METHODS OF COMPILING CROP STATISTICS IN CHINA*

YOSHIRŌ MATSUDA

Introduction

In this paper the author wishes to review the statistical methods used for crop statistics in Mainland China and the changes in the manner of compilation which occurred during the years 1952 to 1959. He intends also to explore the accuracy of Chinese crop statistics in so far as they were affected by those changes.

In the First Five-Year Plan period (1953–1957), agricultural statistics in China, especially those of crops, appear to have been compiled essentially according to two parallel methods: the local government administrative officers' "complete-enumeration" and the local statistical officers' "typological survey." But towards the end of this period, side by side with a nation-wide transformation of individual farms into Agricultural Production Co-operatives, the State Statistical Bureau issued a directive called the "Act on Methods of Collecting Crop Statistics for 1956," and required each of the Agricultural Production Co-operatives to submit the Standard Returns of the Co-operative. While this Act did not intend to discontinue the typological survey, there appears to have been a widespread misunderstanding on the side of the local governments which resulted in its discontinuance.

The Sixth National Statistical Work Conference convened in 1957 urged with much more vigour the continued enforcement of this typological sampling survey along with the complete-enumeration survey.

* This is a revised version of a part of the writer's earlier paper published as Chap. 3 of S. Ishikawa (ed.), *Chūgoku Keizai Hatten no Tōkei-teki Kenkyū* -III- (Statistical Studies on Economic Development in China -III-), Tokyo, The Institute of Asian Economic Affairs, 1962, pp. 208–229. Acknowledgements are due to Professor Shigeru Ishikawa, who wrote the excellent seminal paper in this field, "Chūgoku no Nōsakubutsu Shūkaku Tōkei-hō" (Method of Compiling Agricultural Outputs), *ECAFE Tsūshin* (ECAFE Information), No. 208, for his kind permission to use his collection of official Chinese publications and for his painstaking scrutiny of the earlier versions at each stage of rewriting. The author is also grateful to Professor Mataji Umemura, Mr. Yuichi Kanda, Dr. Yoshio Tsumura, Professor Tairoku Kose, and Dr. John Lossing Buck who read the manuscripts and gave valuable comments and criticism. But the writer is solely responsible for any possible error.

Although the results of this important reform had been a matter of common concern, what actually followed were well-known "erroneous reports" of the agricultural output figures for 1958. The extraordinarily abundant harvest of that year was accompanied by an overall change in the management system of farms, i. e., the reorganization of the Agricultural Production Co-operatives into People's Communes. During the succeeding three years continued nation-wide natural disasters caused unusually poor harvests; People's Communes were substantially reorganized into smaller units. Concerning statistical activities during these years we know very little except that an attempt was made in 1959 to obtain nation-wide crop surveys, which again used dualistic complete-enumeration and typological sampling.

In the writer's opinion, the starting-point of any systematic appraisal of China's crop statistics should be an attempt at the clarification of this dualistic method. This seems to be especially the case in order to evaluate the nature of the erroneous 1958 report, although such an overall appraisal is impossible, due to limited official data.[1]

I. OUTLINE OF CHANGES IN STATISTICAL METHODS

This section will examine chronologically the development of the statistical methods used for the compilation of crop statistics in China.

1. Pre-Collectivization Period

The first period (1953–1956) may be characterized by the nation-wide enforcement of complete enumeration, and the gradual introduction of a typological sampling survey required as a check of the former.

The complete-enumeration survey was enforced at the "Periodic Report System" in agriculture. It was tentatively brought into operation at the time of the establishment of the State Statistical Bureau in October, 1952, and has been continued regularly since then.[2] The accuracy of the statistical information obtained by it depends on both the smallness of the basic unit of enumeration, the ultimate identifiable

1 As to the details, see the writer's mimeographed paper, "Methods of Compling Crop Statistics in China," issued by the Computing Centre of Otaru University, 1965.

2 This system was established under the influence and guidance of experts from the Soviet Union. Chou Hung, "Agricultural Statistics Have Basically Reflected Marked Changes in Rural Economy," *T'ungchi Kungtso* (*TCKT*, Statistical Work), No. 18, 1957; Hsieh Mu-ch'iao, "Preliminary Experiences of Our Country's Statistical Activities during the Period of the First Five-Year Plan and Their Future Tasks," *TCKT*, No. 21, 1957.

unit to which the information can be traced back, and the technique of enumeration used by it. With respect to the smallness of the basic enumeration unit, in fact it varied from Province to Province. Thus, Hsieh Mu-ch'iao confirmed in 1957 that the Periodic Reports to the State Statistical Bureau were required in the past only to the Provincial level, and not to the governments at a lower level than Province. To what level of government and according to what methods the basic enumeration was required were determined by each of the Provincial Governments.[1]

The aggregate statistics, therefore, varied in reliability according to the difference with which the local statisticians obtained access to the individual farms.[2] In this respect, the Chinese authorities explained to the Indian Delegation to China on Agricultural Planning and Techniques, which visited China in July-August, 1956, that the primary statistics were collected by the statistical officers at "*hsien*," or county, and provincial levels.[3] This suggests that the aggregate statistics are at best the summation of reports by the officers in *hsien* and not in "*hsiang*," or villages. This appears to reflect the fact that, though the statistical organization of the *hsien* was comparatively well organized, those of "*ch'ü*," or district, and *hsiang* had not yet been universally organized.[4]

With respect to the technique of enumeration, no detailed information is available. However, the Indian Delegation made a remark that as the officers' reports were based on hearsay information or guesswork at best through eye-surveys in the villages, their reliability is not so high.

Although the above examination may suggest the degree of backwardness in the statistical system at this stage, it should be noted that, compared to past performances, it nevertheless represents an improvement. First, the coverage of statistical reporting was extended to complete enumeration, though not at *hsiang* level. "Crop Reports," the most reliable statistics in pre-war days, were based on the reports of only about 60% *hsien* of the whole of old China excluding Man-

1 Hsieh Mu-ch'iao, "Make Big Effort to Improve Agricultural Statistics," *TCKT*, No. 22, 1957; M. V. Krishnappa et al., *Report of the Indian Delegation to China on Agricultural Planning and Techniques*, New Delhi, 1956, p. 88.

2 Hsieh Mu-ch'iao, "Pay Big Effort...."

3 M. V. Krishnappa et al., pp. 81–86.

4 "Directions for the National Statistical Activities during 1956, Proclaimed on the 20th of February, 1956," *T'ungchi Kungtso T'unghsin* (*TCKTTH*, Reports of Statistical Work), No. 5, 1956.

churia.[1] Besides, in this stage the cultivated areas were surveyed and measured by the unified scale unit which had long been locally diversified in pre-1953 days. The resulting map of holdings or farms made easier the summation of output statistics and, despite a primitive and inaccurate technique of estimating the yield per acre, made them relatively more accurate than before. The mapping of cultivated lands was done side by side with the land reform enforced during 1950–1952. The method, however, was applied in general with the aid of questionnaire observations from each farmer, not on a geographical base but on the base of farmers' holdings. Though it is stated that in the North-East District an original survey was attempted, the actual measurement of the farms by field workers was conducted only where a dispute arose due to conflicting interests between, for instance, tenant-farmers and landowners.[2] Consequently, some leakage in areas of arable land might have existed even after the land reform, though the statistics certainly improved as compared with former times.[3]

In regard to the typological sample survey, the exact date when it started has not yet been confirmed. The earliest known case of the typological sample survey was that of Ch'ang-shu *Hsien* in Kiangsu Province in 1953. At the 1956 National Symposium of Agricultural Statistical Work, Liu Jui-lung, the Vice-Minister of Agriculture, stated that in the preceding two or three years almost all Provinces had adopted the stratified, purposive sampling survey along with crop-cutting.[4]

The function of such a typological survey is to use it as a check on the reliability of the statistics obtained by complete enumeration. Therefore, its spread may also confirm a gradual improvement of the crop statistics based on complete enumeration.

2. *Post-Collectivization Period*

When the transformation of individual farmers into Agricultural Production Co-operatives was completed during the latter half of 1955,

1 Central Agricultural Laboratory, Ministry of Industry, *Crop Report*, Vol. 3, No. 8 1934, pp. 159 ff.

2 Motonosuke Amano, *Chūgoku no Tochi-Kaikaku* (Land Reform in China), Asian Economic Study Series No. 34, Tokyo, The Institute of Asian Economic Affairs, 1962, pp. 55–56, 65–66, 98.

3 Chou Hung, "Agricultural Statistics...."

4 Liu Jui-lung, "Reports at the Symposium on National Agricultural Statistical Activities," *TCKT*, No. 23, 1957. In 1955, Anhui and Chianghsi Provinces executed a sample survey of this sort. See Fang Ch'a, "Several New Schedules Executed for the Present Agricultural Statistical Work," *TCKT*, No. 8, 1957.

the method of crop statistics underwent a profound change. First, in 1956 periodic reporting on agricultural production was required by the State Statistical Bureau and down to the level of *hsien* government. The Bureau also issued to Provincial Governments "Basic Returns of Agricultural Production in Agricultural Production Co-operatives," which was to be used as experiment in 1956 in each Province.[1] Second, the Bureau issued to Provincial Governments a uniform method for the typological sampling survey under the name of "1956 Measure for Surveying Agricultural Crops," though this was to be enforced on an experimental basis.[2] However, the latter measure does not seem to have been successful. It was officially stated that in some regions the typological sampling survey was discontinued in 1956, since the government in these regions considered that the forthcoming adoption of the periodic reporting system on the basis of the Agricultural Production Co-operatives would make the typological sampling survey overlapping and wasteful.[3]

A post-collectivization system of compiling crop statistics was finally established as a result of the Sixth National Statistical Work Conference and the National Symposium of Agricultural Statistical Work held towards the end of 1957, the latter co-sponsored by the joint auspices of the Ministry of Agriculture and the State Statistical Bureau. According to the report made by Hsieh Mu-ch'iao, the new system consisted of the following two methods.[4]

(i) Periodic Reporting System. The State Statistical Bureau issued to the Provincial Governments "Basic Standard Forms of Statistical Returns of Agricultural Production in the Agricultural Production Co-operatives," which were to be revised by the Provincial Government according to the local conditions and to be sent to the *hsien* government for enforcement. Within 3 or 5 years, this reporting system was to be established on the basis of those unified forms of returns for all the Agricultural Production Co-operatives of the country.

(ii) Typological Sampling Survey. The importance of the typological sampling survey based on crop-cutting was emphasized as a supplement to the above system.[5] (Though no precise expression was found, it

1 Editorial of *TCKTTH*, "Earnestly Enforce the 1956 System of Various Statistical Reporting Tables on Agriculture," *TCKTTH*, No. 6, 1956.

2 Chou Hung, "Agricultural Statistics..."; Statistical Bureau, Szechuan Provincial Government, "Method of Stratification and Sampling in the Survey of Agricultural Crops," *TCKT*, No. 2, 1957.

3 Hsieh Mu-ch'iao, "Make Big Effort...."

4 Hsieh Mu-ch'iao, "Make Big Effort..."; *ditto*, "Preliminary Experiences of...."

appears that "every region" was supposed to enforce this survey strictly.[1] As regards the random sampling method, it was considered desirable for those provinces where satisfactory conditions existed to experiment with it gradually. For the time being the enforcement of typological sampling was supposed to continue.

As a comment it may be rightly said that the decision of the State Statistical Bureau to popularize the sampling method has been inspired by the success of the Indian National Sample Survey. The State Statistical Bureau's officials learned much from exchange of their experience with the Indian statisticians in 1956–1957, and the typological sample survey as adopted by the State Statistical Bureau was considered as a step to a scientific method of probability sampling.[2]

During 1958, when this new method was first enforced, there occurred ironically an incident of an erroneous report on the agricultural harvest. These erroneous figures were made public in April, 1959, and were revised in August, 1959. The April 1959 figures were identical with the estimates forecast in November, 1958. From this fact the

5 Hsieh Mu-ch'iao mentioned in the above-cited reports two defects of the periodic reporting system: (a) since it requires a lot of work for enumeration, the State Statistical Bureau cannot receive reports before the year-end and, in case where reporting is made on the basis of the year-end account of the Co-operatives, before the second quarter of the following year. Thereby its practical importance as reference material for drawing up plans is reduced; (b) there is underestimation of output not only through a leakage of output in the Co-operative member's "private plot" and the inaccuracy of the independent farmer's outputs but also through a fairly widespread practice of under-reporting on the side of the Co-operatives. As for the state of delay of the *Periodic Reports* from the Provinces, a responsible official of the State Statistical Bureau disclosed the following percentage figures:

	Delayed Reports	More than One Month's Delay	No Reports
1955	60.3%	(23.7%)	9.3%
1956	46.3%	(10.9%)	8.4%

See, Huang Chien-t'o, "Realistically Summarize Past Experiences and Make Efforts to Improve the Agricultural Statistical Activities," *TCKT*, No. 8, 1957.

1 Editorial of *TCKT*, "Pay Effort to Improve the Activities of Agricultural Crop Survey," *TCKT*, No. 4, 1958.

2 Dr. P. C. Mahalanobis and others visited China to give lectures on sampling survey (See Chu Cheng, "Professor Mahalanobis, a Famous Indian Statistician, Gives Lectures on Statistics in Our Country," *TCKT*, No. 15, 1957). An Acting Director of the State Statistical Bureau, Wang Szu-hua, went to India for inspection from December, 1956 to January, 1957, and on returning to China he advocated positively in his reports the adoption of the sampling survey methods. (Wang Szu-hua, "Introducing the Statistical Activities used in India and Recommending the Enforcement of the Sample Survey on a Nation-Wide Scale," *TCKT*, No. 6, 1957.)

writer infers that these first official figures may probably be the estimates obtained by the new purposive survey method that was designed in line with the suggestions of the Sixth National Statistical Conference. Moreover, the writer considers that the basis of the revised figures was complete enumeration. However, this does not imply that the purposive sample survey as such is inappropriate.

With respect to the causes of the erroneous reporting there are conceivably many sources. However, one should take note of the fact that 1958 was a year of extraordinary bumper crops in the midst of the People's Communication Movement. Cultivation techniques were also drastically changed, including the adoption of a close planting method and a multiple transplanting method as well as alteration of the crop rotation system.[1] It was quite natural that these changes resulted in the impossibility of precise crop forecasts and produced a marked discrepancy between the sown area and the harvested area. In addition, the lack of a land utilization survey aggravated the situation.

As regards land survey, it must be noted that the overall measurement of arable land was conducted in the process of the Collectivization Movement. A lot of cultivated land that had leaked in the questionnaire survey at the time of the land reform was registered under this measurement.[2] However, the land map, though improved, cannot be utilized as a land utilization map when the crop rotation system is undergoing a rapid change.[3]

1 Shigeru Ishikawa, "Chūgoku ni okeru Saikin no Nōgyōgijutsu-Kaikaku ni tsuite" (Latest Changes in Agricultural Techniques in China), *Ajia Keizai*, Vol. III, No. 1, 1962.

2 It was stated that out of the total increase of cultivated areas during the First Five-Year Plan period, about 198 million acres should have been attributed to the disclosure of past leakages before collectivization. Cf. Hsiao Yü, "Reclaim Waste Land to Enlarge Farmland," *Chihua Chingchi* (*CHCC*, Planned Economy), No. 2, 1958.

3 The cultivated areas of each crop which should have been provided for by a land utilization survey were very roughly estimated. For, in case of minor crops, especially when they are sown between the ridges of other crops or sown mixed, the sown areas were converted into figures comparable to other major crops using very rough conversion coefficients. (Wang Kuang-sên, "Preliminary Treatise on How to Define the Indicators of the Average Annual Output of Food-Grains per *Mou* and How to Calculate Them," *Tungchi Yenchiu* (Statistical Studies), No. 1, 1958, pp. 33–38, esp., p. 38; Huang Mêng-fan, "Agricultural Output Statistics," in Hsü Ch'ien et al., *Lectures on Economic Statistics, 1957*, Peking, The Statistical Publishing Company, pp. 111–122, esp., p. 113.) Meanwhile, the margin of error is likely to be small and at least constant, if the cultivation method as well as the rotation system remains unchanged. But, in the period of the Great Leap Forward in 1958 which brought about the expansion of the areas sown with rice and changes in the crop-rotation system, the areas sown with minor crops seem to have undergone a marked change and the errors of sown area might have been concentrated in minor crop areas. This inference may explain the

Despite the statistical disaster in 1958, the same sample survey was carried out again in 1959, and the survey in 1960 was to be carried out in a similar fashion.[1]

II. "TYPOLOGICAL" SAMPLE SURVEY

The analysis of the last section seems to imply that the crucial aspect of the development of China's crop statistics lies in the improvement of the typological sample survey. This section and the next, therefore, examine in further detail this survey method as it was enforced up to the summer and autumn harvest seasons of 1959. Although the statistical literature published in Mainland China appraised various aspects of the survey methods according to their practical use, emphasis in this section will be placed rather on an evaluation from a theoretical viewpoint. For this purpose, a theoretical rearrangement, and reinterpretations of the issues involved, are made according to the following framework:

(1) stratification, (2) sampling procedures, (3) differences between sown areas and harvested areas, (4) crop-cutting techniques, (5) field crop statistics and barn crop statistics.

Of these five phases, (1) and (2) are problems of sample design, while (3) is closely related not only to the frame of sampling, which will have a decisive rôle as weight in estimating total output, but also to the sample survey of land utilization, which is required to get an unbiased estimate. The last two problems, (4) and (5) mainly concern techniques of measuring output.

1. Main Features of Sample Design

The so-called "typological" or "typical" survey corresponds to a sort of purposive sampling survey, but its usage had not settled down before 1957. Thus Hsieh Mu-ch'iao stated in 1957 that "formerly, we did not fully recognize the difference between the typological survey and the 'random sample survey' and used the typological survey as a generic term for both of them... the [random] sample survey is the most scientific one among non-census surveys.... As regards crop survey, we should adopt as method the [random] sample survey side by side with crop-cutting."[2]

higher rates of error in minor crops and potatoes.

1 Tao Jan, "Concluding Report at the Forum on the National Crop Survey at Chengtu, December 25, 1958," *Chihua yü T'ungchi* (*CHTC*, Plan and Statistics), No. 11, 1959.

2 Hsieh Mu-ch'iao, "Preliminary Experiences of...."

However, the survey method which we encounter in the published documents since 1956 is mostly the typological sample survey called "*hua-lei hsientien,*" and the method of random sampling was only found in material published in 1959, which, however, was used on an experimental basis.

The typological survey adopted in China throughout the period in question consists of two processes; namely, the stratification of cultivated land into strata of similar productivity and the selection of the most representative samples from each stratum. As is easily seen, the method of stratification depends on what should be estimated, for example, the output of the whole country, or the output of each Province, *hsien* or *hsiang*. In this regard, a complicated situation existed in China. It was the output of the whole country in a short time-span that the central government wanted to estimate, and for this purpose it would have been desirable to stratify the whole country into several socio-geographical regions independent of the Provincial Government boundaries, as was done in J. Lossing Buck's survey, one of the most reliable surveys of pre-war days.[1] However, as Hsieh Mu-ch'iao admitted, the Provincial Government also required detailed regional data for the purpose of planning the production and procurement of agricultural products. This is the reason why post-war surveys were, in most cases, carried out within the Provincial Government boundaries. Attainment of these two objectives at the same time requires an allocation of many samples.[2] But both the limitation of survey costs and the shortage of statistical technicians restricted the scale and the design of the survey, i. e., the number of stages and the number of samples in each stage, if they wanted to carry out random sampling on a nation-wide scale.[3]

1 John Lossing Buck divided China into a wheat zone and a lowland rice zone. He subdivided them further into eight areas: spring wheat area, winter wheat and millet area, winter wheat and kaoliang area, Yangtze River rice and wheat area, rice and tea area, Szechwan rice area, double-cropping rice area and Hsinan rice area. Each region extended over several provinces and did not coincide with the boundaries of administrative divisions (*Land Utilization in China*, Shanghai, The Commercial Press, Ltd., 1937, Chap. II, pp. 23 ff.). Meanwhile, the appropriate stratification is to stratify the heterogeneous universe into several strata of normally distributed homogeneous populations and to make smaller the intra-strata variances. The more appropriate the stratification, the higher is the efficiency of estimates. In the case of China, J. L. Buck's survey showed that the variance of output per acre within an *hsien* was as large as that within provincial boundaries.

2 Hsieh Mu-ch'iao, "Preliminary Experiences of...."

3 In other words, under the given cost function of the survey, what is aimed at is to determine optimum allocation of samples so as to minimize the variance.

The purposive survey seems to have been enforced as a compromise between these conflicting ends.

An example of the same conflicting ends may be presented in the case of sample design for the First Farm Household Budget Survey which the State Statistical Bureau conducted in 1954 as a probability sample survey on a nation-wide scale. In this survey, each Province lists the clusters of farm households, *hsiang*, as the primary stage, and from this list the primary sample units were selected at equi-intervals. Then they were sub-sampled randomly according to a classification of the characters of the farmers; for example, Co-operative members or non-Co-operative members, size of the farm holdings, etc.[1]

One of the shortcomings ascribed to the State Statistical Bureau was that, with such a sample design, most of the *hsien* were allocated only one sample *hsiang*.[2] Though this is undeserved from the aim of the survey, it may be said that the State Statistical Bureau should have taken into account that the local government had to be informed of the output condition in the locality. For this purpose, they might have been able to allocate more samples to each *hsien* instead of allocating many secondary samples (for example, farm household) to the primary sampling unit. However, given the restricted cost and manpower conditions, this double aim would have been served only by adopting a representative sampling technique with multi-stage stratification although this makes an estimation of sample variance impossible and the correctness of stratification or efficiency of estimates cannot be inferable.

The biases caused by the Chinese methods will be discussed with concrete examples in the next section.

1 Agricultural Statistical Branch, State Statistical Bureau, "Several Basic Experiences from the National Farm Household Budget Survey," *TCKTTH*, No. 10, 1956. According to the examples of Hunan Province where the statistical field work has been active, 247 households were chosen out of 3,575 in 9 villages among 9 *hsiang* during 1952–1954. (This province had about 8,078 thousands of households in 1957. See Hunan Agricultural College, *Hunan Agriculture*, Peking, Higher Education Publishing Company, 1958, p. 55.) Since 1956, 510 households have been allocated to 34 master samples. See Committee for Compiling Economic Data, *Typological Rural Economic Surveys of Eight Provinces*, Peking, The Financial and Economic Publishing Company, 1957; Bureau of Statistics of Hunan Province, "We Developed and Utilized the Farm Household Budget Survey," *TCKT*, No. 1, 1958. On this survey, see Yoshirō Matsuda, "Chūgoku Nōgyō-seisan-tōkei no Ichi Gimmi" (A Review of Chinese Agricultural Production Statistics), in Shigeru Ishikawa (ed.), op. cit., esp., pp. 229–233.

2 Ho Kan, "Review of the Biases of the Farm Household Budget Survey," *TCKTTH*, No. 18, 1956.

2. Selected Examples of the Typological Survey

(1) Survey Methods in Ch'ang-shu *Hsien* from 1953 to 1956[1]

Examples of Ch'ang-shu *Hsien* in Kiangsu Province are important because the changes in survey method can be documented continuously for the period from 1953 to 1956.

(i) From 1953 up to 1955, when Co-operativization was completed, stratification for typological sampling in Ch'ang-shu *Hsien* was based mainly on information concerning natural conditions, farming methods and crop distribution, with supplementary consideration given to the data of the annual output series and the normal output assessed.

In 1953, the cultivated areas were stratified according to the following districts delineated by crop patterns.

Summer Crop Season	Autumn Crop Season
Tung Hsiang Cotton District	The same as for summer
Hsi Hsiang Rice and Cotton District	The same as for summer
Rice Districts:	The same but High-land District subdivided into:
Low-land District with Early-ripening Rice	High Productivity District
High-land District with Late-ripening Rice	Middle Productivity District
	Low Productivity District
4 Districts	6 Districts

In 1954, *hsiang* with a similar crop pattern and similar productivity levels were regrouped to form strata regardless of whether they belonged to different zones as mentioned above. The underlying hypothesis seems that the average level of productivity would be approximated by that of the representative *hsiang*, since the yields of *hsiang* within the same stratum are normally distributed.

The sampling process was as follows: *hsiang* within the same stratum were classified into three grades of high, middle, and low productivity and a representative *hsiang* was drawn from the middle class *hsiang*. Before choosing such a representative *hsiang*, the crops which would be surveyed in it were determined. The natural villages "*ts'un*" in this representative *hsiang*, which is the administrative village, were also classified into three grades—high, middle, and low—according to a forecast of their output level. Then a representative village was chosen from the middle class villages. In such a representative village those fields, where crops to be surveyed are cultivated, were classified

1 Division of Statistics of Ch'ang-shu *Hsien* in Kiangsu Province, "Method and Experiments for the Estimation of Agricultural Outputs in Ch'ang-shu *Hsien*," *TCKT*, Nos. 3-4, 1958.

according to the forecast productivity level into three grades : high, middle, and low. Crop-cutting was done experimentally in several typical "*t'ien-k'uai*" or plots of cultivated land selected from the middle class fields.

(ii) Since 1956, the frame for sampling consisted of a classified list of Agricultural Production Co-operatives. The stratification was made as follows according to the dual criteria, crop patterns, and forecast productivity levels.

1957 Summer Crop Season

Cotton District	*Tung Hsiang* District	(No subdivision is made according
	Hsi Hsiang District	to the rate of forecast output.)
Rice District	High Productivity District	(Without regard to natural condi-
	Middle Productivity District	tions, a subdivision is made accord-
	Low Productivity District	ing to the rate of forecast output.)

5 Districts

In each of the above districts, one Co-operative whose productivity level was considered to be close to the forecast average yields of the unit area was selected as representative Co-operative. From such a Co-operative sample, Production Teams were selected either from each of the three or four sub-classified Production Teams—when the difference of productivity between Production Teams was large—or from the non-sub-classified Production Teams, taking into consideration the natural conditions, crop varieties usually cultivated, and the degree of progress of Team activities.

Curiously, however, the crop-cutting survey was made with respect not only to two or three representative *t'ien-k'uai* selected from each sample Production Teams but also with respect to one or two *t'ien-k'uai* from non-sample Production Teams.

Conclusions derived from these methods could be summarized as follows : (a) the smaller the basis unit of stratification, the more accurate is the result of the estimates ; (b) the farming methods were so much affected by topographical complexities that the actual productivity levels vary greatly among those samples for which the forecasted productivity levels were similar ; (c) too early forecasting of the produce sometimes produced errors—less than half a month before harvesting was found most suitable ; (d) selection of representative *t'ien-k'uai* was adequate only from the middle class *t'ien-k'uai*, for the samples from lower or upper class *t'ien-k'uai* are affected by the existence of extreme cases.

(2) Survey Methods Used in Szechwan Province during 1956[1]

1 Bureau of Statistics of Szechwan Province, "Techniques of the Typological Sampling Survey of Agricultural Outputs," *TCKT*, No. 2, 1957.

The survey methods used in Szechwan Province during 1956 can also be documented as regards to processes, stratification, and sampling.

Stratification Methods

The procedures of stratification had three variants.

According to the first variant, each of the *hsien* was stratified into several natural districts; for example, the plain districts, the high-land districts and the mountain districts according to the natural topography, or alternatively the food-grains cultivation zone and the industrial crops cultivation zone according to crop distribution.

The second variant was based on productivity without regard to the boundaries of the topographical regions. The productivity was rated according to the forecast productivity level of major crops (such as ordinary ripening rice or maize) of 1956 and that of the preceding year. The distribution of crops and the labour employed at the harvest season were also considered as supplementary data. Thus the ordering of all *hsiang* was made and then grouped into several strata.

The third is a combination of the former methods. First, the districts were put into a frame of sampling as in the first method and then within each district *hsiang* were rated as in the second method.

Only one representative *hsiang* was selected from each group in every variant. But each variant appears to have had its own defects. The strata based on the topographical conditions used in the first variant contained areas of heterogeneous productivities that could not be represented by only one representative *hsiang*.

Though the second variant overcame that defect, it still possessed two defects. First, the forecast productivity levels were less reliable as stratification was based on eye-surveys concerning the state of shoots, maturity of ears, irrigation conditions, and so on. Those eye-surveys would be influenced by the observer's subjective attitude, and the results cannot be compared with each other as the surveys were carried out by different observers in vast areas. Second, the ordering of the forecast output levels with no regard to natural conditions becomes quite different from that of the actually harvested output levels, for the small variance of forecast figures of each stratum does not guarantee a small variance of realized figures in heterogeneous natural conditions. The third variant was free from these defects but required high costs and its application was recommended only for areas with complicated natural conditions.

Sampling Procedures

Sampling procedures consist of two stages: (a) selecting first a

representative · *hsiang* from each stratum and then a representative Co-operative from this *hsiang* and (b) selecting a representative *t'ien-k'uai* for crop-cutting experiments. As regards the selection methods of a representative Co-operative from the selected *hsiang*, two variants were mentioned: one was of a subjective nature, taking account of the natural conditions; and the other was more objective, selecting the Co-operative which seemed to be close to the average yield of similar output levels, or the median of the whole harvested area.

The selection of the plot for crop-cutting, the second stage of the sampling procedures, was said to be made according to two methods

Table 1. SCALE OF THE SAMPLE SURVEY IN 1959

	Survey during Autumn Crop-Season in 25 Provinces and Autonomous Districts in the Whole Country	Shantung Province	Anhwei Province		Heilungkiang Province	Yünnan Province
			Summer Crop Season	Autumn Crop Season		
Total Number of Persons mobilized for the Survey	16,000,000		300,000	440,000	16,000,000 (Including Experienced Farmers)	56,430
Field Workers belonging to Administrative Organs Higher than Those of *Hsien* Level	50,000	1,300	51,106	68,858	8,000	1,261
Places Where Field Surveys were Carried out	8,360 (People's Commune) 30,255 (Production Brigade)	374 (Places) 794 (1.6% of Whole Province)	247 (People's Communes) 545 (Production Brigade)	947 (90% of All Communes)	55 (People's Commune) 4,500 (Production Teams)	
Basic Enumeration Unit	74,180 (Production Teams)		1,804 (Production Teams)	5,696	1,800 (Management Blocks)	(Directly surveyed by the Province's Staff—1,098 Cases)
Sample Allocated Areas [to be Crop-cut]	3,300,000 (*k'uai*) 4,000,000 (*mou*)	154,000 (*mou*) [1,018,000 acres]	wheat—6,707 (*k'uai*) main crops—[178,012 acres]	26,923 (*mou*)	67,000 (*k'uai*) 200,000 (*tuan*)	(Directly surveyed by the *hsien's* Staff —12,001 Cases)
(Directly surveyed by *Hsien* or Upper Organs' Staff)			8,776	51,612		
(Directly surveyed by Commune's Staff)			18,916	77,014		
(Directly surveyed by Production Brigade Staff)			64,333	165,130 [0.14% of Total Area was Crop-cut]		

similar to the first stage.

(3) 1959 Survey at the Time of the Autumn Crop Season[1]

The scale of the 1959 typological survey was unprecedentedly large, and 28 Provinces and Autonomous Districts were reported to have executed the sample survey. Table 1 indicates the scale of the (typological) sample survey in 1959.[2]

Of these 28 Provinces and Autonomous Districts, documents of the survey methods are available only for the Provinces of Shantung, Anhui, Heilungkiang, and Yünnan and for the Chuangtsu Autonomous District in Kwangsi Province. The surveys in these Provinces had two common characteristics.

(i) The survey consisted of three stages: a 'preliminary forecast of output', 'typological sampling' and 'crop-cutting experiments.' With regard to the forecast, the importance of checking it by eye-survey methods was stressed.

(ii) There were many cases where the non-subjective sampling methods, such as 'equi-interval sampling' and 'randomized sampling,' were experimented with. As a result, the number of samples in each stratum increased from one representative sample in the previous survey to several samples.[3]

Forecast of Output

This preliminary survey aimed at obtaining tentative figures of output before harvest so as to use them as data for the planning of government purchases and for distribution of farm products as well as for classification criteria of typological sampling. For this purpose, field workers carried out the investigation in co-operation with "experienced old farmers" by way of either eye-survey, or by testing some ears of the crops by hand, or a combination of both. As can easily be seen, these estimations are not so accurate in character. In Anhui Province, for example, the forecast by these methods was said to have led to an overestimation of the output harvested.[4] On the other hand,

1 Agricultural Statistical Branch, State Statistical Bureau, *Experiences of the Agricultural Output Survey*, Peking, The Statistical Publishing Company, 1960 and Hsü Hsiang-hsin, "Reconsiderations of the Agricultural Output Survey—Information from the National Conference on the Agricultural Output Survey," *CHTC*, No. 11, 1959.

2 The figures of each province were collected from scattered explanations, but the units of measure used were so divergent that the present writer has reproduced them without any conversion.

3 See also Wu Kuang-t'ang, "Experiences of the Agricultural Output Survey," *CHTC*, No. 11, 1959.

4 For example, the exaggeration of the forecasts amounted to 1.3% at 11 representative plots, 58 *mou* in area, which were selected out of 5 Production Brigades of 5 Communes

however, there is a report stating that in Shansi Province during the 1958 summer crop season an experienced surveyor committed an underestimation, due to a changed situation caused by the introduction of close planting techniques.

Typological Sampling

The main issue concerning stratification after 1958 was whether the cultivated area to be stratified is to be within or beyond the Commune boundary. For example, in Heilungkiang Province in 1959, each Commune was taken as a base of stratification, while in the Chuangtsu Autonomous District of Kwangsi Province, in the Yünnan and Shantung Provinces, both a Commune base and a base regardless of the Commune boundaries were experimented with. About Anhui Province no indications were made in this respect.

In Anhui Province, the cultivated areas in thc Hwaipei Plain—for example, Fuyang District—are fairly homogeneous in productivity, so that they were listed in the order of forecast output levels, and the sample plots were selected at equi-intervals. In southern districts with diversified cultivating conditions due to complicated geographical features, they were classified in different ways (i. e., according to geographical features—mountainous lands, lowlands near rivers and plains; according to crop-growing status—high, middle, and low; according to geological features—black, yellow, sandy soil and white clay) and then one representative plot was selected after an open discussion by farmers.

Equi-interval selection was so complicated that it could not be used by untrained people, and the forecast of the output level was not only inaccurate but also went through remarkable changes in the process. The latter representative selection technique is reported to have been dictated by subjective bias. For instance, the equi-interval selection estimation had a 99.5% accuracy in the Ch'ang-ch'i Production Brigade of Chintu Commune while representative selection estimation had an 88% accuracy in Paiyang Commune. Thus the difference in accuracy between these two values amounts to 10%, though they could not be compared because the plots were not the same.[1]

In Heilungkiang Province, 19 *hsien* selected 55 People's Communes

in Fuyang *Hsien* (Agricultural Statistical Branch, the State Statistical Bureau, *Experiences of...*, p. 16).

Forecast Yield per Acre (catty/*mou*) 194
Crop-cutting Area's Output* (catty/*mou*) 191.5
* Some of the areas which were used for crop-cutting experiments were not excluded, so that this figure might have been lower than the actual production.

1 Agricultural Statistical Branch, the State Statistical Bureau, *Experiences of...*, p. 18.

which were considered as representative.

(i) Most Communes first investigated the average yields of the 'Administrative Blocks'—corresponding to the 'Production Brigade' in most other Provinces—and classified them into 3 or 5 groups accordingly; then from each group were selected 3 or 5 representative Production Teams of high, middle, and low productivity, from which 3 to 7 representative *k'uai* or plots were selected, in turn, for the crop-cutting experiments.

(ii) A few Communes selected 3 middle-class Administrative Blocks and made them the typical Administrative Blocks for crop-cutting experiments. Following the same method as indicated in (i), they selected from these Blocks one to three Production Brigades, and then from each three to five typical *k'uai* and 3 to 5 *tuan* (subdivided plots of farm land).

(iii) The third method consists in that after Administrative Blocks were selected in accordance with Method (i), the areas of each crop are listed without regard to the jurisdiction boundaries of Production Brigades, and three to five *k'uai* or plots of upper, middle, and lower classes of each crop were selected, and then from these *k'uai* typical *tuan* were selected.[1]

In the Chuangtsu Autonomous District of Kwangsi Province, two methods were adopted.

(i) Taking into account the natural conditions (such as geological, topographical, and hydrological conditions), People's Communes throughout the District were classified into three large groups: plains, highlands and mountainous areas, and from each group a Commune whose rate of forecast output came close to the average rate of each group

1 Examples illustrating the accuracy of the three methods are shown below.

Methods	Estimates (ton)	Actual Output (ton)	Rate of Error (%)	Place
(i)	34,365	35,647	−4	Fengtung Commune in Chingan *Hsien*
(ii)	2,400	4,000[1)]	−67	Changan Commune in Fuchin District[2)]
(iii)	n.a.	n.a.	20	Shangchia Commune in Chaotung *Hsien*[3)]

Notes: 1) Average of the People's Commune.
2) Output of maize per *shang* of one *k'uai*.
[1 *shang* equals 15 *mou*, and 1 *k'uai* equals 3 *tuan*.]
3) 4 Production Parties.

Source: Agricultural Statistical Branch, the State Statistical Bureau, *Experiences of* ..., pp. 41–42.

was selected as a typical Commune. The Production Brigades in this Commune were classified into upper, middle, and lower classes according to natural conditions as well as to the rate of forecast output. From each class a representative Production Brigade was selected; furthermore, from this Production Brigade representative Production Teams were selected. The cultivated land of this Teams was divided into three homogeneous groups of "*li*" (strips of land) with similar rates of output and a typical *li* from each *li* was selected. From each *li*, upper, middle, and lower representative *k'uai* (thus totalling nine each in a small *li*) were selected for crop-cutting.

(ii) Within the *hsien* boundaries every Production Brigade was listed according to its rate of forecast output without being restricted to the jurisdiction boundaries of the People's Communes. Then it was classified into several strata with similar rates of forecast output and the number of strata was decided by the degree of complexity of the natural conditions in a *hsien*. This method compared with the first method was considered as being of superior accuracy.

(iii) Besides these two methods, one more, equi-interval sampling was experimented with in Liucheng *Hsien*. This sampling method was carried out as follows.

Every Production Brigade was listed according to its forecast output level, and then sample plots were selected at equal intervals from the cumulated list of areas in order of productivity. Chinese statisticians stated that the selected Production Brigades were regarded as representative when the margin of error of the weighted average of their forecast outputs to the average forecast output level of total *hsien* was less than 3%. (This procedure seems doubtful because the accuracy of representativeness was measured using what was to be measured, i. e., the forecast output in the *hsien*.)

The plot or *t'ien-k'uai* to be crop-cut was selected in principle by a method similar to the former. Liucheng *Hsien* is, however, so complicated in its geographical features and contains so many small *t'ien-k'uai*, for example 100 to 1,000 in one Production Team, that enough data for forecasting the normal output level were not available for every *t'ien-k'uai* and, therefore, at this stage representative selection was adopted in some cases which went beyond the jurisdiction boundaries of the Production Teams.

The results of these experiments were as follows: Method (i), or the selection of the representative Commune, showed the largest error: 5.6%; Method (ii), or representative selection without regard to the

Communal boundarics, showed an error of 0.2%, and the third Method (iii), or equi-interval sampling, 0.6%.[1]

From these reports the advantage of stratifying the areas without restriction of Communal boundaries seems to have been verified. In fact, in Shantung Province it was officially recommended that the area be stratified without restriction of Communal boundaries because, when a small number of Communes were regarded as representative in case of a specified crop, they often lost their representativeness in other crops.

Crop-cutting Experiments

Actual crop-cutting was carried out by three methods: (a) Crop-cutting of a small plot; (b) Crop-cutting of ridges chosen at equal distances; (c) Reaping several 'hills.' When the field workers used Method (a) for a large cultivated area in a plain, the sample plot to be crop-cut could be reduced to a small area selected at random, but in mountainous areas with complex geographical features the allocated sample area should have been wholly crop-cut. Method (b) saves without losing accuracy, but it is applicable only to crops sown on ridges such as wheat, maize, etc. The computation of average yield

Table 2. CROP-CUTTING SURVEY OF RICE IN KWANGTUNG PROVINCE

Districts	Estimates by Crop-cutting Survey (A)		Actual Results (B)		Ratio of Actual Results to Estimates (B/A)	
	Output per *mou**	Total Output (10,000 *catty*)**	Output per *mou* (*catty*)	Total Output (10,000 *catty*)	Output per *mou* (%)	Total Output (%)
Ts'ung-hua	220	8,821	213	8,562	96.8	97.06
Nankai	249	10,341	247	10,863	99.2	99.8
Hsingning	378	18,216	380	16,270	99.5	99.66
Ch'aoyang	381.2	17,074	373.5	16,730	97.88	98.84

* 1 *mou* is about 0.067 hectare.

** 1 *catty* is 0.5 kg.

Source: Agricultural Statistical Branch, State Statistical Bureau, *Experiences of*..., p. 60.

1 The three methods were applied to the late-ripening rice harvested in autumn.

	Output per Unit (*catty* per *mou*)	Rate of Error (%)
Typological sampling (i)	442	5.6
Typological sampling (ii)	469	0.2
Random sampling (iii)	471	0.6
Actual output	468	

Source: Agricultural Statistical Branch, the State Statistical Bureau, *Experiences of* ..., pp. 30 ff.

per acre involved in this method is cumbersome.

The accuracy of crop-cutting is illustrated in Table 2.

A characteristic of this example is that the figures estimated by crop-cutting were overestimated at every place (this is contary to the example by eye-survey used experimentally in Shanshi Province during the 1959-summer crop season shown above). Furthermore, the fact that the ratios of actual results to estimate differ, though only slightly, per *mou* and also for the total output indicates that there may be some bias in the area used as weight (or multipliers to the output per *mou*) in estimating the total output.[1]

1 The area to be crop-cut is shown in Table 1, but it is not certain whether the area was the crop-cut area or total sample allocated area. Prof. M. Umemura comments that from his experience in Japan, the area was so large as crop-cutting areas that it seems to correspond to the total sample allocated area. Dr. Y. Tsumura, formerly in charge of framing the sample design for Japanese agricultural output statistics at the Division of Statistics, Ministry of Agriculture and Forestry, suggests that in developing countries the crop-cutting area usually becomes wider due to the low level of statisticians' skill and knowledge, therefore, there still remains the possibility that the above-mentioned area was the area to be crop-out. This problem should be left for future study.