Analysis of Indonesian Rice Yield Gaps and Farmer Efficiency

Ray Trewin
Australian National University

Lu Weiguo
Australian National University

Erwidodo
Centre for Agro-Socioeconomic Research

and

Sjaiful Bahri
Centre for Agro-Socioeconomic Research

37th Annual Conference of the Australian Agricultural Economics Society
University of Sydney, Sydney
9-11 February 1993

Australia-Japan Research Centre, Australian National University, GPO Box 4, Canberra 2601

ABSTRACT

Because of questions concerning the high costs and effectiveness of Indonesia's current mix of policies aimed at promoting rice self-sufficiency, attention has turned to developing more efficient policies directed towards achieving self-sufficiency through increases in farmers' yields. The main issue addressed in this paper is whether existing yields can be improved. When a yield gap exists, either between farms and experimental trials or between groups of farms, then the issue becomes how to explain the gap and what policy action should be taken. The robustness of conclusions is examined in view of the fact that conclusions obtained in past analysis of the issues have often been inconsistent.

ACKNOWLEDGEMENTS

Funding for this project was provided by the Australian Centre for International Agricultural Research. The authors gratefully acknowledge assistance with concepts and analysis, without implications for the final content of the paper, given by Phil Kokic (Australian Bureau of Agricultural and Resource Economics), K. P. Kalirajan (Australian National University) and Ric Shand (Australian National University).

INTRODUCTION

A key issue for Indonesian policy-makers is how to maintain rice self-sufficiency, first achieved in 1984. The achievement of self-sufficiency was due to a mixture of policies promoting extensification (increases in the area harvested) and intensification (increases in yields). Extensification took place mainly as a result of increases in the area of land under irrigation. Intensification was the result of the introduction of high-yielding seed varieties and the application of highly subsidised associated inputs, such as water and fertilisers. These have been high-cost policies and attention has turned towards developing more efficient ways of achieving self-sufficiency through increases in yields and better use of existing irrigation infrastructure, for example.

The main issue to be covered in this paper concerns the existence of a yield gap within Indonesian rice production, that is, whether the existing yields of some farmers could be improved. If such a gap exists, either between farmers' yields and experimental results or between groups of farmers, then additional questions arise: what are the explanations for the gap; and can the gap be closed through policy action? Farmer efficiency over time will be specifically considered as a factor in any yield gap.

Past analysis of these issues has often resulted in inconsistent conclusions. This paper analyses the robustness of the conclusions obtained with regard to choice of models, methods of analysis and form of applied data.

The next section discusses the models that can be used to determine whether a yield gap exists and possible explanations for a yield gap along with alternative methods and forms of data for analysing the model. Models and methods are then applied to various sets of data from the Cimanuk Basin, Indonesia. Finally, the paper examines the policy implications of the analysis.

THE ECONOMETRIC MODEL AND ANALYTIC METHODS

Model functional forms

There are various ways of representing the efficiency of a farm's operations, for example via production, profit or cost functions. The production function, which describes the technical relationship that transforms inputs into outputs, is the traditional way of representing farm operations in analysis of farm efficiency. Also often used is the profit function, the complement of the traditional production function approach. Both approaches will be applied in the analysis that follows.

Whether production or some other general function is chosen to represent farm operations, the choice of specific functional form is important. With regard to the production function, examples of specific functional forms relating inputs to outputs include the linear, Cobb-Douglas (linear in logs) and Constant Elasticity of Substitution forms, and various flexible functional forms such as the translog, generalised quadratic and generalised Leontief forms (see Kopp and Smith 1980 for a general discussion of the various forms). The parameters in such models may be constant or varying in some specified manner. The choice of specific functional form is mainly an empirical issue although economic theory does impose some constraints.

Frontiers and envelopes

In analysis of yield gaps and farmer efficiency, it is not the average of observed relationships between farmers' inputs and outputs that is of interest but the maximum possible output that is obtainable from a given combination of inputs — the frontier production function.

A related concept, but one that will not always correspond, involves the envelope encompassing all the input–output combinations contained in the sample data. The distinction is a little like that made by Forsund et al. (1980) between a best-practices frontier (maximum output obtained with respect to the sample) and an absolute frontier (maximum output obtained with respect to all conceivable observations embodying the current technology). These frontiers are distinguished by Forsund et al. as being, respectively, 'non-statistical' (no one-sided error distribution and typically 100%-efficient observation(s)) or 'statistical', but Forsund et al. state that these frontiers would be expected to converge as the sample size grows. However, there is a time aspect to such concepts as well. Farms will invariably never adopt the technology being applied in experimental trials or on experimental stations completely or immediately. Yield gaps will be apparent when comparing farm and experimental outcomes in the one year, regardless of the sample size. A frontier estimated from a sample or the population of farms, whether it be non-statistical or statistical, will invariably lie below that encompassing the experimental observations but will approach it over time as the technology is completely adopted. Rather than making a 'statistical' distinction between a best-practices and absolute frontier, the issue here is to make a distinction in relation to time and adoption. More appropriate terms would be the current best-practices and the long-term absolute frontiers.

These concepts could have important policy implications. These will be looked at in more detail later but an illustration of their importance can be obtained from considering extension policies. At issue is whether extension policies need to be targeted to individual needs or to be more generally based. Analysing the distribution of farmer efficiencies relative to a current best-practices frontier and the best-practices frontier relative to the long-run absolute frontier gives useful information on these important policy options that would not be apparent from a traditional single frontier concept.

Inefficiencies and yield gaps

The distance a farm lies below its frontier measures the degree of technical inefficiency, that is, it is a residual measure. The existence of technical inefficiency of farms has been questioned. For example, Mueller (1974) states that 'little is known about the role of non-physical inputs, especially information and knowledge, which influence the firm's ability to use its available technology set fully... Once all inputs are taken into account, measured productivity differences should disappear except for random disturbances'. This seems to be a question of what constitutes an appropriate input. In terms of policy analysis, it is somewhat irrelevant whether extension advice, for example, improves the level of a 'non-core' input such as information and hence output under Mueller's view, or addresses inefficiencies due to a lack of information under a frontier function approach.

Two concepts have been introduced, yield gap and technical inefficiency, and the distinctions between them need to be discussed and emphasised. These concepts are also represented diagrammatically at the end of this section (Figure 1). A yield gap is the difference in yields achieved by farmers with their inputs and what could be obtained as a result of better developed application of inputs, either in experimental trials in the same fields or in neighbourhood experimental stations. Differences between this concept and that of technical inefficiency can occur for a number of reasons, some statistical and some conceptual. Statistically, comparison is often made between the average yield of a heterogeneous group of farmers, farms and environments and the best yield from a controlled experimental situation. A truer comparison would be obtained by using experimental trial results based on farmers' actual practices (apart from certain input use) and environments, rather than experimental station results. Although some experimental results take cost–benefit aspects into account many experimental station results do not, so the comparison is often between a result achieved through trying to maximise production and one achieved through maximising profits. In fact, it is unlikely that the population of farms will ever emulate experimental stations, and so comparisons between the two will always be highly qualified. Allocative inefficiencies, which result from failure to apply inputs at profit-maximising levels, can contribute positively or negatively to any yield gap depending on whether the inputs are under- or overallocated relative to the profit-maximising level. Regardless of the sign of allocative inefficiency, technical inefficiencies have tended to be the predominant factor in any yield gap. (*IRRI Research Highlights* 1978).

Variables used to explain a yield gap or inefficiencies could relate to direct causes such as profit-seeking behaviour or the use of improved technologies. (Sometimes little distinction is made between the introduction of a new technology and better application of an existing technology; for example, the specific means of applying fertilisers, such as briquettes or deep siting, could be classified as either.) Alternatively, these variables could relate to secondary factors underlying the direct causes, such as farm size, family size, land tenure, varieties planted, method of input application, mechanisation, access to irrigation, access to credit and extension advice and management proxies such as age and education. Some of these are affected by policy while others are structural in nature. Explanators have been determined by regression or other multivariate analysis, such as discriminant analysis, ensuring that appropriate transformations are undertaken so that the gap or efficiency measures satisfy the assumptions required of the analytic technique.

Forms of frontiers and methodological approaches

One form of frontier function is the stochastic production frontier, developed independently by Aigner, Lovell and Schmidt (1977) and Meeusen and van den Broeck (1977). Functions, being simplified representations of actual operations, require assumptions regarding the distribution of random errors before they can be used in empirical estimation. The usual assumption made in relation to average functions is a Normal distribution which is symmetric and bell-shaped. The stochastic frontier incorporates two random components: a traditional random error component, and a component representing the degree of technical inefficiency. Various distributional assumptions have been made with regard to this additional component, including the half-Normal and truncated Normal. In the deterministic frontier, any variation in firm performance is relative to a single frontier and attributed purely to inefficiency. This ignores the possibility of variation due to specific factors not under the firm's control, such as the socioeconomic and physical environment, which are usually incorporated as random error.

Another frontier approach developed by Kokic et al. (1992) applies the robust regression technique of M-quantile regression (Breckling and Chambers 1988) to the function representing farm operations. Basically the technique is a generalisation of M-regression (Huber 1981) and weighs positive residuals by a factor $2p$ and negative residuals by a factor $2(1-p)$. For p close to 1, the M-quantile production function represents the average performance of efficient farms, and for p close to 0 the performance of inefficient farms. A measure of the jth farm's performance, p_j, with the desirable property of not being dependent on the level of inputs can be determined using the technique. Because this approach makes different assumptions to the stochastic frontier approach (for example, in relation to error distributions), similar conclusions will suggest that these assumptions are robust. Differing conclusions should point to assumptions, possibly critical, that require greater information or more careful choice, for example whether an observation is truly an outlier or not.

One of the specific stochastic frontier models used in this paper was developed by Battese and Coelli (1991). This model can accommodate unbalanced panel data associated with a sample of N firms over T time periods and incorporates a simple exponential specification of time varying firm effects. The associated computer program, FRONTIER 2, was used to obtain maximum-likelihood estimates of the model parameters and predictors of the efficiencies of individual firms. However, this program has limitations. For example, when used in conjunction with the full panel of data, the ordering of farms according to efficiency levels does not change from that obtained for the first year whereas using the program one year at a time results in considerable changes in the order. The computer program LIMDEP was also used to estimate the various frontiers when the limitations of the FRONTIER 2 program were thought to be important. The M-quantile frontiers were estimated from an SAS program developed at ABARE by Phillip Kokic.

Data aspects

Frontier functions have been applied both to cross-sectional and to panel data such as the Cimanuk Basin data detailed later in the paper. The use of panel data has a number of advantages: consistent estimates of technical efficiency are more readily obtainable and fewer distributional and independence assumptions are necessary, for example. However, the use of panel data also introduces a number of complications such as whether or not any inefficiencies are time invariant.

In general, farm survey data on production is used for the analysis. However, it may be more appropriate to collect data that is aimed directly at the efficiency issue, for example by asking farmers directly whether price is a key determinant in the level of fertiliser use, if fertiliser use was found to be a cause of a yield gap. Also, consideration should be given to the use of field trials data, for example as a check of the estimated frontier or in the estimation, as greater accuracy could result from having access to the wider spread of input values used in such trials. A difficulty with the use of such data in the past has been the inequitable comparisons made between experimental stations, which operate under ideal conditions and with few constraints on inputs, and cross-regional averages of farms operating under real economic conditions (Pingali et al. 1990).

The earlier concepts of current best-practices and long-term absolute frontiers are useful in interpreting panel data. Analysing the data as a panel might suggest significant degrees of inefficiency, with or without a significant uniform trend over time. On then analysing the panel data a year at a time, it might seem that farmers are very efficient and that the frontier itself varies each year (see, for example, Battese and Coelli 1991). These outcomes would appear to be internally inconsistent given the traditional concept of a stable frontier. They would also seem to be inconsistent both with experimental trials suggesting that significantly higher and more profitable yields are achievable and with prior information showing that no technological change (as distinct from the adoption of technology) has taken place over the period. The current best-practices and long-term absolute frontier concepts enable a consistent interpretation of the analytic evidence. They suggest that in any year farmers are a homogenous group in terms of efficiency and that over time they move, although not always in a smooth fashion, towards a higher stable frontier as they adopt new technology as a group.

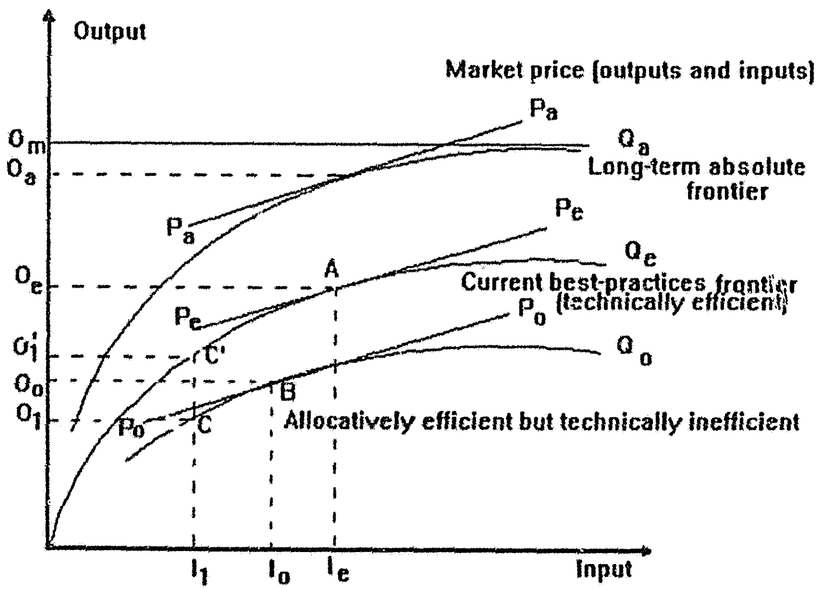
Current best-practices frontiers could be estimated from the panel data, one year at a time. The long-term absolute frontier could be estimated from the panel data as a whole as long as there was some modelling of the adoption rate of the new technology. Alternatively, it could be estimated using experimental trial data, although this information would need to be comparable with farm data, with fertiliser application rates being determined on a profits basis, for example. If the best-practices frontier varies randomly from year to year then this time variation could be used in the form of the pooled panel data to estimate a corresponding long-term absolute frontier. In this case the long-term absolute frontier would be conceptually like the meta-frontier discussed by Pitt (1983), except that instead of encompassing individual technologically specific frontiers it would encompass individual time-specific frontiers.

Another point to be made with regard to data is that in regression analysis often used to estimate the frontiers some points can have greater influence than others. Regression data diagnostic analysis (Belsley et al.1980) has been developed to ascertain which data points are influential in determining the estimated coefficients, forecasts, etc. by observing the responses of these estimates to changes in the data. This does not mean that these data points should be omitted; these points may be the only ones containing certain information which needs to be identified and judiciously used in the model development. Two basic component diagnostics are the diagonals of the least squares projection matrix (the 'hats') and the studentised residuals. The 'hat' matrix identifies points of high leverage that may be influential depending on the *y* values. Two diagnostic measures, DFBETAS and DFFITS, are respectively the scaled change in estimated coefficients, and fit due to deleting an observation, and although these are affected by the basic components it is invariably necessary to consider a suite of diagnostic measures to obtain a full picture.

Diagrammatic representations of various concepts

Figure 1 represents diagrammatically the concepts discussed earlier in this section. A production process is technically inefficient if maximum output is not produced from a given bundle of inputs. This concept is measured by the ratio of expected output to maximum output, for example O_1/O'_1 in the diagram; that is, a comparison of output at points *C'* and *C*, each with the same level of inputs but *C'* lying on the frontier function.

Figure 1 Current best-practices and long-run absolute frontiers, and various measures of inefficiency



Allocative inefficiency occurs when the marginal revenue product of an input is not equal to its marginal cost, implying that inputs are being used in the wrong proportions given input and output prices and technology. This can be defined alternatively as the inability to obtain maximum profit from the application of inputs with a given set of prices and technology. Allocative inefficiency can be measured by the ratio of expected profit to maximum feasible profit. Maximum feasible profit can be measured at the technically efficient level or the (possibly technically inefficient) current level. The latter gives a 'purer' measure of allocative inefficiency and in the diagram is represented by the ratio of the profits at C on Q_0 to the maximum profits given P_0P_0 on Q_0, that is the profits at point B (Q_0 represents a locus that is a neutral shift of the frontier Q_e and passes through the point C). The former is represented in the diagram by the ratio of the profits at C' on Q_e to the maximum profits given P_eP_e on Q_e, that is, the profits at point A. It is important to note the statement made earlier that points that are allocatively inefficient may have either higher or lower levels of production or yields than the allocatively efficient point. In other words, in the case of an overallocation of inputs, an improvement in allocative efficiency could widen any yield gap.

The combination of allocative and technical inefficiency components is generally referred to as economic inefficiency. It is measured as the ratio of predicted profit at the frontier for the actual level of inputs to the maximum feasible profits, obtained by simultaneously solving the frontier function and the first order profit maximisation conditions at the given input and output prices. This ratio is represented in the diagram as the ratio of profits at C to the profits at A. In this paper, only technical inefficiency will be considered in any detail.

The current best-practices and long-term absolute frontiers are also represented in the diagram. The long-term absolute frontier, the maximum output obtained with respect to all conceivable observations embodying the current technology, including experimental observations, is represented by Q_a. The current best-practices frontier, the maximum obtained with respect to the sample in a particular year, is represented by Q_e. Over time, there would be a sequence of Q_es and associated levels of technical and allocative efficiency approaching Q_a and associated points.

The relationship between the yield gap and measures of inefficiency can also be appreciated from the diagram. The gap, $O_a - O_l$, is made up of components due to technical inefficiency ($O_l' - O_l$), and allocative inefficiency ($O_e - O_l'$) and a component representing the differences between current best practices and long-term absolute frontier outputs ($O_a - O_e$). If the yield gap were to be considered in relation to experimental station results then it would be wider ($O_m - O_l$), incorporating an additional component associated with the non-profit-seeking behaviour of experimental stations that makes their yields closer to the maximum than profit-maximising yields ($O_m - O_a$).

APPLICATION TO PANEL DATA FROM THE CIMANUK BASIN

Cimanuk Basin data

The data set used in this study was obtained from the Centre for Agro-Socioeconomic Research (CASER) and was collected as part of a Rural Dynamic Study in the rice production area of the Cimanuk River Basin, West Java, Indonesia. The rice production area of the Cimanuk River Basin is characterised by irrigated rice farms set in an almost uniform agro-climate. Six *desa* (villages) located in five *kabupaten* (the administrative unit between district and province level) were covered in the survey. These villages are listed among the dummy variables defined in the appendix.

The survey was conducted twice in 1977, collecting information on farming practices in the wet season of 1975/76 and the dry season of 1976 and then on farm household activities in the wet season of 1976/77. In 1978 a similar survey was undertaken to cover farm management activities in the dry season of 1977. A follow-up survey for the 1981/82 and 1982 seasons was conducted in 1983 for the same farms/farmers with the emphasis on labour utilisation, asset-holding and land tenure arrangements. Altogether a balanced panel of 171 respondents spread reasonably evenly over the six villages were continuously surveyed over six seasons.

One difficulty with panel data is that, being obtained from farms that have remained in the population over a period of time, they may not be representative of the population of farms at a particular point in time. To address this issue, panel sample estimates of various farm characteristics derived from the Cimanuk Basin data were compared with those obtained from other more general samples. Pingali et al. (1990) include farmer field data for 1980 and 1988 obtained from a sample of 71 households in West Java, Indonesia: the same province as the Cimanuk Basin. Farm cost structure information for West Java is also available for 1982 from CBS. These sets of data are compared in Table 1.

Compared with CBS data, CASER panel data for 1982 for the Cimanuk Basin suggest a slightly smaller yield and slightly higher seed and fertiliser use, although these differences are within the realm of sample errors. Compared with Pingali et al. data, CASER panel data show on average a smaller yield, larger fertiliser use and roughly equal labour use in the Cimanuk Basin, although the Pingali et al. yields look high when compared to CBS data. Thus it would appear that the panel data is reasonably representative of the population of farms at particular points in time.

Aggregate evidence for a yield gap in Indonesia

Aggregate evidence for certain yield gap differences in Indonesia can be obtained by comparing the average of farm level data obained from CBS with multilocational trials data available from the Director General of Food Crops. Farm level performances will be distributed around the reported averages; thus the comparison will say nothing about the performance of the more efficient group of farms compared to the trials or compared to less efficient farms. Farm level comparisons as undertaken later in the paper will show whether a gap exists between all farms and scientific trials or between more and less efficient farms.

Multilocational farm trials undertaken in the 1989 season had the following features:

- trials were conducted in central production areas of mono-culture farming systems;
- the highest quality seed, certified HYV and already adapted to the area, was used;
- fertiliser use was based on the highest incremental yield recommendation in terms of dosage and timing;
- type and dosage of pesticides followed recommendations for the area; and
- farming systems as intensive as possible in terms of land preparation and nursery, seed and fertiliser use were implemented.

The trials were designed to ascertain whether there were any differences between their outcomes and those of actual farms due to different input usage. It is important to note that the usage of inputs and outcomes of the trials were driven by the aim to maximise the value of production in relation to fertiliser costs whereas farm usage of all inputs and outcomes was driven by the aim to maximise profits and minimise costs.

CBS data were obtained from a survey covering all of Indonesia excluding Irian Jaya and East Timor with provincial results calculated by a weighted average method and harvesting area used as the weight. Eight major rice production provinces both on and off Java, together contributing 97 per cent of national production, were compared. The provinces were West Java, Central Java, East Java, West Sumatra, Lampung, South Kalimantan, South Sulawesi and Bali. Results are available for production from both intensification and non-intensification areas.

It can be seen from Table 2 that in intensification areas there are large yield gaps in all provinces except Bali which has a gap of only 7.8 qu/hectares (or 780 kilograms per hectare). Central Java meanwhile has the largest yield gap at 25.4 qu/hectares (or 2540 kilograms per hectare). Yield gaps are even greater for non-intensification areas, although the area affected is much smaller (Table 3).

One reason for these yield gaps, alluded to above, could be differences in fertiliser use, with usage in trials and on farms being driven by different factors. It can be seen from Table 4 that the fertiliser usage recommended and applied in the trials is significantly greater than average fertiliser usage on farms. Other reasons for the gap could be the availability of extension advice or of inputs such as seed, water, fertiliser and credit, as well as statistical factors such as comparing a best estimate with an average estimate.

Further aggregate information on the existence of a yield gap can be obtained from Pingali et al. (1990) who analyse the gap between individual farms and experimental stations in West Java for the years 1980 and 1988. It would be expected that such a gap would be larger than that between farmers and experimental trials due to the more ideal conditions experienced on experimental stations and the fact that their yields are achieved without input constraints. Nevertheless, long-term declines in yields on experimental stations have been observed, mainly because of a decline in the paddy environment due to, for example, increased pest pressure and loss of nutrients. A decline in the generic potential of breeding materials has also been hypothesised as a cause. Looking at farmers' yields, national averages have continued to rise although yields in traditional rice-producing provinces have levelled off. Relevant data on average values and coefficients of variation are given in the following table:

Year	Farmers' yield	Yield frontier	Ratio
	kg/ha (%)	kg/ha (%)	kg/ha (%)
1980	4,897(20.7)	10,062(1.8)	0.49(19.9)
1988	6,335(21.8)	10,006(3.8)	0.63(22.0)

In contrast to the conclusions reached above, Pingali et al. conclude that if any gap exists it is mainly between farm (yields of top-ranking farms matching those of experimental stations), with the gap stemming from different farming abilities and access to irrigation rather than from different usages of inputs. These issues will be analysed further using the Cimanuk Basin data.

Application to Cimanuk Basin panel data

Earlier applications

The stochastic frontier approach has been applied previously to Cimanuk Basin panel data (Erwidodo 1990). Erwidodo fitted Cobb-Douglas production and profit functions to pooled data with total output per farm as the dependent variable and total quantity of seed, fertiliser, labour, farm size and a number of dummies (pesticide use, seed varieties, season, village) as the independent variables. (The appendix contains a detailed definition of the variables used in the analysis.) Various tests of model specification and estimates were performed before analysis of the selected model was undertaken . The coefficient estimates for non-dummy variables in the Cobb-Douglas production function were significant and had the expected sign. The coefficient estimates for the dummy variables suggest that:

- pesticide use had no significant effect on production (no significant crop damage due to pests or diseases occurred in the area under study during the survey period);
- production was greater in the wet season; and
- farm size and region made no significant difference to production (the nationwide rice intensification program was established at the time of the survey).

Estimates of the technical inefficiency of individual farms were obtained using the method suggested by Battese and Coelli (1991). The range of estimates was quite small, ranging from 3.4 to 12 per cent, with a quite low mean level of 6.5 per cent. No significant difference between large and small farms was determinable in the level of technical inefficiency. Analysis of the value of the marginal product of a particular input to its marginal factor cost suggests underutilisation of seed and fertiliser and overutilisation of labour.

The first step in the analysis was to duplicate the model used by Erwidodo. Initially, a hybrid form of the Cobb-Douglas stochastic production function was estimated. This form, additive in logs apart from the potassium fertiliser variable, was used to overcome the difficulty caused by many individual farmers not using any potassium fertilisers. Zeros in a Cobb-Douglas production function can be handled in a number of other ways, for example by adding the individual fertilisers (weighted or unweighted), or by converting the zeros to a small positive value or to unity. The appropriateness of the approach will depend on the need for separate fertiliser estimates and the structure of the untransformed data. Various approaches were used in the analysis.

Alternative models

A general form of the hybrid model was estimated using FRONTIER 2 in conjunction with the panel data, and then various restricted forms were tested. The preferred model was one in which farm technical efficiency was time-invariant and the stochastic distribution had mean zero, both aspects assumed by Erwidodo (1990). Parameter estimates were similar to those of Erwidodo (1990) apart from the hybrid parameter associated with potassium fertiliser. The estimates of technical inefficiency were also similar although the mean value was slightly higher at 9.6 per cent (Table 5).

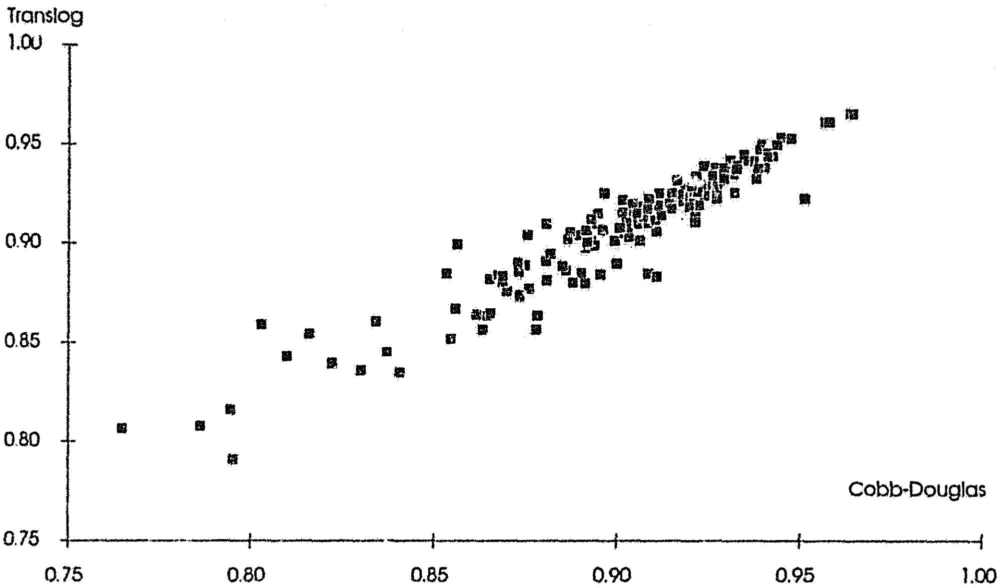
A similar outcome was observed when zero observations for potassium fertiliser were replaced by unity, effectively resulting in zero entries in a log format, or when the individual fertilisers were summed.

At this stage some preliminary analysis was undertaken relating the individual measures of technical inefficiency to a socioeconomic variable, namely farm size. As may have been expected from the earlier results on production, farm size was not a significant explanator of technical inefficiency.

The next step was to estimate a stochastic translog production frontier. Fertiliser variables were aggregated in this model, the weights being determined from a regression of the fertiliser variables on yield. The form of this model encompasses the Cobb-Douglas form so preference for one form over the other can be tested via the significance of the cross-terms in the translog form. The F-test of these terms suggested that the translog form is preferred. (The R^2 for the translog model was 0.885 compared to 0.86 for the Cobb-Douglas model giving an F value of 7.5 which is significant at the 1 per cent level.)

This preferred model was estimated using FRONTIER 2 and the panel data. Again the preferred form of this model was one in which farm technical inefficiency was time-invariant and the stochastic distribution had mean zero. Of the Cobb-Douglas terms, labour was not significant and fertiliser was significant but negatively signed. However, these aspects were balanced by the cross-terms, with the fertiliser-by-fertiliser term being highly significant and positive, and labour significant in conjunction with seed though negative. The estimates of technical inefficiency were slightly lower, with a mean value of 9.1 per cent (see Table 6).

Figure 2 Frontier performance measure: Cobb-Douglas versus translog production function (panel data)



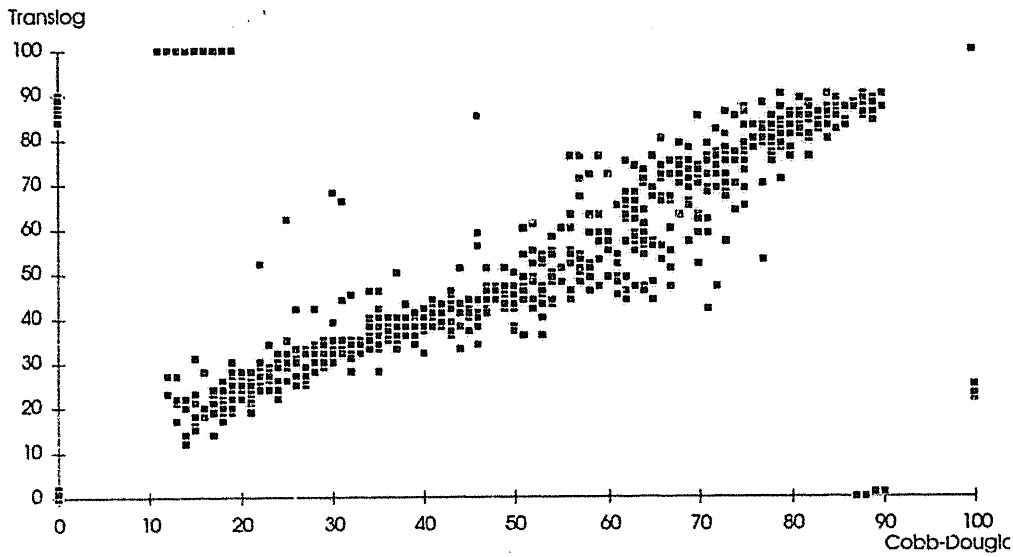
The choice of model would not appear to make a significant difference to the general conclusions of the analysis. In fact, a mapping of individual farm technical inefficiency measures from the Cobb-Douglas model against those from the translog model closely follows a straight line 45 degrees from the origin (Figure 2).

Alternative methods

The previous analysis considers various constant parameter model forms within a stochastic production frontier approach. The M-quantile approach is now applied to Cimanuk Basin data to estimate constant parameter Cobb-Douglas and translog functional forms in the case of both production and profit. In the case of the translog functional form, multicollinearity was found to be present; to overcome this problem the first few principal components explaining most of the variability were estimated and used in the M-quantile approach. Panel data were pooled in the analysis. This would seem appropriate given earlier evidence of time invariance. Sensitivity analysis was undertaken for the key parameters in the M-quantile approach to ensure its robustness. Individual estimates of farm performance in relation to both production and profits were compared with each other and with inefficiency measures obtained from the stochastic frontier approach. The M-quantile approach will always rank farms in terms of their performance even if the best and worst farms are not significantly different. Boot-strapping methods would be required to determine any significant differences. Under the stochastic frontier approach the lack of any significant difference between farms is indicated by a failure to estimate any significant frontier. Key results of the M-quantile approach were that:

- there was marked variation in individual farm performance measures over time, although in no uniform manner;
- there were differences in individual farm performance measures for different functional forms, including the various principal component forms, although this would appear to be more a consequence of

Figure 3 M-Quantile performance measure: Cobb-Douglas versus translog production function (panel data)



the principal components approach to multicollinearity than of the functional forms themselves (Figure 3); and

- there were marked differences in individual farm performance measures for the M-quantile approach and in inefficiency measures for the stochastic frontier approach. (In fact, a number of farms had upper extremes for one measure and lower extremes for the other, suggesting that the treatment of outliers could be critical to the measure obtained.

Year-by-year analysis

The result obtained from M-quantile analysis that individual farm performance measures differed markedly from year to year — something unable to be ascertained from analysis using FRONTIER 2 — leads to reconsideration of earlier analysis of the current best-practices and long-term absolute frontiers concepts. Estimates for individual years were undertaken for a Cobb-Douglas stochastic production function with half-Normal errors after testing this specification against translog and exponential errors specifications (see Table 7).

Stochastic frontiers could not be estimated for the middle four 'years'; that is, farmers as a group appear fully efficient during the dry seasons of 1976 and 1977 and the wet seasons of 1976-77 and 1981-82. The only consistent trend in the estimated coefficients was for the fertiliser variable to increase year by year (see Table 8).

Estimates of individual farm inefficiencies where available ranged from 6 to 40 per cent with a mean value of around 15–16 per cent Given the earlier gap between average farm and experimental station yields of around 50 per cent in the early 1980s, this range suggests a significant gap would exist between best farm and experimental station yields contrary to the conclusion reached by Pingali et al.

Figure 4a M-Quantile versus Frontier performance measure (Cobb-Douglas) (1975/1976 wet season)

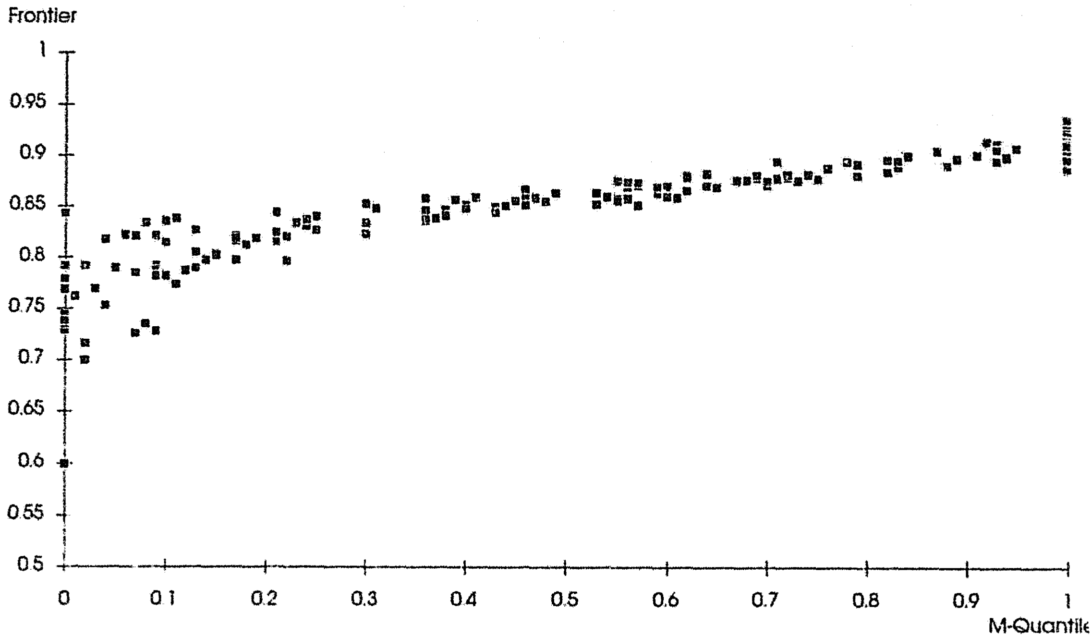
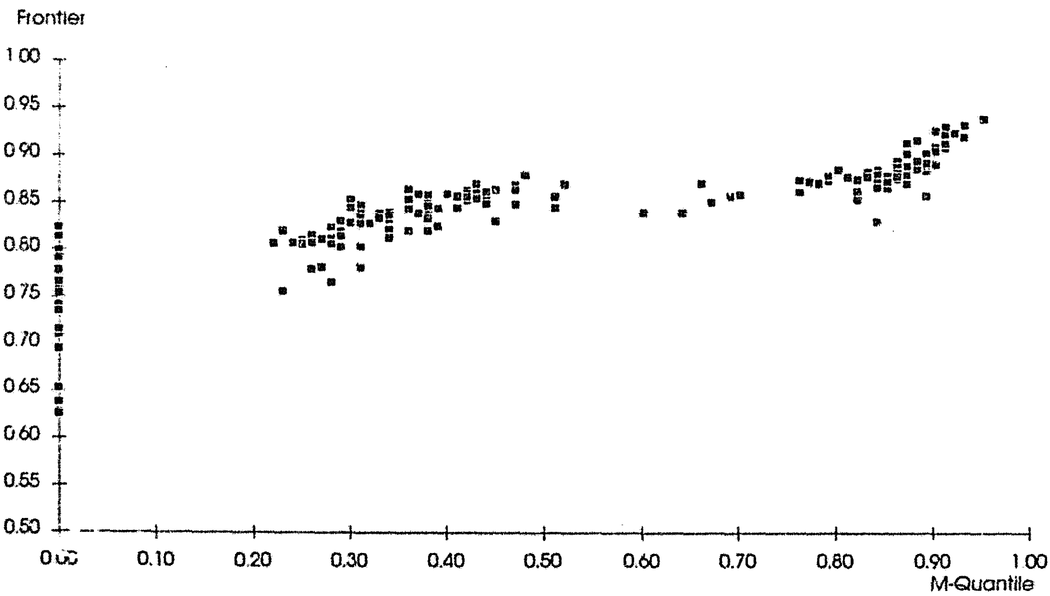


Figure 4b M-Quantile versus Frontier performance measure (Cobb-Douglas) (1983 dry season)



M-quantile analysis was also undertaken on an annual basis for the Cobb-Douglas production function, and estimates of the 95 per cent M-quantiles were produced (see Table 9). Although there are some marked variations in these estimates over time, on average they correspond reasonably well with estimates from the stochastic frontier approach. A number of tests showed that there were significance differences between these estimates over time. (For example, the Kruskal-Wallis test, which is a non-parametric version of a one-way analysis of variance, was highly significant.) Individual farm performance measures were plotted against individual farm efficiency measures for those years in which a frontier could be estimated (Figures 4a and 4b).

As with similar analysis of pooled data, there were some marked differences in individual farm performance measures for the M-quantile approach and in inefficiency measures for the stochastic frontier approach, with a number of farms lying in the upper left corner of the figure especially in the 1983 dry season. Generally though, the situation for both the 1975/76 wet season, and the 1983 dry season is that there is a good overall correspondence between the two sets of measures.

Data diagnostics

Comparison of the M-quantile and stochastic frontier approaches suggests the need to undertake some outlier or regression data diagnostic analysis. Such diagnostics currently do not exist for stochastic frontier models, although these could be derived in a basic form by noting the effect on parameter estimates and forecasts of dropping each data point. Because the majority of the annual models were average production functions estimated by ordinary least squares, or close to it, the diagnostics applicable to such models were estimated as preliminary analysis of this issue. There were 81 points with large DFBETAS and DFFITS. These points were not consistent over the years, however, 52 farms having only one outlying point across the six seasons and only one farm having more than three outlying points across the six seasons. Farms that had high levels of efficiency were usually associated with upper tail outliers in the DFFITS (a residual measure similar to efficiency measures), and vice versa. This situation appeared to apply uniformly when both M-quantile and frontier measures were available (see Table 10).

The correspondence between the two measures and the appropriate data diagnostic was good. The points where the correspondence was not good between the two measures were not influential points. Whether the point is a true outlier and should therefore be excluded from the analysis is important as both approaches treat these points as influential. Detailed analysis of the characteristics of the individual data points (Seaver and Triantis 1989) would be required before a decision on the outlier status of an individual data point could be made.

Future analysis

The above analysis shows the importance in analysing yield gaps and inefficiency measures of considering general models and all available data, either directly or via the analytical method. These aspects have not been covered fully in the analysis undertaken in this paper. For example, consideration should be given to models representing non-neutral shifts in technology (for example, a random coefficients model) to see if these more general specifications would give greater insight into the efficiency of individual farms over time. If these models fit the data significantly better than constant parameter models then significant frontiers that had been swamped by the larger residual noise of the constant parameter models could be revealed. The residual noise in the constant parameter models appears to be large, with R^2s of around 90 per cent and the inefficiency of farms ranging from 6 to 40 per cent. Consideration should also be given to incorporating relevant experimental trial information into the estimation of the frontier functions and inefficiency measures. The characteristics of influential data points in the estimation of frontier functions and inefficiency measures should be analysed to determine whether these points are true outliers and need to be excluded from the subsequent analysis or not (see Seaver and Triantis 1989). Finally, if progress in closing any yield gap is to be made, then individual measures of efficiency need to be explained in terms of variables that may be affected by policies. The greater robustness of measures obtained by the above approach may lead to the identification of more significant and stal . · ·xplanators of the yield gap and inefficiency measures than have been observed in the past.

CONCLUSION

It was stated in the introduction that the main purpose of this paper was to consider the existence of a yield gap in Indonesian rice production, giving specific consideration to changes over time and to robustness. This was addressed by analysing panel data from the Cimanuk Basin in West Java.

Aggregate analysis suggested that there was a yield gap, at least between the average yields of farms and the best yields obtained from experimental trials using more advanced technologies. This outcome seemed at odds with initial analysis of farm inefficiency using a stochastic frontier approach which suggested that any inefficiencies were small. These outcomes were reconciled by introducing concepts that distinguished between the current best practices of farms and the long-term absolute frontiers achieved in experimental trials using more advanced technologies. These concepts have important policy implications. For example, the above results suggest that inefficiency is important, but in a general rather than a farm-specific sense. The yield gap could be closed by providing better extension to transfer the technologies applied in the experimental trials rather than extension advice targeted at particular groups of farms.

The robustness of the analysis was investigated by comparing a Cobb-Douglas and a translog model specification as well as stochastic frontier and M-quantile regression approaches, and by considering the influence of individual data points. There seemed to be little difference between the measures obtained from either model specification. However, as only constant parameter forms were considered, the choice of specifications was somewhat limited. The yield gap could be viewed as being due to different technologies even when the basic technology, for example high-yielding varieties, is the same. This is because the specific implementation of technology, for example the application of fertilisers (broadcasting, briquette, deep siting, etc.), is evolving. In such cases a specification with time varying parameters could be more appropriate. The use of different methods did not illustrate major variations in efficiency measures but did illustrate the importance of considering the analysis on a year-by-year basis. The differences were investigated in terms of the influential data points identified from regression data diagnostics and it was suggested that the treatment of influential data points was fundamental to the farm efficiency or performance measure obtained.

The analysis in the paper is incomplete. A fuller analysis would include more flexible model specifications, such as those with varying parameters. Relevant experimental trial information reflecting the long-term potential yield of farms should be incorporated into frontier analysis and measures of inefficiency. Influential data points in the estimated frontiers should also be assessed to determine whether or not they are true outliers and thus should be excluded from the analysis. Once robust measures of efficiency have been obtained these should be analysed in terms of possible explanators, especially those that can be influenced by policies.

Key policy options that have been used or considered as vehicles for closing yield gaps or addressing inefficiencies include input subsidies, infrastructure investment, extension advice and research. The most likely candidate for addressing the gap between farmers' yields and those obtained from experimental trials is extension policy. The analysis in this paper suggests that improved extension should be considered from a general perspective rather than being targeted towards groups of farms, as no one group stands out as being more efficient than any other.

References

Aigner, D.J., Lovell, C.A.K. and Schmidt P. (1977), 'Formulation and estimation of stochastic frontier production function models', Journal of Econometrics, 6: 21–37.

Battese, G. E. and Coelli, T. J. (1991), 'Frontier production functions, technical efficiency and panel data: with application to paddy farmers in India', Working Papers in Econometrics and Applied Statistics, No. 56, Department of Econometrics, University of New England, Armidale.

Belsley, D.A. et al. (1980), Regression Diagnostics: Identifying Influential Data and Sources of Collinearity, Wiley, New York.

Breckling, J. and Chambers, R. (1988), 'M-quantiles', Biometrika, 75:761–71.

Erwidodo (1990), 'Panel data analysis on farm level efficiency, input demand and output supply of rice farming in West Java, Indonesia' Unpublished PhD thesis, Michigan State University, East Lansing, Michigan.

Forsund, F. R. et al.(1980), 'A survey of frontier production functions and of their relationship to efficiency measurement'. Journal of Econometrics 13:5–25.

Huber, P. J. (1981). Robust Statistics, Wiley, New York.

IRRI Research Highlights (1978), IRRI, Los Banos, Laguna, Philippines.

Kokic, P. et al. (1992), 'A measure of production performance', paper submitted to American Journal of Agricultural Economics.

Kopp, R. J. and Smith, V. K. (1980), 'Frontier production function estimates for steam electric generation: a comparative analysis', Southern Economic Journal, 47: 1049–59.

Meeusen, W. and van den Broeck, J. (1977), 'Efficiency estimation from Cobb-Douglas production functions with composed error', International Economic Review, Vol. 18, No. 2: 435–444.

Mueller, J. (1974), 'On sources of measured technical efficiency: the impact of information', American Journal of Agricultural Economics, 56, 730-38.

Pingali, P.L., et al. (1990), 'The post-green revolution blues in asian rice production — the diminished gap between experiment station and farmer yields', International Rice Research Institute, Manila, Philippines.

Pitt, M. M. (1983), 'Farm-level fertilizer demand in Java: a meta-production function approach', American Agricultural Economics Association.

Seaver, B. L. and Triantis, K. P. (1989), 'The implications of using messy data to estimate production-frontier-based technical efficiency measures', Business Economic Statistics, 7: 49–59

Appendix

Definition of variables

GRKG:	Gross output of rice (kg).
KGS:	Seed (kg)
KGN:	Urea (kg)
KGP:	Phosphate (TSP, kg)
FERT:	KGS plus KGP
LAB:	Total labour (including family and hired labour, hours)
LAND:	Cultivated farm size (hectares)
D1:	dummy variable of pesticide use, 1 if farmer uses pesticides and 0 otherwise
D2:	dummy HYV varieties, 1 if HYV and 0 otherwise
D3:	dummy MV varieties, 1 if MV and 0 otherwise
D4:	dummy variable of season, 1 if wet season and 0 otherwise
D5:	dummy variable of farm size, 1 if farm size greater than 0.5 ha and 0 otherwise
D6:	dummy village, 1 if desa Lanjan kabupaten Indramayu and 0 otherwise
D7:	dummy village, 1 if desa Gunung Wangi Kabupaten Majalengka and 0 otherwise
D8.	dummy village, 1 if desa Malausma kabupaten Majalengka and 0 otherwise
D9:	dummy village, 1 if desa Sukaambit kabupaten Sumedang and 0 otherwise
D10:	dummy village,

Table 1: Comparison of farm level estimates

	Cimanuk Basin data					CBS	Pingali et al. [a]
	1976/77	1977	1981/82	1982	Av.	1982	1980
Yield (kg/ha)	2,513	2,350	4,197	3,969	3,207	4,134	4,897 (20.7)
Seed (kg/ha)	40.6	37.2	42.2	37.8	39.4	35.4	
N (kg/ha)	220.0	192.7	268.1	250.0	232.7	193.3	94.3 (27.6)
P (kg/ha)	63.0	55.9	119.7	110.8	87.4	82.7	
Labour (8 hr days)	103.2	108 1	122.3	113.4	111.7		104.6 (29.5)

[a] Coefficients of variation in brackets

Table 2 Yield gap with BIMAS (intensification) wetland rice in eight major Indonesian rice production provinces, 1989

Province	Yield (t/ha) Trials a)	Farms b)	Per cent c)	Yield gap (t/ha)
Java				
West Java	7.32	5.03	68.7	2.29
Central Java	7.67	5.13	66.8	2.54
East Java	7.32	5.27	72.0	2.05
Off Java				
West Sumatra	6.51	4.52	69.4	1.99
Lampung	6.45	4.18	64.8	2.27
South Kalimantan	5.11	2.83	55.4	2.29
South Sulawesi	6.09	4.28	70.3	1.81
Bali	5.83	5.05	86.7	0.78

Notes :
a DG of food crops.
b CBS's cost structure.
c Percentage of farms to trials.

Table 3 Yield gap for non-intensification areas in eight major Indonesian rice production provinces

Province	Yield (t/ha)			Yield gap
	Trials a)	Farms b)	Per cent c)	(t/ha)
Java				
West Java	7.32	3.21	43.9	4.11
Central Java	7.67	3.36	43.8	4.31
East Java	7.32	2.93	40.0	4.40
Off Java				
West Sumatera	6.51	3.49	53.6	3.02
Lampung	6.45	3.03	47.0	3.42
South Kalimantan	5.11	2.31	45.2	2.80
South Sulawesi	6.09	2.69	44.2	3.40
Bali	5.83	3.75	64.3	2.08

Notes:
a DG of Food Crops.
b CBS's Cost Structure.
c Percentage of famrers to trials.

Table 4. Fertiliser recommended and used on wetland rice (intensification) in eight major rice production provinces

	Recommendation (kg/ha)a)				Used (kg/ha)b)				Gap (kg/ha, %)			
	Urea	TSP	Other	Total	Urea	TSP	Other	Total	Urea	TSP	Other	Total
West Java	250	125	113	488	228	141	65	434	22 (8.8)	-16 (12.8)	48 (42.4)	54 (11.1)
Central Java	250	125	125	500	237	113	38	388	13 (5.2)	12 (9.6)	87 (69.6)	111 (22.2)
East Java	300	113	125	538	292	100	35	427	8 (2.7)	13 (11.5)	90 (72.0)	111 (20.6)
W. Sumatra	200	100	113	413	141	116	51	308	59 (29.5)	116 (16.0)	62 (54.9)	105 (25.4)
Lampung	192	83	113	388	167	134	60	361	25 (13.0)	-5 (61.4)	53 (46.9)	27 (7.0)
S. Kalimantan	142	108	50	300	106	72	19	197	36 (25.4)	36 (33.3)	31 (62.0)	103 (34.3)
S. Sulawesi	225	88	126	439	167	67	25	259	58 (25.8)	21 (23.9)	101 (80.2)	180 (41.0)
Bali	250	100	125	475	250	77	36	363	0 (0.0)	23 (23.0)	89 (71.2)	112 (23.6)

Notes: a) DG of food crop Fertiliser dosage recommended in 1986.
b) CBS's cost structure, 1989.

Table 5a Maximum likelihood estimates of stochastic Cobb-Douglas production function (panel data)

	Coefficients	t-ratio
Constant	4.9697	26.3531
ln KGS	0.1551	5.7144
ln KGN	0.1257	6.9972
ln KGP	0.0711	6.0616
ln LAB	0.2289	7.9315
ln LAND	0.4271	13.7819
D1	0.0138	0.4756
D2	0.1403	2.6215
D3	0.1735	4.4328
D4	0.0444	2.0330
D5	0.0315	0.8431
D6	-0.0334	-0.6675
D7	-0.0254	-0.4300
D8	-0.0647	-1.0363
D9	0.0260	0.4467
D10	0.0877	1.4526
$\sigma_s^2 \equiv \sigma_v^2 + \sigma_u^2$	0.1314	16.0818
$\gamma \equiv \sigma_u^2 / \sigma_s^2$	0.1285	2.3580

Log-likelihood function = -367.60

Chi-square test of one sided error (σ_u^2) = 6.00 with one degree of freedom

Note: The dependent variable is GRKG in log form. σ_v^2 is iid normally distributed random errors and σ_u^2 is non-negative truncated normally distributed random errors.

Table 5b Frequency distribution of farmers based on level of technical inefficiency from Cobb-Douglas production frontier.

Technical inefficiency	Number of farms	Frequency distribution (%)
$\leq 5\%$	5	2.92
$5\% < u \leq 10\%$	104	60.82
$10\% < u \leq 15\%$	50	29.24
15% over	12	7.02
Mean	9.58	
Minimum	3.50	
Maximum	21.98	
Total number of farms	171	100.00

Table 6a Maximum likelihood estimates of stochastic translog production function (panel data)

	Coefficients	t-ratio
Constant	6.4539	3.8560
ln KGS	1.0244	3.0505
ln FERT	-0.5138	-2.6624
ln LAB	-0.0366	-0.0693
ln LAND	0.8534	2.4049
ln KGS*ln KGS	-0.0479	-1.9544
ln KGS*ln FERT	0.0502	1.7918
ln KGS*ln LAB	-0.1452	-2.4172
ln KGS*ln LAND	0.1286	2.5754
ln FERT*ln FERT	0.0422.	6.6480
ln FERT*ln LAB	-0.0092	-0.3325
ln FERT*ln LAND	-0.1274	-4.0510
ln LAB*ln LAB	0.0596	1.2721
ln LAB*ln LAND	0.0245	0.3899
ln LAND*ln LAND	0.0201	0.8160
D1	0.0195	0.6986
D2	0.1072	2.0238
D3	0.1826	4.7590
D4	0.0391	1.8377
D5	0.0121	0.2913
D6	-0.0281	-0.5864
D7	0.0531	0.9533
D8	-0.0109	-0.1925
D9	0.0709	1.2832
D10	0.1444	2.5926
$\sigma_s^2 \equiv \sigma_v^2 + \sigma_u^2$	0.1239	16.0622
$\gamma \equiv \sigma_u^2 / \sigma_s^2$	0.1212	2.2142

Log-likelihood function = 340.56

Chi-square test of one sided error = 5.4271 with one degree of freedom

Note: The dependent variable is GRKG in log form. σ_v^2 is iid normally distributed random errors and σ_u^2 is non-negative truncated normally distributed random errors.

Table 6b Frequency distribution of farmers based on the level of technical inefficiency from translog production frontier

Technical inefficiency	Number of farms	Frequency distribution (%)
≤ 5%	8	4.68
5% < u ≤ 10%	115	67.25
10% < u ≤ 15%	39	22.81
15% over	9	5.26
Mean	9.08	
Minimum	3.470	
Maximum	20.89	
Total number of farms	171	100.00

Table 7 Specification test: translog versus Cobb-Douglas production function

Season	Restricted model (Cobb-Douglas) R^2	Unrestricted model Translog R^2	Calculated F values	Translog preferred?
1975/76 wet	0.9209	0.9275	1.3382	No
1976 dry	0.8663	0.8976	4.4933	Yes
1976/77 wet	0.9134	0.9241	2.0723	No
1977 dry	0.8834	0.8948	1.5930	No
1982/83 wet	0.9372	0.9444	1.9036	No
1983 dry	0.9356	0.9393	0.8960	No

Note : The critical value for the F distribution at 1% significance level is 2.32.

Table 8a Maximum likelihood estimates for parameters of stochastic Cobb-Douglas production function (half-normal)

	1975/76 Wet	1976 Dry (OLS)	1976/77 Wet (OLS)	1977 Dry (OLS)	1982/83 Wet (OLS)	1983 Dry
ln KGS	0.0737	0.1321	0.1167	0.1672	0.1776	0.0999
	(1.2043)	(2.3670)	(2.1650)	(2.7210)	(2.8320)	(1.5348)
ln FERT	0.0785	0.1440	0.1696	0.165	0.2595	0.2849
	(3.0818)	(4.0280)	(6.1120)	(5.3850)	(4.5950)	(5.5700)
ln LAB	0.1875	0.2935	0.3459	0.2049	0.0267	0.1755
	(2.9052)	(3.9040)	(6.2200)	(2.9560)	(0.4160)	(2.8037)
ln LAND	0.6823	0.3980	0.3380	0.4408	0.5895	0.5514
	(7.9974)	(6.0550)	(6.2890)	(6.1280)	(8.0390)	(7.2916)
Constant	6.1961	4.4354	4.4559	3.4112	6.235	6.1745
	(12.4259)	(5.7040)	(7.1570)	(5.7770)	(10.6070)	(11.8815)
R^2 based on OLS	0.9209	0.8663	0.9134	0.8834	0.9372	0.9356
Breusch-Pagan-Godfrey test based on OLS						
χ^2 test with 13 d.f.	10.6510	27.0400	31.3860	25.5280	18.0090	23.1350
Frontier 2 diagnostics						
$\sigma_s^2 \equiv \sigma_v^2 + \sigma_u^2$	0.0954 (2.9120)					0.1055 (3.0826)
$\gamma \equiv \sigma_u^2 / \sigma_s^2$	0.4875 (1.4388)					0.4953 (1.5648)
Log-likelihood function	-9.7176					-17.7291
Chi-square test of one sided error with 1 d. f.	0.4813					0.4674
LIMDEP Diagnostics						
$\lambda \equiv \sqrt{\sigma_v^2 + \sigma_u^2}$	0.3089 (5.7200)					0.3248 (5.5680)
$\gamma \equiv \sigma_u^2 / \sigma_v^2$	0.9753 (1.3830)					0.9907 (1.3900)

Notes: 1) The dependent variable is GRKG in log form. The coefficients reported for the stochastic model are the estimates from FRONTIER 2, which are found to be quite similar to those from LIMDEP. Dummy coefficients are not reported. Figures in parentheses are t values.

2) The critical value for $\chi^2(13)_{0.05}$ is 22.4 and for $\chi^2(1)_{0.05}$ is 3.84.

Table 8b Frequency distribution of farmers based on the level of technical inefficiency from Cobb-Douglas production frontier (1975/76 wet season)

Technical inefficiency	Number of farms	Frequency distribution (%)
≤ 5%	0	0
5% <u ≤ 10%	27	15.79
10% < u ≤ 15%	72	42.11
15% <u ≤ 20%	41	23.98
20% < u ≤ 25%	22	12.87
25% over	9	5.25
Mean	15.12	
Minimum	6.32	
Maximum	40.05	
Total Number of Farms	171	100.00

Table 8c Frequency distribution of farmers based on the level of technical inefficiency from Cobb-Douglas production frontier (1983 dry season)

Technical inefficiency	Number of farms	Frequency distribution (%)
≤ 5%	0	0
5% <u ≤ 10%	18	10.53
10% < u ≤ 15%	73	42.69
15% <u ≤ 20%	54	31.58
20% < u ≤ 25%	18	10.53
25% over	8	4.67
Mean	15.88	
Minimum	6.07	
Maximum	37.42	
Total Number of Farms	171	100.00

Table 9: 95% M-quantile production frontier parameter estimates a

Season	Seed	N fertiliser	P fertiliser	Labour	Land	Constant
1	0.04	0.10	0.06	0.21	0.78	6.48
2	0.29	0.19	-0.05	0.39	0.04	3.74
3	0.13	0.18	0.04	0.22	0.42	5.40
4	0.04	0.19	-0.04	0.00	0.71	7.16
5	0.04	0.18	0.09	0.06	0.59	6.56
6	-0.00	0.18	0.06	0.27	0.64	6.03

a Dummy variable estimates not included.

Table 10 a Performance measures and outliers (1975/76 wet season)

(U=Upper tail, L=Lower tail)

ID	Efficiency Measures		HAT	DFFITS		DFBETAS									
	M-quantile	Frontier	Matrix	Upper	Lower	Intercept		KGS		FERT		LAB		LAND	
5	1.00	0.92		U		U			L		L			U	
6	1.00	0.93	outlier	U		U			L	U					
16	0.02	0.70			L		L			[illegible]		U			L
28	0.00	0.74			L			U			L				
35	0.00	0.79	outlier		L		L								
37	1.00	0.91	outlier	U							L				
39	0.93	0.91							L					U	
47	1.00	0.92					L			U					
95	1.00	0.93					L	U							L
101	1.00	0.92				U					L			U	
106	0.03	0.77					L	U							L
114	0.00	0.74					L					U			L
118	1.00	0.93		U				U		U					
119	0.92	0.92	outlier	U											
141	0.08	0.74					L	U							L
142	0.11	0.77					L					U			
143	0.00	0.60			L				L	U					
144	0.07	0.73				U							L	U	
145	0.02	0.72				U			L				L	U	
151	0.95	0.91				U							L	U	
162	0.57	0.87	outlier												
164	1.00	0.94		U						U		U			

Table 10 b Performance measures and outliers (1983 dry season)

(U=Upper tail, L=Lower tail)

ID	Efficiency Measures		HAT	DFFITS		DFBETAS									
	M-quantile	Frontier	Matrix	Upper	Lower	Intercept		KGS		FERT		LAB		LAND	
7	0.91	0.91					L			U		U			L
18	0.00	0.63			L		L								L
24	0.90	0.91				U					L			U	
38	0.00	0.77					L	U							L
51	0.90	0.93	outlier	U		U							L	U	
61	0.00	0.75					L					U			L
73	0.87	0.90				U								U	
86	0.00	0.64			L						L	U			
110	0.00	0.63			L	U			L		L	U		U	
113	0.91	0.91		U										U	
125	0.92	0.92		U											
139	0.00	0.71			L		L					U			L
152	0.91	0.93		U				U			L		L		L
162	0.00	0.69			L	U		U			L				
163	0.95	0.94		U			L	U		U					L
167	0.93	0.93		U					L	U					
170	0.00	0.65			L		L	U							L
171	0.91	0.91							L			U			

Note: The SAS program was used to estimate the various diagnostic statistics for outliers. For a detailed definition of HAT matrix, DFFITS and DFBETAS, please see Besley et al. (1980).