Field Crops Research 110 (2009) 272–281

Contents lists available at ScienceDirect

Field Crops Research

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

Determination of efficient test sites for evaluation of peanut breeding lines using the CSM-CROPGRO-peanut model

C. Putto [a], A. Patanothai [a,*], S. Jogloy [a], K. Pannangpetch [a], K.J. Boote [b], G. Hoogenboom [c]

[a] *Department of Plant Science and Agricultural Resources, Faculty of Agriculture, Khon Kaen University, Mitraphap Road, Amphur Maung, Khon Kaen 40002, Thailand*
[b] *Department of Agronomy, University of Florida, Gainesville, FL 32611-0500, USA*
[c] *Department of Biological and Agricultural Engineering, The University of Georgia, Griffin, GA 30223-1797, USA*

ARTICLE INFO

Article history:
Received 18 July 2008
Received in revised form 30 September 2008
Accepted 30 September 2008

Keywords:
Multi-environment trials (METs)
Cluster analysis
GGE biplot
Crop simulation model
Peanut breeding

ABSTRACT

Efficient testing of crop breeding lines requires a set of complementary test sites that adequately sample the target environments with minimal duplication. Such test sites have been derived from actual multi-environmental trial (MET) data, which often have a limitation with respect to their environmental coverage. However, this limitation can be overcome using a crop simulation model. The goal of this study was to determine the efficient test sites for METs of peanut breeding lines in Thailand using the CSM-CROPGRO-Peanut model. The model was used to simulate pod yield for 17 peanut lines at all peanut production areas in Thailand that included 76 locations in the early-rainy season, 39 locations in the mid-rainy season and 47 locations in the dry season for 30 years. The simulated data were used to sub-divide the locations for each season into groups using cluster analysis and the genotype plus genotype × environment (GGE) biplot method. Six sets of test sites were obtained based on different scenarios for site selection that included combinations of geographical distribution and representation of location-groups as determined by the two methods. Set 1 was based on geographic distribution. Sets 2–4 were based on location grouping by cluster analysis, but with the sites distributed in all regions (Set 2), or only in the north (Set 3) or northeast (Set 4). Set 5 consisted of the sites currently used, and Set 6 was based on location grouping by the GGE biplot. Although Sets 2 and 6 appeared to capture more genotype × location interaction than the others, performance rankings of the test genotypes were almost the same for all sets. They were, therefore, considered equally effective for breeding line evaluation. The final selection was then based on the convenience, and consequently the cost, for conducting the METs. Set 4 was considered most preferable in this regard. This study demonstrated the usefulness of a crop simulation model as a tool in determining the most efficient test sites for the evaluation of peanut breeding lines.

1. Introduction

Crop breeding lines are generally evaluated in multi-environment trials (METs) over a set of sites and years as the performance ranking might change for different environments due to genotype × environment interaction (Fehr, 1987; Kang, 1990). Generally, the sites for conducting METs are located in different geographical areas or agro-climatic zones (Navabi et al., 2006; Roozeboom et al., 2008). However, they are often determined by arbitrary factors, such as convenience, administrative region, etc., and sometimes by chance (Fan et al., 2001). It has been well recognized that only the crossover type of G × E interaction is associated with significant genotypic rank change, and, thus, has significant implication on the multi-environment evaluation of crop breeding lines (Baker, 1988). If two or more test sites exhibit no or non-crossover G × E interaction, performance ranking of the test genotypes would be the same and testing would only be needed at one of these sites (Fehr, 1987; Annicchiarico, 2002). An efficient testing of crop breeding lines, therefore, requires a set of complementary test sites that adequately sample the environments of interest with minimal duplication (Hambling et al., 1980). The use of inappropriate test sites will not only lower the effectiveness of breeding line evaluation, but also waste valuable time and resources. Thus, the choice of test sites is an important breeding decision.

Efforts have been made to determine appropriate test sites for breeding line evaluation. The most widely used approach is to first group the sites based on their similarity in genotypic responses, and then select a representative site from each group. Clustering techniques using squared Euclidean distance as the dissimilarity measure and incremental sum of squares or Ward's strategy

---

* Corresponding author. Fax: +66 43 364636.
E-mail address: aran@kku.ac.th (A. Patanothai).

0378-4290/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.fcr.2008.09.007

(Ward, 1963) as the clustering strategy are the methodologies that have been used extensively in subdividing the test locations (Delacy and Cooper, 1990; Abdalla et al., 1996). Two types of models that are suitable for identifying subsets of locations with negligible crossover G × E interaction are the shift multiplicative model (SHMN) (Seyedsadr and Conelius, 1992) and the sites regression model (SREG) (Conelius et al., 1996). Both models have been found to identify similar groups of locations with low levels of crossover G × E interaction (Crossa et al., 2002). The genotype and genotype × environment interaction (GGE) biplot method developed by Yan et al. (2000) is equivalent to the SREG method (Crossa et al., 2002). Recently, this method has become popular, including its use in environmental grouping (Navabi et al., 2006; Roozeboom et al., 2008).

Test sites grouping based on these methods has been conducted with actual MET data, but has often failed to provide an adequate coverage of the target population of environments, because there is a practical limit in both the number of sites and years for which METs can actually be conducted. Consequently, the results of such location groupings were constrained by this limitation. Presently, crop simulation models have been developed to the level that they can simulate growth and development of cultivars and varieties for different environments and agronomic practices (Boote et al., 1998, 2003). These models, therefore, can be used to provide the required data for a legitimate classification of target environments for a breeding program.

For peanut, the process-oriented Cropping System Model (CSM)-CROPGRO-Peanut has been developed and is part of the Decision Support System for Agrotechnology Transfer (DSSAT) (Jones et al., 2003; Hoogenboom et al., 2004). This model has been evaluated extensively for various breeding applications in Thailand (Banterng et al., 2006; Phakamas et al., 2008; Putto et al., 2008; Suriharn et al., 2008). It is, thus, a potential tool for obtaining yield data for peanut genotypes for a large number of sites and years that could provide a full coverage of the target population of environments in order to be able to derive efficient test sites.

Sufficient METs, especially in developing countries are rare, mainly due to the lack of sufficient resources to support a comprehensive breeding program. In Thailand, the METs of peanut breeding lines that are conducted by the peanut breeding program of Khon Kaen University (KKU) are mainly located in the northeast region with some in the northern region. These sites had been arbitrarily determined with some consideration on environmental coverage and geographical distribution, but they were mainly selected based on the convenience, and thus the cost for conducting the trials. In addition, the number of test sites is still limited and does not provide an adequate coverage of the target population of environments for the assessment of effective test sites. As the identification of efficient test sites has never been investigated, the objective of this study was to determine the efficient test sites for METs of peanut breeding lines in Thailand using the CSM-CROPGRO-Peanut model.

2. Materials and methods

The main procedure used in this study consisted of simulating pod yield of a set of peanut lines for all peanut production areas in Thailand for 30 years. The simulated yield data were then used to subdivide the identified locations in each of the three growing seasons, i.e., the early-rainy, mid-rainy and dry season, into groups based on their similarity in genotypic responses. Sets of test sites were determined for the individual seasons based on different criteria. These sets of test sites were evaluated and compared for their efficiency in performance evaluation of peanut breeding lines.

2.1. Simulation of multi-environment trial (MET) data

To determine the specific peanut production locations in each season, data for peanut production organized by district for the 2002–2003 crop year were obtained from the Thai Department of Agricultural Extension. Only districts that had a considerable peanut acreage (>160 ha) were selected. This resulted in a total of 45 districts in 26 provinces, with the latitude ranging from 12°27′N to 19°57′N and the longitude ranging from 98°33′E to 104°43′E. Eleven of these provinces were in the Northeast, one in the West, eight in the North, two in the East and four in the Central region of Thailand. Questionnaires were then sent to the district extension agents requesting identification of the main peanut-producing villages in their districts and information about the agronomic management practices of the farmers in these villages. The information that was requested included the various growing seasons, range of planting dates, local soil characteristics and irrigation practices. Once the production villages in the individual districts were identified, soil types for the villages were determined based on the soil map and associated database of the Thai Department of Land Development. The weather station located in or adjacent to each growing area was also determined. In total, 24 weather stations were identified. The basic units for model simulation, designated as locations in this study, were then

Table 1

Peanut breeding lines and cultivars used in this study and the source of the cultivar coefficients.

Entry no.	Line/cultivar	Seed type	Maturity duration	Source of cultivar coefficients
1	(Luhua 11 × KK60-3) F6-22	Large-seeded	Early	Suriharn et al. (2007)
2	KK 5	Small-seeded	Early	Suriharn et al. (2007)
3	KKU 1	Small-seeded	Early	Suriharn et al. (2007)
4	((Nc Ac 17090 × B1) –25 × Luhua 11) F5-14-2	Large-seeded	Medium	Suriharn et al. (2007)
5	(Luhua 11 × China 97-2) F6-11-3	Large-seeded	Medium	Suriharn et al. (2007)
6	(China 97-2 × Singburi) F6-13-1	Large-seeded	Medium	Suriharn et al. (2007)
7	((Nc Ac 17090 × B1)–25 × China 97-2) F5-11-2	Large-seeded	Late	Suriharn et al. (2007)
8	((Nc Ac 17090 × B1)–25 × KK60-3) F6-2-2	Large-seeded	Medium	Suriharn et al. (2007)
9	KKU 72-1	Large-seeded	Late	Suriharn et al. (2007)
10	KK 60-3	Large-seeded	Late	Suriharn et al. (2007)
11	(China 97-2 × KK 60-3) F6-9-1	Large-seeded	Late	Anothai et al. (2007)
12	KKFC 4008-5	Large-seeded	Late	Anothai et al. (2007)
13	(KK 60-3 × (MGS9 × Chico)-12-16-1)26-5-33	Small-seeded	Medium	Banterng et al. (2004)
14	(KK60-3 × (Ah 65 × NC Ac 17090))-3-11-7	Small-seeded	Medium	Banterng et al. (2004)
15	(KK 4 × NC8C)-33-6-21	Small-seeded	Medium	Banterng et al. (2004)
16	A progeny of Tainan 9 × Moket	Small-seeded	Medium	Banterng et al. (2004)
17	(Chico × GA 119-20)8-3-12	Small-seeded	Early	Sujariya (2004)