

**“BOUND TO BE UNCERTAIN”: A NOTE ON
PROBLEMS OF ESTIMATING WET PADI YIELDS IN
PENINSULAR MALAYA DURING THE DEPRESSION,
1929-33**

by
A. Azmi Abdul Khalid

... the statistical data are by no means reliable. It is important to bear this in mind. In general it is our belief that the data are adequate for the purposes for which we have used them. But the possibility cannot be ruled out that future data revisions will alter some of our conclusions qualitatively, not just in detail. This possibility exists with all types of evidence. (R.C.O. Matthews, C.H. Feinstein, and J.C. Odling - Smee, *British Economic Growth 1856-1973* (Oxford U. Press, 1982), p. 8).

Although written for a different purpose, the observation by Matthews, Feinstein and Odling-Smee has a remembrance of universal validity. Repositioned in the context of Malaysian economic history, the statement comes to life particularly as regards wet-padi statistics pertaining to yields.¹ Statistics concerning wet-padi yields before the Second World War have been notoriously problematic. Consequently, it takes great effort to make some sense out of the maze of yield figures, often contradictory in nature. The purpose of this note is to look at the problems involved and also the ensuing steps taken by the colonial state to correct the discrepancies of wet-padi yields in the Malay States, especially during the Depression.

¹In this note, yield figures refer to yield per planted acre (as opposed to yield per harvested acre). According to Lim Chong Yah, "... the pre-1948 data are in fact per planted acre". (Lim Chong Yah, *Economic Development of Modern Malaya* [Kuala Lumpur: Oxford University Press, 1967], 149. This view appears to be confirmed by one official who wrote: "... while the figure relative to yields refers to yields from planting...." (District Office Sitiawan File (henceforth, DOS) 201/1932. Minute dated 13.4.32.

It is not clear what is the earliest date of published figures on yield production of wet rice in Malaya. As it stands there are two conflicting opinions. The Rice Cultivation Committee (RCC) noted that prior to 1915 no statistical data on yields or production were available for Malaya.² On the other hand Lim Chong Yah suggests that Malaya did not have production (presumably yield) and acreage statistics well until 1922.³ Furthermore, both sources indicate the unreliability of the early statistics on padi/rice production. Whilst the actual date of the first published rice yield statistics for Malaya is still uncertain, it is pertinent to note that in 1913, for instance, yield estimates for the Federated Malay States (FMS) are available.⁴ Again no specific reference is mentioned as to whether Malaya refers to the FMS and the Unfederated Malay States (UMS) or whether it also includes Penang (and Province Wellesley) as well as Malacca.

Whether 1915 or 1922 is tentatively accepted, the interim period until the beginning of the Depression in 1929 is still clouded by the understatement of rice yield figures. This prompted Lim to remark that as the quality of statistical data appears to improve over time, "there is ample evidence that statistics from 1932 onwards are much more dependable than the pre-1932 figures".⁵ Although this statement has yet to be tested rigorously, there were serious attempts made by the colonial government in Malaya to deal with the problems of inaccuracies in rice yield statistics, particularly after the RCC Reports were made public in 1931. The RCC was established and appointed by the High Commissioner on 14th July 1930 to investigate "what are the best steps to be taken in order to encourage rice cultivation in Malaya"⁶ One of the important reasons for the setting up of the Committee was the beginning of the Depression in the heavily export-dependent Malayan economy. The rationale was that in

²Report of the Rice Cultivation Committee (henceforth, RRCC), Volume 1 (Kuala Lumpur: 1931), 17.

³Lim Chong Yah, *Economic Development of Modern Malaya*, 149.

⁴See, for example, I. Lewton-Brain, *Agriculture in Malaya in 1913* (Kuala Lumpur: Department of Agriculture 1914), Bulletin No. 20, 22.

⁵Lim Chong Yah, *Economic Development of Modern Malaya*, 149.

⁶RRCC, Volume I, 13. For a critical appraisal of the RCC's findings, see Lim Teck Ghee, *Peasants And Their Agricultural Economy In Colonial Malaya 1874-1941* (Kuala Lumpur: Oxford University Press, 1977), 182-187.

order to lessen Malaya's dependence on rice imports⁷ and to alleviate the perceived and anticipated hardship of Malay peasants during the course of the Depression, there was an urgent need to revitalize the rice economy. Among the various recommendations contained in the RCC Reports, the problem of the statistical unreliability of rice yields was mentioned. This aspect was put in perspective by the Report:

As a most important step towards the improvement and extension of rice cultivation in Malaya, better provision than at present exists for the collection of statistics both in relation to areas and yields is essential; without this, progress is bound to be uncertain in as much as it is impossible to accurately gauge the effects of measures for improvements which are undertaken.⁸

The urgency of finding a more reliable method of estimating rice yields presumes a high degree of unreliability of yield figures before 1931. It could also be due to the need to coordinate methods of statistical compilation for rice yields in every state in Malaya.

But what were the problems associated with the unreliability of yield figures? One problem was that there was no synchronization between acreage and yield figures in relation to a particular year. To illustrate this it would be relevant to note the comment by an official in the district of Sitiawan in the state of Perak:

In the annual returns for 1931, the figure for acreage relates to the area under padi crops as on 1.1.32 while the figure for yield refers to the total yield from planting done during the latter part of 1930⁹

This means that in practice the yield statistics were behind time by at least a year. It prompted the same official to stress that "the yields shown in the return does not represent the yields from the area shown in the same return".¹⁰ A more serious problem was rooted in the absence of a standard or uniform method of

⁷For more details on this especially with reference to the FMS, see A. Azmi Abdul Khalid, "The Federated Malay States: Growth and Development of the Import Trade, 1900-1940", *HISTORIA* (Kuala Lumpur: 1984), 334-337.

⁸RRCC, Volume I, 18.

⁹Minute by an official, 13.4.32, DOS 201/1932.

¹⁰*Ibid.*

estimating the rice yields from the sawah in particular during the *menuai* (harvesting) seasons. In practice there were two distinct methods used: i) by focussing on sample plots and ii) by relying on the estimates of the *ketua kampung* (village headman) who relayed them through the *penghulu* (head of a *mukim* or sub-district) to the Land Offices.¹¹ Out of the two methods, the first one was considered more reliable and practised in Province Wellesley and Malacca. Many complaints were made regarding the second method. For example, in Johore where rice cultivation was insignificant because of the dominance of rubber growing, one official observed that:

Whether these [yield] figures are based on personal knowledge of cultivated returns or are only matter of guesswork on the part of the Penghulus, I am unable to say.¹²

At a higher administrative level, the dissatisfaction with the existing methods increased markedly. F.M. South remarked that it was difficult and confusing to match or reconcile the yield estimates compiled by the various Land Offices with those submitted by the Agricultural Field Officers.¹³

These conflicting yield estimates urged the RCC to recommend the method of estimation based on sample plots. It proposed that the system should be tried out in "all padi areas with the least possible delay."¹⁴ The Department of Agriculture (FMS and Straits Settlements) had tried to find out a suitable method of sample plot measurements based on a technique devised by Hubback which was used in India with some measure of success. Caution in adopting the sample plot method was voiced by F.W. South. In response to the RCC's recommendation he emphasized that the method should be carefully thought out before it was used. Otherwise, according to South, "grave errors" could appear as was the case in previous trials. Summing up his views,

¹¹It is interesting to note that an almost similar practice was followed in Thailand - see James C. Ingram, *Economic Change In Thailand 1850-1970* (Kuala Lumpur: Oxford University Press, 1971), 240.

¹²Minute by F.R. Mason, Acting Principal Agricultural Officer to the General Adviser Johore, 16.9.93, Commissioner of Land Revenue, Segamat (henceforth, CLRS) 539/1933. See also General Adviser Johore 543/1933 and Commissioner of Land Revenue, Johore Baru (henceforth CLRIB) 408/1033.

¹³F.W. South, Acting Director-General of Agriculture, Straits Settlements and FMS to the Resident, Perak, 27.1.32, DOS 201/1932.

¹⁴RRCC, Volume I, 19.

he concluded that "it seems very desirable to make the best of present methods" before a better and more reliable method could be implemented.¹⁵

The search for the most reliable and practical method of yield estimation continued well throughout the Depression. However, an important and thought-provoking memorandum by W.N. Sands appeared to guide the search to a compromise method. In a memorandum to the Resident of Perak, Sands reported that he had tested four varying methods of sample plot estimations viz. the Hubback method (with triangular 1/45000 acre plots), the 1/360 acre square plot method, the 1/40 acre square plot method and the Kedah method (1/45 acre oblong plot method). From these methods, Sands chose the Kedah method. It began with the random selection of various plots with an area of 1/32nd *relong* (44 ft. long and 22 ft. wide) or 1/45th of an acre. These plots were chosen randomly on the basis of "five widely scattered plots in each mukim wherever possible" and assuming that in each *mukim*, the planted area averaged out to about 1,500 acres. Due to many strains of padi grown with different individual characteristics, sample plots were carved out of the portions with the most extensive acreage. In practice, "two to five different varieties were sampled in each *mukim*". The yield were then averaged. In order to calculate the total yield of a *mukim* on the basis of the yields from the sample plots, the total unharvested area was excluded from the total planted area. These unharvested areas could be due to damage by pests, drought or floods. Next, Sands used the following method of calculation: the "mean yields of the five plots multiplied by 32 and less 10% represented the calculated yield per *relong*, which figure multiplied by the net area harvested, gave the total calculated yield of the mukim."¹⁶

As neatly as Sands had presented the formula based on the Kedah system, there were also attendant difficulties. It would appear that manpower was one of the constraints; in the case of Kedah, all the field officers were fully occupied during

¹⁵F.W. South, Acting Director-General of Agriculture, Straits Settlements and FMS to the Resident, Perak, 27.1.32, DOS 201/1932.

¹⁶Memorandum Concerning Experimental Work For The Estimates of Padi Yields by W.N. Sands (henceforth, Memo. by Sands), Principal Agricultural Officer, Kedah, 15th July 1933, CLRJB 408/1033. As regards the Kedah System of Estimating Yields (a part of the memorandum above), see also CLRS 539/1933.

harvest time and sometimes “they had to enlist the help of *penghulus* in order to measure yields for them.”¹⁷ The method could be impractical in areas where padi areas were small and scattered as was the case in Johore. This contrasted with the large and often continuous padi areas like those in Kedah.¹⁸ In fact, the deviation from the past practice of yield estimation inspired one official to state that “these new methods are all very bewildering”.¹⁹

But how reliable was the Kedah method? T.S. Adams who was the Adviser of Lands in Kedah thought that by late 1930 “yields were definitely under estimated.”²⁰ When Sands assumed work as the Principal Agricultural Officer for Kedah in the same year, he found out that padi yields recorded by the *penghulus* “were much lower than those actually obtained.”²¹ This observation is interesting because what Sands had implicitly hinted was the existence or awareness of another method which was used to counter - check the estimates given by the *penghulu*. But unfortunately no mention was made regarding the results that were “actually obtained”. In view of this it can be said that the Kedah system was initially widely used only from 1930.²² In his memorandum Sands clearly stated that sample plot measurements were conducted annually as follows:²³

1930 :	31,480 sample plots in 96 mukims
1931 :	32,578 sample plots in 119 mukims
1932 :	33,555 sample plots in 133 mukims

No information was supplied before 1930. As the Kedah system was gradually extended in the state, there were more opportunities to test the results obtained through this method.

¹⁷Memo. by Sands, 6.

¹⁸Minute by F.R. Mason, 16.9.1933, CLRS 539/1933.

¹⁹Minute by G. Hawkins, 23.9.33, CLRS 539/1933.

²⁰RRCC, Volume II, 113.

²¹Memo. by Sands, 4.

²²Before the Kedah method was highlighted and widely discussed in official circles, the reference was made to the method as practiced in Melaka. By this latter method, yield estimates were made by Padi Inspectors in the Department of Agriculture who reaped some sample plots in the *sawah* (padi fields). H.A. Tempany, who was the Chairman of the RCC stated that “this is the most satisfactory system at present capable of being devised and its extension on a uniform basis throughout Malaya is strongly advocated” (RRCC, Volume 2, 8).

²³Memo by Sands, 5.

To illustrate this, Sands compared his results with those obtained from the *penghulus* between 1930-31 to 1932-33. It was discovered that the margin results showed 15% under estimation by the *penghulus* in 1930-31 and decreased significantly to 8% and 2% in 1931-32 and 1932-33 respectively. The larger divergences were traced to valleys and inland areas for example in the districts of Kulim and Bandar Bahru where there were many small and scattered padi fields. In the plains, covering the Kota Star and Kubang Pasu districts, the results "nearly corresponded with our own."²⁴ This range of margins appears to confirm Mason's earlier point that the Kedah system cannot be efficiently and successfully operated in areas where padi plots were widely dispersed.

It is apparent that Sands was sensitive to the problems encountered by the Kedah method. In fact he stressed that in order to arrive at a more plausible yield estimate it is required that a thorough scrutiny should be made about a district's crop reports, the estimates by *penghulus* and the sample plot figures. Sands doubted that "the results can be dealt with successfully by the usual statistical methods because the necessary requisites for reliability are absent." There were other important variables like the variations in the ecological setting of padi areas which could alter the purely statistical estimates.²⁵ Eventually what prevailed was the Kedah method of sample plot measurement which also took into account the *penghulu's* returns as well as other variables which influenced the padi crop at the time of harvesting. In September 1933 Tempany recommended that the Kedah method should be tried out "throughout the padi growing areas of Malaya."²⁶

In retrospect, the desire to create a uniform method of yield measurement was justified in the context of the RCC's larger task of persuading the predominantly Malay peasants to grow more rice. Yield figures could throw light on the changes in productivity and hence in this connexion, it would be more meaningful if the data base for rice yields is founded on more

²⁴Ibid., 6.

²⁵Ibid., 7.

²⁶H. Tempany, Director of Agriculture, S.S. and FMS to the Secretary to the High Commissioner for the Malay States, 4th September 1933, CLRS 539/1933.

organized and systematic yield statistics. There could be some truth in Lim's earlier comment that post - 1932 yield figures contain less discrepancies as compared to the preceding years, especially when it is viewed in the context of the official endorsement of the Kedah method in 1933.

However there is another dimension to the underreporting of rice yields. It was claimed that the main reason for the persistent under estimates before 1930 was the intention, among padi cultivators, to eschew the payment of more taxes to the colonial state.²⁷ This viewpoint is, at best, tentative. For it to be empirically confirmed or refuted, more research has to be undertaken in the principal padi growing areas before 1930. On the contrary it is probable that the peasants' underreporting was actually based on their own version of measurement or estimation grounded on the different cultural foundation and world-view. Furthermore, the problem of underreporting was not peculiar to Malay peasants alone. In late 19th century Japan it was also rampant, especially in the early Meiji period.²⁸ The fact of underreporting of rice yields in Malaya before 1930 should pose a challenge to historians and social scientists intending to infuse quantitative techniques into history to the point of reconstructing the yield figures by proper statistical methods. Perhaps, as Nakamura has done for 19th century Japan, a start can be made by constructing a padi rice yield index and testing out the revised yield estimates.²⁹ Whatever way one sets out to perform the actual task, it is well worth remembering that statistical methods in history are only useful tools of analysis which could be applied to clarify certain doubts within certain limits. It is indeed instructive to end this note by echoing Ohlin's polite end remark:

Statistical estimates are not 'facts' of history, first to be ascertained and then to be interpreted; the estimation of historical statistics is merely an aspect of historical interpretation.³⁰

²⁷Memo. by Sands, 4.

²⁸James I. Nakamura, *Agricultural Production And The Economic Development of Japan 1873-19122* (Princeton: Princeton University Press, 1966), Chapter 4.

²⁹Ibid., 87-104.

³⁰G. Ohlin, "No Safety in Numbers: Some Pitfalls of Historical Statistics", in H. Rosovsky, (ed.), *Industrialization in Two Systems: Essays in Honour of Alexander Gerschenkron* (New York: J. Wiley and Sons, 1966), 90.