REGIONAL STUDIES AND RESEARCH FOR CONSISTENT AND OPTIMAL PLAN FORMULATION : THE NEED FOR A RIGHT KIND OF ORIENTATION

M. N. Pal*

1. Introduction :

A survey of regional resources, regional problems, etc., in different regions forms a basis for any consideration for regional development plan formulations. Yet the delineation of mere regions and sub-regions, showing common features and problems mostly at a static point of time in the past, can be considered as a stage that is far behind the actual plan formulation. This approach for delineation might, as claimed by Learmonth, be *holistic*, provided it could really incorporate relevant economic and social variables, without depending excessively on the readily available physical variables. But it goes without saying that merely identifying the problems is not sufficient, unless the regional studies evaluate the ways and means to solve them. In a regional study intended for planning purposes, we should not avoid this task of finding solutions explicitly.

2. Problem-Solving Procedure :

Following S. L. Optner [25], a problem can be defined as a situation in which there are two states : one is characterized by the present (existing) state ; the other by a proposed (or desired) state obtained as a by-product of pre-assigned goals, objectives and related constraints. In both states, there is a set of objects, attributes and relationships (interactions) that form what is called a system. Objects, attributes and relationships of a system are inter-locked in such a fashion that they generate an on-going state in the system, called a process. The processes are always at work, transforming a system from one of its states to another state at differential fashions. The purpose of planning is to condition the processes in such a way as to effect a certain desired transformation in the state of a system. The solution states how the gap between existing and desired states will be closed. When the gap can be fully closed, the solution is total, otherwise it is partial. A total

*The author is an Associate Professor, Indian Statistical Institute, New Delhi.

solution can also be tentative, if its reliability remains unconfirmed. The problem-solving procedure should consider the following major steps :

(a) fix the objective and its related constraints to identify the desired state of a system and the problem,

(b) state the condition of the problem and stipulate the necessary assumptions,

(c) establish the alternative processes and limitations to be valued,

(d) set the consistent criteria for comparative measurements of alternatives,

(e) determine the nature of risk (uncertainty of obtaining the desired state),

(f) select the optimal alternative, the probable consequences of which would be most preferable,

(g) make a plan for implementation of the solution through appropriate implementing authority or authorities,

(h) implement the solution on a pilot basis, if possible

(i) evaluate the results of implementation,

(j) modify the solution and iterate, if necessary, from an appropriate step stated above,

(k) implement the final solution.

Any regional study should cover at least the steps (a) to (f) for a plan formulation, if the plan implementation, covering all steps from (a) to (k), is not readily in sight. Thus the regional study that stops at step (a) without the identification of and suggestion on processes to overcome the problem, could be an academician's study, but not a planner's regional study, because of the incompleteness.

3. Systems Terminologies :

3.1. A systems view of planning is implicit in the preceding discussions. A brief discussion on the systems terminologies would therefore be necessary. Following Klir and Valach [17] and improving upon their mathematical statement, the definition of a *system* is contained in the statement that every set $S = \{ R_1, R_2, \dots R_n \}$ is a system with reference to a context, where the set of n objects with attributes a_i's, $i = 1, 2, \dots n$, are identified with reference to the context and represented by the set $R_1 = \{a_1, a_2, \ldots a_k\}$ and R_k's are sets containing interactions or relationships of any k elements of R_1 for $k = 2, 3, \ldots n$. A few or all elements of a set R_k can be zero representing no interaction or relationship. Mathematically speaking, a system is generally closed, meaning thereby that all relevant objects with reference to a context are identified in R_1. But in a real world problem-solving situation there exists a thing like *environment to a system.* Following Lewis Goslin [11] an enviroment to a system can be defined as follows : It is the set of all objects, a change in whose attributes affects the system and also of those objects whose attributes are affected by the behaviour of the system and at the same time the changes in attributes of those objects are not subject to control from within the system. On the other hand, if the changes could be controlled from within the system, the objects are to be treated as part of the system. As the scope of control from within an identified system is important for decision-problems, the distinction between a system and its uncontrollable environment is perforced. If the environment to a system exists alongwith the system, it is called an open system. If the environment does not exist, the system is closed. Now it looks quite logical to conclude that the decision-making, only through an open system analysis in the context of a real world situation, can not become comprehensive when the environment to the system remains unknown or is ignored. But a conditional sort of comprehensiveness is accomplished through the assumption that the behaviour of the system is affected only by given conditions previously established by the environment, provided the conditions themselves do not alter by the current change in the behaviour of the system. Many of our regional studies suffer from atleast the following two limitations . (i) non-consideration of all relevant interaction sets R_k's contained by a regional system, (ii) over-looking the difficult task of incorporation of the possible environment to a regional system.

3.2. Again many regional studies are in the form of a *sub-system* analysis. By definition, a sub-system is a system (open system) of lower order contained by some parent system. The element set R_1 of the parent system can be, for example, the set of sub-systems. The same parent system can be visualised as constituted by different collections of sub-systems simultaneously. Theoretically speaking, the decision-making should be within the framework of a total parent system. But practical limitations often perforce one to make a sub-system analysis and arrive at partial decisions. The partial decision-makings for sub-systems of a parent system could be permitted provided one is in a position to examine the relationships or interactions between sub-systems for final decision-making. If one is interested in analysing only one particular sub-system, it should not be ignored that other sub-systems of the parent system can at least be treated as environment to the sub-system in the analysis. This point is particularly important for those who are interested in a one-region or a one-sector analysis.

3.3. Descriptions, facts, data that are related to the context of a system. both quantitative and qualitative, need to the manipulated into information and channelled their flow for use in the evaluation of the states and processes of the system. All classes of information together with their flows and linkages, that tend to provide a comprehensive framework for the evaluation of the system, themselves form a system. As this system always enters in the analysis of corresponding contextual system, it can be called the *embedded information system* of the system. A real world information system, based on accessible information, may be a sub-set of the total information system, based on information required for a system. Depending upon the extent of *information gap* between required and accessible ones, various consistent assumptions are made on the gap and therefore techniques of analysis are modified or simplified for a real world decision problem. As there is a definite relation between the characteristics of decision problem and the techniques of analysis, the decision-making becomes less comprehensive with the increase in assumptions on information gap. If a large information gap is ignored and also if accessible information is unreliable, the decision-making can become less precise to erroneous. In fact, looking into a decision problem through an appropriate system analysis gives a promise of better information flow and more reliable information. However, the use of system analysis has not yet obtained a fair operational status in the socio-economic development problems. To achieve a momentum for operational status, complexities of total systems may need a simplification. This simplification can be done through some sort of physical or analytical *model systems*. The model systems are very often *homomorphic*, not *isomorphic*, to the corresponding total system. According to Klir and Valach [17], a system S' is homomorphic to another system S when elements in S' can be assigned uniquely to elements in S but not vice versa and the relationships in S' can be assigned uniquely to relationships or interactions in S but not vice versa ; two systems are isomorphic to each other if elements in S' can be uniquely assigned to the elements in S and vice versa and if for every relationship or interaction in S' there exists an exactly similar relationship of interaction in S and vice versa. If isomorphic model system can be identified, the problem of control for comprehensive decision-making is eliminated. But an analysis with a homomorphic model system, which is more commonly in use, may sometimes yield a heuristic or tentative solution with a high degree of uncertainty in obtaining the desired state of the system.

3.4. A nation is treated usually as a regional system, but, in general, a regional system can be as small as a metropolitan centre or as large as the entire world. Clearly a region within a regional system can itself be a sub-system. The context of regional systems, in the present discussion is the socio-economic development for the people of regions and of the system to which the regions belong. The environment of a regional system can contain, for example, the exogeneous international situation to which the system can be subjected to. At times a national planning system without regional dimension can belong to the environment of the corresponding regional system. By now it is clear that an identification of the on-going processes of a regional system is essential for planning purposes. But most of the regional studies in India are on regional delineation with some sort of decentralised analysis that captures usually a particular static state of the processes, very often in incomplete form, ignoring its dynamism. Regional delineations or regionalisations are based on the following three concepts of "region" [see Richard Hartshorne 13]:

Ad-hoc region : or a region in the general sense, is simply a particular piece of area which may be in some way distinctive from other areas,

Formal region : or a homogeneous or uniform region, is an area within which the variations of one or more selected features fall within a certain narrow range,

Functional region : or a region of coherent organisation or a nodal region, is an area in which one or more selected phenomena of movement connect the diverse localities within into a functionally organised unit.

According to these concepts it could be noticed that the formal and functional regions together can be taken to form a kind of regional system with sub-systems of regions inter-connected by spatial interactions only, on the assumption that the constituent diverse localities in a functional region are identified before hand by some sort of uniformity criteria of the formal regionalisation. But this form of regional system is incomplete in the sense that the system analysis for process evaluation is still to be incorporated into it. The ad-hoc regions, though have none of the defining features of a regional system, can be accepted some times to form regional sub-systems when we have identified their role, characteristics and linkages with reference to the main purpose of analysis in a particular system. For example, States of India are considered as ad-hoc regions for planning purposes, but they really form sub-systems for plan implementation under the compulsion of administrative controls. Thus, when the systems view of planning is imposed on the usual regional studies, more meaning could be attached and deficiencies would come to light [see also Chadwick 4, and Harvey 14]. In order to extend the usual analysis from the stage of delineation to the stage of integrated decision-making, the unidentified objects and interactions implicit in on-going processess should have to be traced and analysed through the use of certain techniques and stages of system analysis and also through a reduction of information gap in a recurrent manner.

4. Rational Looks to Regional Analysis Procedures :

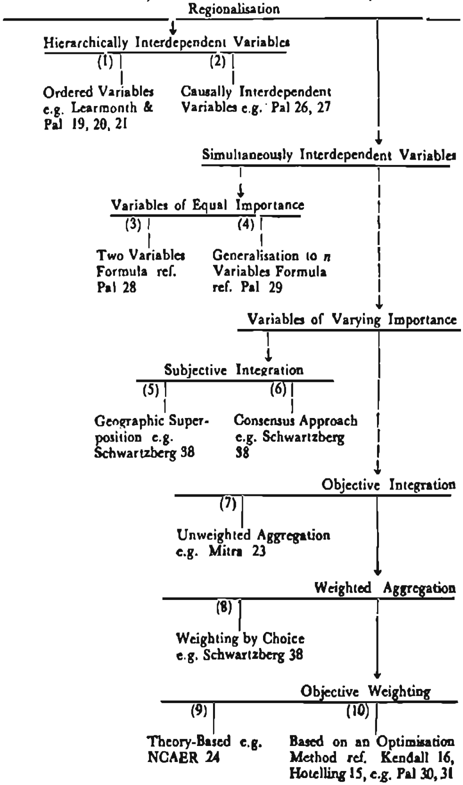
4.1. Attempts will be made now to search for rational orientations in the typical regional analysis procedures. We start, in this sub-section, from the concept of regionalisation. Whenever a normative type of regional planning measures is to be applied differentially on account of a spatial variation of certain relevant characteristics or variables, the concept of formal regionalisation, even without the consideration of interaction aspects at the initial stage, can be useful. But for this, we have to work very often for a composite formal regionalisation, involving many inter-related features of variables, rather than be content with simple formal regionalisations, involving only one variable in each. Analysis procedures for the evaluation of composite regional character can be classified as that in diagram 1. In this diagram, some ten types could be recognised that are shown in rectangular blocks, numbered (1) to (10) In each block, references are given after the authors of the method, or some typical Indian illustrations are quoted. A detailed review of these types is available in another study made by the present author [32] very brief comments are given below :

Type (I) :

Some times variables or groups of variables should be considered in order of their importance in the analysis. For example, the normal

DIAGRAM 1

Classification of Analysis Procedures Available for Composite Formal Regionalisation



rainfall may be considered as a basic variable, while the temporal variability of rainfall can be considered as an inverse measure of the reliability of normal rainfall and as such, it is useful when analysed together with the normal rainfall, without losing the identity of individual variables. Here a cartographic presentation of ordered variables is suitable and any quantitative aggregation is meaningless.

Type (2) :

To take decision on certain problem variables, the cause and effect analysis may be needed, in which certain other variables explain the change in the problem variables. For example, Engel's law of consumer behaviour may explain the change in demand of certain commodity with rise in income at regional level. Here statistical regression methods may be useful and any aggregation of cause and effect variables, losing the identity of individual variables will be meaningless.

Type (3) :

More often in regional studies, the variables can be such that they are reflective of certain composite character and thus they are interdependent without having any direction of interdependence. If the constituent variables are to have equal representativeness in the sense that they are equally correlated with the composite character, the formula developed by Pal [28] for two variables is useful. This composite index formula is also optimal in the sense that the aggregate correlation, that is accounted by such a linearly combined index with its constituent variables, is the maximum. This can be used as a very effective single measure of areal concentration that moderates the non-concommitant extreme values of the constituent variables.

Type (4) :

The composite index formula as mentioned in type (3) has been generalised by Pal [29] from two to any number of constituent variables. Here also the constituent variables are equally correlated with the composite index, but the optimality mentioned for the two variables case is not satisfied by this kind of index with three or more constituent variables. The situation in which it can be used have been discussed in the above reference and further discussions on a comparative advantage of this index over an optimal index of type (10) will be discussed later.

Type (5) :

In a situation similar to types (3) and (4), but with an unknown or a varying nature of representativeness of constituent features in detecting the composite character, geographers usually apply the technique of superimposition of varieties of both commensurable and non-commensurable information obtained through field experiences, opinion surveys, and on the basis of available quantitative and non-quantitative data. This mode of differentiating areas is not according to any strict quantitative formulation and as such the ranking of each area in comparison with all others, that may be needed in evolving the norms for decision-making, is not possible even though the areas could be differentiated qualitatively. As a typical example, Schwartzberg's qualitative differentiation of areas for levels of economic development in India can be quoted [28].

Type (6) :

In this approach, consensus is sought to grade different areas in respect of a composite characteristic from among a selected group of individuals. Here, individuals' opinions may be recorded in qualitative grades that can be quantified, averaged and used in ranking the areas in respect of the characteristic. Subjective biases enter here, because the individuals' judgment on quantitative grades are not based usually on a well-defined and common notion of the characteristic itself. Schwartzberg [39] attempted to rank the different areas in India for grades of development by opinion surveys from a handful of selected Indian economists and geographers, without stating any common definition of development.

Type (7) :

Here the gradings of areas are based not on the individuals' opinions as in type (6), but on the ordered ranks reflected by several variables that either influence or are influenced by a composite characteristic. Obviously, the differential degrees of influence that can be possible have been ignored here while taking some kind of unweighted average of gradings from different variables for composite rank of areas. Moreover, there can be mistakes in the judgment between direct and inverse influences of a variable that can exist with the increasing values of the composite characteristic. Ashok Mitra [23] has used a procedure of this type to obtain a composite development index for ranking Indian districts, based on gradings by 33 variables in five sub-groups. His study suffers from mistakes referred to above. For example, he presumed a direct variation of variables, like the proportion of persons belonging to depressed scheduled caste and scheduled tribe, with the increasing gradings of development. But these variables are really associated inversely with development at present; their potential value as "development assets" is yet controversial.

Type (8) :

In this type, a possible existence of the differential degrees of influence of variables as referred to in type (7) has not been ignored and some weights have been attached, purely by a subjective choice, to the rank-orders of different variables while these are averaged for composite gradings of areas. Apart from this limitation of subjective choice of weights, this procedure suffers from mistakes referred to already in type (7). Schwartzberg [38] has used this procedure of weighted aggregation by subjective choice of weights to some 33 variables for development gradings of 18 old Indian States (or State-groups).

Type (9) :

Economists attach objectively the price-weights to quantity variables usually. Weighting can also be done by some similar kind of theory-based objective judgment. Thus the attempt to estimate the per capita income of districts by the National Council of Applied Economic Research (NCAER) [24] in India is an illustration of price-weighted aggregation of economic activities of production and services. There exists, however, an estimation problem for areal incomes by both originating and accruing concepts for highly open economics of areas. The income, particularly by the accruing concept, is usually considered as a good measure of economic development. But, better, we consider this as one of the many possible measures of composite development.

Type (10) :

Let a composite characteristic be reflected by the variations in n variables, each varying over N areal units of observation. As a common characteristic is influencing, or be influenced by, all n variables, it is likely that the variables are strongly inter-related. If the inter-correlations were all perfect (having an absolute value of unity for a correlation coefficient), the areal units identified by the vectors of the n values of variables would approximate to a straight line when the statistical distributions of variables are similar. In practice, however, the correlations are not expected to be perfect, We, therefore determine the straight line of closest fit to the cluster of N vectors (or points in the Euclidean n-space) and rank the areal units by a related formula that gives the order of points of projection onto the line of closest fit. This index formula developed by Kendall [16] is optimal in the sense that the aggregate correlation that is accounted by such a linearly combined index is the maximum. By this procedure, he constructed a composite agricultural productivity index for his studies by countries in England as early as in 1939. It should be noted that the composite index by Kendall's procedure is formally equivalent to what Hotelling [15] called earlier in 1933 a "first principal component". Hotelling's principal component analysis is not restricted in the determination of only the first principal component, nor it starts with a choice of constituent variables that should be sufficiently inter-correlated. In principal component analysis we compute a series of principal components which are mutually taken as orthogonal (independent) for a mathematical convenience and second or subsequent principal component explains the maximum of residual variations (i.e. aggregate squared correlation) that remain unexplained by the preceding component or components. While the optimal property of a first principal component for a set of inter-dependent variables (resulted for the influence of a single characteristic) is useful, the mathematical convenience of the breaking into orthogonal principal components for a set of heterogeneous variables with and without interdependence (possibly caused by the mixed influence of several characteristics or sub-characteristics) is not acceptable for construction of several composite regional indices, since the regional indices are mutually oblique (dependent) almost as a rule. Thus, in a situation when we are to handle a group of heterogeneous variables that could be considered as to influence or be influenced by several characteristics or sub-characteristics, the construction of several principal components by Hotelling's procedure is not the answer. In such a situation Pal [30, 31] divided the group of variables into sub-groups in such a way that variables in a sub-group have high inter-correlations, while the cannonical correlation between a pair of sub-groups is low on the average. He then constructed sub-group indices which are first principal components of different sub-groups, or equivalently the Kendall's sub-group indices. It should better be called by Kendall's index formula, rather than a first principal component, since the principal component analysis has one kind of statistical implications while the Kendall's procedure has another kind. For example, in a full-fledged principal component analysis, there is no need for any assumption on the statistical distributions of the variables and in that the required number of principal components, explaining almost the total variation, must be brought into the analysis. On the other hand, it is already mentioned that variables considered in Kendall's procedure must have similar statistical distributions, since this assumption on distributions is essential for a validity of hypothesis on the linear type of correlations between variables as implicit in the inter-correlation matrix. Thus in Kendall's formula one should make suitable mathematical transformations of initial variables in order to have approximately similar distributions for transformed variables with the results of getting an improved inter-correlation matrix and consequently an improved magnitude of aggregate correlation of the Kendall's index with its constituent variables. Pal's illustrations for his studies in South India [30] and India [31] were on the construction of several indices for sub-characteristics of over-all development by the procedure of type (10); the sub-indices were then combined to obtain a composite index of development by one of the procedures mentioned in types (3) or (4), (9) and (10), that suits the particular situation.

4.2. The type (10) can be considered as to be the best among the types (5) to (10) that could be applied in a more or less similar situation with inter-dependent variables. If a proper theory is available, the type (9) can also become useful, particularly when the inter-dependence of variables is very poor. However the type (10) has a greater generality in the sense that more variables could be handled in it, including even a type (9) index as one of the variables for type (10). Most of regional studies that are related to the procedure of type (10) have, however, taken the principal component factorisation of Hotelling, extracting two or more mutually orthogonal principal components, rather than the more appropriate procedure of Kendall's formula for a single factor over a properly chosen group of variables, despite the fact that the present author has repeatedly cautioned against the use of second and subsequent principal components, in connection with his studies [30, 31]. For example, the studies made by Prakasa Rao, Berry, Das-Gupta, Kulkarni and others [36, 1, 2, 9, 18] with Indian data are of this type. In one illustration of Berry and Prakasa Rao [2], they have gone a step further by rotating the extracted principal components by "normal varimax" criterion [for details of this method, see Harmon 12]. The present author has already pointed out [see Pal 33] that this way of handling principal component analysis do not lead to the "simple structure" [see Thurstone 43] for meaningful interpretation of components, even after rotation by the varimax criterion; moreover the optimal property of principal components is lost in rotated components. In a recent application Das-Gupta [9] has extracted first five principal components, explaining 90 per cent of total variation of some 15 variables. But surprisingly he has named the first principal component as the index of socio-economic development that explain only the 44 per cent of total variation, while there is no mention of why other four principal components are not brought into analysis. He has also wrongly stressed that the first components is "known as the principal component", while, in fact, all components are called principal components His use of first principal component only is also not identical with the Kendall's procedure when he stressed again that "this analysis is independent of any distributional assumption", meaning thereby that he did not care for suitable transformations of variables as referred to in Kendall's procedure. Again he has made pairwise comparisons between his socio-economic indices A and B and Mitra's [23] development index C, say, where indices are as follows:

A : first principal component explaining a variation of only 28 per cent of total variation from 24 variables,

B : first principal component explaining a variation of 44 per cent of total variation from 15 variables,

C : type (7) index over 33 variables.

These comparisons, intended to measure the extent of departures of C from A or B and of A from B, were made through proportions of areal units that do not have the same classified ranks. A better way to make these comparisons would have been by a calculation of correlation coefficients which could have been tested statistically to say whether the departures were significant or not. But what he did finally in the name of discriminant analysis looks like a non-statistical or a mis-appropriate handling of statistical tools. Here he has constructed several linearly combined indices called discriminant functions with several sets of weights for variables, depending upon the number of classes obtained in the classification made by any of above mentioned indices A, B and C; the discriminant function for a class has that set of weights which maximises the discrimination between the class and the rest; finally he classified an areal unit afresh to that class whose discriminant function attained the maximum value for the areal unit and counted the number of misclassified areal units by any of indices A, B and C. It is difficult to understand how he could assert that this mode of discriminant analysis is "a more rigorous test" to the classification made by any of the indices A, B and C. It can be clearly seen that in this analysis one is deliberately changing the set of weights for variables of an areal unit with the sole purpose that its difference from other areal units in other classes become as maximum as possible; this deliberate change of weights has no justification in terms of degrees of influence between the index and its constitutent variables and also there has been different standards of discrimination for different areal units, as sets of changed weights can be at most as many as the number of classes. This should be emphasized here that the degree of representativeness of variables in the index constructed is important, particularly in a situation where our selection is such that the variables are strongly induced or influenced by the composite characteristic to be depicted by the index. If there are reasons to believe that a particular constituent variable is not duly represented in any linearly combined index, the use of such an index would raise controversy. It has been illustrated in the paper by Chattopadhyay and Pal [5] that even the optimal index mentioned in type (10) need not necessarily be the one that is most suited for a proper composite regionalisation, in the absence of a due representativeness of certain key variables. In this illustration, it has been proved statistically that a type (4) index by Pal's procedure can be more useful than even a type (10) index by Kendall's procedure.

4.3. It has been already mentioned that a normative type of decision can be made on the basis of results obtained in a composite regionalisation. For example, in all the illustrations referred to under types (5) to (10), some kind of composite index of development was constructed, whatever be the procedure, for ranking or distinguishing different areal units in India or in a part of India. Now the question arises, why there is so much of attention to identify differential development ranks of areal units in regional studies? Its answer lies in the fact that a reduction in regional disparities in development has been accepted universally as a consistent regional policy. As the extent of development activities is directly related to the amount of planning investment, one can take an appropriate composite index of development as the inverse measure or norm for planning investments. Thus Pal [30] calculated the allocation norms of investment for different regions in South India wherein a higher share of investment was proposed for a less-developed area, consistent with the above mentioned regional policy. This mode of allocation of regional investment, though appears to be consistent with certain regional policies, need not necessarily be the most rational procedure, as there has been an absence of consideration for other relevant techno-economic parameters. It is generally noticed that a less-developed or less-productive region has a higher value of the incremental investment-income ratio k, i.e., if the region 1 is more developed than the region 2, we expect to get $k_1 < k_2$, where k_1 and k_2 are values of k in regions 1 and 2 respectively. Thus to generate one unit of additional income, we need to invest more in region 2 than in region 1. So, it may look uneconomic, from the angle of optimal income generation, to invest more in region 2, contradicting our earlier stand under the policy of reduction of regional disparity in development. But the case of a less-developed region could still be established if we could show that it has a higher savings rate s, than that of a more-developed region. Following Rahman's [37] analytic proof, one can argue that an optimum income generation programme, satisfying even an appropriate regional disparity constraint, can still favour the less-productive region 2 for investment allocation in a number of initial years of a sufficiently large plan period, provided we have $s_2/k_2 > s_1/k_1$. On the other hand, there would be no case for the less-developed region 2, if $s_2/k_2 < s_1/k_1$. Thus when the savings rate is sufficiently higher in a less-developed region, it might get a more favourable allocation of investment initially, which is consistent with the preceding normative decision. But if the savings rate is very low in a region as compared to that of other more-developed regions, a favourable allocation of investment in the less-developed region under the preceding normative decision would lead to controversy and criticism, particularly in a federal set-up like that in India. That is why we notice that a more developed State like West Bengal finds injustice in the central allocation of planning investment, following a somewhat similar kind of arguments. But it is more surprising to note that our national planners take decisions on regional investments without an assessment of regional saving rates. We should focus our attention to remove this regional information gap at the earliest possible time.

4.4. The importance of functional regionalisation lies in the fact that it is based on an assessment of spatial interactions induced by productive activities that require the flow of factor inputs including labour and investment and also generate the movements of products to the markets or sources of demand and the travels of consumers to the markets or sources of supply. Thus for functional regionalisation one needs an examination of the flows of various commodities to feed the functional activities and the movements of people to satisfy their functional needs, besides the knowledge of various functional activities on space. Certain selected flow mappings are in use for determining inter-locational and inter-regional linkages or relationships [20]. But there is a terrible information gap in flow data so that these maps do not give the full story of the inter-locational relations. Because of this information gap, many studies on functional regionalisation have taken the course of indirect delineation, for examples, by gravity models [for a brief introductory account, see Pal 32], or by Christaller's central place theory [6, also see Losch 22]. But the delineation of functional regions by gravity models is very crude and it does not tell anything on the linkages between localities. The central place model tries to fit a hierarchy of linkages between localities on the basis of an assumption of regular hexagonal market areas. This assumption seems to be inconsistent, more often for localities at the higher order of the hierarchy. However, a procedure developed by Pal [29] and illustrated for a study in Muzaffarnagar district of Uttar Pradesh [34] establishes the hierarchical linkages among higher order localities (or central places) including metropolises, cities, towns and large villages, without the above mentioned regularity assumption, on the basis of following two consistent hypotheses:

(i) people will try to move a route distance as minimum as possible for the satisfaction of their functional needs,

(ii) people of lower order central places will move to higher order central places only when the functional rank of higher order central places is sufficiently higher as compared to that of lower order central places.

Details of the procedure are available in the references quoted above. This procedure also gives an optimal pattern in the sense that the linkages are established with the objective of minimisation of route-distances under certain constraints on functional pull between localities of different functional ranks. Now question arises how the studies in functional regionalisation obtained on the basis of either flow data directly or other procedures indirectly could be used for decision-making? The usual regional studies that depict a present state of functional linkages between localities are never sufficient in the absence of any knowledge on its future state. The future state is bound to be different mainly for two reasons. Firstly, for a given distribution of functional activities, the spread and hierarchy of central places that exist at present need not necessarily be rational always or desirable everywhere. Secondly, the distribution, nature and intensity of functional activities are likely to change in the course of time, particularly in a developing country like India, with the result of a changed pattern of spatial interactions. Unfortunately, in the usual regional planning practices in India, there has not been much attempt to visualise beforehand such potential changes in the spatial interactions. Naturally, much needed planning measures towards providing the infra-structural requirements for such changes are lacking there, resulting into bottlenecks in spatial flows and also into retardations of possible growth prospects in different areas. As the procedure put forward by Pal is based essentially on the pattern of functional activities at different localities, the procedure can also establish a future state of spatial interactions on the basis of a predetermined functional activity pattern for a future date. The present and future states could then be compared to evaluate the dimension of change for planning considerations. In this procedure, an aggregate intensity of spatial interactions could be anticipated without the details of specific route flows. A better procedure would have been (i) to study the distribution of a specific functional activity or industry, (ii) to apply the well known linear programming procedure for an optimal determination of the specific route flows and (iii) to make a subsequent synthesis over all such individual functional activities for a final evaluation of all spatial interactions together. A typical linear programming model of the above type has been presented by Pal [29] elsewhere which gives the types of economic parameters and information needed for its application. But even such a simple application of linear programming procedure is yet difficult, because of a substantial kind of information gap. For this reason the procedure put forward by Pal is more operational. There can be other kinds of decision-making for spatial interactions, all of which can hardly be discussed within the limited space or time available at present. Let me mention only a small example of interest to a city planner. A city can also be considered as a fuuctional region with a more densely populated CBD (Central Business District) linked with neighbouring areas and suburbs having gradually falling densities. There can be many complicated models that may fit the falling density pattern in a city; one simple such model used by Colin Clark [6, 8] can be presented as follows:

$$y = A\,e^{-bx},$$

where

y = residential population density in thousands per square mile,

x = the distance from the CBD (or the centre of the city where a hypothetical density can be extrapolated by putting $x = 0$), and

b = coefficient of sprawl or spread of the city.

It should be noticed that the coefficient b varies greatly between cities. A high vale of b indicates compactness. With low values of b, the density falls off gradually and the city spreads out over a considerable distance before a minimum, or rural, density is reached. This *b* is obviously related to the intensity of public transport system available in the city. An empirical investigation over Indian cities would help in establishing the norms for public transport system required in relation to various values of *b*. These norms could be used for growing cities with changed coefficient of sprawl *b* for finding the additional transport requirements. But these norms are yet to be established for Indian cities.

4.5. With all these discussion on formal and functional regionalisations and related decision-makings that are possible with them, we have not yet identified the processes of growth that should be conditioned under a development planning. The future state of functional activities or industries yet remain unpredicted, even though we have mentioned on the optimal regional allocation of investment, consistent with the regional policy of reducing disparities in development and on the planning for infra-structure to take care of a changing spatial interaction pattern. For conditioning the processes of development, one should indentify the existing processes, particularly those combinations that caused an under-developed or less-developed state and also the nature of conditioning needed to take the present state to a desired state. The real cause of an under-developed state could be attributed very often to a disproportionate share of such activities that are of low productivity. Thus, vast under-developed areas of India have the low productive agricultural occupation as the only or major means of livelihood. An industrial area can also remain in an under-developed state, because of a high concentration of such industries that have generally a low productivity. The absolute magnitude of different industrial activities at different locations, together with the productivity ratings of these industries, would help in ascertaining the fact whether or not an area has considerable magnitude of highly productive activities. Areas having a considerable magnitude of highly productive activities are likely to have a sustained growth by multiplier effect. Areas having a disproportionate share of low productive activities, including agricultural activities, will become relatively less and less developed in the course of time as compared to the national level of development. As such a reorientation (or a different combination) of activities is necessary in order to check the growing regional disparity with the growth of the national economy. To understand the rationale of reorientation, it should be noted that the productivity of an industry may be improved by increasing productive capacity in conformity to the consumer demand for the outputs of the industry and by decreasing the factor inputs per unit output of the industry. Thus the growth of a region could be effected by that feasible combination of activities which would generate a speedier growth of effective demand for outputs from the combination. The growth of a region could also be effected by an appropriate conditioning of existing processes which would require a series of background studies, suggesting the technological improvement possibilities with due care for the full utilisation of region's resources and labour. Further details on this line of arguments have been discussed by the present author elsewhere [29, 32]. However, there has not been many regional studies along this line of arguments in India. Contrary to this line of arguments, regional studies often conclude with the suggestion of maintaining a status quo in the combination of activities rather than of visualising a feasible reorientation of activities for a speedier growth. Thus in India, agricultural regions are allowed to remain the same with marginal improvements in the factor inputs combination only, knowing fully well that the non-agricultural productivity is likely to remain always at a higher level. Regional studies on technological improvement possibilities are virtually absent in India, except in agriculture, because of partly the information gaps and partly the lack of a proper dissemination of information already available for industries at various locations.

4.6. When the conditioning of processes required for a desired state could be visualised with the knowledge of substitution possibilities between factor inputs for different industrial firms over differnt regions at different scales of returns of output, one can think of applying a homomorphic model system as used by Bhatia [3] for India through linear programming procedure for determining the complex optimal flows among regions for all productive sectors over different periods of a planning time horizon. But Bhatia thought of conditioning the processes by a borrowed technology from the U.S.A. for India without examining for a feasible technology that would be most suited at different regions, nor did he divide a productive sector by size of firms to take care for the diminishing returns of scales. It is agreed that an application of linear or other programming procedure in this manner would pose a tremendous amount of information gap in India. But if this information gap can not be overcome in an initial under-developed economy, one should only be content with simple applications of this procedure over a single productive sector only, treating the others as an environment to the regional system of the productive sector. Such an one-sector application should however have to be followed for all sectors separately and final adjustments are to be done by iterative corrections for over-all consistency. Dutta-Chaudhuri [10] made such a simple application for the sector of metallurgical and machine building industries in India, taking other sectors as parts of environment to the regional system of that sector, though, however he also could not take care for the scales of returns and did not visualise a technology with substitution possibilities. It should be emphasized that, in a development economy, a future technology is likely to be different from the existing technology, because of certain more advantageous substitution possibilities. Thus any attempt to predict the future, as is usually in vogue for consistency checking, on the basis of a single year observation of inter-industrial transactions, whether by regions or for the nation, would be highly inconsistent. Before using a single year inter-industrial transactions for consistency checking in predicted values, we should analyse them over a number of years for trends of change in inter-industrial coefficients over different regions.

5. Conclusion :

Preceding discussions have not exhausted all possible lines of orientations of the regional studies and research, that can occur in connection with consistent and optimal plan formulations. But whatever have been discussed are possibly sufficient to point out that there is enough scope for improved analyses with better results and decision-making in the context of regional planning. The improved analyses would however, require a thorough knowledge of various fields including statistics, econometrics and other technical subjects, besides the recurring efforts for reduction of information gaps. But unfortunately statisticians, econometricians and other technical people working in this field are very limited in number. This is partly due to the complexities inherent in regional research, which have been aggravated more by tremendous information gaps, concommitant with the degree of under-development of the national economy. Unless this limitation of personnel could be removed, the prospect of a right kind of regional planning is bleak in India.

Acknowledgement :

I gratefully acknowledge the benefit of discussion with Dr. C. R. Rao, Sc.D., F.R.S., Director of the Research and Training School, Indian Statistical Institute, on certain statistical aspects referred to in this paper.

References :

1. Berry, B. J. L. and Others : "Essays on Commodity Flows and the Spatial Structure of Indian Economy." The University of Chicago Press, Chicago, Dept. of Geography, Research Paper 111, 1966.

2. Berry, B. J. L. and V. L. S. P. Rao : "Urban-Rural Duality in the Regional Structure of Andhra Pradesh : A Challenge to Regional Planning and Development." Paper Contributed to the Symposium on Regional Planning, 21st International Geographical Congress, New Delhi, India, November-December, 1968.

3. Bhatia, V. G. : "Inter-Regional Allocation of Investment in Development Programming." Doctoral Dissertation, Harvard University, Cambridge, 1961.

4. Chadwick, G. : "A System View of Planning." Pergamon Press, Oxford, 1971.

5. Chattopadhyay, R. N. and Pal, M. N. : "Some Comparative Studies on the Construction of Composite Regional Indices." Indian Journal Regional Science, Vol. 4, No. 2, 1972.

6. Christaller, W. : "Central Places in Southern Germany." Englewood Cliffs, New Jersey, 1966, (English Translation).

7. Clark, C. : "Urban Population Densities." Bulletin De L' Institute International De Statistique, 1958, Stockholm.

8. Clark, C. : "The Economic Function of a City in Relation to its Size." Econometrica, 13, 1945, p. 112.

9. Das-Gupta, B. : "Socio-Economic Classification of Districts : A Statistical Approach." Economic and Political Weekly, Vol. 6, No. 33, August 14, 1971, pp. 1763-74.

10. Datta-Chaudhuri, M. : "Rlegaion Specialisation in Metallurgical and Machine-Building Industries in India in the Framework of a Planning Model for Optimum Use of National Resources." The Indian Economic Review, Vol. 2, (New Series), No. 2, 1967, pp. 219-270.

11. Goslin, L. N. : "The Product Planning System." Homewood (Illinois), Richard D. Irwin Inc., 1967.

12. Harmon, H. H. : "Modern Factor Analysis." The Chicago University Press, Chicago, 2nd edition (revised), 1967.

13. Hartshorne, R. : "What Do We Mean By Regions?" Contribution to the 54th Annual Meeting of the Asson. of Amer., Geogr. Los Angeles, California, August 18-22, 1958.

14. Harvey, D. : "Explanation in Geography." Edward Arnold, London, 1969, pp. 447-480.

15. Hotelling, H. : "Analysis of a Complex of Statistical Variables into Principal Components." Journal of Educational Psychology, 24, 1933, pp. 417-420 and 498-520.

16. Kendall, M. G. : "The Geographical Distribution of the Crop Productivity in England." Journal of Royal Statistical Society, 102, 1939, pp. 21-48.

17. Klir, J. and M. Valach : "Cybernetic Modelling," English edition, Iliffe Books Ltd., London, 1967.

18. Kulkarni, G. S. and M. P. Kharo : "Effect of Data Transformation on Regional Delineation by the Principal Component Method," Paper Contributed to the Symposium on Regional Planning, 21st International Geographical Congress, New Delhi, India, November-December, 1968.

19. Learmonth, A. T. A. and M. N. Pal : "A Method of Plotting Two Variables on the Same Map Using Isopleths," *Erdkunde*, 13, 1959, pp. 145-150.

20. Learmonth, A. T. A. and His Team : "Mysore State Vol. I : An Atlas of Resources," Asia Publishing House Bombay, 1960.

21. Learmonth, A. T. A. and His Team : "Mysore State Vol. II : A Regional Synthesis, Asia Publishing House, Bombay, 1962.

22. Losch, A. : "The Economics of Location," Yale University Press, New Haven, 1954.

23. Mitra, A. : "Levels of Regional Development in India," Census of India 1961, Volume I, Part I-A(i) : Text, 1965, Govt. of India Press, New Delhi, Part I-A(ii) : Tables, 1966, Govt. of India Press, New Delhi.

24. NCAER : "Inter-District and Inter-State Income Differentials—1955-56," National Council of Applied Economic Research, New Delhi, 1963.

25. Optner, S. L. : "Systems Analysis for Business and Industrial Problem Solving," Prentice-Hall Inc. Englewood Cliffs (New Jersey), 1965.

26. Pal, M. N. : "Differential Regional Pattern of Consumption of Commodities in India," AICC Economic Review, 18, No. 5, 1966, pp. 7-12.

27. Pal, M. N. : "Regional Demand Projection for Different Commodities in India : 1970-71 and 1975-76," Indian Economic Journal, 14, 1967, pp. 454-466.

28. Pal, M. N. : "Zur Berechnung eines Kombinierten Konzentrationsindexes," *Raumforschung und Raumordnung*, 21, 1963, pp. 87-93.

39. Pal, M. N. : "Quantitative Techniques for Regional Planning," Proceedings of the Conference of the Commission on Quantitative Methods, International Geographical Union. London School of Economics, August 18-20, 1969. Published in Indian Journal of Regional Science, 3, 1971, pp. 1-33.

30. Pal, M. N. : "A Method of Regional Analysis of Economic Development with Special Reference to South India," Journal of Regional Science, 5, 1963, pp. 41-58.

31. Pal, M. N. : "Regional Disparities in the Level of Development in India," Paper Presented to the Fifth Econometric Conference of India, Delhi School of Economics, Delhi, December 20-22, 1965.

32. Pal, M. N. : "Information System and Regional Development in India," United Nations Research Institute for Social Development, Geneva, UNRISD/70/C.74, GB 71-3383, October 1970.

33. Pal, M. N. : "Use of Principal Component in Regional Analysis," Paper Contributed to the Symposium on Regional Planning, 21st International Geographical Congress, New Delhi, India, Nov.-Dec., 1968.

34. Pal, M. N. : "On the Evaluation of Spatial Configuration of Functional Linkages," Paper Presented to the Seminar Organised by the University Department of Human Geography, Delhi, on 11th February 1970.

35. Pal, M. N. and R. N. Chattopadhyay : "Some Comparative Studies on the Construction of Composite Regional Indices," Forth-coming in the Indian Journal of Regional, Science, Vol. 4, 1972.

36. Prakasa Rao, V. L. S : "Rational Groupings of the Districts of the Madras State." Indian Geographical Journal, 28, 1953, pp. 33-43.

37. Rahman, M. A. : "Regional Allocation of Investment," Quarterly Journal of Economics, 77, 1963, pp. 26-39.

38. Schwartzberg, J. E. : "Three Approaches to the Mapping of Economic Development in India," Annals of Asson. of Amer. Geogr., 52, 1962, 455-468.

39. Thurstone, L. L. : "Multiple Factor Analysis." The Chicago University Press, Chicago, 1947.