Quantifying the effects of varietal types × management on the spatial variability of sorghum biomass across US environments

Jonathan J. Ojeda¹ | Graeme Hammer¹ | Kai-Wei Yang² | Mitchell R. Tuinstra² | Peter deVoil³ | Greg McLean¹ | Isaiah Huber⁴ | Jeffrey J. Voleneç² | Sylvie M. Brouder² | Sotirios Archontoulis⁴ | Scott C. Chapman⁵

¹Queensland Alliance for Agriculture and Food Innovation, The University of Queensland, Brisbane, Queensland, Australia
²Department of Agronomy, Purdue University, West Lafayette, Indiana, USA
³Queensland Alliance for Agriculture and Food Innovation, The University of Queensland, Gatton, Queensland, Australia
⁴Department of Agronomy, Iowa State University, Ames, Iowa, USA
⁵School of Agriculture and Food Sciences, The University of Queensland, Gatton, Queensland, Australia

Correspondence
Scott C. Chapman, The University of Queensland, School of Agriculture and Food Sciences, Gatton, Qld, Australia. Email: scott.chapman@uq.edu.au

Present address
Jonathan J. Ojeda, Regrow Ag, Brisbane, Queensland, Australia

Funding information
Agriculture and Food Research Initiative Competitive, National Institute of Food and Agriculture, Grant/Award Number: 2011-68005-30411; Advanced Research Projects Agency - Energy, Grant/Award Number: DE-AR0001135

Abstract
Regional-scale estimations of sorghum biomass production allow identification of optimum genotypexenvironmentxmanagement (GxExM) combinations for bioenergy generation. The objective of this study was to determine the degree of contributions of G, E, and M toward variability in sorghum biomass in the United States. Using the Agricultural Production Systems sIMulator in a grid computing platform, biomass was simulated for irrigated and rainfed conditions for 30 years across the United States for four sorghum varietal types (grain—GS, sudangrass—SS, photosensitive—PS, and photo-insensitive—PI). Simulated biomass was assessed by environments clustered using the sum of intercepted solar radiation (ir), mean of temperature stress factor (tp) and water stress factor (sw). Simulated biomass ranged from 5.8 t ha⁻¹ (GS-rainfed) to 27.5 t ha⁻¹ (PI-irrigated). Under high-temperature environments (mean annual temperature = 25°C), rainfed biomass between 40 and 80 days after planting (DAP) was strongly correlated with sw (r = 0.64–0.86) and irrigated biomass with ir (r = 0.68–0.81). Under low-temperature environments (mean annual temperature = 18°C) after 40 DAP, tp and ir had greater effects than sw (r = 0.55–0.82). Biomass variance was mainly explained by varietal type (50%–76%) in all environmentsxirrigation combinations, except in the high- and mid-temperature environments under rainfed conditions where rainfall had the major effect (25%–45%). However, when mean temperature during the growing season decreased from 25°C (high environments) to 18°C (low environments), the contribution of mean temperature to biomass variance increased from 7% to 34% (rainfed) and from 4% to 36% (irrigated). Varietal type had the larger interactions with other factors independently of the environment and irrigation. We demonstrated a need to quantify (i) the main GxExM drivers of biomass variability based on environmental stress factors and (ii) the variance contribution of these drivers on sorghum biomass. Our regional-scale estimations...
are key inputs for future robust biomass projections of energy sorghum genotypes integrating G×E×M under climate change scenarios.

**KEYWORDS**
APSIM, bioenergy, feedstock, genotype, irrigation, model upscaling, pSIMS, regional modeling, spatial variability, variance partitioning

1 | INTRODUCTION

Sorghum [*Sorghum bicolor* (L.) Moench] is the fifth most important cereal worldwide, with grain production of 59.3 million tons in 2018 (FAO, 2020). In the United States, the sorghum grain demand for ethanol production systems have considerably increased during the last years (Duff et al., 2019). Sorghum has cultivated varieties distributed among five races and over 25 species of wild relatives, offering enormous genetic diversity for crop improvement (Bhattacharya et al., 2011). Forage sorghum varieties maximize biomass at nitrogen (N) fertilizer rates less than those needed to maximize biomass for corn (*Zea mays* L.; Thivierge et al., 2015). Forage sorghum varieties have been recognized to produce more biomass per unit of irrigation (Rooney et al., 2007). Several sorghum varieties also have photoperiod sensitivity that prevents the crop flowering under long photoperiods of summer in higher latitudes, which confers advantage for biomass production as the crop continues as vegetative until killed by frost (Olson et al., 2012). Sorghum is also known for its drought tolerance, allowing it to produce high biomass in water-limited environments in part due to its lower transpiration compared to corn, which allows the crop to conserve soil moisture and maintain growth for a longer period (Borrell et al., 2000; Carcedo & Gambin, 2019; Chapman et al., 2000; Mutava et al., 2011).

Crop biomass is influenced by genetics (G), environment (E), crop management (M), and their interactions (Hammer et al., 2020; Rotili et al., 2020; Seyoum et al., 2018). Crop modeling has been extensively used to quantify the effect of G × E × M on sorghum grain yield using field-scale simulations (i.e., from a climate station, single soil profile data from sampling, and crop management based on field measurements; Hammer et al., 2010; Liang et al., 2020; MacCarthy et al., 2009; Raymundo et al., 2021). However, to better explore the main drivers of sorghum biomass for different genotypes, regional modeling approaches that consider the spatial heterogeneity of E and M are helpful (Antle et al., 2017). Such information will assist the agricultural industry in assessing crop management options, harvest methods, and future infrastructure for sorghum energy biomass (Daly et al., 2018). Equally important is the need for this spatial information to identify the best G×E combinations that increase sorghum biomass and support the potential expansion of bioenergy production to 2040 in US regions where sorghum is not currently grown (Langholtz et al., 2016).

Several studies reported observed sorghum biomass for a variety of genotypes and regions from field experiments (Buxton et al., 1999; Gill et al., 2014; Hoffmann & Rooney, 2014) and simulated biomass from crop model exercises at field scale (Kent et al., 2020; Lopez et al., 2017; Yang et al., 2021). There were also efforts to develop maps of potential sorghum biomass under rainfed (Daly et al., 2018; Lee et al., 2018) and irrigated conditions (Huntington et al., 2020) based on statistical regression models and machine learning across the United States. However, these models have two limitations: (i) they do not consider G (e.g., crop phenology, morphophysiological features) and M (e.g., irrigation strategy, planting date) as drivers of the spatial variability of sorghum biomass and (ii) they had only limited validation (e.g., seven biomass observations in Daly et al., 2018).

Here we applied a biophysical crop model (the Agricultural Production Systems IMulator [APSIM]) to generate G×E×M combinations and assessed the effects of the G×E×M interactions on rainfed and irrigated sorghum biomass and its variability at regional scale in the United States. The quantification of this variability and its partitioning (hereafter referred to as variance partitioning) is key to distinguishing the main drivers of spatial biomass variability (Ramirez-Villegas & Challinor, 2012; Teixeira et al., 2017). The identification of the relative importance of those drivers on sorghum biomass helps inform how we might use insights from regional model applications.

The objective of this study was to distinguish which of G, E, and M are dominant factors generating variability in sorghum biomass in the United States. We specifically investigated the relative variance contribution on simulated biomass from four sorghum varietal types, grain (GS) sudangrass (SS), photosensitive (PS), and photo-insensitive (PI), under two irrigation scenarios (rainfed and irrigated) across the potential production areas for energy sorghum in the United States. We choose varietal type and irrigation strategy as factors of analysis in this study due to their potential contribution to increase sorghum biomass for bioenergy. Although irrigation is not a current common practice in sorghum production systems of the United States, this
practice is expected to increase in order to reduce the gap between actual and potential biomass production to the extent that this is economically viable (Ciampitti et al., 2020; Huntington et al., 2020). First, we tested APSIM using aggregated field data from a diverse set of previous biomass experiments. Second, we set up a national-scale modeling experiment to estimate the variability in biomass across the sorghum-growing regions and its spatial drivers.

2 | MATERIALS AND METHODS

2.1 | Study region

This study was undertaken across the extent of the potential production areas for energy sorghum in the United States at a 30 arc-minute resolution (~35 km grid cell resolution in the North to ~50 km grid cell resolution in the South). This included 1860 grid cells that were selected based on the potential economic availability of biomass resources from agricultural lands described at the farmgate in the Department of Energy 2016 Billion-Ton Report (Langholtz et al., 2016). The geographical area has considerable heterogeneity in topography, resulting in several environments with a variety of climate regimes (Figure S1). The potential areas for biomass sorghum described by Langholtz et al. (2016) cover 100% of the state areas without discrimination between croplands. In order to better represent potential agricultural areas for biomass sorghum, we masked our data analysis to cropland areas (1325 grid cells at 30 arc-minute resolution) reported in the parallel System for Integrating Impact Models and Sectors (pSIMS; Elliott et al., 2014).

2.2 | Study workflow

Figure 1 shows the data sources, the source of biomass variance, the models and spatiotemporal scales, and the methods of analysis and tools used in this study. Initially, we conducted a field-scale model validation to test the ability of a crop model to predict in-season and final sorghum biomass (Sections 3.1 and 4.1) for locations where field data were available. Then, using climate and soil gridded data at 30 arc-minute resolution, the sorghum biomass was simulated across the potential production areas for energy sorghum in the United States (Sections 3.2 and 4.2).

2.3 | Field-scale modeling (APSIM validation)

2.3.1 | APSIM-Sorghum description

The APSIM-Sorghum model (hereafter referred to as APSIM or the model; Hammer et al., 2010, 2019) within...
APSIM Classic 7.9 r4132 (Holzworth et al., 2014) was initially validated and then used to simulate sorghum biomass across the potential production areas for energy sorghum in the United States (Section 2.4). The sorghum model in APSIM has been previously described (https://www.apsim.info/documentation/model-documentation/crop-module-documentation/sorghum) and tested across a large range of environments (e.g. Akinseye et al., 2020; Carcedo & Gambin, 2019; Clarke et al., 2019; Singh et al., 2017). A complete description of the model structure and parameters was provided by Hammer et al. (2010 and 2019). The model contains algorithms that simulate crop phenology, growth, and soil-plant N dynamics. Crop phenology is simulated using thermal time thresholds for each phenological stage. When water and N conditions are optimum, crop growth rate on a daily step is only limited by solar radiation and is calculated as the product of intercepted solar radiation and radiation use efficiency (RUE). When the crop is under water stress, biomass accumulation is calculated as the product of potential crop transpiration (limited by soil moisture, root extent, and water uptake capacity) and transpiration efficiency, adjusted for atmospheric vapor pressure deficit (i.e., relative humidity; Hammer et al., 2010).

2.3.2 | Observed field data

Observed data from 439 site\times year\times management/genotype combinations reported in published (Buxton et al., 1999; Gill et al., 2014; Heitman et al., 2017; Maw et al., 2016; Roby et al., 2017; Wight et al., 2012) and other datasets (J.J. Volenc and S.M. Brouder, unpublished datasets) were assembled to test the APSIM performance to predict aboveground biomass (hereafter referred to as biomass; Table 1). These experiments were conducted from 1988 to 2015 and comprise an unbalanced set of 21 sorghum genotypes across eight US states (17 locations; Figures S2 and S3). The experimental data were generated by three research projects. (1) The Regional Feedstock Partnership between the US Department of Energy and the Sun Grant Initiative (Gill et al., 2014; Lee et al., 2018) which aims to assess the yield potential and stability of sorghums grown across diverse production environments in the United States. This dataset included a combination of 6 sorghum genotypes (one cultivar and five hybrids)\times 7 locations (six states)\times 5 years (2008–2012). (2) The second dataset was generated by Purdue University under the CenUSA Bioenergy Project (https://nifa.usda.gov/censusa-bioenergy) which included 3 genotypes\times 3 Indiana locations\times 4 years. That project compared sorghums to perennial grasses for producing biofuels and bioproducts on marginal lands. A previous detailed APSIM validation (third dataset; Yang et al., 2021) was used to compare model biomass estimations for final (i.e., at the end of the season) versus in-season harvests. Yang et al. (2021) calibrated APSIM for in-season ([from 31 to 94 days after planting [DAP]) biomass estimations under the TERRA Project (https://www.purdue.edu/terra) for 18 genotypes at West Lafayette (IN) in 2 years (2015 and 2017). A description of the TERRA Project dataset was not included in this paper, for details see Yang et al. (2021). All experiments have been carried out to study biomass response to various treatments, including N fertilizer rates, time of planting, population, and the adaptability of genotypes across locations and years (Table 1). The crop data collected in these experiments vary with the environment and treatment applied; however, crop biomass at harvest was the only common variable between them. Therefore, we only used this crop response variable to validate APSIM estimations.

2.3.3 | Model inputs for field simulations

Historical daily climate data (daily rainfall, maximum and minimum temperature, and global solar radiation) from the AgMERRA climate forcing dataset (Ruane et al., 2015) were used as inputs to the model. This dataset extends from 1980 to 2010. For years after 2010, climate data were retrieved using the bestiapop Python package (https://github.com/JJguri/bestiapop) (Ojeda et al., 2020; Ojeda et al., 2021a) from the NASA Prediction of Worldwide Energy Resource (POWER)—Climatology Resource for Agroclimatology database (https://power.larc.nasa.gov). These climate datasets have been widely tested with weather station data across several environmental regions (Bender & Sentelhas, 2018; Merlos et al., 2015; Ojeda et al., 2017).

We used the FAO Harmonized World Soil Database (Nachtergaele et al., 2010) to setup the soil profile properties for each location. The soil data included bulk density, organic carbon, pH, air dry (corresponding to the moisture limit for dry evaporation of the soil), drained lower limit (LL15), drained upper limit (DUL), saturated (SAT) volumetric water contents, KL (potential rate of water extraction), and XF (root exploration factor). Maximum plant available water capacity (PAWC) is calculated as the difference between DUL and LL15. A complete description of soil parameters used for the validation of APSIM is provided in Table S1 for each environment.

Initial soil moisture data were not available for all field experiments. Therefore, initial moisture in the entire soil profile was assumed to be full (100% of PAWC) at planting in all field experiments except for the TX locations. This is typical for most corn/soybean [Glycine max (L.) Merr.]-growing areas in the Midwest and Eastern United States
**TABLE 1**  Description of final harvest datasets from 439 site×year×management/genotype combinations used for the APSIM model testing across the potential areas for energy sorghum in the United States

| State     | Location  | Latitude | Longitude | Year/s | M/GC² | Planting date | Fert rate (kg N ha⁻¹) | Harvest date or month/s | Genotype | VT | Population (pl m⁻²) | Reference                   |
|-----------|-----------|----------|-----------|--------|-------|---------------|-----------------------|-------------------------|----------|----|---------------------|----------------------------|
| Indiana (IN) | Butlerville (Bu) | 39.03    | −85.52    | 2011   | 60    | June 7        | 0, 50, 100, 150, 200 | January                 | Dual Purpose           | PS | 10.9               | JJ Volencic and SM Brouder (unpublished) |
| 2012      | May 11    | 0, 50, 100, 150, 200 | October 29 | Dual Purpose | PS | 12.6               |                       |                         | Dale                  | PI | 11.1               |                              |
| 2013      | June 4    | 0, 50, 100, 150, 200 | November 5 | Dual Purpose | PS | 12.6               |                       |                         | Photosensitive         | PS | 11.1               |                              |
| 2014      | May 28    | 0, 50, 100, 150, 200 | November 7 | Dual Purpose | PS | 10.9               |                       |                         | Dale                  | PI | 11.1               |                              |
| Columbia City (CC) | 41.11 | −85.40  | 2011 | 59 | June 14 | 0, 50, 100, 150, 200 | November 4 | Dual Purpose | PS | 10.3               |                              |
| 2012      | April 18  | 0, 50, 100, 150, 200 | November 5 | Dual Purpose | PS | 10.5               |                       |                         | Photosensitive         | PS | 10.5               |                              |
| 2013      | May 29    | 0, 50, 100, 150, 200 | November 20 | Dual Purpose | PS | 10.0              |                       |                         | Dale                  | PI | 10.0               |                              |
| 2014      | May 31    | 0, 50, 100, 150, 200 | December 2 | Dual Purpose | PS | 10.8              |                       |                         | Photosensitive         | PS | 10.8               |                              |
| Lafayette (La) | 40.29 | −86.90  | 2011 | 60 | June 17 | 0, 50, 100, 150, 200 | November 2 | Dual Purpose | PS | 10.8               |                              |
| 2012      | May 30    | 0, 50, 100, 150, 200 | October 16 | Dual Purpose | PS | 10.8               |                       |                         | Photosensitive         | PS | 10.8               |                              |
| 2013      | June 7    | 0, 50, 100, 150, 200 | October 28 | Dual Purpose | PS | 10.8              |                       |                         | Dale                  | PI | 10.8               |                              |
| 2014      | June 7    | 0, 50, 100, 150, 200 | December 1 | Dual Purpose | PS | 10.8              |                       |                         | Photosensitive         | PS | 10.8               |                              |

(Continues)
| State       | Location | Latitude | Longitude | Year/s               | M/GC<sup>a</sup> | Planting date | Fert rate (kg N ha<sup>−1</sup>) | Harvest date or month/s | Genotype | VT | Population (pl m<sup>−2</sup>) | Reference                  |
|-------------|----------|----------|-----------|----------------------|-------------------|----------------|-----------------------|--------------------------|-----------|----|--------------------------------|----------------------------|
| Iowa (IA)   | Ames (Am)| 42.06    | −93.68    | 2009, 2010, 2011, 2012 | 24               | May 21         | 200<sup>b</sup>         | September                | Photosensitive          | PS       | 12.5                        | Gill et al. (2014)          |
|            |          |          |           |                      |                   |                | 22053BMR               |                          | Dale           | PI      | 12.5                        |                            |
|            |          |          |           |                      |                   |                | GrazeAll               |                          |                           | SS       | 20.0                        |                            |
|            |          |          |           |                      |                   |                | GrazeNBale            |                          |                           | SS       | 20.0                        |                            |
|            |          |          |           |                      |                   |                | ES5200                 |                          |                           | PS       | 16                         |                            |
| Boone (Bo) | 42.06    | −93.68   |           | 2014, 2015           | 2                 | May 21         | 200                  | September                | Garst753BMR            | PS       | 16                         | Roby et al. (2017)         |
| Chariton (Ch)| 40.97   | −93.42   |           | 1988, 1989, 1990, 1991, 1992 | 10               | May 21         | 200                  | September                | M81E               | PI      | 16                         | Buxton et al. (1999)       |
| Kansas (KS) | Manhattan (Ma)| 39.19 | −96.58 | 2008, 2009, 2010, 2011 | 22               | June 1          | 200<sup>b</sup>         | October                  | SugarT              | PI      | 12.5                        | Gill et al. (2014)          |
|            |          |          |           |                      |                   |                | M81E                  |                          |                           | SS       | 15.0                        |                            |
|            |          |          |           |                      |                   |                | 22053BMR               |                          |                           | SS       | 20.0                        |                            |
|            |          |          |           |                      |                   |                | GrazeAll               |                          |                           | SS       | 20.0                        |                            |
|            |          |          |           |                      |                   |                | GrazeNBale            |                          |                           | SS       | 20.0                        |                            |
|            |          |          |           |                      |                   |                | ES5200                 |                          |                           | PS       | 16                         |                            |
| Missouri (MO) | Columbus (Co) | 38.88 | −92.2 | 2010, 2011            | 10               | June 9          | 0, 56, 112, 168, 224 | October                  | ES5200            | PS      | 12.5                        | Maw et al. (2016)          |
| Kentucky (KY) | Lexington (Le) | 38.08 | −84.55 | 2010, 2011, 2012      | 8                | June 5          | 200<sup>b</sup>         | September–October    | SugarT              | PI      | 12.5                        | Gill et al. (2014)          |
|            |          |          |           |                      |                   |                | M81E                  |                          |                           | SS       | 15.0                        |                            |
|            |          |          |           |                      |                   |                | 22053BMR               |                          |                           | SS       | 20.0                        |                            |
|            |          |          |           |                      |                   |                | GrazeAll               |                          |                           | SS       | 20.0                        |                            |
|            |          |          |           |                      |                   |                | GrazeNBale            |                          |                           | SS       | 20.0                        |                            |
|            |          |          |           |                      |                   |                | ES5200                 |                          |                           | PS       | 16                         |                            |
| North Carolina (NC) | Roper (Ro) | 35.87 | −76.66 | 2009, 2010, 2011, 2012 | 24               | June 9          | 200<sup>b</sup>         | October                  | SugarT              | PI      | 12.5                        | Gill et al. (2014)          |
|            |          |          |           |                      |                   |                | M81E                  |                          |                           | SS       | 15.0                        |                            |
|            |          |          |           |                      |                   |                | 22053BMR               |                          |                           | SS       | 20.0                        |                            |
|            |          |          |           |                      |                   |                | GrazeAll               |                          |                           | SS       | 20.0                        |                            |
|            |          |          |           |                      |                   |                | GrazeNBale            |                          |                           | SS       | 20.0                        |                            |
|            |          |          |           |                      |                   |                | ES5200                 |                          |                           | PS       | 16                         |                            |
| Wallace (Wa) | 34.76    | −78.1    |           | 2012, 2013, 2014, 2015 | 25               | June 9          | 0, 67, 134, 201, 268  | October                  | SugarT              | PI      | 12.5                        | Heitman et al. (2017)      |
|            |          |          |           |                      |                   |                | M81E                  |                          |                           | SS       | 15.0                        |                            |
|            |          |          |           |                      |                   |                | 22053BMR               |                          |                           | SS       | 20.0                        |                            |
|            |          |          |           |                      |                   |                | GrazeAll               |                          |                           | SS       | 20.0                        |                            |
|            |          |          |           |                      |                   |                | GrazeNBale            |                          |                           | SS       | 20.0                        |                            |
|            |          |          |           |                      |                   |                | ES5200                 |                          |                           | PS       | 16                         |                            |
| State        | Location      | Latitude | Longitude | Year/s                        | M/GC<sup>a</sup> | Planting date | Fert rate (kg N ha<sup>−1</sup>) | Harvest date or month/s | Genotype | VT | Population (plants m<sup>−2</sup>) | Reference               |
|--------------|---------------|----------|-----------|-------------------------------|-------------------|---------------|------------------------|-------------------------|-----------|-----|----------------------------------|-------------------------|
| Mississippi  | Raymond (Ra)  | 32.21    | −90.51    | 2008, 2009, 2010, 2011, 2012 | 27                | May 2         | 200<sup>b</sup>          | October                 | SugarT   | PI  | 12.5                             | Gill et al. (2014)        |
|              |               |          |           |                               |                   |               |                        |                         | M81E      | PI  | 12.5                             |                         |
|              |               |          |           |                               |                   |               |                        |                         | 22053BMR | SS  | 15.0                             |                         |
|              |               |          |           |                               |                   |               |                        |                         | GrazeAll | SS  | 20.0                             |                         |
|              |               |          |           |                               |                   |               |                        |                         | GrazeNBale | SS  | 20.0                             |                         |
|              |               |          |           |                               |                   |               |                        |                         | ES5200   | PS  | 16                               |                         |
| Texas (TX)   | Bushland (Bu)<sup>c</sup>  | 35.19    | −102.07   | 2000, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2014, 2016, 2017 | 59                | May 19 to June 13 | 200          | September–October       | SPSS405 | PI  | 20                               | Yang et al. (2021)        |
|              |               |          |           |                               |                   |               |                        |                         | SPHIKANEII | PI  |                                 |                         |
|              |               |          |           |                               |                   |               |                        |                         | SPNK300   | PI  |                                 |                         |
|              |               |          |           |                               |                   |               |                        |                         | SPSH      | PS  |                                 |                         |
|              |               |          |           |                               |                   |               |                        |                         | SPTH      | PS  |                                 |                         |
|              |               |          |           |                               |                   |               |                        |                         | SPSordan79 | PI  |                                 |                         |
|              |               |          |           |                               |                   |               |                        |                         | SPTRudan8 | PI  |                                 |                         |
|              |               |          |           |                               |                   |               |                        |                         | PH877F    | PI  |                                 |                         |
|              |               |          |           |                               |                   |               |                        |                         | PH849F    | PI  |                                 |                         |
| College Station (CS) | 30.62 | −96.36 |          | 2008, 2009, 2010, 2011, 2012 | 29                | April 27      | 200<sup>b</sup>          | September               | SugarT   | PI  | 12.5                             | Gill et al. (2014)        |
|              |               |          |           |                               |                   |               |                        |                         | M81E      | PI  | 12.5                             |                         |
|              |               |          |           |                               |                   |               |                        |                         | 22053BMR | SS  | 15.0                             |                         |
|              |               |          |           |                               |                   |               |                        |                         | GrazeAll | SS  | 20.0                             |                         |
|              |               |          |           |                               |                   |               |                        |                         | GrazeNBale | SS  | 20.0                             |                         |
|              |               |          |           |                               |                   |               |                        |                         | ES5200   | PS  | 16                               |                         |
| Corpus Christi (CC) | 27.78 | −97.56 |          | 2008, 2010, 2011, 2012 | 18                | April 27       | 200<sup>b</sup>          | July                    | SugarT   | PI  | 12.5                             | Gill et al. (2014)        |
|              |               |          |           |                               |                   |               |                        |                         | M81E      | PI  | 12.5                             |                         |
|              |               |          |           |                               |                   |               |                        |                         | 22053BMR | SS  | 15.0                             |                         |
|              |               |          |           |                               |                   |               |                        |                         | GrazeAll | SS  | 20.0                             |                         |
|              |               |          |           |                               |                   |               |                        |                         | GrazeNBale | SS  | 20.0                             |                         |
|              |               |          |           |                               |                   |               |                        |                         | ES5200   | PS  | 16                               |                         |
| Weslaco (We) | 26.22 | −98.23 |          | 2008 | 2 | April 23 | 0–336 | August | 4-EverGreen | PS  | 16.0 | Wight et al. (2012) |             |

Abbreviations: PI, photo-insensitive; PS, photosensitive; SS, sudangrass; VT, varietal type.

<sup>a</sup>M/GC indicates the number of management/genotype combinations per location.

<sup>b</sup>An optimum regional N fertilizer application of 200 kg N ha<sup>−1</sup> was considered in the experiments reported by Gill et al. (2014) based on the lack of data reported by the authors.

<sup>c</sup>All locations were conducted under rainfed conditions, except Bushland (TX).
during the planting window (i.e., late April to early June; Ojeda et al., 2017). Under the TX locations (Bushland, College Station, Corpus Christi, and Weslaco), the entire soil profile was assumed to be 50% of PAWC at planting. Crop management was configured to reproduce those applied in the experimental plots based on recorded information (Table 1).

Four varietal types (GS, SS, PS, and PI) were created based on observed data from 18 hybrids in west-central IN, United States (Yang et al., 2021). The genotypes used by Yang et al. (2021) were clustered in four groups based on thermal time between the end of juvenile stage (100°Cd after emergence) and floral initiation stage (which affects maturity, final leaf number, and potential canopy area), extinction coefficient, RUE, and response to photoperiod (which also affects final leaf number) (Table S2). We compared the varietal characteristics of the 21 genotypes from the validation dataset with these four groups and we assigned to them one varietal type to be used in the model simulations. All varietal types were used in the model validation except for grain sorghum (GS; Table 1).

2.3.4 Evaluation of model performance

The model fitness was evaluated by assessing the concordance correlation coefficient (CCC) (Lawrence & Lin, 1989) and the root mean squared error (RMSE) between observed and simulated (Pifeiro et al., 2008) values for biomass. The CCC combines precision through Pearson’s correlation coefficient, which represents the proportion of the total variance in the observed data that can be explained by the model, and accuracy by bias which indicates how far the regression line deviates from the concordance line (y = x). This coefficient has been previously used in several crop modeling studies to evaluate model performance (Ojeda et al., 2017; 2018; Pembleton et al., 2013; Tedeschi, 2006).

2.4 Regional-scale modeling

2.4.1 pAPSIM simulations

The pSIMS (Elliott et al., 2014) synthesizes crop management, weather, and soil data per grid cell and allows for parallel computing crop model simulations. pSIMS is a Python-based framework designed to support integration and high-resolution application of any site-based climate impact model that can be compiled in a Unix environment (Elliott et al., 2014). This framework has been previously used to assess climate change impacts on corn grain yield using the Decision Support System for Agrotechnology Transfer (DSSAT; Glotter et al., 2016) and APSIM (Baum et al., 2020). We used a parallelized version of APSIM (pAPSIM) implemented in pSIMS to simulate sorghum biomass at 30 arc-minute resolution over the potential production areas for energy sorghum in the United States. pAPSIM allows for regional-scale simulations by using the Swift parallel scripting language (Wilde et al., 2011) to process APSIM concurrently on several clusters and computer cores. Inputs included historical daily weather data (1980–2009) from the AgMERRA climate forcing dataset (Ruan et al., 2015) and soil data from the FAO Harmonized World Soil Database (Nachtergaele et al., 2010) using the parametrization described in Section 2.3.3. Spatial variability of PAWC (from 150 to 400 mm) across the study region is shown in Figure S4.

2.4.2 Experimental design and sources of variance

A regional simulation experiment was designed as a factorial combination of three sources of variance (i) 1325 climate×soil grids (30 arc-minute resolution), (ii) four varietal types (GS, SS, PS, and PI), and (iii) two irrigation strategies (rainfed and irrigated). The combination of the three sources of variance resulted in 10,600 combinations of factors. These combinations were simulated over 30 years (1980–2009) which resulted in a total of 318,000 biomass simulations (Figure 1). We fixed all crop management factors over time and space (except planting date, see Section 2.4.3) to separately evaluate the varietal type and irrigation strategy effects on biomass variability from other factors. The varietal types were parametrized using the same clustering explained in Section 2.3.3 for model validation. The automatic irrigation strategy during the crop season was configured to maintain nonlimiting water conditions for the crop under irrigation. A soil water availability factor was calculated by dividing actual soil water supply [SW (actual soil water) − LL15] by the potential soil water supply (DUL − LL15) every day. Irrigation was applied when this ratio was less than 1 (i.e., the difference between the supply and demand). Automatic fertilizer application was set in the model to supply N in response to crop demand and minimize the degree of stress on growth rates to 0. Soil water and N were reset to their default values at planting every year. The crop was harvested when it achieved physiological maturity (1204, 1230, 1266, and 1363°Cd from emergence to physiological maturity for GS, PI, SS, and PS, respectively), or it was killed by frost, or at 120 DAP. We used the fixed harvest date (120 DAP) to end the growing season for seasons where the photosensitive varietal type did not achieve physiological maturity or was not killed by frost.
2.4.3 | Planting date estimation

Sorghum planting dates in the United States are primarily governed by crop rotation and are adjusted according to thermal conditions (minimum threshold for soil temperature) (Ciampitti et al., 2019). The recommended mean air temperature for optimal germination ranges from 15 to 23°C. Minimum temperatures below 10°C can delay sorghum germination (Anda & Pinter, 1994) and, after germination, the crop does not grow and develop with mean temperatures below 11°C (Hammer et al., 1993). Farmers in the Northern Great Plains/Late Production region (OK, KS, AR, and MO) sow sorghum after the corn (Z. mays L.) planting ends, corresponding to mean temperatures >11°C (Figure S5a). On the other hand, planting dates in the Southern Great Plains/Early Production region (TX, LA, and MS) are also driven by soil moisture and rainfall events around the planting window. For this study, we set up in pAPSIM a variable planting rule at 30 arc-minute resolution based on the median corn planting date, mean air temperature, and rainfall because there is lack of spatial data for sorghum planting dates at high resolution in the United States. The USDA Corn Crop Progress Maps are gridded geospatial weekly datasets (9 km resolution) which are fully synthetic representations of confidential, county corn planting dates (Rosales, 2021; Figure S5b).

We used the 2020 USDA Corn Crop Progress Map to estimate the planting date windows for sorghum at 30 arc-minute resolution in the study region. A sorghum window planting date from 15 to 45 days after the median corn planting date was applied for both irrigation strategies (rainfed; Figure S5a and irrigated; Figure S5b). The planting was executed when the mean air temperature was, at least, 11°C for 7 consecutive days. Under rainfed conditions, there was also a rainfall requirement of, at least, 30 mm of cumulative rainfall during 14 consecutive days to sow the crop.

2.4.4 | Environmental clustering

In order to study the G×E×M interactions, environmental clusters were created by the algorithm K-means (Kumar et al., 2011; Vassilvitskii & Arthur, 2006) applying scikit-learn Python 3.8.5 (https://www.python.org). K-means groups similar data points together to discover underlying patterns. The clustering was applied using the sum of intercepted solar radiation \(ir\) (Equation 1; Figure 2a–c), mean of temperature stress factor \(tp\) (Equation 2; Figure 2d–f), and mean of water stress factor \(sw\) (Equation 3; Figure 2g–i) from planting to 40 DAP (P–40), from 40 to 80 DAP (40–80), and from 80 DAP to harvest (80–H) defined as:

\[
ir = \sum_{k=0}^{r1} radn_k \times I_k/I_0, \tag{1}
\]

where \(radn\) is the daily solar radiation (MJ m\(^{-2}\)) and \(I/I_0\) is the fraction of solar radiation that the crop intercepts each day,

\[
 tp = \sum_{k=0}^{r1} f t_k, \tag{2}
\]

where \(ft\) is the temperature response of photosynthesis (i.e., RUE in the model) which varies from 1 between 20 and 35°C to 0 when the mean air temperature is <8 or >50°C.

\[
 sw = \sum_{k=0}^{r1} f w_k, \tag{3}
\]

The soil water availability factor (Section 2.4.2) was used to derive the stress factor for leaf expansion \(fw\). When \(fw = 0\), the crop was under complete water stress and when \(fw = 1\), there was no crop water stress. In all cases, \(t0\) and \(r1\) indicate the start and end of each period considered in the calculation (planting, 40 DAP, and 80 DAP for \(t0\) and 40 DAP, 80 DAP, and harvest for \(r1\). The cluster number \((k = 3; \text{Figure 2j})\) was determined based on the Elbow method for optimal clustering (Bholowalia & Kumar, 2014; Ding & He, 2004; Figure S7) based on the patterns of \(ir\), \(tp\), and \(sw\) during the crop growing season across the study region (Table 2). The clusters were named accordingly to the mean temperature during the three periods assessed (P–40, 40–80, and 80–H) in high (high-temperature), mid (mid-temperature), and low (low-temperature; Figure 2k).

2.4.5 | Spatial visualization, summary statistics, and correlation analysis

The simulated biomass output by pAPSIM was extracted from NetCDF files. Only data above the 2.5% percentile (5.8 t ha\(^{-1}\)) of the entire dataset was considered for the analysis. This biomass threshold compensated the minimum operational costs of sorghum production in the United States (315 USD ha\(^{-1}\); USDA NASS, 2021a) considering a reference sorghum biomass market price of 65 USD t\(^{-1}\) (Sun et al., 2020). The selected cutoff point is also aligned with previous crop failure definitions for grain sorghum of 3 t ha\(^{-1}\) (Ciampitti et al., 2020; Grains Research & Development Corporation, 2017; Whish et al., 2005). Simulated biomasses below this threshold were considered as crop failures and quantified by the failure risk as the number of years with crop failure/30...
(total of years that the crop was sown) × 100. We mapped the spatial patterns of simulated biomass (mean and SD) for each varietal type and irrigation strategy at a 30 arc-minute resolution across the study region using xarray, Cartopy, GeoPandas, GDAL, Fiona, and Shapely Python packages. Summary statistics and plotting by environments (high, mid, and low) were also developed using scipy.stats.linregress, pandas, NumPy, and Matplotlib Python packages. We calculated the Pearson’s correlation coefficient (across years [1980–2009] for a given grid cell) between the simulated biomass and climate variables (global solar radiation sum, rainfall sum, and mean daily temperature from planting to harvest) and stress factors (sum of ir, mean tp, and mean sw for P–40, 40–80, and 80–H) for each varietal type × irrigation combination.

2.4.6 | Variance partitioning

To determine the contribution of each source of variance to the total biomass variability, two variance-based sensitivity indices were computed for each grid cell and growing season (Monod et al., 2006). According to this method, the variance of the simulated crop biomass is partitioned into fractions which can be attributed to each source. Specifically, variance sources were global solar radiation, mean temperature, rainfall, varietal type, irrigation strategy, planting date, and PAWC. Two indices, namely Main Effect (Equation 4) and Total Effect (Equation 5), were used to separate the variance caused by one source from the variance caused by the interaction.

\[
\text{Main effect}_i = \frac{\text{Variance}(E[Biomass X_i])}{\text{Variance}(Y)}, \tag{4}
\]

and

\[
\text{Total effect}_i = 1 - \frac{\text{Variance}(E[Biomass X_{-i}])}{\text{Variance}(Y)}, \tag{5}
\]

where \(E[Biomass X_i]\) denotes the expected value of sorghum biomass across all sources \(X_i\), while \(E[Biomass X_{-i}]\)
is the expected value of sorghum biomass across all sources except $X_i$. In other words, Main Effect explains the share of the components to crop biomass variability without interactions, that is, if Main Effect = 1, the assessed factors explain the entire proportion of sorghum biomass variability, but if Main Effect < 1, residuals exist which means additional factors are needed to explain this variability. Total Effect represents the interaction of a given factor with other factors, that is, high Total Effect values for a given factor denote high interactions of that factor with other factors, therefore, Total Effect does not include residuals. We used percentile threshold to classify environmental factors (global solar radiation, mean temperature, rainfall, and PAWC) into three categories of low, medium, and high, that is, the low class is defined as below the 33% percentile, medium as between the 33% and 66% percentiles, and high above the 66% percentile of the values distribution.

3 | RESULTS

3.1 | Field-scale modeling

The observed biomass varied between environments and genotypes from 0.5 t ha⁻¹ in Columbia City, IN (genotype Dale) to 34.7 t ha⁻¹ in Roper, NC (E5S2000; Figure 3a,b). The analysis of the field experimental data showed that observed biomass strongly varied for a given genotype (SD from 1.7 t ha⁻¹ for SPSordan79 to 6.0 t ha⁻¹ for M81E) or environment (SD from 2.3 t ha⁻¹ in Lexington (KY) and Corpus Christi [TX] to 6.1 t ha⁻¹ in Roper [NC]; Table S3). The performance of APSIM to simulate final sorghum biomass (harvests between 80 and 176 DAP) for 21 genotypes (including grain and forage sorghum) under 16 environments was reasonable, as evidenced by the values obtained for RMSE = 4.4 t ha⁻¹ and CCC = 0.68 (Figure 3b), although there were still some situations where simulations were greatly under or overestimated (Figures S8 and S9; Table S4). It is important to note this summary performance of the model may be affected by the imbalance of the presence of each genotype in each environment (Figure S3). The SD of simulated biomass (5.4 t ha⁻¹) was close to the SD in the observed field data for final harvests (5.5 t ha⁻¹). The accuracy of APSIM was higher when simulating in-season sorghum biomass of 18 genotypes at West Lafayette, IN (i.e., harvests within the growing season; RMSE = 3.4 t ha⁻¹, CCC = 0.86; Figure 3c; Table S5; Yang et al., 2021) than final biomass in other environments (Figure 3b). The genotypes had a bigger effect on the deviation between simulated and observed values than environments (Figure 3d,e).

3.2 | Regional-scale modeling

3.2.1 | Spatial variability of sorghum biomass

The mean, SD, and coefficient of variation (CV) of simulated biomass were markedly influenced by environment, varietal type, and irrigation strategy (Table 3). Density distributions consistently showed differences in sorghum biomass between varietal types (negative skew for GS and PS) across environments and illustrated the reduction of CV of biomass under the irrigated strategy (Figure 4). Although PI showed the highest simulated sorghum biomass under the irrigated high environments (27.5 t ha⁻¹ in the Northwest TX), this GxE combination showed the highest CV of biomass under rainfed conditions (26.1%) (Figure 5). PS exhibited the lowest simulated sorghum biomass under the rainfed low environments (5.8 t ha⁻¹, center WI; Figure 5).

Under rainfed conditions, the mean simulated biomass ranged from 11.1 t ha⁻¹ for GS under the low environments to 15.7 t ha⁻¹ for PI under the mid environments (Figure

### TABLE 2

| Factor   | Period | Mean | Mid | Low | Mean | Mid | Low |
|----------|--------|------|-----|-----|------|-----|-----|
| $ir$     | P–40   | 101  | 58  | 62  | 80   | 56  | 47  |
| MJ m⁻²   | 40–80  | 756  | 688 | 640 | 124  | 139 | 79  |
| $tp$     | P–40   | 0.98 | 0.93| 0.91| 0.12 | 0.15| 0.10|
| 0–1      | 40–80  | 1.00 | 0.99| 0.96| 0.00 | 0.02| 0.03|
| $sw$     | P–40   | 0.89 | 0.85| 0.90| 0.04 | 0.07| 0.10|
| 0–1      | 40–80  | 0.87 | 0.92| 0.98| 0.12 | 0.15| 0.10|
| 80–H     | 0.85  | 0.91| 0.98| 0.24 | 0.17| 0.06|
On average, SD of simulated biomass was less for the GS (2.0 t ha\(^{-1}\)) in comparison with SS (2.3 t ha\(^{-1}\)), PS (2.2 t ha\(^{-1}\)), and PI (3.0 t ha\(^{-1}\)); Figure 5). Under irrigated conditions, the mean simulated biomass ranged from 11.7 t ha\(^{-1}\) for GS under the low environments to 21.6 t ha\(^{-1}\) for PI under the high environments (Figure 5; Table 3). The range of SD of biomass was reduced under irrigated conditions and it varied from 1.5 t ha\(^{-1}\) (GS) to 1.9 t ha\(^{-1}\) (PS). Under the mid environments, rainfed biomass was highest in the states of NE, IL, and IN (Figure 5). Spatial variance of flowering date and harvesting date changed mainly due to varietal type and environment (Figure S10). On average, we found a negative correlation for biomass versus flowering date and a nonsignificant correlation between biomass versus harvesting date (Figure S11). Although PS showed the maximum flowering date and cycle length, on average this varietal type did not show the maximum sorghum biomass (Table 3).

Using the 2.5th percentile across all environments as a “failure” cutoff (seasonal crop biomass < 5.8 t ha\(^{-1}\)), it was found that most “failures” occurred in the low environments (82.4% of the total failed seasons including rainfed and irrigated scenarios; Figure 6). For low environments, the spatial distribution of crop failures reflected the incidence of mean temperature less than \(-20^\circ C\) and the effect was independent of irrigation status. However, under rainfed conditions, it was related to soil moisture condition around planting in the mid- (9.1% of the total failures)
and high environments (8.5% of the total failures) mainly in the west areas of TX, OK, and KS (Figure 6). Maximum failure risks (>80%, i.e., the crop failed in 24 years out of 30) were found for PS in the low environments under high latitude areas such as west NY (Figure 6).

### 3.2.2 | G×E×M drivers of simulated sorghum biomass

Overall, simulated biomass showed different responses to spatial variability in climate factors (global solar radiation, mean temperature, and rainfall) for different varietal types and irrigation strategies across the study region (Figure S12). Main spatial differences were found between the low environments of the northern regions and the other regions and their differential response to irrigation strategies (Figure S12).

In the high environments (i.e., high mean temperature environments, most of TX, OK, KS, AR, MO, KY, TN, FL, and South IL and IN and West of SC, NC, and VA), there was a positive correlation between irrigated biomass and global solar radiation ($r = 0.57–0.76$), maximum temperature ($r = 0.41–0.66$), and irrigation ($r = 0.54–0.75$), while the biomass was negatively correlated with planting date (no correlation—$-0.43$; Figure 7a). Under rainy conditions, the correlation between biomass and all factors was negative ($r$ up to $-0.7$ for maximum temperature), except for rainfall ($r = 0.66–0.75$) and PAWC ($r = 0.26–0.33$). Under the mid environments (i.e., mid mean temperature...
environments, most of NE, IA, AL, GA, West SC, NC, VA, and OH, and Central and North IN and IL), the pattern of correlation between biomass and environmental and crop management factors were similar to the high environment, except for the effect of minimum temperature which was positive ($r = 0.33–0.74$) and PAWC which was negligible ($r = 0$ correlation–0.22). In contrast, the correlation of simulated biomass to global solar radiation and to maximum and minimum temperatures were positive for the low environments (i.e., low mean temperature environments, North MN and WI and most of MI, OH, PA, and NY) under rainfed and irrigated scenarios ($r$ up to 0.77). For all environments and varietal types, the correlation between biomass and irrigation was positive ($r = 0.51–0.75$). The drivers of biomass variability differed between varietal types mainly in the mid environments under rainfed conditions. For example, the correlation between biomass and minimum temperature was positive for SS and PS, and negligible for the other varietal types under these environments (Figure 7a).
Simulated biomass showed different responses to variations in stress factors \( (ir, tp, \text{and } sw) \) for different varietal types, irrigation strategies, periods of the growing season, and environments (Figure 7b). Under the high environments, rainfed biomass was positively correlated with \( sw \) and irrigated biomass with \( ir \) during the complete crop growing season although, between 40 and 80 DAP, the correlations were the highest for all varietal types \( (r = 0.64-0.86; \text{Figure 7b}) \). Similarly, correlations between biomass and stress factors were found in the mid environments, except for \( tp \) which showed positive correlation for all varietal types, particularly for PS \( (r = 0.71) \). Under the low environments, biomass was positively correlated with \( ir \) and \( sw \) for all varietal types except PS.
environments, $ir$ and $tp$ outweigh the effect of $sw$ for all varietal types and irrigation strategies ($r = 0.31–0.82$; Figure 7b). A complete list of linear correlation coefficients (including $p$-values) between biomass, environmental and crop management variables can be found in Table S6.

### 3.2.3 | Partitioning of the sorghum biomass variance

To identify which components were associated with crop biomass variance, we partitioned simulated biomass into different GxExM factors (Figure 8). When we compared the results by environment and irrigation strategy, varietal type explained the greatest variance in biomass, but it differed depending on environment mainly due to differences in their climatic patterns (Figure S1). Biomass variance was mainly explained by varietal type (50%–76%) in all environments xirrigation strategies combinations, except in the high- and mid environments under rainfed conditions where rainfall had the major effect (25% and 45% for mid- and high environments; Figure 8a). However, when mean temperature during the growing season decreased from 25°C (high environments) to 18°C (low environments), the contribution of mean temperature to biomass variance increased from 7% to 34% (rainfed) and from 4% to 36% (irrigated). Main effect of planting date never was higher than 4% (Figure 8a) although a stronger negative correlation was found for the high environments in comparison with the other environments (Figure S13). Particularly for rainfed environments, we found higher residual components (5%–14%) than irrigated environments (4%), indicating there were additional factors that explain some of the biomass variance. Varietal type had the larger interactions with other factors independently of the environment and irrigation strategy (Figure 8b).

### 4 | DISCUSSION

Understanding the sources of sorghum biomass variability in the United States assists in developing confidence in bioenergy projections to achieve the 58 million tons of biomass sorghum in 2040 targeted by the Department of Energy 2016 Billion-Ton Report (Langholtz et al., 2016). This study quantified the spatial variability of sorghum biomass, the associated G, E, and M drivers and main sources of variance under contrasting varietal types and irrigated conditions in the potential areas for energy sorghum in the United States. With a better understanding of the relative contribution of each biomass driver, specific crop practices can be implemented to ensure the biomass variability is adequately reduced and potential production can be achieved. Further, the environmental classification and simulation analysis provide a baseline to understand the impacts of climate change on sorghum biomass production.

#### 4.1 | Field-scale modeling

Previous evaluations of APSIM reported RMSE for sorghum biomass estimations in the range of 1.5–1.6 t ha$^{-1}$ (Akinseye et al., 2017; Carcedo & Gambin, 2019). The results reported by Yang et al. (2021) showed a better accuracy for in-season biomass estimations at West Lafayette, IN (RMSE = 3.4 t ha$^{-1}$; CCC = 0.86) compared to the final biomass estimations we found under the other
environments (RMSE = 4.4 t ha\(^{-1}\); CCC = 0.68; Figure 3). The reason for these differences could be related to the fact that the fertility and the initial conditions of the soil at West Lafayette considered excellent for corn production, were also optimal for sorghum growth and the measurements, and crop management records of high quality. By comparison, the other Indiana sites would be considered unsuitable for corn production (Bu and CC) or moderately good for corn (La). For example, Kent et al. (2020) reported RMSE = 4.36 t ha\(^{-1}\) and CCC = 0.41 comparing rainfed biomass from DAYCENT with experimental data, while Gautam et al. (2020) found RMSE = 5.5 t ha\(^{-1}\) comparing simulated rainfed biomass with county observed datasets from the US National Agricultural Statistics Service with the same model. APSIM responded satisfactorily to multiple genotypes and environments (climate–soil–management combinations) across a wide range of biomass observations (0.5–34.7 t ha\(^{-1}\)). The model’s fitness-for-purpose for this paper was further supported by the spatial patterns of sorghum biomass in the potential production areas for energy sorghum in the United States that align with field experimental data across the country (Figure 5).

Lack of specificity in model inputs is a crucial source of uncertainty in crop yield estimations (Fleisher et al., 2017; Ojeda et al., 2021a, 2021b, 2021c). Part of the prediction error found in this study (Figure 3) is explained by the fact that field experiments used for model validation were not specifically designed for modeling purposes, and therefore, there were gaps for model inputs such as a lack of measured soil data to initialize the model or crop management records. This highlights the need for minimum dataset guidelines be developed for future studies that ensures measurement of key soil attributes (e.g., soil moisture, NO\(_3\), and NH\(_4\) before planting) that will enhance the reuse of biophysical data in calibration and validation of crop models.

4.2 | Regional-scale modeling

4.2.1 | Spatial variability of sorghum biomass

Sorghum biomass was simulated over a period of 30 years at 1325 locations (climatemxsoil combinations) across the potential sorghum cropping areas in the United States using present planting dates, nonlimiting N fertilization, and recommended crop management practices (Ciampitti et al., 2019, 2020). The outcomes of this research quantify the sensitivity of sorghum biomass to varietal type and irrigation strategy in different environments. For potential sorghum cropping areas in the United States (Langholtz et al., 2016), in average, varietal type (first), rainfall, and mean temperature (second) were the most important drivers of sorghum biomass variability (Figure 8). This was also reflected in the probability density distribution (Figure 4) and in the spatial pattern of simulated biomass across the high-, mid-, and low environments evaluated (Figure 5). These findings emphasize that potential sorghum biomass can be achieved (up to 27.5 t ha\(^{-1}\)) when (i) the crop is fully irrigated (i.e., preventing water stress) and (ii) varietal types with high RUE (RUE > 1.3 MJ m\(^{-2}\)), high extinction coefficient (ec > 0.56), and relatively short timing to flowering (<80 DAP; e.g., PI and SS) are used in water-limited environments, for example, West TX, OK, or KS.

Aside from Huntington et al. (2020), most regional biomass estimations in the United States have been generated under rainfed conditions (Daly et al., 2018; Gautam et al., 2020; Langholtz et al., 2016). As an opportunity crop, sorghum (predominantly for grain) is mainly grown under nonirrigated conditions (<38% of the area under irrigation; USDA NASS, 2021b) mostly in Texas (Ciampitti et al., 2019). Field experiments have demonstrated that irrigation can increase energy biomass production in several water-limited regions of the United States such as the Southern Great Plains (Yimam et al., 2015), the Midwest (Roby et al., 2017), and the Southeast (Rocatelli et al., 2012). Huntington et al. (2020) recently projected irrigated biomass production to 2099 using machine learning algorithms based on climate variables. They found a range of 25–27.5 t ha\(^{-1}\) of sorghum biomass for West TX, OK, and KS where irrigation was the most important driver of biomass variability. Water is becoming more limiting in many of these regions where sorghum is now grown (TX, OK, KS, and NE) as the Ogallala aquifer becomes depleted. This will limit irrigation options in these regions or make it uneconomical, including for biomass production. In this scenario, our research can help to identify suitable agro-ecozones where sorghum will be adapted for rainfed production. Gautam et al. (2020) reported higher rainfed biomass estimations (up to 19 t ha\(^{-1}\)) than our analysis for similar US environments (mainly in LA, AR, and MO) using the DAYCENT model. The higher biomass estimations of that study compared to our results could be related to the fact that they used (i) a biogeochemical model instead of a crop model carrying differences in model structure, (ii) different climate (Global Historical Climatology Network, GHCN) and soil (US National Cooperative Soil Survey, SSURGO) data sources as input to the model which may generate differences in the data resolution and parameter estimations, and (iii) uniform genotype parameters across all environments. This contrast reinforces the value of exploring a wide range of G×E×M combinations when actual and future regional...
simulations of sorghum biomass are generated at regional scale, especially for irrigated conditions.

Sorghum crop failures can be driven by different factors associated with the growing environment (Hammer et al., 2014). For the low environments, crop failures were associated with low minimum temperatures that killed the crop (minimum temperatures less than 0°C) or delayed the crop development due to a slow thermal time accumulation. For example, when the crop was harvested before the stage of flag leaf biomass was significantly reduced and the crop failed. This effect was higher for photoperiod-sensitive types (Figure 6) due to greater thermal time requirement from the end of juvenile to floral initiation with respect to other varietal types (Table S2). On the other hand, all crop failures occurred under rainfed conditions (Figure 6) and mainly in soils with low PAWC in the first soil layers (data not shown) in the mid- and high environments. The main drivers of these crop failures were the low temperatures and soil moisture around planting and small and less frequent rainfall events during the beginning of the growing season. APSIM has a time threshold after planting which limits the crop emergence (150°Cd−1). Therefore, when these environmental conditions occurred and a delay exists between planting and emergence, the crop did not emerge.

4.2.2 | G×E×M drivers of simulated sorghum biomass

Spatial gradients in global solar radiation, mean temperature, and rainfall, and their corresponding stress factors (ir, tp, and sw) influenced sorghum biomass under both irrigated and rainfed conditions (Figure 7; Figure S12). Similar multiscale environmental crop response characterizations have been previously reported for corn in the United States (Jin et al., 2017) and New Zealand (Teixeira et al., 2017), for sunflower (Helianthus annuus L.) in France (Casadebaig et al., 2020), and for wheat (Triticum aestivum L.; Chenu et al., 2013) and potatoes (Solanum tuberosum L.; Ojeda et al., 2021b; Ojeda et al., 2020) in Australia. However, in this manuscript we combined environmental clustering based on three crop stressors (ir, tp, and sw), correlation and variance partitioning analysis to assess the biomass sorghum variability at a large scale in potential growing environments of the United States. For the high- and mid environments (Figure 2), our findings suggest that irrigation is needed to achieve potential biomass production and, under rainfed conditions, the right selection of varietal type and planting date are key to start the season with high moisture levels and reduce the interannual variability of biomass (Figure 5; Table 3) and failure risk (Figure 6). However, under the low- and mid environments (e.g., irrigated areas in NE), the effect of temperature on RUE and the radiation interception after 40 DAP play the most important role to achieve high crop biomass (Figure 7b). Therefore, early planting dates (though avoiding last spring frost) and the use of varietal types with shorter grow cycles (such as PI; Figure S10) would benefit the crop growing conditions under these environments.

The variance partitioning method is a sensitivity analysis that allows quantification of the variance contribution of selected input variables on model outputs such as crop grain yield or biomass (Monod et al., 2006; Ruget et al., 2002). This method has been applied mainly to quantify the contribution of various climatic factors to corn and wheat grain yield under European environments (Webber et al., 2018). Based on our results, we identified that for environments where the crop is not constrained by temperature (high environments) and water (i.e., when it is irrigated), varietal type [individually (Figure 8a) and combined with other factors (Figure 8b)] outweighed the contribution of other crop management factors (planting date) and environmental factors (solar radiation, mean temperature, rainfall, and PAWC) to crop biomass variability. However, when mean temperature had a major role on crop growing conditions (low- and mid environments), mean temperature starts to play an important role on biomass variance (Figure 8a). These results highlight the importance of including the genotype as a factor when a factorial simulation analysis is configured to assess biomass variability at large scales where results may vary across environments.

4.2.3 | Limitations of this study and future research

The study region used in this paper (Figure 2) included environments where sorghum is not currently grown in the United States (Ciampitti et al., 2019, 2020). The regional modeling analysis was implemented over the potential areas of energy sorghum production across the United States (Langholtz et al., 2016) where low mean temperature environments are included which are not ideal for sorghum biomass production (Hammer et al., 1993). Surprisingly, irrigated biomass productions up to 22.2 t ha−1 were predicted with PI genotypes under these environments (e.g., central-south MN and south MI; Figure 5). However, no validation data were found for model testing under these environments (Figure S2). Further field datasets for crop model validation are required to improve the confidence of APSIM estimations from these environments.
Planting date is a major factor in determining the potential biomass production of any crop across environments (Baum et al., 2020) and usually information about this factor is scarce at high-resolution spatial levels (Minoli et al., 2019). The objective of this study was not to assess the effect of planting dates on biomass variability. However, this factor is an essential input for the model which defines the environment when the crop will grow.

In addition to climate, soil, genotype, and irrigation strategy, other crop management factors such as N fertilization (form, timing, method of application), initial soil moisture, other plant nutrients (P, K, and S) and pH, and previous crop may affect biomass differently and the influence of these factors has not been investigated in this study. For example, a reset rule for soil water and N was applied every year to avoid the effects of previous year on following years. However, sorghum farmers use different crop rotations (single or double crops) based on their cropping system (grain, silage, dual-purpose, and energy; Olson et al., 2012; Rooney, 2004) which vary accordingly with the region, farm history, or land tenure. For example, over 70% of the farmland in IN is not farmed by the landowner; instead, it is rented often on annual contracts which makes difficult to record information about previous crops (NASS, 2014).

This study was conducted using only one single crop model (APSIM Classic 7.9) under the pSIMS framework (pAPSIM) and findings may be different using other crop models and gridded platforms, which vary in structure and complexity (Chapagain et al., 2020; Thorburn, 2017). Our results were presented at only one spatial resolution level (30 arc-minutes) due to the availability of long-term climate and soil data. However, pAPSIM simulations could increase the spatial detail in future studies where climate, soil, and crop management information are available at higher resolution levels than the used in this study.

5 CONCLUSIONS

This work was the first attempt to validate and apply the APSIM sorghum module across a wide range of G×E×M for predicting sorghum biomass at regional scale in the United States. This study demonstrates: (1) Maximum sorghum biomass can be achieved combining varietal types with high RUE and ec and potential irrigation management (non-water stress) in the high-temperature environments of the United States. (2) To expand the energy sorghum-growing areas to low-temperature environments, special attention needs to be paid to crop management practices that improve the environmental conditions during the growing season (optimum temperatures, availability of solar radiation, and soil water under rainfed conditions) such as the use of varietal types with short period growth time or the selection of planting dates accordingly to the regional frost calendar. (3) Varietal type parameterization was critical for explaining differences in biomass for both irrigation strategies under contrasting environments. (4) The environmental clustering based on stress factors and the variance contribution analysis is needed to discriminate the main G×E×M drivers of biomass variability and to prioritize farm decisions.

ACKNOWLEDGMENTS

This work was supported by the US Department of Energy ARPA-E Program Award Number DE-AR0001135 and by the Agriculture and Food Research Initiative Competitive grant no. 2011-68005-30411 from the National Institute of Food and Agriculture. The authors gratefully acknowledge the provision of field experimental data for model validation from the University of Idaho (Dr. Kent Jeffrey). They also thank Chrisbin James (School of Agriculture and Food Sciences, The University of Queensland) and Diego Perez (Bupa) for technical support on Python and Linux coding to set up pAPSIM. APSIM is provided free for research and development use (see www.apsim.info for details). They thank Maria Salas-Fernandez (Iowa State University) for useful discussion on initial design of simulations for this paper.

CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding authors upon reasonable request.

ORCID

Jonathan J. Ojeda https://orcid.org/0000-0002-9172-0059
Graeme Hammer https://orcid.org/0000-0002-1180-7374
Kai-Wei Yang https://orcid.org/0000-0002-2234-3747
Mitchell R. Tuinstra https://orcid.org/0000-0002-5322-6519
Peter deVoil https://orcid.org/0000-0001-6584-4376
Greg McLean https://orcid.org/0000-0002-8491-2984
Isaiah Huber https://orcid.org/0000-0003-4721-1406
Jeffrey J. Volenec https://orcid.org/0000-0002-5455-6157
Sylvie M. Brouder https://orcid.org/0000-0002-7785-5336
Sotirios Archontoulis https://orcid.org/0000-0001-7595-8107
Scott C. Chapman https://orcid.org/0000-0003-4732-8452
Hammer, G., McLean, G., Doherty, A., van Oosterom, E., & Chapman, S. (2013). Sorghum: State of the art and future perspectives (Vol. 58, pp. 215–239). ASA and CSSA. https://doi.org/10.2134/agronmonogr58.c10

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 and Pasture Science, 65*(7), 614–626. https://doi.org/10.1002/cpsc.20110

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. *Journal of Experimental Botany, 61*(8), 2185–2202. https://doi.org/10.1093/jxb/erq095

Heitman, A. J., Castillo, M. S., Smyth, T. J., Crozier, C. R., Wang, Z., Heiniger, R. W., & Gehl, R. J. (2017). Nitrogen fertilization effects on yield and nutrient removal of biomass and sweet sorghum. *Agronomy Journal, 109*(4), 1352–1358. https://doi.org/10.2134/ajron2016.12.0710

Hoffmann, L., & Rooney, W. L. (2014). Accumulation of biomass and compositional change over the growth season for six post-1950 sorghum lines. *BioEnergy Research, 7*(3), 811–815. https://doi.org/10.1007/s12155-013-9405-5

Holzworth, D. P., Huth, N. I., deVoil, P. G., Zurcher, J. E., Herrmann, N. I., McLean, G., Chenu, K., van Oosterom, E. J., Snow, V., Murphy, C., Moore, A. D., Brown, H., Whish, J. P. M., Vorrall, S., Fainges, J., Bell, L. W., Peake, A. S., Poulton, P. L., Hochman, Z., ... Keating, B. A. (2014). APSIM – Evolution towards a new generation of agricultural systems simulation. *Environmental Modelling & Software, 62*, 327–350. https://doi.org/10.1016/J.ENVSOFT.2014.07.009

Huntington, T., Cui, X., Mishra, U., & Scown, C. D. (2020). Machine learning to predict biomass sorghum yields under future climate scenarios. *Biofuels, Bioproducts and Biorefining, 14*(3), 566–577. https://doi.org/10.1002/bbb.2087

Jin, Z., Zhuang, Q., Wang, J., Archontoulis, S. V., Zobel, Z., & Kotamartih, V. R. (2017). The combined and separate impacts of climate extremes on the current and future US rainfed maize and soybean production under elevated CO2. *Global Change Biology, 23*(7), 2687–2704. https://doi.org/10.1111/gcb.13617

Kent, J., Hartman, M. D., Lee, D. K., & Hudiburg, T. (2020). Simulated biomass sorghum GHG reduction potential is similar to maize. *Environmental Science & Technology, 54*(19), 12456–12466. https://doi.org/10.1021/acs.est.0c01676

Kumar, J., Mills, R. T., Hoffman, F. M., & Hargrove, W. W. (2011). Parallel k-means clustering for quantitative ecoregion delineation using large data sets. *Procedia Computer Science, 4*, 1602–1611. https://doi.org/10.1016/j.procs.2011.04.173

Langholtz, M. H., Stokes, B. J., & Eaton, L. M. (2016). 2016 billion-ton report: Advancing domestic resources for a thriving bioeconomy, Volume I: Economic availability of feedstock (Vol. 2016, pp. 1–411). Oak Ridge National Laboratory, Oak Ridge, Tennessee, Managed by UT-Battelle, LLC for the US Department of Energy.

Lawrence, I., & Lin, K. (1989). A concordance correlation coefficient to evaluate reproducibility. *Biometrics, 255–268*. https://doi.org/10.2307/2532051

Lee, D. K., Aberle, E., Anderson, E. K., Anderson, W., Baldwin, B. S., Baltensperger, D., Barrett, M., Blumenthal, J., Bonos, S., Bouton, J., Bransby, D. I., Brummer, C., Burks, P. S., Chen, C., Daly, C., Egenolf, J., Farris, R. L., Fike, J. H., Gaussoin, R., ... Owens, V. (2018). Biomass production of herbaceous energy crops in the United States: Field trial results and yield potential maps from the multiphase regional feedstock partnership. *GCB Bioenergy, 10*(10), 698–716. https://doi.org/10.1111/gcb.12493

Liang, X., Haooghoenoom, G., Voulgaragki, S., Boote, K. J., & Vellidis, G. (2020). Deriving genetic coefficients from variety trials to determine sorghum hybrid performance using the CSM-CERES-Sorghum model. *Agronomy Journal, 102*, 1–12. https://doi.org/10.1002/agj2.20644

Lopez, J. R., Erickson, J. E., Asseng, S., & Bobeda, E. L. (2017). Modification of the CERES grain sorghum model to simulate optimum sweet sorghum rooting depth for rainfed production on coarse textured soils in a sub-tropical environment. *Agricultural Water Management, 181*, 47–55. https://doi.org/10.1016/j.agwat.2016.11.023

MacCarthy, D. S., Sommer, R., & Vlek, P. L. G. (2009). Modeling the impacts of contrasting nutrient and residue management practices on grain yield of sorghum (Sorghum bicolor (L.) Moench) in a semi-arid region of Ghana using APSIM. *Field Crops Research, 113*(2), 105–115. https://doi.org/10.1016/j.fcr.2009.04.006

Maw, M. J. W., Hooux III, J. H., & Fritschi, F. B. (2016). Sweet sorghum ethanol yield component response to nitrogen fertilization. *Industrial Crops and Products, 84*, 43–49. https://doi.org/10.1016/j.indcrop.2016.01.038

Merlos, F. A., Monzon, J. P., Mercau, J. L., Taboada, M., Andrade, F. H., Hall, A. J., Jobbag, E., Cassman, K. G., & Grassini, P. (2015). Potential for crop production increase in Argentina through closure of existing yield gaps. *Field Crops Research, 184*, 145–154. https://doi.org/10.1016/j.fcr.2015.10.001
Minoli, S., Egli, D. B., Rolinski, S., & Müller, C. (2019). Modelling cropping periods of grain crops at the global scale. *Global and Planetary Change*, 174, 35–46. https://doi.org/10.1016/j.gloplacha.2018.12.013

Monod, H., Naud, C., & Makowski, D. (2006). Uncertainty and sensitivity analysis for crop models. *Working with Dynamic Crop Models: Evaluation, Analysis, Parameterization, and Applications*, 4, 55–100.

Mutava, R. N., Prasad, P. V. V., Tuinstra, M. R., Kofoid, K. D., & Yu, J. (2011). Characterization of sorghum genotypes for traits related to drought tolerance. *Field Crops Research*, 123(1), 10–18. https://doi.org/10.1016/j.fcr.2011.04.006

Nachtsgaele, F., van Velthuizen, H., Verelst, L., Batjes, N. H., Dijkshoorn, K., van Engelen, V. W. P., Fischer, G., Jones, A., & Montanarella, L. (2010). The harmonized world soil database. In *Proceedings of the 19th World Congress of Soil Science, Soil Solutions for a Changing World*, Brisbane, Australia, 1–6 August 2010, pp. 34–37.

NASS. (2014). *Census of agriculture. 2014 Tenure, Ownership, and Transition of Agricultural Land (TOTAL)*. https://www.nass.usda.gov/Publications/AgCensus/2012/Online_Resources/TOTAL/index.php

Ojeda, J. J., Caviglia, O. P., Volenc, J. J., Brouder, S. M., & Agnusdei, M. G. (2018). Modelling stover and grain yields, and subsurface artificial drainage from long-term corn rotations using APSIM. *Agricultural Water Management*, 195, 154–171. https://doi.org/10.1016/j.agwat.2017.10.010

Ojeda, J. J., Rezaei, E., Kamali, B., McPhee, J., Meinke, H., Siebert, S., Webb, M., Ara, I., Mulcahy, F., & Ewert, F. (2021a). Impact of crop management and environment on the spatio-temporal variance of potato yield at regional scale. *Field Crops Research*, 270, 108213. https://doi.org/10.1016/j.fcr.2021.108213

Ojeda, J. J., Rezaei, E. E., Remenyi, T. A., Webb, M. A., Webber, H. A., Kamali, B., Harris, R. M. B., Brown, J. N., Kidd, D. B., Mohammed, C. L., Siebert, S., Ewert, F., & Meinke, H. (2020). Effects of soil- and climate data aggregation on simulated potato yield and irrigation water requirement. *Science of the Total Environment*, 710, 135589. https://doi.org/10.1016/j.scitotenv.2019.135589

Ojeda, J. J., Rezaei, E. E., Remenyi, T. A., Webber, H. A., Siebert, S., Meinke, H., Webb, M. A., Kamali, B., Harris, R. M. B., Kidd, D. B., Mohammed, C. L., McPhee, J., Capuano, J., & Ewert, F. (2021b). Implications of data aggregation method on crop model outputs – The case of irrigated potato systems in Tasmania, Australia. *European Journal of Agronomy*, 126. https://doi.org/10.1016/j.eja.2021.126276

Ojeda, J. J., Huth, N., Holzworth, D., Raymundo, R., Zyskowski, R., Sinton, S., Michel, A., & Brown, H. (2021c). Assessing errors during simulation configuration in crop models – A global case study using APSIM-Potato. *Ecological Modelling*, 458. https://doi.org/10.1016/j.ecolmodel.2021.109703

Ojeda, J., Volenc, J. J., Brouder, S. M., Caviglia, O. P., & Agnusdei, M. G. (2017). Evaluation of agricultural production systems simulator as yield predictor of *Panicum virgatum* and *Miscanthus x giganteus* in several US environments. *GCB Bioenergy*, 9(4). https://doi.org/10.1111/gcbb.12384

Olson, S. N., Ritter, K., Rooney, W., Kemanian, A., McCarl, B. A., Zhang, Y., Hall, S., Packer, D., & Mullet, J. (2012). High biomass yield energy sorghum: Developing a genetic model for C₄ grass bioenergy crops. *Biofuels, Bioproducts and Biorefining*, 6(6), 640–655. https://doi.org/10.1002/bbb.1357

Pembleton, K. G., Rawnsley, R. P., Jacobs, J. L., Mickan, F. J., O’Brien, G. N., Cullen, B. R., & Ramlain, T. (2013). Evaluating the accuracy of the Agricultural Production Systems Simulator (APSIM) simulating growth, development, and herbage nutritive characteristics of forage crops grown in the south-eastern dairy regions of Australia. *Crop and Pasture Science*, 64(2), 147–164. https://doi.org/10.1071/CP12372

Piñeiro, G., Perelman, S., Guerschman, J. P., & Paruelo, J. M. (2008). How to evaluate models: Observed vs. predicted or predicted vs. observed? *Ecological Modelling*, 216(3–4), 316–322. https://doi.org/10.1016/j.ecolmodel.2008.05.006

Ramirez-Villegas, J., & Challinor, A. (2012). Assessing relevant climate data for agricultural applications. *Agricultural and Forest Meteorology*, 161, 26–45. https://doi.org/10.1016/j.agrformet.2012.03.015

Raymundo, R., Ciampitti, I., & Morris, G. P. (2021). Crop modeling defines opportunities and challenges for drought escape, water capture, and yield increase using chilling-tolerant sorghum. *BioRxiv*. https://doi.org/10.1101/2021.01.27.428532

Roby, M. C., Fernandez, M. G. S., Heaton, E. A., & VanLoo, A. (2017). Biomass sorghum and maize have similar water-use efficiency under non-drought conditions in the rain-fed Midwest US. *Agricultural and Forest Meteorology*, 247, 434–444. https://doi.org/10.1016/j.agrformet.2017.08.019

Rocatelli, A. C., Raper, R. L., Balkcom, K. S., Arriaga, F. J., & Bransby, D. I. (2012). Biomass sorghum production and components under different irrigation/tillage systems for the southeastern US. *Industrial Crops and Products*, 36(1), 589–598. https://doi.org/10.1016/j.indcrop.2011.11.007

Rooney, W. L. (2004). Sorghum improvement-integrating traditional and new technology to produce improved genotypes. *Advances in Agronomy*, 83, 37–109. https://doi.org/10.1016/S0065-2113(04)83002-5

Rooney, W. L., Blumenthal, J., Bean, B., & Mullet, J. E. (2007). Designing sorghum as a dedicated bioenergy feedstock. *Biofuels, Bioproducts and Biorefining*, 1(2), 147–157. https://doi.org/10.1002/bbb.15

Rosales, A. (2021). *Crop progress and condition layers*. https://www.nass.usda.gov/Research_and_Science/Crop_Progress_Grid/CropProgressDescription.pd

Rotili, D. H., de Voil, P., Eyre, J., Serafin, L., Aisthorpe, D., Maddonni, G. Á., & Rodríguez, D. (2020). Untangling genotype x management interactions in multi-environment on-farm experimentation. *Field Crops Research*, 255. https://doi.org/10.1016/j.fcr.2020.107900

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

Ruget, F., Brisson, N., Delécolle, R., & Fairv, R. (2002). Sensitivity analysis of a crop simulation model, STICS, in order to choose the main parameters to be estimated. *Agronomie*, 22(2), 133–158. https://doi.org/10.1051/agro:2002009

Seyoum, S., Rachaputi, R., Chauhan, Y., Prasanna, B., & Fekybelu, S. (2018). Application of the APSIM model to exploit G×E×M interactions for maize improvement in Ethiopia. *Field Crops Research*, 217, 113–124. https://doi.org/10.1016/j.fcr.2017.12.012
Singh, V., Nguyen, C. T., McLean, G., Chapman, S. C., Zheng, B., van Oosterom, E. J., & Hammer, G. L. (2017). Quantifying high temperature risks and their potential effects on sorghum production in Australia. *Field Crops Research*, 211, 77–88. https://doi.org/10.1016/j.fcr.2017.06.012

Sun, F., Sarin, S. C., Cundiff, J. S., & Sert, I. O. (2020). Design of cost-effective sorghum biomass feedstock logistics-A comparison of different systems. *Biomass and Bioenergy*, 143. https://doi.org/10.1016/j.biombioe.2020.105823

Tedesci, L. O. (2006). Assessment of the adequacy of mathematical models. *Agricultural Systems*, 89(2–3), 225–247. https://doi.org/10.1016/j.agsy.2005.11.004

Teixeira, E. I., Zhao, G., de Ruijer, J., Brown, H., Ausseil, A. G., Meenken, E., & Ewert, F. (2017). The interactions between genotype, management and environment in regional crop modelling. *European Journal of Agronomy*, 88, 106–115. https://doi.org/10.1016/j.eja.2016.05.005

Thivierge, M.-N., Chantigny, M. H., Bélanger, G., Seguin, P., Bertrand, A., & Vanasse, A. (2015). Response to nitrogen of sweet pearl millet and sweet sorghum grown for ethanol in eastern Canada. *Bioenergy Research*, 8(2), 807–820. https://doi.org/10.1007/s12155-014-9558-x

USDA NASS. (2021a). *Commodity costs and returns*. Economic Research Service. https://www.ers.usda.gov/data-products/commodity-costs-and-returns/commodity-costs-and-returns/#Cost-of-ProductionForecasts

USDA NASS. (2021b). *Irrigated sorghum area Quick Stats*. https://quickstats.nass.usda.gov/

Vassilvitskii, S., & Arthur, D. (2006). k-means++: The advantages of careful seeding. In *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 1027–1035.

Wallach, D., & Thorburn, P. J. (2017). Estimating uncertainty in crop model predictions: Current situation and future prospects. *European Journal of Agronomy*, 88, A1–A7. https://doi.org/10.1016/j.eja.2017.06.001

Webber, H., Ewert, F., Olesen, J. E., Müller, C., Fronzek, S., Ruane, A. C., Bourgault, M., Martre, P., Ahabaei, B., Bindi, M., Ferrise, R., Finger, R., Fodor, N., Gabaldón-Leal, C., Gaiser, T., Jabloun, M., Kersebaum, K.-C., Lizaso, J. I., Lorite, I. J., ... Wallach, D. (2018). Diverging importance of drought stress for maize and winter wheat in Europe. *Nature Communications*, 9(1), 1–10. https://doi.org/10.1038/s41467-018-06525-2

Whish, J., Butler, G., Castor, M., Cawthray, S., Broad, I., Carberry, P., Hammer, G., McLean, G., Routley, R., & Yeates, S. (2005). Modelling the effects of row configuration on sorghum yield reliability in north-eastern Australia. *Australian Journal of Agricultural Research*, 56(1), 11–23. https://doi.org/10.1071/AR04128

Wight, J. P., Hons, F. M., Storlien, J. O., Provins, T. L., Shahandeh, H., & Wiedenfeld, R. P. (2012). Management effects on bioenergy sorghum growth, yield and nutrient uptake. *Biomass and Bioenergy*, 46, 593–604. https://doi.org/10.1016/j.biombioe.2012.06.036

Wilde, M., Hategan, M., Wozniak, J. M., Clifford, B., Katz, D. S., & Foster, I. (2011). Swift: A language for distributed parallel scripting. *Parallel Computing*, 37(9), 633–652. https://doi.org/10.1016/j.parco.2011.05.005

Yang, K.-W., Chapman, S., Carpenter, N., Hammer, G., McLean, G., Zheng, B., Chen, Y., Delp, E., Masjedi, A., Crawford, M., Ebert, D., Habib, A., Thompson, A., Weil, C., & Tuinstra, M. R. (2021). Integrating crop growth models with remote sensing for predicting biomass yield of Sorghum. *In Silico Plants*. https://doi.org/10.1093/insilicoplasts/diab001

Yimam, Y. T., Ochsner, T. E., & Kakani, V. G. (2015). Evapotranspiration partitioning and water use efficiency of switchgrass and biomass sorghum managed for biofuel. *Agricultural Water Management*, 155, 40–47. https://doi.org/10.1016/j.agwat.2015.03.018

**SUPPORTING INFORMATION**

Additional supporting information may be found in the online version of the article at the publisher’s website.

**How to cite this article:** Ojeda, J. J., Hammer, G., Yang, K.-W., Tuinstra, M. R., deVoil, P., McLean, G., Huber, I., Voleneck, J. J., Brouder, S. M., Archontoulis, S., & Chapman, S. C. (2022). Quantifying the effects of varietal type x management on the spatial variability of sorghum biomass across US environments. *GCB Bioenergy*, 14, 411–433. https://doi.org/10.1111/gcbb.12919