSIM-wheat over 84 environments in the Australian wheat belt was simulated [156] (a subset of the environments shown in [138,157]). The output of these simulations allowed to evaluate the potential of biomass measured during the growing season to improve yield predictions with a multi-trait genomic prediction model. A further opportunity offered by the combination of statistical genetic models and CGMs is the evaluation of the impact of phenotyping frequency and the size of measurement error on trait heritability and prediction accuracy for the target trait. In such a way, an approach that combines CGMs like APSIM-wheat with a quantitative genetic basis potentially allows to evaluate phenotyping and selection strategies across environments.

5. Challenges ahead

This paper has discussed a number of G2P modelling approaches to take into consideration the different sources of phenotypic information and their underlying G × E structure. Unfortunately, the fast implementation of phenotyping technologies has not necessarily been [illegible] and [illegible]) will play a central [illegible] ssemination of phenotyping [illegible] mes by standardizing the pheno- [illegible] envirotyping protocols and data storage.

[illegible] per, we have discussed a number of G2P models [illegible] to predict phenotypes across environments and will [illegible] btain a larger response to selection. Although the central role [illegible] G2P models for genetic gain is not under discussion, little will be gained from these models, without their implementation in an effective pipeline that facilitates the integration of data pre-processing, prediction and decision support tools [158]. This aspect of an effective implementation of G2P models is still at an early stage, but its development promises to play a central role in the years to come.

6. Concluding remarks

- Additional phenotyping can be done at growth chambers, platforms, managed environment trials or a sample of (field) multi-environment trials.
- Genotype-to-phenotype modelling (G2P) plays a central role in enabling breeders to combine the different sources of phenotypic and genotypic information and assisting them in the design of phenotyping strategies.
- Modelling steps involve image pre-processing, correcting for spatial trends within each trial (low level traits), dynamical modelling of genotypic responses within environments against time thereby producing intermediate order traits, modelling dynamics parameters in their dependence on environmental gradients and calculate sensitivities (higher order traits), and combining the information of multiple higher order traits with environmental data to predict the target trait.
- G2P models for the target trait from underlying component traits involve statistical models like multi-trait models, factorial regression models, and crop growth models.
- Design issues that can be addressed with G2P models involve the decision of which environments and genotypes to use for phenotyping, which traits to prioritize in which environment type (ideotype design) and what kind of phenotyping schedule to use in terms of measurement frequency and precision to increase prediction accuracy for the target trait.

Funding

This paper has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement n° 731013, EPPN^{2020}; from the European Union's Seventh Framework Programme (FP7/ 2007-2013) under the grant agreement n° FP7-613556, Whealbi; from the Grains Research and Development Corporation (GRDC CSP00179 grant to CSIRO), and from the US DoE ARPA-E TERRA program grant supporting research at Purdue University and the University of Queensland. Open access to this manuscript was funded by VSNUGOLD.

Acknowledgements

We thank both anonymous reviewers for their comments as these comments contributed significantly to improve the final version of this paper.

[redacted]

dy breeding 61–572, https://

Baumgarten, N.J. Hausmann, of plant breeding: complementing prediction, Crop Pasture Sci. 65 (2014) 0.1071/CP14007.

odruff, R.L. Eisemann, P.S. Brennan, I.H. DeLacy, A selection ccommodate genotype-by-environment interaction for grain yield of managed-environments for selection among genotypes, Theor. Appl. Genet. 90 (1995) 492–502, https://doi.org/10.1007/bf00221995.

[10] G.J. Rebetzke, K. Chenu, B. Biddulph, C. Moeller, D.M. Deery, A.R. Rattey, D. Bennett, E.G. Barrett-Lennard, J.E. Mayer, A multisite managed environment facility for targeted trait and germplasm phenotyping, Funct. Plant Biol. 40 (2012) 1–13, https://doi.org/10.1071/FP12180.

[11] G.L. Hammer, S. Chapman, E. van Oosterom, D.W. Podlich, Trait physiology and crop modelling as a framework to link phenotypic complexity to underlying genetic systems, Aust. J. Agric. Res. 56 (2005) 947–960, https://doi.org/10.1071/AR05157.

[12] F. Tardieu, L. Cabrera-Bosquet, T. Pridmore, M. Bennett, Plant phenomics, from sensors to knowledge, Curr. Biol. 27 (2017) R770–R783, https://doi.org/10.1016/j.cub.2017.05.055.

[13] J.L. Araus, S.C. Kefauver, M. Zaman-Allah, M.S. Olsen, J.E. Cairns, Translating high-throughput phenotyping into genetic gain, Trends Plant Sci. (2018) 1–16, https://doi.org/10.1016/j.tplants.2018.02.001.

[14] A.-R. Ballester, Y. Tikunov, J. Molthoff, S. Grandillo, M. Viquez-Zamora, R. de Vos, R.A. de Maagd, S. van Heusden, A.G. Bovy, Identification of loci affecting accumulation of secondary metabolites in tomato fruit of a Solanum lycopersicum × Solanum chmielewskii introgression line population, Front. Plant Sci. 7 (2016) 1–14, https://doi.org/10.3389/fpls.2016.01428.

[15] Y. Xu, C. Xu, S. Xu, Prediction and association mapping of agronomic traits in maize using multiple omic data, Heredity (Edinb.) (2017) 1–11, https://doi.org/10.1038/hdy.2017.27.

[16] V.S. Weber, J.L. Araus, J.E. Cairns, C. Sanchez, A.E. Melchinger, E. Orsini, Prediction of grain yield using reflectance spectra of canopy and leaves in maize plants grown under different water regimes, Field Crops Res. 128 (2012) 82–90, https://doi.org/10.1016/j.fcr.2011.12.016.

[17] L. Cabrera-Bosquet, C. Grieder, S. Alvarez Prado, C. Sánchez, J.L. Araus, Kernel δ18O reflects changes in apical dominance and plant transpiration in tropical maize, J. Agron. Crop Sci. 203 (2017) 277–285, https://doi.org/10.1111/jac.12196.

[18] T. Duan, S.C. Chapman, Y. Guo, B. Zheng, Dynamic monitoring of NDVI in wheat agronomy and breeding trials using an unmanned aerial vehicle, Field Crops Res. 210 (2017) 71–80, https://doi.org/10.1016/j.fcr.2017.05.025.

[19] E. Millet, C. Welcker, W. Kruijer, S. Negro, S. Nicolas, S. Praud, N. Ranc, T. Presterl, R. Tuberosa, Z. Bedo, X. Draye, B. Usadel, A. Charcosset, F. van Eeuwijk, F. Tardieu, A. Coupel-Ledru, C. Bauland, Genome-wide analysis of yield in Europe: allelic effects as functions of drought and heat scenarios, Plant Physiol. 172 (2016) 00621, https://doi.org/10.1104/pp.16.00621.

[20] G.A. Slafer, R. Savin, V.O. Sadras, Coarse and fine regulation of wheat yield components in response to genotype and environment, Field Crops Res. 157 (2014) 71–83, https://doi.org/10.1016/j.fcr.2013.12.004.

[21] F.A. van Eeuwijk, M.C.A.M. Bink, K. Chenu, S.C. Chapman, Detection and use of QTL for complex traits in multiple environments, Curr. Opin. Plant Biol. 13 (2010) 193–205, https://doi.org/10.1016/j.pbi.2010.01.001.

[22] F.A. van Eeuwijk, D.V. Bustos-Korts, M. Malosetti, What should students in plant breeding know about the statistical aspects of genotype × environment interactions? Crop Sci. 56 (2016) 2119–2140, https://doi.org/10.2135/cropsci2015.06.0375.

[23] P.H.C. Eilers, B.D. Marx, Flexible smoothing with B-splines and penalties, Stat. Sci. 11 (1996) 89–102, https://doi.org/10.2307/2246049.

[24] P.H.C. Eilers, B.D. Marx, M. Durbán, Twenty years of P-splines, SORT-Stat. Oper. Res. Trans. 39 (2015) 149–186.

[25] F. Perez-Sanz, P.J. Navarro, M. Egea-Cortines, Plant phenomics: an overview of image acquisition technologies and image data analysis algorithms, Gigascience 6 (2017) 1–18, https://doi.org/10.1093/gigascience/gix092.

[26] S. Liu, F. Baret, B. Andrieu, P. Burger, M. Hemmerlé, Estimation of wheat plant

winter: a pow- response to temperature, doi.org/10.1071/FP14226.

multi-leaf segmentation, alignment and videos, arXiv:1505.00353v2. (2017). http:// .

, S. Mondal, E. Autrique, L.G. Pérez, J. Crossa, M. Reynolds, temperature and vegetation indices from high-throughput phe- improve accuracy of pedigree and genomic selection for grain yield in wheat, G3 Genes Genomes Genet. 6 (2016) 2799–2808, https://doi.org/10.1534/g3.116.032888.

[31] A.B. Potgieter, B. George-Jaeggli, S.C. Chapman, K. Laws, L.A. Suárez Cadavid, J. Wixted, J. Watson, M. Eldridge, D.R. Jordan, G.L. Hammer, Multi-spectral imaging from an unmanned aerial vehicle enables the assessment of seasonal leaf area dynamics of sorghum breeding lines, Front. Plant Sci. 8 (2017) 1–11, https://doi.org/10.3389/fpls.2017.01532.

[32] H.S. Naik, J. Zhang, A. Lofquist, T. Assefa, S. Sarkar, D. Ackerman, A. Singh, A.K. Singh, B. Ganapathysubramanian, A real-time phenotyping framework using machine learning for plant stress severity rating in soybean, Plant Methods 13 (2017) 23, https://doi.org/10.1186/s13007-017-0173-7.

[33] B. Cullis, B. Gogel, A. Verbyla, R. Thompson, Spatial analysis of multi-environment early generation variety trials, Biometrics 54 (1998) 1–18, https://doi.org/10.2307/2533991.

[34] J.G. Velazco, M.X. Rodríguez-Álvarez, M.P. Boer, D.R. Jordan, P.H.C. Eilers, M. Malosetti, F.A. van Eeuwijk, Modelling spatial trends in sorghum breeding field trials using a two-dimensional P-spline mixed model, Theor. Appl. Genet. 130 (2017) 1375–1392, https://doi.org/10.1007/s00122-017-2894-4.

[35] M.X. Rodríguez-Álvarez, M.P. Boer, F.A. van Eeuwijk, P.H.C. Eilers, Correcting for spatial heterogeneity in plant breeding experiments with P-splines, Spat. Stat. 23 (2018) 52–71, https://doi.org/10.1016/j.spasta.2017.10.003.

[37] B.R. Cullis, A.B. Smith, N.E. Coombes, On the design of early generation variety trials with correlated data, J. Agric. Biol. Environ. Stat. 11 (2006) 381–393, https://doi.org/10.1198/108571106x154443.

[38] P.J. Flood, W. Kruijer, S.K. Schnabel, R. van der Schoor, H. Jalink, J.F.H. Snel, J. Harbinson, M.G.M. Aarts, Phenomics for photosynthesis, growth and reflectance in Arabidopsis thaliana reveals circadian and long-term fluctuations in heritability, Plant Methods 12 (2016) 14, https://doi.org/10.1186/s13007-016-0113-y.

[39] R. van Rooijen, W. Kruijer, R. Boesten, F.A. van Eeuwijk, J. Harbinson, M.G.M. Aarts, Natural variation of YELLOW SEEDLING1 affects photosynthetic acclimation of Arabidopsis thaliana, Nat. Commun. 8 (2017) 1421, https://doi.org/10.1038/s41467-017-01576-3.

[40] C.X. Ma, G. Casella, R. Wu, Functional mapping of quantitative trait loci underlying the character process: a theoretical framework, Genetics 161 (2002) 1751–1762 http://www.ncbi.nlm.nih.gov/pubmed/12196415%5Cnhttp://www.ncbi.nlm.nih.gov/pmc/articles/PMC1462199/pdf/12196415.pdf.

[41] R. Wu, C.-X. Ma, R.C. Littell, S.S. Wu, T. Yin, M. Huang, M. Wang, G. Casella, A logistic mixture model for characterizing genetic determinants causing differentiation in growth trajectories, Genet. Res. 79 (2002) 235–245, https://doi.org/10.1017/s0016672302005633.

[42] M. Malosetti, R.G.F. Visser, C. Celis-Gamboa, F.A. Eeuwijk, QTL methodology for response curves on the basis of non-linear mixed models, with an illustration to senescence in potato, Theor. Appl. Genet. 113 (2006) 288–300, https://doi.org/10.1007/s00122-006-0294-2.

[43] J.A. Bac-Molenaar, C. Granier, J.J.B. Keurentjes, D. Vreugdenhil, Genome wide association mapping of time-dependent growth responses to moderate drought stress in Arabidopsis, Plant Cell Environ. (2015) 88–102, https://doi.org/10.1111/pce.12595.

[44] G.B. West, J.H. Brown, B.J. Enquist, A general model for ontogenetic growth, Nature 413 (2001) 628–631.

[45] P. Hurtado, S. Schnabel, A. Zaban, M. Veteläinen, E. Virtanen, P.C. Eilers, F. Eeuwijk, R.F. Visser, C. Maliepaard, Dynamics of senescence-related QTLs in potato, Euphytica 183 (2012) 289–302, https://doi.org/10.1007/s10681-011-0464-4.

[46] J. Sun, J.E. Rutkoski, J.A. Poland, J. Crossa, J.-L. Jannink, M.E. Sorrells, Multitrait, random rgression, or simple repeatability model in high-throughput phenotyping data improve genomic prediction for wheat grain yield, Plant Genome 10 (2) (2017), https://doi.org/10.3835/plantgenome2016.11.0111.

[47] J.R. Stinchcombe, M. Kirkpatrick, Genetics and evolution of function-valued traits: understanding environmentally responsive phenotypes, Trends Ecol. Evol. 27 (2012) 637–647, https://doi.org/10.1016/j.tree.2012.07.002.

[48] R. Wu, M. Lin, Functional mapping-how to map and study the genetic architecture of dynamic complex traits, Nat. Rev. Genet. 7 (2006) 229–237, https://doi.org/10.1038/nrg1804.

[49] Z. Li, M.J. Sillanpää, Dynamic quantitative trait locus analysis of plant phenomic data, Trends Plant Sci. 20 (2015) 822–833, https://doi.org/10.1016/j.tplants.2015.08.012.

[50] S. Macgregor, S.A. Knott, I. White, P.M. Visscher, Quantitative trait locus analysis of longitudinal quantitative trait data in complex pedigrees, Genetics 171 (2005) 1365–1376, https://doi.org/10.1534/genetics.105.043828.

[redacted]

e with con- ontrol of growth and t stress in wheat, J. Exp. .1093/jxb/erv320.

Millet, F. Tardieu, B. Parent, Distinct ongation by light and evaporative demand in n. 40 (2017) 2017–2028, https://doi.org/10.1111/pce.

C. Klukas, S. Friedel, P. Rischbeck, D. Chen, A. Entzian, N. Stein, Graner, B. Kilian, Dissecting spatiotemporal biomass accumulation in barley under different water regimes using high-throughput image analysis, Plant Cell Environ. 38 (2015) 1980–1996, https://doi.org/10.1111/pce.12516.

[59] L. Cabrera-Bosquet, C. Fournier, N. Brichet, C. Welcker, B. Suard, F. Tardieu, High-throughput estimation of incident light, light interception and radiation-use efficiency of thousands of plants in a phenotyping platform, New Phytol. 212 (2016) 269–281, https://doi.org/10.1111/nph.14027.

[60] R.C. Muchow, T.R. Sinclair, J.M. Bennett, Temperature and solar radiation effects on potential maize yield across locations, Agron. J. 82 (1990) 338–343, https://doi.org/10.2134/agronj1990.00021962008200020033x.

[61] F. Technow, C.D. Messina, L.R. Totir, M. Cooper, Integrating crop growth models with whole genome prediction through approximate Bayesian computation, PLoS One 10 (2015) 1–20, https://doi.org/10.1371/journal.pone.0130855.

[62] M.X. Rodríguez-Álvarez, D.J. Lee, T. Kneib, M. Durbán, P. Eilers, Fast smoothing parameter separation in multidimensional generalized P-splines: the SAP algorithm, Stat. Comput. 25 (2015) 941–957, https://doi.org/10.1007/s11222-014-9464-2.

[63] M.P. Boer, D. Wright, L. Feng, D.W. Podlich, L. Luo, M. Cooper, F.A. van Eeuwijk, 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 (2007) 1801–1813, https://doi.org/10.1534/genetics.107.071068.

[64] M. Malosetti, J.-M. Ribaut, F.A. van Eeuwijk, The statistical analysis of multi-environment data: modeling genotype-by-environment interaction and its genetic basis, Front. Physiol. 4 (2013) 1–17, https://doi.org/10.3389/fphys.2013.00044.

[65] M.P.M. Thoen, N.H. Davila Olivas, K.J. Kloth, S. Coolen, P.P. Huang, M.G.M. Aarts, J.A. Bac-Molenaar, J. Bakker, H.J. Bouwmeester, C. Broekgaarden, J. Bucher, J. Busscher-Lange, X. Cheng, E.F. Fradin, M.A. Jongsma, M.M. Julkowska, J.J.B. Keurentjes, W. Ligterink, C.M.J. Pieterse, C. Ruyter-Spira, G. Smant, C. Testerink, B. Usadel, J.J.A. van Loon, J.A. van Pelt, C.C. van Schaik, S.C.M. van Wees, R.G.F. Visser, R. Voorrips, B. Vosman, D. Vreugdenhil, S. Warmerdam, G.L. Wiegers, J. van Heerwaarden, W. Kruijer, F.A. van Eeuwijk, M. Dicke, Genetic architecture of plant stress resistance: multi-trait genome-wide association mapping, New Phytol. 213 (2017) 1346–1362, https://doi.org/10.1111/nph.14220.

[66] S. Ceccarelli, E. Acevedo, S. Grando, Breeding for yield stability in unpredictable environments: single traits, interaction between traits, and architecture of genotypes, Euphytica 56 (1991) 169–185, https://doi.org/10.1007/BF00042061.

[67] V.O. Sadras, D. Calderini, Crop Physiology: Applications for Genetic Improvement and Agronomy, Elsevier Science, 2014.

[68] X. Yin, S.D. Chasalow, C.J. Dourleijn, P. Stam, M.J. Kropff, Coupling estimated effects of QTLs for physiological traits to a crop growth model: predicting yield variation among recombinant inbred lines in barley, Heredity (Edinb.) 85 (2000) 539–549, https://doi.org/10.1046/j.1365-2540.2000.00790.x.

[69] X. Yin, P. Stam, M.J. Kropff, A.H.C.M. Schapendonk, Crop modeling, QTL mapping, and their complementary role in plant breeding, Agron. J. 95 (2003) 90–98, https://doi.org/10.2134/agronj2003.9000.

[70] X. Yin, P.C. Struik, F.A. van Eeuwijk, P. Stam, J. Tang, QTL analysis and QTL-based prediction of flowering phenology in recombinant inbred lines of barley, J. Exp. Bot. 56 (2005) 967–976, https://doi.org/10.1093/jxb/eri090.

[71] M. Reymond, B. Muller, A. Leonardi, A. Charcosset, F. Tardieu, Combining quantitative trait loci analysis and an ecophysiological model to analyze the genetic variability of the responses of maize leaf growth to temperature and Water deficit, Plant Physiol. 131 (2003) 664–675, https://doi.org/10.1104/pp.013839.

[72] F. Tardieu, M. Reymond, B. Muller, C. Granier, T. Simonneau, W. Sadok, C. Welcker, Linking physiological and genetic analyses of the control of leaf growth under changing environmental conditions, Aust. J. Agric. Res. 56 (2005) 937–946, https://doi.org/10.1071/AR05156.

Bot. 65 (2014)

ments of Statistical Learning, second

ulloch, Variance Components, Wiley, 2009, ooks?id = mkCSOOzHMngC.

sina, E. van Oosterom, S. Chapman, V. Singh, A. Borrell, Cooper, Molecular breeding for complex adaptive traits: how in- crop ecophysiology and modelling can enhance efficiency, in: X. Yin, P.C. Struik (Eds.), Crop Systems Biology: Narrowing the Gaps between Crop Modelling and Genetics, Springer International Publishing, Cham, 2016, pp. 147–162, , https://doi.org/10.1007/978-3-319-20562-5_7.

[78] P. Jackson, T. McRae, M. Hogarth, Selection of sugarcane families across variable environments I. Sources of variation and an optimal selection index, Field Crops Res. 43 (1995) 109–118, https://doi.org/10.1016/0378-4290(95)00039-S.

[79] P. Jackson, X.M. Wei, F. Atkin, Optimal selection indices in early stage trials in sugarcane breeding programs, ASSCT 2015, 37th Annu. Conf. Aust. Soc. Sugar Cane Technol. vol. 37, 2015, pp. 244–251 https://www.scopus.com/inward/record.uri?eid = 2-s2.0-84978525380&partnerID = 40&md5 = 839ac63638643eece05eb3a5817ebb0f.

[80] J. Basnayake, P.A. Jackson, N.G. Inman-Bamber, P. Lakshmanan, Sugarcane for water-limited environments. Variation in stomatal conductance and its genetic correlation with crop productivity, J. Exp. Bot. 66 (2015) 3945–3958, https://doi.org/10.1093/jxb/erv194.

[81] H. Poorter, F. Fiorani, R. Pieruschka, T. Wojciechowski, W.H. van der Putten, M. Kleyer, U. Schurr, J. Postma, Pampered inside, pestered outside? Differences and similarities between plants growing in controlled conditions and in the field, New Phytol. 212 (2016) 838–855, https://doi.org/10.1111/nph.14243.

[82] G. Dignat, C. Welcker, M. Sawkins, J.M. Ribaut, F. Tardieu, The growths of leaves, shoots, roots and reproductive organs partly share their genetic control in maize plants, Plant Cell Environ. 36 (2013) 1105–1119, https://doi.org/10.1111/pce.12045.

[83] C. Welcker, W. Sadok, G. Dignat, M. Renault, S. Salvi, A. Charcosset, F. Tardieu, A Common genetic determinism for sensitivities to soil water deficit and evaporative demand: meta-analysis of quantitative trait loci and introgression lines of maize, Plant Physiol. 157 (2011) 718–729, https://doi.org/10.1104/pp.111.176479.

[84] R. Chapuis, C. Delluc, R. Debeuf, F. Tardieu, C. Welcker, Resiliences to water deficit in a phenotyping platform and in the field: how related are they in maize? Eur. J. Agron. 42 (2012) 59–67, https://doi.org/10.1016/j.eja.2011.12.006.

[85] G.J. Rebetzke, R.A. Richards, Genetic improvement of early vigour in wheat, Aust. J. Agric. Res. 50 (1999) 291–301, https://doi.org/10.1071/A98125.

[86] T. Duan, S.C. Chapman, E. Holland, G.J. Rebetzke, Y. Guo, B. Zheng, Dynamic quantification of canopy structure to characterize early plant vigour in wheat genotypes, J. Exp. Bot. 67 (2016) 4523–4534, https://doi.org/10.1093/jxb/erw227.

[87] Y. Jia, J.-L. Jannink, Multiple-trait genomic selection methods increase genetic value prediction accuracy, Genetics 192 (2012) 1513–1522, https://doi.org/10.1534/genetics.112.144246.

[88] J. Burgueño, G. de los Campos, K. Weigel, J. Crossa, Genomic prediction of breeding values when modeling genotype x environment interaction using pedigree and dense molecular markers, Crop Sci. 52 (2012) 707–719, https://doi.org/10.2135/cropsci2011.06.0299.

[89] D. Jarquín, J. Crossa, X. Lacaze, P. Cheyron, J. Daucourt, J. Lorgeou, F. Piraux, L. Guerreiro, P. Pérez, M. Calus, J. Burgueño, G. Campos, A reaction norm model for genomic selection using high-dimensional genomic and environmental data, Theor. Appl. Genet. 3 (2013) 1–13, https://doi.org/10.1007/s00122-013-2243-1.

[90] N. Heslot, D. Akdemir, M. Sorrells, J.-L. Jannink, Integrating environmental covariates and crop modeling into the genomic selection framework to predict genotype by environment interactions, Theor. Appl. Genet. (2013) 1–18, https://doi.org/10.1007/s00122-013-2231-5.

[91] M. Malosetti, D. Bustos-Korts, M.P. Boer, F.A. van Eeuwijk, Predicting responses in multiple environments: issues in relation to genotype × environment interactions, Crop Sci. 56 (2016) 2210–2222, https://doi.org/10.2135/cropsci2015.05.0311.

[92] M. Malosetti, J. Ribaut, M. Vargas, J. Crossa, F. Euewijk, A multi-trait multi-environment QTL mixed model with an application to drought and nitrogen stress trials in maize (Zea mays L.), Euphytica 161 (2008) 241–257, https://doi.org/10.1007/s10681-007-9594-0.

[93] F.N. Gumedze, T.T. Dunne, Parameter estimation and inference in the linear mixed model, Linear Algebra Appl. 435 (2011) 1920–1944.

[94] J. Crossa, G. de los Campos, P. Perez, D. Gianola, J. Burgueño, J.L. Araus, D. Makumbi, R.P. Singh, S. Dreisigacker, J.B. Yan, V. Arief, M. Banziger, H.J. Braun, Prediction of genetic values of quantitative traits in plant breeding using pedigree and molecular markers, Genetics 186 (2010), https://doi.org/10.1534/genetics.110.118521 713-U406.

[95] A. Legarra, I. Aguilar, I. Misztal, A relationship matrix including full pedigree and genomic information, J. Dairy Sci. 92 (2009) 4656–4663, https://doi.org/10.3168/jds.2009-2061.

[96] I. Misztal, Properties of random regression models using linear splines, J. Anim. Breed. Genet. 123 (2006) 74–80, https://doi.org/10.1111/j.1439-0388.2006.00582.x.

[redacted]

ron. 31

.x.

lante et de la culture,

Res. (1921) 557–585.

atent Variables, 1st ed., Wiley-

.worldcat.org/isbn/0471011711.

orrelation in Biology: A User's Guide to Path Analysis,

ons and Causal Inference, 2nd ed., Cambridge University Press,

ge, USA, 2016.

Onogi, O. Ideta, T. Yoshioka, K. Ebana, M. Yamasaki, H. Iwata, Uncovering a nuisance influence of a phenological trait of plants using a nonlinear structural equation: application to days to heading and culm length in Asian cultivated rice, PLoS One 11 (2016) 1–17, https://doi.org/10.1371/journal.pone.0148609.

[108] J. Pearl, Causality, March ISBN 0521773628 Cambridge University Press, Cambridge, UK, 2000, p. 400 http://adsabs.harvard.edu/cgi-bin/nph-bib_query?bibcode=2000caus.book.

[109] P. Spirtes, C. Glymour, R. Scheines, Causation, Prediction, and Search, Second Edition (Adaptive Computation and Machine Learning), second ed., A Bradford Book, 2001.

[110] S.M. Welch, J.L. Roe, Z. Dong, A genetic neural network model of flowering time control in, Agron. J. 95 (2003) 71–81, https://doi.org/10.2134/agronj2003.7100.

[111] S.M. Welch, Z. Dong, J.L. Roe, S. Das, Flowering time control: gene network modelling and the link to quantitative genetics, Aust. J. Agric. Res. 56 (2005) 919–936, https://doi.org/10.1071/AR05155.

[112] E.C. Neto, C.T. Ferrara, A.D. Attie, B.S. Yandell, Inferring causal phenotype networks from segregating populations, Genetics 179 (2008) 1089–1100.

[113] E.C. Neto, M.P. Keller, A.D. Attie, B.S. Yandell, Causal graphical models in systems genetics. A unified framework for joint inference of causal network and genetic architecture for correlated phenotypes, Ann. Appl. Stat. 4 (2010) 320–339.

[114] M. Scutari, P. Howell, D.J. Balding, I. Mackay, Multiple quantitative trait analysis using bayesian networks, Genetics 198 (2014) 129–137, https://doi.org/10.1534/genetics.114.165704.

[115] H. Wang, F.A. van Eeuwijk, A new method to infer causal phenotype networks using QTL and phenotypic information, PLoS One 9 (2014) e103997, https://doi.org/10.1371/journal.pone.0103997.

[116] H. Wang, J. Paulo, W. Kruijer, M. Boer, H. Jansen, Y. Tikunov, B. Usadel, S. van Heusden, A. Bovy, F. van Eeuwijk, Genotype-phenotype modeling considering intermediate level of biological variation: a case study involving sensory traits, metabolites and QTLs in ripe tomatoes, Mol. Biosyst. 11 (2015) 3101–3110, https://doi.org/10.1039/C5MB00477B.

[117] Z. Dong, O. Danilevskaya, T. Abadie, C. Messina, N. Coles, M. Cooper, A gene regulatory network model for floral transition of the shoot apex in maize and its dynamic modeling, PLoS One 7 (2012), https://doi.org/10.1371/journal.pone.0043450.

[118] B.D. Valente, G.J.M. Rosa, G. De Los Campos, D. Gianola, M.A. Silva, Searching for recursive causal structures in multivariate quantitative genetics mixed models, Genetics 185 (2010) 633–644, https://doi.org/10.1534/genetics.109.112979.

[119] K. Töpner, G.J.M. Rosa, D. Gianola, C.-C. Schön, Bayesian networks illustrate genomic and residual trait connections in maize (Zea mays L.), G3 Genes Genomes Genet. 7 (2017) 2779–2789, https://doi.org/10.1534/g3.117.044263.

[120] M.H. Maathuis, D. Colombo, M. Kalisch, P. Bühlmann, Predicting causal effects in large-scale systems from observational data, Nat. Methods 7 (2010) 247–248.

[121] N. Meinshausen, A. Hauser, J.M. Mooij, J. Peters, P. Versteeg, P. Bühlmann, Methods for causal inference from gene perturbation experiments and validation, Proc. Natl. Acad. Sci. U. S. A. 113 (2016) 7361–7368.

[122] J. Peters, P. Bühlmann, N. Meinshausen, Causal inference using invariant prediction: identification and confidence intervals, arXiv Prepr. arXiv1501.01332v3. (2015).

[123] J. Peters, D. Janzing, B. Schölkopf, Elements of Causal Inference: Foundations and Learning Algorithms, MIT Press, 2017.

[124] S. De Vos, K.J. Wardenaar, E.H. Bos, E.C. Wit, M.E.J. Bouwmans, P. De Jonge, An investigation of emotion dynamics in major depressive disorder patients and healthy persons using sparse longitudinal networks, PLoS One 12 (2017) 1–18, https://doi.org/10.1371/journal.pone.0178586.

[125] N.A. Alimi, M.C.A.M. Bink, J.A. Dieleman, J.J. Magán, A.M. Wubs, A. Palloix, F.A. van Eeuwijk, Multi-trait and multi-environment QTL analyses of yield and a set of physiological traits in pepper, Theor. Appl. Genet. 126 (2013) 2597–2625,

ternational,

C. Gho, M. Cooper, Leveraging

ariation to improve phenotypic prediction:

GM) with whole genome prediction (WGP), Eur.

oi.org/10.1016/j.eja.2018.01.007 in press.

ean, A. Doherty, E. van Oosterom, S. Chapman, Sorghum crop

its utility in agronomy and breeding, Sorghum State Art Future

ect, American Society of Agronomy and Crop Science Society of America,

nc., Madison, WI, 2016, https://doi.org/10.2134/agronmonogr58.2014.0064.

[130] N.G. Inman-Bamber, P.A. Jackson, C.J. Stokes, S. Verrall, P. Lakshmanan, J. Basnayake, Sugarcane for water-limited environments: enhanced capability of the APSIM sugarcane model for assessing traits for transpiration efficiency and root water supply, Field Crops Res. 196 (2016) 112–123, https://doi.org/10.1016/j.fcr.2016.06.013.

[131] R. Quiroz, H. Loayza, C. Barreda, C. Gavilán, A. Posadas, D.A. Ramírez, Linking process-based potato models with light reflectance data: does model complexity enhance yield prediction accuracy? Eur. J. Agron. 82 (2017) 104–112, https://doi.org/10.1016/j.eja.2016.10.008.

[132] M. Van Oijen, J. Rougier, R. Smith, Bayesian calibration of process-based forest models: bridging the gap between models and data, Tree Physiol. 25 (2005) 915–927, https://doi.org/10.1093/treephys/25.7.915.

[133] T. Iizumi, M. Yokozawa, M. Nishimori, Parameter estimation and uncertainty analysis of a large-scale crop model for paddy rice: application of a Bayesian approach, Agric. For. Meteorol. 149 (2009) 333–348, https://doi.org/10.1016/j.agrformet.2008.08.015.

[134] A. Onogi, M. Watanabe, T. Mochizuki, T. Hayashi, H. Nakagawa, T. Hasegawa, H. Iwata, Toward integration of genomic selection with crop modelling: the development of an integrated approach to predicting rice heading dates, Theor. Appl. Genet. 129 (2016) 805–817, https://doi.org/10.1007/s00122-016-2667-5.

[135] M. Cooper, F. Technow, C. Messina, C. Gho, L. Radu Totir, Use of crop growth models with whole-genome prediction: application to a maize multienvironment trial, Crop Sci. 56 (2016) 2141–2156, https://doi.org/10.2135/cropsci2015.08.0512.

[136] H. Campos, M. Cooper, G.O. Edmeades, C. Löffler, J.R. Schussler, M. Ibañez, Changes in drought tolerance in maize associated with fifty years of breeding for yield in the U.S. corn belt 1, Maydica 51 (2006) 10–11.

[137] G.J. Rebetzke, A.R. Rattey, G.D. Farquhar, R.A. Richards, A. Tony, G. Condon, Genomic regions for canopy temperature and their genetic association with stomatal conductance and grain yield in wheat, Funct. Plant Biol. 40 (2012) 14–33, https://doi.org/10.1071/FP12184.

[138] K. Chenu, R. Deihimfard, S.C. Chapman, Large-scale characterization of drought pattern: a continent-wide modelling approach applied to the Australian wheatbelt-spatial and temporal trends, New Phytol. 198 (2013) 801–820, https://doi.org/10.1111/nph.12192.

[139] M.T. Harrison, F. Tardieu, Z. Dong, C.D. Messina, G.L. Hammer, Characterizing drought stress and trait influence on maize yield under current and future conditions, Glob. Change Biol. 20 (2014) 867–878, https://doi.org/10.1111/gcb.12381.

[140] R. Rincent, E. Kuhn, H. Monod, F.-X. Oury, M. Rousset, V. Allard, J.Le Gouis, Optimization of multi-environment trials for genomic selection based on crop models, Theor. Appl. Genet. (2017) 1–18, https://doi.org/10.1007/s00122-017-2922-4.

[141] G.L. Hammer, G. McLean, S. Chapman, B. Zheng, A. Doherty, M.T. Harrison, E. van Oosterom, D. Jordan, Crop design for specific adaptation in variable dryland production environments, Crop Pasture Sci. 65 (2014) 614–626, https://doi.org/10.1071/CP14088.

[142] R. Rincent, D. Laloë, S. Nicolas, T. Altmann, D. Brunel, P. Revilla, V.M. Rodríguez, J. Moreno-Gonzalez, A. Melchinger, E. Bauer, C.-C. Schoen, N. Meyer, C. Giauffret, C. Bauland, P. Jamin, J. Laborde, H. Monod, P. Flament, A. Charcosset, L. Moreau, Maximizing the reliability of genomic selection by optimizing the calibration set of reference individuals: comparison of methods in two diverse groups of maize inbreds (Zea mays L.), Genetics 192 (2012) 715–728, https://doi.org/10.1534/genetics.112.141473.

[143] D. Bustos-Korts, M. Malosetti, S. Chapman, B. Biddulph, F. van Eeuwijk, Improvement of predictive ability by uniform coverage of the target genetic space, G3 Genes Genomes Genet. 6 (2016) 3733–3747, https://doi.org/10.1534/g3.116.035410/-/DC1.

[144] J. Isidro, J.-L. Jannink, D. Akdemir, J. Poland, N. Heslot, M. Sorrells, Training set optimization under population structure in genomic selection, Theor. Appl. Genet. 128 (2015) 145–158, https://doi.org/10.1007/s00122-014-2418-4.

[145] D.A. Ramírez, C. Gavilán, C. Barreda, B. Condori, G. Rossel, R.O.M. Mwanga, M. Andrade, P. Monneveux, N.L. Anglin, D. Ellis, R. Quiroz, Characterizing the diversity of sweetpotato through growth parameters and leaf traits: precocity and light use efficiency as important ordination factors, South Afr. J. Bot. 113 (2017) 192–199, https://doi.org/10.1016/j.sajb.2017.08.009.

[146] J.A. Postma, C. Kuppe, M.R. Owen, N. Mellor, M. Griffiths, M.J. Bennett, J.P. Lynch, M. Watt, OpenSimRoot: widening the scope and application of root architectural models, New Phytol. 215 (2017) 1274–1286, https://doi.org/10.1111/nph.14641.

ourt, E. Costes,
pace to design
926–1939, https://doi.org/

er, Using crop simulation to generate gen-
n effects for sorghum in water-limited environ-
s. 53 (2002) 379–389, https://doi.org/10.1071/

.A. van Eeuwijk,
Modelling to Evaluate
nder Different Drought Patterns,
doi.org/10.18174/421321.

man, N. Huth, R. Faivre, K. Chenu, Assessment of
plant traits across environments by combining global
and dynamic modeling in wheat, PLoS One 11 (1) (2016)
tps://doi.org/10.1371/journal.pone.0146385.

Janssen, C.H. Porter, A.D. Moore, I.N. Athanasiadis, I. Foster, J.W. Jones, J.M. Antle, Towards a new generation of agricultural system data, models and knowledge products: information and communication technology, Agric. Syst. 155 (2017) 200–212, https://doi.org/10.1016/j.agsy.2016.09.017.

[159] A.B. Smith, B.R. Cullis, R. Thompson, The analysis of crop cultivar breeding and evaluation trials: an overview of current mixed model approaches, J. Agric. Sci. (2005) 449–462.

[160] D. Reynolds, F. Baret, C. Welcker, et al., What is cost-efficient phenotyping? Optimizing costs for different scenarios, Plant Sci. (2018).

[161] RCP, Kuijken, F.A. van Eeuwijk, L.F.M. Marcelis, H.J. Bouwmeester, Root phenotyping: from component trait in the lab to breeding, J. Exp. Bot. 66 (18 2015) 5389–5401.