DISCUSSION

B. V. Sukhatme, Iowa State University

The availability of timely and relatively inexpensive remote sensing data through NASA land observatory satellite has opened up new possibilities to design agricultural surveys to obtain reliable acreage and production estimates for different crops. The organization of this session is timely and the organizers are to be congratulated for devoting this session entirely to satellite agricultural surveys. I wish to congratulate the authors for their contributions and thank the chairman and the organizers for giving me an opportunity to participate in the discussion.

I shall first consider the paper by Chhikara and Feiveson. The paper discusses various problems associated with the design of the survey and analysis of the data collected in an experimental study carried out to demonstrate the applicability of remote sensing technology to large-scale wheat acreage estimation in the U.S.A. and other wheat growing countries. Based on their results, the authors conclude that remote sensing technology can be successfully used for the estimation of large area crop acreage at least for the combined set of crops such as wheat, barley and oats which have the same growing pattern. This is commendable and is likely to be appreciated by all, especially those countries whose acreage statistics are unsatisfactory and are not reported or published regularly. Since I am not conversant with the LACIE program, my comments may sometimes be in the nature of questions and may be even naive.

i) It has been remarked that there is some difficulty in distinguishing wheat, barley and oats from one another. In certain situations, it may be desirable to have separate acreage estimates for the three crops. It would be necessary to develop proper techniques for this purpose. A two phase design similar to the one used by Thomas suggests itself.

ii) The authors have mentioned several problems giving rise to measurement error and non-response. It would be worthwhile to have some idea of their contribution to the bias and mean square error and to investigate whether it is possible to have a built-in device in the sampling design to take care of problems arising due to non-response and measurement errors.

iii) An important aspect of the design of the survey is the allocation of the sample sizes to the different strata and the authors have devoted considerable attention to this problem. Several questions arise:

a) In minimizing the apriori estimate of variance, the within stratum variance S_K^2 has been assumed to be proportional to $P_k(1-P_k)$ where P_k is the historical proportion of wheat in the SPD corresponding to the k-th stratum. It may be desirable to investigate whether this assumption is justified. If the assumption cannot be justified, it may be desirable to investigate the efficiency of this allocation with respect to proportional allocation.

b) The t_k as determined by optimal allocation are not necessarily integers. For this purpose the strata were categorized into three groups. The t_k's for group I with $t_k > 1$ were rounded to the nearest integers. For group II with $0{\cdot}1 \leq t_k < 1{\cdot}0$, a certain number of units equal to the integer closest to Σt_k were drawn using a two-stage design while for group III with $t_k < 0{\cdot}1$, no sampling was carried out. This is certainly a valid solution. However, it is questionable whether the allocation so determined will in fact minimize the variance of the estimated acreage.

c) Several solutions suggest themselves. One possible solution is to reduce the number of strata. Yet another solution would be to group all the group II strata in to Σt_k new strata and draw one unit from each one of them. A somewhat different approach will be to determine the allocation by minimizing the variance of the acreage estimate subject to the condition that t_k is an integer. It may be desirable to investigate the efficiency of the different approaches to the problem of allocation.

d) The authors have considered at great length the problem of estimation of variance. The formulae are necessarily complicated for group II. Some of the earlier suggestions may not only help in simplifying the design but also simplify considerably the problem of estimation of variance.

e) To compare LACIE and SRS acreage estimates month by month, it would be useful to have the corresponding mean square cerrors.

I shall now consider the paper by Sigman, Hanuschak, Craig, Cook and Cardenas. ESCS conducts every year a nationwide agricultural survey called the June Enumerative Survey collecting all types of agricultural data besides the acreage data. This paper proposes using Landsat-data as auxiliary data to provide improved crop acreage estimates for multi-county areas such as districts and states.

i) As remarked by the authors, the data collected in JES survey is not only used to develop discriminant functions for the Landsat-data but also in subsequent regression estimators. It is therefore likely that the square of the estimated correlation coefficient has a large positive bias resulting in overestimating the gain in precision

obtained by using Ladsat-data as auxiliary data. It is therefore important to develop estimates of R_h^2 which are less biased especially when the sample size is small.

ii) It may also be desirable to investigate the validity of the normality assumption needed in developing the discriminant function.

iii) The parameters of the multivariate normal distribution are not known and will have to be estimated. If the sample size is small, the estimated parameters may not be reliable introducing additional errors in the classification based on the discriminant function.

iv) A question also arises concerning the choice of the prior probabilities for the various categories. If the assigned probabilities differ considerably from the true probabilities, the extent of misclassification of pixels in various categories may be considerable.

v) In view of the various problems associated with the discriminant function, it may be desirable to consider some sort of multivariate ratio or regression type estimators with M.S.S. measurements from the satellite as the auxiliary variables. I understand that this was in fact suggested by Professor Wayne Fuller.

vi) Another point worth considering is the effect of discriminant function on the variance of the regression estimator. Since the auxiliary variable based on discriminant function is a random variable, the expression for the variance of the estimator may have to be suitably modified.

Finally, I shall consider the paper by Thomas of the University of California, Berkeley. The paper suggests a two-phase design using data obtained from image analysis of Landsat-data as auxiliary data to provide improved crop acreage estimates at the county level. Besides, the paper has also proposed a three-phase design using image analysis data along with aerial photography and conventional ground survey data to provide precise estimates of irrigated acreage at state level.

i) It is not clear how the sample was selected at each phase of sampling. It is remarked that sampling was carried out with probability proportional to size and with replacement while the formulae for regression estimation appear to assume simple random sampling without replacement.

ii) The standard error of the estimated wheat proportion based on pps estimator is much smaller than that based on the regression estimator. The author has also raised questions concerning the bias and stability of the estimated regression coefficients. It is not at all clear why the regression estimator is preferred to pps estimators especially when the sample size at the ground level is so small. If pps estimator is in fact superior to regression estimator, further gains in precision may be possible by using pps sampling without replacement.

iii) The phase one estimator of irrigation proportion is clearly the usual ratio estimator. However, the expression for its estimated variance is different than the one normally used in practice. Some justification would be desireable.

iv) The two phase and three phase estimators for irrigation proportion proposed by the author appear to be somewhat different than the estimators likely to be used in such situations and it would be interesting to know the motivation and possibly some justification for considering such estimators.