Performance evaluation [illegible] / indicators / irrigation [illegible] / irrigation management

Tilak

IIMI Colombo 91 May.

PERFORMANCE EVALUATION MEASURES: A TEST ON THEIR CONSISTENCY, VALIDITY AND RELEVANCE

Khin Maung Kyi
Senior Management Specialist

International Irrigation Management Institute
P.O. Box 2075, Colombo,
Sri Lanka.

May 1991

TABLE OF CONTENTS

?

EXECUTIVE SUMMARY

This paper *is* an attempt to operationalize and use the performance measures such as Relative Water Supply, Delivery Performance Ratio, various equity ratios, in measuring the performance of the systems about which some data have been already compiled by IIMI Research Projects in the immediate past. The purpose of this paper is not to create new indices but to examine the reliability, consistency, validity, and relevance of the measures already suggested. The paper poses three important cardinal questions:

a) How reliable, repeatable or consistent are the results of the measures we have suggested.

b) How valid is the instrument we have proposed ? or Alternatively, does the instrument do measure what it purports to measure ?

c) How relevant is the measure we have devised to the needs and requirements of the various users under varying conditions.

The paper then analyzes the consistency between sub-system and main system results, repeatability of internal consistency of the instruments and "constructs" validity of instruments through inter comparison between various related measures. Statistical tests are used and results are summarized in the 'Conclusion.' A methodological framework for the'analysis of reliability, consistency, validity, and relevance of these instruments is posited in the Appendix I as a tentative model.

PERFORMANCE EVALUATION MEASURES: A TEST ON THEIR CONSISTENCY, VALIDITY AND RELEVANCE

Khin Maung Kyi*

This is a report on an aspect of the IIMI/IFPRI performance Evaluation Project which deals with the testing of performance measures, developed or compiled under the project during the last two years, as to their applicability, consistency and relevance, on the basis of data previously collected by IIMI research projects in various countries. The Performance Evaluation Project initiated and implemented by the International Irrigation Management Institute and the International Food Policy Research Institute in its first stage includes development of both a conceptual framework and methodology for measuring the performance of irrigation systems and pre-testing of these measures for their consistency, relevancy and applicability.

IIMI as part of its own research activities has conducted the various field studies in the Philippines, Indonesia, Sri Lanka, Malaysia and Pakistan, which all include measurement of such basic data as volume of water delivered at various points, time periods of delivery, areas irrigated, crop cropped etc., in addition to other supporting data such as weather, rainfall and type of soil. It is true' that these data are. collected by various projects for its own purposes. No uniform procedure for collecting the information was followed in these projects: Location of measurement points, type of measurements used and the number of observations made are dissimilar, depending upon the need of a particular research. However, for our purpose this available information is useful enough to find out how good various measures that we have developed in this project in the last two years are in respect of their consistency and validity.

---

* The author gratefully acknowledges statistical and computational assistance provided by Tilak Kuruppuarachchi, Mrs. R. Sooriyaarachchi. Mr. Adriza and Ronald Chellappah and also gives many special thanks to his Secretary, Miss Mala Ranawake who turned his jumbled utterances from the dictaphone to a recognizable prose or always had patience to decipher hardly legible handwriting. Without her unfailing assistance, his work at IIMI would have been much restricted. He would also express his gratitude to various research project leaders who have allowed him to use their much valuable records and especially to Fred Valera and Senen Miranda who shared with him their sightful knowledge on irrigation management. He is also indebted to all those colleagues at the Institute, particularly Drs. P.S. Rao and Mark Svendsen who taught him various aspects of irrigation management. It should, however, be noted that any error, omissions or misunderstanding in the paper is entirely the author's.

PURPOSE AND METHODS

Using these data sets, various tests will be performed to find out usability of data, reliability of data inputs, consistency of results obtained. We shall first construct various measures representing the adequacy, equity and other related concepts on the basis of the data set available, and then examine the validity and consistency of these measures through suitable statistical tests. We shall also test the consistency of results thus obtained by comparing between systems and sub-systems, between systems themselves, and also between seasons and periods. This paper also will attempt, when possible, testing of validity of results by comparing with the conclusions previously reached by independent sources about the degree of achievement of results in these sampled systems. It was intended that this data will provide a basis for discussing the usability of these tests under varying conditions in different systems. For this purpose, we have selected data available from Indonesia, Malaysia and the Philippines. The data on Sri Lanka, collected in 1988 was not made available to us and, therefore, we were unable to apply the same tests to the Sri Lankan data.

This paper is essentially concerned with testing suggested measures included in the earlier work in this project, a framework or as a methodology for analyzing the performance of the irrigation systems. Among all possible measures, such as output, outcome, impact and effects,. we have proposed that for the purpose of this project we should restrict ourselves to measuring outcomes of the systems which could easily be identified or seen as directly resulting from the operation of the system, such as, adequacy of supply of water, reliability of water supply, equity in distribution of water (Svendsen). It should also be noted that the theoretical rationale and conceptual limitations of specific indices for measuring these proposed outcomes have been already discussed in a previous paper (Abernethy). The question this paper is addressing is "are these measures we have suggested or proposed consistent, relevant and useful in their actual or practical application."

As soon as we talk about measures and measurement we have to be clear about what we are discussing. Measures or indicators are devised to represent certain concepts or prepositions; for instance, the Gini index will test the level of equity achieved in distribution of water. Why is it that the equitable or equal distribution of water between different plots according to their size is considered important ? That will be the question of judgement or preference. In this paper, essentially, we will be discussing how well these indicants or measures, suggested, are related to the concepts we are supposed to measure. In other words, we are not dealing here with the correctness or otherwise of a particular preference or value judgement such as choice of the problem itself or a particular definition of equity or definition of what is adequate level of water for the crop. We are here more concerned with operationalization of the concepts already defined, problems involved in constructing measures, and also whether intended results are achieved in the process. In this regard two questions will generally follow : one is, 'how reliable/consistent the indicants or measures are' and the other is "does an indicant measure what it purports to measure'. The first, social scientists will call as the questions of reliability,, that is, replicability or repeatability, and the second, they will identify as validity of the measures. The reliability, in fact, is dealing with the consistency of the results of the measurements made. Essentially, reliable measure is supposed to give consistent or similar results over repeated observations, either in time or in space. But we will give here the test of reliability a broader meaning. To be reliable or to have consistent results in our context, the quality of data or the way it is collected, organized and presented should be taken into account. The quality of sources, consistency of coding, appearance of some random errors will all affect the reliability of the results that we are dealing with.

As we have noted, 'stability over time' is one of the important aspects of consistency. Of course, no results will be completely the same among repeated observations. For instance, we do not expect that the results of a particular measure, such as adequacy measure, between the seasons will always be the same. However, if a particular indicant or measure is reliable we should discover a fairly consistent pattern appearing. If observations are completely random, no patterns could possibly emerge. Likewise, consistency in our context also includes consistency between different presentations of the same situation, understanding'or generalizing the same phenomena spacewise or timewise. For instance, when we are talking about the adequacy of water of a particular system, there are different ways in which we can represent the situation. We can represent the situation of the whole system. For instance, that may mean the relative water supply of the whole system and calculated in one whole aggregate. Alternatively, we can also calculate a sample of sub-systems from which the total situation will be deduced. But very often, such calculations of the availability or the adequacy of water for a sample of sub-systems may not be possible because of scarcity of information. In that case, the question will be 'whether the total aggregate calculated for the whole system can be used to represent various sub-systems as well' or 'is the total system calculation a fair presentation of the consensus or the average situation of the various sub-systems.' Likewise, we may also be in a situation where we are making judgement on the adequacy of water in the whole system on the basis of a few sampled observations of parts of the system. In other words, we are dealing with the problem of consistency of results between sub-systems and the main system, inferring from a few about the whole or deducing from the whole about the parts. These we will define as a part of the test of consistency, the test of consistency over time and test of consistency between the main system and sub-system.

Validity, on the other hand, is related to how well an indicant or an indicator actually measures what it purports to measure. In other words, it is the relationship between the measurement and reality or its relevance to the reality. Do indicants really measure the reality as intended or are they measuring something else ? This is the question validity will deal with.

Theoretically, validity is represented by three different methods or three different tests. One is called a 'criterion validity.' Whether a particular measure or indicant is really measuring what it is supposed to measure can be tested by relating it to an independent criteria. For instance, the results of a psychological test of schizophrenia applied to a group of patients should be related to or verified by the independent diagnoses conducted on the same set of patients by the competent medical experts in the mental hospital. In our case, we should be able to check whether the indicator results represented by the relative water supply is actually tallying with other independent observations that can be made, such as how many days the crop goes without water, any information on the sufficiency of water by competent observers. For any measurement, one could suggest that there should be one or more numbers of independent criteria with which the results should tally. In this paper we will not be able to apply this type of test because we are not making any field observations and using the date which is only existing in the files. We are unable to, for each test we are using, provide any independent information to rest the results of ratios or indices.

The next one is 'content validity.' Content validity relates to the definition and theoretical content of the elements included in the measures, that is, whether each element in the calculation of the measures are theoretically justifiable. This, of course, can be done through conceptual discussions and wherever possible this kind of exercise will be applied. For instance, Charles Abernethy's objection on inclusion of percolation losses in the demand calculation in the relative water supply that if these are included the meaning of the relative water supply as management indicator will be changed is an example how content validity is applied.

The next is 'construct validity.' This is concerned with, essentially, whether the results of the construct that we are measuring agree with theoretically derived hypotheses concerning these concepts. In other words, any prediction that we can make, theoretically, on the basis of this concept can be verified from the information resulting from the related research studies. Such verification, if forthcoming, will strengthen the theoretical foundation of a Particular measure. At this stage, we do not have enough researches to test the agreement between theoretically derived hypothesis and empirical findings. However, one alternative method will be to find agreement or consistency between a battery of measures which are theoretically close. In other words, we could assume that if there are batteries of measures which are theoretically close to each other, they should be able to give the same kind of results. This is possible in our case because we will be testing a series of measures which are theoretically close. By testing how closely related a number of measures are in terms of their results, we should be able to add which of the measures are empirically close to each other and which are not and, we could possibly also infer whether elements other than we intend to measure are included in these indicators.

We have mentioned also that the usefulness of the measure is an important aspect of this kind of test. But, however, in the present context of this paper, we cannot deal with this question as usefulness can only be tested with the reaction of the users, information of which we do not possess. Only through accumulating the experience of .applying the suggested measures in the real situation, usefulness can really be verified. Because of this limitation, we deal mostly with relevance and consistency of the measures suggested in this paper. This relevance will be tested through either a content analysis or the construct validity, whereas consistency of the result will be tested through agreement between the different spatial conditions or the consistency between the time period.

THE FRAMEWORK FOR ANALYSIS

In this section, we will be describing the procedures for testing the reliability and validity of measures already mentioned. Here we will be dealing with two sets of measures. One set of measures concerns with the adequacy of water in terms of requirements. Requirement means getting the required volume of water at the required time. In other words, this concept will include both volume as well as timeliness. If water is delivered to satisfy the crop requirement, the elements of both time and volume is implied. The other concept we deal with in this paper is equity. The concept of equity is defined as differentials in the distribution of water which varies according to the spatial, locational, ownership type, size of the farm etc. This concept, equity, is highly value loaded. This assumes that'each farm plot has an optimum need for water and, on the basis of size, each farm plot should get their optimum requirement. Here, what is optimal and what is actually the requirement, even if these requirement calculations use the agronomic estimation of the plant needs, are still judgmental. In addition, this concept also assumes that the optimal for the farm plot override the other considerations. It should also be realized that what is optimal for each farm plot may not necessarily be optimal for the whole unit. The best arrangement of the whole unit, if its objective is to obtain the highest production, may be achieved by less equitable distribution of water between different plots. These questions involve evaluative judgement and, therefore, these measures are named as 'evaluative measures.'

In handling these various measures, we propose to classify these measures by three categories. First measure, we shall call as 'evaluative measures.' These measures will include such indicants as relative water supply *or* delivery performance ratio or planned vs. actual performance ratio and also the equity measures including various spatial uniformity indices. Either in the relative water supply or the planned vs. actual comparison of water delivery, there is always elements of judgement and evaluation. The planned water supply is judgmental on the part of the planner, *so* also is the calculation of demand in the relative water supply. In the similar vein all the equity measures are judgmental and evaluative. Spatial uniformity in distribution of water assumed as a criteria in the equity measures is judgmental and evaluative in the sense that it implies that water should go according to the area irrigated. On the other hand, there are also measures which are of a general statistical category which does. not necessarily imply any evaluative judgement. For instance, various variability measures such as averages, variances, are the measures which only tell about the observations as they exist or vary between themselves or within themselves or how variable their distributions are, and which do not necessarily imply any evaluative preference. There are also a number of such measures which fall into this category. For instance, coefficient of variation of the actual distribution of water is by itself not evaluative. However, Gini index which already assumes that the amount of water should be proportionate to the area earmarked for cultivation implies a value preference.

The third measure which you would classify, we call 'composite measures.' So far, indicators that we have been discussing, either evaluative measure or variability measures, are intended to measure a single concept, either adequacy of water or equity of water, whereas there are other measures which go into the composite effect of the water distribution. For instance, productivity measure may be named as such a concept. Of course, productivity results from water and other inputs and the measures of productivity include composite effects of these inputs. Even if this definition is restricted to causes of variation in distribution of water only, various aspects of water distribution, such as timeliness, volume, the area where it is distributed, the way this water is used will all be reflected in the productivity measure. Theil's index of performance ratio, the sum of the square of the deviation between the expected and the actual is another example of such an index as could be broken down into the various possible causal factors. The measure of profitability of the water management unit or a system also will be called as a composite measure though the effects of water distribution figure prominently in such a measure. As in the case of the Philippines, where the water distribution agencies or units or systems are evaluated on the basis of ability to collect fees and cover the costs, the profitability indices include effects of many variables other than the physical distribution of water. The reason that we should include composite measure is that all systems are not accountable for or evaluated on water distribution alone. If the system operates on the basis of the composite criteria or composite objectives, the measures themselves will also become composite measures. In other words, water distribution measure or the effect of water distribution will be. subsumed in the composite measure. Under such situation how the composite measure actually operates should be emphasized. In this paper we shall be discussing the three types of measurement and for each category of tests.

INFORMATION ON DATA SET

The data from the four studies made by IIMI Researchers in various Southeast Asian countries are used in this paper. The first two are from Indonesia each of which includes very detailed information about a particular canal and operation about its various blocks. The first one is Cikeusik system, located in West Java. Information in this data set includes a record of one canal and measurements at eleven blocks. The command area covered in this study amounts to 7,318 ha. and the principal crop types of this system are rice, sugar and palawija. The information covers three seasons from wet season of 1987/1988 to wet season in 1988/1989. Each season has twelve observation periods and each period is for fourteen days. Another system, Ciwaringin, also located in West Java, covering an irrigated area of 2,410 ha., has one canal head and six blocks. The point on the six blocks are measured for a period of three seasons as in the case of Cikeusik systems. The crops grown include mainly rice and palawija. It should also be noted that some information on the other four systems also are available but these information are not comprehensive as the two observed systems. All available information will be utilized for appropriate comparison and tests.

The next system is Kerian system in West Malaysia mainly designed and devoted to rice production. It has an irrigated area of 24,000 ha. which includes eight compartments. Both canal head and compartments are measured for the research period of six seasons from off season 1987 to main season 1989/1990. Each season covers a period of twenty two weeks.

In the case of Utris System in the Luzon island in the Philippines, the study covers one main canal which has four laterals; these points, both lateral gates and main canal gates were measured for five seasons, dry season 1986/1987 to dry Season 1988/1989. The crop type is mainly rice in wet season and rice and onion in the dry season and it has a command area of 3,900 ha. It again has the average of 25 weeks in each season.

This information is summarized in Table 1.

TABLE 1

INFORMATION ABOUT THE DATA SETS

NAME OF SYSTEM OR SUB-SYSTEM	LOCATION	COMMAND AREA COVERED Ha.	NUMBER OF OBSERVATION POINTS	CROP TYPE	PERIOD OVER
Cikeusik	West Java	7,318	11 Blocks	Rice Sugarcane Palawija	3 Seasons Wet 1987/88 to Wet 1988/89
Ciwaringin	West Java	2,410	6 Blocks	Rice Palawija	3 Seasons Wet 1987/88 to Wet 1988/89
Kerian	Malaysia	24,000	8 Compartments	Rice	6 Seasons Off 1987 to Main 1989/90
Utris	Philippines Luzon	3,900	4 Laterals	Rice Onion	5 Seasons Dry 1986/87 to Dry 1988/89

In the case of the Indonesian systems, records also include planned water delivery target which is based on the formula used by the Provincial Irrigation Departments themselves and, therefore, two sets of calculations are made, one, a comparison between the actual water delivery and the planned target, the other, the calculation of Relative Water Supply on the basis of records available. In the case of the other two systems, in Malaysia and the Philippines, no record other than records of estimated demand and the actual water supply and rainfall is available and the planned and actual water delivery performance as in the Indonesian case does not exist. Records on rainfall, acreage, duration of the seasons and records of the water flow are available and, therefore, the relative water supply ratios are calculated from these two systems. Regarding equity-related measures, various indices are calculated on the basis of available information. The records of the Cikeusik system, which includes more information points have been fully utilized. In other words, this paper uses records available to the fullest extent and the various performance measures are applied and comparisons are made between these measures.

MEASURES OF PERFORMANCE USEP

We may briefly describe the measures of performance used in this study. Detailed descriptions of some of the methods which are not commonly used is included in an appendix to explain some of the important points of these measures.

ADEQUACY MEASURES

Regarding adequacy of water distribution, two important evaluative indices are used: Delivery Performance Ratio, Relative Water Supply and basic statistical measures of variability and stability.

Delivery Performance Ratio is derived from the formula suggested and used by the Provincial Department of Irrigation itself. This is, essentially, the actual water supply divided by the planned water supply. The planned water supply is, essentially, derived from the calculation of crop water requirements as used by the Provincial Irrigation Departments and modified by local experiences, including canal management experience accumulated by respective users' departments.

Relative Water Supply measure is the simplified version of the relative water supply formula by Levine. The relative water supply equals the actual supply divided by the actual demand. In this case, the actual supply includes only irrigation and rainfall. The demand is essentially made of two elements, one needed for land preparation, the other for the crop growth, and it assumes that land preparation will take four weeks and the crop growth will take the rest of the period in Indonesia and details of the formulas are given in the appendix.

EQUITY MEASURES

In equity measures we use quite a few of well recommended measures such as Gini index and also less well known measures, such as Kolmogorov-Smirnov test (KS test). This KS test is a distribution free measure which can examine how any small sample or observation can differ from an expected or assumed distribution. Expected value may be uniform, chi square or normal distribution. In applying this measure, we use uniform or equal proportion distribution as the norm to compare with the actual. The method actually tests how different the cumulative expected frequency distribution. is from 'the cumulative actual distribution. There are'three advantages in this test. It is distribution free test in which any assumed distribution can be compared with the actual. Secondly, the test provides a statistical test of significance, and thirdly, not only it can be applicable to a small sample but also is much more sensitive than the chi-square in small sample tests. Therefore, it is well suited to our case in which we generally have only a few observation points, varying from six to eleven blocks or laterals.

In addition, we also use Gini index which measures the inequality between the different geographical units. Other measures used in equity measure include Rao's Relative. Equity Ratio and Theil's formula relating to the differences between blocks and areas, and measures suggested by Abernethy, such as coefficient of variation, interquartile and modified interquartile ratios. Quartile and coefficients of variations are well known statistical measures. Rao's index is a modification of an interquartile measure in which the pattern of distribution is also taken into account. Rao's measure divides the sample into three portions, head, tail and the middle. Dividing point depends on how precisely one wishes to define a band in which the equal ratio between the amount of water distributed and the area covered ranges. The basis of this calculation is that for each plot of land, the size of its area covered divided by the total area indicates what proportion of land each area covers. In the same way, amount of water received by one particular plot divided by the total amount of water received by the whole set of observations gives the portion of water received by a particular area'. The ratio of these two proportions provide the level of the equity in receipt of water for each land area. Assumption here is that the ratio of water received divided by the ratio of area covered should equal one. That means that everyone should receive proportionately the same ratio of water irrespective of the size or location. Rao suggests that the middle or the head or tail can be demarcated on the basis of this ratio; the area which enjoys the ratio of one being the middle and the head and the tail areas above and below this ratio respectively. The ratio is calculated for all the plots, and where this particular dividing line between head, tail or middle, falls depends upon the type of distribution of water. This, in fact, is superior to the other measure because it takes into account the shape of distribution existing in actual situation itself. And it also is simple enough to calculate using any small sets of observations. In addition to this Rao's measure; we also use the measures suggested by Theil. For each period, if we calculate Theil index, across different blocks or compartments, we can also get the variation between different blocks. This Theil ratio also can be decomposed to discover possible finer causes for variation. For testing the validity and consistency, these seven measures are introduced and compared between them for each of the systems in which the required data are available.

COMPOSITE **MEASURES**

The first measure under this category is the application of Theil's Performance Index to available data from all these countries. Theil's formula,' essentially includes'two elements, expected quantity and actual quantity. Theil's formula is a generalized formula, named Root Mean Square Prediction Error, which tests the seriousness of deviation of the actual from the forecasts by its mean squares. The formula is based on the summation of the squares of difference between the actual and the expected for a series of observations. This formula can also be decomposed to find out to what extent various possible causal factors are responsible for the variation between the actual and the expected. In this analysis, we will use not only the main formula, which is mean square prediction error, but also its decomposition. By decomposition we shall also be able to investigate the causes of variation between the blocks, that means the variation between the spatial locations, and also those of variations between different time periods which indicate ability to follow the plan time wise. This will be a very useful analysis, especially when we wish to find out both timeliness and the adequacy of water distributed. In addition to this index, the aggregate sum of Theil's indices measuring variations of each of the blocks also will be utilized here to see the usefulness of his model as a whole.

In addition, as we have explained, we may also use as a composite measure profitability index as defined in the studies on the Philippines which essentially is total revenues divided by the total cost of the systems in question. Revenue will include both collection of fees as well as other revenues such as hire of vehicles, equipments etc., whereas cost includes out of pocket maintenance costs as well as salary costs of the maintenance employees.

. Some of the well known measures, such as Gini index are not explained but some of the less well known measures, such as Theil's index and KS index are explained briefly in the appendix.

TEST METHODOLOGY

In developing the tests of consistency and validity, we use various statistical tests, such as difference between means and analysis of variances and also other measures which test the variability of the distribution, such as coefficient of variation which does not imply any evaluative standard as in the case of relative water supply. These measures do not imply any preset standards. However, it tells the characteristics of the distribution or variability of each set of data, or comparison between sub sets of data and, therefore, in many cases these tests also will provide an analysis which can test the variability and reliability of each data set as well as consistency between data sets or sub sets of the same data set. When different measures are compared, we also utilize a series of simple of correlation or inter-correlation table among these measures to test how well each measure relates to the other. Also applied are well known tests of consistency such as Alpha efficients and other indices. It should also be noted that all tests of reliability and consistency are derived from the classical test score theory.

FINDINGS ON THE TESTS

ADEQUACY AND TIMELINESS MEASURES

Regarding these two concepts, Relative Water Supply and Delivery Performance Ratio are mainly discussed in this section. However, it must be realized that both the relative water supply and the delivery performance ratio essentially present the adequacy or otherwise of water for satisfying crop water requirements for particular time periods. If it is the balancing of the demand and supply for a whole time period such as for a season, it could only tell a crude approximation of the adequacy. To answer the timeliness, more refined period by period or week by week data will be needed. In addition, while relative water supply is easily calculable with existing data, though with some degree of approximation, delivery performance ratios are available in Indonesia because the respective irrigation agencies actively plan their own water allocation in these observed systems. Water management agencies, provincial irrigation agencies, actively make seasonal planning and estimation of water for each planned system. The ratio, the planned vs. the actual, provides an alternative measure to compare with the relative water supply. The reason for the use of relative water supply for measuring the adequacy of water is well justified. This measure takes the crop water requirement into account as a basis of demand'. Depending on the level of refinements of the formula as well as the quality and extent of information available, both demand and supply could be estimated. When this calculation is broken into week by week basis, the timeliness factor can also be taken into account. These two measures are available in Cikeusik and Ciwaringin systems in Indonesia. Only relative water supply could be calculated in the case of the Philippine and Malaysian systems. The question of consistency between various tests and the representativeness and interpretability at different levels of analysis, such as system or sub system level, will be discussed by applying various tests to these two measures.

A. CONSISTENCY BETWEEN DIFFERENT MEASURES

This question as we mentioned earlier answers partially the validity as well as relevancy of these measures in actual situations. We have hypothesized that conceptually close tests designed to measure the similar or same criterion should definitely give similar results. The affirmation of this hypothesis will strengthen the validity or the purpose of the tests themselves. Table 2 presents the correlations between relative water supply and OPR measures for six irrigation systems in Indonesia.

TABLE 2

CORRELATION BETWEEN RELATIVE WATER SUPPLY AND DELIVERY PERFORMANCE RATIO IN SIX DIFFERENT SYSTEMS IN SYSTEMWISE COMPARISON

SYSTEM	Command Area	No. of Periods Observed	Correlation Coefficient	Level of Significance
Bantarwangi	535	*14*	0.46	.001
Cikeusik	7,318	76	0.99	.001
Ciparigi	295	74	0.91	.001
Cipikul	489	71	0.88	.001
Cittanggulun	873	76	0.86	.001
Kunigan	517	66	0.93	.001

We are fortunate to have observations of 76 weeks each for these systems. Each point of observation represents total system performance for each indicator and the detailed breakdowns from the sub-systems on laterals for all systems included are not available. The correlation between RWS and DPR is calculated systemwise for each respective set of observations. The results strongly suggests that both measures are similar and consistent between themselves, represented by very strong correlation efficients. It should also be noted that though DPR ratio is a comparison between the planned and the actual in contrast with the RWS, which is the actual supply divided by demand based on actual water requirements, elements such as acreage, formulae for crop water requirements, stage days etc. used in the calculations of the demand in the RWS and planned targets in the DPR very similar. The estimates themselves on the other hand are

BLOCK	Command Area (Ha.)	No. of Periods Observed	Correlation Coefficient	Level of Significance
BL I Kr	85	36	0.972	.001
LS II Kr	35	36	0.960	.DO1
MTR V Kr	43	36	0.410	.001
MTR 7 Tg	71	36	0.425	.001
PB 3 Kn	39	36	0.944	.001
PB 5 Kr	72	36	0.932	.001
PB 8 Kn	63	36	0.724	.001
PG 1 Kn	239	36	0.963	.001
BAT III Kr	12	34	0.980	.001
JTS 3 Kr	100	35	0.958	.001
SR 3 Kr	16	36	0.941	.001

TABLE 4

CORRELATION BETWEEN RELATIVE WATER SUPPLY AND DELIVERY PERFORMANCE RATIO IN CIWARINGIN SYSTEM FOR EACH BLOCK FOR WET SEASON 1987/1988, DRY SEASON 1988, WET SEASON 1988/1989

BLOCK	Command Area (Ha.)	No. of Periods Observed	Correlation Coefficient	Level of Significance
MC 2 Kn	123	36	0.953	.001
NP 4 Kn	30	36	0.972	.001
W 2 Kn	15	36	0.925	.001
W 4 Kn	45	36	0.923	.001
WLN 1 Kr	33	36	0.983	.001
WLN 3 Kr	34	36	0.977	.001

tests of consistency as used in social sciences: tests of item consistency or item convergence so routinely applied in social studies will not be applicable. However, if the repeated measurement of water adequacy ratios over a number of periods for different outlets could be treated as a single test and the results of RWS or DPR tests between different outlets for a particular period or between different periods for an outlet can be considered as sub-set of a single test, the consistency between various sub-sets of the measure, e.g. the consistency of results of different periods, when compared between blocks or of those of different blocks compared between different periods, will in essence approximate Cronbach's idea of split-half method or the test of equivalence. It may also be taken as a test of repeatability over a number of observations. With that idea in mind, we attempt to apply measures of internal consistency for each of the two indices (DPR and RWS) in this section.

The first statistical test we suggest for that purpose is coefficient of variation. The coefficient of variation will test the degree or the variability of a particular population with respect of a certain characteristic, in this case, the distribution of amount of water delivered over a number of outlets. By comparing the coefficient of variation of the relative water supply or the delivery performance ratio between time period or between sub-systems for a season we shall be able to

1. Draw the conclusion if patterns of distribution between these two measures (DPR and RWS) differ from each other.

2. See whether the results of each measure by itself differ between sub-systems or between periods in one season or across the seasons.

These will be very useful because by knowing the comparative patterns of distribution between two measures for a particular season or patterns of distribution of a single measure across the seasons between sub-systems or between periods, we should be able to see how stable or how variable is the distribution of the results of these measures. In a way, this will give us the limits within which we have to interpret the results of these two measures. Table 5 indicates the comparison of coefficients of variation of Relative Water Supply and Delivery Performance Ratios by season and by block for Cikeusik system, averaging period results for each block.

TABLE 5

COMPARISON OF COEFFICIENT OF VARIATION OF RELATIVE WATER SUPPLY AND DELIVERY PERFORMANCE RATIO RESULTS BY SEASON AND BY BLOCK
CIKEUSIK SYSTEM

SEASON		BL1Kr	KTR5Kr	LSIIKr	KTR7Tg	PB3Kn	PB8Kn	PB5Kr	PG1Kn	BATIIIKr	SR3Kr	JTS3Kr
Yet 1987/1988	RWS	0.126	0.233	0.166	0.238	0.081	0.043	0.193	0.270	0.086	0.093	0.147
(Dec-May)	DPR	0.087	0.418	0.930	0.194	0.170	0.038	0.331	0.236	0.079	0.911	0.093
Dry I 1988	RYS	0.243	0.129	0.132	0.248	0.265	0.210	0.182	0.471	0.124	0.154	0.068
(May-Sept)	OPR	0.198	0.938	0.156	0.518	0.135	0.296	0.214	0.545	0.193	0.114	0.632
Dry II 1988	RWS	0.158	0.130	0.171	0.483	0.013	0.091	0.161	0	0	0.385	0.111
(Sept.-Nov)	DPR	0.497	0.074	0.396	0.105	0.586	0.135	0.105	0.587	0.412	0.250	0.149
Dry 1988	RWS	0.214	0.649	0.159	0.423	0.231	0.160	0.187	0.411	0.150	0.279	0.093
[Hay-Nov)	OPR	0.339	0.115	0.213	0.516	0.519	0.288	0.215	0.646	0.352	0.184	0.588
Wet 1988/1989	RWS	0.053	0.078	0.782	0.187	0.117	0.059	0.118	0.213	0.277	0.113	8.068
(Dec-April)	DPR	0.095	0.163	0	0.131	0.103	0.035	0.226	0.266	0.302	0.145	0.351

In Table 5, by looking at the data for the 11 blocks for 4 seasons, we find that large coefficients of variation appears mostly for the DPR in the dry seasons. In other words, the DPR in dry season is likely to have a wide variability, whereas the relative water supply is more or less stable measure and its variability Is within a very narrow range. The reason also could be inferred from the nature of the two measures: whereas relative water supply is based on the events that have already happened, that is, the supply that has been made and the demand calculated on the basis of actual acre grown and actual crop grown. The DPR ratio, on the other hand, is a comparison between what it was planned a few months ago and what actually happened in the real situation. Therefore, there could be a wide variation between the planned and actual supply when there is a shortage of water in a particular dry season and consequently what was expected can be very much different from what actually has happened. This observation is confirmed in Table 6 as well, which indicate a similar comparison for Ciwaringin system.

TABLE 6

COMPARISON OF COEFFICIENT OF VARIATION OF RELATIVE WATER SUPPLY AND DELIVERY PERFORMANCE RATIO RESULTS BY SEASON AND BY BLOCK
CIWARINGIN SYSTEM

SEASON		MC5Kn	NP4Kn	W2Kn	W4Kn	WLN1Kr	WLN3Kr
Wet 1987/1988	RWS	0.156	0.115	0.084	0.101	0.094	0.123
(Dec-April)	OPR	0.208	0.146	0.120	0.109	0.129	0.247
Dry 1988	RWS	0.177	0.172	0.160	0.087	0.111	0.072
(May-Nov)	DPR	0.688	0.141	0.256	0.109	0.170	0.318
Wet 1988/1989	RWS	0.163	0.039	0.098	0.182	0.314	0.665
(Oec-April)	DPR	0.259	0.158	0.276	0.121	0.220	0.057

Therefore, we may conclude that the wide variability of the DPR ratio in the dry season is caused by the very nature of the measure itself and also the likelihood of water shortage in the dry season as different from what was expected. Apart from this variation, we find that for both systems, the coefficient of variation are mostly below 0.5 or 0.4 or below 40% in most cases. It could be concluded that the average or means move within a a narrow range, indicating the stability of representativeness of the means.

On the other hand we can analyze the variability of the water distribution between different blocks or outlets over the various seasons by looking at the data for each type of measure only. In the same Table 5, we find that for Relative Water Supply, coefficients of variations are higher for blocks MTR5KR, MTR7Tg, PGlKn for three seasons in contrast to other blocks which have relatively much smaller variations over these seasons. This suggests that coefficients of variation of the means of these blocks comparatively follows a more or less similar pattern over the seasons. Likewise, the coefficients for variation of the DPR for each block when compared against those of other blocks between different seasons, blocks MTR5Kr, MTR7Tg, PG1Kn, have relatively much higher coefficient of variation in three or four seasons. Of course, as noted before, the DPR values have more volatile variations than the RWS values in every season, more pronouncedly in the dry season. This whole finding clearly indicates that the seasonal means values of each test or measure for each block follow the consistent patterns of distribution or have the same pattern of variabi1ity.

As noted earlier, the test of variability of the distribution does not guarantee the internal consistency of the test but only give a partial support for the comparative stability of results (Mean Values) obtained for each type of measure. Another alternative test of internal consistency of a particular measure will be the inter-correlations of means values of various blocks between different time periods. The assumption here is that the pattern of water received to 11 blocks for one period should be similar to the pattern of water received for these 11 blocks in another period of the same system for another period. In other words, a high average correlation between different time periods will suggest the same pattern of distribution between blocks irrespective of time periods. Conversely, this test of stability could be calculated between different time periods for a number of blocks. It is assumed that each set of mean values of DPR or RWS for one period is a subset or a part of the whole test, a strong average correlation between observations of various sets of values will approximate the Cronbach's test of equivalence or stability. We apply this test of internal consistency to Cikeusik System which has sufficient number of subsets by blocks (11 in number) and by periods (12 in number). Assuming that 11 blocks have the same pattern of water distribution between them for a number of periods, we could compare or correlate the DPR or RWS of 11 blocks of one period with those of these 11 blocks of another period. The resulting inter-correlations can be averaged and the Cronbach's Alpha could be calculated. The internal consistency of the whole measure could be examined. Table 7 indicates number of average of inter-correlations between periods for each season for System Cikeusik. The results indicated that at least for one season, Wet Season 1987-88, both DPR and RWS have a strong average inter-correlation with high enough Alpha scores between periods. That means the water distribution between 11 blocks follow the consistent patterns in 8 and 6 observed periods respectively for DPR and RWS ratios. Treating one season results as representing the whole test, in Cronbach's sense, the consistency between subsets of the season indicates a strong reliability of the whole test.

TABLE 7

NUMBER AND AVERAGE OF INTER-CORRELATIONS BETWEEN PERIODS FOR EACH SEASON FOR CIKEUSIK SYSTEM

	Wet Season 1987/1988 Dec-May		Dry Season 1988 Jun-Nov		Wet Season 1988/1989 Dec-Apr	
	DPR	RWS	DPR	RWS	DPR	RWS
1. Number of periods compared. *	8	6	8	5	11	7
2. Number of inter-correlations included. **	28	15	28	10	55	21
3. Number of significant inter-correlations.	16	9	4	0	4	3
4. Average values of inter-correlations.	0.66	0.61	0.16	0.31	-+	-+
5. Coefficient Alpha. ***	0.93	0.72	0.61	0.55	-+	-+

* Excluding incomplete sets (periods).

** Excluding outliers or including only periods forming a correlated diagonal.

*** Chronbach's Alpha:

$$\alpha = \frac{a}{a-1}\left[1 - \frac{a}{a+2b}\right]$$ when the correlation matrix is used.

a = number of indicants (subsets) in the composite.
b = sum of the correlations among the indicants.

\+ No pattern is discerned.

However the results from other two seasons are not significant for both DPR and RWS measures. Unlike item analysis, the results of one subset (one period) could differ from those of another subset (another period) for a number of reasons, such as, managerial interference, failures of physical structure etc. In spite of this limitation, a high positive inter-correlation, when it appears, lends a strong support to the internal consistency of the results.

This test was not applied to Ciwaringin System as there are not enough blocks - only six available in that system. It should also be noted that we have excluded some period results and also outliers which .show very much divergent results from the emerging main pattern, particularly the early or late periods in the season in which water distribution may not be very uniform or regular. What is finally included is a diagonal in the inter-correlation matrix which clearly forms a distinct and correlated set.

The results of the present test could not be considered as conclusive with regards to the internal consisistency. However, with strong positive set of correlations at least in one season, the potential usefulness of this approach could easily be recognized.

This kind of comparison is not possible for the other two systems because we could only calculate the relative water supply for the two systems, Kerian system in Malaysia and the Utris system in the Philippines. As far as we could judge from the Indonesian experience, we should conclude that relative water supply and the DPR give fairly consistent results. We also observe that the patterns of variability between the two measures are more or less the same except that the DPR ratio is likely to vary widely in dry season where the water shortage is likely to occur. The low variability as well as the same pattern of variability of means values of each measure over the seasons probably suggests a strong internal consistency of these measures, as further supported by inter-correlation matrix.

C. THE CONSISTENCY BETWEEN THE MAIN AND SUB-SYSTEM RESULTS

If we were to design the measures which would be useful to the managers, any other set of practical users, the first thing that one would realize is that detailed information is not available in many systems either for the relative water supply'or the planned and actual delivery. Only the system level information is easily available. The total acreage of the system, the average of the number of cropping days, the cropping period etc., will be available at the system level only. But as soon as one goes down to the sub-system level, literal or distributory level, the information becomes scanty: If we were to develop intended measures with the available information only, we may be able to calculate them for the system level only. The question is 'how useful is the system level information in representing the total situation of the system ?' Many will say that the system level aggregation is too crude to tell what is happening down at the sub-system or distributory levels. One may even question the wisdom of using the aggregate information to calculate the system levels ratio without even checking its consistency at least with a sample of calculations of the same at sub system levels. These questions are serious. How representative is the system level calculation in comparison with the situation in the sub-systems ?' In other words, both the consistency of results between the main system and the sub-system level calculations and the subsumeability of sub-system results in the main system results become one of the important issues in .evaluating the usefulness or the relevance of the measures that we suggest.

In the next few tables we are analyzing this particular question. First, let us look at Table 8 which is intended to test whether the sub system averages differ very greatly from the overall system calculation.

TABLE a

COMPARISON OF OVERALL SYSTEM WATER DELIVERY VALUES, AVERAGE OF BLOCK MEANS AND EACH BLOCK MEANS FOR CIKEUSIK SYSTEM

SEASON	Overall System Valuer	Avg. of all block means	C.V	S.D.	MEANS OF TERTIARY BLOCK BL1Kr	SR3Kr	JTS3Kr	MTR5Kr	MTR7Kr	PB3Kn	PB5Kr	PB8Kn	LS2Kr	BAT3Kr	PG1Kn
Yet 87/88															
RYS	1.146	1.052	10.8	0.11	1.037	0.99	1.066	1.016	1.376	0.987	1.087	0.995	0.997	1.061	0.968
OPR	1.155	1.211	20.2	0.24	1.01	1.45	1.000	1.242	1.338	1.624	1.048	0.991	1.576	1.022	0.975
Dry I 1988															
RWS	0.865	0.987	16.6	0.16	0.859	0.879	0.991	1.093	0.993	0.884	1.094	0.873	0.923	0.866	1.407
OPR	0.859	0.917	27.5	0.25	0.815	0.913	1.276	0.324	0.683	0.986	1.069	0.822	0.993	1.003	1.142
Dry II 1988															
RYS	0.924	0.767	49.9	0.38	0.91	0.951	0.925	0.988	1.013	0.912	0.915	C.822	1.004		
DPR	0.928	0.952	32.8	0.31	0.68	0.983	0.864	0.882	1.116	1.588	0.88	1.226	1.160	0.589	0.503
Wet 88/89															
RWS	0.973	1.022	5.1	0.05	0.993	1.004	0.994	0.944	1.096	1.035	0.954	0.948	1.062	1.105	1.033
OPR	0.912	1.096	14.4	0.15	0.99	1.073	0.992	1.147	1.349	1.033	0.948	0.949	1.431	1.049	1.086

C.V.	Coefficient of Variation
S.D.	Standard Deviation

In this table, three elements are shown. One is the calculation of relative water supply and DPR of the system as a whole. In this calculation, the total system acreage for particular crops, requirements of water for each of the crops grown in the system and the total supply of water for the system are included whereas means of the each tertiary block is based on the calculations of the relative water supply and the OPR of various periods for a particular block. An average of all the block means is also shown. .In addition, for each season and for each category, such as the relative water supply, its respective standard deviations and coefficients of variation also are shown. From this table we could see that the overall system values and the average of the block means are very much similar. Values of these two categories, the total system values and average of the block means for each of the season and for each category of measures are very close. Again, the coefficients of variation of these eleven blocks' means, and their respective averages show that there is very small variability in the distributions of block means for each season and for each indicator type. This strongly suggests that, as far as this system is concerned, the overall system value fairly represents the individual sub-system results with regard to relative water supply and the delivery performance ratio. It is certainly reassuring that the overall system value calculated on the total acreage and total crop water requirements and total water supply is very much similar to what is obtained as similar values at the sub-system level.

Another test of the consistency between the main system results and sub-system results is the one way analysis of variance which tests the means of blocks in relation to the total system means.

Tables 9 and 10, showing the results of one way analysis of variance between blocks for each system for each indicator type, make it possible to examine whether means of the different blocks are differing from each other very significantly and also whether block means are different from the grand mean statistically.

TABLE 9

ONE WAY ANALYSIS OF VARIANCE OF RELATIVE WATER SUPPLY AND DELIVERY PERFORMANCE RATIO BETWEEN BLOCKS FOR EACH SEASON FOR CIKEUSIK SYSTEM

RELATIVE WATER SUPPLY

SEASON	BETWEEN BLOCKS		WITHIN BLOCKS		F VALUES
	Degree of Freedom	Mean Square	Degree of Freedom	Mean Square	
Wet 87/88	10	0.139	98	0.045	3.068**
Dry 1988	10	0.147	94	0.117	1.258
Wet 88/89	10	0.019	86	0.030	0.647

SEASON	BETWEEN BLOCKS		WITHIN BLOCKS		F VALUES
	Degree of Freedom	Mean Square	Degree of Freedom	Mean Square	
Wet 87/88	10	0.735	120	0.832	0.882
Dry 1988	10	0.484	121	0.209	2.314*
Wet 88/89	10	0.299	120	0.268	1.114

RELATIVE WATER SUPPLY

SEASON	BETWEEN BLOCKS		WITHIN BLOCKS		F VALUES
	Degree of Freedom	Mean Square	Degree of Freedom	Mean Square	
Wet 87/88	5	0.037	63	0.014	2.580*
Dry 1988	5	0.051	40	0.022	2.269
Wet 88/89	5	0.094	46	0.213	0.444

DELIVERY PERFORMANCE RATIO

SEASON	BETWEEN BLOCKS		WITHIN BLOCKS		F VALUES
	Degree of Freedom	Mean Square	Degree of Freedom	Mean Square	
Wet 1987	5	0.055	66	0.111	0.496
Dry 1988	5	0.290	66	0.185	1.565
Wet 1988	5	0.071	66	0.052	1.369

the individual block means do not differ from each other. For the Water Delivery Performance Ratio, except for the dry season 1988, all the rest of analysis variance results indicate that the individual block means in both systems are not statistically different between themselves. A similar finding follows in Ciwaringin System as well.

Both of these findings strongly suggest that as far as the Ciwaringin and Cikeusik systems are concerned the calculations of the relative water supply or the DPR on the basis of total system basis will represent what is attained in each of the blocks. This provides a strong reason that total system calculation can be utilized. However, it is still difficult to make very generalized statements without knowing in each case the average or the means of each of the blocks and its variability of the observations for each season between these blocks. It will be difficult to generalize form a single piece of evidence. This finding will have to be considered true only to these particular systems for particular periods. However, one of the advantages of this test is that with a sufficient number of observations in each block, this test could be performed for a small. number of blocks or distributories, which will be the case in the most of our available data sets.

BY calculating these tests on the available information on a few number of selected blocks in the total system, we could make a judgement on the useability of these measures as a representative of the system situations. As we are using only a few sample blocks in each of the systems, a fair guess about the representativeness of this finding can be made by analysis of variance or the test of difference between means.

EQUITY MEASURES

To test whether any water distribution pattern is "equitable or fair" relative to the size of each farm area, a number of statistical procedures are available. Aberenthy (1986) compiled a number of equity measures and tested their relative merits. Rao (1987) modified the test of tail and head differences, the modified interquartile ratio, by Abernethy by taking account the patterns of the distribution itself. All these measures are essentially the equality measures, based on the concept that each hectare in a system should ought to receive the same amount of water. Without going to the rationale of concept itself, this paper will attempt to examine all these equality measures in relation to their consistency, validity and relevancy.

CONSISTENCY AND VALIDITY

Assuming that all these measures" represent the test of spatial equality in the distribution of water, the concurrence or relatedness of results of these tests in a number of data sets will make a strong case for the consistency of the results and, implicitly, the conceptual validity. In this exercise the following measures are included -

1) Coefficient of variation.
2) Interquartile ratio.
3) Modified interquarter ratio.
4) Rao's Relative Equality Ratio.
5) Gini index.
6) Kolmogorov-Smirnov test.
7) Theil's index.

All these tests are very similar to each other in their essential characteristics or are derived from the similar principles. Both interquartile ratios, Rao's Relative Equality Ratio, are different ways of apportioning into low and high groups and comparing their relative importance. Rao's ratio takes into account the character of distribution itself when apportioning into high, middle and low groups. In his case, the middle is one which enjoys the exact proportion of water that ought to be. Gini index is similar in nature or very close to above three ratios because it is comparing the ratio of the cumulative area to the cumulative volume, that is, what it should be if everyone is to receive on equal proportion basis, with the actual ratio of cumulative area and volume at each successive plot arrayed from low to high ratio. The Kolmogorov-Smirnov test is a non-parametric test of actual cumulative proportions against theoretical cumulative proportions. The concepts are derived from the same principles. The coefficient of variation tests the variability of the average or the means. In Theil's index of Root Mean Square Prediction Error, the sum of the square of the difference between the planned or expected and the actual is the basis of calculation. Rao's Relative Equity Index, the desirable proportion is one, that is, the proportion of area and the proportion of water received should be equal so that the volume of water distributed be proportional to the area of the plot. The relative equity ratio being the relative proportion of water received at the head divided by the relative proportion of water received at the tail. Anything beyond one moves away from the desirable ratio. On the other hand, Gini index will be shown as proportion of one, that is, the extent of degree of inequity that exists. If all are receiving the same proportion, then Gini index will equal one as in the case of Rao's ratio. D statistics also is calculated similarly but as it takes only the largest value and applies the significant test on the largest value, it is usually different from the other values. In the case of Gini index it is the summation of various proportion of water received relative to the area deducted from the most desirable distribution.

TABLE 11

SIMPLE STATISTICS OF VARIOUS EQUITIVE MEASURES FOR CIKEUSIK SYSTEM

VARIABLE	N	MEAN	STD. DEV	SUM	MINIMUM	MAXIMUM
COVAR	36	78.963956	23.155175	2878.702410	46.238900	126.208400
INTQR	36	5.858247	6.813229	210.896901	1.821918	39.000000
MINTPR	36	5.515720	6.454606	198.565937	0.066667	28.666667
RAOS	36	6.947417	3.900001	250.107000	2.705000	17.086000
GINI	36	0.481139	0.110783	17.321000	0.322000	0.666000
OSTAT	36	0.216583	0.098953	7.797000	0.103000	0.450000
THEIL	36	0.216882	0.194971	7.807740	0.022500	0.812290

LEGENDS

COVAR - COEFFICIENT OF VARIATION
INTQR - INTER QUARTER RATIO
MINTQR - MODIFIED INTER QUARTER VALUES
THEIL - ROOT MEAN SQUARE ERROR TEST

RAOS - RAO'S RELATIVE EWITIVE RATIO
GINI - GINI COEFFICIENT VALUES
DSTAT - KOLMOGOROV-SMIRNOV TEST

A comparison of the characteristics of seven categories of measure shown in Table 11 for the Cikeusik system indicates that, as noted by Abernethy, coefficient of variation, interquartile and the modified interquartile ratio behave more or less in the similar way. Rao's index and Gini index also follow the same principle – the equal proportion of water. Gini varies between 0 to 1, while the Rao's index will move beyond one or times of one. The table indicates the results of seven indices calculated for Cikeusik system. To appreciate the consistency of these results, we have calculated inter-correlation between the seven sets of calculations for Cikeusik system. This inter-correlation is shown in Table 12.

TABLE 12

THE INTER CORRELATION BETWEEN DIFFERENT EQUITIVE MEASURES FOR CIKEUSIK SYSTEM

Pearson Correlation **Coefficients/Prob** > |R| under Ho: Rho=0 / N = 36

	COVAR	INTPR	MINTPR	RAOS	GINI	OSTAT	THEIL
COVAR	1.00000 0.0	0.45216 0.0056	0.80291 0.0001	0.11319 0.0001	0.94285 0.0001	0.79940 0.0001	0.65100 0.0001
INTPR	0.45216 0.0056	1,00000 0.0	0.62821 0.0001	0.80978 0.0001	0.53711 0.0007	0.30713 0.0679	0.26134 0.1237
MINTQR	0.80291 0.0001	0.62821 0.0001	1.00000 0.0	0.79581 0.0001	0.76234 0.0001	0.78268 0.0001	0.56384 0.0003
RAOS	0.11319 0.0001	0.80978 0.0001	0.19581 0.0001	1.00000 0.0	0,84610 0.0001	0.59235 0.0001	0.59138 0.0001
GINI	.0. 94205 0.0001	0.53711 0.0007	0.76234 0.0001	0.84670 0.0001	1.00000 0.0	0.77517 0.0001	0.70116 0.0001
DSTAT	0.79940 0.0001	0.30173 0.0619	0.78268 0.0001	0.59235 0.0001	0.71511 0.0001	1.00000 0.0	0.48388 0.0028
THEIL	0.65100 0.0001	0.26134 0.1237	0.56384 0.0003	0.59138 0.0001	0.10716 0.0001	0.48398 0.0028	1.00000 0.0

LEGENDS

COVAR - COEFFICIENT OF VARIATION
INTPR - INTER QUARTER RATIO
MINTPR - MODIFIED INTER QUARTER VALUES
THEIL - ROOT MEAN SQUARE ERROR TEST

RAOS - RAO'S RELATIVE EQUITIVE RATIO
GINI - GINI COEFFICIENT VALUES
DSTAT - KOLMOGOROV-SMIRNOV TEST

In this table it is very striking that correlations between the coefficient of variation and four measures, interquartile, Rao's Relative Equity Ratios, Gini index and D statistics are very strong. In other words, coefficient of variation itself is a very good measure, consistent with the other four measures. **Rao's** index, on the other hand, is correlated with coefficent of variations, interquartile and modified quartile and Gini index very strongly. Again, Gini index is another very focal,measure which correlated very strongly with coefficient of variations, interquartile ratio, Rao's index, D statistics and Theil's index. D statistics is strongly correlated with only three measures; these are coefficient of variation, interquartile and Gini index. Theil's measure correlates *only* with the Gini index. On the basis of the number of high correlation attained by each of these measures, we made the following distribution in order of their importance.

NUMBER OF CORRELATIONS WITH OTHERS HAVING THE COEFFICIENT OF .7 OR ABOVE BY EACH MEASURE

Gini	5
Rao's	4
MINTQR	4
Coefficient of Variation	4
KS Test	3
Theil's index	1
Interquartile	1

You will note that the Gini ratio as having the highest number of inter correlations, having five correlation, with coefficient of 0.7 and above with others, whereas coefficient of variation, the Rao's index and the modified interquartile ratio, each have four high correlations and D statistics have three high correlations and Theil's index and interquartile ratio has one each. In other words, in terms of high inter correlations, the Gini index tops the list, trailed by Theil's and the interquartile ratio as the last. On the whole, this table indicates that internal consistency between these measures, except Theil and interquartile indices, is very high. Among them, modified interquartile, Gini, coefficient of variation, Rao's are the most inter correlated indices. The internal consistency test of coefficient alpha is performed as shown in Table **13.** The coefficient Alpha or Cronbach's Alpha tests the internal consistency or equivalence of tests involved in the measurements. **It** is equal to the average of all the possible split-half correlations. With all the indices included the alpha index reaches .93. When we exclude less correlated items like D statistics and Theil index, the rest of the items still have an alpha value of .9 and this strongly suggests that use of Gini, Rao or modified interquartile range will give a satisfactory result or very similar results in measuring the equity of water distribution. **It** should be added that similar but slightly weaker results are obtained when the same test is applied to the Ciwaringin System.

A strong inter correlation between Gini, Rao's index, and the modified interquartile ratio and the coefficient of variation probably results from a conceptual closeness of these measures. Except the coefficient of variations, all others are derived from the principle that each plot should enjoy the same proportion of water as others relative to the size of the area. Each ratio indicates the degree of relative inequity that exists in the system. As we have already hypothesized, a high correlation between conceptually close measures lends support to the validity of these measures. In addition, the high consistency of all results over 36 periods between each set of observations or ratios suggests some regularity or stability of the results of these measures between different time periods.

TABLE 13

TEST OF INTERNAL CONSISTENCY OR EQUIVALENCE: RESULTS OF COEFFICIENT ALPHA

	NUMBER OF TESTS INCLUDED	RESULTS OF COEFFICIENT ALPHA
All of the tests	7	.931
Excluding Theil's	6	.935
Excluding Theil's and Dstat	5	.933
Excluding Theil's, Dstat and Gini	4	.907
Excluding Theil's, Dstat, Gini and Rao's	3	.835
Excluding Theil's, Dstat, Gini, Rao's and MINTQR	2	.622

(1) $$1/n \sum_{i=1}^{n} (P_i - A_i)^2$$

Where P = the predicted; A = the actual and n = the number of pairs observed.

When we take the square root (Root Mean Square Prediction Error), the formula becomes -

(2) Inequality coefficient = $\sqrt{\dfrac{\Sigma (P_i - A_1)^2}{\qquad}}$

This gives a variation which began from zero and one. This value would be zero if and only if the forecast of the plan is perfect, that means, the plan equals the actual. In that case, the ratio will be zero, that means a perfect performance. On the other hand, the value will be equal to one, if naive no-change extrapolation is assumed. In other words, no change in the actual situation is expected and the plan is not made. In that case the numerator and the denominator will be the same and, therefore, the value of the index will equal one. However, this index will have no upper bounds. If a planned amount or expected amount is relatively very large to the actual, the value can be more than one. The main advantage of this ratio is that it will give, for a series of observations, the average of the difference between the sum of the square of the difference between the actual and the planned or the degree of attainment of planned targets. As long as there is an assumed or planned target, comparing with the actual and the target through this index will give us a very good average situation of the system performance.

In addition, one of the main advantages of this method is that this value can also be decomposed into the various elements. The formula is written below, in which the left hand side of the formula is the average of the sum of the difference between the square of the difference between the actual and the planned.

(3) $1/n\ I\ (P_i - A_i i)^2 = (\bar{P} - \bar{A}) + (S_P - S_P) + 2(1 - r)\ S_P S_A$

where $\bar{P}$ and $\bar{A}$ are the means: $\bar{P} = l/n\ \Sigma\ P,$ and $\bar{A} = 1/n\ \Sigma\ A,,$

S_P and S_A the standard deviations:

$$S^2_P = 1/n\ I\ (P_i - \bar{P})^2,\ S^2_A = 1/n\ \Sigma(A_i - \bar{A})^2$$

and r the correlation coefficient of predicted and realized changes:

$$r = 1/n\ I\ (P_i - \bar{P})\ (A_i - \bar{A})/S_P S_A$$

If we have a set of observations between the planned and the actual we shall get on the left hand side the average of how the system as a whole performed in respect of difference between the actual result and the planned. On the right hand side this formula equals two three terms, the first term is the square of the difference between the means of planned and the actual. And the next term will be the square of difference between standard deviations. The third term will be the one minus r in which r is the correlation coefficient between the predicted and the realized or between the planned and the actual. Others in the third terms are constant. We could interpret these three terms for the management point of view. Suppose if we have an array or table in which as we have as in the Cikeusik, the columns represent 11 blocks and the rows represent the number of weeks that we have observed in a season. If we have, in each cell, the planned and the actual for each block and for each period, this formula can be calculated across the blocks. That means the difference of the planned and the actual and the square of these differences will be added across the eleven blocks for each week and averaged. This will give us an average of prediction error for each period between the eleven blocks. It will give us how actual and planned are achieved for a particular week between eleven blocks. This average achievement for 11 blocks is the first approximation of the total of performance blockwise.

If we decompose this figure, the first term on the right hand side will be the square of the difference between means. Since we are summing up the blocks or the geographical area, the difference between the means, the mean of the actual and the mean of the planned will give us the average of the situation regarding water distribution. That will tell us roughly for eleven blocks how adequate was the average availability of water. This can be interpreted as the adequacy of water for that particular period between the eleven blocks. The next term is difference between two standard deviations. Each standard deviation tells us the variability of observation between eleven blocks respectively for each of the planned targets and the actual performance. Difference between these two variabilities will tell us how planned inequity differs from the actual inequity. The standard deviation of the planned can be taken as the planned variability or the planned inequity. Likewise, the standard deviation of the actual can be interpreted as the actual variability or the actual inequity. On the other hand, the third term, one minus r will become zero if the planned and the actual vary completely proportionately. In that case, r will be equal to one and, therefore, the third term will become zero. If the third term is zero, that means that management follows its original plan very closely. Although the amount of water may be less, proportionately each block will receive the same proportion as planned. This will tell the management ability to follow how closely its own plan is followed.

This gives us a very useful indicator for the management performance. It includes the adequacy of water on the average, it indicates how the actual inequity or variability between the block differs from the expected variability or originally planned inequity. In one formula, this will take into account all three problems, management ability, the equity problem and also the average adequacy of water. In the same way, if weeks are added instead of blocks, then we will use the same measure to interpret week wise. It will tell us the variability between the weeks instead of the variability or the variation between the blocks. In other words, by looking at the variation between the weeks instead of between the blocks one can see how the planned and actual vary across the time while keeping the blocks constant. This will give us a similar analysis which interprets performance, timewise, instead of location wise. How closely management can follow its plan timewise (the third term), how different the averages of the planned and actual are between weeks (the first term) and also how variable is the actual and the planned between the weeks (the second term). Calculating differently, but using the same data this again gives us another perspective of the management performance. This formula is applied to all the four systems that we have. For two Indonesian systems, the planned target and the actual performance of the DPR ratios are utilized to calculate Theil's index and also its decompositions. In the case of Kerian system in Malaysia and the Utris system in the Philippines, we only have data for the relative water supply. If we assume that demand as the expected and the supply as the actual, then we can calculate this Theil's index and its decompositions.

For Cikeusik system, referring to Table 14, the difference between the actual and the planned for the system as a whole, that is the inequality ratio, the last item in the table, is fairly small in the wet season. The ratio is very much below one. However, in the case of dry season, there are five periods, in which the error ratio is very high. This again tallies with the results of adequacy measures we have observed earlier. It is also noted that the variability between the planned and the actual is more pronounced in the dry seasons in the case of Indonesian systems. This generally tells us how we can utilize these indices for both the total results and decomposition. Similar results also are obtained for the other systems. This gives new information for the managers to use. If we have had other collaborative information, we could have judged the reasonableness of these results. We could use the blockwise variation of Theil's index as an indication of its equity (its second term on the right hand side). However, when we correlate each of the results of this decomposition of the Theil's index with other equity measures, the correlations are significant but relatively low. The reasons may be due to the nature of causality. In the case of Gini index and related family index of equity, the assumption is that each plot should get the proportionate water on the basis of its area and accordingly the formula are constructed on this principle. In the case of Theil's index this assumption is not there. This is the total summation of difference between the actual and the results. This total summation of variations between the prediction and the actual performance can be caused by many reasons. In other words principles or assumed causes are not the same and, therefore, they cannot possibly be correlated very high with the other equity performance measures. Another problem in this formula is that in decomposing into three items, the last term occupies a very important position. In terms of proportion, the last term accounts for a very large part of the difference between the actual and the planned. In other words, the correlation, r, between the planned and the actual for each sets of observation figures very prominently in the decomposition and the difference between the planned and actual inequity and the adequacy indicators form a very small part of the proportion. We are still not sure how this information can be utilized because the additional information r in the last term, management ability to follow the plan very closely, will still have to be examined in relation to the system manager's actual performance, other related concepts or actual field experiences.

TABLE 14

RESULTS OF ALL TERTIARY BLOCKS (11 TERTIARY BLOCKS) ON THREE SEASONS (WET SEASON 1987/1988, DRY SEASON 1988, WET SEASON 1988/89)

CIKEUSIK SYSTEM

			w	x	y	z					RI	
		Weeks	S(P-A)2/11	(MP-MA)2	(sdP-sdA)^2	2(1-r)*sdP*sdA	1/w	x/w	y/w	z/w	S(P-A)2/S(P)2	
Dec	I	1	1.66883	0.04069	0.05130	1.57085	1	0.024	0.034	0.941	0.40632	
				1.66883				I				
	II	2	0.63399	0.02041	0.00049	0.61303	1	0.032	0.000	0.966	0.04292	
				0.63399				1				
Jan	I	3	3.11184	0.10346	0.24682	3.42151	1	0.027	0.065	0.907	0.09126	
				3.11186				1				
	11	4	2.04170	0.22955	0.02141	1.19066	1	0.112	0.010	0.877	0.06463	
				2.04170				1				
feb	I	5	2.18860	0.15652	0.21356	1.81852	I	0.071	0.091	0.830	0.04951	
				2.18860				I				
	II	6	1.91506	0.02683	0.14573	1.74250	1	0.014	0.016	0.909	0.04191	
				1.91506				1				
Mar	I	7	1.55930	0.00000	0.08153	1.11111	1	0.000	0.052	0.941	0.03767	
				1.55930				I				
	II	8	1.11494	0.00016	0.12941	0.98537	1	0.000	0.116	0.883	0.02895	
				1.11494				1				
Apr	I	9	1.81920	0.02905	0.14939	1.64016	I	0.015	0.082	0.901	0.06981	Average
				1.81920				I				ke:
	II	10	4.27032	0.26503	0.16643	3.83886	I	0.062	0.038	0.898	0.19022	Season
				4.21032				1				(Dec 1987/
Hay	I	11	1.14812	0.18556	0.00163	1.56153	1	0.106	0.000	0.692	0.15823	Way 1988)
				1.14812				1				
	II	12	0.64849	0.02914	0.00735	0.61140	1	0.045	0.011	0.942	0.05347	0.1025851
Jun	I	13	1.09504	0.26151	0.37043	0.45710	1	0.244	0.338	0.417	0.07760	
				1.09504				1				
	II	I4	5.22344	0.91111	0.00005	4.24561	1	0.187	0.000	0.812	0.28226	
				5.22344								
Jul	I	I5	4.45029	0.11921	0.10031	3.57075	1	0.175	0.022	0.802	0.25098	
				4.45029				1				
	II	16	6.05653	1.22520	0.00048	4.83085	1	0.202	0.000	0.797	0.32447	
				6.05653				1				
Aug	I	17	6.33908	1.13111	0.17243	4.99288	1	0.185	0.027	0.787	0.37430	
				6.33908				1				
	II	18	2.11442	0.66692	0.31623	1.43127	1	0.276	0.130	0.592	0.29418	
				2.41442				1				
Sep	I	I9	0.08042	0.01417	0.00510	0.06055	1	0.176	0.070	0.752	0.02250	
				0.08042				1				
	II	20	0.48807	0.16688	0.00621	0.31498	1	0.341	0.012	0.645	0.12500	
				0.48807				1				
Oct	I	21	4.92320	1.21812	0.43487	3.21021	1	0.259	0.028	0.652	0.53957	Average
				4.92320				1				Dry

TABLE 14 (Cont'd.)

	Weeks	Y S(P-A)2/11	X (MP-MA)2	Y (sdP-sdA)^2	Z 2(1-r)*sdP*sdA	1/w	x/w	y/w	z/w	RI S(P-A)2/S(P)2	
II	22	5,46135	0.31300 5.46135	0.00015	5.15419	1	0.057 1	0.000	0.942	0.57561	Season (Jun 1988/
Nav I	23	5.30585	0.58613 5.30585	1.36942	3.34910	1	0.110 1	0.258	0.631	0.81229	Nov 1988)
II	24	3.56310	0.36900	0.42500	2.16910	1	0.103	0.119	0.777	0.68141	0.3633462
Dec I	25	2.22130	0.26124 2.22130	0.07098	1.88609	1	0.118 1	0.031	0.849	0.18507	
II	26	2.52018	0.01823 2.52018	0.09022	2.41114	1	0.007 1	0.035	0.956	0.13023	
Jan I	21	1.11621	0.85893 1.11621	0.01212	6.30515	1	0.119 1	0.001	0.878	0.23214	
II	28	12.90316	0.35033 12.90316	0.32565	12.22118	1	0.027 1	0.025	0.941	0.31527	
Feb I	29	3.64099	0.19134 3.64099	0.63621	2.81311	1	0.052 1	0.174	0.772	0.08781	
11	30	4.13189	0.16643 4.13189	0.42316	3.54830	1	0.040 1	0.102	0.857	0.10084	
Mar I	31	3.21094	0.07384 3.21094	0.41220	2.12490	1	0.022 1	0.128	0.848	0.09431	
II	32	2.44715	0.09489 2.44115	0.31349	1.91816	1	0.038 1	0.152	0.808	0.01338	
Apr I	33	4.16215	0.22361 4.16215	1.21618	2.72290	1	0.053 1	0.292	0.654	0.13957	Average Yet
II	34	5.11292	0.00019 5.17292	1.50812	3.66461	1	0.000 1	0.291	0.708	0.33833	Season (Dec 1988/
May I	31	5.11999	0.18335 5.11999	1.14150	3.19513	1	0.035 1	0.222	0.741	0.35735	May 1989)
II	36	2.21152	0.03100	0.00658	2.16595	1	0.017	0.002	0.979	0.11169	0.1847121

LEGEND

W = The average of the square of the differences between the planned and the actual.

x = The square of the difference between the means of the planned and the means of the actual.

Y = The square of the difference between the standard deviations of the planned and the actual.

z = The third term including the r where r equals the correlation coefficient of the planned and the actuals.

RI = Inequality Ratio of the 11 blocks by week.

PROFITABILITY INDEX: ITS RELEVANCE AND DYSFUNCTIONS

When we are discussing the reliability and validity of various performance measures relating to the adequacy, timeliness and equity of water distribution, we in fact are assuming the existence of an administrative-type irrigation' organization which uses various criteria of performance relating .to water distribution, client's satisfaction, cost recovery, good maintenance, and other aspects of management. It is taken for granted that under such type of organizations the efficient management of water distribution will be one of the important aspects of the performance evaluation.

However, we must realize that the need for a particular set of performance measures is conditional on the type of the organizations and the kind of operating rules exercised. As soon as these conditions or the operating rules change, a new need for different kinds of measures arises. A case in point is the introduction of corporate-type organization in the National Irrigation Administration in the Philippines. In the early eighties, this organization was transformed from the administrative type organization into the public corporation with full relative financial autonomy or independence. As a corporation, it now operates on the basis of its own profit and loss account, with most of the traditional subsidies reduced or done away with, and as such, the need for financial responsibility and contribution by each 'and every unit of the operating system become acute. As a result, the system of financial accountability of every operating unit such as River Irrigation System or Provincial Irrigation Office is introduced.

This system of management consequently ushered in a new type of control 'measures such as viability index and collection efficiency index. Operating indices relating to water distribution and maintenance still exist but with the overwhelming importance of financial success, as largely on this index, financial reward is given, the managers tend to direct their activities towards achieving these financial objectives, in some cases to the detriments of other objectives. In other words, a change in the organizational set up and the rule of operations largely determines the importance or the prominence of different sets of performance measures with attendant consequences.

As a composite measure, the profitability index embraces a number of factors contributing to the enhancement of financial results. The profitability or the financial variability as it is so named in the NIA is a function of both revenue and expenses. The revenue in turn is derived from both the amount of collections of irrigation fees and the income from other sources such as rentals. The collectibility is theoretically dependant to a large extent on the service provided to the farmers and their satisfaction. The expenditure on the other hand includes both operating cost and maintenance cost. Theoretically good financial results will be derived from or achieved through the improvement of various component factors. The improvement in operating efficiency will enhance the collectibility of fees while the cost reduction will improve the financial ratio. In other words, theoretically the composite index such as profit index has a potential for all round improvement in management. But in the case of the NIA, though the financial viability of many irrigation systems have improved, its overall impact brings in some undesirable consequences.

As the operating ratios including water distribution efficiency index do not have proper standards or some of the ratios are hard to discriminate the efficiency between the different systems, they become ineffective in evaluating the managers. A strong drive for collections and expense reduction has led to managers neglecting maintenance and paying undue attention to collection activities. Though the short-run financial results have improved, the long term viability in both physical and financial terms is very much in doubt.

This experience indicates both the usefulness of the composite measures and the dysfunctions of this measure when the ultimate financial results and its contributory component factors are not properly or tightly related. It also indicates the inherent weakness of the profit measures as it emphasizes the short-term or yearly results than the longterm impacts.

CONCLUSION

This paper is primarily concerned with testing how sound the measures we have suggested in our Performance Assessment Project are in terms of instrumentation methodology, relevance and usefulness. In this connection, we propose to ask the following questions.

1. How reliable are the results of the measures we have suggested, that is, are the results of any measure suggested repeatable or consistent ? If a well-anchored rifle produces hits at one single spot, we may say it is reliable: on the other hand, if this rifle makes widely scattered hits, we may say that the rifle is unreliable. In other words, the reliable instrument is one that produces repeatable or consistent results.

2. How valid is the instrument or measure we propose ? Or alternatively, we may ask: "Does the measure do what it is supposed to do ?' The question of the validity is the crux of the problem in this stage of development of instruments or measures. Making an instrument reliable is only fulfilling a necessary condition in the process of developing instruments or measures. The instrument ultimately must serve the purpose for which it is intended. A measure for assessing the adequacy of water in an irrigation system must represent or indicate the true conditions of the adequacy. If it only measures the availability on certain days only, or if it only indicates the agronomist's standards which is at variance with what the farmers feel they need under the circumstances, it could not be a valid or true measure of the adequacy. In our example of a rifle, if it hits repeatedly at a spot away from the target, we may say that the rifle is not serving the purpose. It is hitting at the wrong spot. The hits are consistent but not correct or valid, not serving the intended purpose.

3. Even if the valid instrument is constructed, the effectiveness of this instrument depends on under what circumstances or for what conditions the instrument is used. Many types of instruments may be available but unless conditions or circumstances under which each instrument or measure should be used are specified, its effectiveness is limited. In other words, limiting conditions and circumstances within which instruments must be used should be properly defined. The question of relevance begins with

the first appearance of a measure and will continue as long as it is being actively used. This question could be tackled first on the basis of previpus knowledge and then with the actual experience of repeated applications. This is a matter of continual assessment, redefinition and specification.

Regarding the first question, that of repeatability and consistency, we recognize the special nature or peculiarity of our instruments appropriate for irrigation management. Our suggested instruments are not additive indices of similar items as in the science in which item-analyses are usually performed. In our case, each result of an adequacy measure or an equity measure is a product of a mathematical formula based on certain rationale, not a summation of responses to a set of similar items. Therefore, the tests of reliability usually applicable to social studies may not be appropriate here. However, the basic rationale or theory of reliability or consistency could still be utilized as the test of reliability is essentially derived from the classical test theory.' Reliability is a measure of how contaminated is the observed variance with random error; if the random error is affecting the observed variance very severely, the reliability will be approaching zero. The greater the error variance, relative to observed variance, the closer the reliability is to zero. Some of the basic equations of the test theory states

$Px = 1 - [var(e)/var(x)]$
where
Px = Reliability
var(e) = error variance
var(x) = observed variance

In any test of reliability will be the test of the extent of random error involved and whatever we devised will be derived from this principle.

The measures we have proposed for measuring the performance of irrigation systems include essentially a quantum of water delivered at a particular designated or selected point at a particular time: they are not the composite of responses to fifty items or so by particular respondents. Other categories such as the demand for water or the planned are the calculations based on already existing data. Either the test of repeatability as done by the test-retest method or the test of equivalence as measured by split-half method or similar devices are not meaningful. Each observation in our case is time-specific and place-specific. No two observations could be the'same and none of the above methods are appropriate.

We have mentioned earlier that the reliabil'ity is the function of the size of error variance relative to the observed variance. Random errors in this regard include errors in observation, breakdown of the instruments, errors in recording, errors in transfer or manipulation of records or errors in calculations, etc. One practical assessment of the reliability of records will be the examination of field records and also making the selected field observations so as to see the possible sources as well as the extent of errors in recording process. That we have done with regard to both .the availability and reliability of recorded data and a few selected countries, India, Sudan, and the Philippines. Our conclusion in this paper is that the availability of the right kind of information is more problematic than the reliability of the data collected. As the data are usually not collected for the purpose of managing the system but for other purposes such as revenue collection, economic planning or for routine record keeping, they could in most cases satisfy the requirements of the system assessment and system management. In addition, we observed that there are errors, both random and non-random (or biased) in the existing records. In fact, non-random errors than random errors seem to be more serious in the current recording procedures and processes. When collecting as much revenue as possible is the prime concern, the upward bias in recording of irrigated area will be there. When the irrigation departments plan water requirements and compare the planned targets as against the actual distribution, the tendency exists to make planned targets not too distant from the expected delivery. There would probably be a downward pressure on setting targets and the upward.pressure on recording the actual performance. The random errors exist in all organizations though the degree of errors will vary depending on the degree of discipline and the extent of control and motivation of the management to keep the records correct.

Once' records are collected and used, the evidence of random errors will have to be examined through various tests of consistency and that of biases or non-random errors through the tests of validity. With regard to random errors or the reliability, we see the two possible tests as significant for our purpose. Since each measurement or recording is identified by both its location and its relevant time period, irrigation records could be treated as a matrix, column describing the different locations and now indicating different time periods. Unlike observations under classical test theory, these observations could not be treated as random observed values and any deviations from true values of these observations will be assumed to be cancelled out by each other over the long run. Each one is distinct entity and one observation in a particular location cannot be the same timewise as any other observation. It is already pre-arranged or planned that more water will be delivered at certain periods than other. On the other hand, even for a particular period, different locations will receive unequal proportions of water than intended or because of typology of land or normal deterioration *of* canals. However, one could assume that under the circumstances, the pattern of distribution of water at different points in one occasion or one time period should be similar to that of another period. Any deviation from consistent pattern will be considered as the effects of random errors. However, keeping the location constant or at one location, the pattern of distribution water received between different time periods may not be the same from the pattern received from another location between different periods. The preponderance of the location in affecting the amount of water received is the reason for this phenomenon. The location near to the main source or water delivery gate will be likely to get more water even if the total amount delivered is less than that of other period. Thus, the pattern of the distribution at a point near the delivery gate will be more equalized between different periods than those other points far away from the gates.

Therefore, the test we propose here for testing the consistency of the delivery of water is the comparison of the relative level of water received at different locations at one period with those of other periods. The average of correlations of water distribution patterns at different locations between periods will approximately equal the Conbach's test of consistency.

Coefficient Alpha. However, in our data analysis we have done only preliminary idea, testing of the coefficient of variations of observations at one point for different works have been calculated to see the extent of random factors that may have played in influencing the distribution patterns of water. The results show that coefficient of variation of DPRs and RWSs ratios do not have very pronounced differences in staying within reasonable range. Of course, this type of analysis does not show the similarity of patterns of distribution but only does indicate the variability of means only. The test of consistency through inter-correlation between different sets of observations must be made to test the consistency of these patterns.

The second important test of consistency is testing the consistency between main system and sub-system results. The need for this kind of results arises from the problem of cost and time in collecting the complete enumeration of data at all locations in the system. For the immediate actions or for getting a general picture, the manager has to rely on the observations at certain designed or sampled points or the calculation based on the aggregates .of the system. Under these circumstances, there is always a nagging question of whether the use of aggregates such as the relative water supply of the whole system on the basis of the total system aggregate figures or generalizing on the basis of a few observations with the use of the average of the results at a few observation points is reasonable or defensible in measuring the performance of irrigation systems.

Regarding this question, we introduced various devices to test whether the .main system results are consistent with the sub-system results or whether the difference; between sub-systems average or between the sub-system means and the grand means are significant. First, we compared the overall system values with block averages, and then the variability of block means by looking at their coefficients of variations. We find that as far as the Indonesian data is concerned the average of the block means and the total system values independently calculated are similar and close to each other indicating the consistency between the sub-system and the main system result.

The same data **is** submitted to analysis of variance appropriate for handling questions of such nature. The similar results are obtained. The grand means and the block means are not significantly different. Though these results are satisfactory for two systems in Indonesia, thus indicating the consistency of data collected, we could not be sure that this applies to all situations. **As** another system is studied, the same test or the test of similar nature will have to be performed. Only after repeated experience, we may be able to specify circumstances which contribute towards the consistency between the main system and sub-system results.

The next we have dealt with, in this paper, is the question of validity through an extensive analysis of consistency between theoretically close measures - construct validity. **It** should be noted that the real validity or goodness of a measure could only be vouched by the users in the field only. The validity too is a continuing process that must be settled only by the actual usage in the field. For instance, in this paper we could not do much about the testing validity through independent criterion. This will require an additional information about how the system is functioning. An independent observation or a manager's observation about how well the system is doing with regard to the adequacy, timeliness or equity of water distribution may be needed to compare with the results of the degree of equity or adequacy as indicated by our suggested measures. The information from various independent sources such as farmers' reactions will have to be gathered for this purpose. It should be noted that the validity is never absolute and only with better techniques of independent observation and the availability of data, we could be more assured of their goodness of these suggested measures.

The question of relevance is dealt with in this paper when we discuss the meaningfulness of water distribution measures under the profit driven situations. A need for further measures to countermand the dysfunction of the profitability index is also indicated. Circumstances under which any particular measure be meaningful or usable will be a further area of research in this endeavour.

The purpose *of* this paper is not to give the ultimate answer to the usability of measures suggested in our research. Its main purpose is to apply and learn from these applications about the limits and usefulness of the measure for further investigation. Toward this end we have developed a methodology or a framework for further analysis. We have identified what are the important questions we must raise, discussed what methods are relevant in tacking this question, developed both a logical basis or a framework for analyzing the question of reliability and validity, however tentative these ideas still are. The author hopes that we are now a bit more sure about further testing the usability of performance measures in varying environments and conditions.

REFERENCES

Abernethy, C. (1989): Indicators of the Performance of Irrigation Water Distribution System. (Performance Symposium Report 1989):

Abernethy, C.L. (1989): Performance Criteria for Irrigation Systems. Conf. on Irrigation Theory and Practices, Southampton, U.K.

Abernethy, C.L. (1986) : Performance Measurement in Canal Water Management. ODI/IIMI network paper 86/2d.

Bos, M.G. and J. Nugteren (1982): On Irrigation Efficiencies. International Institute for Land Reclamation and Improvement, Wageningen.

Doorenbos, J. and Kassam, A.H. (1979) : Yield Response to Water. FAO Irrigation and Drainage Paper 33.

Kyi, K.M., C.M. Wijayaratna and C. Nijman (1990): Study of Organizational Dynamics in Corporate-type Irrigation Organization: An Organizational Analysis of the National Irrigation Administration in the Philippines. (IIMI Internal Programme Review Report 1990).

Lenton, R.L. (1982): A Note on Monitoring Productivity and Equity in Irrigation Systems.

Levine, G (1982): Relative Water Supply : An Explanatory Variable for Irrigation Systems. Cornell University Technical Report No.6.

Mathotra, S.P., S.K. Raheja and D. Seckler (1984): A Methodology for Monitoring the Performance of Large-Scale Irrigation Systems : A Case Study of the Warabandi System of Northwest India. Ag. Admin., V.17, p.231.

Rao, P.S. (1987): Relative Equity Ratio : Concept and Method of Computation. Unpublished.

Sampath, R.K. (1986) : Inequity Measures for Irrigation Policy and Performance Evaluation Analyses.

Sampath, R.K. (1988) : Equity Measures for Irrigation Performance Evaluation. Water International, v.13, p.25.

Seckler, D., R.K. Sampath and S.K. Raheja (1988) : An Index for Measuring the Performance of Irrigation Management Systems with an Application. Wat. Res. Bull., V.24, no. 104, p.855

Sharma, D.N., Ramchand Dad and R.K. Sampath (198_): Performance Measure for Irrigation Water Delivery Systems. Unpublished Paper, Department of Agri. Economics and Natural Resources, Colorado State University, Fort Collins, Colorado, U.S.A.

Siegel, Sidney (1956) : Nan-parametic Statistics, McGraw, N.Y

Small, Leslie E. and Mark Svendsen (1989): A Framework for Assessing Irrigation Performance. (Read in the 1989 Symposium). Revised version.

Theil, Henri (1966): Applied Economic Forecasting. North-Holland Publishing Co Amsterdam.

Wijayaratna, C.M. (1986) : Assessing Irrigation System Performance: A Methodological Study with Application to Gal Oya Scheme, Sri Lanka. Ph.D. thesis, Cornell Univ.

Zeller, Richard A., and E.C. Carmines (1980): Measurement in the Social Sciences. Cambridge University Press, Cambridge.

APPENDIX I

A METHOWLOGICAL FRAMEWORK FOR ANALYSIS OF RELIABILITIES, VALIDITY AND RELEVANCE OF MEASURES USED IN MEASURING IRRIGATION SYSTEM'S PERFORMANCE

1. Basic Questions

   (a) How reliable, repeatable or consistent are the results of the measures we have suggested or proposed ?

   (b) How valid is the instrument we propose ? Alternatively, does the instrument measure what it purports to measure ?

   (c) How relevant is the measure we have devised to the needs and requirements of various users under varying conditions ?

2. Data 'Matrix

   A matrix of observed values (DPR, RWS, Equity or any other ratios), indicated by periods (irrigation scheduled weeks) and by locations (blocks, canals or laterals etc.) for a cropping season.

METHODOLOGICAL GRID

1 TYPE OF ANALYSIS	2 QUESTIONS DEALT WITH	3 CRITERIA	4 METHODS	5 USES AND LIMITATIONS
1. Pre-check on available data, its sources and methods of computation.	Both random and non-random errors.	Sources of non-random errors. Causes of non-random errors. Consistency with collaborative, parallel and source data.	1. Sample checks on procedures, mistakes. omissions, . commissions. 2. Purpose of data collection, biases involved, comparison with Other data sets or results.	Some exclusion of outliers or abnormal data. Cleaning UP the data set.
2. Internal consistency of measures.	Repeatability and consistency.	Are the subsets of the measure consistent with one another ?	Inter-correlations - (a) Average of inter-correlation. (b) Coefficient Alpha and related test.	The size of subsets (columns or rows in matrix) should not be too small. Non parametric methods such as rank correlation may be used in case of small size of the samples. A number of subset may not conform to the pattern. Reasons for this need to be studied.
3. Consistency between sub-systems and main systems results.	Reasonableness of aggregation, summation, generalization.	Do system and sub-system results differ very widely ?	1. Comparison of basic parameters between units . (Means standard deviations, coefficient of variations). 2. Analysis of variances between results of different sub-systems (laterals). 3. Analysis of variances (nested design).	Data availability is a problem here. A sample design to collect the requisite data may be needed,
4. Consistency between results of various measures.	Validity.	Are the results of theoretically close tests converging ?	1. Inter-correlation tables. 2. Consistency tests (Alpha coefficients). 3. Factor analysis for a large number of tests and a large sire sample.	It only confirms the theoretical closeness of the measures.
5. Test of relevance.	Relevance, usefulness, productability, validity.	‘Are the tests results consistent with known results or independent results ?	1. Correlations with other independent results. 2. Regression with predictive variables.	Independent assessment by experts, managers, users, farmers. may be needed. The prediction of impacts On the basis of results of the measures or the regression of decomposed results may be needed. More importantly, a research design to tap the status of performance through the qualititive assessment is also a must.

APPENDIX II

FORMULAE FOR DELIVERY PERFORMANCE RATIO AND RELATIVE WATER SUPPLY

PHILIPPINE CASE

RWS = Actual Supply/Demand

Actual Supply = Irrigation + Rainfall (total)

Demand = 2 li/sec/ha (land preparation approximately 6 weeks)

= 1.5 li/sec/ha (crop growth until 2 weeks before harvest)

MALAYSIAN CASE

RWS = Actual Supply/Estimated Demand

Actual Supply = Irrigation + Rainfall (total)

Estimated Demand = 2.3 li/sec/ha for 4 weeks presaturation

= 1.15 li/sec/ha for crop growth

All units are either in terms of volume (cubic meter, m^3, cu.m.) or depth of water per unit area (mm, m)

INWNESIAN CASE

DELIVERY PERFORMANCE RATIO (DPR)

Procedures for calculating Delivery Performance Ratio

$$\text{DPR} = \frac{\text{Actual Water Supply}}{\text{Water Supply Planned}} = \frac{\text{Qact}}{\text{Qpla}}$$

Qact = total water delivered from canal (l/sec); **it** calculated width of canal and height of water.

Qact = 1.71 * 6 * $h^{3/2}$; 1.71 is constant for the flume is Romein

add

Qact = 1.86 * 6 * $h^{3/2}$; 1.86 is constant for the flume is Cipollety

b = width of canal
h = height of water

RELATIVE WATER SUPPLY (RWS)

Procedure for calculating Actual Relative Water Supply

$$\text{RWS} = \frac{\text{Actual Water Supply}}{\text{Water Demand}}$$

Actual water supply is total water delivered from canal and actual seasonal rainfall.

RELATIVE WATER SUPPLY (RWS)

Procedures for calculating Actual Relative Water Supply (RWS) $\frac{\text{Supply}}{\text{Demand}}$

* Supply: Total seasonal irrigation water delivered t total seasonal rainfall (usually expressed in depths of water or volume expressed in depths of water delivered to the irrigated area, in mm).

* Demand: Theoretical or Empirical (measured) water used quantity established for a particular crop. For rice, this consists of evapotranspiration and seepage and percolation, and water use for land soaking and preparations. The open-pan evaporation approximates the rice evapotranspiration. While for other crops, crop coefficients are usually used to estimate evapotranspiration. The unit of measure is also expressed in depth of water (mm) or volume.

Example:

For main system RWS ?

Data needed:

Supply -> Ir □ Average daily irrigation flow = cubic meter per second (m3/sec) (delivered into the main canal from the weir).

d = Total number of irrigation days for the seasons

A = Total area irrigated for the season (ha)

Rn. = Total daily rainfall (mm) -> do not include amounts greater than 100mm/day.

Thus:

$$\text{Supply} = \frac{Ir * d * 8.64}{A} + Rn$$

Demand:

For Rice -> (LS & LP) = Land Soaking and Land Preparation
Water measured from the start of soaking the fill until just before transplanting.

Average for heavy clay soils range 300 - 500, for one month of land soaking and preparation.

EV = Water for crop growth.
Total daily open pan evaporation (mm) for the season; or
Average daily open-pan evaporation
* Total number of growing days (until 2 weeks before harvest)

S&P = Total daily seepage and percolation (mm) for the season or average daily seepage and percolation.
* Total number of growing days.

Thus: Demand = (LS & LP) t (EV) + (S & P)

so RWS = Supply/Demand

$$RWS = \frac{[(Ir * d * 8.64)/A] \text{ t } Rn}{(LS \,\&\, LP)} \text{ t } \frac{[(Ir * d * 8.64)/A] \text{ t } Rn}{(EV) \text{ t } (S \,\&\, P)}$$

Land Preparation Stage + Growing Stage

APPENDIX III

THE KOLMOGOROV-SMIRNOV TEST

This test is concerned with the degree of agreement between the distribution of a set of sample values (observed score) and some theoretical distribution (normal, uniform or x^2).

f = number of subjects choosing a particular item or rank.

$F_o(x)$ = Theoretical cumulative distribution of choices under Ho^o.

$F_n(x)$ = Cumulative distribution of observed choices.

D - The largest value *of* $F_o(x) - F_n(x)$

The table of critical value *of* **D** give the probability of occurring the maximum value **D. (Siegel, Sidney, 1956).**