How process-based modeling can help plant breeding deal with G x E x M interactions

Amir Hajjarpoor [a], William C.D. Nelson [b], Vincent Vadez [a,c,*]

[a] *UMR DIADE, Université de Montpellier, Institut de Recherche pour le Développement (IRD), Montpellier, France*
[b] *International Crops Research Institute for Semi-Arid Tropics (ICRISAT), Bulawayo, Zimbabwe*
[c] *International Crops Research Institute for Semi-Arid Tropics (ICRISAT), Patancheru, India*

A R T I C L E I N F O

A B S T R A C T

Genotype-by-Environment-by-Management (GxExM) interactions represent many unknowns for crop improvement programs, which hampers the development of improved varieties, especially for highly variable environments like those limited by rainfall. While breeding programs have traditionally used statistical tools to deal with these interactions, process-based crop modeling has recently become an alternative and powerful approach. Overall, while statistical methods remain the most optimal solution to deal with GxExM interactions when many production datasets across time and space are available from multi-environment trials (MET), *in silico* methods like crop modeling can be used if such data is lacking, or if MET data don't cover the entire target region. Yet, despite several reviews on the potential uses of process-based modeling tools to aid such issues, their practical use in helping breeding programs is still in its infancy. After exposing the pros and cons of process-based modeling, this paper presents the step-by-step process that would allow breeding programs to harness this tool to help guide their breeding decisions. We also argue that the issue of GxExM interactions should be tackled in a co-construction process, involving breeders, agronomists, extensionists, and modelers from the beginning, and this would bring crop models one step closer to being used to help make plant breeding decisions.

1. Introduction

Traditional varietal improvement has focused on high yield heritability (Chenu, 2015), although it has not provided generic solutions to the issue of environment and management interactions. High heritability increases selection accuracy and breeding efficiency, but heritability declines when Genotype-by-Environment-by-Management (GxExM) interactions increase, which becomes a primary constraint for the selection of germplasm with consistent performance across environments (Hammer et al., 2014). Understanding and quantifying the causes of such interactions is critical for breeding progress to develop GxM packages that maximize production in specific E's (environments) (Chauhan and Rachaputi, 2014; George, 2014) and the most frequent E types from year to year (Hammer et al., 2014). The presence of GxExM interactions, if known, could even create opportunities for new prediction-based crop improvement strategies. However, there is often a major knowledge gap in most crop improvement programs (Cooper et al., 2020). Even though successful crop improvement in terms of crop productivity depends on the complementary mix of cultivar

improvement (G) coupled with improved agronomy (M) (Cooper et al., 2020, 2021a,b; Messina et al., 2020b; Fischer et al., 2014), the term crop improvement is often associated with genetic improvement alone (Kholová et al., 2020). A more holistic, systems-based approach is needed that includes all aspects of crop performance improvement. A focus on genetics alone undermines crop yield improvement potential (Van Bueren et al., 2018). However, this requires significant changes to the current crop improvement pipelines and necessitates the integration of new tools, ways of thinking, and principally, collaboration.

Yield-based approaches that classify crop production regions have been used to reduce GxE interactions (Cooper and Woodruff, 1993) but often lack detailed soil and weather measurements to help us understand these interactions. Some approaches are based on environmental characteristics that identify mega-environments, which are sometimes called adaptation zones or ecological zones (Van Eeuwijk et al., 2016). The International Maize and Wheat Improvement Center (CIMMYT) developed the concept of mega-environments (Rajaram et al., 1994) to target maize and wheat germplasm development using similarities in environmental variables and cropping system requirements. Surprisingly,
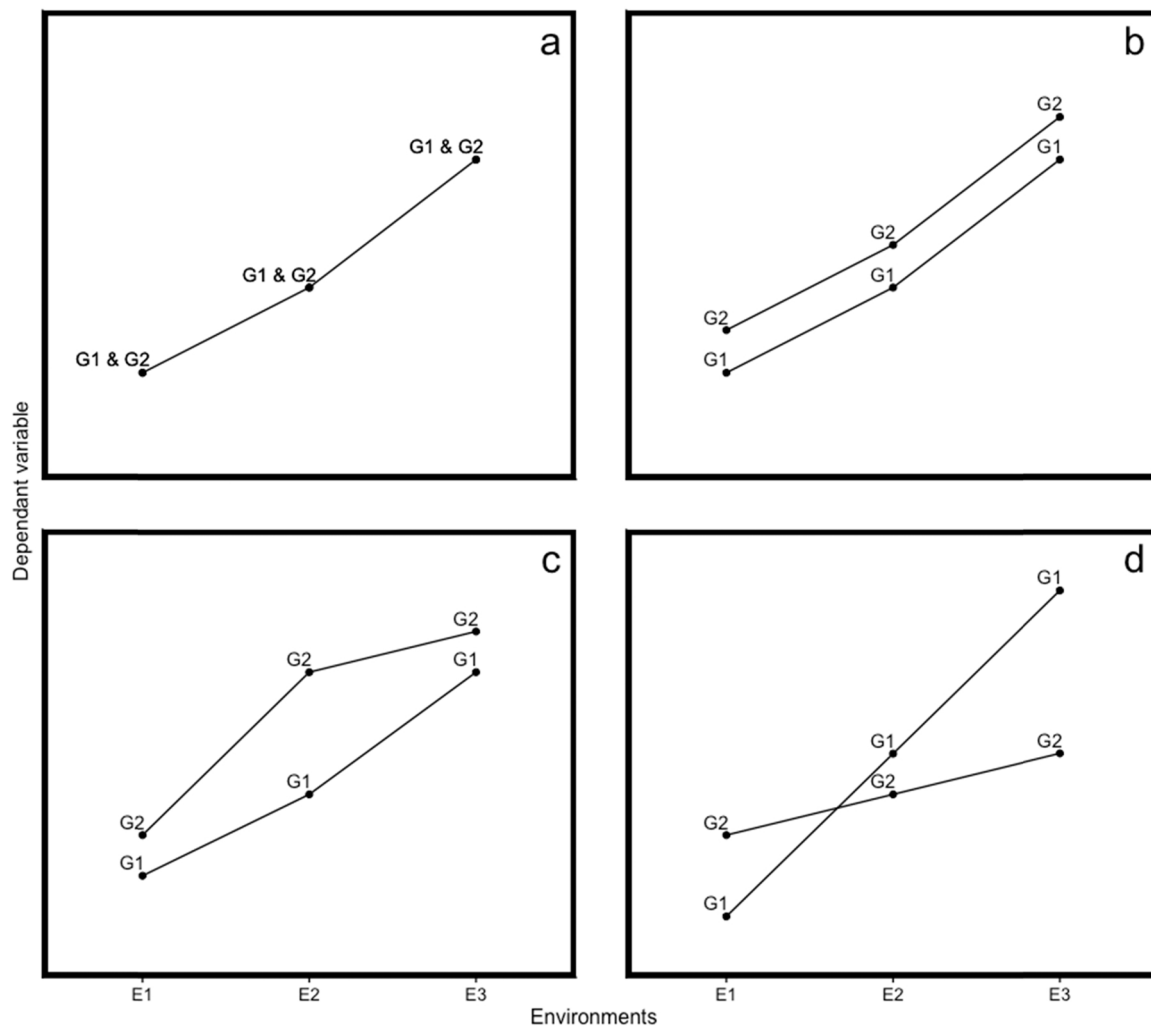
Fig. 1. Genotype by environment (GxE) interaction and reaction norms: (a) clear differences between environments, no GxE interaction; (b) parallel reaction norms, differences between genotypes and environments, no GxE interaction; (c) non-parallel reaction norms, non-crossover GxE interaction; (d) non-parallel reaction norms, crossover GxE interaction
Adapted from Bustos-Korts et al. (2019).

Windhausen et al. (2012) showed that the maize mega-environment groupings did not explain as much of the GxE interactions for grain yield as did simple grouping into low (<3 t/ha) and high (>3 t/ha) yield sites. A comprehensive GxExM analysis permits a nuanced understanding of the factors that lie behind regional differences in yield (Porter et al., 2010), and here a statistical approach to deal with GxExM interaction is simple, provided there is a large enough number of trials and sites. This is done regularly with multi-environment trials (METs). However, METs are generally restricted by the number and the location of trial sites and crop seasons, especially for minor crops (Chenu, 2015; Ramirez-Villegas et al., 2020; Witcombe et al., 1998), which eventually limits the spatio-temporal reach of the GxExM analysis (Kholová et al., 2020). In contrast, an *in-silico* analysis using process-based crop simulation models (CSM) can help overcome spatial and temporal limitations (Cooper et al., 2020; Heinemann et al., 2015; Ramirez-Villegas et al., 2020), and allows testing in future environments or environments that are not covered by METs. Environment types are then based on specific site–year–management–genotype combinations rather than on locations (Chenu, 2015; Hammer et al., 2014), leading to improvements in the sustainability and productivity of diverse production systems (Messina et al., 2020). Various approaches have been developed and applied to classify crop production regions for breeding purposes within the concept of the target population of environments (TPE; Comstock,

1977). A TPE is a set of key environmental conditions that are expected to occur within the breeding program's target geography (Cooper et al., 2016; Van Eeuwijk et al., 2016), e.g. different types of water-limited environments (Hammer et al., 2014; Kholová et al., 2013). Any research and breeding program that encounters strong GxE interactions must define target environment types (Cooper et al., 2021b). As a result, the TPE represents the range of soil, meteorological, and agronomic conditions in which the varieties of a breeding program will be grown (Bustos-Korts et al., 2019).

Two schools of thought address GxExM interactions, using either statistical methods or process-based modeling. Most statistical models used for GxE analysis are phenotypic response functions for each genotype to environmental variables, which largely consist of data from MET (Malosetti et al., 2013). Environmental variables in MET are limited to soil and weather. Process-based crop simulation modeling helps disentangle GxExM interactions by classifying abiotic stress patterns (frequency or intensity) within a geographical space (e.g. Battisti and Sentelhas, 2019; Chapman et al., 2000; Chenu et al., 2011; Cooper et al., 1997; Hajjarpoor et al., 2021, 2018; Heinemann et al., 2015; Kholová et al., 2013; Sciarresi et al., 2019; Wang et al., 2019). This offers the opportunity to balance risks and extend the growing season to their full potential. Recent reviews highlighted the potential of CSM in accelerating crop improvement program achievements in targeted areas
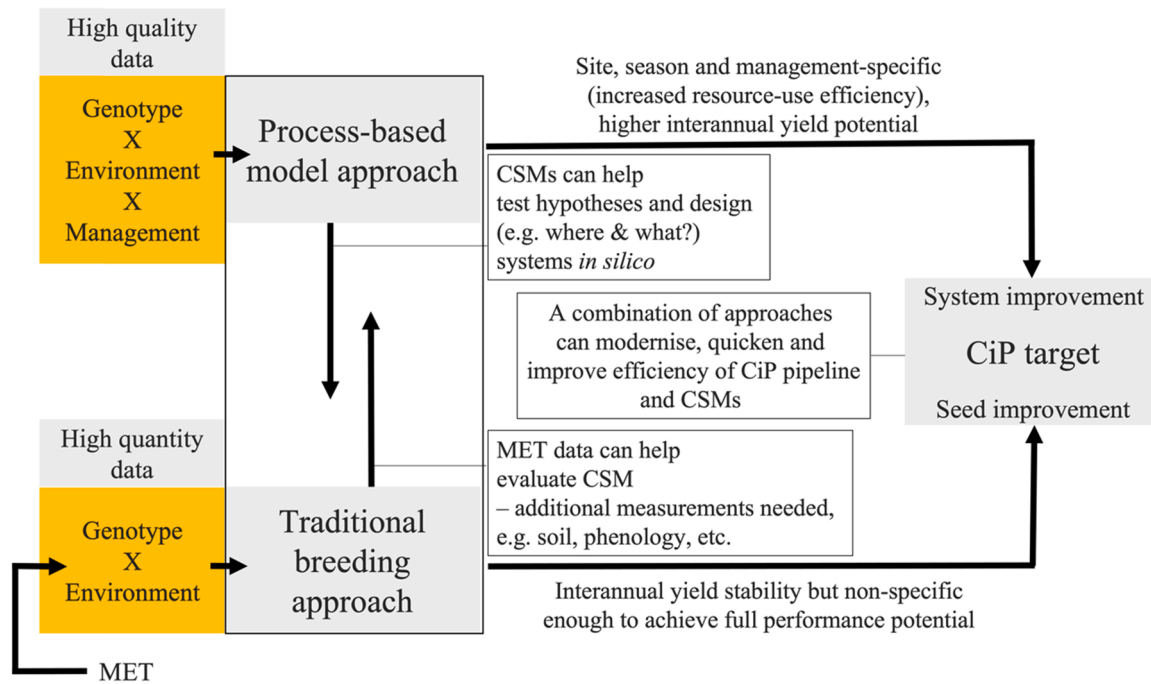
Fig. 2. Two schools of thoughts for crop improvement programs (CiP) and how they are connected: Crop simulation model (CSM) and statistical approach using multi-environment trials (MET).

(e.g. Boote et al., 2021; Ramirez-Villegas et al., 2020; Reynolds et al., 2018). However, their use in guiding the design of such programs is still in its infancy and some key questions remain unanswered: What are the advantages and drawbacks for statistical or modeling approaches to deal with GxExM? What is a simple step-by-step process to bring process-based modeling to the doorstep of crop improvement programs?

This paper proposes a simple step-by-step approach to bring the model's capacity to the doorstep of breeding programs, toward a better definition of target population of environment(s), within which the clustering of sub-unit(s) allow for the reduction of GxExM interactions. To that end, the specific objectives of this paper are two-fold: (i) it discusses the pros and cons of using a statistical or a process-based modeling approach to deal with GxExM interactions; (ii) building on our experience, we draw a road map for breeders to use a process-based modeling approach in cases where CSM fits best. More specifically for the second objective, we describe how to provide easy access to modeling tools to better select testing sites and breeding targets. We address the issue of a lack of data repositories (e.g. soil, weather, production) and describe a co-construction process among the various disciplines involved in the crop improvement process, going beyond "breeders only" and including the role of crop modelers (Van Bueren et al., 2018).

2. A challenge to solve: GxExM interaction interferes with genotype selection

Most crop improvement strategies in developing countries aim to produce few "broad-adapted genotypes" with superior performance across a range of agroecologies, despite highly varying climatic characteristics, such as water-stress scenarios, and are mostly built around improving yield potential (Ceccarelli, 1989; Ceccarelli et al., 2010, 2007). This has been for many years the approach of the CGIAR with the concept of "mega-varieties". The GxE interaction is a phenomenon that occurs when the relative phenotypic performance of a pair or a set of genotypes is significantly conditioned by the environment (Bustos-Korts et al., 2019). The phenotypic response of a genotype to the environment is described by a function known as the reaction norm (Fig. 1). When the

reaction norm lines of evaluated genotypes in different environments are not parallel, there is a GxE interaction (Finlay and Wilkinson, 1963). This interaction is considered a challenge due to its implications for genotype selection.

2.1. Statistical analysis or a process-based modeling approach to address the challenge of interaction

There are two schools of thought to deal with GxExM interactions using process-based modeling or statistics. Here, we review scenarios in which one prevails over the other. The flowchart below shows how these schools of thought are connected (Fig. 2).

Several statistical approaches have developed during the last decades to analyze multi-environment trial (METs) data and GxE interactions. An extensive overview of such approaches can be seen through the work of Malosetti et al. (2013), Van Eeuwijk et al. (2016), Bustos-Korts et al. (2019) and Brown et al. (2020) (Fig. 3).

If trials span several seasons and multiple sites, statistical methods can group sites into TPEs (DeLacy et al., 1996; IRRI, 2006). However, in practical terms, even if MET datasets are abundant, testing environments may not represent the range of stress intensity or timing across the TPE. In this instance, selection would not optimally select genotypes needed for the TPE. The degree of mismatch between MET and TPE is expected to be high in highly variable (often abiotic) environments, and could mislead the genetic gain (Cooper et al., 1997, 1995; Qiao et al., 2004). In addition, while climatic indices are used as model covariates in the evaluation of understanding general crop agroecologies (Brown et al., 2020; Van Etten et al., 2019), temporal variation due to inter-annual variation could be ignored in some climatic scenarios (Windhausen et al., 2012). Such limitations are now being addressed by including environment type frequencies (Battisti and Sentelhas, 2019; Chenu et al., 2011; Hammer et al., 2014). Envirotyping can measure how well the samples of environments realized in any MET represent the TPE in question (Cooper et al., 2020).

Considerable advances are being made in improved TPE characterization, environment types, and frequencies of environment types, thanks to the coupling of CSM with long-term weather records in order
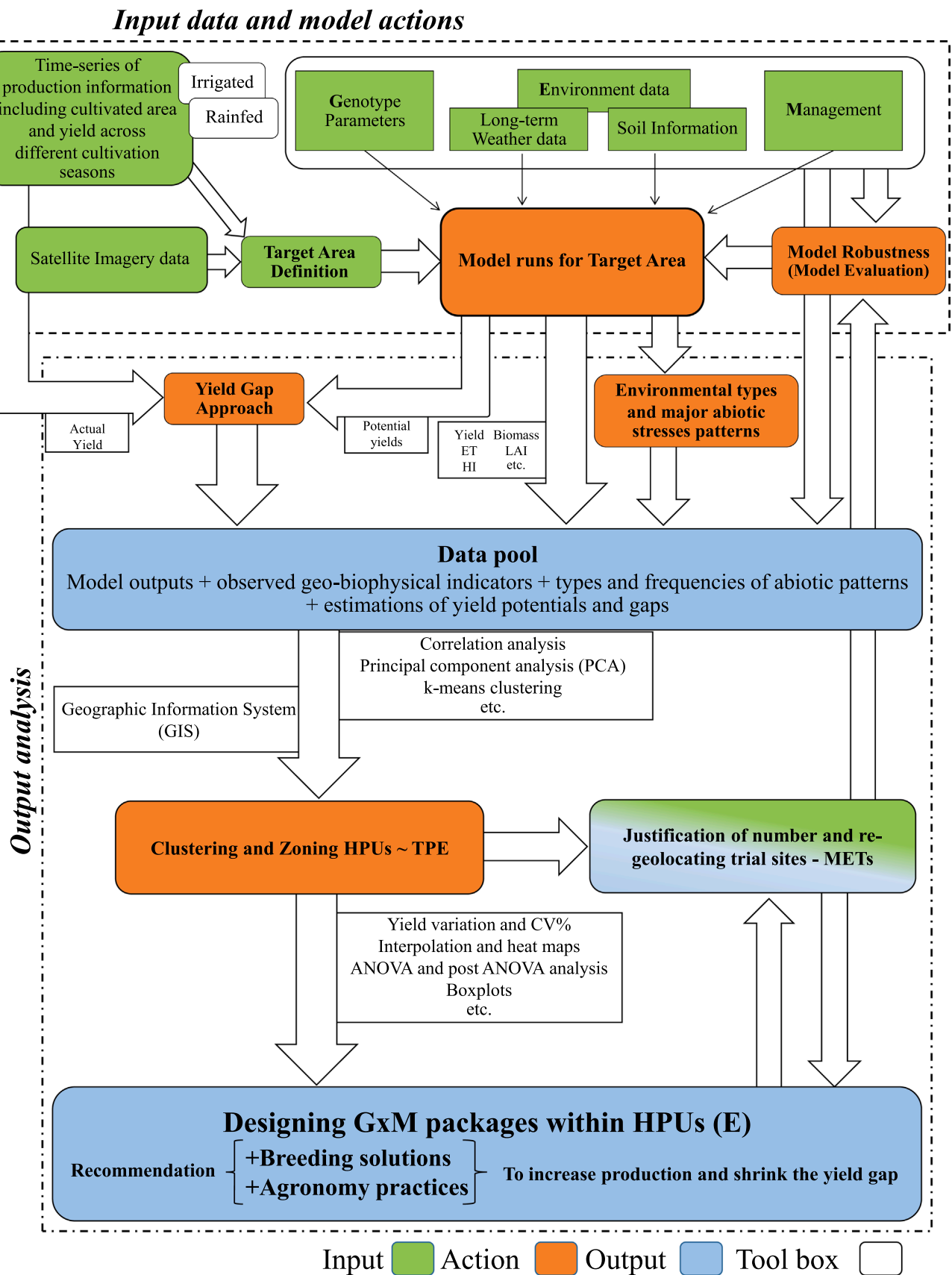
Fig. 3. Flowchart of step-by-step process-based modeling for breeders, which can result in the design of GxM packages in the target E. Input requires a team effort and close interaction with breeders. The action step is typically the work of crop modelers. The output steps revolve around collaboration and interaction with breeders. The toolbox is a set of existing methods. Homogenous production units (HPU); target population of environment (TPE); multi-environment trail (MET).

to generate seasonal sequences of stress that can be used to determine frequencies of stress environment types (Battisti and Sentelhas, 2019; Chapman et al., 2000; Chenu et al., 2011; Kholová et al., 2013; Seyoum et al., 2017). In combination with MET data, this type of information can use data from different trials according to how representative they are of the TPE (Chapman et al., 2000).

2.2. Which way to go: statistical analysis or process-based modeling

Clearly, both statistical and process-based modeling methods have limitations, mainly because they originated from different disciplines and were therefore designed to answer different questions - both have advantages and drawbacks. With ample MET datasets, a statistical approach has the advantage of being routinely implemented by most breeding programs. This would likely apply to widely grown crops receiving lots of funding support, such as "the big three", i.e. maize, rice and wheat. Both statistical and process-based model approaches lack quality datasets for minor crops, including many legumes and crop that are grown in regions characterized by abiotic stresses, such as low precipitation. The two main drawbacks of the statistical approach are: (i) the fact that sites of the MET may not well represent the frequency of environments faced by the crops in its TPE, implying extra costs incurred by over-testing in certain environment types. This is especially the case where inter annual variability is large and for instance regions limited by water. MET networks also are often old and their relevance has not been updated (Hajjarpoor et al., 2021). Consequently, this approach offers little insight into the source of GxExM interactions; (ii) while the statistical approach has fewer 'start-up' costs when MET data is available, one must acknowledge the high field trial running costs of existing plant breeding programs. Field trials will always be a necessity, for instance to test innovation, but there are currently many missed opportunities for data collection and CSM improvement in existing MET infrastructures (e.g. agronomic data is often missing from MET datasets). Here, the digitization of data collection, transfer and translation/ analysis can help improve efficiencies substantially from current systems the world over. The clear advantage of the process-based modeling approach is that it requires much less MET data. Provided there is an existing model for the chosen crop and that this model has already been positively evaluated for the TPE in question with a fairly limited and small amount of MET data, it would then become a prime choice for minor crops facing a lack of MET data. The main drawbacks of the process-based modeling approach are: (i) that different types of datasets (climate, soil, crop coefficients) are necessary to run models and they are sometimes hard to get. When available, the lack of streamlined organization between data collection and use means data wrangling tasks are not trivial. Enhanced collaboration and structure on this front could help increase efficiencies substantially; (ii) high-quality production data, i.e. highly detailed, is necessary to test model performance; (iii) drawbacks (i) and (ii) lead to model uncertainty; (iv) lastly, crop models are not always intuitive to use, which results in limited user adoption - breeding programs still need to harness their potential.

Therefore, a 'way to go' is not statistical or process-based modeling, but a collaborative way forward seeking mutual benefit for the joint cause of crop selection accuracy and speed. Ultimately, CSM-specific data collection from current breeding MET can improve CSM, which in turn can help refine MET increasing breeding program efficiency. The two disciplines are clearly highly complementary. The following subsections address a necessary process as well as step-by-step technical aspects that could help plant breeding fully harness the advantages of process-based modeling.

3. Practical examples of using process-based modeling to deal with TPE

Process-based crop simulation models can classify stress patterns based on their frequency or intensity within a geographic location (Sinclair et al., 2020). For instance, the post-rainy sorghum production environment in India was characterized and revealed five main types of water limitation that occurred at different times throughout the cropping cycle. This led to varying levels of yield reduction (Kholová et al., 2013), possibly requiring different breeding and/or agronomic packages across the TPE. Chauhan and Rachaputi (2014) demonstrated this for mung bean in Australia, using APSIM-generated output to identify distinct regions based on clustering water-stress index values and seasonal yield variation at different locations.

Model-defined regions can be used as selection environments for TPEs, linking GxM to the E for optimal yield under variable conditions. The TPE approach has been successfully used by wheat breeding programs in Australia (Chenu et al., 2011; Lobell et al., 2015) and applied to maize in Europe (Harrison et al., 2014). A similar TPE analysis was conducted for chickpea in India, in which six subregions were identified as homogenous production units (HPUs) (Hajjarpoor et al., 2018). They declared that within each HPU, similar system responses to genetic or management interventions would be expected. This study revealed, for example, that drought was not a problem in all units, contrary to common belief in chickpea production regions in India and might require specific breeding and agronomic recommendations. A similar approach was used for groundnut in India (Hajjarpoor et al., 2021), where subdividing TPE to HPUs reduced yield variation. Consequently, limitations and recommendations specific to each HPU were provided. Harrison et al. (2014) is an excellent example of how the TPE approach can develop our understanding of crop traits needed under projected drought stress, including the effects of elevated CO_2 on maize as a C4 crop. Based on simulations and cluster analysis, the results revealed four dominant drought-stress seasonal patterns. The authors concluded that some common maize breeding targets in drought conditions, such as shifting phenology (maturity), better anthesis-silking synchrony and higher kernel number when subjected to one drought-stress pattern, may not necessarily have the same favourable effects as for other drought-stress patterns. Kaloki et al. (2019) presented an alternative approach that assessed the suitability of chickpea ideotypes in Australia. The study incorporated key traits and targeted specific environments, which were grouped into three major clusters using the soil water deficit method with varying water stress levels. A chickpea ideotype, designed to grow under water-limited conditions, outperformed commercial varieties *in silico*, therefore providing chickpea breeders with a blueprint for crop improvement targets. There are many examples of where crop models have been convincingly used to offer guidance to breeders and agronomists in setting priorities for related interventions (Cooper et al., 2014; Hammer et al., 2006; Messina et al., 2020a, 2015; Sinclair et al., 2010).

4. Putting in place the process to use crop simulation to assist breeding

4.1. Team assembly

In 1981, Byth argued that a basic objective of crop improvement research was to improve the efficiency of MET. This, he believed, could be accomplished in two ways: (i) more effective exploitation of existing investment in MET through multidisciplinary involvement in their implementation, and (ii) objective reduction in their size (genotypes and/or environments) based on a better understanding of TPE adaptation. Therefore, an active collaboration network amongst agricultural disciplines is necessary to add value to MET data and fully understand GxExM interactions. The challenges are no longer for plant breeders alone, but are highly relevant for a broader group of people who require multidisciplinary solutions.

4.2. Model selection

While there are different types of CSM, here we opted for

Table 1

Minimum and desirable data to be collected from breeding trials for process-based modeling.

Data to collect	Minimum	Desirable
Plot-wise data		
Dates	Sowing	
	Flowering (50%)	Mention if 50% or 1st flower (or 1st pod for legumes)
	Physiological maturity	
Agronomy	Density (final) - At least row spacing and hill spacing and number of plants per hill	Density at 2–3 weeks after sowing
		Plot Heterogeneity (1–5: good-bad)
		Leaf area index (LAI)
		Leaf N concentration
		Phyllochron
Harvest	100-seed weight	Biomass/Stover yield (stems+leaves)
	Productive tiller count or	
	Panicle count	
	Grain yield	
Location / Field data		
Location	Latitude	Elevation
	Longitude	
Weather	Rainfall	Sunshine hours or Radiation
	Min Temperature	
	Max Temperature	
Irrigation	Date	
	Quantity	
	Type (furrow, sprinkler, etc.)	
Fertilization	Date	
	Quantity	
	Type (DAP, urea, etc.)	
Soil	Texture (sandy, clay, etc.)	Initial moisture content (perceived)
	Type (Vertisol, Alfisol, Arenosol, etc.)	Initial N content
	Depth	Previous crop and cultivation period (e.g. Groundnut, Jun-Oct 2020)

Table 2

List of important crop parameters that are common among crops and crop models.

Parameter	How to estimate, where to look, and what to look for? Parameter-specific information
Phenology phases	Thermal time (TT) or biological days (BD) are required to predict critical phenological stages in all crop models. These can be estimated based on cardinal temperatures, average daily temperature, and day to the occurrence of individual phenological stages. Depending on the crop and the model, the phenological stages may differ. However, it is particularly important to record **day to flowering** and **day to maturity** (from sowing); these two records can be the basis to define TT or BD requirements for different phenological stages.
Photoperiod (PP) reaction	Some crops (or genotypes) are sensitive to PP in part of their cycle. With photoperiod-sensitive genotypes, some parameters need to be estimated. Often, **critical photoperiods** can be relatively constant among cultivars within a species. Models use a function that affects phenology phase occurrence. **Photoperiod sensitivity** slope (PPsen) is a parameter used in this function and it can be obtained from growth-chamber and/ or field experiments via iterative optimization methods (programs exist). **Stages sensitive to PP** can be found in literature or be obtained from experiments using reciprocal transfers between photoperiod treatments.
Leaf area development	All leaf area development variables can be obtained directly from experimental observations of the crops or indirectly via ceptometers. The **phyllochron** can be obtained by fitting a linear regression line to the main-stem leaf/ node number versus the temperature unit. The **allometric relationship coefficients** between plant leaf area and main-stem leaf number can be found by fitting a power equation to the respective data, i.e. plant leaf area versus main-stem leaf number. If different plant density data is available, their relationship to plant density can also be found. To do this, a power equation is fitted to the data of each plant density. The relationship between the obtained coefficients and plant density can then be found. In some models like EPIC and SWAT, a common exponential regression function describes LAI expansion as a function of normalized temperature units. In this case, the maximum expected LAI (LAIMX) value is an input that growing conditions and plant density may influence. The slope of an LAI plot versus leaf dry weight estimates a **specific leaf area (SLA)**. As early-stage leaves are thin, data obtained at higher LAIs should be used to avoid an overestimation of SLA.
Biomass accumulation	Different methods and parameters are used to predict daily dry matter production depending on the selected model, however, most parameters are constant or can be found in the literature. For example, some models like ORYZA2000, CROPGRO and HybridMaize, which simulate gross photosynthesis and respiration separately, may need extra parameters. At the same time, most CSMs use temperature-adjusted radiation use efficiency (RUE) to convert photosynthetically active intercepted radiation into dry matter (van Wart et al., 2013). To estimate RUE experimentally, dry matter production by a crop canopy needs to be measured in conjunction with simultaneous measurements of PAR interception by the canopy. RUE can be constant among genotypes of a species and can be found in literature like Soltani and Sinclair (2012). Another parameter that is used in the model estimation of biomass production is the **extinction coefficient (KPAR).** Plant and canopy characteristics are combined into a single composite property, KPAR, which can be obtained experimentally from PAR interception and crop LAI measurements.
Yield Formation	Three parameters are common in crop models to estimate final yield, including rate of harvest index (HI) increase, maximum expected HI and fraction of remobilizable dry mass. Among them, **maximum HI** is a priority to be estimated and can be obtained directly from experimental crop observations. The rate of HI increase can also be estimated based on the ratio of grain to total shoot dry matter increase as a measure of reproductive efficiency.

mechanistic or process-based models, as opposed to empirical models. Process-based models simulate biological processes and their responses to the environment, represented by equations in their algorithms (Boote et al., 2021; Tardieu et al., 2020). Of course, even process-based models become empirical at a certain level of complexity due to limited knowledge of biological processes (Soltani and Sinclair, 2012). The predictive capacity of process-based models depends on our understanding of system interactions, often dictated by the quality of available data (Kholová et al., 2020) - not all yield-reducing factors are currently considered (De Wit et al., 2019). One example is that many widely used crop models lack certain details, such as the effects of pests and diseases (Donatelli et al., 2017) or the fact that CSM must include the appropriate modules in order to simulate extreme events (Ababaei and Chenu, 2019; Chenu, 2015; Lobell et al., 2015). According to a review of Barlow et al. (2015) on wheat production, while frost damage is taken into account in a number of CSM, heat shock is rarely considered. As our cropping systems evolve to accommodate climatic changes but also societal demands, so do the needs of process-based models and plant breeding. This highlights the importance of model choice in areas where extreme events are likely as well as the improvement of existing models and increased harmonization (e.g. insight and protocol exchange) and discussion within the CSM community itself (e.g. through initiatives such as AgMIP; Rosenzweig et al., 2013).

In conclusion, process-based crop models are essential to undertake geospatial assessments, such as TPE, that account for the temporal and spatial variability necessary to resolve crop production challenges. Methods with clear documentation describing how to run robust

TableA1

Few example of the climate databases for modeling purposes in daily basis and with all parameters available including temperature, precipitation and radiation.

Database	Resolution	Time frame	Source type	References / Source URL
POWER	0.50°	1981-Present	Satellite data	https://power.larc.nasa.gov/
ERA5	0.25°	1979-Present	Reanalysis	(Hersbach et al., 2020)
				https://doi.org/10.24381/cds.adbb2d47
WFDEI	0.50°	1979–2016	Bias corrected reanalysis	(Weedon et al., 2014)
				https://rda.ucar.edu/datasets/ds314.2/
Princeton	0.25°	1948–2016	Bias corrected reanalysis	(Sheffield et al., 2006)
				http://hydrology.princeton.edu/data/pgf/v3/
AgMERRA	0.25°	1980–2010	Reanalysis with in situ and remotely-sensed observational datasets	(Ruane et al., 2015)
				https://doi.org/10.1016/j.agrformet.2014.09.016

TableA2

Few example of the available soil database for modeling purposes.

Database	Resolution	URL / References
HC27	n/a	https://doi.org/10.7910/DVN/90WJ9W
WISE	n/a	https://www.isric.org/explore/wise-databases
ISRIC	1 km/250 m	https://doi.org/10.5194/soil-7–217–2021
FAO soil maps (DSMW)	1:5 m	http://www.fao.org/land-water/land/land-governance/land-resources-planning-toolbox/category/details/en/c/1026564/
SoilGrids1km	1 km	https://soilgrids.org/
AfSYS-GYGA rooting depth	1 km	https://www.isric.org/documents/document-type/isric-report-201502-root-zone-plant-available-water-holding-capacity-sub
Global maps of soil hydraulic properties	1 km	https://www.futurewater.eu/2015/07/soil-hydraulic-properties
Global High-Resolution Soil Profile Database for Crop Modelling Applications	10 km	https://doi.org/10.7910/DVN/1PEEY0

experiments for parametrization and model evaluation are necessary (Kholová et al., 2020; Silva and Giller, 2021; Sinclair et al., 2020). Models with a smaller number of parameters are likely to be more transparent and easier to use (Sinclair et al., 2020).

4.3. Data collection and limitation

Generally, little information is collected from MET besides crop performance data, to understand the conditions experienced by the crop. This highlights part of the knowledge gap in understanding data requirements across disciplines. For instance, climate data is often not collected as part of MET, even in water-limited environments where rainfall and irrigation (water supply) would be critical to collect. Soil properties (e.g. soil depth, water holding capacity (WHC), nutrients and texture), are seldom collected. The evaluation of crop models for a single location or spatially explicit models covering large expanses of land requires such data, including climate and soil data as well as detailed agronomic information (Hyman et al., 2017). Other than environmental data, measurements advised by stakeholders from other disciplines can enrich the quality of MET data. In traditional phenotyping, most resources are spent on MET to measure yield only (Brown et al., 2020). Secondary phenotypes that could be used as genotype-specific covariables in yield prediction models are still being investigated (Van Eeuwijk et al., 2019). Among these, plant phenology data, in particular flowering and maturity time are of prime importance (Soltani and Sinclair, 2012), and care should be taken that methods often vary among disciplines, e.g. the definition of stem height. Therefore, discussion is required between breeders and other disciplines to ensure additional data is collected from

MET (see Table 1 for minimum and desirable data for modeling purposes).

4.4. Data sharing and limitations

As in other scientific domains, plant scientists worldwide generate data of various types (e.g. quantitative, qualitative, text, computed values), in diverse formats, and in ever-increasing abundance (Leonelli et al., 2017). Although data availability and access to climate and soil data have expanded over the last decade (Hyman et al., 2013), trial and production data are still rarely shared for reuse (Brown et al., 2020). With the variability and uncertainty of field research, promoting open access MET data via repositories and publishers could be key (Brown et al., 2020; Leonelli et al., 2017; Williams, 2012). Data standardization in terms of syntax, semantics, and structure is required to ensure datasets are efficiently usable. An initial effort to link trial data to CSM was developed by researchers of the Agricultural Model Intercomparison Project (AgMIP; Rosenzweig et al., 2013). Below, we make implementation suggestions and provide links to some of the discussed data repositories.

5. A step-by-step process-based modeling approach for breeders

Here we aim to illustrate the systematic characterization process for the development of TPEs using a process-based model.

5.1. Interact with breeders, get input data, run modeling simulations

5.1.1. Define the target area

The analysis should cover, preferably, at least 80% of the current growing area of the target crop within a given TPE (Hajjarpoor et al., 2018, 2021). Ideally, recent time-series production data is used, often provided by Ministries of Agriculture, agriculture organizations, or institutes that offer web-based tools like the District Level Database (DLD). Satellite imagery data is another source of data that can be used to define target regions. It is important to consider regions where there is an increasing trend in the area under cultivation for a specific crop, as these regions may be included in future TPEs. Remote sensing data, censuses, surveys and local expert interviews, as well as combinations of them, can be used to analyze crop geography over time (Hyman et al., 2013). In the absence of data, TPEs can be geographically defined, e.g. within country or state boundaries, without referring to specific cropping areas.

5.1.2. Get genotype parameters

Genotype parameters are coefficients that are part of mathematical equations that represent essential biological/ plant production functions, and are the basis of model functioning. Much of this is readily available in the literature (e.g. Hajjarpoor et al., 2021) or borrowed from similar mechanistic models (e.g. Hajjarpoor et al., 2018). If the value of a parameter has been determined in published research, this data is a good starting point (Soltani and Sinclair, 2012). If not available, parameters must be estimated using standardized protocols and translated into model coefficients. In practice, many of these parameters are fixed,

so the genetic variability is represented by a few tens of parameters (Tardieu et al., 2020). Models that require a step of "calibration" to better fit the model prediction to the observations are considered unsuitable and genotype parameters from these should be avoided.

Getting genotype parameters, if not available, requires field experiments with crops grown without nutrient limitations or biotic adversities, together with weather, soil, and management data. Discussion with breeders is important at that stage to ensure genotype parameters are generated for genotypes that they try to improve, popular cultivars for instance. Records of different phenology stages, time series of aboveground biomass, leaf area index over time, and grain yield are required at a minimum for parametrization. From this, phenology prediction, leaf area development, biomass accumulation, harvest index and yield formation parameters can be estimated (Table 2).

5.1.3. Get environmental data

In plant breeding studies, the first driver of yield variability is the environment (Chenu, 2015), including the climate and the soil. Such data should be obtained for the TPE in question. Many agricultural areas have scarce and poor quality climate data and high soil variability (Brown et al., 2020; Hyman et al., 2013). Recent progress has been made in creating global or continent-specific databases of observed and generated datasets. Many of these databases are freely available online and could support high-resolution simulations (Appendix I and II with web links included). Estimates of crop yield responses to climate variability and change are also subject to the choice of weather data (Parkes et al., 2019). The use of synthetic data sources should be done with caution, and their relevance should always be checked against observed data where possible (Dias and Sentelhas, 2021; Hajjarpoor et al., 2021, 2018). This is particularly important in regions with high topographic variability. Generating a virtual plant, CSM can serve as a quality indicator for such data, whereby an investigator can compare the different outputs of the model when running it with different sources of measured (MWD) and gridded weather data (GWD) (see Hajjarpoor et al., 2018; Mourtzinis et al., 2017). For a given weather or crop parameter, the correlation coefficient (r), the absolute mean error (ME), root mean square error (RMSE) and normalized RMSE (RMSE$_n$%), as well as the coefficient of determination (r^2) according to linear regression, are common measures used to assess agreement and biases between GWD and MWD.

Soil data is somewhat more complex to acquire because of the large spatial variation (Hyman et al., 2013). However, some soil databases like HC27 of IFPRI (Koo and Dimes, 2013) have simplified soil profiles based on three criteria of soil texture, root depth, and organic carbon, to which crop models are the most responsive. The HC27 generic soil map is the result of the collaboration between FAO, IIASA, ISRIC, ISSCAS and JRC, which categorizes soils into 27 profiles (see Appendix 2 for web links). Nehbandani et al. (2020) used a simple CSM to assess the quality of the HC27 soil information, comparing the output of the model using observed soil information. The comparison provided statistically similar results to observed soil data, concluding that HC27 can be used to simulate potential yield and water-related factors. If gridded or generic soil data is not available or accurate enough for a region, information on common soil types and effective rooting depths for the target crop should be gathered through surveys and consultations with local experts.

5.1.4. Get management data

Process-based models require information on agricultural management practices, such as sowing windows, plant density, irrigation and fertilizer use, as well as weed management. The detail of the management information depends on the model. For example, not all are designed to simulate weed competition and management. Such management information can be obtained through close consultation with local agronomists and breeders, or found on global databases, platforms, or scientific literature. Various platforms have been developed to support global-scale assessments, such as the FAO crop calendar tool for 130 crops or a global data set of monthly irrigated and rainfed growing areas for 26 irrigated and rainfed crops (MIRCA2000, Portmann et al., 2010), although management data other than sowing date is not provided. While Sacks et al. (2010) also provide assembled datasets of global crop planting and harvesting dates for 19 major crops, other platforms like the GYGA data sheet are used to collect local crop management data. Specification is needed for dominant water regimes (i.e. rainfed, partially- or fully-irrigated) and the percentage of water availability, which can be found in FAO AQUASTAT, official country-level irrigation data collected by national agricultural research centers (NARC) (e.g. area under irrigation). Recommended rates of fertilizer can also be obtained through NARC, seed companies, agriculture ministries or variety management recommendations by breeders (e.g. Table 1). To reiterate, close interaction with breeding groups is critical here, as they are often in direct contact with the various local agencies.

5.1.5. Evaluate the model

It is necessary to rigorously evaluate the model predictions (phenological stage, grain yield, aboveground dry matter, crop evapotranspiration) against quality observations from different trials under various management practices across environments (Sinclair et al., 2020; van Ittersum et al., 2013). Statistical approaches can be used to quantify such comparisons (e.g. Seidel et al., 2018). Of course, an important feature of robustness is not simply an acceptable prediction of final yield but a realistic representation of the temporal dynamics of the crop growth stages (Soltani and Sinclair, 2012). Another important aspect is the sensitivity of the crop model, i.e. its capacity to reliably predict crop performance across a range of conditions. For instance, water was found to be a main factor affecting groundnut yield throughout India. This led to a fairly large range of observed yields, which the model was able to sensibly predict (see Fig. 2 in Hajjarpoor et al., 2021). However, when it comes to extreme events like frost or heat shocks, more accurate evaluation is needed (Vadez et al., 2016). Most models have a similar trend in simulated crop yields as temperatures increase, but not all of them take direct heat stress effects into account, which could lead to further yield variation (Eitzinger et al., 2013). Not all CSM have been tested for all possible stresses or combined stresses, far from it.

The quality of the observed data, on the other hand, determines the relevance of model outcomes (Kholová et al., 2020; Wing and De Cian, 2014), and the statistical assessment of robustness evaluations help document the expected reliability of model assessment results (Sinclair et al., 2020). When properly designed, used and tested, crop models provide a unique framework to capture impacts within untested environments, such as projected climates.

5.1.6. Generate model runs across a target area

Once the model has been validated, crop model simulations must be run for a specific number of seasons to provide an accurate estimate of growth and development parameters as well as yield at a given location. For valid estimations of yield potential (irrigated) or water-limited yield potential (rainfed) and their variability, 10–30 years of daily weather data are required - the more the better when year-to-year variability in water availability is high. Under less variable conditions, such as irrigated or favorable rainfed environments, the required number of years is lower than in unfavorable environments. Soltani and Hoogenboom (2003) found that statistical characteristics similar to observed data required at least 15 years of simulated weather data. However, based on several studies, Sinclair et al. (2020) stated that up to 30 seasons may be required to capture the full impact of weather variability in simulation results at each location.

5.2. Analyze the output

5.2.1. Measure the yield gaps

The yield gap is the difference (the gap) between yield currently

achieved on farms and the yield that can be achieved by using the best agronomic practices on-station (*in vivo*) or simulated (*in silico*) (Lobell et al., 2009; van Ittersum et al., 2013). The yield gap analysis is a technique to navigate and understand crop production system limits, and to explore strategies to improve yield. For that, multi-year production data in a fairly disaggregated way (e.g. district) is necessary across the TPE. Various sources of gathering production data are covered in section 4.1.1. Here, 10–30 years of production data is usually required depending on the system in question and environment. Key is that it must be sufficient to account for district-to-district and year-to-year variability in actual yield due to weather, especially in water-limited environments (Grassini et al., 2015), while avoiding bias due to the previous technological trends (van Ittersum et al., 2013). Global Yield Gap Atlas also provides a protocol for estimating average actual yields and yield gaps when long-term data is lacking (GYGA protocol).

5.2.2. Delimit environmental types and major abiotic stresses patterns

The types and frequency of major abiotic stresses such as heat, cold, drought, and waterlogging can be obtained from model output. By contrast, pest and disease sub-models have not been well established (Donatelli et al., 2017; Ramirez-Villegas et al., 2020; Reynolds et al., 2018), and their use in defining environmental types remains qualitative. Drought is the most common stress in the semi-arid tropics, and environmental types (level and point in time) can be classified using a water supply/demand ratio (see Kholová et al., 2013; Seyoum et al., 2017) or the simulated water deficit index (see Battisti and Sentelhas, 2019; Hajjarpoor et al., 2021). Lobell et al. (2015) highlighted the impacts of shifting drought and heat stress for sorghum and wheat in northern Australia and emphasized the need to work on multiple-stress resilience. Further examples can be seen globally (Moeletsi et al., 2011; Nelson et al., 2021b). These indices are then clustered towards grouping scenarios. Locations, genotypes, and management all influence the frequency of these patterns.

In terms of identifying areas for specific genotype development, the duration of temperature extremes, the cumulative effects of multiple abiotic stresses, such as heat and frost events, and the interactions with other limiting factors like drought are critical (Ababaei and Chenu, 2019; Chenu, 2015). Therefore, agronomic output from the model, the types and frequency of major abiotic stresses combined with observed geo-biophysical indicators and estimates of yield potentials and gaps, are indicators that can be added to cluster analyses, enriching the characterization of main clusters in the target region. Prior to doing so, as a number of these indicators can be strongly correlated to one another - clustering among them would over-represent some variables (Hajjarpoor et al., 2021) - correlation analysis is recommended to select only non-correlated indicators for the cluster analysis.

5.2.3. Cluster and define homogenous production units (HPUs) ∼ TPE

All non-correlated indicators collected or generated during the previous steps are used as input for the principal component analysis (PCA) and clustering (Chauhan and Rachaputi, 2014; Hajjarpoor et al., 2021, 2018). Principal components (PCs) are linear transformations of original data into uncorrelated variables. Subsequently, a set number of components that explain most of the variability in the dataset are used for the cluster analysis. Geospatial settings with similar characteristics are mathematically put into groups. Settings could be classed as similar based on an area the model is run for, be that at a district or grid level. Individual clusters should have a high level of internal homogeneity (within clusters) and exterior heterogeneity (between clusters) (Priyadarshan, 2019). Cluster analysis usually provides many clusters, although R packages like NbClust (Pašiaková et al., 2013) can determine a cluster number range that most reliably represent variation in the dataset. This is the stage at which a qualitative interaction with plant breeders is necessary to confront the output of the cluster analysis with their understanding and experience of the variation within the TPE. It is

from such qualitative assessments and the cluster number range from the NbClust output that the final number of clusters is optimized as a mix of statistical indices (NbClust) and qualitative interactions (e.g. expert interviews) (Hajjarpoor et al., 2021). The results can be visualized using Geographic Information System software like ArcMap. The same GxExM interactions and production limitations are expected in homogenous production units (HPU) developed by this step-by-step analysis.

5.2.4. Justification of numbers and location optimization for MET

Grouping by HPU can reduce GxExM interactions and help refine breeding targets. In contrast to the conventional strategy of "testing everything everywhere", the HPU insight helps avoid testing genotypes in less relevant sites and the selection of genotypes in abnormal years, which do not sufficiently represent the most frequent environmental scenarios occurring at a particular HPU (Kholová et al., 2020).

A modeling study on groundnut production in India showed how the number and location of testing sites could be optimized according to the homogeneity and heterogeneity of a given area (Hajjarpoor et al., 2021). The study concluded that year-to-year variation at a particular site should be balanced with spatio-temporal variation among sites to find the optimum number and location of sites within an identified HPU. The study also estimated exploitable production gaps where larger ones justify more testing sites. Geolocating trial sites allows for the enrichment of datasets with existing environmental data from different repositories (Brown et al., 2020). The correct design and analysis of METs to reliably represent the reference cultivar populations and TPEs is fundamental to the design of CSM-based prediction methods for plant breeding (Cooper et al., 2020), where a "which-wins-where" pattern across environments can be simulated among crops and cultivars. While "which-wins-where" is important for breeding programs, the inclusion of CSM can help inform us "why" "which-wins-where". In return, existing MET - when complimented with additional data collection to suit CSM use - can help enhance CSM development, evaluation, and robustness.

5.2.5. Design GxM interventions within E

The same simulation setup that is used to establish HPUs may now be utilized to find GxExM interventions that have a high chance of improving crop production in the system. Essentially, coefficient of variation (CV %) of actual and potential yield uses an estimate of the GxM interactions within and among grouped units in comparison to the whole region. Boxplots are common to visualize the distributions of different parameters within each HPU and the entire area. The HPU approach establishes precise geographic regions. Within these regions, specific GxM combinations can be designed by crop improvement programs to boost productivity. Battisti and Sentelhas (2019) is an example how water-deficit patterns can define crop management and breeding strategies by zones. The authors divided Brazil into ten soybean-growing zones based on attainable yield and water-deficit patterns, claiming this to help in the development of better strategies to improve soybean yield in the regions studied. Hajjarpoor et al., (2018, 2021) provide excellent examples of GxM package design within HPUs, which were obtained through clustering geo-biophysical indicators. Nelson et al. (2021a) also demonstrate how process-based models can aid the design of climate-smart production systems such as intercropping with a focus on soil-water use and planting density. Such models allow us to test specific plant types with specific management options in multiple locations and can therefore be used as decision-making tools to quantitatively specify agronomic and breeding objectives.

6. Conclusion

Crop improvement programs need to find more efficient ways to deal with GxExM interactions, especially in highly variable environments. The 'way to go' is not statistical or process-based modeling, but a collaborative way forward for the joint cause of crop performance

improvement. To achieve this, plant breeding research and related research disciplines should form multidisciplinary teams that can improve our understanding of GxExM interactions, and then help breeding decisions. Regardless of the approach, data requirements are high and therefore require multidisciplinary collaboration. While statistical methods demand a large amount of MET data, process-based models require high-quality, detailed data on soil, weather, crop production, and genotype parameters. Which one to choose depends on data availability. This paper proposes a simple step-by-step approach to bring the capacity of process-based models to the doorstep of breeding programs, towards a better definition of target population of environment, within which the clustering of subunits allow for the reduction of GxExM interactions.

Declaration of Competing Interest

Acknowledgments

Appendix I. . Climate data

See appendix TableA1,TableA2.

A major limitation in attempting relevant simulations across a wide geographical area is assembling a weather database of sufficient geographical resolution (Mourtzinis et al., 2017; Van Wart et al., 2015). However, the accuracy of weather-yield relationships derived from crop models is dependent on the quality of the underlying input data used to run these models. In this context, a major challenge in many developing countries is the lack of free accessible and reliable meteorological datasets (Hajjarpoor et al., 2018; Parkes et al., 2019; Vadez et al., 2017). However, in some other countries, there may be plenty of sources with national and meteorological systems. A possible surrogate for regions where weather station network is irregular is the use of gridded weather data (GWD).

Phenotyping programs and G x E interaction assessments can benefit from broad-scale climate analysis to assess the extent to which sites represent target environments. An important recent advancement in climatic analysis is the availability of ready-to-use climate data available via the Internet (Table A1) or in software applications like MARKSIM (Jones et al., 2002; Jones and Thornton, 2000). Within a diverse topography, selecting a GWD with higher spatial resolution (smaller number/grids in Table A1) would be preferable. In addition, going for point data in a heterogeneous environment can be a solution. Historical point data (weather station data) is provided by some inventories like the integrated Agricultural Information and Management System (iAIMS) and the Global Historical Climate Network Daily (NOAA).

Some other databases are also available, but to our knowledge there is a lack of some essential parameters for modeling purposes like the CHIRPS database that was initially developed based on estimates from rain gauge and satellite observations to build high resolution (0.05°) gridded precipitation data (Funk et al., 2015). This type of data is very helpful for hydrometeorology, water resource management, trend analysis, flood and seasonal drought monitoring. In addition, some

modelers use a combination of CHIRPS precipitation and other surface meteorology databases. Also, we only listed databases with daily temporal resolution, however, some databases like Land Data Assimilation System (FLDAS) (Amy McNally, 2018) and Climate Research Unit gridded Time Series (CRU ST) (Harris et al., 2020) are available at a monthly temporal resolution. In this case, it is possible to use some weather generators integrated into crop models like WeatherMan (in DSSAT, Pickering et al., 1994) or standalone models such as WGEN and LARS-WG to generate daily time series data (Semenov et al., 1998). These weather generators can also be used to downscale large-scale climate projections from global climate models (GCMs) to local-scale climate scenarios for impact assessments (Semenov and Stratonovitch, 2015).

However, these reference datasets are also known to contain important biases and uncertainties according to the source type and method to propagate data (Reanalysis of climate model data, remote sensing and extrapolation method of in situ gauges). The user, then, has to be cautious using these synthetic data sources and always crosscheck their relevance against ground-truthed information for the intended modeling exercise (Dias and Sentelhas, 2021; Hajjarpoor et al., 2021, 2018). Parkes et al. (2019) reviewed and checked the uncertainty of estimating crop yield responses to different weather dataset. Ruane et al. (2015) also compared different databases of weather data for agricultural modeling.

Appendix II. . Soil data

Data on soil properties are a key category of information for agroecological assessments (Table A2). However, advances in the development of soil datasets are hindered by the difficulty of mapping the entire world. The main problem is that soils can be highly variable even across short distances. Moreover, not all countries use the same soil classification systems. The concept of the likelihood or probability of finding a given soil property has been used to reflect data uncertainty at a particular point when using maps like the FAO 1:5 million soil map of the World. This map remains the most widely used soil map for continental and global applications.

According to van Ittersum et al. (2013), simulations should be run for the dominant soils in a defined area and expert opinion from country agronomists should be used.

References

Ababaei, B., Chenu, K., 2019. Recent trends in drought, heat and frost-induced yield losses across the Australian wheatbelt. Proceedings 36, 5. https://doi.org/10.3390/proceedings2019036005.

Barlow, K.M., Christy, B.P., O'Leary, G.J., Riffkin, P.A., Nuttall, J.G., 2015. Simulating the impact of extreme heat and frost events on wheat crop production: A review. F. Crop. Res. 171, 109–119. https://doi.org/10.1016/j.fcr.2014.11.010.

Battisti, R., Sentelhas, P.C., 2019. Characterizing Brazilian soybean-growing regions by water deficit patterns. F. Crop. Res. 240, 95–105. https://doi.org/10.1016/j.fcr.2019.06.007.

Boote, K.J., Jones, J.W., Hoogenboom, G., 2021. Incorporating realistic trait physiology into crop growth models to support genetic improvement. Silico Plants 3. https://doi.org/10.1093/insilicoplants/diab002.

Brown, D., Van Den Bergh, I., De Bruin, S., Machida, L., Van Etten, J., 2020. Data synthesis for crop variety evaluation. A review. Agron. Sustain. Dev. 40. https://doi.org/10.1007/s13593-020-00630-7.

Bustos-Korts, D., Romagosa, I., Borràs-gelonch, G., Casas, A.M., Slafer, G.A., Eeuwijk, F. Van, 2019. Genotype Environ. Interact. Adapt. https://doi.org/10.1007/978-1-4939-2493-6_199-3.

Bustos-korts, D., Romagosa, I., Borràs-gelonch, G., Casas, A.M., Slafer, G.A., Eeuwijk, F. Van, 2019. Genotype by Environment Interaction andAdaptationGenotype Environ. Interact. Adapt. https://doi.org/10.1007/978-1-4939-2493-6_199-3.

Ceccarelli, S., 1989. Wide adaptation: How wide? Euphytica 40, 197–205. https://doi.org/10.1007/BF00024512.

Ceccarelli, S., Grando, S., Baum, M., 2007. Participatory plant breeding in water-limited environments. Exp. Agric. 43, 411–435. https://doi.org/10.1017/s0014479707005327.

Ceccarelli, S., Grando, S., Maatougui, M., Michael, M., Slash, M., Haghparast, R., Rahmanian, M., Taheri, A., Al-Yassin, A., Benbelkacem, A., Labdi, M., Mimoun, H., Nachit, M., 2010. Plant breeding and climate changes. J. Agric. Sci. 148, 627–637. https://doi.org/10.1017/s0021859610000651.

Chapman, S.C., Hammer, G.L., Butler, D.G., Cooper, M., 2000. Genotype by environment interactions affecting grain sorghum. III. Temporal sequences and spatial patterns in the target population of environments. Aust. J. Agric. Res. 51, 223. https://doi.org/10.1071/AR99022.

Chauhan, Y.S., Rachaputi, R.C.N., 2014. Defining agro-ecological regions for field crops in variable target production environments: A case study on mungbean in the northern grains region of Australia. Agric. . Meteorol. 194, 207–217. https://doi.org/10.1016/j.agrformet.2014.04.007.

Chenu, K., 2015. Characterizing the crop environment - nature, significance and applications, in: Crop Physiology. Applications for Genetic Improvement and Agronomy, Second edition.,. Academic Press,, pp. 321–348. https://doi.org/10.1016/B978-0-12-417104-6.00013-3.

Chenu, K., Cooper, M., Hammer, G.L., Mathews, K.L., Dreccer, M.F., Chapman, S.C., 2011. Environment characterization as an aid to wheat improvement: interpreting genotype-environment interactions by modelling water-deficit patterns in North-Eastern Australia. J. Exp. Bot. 62, 1743–1755. https://doi.org/10.1093/jxb/erq459.

Comstock, R.E., 1977. Quantitative genetics and the design of breeding programs. In: Pollak, E. (Ed.), Proceedings of the International Conference on Quantitative Genetics. Iowa State University Press, Ames, USA, pp. 705–718.

Cooper, M., Gho, C., Leafgren, R., Tang, T., Messina, C., 2014. Breeding drought-tolerant maize hybrids for the US corn-belt: Discovery to product. J. Exp. Bot. 65, 6191–6194. https://doi.org/10.1093/jxb/eru064.

Cooper, M., Powell, O., Voss-Fels, K.P., Messina, C.D., Gho, C., Podlich, D.W., Technow, F., Chapman, S.C., Beveridge, C.A., Ortiz-Barrientos, D., Hammer, G.L., 2021a. Modelling selection response in plant-breeding programs using crop models as mechanistic gene-to-phenotype (CGM-G2P) multi-trait link functions. Silico Plants 3. https://doi.org/10.1093/insilicoplants/diaa016.

Cooper, M., Stucker, R.E., DeLacy, I.H., Harch, B.D., 1997. Wheat breeding nurseries, target environments, and indirect selection for grain yield. Crop Sci. 37. https://doi.org/10.2135/cropsci1997.0011183X003700040024x.

Cooper, M., Tang, T., Gho, C., Hart, T., Hammer, G., Messina, C., 2020. Integrating genetic gain and gap analysis to predict improvements in crop productivity. Crop Sci. 1–23. https://doi.org/10.1002/csc2.20109.

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

Cooper, M., Voss-Fels, K.P., Messina, C.D., Tang, T., Hammer, G.L., 2021b. Tackling G × E × M interactions to close on-farm yield-gaps: creating novel pathways for crop improvement by predicting contributions of genetics and management to crop productivity. Theor. Appl. Genet. https://doi.org/10.1007/s00122-021-03812-3.

Cooper, M., Woodruff, D.R., 1993. Predicting grain yield in Australian environments using data from CIMMYT international wheat performance trials 3. Testing predicted correlated response to selection. F. Crop. Res. 35, 191–204. https://doi.org/10.1016/0378-4290(93)90153-E.

Cooper, M., Woodruff, D.R., Eisemann, R.L., Brennan, P.S., Delacy, I.H., 1995. A selection strategy to accommodate genotype-by-environment interaction for grain yield of wheat: managed-environments for selection among genotypes. Theor. Appl. Genet. 90, 492–502. https://doi.org/10.1007/bf00221995.

De Wit, A., Boogaard, H., Fumagalli, D., Janssen, S., Knapen, R., Van Kraalingen, D., Supit, I., Van Der Wijngaart, R., Van Diepen, K., 2019. 25 years of the WOFOST cropping systems model. Agric. Syst. 168, 154–167. https://doi.org/10.1016/j.agsy.2018.06.018.

DeLacy, I.H., Basford, K.E., Cooper, M., Bull, J.K., McLaren, C.G., 1996. Analysis of multi-environment trials–an historical perspective. Plant Adapt. Crop Improv 39124.

Dias, H.B., Sentelhas, P.C., 2021. Assessing the performance of two gridded weather data for sugarcane crop simulations with a process-based model in Center-South Brazil. Int. J. Biometeorol. https://doi.org/10.1007/s00484-021-02145-6.

Donatelli, M., Magarey, R.D., Bregaglio, S., Willocquet, L., Whish, J.P.M., Savary, S., 2017. Modelling the impacts of pests and diseases on agricultural systems. Agric. Syst. 155, 213–224. https://doi.org/10.1016/j.agsy.2017.01.019.

Eitzinger, J., Thaler, S., Schmid, E., Strauss, F., Ferrise, R., Moriondo, M., Bindi, M., Palosuo, T., Rötter, R., Kersebaum, K.C., Olesen, J.E., Patil, R.H., Şaylan, L., Çaldağ, B., Çaylak, O., 2013. Sensitivities of crop models to extreme weather conditions during flowering period demonstrated for maize and winter wheat in Austria. J. Agric. Sci. 151, 813–835. https://doi.org/10.1017/S0021859612000779.

Finlay, K., Wilkinson, G., 1963. The analysis of adaptation in a plant-breeding programme. Aust. J. Agric. Res. 14, 742. https://doi.org/10.1071/ar9630742.

Fischer, T., Byerlee, D., Edmeades, G., 2014, Crop yields and global food security: will yield increase continue to feed the world?, Australian Centre for International Agricultural Research.

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., Michaelsen, J., 2015. The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. Sci. Data 2, 150066. https://doi.org/10.1038/sdata.2015.66.

George, T., 2014. Why crop yields in developing countries have not kept pace with advances in agronomy. Glob. Food Sect. 3 49–58. https://doi.org/10.1016/j.gfs.2013.10.002.

Grassini, P., van Bussel, L.G.J., Van Wart, J., Wolf, J., Claessens, L., Yang, H., Boogaard, H., de Groot, H., van Ittersum, M.K., Cassman, K.G., 2015. How good is good enough? Data requirements for reliable crop yield simulations and yield-gap analysis. F. Crop. Res. 177, 49–63. https://doi.org/10.1016/j.fcr.2015.03.004.

Hajjarpoor, A., Kholová, J., Pasupuleti, J., Soltani, A., Burridge, J., Degala, S.B., Gattu, S., Murali, T.V., Garin, V., Radhakrishnan, T., Vadez, V., 2021. Environmental characterization and yield gap analysis to tackle genotype-by-environment-by-management interactions and map region-specific agronomic and breeding targets in groundnut. F. Crop. Res. 267, 108160 https://doi.org/10.1016/j.fcr.2021.108160.

Hajjarpoor, A., Vadez, V., Soltani, A., Gaur, P., Whitbread, A., Suresh Babu, D., Gumma, M.K., Diancoumba, M., Kholová, J., 2018. Characterization of the main chickpea cropping systems in India using a yield gap analysis approach. F. Crop. Res. 223, 93–104. https://doi.org/10.1016/j.fcr.2018.03.023.

Hammer, G., Cooper, M., Tardieu, F., Welch, S., Walsh, B., van Eeuwijk, F., Chapman, S., Podlich, D., 2006. Models for navigating biological complexity in breeding improved crop plants. Trends Plant Sci. 11, 587–593. https://doi.org/10.1016/j.tplants.2006.10.006.

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

Harris, I., Osborn, T.J., Jones, P., Lister, D., 2020. Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. Sci. Data 7. https://doi.org/10.1038/s41597-020-0453-3.

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

Heinemann, A.B., Barrios-Perez, C., Ramirez-Villegas, J., Arango-Londoño, D., Bonilla-Findji, O., Medeiros, J.C., Jarvis, A., 2015. Variation and impact of drought-stress patterns across upland rice target population of environments in Brazil. J. Exp. Bot. 66, 3625–3638. https://doi.org/10.1093/jxb/erv126.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., Thépaut, J., 2020. The ERA5 global reanalysis. Q. J. R. Meteorol. Soc. 146, 1999–2049. https://doi.org/10.1002/qj.3803.

Hyman, G., Espinosa, H., Camargo, P., Abreu, D., Devare, M., Arnaud, E., Porter, C., Mwanzia, L., Sonder, K., Traore, S., 2017. Improving agricultural knowledge management: The AgTrials experience. F1000Research 6, 317. https://doi.org/10.12688/f1000research.11179.2.

Hyman, G., Hodson, D., Jones, P., 2013. Spatial analysis to support geographic targeting of genotypes to environments. Front. Physiol. 4. https://doi.org/10.3389/fphys.2013.00040.

IRRI, 2006, Defining the target environment [WWW Document]. URL ⟨http://www.knowledgebank.irri.org/ricebreedingcourse/Lesson_2_Setting_goals_and_identifying_the_target_environment.htm⟩ (accessed 16 September 2020).

Jones, P.G., Thornton, P.K., 2000. MarkSim: software to generate daily weather data for Latin America and Africa. Agron. J. 92, 445–453.

Jones, P.G., Thornton, P.K., Diaz, W., Wilkens, P.W., Jones, A.L., 2002. MarkSim: A computer tool that generates simulated weather data for crop modeling and risk assessment. Cent. Int. Agric. Trop. Cali, Colomb.

Kaloki, P., Luo, Q., Trethowan, R., Tan, D.K.Y., 2019. Can the development of drought tolerant ideotype sustain Australian chickpea yield? Int. J. Biometeorol. 63, 393–403. https://doi.org/10.1007/s00484-019-01672-7.

Kholová, J., Adam, M., Diancoumba, M., Marrou, H., Hammer, G., Hajjarpoor, A., Chenu, K., 2020. Sorghum; General crop modelling tools guiding principles and use of crop models in support of crop improvement programs in developing countries. In: Tonapi, V.A., Talwar, H.S., Kumar, A., Bhat, B.V., Reddy, C.R., Dalton, T.J. (Eds.), Sorghum in the 21st Century: Food – Fodder – Feed – Fuel for a Rapidly Changing World. Springer, Singapore, pp. 189–207. https://doi.org/10.1007/978-981-15-8249-3.

Kholová, J., McLean, G., Vadez, V., Craufurd, P., Hammer, G.L., 2013. Drought stress characterization of post-rainy season (rabi) sorghum in India. F. Crop. Res. 141, 38–46. https://doi.org/10.1016/j.fcr.2012.10.020.

Koo, J., Dimes, J., 2013. HC27 Generic Soil Profile Database. https://doi.org/10.7910/DVN/90WJ9W.

Leonelli, S., Davey, R.P., Arnaud, E., Parry, G., Bastow, R., 2017. Data management and best practice for plant science. Nat. Plants 3. https://doi.org/10.1038/nplants.2017.86.

Lobell, D.B., Cassman, K.G., Field, C.B., 2009. Crop Yield Gaps: Their Importance, Magnitudes, and Causes. Annu. Rev. Environ. Resour. 34, 179–204. https://doi.org/10.1146/annurev.environ.041008.093740.

Lobell, D.B., Hammer, G.L., Chenu, K., Zheng, B., Mclean, G., Chapman, S.C., 2015. The shifting influence of drought and heat stress for crops in northeast Australia. Glob. Chang. Biol. 21, 4115–4127. https://doi.org/10.1111/gcb.13022.

Malosetti, M., Ribaut, J.-M., Van Eeuwijk, F.A., 2013. The statistical analysis of multi-environment data: modeling genotype-by-environment interaction and its genetic basis. Front. Physiol. 4. https://doi.org/10.3389/fphys.2013.00044.

Messina, C.D., Cooper, M., Hammer, G.L., Berning, D., Ciampitti, I., Clark, R., Diepenbrock, C., Gho, C., Jines, M., Lee, T., McCormick, R., Mihura, E., Podlich, D., Rotundo, J., Smalley, M., Tang, T., Truong, S., van Eeuwijk, F., 2020a. Two decades of creating drought tolerant maize and underpinning prediction technologies in the US corn-belt: Review and perspectives on the future of crop design. bioRxiv. https://doi.org/10.1101/2020.10.29.361337.

Messina, C.D., Cooper, M., Reynolds, M., Hammer, G.L., 2020b. Crop science: A foundation for Advancing Predictive Agriculture. Crop Sci. N./a. https://doi.org/10.1002/csc2.20116.

Messina, C.D., Sinclair, T.R., Hammer, G.L., Curan, D., Thompson, J., Oler, Z., Gho, C., Cooper, M., 2015. Limited-Transpiration Trait May Increase Maize Drought Tolerance in the US Corn Belt. Agron. J. 107, 1978–1986. https://doi.org/10.2134/agronj15.0016.

Moeletsi, M.E., Walker, S., Landman, W.A., 2011. ENSO and implications on rainfall characteristics with reference to maize production in the Free State Province of South Africa. Phys. Chem. Earth, Parts A/B/C. 36, 715–726. https://doi.org/10.1016/j.pce.2011.07.043.

Mourtzinis, S., Rattalino Edreira, J.I., Conley, S.P., Grassini, P., 2017. From grid to field: Assessing quality of gridded weather data for agricultural applications. Eur. J. Agron. 82, 163–172. https://doi.org/10.1016/j.eja.2016.10.013.

Nehbandani, A., Soltani, A., Taghdisi, R., Amir, N., Seyyed, D., Alimagham, M., 2020. Assessing HC27 Soil Database for Modeling Plant Production. Int. J. Plant Prod. https://doi.org/10.1007/s42106-020-00114-4.

Nelson, W.C.D., Hoffmann, M.P., Vadez, V., Rötter, R.P., Koch, M., Whitbread, A.M., 2021a. Can intercropping be an adaptation to drought? A model-based analysis for pearl millet–cowpea. J. Agron. Crop Sci. 1–18. https://doi.org/10.1111/jac.12552.

Nelson, W.C.D., Siebrecht-Schöll, D.J., Hoffmann, M.P., Rötter, R.P., Whitbread, A.M., Link, W., 2021b. What determines a productive winter bean-wheat genotype combination for intercropping in central Germany? Eur. J. Agron. 128, 126294 https://doi.org/10.1016/j.eja.2021.126294.

Parkes, B., Higginbottom, T.P., Hufkens, K., Ceballos, F., Kramer, B., Foster, T., 2019. Weather dataset choice introduces uncertainty to estimates of crop yield responses to climate variability and change. Environ. Res. Lett. 14, 124089 https://doi.org/10.1088/1748-9326/ab5ebb.

Pašiaková, M., Gajdoš, V., Bučová, J., Brixová, B., Bielik, M., 2013. NbClust: An R Package for Determining the Relevant Number of Clusters in a Data Set. Acta Geol. Slov. 5, 45–54.

Pickering, N.B., Hansen, J.W., Jones, J.W., Wells, C.M., Chan, V.K., Godwin, D.C., 1994. WeatherMan: A utility for managing and generating daily weather data. Agron. J. 86, 332–337. https://doi.org/10.2134/agronj1994.00021962008600020003x.

Porter, J.R., Challinor, A., Ewert, F., Falloon, P., Fischer, T., Gregory, P., Van Ittersum, M. K., Olesen, J.E., Moore, K.J., Rosenzweig, C., 2010. Food security: focus on agriculture. Sci. (80-.) 328, 172–173.

Portmann, F.T., Siebert, S., Döll, P., 2010. MIRCA2000-Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling. Glob. Biogeochem. Cycles 24, N./a-N./a. https://doi.org/10.1029/2008gb003435.

Priyadarshan, P.M., 2019. PLANT BREEDING: Classical to Modern, 1st ed..,. Springer,, Singapore, Singapore. https://doi.org/10.1007/978-981-13-7095-3.

Qiao, C.G., Basford, K.E., Delacy, I.H., Cooper, M., 2004. Advantage of single-trial models for response to selection in wheat breeding multi-environment trials. Theor. Appl. Genet. 108, 1256–1264. https://doi.org/10.1007/s00122-003-1541-4.

Rajaram, S.M., van Ginkel, M., Fischer, R.A., 1994, CIMMYT's wheat breeding mega-environments, Proceedings of the 8th International Wheat Genetics Symposium.

Ramirez-Villegas, J., Molero Milan, A., Alexandrov, N., Asseng, S., Challinor, A.J., Crossa, J., Eeuwijk, F., Ghanem, M.E., Grenier, C., Heinemann, A.B., Wang, J., Juliana, P., Kehel, Z., Kholova, J., Koo, J., Pequeno, D., Quiroz, R., Rebolledo, M.C., Sukumaran, S., Vadez, V., White, J.W., Reynolds, M., 2020. CGIAR modeling approaches for resource-constrained scenarios: I. Accelerating crop breeding for a changing climate. Crop Sci. 60, 547–567. https://doi.org/10.1002/csc2.20048.

Reynolds, M., Kropff, M., Crossa, J., Koo, J., Kruseman, G., Molero Milan, A., Rutkoski, J., Schulthess, U., Singh, B., Sonder, K., Tonnang, H., Vadez, V., 2018. Role of modelling in international crop research: Overview and some case studies. Agronomy 8. https://doi.org/10.3390/agronomy8120291.

Rosenzweig, C., Jones, J.W., Hatfield, J.L., Ruane, A.C., Boote, K.J., Thorburn, P., Antle, J.M., Nelson, G.C., Porter, C., Janssen, S., Asseng, S., Basso, B., Ewert, F., Wallach, D., Baigorria, G., Winter, J.M., 2013. The Agricultural Model Intercomparison and Improvement Project (AgMIP): Protocols and pilot studies. Agric. . Meteorol. 170, 166–182. https://doi.org/10.1016/j.agrformet.2012.09.011.

Ruane, A.C., Goldberg, R., Chryssanthacopoulos, J., 2015. Climate forcing datasets for agricultural modeling: Merged products for gap-filling and historical climate series estimation. Agric. . Meteorol. 200, 233–248. https://doi.org/10.1016/j.agrformet.2014.09.016.

Sacks, W.J., Deryng, D., Foley, J.A., Ramankutty, N., 2010. Crop planting dates: an analysis of global patterns. Glob. Ecol. Biogeogr. no-no. https://doi.org/10.1111/j.1466-8238.2010.00551.x.

Sciarresi, C., Patrignani, A., Soltani, A., Sinclair, T., Lollato, R.P., 2019. Plant traits to increase winter wheat yield in semiarid and subhumid environments. Agron. J. 111, 1728–1740. https://doi.org/10.2134/agronj2018.12.0766.

Seidel, S.J., Palosuo, T., Thorburn, P., Wallach, D., 2018. Towards improved calibration of crop models – Where are we now and where should we go? Eur. J. Agron. 94, 25–35. https://doi.org/10.1016/j.eja.2018.01.006.

Semenov, M.A., Brooks, R.J., Barrow, E.M., Richardson, C.W., 1998. Comparison of the WGEN and LARS-WG stochastic weather generators for diverse climates. Clim. Res. 10, 95–107. https://doi.org/10.3354/cr010095.

Semenov, M.A., Stratonovitch, P., 2015. Adapting wheat ideotypes for climate change: Accounting for uncertainties in CMIP5 climate projections. Clim. Res. 65, 123–139. https://doi.org/10.3354/cr01297.

Seyoum, S., Chauhan, Y., Rachaputi, R., Fekybelu, S., Prasanna, B., 2017. Characterising production environments for maize in eastern and southern Africa using the APSIM Model. Agric. . Meteorol. 247, 445–453. https://doi.org/10.1016/j.agrformet.2017.08.023.

Sheffield, J., Goteti, G., Wood, E.F., 2006. Development of a 50-Year High-Resolution Global Dataset of Meteorological Forcings for Land Surface Modeling. J. Clim. 19, 3088–3111. https://doi.org/10.1175/jcli3790.1.

Silva, J.V., Giller, K.E., 2021. Grand challenges for the 21st century: what crop models can and can't (yet) do. J. Agric. Sci. 1–12. https://doi.org/10.1017/S0021859621000150.

Sinclair, T.R., Messina, C.D., Beatty, A., Samples, M., 2010. Assessment across the United States of the Benefits of Altered Soybean Drought Traits. Agron. J. 102, 475–482. https://doi.org/10.2134/agronj2009.0195.

Sinclair, T.R., Soltani, A., Marrou, H., Ghanem, M., Vadez, V., 2020. Geospatial assessment for crop physiological and management improvements with examples using the simple simulation model. Crop Sci. N./a. https://doi.org/10.1002/csc2.20106.

Soltani, A., Hoogenboom, G., 2003. Minimum data requirements for parameter estimation of stochastic weather generators. Clim. Res. 25, 109–119.

Soltani, A., Sinclair, T.R., 2012. Modeling Physiology of Crop Development, Growth and Yield. CABI,, Wallingford. https://doi.org/10.1079/9781845939700.0000.

Tardieu, F., Granato, I.S.C., Van Oosterom, E.J., Parent, B., Hammer, G.L., 2020. Are crop and detailed physiological models equally 'mechanistic' for predicting the genetic variability of whole-plant behaviour? Nexus Mech. Adapt. Strateg. In Silico Plants 2. https://doi.org/10.1093/insilicoplants/diaa011.

Vadez, V., Halilou, O., Hissene, H.M., Sibiry-Traore, P., Sinclair, T.R., Soltani, A., 2017. Mapping Water Stress Incidence and Intensity, Optimal Plant Populations, and Cultivar Duration for African Groundnut Productivity Enhancement. Front. Plant Sci. 8, 432. https://doi.org/10.3389/fpls.2017.00432.

Vadez, V., Jana, K., Suresh, D., 2016, Probability of Temperature exceedance in Groundnut during floweirng period at its threshold temperature (heat stress) in India. https://doi.org/doi:10.21421/D2/XQ2MTV.

Van Bueren, E.L., Struik, P.C., Van Eekeren, N., Nuijten, E., 2018. Towards resilience through systems-based plant breeding. A review. Agron. Sustain. Dev. 38. https://doi.org/10.1007/s13593-018-0522-6.

Van Eeuwijk, F.A., Bustos-korts, D., Millet, E.J., Boer, M.P., Kruijer, W., Thompson, A., Malosetti, M., Iwata, H., Quiroz, R., Kuppe, C., Muller, O., Blazakis, K., Yu, K., Tardieu, F., Chapman, S.C., 2019. Modelling strategies for assessing and increasing the effectiveness of new phenotyping techniques in plant breeding. Plant Sci. 282, 23–39. https://doi.org/10.1016/j.plantsci.2018.06.018.

Van Eeuwijk, F.A., Bustos-Korts, D.V., Malosetti, M., 2016. What Should Students in Plant Breeding Know About the Statistical Aspects of Genotype × Environment Interactions? Crop Sci. 56, 2119–2140. https://doi.org/10.2135/cropsci2015.06.0375.

Van Etten, J., De Sousa, K., Aguilar, A., Barrios, M., Coto, A., Dell'Acqua, M., Fadda, C., Gebrehawaryat, Y., Van De Gevel, J., Gupta, A., Kiros, A.Y., Madriz, B., Mathur, P., Mengistu, D.K., Mercado, L., Nurhisen Mohammed, J., Paliwal, A., Pè, M.E., Quirós, C.F., Rosas, J.C., Sharma, N., Singh, S.S., Solanki, I.S., Steinke, J., 2019. Crop variety management for climate adaptation supported by citizen science. Proc. Natl. Acad. Sci. 116, 4194–4199. https://doi.org/10.1073/pnas.1813720116.

van Ittersum, M.K., Cassman, K.G., Grassini, P., Wolf, J., Tittonell, P., Hochman, Z., 2013. Yield gap analysis with local to global relevance—A review. F. Crop. Res. 143, 4–17. https://doi.org/10.1016/j.fcr.2012.09.009.

Van Wart, J., Grassini, P., Yang, H., Claessens, L., Jarvis, A., Cassman, K.G., 2015. Creating long-term weather data from thin air for crop simulation modeling. Agric. . Meteorol. 209–210, 49–58. https://doi.org/10.1016/j.agrformet.2015.02.020.

van Wart, J., Kersebaum, K.C., Peng, S., Milner, M., Cassman, K.G., 2013. Estimating crop yield potential at regional to national scales. F. Crop. Res. 143, 34–43. https://doi.org/10.1016/j.fcr.2012.11.018.

Wang, E., Brown, H.E., Rebetzke, G.J., Zhao, Z., Zheng, B., Chapman, S.C., 2019. Improving process-based crop models to better capture genotype×environment×management interactions. J. Exp. Bot. 70, 2389–2401. https://doi.org/10.1093/jxb/erz092.

Weedon, G.P., Balsamo, G., Bellouin, N., Gomes, S., Best, M.J., Viterbo, P., 2014. The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data. Water Resour. Res. 50, 7505–7514. https://doi.org/10.1002/2014WR015638.

Williams, S.C., 2012. Data practices in the crop sciences: a review of selected faculty publications. J. Agric. Food Inf. 13, 308–325. https://doi.org/10.1080/10496505.2012.717846.

Windhausen, V.S., Wagener, S., Magorokosho, C., Makumbi, D., Vivek, B., Piepho, H.-P., Melchinger, A.E., Atlin, G.N., 2012. Strategies to Subdivide a Target Population of Environments: Results from the CIMMYT-Led Maize Hybrid Testing Programs in Africa. Crop Sci. 52, 2143–2152. https://doi.org/10.2135/cropsci2012.02.0125.

Wing, I.S., De Cian, E., 2014. Modelling agricultural adaptation. Nat. Clim. Chang. 4, 535–536. https://doi.org/10.1038/nclimate2287.

Witcombe, J.R., Virk, D.S., Farrington, J., 1998, Seeds of choice: Making the most of new varieties for small farmers.