Vol XXXV
No. 2

APRIL-
JUNE
1980

ISSN 0019-5014

INDIAN JOURNAL OF AGRICULTURAL ECONOMICS

INDIAN SOCIETY OF
AGRICULTURAL ECONOMICS,
BOMBAY

MEASUREMENT OF GROWTH AND FLUCTUATIONS IN CROP OUTPUT—AN APPROACH BASED ON THE CONCEPT OF NON-SYSTEMATIC COMPONENT

V. M. Rao, M. V. Nadkarni and R. S. Deshpande*

Reliable estimation of trends in agricultural output is more than a matter of mere academic interest. Their use in predicting the future may hardly be dependable; but as measures of performance during a period, they have considerable significance to policy-makers. Estimation of trends, however, has been plagued by the problem of fluctuations, which are also as important a part of time-series as trends themselves. Both from an academic angle as also from that of a policy-maker, it will be wasteful to lose information on the nature and magnitude of fluctuations by concentrating on trends alone. Thus, either the method of moving average or exclusion of extreme years of high or low production, is a 'wasteful' method of estimating trends. It amounts to throwing away information on precisely the years marked by large weather-effects which should be of as much interest as trends themselves.

Fitting a trend to raw data and calculating coefficient of variation of residuals from the fitted trend apparently take note of both the trend and fluctuations. Though normally it may be an adequate procedure, it may not be workable when fluctuations are large and frequent. This is because the very estimation of trend is distorted by fluctuations, and neither the trend nor the fluctuations derived here may adequately reflect the reality involved.

An explicit inclusion of weather in trend equations as a proxy for fluctuations may appear to provide an alternative. This holds no appeal to us either. Weather is too complex a matter to be specified in terms of one or two variables, and is, in any case, not within the competence of mere economists ! The matter looks particularly hopeless when one is concerned with a fairly large region like a State as a whole or a big sub-division such as northern plateau or southern plateau in Karnataka. Weather cannot be averaged over such vast regions as a variable for explaining fluctuations in the regional output of various crops. Besides, the procedure also amounts to assuming that weather is the only source of fluctuations and that once it is included in a model all the fluctuations are taken care of. In any case, such a model affords no opportunity of studying fluctuations separately as they are identified with weather.

Intuitively, the following would be a more fruitful strategy which could segregate the extreme years without losing the information contained in them:

(*a*) Separate the years included in the time-series into three groups—'normal' years, 'peak years' and 'trough' years.

(*b*) Use only the 'normal' years to estimate the trend line.

* Senior Fellows and Research Analyst, respectively, Institute for Social and Economic Change, Bangalore. Thanks are due to Dr. V. G. Rao for comments and discussion on an earlier draft. We have also benefited greatly from the comments made at Lonavla Seminar in February 1980 where this was earlier presented.

(*c*) Having obtained the trend line, bring back the 'peak' and 'trough' years into the picture to see how weather-effects pull production away from the trend line.

While step (*b*) ensures undistorted estimation of trend, step (*c*) provides a line of investigation covering the information contained in 'peak' and 'trough' years. Though uncertainty in agriculture has received some attention in the past, there has been a tendency to assume that residuals from trend are randomly distributed. This, however, needs to be investigated. This is not to satisfy only an academic curiosity. A proper understanding of the nature of distribution of residuals may be crucial in policy issues like feasibility of crop insurance and planning a scheme of buffer stocks.

The purpose of this paper is to present the results of an exercise undertaken to show the feasibility of this strategy in practice. Before describing the exercise, it is necessary to say a word about the specification of the model for estimating the trend which could provide theoretical justification for separating the years into 'normal', 'peak' and 'trough' years. Unless supported by a plausible model, such separation and procedures based on it would only remain at the level of informal 'rule-of-thumb' practice. On the other hand, it is our hope that our findings have implications going beyond the immediate purpose of better estimation of trend. We write the trend equation as follows:

$O_t = a + bt + e_t$ .... (in 'normal' years)

$O_t = a + bt + n_t$ .... (in 'peak' and 'trough' years)

where O = output, t = year, e_t is assumed to have all the properties attributed to a random disturbance term and n_t is assumed to a term about which nothing can be specified *a priori* except that it is neither fully systematic nor fully random. We call n_t a 'non-systematic' component. It is assumed that the trend coefficient and the functional form of trend remain the same throughout the period of observation: specifically, the trend estimated for normal years is assumed to hold for peak and trough years as well.

It is seen that the conventional regression model using 'normal' years would yield a valid estimate of the trend coefficient; this would not be true if the regression covers all the years. More important, the assumption of relative ignorance about n_t underlines the need to discover its behaviour and properties *empirically* and, hence, the need to extend data analysis to bring n_t under observation. It is obvious that the model given above could be modified to include *a priori* specifications on n_t and, equally, to try out alternative functional forms for the trend. However, this would need a much larger paper than we have planned and we have preferred to keep this paper confined to a simple demonstration of the usefulness of the concept of non-systematic component in studying agricultural growth and fluctuations.

The key step involved in applying the model suggested above is the separation of the period of observation into 'normal', 'peak' and 'trough' years. Instead of using *a priori* information on climatic variables like rainfall, temperature, etc., to determine good and bad years, we prefer to use data on output itself for this purpose. Climatic data are of interest because of their impact on crop output and not *per se*. We have already stated our reservations against using weather as a proxy for fluctuations in estimating an undistorted trend. They would hold in this context too. Output, moreover, can be influenced by variables other than weather, including man's capacity to cope with droughts, which is expected to improve through time.

We, therefore, feel persuaded that the approach feasible to economists can only be along using evidence contained within production data on the fluctuations that have occurred over the period of observation. We have tried two simple alternatives based on this approach of relying on the internal evidence. The first alternative, which we call Method I, proceeds by separating 'local' peaks and troughs and treats the remaining years as normal. Specifically, a year is taken as a peak year if the output of a crop in that year is *higher* than the output in both the preceding and succeeding years and the excess in each comparison is *more than 5 per cent*; similarly, a year is taken as a trough year if it has lesser output than both the adjacent years with the deficit in each case being more than 5 per cent. It may be added that a peak or trough year identified by Method I is not necessarily an extreme year over the period of study as a whole; nor, in strict logic, would Method I necessarily identify an extreme year as a peak or trough year. In other words, Method I considers weather-effects in terms of year-to-year fluctuations and not in terms of fluctuations around an average or trend over a period.

The second alternative, which we call Method II, looks at the fluctuations around the trend. Specifically, Method II proceeds by fitting a trend line to all the years and, then, drawing upper and lower band lines to cover an interval of *10 per cent* around the fitted trend line.[1] The years above the upper line and below the lower line are identified respectively as the peak and trough years, with the years falling within the 10 per cent interval being taken as the normal years.[2]

It is hoped that the two ways of looking at weather implied in our methods — *viz.*, the 'year-to-year' view of Method I and the 'around-the-trend-over-a-period' view of Method II—are meaningful and interesting to explore. Equally important, the methods themselves are easy to understand and work with. It is needless to add that an element of subjectivity and experimentation is unavoidable in choosing the intervals to get a 'plausible' separation of years.

In what follows, we report the results of applying the two methods to annual crop output data for the *State of Karnataka* (formerly known as Mysore State) over a twenty-one-year period, 1955-56 to 1975-76. The crops covered in the study are shown below. These are the principal crops of the State covering about 80 per cent of the total cropped area in the State.

1. When a 5 per cent band around the trend was taken as in Method I, it did not yield adequate number of normal years.

2. The question of why we have taken 5 or 10 per cent bands instead of larger or smaller bands does not worry us any more than the question of why we take 5 or 10 per cent confidence levels in statistical testing. A band of one per cent would obviously appear as too small in our context and a band of more than 10 per cent unreasonably large.

Individual crops	Crop groups
Rice	Total jowar (*kharif* + *rabi*)
Ragi	
Kharif jowar	
Rabi jowar	Total cereals (including minor cereals not considered as individual crops)
Bajra	
Maize	
Wheat	Total pulses (including pulses other than gram and *tur*)
Gram (pulse crop)	
Tur (pulse crop)	
Groundnut	Total foodgrains (including minor cereals and pulses not considered as individual crops)
Cotton	
Sugarcane	

Table I indicates the number of normal, peak and trough years obtained by the two methods.

TABLE I—NUMBER OF NORMAL, PEAK AND TROUGH YEARS SHOWN BY THE TWO METHODS

Crop	Method I			Method II		
	Normal years (No.)	Peak years (No.)	Trough years (No.)	Normal years (No.)	Peak years (No.)	Trough years (No.)
Rice	9	3	7	15 (7)	3 (1)	3 (3)
Ragi	12	3	4	11 (9)	5 (2)	5 (2)
Kharif jowar	11	4	4	12 (9)	4 (3)	5 (3)
Rabi jowar	12	4	3	10 (6)	5 (2)	6 (1)
Total jowar	14	2	3	15(12)	4 (2)	2 (2)
Bajra	7	6	6	6 (2)	7 (3)	8 (5)
Maize	12	4	3	4 (2)	9 (2)	8 (1)
Wheat	9	4	6	7 (3)	8 (1)	6 (4)
Total cereals	11	3	5	14(11)	3 (1)	4 (3)
Gram	13	3	3	9 (9)	8 (3)	4 (3)
Tur	6	6	7	10 (5)	5 (3)	6 (5)
Total pulses	12	3	4	10 (6)	5 (1)	6 (2)
Total foodgrains	12	2	5	14(12)	3 (1)	4 (3)
Groundnut	15	2	2	9 (6)	5 (1)	7 (2)
Cotton	15	3	1	5 (4)	9 (3)	7 (1)
Sugarcane	17	1	1	17(14)	3 (1)	1 (Nil)

Notes:—1. Method I excludes the two end years.
2. Figures in brackets in Method II indicate the number of common years in both methods.

As regards the estimation of trend, we proceed as follows. For each crop and crop group, a trend line of the form $O_t = a + bt$ is fitted separately by the Least Squares method to (*i*) all years, (*ii*) normal years of Method I and (*iii*) normal years of Method II. Since the coefficient b, taken by itself, is difficult to compare across crops, it is converted into growth rate (per cent) by dividing it by the mean output level of the crops over the period of study. Note is also taken for each fitting of trend line of the R^2 value and of the level of significance of b. The trend line fitted to 'All years' and its associated R^2 value and level of significance provide the base to judge the performance of the two methods in giving a better estimate of the trend in crop output. Table II, which gives these results, yields interesting insights.

TABLE II—GROWTH RATES IN CROP OUTPUT

Crop	All years		Method I Normal years		Method II Normal years	
	Rate of growth (per cent per annum)	R^2	Rate of growth (per cent per annum)	R^2	Rate of growth (per cent per annum)	R^2
Rice	3·22	0·75*	4·23	0·96*	3·46	0·97*
Ragi	0·83	0·04NS	0·03	0·003NS	0·36	0·23NS
Kharif jowar	3·95	0·73*	3·50	0·77*	3·63	0·95*
Rabi jowar	1·14	0·10NS	2·47	0·68*	1·34	0·70*
Total jowar	2·75	0·68*	2·59	0·77*	2·57	0·87*
Bajra	4·75	0·49**	4·69	0·71*	5·15	0·96*
Maize	17·44	0·76*	22·04	0·62*	Not	fitted@
Wheat	7·55	0·65*	8·38	0·74*	8·59	0·96*
Total cereals	3·10	0·75*	3·40	0·98*	3·28	0·97*
Gram	2·74	0·33***	3·83	0·89*	3·63	0·97*
Tur	3·93	0·57**	4·77	0·97*	3·83	0·98*
Total pulses	3·60	0·66*	4·16	0·81*	3·61	0·94*
Total foodgrains	3·15	0·76*	3·44	0·98*	3·13	0·90*
Groundnut	0·69	0·06NS	0·14	0·14NS	0·42	0·21NS
Cotton	3·41	0·34***	3·29	0·33**	2·76	0·67***
Sugarcane	5·24	0·94*	5·37	0·95*	5·10	0·98*

Notes:—NS = Not significant.
* Significant at 1 per cent.
** Significant at 5 per cent.
*** Significant at 10 per cent.
@ In this case the number of normal years is less than five, and hence the trend was not fitted.

It is seen that, despite a sizable reduction in the number of observations owing to elimination of peak and trough years, the trend lines fitted by the two methods do better in most cases as reflected in the higher R^2 values as compared to the R^2 values for 'All years'. In the case of five crops—*rabi* jowar, bajra, gram, *tur* and cotton—Method I leads to an improvement in the level of significance, which is also true of Method II excepting cotton. As regards the value of growth rates, there is no evidence of any consistent over- or under-estimation of growth rate by the 'All years' regressions. However, the cases of under-estimation (rice, *rabi* jowar, maize, gram and total cereals) are seen to be more than those of over-estimation (*kharif* jowar, total jowar and cotton). The important point is that a researcher or policy-maker interested in a comparative perspective on growth rates in different crops could get a distorted picture by relying on 'All years' regressions. While it is not easy to put a value on the extent of distortion, Table II is of help in seeing the possibility of distortions taking place when the growth rates of crops are estimated without separating peak and trough years.

A further insight provided by Table II relates to the range of conditions within which the model suggested in this paper is of help in assessing the growth rates of crops. At one end can be placed sugarcane, a crop grown under conditions of adequate and assured perennial irrigation. Sugarcane output has few peak and trough years (see Table I) and its growth rate and level of significance remain stable over the three alternative estimates presented in Table II. Sugarcane output, in a word, is sufficiently free from fluctuations to make the 'All years' regression indistinguishable from the results given by Methods I and II. At the other end can be placed *ragi* and groundnut having growth rates too weak to be audible above the noise of fluctuations; in these cases, no amount of ingenuity in data analysis might really be of much help. As can be seen from Table II, there is a substantial intermediate range between the two extremes in which the separation of peak and the trough years could lead to a clearer perception of the growth rate. A broader proposition which we are tempted to suggest is that a large part of Indian agriculture lies precisely in the intermediate range. While the developmental efforts are substantial enough to push up crop growth rates, the era of freedom from weather-induced fluctuations would be judged by the experts as being several decades away.

There is a hint in Table II that the growth rates are somewhat easier to read in the case of crop groups than the individual crops. The largest crop group that we have is 'Total foodgrains' in whose case the 'All years' regression performs as well as those of Methods I and II; also, for the other crop groups, *viz.*, 'Total jowar', 'Total pulses' and 'Total cereals', the level of significance is already high in the 'All years' regression and no improvement in this respect is brought about by Methods I and II. This is understandable since a crop group benefits from the evening out of fluctuations across the component crops included in the group. On this reading, the need to separate the peak and trough years is particularly pressing in the case of individual crops. On the other hand, it is precisely for the individual crops that the direct approach based on data on weather variables like rainfall, temperature, etc.,—which are year- and region-specific and not crop-specific—is difficult to use to characterize a year as peak, trough or normal. In other words, the indirect approach underlying our methods seems to be particularly appropriate to use for the individual crops. On the whole, we hope that the concept of non-systematic component and the methods it suggests for estimating growth rates of crops do derive a measure of positive support from Tables I and II.

It is time to bring the non-systematic component into the open. We proceed as follows to isolate it for observation. First, for each crop and for each fitted trend line, the residuals are measured as the difference between the actual output and the output as estimated by the trend line. This gives us three sets of residuals for each crop corresponding to the three regressions—All years, Method I and Method II. Second, the standard deviation of residuals is calculated for each set in each crop—based on *all the residuals* for the 'All years' regression and based on the *normal year residuals* for the other regressions. Third, the residuals in each set are divided by the corresponding standard deviation to express them in units of standard deviations. The percentage distribution of the standardised residuals is presented in Table III. The size classes in Table III are in terms of units of standard deviations. The 'All years' residuals and both the normal year residuals are seen to lie within the range +3 and —3, excepting a fractional percentage in the first case. Judging by the rule of thumb property of the standard normal distribution (66 per cent in the range —1 to +1, 95 per cent in —2

TABLE III—PERCENTAGE DISTRIBUTION OF STANDARDISED RESIDUALS

Size class (in units of standard deviation)	All years	Normal years		Peak and trough years	
		Method I	Method II	Method I	Method II
More than 5	—	—	—	5·6	13·3
5 to 4	—	—	—	2·2	10·8
4 to 3	—	—	—	6·7	10·0
3 to 2	2·3	2·2	0·9	7·8	10·0
2 to 1	7·9	10·9	13·5	10·0	6·7
1 to 0	42·9	38·4	41·4	16·7	0·8
0 to —1	31·3	32·6	27·0	13·3	1·7
—1 to —2	11·9	14·5	16·2	16·7	2·5
—2 to —3	3·2	1·4	0·9	6·7	6·7
—3 to —4	0·4	—	—	4·4	8·3
—4 and less	—	—	—	10·0	29·2
Total	100	100	100	100	100

Note:—This table is based on the residuals of individual crops excluding maize in Method II.

to +2 and 99 per cent in —3 to +3), all the three distributions would seem to be approximately normal. In fact, one calculating only the 'All years' regression and checking the distribution of residuals for normality may have little reason to suspect the specification of the disturbance term. It is only the further probing of the residuals permitted by Methods I and II which brings out the striking contrast between the distribution of normal year residuals, on the one hand, and the peak and trough year residuals, on the other. It is from this contrast that the model suggested in this paper derives its plausibility.

Under the assumptions of our model, the normal year residuals are in the nature of a random variable having a normal distribution while the peak and trough year residuals form a non-systematic term not interpretable as a random variable. The segregation of these two components gives a random term with a smaller standard deviation as compared to the standard deviation of the 'All year' residuals (see Table IV). This fall in the value of standard deviation permits a clearer perception of crop growth rates in 'normal year' regressions as compared to 'All years' regression.

TABLE IV—STANDARD DEVIATIONS OF RESIDUALS

Crop (units*)	All years	Normal years	
		Method I	Method II
Rice	1877	703	747
Ragi	1931	889	353
Kharif jowar	1152	757	432
Rabi jowar	1193	594	313
Total jowar	1611	982	817
Bajra	478	294	86
Maize	602	447	@
Wheat	405	362	68
Total cereals	4792	1148	1149
Gram	135	45	25
Tur	259	44	47
Total pulses	691	443	237
Total foodgrains	5276	1252	1359
Groundnut	1064	986	368
Cotton	1283	1076	222
Sugarcane	5746	4850	3531

* 100 metric tonnes for all crops except cotton; for cotton 100 bales of 180 kgs. each.
@ Since here the normal years are less than five, the regression was not fitted.

We are in the process of studying further the peak and trough residuals. An obvious line of investigation to pursue is to see their usefulness in studying crop-specific and aggregate crop-weather relationships, such as to have predictive value. It is difficult to think of an econometric researcher in India who does not pine for a serviceable weather index. Attempts have already been made by researchers to use the residuals around the trend line to get a weather index.[3] It is our feeling that the peak and trough year residuals are an even richer source of raw material for deriving crop-weather relationships. It is worth mentioning here a suggestive clue noticed during the preliminary look at the residuals. Taking the broadest crop group, *viz.*, Total foodgrains, we located years having residuals exceeding three standard deviations in value, plus or minus. Plausibly, these years are, weatherwise, 'very good' or 'very bad' years. Methods I and II indicated the following years while none of the years had the residual value exceeding the prescribed limit in the case of 'All years' residuals.

'Very Bad' years	'Very Good' years
1965-66	1973-74
1966-67	1975-76
1967-68	
1972-73	

Interestingly, this categorisation of extreme years appears to be in conformity with the rainfall data. Karnataka State consists of 19 districts. Table V presents the distribution of these districts according to their rainfall status in the years noted above; any comment on the table would seem to be superfluous.

TABLE V—DISTRICTS IN KARNATAKA STATE (19) CLASSIFIED ACCORDING TO THE LEVEL OF RAINFALL RECEIVED DURING THE SPECIFIED YEARS

Year	Below normal			Normal	Above normal		
	—50% and below	—30% to —50%	—10% to —30%	—10% to +10%	+10% to +30%	+30% to +50%	Above +50%
	(No. of dist-ricts)	(No. of dist-ricts)	(No. of dist-ricts)	(No. of dist-ricts)	(No. of dist-ricts)	(No. of dist-ricts)	(No. of dist-ricts)
1965-66	—	8	10	—	—	—	—
1966-67	—	—	6	7	4	2	—
1967-68	—	—	12	6	1	—	—
1972-73	—	5	5	5	4	—	—
1973-74	—	—	2	12	4	1	—
1975-76	—	—	—	1	7	6	5

Note:—Normal rainfall is based on 75 years' average.

3. James L. Stallings, "Weather Indexes", *Journal of Farm Economics*, Vol. 42, No. 1, February 1960, pp. 180-186.

Being exploratory in nature, it would be appropriate to conclude this paper on a speculative note. First, one may ask whether the model suggested in the paper retains relevance in the setting of modernized agriculture. In relation to agriculture, modernization refers to two principal changes: (*i*) diminishing weather-domination and fuller control of the producer over crop output and (*ii*) greater sensitivity of crop output to changes in the price-technology-policy environment. It may be recalled that our model is tuned not so much to weather as to fluctuations in output. In this sense, it is neutral to the factors making for fluctuations. As modernization proceeds, the non-systematic component would only begin to register the effects of a new set of factors replacing weather as the dominant influence on crop output; in fact, it could be an important witness to the pace and extent of modernization in agriculture. An interesting line of investigation here would be to see if the trough years (cropwise or for total crop output) show an increasing trend and if the gap between peak and trough years gets narrowed with technological advance and greater human control on output.[4] Second, separating peak and trough years from normal years amounts, in effect, to looking for what may be called 'single-point discontinuities' in the time-path of a variable. Econometricians often visualise a period as having structural breaks in the sense of changes in the parameters of relationships, *e.g.*, a shift in the line or change in its slope. It would be awkward to regard a 'single-point discontinuity' as a structural break; it would be equally awkward to shove it off into the catchall disturbance term. The concept of non-systematic component provides a fruitful way to bring such points under observation and to decipher their message. In this sense, the weather-dominated agriculture need not be the only field of application for this concept. After all, the world in which we live has many things besides weather which are equally fitful and unpredictable.

4. A detailed investigation into several aspects of uncertainty in crop yields, including this, is in progress at our Institute.