Characterization of seasonal pattern of drought stress using modeling approach. I. An environmental characterization reveals that water stress is currently not a major constraint to sorghum production in Mali

M. Diancoumba, …

Abstract

Drought stress is a prevalent feature of semi-arid tropical climates but when, where, how much and how often drought stress reduces yield remains largely an unanswered question. In West Africa, there is indeed significant concern that the large genotype x environment interactions remain a limitation for targeting genotypes to their appropriate environments. Assessing the extent and nature of the complexity of such interactions over time and space remains a challenge that this study tackles using crop simulation modelling. In the present study we i) parameterized two genotypes representative of the sorghum types typically grown in Mali sorghum production belt into the APSIM-sorghum platform, ii) identified and analyzed the type and frequency of each drought stress patterns experienced by these two genotypes during their life cycle, iii) and evaluated the effect of each drought stress type identified on grain yield. Two genotypes (CSM63E and CSM335) representing the sahelian and sudano-sahelian production belt in Mali were parameterized in the APSIM sorghum-platform, , i.e. novel functions were introduced into the APSIM and coefficients describing the growth and development of these two genotypes were computed using several years of observed data. Thereafter the validated model was used to characterize the seasonal variation in drought stress index (water supply/demand ratio) occurring at 22 weathers stations selected across Mali. The observation period varied greatly across sites. The seasonal water deficit patterns identified in each plant type, site, soil and year were then clustered in order to determine the major drought scenarios occurring during the cropping cycles. The frequency of occurrence of the identified environmental scenario and their impact on sorghum production was estimated for each geographical location across the sorghum production belt.

The clustering analysis revealed 3 major stress scenarios were occurring across the locations; namely a scenario with no-stress (73%, 83%), an early pre-flowering terminal stress (19%, 9%) and a flowering stress (9%, 7%) respectively for CSM335 and CSM63E. The no-stress scenario was the most dominant independently of the years and agro-ecological zones. For CSM63E and CSM335 genotypes, the early pre-flowering stress caused yield loss of 58% and 64% with respect to the no-stress scenario while the flowering stress caused only 8% and 19% yield loss. The higher frequency of occurrence of the no-stress scenario particularly in the southern region indicates that water deficit is not the major limiting factor to

33   sorghum production in this region. Also, the still relatively low yield under the no-stress scenario implies

34   that other factors (e.g. plant density, sowing date, N application) are limiting the productivity of sorghum

35   in Mali. The established framework can be further elaborated to assist sorghum improvement programs

36   to set priorities for guiding agri-interventions.

37

38   Keywords: interaction, parameterization, characterization, drought stress pattern, sorghum

39

40   **Introduction**

41   Sorghum (*Sorghum bicolor Moench*) is one of the most important multi-purpose and staple-food crops for

42   many rural communities in the drier regions of Sub-Saharan Africa. In Mali, sorghum, the second major

43   crop, is cultivated with millet on 80% of cereals grown surfaces with an average grain yield stagnating at

44   ~923 kg/ha since 26 years (USAID 2011). Sorghum low yields are explained by the existing limitations of

45   production environments where this crop is being cultivated; i.e. rain-fed conditions in marginal land with

46   low input management practices. In addition, the weather conditions vary largely across time and

47   geographical scale, making the breeding of improved genotypes for these regions a difficult task.

48

49   Initial attempts to improve crop adaptability to different environmental conditions and so crop

50   production, has primary concentrated on the elimination in modern sorghum varieties of photoperiod

51   sensitivity character. This choice was motivated by the shortening of plant height and crop life cycle and

52   the enhancing of cultivars' general adaptation to different environments. However, the removal of the

53   photoperiodicity has led to a non-adaptation of the new varieties to West African conditions and thus to

54   their non-adoption by farmers (Summerfield et al. 1991; Vaksmann et al. 1996; Folliard et al. 2004).

55   Photoperiod sensitivity has been found to be a particular adaptation trait of West African genotypes to

56   an environment characterized by its seasonal variability and its climatic risk. It is one of the key driver of

57   sorghum cultivation that allows the synchronization of genotypes flowering time with the end of the

58   growing period (Vaksmann et al. 1996). Understanding the importance of responsiveness of Malian

59   genotypes to photoperiod has been a major advance in breeding. However, adaptation to drought

60   requires more than photoperiod sensitivity and a thorough understanding of the set of adaptive traits and

61   their interactions with varying scenarios of drought is needed to overcome the penalty in yield potential

62   in this region (Blum 2005). And first of all, the question of how much drought is an issue in the region

63 needs to be answered to better target breeding efforts, addressing especially its frequency and intensity

64 across time and geographical scales.

65 To guide breeding efforts towards the development of environment-adapted genotypes fitted to the most

66 likely weather conditions, crop modeling tools can be used to capture the highly fluctuating GxE

67 interactions (Chapman 2000b; Hall and Richards, 2013; Grassini et al., 2015). Soumare et al. (2008) carried

68 out a characterization study where they used a simple water balance model and considered that a

69 genotype was adapted to an environment when it flowered 20 days before the average date of cessation

70 of the rains. They showed that early maturing genotypes had a wider geographical adaptation but were

71 more sensitive to when the sowing start. They also showed that photoperiod sensitive genotypes were

72 adapted to specific climatic regions but tolerated a wider range of sowing dates.

73 Several other studies have been conducted in characterizing the production environment of maize,

74 sorghum, rice, and wheat in different parts of the world using the proposed approach and models such as

75 APSIM. One such study in India (Kholová et al. 2013) simulated sorghum phenology, yield and the type of

76 stress scenarios experienced by this crop across each growing season studied using available weather

77 data. They identified 5 main stress scenarios and among them, a severe stress type that occurred in 25%

78 of the studied seasons and resulted in a considerable reduction of grain yield but not in biomass yield. We

79 use this approach to estimate the production potential in Mali, to understand major stress scenarios and

80 then guide on the most suitable breeding interventions that would more likely lead to a production lift by

81 testing these GxE combinations in-silico (van Oosterom et al., 2013; Anderson et al., 2016).

82

83 This paper uses the sorghum-APSIM model to characterize drought scenarios in Mali. Specifically, we aim

84 to i) parameterize the representative genotypes of sorghum grown in the main production regions of Mali

85 into the APSIM model, by extracting coefficient characterizing the growth and development of these two

86 genotypes using data from several years of trials ii) identify the type and frequency of water deficit

87 patterns experienced by these genotypes during their life cycle across the sorghum production belt in

88 Mali, iii) evaluate the effect of each drought stress type identified on grain yield.

89

90 **Materials and Methods**

91   1) Overview:

92 Two representative genotypes grown in Mali (CSM63E and CSM335), respectively predominant in the

93 Sahelian and the Sudano-Sahelian agro-ecological zones of Mali, and differing in life cycle duration and

94 response to photoperiod (Pp) were parameterized in the improved sorghum crop model within the APSIM

95 platform (version 7.6). Crop canopy growth data and total grain and biomass yield used in this study were

96 extracted from unpublished reports (Benoit Clerget, unpublished data).

97 The sorghum production area of Mali was divided into 5 isohyets based on average annual rainfall. In each

98 of these isohyets, 1 to 7 locations (making a total of 22 locations) were selected for the characterization.

99 In each location, weather data was available for a range of 10-72 years, giving a total number of 1059

100 year-location combinations. The APSIM crop model was used to perform simulation for each of these 1059

101 combinations with 2 representative soil types for two genotypes (CSM63E and CSM335). For these

102 simulations, the model was set to decide on the optimal sowing time within a window of time between

103 1st June and 15th July, which is representative of farming decisions for sowing sorghum. The criteria for

104 sowing were accumulation of 20mm rainfall over a 7-days period and a stored soil moisture of at least

105 10mm. A characterization of seasonal drought patterns experienced by CSM63E and CSM335 at each

106 location was performed using a water status index (water supply/demand ratio) output by the model,

107 which provides a daily calculation of the ratio of how much water is available to the crop to how much

108 water is needed by the crop for each location-year combination. Values of the S/D ratio equal or above 1

109 indicate the crop experiences no stress whereas values below 1 indicate some stress. A clustering analysis

110 was thereafter undertaken to group these S/D ratio profiles into distinct scenarios experienced by each

111 genotype through their life cycle. Finally, we evaluated the effect of these drought patterns on sorghum

112 grain yield.

113

114 *2) Sorghum production regions in Mali, main characteristics*

115 In Mali, sorghum is produced in the cotton-cereal production basins and belongs to the traditional cotton-

116 cereal farming systems (Kaminski et al. 2013). The climate in this area is characterized by a long dry season

117 and a rainy season whose duration increases from north to south (Lys 2010). Average annual rainfall varies

118 from an average of 200 mm in the Northern part of the production belt to more than 1100 mm in the

119 south. Generally, the rainy season stretches from June to October with the peak rains falling around July-

120 August. The average annual maximum and minimum temperature varies from 30.6°C to 37°C and from

121 18.6°C to 22.7°C respectively.

122 The long term daily weather (rainfall, maximum and minimum temperatures, solar radiation) records for

123 22 stations selected across 3 agro-ecological zones within the country, were used. The agro-ecological

124 zones were divided into 5 isohyets on annual average rainfall-basis (Table 1). This daily weather data were

125 obtained from Mali-Meteo. In the case only rainfall data was collected, the other parameters were

126 generated from the neighboring using` NewLocclim (Grieser et al. 2006) data. The number of observed

127 years during the period 1935-2015, varied across sites from 10 to 72 seasons (table 1). These sites were

128 selected first to span a North-South gradient of rainfall conditions covering the whole sorghum production

129 area in Mali, on the basis of consistent data availability at potential locations.

130 The soil data of "Projet Inventaire des Ressources Terrestres au Mali" (PIRT, 1986) combined with Wise

131 soil database (Batjes, 1995, 2002) were used in this study. The major soil types on which each site is

132 located were determined and reclassified into 2 main soil types based on the level of soil water holding

133 capacity (SWHC) (Loffler et al, 2005; Chapman et al, 2000), the lower SWHC (62mm) being for the

134 Plinthaquic Kandustalf soil and the higher SWHC (156mm) being for the Alfisols soil. These 2 soils were

135 mainly characterized by a low organic carbon content (0.2% for the lower SWHC and 0.37 for the higher

136 SWHC) with a soil depth of 101cm and 120cm respectively for the lower and higher SWHC.

137

138 *3) Crop information*

139 Comprehensive and reliable phenological data sets, handy for crop models parameterization in Mali, are

140 rare and sometimes difficult to access. The crop data used in this study were obtained from thoroughly

141 conducted experiments by the ICRISAT sorghum program (data not published) over consecutive years at

142 ICRISAT, Samanko (12°32N, 8°04W) and at the research center of Cinzana (13°15N, 5°57W) in Mali. The

143 complete description of the experiments is available in reports provided by ICRISAT sorghum program

144 (2000-2008; available upon request and loaded on the Dataverse database at ICRISAT). Data such as final

145 grain (kg/ha) and total biomass yield (kg/ha), planting date, growth stages dates and number of green

146 leaves for 8 years and for about 20 genotypes covering the period of 2000-03 to 2005-08 were compiled

147 in a structured excel file and used for the model parameterization.

148 This study focused on two (CSM63E and CSM335) of the overall genotypes evaluated by the ICRISAT

149 sorghum program. These genotypes are guinea type landrace and originated from Mali. They were

150 selected in this study because of the differences in their phenological characteristics, their contrasting

151 response to photoperiod (Pp) in addition to their adaptation to different agro-ecological zones. CSM63E

152 named "Jacumbe" by farmers, (meaning "drought avoidance" due to its short life cycle), is an early

153 maturing genotype less sensitive to Pp, identified to be well adapted to the sahelian region. CSM335 also

154 called "Tieblen" (meaning "red man" due to the redness of its glumes) is an average maturing genotype,

155 averagely photosensitive known to perform well in the soudanian-sahelian region.

156 Data on about ten agronomic trials, extracted from the structured excel file mentioned above, were used

157 for CSM335 and CSM63E parameterization in APSIM - sorghum model. A separated dataset collected

158 from field experiment conducted in ICRISAT, Samanko, Mali in 2013 was used for model evaluation. The

159 experiment was laid out as a randomized complete block design with 4 replications. The experiment was

160 sown on June 14, July 9 and August 5 at 67, 000 hills/ha. Details on the overall experiment are given by

161 (Akinseye et al. 2017).

162

163 *4) Crop model description*

164 APSIM is a crop growth and development model that combines, individual module (crops, soils, N,

165 carbone) to simulate biophysical processes in a cropping system context. The Sorghum module that was

166 developed in the APSIM cropping system simulation model operates on a daily time step.

167 Phenology in Sorghum module is calculated using a thermal time approach through developmental stages

168 (emergence to flowering phase, flowering to maturity phase). The duration of each stage is estimated by

169 cumulating the daily thermal time from the beginning to the end of each of those stages. The thermal

170 time of each day was calculated from a broken linear function of temperature (Hammer et al., 1993):

171 $Dtt = T - T_b$ if $T_b < T < T_{opt}$ (1)   (1)

172 $Dtt = (T_{opt} - T_b)[1 - (T - T_{opt})/(T_{max} - T_{opt})]$ if $T_{opt} < T < T_{max}$ (2)

173 Were Tb is the base temperature, Topt the optimum temperature and Tmax the maximum temperature.

174 Values for these cardinal temperatures were the default values existing in APSIM sorghum-model: Tb

175 (11°C before flowering and 5°C from flowering to maturity), Topt (30°C before flowering and 22.5°C from

176 flowering to maturity) and Tmax (42°C).

177 The duration from flowering to maturity has been shown to be relatively constant in many cereals crops

178 including sorghum. The time from emergence to flowering is influenced by two main factors: temperature

179 and photoperiod and can be divided into two phases: emergence to end of juvenile phase and end of

180 juvenile to panicle initiation (end-juv to PI) phase. The duration of end-juv to PI is controlled by the

181 genotype sensitivity to photoperiod and depended on the positive deviation of daylength from a critical

182 daylength (Ppcrit1) value (Abamu et al, 2001). Photoperiod sensitivity is then defined as the derivative of

183 the function relating time to flowering to daylength in non-optimal photoperiod conditions and will have

184 units of days delay per hour increase of daylength. Photoperiod (Pp) was computed by APSIM-sorghum

185 module using the latitude, day of the year, and twilight time, i.e. when the sun is 2.2° below the horizon

186 (Keating et al. 2003), which was used to calculate the daily photoperiod.

187 Crop response to photoperiod was estimated from a triple broken-linear response function which defines

188 the response to photoperiod in terms of a minimum (Ppcrit1) photoperiod and maximum (Ppcrit2)

189 photoperiod, above and below which, respectively increase in photoperiod delays flowering and the slope

190 between these values defines plant sensitivity to photoperiod (Ppslope) (Roberts and Summerfield 1987;

191 Hammer et al. 1989, 1993 cited by Crauford et al, 1999). The default PPcrit1&2, existing in the model

192 within which sorghum reacts to Pp (11.5 and 13.5) were used in this study (Hammer and Muchow, 1991).

193 The Ppslope as well as the thermal time from end-juv to PI (tt_end-juv-ini) phase were derived using a

194 linear relationship between time from end of juvenile phase to panicle initiation (end-juv to PI) and the

195 average Pp for that period was derived (Van Oosterom et al. 2001):

196 (tt_end-juv-ini) = Ppcrit1 x Ppslope - a   (3)

197

198 Where Ppslope is the slope of the regression and a the intercept of the regression.

199 Leaf area index (LAI) was computed as the product of daily increase in individual plant leaf area and plant

200 density and was manually adjusted as to fit the simulated values obtained with the observed.

201 Commonly APSIM-sorghum model simulates the leaf area development using the total plant leaf area

202 (TPLA) approach, but in this study we used an improved version (not yet released) of APSIM-sorghum that

203 applies an individual leaf area approach that functions as follows.

204 In the first place, it calculates the relationship between the position and the area of the largest leaf using

205 the function:

206        $Y_0 = MaxLNo*aMaxSlope + aMaxIntercept$ (4)

207 Where Y_0 is the area of the largest leaf, aMaxSlope is the largest leaf area factor and aMaxIntercept the

208 intercept for largest leaf calculation.

209 Then based on that relation, it uses a bell-shaped function to relate the size of each leaf to its ordinal

210 position on the main stem using the equation:

211        $Y = Y0*exp (a*(X-X0)^2 + b*(X-X0)^3)$ (5)

212 Where X_0 is the position of the largest leaf, Y_0 the area of the largest leaf, X the position of each individual

213 leaf, 'a' is an empirical constant determining the breadth of the bell-shaped curve and 'b' is an empirical

214 constant determining the skewness of the bell-shaped curve. Each of the parameters X_0, Y_0, a, and b used

215 for ILA parameterization are function of total leaf number.

216 $X0 = aX0*MaxLNo$;   (6)

217 a=a0-exp(a1*MaxLNo), (Birch 1998)   (7)

218 b=b0-exp(b1*MaxLNo), (Birch 1998)   (8)

219 empirical constants: a0=-0.009; a1=-0.2; b0=0.0006; b1=-0.43 (Birch et al 1998)

220 Data on grain size and weight were obtained for CSM63E from an Indian genotypes already parameterized

221 in APSIM for which observed detailed data were available. For CSM335, the standard values of the model

222 were used.

223   *5) Genotype parameterization and evaluation*

224 Data on the ten agronomic trials were used for genotypes parameterization. After informing the model

225 with the needed available mentioned above, it was run and checked whether the results were in

226 agreement with the observed data. The approach used to estimate model parameters that differ between

227 genotypes and to which the model is sensitive was to derive genotype-specific parameter estimations

228 when experimentally measured data were available followed by an automated adjustment of parameters

229 that varied between genotypes. Observed variables were regressed against variables simulated by the

230 model. The parameterization process was divided into 2 steps: phenology and photoperiod

231 parameterization, and canopy size and development parameterization. The grain parameters (grain

232 number and size) were used from previous studies with Indian genotypes with similar characteristics to

233 ours in terms of duration. Equation (1), (2) and (3) were used to estimate the phenology, photoperiod

234 sensitivity and canopy parameters respectively.

235 The evaluation of the model consisted in drawing a scatter plot of simulated phenology, grain and biomass

236 yield, leaf number and LAI output by the model versus the observed phenology, yield, leaf number and

237 LAI measured in the trials. For the parameterization as well as for the evaluation of the model, a 1:1 line

238 was added to the scatter plot to test the robustness of the model. If the predictions were closed to the

239 measurements, all the data would lie next to the line otherwise they would be scattered around the line

240 (Soltani et Sinclair 2012). A regression line was also included in the plot along with the R^2 in order to

241 determine the correlation between simulated versus measured results. The root mean square of error

242 was used to evaluate the goodness of fit between observed and predicted data.

243

244   *6) characterization of stress patterns*

245 Once the parameters for the two genotypes were obtained, the APSIM-sorghum model was run to

246 characterize the 22 sites selected for seasonal patterns of water deficit analysis using an approach

247 reported earlier (Muchow et al, 1996; Chapman et al, 2000; Chenu et al, 2011; Kholova et al, 2013). In this

248 study, simulations were performed for two soils (low and high level of soil water holding capacity) for all

249 locations and years and for CSM63E and CSM335 genotypes to set up a baseline of stress scenario.

250 To represent the real farming conditions in the study area, the soil moisture in the model was re-set to

251 20% of the fractional available soil water (Chenu et al. 2013; Chauhan et al. 2014) at the beginning of each

252 growing season, representing an almost dry soil. This initial condition was reset each year of the simulation

253 in order to homogenize the yearly onset conditions occurring in each location. Planting was estimated to

254 occur within a constant sowing window, 1 June – 15 July, for all locations during each season observed.

255 Nevertheless, this planting was assumed to take place if an accumulative rainfall of 20mm were received

256 within 7days and the stored soil moisture was at least 10mm. The latter was set to 10mm in order to

257 mimic the "empty soil profile" in these areas at the beginning of the rainy season after a long dry period.

258 The sowing was set to start the last day of the window even though the defined conditions were not met.

259 All other parameters including plant population (5.7plants m^{-2}) and row spacing (86 cm) were held

260 constant throughout the simulations. The plant population and row spacing values were obtained from

261 data collected from farmer's fields (unpublished data). Also the fertilizer was applied as per the

262 recommendation (100 kg ha^{-1} DAP at sowing, 50 kg ha^{-1} Urea 45 days after sowing) and was the same for

263 all simulations.

264

265 *7) Identification of the type and frequency of water deficit patterns*

266 A stress index was used to identify seasonal drought stress patterns in the study areas. This index,

267 computed as the ratio of canopy water demand to root supply capacity (Chenu et al. 2011; Kholová et al.

268 2013) was output by the model on a daily basis along with crop growth and development, grain yield and

269 thermal time. It was averaged for each 100°Cd from 400°Cd before flowering to 400°Cd after flowering

270 over different growing seasons. The main drought stress patterns were then determined by clustering the

271 seasonal simulated patterns using partitioning clustering K-means (Partitioning Around Means) and PAM

272 (Partitioning Around Medoids) in R software (R Development Core team, 2008).

273 In order to enable the comparison of similar stress patterns experienced by different genotypes within

274 and between locations, CSM335 and CSM63E simulations clustering were first combined and then the

275 results of the clustering were disaggregated for individual genotype to study the potential differences

276 between the 2 genotypes. Additional information regarding this approach can be found in Chapman

277 (2000); Chenu et al. (2011); Kholová et al. (2013).

278 The clustering of each year-site combination was performed using unbalanced data (different number of

279 years observed from site to site) of historical weather data. This can lead to a bias estimation of the

280   frequency of occurrence of the different environment types. Therefore, the occurrence of each stress

281   patterns types was quantified based on weighted frequency correlated with the number of seasons

282   observed in the entire production belt. The following equation was used for the weighting frequency

283   calculation:

284   $x = \frac{\sum_{i=1}^{N} wi(xi - \mu*)2}{V1-1}$ (4)

285

286   *Grain yield evaluation*

287   The effect of the main stress patterns on grain yield was evaluated per agro-ecological zones/isohyets and

288   across seasons.

289

290   **Results**

291   *1.   Genotypes parameterization*

292   The comparison of model output versus the observed variables (Fig.1A) showed that the model was able

293   to simulate flowering time as a response to photoperiod for the 2 genotypes with a R^2 ranging from 0.7

294   for CSM63E to 0.9 for CSM335 with the RMSE for both genotypes less than 10% of the overall flowering

295   estimated time. The parameterized thermal time accumulated from the end of juvenile phase to panicle

296   initiation varied from 166°Cd for CSM63E to 116 °Cd for CSM335. The linear relationship between average

297   Pp from end-juv to PI and the accumulated thermal time of that period showed Ppslope ranging from 31

298   for CSM63E to 198 for CSM335, i.e. CSM63E was less photoperiod-sensitive than CSM335.

299   The duration from flowering to maturity was also well captured (R2=.7 for CSM63E and .9 for CSM335)

300   with an accumulated thermal time for that period ranged from 434°Cd to 518 °Cd for CSM63E and CSM335

301   respectively.

302   Fig.1B illustrated the goodness of fit for the total leaf number and the LAI, across a range of years and

303   environmental conditions with R^2 >0.90 in both cases. The close agreement between fitted, simulated and

304   observed leaf number and LAI dynamics indicated that individual leaf area approach used to estimate the

305   leaf area development was effective.

306   In general, the model simulated the trend of total leaf area and biomass yield (Fig1.C) reasonably well.

307   But for biomass yield, the model performed well in a range of field measurements and demonstrated

308   some well below or above the 1:1 line respectively for CSM63E and CSM335. Although this reflected the

309   prediction of total leaf number particularly with CSM335, the coefficient of determination (R^2) in all cases

310   exceeded 0.90. The RMSE for CSM63E and CSM335 biomass are 2303 and 1959 respectively.

311   As a whole, the distribution of the grain yield data around the 1:1 line indicated that there was an

312   agreement between observed and simulated yield (Fig1.D). The model predicted the grain yield of

313   CSM63E and CSM335 with an RSME of 646 and 858 respectively.

314

315   *2. Genotypes evaluation*

316   The sorghum module in APSIM was evaluated for its ability to closely reproduce general trend of growth

317   duration, total biomass and grain yield of CSM63E and CSM335 that matched the observed. Fig.2 showed

318   an agreement between predicted and observed days from sowing to flowering and maturity (Fig.2A), leaf

319   number (Fig2B), leaf area index (Fig2.C) and grain and biomass yield. In general, the model simulated the

320   days to flowering for both genotypes with an R2 above 0.85. The model prediction of total leaf number

321   and LAI was good with an R2 of 0.78 and 0.71 for CSM63E and 0.88 and 0.64 respectively for CSM335. A

322   good correlation was also observed between simulated and observed biomass yield this in spite the low

323   number of field measurements data. The negative correlation observed with the fitting of the grain yield

324   data may be due to the low observed values reported particularly for CSM63E.

325

326   *3. Water deficit patterns*

327

328   *a. Stress patterns and frequency occurrence*

329

330   Simulations and thermal time computation at every 100°Cd from (to) 400°Cd before (and after) flowering

331   revealed all water deficit scenarios simulated from each season (Fig.3).

332   The clustering of these water deficit scenarios across all locations and years resulted in 3 major and

333   distinct water deficit scenarios for CSM63E and CSM335 (Fig.4:A&B). The first cluster (Water supply and

334   demand ratio: WSD 1) was denominated as a "no-stress" scenario for both genotypes and could be

335   described as a condition where plants mostly did not experienced water stress. The second cluster (WDS

336   2) was called an "early pre-flowering terminal stress" and represented conditions where water stress

337   started around -300°Cd before flowering for CSM63E and from the beginning of the cycle for CSM335,

338   without any recovery till the end of the season. The third cluster (WDS 3) was defined as "flowering

339 terminal stress" scenario for both genotypes and stress began around 0 °Cd which corresponded to
340 flowering time and continued until physiological maturity.
341 For both genotypes CSM63E and CSM335, the dominant stress pattern was WSD1 with a frequency of
342 occurrence of 83% and 73% respectively. The less dominant stress that occurred with both genotypes was
343 WSD3, with a frequency of 8% and 9%, followed by WSD2 with a frequency of 9% and 19%, for CSM63E
344 and CSM335 respectively (Fig.4).

345

346   *b. Stress patterns occurrence across agro-ecological zones*

347

348 Among the three stress scenarios, only WSD1 was observed in all five agro-ecological zones. All scenarios
349 occurred in the northern region characterized by a low rainfall combined with poor sandy cropping soils.
350 In that region, the frequency of occurrence of WSD3 was higher for CMS63E compared to CSM335 with
351 32% and 20% of occurrence, respectively. There was 75% of chance to have WSD2 stress in that same
352 region when growing CSM335. For both genotypes, WSD3 did not occur in the far southern region where
353 the annual average rainfall was higher (Fig. 5), and WSD2 was not observed for CSM335 and marginally
354 present for CSM63E (0.4%). With CSM63E, WSD2 was not observed in the central regions considered as
355 the main sorghum production belt (600-1000mm) (Fig.6).
356 Figure 6 shows the trajectory, in the 5 agro-ecological zones, of (A) WSD1, (B) WSD2 and (C) WSD3 for
357 CSM335 and CSM63E. The response of the different genotypes greatly varied from one isohyet to another.
358 However, for both genotypes, the three stress scenarios were more severe in the <400 isohyet compared
359 to others. For instance CSM63E suffered from mid-season stress under WSD1 with a recovery at 200°Cd
360 after flowering. Conversely, CSM335 experienced a late season stress under the same WSD1 in the same
361 region. In the 600-800 isohyet and with CSM63E, WSD3 tended to be a post-flowering stress as it started
362 at 200°C after flowering. Similarly, with CSM335 genotype, WSD3 started at 300°Cd after flowering in the
363 800-1000 region and didn't recover until the end of the observed period. In that same region and with
364 CSM335, the WSD2 started far later at -200°C before flowering compared to the other regions where it
365 was observed. With CSM63E, WSD3 occurred more precociously in the 800-1000 region compared to the
366 other areas and slightly recovered from flowering time to the end of the season.

367

368   *c. Stress patterns occurrence across seasons, sites and soils*

369　The two soil types that were used in the simulation represented a range of possible agricultural soils in

370　Mali and had large differences in the soil water holding capacity.

371　Fig.7 showed that the 39% larger soil water holding capacity in the lower SWHC resulted in the WSD2

372　scenario disappearing in Kenieba for CSM63E, in Samanko, Segou, Nioro and Mopti for CSM335. Higher

373　water holding capacity in the alfisol also resulted in the WSD3 scenario disappearing in Mopti and Nioro

374　with CSM63E. It also caused a reduction of the frequency of occurrence of WSD2 in the <400 isohyet for

375　the benefit of WSD3 that increased from 19% to 45% for CSM63E and from 3% to 38% for CSM335.

376

377　　　*4.　Grain yield*
378　　　*a.　Grain yield variability across stress scenarios*

379　The difference in average grain yield was mostly significant between genotypes. The highest grain yield

380　was obtained from far with CSM63E and across all scenarios. For both genotypes, the WSD 1 showed the

381　highest average yield: 1229 kgha^{-1} and 806 kgha^{-1} for CSM63E and CSM335 respectively, (Fig 8). WSD3

382　showed the second highest yield for CSM63E (1125 kgha^{-1}) and for CSM335 (654 kgha^{-1}). The lowest

383　average yield for CSM63E (520 kgha^{-1}) and CSM335 (290 kgha^{-1}) was observed under WSD2. CSM335 has

384　a wider variability for WSD2 than for WSD3, while CSM63E shows the opposite behavior (Fig.8).

385

386　*B. grain yield across agro-ecological zones*

387　Water deficit effect on yield was observed across all isohyets. For CSM63E and CSM335, the highest yield

388　was observed under no-stress in all isohyets except in 600-800mm isohyet where CSM63E performed well

389　under WSD3. The lowest average yield was observed under early pre-flowering terminal stress in the far

390　northern region for CSM63E (430kgha) and in the southern isohyet 800-1000mm for CSM335 (113 kgha^{-}

391　^{1}) (Fig.9). Even though CSM335 was not exposed to any stress type in the far southern isohyet, probably

392　due to an excess in rainfall, the highest yield was not observed in that region. Overall, results showed that

393　drought was not the major limiting factor to sorghum production in this region. In addition, it showed that

394　overall the yield of CSM63E were above those of CSM335 across isohyets and environmental scenarios.

395　　　*b.　Grain yield variability across sites and soils*

396　The 39% higher soil water holding capacity of the Plinthaquic Kandustalf soil resulted in a yield advantage

397　of about 50% (Fig.7) in Cinzana, Samanko and Kenieba while it caused a yield loss of 29% in Gao with

398　CSM335. Comparatively, the positive effect of a higher SWHC on the yield of CSM63E was more marked

399 in southern regions than in northern areas. For instance in Kenieba, about 43% yield increase was
400 observed with CSM63E grown on higher soil water holding capacity. The location mean yield variability
401 was higher for CSM63E compared to CSM335 even though the highest yield were observed with that
402 genotype. And the average yield of CSM335 tended to be more stable on the higher SWHC.

404 **Discussion**

406 - *Identified drought patterns, their frequency and their occurrence across agro-ecological zones*
407 *in Mali*
408 The characterization results revealed three major stress scenarios with the no stress dominating with 83%
409 and 73% for CSM63E and CSM335 respectively. This point out that drought is not the major factor limiting
410 sorghum production particularly in the Southern region where the frequency of occurrence of WSD2 and
411 WSD3 were very low (WSD2 and WSD3: 3%, 1 to 2% and 1 to 9%, 8 to 9% for CSM63E and CSM335). It
412 also confirm the drought resistant character of CSM63E previously reported by Smale et al (2016). The
413 10% difference observed between the frequencies of both genotypes can be confer to the relatively longer
414 duration of CMS335 compare to CSM63E. WSD3 was shown to be less dominant for both genotypes in all
415 year-locations combinations but its frequency was higher for CSM63E in the northern region compare to
416 CSM335. This means that the avoidance of stress reported for CSM63E might be only due to its faster
417 maturity but not its physiological ability to avoid or escape stress. Nevertheless, CSM335 was revealed to
418 be more exposed to WSD2 in that same region with 75% of chance most of which is observed on the lower
419 SWHC. Chapman et al (2000b) reported in a characterization study conducted in Australia that the low
420 unreliable rainfall associated with the lower SWHC can result to a higher frequency of occurrence of
421 severe terminal stress. Both genotypes when sown at the same moment in the same site, the reproductive
422 stage of CSM63E will coincide with the vegetative stage of CSM335; CSM63E being a fast-growing, non-
423 photosensitive genotype will extract and consume the soil available moisture and thus avoid WSD2 and
424 most likely mature before WSD3 become more severe that is in the case WSD3 occurred later in the
425 season. In the case WSD3 is early, CSM63E will be affected as the extracted moisture will be consumed.
426 In contrast CSM335, that is averagely photosensitive will grow comparatively slowly, extract and consume
427 probably less water during vegetative stage and end up being exposed to WSD2 when that occur early in
428 its growth cycle. And in the case it escaped, part of the extracted moisture will be consume during
429 reproductive stage which will allow it to outstand the at-flowering stress. Chapman et al (2008) observed

430 a late maturing hypothetical genotype to perform well in regions dominated by early pre-flowering stress

431 in Australia with water conservative traits such as transpiration efficiency.

432 Our results also delimited an optimal performance region of CSM63E between 400-800mm, which was

433 found to be slightly different from its adaptation zone previously defined by Soumare et al (2004) that

434 stated that CSM63E could perform well only in isohyets below 700mm.

435 In any case, the fact that drought was not a major issue across Mali for sorghum production suggest that

436 the requirement for photoperiod sensitive material may not have been driven by a "drought-related"

437 need, but rather by the need to have crops maturing at a similar time to avoid bird attacks if the crops

438 matures too early and ahead of photoperiod sensitive lines, or to avoid the potential effect of nursing

439 midge pest that would affect later maturing crops. In our case, further physiological evaluation are needed

440 in order to assess the response of contrasted genotypes response to photoperiod under the 3 stress

441 scenarios.

442   - ***Grain yield variability across stress scenarios, ecological zones and Soil type***

443

444 Results reported here showed that yield of CSM63E were above those of CSM335 across all zones and

445 always higher under WSD1 than WSD2&3 except in isohyet 600-800mm where CSM63E performed well

446 under WSD3. This could be related to the late occurrence of that stress type (200°Cd after flowering) in

447 this region. Seyoum et al (2017) found similar results with maize in Eastern and Southern Africa and they

448 explained that the post-flowering stress while occurring after a good flowering and silking

449 synchronization, would have a slight effect on the grain as the seed set would have been secured due to

450 the absence of stress around flowering time. In contrast, yield of CSM335 were lower compare to CSM63E

451 but least variable particularly in isohyet 400-1000mm particularly on higher SWHC. CSM335 was reported

452 by Akinseye et al, (2017) to produce more biomass but less grain yield. This could be benefic for animal

453 feeding in regions where the mobile animal husbandry is disappearing due to extensive farming. The

454 higher biomass production trait can be used in the development of dual purpose genotypes.

455 In general WSD2 affected substantially the grain yield of both genotypes. The yield reduction caused by

456 this stress type is still lower than that reported by Seyoum et al (2017) in East-South Africa on maize (80%)

457 but higher than that obtained by Heinemann et al. (2016) (36%) in a drought characterization study across

458 the rice upland in Brazil.

459 The low yield of both genotypes observed in the southern region, this, in spite of the recommended

460 fertilizer applied can be explained by the fact that first, they are out of their defined adaptation region

461   and second because a high rainfall can lead to a rapid mineralization of soil nitrogen and eventually to its

462   leaching. Further, soil pH also decreases with increasing rainfall (Clerget, 2000 unpublished data).

463   The stability of CSM335 grain yield across sites on higher SWHC can guide breeding efforts to focus more

464   on the management of soil water available for the plant during the growing season.

465   It should be noted that this study was conducted with a limited number of varieties over a limited number

466   of sites. This study has to be repeated with more set of varieties that have contrasted phenology and yield

467   characteristics over a larger scale.

468
469
470   - *Conclusion*

471   The use of this approach was important as it helps:

472   - Identifying the type and dynamic of the 3 major drought stress scenarios experienced by 2
473   representatives sorghum genotypes in Mali
474   - Delimitating the area of performance of each genotype studied
475   - Evaluating the impact of each stress type on yield across ecological zones

476   That will not only help crop breeding programs in guiding in their selection process but also agronomists
477   in the phenotyping strategies

478   This approach can be used under future climate conditions in order to help in predictions.

479   Other stresses rather than drought should be taking into account in further studies.

480

481

482 **References**

483 Akinseye, F.M. et al., 2017. Field Crops Research Assessing crop model improvements through
484   comparison of sorghum (sorghum bicolor L . moench) simulation models : A case study of West
485   African varieties. *Field Crops Research*, 201, pp.19–31. Available at:
486   http://dx.doi.org/10.1016/j.fcr.2016.10.015.

487 Blum, A., 2005. Drought resistance, water-use efficiency, and yield potential—are they compatible,
488   dissonant, or mutually exclusive? *Australian Journal of Agricultural Research*, 56, pp.1159–1168.
489   Available at: www.publish.csiro.au/journals/ajar [Accessed March 29, 2017].

490 Batjes N.H. (Ed.) 1995. A homogenized soil data file for global environmental research: a subset of FAO,
491   ISRIC and NRCS profiles (version 1.0). Working Paper and Preprint 95/10b.

492 Batjes, N.H. 2002. A homogenized soil profile data set for global and regional environmental research.
493   (WISE, version 1.1). Report 2002/01. International Soil reference and Information Centre (ISRIC),
494   Wageningen, The Netherlands.

495 Chapman, S.C., 2000. Genotype by environment interactions affecting grain sorghum. II. Frequencies of
496   different seasonal patterns of drought stress are related to location effects on hybrid yields.
497   *Australian Journal of Agricultural Research*, 51, pp.209–21.

498 Chauhan Y.S., R.R.C.N., 2014. Defining agro-ecological regions for field crops in variable target
499   production environments: A case study on mungbean in the northern grains region of Australia.
500   *Agricultural and Forest Meteorology*, 194, pp.207–217. Available at: http://ac.els-
501   cdn.com/S0168192314001075/1-s2.0-S0168192314001075-main.pdf?_tid=9f727f92-67c7-11e7-
502   bf7e-00000aab0f6c&acdnat=1499949455_3a0c7337cfb13e4a8694e0b84c48ce5f [Accessed July 13,
503   2017].

504 Chenu, K. et al., 2011. Environment characterization as an aid to wheat improvement: Interpreting
505   genotype-environment interactions by modelling water-deficit patterns in North-Eastern Australia.
506   *Journal of Experimental Botany*, 62(6), pp.1743–1755.

507 Chenu, K., Deihimfard, R. & Chapman, S.C., 2013. Large-scale characterization of drought pattern : a
508   continent- wide modelling approach applied to the Australian wheatbelt – spatial and temporal
509   trends. *New Phytologist*, 198, pp.801–820.

510 Folliard, A. et al., 2004. Modeling of sorghum response to photoperiod: A threshold-hyperbolic
511   approach. *Field Crops Research*, 89(1), pp.59–70.

512 Grieser, J., Gommes, R., Bernardi, M., 2006. New LocClim -the Local Climate Estimator of FAO.
513   *Geophysical Research Abstracts*, 8(08305). Available at: http://www.juergen-
514   grieser.de/publications/publications_pdf/EGU06-J-08305.pdf [Accessed September 7, 2017].

515 Kaminski, J., Elbehri, A. & Samake, M., 2013. An assessment of sorghum and millet in Mali and
516   implications for competitive and inclusive value chains*. *Rebuilding West Africa's Food Potential,
517   A. Elbehri (ed.), FAO/IFAD*. Available at: http://www.fao.org/3/a-i3222e/i3222e15.pdf [Accessed
518   September 7, 2017].

519 Keating, B. et al., 2003. An overview of APSIM, a model designed for farming systems simulation. *Europ.
520   J. Agronomy*, 18, pp.309–332.

521 Kholová, J. et al., 2013. Drought stress characterization of post-rainy season (rabi) sorghum in India.

Field Crops Research, 141, pp.38–46. Available at: http://dx.doi.org/10.1016/j.fcr.2012.10.020.

Lys, P., 2010. The Current State of Cost of Production Estimates A review of Country Practices in Ethiopia, Mali and Zambia. Available at: http://www.fao.org/fileadmin/templates/ess/ess_test_folder/Workshops_Events/production_cost /Lys_Report_on_country_case_studies_on_CoP.pdf [Accessed September 7, 2017].

Van Oosterom, E.J., Carberry, P.S. & O'Leary, G.J., 2001. Simulating growth, development, and yield of tillering pearl millet I. Leaf area profiles on main shoots and tillers. *Field Crops Research*, 72(1), pp.51–66.

Soltani, A., Sinclair, T.R., 2012. *Modeling Physiology of Crop Development, Growth and Yield*,

Soumare, M. et al., 2008. Prévision de l'aire de diffusion des sorghos photopériodiques en Afrique de l'Ouest Mamy. *Cahiers Agricultures*, 17, pp.160–164.

Summerfield, B.R.J., Roberts, E.H. & Ellis, R.H., 1991. Towards the reliable prediction of time to flowering in six annual crops. I. The development of simple models for fluctuating field environments. , 27, pp.11–31.

USAID, 2011. *Etude Nationale sur les Bassins de Production des Speculations Cerelieres au Mali*,

Vaksmann, M., Traore, S.B. & Niangado, O., 1996. Le photopériodisme des sorghos africains. *Agriculture et developpement*, (No9), pp.13–18.

541 Table 1: Description of the site chosen across the five agro-ecological zones and used in the
542   characterization analysis, providing location, range of rainfall in each site, geographical
543   coordinates, start and end of years when weather data were collected, mean rainfall observed at
544   these site during the observation period, rainfall range during the cropping season.

Agro-ecological zones	Isohyets	Sites	Long (°W)	Lat (°N)	Start	End	No. of years	Mean Rainfall (mm)	Range in growing season rainfall (mm)
	< 400	Gao	0.05	16.27	1950	1980	30	243	128-430
		Nara	7.28	15.17	1950	2003	53	426	198-648
Sahelian zone	400-600	Mopti	4.10	14.52	1950	2005	55	492	281-963
		Nioro	9.60	15.23	1950	1980	30	567	298-965
		Kayes	11.43	14.43	1950	2004	54	651	358-936
		Segou	6.15	13.40	1950	2008	58	661	391-954
		Cinzana	5.95	13.25	1961	2010	49	697	522-1006
	600-800	Tominian	4.60	13.28	1950	1990	40	715	421-977
		Bafoulabe	10.83	13.80	1971	1980	9	734	469-858
		Kolokani	8.03	13.58	1935	2006	71	753	381-1225
Soudanian zone		Ntarla	5.75	12.70	1967	1983	16	846	639-1139
		Dioila	6.80	12.48	1939	2008	69	860	429-1285
		Samanko	8.07	12.53	1968	2015	47	920	538-1366
	800-1000	Koutiala	5.47	12.40	1950	2008	58	922	516-1433
		Sotuba	7.93	12.66	1951	2015	64	942	591-1425
		Massigui	6.77	11.87	1960	1990	30	943	639-1489
		Kita	9.45	13.07	1938	2008	70	986	398-1515
		Kangaba	8.42	11.93	1940	2006	66	1040	422-1674
		Fakola	6.90	10.53	1966	1992	26	1054	639-1542
Guinean zone	> 1000	Kenieba	11.35	13.10	1950	1997	47	1159	256-1913
		Bougouni	7.50	11.42	1950	2007	57	1161	718-1661
		Sikasso	5.68	11.35	1950	2010	60	1170	754-1558

545

546

547

548

549

550   Table 2. Frequency of occurrence of the 3 stress scenarios across isohyets

Genotypes	Stress scenarios	<400	400-600	600-800	800-1000	>1000
	1	4.8	61.1	99.0	98.8	1.0
CSM335	2	74.2	18.2	0.2	0.3	
	3	21.0	20.7	0.9	1.0	
	1	25.8	90.4	99.8	99.0	99.6
CSM63E	2	41.9	3.2			0.4
	3	32.3	6.4	0.2	1.0	

551

552