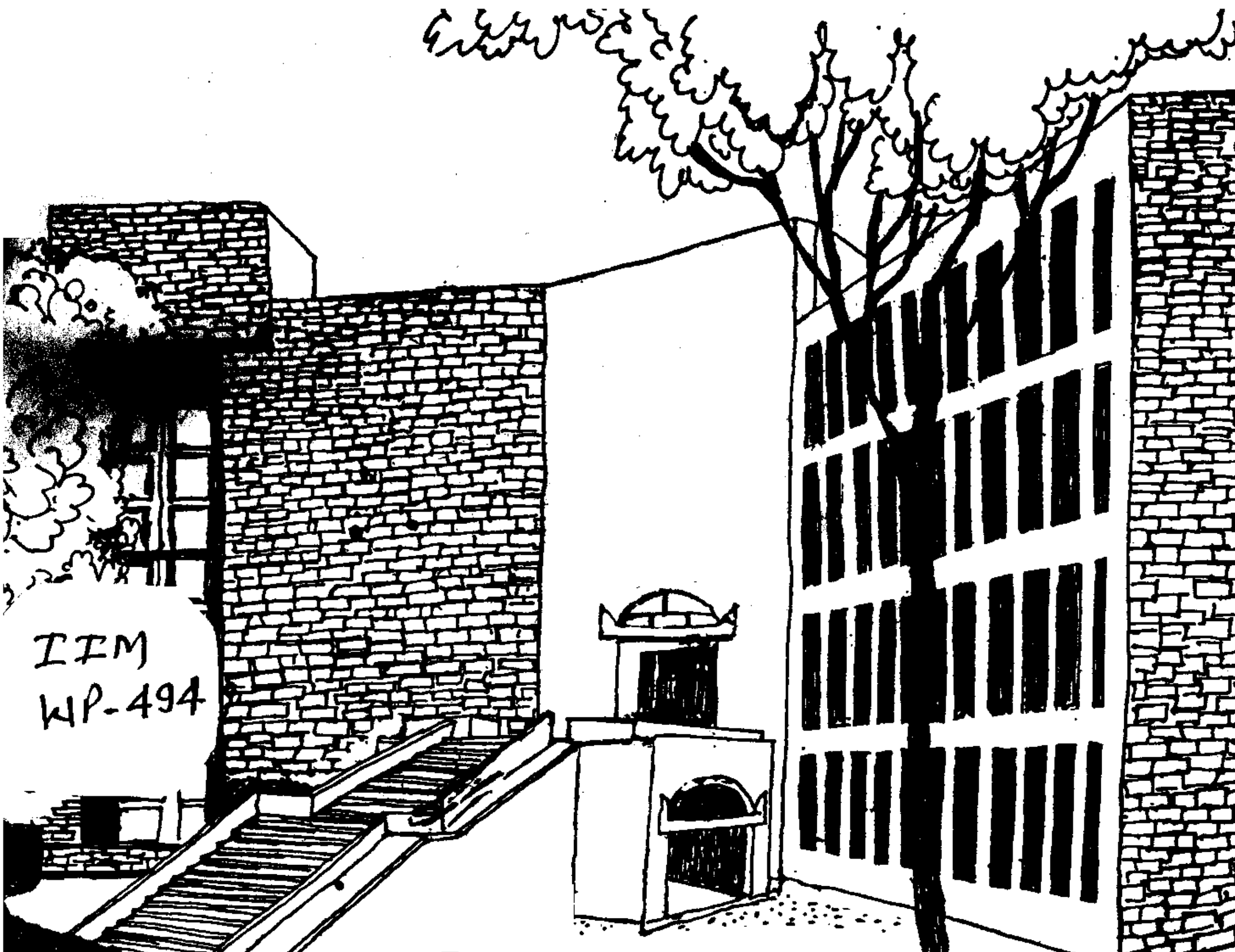
विद्याविनियोगाद्विकासः
IIM
AHMEDABAD

W. P. 494

Working Paper



PLANNING FOR RURAL ROADS IN INDIA

BY

Nitin R. Patel

&

T. Madhavan

W P No. 494

February 1984

The main objective of the working paper series of the IIMA is to help faculty members to test out their research findings at the pre-publication stage.

INDIAN INSTITUTE OF MANAGEMENT
AHMEDABAD-380015
INDIA

PLANNING FOR RURAL ROADS IN INDIA*

Nitin R. Patel & T. Madhavan

1. INTRODUCTION

This paper summarizes our experience in developing an interactive decision-support system for planning rural roads in one district in India. The aim of our research was to explore the potential benefits to be obtained from use of computer-based quantitative models in designing rural road networks. The cost per km. of rural roads varies from Rs.50,000 to Rs. 270,000 (US $ 5,000 to US $ 27,000) and estimates made by the National Transport Policy Committee [1] suggest that an overall investment of Rs. 110,000 million (US $ 11 billion) would be required to connect all villages in India with roads. In view of the very high amount involved, a project was mounted to examine if computer-based approaches which view connectivity from a network standpoint would provide significant scope for reducing rural road construction costs. The district of Kheda in the state of Gujarat was chosen to provide the setting against which the computer approach could be tested. This district with an area of 7,194 sq. km. and a population of 3.015 millions is in the state of Gujarat and includes some of the most fertile land in the state. It has 18 towns and 962 villages. An important feature for our purposes is that: 'Except for a small hilly area in the northern part, the district is an unbroken plain sloping gently from the north-east towards the south-west' (Kheda District Gazeteer [2].

* This paper has been prepared for the invited session on Transport at the International Federation of O.R. Societies Conference, IFORS'84, August 6-10, 1984, Washington, U.S.A.

This is also evident from road maps of the region. Even at a scale of 8" to a mile virtually all roads are straight lines joining pairs of villages. Kheda District consists of 10 sub-units called talukas* which vary in extent from 474 sq.km. to 1,195 sq.km. and in population density from 165 to 600 sq.km.

Since our concern was only for rural roads, the basic question that provided our research focus was the following:

Taking the system of national highways, state highways and district roads as given, could the same level of access be provided with a rural road network that is of shorter length than the existing one?

The question was addressed at the taluka level for the talukas of Anand, Balasinor, Cambay and Nadiad in Kheda district.

2. CRITERIA FOR MEASURING EFFECTIVENESS OF A RURAL ROAD NETWORK

The functions of rural roads, as set out in the Report of the National Transport Policy Committee [1] are to "serve as feeders linking villages with each other as well as with the nearest district roads, state or national highways, railway stations and market centres". These roads are viewed as supporting socio-economic development by providing access to facilities that serve both economic activities such as agricultural production as well as social development in sectors such as education and health.

---

* A taluka is an administrative unit corresponding to a county in the U.S.

Examples of sectors and facilities of importance for rural roads are given below:

	Sector	Facilities
1.	Education	Primary and Secondary Schools, Colleges
2.	Health	Clinics, Primary Health Centres, Sub-Centres, Hospitals.
3.	Agricultural Outputs	Government regulated markets, local area markets
4.	Agricultural Inputs	Hybrid Seed Stores, Fertilizer outlets, Veterinary Services
5.	Household needs	Hats (small periodic markets), shops
6.	Legal/Administrative	Land records office, Civil Courts
7.	Commercial	Banks, Credit Co-operatives, Offices (employment)
8.	Communications	Telephone, Post Office, Telegraph Office
9.	Industrial	Factory, mill, industrial estate (employment)
10.	Recreation	Cinemas, Fairs, Folk-shows.

A key issue in quantitative approaches to the design of rural road systems is the development of a suitable index to serve as a criterion for comparison of alternative configurations. Such an index should reflect the extent to which a given configuration achieves the basic purposes of rural roads mentioned above. It must take into account the relative ease with which the various socio-economic facilities can be accessed, since utilization of such facilities depends to a large extent on ease of access (Hine [3], Massam [4], Owen [5] Patel [6]). To be useful it must have the following properties:

1. Its value should depend on the location of socio-economic facilities in the area. In other words, whereas network A may be the best amongst a number of alternative networks for a given configuration of facilities, for a different configuration of facilities another network B may be superior to A.

2. It should be a measure that provides comparability between different regions. In other words it should be possible to say that network A has an effectiveness in Taluka X that exceeds that of network B in Taluka Y.

3. It should be computable without excessive data collection requirements. In otherwords it should depend on readily available data.

The commonly used indices of connectivity employed in the literature such as the cyclomatic number, the α, β and γ indices and the Shimbel-Katz index (See Garrison and Marble [7] for definitions) are all based on graph-theoretic considerations. Thus, while they provide reasonable aggregate measures for connectivity in a network they are inadequate for specific decisions such as the inclusion or exclusion of a particular link. This has been pointed out by a number of authors (e.g. Hay [8], Sharma [9]). The major drawback is their graph-theoretical nature which ignores important network information, such as the lengths and orientation of links. In our context accessibility is highly dependent on the distance of travel and a more appropriate criterion would be what Shimbel [10] has called the dispersion of the network. This is the sum of all shortest paths between every pair of nodes in the network. However, this is inadequate for us as it gives equal weights to all node-pairs. In our context lengths of paths from a village to other villages or towns are important only if they have economic or social facilities unavailable at the village itself. What we are proposing as an index can be viewed as a weighted dispersion index with weights chosen in a particular way.

Suppose we concern ourselves with only one type of service facility (for concreteness take, say, secondary schools) what would be an index that will satisfy our requirements? Let us assume that we know over the planning horizon (a popular horizon in India for road planning is 20 years*) where the secondary schools will be located. If there were no budget limitations we would build a road from each village to the nearest village having a secondary school. In a flat terrain such as Khoda this would be a straight line road. The ideal network would be the collection of all such roads. Let us suppose, under this ideal network, village is is at a distance d_i from the facility. Now consider any other network A. We can develop a measure which indicates the extent of deviation from the ideal of this network by considering the shortest distance a_i using network A to get to a secondary school. Then an index of the effectiveness of network A for village i in terms of access to secondary schools is $r_i = a_i/d_i$. Notice that by definition $a_i/d_i \geq 1$.

An overall measure $\bar{r}$ across all villages would be the population weighted average ratio a_i/d_i i.e.

$$\bar{r} = \sum_i (p_i/P)\, r_i$$

Where p_i = Population of village i,

r_i = 0 if the facility is in location i, and

$P = \sum_i p_i$ = total population in the area.

The measure $\bar{r}$ has a nice intuitive interpretation: 100 ($\bar{r}$ - 1) is simply the average extra distance (in percentage terms) travelled by the people in the area in using network A instead of the ideal network.

Alongside this measure we propose to carry along another measure r^{max} which is the highest r_i ratio amongst all villages. This will indicate the worst service-level provided by network A.

Using the above approach, we could, conceptually compute $\bar{r}_j$ and r_j^{max} for each type of service facility j. These could be aggregated into overall indices $\bar{R}$ & R^{max} for road network A by using a system of weights W_j which reflect the importance, frequency and cost of movement for each service.

$$\bar{R} = \sum W_j\ \bar{r}_j,\quad R^{max} = \sum W_j\ r_j^{max}$$

However, we do not see any sound practical scheme which would be robust to errors in estimation of W_j. Therefore we propose a different approach to computing the overall effectiveness of a network.

This approach is based on the observation that although there are a variety of services there is a clear-cut heirarchy of settlements with respect to the service facilities available in them. A considerable body of theoretical and empirical justification for this assumption of functional heirarchy exists in the classical central place theory and its modern variants (See, for example, Berry [11] , Lloyd and Dicken [12]). According to Walter Christaller, the originator of Central Place Theory, "a place on a particular level in the heirarchy provides not only goods and services which are specific to its level but also all other goods and services of lower usor" (quoted in Lloyd and Dicken [12]). This theory has also found substantial empirical support in India (Wanmali [13] , Mahendru et.al. [14]). Further, there is a clear relationship between population of a settlement (village or town) and its level in the heirarchy. The higher the population, the higher is the level in the heirarchy. (Berry [11] , Lloyd and Dicken [12]). Support for this assumption is provided by Table I below for Kheda.

Kheda 1981

Census Cat.	1	2	3	4	5	6	7	Total
No.	132	196	296	263	65	9	1	962
Primary School	108	196	296	263	65	9	1	938
Secondary School	2	3	40	136	58	8	1	248
Phone	4	28	98	175	59	9	1	374
Post Office	8	56	165	239	61	9	1	539
Regd. Medical Practitioner	3	6	50	116	47	6	1	229
Child Health Worker	62	135	242	233	61	6	1	740
Population	< 500	-< 1000	-< 2000	-< 5000	-< 10000	-< 20000	50000	

Further statistical support was obtained from regression of a weighted index of facilities on population. Weights were given according to the 'scar city principle' (Mahendru et. al., [14] , Desai [15]) which gives a weight to each facility which in the reciprocal of the fraction of villages having the facility. The correlation coefficient was 0.71. Desai [15] obtains comparable results in a similar exercise done for the nearby district of Mahesana where correlation coefficients of 0.70 and 0.81 were observed. We also used 'threshold weights' as formulated by Wanamali [13] and obtained a correlation coefficient of 0.75.

In view of the above, we have classified the villages and towns in Kheda into four classes based on population: (1) Below 1000, (2) 1001 to 5000, (3) 5000 to 20,000 and (4) above 20,000. The break-up was made to maintain reasonable uniformity in facilities available in a class while limiting the number of classes to a reasonable number. Now instead of considering individual facilities we can think of four bundles of facilities, each bundle being available in one of the four classes. Further, the bundle belonging to class j is included in the bundle belonging to class j + 1.

One majoe advantage in using population instead of actual facility locations is that one does not need to know the locations of relevant facilities likely in the next twenty years. Most location choices of the government are not made this far in advance. Where the facilities arise from private initiative it is even more difficult to predict their location. By using population as a surrogate, we are implicitly saying that in the long run, villages with similar populations in the same taluka will be endowed with similar facilities. This seems to us a very defendable assumption. Often, in fact government planning uses population as a basis for locational decisions - e.g. the Fifth Five Year Plan (1978-83) for India has as a target the building of roads in tribal areas to connect all villages with a population of 1500 and above and 50% of villages with a population of 1000 to 1500. Our assumptions, then, will lead us to propose rural road networks which are near optimal in the long run, as indeed, road investments need to be.

The criterion used to measure network effectiveness was a vector $\underline{R}$ consisting of six elements $(\bar{r}_2, \bar{r}_3, \bar{r}_4, r_2^{max}, r_3^{max}, r_4^{max})$ where subscript j (j = 2,3,4) represents the j^{th} bundle of facilities. The vector nature of the criterion does have the difficulty that two arbitrary networks may not be comparable. However, it avoids an artificial weighting approach and is adequate for our purposes since we propose to examine networks which are close in effectiveness to the existing network along all facility bundles. It also has the desirable properties mentioned earlier in this section. One minor modification was made to $\underline{R}$ as defined above. In the case where $d_i \leq 3$ km., a_i/d_i was set to one for two reasons. Firstly, if a facility is so close, it is practically in the village since the average village diameter is 3 to 4 km. in size and secondly, if a facility is reachable in a walk of about half an hour, it is possible to comfortably do the distance without roads.

3. INTERACTIVE DECISION-SUPPORT SYSTEM

A computer-based decision support system using a graphics terminal was developed to assist in the task of generating networks with road-length less than the existing networks while matching the existing network in effectiveness (as measured by the criterion R developed above). The system operated on a data-base of villages and roads to display a taluka map showing these on a C.R.T. The user could interactively add or delete links (straight line segments joining villages) to test the effectiveness of alternate networks. The system rapidly re-computes rural road length, shortest paths and R for the network selected. The advantage of this approach is that the judgement of the user can be incorporated in the solution. Thus special aspects of the situation not easily reflected in models can be treated.

Owing to the large number of villages in a taluka (around 100) obtaining a purely inter-active evaluation can be very cumbersome. To provide further computational support to the user, a heuristic was developed to provide a good starting solution around which the user can conduct exploratory searches. The heuristic was designed to generate a series of different networks depending on the choice of a parameter α. This parameter is an angle in the range 0 to $\pi/2$. For $\alpha = 0$, the heuristic comes up with a network which is high on road length but low on R, whereas with $\alpha = \pi/2$ the reverse is true. For values in between it generates networks falling between these extremes.

The heuristic consists of the following steps:

1. Set $j = 1$
2. Allocate each village in class j to its nearest village in class $j + 1$ (on a straight line distance basis)

3. For each group of villages in class j allocated to the same village in class (j+1) solve the 'Centre Connection Problem' (described below) with the class j+1 village as the centre. The solution gives roads to be constructed between class j and class j villages and between class j and class j+1 villages.

4. Increment j by one. If $j < 4$ repeat steps 2 & 3.

5. Calculate the rural road length and R. Display the solution for interactive response.

It is in step 3, the solution of the 'Centre Connection Problem', that the parameter α plays a role. We next discuss this sub-problem and the algorithm used to solve it.

4. THE CENTRE CONNECTION PROBLEM

Given points $P_1, P_2, \ldots, P_n$, in the Euclidean plane and a point P_0 called the centre, find the shortest length network consisting of straight line segments connecting pairs of those points such that all points are connected subject to the following constraints:

(a) For any path from a point P_i to the centre P_0 whose first segment is $P_i P_j$, the angle $P_j P_i P_0$ must be less than α.

(b) The shortest path from P_i to P_0 must be via villages nearer to P_0. (i.e. if the path is $P_i \ P_j \ P_k \ldots P_1 \ P_0$ then $P_i P_0 > P_j P_0 > P_k P_0 \ldots > P_1 P_0$.

The reason for constraint (a) is to provide parameterization. Constraint (b) is a reasonable requirement that rules out solutions with long rambling paths from villages to the centre.

Two remarks are in order :

1. An optimum solution must be a tree (otherwise by finding a tree of shortest paths from P_o to other points and deleting all line segments not in this tree, we can reduce the length of the network).

2. The network length is a monotone decreasing function of α. For $\alpha = 0$, we have a pure radial network. For $\alpha = \pi/2$ we have a tree that is near the minimum spanning tree in length (In fact often it is the minimum spanning tree, but it can be shown that this is not necessarily the case).

An $O(n^2)$ algorithm that solves the Centre Connection Problem consists of the following step :

For each i $(i = 1,2, \ldots, n)$ choose $P_i P_j$ which is the shortest line segment amongst those satisfying constraint (a) and with $P_o P_j < P_o P_i$.

Proof of the optimality of this algorithm is given in the Appendix.

In using this algorithm in problem, we have modified it to take cognisance of existing district, state and national roads in the following way. If there is a line segment satisfying constraint (a) connecting to an existing non-rural road linked to P_o and if it is shorter than any other segment considered, then this link is chosen.

5. MAGNITUDE OF LIKELY SAVINGS

Representative results from applying the heuristic to talukas are given in Table II below (The computing time, as might be expected from the simplicity of the heuristic is very small).

Table II - Savings in Rural Road Length

Taluka	Solution Type	R						Rural Road Length (Km)	% Saving*
		$\bar{r}_2$	$\bar{r}_3$	$\bar{r}_4$	r_2^{max}	r_3^{max}	r_4^{max}		
Balasinor	Existing(1981)	1.02	1.14	1.29	1.62	1.73	2.00	179	25%
	Heuristic	1.01	1.16	1.29	1.61	1.67	1.64	134	
Cambay	Existing(1981)	1.08	1.17	1.19	2.02	1.87	1.86	252	41%
	Heuristic	1.04	1.16	1.20	1.68	1.67	1.60	149	
Nadiad	Existing(1981)	1.00	1.07	1.21	1.00	2.09	2.20	165	27%
	Heuristic	1.00	1.07	1.21	1.00	2.09	2.13	121	
Anand	Existing(1981)	1.00	1.00	1.14	1.00	1.05	1.76	136	44%
	Heuristic	1.00	1.01	1.18	1.00	1.18	1.87	76	

The results convincingly show that large savings can accrue from using decision support models of the type described in this paper. Translated into cost terms the savings for these four talukas alone amounts to Rs. 25 million (US $ 2.5 million).

ACKNOWLEDGEMENTS

We would like to thank Mr. Dipankar Basu, I.A.S. and Prof. S.C. Bhatnagar for many stimulating discussions on this project. Work on the Centre Connection Problem is the outcome of discussions that the first author had with Dr. Heinz Speiss during his visit to the Centre de Recerche Sur les Transport, University of Montreal. Mr. B.H. Patel very ably handled both field data collection and computer programming. We are grateful to the Indian Road Congress for funding our work and to the Government of Gujarat - in particular the Department of Roads and Buildings - for their co-operation in data collection.

APPENDIX

We prove the optimality of the algorithm to solve the Centre Connection Problem via Theorems 1 and 2 below:

Let G = (N,A) be a directed graph with each node i in N representing village P_i with an arc of length $P_i P_j$ in A directed from node i to node j if and only if $P_o P_i > P_o P_j$ and if $P_o P_i P_j \leq \alpha$ $(0 \leq \alpha \leq \pi/2)$. Then it is clear that solving the Centre Connection problem is equivalent to finding a minimum spanning directed tree rooted to node 0. (See Lawler [16] for definitions).

Theorem 1: Graph G is acyclic (has no directed cycles)

Proof: If G contains a directed cycle there is a node i such that there is a directed path i, j_1, j_2, ..., j_k, i. But this implies $P_i P_o > P_{j_1} P_o > P_{j_2} P_o \ldots > P_i P_o$ which is impossible.

Theorem 2: If we choose a minimum length out-ward arc at each node i (except i = 0), the resulting sub-graph is a minimum spanning directed tree rooted to node 0.

Proof: Since the sub-graph has (n-1) arcs, and each node (except node o has out-degree = 1, if it is not a directed tree rooted to 0 it must have at least one cycle. But it cannot have a directed cycle since the graph is acyclic. Also it cannot have an undirected cycle, since this would imply that at least one node on the cycle has out degree > 1. So the sub-graph is a directed tree rooted to node 0.

Let the length of the arc chosen from node i be l_i. Consider any other directed tree rooted to node 0 with arc length l_i' from node i. Then since $l_i' \geq l_i$, it follows that the sub-graph is of minimal length.

REFERENCES

1. Planning Commission, Report of the National Transport Policy Committee (Government of India, New Delhi, 1980).

2. Rajyagor, S.B. (ed.) Kheda District Gazetteer (Government of Gujarat, Ahmedabad, 1977).

3. Hine, J.L., The appraisal of rural 'feeder' roads in developing Countries, Proceedings of P.T.R.C. Annual Meeting, University of Warwick, England, (July, 1976).

4. Massam, B., Location and Space in Social Administration (Edward Arnold, London, 1975).

5. Owen, W., Distance and Development - Transport and Communications in India (The Brookings Institution, Washington, 1968).

6. Patel, N.R., Locating Rural Social Service Centres in India, Management Science, 25 (1979), pp. 22-30.

7. Garrison, W.L. and Marble, D.F., Graph Theoretic Concepts, in: Hurst, M.E.E. (ed.), Transportation Geography - Comments and Readings (McGraw-Hill, New York, 1974).

8. Hay, A., Transport for the Space Economy (Macmillan, London, 1973).

9. Sharma, A.N., Road Network Planning for Rural Development, in: Jaiswal, N.K. (ed.), Scientific Management of Transport Systems (North-Holland, Amsterdam, 1981).

10. Shimbel, A., Structural Parameters of Communication Networks, Bulletin of Mathematical Biophysics, 15 (1953) pp. 501-507.

11. Berry, B.J.L., Geography of Market centres and retail distribution (Prentice-Hall, Englewood Cliffs, 1967).

12. Lloyd, P.E. and Dicken, P., Location in Space (Harper & Row, New York; 1977).

13. Wanmali, S., Service Centres in Rural India (B.R. Publishing Corpn., New Delhi, 1983).

14. Mahendru, A.K., Sikdar, P.K. and Khanna, S.K., Nodal Points in Rural Road Network Planning, Indian Highways (April 1982) pp. 5-10.

15. Desai, A.P., A Study of Settlement Pattern for Six Talukas in the Narmada Command Area of Mahesana District of Gujarat, Research Report to the Narmada Planning Group, Government of Gujarat (December 1982).

16. Lawler, E., Combinatorial Optimization - Networks and Matroids (Holt, Rinehart and Winston, New York, 1976).