SAMPLING ISSUES FOR COTMAN DATA COLLECTION

Mark Cochran, Kieth Vodrazka, N. P. Tugwell, Diana Danforth and Sha Mi
University of Arkansas
Fayetteville, AR

Abstract

This study was conducted to determine the appropriate sampling design for COTMAN data collections. Design issues focus on the number of sites per field and the number of plants per site that should be sampled. Intensive sampling of 84 commercial fields in Northeast Arkansas (Mississippi County) and 93 commercial fields in east Central Arkansas (Lee and Phillips Counties) were collected in 1995. Data were analyzed for plant height, first position square retention, squaring nodes and nodes above white flower. A nested ANOVA was used to assess the percent contribution of to total variance attributable to sites and to plants per sites. The sample size for a given cost per field that minimized the variance was also calculated. These optimal sample sizes were contrasted with current COTMAN sampling recommendations. With the possible exception of plant height measures, the optimal sample sizes fell within the recommended ranges.

Introduction

COTMAN is a computerized expert system that interprets in-season crop monitoring data to detect plant stress, evaluate plant development, and identify the last effective boll population for end-of-season management decisions such as insecticide termination and harvest initiation (Bourland et. al., 1994). Appropriate sampling designs for the collection of the plant monitoring data has long been a concern.

The COTMAN program is divided into two major components: SQUAREMAN for managing the squaring period and BOLLMAN for the late and end of season management. Each component has its own unique data requirements. The general data requirements for the overall program included farm level information such as total cotton acreage, long term weather records, harvest picker capacity and a target harvest completion date. Required information for each field included acreage and planting date with optional data on cultivar, soil type, irrigation and percent replant. Data for the SQUAREMAN component is divided into two categories: data collected once per season and data collected once or twice per week. Stand count densities and the first fruiting node number comprise the first category while average plant height, number of fruiting nodes and first position fruit retention comprise the former category. This data is collected between first square and first flower. The BOLLMAN plant monitoring data consists of nodes above whiter flower (NAWF) collected from first flower to cutout. The key plant monitoring information collected via sampling for SQUAREMAN and BOLLMAN includes plant height, number of squaring nodes, first position fruit retention and nodes above white flower. Current sampling procedures proscribe visits to a number of sites in each field and the sampling of a number of plants at each site. Sampling concerns focus on the appropriate number of site to visit per field and the appropriate number of plants to monitor per site.

The objectives of this study were to: 1) measure the variability in relevant plant monitoring data and to establish the contribution to total variance of variability related to plants per site relative to sites per field; 2) determine appropriate sampling design of number of plants per site and the number of sites per field. The results of this study will be contrasted with current COTMAN recommended sampling procedures and examined as evidence for revisions in the sampling scheme.

The current COTMAN sampling recommendations are as follows: plant
height – 2 measurements per site with 4 to 8 sites per field; number of squaring nodes – 10 plants per site with 4 to 8 sites per field; first position square retention – 10 plants per site with 4 to 8 sites per field; and NAWF – 10 plants per site with 4 - 8 sites per field.

Data

Data were collected from two locations in Arkansas during the 1995 growing season. In northeast Arkansas (Mississippi county), 84 commercial fields with an average field size of 57 acres were intensively sampled. In east central Arkansas, data were collected from 93 commercial fields with average filed size of 62 acres. For each field, the relevant plant monitoring indicators were intensively monitored at least once a week and sampling times were carefully noted. Additional details may be found in Vodrazka.

Results

By examining the data as a nested ANOVA, the percent contribution to total variance was calculated for weekly observations of plant height, squaring nodes, square retention and NAWF. The weekly calculations for each monitoring indicator did not vary much so an average across all weeks was estimated. These estimates provide an indication of the relative importance of the variability experienced from plants at a site to the variability observed between sites. This relative importance is critical since the marginal costs of sampling an additional plant at a given site are much lower than sampling an additional site in a field.

Reprinted from the *Proceedings of the Beltwide Cotton Conference*
Volume 1:463-464 (1999)
National Cotton Council, Memphis TN

The percent contribution to total variance for square retention, squaring nodes and NAWF is much higher for plants at a site than for sites in a field. For these three plant monitoring indicators, the variability between plants is 3 to 14 times greater than the variability between sites. These results were very consistent across weeks and locations. This implies that marginal efficiency from increasing sample size in the number of plants per site will likely be greater than increasing the number of sites per field. However, this pattern was not consistently observed for plant height. In this case, the percent contribution to total variance was greater at the site level.

The relationships between sample size and percent contribution to total variance were analyzed with the marginal sampling costs to determine an optimal sampling design. Marginal sampling time was used as a proxy for marginal sampling costs using data collected by King. The optimal sampling design was determined by selecting the number of plants per site and the number of sites per field that minimized total variance for a given number of minutes to sample a field. Optimal sampling designs were calculated for 15, 20, 25 and 30 minutes per field and for each week of the relevant sampling period. These are summarized and contrasted with the COTMAN current recommendations in Table 1. For plant height, the optimal sampling design ranged from 2 to 4 measurements at 5 to 11 sites. For squaring nodes the optimal samples ranged from 7 to 14 plants at 3 to 10 sites. For square retention, the optimal ranges were 11 to 27 plants at 2 to 7 sites. For NAWF, the optimal results were 7 to 23 plants at 3 to 8 sites.

Conclusion

With the exception of plant height, the ranges for the optimal sampling designs corresponded with the current COTMAN sampling recommendations. This is particularly true when one recalls that plant height, squaring nodes and square retention are sampled in conjunction with one another. With plant height, this study provides some evidence that perhaps the recommendation should be revised to include more sites. Estimates for square retention could also be improved by sampling more plants at a given site.

References

Bourland, F. , N. P. Tugwell, D. M. Oosterhuis, and M. Cochran. 1994. Cotton plant monitoring: The Arkansas system. Proceedings Beltwide Cotton Conference. p. 1280-1281.

King, W. H. 1997. Sampling the cotton growth and developmental process: Implications in Heliothine larval survival and late-season insect control costs. U. of Arkansas M.S. Thesis.

Vodrazka, K. 1998. Optimal Number of plants and sample sites for monitoring main-stem nodal development, square retention and plant height in cotton. U. of Arkansas M.S. Thesis.

Table 1. Comparison of Estimates that Minimize Variance with COTMAN Recommendations.

	COTMAN	Minimum Variance Ranges
Plant Height	2 measures (4-8 sites)	2-4 measures (5-11 sites)
Squaring Nodes	10 plants (4-8 sites)	7-14 plants (3-10 sites)
Square Relation	10 plants (4-8 sites)	11-27 plants (2-7 sites)
Nodes-Above-White-Flower	10 plants (4-8 sites)	7-23 plants (3-8 sites)