Field Crops Research 277 (2022) 108422

Contents lists available at ScienceDirect

Field Crops Research

journal homepage: www.elsevier.com/locate/fcr

An APSIM-powered framework for post-rainy sorghum-system design in India

Swarna Ronanki [a,1,2], Jan Pavlík [b,1,3], Jan Masner [b,*,4], Jan Jarolímek [b,5], Michal Stočes [b,6], Degala Subhash [c], Harvinder S. Talwar [a], Vilas A. Tonapi [a], Mallayee Srikanth [c], Rekha Baddam [c], Jana Kholová [c,*,7]

[a] ICAR, Indian Institute of Millets Research, Hyderabad 500 030, Telangana, India
[b] Department of Information Technologies, Faculty of Economics and Management, Czech University of Life Sciences Prague, Kamýcká 129, Prague 165 00, Czech Republic
[c] GEMS team, International Crops Research Institute for the Semi-Arid Tropics, Patancheru, Hyderabad 5023204, Telangana, India

ARTICLE INFO

Keywords:
APSIM
Post-rainy
Sorghum
GxExM
Agri-system design

ABSTRACT

Sorghum contributes to the livelihoods of millions of food-insecure households in semi-arid agri-systems. Annual production widely fluctuates throughout India due to erratic rainfall and suboptimal agronomic practices. Our novel approach utilizes the digital reflection of post-rainy (rabi) sorghum production systems in India to help better understand its spatio-temporal variations and enable the designing of geography-specific, climate-responsive system interventions (Genotype × Management; GxM). For this, we evaluated a range of farmer-relevant agronomic management practices across three soil types (shallow, medium, and deep vertisols) in combination with observed ranges of biological variability in sorghum cultivar characteristics. We used the crop growth simulation model Agricultural Production Systems sIMulator (APSIM) to identify GxM combinations that can support the enhancement/ stability of post-rainy sorghum production systems in India. In general, we found the post-rainy sorghum systems would benefit from early-season sowing (16th - 23rd September), short crop duration (compared to Maldandi (M35–1), commonly grown crop type), and medium fertilizer inputs (70–70 kg urea ha^{-1} as basal and top-dress application). In addition, site-specific crop management (M) and crop characters (G) optimizations would further enhance/ stabilize sorghum production. Simulations highlighted that in the poorly-endowed environmental context (i.e. shallow soils and low-rainfall areas), optimal G×M targets might involve water conservation GxM combinations, such as low plant populations and low fertilization along with low crop vigor and limited transpiration responsiveness. Details on site-specific optimum GxM are available in a web application at https://ls40.pef.czu.cz/maps/. To enable the use of the study outputs for certain applications (e.g. breeding), we separated the examined geographies based on similarities in optimum production characteristics and similarities in system response to GxM interventions into four "homogeneous system units" (HSU; i.e. geographical units within which reduced GxM interactions are expected). These HSUs intended to offer geography-specific targets to prioritize, test, and introduce distinct G×M interventions. We conclude that the APSIM-powered framework presented provides region-specific Genotype × Management options that could become a blueprint for defining quantitative breeding targets that achieve enhanced productivity/ stability of dry-season sorghum cultivation throughout India.

* Corresponding authors.
[1] These authors have contributed equally
[2] ORCID: 0000-0003-1606-5522
[3] ORCID: 0000-0002-6136-0785
[4] ORCID: 0000-0003-4593-2306
[5] ORCID: 0000-0001-7194-3055
[6] ORCID: 0000-0001-7128-1071
[7] ORCID: 0000-0001-7133-1382

https://doi.org/10.1016/j.fcr.2021.108422
Received 15 April 2021; Received in revised form 13 December 2021; Accepted 19 December 2021
Available online 10 January 2022
0378-4290/

1. Introduction

Sorghum (*Sorghum bicolor* (L.) Moench) is a largely variable crop species adapted to cultivation across tropical to temperate climates and grown primarily for human food and animal feed as well as for the production of biofuels (Habyarimana et al., 2019). In India, sorghum is one of the few multi-purpose, resilient crops suitable for marginal lands during the post-rainy (*rabi*) season (typically September - January) and supports the livelihoods of millions. Frequent droughts caused by climatic variability combined with low input agronomic practices are the main reasons why the farmer's yields across the rabi sorghum tract fluctuate and the average grain yields stagnate at ~800 kg ha^{-1} despite yield potential being much higher (~3500 kg ha^{-1}; Ambadi et al., 2018; Dayakar Rao et al., 2009). A sensible way to bridge this yield gap is to analyze the major constraints of the production system (Kholová et al., 2013) and design the appropriate Genotype × Management (G×M) interventions to lift current yields closer to their potential (Soltani et al., 2016; Pradhan et al., 2015; Chauhan and Rachaputi, 2014; Kholová et al., 2014). Traditionally, multi-location field trials are used to evaluate cultivar, management, and environment interactions (GxExM) in-situ. However, field trials are time and resource-consuming and results are often difficult to extrapolate to other sites and seasons. In this situation, validated crop modeling set-ups in conjunction with field data can extrapolate the GxExM analyses across the spatio-temporal scales and can be used to capture the system's behavior and fluctuating GxExM interactions. By doing so, this approach can complement in-vivo field observations with in-silico predictions which could not be covered experimentally (Jones et al., 2017). For sorghum, several crop models have been implemented in simulation frameworks such as the Decision Support System for Agrotechnology Transfer (DSSAT) (Jones et al., 2003), Agricultural Production Systems sIMulator (APSIM) (Holzworth et al., 2015) or Samara (Dingkuhn et al., 2011). These models differ in the implementation of algorithms to capture the soil-crop-atmosphere interactions.

In prior studies, we found that an APSIM based set-up can reliably reflect the agronomy of post-rainy season sorghum production systems (Kholová et al., 2013, 2014). Therefore, in our present work, we aim to expand the existing post-rainy sorghum simulation set-up and deploy the structure in order to identify the optimum Genotype × Management options to improve/ stabilize sorghum production. This analysis is intended to support crop improvement program decision making on region-specific crop and management interventions that can potentially improve/ stabilize production across the rabi sorghum tract in India. This is presented in the form of an open-access interactive web-based tool to ensure stakeholders access and use.

2. Materials and methods

2.1. Overview

The majority of the Indian rabi sorghum grain is produced in Maharashtra, Karnataka, Andhra Pradesh, and Telangana (as per Kholová et al. (2013) and is the unique production system prioritized for this study. The parameters for three soil types characteristic of these major production areas were collated from available databases (National Bureau of Soil Survey and Land Use Planning, International Soil Reference and Information Centre). The gridded meteorological information was obtained from Agricultural Modern-Era Retrospective Analysis for Research and Applications (AgMERRA) – National Aeronautics and Space Administration (NASA) and evaluated as most suitable when tested against the observed meteorological information (also in Hajjarpoor et al., 2018). Crop simulations were developed using the sorghum model in APSIM (see Holzworth et al., 2015; Keating et al., 2003; Hammer et al., 2010). The rabi sorghum crop type M35–1 and its validated genotypic coefficients were used as a base for the agri-system evaluation (Hammer et al., 2010; Ravi Kumar et al., 2009; Kholová et al., 2013, 2014). This base was further expanded with system-relevant combinations of management practices and rabi-sorghum relevant cultivar parameters. The spatio-temporal information on optimum Genotype x Environment and production parameters were finally used to separate the region into clusters with higher levels of similarities in these characteristics.

2.2. APSIM sorghum module

APSIM set-ups from previous work (Ravi Kumar et al., 2009; Kholová et al., 2013, 2014) were used in this study to simulate sorghum growth and development with a range of weather and soil information, management, and genetic coefficients representing the major post-rainy sorghum production regions. Altogether, we ran 4,299,264 simulations to analyze the post-rainy sorghum production system in India. A detailed description of the APSIM model is available in Holzworth et al. (2014, 2015) and Hammer et al. (2010). In short, the APSIM sorghum module algorithms process the interactions between the daily weather (rainfall, minimum and maximum temperature, solar radiation) and soil inputs considering the crop management practices and crop genetic coefficients to arbitrate the daily status of the soil-crop-atmosphere continuum and integrates this information into comprehensive outputs on crop development, growth and the production used for further analysis in this study.

2.3. Model inputs

2.3.1. Soil information

Throughout the main rabi sorghum production tract in India (Maharashtra, Karnataka, Andhra Pradesh, and Telangana), sorghum is usually grown on vertisols (International Soil Reference and Information Centre; Kumar et al., 2017). The variation in soil depth and water holding capacity significantly influences crop production. Accordingly, for each simulation unit, the Genotype × Management options were tested in the context of the 3 vertisol composites (bulk density ~1.4 g cm-3; ~0.7% organic carbon; C:N ~14.5) with varying soil depth and plant available water (PAW); i.e. shallow (70 cm depth; 94 mm PAW); medium (105 cm depth; 132 mm PAW); deep (150 cm depth; 144 mm PAW). In all soils, the soil nitrogen content was set for 50 kg ha^{-1} NO3 and 10 kg ha^{-1} NH4. The soil conditions were automatically re-initialized before each season's simulation. These soil parameters were compiled from the reports of measured soil parameters gathered by the International Soil Reference and Information Centre (ISRIC) and the National Bureau of Soil Survey and Land Use Planning (NBSS & LUP) in Bangalore.

2.3.2. Weather information

As there is a general lack of quality weather information accessible in India, we tested several commonly used synthetic weather data (daily Tmin, Tmax, rainfall) from different sources (Agricultural Modern-Era Retrospective Analysis for Research and Applications (AgMERRA; https://data.giss.nasa.gov/impacts/agmipcf/agmerra/), NASA-POWER (https://power.larc.nasa.gov) and MarkSim (MarkSim® GCM - DSSAT weather file generator (cgiar.org)). To complement these datasets, solar radiation was estimated using an algorithm based on sunshine hours and extraterrestrial radiation (Soltani and Hoogenboom, 2003; Soltani and Sinclair, 2012). The synthetic weather data was then compared with the observed weather information according to i) their agreement with observed Tmax and Tmin and sum of rainfall and ii) the agreement between the mean simulated yields using observed weather data were compared against yields using synthetic weather data from the same locations. The distribution of meteorological stations and records used for comparison with synthetic data is described in Supplementary Fig. 1.

We used standard metrics to indicate the goodness of fit; i.e. correlation coefficient (R2), root mean square error (RMSE - Eq. 1), and index of agreement (D-index). The D-index was proposed by Willmott et al.

(1985) specifically for modeling studies. D-index value range is −1–1 (Eq. 2) and, accordingly, a D-index value closer to one indicates closer agreement between the two variables compared.

Equations are listed below:

$$RMSE(root\ mean\ square\ error) = \left[\sum_{i=0}^{n} \frac{(p_i - \quad o_i)^2}{n}\right]^{0.5} \quad (1)$$

$$Index\ of\ agreement(d) = 1 - \left(\frac{\sum_{i=1}^{n}(P_i - O_i)^2}{\sum_{i=1}^{n}\left(|P_i'| + |O_i'|\right)^2}\right) \quad (2)$$

where n is the number of observations, Pi is the predicted observation, Oi is a measured observation, and P_i'= Pi - M and O_i' = Oi - M (M is the mean of the observed variable).

2.3.3. Crop specific coefficients used in the APSIM model

Crop coefficients need to be specified in APSIM-sorghum to reflect the growth and developmental characteristics of a crop cultivar. In our case, we used rabi sorghum – Maldandi, M35–1 (Kholová et al., 2013; Ravi Kumar et al., 2009; tt_endjuv_to_ini = 250; TPLA max = 2.8; VPD responsiveness = 0.95) crop coefficients to define a rabi sorghum "template". The synthetic cultivars were created by altering the M35–1 genotypic coefficients (Kholová et al., 2014) which had previously been found relevant for the development of improved rabi-sorghum cultivars, extensively studied, quantified, and tested in-vivo and in-silico using the APSIM platform (Kholová et al., 2013, 2014; Ronanki et al., 2018; Vadez et al., 2011). These include 1) "tt_endjuv_to_ini" which corresponds to a duration of end-of-juvenile to panicle-initiation developmental phase [thermal time units; TT] and specify the crop cycle duration; 2) coefficient of "TPLAmax" function which specifies the growth of total plant leaf area during the plant development, and 3) "VPD responsiveness" which defines the crop transpiration responsiveness to vapor pressure deficit (Hammer et al., 2010; https://www.apsim.info/documentation/model-documentation/crop-module-documentation/sorghum/). The parameters and their ranges used in this study are reported in Table 1.

Table 1
Overview of the variation in crop management (M) practices and in crop genetic (G) coefficients (representing a range of biologically relevant crop variants) tested by the crop growth model in the context of different soils. These resulted in 13,824 GxM combinations that were simulated within 311 grids to evaluate the optimum GxM supporting crops production/ resilience across post-rainy sorghum production systems in India. Here the tt_endjuv_to_ini corresponds to the duration of the end of juvenile to panicle initiation phase [thermal time units]; TPLA stands for total plant leaf area and corresponds to the power coefficient of the TPLAmax function, VPD stands for vapor pressure deficit of two crop types (VPD responsive and non-responsive crop types were created as detailed in Kholová et al., 2014).

M/ G/ soil	G/M/soil variation (APSIM coefficient/ module used)	Range of variation in G/M/soil (varied unit)
soil	Soil	Shallow soil (70 cm); Medium soil (105 cm); Deep soil (150 cm)
M	Sowing window	16th September – 23rd September; 23rd September – 30th September; 30th September – 7th October; 7th October - 14th October; 14th October – 21st October; 21st October – 28th October
M	Planting density	6; 8; 10; 12; 14; 16 plants m^{-2}
M	Nitrogen fertilization (Urea application schedule)	0–0; 20–20; 50–50; 100–100 kg ha^{-1}
G	Crop duration (tt_endjuv_to_ini; [TT])	Very Early (150); Early (200); Medium (250); Late (300)
G	Rate of canopy growth, vigor [power coefficient for TPLA max function in APSIM]	Low (2.4); Medium (2.6); High (2.8); Very high (3.0)
G	Transpiration responsiveness [Capacity of the canopy to limit transpiration in high VPD]	Low (0.95); High (0.80)

2.4. Simulations setup

APSIM is a process-based cropping systems simulation tool capable of reproducing the range of agronomic interventions and the base of several commercial applications; e.g. YieldProphet ® (Yield Prophet) (Hochman et al., 2009), WhopperCropper (The Regional Institute - J. Managing Climate Variability - Crops), CropARM (Decision support tools and modeling | Tasmanian Institute of Agriculture (utas.edu.au) (Richter et al., 2017). The APSIM sorghum module v. 7.6, with incorporated algorithms enabling simulations of crop transpiration responsiveness to atmospheric drought (details in Kholová et al., 2014), was set-up for each of the 311 gridded weather time-series (31 seasons; AgMERRA-NASA series), soils typically sown to rabi sorghum in the region (shallow, medium and deep vertisol; Trivedi, 2009; Jirali et al., 2010; https://www.millets.res.in/farmer/Recommended_packages_of_practices_Rabi_sorghum.pdf), 3 cultivar-specific parameters representing the biological variation in sorghum crops (G) and a range of management practices (M) relevant for the region. This resulted in 4, 299,264 simulations (Table 1). The baseline for the simulations was inspired by the recommended management practices for growing post-rainy season sorghum documented by Rooney et al. (2007), Trivedi (2009), and Olson (2012). The range of variation in the crop management practices (Ravi Kumar et al., 2009) relevant for the region was used as per the discussion with experts from the Indian Institute of Millets Research (IIMR), and International Crops Research Institute for Semi-Arid Tropics (ICRISAT) crop improvement teams and farmers (Table 1). Dimes and Revanuru (2004) previously tested the suitability of APSIM to reproduce these M interventions (Nitrogen), Turner and Rao (2013) looked at plant density and cultivar duration and Akinseye et al. (2020) sowing dates. The sowing within each of the specified sowing windows (Table 1) was triggered by a minimum of 9 mm of rainfall within 5 days. Upon meeting these requirements, APSIM initiated the sowing with the specified combination of inputs. The soil carbon and nitrogen were re-initialized before each sorghum season. The soil moisture profile at sowing was assumed to be fully saturated after the rainy season in all grids and, in addition, farmers often use irrigation after sowing to ensure germination (Trivedi, 2009). After setting up the simulation runs, all the Genotype × Management combinations were evaluated in-silico in all grids covering the rabi sorghum production regions.

2.5. Automation of APSIM runs with C#

The APSIM sorghum model was run using environmental data spanning 31 years with a total of 13,824 Genotype × Management combinations in 311 grids. In total 4,299,264 simulations were performed to generate the complete set of output files. This is a time-consuming task and a single commodity computer would take years to run this amount of simulations (Jarolímek et al., 2019). Additionally, the system requires huge storage capacities to hold the generated output files. Despite the fact that APSIM is designed to perform intended runs, the number of simulations exceeded the capacity of its in-built features. Therefore, we used supporting software tools for automation and scheduling of the designed factorial runs to tackle the large number of simulations. Simulations were generated by a custom software solution developed using the.NET framework and C# programming language. This resulted in a special software application that scheduled and generated simulation runs in batches as per the computational capacity of the available high-performance computing facility at the Czech University of Life Sciences Prague (128 GB RAM, 16 core AMD EPYC 7281 2.7 GHz CPUs). Therefore, the batches were run in parallel on 7

Table 2
The table shows the weightage of simulated production quantity (grain yield, stover yield) and production stability (biomass deviation, years with grain yield failure) indicators for the construction of the simulation weighting index. The value of each of the indicators was weighted (%) for the construction of a single simulation index (Eqs. 3–7). This was used to evaluate a particular GxM combination across 31 seasons of simulations within a particular grid and separately for both scenarios.

Scenario/ parameter weightage	Grain Yield, Eq. (3) (average of 31 simulated years)	Stover Yield, Eq. (3) (average of 31 simulated years)	Biomass deviation; Eq. (3) (across 31 simulated years)	Frequency of years with grain yield failure; Eq. (4) (across 31 simulated years)
Production	70%		30%	
	40%	30%	15%	15%
Stability	30%		70%	
	17.14%	12.86%	35%	35%

high-performance computers. Further technical details on this process are documented in Jarolímek et al. (2019).

2.6. Output analysis and visualization using interactive online tools and maps

Each of the simulation output files containing a particular Genotype × Management scenario generated for 31 seasons within each grid was evaluated for the main agronomically important parameters linked to the production quantity (mean of grain yield, stover yield, Eq. 3), and production stability (frequency of years with yield failure, Eqs. 4–6 and standard deviation of total biomass - grain and stover, Eq. 3). This was achieved by creating the index to weigh each of these outputs according to its anticipated importance of the farmers' demands on sorghum production in two scenarios: "production" and "stability" scenarios (Table 2, Eqs. 3–7; similarly in Thornton et al., 2018). The range of approaches to quantify agri-system production and stability were comprehensively reviewed in Zampieri et al. (2020) (used in e.g. Descamps et al., 2018, Thornton et al., 2018). In this work, we adapted some of these simple concepts to evaluate production and stability based on the available crop model outputs.

The "production scenario" intended to reflect the likely demand of the more economically secure sorghum farmers and increased the weightage of production indicators (mean of grain yield, stover yield, Eq. 3, Table 2). The "stability scenario" was designed to reflect the likely needs of economically vulnerable sorghum farmers and so the proportionally higher weightage was introduced to production stability indicators, minimizing the probability of grain yield failure and year-to-year total biomass fluctuations, (Eqs. 4–6, Eq. 3; Table 2). Accordingly, in the production scenario, the index considered the production factor weight of 70% (grain and stover yield; 40%; 30%) and 30% weightage of stability indicators that penalized production fluctuations and yield failure, i.e. frequency of years where crops failed to reach the grain filling stage with grain yield 0 and biomass deviation; 15%, 15% respectively (Table 2, Eqs. 3–7). In the "stability" scenario, the higher weightage was introduced to the stability indicators, i.e. weightage of production factors was only 30% (grain and stover yield; 17.14%; 12.86% respectively) with 70% weightage on stability indicators, i.e. frequency of years with yield failure and biomass deviation; 35%, 35%, respectively (Table 2, Eqs. 3–7).

The simulation index was calculated for each grid and combination of GxM and is an aggregated value of several features representing the simulation's time series. A simulation index consisted of grain yield, stover yield, and biomass deviation scores which were calculated as differences from the mean normalized by dividing by standard deviation, which is often referred to as "standard score" or "z-score" in statistics. The simulation average was calculated from the 31 years of simulation data for each particular GxM combination, while the overall average and standard deviation was calculated from all simulations within the same soil group in the given grid (Eq. 3).

Grain yield, stover, and biomass deviation scores were calculated as standard score (calculated as difference from the mean divided by standard deviation):

$$\textit{standard score of } X = \frac{\textit{average value of } X - \textit{average value of ALL}}{\textit{standard deviation of ALL}} \tag{3}$$

where X is a currently evaluated simulation and ALL are all simulations within the same soil group in that particular grid.

Production failure score was calculated to penalize simulations that contained the years with grain yield failure (Eq. 4) or simulations where the ratio of successful growth years to ALL years was below 80% (Eq. 5). The final failure score (Eq. 6) considered the higher of the two values (which increase quadratically):

$$\textit{yield failure score} = (\textit{ratio of yield failure} * 10)^2 \tag{4}$$

$$\textit{simulation failure score} = \max(((0.8 - \textit{ratio of successful simulation years}) * 10)^2, 0) \tag{5}$$

$$\textit{total failure score} = \max(\textit{yield failure score}, \textit{simulation failure score}) \tag{6}$$

The final simulation index (production and stability) was then calculated by multiplying the scores above (Eqs. 3, 6) by weights depending on the scenario (see Table 2). For example, the calculation of simulation weighting index for the production scenario was:

$$\begin{aligned}\textit{simulation index} &= 0.4 * \textit{grain score} + 0.3 * \textit{stover score} \\ &\quad - 0.15 * \textit{biomass score} - 0.15 * \textit{total failure score}\end{aligned} \tag{7}$$

For each grid, the simulation index for production and stability scenarios was generated and the 10 simulations resulting in the highest index within each scenario selected. Within each scenario, 10 simulations with the maximum simulation index score were evaluated for the occurrence of particular Genotype × Management combinations. These were then stored in a database and can be use for visualization using an interactive web application available at https://ls40.pef.czu.cz/maps/; Source code is available at https://github.com/culs-fem-dit/APSIM-maps). The frontend uses React JavaScript framework and Google Maps API and the backend Nette PHP framework. Users can choose four parameters to be shown in the map - main variable, soil type, cultivation scenario, and cultivar (optimal G combinations or M35–1 representing Maldandi crop). Results for an entire grid are visualized in the form of a discrete heatmap. Additionally, users can show a second map for the comparison of different interventions and scenarios. In this way, the user can easily visualize the maximum attainable agronomically important parameters (grain and stover yield) with optimized Genotype and Management (or optimized M for currently grown M35–1 crop type) while understanding which Genotype and Management combinations lead to this outcome within specific geographic units.

Table 3
Statistical metrics used for evaluation of agreement between the three sources of gridded meteorological characteristics (AgMERRA-NASA, NASA-POWER, MarkSIM) with observed meteorological characteristics; R2 (Pearson's correlation coefficient), RMSE (root mean squared error), D-index. The actual correlations are visualized for AgMERRA-NASA data on Fig. 1a, b and Fig. 2a, b.

Weather Source	Meteorological/ agronomic characteristic	R2	RMSE	D-index
AgMERRA-NASA	Maximum Temperature (monthly means)	0.88	1.06	0.96
	Minimum Temperature (monthly means)	0.83	2.03	0.89
	Rainfall (monthly in-season mean)	0.86	0.61	0.96
	Grain Yield (site average)	0.88	691	0.96
	Biomass (site average)	0.88	1857	0.74
NASA-POWER	Maximum Temperature (monthly means)	0.68	1.74	0.89
	Minimum Temperature (monthly means)	0.79	1.46	0.94
	Rainfall (monthly in-season mean)	0.96	0.34	0.98
	Grain Yield (site average)	0.36	1520	0.27
	Biomass (site average)	0.27	2708	0.36
MARKSIM	Maximum Temperature (monthly means)	0.48	2.36	0.81
	Minimum Temperature (monthly means)	0.71	1.85	0.90
	Rainfall (monthly in-season mean)	0.74	0.85	0.95
	Grain Yield (site average)	0.25	1988	0.19
	Biomass (site average)	0.32	3047	0.19

2.7. APSIM output file analysis

2.7.1. Identification of GxExM for optimal production and resilience scenarios

Within each grid and soil type (representing a particular E) the output files representing particular Genotype × Management combinations were evaluated using the production and stability scenario index (see Section 2.6). For each simulation grid, the obtained scenario-specific indexes were sorted and the resulting distribution evaluated using interquartile range and z-score to detect possible outliers (details in Suppl. Fig. 2). This approach revealed that the top-end of the distribution does not contain any obvious outliers and that the approach of penalizing simulations with high yield failure or simulation failure rates (Eqs. 4 and 5) was sufficient to disqualify many of the simulations that appeared on the lower tail of the distribution. Additionally, the k-means clustering method was applied to visually inspect the proportion of the data that should be utilized for further analyses (Suppl. Fig. 2). After several manual iterations, we decided that the 10 simulations attaining the highest index for each scenario would be a sufficiently large sample to provide insight on the main characteristics of the sorghum system for a particular grid (Suppl. Fig. 2; Table 4a, b); These "10 best" simulations would now include the particular combinations of Genotype × Management leading to superior agronomic performance of the system (E in the particular grid) in the "production" and "stability" scenarios. The analysis outputs within the "10 best" simulations are summarized in Suppl. Fig. 3a, b, and 4 - further grid details can be visualized and dissected using the web application.

2.8. Identification of geographically homogeneous system units

For this task, the Principal Components Analysis (PCA) was run for each combination of grid, soil, and scenario (production, stability) for the characteristics that define the 10 simulations attaining the highest scenario-related index (see above). The loadings for 3 Principal Components (explaining altogether >80% of dataset variation) of each scenario and soil have been averaged across each grid. Resulted average loadings of 311 grid items were initially separated into 3, 4 and 5 clusters (R package; https://www.r-project.org/, Table 4), visualized and the cluster-specific production and stability characteristics calculated. Considering the 3–4–5 cluster characteristics and after consultation with experts (ICRISAT and IIMR sorghum breeding teams), 4 clusters appeared the most sensible to be effectively used in crop improvement programs (discussed in 4.3). Subsequently, geographical distribution of the cluster associated with each grid item was visualized using ArcGIS software v.1.0 and the main characteristics within each cluster summarized (Fig. 3; optimal Genotype × Management and agronomic characteristics of cropping system). This approach allowed us to separate the geographies with relatively similar responses to the cultivars and management interventions (i.e. "homogeneous system units").

3. Results

3.1. Selection of an appropriate source of meteorological input

To identify the most reliable source of meteorological input, three data sources were obtained and tested (AgMERRA-NASA, NASA-POWER, MarkSIM (these are described in detail in Ruane et al., 2015; Thornton et al., 2018; Jones et al., 2002; Rienecker et al., 2011). In all three datasets, there was a good agreement with the observed monthly temperature averages (Tmin, Tmax; Table 3). The monthly in-season

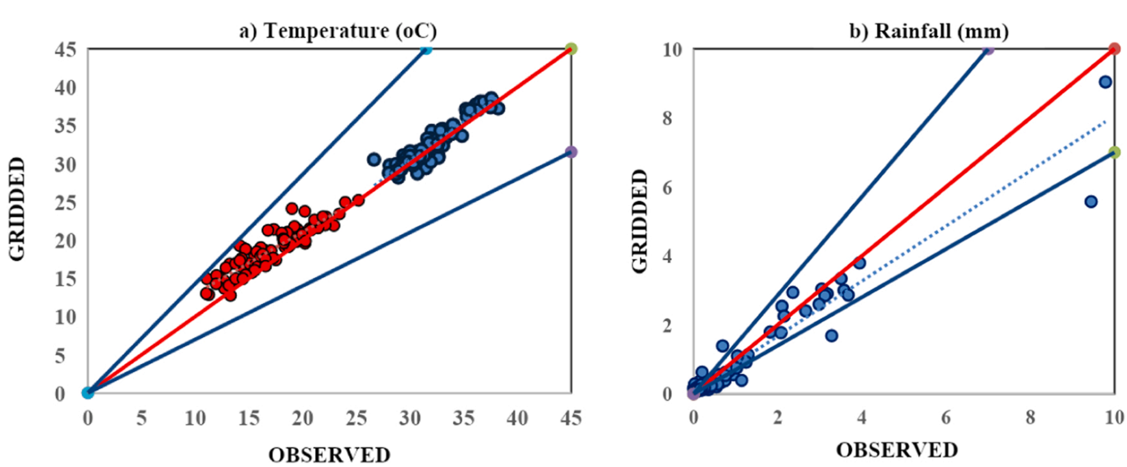

Fig. 1. a, b. Comparison of minimum (red circles) and maximum (blue circles) temperature from the gridded AgMERRA weather dataset with observed temperature (a) and comparison of in-season rainfall (October - March) from the gridded AgMERRA weather dataset with observed rainfall of the same period (b). In each graph, the middle (red) line represents 1:1 relation and the other (blue) lines represent the 30% divergence percentile of the 1:1 line. (For interpretation of the references to colour in this figure, the reader is referred to the web version of this article.)

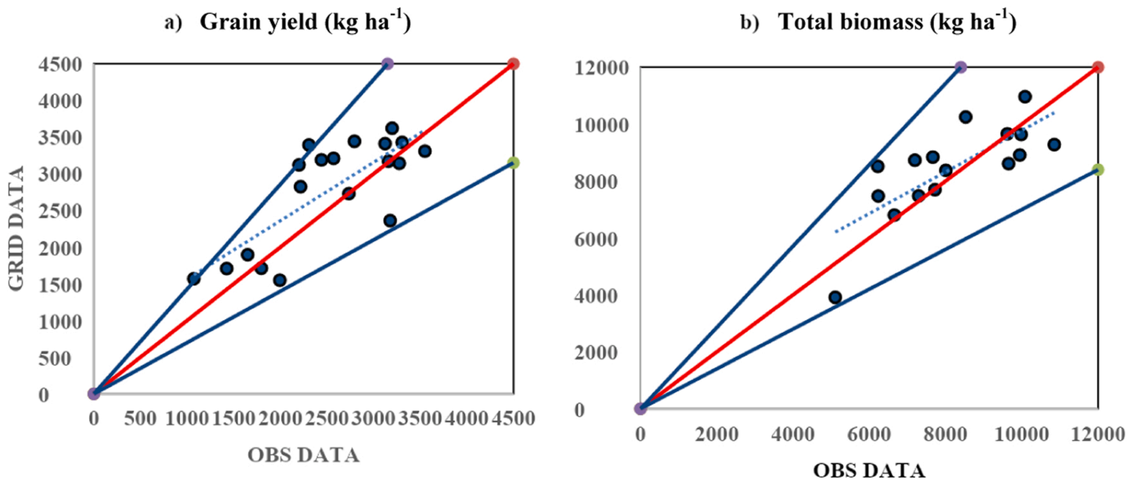

Fig. 2. a, b. Comparison of (a) grain yield and (b) total biomass simulation output from the APSIM model with observed weather data versus running APSIM with the synthetic data of AgMERRA (from Fig. 1a, b). Middle (red) lines represent a 1:1 line and the other (blue) lines denote a 30% divergence percentile of the 1:1 line. (For interpretation of the references to colour in this figure, the reader is referred to the web version of this article.)

rainfall averages were still in good agreement for AgMERRA-NASA and NASA-POWER datasets but the statistical metrics were notably lower for the MarkSIM data (Table 3). Consequently, the statistical metrics describing the relation between agronomic parameters simulated with the observed meteorological datasets and the three tested sources of gridded meteorological information were considerably better for AgMERRA-NASA (Table 3; visualized in Fig. 1a, b, Fig. 2a, b). Based on these results, gridded AgMERRA-NASA data was used to expand the spatio-temporal dimensions of simulations. Consequently, 311 gridded meteorological records (each grid size 0.5°x 0.5° encompassing 31 years of weather records (1980–2010)) were used to cover the major rabi sorghum production tract in India.

3.2. Genotype × Management runs across the grid; main characteristics of the processes, generated data and maps

The APSIM sorghum model was run across the Indian rabi sorghum production tract (311 grid items) producing a total of 4,299,264 simulations. Computation took approximately 14 days including several downtime periods and generated 14.6 TB of output data. APSIM is natively set to generate the data in raw text format. Therefore, specifically for our study, follow-up processing was necessary to extract and parse the relevant pieces of information. This text information was transformed into.csv file format to ease the calculations required for the study (a program was written in C# language to select only the relevant data. All this data is available at DOI:10.5281/zenodo.5256068 (https://zenodo.org/record/5256068#. YSYy_S0Rpf0). Statistical

Table 4a
The overview of crop production parameters (grain, stover and total biomass yield [kg ha^{-1}]) and parameters linked to crop stability (proportion of seasons with grain yield failure and standard deviation in total biomass production) averaged across “10 best” simulations attaining highest scenario-weighting index within each scenario and for each of the 3 soils. These parameters were evaluated for site-specific optimal G combinations with optimized M practices (Table 4a) and Maldandi-specific G parameters with optimized M practices (M35–1; Table 4b).

Scenario	Soil	Grain yield [kg ha^{-1}]	Stover yield [kg ha^{-1}]	Proportion of seasons with grain yield failure [%]	Deviation in total biomass production [kg ha^{-1}]
Production	All soils	2690	4911	16	986
	Shallow	2166	4185	35	1146
	Medium	2819	5279	14	1104
	Deep	3085	5270	0	707
Stability	All soils	2455	4318	0	551
	Shallow	2032	3459	1	556
	Medium	2718	4667	0	644
	Deep	2615	4827	0	452

Table 4b
The overview of crop production parameters (grain, stover and total biomass yield [kg ha^{-1}]) and parameters linked to crop stability (proportion of seasons with grain yield failure and standard deviation in total biomass production) averaged across “10 best” simulations attaining highest scenario-weighting index within each scenario and for each of the 3 soils. These parameters were evaluated for site-specific optimal G combinations with optimized M practices (Table 4a) and Maldandi-specific G parameters with optimized M practices (M35–1; Table 4b).

Scenario	Soil	Grain yield [kg ha^{-1}]	Stover yield [kg ha^{-1}]	Proportion of seasons with grain yield failure [%]	Deviation in total biomass production [kg ha^{-1}]
Production	All soils	2298	4607	22	830
	Shallow	1753	3927	50	982
	Medium	2445	4895	14	914
	Deep	2695	5000	0	595
Stability	All soils	2136	4217	5	570
	Shallow	1649	3479	12	638
	Medium	2368	4543	2	644
	Deep	2391	4630	0	429

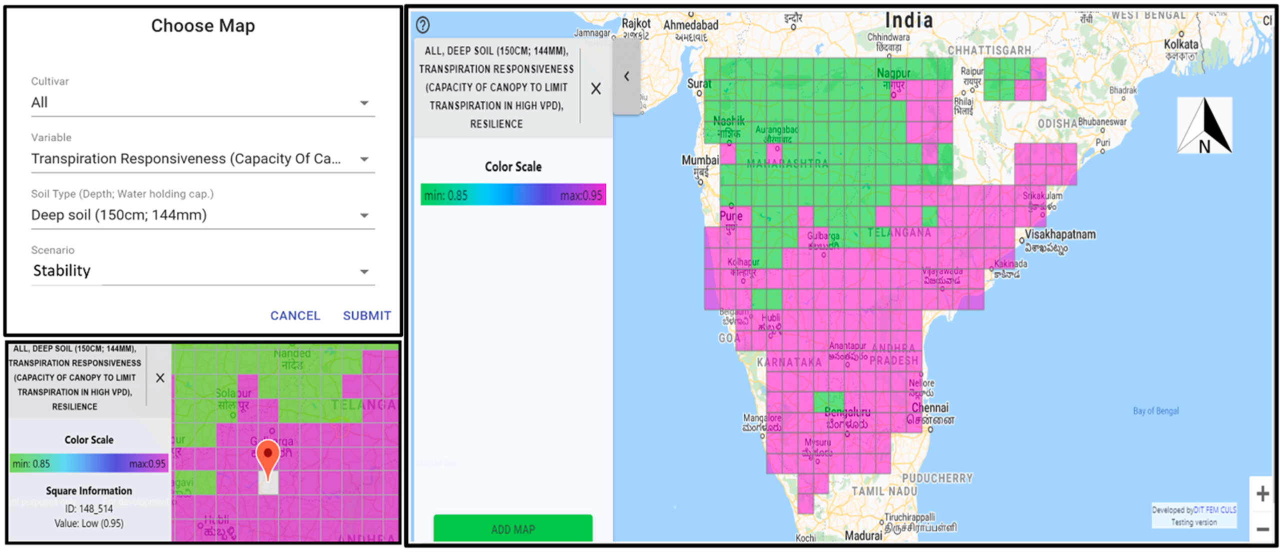

Fig. 3. The visualization of the APSIM simulations outputs via the web application (https://ls40.pef.czu.cz/maps). Each of the panels shows the target post-rainy sorghum growing region in the peninsular part of the Indian sub-continent. The coloured grids (0.5°x 0.5°) signify the geographical variation in management (M) practices and crop characters (G) expected to contribute to the improvement of post-rainy sorghum production/ stability. The user can choose to visualize the grain and stover yield (under "variable") potentially achievable for the currently grown maldandi crop type (by choosing "M 35–1") or optimized cultivar ("All") for a particular soil type ("Deep"/"Medium"/"Shallow") and scenario ("Production"/"Stability"). Furthermore, users can visualize which level of M (density, sowing window, fertilization) and G (vigour, crop duration, transpiration responsiveness) contributes towards optimal production for the particular cultivar, soil, and scenario and in which region. The actual level of the chosen variable can be visualized by clicking on the grid of interest. The tool has a feature to visualize two maps at the time (green box "Add map") for comparison. (For interpretation of the references to colour in this figure, the reader is referred to the web version of this article.)

analysis was done using the MS Excel environment in combination with basic tools provided by excel and specialized macros written in Visual Basic specifically for this task.

3.3. Production gains achievable by optimizing crop management and cultivar choice

Table 3a, b summarize the agronomic performance indicators of the top 10 best-simulated scenarios (i.e. attaining the highest index under production and stability scenario) for optimal cultivar (i.e. optimal combination of G-factors for each environment (grid Table 4a) and M35–1 (i.e. G-factors specific for the M35–1 maldandi crop type; Table 4b) across and within each of the tested soils and across all tested geographies (i.e. simulation units; environments represented by grids). As expected, crop production was predicted to be higher on deeper soils for optimal cultivars and M35–1. Furthermore, the simulations revealed that the site-specific optimization of the cultivar is expected to enhance the grain yields by around 10% (~350 kg ha^{-1}) and stover yields around 5% (~200 kg ha^{-1}). The optimal cultivar was further expected to minimize the proportion of years with grain yield failure compared to M35–1 across the tested conditions. On the other hand, site-specific cultivar optimization was predicted to cause more fluctuations in production across the years (higher biomass deviation indicating lower system stability) compared to M35–1 in the production scenario (compare Table 4a, b).

3.4. Identification of "homogeneous system units" based on the geographical distribution of optimal Genotype × management

The Genotype × Management interventions leading to maximum index of system production and stability are summarized in Table 4a (for optimal G combinations, i.e. optimal cultivars) and 4b (for the G combination specifying the maldandi M 35–1 cultivar). The spatial variability in the optimum G and M intervention can be found in Suppl. Fig. 3a (for optimal cultivars in certain production scenarios), Suppl. Fig. 3b (for the optimal cultivar in the "stability" scenario), and could be further geo-spatially explored using the web application (Fig. 2). Here it was apparent that the optimum Genotype × Management would be location and soil specific. Nevertheless, from Suppl. Fig. 3a, b, and the web application (Fig. 3), we could visually observe distinct geographical North-West to South-East patterns that changed with the soil depths and scenarios. Generally, there was a trend favoring interventions with high doses of N-fertilizer and earlier planting windows (except in the North-West regions, which were predicted to benefit from later sowing windows). In the most stringent scenarios (i.e. stability, shallow soils, North-West geographies), the optimal M tended to favor combinations with lower planting densities. Across all of the investigated geographies and scenarios, the optimal cultivars most frequently involved combinations of short-duration and high vigor characters. The North-Western regions would specifically benefit from the introduction of crops with tight canopy transpiration control (transpiration responsiveness 0.85, Suppl. Fig. 3a, b).

Similarly, we visualized the geographical distribution of optimum management practices for the maldandi crop type (M35–1; Suppl. Fig. 4, the web application). Output highlighted that there was an apparent North-West to South-East gradient in optimal M combinations. Most of the optimal M combinations generally favoured much lower planting densities and lower fertilizer inputs compared to the optimized crop type (i.e. compared to site-specific G combinations, compare optimal M from Suppl. Fig. 3a, b with Suppl. Fig. 4). Also, similarly to the site-specific optimized crop types analysis, the North-West part of the investigated region would benefit from later planting windows more than the rest of the production region (compare Suppl. Fig. 3a, b with Suppl. Fig. 4).

Using PCA, the outputs from each of the simulation units ("grids") were clustered into the four geographical units based on their similarities in production/ stability system characteristics as well as optimized combinations of G and M parameters. Such analysis, in principle, seperated the grids into the geographical regions with the similarities in system response to G and M interventions (Fig. 4). When visualized, these four geographical units formed the pattern of concentric layers around the "core" of the North-Western part of the sorghum production area (HSU_1; light blue, Fig. 3). The main system characteristics along

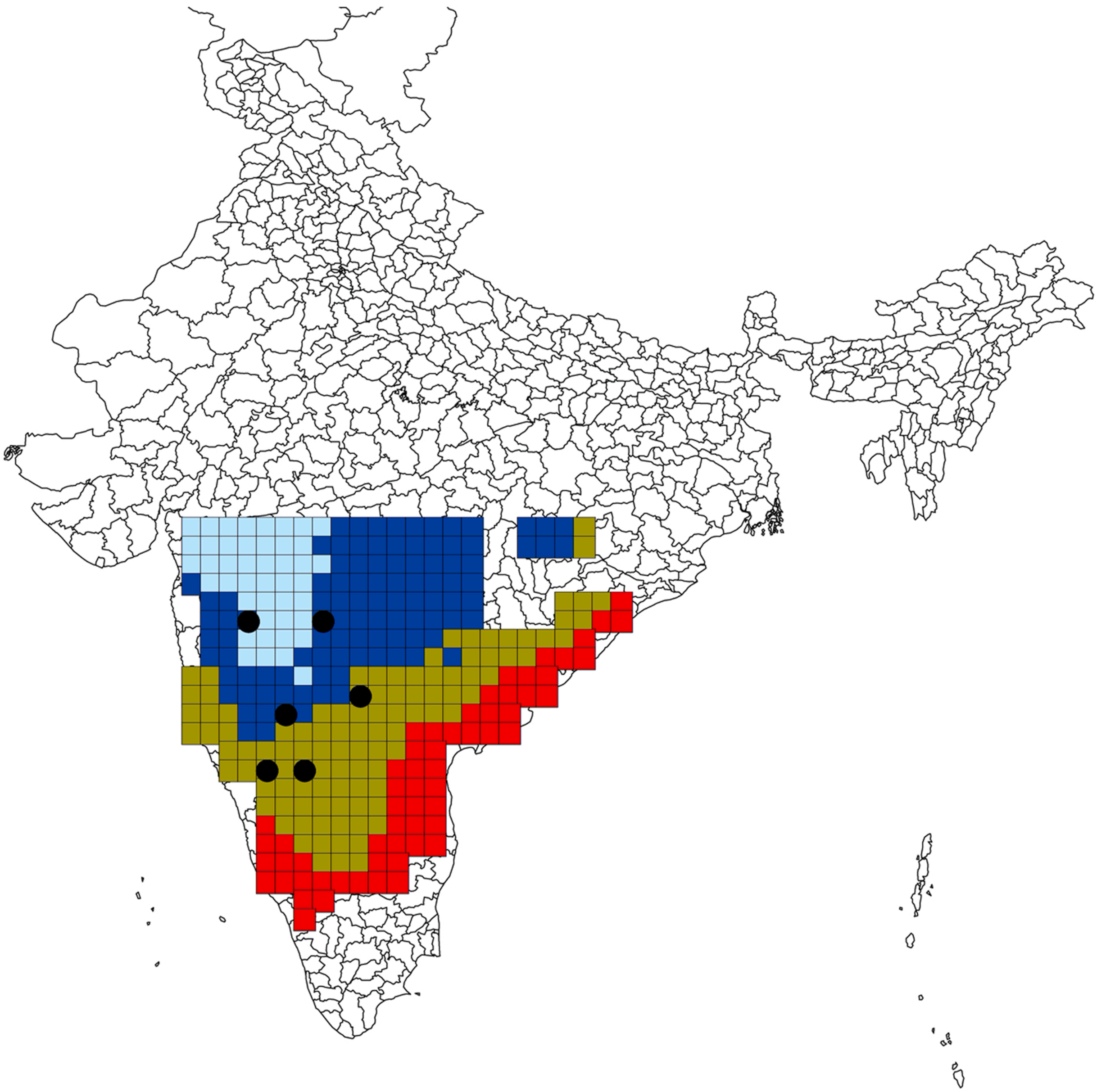

Fig. 4. The map of India over-layed with the four identified "homogeneous system units" (HSU, highlighted in different colors; summary system characteristics of the HSU clusters are in Table 5). Each of the grids (0.5°x 0.5°) is expected to respond more homogeneously to particular GxM interventions than the remaining grids within one HSU compared to the grids from a different HSU. The discrete black circles on the map highlight the current post-rainy sorghum testing sites within an All India Coordinated Research Project (AICRP; http://www.millets.res.in/aicrp.php).

with the optimal Genotype × Management combinations within these geographical units were summarized in Table 5. Generally, Table 5 illustrates that with the increasing distance from this core (HSU_1) the production potential and system stability would increase in the direction towards HSU_4 (stover and grain yield production as well as system stability indicators). This was also well-reflected in optimmalized G and M parameters within each HSU; i.e. planting density, crop duration, and plant vigor. A parameter indicating plant responsiveness to VPD also increased with increasing distance from the production core (from HSU_1 towards HSU_4). Across all HSUs, most of the optimal Genotype × Management combinations leaned towards the early sowing windows and stable fertilizer doses ~ 70–70 kg ha^{-1} (basal dose - top-dressing). Table 5.

4. Discussion

4.1. Deployment of high-performance computations for effective APSIM-runs

The technical details and challenges involved in this computational exercise were described in Jarolímek et al. (2019). In principle, we used a cluster of seven high-performance computers and tested several options to distribute the computations across the cluster effectively. This involved separation of the simulation set-ups into batches, which were consequently scheduled and run manually. This exercise enabled us to design the structure of the software tools for the computational facility used. This allows for further process automation should similar exercises be required with the cluster in the future. Alternative software resources available in the public domain, specialized for computational distribution such as HTCondor (https://research.cs.wisc.edu/htcondor/) might be also considered. Another option to be tested would include running

Table 5
The summary statistics of the system production and stability indicators resulting from optimized management (M) and genetic crop characters (G) within the identified Homogeneous System Unit (HSU) clusters as spatially defined in Fig. 4.

Cluster number (color in Figure 4)	HSU_1 (light blue)	HSU_2 (dark blue)	HSU_3 (green)	HSU_4 (red)
(M) Sowing window	16 - 23sep	23 - 30sep	16 - 23sep	16 - 23sep
(M) Plant density (plant m^{-2})	12.54	12.58	13.18	13.66
(M) Nitrogen fertilization (kg ha^{-1} Urea)	74-74	72.8-72.8	71-71	71.3-71.3
(G) Crop duration (tt_endjuv_to_ini)	181	178	197	220
(G) Rate of canopy growth, vigor (power coefficient for TPLA max)	2.88	2.882	2.926	2.96
(G) Transpiration responsiveness (Capacity of canopy to limit transpiration in high VPD)	0.9065	0.893	0.941	0.9485
Grain yield [kg·ha^{-1}]	2231	2182	2675	3308
Stover yield [kg·ha^{-1}]	4026	3982	4692	5986
Proportion of seasons with grain yield failure [%]	24	12	2	1
Deviation in total biomass production [kg ha^{-1}]	767	747	740	854

the simulations using commercial cloud services, such as Azure or Amazon Web Services.

It is important to note that the new version of APSIM is being developed (NextGen APSIM; https://apsimnextgeneration.netlify.app). NextGen APSIM modules are already capable of running multiple simulations much more efficiently compared to classic APSIM software (https://www.apsim.info). However, transiting this massive work to the NextGen APSIM-based framework would require rigorous cross-validation of the model functions that are key for the presented study as well as the model set-up. Nonetheless, the necessity of transitioning to the NextGen system must be seriously considered, especially in the context of rising demand for a similar types of analysis (e.g. in the context of CGIAR crop improvement program modernizations; https://excellenceinbreeding.org/; https://bigdata.cgiar.org/event/webinar-target-population-of-environments-tpe-beyond-helping-make-better-crop-improvement-practice/).

4.2. Cultivar x Management effects and their optimal combinations for higher and stable sorghum production

APSIM has been used to model management and genetic interventions for various cropping systems (e.g. wheat, chickpea, maize, potato; Chenu et al., 2011, 2013; Chapman et al., 2000a, 2000b, 2000c; Lobell et al., 2015, Chauhan et al., 2008; Chauhan et al., 2013; Beah et al., 2021; De Silva et al., 2021; Ojeda et al., 2020, 2021). Currently, APSIM is a base for several commercial applications used by different stakeholders such as farmers or breeders (YieldProphet ® (Yield Prophet), WhopperCropper (The Regional Institute - J.Managing Climate Variability - Crops), CropARM (Decision support tools and modeling | Tasmanian Institute of Agriculture (utas.edu.au)). In the case of sorghum, the APSIM model has been used to evaluate the production regions assessed in this study and the effect of management and crop genetic interventions (Ravi Kumar et al., 2009 - maldandi sorghum parameterization; Kholová et al., 2013 - rabi sorghum systems characterization; Kholová et al., 2014 - evaluation of genetic interventions; Dimes and Revanuru, 2004 - nitrogen application; Turner and Rao, 2013 - effect of planting density & duration of cultivars; Akinseye et al., 2020 - effect of sowing dates).

In the context of this study, where accessing observed meteorological information is problematic, we transited the entire modeling framework into the AgMERRA-NASA-based gridded framework (Ruane et al., 2015) to allow for more geographically precise and spatially balanced analysis. Out of the tested options (NASA-POWER, MarkSIM, AgMERRA-NASA), the AgMERRA-NASA-based set-up (31 years of daily weather records) was found sufficient to represent the historical weather variability across the rabi-sorghum production region. The sorghum simulation outputs were, additionally, cross-compared with the ranges of agronomic parameters reported from multi-year, multi-location agronomic trials conducted in post-rainy seasons within the Indian national sorghum evaluation network (AICRIP project; http://www.millets.res.in/aicrp.php). Information from these trials could not be closely compared with our simulations, as, for example, the tested genotypes are usually Maldandi types but not exactly M35–1 and the management practices in these trials usually involve "life-saving irrigation'' or other agronomic practices that are rarely sufficiently documented to allow strict comparison. However, the crop production ranges and responses to the management practices (fertilization, planting density) reported from AICRIP sorghum testing trials (http://www.millets.res.in/aicsip13.php) were in reasonable agreement with the model outputs. Therefore, the AICRIP trials would be an interesting data source to further emulate the presented modeling framework.

The modelling approach is one of the very few options that allows us to disentangle, quantify and optimize the effects of cultivar and crop management interventions (Jeuffroy et al., 2014; Lecomte et al., 2010; Chapman et al., 2002). This information is critical to guide any efforts for agri-system improvement and breeding. The vast amount of information generated in this simulation exercise (14.6 TB) allowed us just this - i.e. to separate and quantify the effects of particular cultivar X management interventions on the important characteristics of the post-rainy sorghum cropping system. We found that the magnitude of any GxM intervention effect depended on the soil properties. Our findings further indicated that, in general, site-specific fine-tuning the crop and crop agronomic practices within the breeder- and farmer- relevant ranges would have an important effect on crop production/ stability in the rain-fed rabi sorghum belt of India. Despite the fact that we did not investigate the specific factors leading to production constraints, as in Kholová et al. (2013), our study emphasizes that crop products (accompanied by well-designed agronomic practices) for rain-fed systems have to be carefully tailored to the variability of production environments. We showed that the commonly occurring genetic variability in sorghum species is sufficient to enhance rabi-systems (i.e. duration, vigor and canopy conductance; Kholová et al., 2014; Bodner et al., 2015). Significant improvements can be achieved by fitting existing cultivars with suitable management to the particular context of the rabi-production system or by generating new crop products using the resources generated in this study as a guideline.

For this purpose, we provide the data generated along with the tool designed for further exploration by any stakeholders for free (e.g. post-rainy sorghum breeding programs, policy-makers or on-ground farmer advisory services). This tool is now ready to explore the spatial distributions of optimal management practices for the Maldandi crop type (M35–1) as well as the production potentially achievable with other crop types (details in Fig. 2). As the map shows, these are dependent on the location, soil type, and scenarios considered. The estimated site-specific genetic enhancement of Maldandi has the potential to increase ~10% (± 7%) grain and 5% (± 2%) stover production across all locations. The same intervention would stabilize grain production across seasons. We envision that a similar type of IT tool, could complement the recommendation packages (e.g. https://www.millets.res.in/farmer/Recommended_packages_of_practices_Rabi_sorghum.pdf; periodically released by the Indian government) in order to enable the site-specific recommendations through on-ground agencies who can operate the simple interactive map. This framework should also serve as a base upon which further enhancements required by the different users can be built, such as crop improvement teams, policy makers, and farmer advisory services, which would enable its broader deployment and impact. Similar principles have been reflected in CropArm (http://www.armonline.com.au/#/wc), an APSIM-based tool developed to support decisions in Australian farming systems. To our knowledge, the presented online tool is the first of this kind that is able to support effective system design for climate risk-prone agri-systems in developing countries. The tool's development demonstrates a diligent approach on how to condense the vast amount of data typically produced by crop modelers into digestible information for non-experts. This approach will be further expanded for other regions and crops, evolved and sensitized to the indigenous agri-system requirements and contexts.

4.3. *Identification of homogeneous system units (HSUs)*

The primary beneficiaries in mind while developing the study were crop breeders. Breeders typically require crop modelers to identify geographies for which a particular crop product and agronomic management can be developed and where it is best tested (e.g. BPAT review; https://plantbreedingassessment.org/bpat-project/bpatmission/) (Kholová et al., 2021). This kind of analysis required the further stratification of the information held by the generated dataset. Firstly, we reduced the data dimensionality using principal component analysis (PCA) and consequently deployed the clustering approach to form geo-spatially distinct classes - the "homogeneous system units" (HSUs). This allowed us to separate the tested geographies (grids) into four HSUs based on similarities in optimum production characteristics and system responses to GxM interventions. Such novel assessments considerably extended the previously used approaches (environmental characterization/ target population of environments) e.g. in Kholová et al. (2013), Hajjarpoor et al., (2018, 2021), Chauhan et al. (2013), Chenu et al. (2011), and Chapman et al. (2000a, 2000b). The "HSU" analysis allowed us to differentiate the geographies with maximum similarities within a geographic group and dissimilarities between the groups not only based on the modeled interactions of crop, environment, and management but also on system responsiveness to the Genotype × Management interventions. We suggest that such geospatial classification enable breeding programs to, for instance, optimize the distribution of the multi-location testing sites, improve the statistical treatment of the data generated in different geographies and precisely design and target crop product development efforts. For instance, in typical crop improvement programs, the crop is tested with very limited management options or the management is adapted only "*post-mortem*" when the genotype is already fixed. These circumstances inevitably stagnate the crop production improvement in these complex systems. To overcome this gap, we provided a unique tool that allows for the simultaneous prediction of optimal crop management along with the suitable crop cultivar, which is otherwise impossible. We conclude that the presented APSIM-powered framework enables the improvement of breeding targets, empowering breeding programs to design region-specific Genotype × Management options *ex-ante* that could significantly accelerate efforts to improve productivity/ resilience of dry-season sorghum cultivation.

4.4. *Possible limitations of the study and continuous improvement of the framework*

Models are reflections of our imperfect knowledge, which is why it's important to acknowledge the assumptions and other possible limitations of the acquired modeling outputs. In our case, we need to mention the use of a gridded data source (AgMERRA-NASA) instead of the actual meteorological observations that may have been preferable. Although meteorological information is becoming more available as standard across the globe, many countries, like India, are still not well covered with accessible, high-quality, and up-to-date information. Since our study required homogeneous coverage of key regions, we chose to use NASA-generated information that has supported modeling of agri-systems similar to ours (Table 3, Fig. 1a, b, 2 a, b). In the ideal case, we would have had detailed agronomic evaluations of sorghum production across locations to cross-validate the simulation set-up responsiveness to major system limitations (e.g. agronomic practices). As mentioned above, these datasets are very rare in the local context and their generation is cost- and time-intensive. While we work on such dataset generation, we do have numerous studies and even commercial products based on the APSIM sorghum module responsiveness to a range of M and G contexts (e.g. Akinseye et al., 2020; Dimes and Revanuru, 2004; Turner and Rao, 2013).

In future, we plan to use this sorghum modeling framework to support broader socio-economic modeling studies. Here we presented our attempt to demonstrate the generic approach, i.e. the scenario weighting index which is based on an educated guess founded on literature surveys (e.g. Blümmel and Rao, 2006, Tesfaye, 1998, Ravi et al., 2003, Reddy et al., 2005, Rao et al., 2017, Blümmel et al., 2015) and discussions with experts. Such estimates are to be improved as we progress in understanding and interlinking this work with socio-economic studies of the target population of stakeholders in particular regions. The understanding of community demands or particular user cases should, in principle, guide further simulation exercises and, among others, the resolution of simulations, the GxM scenarios tested, further assumptions made, tool co-creation, and design.

5. Conclusion

The presented work aims to translate current advances in crop modeling science into a quantitative understanding of crop production systems for the key pool of beneficiaries (e.g. breeding programs, farmer advisories, decision-makers, etc.) via a simple visualization tool. The presented framework simplifies the complex modeling data (~0.5 million simulations, ~14 TB) and utilizes them to understand the context-dependencies of post-rainy sorghum agricultural systems even by a community of non-experts. Although numerous on-line tools have been developed, primarily to provide advice to large-scale agricultural producers in developed countries, these might not fit the requirements of stakeholders in countries like India. We argue that to enable an effective understanding of the diversity of small-scale agriculture systems, the tool has to be tailored: (i) to be easily accessible (possibly free of cost) (ii) simple enough and sufficiently interactive, and (iii) encompass a valid range of the farming scenarios. We have developed draft tool and analytics with the example of post-rainy sorghum production systems in India and will continue the customization and evolution of the presented tool to serve the particular needs of various end-users. A similar approach can be now adapted to other agricultural production systems, especially those that are small-scale and low input.

CRediT authorship contribution statement

Swarna Ronanki: Conceptualization, Formal analysis, Investigation, Writing – original draft, Writing – review & editing. **Jan Pavlik:** Data curation, Investigation, Resources, Software, Writing – original draft, Writing – review & editing. **Jan Masner:** Resources, Software, Writing – original draft, Writing – review & editing. **Jan Jarolimek:** Funding acquisition, Project administration, Resources, Supervision. **Michal Stoces:** Software. **Degala Subhash:** Data curation, Visualization. **Harvinder S. Talwar:** Resources. **Vilas A. Tonapi:** Funding acquisition, Resources. **Mallayee Srikanth:** Visualization. **Jana Kholová:** Conceptualization, Data curation, Funding acquisition, Methodology, Project administration, Supervision, Writing – original draft, Writing – review & editing. **Rekha Badam:** Writing – original draft.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The results and knowledge included in this article have been obtained withsupport from the following grants; Internal grant agency of the Faculty of Economics and Management, Czech University of Life Sciences Prague, grant no. 2019B0009 – Life Sciences 4.0, the CGIAR Research Program on Grain Legumes and Dryland Cereals (GLDC) and a mini-grant from the CGIAR Community of Practice on Modelling (https://bigdata.cgiar.org/communities-of-practice/crop-modeling/), and the core funding of ICAR- Indian Institute of Millets Research. A Global Challenges Research Fund project - Transforming India's Green Revolution by Research and Empowerment for Sustainable Food Supplies (TIGR2ESS, BB/P02797/01). Authors are grateful to Dr. Amir Hajjarpoor (UMR DIADE, Université de Montpellier, Institut de Recherche pour le Développement) for compilation and analysis of meteorological information from several sources. Authors further acknowledge the contribution of Dr. Sunita Choudhary (International Crops Research Institute for the Semi-Arid Tropics, Patancheru, Hyderabad) for resource mobilization as well as dissemination and promotion of scientific findings reported in this work to the key stakeholders.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.fcr.2021.108422.

References

Akinseye, F.M., Ajeigbe, H.A., Traore, P.C., Agele, S.O., Zemadim, B., Whitbread, A., 2020. Improving sorghum productivity under changing climatic conditions: a modelling approach. Field Crops Res. 246, 107685.

Ambadi, A., Krishnamurty, D., Rao, S., Desai, B.K., Ravi, M.V., Shubha, S., 2018. Yield potential and economics of rabi sorghum (Sorghum bicolor L.) as influenced by different crop residues and green biomass composts. J. Appl. Nat. Sci. 10 (1), 128–132.

Beah, A., Kamara, A.Y., Jibrin, J.M., Akinseye, F.M., Tofa, A.I., Ademulegun, T.D., 2021. Simulation of the optimum planting windows for early and intermediate-maturing maize varieties in the Nigerian savannas Using the APSIM model. Front. Sustain. Food Syst. 5, 624886 https://doi.org/10.3389/fsufs.

Blümmel, M. , Rao, P.P. 2006. Economic value of sorghum stover traded as fodder for urban and peri-urban dairy production in Hyderabad, India. International Sorghum and Millets Newsletter (47):97–100.

Blümmel, M., Deshpande, S., Kholova, J., Vadez, V., 2015. Introgression of staygreen QLT's for concomitant improvement of food and fodder traits in Sorghum bicolor. Field Crops Research. Elsevier BV, pp. 228–237. https://doi.org/10.1016/j.fcr.2015.06.005.

Bodner, G., Alireza, N., Hans-Peter, Kaul, 2015. Management of crop water under drought: a review. Agron. Sustain. Dev. 35 (2), 401–442.

Chapman, S.C., Cooper, M., Butler, D.G., Henzell, R.G., 2000a. Genotype by environment interactions affecting grain sorghum. I. Characteristics that confound interpretation of hybrid yield. Aust. J. Agric. Res. 51, 197–207.

Chapman, S.C., Cooper, M., Hammer, G.L., Butler, D.G., 2000b. Genotype by environment interactions affecting grain sorghum. II. Frequencies of different seasonal patterns of drought stress are related to location effects on hybrid yields. Aust. J. Agric. Res. 51 (2), 209–221.

Chapman, S.C., Cooper, M., Butler, D.G., Hammer, G.L., 2000c. Genotype by environment interactions affecting grain sorghum. III. Temporal sequences and spatial patterns in the target population of environments. Aust. J. Agric. Res. 51, 223–233.

Chapman, S.C., Cooper, M., Hammer, G.L., 2002. Using crop simulation to generate genotype by environment interaction effects for sorghum in water-limited environments. Aust. J. Agric. Res. 53, 379–389. https://doi.org/10.1071/AR01070.

Chauhan, Y., Wright, G., Rachaputi, N., McCosker, K., 2008. Identifying chickpea homoclimes using the APSIM chickpea model. Aust. J. Agric. Res. 59 (3), 260–269.

Chauhan, Y.S., Solomon, K.F., Rodriguez, D., 2013. Characterization of north-eastern Australian environments using APSIM for increasing rainfed maize production. Field Crops Res. 144, 245–255.

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 (6), 1743–1755.

Chenu, K., Deihimfard, R., Chapman, S.C., 2013. Largescale characterization of drought pattern: a continent-wide modelling approach applied to the Australian wheatbelt - spatial and temporal trends. New Phytol. 198, 801–820.

Dayakar Rao B., Shashidhar Reddy Ch, Nirmal Reddy K., Ratnavathi CV, Shyamprasad G. Seetharama N. , 2009. Package of Practices for Improved Rabi Sorghum Cultivation (English), National Agricultural Innovation Project (NAIP), ITC, Secunderabad, 500 003 and NRCS, Rajendranagar, 500 030, Andhra Pradesh, India. Technical Bulletin number NAIP/ITC/NRCSTECH/4/2009, 14pp.

Descamps, C., Quinet, M., Baijot, A., Jacquemart, A.L., 2018. Temperature and water stress affect plant–pollinator interactions in Borago officinalis (Boraginaceae). Ecol. Evol. 8 (6), 3443–3456.

De Silva, S.H.N.P., Takahashi, T., Okada, K., 2021. Evaluation of APSIM-wheat to simulate the response of yield and grain protein content to nitrogen application on an Andosol in Japan. Plant Prod. Sci. 1–12.

Dimes, J.P., Revanuru, S., 2004. Evaluation of APSIM to simulate plant growth response to applications of organic and inorganic N and P on an Alfisol and Vertisol in India. In *Aciar Proceedings* (pp. 118–125). ACIAR; 1998.

Dingkuhn, M., Soulié, J.C., Lafarge, T., 2011. Samara V2: A cereal crop model to study G x E x M interaction and phenotypic plasticity, and explore ideotypes. In: AgMIP Rice International Workshop, 28–30 August, Beijing, China.

Habyarimana, E., Piccard, I., Catellani, M., De Franceschi, P., Dall'Agata, M., 2019. Towards predictive modeling of sorghum biomass yields using fraction of absorbed photosynthetically active radiation derived from sentinel-2 satellite imagery and supervised machine learning techniques. Agronomy 9 (4), 203.

Hajjarpoor, A., Vadez, V., Soltani, A., Gaur, P., Whitbread, A., Babu, D.S., Gumma, M.K., Diancoumba, M., Kholová, J., 2018. Characterization of the main chickpea cropping systems in India using a yield gap analysis approach. Field Crops Res. 223, 93–104.

Hammer, G.L., van Oosterom, E., McLean, G., Chapman, S.C., Broad, I., Harland, P., Muchow, R.C., 2010. Adapting APSIM to model the physiology and genetics of complex adaptive traits in field crops. J. Exp. Bot. 61 (8), 2185–2202.

Hochman, Z., Van Rees, H., Carberry, P.S., Hunt, J.R., McCown, R.L., Gartmann, A., Holzworth, D., Van Rees, S., Dalgliesh, N.P., Long, W., Peake, A.S., 2009. Re-inventing model-based decision support with Australian dryland farmers. 4. Yield Prophet® helps farmers monitor and manage crops in a variable climate. Crop Pasture Sci. 60 (11), 1057–1070.

Holzworth, D.P., Snow, V., Janssen, S., Athanasiadis, I.N., Donatelli, M., Hoogenboom, G., White, J.W., Thorburn, P., 2015. Agricultural production systems modelling and software: current status and future prospects. Environ. Model. Softw. 72, 276–286.

Holzworth, D.P., Huth, N.I., deVoil, P.G., Zurcher, E.J., Herrmann, N.I., McLean, G., Chenu, K., van Oosterom, E.J., Snow, V., Murphy, C., Moore, A.D., Brown, H., Whish, J.P.M., Verrall, S., Fainges, J., Bell, L.W., Peake, A.S., Poulton, P.L., Hochman, Z., Thorburn, P.J., Gaydon, D.S., Dalgliesh, N.P., Rodriguez, D., Cox, H., Chapman, S., Doherty, A., Teixeira, E., Sharp, J., Cichota, R., Vogeler, I., Li, F.Y., Wang, E., Hammer, G.L., Robertson, M.J., Dimes, J.P., Whitbread, A.M., Hunt, J., van Rees, H., McClelland, T., Carberry, P.S., Hargreaves, J.N.G., MacLeod, N., McDonald, C., Harsdorf, J., Wedgwood, S., Keating, B.A., 2014. APSIM - evolution

towards a new generation of agricultural systems simulation. Environ. Model. Softw. 62, 327–350. https://doi.org/10.1016/j.envsoft.2014.07.009.
Jarolímek, J., Pavlík, J., Kholová, J., Ronanki, S., 2019. Data pre-processing for agricultural simulations. AGRIS Line Pap. Econ. Inform. 11 (1), 49–53.
Jeuffroy, M.H., Casadebaig, P., Debaeke, P., Loyce, C., Meynard, J.M., 2014. Agronomic model uses to predict cultivar performance in various environments and cropping systems. A review. Agron. Sustain. Dev. 34 (1), 121–137.
Jirali, D.I., Biradar, B.D., Rao, S.S., 2010. Performance of Rabi sorghum genotypes under receding soil moisture conditions in different soil types. Karnataka J. Agric. Sci. 20 (3).
Jones, J.W., Antle, J.M., Basso, B., Boote, K.J., Conant, R.T., Foster, I., Godfray, H.C.J., Herrero, M., Howitt, R.E., Janssen, S., Keating, B.A., 2017. Toward a new generation of agricultural system data, models, and knowledge products: state of agricultural systems science. Agric. Syst. 155, 269–288.
Jones, P.G., Thornton, P.K., Díaz, W., Wilkens, P.W., Jones, A.L., 2002. MarkSim: A computer tool that generates simulated weather data for crop modeling and risk Assessment: version 1 [CD-ROM].
Jones, J.W., Hoogenboom, G., Porter, C.H., Boote, K.J., Batchelor, W.D., Hunt, L.A., Wilkens, P.W., Singh, U., Gijsman, A.J., Ritchie, J.T., 2003. The DSSAT cropping system model. Eur. J. Agron. 18 (3–4), 235–265.
Keating, B.A., Carberry, P.S., Hammer, G.L., Probert, M.E., Robertson, M.J., Holzworth, D., Huth, N.I., Hargreaves, J.N.G., Meinke, H., Hochman, Z., McLean, G., Verburg, K., Snow, V., Dimes, J.P., Silburn, M., Wang, E., Brown, S., Bristow, K.L., Asseng, S., Chapman, S., McCown, R.L., Freebairn, D.M., Smith, C.J., 2003. An overview of APSIM, a model designed for farming systems simulation. Eur. J. Agron. 18, 267–288.
Kholová, J., McLean, G., Vadez, V., Craufurd, P., Hammer, G.L., 2013. Drought stresscharacterization of post-rainy season (*rabi*) sorghum in India. Field Crops Res. 141, 38–46.
Kholová, J., Tharanya, M., Sivasakthi, K., Srikanth, M., Rekha, B., Hammer, G.L., McLean, G., Deshpande, S., Hash, C.T., Craufurd, P., Vadez, V., 2014. Modelling the effect of plantwater use traits on yield and stay-green expression in sorghum. Funct. Plant. 41, 1019. https://doi.org/10.1071/FP13355.
Kholová, J., Urban, M.O., Cock, J., Arcos, J., Arnaud, E., Aytekin, D., Azevedo, V., Barnes, A.P., Ceccarelli, S., Chavarriaga, P., Cobb, J.N., Connor, D., Cooper, M., Craufurd, P., Debouck, D., Fungo, R., Grando, S., Hammer, G.L., Jara, C.E., Xu, Y., 2021. In pursuit of a better world: crop improvement and the CGIAR. J. Exp. Botany 72 (14), 5158–5179. https://doi.org/10.1093/jxb/erab226.
Kumar, S.R., Bhat, P., Rajappa, P.V., 2017. Management strategy to improve input use efficiency and enhance sorghum productivity per stored rain drop in vertisols during rabi season. Curr. Sci. 30, 304–307.
Lecomte, C., Prost, L., Cerf, M., Meynard, J.M., 2010. Basis for designing a tool to evaluate new cultivars. Agron. Sustain Dev. 30, 667–677. https://doi.org/10.1051/agro/2009042.
Lobell, D.B., Hammer, G.L., Chenu, K., Zheng, B., McLean, G., Zheng, B., Chapman, S.C., 2015. The shifting influence of drought and heat stress for crops in Northeast Australia. Glob. Change Biol. 21, 4115–4127.
Olson, S. , 2012. Designing an Ideal Energy Crop: The Case for Sorghum bicolor (Doctoral dissertation).
Ojeda, J.J., Rezaei, E.E., Remenyi, T.A., Webb, M.A., Webber, H.A., Kamali, B., Harris, R. M., Brown, J.N., Kidd, D.B., Mohammed, C.L., Siebert, S., 2020. Effects of soil-and climate data aggregation on simulated potato yield and irrigation water requirement. Sci. Total Environ. 710, 135589.
Ojeda, J.J., Rezaei, E.E., Remenyi, T.A., Webber, H.A., Siebert, S., Meinke, H., Webb, M. A., Kamali, B., Harris, R.M., Kidd, D.B., Mohammed, C.L., 2021. Implications of data aggregation method on crop model outputs–the case of irrigated potato systems in Tasmania, Australia. Eur. J. Agron. 126, 126276.
Pradhan, P., Fischer, G., van Velthuizen, H., Reusser, D.E., Kropp, J.P., 2015. Closing yield gaps: how sustainable can we be? PLoS One 10, e0129487.
Rao, Benhur ,Mukherjee, Deep Narayan, Devi, Y., Tonapi, Vilas, 2017. An Economic Analysis of Improved Rabi Sorghum Cultivars in Rainfed Situation of Maharashtra, India. 4. 7–15.
Ravi, D., Vishala, A.D., Nayaker, N.Y., Seetharama, N. and Blümmel, M. 2003. Grain yield and stover fodder value relations in rabi sorghum. International Sorghum and Millets Newsletter. 44: 28–31.
Ravi Kumar, S., Graeme, L., Hammer, Ian Broad, Peter, Harland, Greg, Mc.Lean, 2009. Modelling environmental effects on phenology and canopy development of diverse sorghum genotypes. Field Crops Res. 111, 157–165.
Reddy, K.G. and Michael, B. and Rao, P.P. and Reddy, B.V. S. and Ramesh, S. and Reddy, K.M. V.P. (2005) Evaluation of farmer-grown improved sorghum cultivars for stover quality traits. International Sorghum and Millets Newsletter, 46. pp. 86–89.
Richter, M.E., Phelan, D., Harrison, M., Dean, G., Pengilley, G., Hinton, S., Mohammed, C., 2017. CropARM: An agronomic support tool assisting Tasmanian farmers for rainfed and irrigated wheat production. In *" Doing More with Less", Proceedings of the 18th Australian Agronomy Conference 2017, Ballarat, Victoria, Australia, 24–28 September 2017* (pp. 1–4). Australian Society of Agronomy Inc.
Rienecker, M.M., Suarez, M.J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M.G., Schubert, S.D., Takacs, L., Kim, G.K., Bloom, S., 2011. MERRA: NASA's modern-era retrospective analysis for research and applications. J. Clim. 24 (14), 3624–3648.
Ronanki, S., Kholová, J., Talwar, H.S. , 2018. Simulation of post rainy sorghum yield response N fertilization in India. ⟨https://21centurysorghum.com/wp-content/uploads/2018/06/09h40-Ronanki-Hall-B-Thurs1.pdf⟩.
Rooney, W.L., Blumenthal, J., Bean, B., Mullet, J.E., 2007. Designing sorghum as a dedicated bioenergy feedstock. Biofuels Bioprod. Bioref. 1 (2), 147–157.
Ruane, A.C., Winter, J.M., McDermid, S.P., Hudson, N.I., 2015. AgMIP climate datasets and scenarios for integrated assessment. In: Rosenzweig, C., Hillel, D. (Eds.), Handbook of Climate Change and Agroecosystems: The Agricultural Model Intercomparison and Improvement Project (AgMIP) Integrated Crop and Economic Assessments. ICP Series on Climate Change Impacts, Adaptation, and Mitigation, Vol. 3. Imperial College Press, pp. 45–78. https://doi.org/10.1142/9781783265640_0003. Part 1.
Soltani, A., Hoogenboom, G., 2003. Minimum data requirements for parameter estimation of stochastic weather generators. Clim. Res. 25 (2), 109–119.
Soltani, A., Sinclair, T.R. 2012. Modelling physiology of crop development, growth and yield. CABI: 322.
Soltani, A., Hajjarpour, A., Vadez, V., 2016. Analysis of chickpea yield gap and water-limited potential yield in Iran. Field Crops Res. 185, 21–30.
Tesfaye A. , 1998. Economics of milk production in and around Hyderabad of Andhra Pradesh. M.Sc. thesis, Acharya NG Ranga Agricultural University, Hyderabad 500 030.
Trivedi, T.P., 2009. Handbook of Agriculture. Directorate of Information and Publications of Agriculture, Indian Council of Agricultural Research, New Delhi, India.
Thornton, P.K., Whitbread, A., Baedeker, T., Cairns, J., Claessens, L., Baethgen, W., Bunn, C., Friedmann, M., Giller, K.E., Herrero, M., Howden, M., Kilcline, K., Nangia, V., Ramirez-Villegas, J., Kumar, S., West, P.C., Keating, B., 2018. A framework for priority-setting in climate smart agriculture research. Agricultural Systems 167, 161–175. https://doi.org/10.1016/j.agsy.2018.09.009.
Turner, N.C., Rao, K.P.C., 2013. Simulation analysis of factors affecting sorghum yield at selected sites in eastern and southern Africa, with emphasis on increasing temperatures. Agric. Syst. 121, 53–62.
Vadez, V., Krishnamurthy, L., Hash, C.T., Upadhyaya, H.D., Borrell, A.K., 2011. Yield, transpiration efficiency, and water-use variations and their interrelationships in the sorghum reference collection. Crop Pasture Sci. 62 (8), 645–655.
Willmott, C.J., Ackleson, S.G., Davis, R.E., Feddema, J.J., Klink, K.M., Legates, D.R., O'Donnell, J., Rowe, C.M., 1985. Statistics for the evaluation of model performance. J. Geophys. Res. 90 (C5), 8995–9005.
Zampieri, M., Weissteiner, C.J., Grizzetti, B., Toreti, A., van den Berg, M., Dentener, F., 2020. Estimating resilience of crop production systems: from theory to practice. Sci. Total Environ. 735, 139378 https://doi.org/10.1016/j.scitotenv.2020.139378.
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. For. Meteorol. 194, 207–217. https://doi.org/10.1016/j.agrformet.2014.04.007.