ASA, CSSA, and SSSA Virtual Issue Call for Papers:

Advancing Resilient Agricultural Systems:

Adapting to and Mitigating Climate Change

Content will focus on resilience to climate change in agricultural systems, exploring the latest research investigating strategies to adapt to and mitigate climate change. Innovation and imagination backed by good science, as well as diverse voices and perspectives are encouraged. Where are we now and how can we address those challenges? Abstracts must reflect original research, reviews and analyses, datasets, or issues and perspectives related to objectives in the topics below. Authors are expected to review papers in their subject area that are submitted to this virtual issue.

Climate Change Virtual Issue

Submit Now!

Deadlines
Abstract/Proposal Deadline: Ongoing
Submission deadline: 31 Dec. 2022

How to submit
Submit your proposal to manuscripts@sciencesocieties.org

Please contact Jerry Hatfield at jerryhatfield67@gmail.com with any questions.

Topic Areas

- Emissions and Sequestration
  » Strategies for reducing greenhouse gas emissions, sequestering carbon

- Water Management
  » Evaporation, transpiration, and surface energy balance

- Cropping Systems Modeling
  » Prediction of climate change impacts
  » Physiological changes

- Soil Sustainability
  » Threats to soil sustainability (salinization, contamination, degradation, etc.)
  » Strategies for preventing erosion

- Strategies for Water and Nutrient Management
  » Improved cropping systems

- Plant and Animal Stress
  » Protecting germplasm and crop wild relatives
  » Breeding for climate adaptations
  » Increasing resilience

- Waste Management
  » Reducing or repurposing waste

- Other
  » Agroforestry
  » Perennial crops
  » Specialty crops
  » Wetlands and forest soils

American Society of Agronomy | Crop Science SOCIETY OF AMERICA | Soil Science Society of America

Optimizing the test locations and replicates in multi-environmental cotton registration trials in southern Xinjiang, China

Naiyin Xu [1,2]*, Yintao Qiao[1,2], Suqin Zhao[3], Xiaoni Yang[1], Jian Li[1], and Michel Fok[4]

[1] Institute of Industrial Crops, Jiangsu Academy of Agricultural Sciences, Nanjing 210014, China

[2] School of Agricultural Engineering, Jiangsu University, Zhenjiang, Jiangsu 212013, China

[3] Xinjiang Uygur Autonomous Region Seed Industry Development Center, Urumqi 830006, China

[4] CIRAD, UPR AIDA, F-34398, Montpellier, France

*For corresponding: naiyin@126.com

1

Abstract

Multi-environment trials are routinely conducted around the world to select superior cultivars; the numbers of replicates and locations remains insufficiently studied despite their effects on trial efficiency and cost. The objective of the present study was to compensate for the mentioned lack by dealing with the specific case of cotton in arid conditions of China and by estimating the optimal numbers of locations and/or replicates in a more comprehensive way than implemented so far, i.e. optimizing with regard jointly to three criteria, namely, genotype ranking, location evaluation and environment investigation. Trial heritability and the optimal numbers of locations and replicates were estimated through adapted genotype plus genotype by environment (GGE) biplot analysis from data of cotton variety registration trials in the fringes of the Taklimakan Desert in Southern Xinjiang in China over the 2011-2020 period when three replicate trials were generally conducted in nine locations. Despite the identification of two mega-environments (MEs) through LG (location grouping) biplot analysis, our study showed robust results of genotype ranking, correlation among test locations, and location grouping that were not affected by reducing the number of replicates. It was concluded that two replicates (instead of three) were sufficient for reliable genotype evaluation, test location evaluation and ME classification in the studied trial scheme. The implied savings justifies similar study on other schemes.

Keywords: cotton; GGE biplot; optimal replicate; location grouping (LG) biplot, mega-environment (ME) plot

2

Core ideas

Core Idea 1: Two replicates suffice for reliable genotype and test location evaluation in multi-

environment cotton variety trial

Core Idea 2: ME plot is an effective complement to LG biplot in mega-environment analysis.

Core Idea 3: The genotypic ranking remain robust or reliable despite the unique mega-environment

assumption was not confirmed.

Abbreviations: GE, genotype by environment interaction; GGE, genotypic main effect plus genotype

by environment interaction; LG, location grouping; ME, mega-environment

3

1. Introduction

Cotton (*Gossypium hirsutum* L.) is one of the extensively cultivated crops in both agricultural and textile industry countries around the world (Ijaz et al., 2019). Its cultivation has benefitted from the release of new varieties based on multi-environment variety trial schemes in which the ever presence of genotype by environment interaction is a challenge to address for proper cultivar evaluation (Yan and Hunt, 2001; Xu et al., 2014; Yan, 2016).

Consequently, multi-location trials generally have three main features or objectives, i.e., genotype rankings, location evaluation and mega-environment (ME) investigation (Yan et al., 2011). Genotype evaluation and selection are the principal target of crop variety trials, the test location evaluation and ME investigation are fundamental prerequisites for meaningful genotype evaluation that can only be properly done within a ME (Baxevanos et al., 2008; Yan, 2021a). In a multi-location trial scheme, a discriminative and representative test location is defined as a desirable test location where selected genotypes are expected to perform as excellently in the target-planting region as in the selecting location (Yan, 2021a). A target region represented by a set of test locations is usually assumed to be a single ME (Yan, 2015), which was identified as a homogenous portion of the target crop growing region for genotypes to perform similarly in all locations within a ME (Gauch and Zobel, 1997), or a cluster of test locations where the similar cultivars performed the best across years (Yan and Rajcan, 2002). When dealing with Mega-environment, the selection of cultivars is based on the average performance and ignore variety by location interaction within a ME. Genotype and test location evaluation must consider genotype main effect (G) and genotype by environment interaction effect (GE) simultaneously, while ME investigation is the first step towards meaningful genotype and test location evaluation (Yan, 2021a). Existing works have ascertain that Genotypic

4

main effect plus Genotype by Environment interaction (GGE) biplot is an effective approach in analyzing multi-location variety trial data to achieve all of the three above mentioned objectives of crop variety trials(Yan and Tinker, 2006).

Another challenge for cultivar development is that crop variety trials are laborious, resource and time consuming. The cost for growing a single field plot is usually budgeted for US$32–40 in Canada (Yan, 2021b), and around US$30 for a plot of cotton variety trials in China. For economic reasons, the improvement of crop varietal trial schemes by optimizing the number of test years, locations and replicates within trials is of great importance to minimize cost and maximize efficiency (Swallow and Wehner, 1989; Yan et al., 2015; Xu et al., 2016; Baxevanos et al., 2017a; Woyann et al., 2020; Zhang et al., 2020; Yan, 2021b). Quantitative equations to estimate the optimal number of locations and replicates could derive from three levels of variety trial data, namely, single trial at a single location(Yan et al., 2015), multi-location trials in a year(Yan et al., 2015), and multi-year multi-location trials (Yan, 2021b), but the three levels were of distinct reliability.

In practice, optimizing the number of replicates often becomes the sole option because there is very little space to adjust the numbers of test years and test locations for most official crop variety trials (Yan, 2021b). The equation for estimating the optimal number of replicates within a single trial was applied in the national cotton (Xu et al., 2016) and winter wheat (Zhang et al., 2020) variety trials in China, wheat and cotton variety trials in the Mediterranean regions (Baxevanos et al., 2017a; Baxevanos et al., 2017b), soybean variety trials in Brazil (Woyann et al., 2020), and winter wheat variety trials in California (George and Lundy, 2019). However, Yan (2021) emphasized that the optimal number of replicates derived from a single trial basis is often over-estimated because decisions on genotypes are made on multi-location trial basis and not on a single trial. Yan (2021b) developed the method to estimate the minimum number of replicates on a multi-location basis and concluded that a single replicate or two replicates would have sufficed for oat variety trials in Canada. In crop variety trials over years in a single location, the estimation approach is made difficult

as entries vary over years by yearly replacing those performing poorly (Xu et al., 2014). Consequently, the analysis of the data of multi-location variety trial over years becomes the best approach to estimate the optimal number of replicates (Yan, 2021b).

The optimal number of replicates has been estimated, but incompletely validated with regard to the three major objectives of multi-location trials, i.e. genotype ranking, location evaluation and ME classification. Yan (2021b) estimated the optimal number of replicates and validated the findings only in terms of genotypic rankings, validation against test location evaluation and ME classification are yet to be done. The objective of this study was to compensate for the mentioned lack by optimizing the numbers of locations and replicates in a more comprehensive way in the specific case of cotton in arid conditions of China, as an extension of the previous report (Yan, 2021b).

2. Materials and Methods

2.1 Dataset Source

The data of cotton lint yield analyzed in this study were from the national cotton variety trials in Southern Xinjiang from 2011 to 2020. The cotton-producing region in Southern Xinjiang is around the fringes of the Taklimakan Desert up to the south of Tianshan Mountains; it is currently the most important cotton-producing region in China, accounting for about 40% in acreage and more than 50% in cotton lint production of the country. The annual precipitation in the region is only about 25-98 mm, so cotton is grown under drip irrigation as well as under plastic mulching to adapt to a short growing season in continental arid conditions. The national cotton variety trials in the region generally consists of 9 test locations in the recent decade, namely Alaer, Shaya, Bazhou, Korla, Kuqa, Makit, Shache, Tumxuk and Tiemenguan. Makit happens to be located in the alluvial fan oasis plain of Yarkant and Tiznavu Rivers with cumulated irrigated silt soil, other locations are of the sandy loam soil type.

The trials were conducted based on randomized complete blocks designs with three replicates and the plot size was 20 m^2 at all locations, in compliance to instructions of the national cotton

registration committee. All trials were conducted under drip irrigation with plastic mulching. The average quantity of nitrogen (N), phosphorous (P) and potassium (K) fertilizer was 450 kg·ha^{-1}, 280 kg·ha^{-1} and 118 kg·ha^{-1} respectively. Fertilization levels varied from location to location, ranging from 308 to 622 kg·ha^{-1} for N, from 95 to 385 kg·ha^{-1} for P and from 68 to 181 kg ·ha^{-1} for K. Trial implementers at test determined the fertilizer rates locations according to field environmental conditions and cotton growing situations. Location Makit came out with the lowest fertilization rates, notably for P and K, at one third of those of some other locations. Other field management measures were similar to production at farmers' level. Details on the latitude, longitude, altitude, soil type, fertilizer and test year of the test locations are presented in Table 1.The number of varieties tested was from 7 to 14, depending on years (Table 2). The sets of genotypes varied each year, only those with good performance one year were tested a second time. By the end of the cropping cycle, the normally opened bolls were manually harvested in each plot to record seed cotton yield. One kilogram of seed cotton was randomly sampled from each of two replicates was picked and ginned in a roller-ginning machine to determine the lint percentage, which was used to calculate the lint yield. For each plot, cotton lint yield was calculated by seed cotton yield multiplying the lint percentage.

2.2 Statistical methods for determining the optimum number of replicates

Analysis of variance (ANOVA) was performed based on yearly datasets of multi-location cotton variety trials by using the statistical package of GGE biplot pattern explorer (Yan, 2001). Each dataset of yearly multi-environment trials (MET) was recomposed into four datasets, namely, three two-replicate combinations and the original three-replicate dataset. Hence, a serial datasets of 40 METs were subjected to ANOVA in all.

The heritability (H) based on multi-location trials in a year was calculated as (Yan et al., 2015; Baxevanos et al., 2017a; Yan, 2021b):

$$H = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_{gl}^2/N_l + \sigma_e^2/N_l N_r} \quad (1)$$

Where σ_g^2 is the genotypic variance, σ_e^2 is the experimental error variance, σ_{gl}^2 is the variance for genotype by location interaction (GL), N_l is the number of locations and N_r is the number of replicates. All are estimated on the yearly multi-location basis.

From Eq. 1, the required number of locations (N_l) in a mega-environment can be estimated by Eq.2. Let the left part in parentheses in Eq.2 represent by Noise-information quotient (Q_l) in Eq.3, and set the target H to 0.75 (Yan and Holland, 2010), then the required number of locations to achieve the target heritability of 0.75 ($N_{L, H75}$) can be estimated by Eq.4.

$$N_l = \left(\frac{\sigma_{gl}^2 + \sigma_e^2/N_r}{\sigma_g^2}\right)\frac{H}{1-H} \quad (2)$$

$$Q_l = \frac{\sigma_{gl}^2 + \sigma_e^2/N_r}{\sigma_g^2} \quad (3)$$

$$N_{l,H75} = 3Q_l \quad (4)$$

From Eq. 1 and analogous to Eq. 2, the optimal number of replicates on the yearly multi-location trial basis can be estimated by:

$$N_r = \max\left(1, \left(\frac{H}{H_{max}-H}\right)\left(\frac{\sigma_e^2}{\sigma_g^2}\right)\left(\frac{H_{max}}{N_l}\right)\right) \quad (5)$$

Where H_{max} is the maximum achievable across-location heritability and is determined by:

$$H_{max} = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_{gl}^2/N_l} \quad (6)$$

Eq 5 is a special case of Eq. 1 when assuming zero experimental error variance or with an infinite number of replicates (Yan, 2021b). The target cross-location heritability should be a certain level of H/H_{max}, rather than the actual value of H. The relative cross-location heritability is the measure for adequate replicate in the multi-location trial framework (Yan, 2021b). If the target cross-location heritability is set such that $H=0.75H_{max}$, then Eq. 5 becomes:

$$N_{r,H75} = \max\left(1, 3\left(\frac{\sigma_e^2}{\sigma_g^2}\right)\left(\frac{H_{max}}{N_l}\right)\right) \quad (7)$$

8

Where $N_{r,H75}$ is the optimal number of replicates given the number of locations N_l, and H_{max} is as defined in Eq.6. Eq.7 shows that the required number of replicates is determined by the relative magnitude of experimental error variance (σ_e^2/σ_g^2), the number of locations (N_l) and the maximum achievable across-location heritability (H_{max}).

2.3 Validation on the estimated optimum number of replicates

The validation of the optimized number of replicates was implemented through the analysis of views resulting from GGE biplots. The mean vs. instability view of GGE biplot (Yan and Tinker, 2005) was to rank the genotypes tested in 2020 based on the mean yield across locations using data from all three replicates. A summary biplot (Yan, 2021b) was used to show the similarity between genotypic rankings based on full data (three replicates, or 3-rep) versus those based on reduced data (any single replicate or any two-replicate combinations, or single or 2-rep). Pearson correlation coefficients between the vectors of genotypic means based on different number of replicates are presented for numerical comparison.

The test location discriminating ability and representativeness view of the GGE biplot (Yan, 2010; Xu and Li, 2014) was used to compare location evaluation based on full data versus reduced data. In such biplots, the vector length of locations graphically displays the location discriminating ability, the longer the vector the more discriminating the location. The line with a single arrow is the average environment axis (AEA), the cosine of a location vector and the angle between AEA indicate the representativeness of the location, the larger the angle the less representative the location (Yan, 2010). The product of the discriminating ability and the representativeness is named as the desirability index, which is a comprehensive measure of the suitability of a location as a test location for the target region (Xu and Li, 2014; Yan, 2015). A least significant difference (LSD) test was performed at 5% significance to compare among the test locations and among the different replicate combinations.

Mega-environment (ME) analysis was conducted based on Location Grouping (LG) biplot analysis (Yan, 2019; Yan et al., 2021). PC1 and PC2 in LG biplot are the first two principal components from singular value decomposition of the location by trial two-way table of correlations among locations. The trials conducted at each location are presented as a cluster of trials, with the location name placed at the center and the individual trials, indicated by the last two digits of the year, placed around it. Trials that overlapped in biplot views belong to a same ME. Mega-environment (ME) biplot is the same biplot with LG biplot, except that all trials considered as belonging to the same ME are connected to the mean coordination of those trials, which defines the placement of the ME. ME plot is proposed for the first time to show the relationship among MEs more clearly and convincingly.

It is worth noting that each of the single replicate or the combination of two replicates (2-rep combinations) represented in biplots corresponds to a random sample of all possible single or 2-rep combinations, because the replicates were nested within test locations and the replicate-labeled "1" at one location was unrelated to that at any other location. The replicate-labeled "1" in all locations formed REP1 for the sake of convenient traceability and reenactment of the analysis in this study. This is only a special sample of many possible combinations of REP1, and so on. The analyses were conducted using the GGE biplot software (Yan, 2001).

3. Results

3.1 Estimated optimal number of locations and number of replicates

The analysis of variance (ANOVA) for cotton lint yield across the 40 METs revealed a high proportion of sum of squares for locations, which accounted on average for 59.7% (ranging from 31.7% to 79.1%) of the total treatment sum of squares (SSTRMT). Both GE and G effects were significant in all METs, with average sums of squares at 24.7% and 15.7% of SSTRMT, respectively. Sum of squares for GE was larger than G in 33 out of 40 METs. The large magnitude of GE relative to G suggests possible existence of different mega-environments. Main parameters derived from ANOVA outputs were

10

listed in Table 2. Table 2 shows the ratio of experimental error variance over genotypic variance (σ_e^2/σ_g^2) and the ratio of genotype by location interaction variance over genotypic variance (σ_{gl}^2/σ_g^2) based on all replicates (3-rep) and those based on two-replicate combinations (2-rep). Both types of ratios were highly consistent in most cases, except for the combination of replicate 1 and 3 (REP13 in brief, the same below) and REP23 in 2015, when the genotypic variance was low.

Table 2 also shows that the trial heritability (H) across locations varied substantially from year to year, ranging from 0.50 to 0.91 on average. The heritability estimated on all replicates in each year was very close to the mean value of all replicate combinations, higher than 0.75 in six years, between 0.60 and 0.75 in three years and less than 0.60 in 2015. As a rare case, the heritability of REP13 and REP23 in 2015 was only 0.42 and 0.34 respectively. The mean heritability across years based on all three replicates (REP123) was 0.75, while that based on REP12 was above 0.75, and those on REP13 and REP23 were higher than 0.70.

The estimated maximum heritability (H_{max}) across locations varied depending on the year, ranging from 0.62 to 0.97 (Table 2). The overall mean of different replicate combinations was approximately 0.8 in all years except 2015. The estimated maximum heritability was lower than 0.60 for REP13 and REP23 in 2015.

The coefficient of variation (CV) of yearly multi-location trials is a measure of trial accuracy independent of genotypic variation (Yan, 2021b); our results (Table 2) showed that it was lower than 13% in all years and even lower than 10% except 2012 and 2015 (higher values in 2012 and 2015 for REP13, REP23 and REP123). The overall mean CV of any 2-rep and 3-rep across years were around 8% in all years. The mean CV based on all replicates across all locations and years were not significantly different from that based on two replicates.

The estimated optimal number of locations ($N_{l,H75}$) varied greatly from year to year (Table 2), ranging from 2 in 2016 to 17 in 2015. The overall mean estimated optimal number of locations based on all

replicates was 8, which was close to the current number of locations (9). The overall mean estimated optimal number of locations based on 2-rep combinations averaged 10.3, ranging from 7.8 to 11.8. It is worth noting that this overall mean of 3-rep would be 7 and those of 2-rep would be 8 ranging from 7.6 to 9.0 when the unusual year 2015 was excluded. These results clearly indicate that there was little room to optimize the trial scheme by adjusting the number of locations for the current trial setup, so the optimization of the number of trial locations will not be further addressed. By definition, the genotype by location interaction within a mega-environment is unrepeatable. The number of locations required for reliable selection is determined by the relative magnitude of genotypic vs. genotype by location interaction variances.

The mean estimated optimal number of replicates on the multi-location basis ($N_{r,H75}$) of 2-rep or 3-rep, referring to the basic number of one replicate on trials, was around 1 across years, and only outliers with slightly greater than 2 for REP13 and REP23 in 2015.

In summary, the ratio of experimental error variance over genotypic variance, the ratio of genotype by location interaction variance over genotypic variance, coefficient of variation, the estimated optimal number of locations and replicates estimated based on all three replicates were not significantly different from those based on any two-rep combinations (Table 2). The estimated optimal number of locations was very close to that actually in use, indicating that there was little room to increase or reduce the number of test locations. It is undoubtedly that if more locations represented environmental contrasts are added to the dataset, the effect of genotype by location interaction will expand and lead to increase the value of the optimal number of locations. However, it is assumed that the test locations are chosen such that they represent the mega-environment well. The estimated optimal number of replicates was equal to 1 in the vast majority of cases, implying that a single replicate of a basic trial replicate would be sufficient for reliable cultivar evaluation in the current multi-location trial scheme. Empirical validation of this observation will be implemented below.

3.2 Cross-location genotypic ranking based on all replicates vs. any one or two replicates

In the case of multi-location trial in 2020, relatively high trial heritability was achieved based on all replicates or any combination of two replicates (ranging from 0.75 to 0.78 for an average of 0.77). As a result, the estimated optimal number of replicates on a multi-location basis was estimated to be around 1 (Table 2), indicating that a single replicate would suffice. Therefore, it is expected that genotypic evaluation based on any two replicates would sufficiently approximate that based on all replicates in this case.

Figure 1 presents the mean vs. instability view of GGE biplot based on the yield data from all three replicates. The following patterns can be drawn from this biplot. Firstly, the red line with a single arrow is the average environment axis (AEA) with the arrow pointing to higher values of mean yield across locations. The magnitude of the difference between two genotypes can be visually assessed by the distance between them. The highest yielding genotype was Ym8, followed by Th11 and a group of other genotypes including Jfy122, Jk1975, Jk1974, H1594, Hx1691 and Ym5 whose yields were higher than the grand mean. In the opposite, Sf15038 was the lowest yielding genotype, followed by D27563, and so on. Secondly, the blue line with double arrows is the instability coordinate with arrows pointing to values of greater instability. In terms of yielding variation among test locations, genotype Th11 and Ym8 were most stable according to their distance to the AEA, followed by Jk1975, Zsmgz60 and Jk1974, and so on. In the opposite, genotype Jfy122 was most unstable, followed by Ym5, D27565 and Hx1691, and so on.

Similarly, genotypes can be ranked based on any single replicate or any 2-rep combinations and the rankings are visually summarized and compared in Figure 2. The following patterns can be drawn from this biplot. Firstly, the genotypic rankings from the seven replicate scenarios were closely correlated, as indicated by the acute angles among the vectors. The cosine of the angle between any two vectors approximates the Pearson correlation between them (Table 3). The angles between the single replicate vectors (REP1, REP2, and REP3) were obviously larger than that between the 2-rep

combinations, hence indicating that the genotype rankings based on single replicates were less reliable than that based on two replicates, a result consistent with expectation. In other words, the genotypic rankings based on 2-rep combinations were more similar, comparatively to that based on a single replicate, to that of all replicates than rankings based on single replicates (Figure 1). Secondly, all single replicates (REP1, REP2, and REP3) and 2-rep combinations (REP12, REP13, and REP23), as well as the full dataset (REP123), identified Ym8 as the highest yielding genotype, as they all fell into the Ym8 sector defined between the two radiate lines labeled "1" and "2." This is consistent to the fact that one replicate would have suffice, but two would be more secured (Yan, 2021b).

The correlation coefficients between any one or two reps on one hand and the three replicates on the other hand determines the similarity between genotypic rankings. Pearson correlation coefficients between mean lint yield based on all three replicates and genotypic means based on any one or two replicate combinations in 2011-2020 were listed in Table 3. The dataset in 2020 was chosen to show it graphically in Figures 1 and 2, as a representativeness of the pattern in most years, because it is the latest dataset. Most of the other years studied are consistent with the results from the 2020 data, except that for 2015. Table 3 indicated that the mean correlation coefficients between single replicates and the full data were around 0.96[**], while that for 2-rep combinations were around 0.99[**].

3.3 Test location evaluation based on all replicates vs. single or two replicates

Figure 3 presents the test location discriminating ability and representativeness biplot based on all three replicates. The following can be seen from it. Firstly, the vectors of the test locations Makit and Korla are apparently shorter than other locations, indicating that their discriminating ability was relatively weak. The vector length of location Tumxuk was the longest, while other location vectors are of the similar length. Secondly, the vectors Shache and Tiemenguan are very close to AEA axis, showing their strong representativeness. The test locations Makit and Korla are in near-right angle

14

with AEA axis, indicating their poor representativeness. Thirdly, Makit and Korla therefore come out as undesirable test locations if the whole region is considered as a single mega-environment. Note that Makit and Korla clearly deviate from the majority cluster of trial locations, the angle between the two locations and the cluster of other locations is visibly obtuse, suggesting that the locations may belong to distinct mega-environments.

Similar analysis can be implemented for six GGE biplots of discriminating ability and representativeness view based on any single replicate and any 2-rep combinations (Figure 4). It can be seen that the patterns observed from Figure 3 remained largely true in each of the six biplots in Figure 4. However, as expected, the biplots based on 2-rep combinations are more consistent with Figure 3 in comparison to those based on single replicate only. This indicates that any single replicate was able to reveal the main patterns of test locations, but two replicates were better.

For further elucidation, the numerical values of discrimination ability, representativeness and desirability index of locations in each biplots in Figures 3 and 4 are presented in a two-way table of location by replicate combination for each parameter. The three two-way tables were then subjected to a two-factor analysis of variance (ANOVA) and multiple comparison. The means of the discrimination ability, the representativeness, and the desirability index for each location and each replicate scenario are listed in Table 4.

The following can be drawn from Table 4. Firstly, the discrimination ability (i.e., the location vector length in Figure 3 and 4) of Makit and Korla were significantly lower than that of other locations, while that of Tumxuk was significantly larger than that of other locations. The representativeness of Makit and Korla were significantly weaker than that of other locations, while Shache and Tiemenguan were significantly more representative than other locations. With regard to desirability index, it was clear that Makit and Korla were the poorest test locations, followed by location Kuqa, while other locations were similar (See upper part of Table 4). (2) Multiple comparison among different replicate scenarios showed that the mean location vector length based on all three

replicates was not significantly different from that of any single replicate except REP3. Moreover, there was no significant difference between various replicate scenarios in representativeness or desirability. In short, the patterns and the characteristics of the test locations based on all replicates (Figure 3) largely held when based on a single replicate or two replicates (Figure 4).

3.4 Mega-environment investigation based on all replicates vs. any one or two replicates

Figure 5 assembles two representations. Figure 5A is the location by trial LG biplot based on the full three replicate dataset as above where the locations appear in blue and the trials in red. Figure 5B is named as mega-environment (ME) plot, it is a variant of Figure 5A in which all trials considered belonging to the same ME are connected to the average coordinate of these trials so that distinct MEs are more clearly shown as well as the trial ranges within the MEs. The most important pattern coming out Figure 5A is that trials at the location Makit constituted an independent ME, while other locations stayed together to form the other and major ME, as it can be seen more clearly in Figure 5B. Apparently, the angle between the two MEs is close to be a right angle, indicating that results obtained at location Makit was unrelated to that in the major ME.

Figure 6 assembles the six ME plots based on any single replicate or two replicate combinations destined to check if the two MEs revealed from data on all replicates could be found by using less replicates. Indeed, the patterns shown in Figure 5B appear in most of the ME plots in Figure 6 except for REP3, where the two MEs appear to overlap.

4. Discussion

Our study was to follow the method of Yan (2021b) to proceed a more comprehensive validation of the optimized number of replicates in the specific case of cotton in arid conditions of China, by considering test location evaluation and ME analysis in addition to genotypic ranking. First, our results showed that there is little room to alter the number of locations to optimize the current multi-location trial scheme in Southern Xinjiang as the estimated optimal number of locations was

16

around 8, which was very close to that in use. Eq. 2 suggests that increasing the number of test locations would be effective to improve the cross-location heritability, hence selection accuracy. Empirically, when heritability is less than 0.75, the effect of increasing the test location number is significant judging by the curvilinear relationship between location numbers and heritability (Yan et al, 2015). Our results are consistent with those of existing works. In the cases of the 2015–2019 Ontario oat trials(Yan, 2021b) and the 2013–2019 Quebec provincial oat trials(Yan, 2021a), the estimated number of optimal locations was also close to the number in use as the cross-location heritability was overall around 0.75 in most years. Using the similar method, Baxevanos et al. (2017a) estimated an optimal number of locations (3) that was much lower than the actual locations used (11) for a high trial heritability (0.95) in the Greek cotton ME. The empirical optimal heritability (H, 0.75) is in fact inevitably modified by the noise-information quotient (Q_l, in Eq.3), the actual point of inflexion on the curve may higher than 0.75 for a larger Q_l. In theory, trial heritability and noise-information quotient are the key factors to determine the feasibility of altering location number in optimizing crop trial scheme. It is worth noting how fortunate the lack of relevance to alter the number of test locations: for most mature crop variety trials, the locations are normally fixed for a while as Yan (2021b) noticed for Canada while in China. The trial schemes of cotton or other crops have seldom changed since at least the last decade. However, in general, the extent of the option or relevance to alter the number of locations depend on the nature of variety trial scheme in different regions and countries with the associated trial heritability levels.

Second, our results indicate that a single replicate of a basic trial replicate would be sufficient for reliable cultivar evaluation in the current multi-location trial scheme. The estimated optimal number of replicates on the multi-location basis ($N_{r,H75}$) using any 2-rep or 3-rep datasets was indeed equal to 1 in almost all years. Our results are consistent with that of the study of the Ontario oat trials over the 2015-2019 period where the optimal number of replicates on the multi-location basis was also around 1.0. For non-Ontario trials, the optimal number of replicates averaged around 1.8(Yan, 2021b). Zhang et al. (2020) reported the optimal number of replicates for winter wheat trials in

Northern China was around 1.4. It is worth emphasizing that the required number of replicates is determined by the relative magnitude of experimental error variance, the number of locations and the maximum achievable across-location heritability as Eq. 7 shows, so results are linked to the empirical heritability threshold retained. The threshold selected in our study (H=0.75) is adapted to the trial scheme in Southern Xinjiang according to the knowledge of the real heritability values, the threshold to be retained in other trials schemes must result from a previous heritability assessment of those schemes. However, because we proceeded a validation against the three objectives and not only that of genotype ranking, our study provides more solid results. Although, we found that a single replicate could respectably express the patterns involved in three replicates with a few exceptions, two replicates were necessary to achieve the target trial heritability (say, 0.75). Cross-location analysis actually verified that two replicates would be needed for reliable genotype evaluation, hence implying the possibility to reduce the number of replicates from three to two in the cotton trial scheme in southern Xinjiang.

This study provides some advancement in mega-environment investigation. This investigation is actually of fundamental importance in the conduct of multi-locational variety trials because genotype or test location evaluation assume that all test locations belong to a single mega-environment (Yan, 2021a). Mega-environment was previously analyzed extensively using "Analysis annually, summarize perennially" method based on GGE biplot (Yan and Holland, 2010; Xu et al., 2014), GGE plus genotype and genotype by location interaction biplot (GGE+GGL) (Yan et al., 2015; Yan, 2021a), and GGE plus genotype and genotype by sub-region interaction biplot (GGE+GGS) (Yan et al., 2015). However, our study is based on the latest location grouping (LG) biplot (Yan, 2019; Yan et al., 2021), which is currently considered to be the most objective and effective method for ME analysis. In LG biplot however, the trial vectors within LG biplot often overlapped and intertwined together making it sometimes ambiguous to identify the relationships among possible MEs. We provided a method to address this issue by representing the mega-environment (ME) plot in new types of figures (Figures 5B and 6) that are clearer, more straightforward, and more convincing to

18

show all trials within a ME by connecting to the average coordinates of the ME, instead of the average of yearly trials in a location.

Our mode of representation also points out of the better relevance of two replicates to identify ME patterns. With only one replicate, we would not have found a minor ME composed of the single location of Makit coexisting in the current trial scheme along the major ME composed of the other locations. Consequently, our result points out that some adjustment of the current varietal trial scheme in Southern Xinjiang might make sense as such a region covers an acreage of over 1000 thousands hm^2, where soil types could differ as well as fertilizing practices. Makit is located in the alluvial fan oasis plain of the Yarkant and Tiznavu rivers, so it is of the cumulated irrigated soil type (Table 1) with lower contents of organic matter and available phosphorous whose impacts were furthermore accentuated by lower P and K fertilization (Table 1). Therefore, the sole location ME in Makit could be named as the "poor phosphate ME". There is no doubt that the discovery of a poor phosphate ME will be helpful in optimizing the current cotton trial scheme and in guiding the improvement of local cotton field management.

In our study, the outcome of an identified minor ME coexisting with a major one might sound strange within a multi-location trial scheme on an assumed unique mega-environment. Our outcome is nevertheless is similar to what was found in Canada where the "crown rust prone ME" was found in the mega-environment delineation using LG biplot based on the yield data of the 2015-2019 Ottawa oat registration test in Canada(Yan et al., 2021). A major reason is that in all countries, multi-location trial schemes date back to decades ago and were set up on empirical basis without a clear understanding of the notion of mega-environment. Besides, the development and application of GGE biplot analysis alleviate the need for or eliminate the requirement of unique mega-environment in multi-location trial schemes as it enables to select varieties adapted to various mega-environments (Xu et al., 2017).

5. Conclusion

Our study was conducted to proceed more comprehensively the optimization of the number of locations and replicates in the conduct of multi-location variety trials by considering the three objectives of genotype ranking, location evaluation and ME classification, instead of only genotype ranking, in line with the concern of reducing cost and improving trial efficiency. Based on the specific case of cotton in arid conditions of China and the method of LG biplot, our study proposes new representations of ME plot to better identify ME and the trial locations concerned. Our results confirms that a single replicate could be sufficient to achieve a satisfactory trial heritability (set at 0.75 in out study) but two replicates are necessary for secure location evaluation and mega-environment investigation. Compared to the current number of three replicates in the studied trial scheme, our results give the prospect of one-third reduction in trial cost, or allowing 30% more genotypes to be tested at the same cost. The implied savings at equal efficiency justifies similar studies in other countries or on other crops. In the specific case studied, the outlined reality of a minor increasks for adjusting the existing trial scheme with adjusted fertilizing among other cultivation practices in the corresponding locations whose number could be increased for better reliability of adapted genotype selection.

AUTHOR CONTRIBUTIONS

Naiyin Xu: Conceptualization, Data curation, Funding acquisition, Methodology, Writing – original draft. Yintao Qiao: Formal analysis, Writing – original draft. Suqin Zhao: Conceptualization, Funding

acquisition, Investigation, Resources. Xiaoni Yang: Data curation, Investigation, Validation. Jian Li: Investigation, Methodology, Project administration. Michel Fok: Conceptualization, Writing-review & editing.

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.

ORCID

Naiyin Xu https://orcid.org/0000-0003-2291-931X

Yintao Qiao https://orcid.org/0000-0002-5191-4297

Suqin Zhao https://orcid.org/0000-0001-6489-9391

Xiaoni Yang https://orcid.org/0000-0002-4935-3873

Jian Li https://orcid.org/0000-0003-0369-9286

Michel Fok https://orcid.org/0000-0001-8709-5687

REFERENCES

Baxevanos, D., Goulas, C., Rossi, J., & Braojos, E. (2008). Separation of cotton cultivar testing sites based on representativeness and discriminating ability using GGE biplots. *Agronomy Journal,* 100, 1230-1236. https://doi.org/10.2134/agronj2007.0363

Baxevanos, D., Korpetis, E., Irakli, M., & Tsialtas, I.T. (2017a). Evaluation of a durum wheat selection scheme under Mediterranean conditions: Adjusting trial locations and replications. *Euphytica,* 213, 82. https://doi.org/10.1007/s10681-017-1871-y

Baxevanos, D., Tsialtas, L.T., Vlachostergios, D., & Goulas, C. (2017b). Optimum replications and locations for cotton cultivar trials under Mediterranean conditions. *The Journal of Agricultural Science,* 155, 1553-1564. https://doi.org/10.1017/S0021859617000648

Gauch, H.G., & Zobel, R.W. (1997). Identifying mega-environments and targeting genotypes. *Crop Science,* 37, 311-326.

George, N., & Lundy, M. (2019). Quantifying genotype × environment effects in Long-Term common wheat yield trials from an agroecologically diverse production region. *Crop Science,* 59, 1960-1972. https://doi.org/10.2135/cropsci2019.01.0010

Ijaz, B., Zhao, N., Kong, J., & Hua, J. (2019). Fiber quality improvement in upland cotton (*Gossypium hirsutum* L.): Quantitative trait loci mapping and marker assisted selection application. *Frontiers in plant science,* 10:1585. https://doi.org/ 10.3389/fpls.2019.01585

Swallow, W.H., & Wehner, T.C. (1989). Optimum allocation of plots to years, seasons, locations, and replications, and its application to once-over-harvest cucumber trials. *Euphytica*, 43, 59-68. https://doi.org/doi: 10.1007/bf00037897

Woyann, L.G., Zdziarski, A.D., Zanella, R., Rosa, A.C., Conte, J., Meira, D., Storck, L., & Benin, G. (2020). Optimal number of replications and test locations for soybean yield trials in Brazil. *Euphytica*, 216, 11. https://doi.org/10.1007/s10681-019-2548-5

Xu, N., Fok, M., Zhang, G., Li, J., & Zhou, Z. (2014). The application of GGE biplot analysis for evaluating test locations and mega-environment investigation of cotton regional trials. *Journal of Integrative Agriculture*, 13, 1921-1933. https://doi.org/10.1016/S2095-3119(13)60656-5

Xu, N., Jin, S., & Li, J. (2016). Design of test location number and replicate frequency in the regional cotton variety trials in China. (In Chinese, with English abstract.) *Acta Agronomica Sinica*, 42, 43-50. https://doi.org/10.3724/SP.J.1006.2016.00043

Xu, N., & Li, J. (2014). Identification of ideal test environments for multiple traits selection in cotton regional trials. (In Chinese, with English abstract.) *Acta Agronomica Sinica*, 40, 1936-1946. https://doi.org/10.3724/SP.J.1006.2014.01936

Xu, N., Fok,M., Li, J., Yang, X., & Yan,W. (2017). Optimization of cotton variety registration criteria aided with a genotype-by-trait biplot analysis. *Scientific Reports*, 7:17237. https://doi.org/10.1038/s41598-017-17631-4

Yan, W. (2001). GGEbiplot - a windows application for graphical analysis of multienvironment trial data and other types of two-way data. *Agronomy Journal*, 93, 1111-1118. https://doi.org/10.2134/agronj2001.9351111x

Yan, W. (2010). Optimal use of biplots in analysis of multi-location variety test data. (In Chinese, with English abstract.) *Acta Agronomica Sinica*, 36, 1805-1819. https://doi.org/10.3724/SP.J.1006.2010.01805

Yan, W. (2015). Mega-environment analysis and test-location evaluation based on unbalanced multiyear data. *Crop Science*, 55, 113-122. https://doi.org/10.2135/cropsci2014.03.0203

Yan, W. (2016). Analysis and handling of G × E in a practical breeding program. *Crop Science*, 56, 2106-2118. https://doi.org/10.2135/cropsci2015.06.0336

Yan, W. (2019). LG biplot: A method for mega-environment investigation using existing crop variety trial data. *Scientific Reports*, 9:7130. https://doi.org/10.1038/s41598-019-43683-9

Yan, W. (2021a). A systematic narration of some key concepts and procedures in plant breeding. *Frontiers in plant science*, 12, 724517. https://doi.org/10.3389/fpls.2021.724517

Yan, W. (2021b). Estimation of the optimal number of replicates in crop variety trials. *Frontiers in plant science*, 11, 590762. https://doi.org/10.3389/fpls.2020.590762

Yan, W., Frégeau-Reid, J., Martin, R., Pageau, D., & Mitchell-Fetch, J. (2015). How many test locations and replications are needed in crop variety trials for a target region? *Euphytica*, 202, 361-372. https://doi.org/10.1007/s10681-014-1253-7

Yan, W., Mitchell-fetch, J., Betties, A., Nilsen, K., Pageau, D., DeHaan, B., Hayes, M., Mountain, N., Cummiskey, A., & MacEachern, D. (2021). Oat mega-environment in Canada. *Crop Science*, 61, 1141-1153. https://doi.org/10.1002/csc2.20426

Yan, W., Pageau, D., Frégeau-Reid, J., Lajeunesse, J., Goulet, J., Durand, J., & Marois, D. (2011). Oat mega-environments and test-locations in Quebec. *Canadian Journal of Plant Science*, 91, 643-649. https://doi.org/10.4141/cjps 10139

Yan, W., & Holland, J.B. (2010). A heritability-adjusted GGE biplot for test environment evaluation. *Euphytica*, 171, 355-369. https://doi.org/10.1007/s10681-009-0030-5

Yan, W., & Hunt, L.A. (2001). Interpretation of genotype × environment interaction for winter wheat yield in Ontario. *Crop Science*, 41, 19-25.

Yan, W., & Rajcan, I. (2002). Biplot analysis of test sites and trait relations of soybean in Ontario. *Crop Science*, 42, 11-20.

Yan, W., & Tinker, N.A. (2005). An integrated biplot analysis system for displaying, interpreting, and exploring genotype-by-environment interactions. *Crop Science*, 45, 1004-1016. https://doi.org/10.2135/cropsci2004.0076

Yan, W., & Tinker, N.A. (2006). Biplot analysis of multi-environment trial data: Principles and applications. *Canadian Journal of Plant Science*, 86, 623-645. https://doi.org/10.4141/P05-169

Zhang, Y., Xu, N., Guo, L., Yang, Z., Zhang, X., & Yang, X. (2020). Optimization of test location number and replicate frequency in regional winter wheat variety trials in northern winter wheat region in China. (In Chinese, with English abstract.) *Acta Agronomica Sinica*, 46, 1166-1173. https://doi.org/10.3724/SP.J.1006.2020.91069

Table 1 Test locations, geographical coordinates and fertilization levels of the national cotton variety trial in Southern Xinjiang in 2011-2020

Location	Test years	Longitude	Latitude	Altitude	Fertilization (kg·ha^{-1})			Soil type
		°W	°N	m	N	P	K	
Alaer	9	81.26	40.56	1010	432	215	167	Sandy loam
Bachou	10	86.12	41.75	934	347	326	144	Sandy loam
Korla	6	86.84	41.69	900	462	322	181	Sandy loam
Kuqa	9	82.90	41.53	1005	433	239	115	Sandy loam
Makit	6	77.65	38.90	1180	308	95	73	Irrigated silt
Shaya	10	82.79	41.21	983	541	312	100	Sandy loam
Shache	10	77.25	38.41	1235	388	288	77	Sandy loam
Tiemenguan	6	85.67	41.80	909	517	385	68	Sandy loam
Tumxuk	10	79.08	39.87	1094	622	337	136	Sandy loam

Table 2 Variance ratios, trial heritability, coefficient of variation, estimated optimal numbers of replicates and trial locations based on lint yield with different replicates

Parameter	Replicate	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	Mean
σ_e^2/σ_g^2	REP12	0.94	2.44	2.80	1.04	1.20	0.79	1.84	1.18	0.48	1.02	1.37a
	REP13	0.80	3.72	4.69	1.27	9.07	1.06	2.18	1.64	0.77	1.26	2.65a
	REP23	0.68	5.03	2.30	0.94	14.48	1.07	1.90	1.58	0.59	1.15	2.97a
	REP123	0.80	3.46	3.00	1.07	5.15	0.97	1.94	1.44	0.60	1.15	1.96a
σ_{gl}^2/σ_g^2	REP12	1.93	2.71	2.52	1.98	2.56	0.17	2.80	1.58	0.91	1.54	1.87a
	REP13	2.46	2.68	3.01	1.75	5.23	0.30	3.51	1.41	1.19	1.70	2.32a
	REP23	1.56	3.85	2.53	1.44	6.35	0.22	2.22	1.61	0.90	1.51	2.22a
	REP123	1.96	2.90	2.56	1.68	3.95	0.23	2.78	1.51	0.97	1.59	2.01a
H	REP12	0.75	0.67	0.67	0.78	0.69	0.93	0.65	0.79	0.86	0.77	0.76a
	REP13	0.71	0.64	0.60	0.79	0.42	0.89	0.60	0.78	0.82	0.75	0.70b
	REP23	0.79	0.56	0.69	0.83	0.34	0.90	0.69	0.77	0.86	0.77	0.72ab
	REP123	0.76	0.66	0.69	0.82	0.55	0.93	0.67	0.80	0.86	0.78	0.75a
H_{max}	REP12	0.78	0.75	0.76	0.82	0.73	0.98	0.71	0.84	0.89	0.82	0.81a
	REP13	0.74	0.75	0.73	0.84	0.57	0.96	0.67	0.85	0.86	0.81	0.78a
	REP23	0.82	0.68	0.76	0.86	0.52	0.97	0.76	0.83	0.89	0.82	0.79a
	REP123	0.78	0.73	0.76	0.84	0.64	0.97	0.72	0.84	0.88	0.82	0.80a
$CV\%$	REP12	6.03	9.57	7.64	7.45	7.75	8.43	7.65	5.93	5.29	5.93	7.17b
	REP13	5.71	10.41	8.26	7.93	12.93	9.97	8.01	6.77	6.12	6.60	8.27a
	REP23	5.58	11.12	7.23	7.36	12.54	9.51	7.94	6.52	5.96	6.48	8.02a
	REP123	5.78	10.39	7.72	7.59	11.34	9.33	7.87	6.42	5.80	6.34	7.86a
$N_{l,H75}$	REP12	7	12	12	8	9	2	11	7	3	6	8a
	REP13	9	14	16	7	29	2	14	7	5	7	11a
	REP23	6	19	11	6	41	2	10	7	4	6	12a
	REP123	7	12	11	6	17	2	10	6	4	6	8a
$N_{r,H75}$	REP12	1	1	1	1	1	1	1	1	1	1	1.0
	REP13	1	1	1	1	2	1	1	1	1	1	1.1
	REP23	1	1	1	1	3	1	1	1	1	1	1.2
	REP123	1	1	1	1	1	1	1	1	1	1	1.0
Number of genotypes		7	12	14	12	7	8	10	11	10	14	11
Number of locations		7	8	8	9	7	7	7	8	7	7	8

REP123, mean yield based on all three replicates; REP12, mean yield based on the combination of replicates 1 and 2, and so on. σ_e^2/σ_g^2, σ_{gl}^2/σ_g^2, H, H_{max}, CV, $N_{l,H75}$ and $N_{r,H75}$ stand for the ratio of experimental error variance over genotypic variance, the ratio of genotype by location interaction

25

variance over genotypic variance, the trial heritability across locations, the estimated maximum heritability, coefficient of variation, the estimated optimum number of locations let H=0.75, the estimated optimum number of replicates with H=0.75H_{max}. Values followed by different small letters are significantly different at the 0.05 probability level.

Table 3 Pearson correlation coefficients between mean lint yield based on all three replicates and genotypic means based on any one or two replicate combinations in 2011-2020

Year	REP1	REP2	REP3	REP12	REP13	REP23	P < 0.05	P < 0.01
2011	0.97	0.97	0.98	0.99	0.99	0.99	0.708	0.849
2012	0.97	0.93	0.94	0.99	0.98	0.99	0.585	0.711
2013	0.92	0.97	0.97	0.99	0.99	0.99	0.539	0.666
2014	0.98	0.99	0.98	1.00	1.00	1.00	0.585	0.711
2015	0.97	0.91	0.66	0.97	0.97	0.98	0.708	0.849
2016	0.95	0.99	0.95	0.99	1.00	0.98	0.714	0.837
2017	0.95	0.97	0.95	0.99	0.99	0.99	0.639	0.769
2018	0.99	0.96	0.96	0.99	0.99	1.00	0.568	0.710
2019	0.98	0.99	0.99	1.00	1.00	1.00	0.639	0.769
2020	0.97	0.94	0.95	0.99	0.98	0.99	0.539	0.666
Mean	0.96	0.96	0.93	0.99	0.99	0.99	0.568	0.710

REP1, yield based on replicate 1; REP12, mean yield based on replicates 1 and 2, and so on. P < 0.05 and P < 0.01 stand for the threshold values for r at 0.05 and 0.01 significant level..

Table 4 Discrimination ability, representativeness, and desirability index by location×replicate combination resulting from Figure 3, 4.

Parameter	Location	REP1	REP2	REP3	REP12	REP13	REP23	REP123	Mean
Discrimination ability	Alaer	1.425	1.346	1.172	1.383	1.290	1.259	1.310	1.312cd
	Shaya	1.499	1.406	1.345	1.464	1.396	1.355	1.382	1.407b
	Bazhou	1.432	1.517	1.216	1.434	1.364	1.387	1.411	1.394bc
	Korla	1.057	0.628	0.954	1.132	0.937	0.857	0.881	0.921e
	Kuqa	1.368	1.320	1.156	1.332	1.262	1.231	1.284	1.279d
	Makit	0.906	0.989	0.724	0.703	0.866	0.914	0.919	0.860e
	Shache	1.233	1.344	1.240	1.201	1.223	1.252	1.247	1.249d
	Tumxuk	1.640	1.504	1.484	1.562	1.556	1.501	1.539	1.541a
	Tiemenguan	1.317	1.304	1.310	1.320	1.313	1.336	1.329	1.318cd
	Mean	1.320 a	1.262ab	1.178c	1.281ab	1.245bc	1.232bc	1.256ab	
Representative-ness	Alaer	0.891	0.994	0.926	0.958	0.915	0.956	0.943	0.94ab
	Shaya	0.871	0.847	0.848	0.812	0.886	0.861	0.869	0.856bc
	Bazhou	0.913	0.846	0.860	0.965	0.848	0.870	0.869	0.882bc
	Korla	0.016	-0.209	0.029	0.024	-0.004	0.004	-0.009	-0.021e
	Kuqa	0.900	0.783	0.824	0.859	0.861	0.783	0.820	0.833c
	Makit	-0.303	0.183	0.450	-0.095	0.060	0.291	0.118	0.101d
	Shache	1.000	0.951	1.000	1.000	1.000	0.996	0.997	0.992a
	Tumxuk	0.801	0.905	0.797	0.875	0.812	0.860	0.854	0.843bc
	Tiemenguan	0.970	0.996	0.999	0.982	0.994	0.997	0.995	0.990a
	Mean	0.673a	0.700a	0.748a	0.709a	0.708a	0.735a	0.717a	
Desirability index	Alaer	1.270	1.338	1.085	1.325	1.179	1.204	1.235	1.234a
	Shaya	1.306	1.191	1.141	1.188	1.237	1.166	1.201	1.204a
	Bazhou	1.307	1.284	1.045	1.383	1.156	1.206	1.227	1.230a
	Korla	0.017	-0.131	0.028	0.027	-0.004	0.003	-0.008	-0.010c
	Kuqa	1.232	1.034	0.952	1.145	1.086	0.964	1.053	1.067b
	Makit	-0.275	0.181	0.326	-0.067	0.052	0.266	0.108	0.084c
	Shache	1.233	1.278	1.240	1.200	1.223	1.247	1.243	1.238a
	Tumxuk	1.313	1.362	1.183	1.367	1.264	1.291	1.315	1.299a
	Tiemenguan	1.278	1.299	1.308	1.296	1.305	1.332	1.321	1.306a
	Mean	0.965a	0.982a	0.923a	0.985a	0.944a	0.964a	0.966a	

REP123, mean yield based on all three replicates; REP1, yield based on replicate 1; REP12, mean yield based on replicates 1 and 2, and so on. Means of same treatment followed by the same letter are not significantly different at 5% level.

27

Figure captions

Figure 1. Mean vs. Instability view of GGE biplot based on the lint yield data with all three replicates for the 2020 cotton variety trials

The line with a single arrow is the average environment axis (AEA) with the arrow pointing to higher mean yield across locations, while the line with double arrows is the instability coordinate with arrows pointing to lower stability. Genotype marks prefixed with asterisk (*) .Test location marks were replaced by plus sign (+) for clarity.
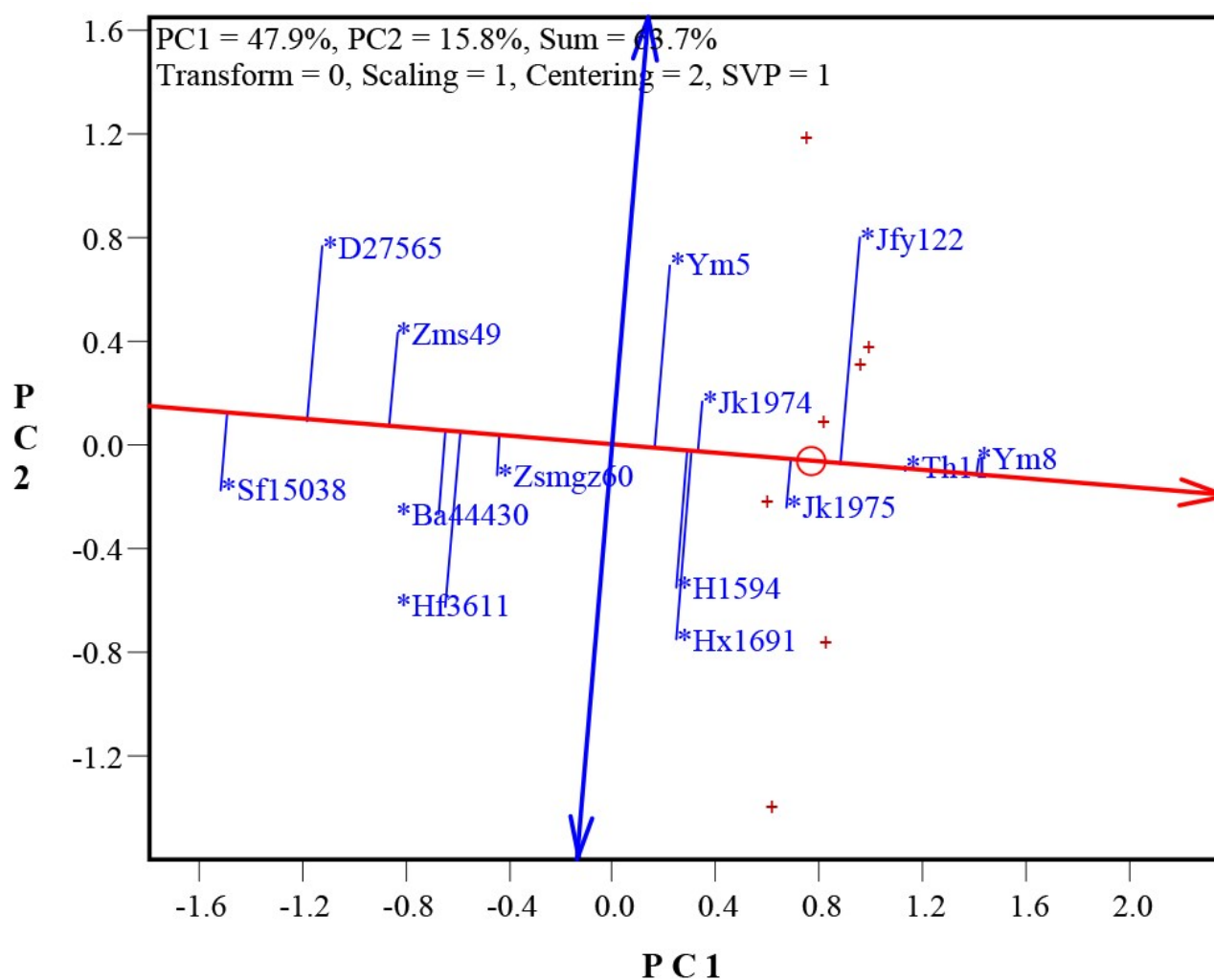

Figure 2. Summary biplot for the 2020 cotton variety trials in Southern Xinjiang to show the similarity/dissimilarity between genotypic lint yield rankings based on one to three replicates.

REP1, genotypic ranking based on replicates 1; REP12, genotypic ranking based on replicates 1 and 2; REP123, genotypic ranking based on all three replicates, and so on. Genotypes within the polygon were replaced by asterisk (*) for clarity.
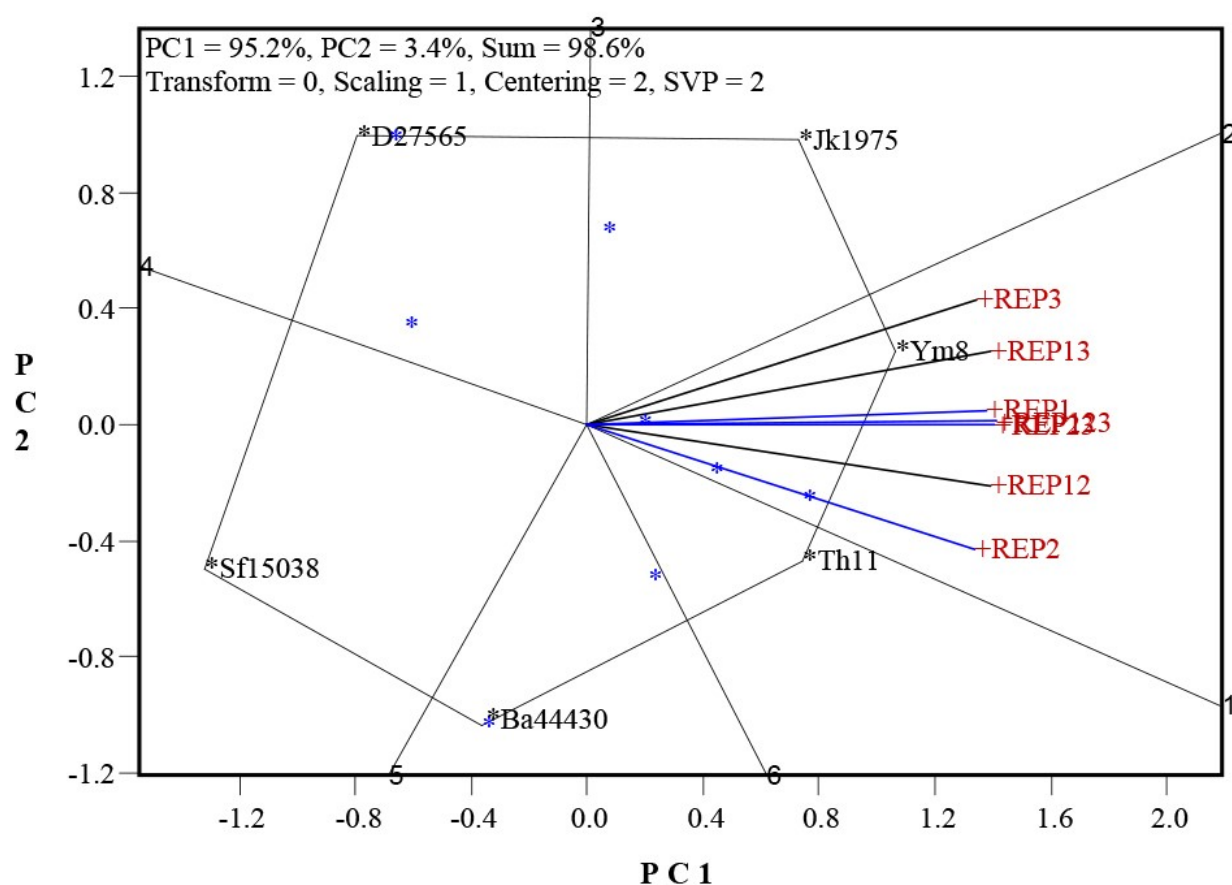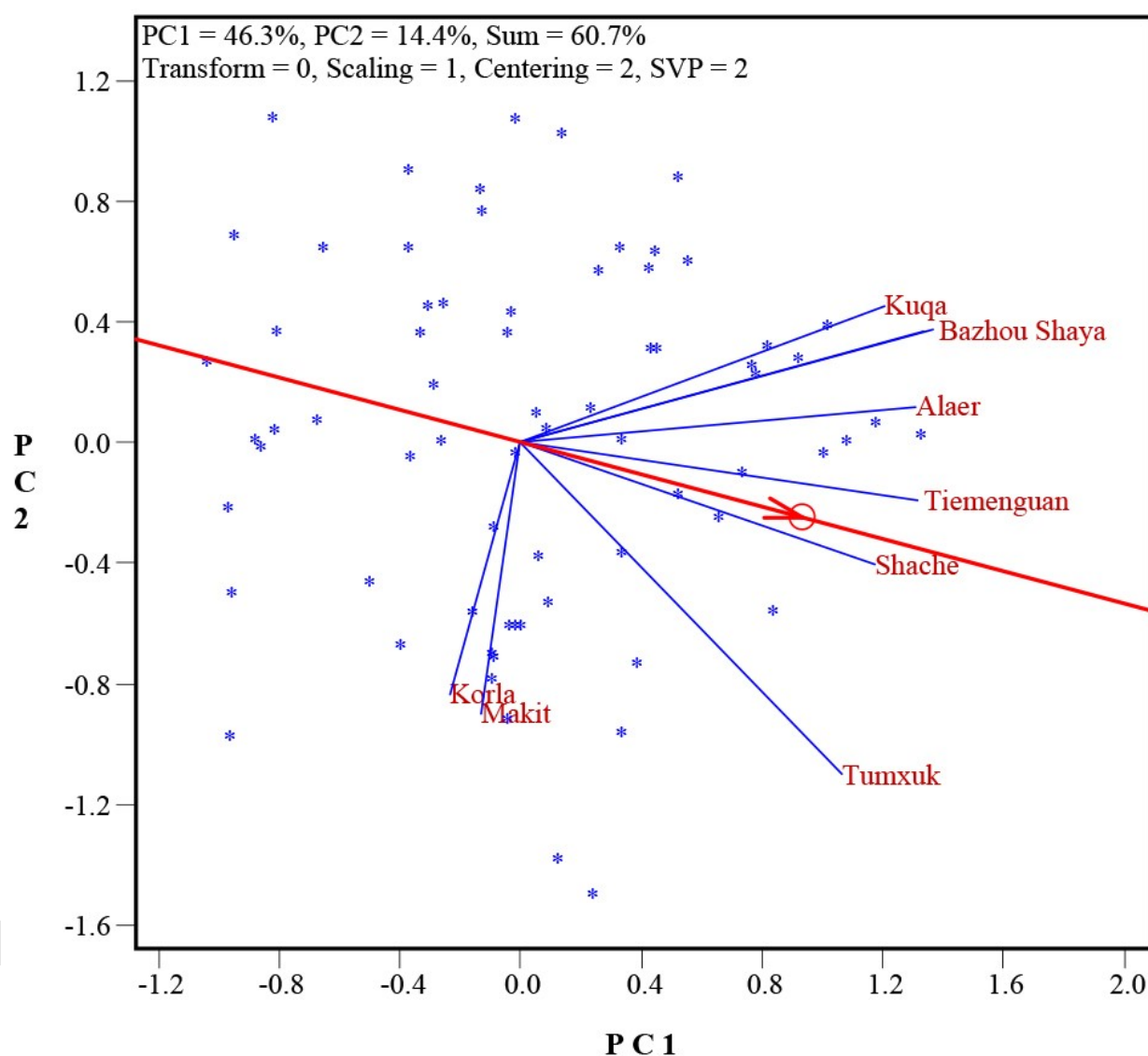

Figure 3. Test location discriminating ability and representativeness view of GGE biplot based on all three replicates for the 2011-2020 cotton trials.

The line with a single arrow is the average environment axis (AEA) showing the average performance of genotypes. Test location marks were prefixed with plus sign (+).The vector length of locations displays the location discriminating ability, the longer the vector the more discriminating the location. The cosine of a location vector and the angle between AEA indicates the representativeness, the larger the angle the less representative the location. Genotype marks were replaced by asterisk (*) for clarity.

Figure 4. Test location discriminating ability and representativeness biplots based on one to three replicates for the 2011-2020 cotton trials.

(A) GGE biplot to show location discriminating ability and representativeness based on lint yield data from replicate 1. (B) Based on replicate 2. (C) Based on replicate 3. (D) Based on replicates 1 and 2. (E) Based on replicates 1 and 3. (E) Based on replicates 2 and 3. See Fig.3 for details.
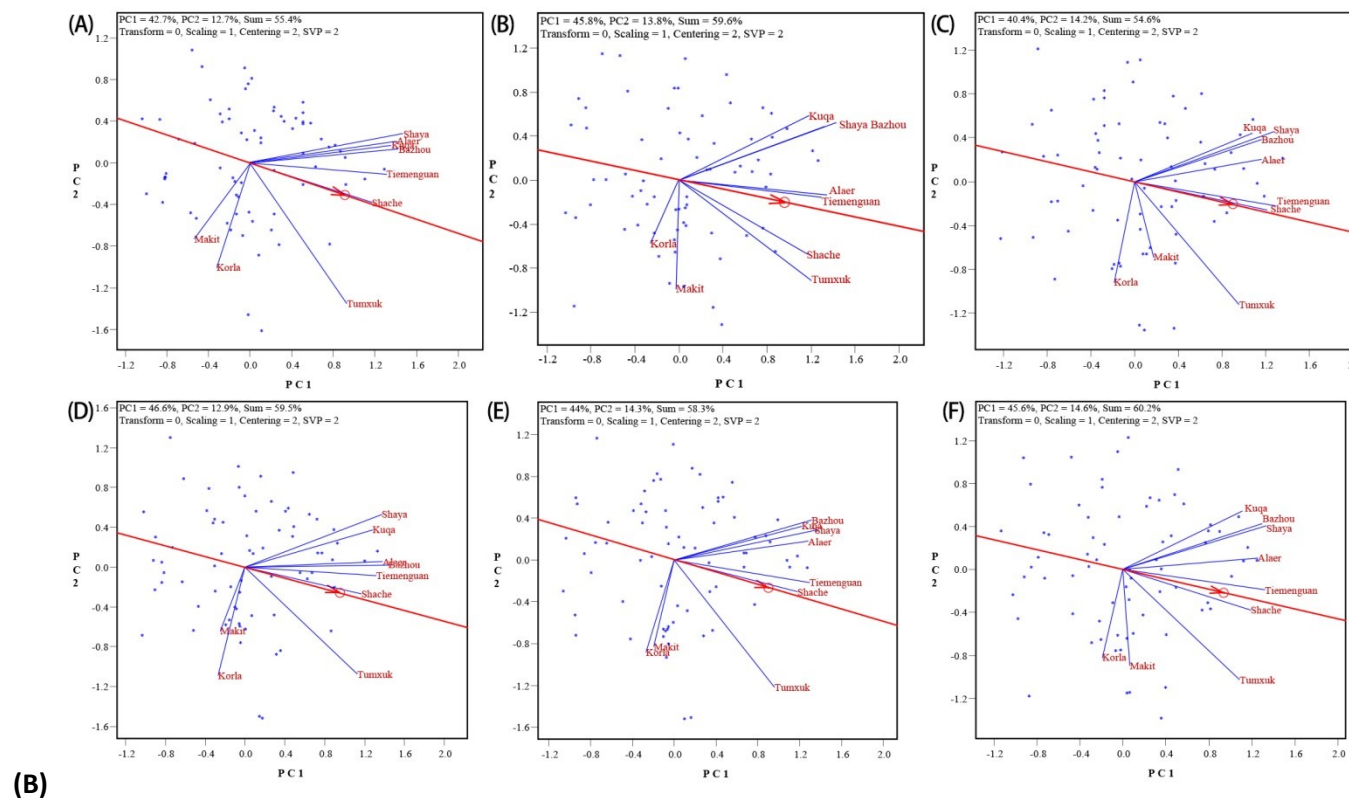
Figure 5. LG biplot (A) and ME plot (B) for mega-environment identification based on three replicates for the 2011-2020 cotton trials.

PC1 and PC2 are the first two principal components from singular value decomposition of the location by trial two-way table of correlations, without centering ("Centering = 0") or scaling ("Scaling = 0"). The singular values were entirely partitioned to the location-year vectors ("SVP = 2"). (A) LG biplot, in which the trials conducted at each location are presented as a cluster of trials, with the location name placed at the center and the individual trials, indicated by the last two digits of the year, placed around it. The trials are connected to the location with straight lines. (B) Mega-environment (ME) plot, in which all trials belonging to the same ME are connected to the mean coordination of those trials to show MEs clearer.
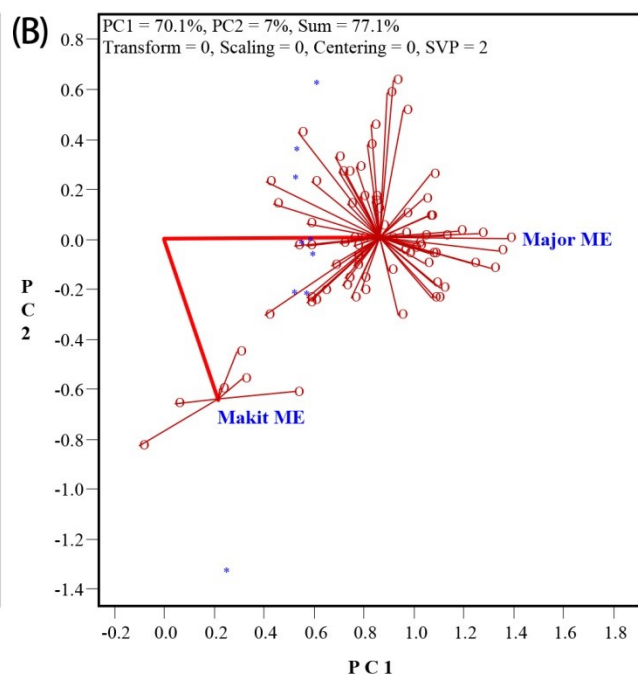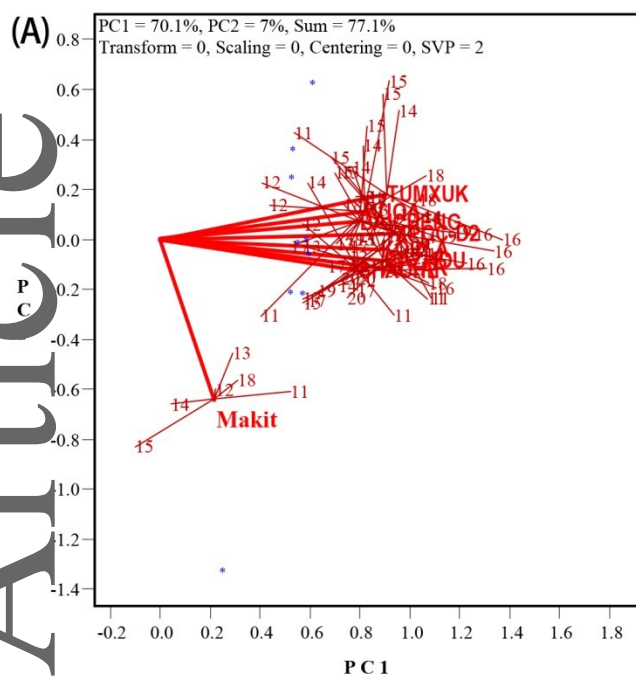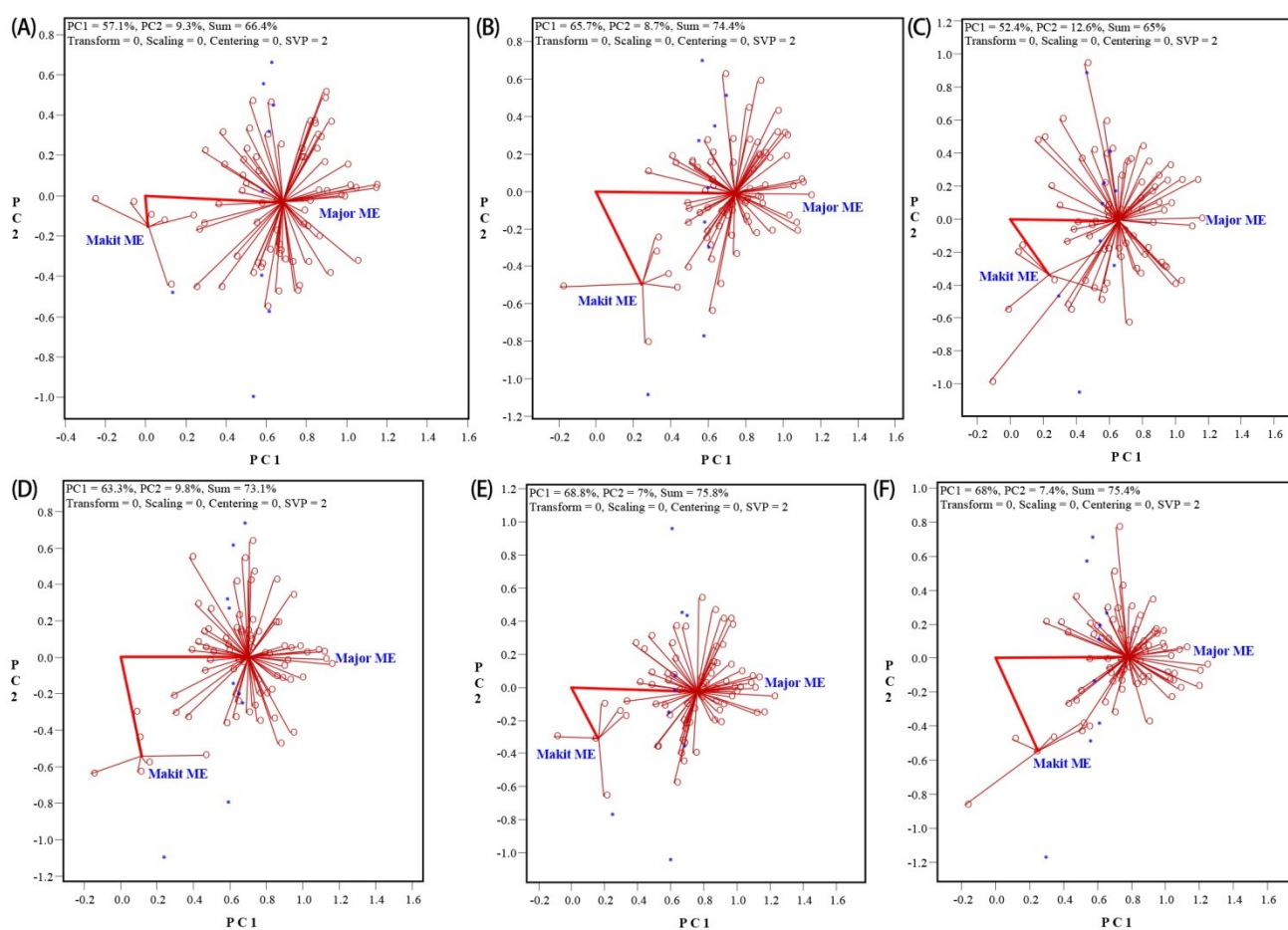
32

Figure 6. ME plots for meg-environment identification based on any one or two replicates for the 2011-2020 cotton trials.

(A) LG biplot, in which the trials conducted at each location are presented as a cluster of trials, with the location name placed at the center and the individual trials, indicated by the last two digits of the year, placed around it. The trials are connected to the location with straight lines. (A) ME plot to show relationship between MEs based on data from replicate 1. (B) Based on replicate 2. (C) Based on replicate 3. (D) Based on replicates 1 and 2. (E) Based on replicates 1 and 3. (E) Based on replicates 2 and 3. See Fig.5 for details.