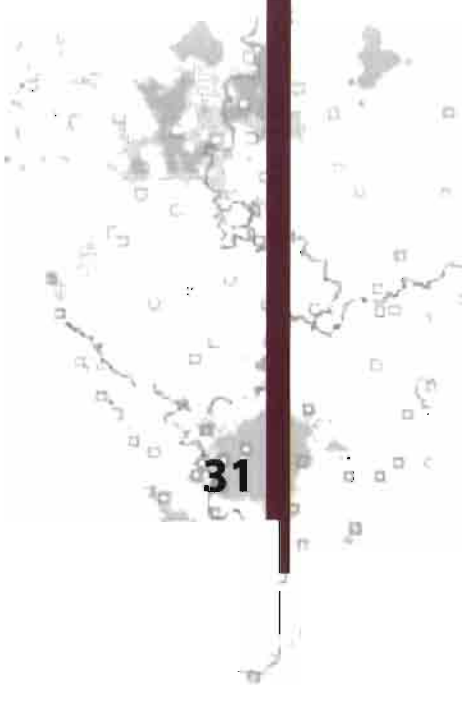
Institut
Français
de Pondichéry

Pondy Papers in Social Sciences

Mapping out Social Change in South India

A Geographic information system and its applications



31

C.Z. Guilmoto, S. Oliveau,
V. Chasles, R. Delage & S. Vella

MAPPING OUT SOCIAL CHANGE IN SOUTH INDIA

The *Institut français de Pondichéry* (IFP, French Institute of Pondicherry) is a multidisciplinary research institute, whose scope embraces Indian civilization, History and Society, Environment and Development in South and South-East Asia, through the Departments of Indology, Ecology and Social Sciences. Created in 1955, the IFP is part of a network of research centres funded and supervised by the Division of Social Sciences, Human Sciences and Archaeology of the French Ministry of Foreign Affairs. Since 1988 the Department of Social Sciences has been carrying out joint Indo-French research programmes in anthropology, history, demography, geography, sociology, by focusing on cultural and social changes.

As the result of these programmes, the Pondy Papers in Social Sciences (PPSS) are a series of working papers to which IFP researchers and associated persons contribute (refer to the list at the end of the issue).

French Institute of Pondicherry, 11, St. Louis Street, P.B. 33, Pondicherry 605001-India, Tel: (413) 2334170/334168, Email: ifpdir@ifpindia.org
Website: www.ifpindia.org

In Delhi, the *Centre de Sciences Humaines* at Delhi (CSH, Centre for Social Sciences and Humanities), is also part of the same network of research centres, coming under the French Ministry of Foreign Affairs. Since 1989, the Centre's research work is primarily oriented towards the study of issues concerning the contemporary dynamics, relative to development in India and South Asia. The activities of the Centre are focused on four main themes, namely: Economic growth and sustainable development, international and regional relations, Institutional structures and political constructions of identity, and Urban dynamics.

Centre de Sciences Humaines, 2, Aurangzeb Road, New Delhi, 110 011, India.
Tel: (91) 11 2301 62 59/ 41 73, E-mail: public@csh-delhi.com
Website: www.csh-delhi.com

MAPPING OUT SOCIAL CHANGE IN SOUTH INDIA
A geographic information system and its applications

Christophe Z. Guilmoto, Sébastien Oliveau,
Virginie Chasles, Rémy Delage and Stéphanie Vella

31

Pondy Papers in Social Sciences

FRENCH INSTITUTE OF PONDICHERRY

ISSN: 0972-3188

Department of Social Sciences
11, Saint Louis Street, PO Box 33
Pondicherry, 605 001, India
Ph.: (91) 413 2334170/168, Fax: (91) 413 2339534
Website: www.ifpindia.org

Composition - Cellule d'édition, IFP
Impression - All India Press, Pondicherry
Front Cover - Cover photography: Laurent Dufy

CONTENTS

LIST OF TABLES AND FIGURES

ACKNOWLEDGEMENTS

The first half of this paper describes the constitution of the large-scale spatial database of the *South India Fertility Project* (SIFP), coordinated by C.Z. Guilmoto. The applications presented in the second half are derived from complementary projects accommodated in the "Population and Space in South India" programme of the Department of Social Sciences of the French Institute of Pondicherry.

These projects have benefited from various sources of support, mainly the French Institute of Pondicherry (FIP), the French Research Institute for Development (IRD) and the Wellcome Trust (London), which financed the SIFP and made possible the whole enterprise. Additional support came from the United Nations Population (New Delhi) for the SIPIS and from the CNRS for the EMIS team, as well as from the Franco-Indian exchange programme (French Ministry of Foreign Affairs and the Indian Council for Cultural Relations) and from grants provided by Aires Culturelles (French Ministry of Education) for the sojourn in India of doctoral students in the team.

We particularly thank S. Vingadassamy and R. Amuda for their crucial support in Pondicherry during the constitution of the database. We have also benefited from the assistance of Bertand Buffière, Véronique Joseph, Karine Keraudy, Sangee Patel and Joe Vimal Raj, as well as from the entire team of the Department of Social Sciences. Finally, our thanks are extended to Frédéric Landy for his attentive reading, enabling us to improve a preliminary version of this text.

INTRODUCTION

If we call to mind in the preliminary remarks to this article the fact that the *modus operandi* of research has to do both with the imperatives of scientific inquiry and with the relative fortuity of available information, it is because while investigating material for a new study concerning the decline in the birth rate in India, the first author discovered virtually by chance the existence of a previously unpublished database which, as it contained information that could be employed in research on fertility in India, was to serve as a source for the projects described here. In fact, several years subsequent to the undertaking in 1991, the administration of the Census of India decided to make a large part of the results available to the public in a digital format. Among the results were files containing information relating to villages and towns that were later chosen for publication in the *Census Handbooks* by district[1]. This new form of diffusion of results of the Census of India itself represents a small revolution in a country where statistical information is undoubtedly as rare as it is difficult to access and, for this reason, is very dear from all points of view[2].

Taking this potential statistical windfall into consideration, the constitution of a large regional data bank was foreseen, to be built up by collecting primary data so as to avail of a base including all the villages and towns of the chosen states, namely, Andhra Pradesh, Karnataka, Kerala and Tamil Nadu, to which the districts of the Union Territory of Pondicherry were added. A few days later, an equally fortuitous discussion with a few colleagues intrepidly setting about an Indian geographic information system (GIS) led to a broadening of the approach to the spatial dimension: if it were possible to georeference these census units, one would avail of a formidable tool for examining the results of the last Indian census. In this way, beyond the study of the spread of the fertility decline in South India, the data and the

1 The *District Census Handbooks* are published on the occasion of each decennial census in English and Hindi for each district in India (466 in 1991). Their publication occurs very belatedly, often just prior to the following census, and the 1991 census did not depart from this rule, not withstanding the new information contained.

2 An illustration of these remarks is the recent introduction of an Indian web site very rich in statistics of all types (www.indiastat.com). Users must, however, pay relatively dearly for access to data that are for the most part taken from government statistics and are thus supposed to be in the public domain (such as statistics from the Census of India).

techniques henceforth available would make it possible to conceive of an exhaustive approach to the socio-economic situation in South India, from both a statistical and a spatial point of view.

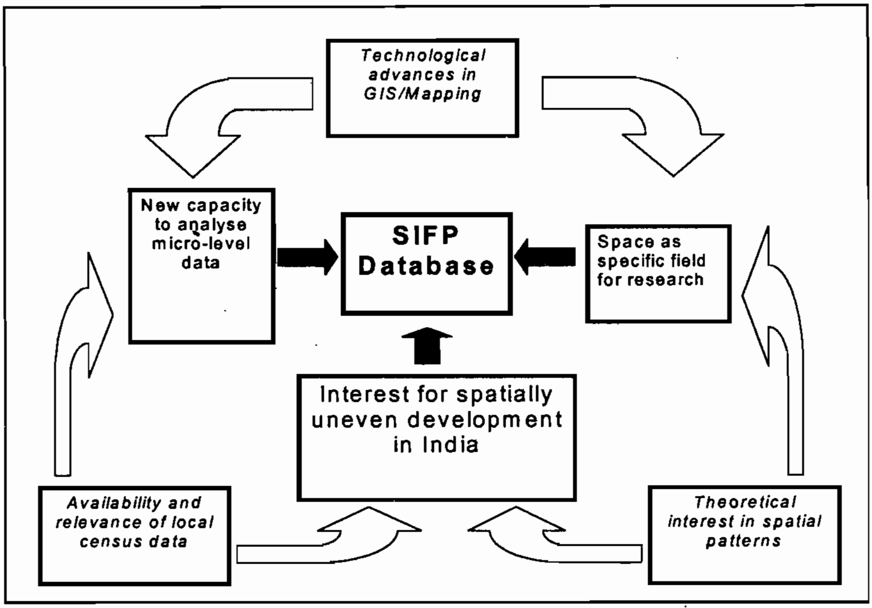


Figure 1: Technological, theoretical and scientific environment of the SIFP project

This initiative did not take place in an adventitious scientific environment. Rather, it is understandable in terms of a progressive development of ideas, techniques and needs that developed in the course of the 1990s. Figure 1 recapitulates this situation in the form of a chart summing up the environment of this period. Three main phenomena come together: the technological advance of geographic information systems in the last ten years[3], the increasingly wide circulation of micro-level data and renewed interest in spatial questions. The first two phenomena have led to the confluence of new processing tools and new data: the data from the Indian census of 1991, henceforth available in computer format, could also be used on the micro-level for cartographic or statistical purposes. This

[3] For general presentations of the application of GIS in the spatial analysis of social data, see in particular Martin (1996) and Longley and Batty (1996).

represented an innovation in India, where cartographic productions are very limited in number and almost never venture below the level of the district.

The GISs have gradually become more widely available, having become less costly and above all more easily accessible for non-specialists in India. The concurrence of the development of GISs and the renewed interest in spatial phenomena (illustrated, for example, by the new economic geography) has led to the consideration of questions that were in the past more or less theoretical and inaccessible for purely technical reasons. GISs, by making cartography and spatial analysis more easily accessible, have made it possible for long-standing or new questions to find application support. In countries such as India, this provides in particular the opportunity to closely analyze inequalities in economic and agricultural development, and even to disaggregate large-scale datasets, which observers can use in order to discover the exact geographic contours of social change. The rapidity and the vigour of contemporary transformations contribute, in fact, to the continuous redrawing of the map of India.

The SIFP thus represented a logical outcome, around of research themes, the existence of previously unused data and the advances in the technology of spatial processing (computer-assisted cartography, GIS, geostatistics, etc.) that is still only in its inception. The impulse given by the theoretical reflection on questions of fertility, notably the debate between the structural factors in fertility decline and the mechanisms of diffusion that occur together with it (see Casterline, 2001), was certainly central, but the extension to other applications derived from economic or geographic discussions has proved to be just as logical, enriched by individual initiatives that are at times removed from the initial demographic issue, as the following illustrations will show. The French Institute of Pondicherry, with a long tradition in the cartography of vegetation in South India, was the ideal site for such a project and all the participants owe much to its support, whether direct, in the form of financial support to researchers and students, or indirect, by providing to all a place of encounter and development.

This volume retraces the main stages of the SIFP project and its diverse avatars, illustrating our topics with numerous maps and documents. Rather than restricting ourselves to the concrete results of the analyses, we have chosen here, in order to make the material more instructive, to present the successive technical phases in greater detail. They bear witness to the difficulties already forgotten by those who design models in the affluent countries, with a wealth of old, and often systematic, spatialized databases, but these very difficulties await all projects concerning the developing countries. While it is true that the statistics may often be lacking or seem to remain under-utilized for manifold technical reasons, the geographic resources are even more deficient and only an infinitesimal part of them is today available in digital format, and consequently utilizable by the GISs.

In absolute terms, it would be easy to set ourselves ambitious goals as regards the geographic coverage, but the tedious work of statistical and geographic data acquisition, which is indispensable for the realization of such projects, sometimes tends to make them truly impracticable: a number of GIS projects which should facilitate the publication of an atlas or serve as the base of a web site ultimately never see the light of day and the money swallowed up in these projects results in nothing more than a few maps which could have been drawn by hand and with less cost, without mobilizing the GIS technology. However, in the era of the study of social transformations, which are sometimes more rapid and more dramatic in the developing countries than elsewhere, the geographic approach often stands out as a key to the reading of these mechanisms and it is urgent that the technological effort be channelled in this direction by setting reasonable objectives. We hope in this document to illustrate the cumulative and diversified advantage that this enterprise can provide when it attains its goal, in this case the establishment of a geo-demographic database of the whole of South India based on the disaggregated census data.

1

A GEOGRAPHIC INFORMATION SYSTEM FOR SOUTH INDIA

Georeferencing South India[4]

From the census, which provides village-level data every ten years, to the large surveys such as the National Family and Health Surveys (dating from 1992 and 1998) or those of the National Council for Applied Economic Research, the statistical data in India cover almost all demographic, social and economic aspects of the society. Notwithstanding this apparent profusion, precise geographic utilization and cartography on a small scale are still rare, or indeed nonexistent.

Thus, one often finds studies made at the district level, but applications on a lower administrative level are more seldom and more limited in space. Specifically cartographic studies on the village level, apart from monographic approaches, are absent from scientific production.

This absence can be explained by the lack of detailed and accessible cartographic coverage. Cartographic sources can be classified according to two major types. On the one hand, there are the sources that provide general coverage of the territory on different scales. These sources originate from public bodies or private publishers. On the other hand, there is the production of town maps stemming from governmental planning bodies or map merchants. Although this second category of maps does not interest us here, the recent high-quality publications by the Eicher Company of maps based on satellite images for the cities of Delhi, Chennai, Bangalore and Mumbai should be mentioned, as they foreshadow a new era of cartography in India.

The production of maps covering the entire territory of India is the responsibility of the Survey of India. The latter, created by the British, is a

[4] Section written by Christophe Z Guilmoto and Sébastien Oliveau. An abridged version was previously published in French (Guilmoto, Oliveau and Vingadassamy, 2002).

derivation of the military administration. The map was a strategic object and its public diffusion was prohibited until the beginning of the nineteenth century[5]. Still today, the production of maps is a matter of the state, and a series of places held to be sensitive continues to be excluded from cartography, among which are hydroelectric projects, dams, steel plants, installations under All India Radio, telecommunication installations and water purification/supply installations, etc. Moreover, for a coastal strip of more than 50 km inland, the obtaining of maps classified as "restricted" is subject to special requests to the administration of the Survey of India[6]. To give an example, a map from 1973 on a scale of 1: 50,000 concerning the area surrounding the town of Dindigul (200,000 inhabitants, Tamil Nadu, located 140 km from the Coromandel Coast) is classified as "restricted" because of its proximity to the coast. For certain littoral states, such as Kerala, topographical maps are not officially available and nearly half of the contemporary maps comprising the South Indian peninsula are therefore "restricted". Voices are today being raised in the Indian milieu of professional cartography to denounce the restrictive policies pertaining to 227 of the 385 "degree Toposheets". Thus the journal *GIS@development* published in Delhi is a forum of professional users that regularly evokes the shortcomings of a policy that can appear outdated. Scientists and professionals are also unanimous in demanding profound changes in the policy of access to topographical information. We will mention in this regard a large meeting held in July 1999 at the Indian Academy of Science in Bangalore devoted to "public access to Indian geographical data". A large number of papers presenting important material for the discussion was then published in the journal *Current Science* (25 July 1999) and subsequently also taken up in the national press.

In the meantime, the production of the Survey of India remains the only official source in India from which the rest of the maps are derived, for example, the atlases of the NATMO (National Atlas & Thematic Mapping Organisation). The maps presented by the Indian census, such as those of the

---

[5] For a general history of cartography in India, one can refer to Madan (1997). For the setting up of the Survey of India, the work by Edney (1997) can serve as reference. Regarding the context surrounding the creation of the Survey of India and the first prohibitions of the diffusion of maps, the recent article by Raj (2003) is of interest.

[6] The request form of the Survey of India – Form 0.57(a) – for "restricted" maps contains notably the following clause: "Clearance of the Ministry of Defence is essential for the issue of Restricted maps to private individuals, organizations and commercial firms". It also specifies that the loss of the map must be notified to the Ministry of Defence, etc.

taluks contained in the *District Census Handbooks*, are also based on the maps of the Survey of India: their distribution is quite wide because of the hundreds of volumes of the census that are published, but the geographic quality of these maps remains mediocre, for they contain neither absolute location[7] nor scale. What is more, much information of a topographic nature is also missing from these maps. On the other hand, they show the individual location of all the towns and villages and, for this reason, are unique in India. In fact, the toposheets, when they are available, are far from containing all the localities included in the census tables and do not always enable the identification of inhabited places, as our work showed us.

The public thus avails on the Indian scale of two official sources: the first is that of the toposheets that present all the necessary details for geography, but suffer from an incomplete coverage and are sometimes not updated or unavailable. The second source is that of the villages represented in the census maps: this covers all the administrative units, but does not make it possible to georeference them so as to be introduced in a GIS. In practice, to exhaustively collect the census maps or the toposheets is often a challenge, the one or the other being very difficult to procure, as a visit to a regional headquarters of the Census of India or of the Survey of India readily shows.

One understands better why the private publishers in their turn have been able to diffuse their products, from specialized web sites to pavement stalls in the large cities, notwithstanding the often very mediocre quality of their maps and the perfunctory character of the information they contain (roads and principal localities)[8]. It is in this context of semi-penury, in which cartographic information exists but is dispersed and of terribly inconsistent quality, that new suppliers have arrived on the market. A web site such as www.mapsofindia.com is an interesting example, for it offers via Internet

---

7 Here one opposes the absolute position in terms of longitude and latitude, which is at the basis of the work with geographic information systems, to that of relative position, which situates the map in a space of vicinity by indicating the boundaries of borderland units.

8 We can quote in this regard several evaluations by the Survey of India concerning the quality of commercial maps: "Some publishers have no idea of map projection and its significance. [...] Spelling of names is incorrect. [...] Colour schemes are crude and unsuitable. [...] Over-generalization has caused the loss of the character of detail. [...] The boundary symbol is drawn so thick that small kinks and bends are lost. [...] Symbols differ in different parts of the same map. Printing is poor. Quality of paper is not good. [...] Some maps do not indicate the scales to which they are drawn". (excerpted from http://www.surveyofindia.gov.in)

district-scale maps of an unfortunately rather questionable quality from a geographic point of view. A general coverage of the country by maps that are less precise than those of the census, but of better quality than the maps offered by other private publishers, facilitated a considerable development of their diffusion, and this enterprise today provides its products in the form of CD-ROMs or offers them for sale on the web (after a long period during which they were gratis). In technical terms, the quality of these maps remains, however, quite uneven.

The demand for good quality cartography is becoming increasingly important in India, all the more so as at the same time a significant tourist activity is developing, which is known to make extensive use of maps. Moreover, Lonely Planet, the well-known tourist guide company, decided to update its *Road Atlas* in 2001 (first edition 1995). The work, which exceeds 100 pages, constitutes the best cartographic source at 1:250,000 currently available on the subcontinent. It combines information on relief and hydrography with the main towns and villages, as well as more general information (cultural sites, beaches, wildlife sanctuaries, airports, etc.).

Coming back to our project, the establishment of a geographic database in 1998 was not exactly an easy matter. As we have mentioned, there were hardly any sources making it possible at the same time to locate all the villages and to georeference them, apart from the fragmented information provided in the toposheets. It was therefore necessary to envisage a strategy adapted to the different available sources. An initial solution thus consisted in resorting to the toposheets, including British maps dating from colonization when they proved to be of easier access than the modern toposheets with a restricted diffusion. It may seem ironic to use old maps for a contemporary cartography, but this paradox illustrates the still outdated character of the production and diffusion of Indian cartography. Hence, a certain number of maps used data from the beginning of the twentieth century: Figure 2, for example, represents a topographic map dating from 1914-15, corresponding to the region of Vellore in the state of Tamil Nadu.

The second solution resides in the utilization of census maps, available at the level of each taluk or tashil, as we mentioned earlier. These maps offer the advantage of providing an exhaustive location of all primary units of the census, namely the revenue villages, administrative units that can comprise several distinct hamlets. They remain, nevertheless, rough (Figure 3) and in particular do not allow a direct georeferencing in the absence of geographic coordinates; for this reason, a meticulous work is required to piece together information. It will be noted that the maps are often published belatedly after the census, the last publications for the 1991 census dating from August 2000, which does not facilitate the task.

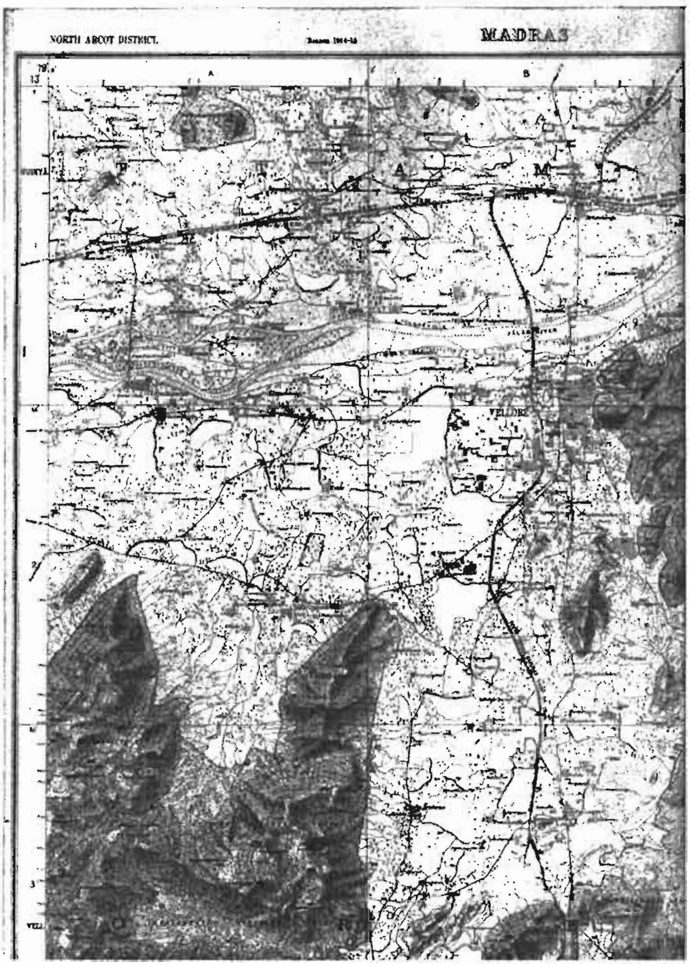


Figure 2: Topographical map of the Vellore region
(Tamil Nadu, India), 1914-15

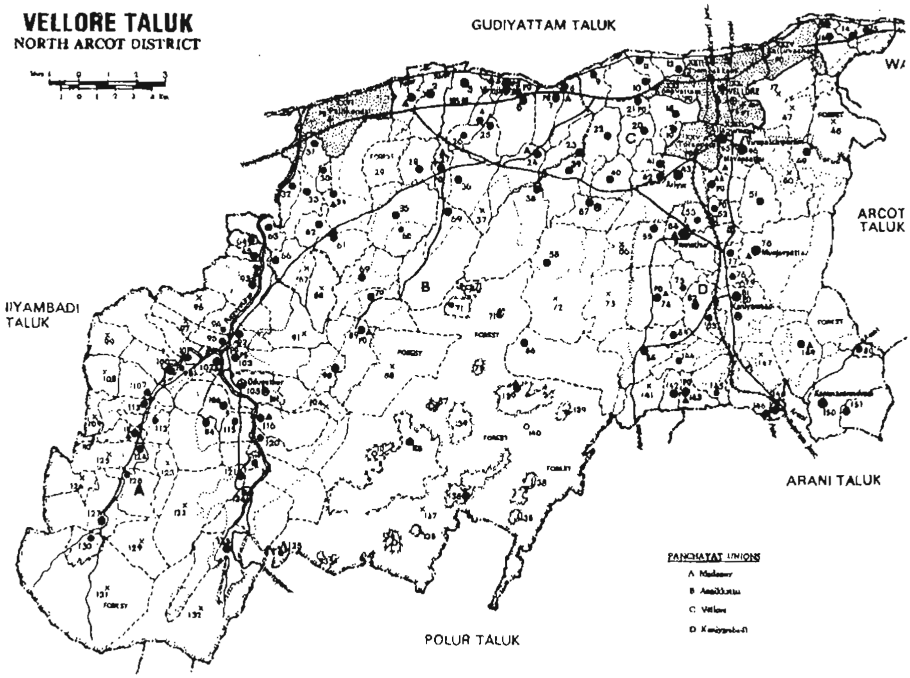


Figure 3: Census villages in Vellore taluk (Tamil Nadu, India) from the 1981 Census map

In view of the scope of the work, the establishment of a georeferenced database proved to be a long-term task, which is summarized in Figure 4. The apparent simplicity of the diagram should not mislead as to the difficulty of the operations, the major stages of which will be presented.

The first stage consisted in capturing, with the help of 712 taluk maps provided by the census, the location of villages and towns. The operation of assembling and keying the digitized maps was undertaken with AutoCad software (computer-aided design software). The absence of maps or of volumes for 1991 often made it necessary for us to utilize the data from 1981, which had to be updated later by the addition of villages newly included in the subsequent census. The 76,366 administrative units (75,523 villages and 843 towns) were entered with the same software. One of the major decisions when drawing up the database concerned the mode of representation of the villages. The primary maps of the census provided the detailed demarcation of the boundaries of each village, corresponding to what one terms in GIS jargon a "polygon". However, it was decided to simplify the data entry and to digitize each village in the simple form of a point, while the boundaries of urban agglomerations would all be captured in polygons. There were several reasons for this choice.

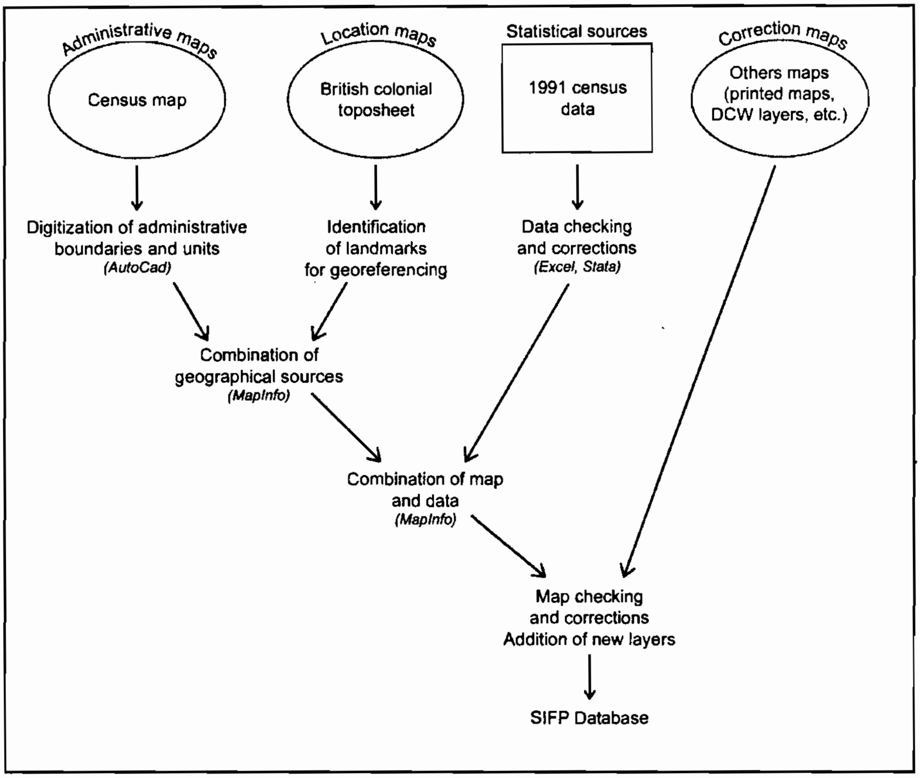


Figure 4: From many sources to the SIFP database

First, the boundaries of villages provided by the census did not appear to be reliable and were sometimes so badly represented on the primary maps that they were not digitizable. Comparisons between the areas of the villages given by the census and measured by the GIS showed, moreover, the unequal quality of this information. Then, the future evolution of the database was judged to be easier if the village units were represented by points than if they were to be represented by areas. In fact, the appearance or disappearance of a new administrative entity is expressed by the appearance or disappearance of a point. Thus, the overall cartography hardly changes. On the other hand, in a polygonal representation, the least punctual change necessitates the modification of all adjacent units. Furthermore, on the scale of the whole of South India, the cartographic representation of village boundaries would not be possible and it would require such a level of generalization (i.e. simplification) that automatic tessellation (i.e. partition into a pattern of interlocking shapes) of point patterns would provide comparable results. And, to conclude in a prosaic manner, the initial capture by polygons rather than points for each village would have required an effort ten to twenty times greater than the capture by points and, on the scale of more than 50,000 villages, this difference is not without consequence in terms of time and budget.

At the same time, the operation of locating control or registration points, which constitutes the second stage, was done manually on the old British toposheets to enable future georeferencing of maps captured with AutoCad. Maps without geographic reference points are, in fact, unusable in a modern GIS. The difficulty thus resided in the identification of registration points that would be present on the different maps at our disposal and, when our sources dated from the colonial period, sometimes would not have changed between the beginning and the end of the century! These points ("points de calage" in French) make it possible to register a map derived from the census maps and therefore initially devoid of geographic coordinates, as shown in Figure 5. They mostly consisted in bridges, intersections of roads and railway lines or other noteworthy and reputedly fixed points. Once the entire set of census villages and towns were digitized and georeferenced for each taluk, it was necessary to assemble the maps of the taluks beside each other in order to reconstitute a coherent overall map without overlap or gap between the taluks. This required global geometric transformations (using a technique know as "rubber sheeting") for an even fit.

The first test was initially conducted on the district scale (that of Vellore appearing on the preceding maps) so as to validate the capture procedure and the quality of statistical information connected with the census of 1991 (Oliveau, 1998). The result having been positive, the data entry of all the villages began. After two years of preparation, the 76,366 units captured with AutoCad were gathered in a GIS integrating the whole. The georeferencing was done with MapInfo, the software that appeared to best perform this meticulous work without needless efforts.

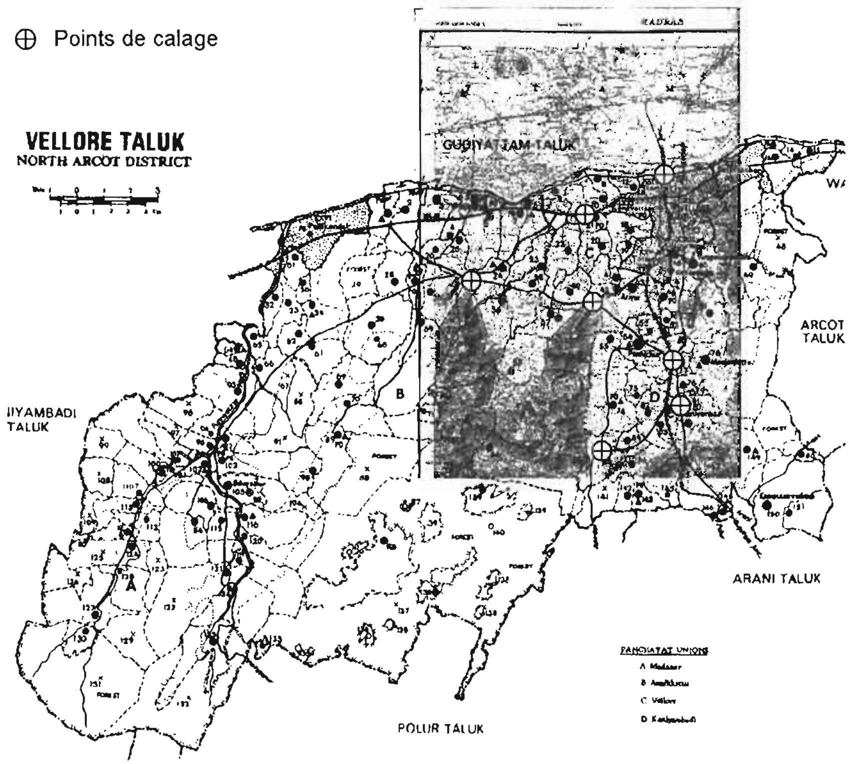


Figure 5: Registration points to join the topographical and Census maps (Figures 2 and 3)

While these capturing operations were going on, work on statistical preparation and verification of census data took place. The data, provided by the Census of India in the form of more than 400 files in Dbase format, were assembled in Stata software. This control work[9] brought to light numerous statistical problems, which had to be corrected: missing data, inconsistencies, repetitions, etc. To this purpose, we used upon the various volumes of the *District Census Handbooks*, including volumes from 1981 when those from 1991 were not yet available. When the data correction had been completed, it was necessary to connect the census statistics with the spatial units in the GIS. Linking up the villages from the census tables with the GIS locations gave rise to further difficulties. Thus, numerous villages

[9] This work was in particular carried out by Bernard Buffière and Véronique Joseph (M. Phil. students in demography) during a period of study in Pondicherry in 1998.

were absent from the maps and this was particularly the case in tribal zones inadequately covered by the administration. What is more, the English transcription of the local names (four principal languages and alphabets co-exist in South India) created numerous ambiguous toponyms. Good knowledge of the field by the technicians and verifications with the regional census offices mitigated part of these problems. In general, 95% of the villages were processed without difficulty on the basis of available information and could be unambiguously located. Among the remaining cases, 95% were in their turn corrected by a more attentive examination of the sources and a specific correction: written form of the name, geographic position, data, etc. However, for the hundred or so villages remaining (5% of 5% of 75,000 villages), the marginal cost of the processing can become truly prohibitive: it would have been virtually necessary to go there with a GPS (global positioning system) and a team of researchers, a proposition considerably beyond the possibilities of our project. Depending to the given case, it was thus necessary to aggregate them with the neighbouring villages, locate them in an imprecise manner, or indeed eliminate them from the base when the data were too defective.

An example of the GIS obtained is seen in Figure 6, which represents the border between the states of Kerala and Karnataka. The unequal geographic distribution of villages will be noticed, very dense in the Mysore region of Karnataka and much more spaced out in Kerala, in which the population is on the contrary characterized by very populated "villages" and a very open settlement. With the help of other cartographic sources (most notably data from the *Digital Chart of the World*, also known as DCW[10]), we were able to verify the quality of the georeferencing and integrate new information. Additional layers of information were progressively incorporated.

The final database, unique in India by this granularity, thus offered an incomparable mine of information covering the four states of South India: Andhra Pradesh, Karnataka, Kerala and Tamil Nadu, as well as the Union Territory of Pondicherry. A few figures summarize the extent of the system: the area covered is of the order of 636,000 km^2, for a population of 223,384,786 persons in 2001, distributed over 75,523 villages and 843 urban

[10] The Digital Chart of the World is a geographic database on a world scale. It was developed by the ESRI enterprise for the cartography agency of the American defence. The primary source is the ONC (Operational Navigation Chart).

units (in 1991). With each administrative unit, 130 attributes are connected concerning socio-economic data (literacy, professional activities, etc.) and some forty attributes concerning village infrastructures and land utilization. In addition to these layers on village and urban scales, the data at the level of taluks and districts were aggregated for wider studies. By the same token, a layer pertaining to road and railway networks was added by crossing available information in the census and in diverse private maps. The hydrographical coverage (notably with the help of the DCW) and data pertaining to altitude were also incorporated.

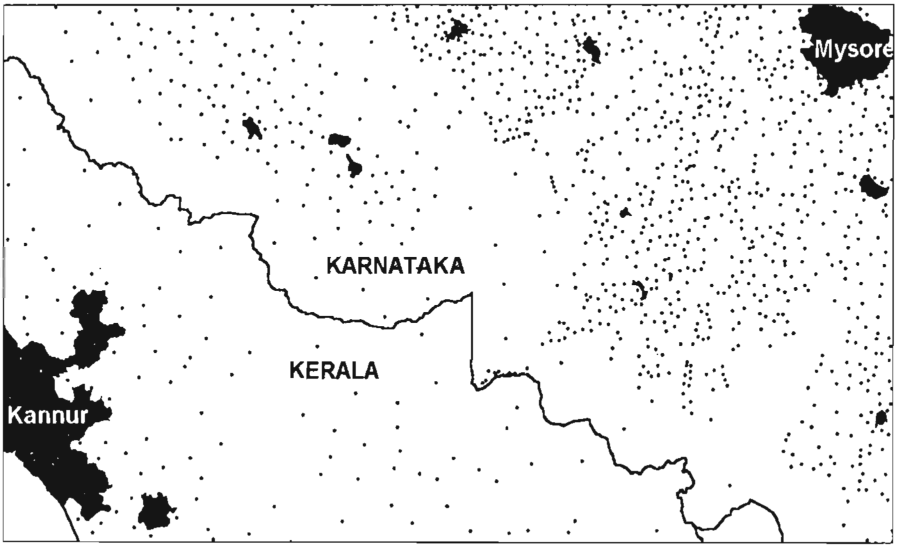


Figure 6: 1991 Census localities (villages) and urban areas along the Karnataka-Kerala border

An important point is worth noting. While the GIS thus created is already relatively large, it was conceived so as to be able to be updated and completed later. A first updating began with the publication of data from the 2001 census. The first results concerning the districts and towns have appeared and have been integrated. But the georeferencing also makes it possible to integrate other sources such as maps in vector format (for example, the DCW), raster or satellite images. However, in conclusion, we should mention several of the limitations of the exercise, which are of three types.

As the information could not be produced in real-time and the sources were all the same limited, the data and the raw census maps were already seven years old at the start of the project. They were over ten years old at the end of the project, and that constitutes a first limitation of their utilization. Local situations have of course evolved, and this is all the more true as infrastructures and demography are particularly dynamic dimensions in South India. Thus, for example, the coastal road that, since 1999, links Chennai and Pondicherry and numerous access roads today bypassing agglomerations are absent from the originally utilized maps.

As the census is the only exhaustive source at the village level, the second limitation resides in the unique source of data. Possible verifications and corrections are limited. We have carried out a series of statistical and logical tests on the data as well as cross verifications between digitized data and printed publications, when they were available. The final statistical base of the *South India Fertility Project*, while it was not perfect, is somewhat better than the census publications. On the other hand, there exists no other statistical data at the village level. It is therefore not possible today to enrich our base with other information, if not by the incorporation of specific data gathered in the field (and consequently geographically very limited).

The third limitation is of a cartographic order. Since the sources are diverse and sometimes of middling quality, the final error in precision, based on verifications made in the field with a GPS, is less than 500 metres for a territory extending over some 1400 kilometres from north to south. It is often less than 250 metres, but unforeseeable occurrences are probable, such as in badly covered forest regions. The general precision of the database remains satisfactory and allows a spatial processing that we perform (cf. *infra*). Moreover, it should be pointed out that computerized cartographic sources pertaining to the Indian population are nearly nonexistent. A source such as the DCW is less precise (its definition is of 1:1,000,000), older (the surveys for India date from the 1970s) and contains more errors: location by means of satellite images does not "see" the infrastructures existing beneath vegetation (roads in forests or bordered by trees), just as it is not able to differentiate abandoned infrastructures from those still in use[11].

[11] As concerns the quality of the DCW and the limitations to its utilization, see Kraak and Ormeling (1996: 207).

Sharing the data: CD-ROM and Internet

Once the database had been built, its use began in a systematic manner in the framework of the scientific projects described below. At the same time, the team soon received frequent requests for geographic and statistical information of all types. Requests came from other researchers, but demand was greater in the Indian society (NGOs, administrations, private enterprises). In addition, as all the data utilized came from the public domain, it seemed fitting to return this information to the citizens. It thus appeared to be necessary to make our research material available to the greatest number, but in a form that could be utilized by everyone. To do this, it was first necessary to envisage a mode of diffusion and tools for exploring the data that could satisfy such varied users.

The reconstitution of this data encountered diverse problems. Thus, the enormous mass of data precluded that it be made available on the Internet in its entirety, unless the consultations were limited to very reduced samples or to very well equipped users. By the same token, making the raw data available did not seem to be satisfactory owing to the fact that the public having the means to process it was restricted, first, because GIS software is costly, and then because it is complicated.

To mitigate the problem of the volume of data to be transmitted, a CD-ROM was chosen as the medium. This makes it possible to store the complete database and to add to it a few supplementary maps prepared for the occasion. A mini-GIS capable of mapping and exploring the data was included on the CD-ROM.

This CD-ROM project, supported by funding from the United Nations Population Fund (UNFPA), was entitled *South India Population Information System* (SIPIS). The state of Tamil Nadu was chosen for this project, as the four states could not be treated of concurrently because the processing capacities of personal computers at that time were not sufficient (and just about are today). Begun in January 2000, the finalization of the CD-ROM and the software lasted less than one year because the database had nearly been completed.

In order that the greatest number of users could employ this tool, efforts were above all made to develop a user-friendly tool enabling one to move easily from the data to the maps. From the point of view of geographic information, this project necessitated the production of new data calculated on the basis of the raw data of the census so that the end-user need only press a button to obtain structured information, such as a literacy index or a map of administrative boundaries. All the statistical information was brought together in large families (social, economic, employment, infrastructure data, etc.). Likewise, a few systematic maps are provided in raster format (literacy, fertility, etc.) to assist the user in identifying regional tendencies.

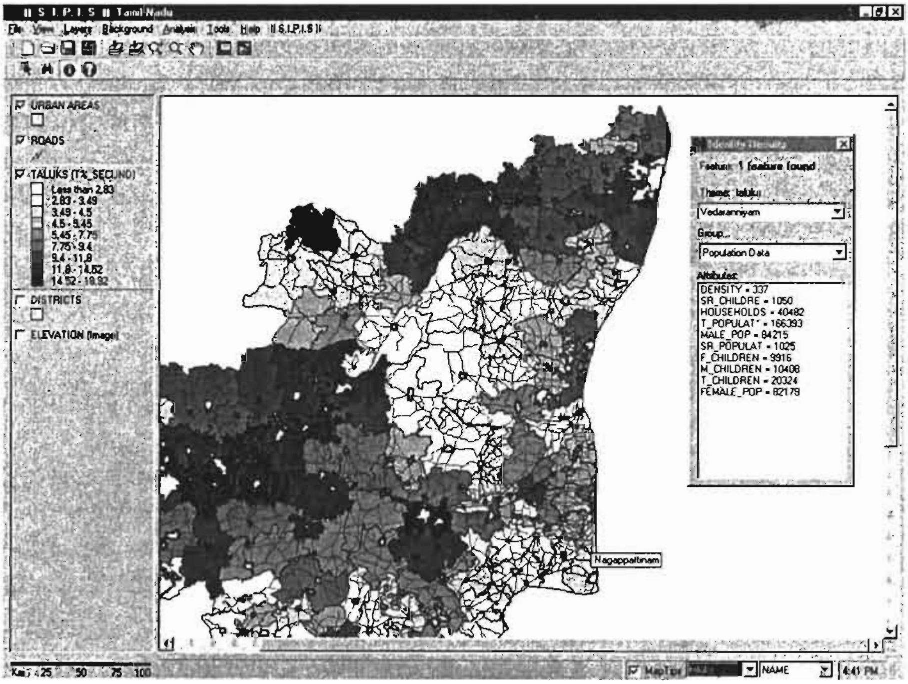


Figure 7: The SIPIS software and database

Figure 7 presents a map drawn by means of the SIPIS. A map of the labour force in the secondary sector at the taluk level is shown, with the layers of urban areas and roads as well as the demographic data for a particular unit, in this instance the taluk of Vedaranniyam (which is incidentally the first place from where Lord Rama is said to have attempted to reach Lanka). The software programming was done with tools developed by ESRI. For this reason, the cartographic functions of the SIPIS interface resemble those of the existing software *Arc Explorer* developed by ESRI. However, the interface was simplified so as to make it usable by untrained persons. A few options were added, notably a help window in Tamil language.

In this connection, it should be noted that the realization of a Tamil version of the SIPIS proved to be impossible, and this was a source of disappointment for the authors. The first reason was due to the difficulty in composing a GIS menu in Tamil because of the total absence of corresponding nomenclature in that language. But more prohibitive was the impossibility of finding an exhaustive list of villages in the Tamil language, something that cannot fail to appear paradoxical in a region characterized by a relative linguistic nationalism. Although some lists of villages exist (the directory of post codes, for example), they do not correspond to that established by the census. The best multilingual lists of villages date from the colonial period (*Alphabetical List of Villages in the Taluks and Districts of the Madras Presidency*, 1933): they of course do not correspond any longer to the present village grid. Moreover, it would be very hazardous to transcribe in Tamil the list of the Census of India (available in English) as ambiguities and errors in transcription are numerous.

The resulting CD (Guilmoto *et al.*, 2000) was very favourably received, all the more so as the subsidy by the UNFPA made a low sales price possible. The magnitude of the orders for the SIPIS from the French Institute of Pondicherry confirms the foreseen demand. The choice for the diffusion of this database would therefore appear to have been judicious, but the medium of diffusion still retains several limitations. First, once the information has been written on the CD-ROM, the product is completely finalized; with the appearance of the next census, an update will be wanted. It would thus be necessary to recommence the operation. The second limitation of the CD-ROM pertains to the physical media of information (such as printed works), that is to say, problems of diffusion: availability at all places at any given time is not possible. Finally, and this is a major point, the software content, although it was made as user-friendly as possible, nevertheless leaves the user alone faced with the interpretation of the data. A further limitation, which is technical, is that the introduction of new versions of Windows, incompatible with the earlier ones, has undermined the diffusion of the software.

Encouraged by the success of this first evaluation by the wider public, we have considered the diffusion of previously processed geographic information. In addition to the raw information diffused by CD-ROM, why not attempt to offer to the public information prepared by the researchers and briefly explained? The idea of producing an atlas naturally arose. The superiority of an atlas over a data bank for the diffusion of geographic information is of two types: graphic and scientific. Graphic, of course, because the processed information is directly readable; the maps are made by researchers who master the tools of cartography and the method of structuring information. And it is scientific because the information is selected; the researchers have chosen the maps to be presented according to their interest and following a thematic classification. Furthermore, the maps are explained, restored to their contexts[12].

This atlas could have seen the light of day on paper, as all traditional atlases, or as a CD-ROM, as is increasingly seen today. But these two forms suffer from the same limitations: impossible to update, cost of production and diffusion. The Internet, a free[13] and decentralized medium, thus seemed to be the perfect tool for our objective: to diffuse free of charge and as widely as possible the information produced by the researchers. We therefore adopted the project of an online atlas.

Other than the lower cost, which we have mentioned, the realization of an online atlas has other obvious advantages. Thus, the accommodation capacities in terms of the number of maps are nearly unlimited. Then, a later update (with data from a new census, for example), as and when new maps are published, is possible. Finally, the comparison of maps is simplified, as one is able to vary the scales (a local phenomenon can be situated in the state in which it appears or viewed on the scale of South India) and the themes (on the same scale, one can move from the fertility map to the literacy map, then to gender inequalities, etc.). Figure 8 shows a screen from the site (www.demographie.net/atlasofindia), representing the distribution of primary sector employees in South India.

[12] Various researchers of the team have joined colleagues (F. Landy, O. Aubriot, M. Thanuja, and C. Garcia) who have undertaken to write commentaries to the maps corresponding to the areas in which they specialize.

[13] Free for the user. Even though relatively low, the producer bears the costs inherent in the production of the maps and their online publication.

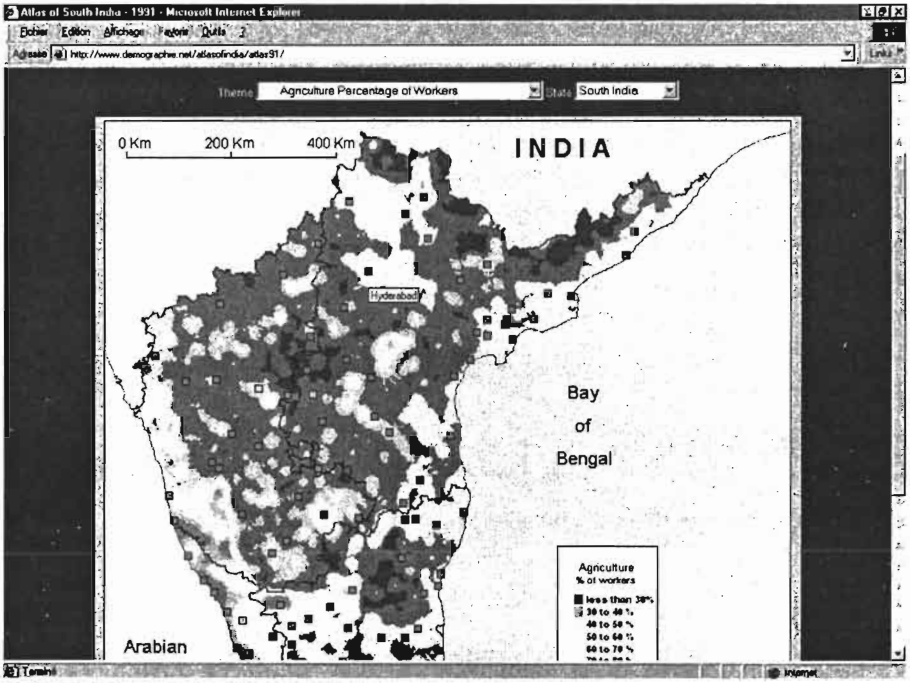


Figure 8: The South India web atlas

But the establishment of this online atlas gave rise to new technical problems. The first concerns the general structure of the atlas. The 100 available maps must be gathered together thematically. In addition, the regional division was retained. The selection of the geographic entity is made by a simple click on a pull-down menu at the chosen level: South India, Andhra Pradesh, Karnataka, Kerala, and Tamil Nadu. A second concern while realizing this atlas was to enable access to persons using an Internet access with a low rate, as is the case in India, without however sacrificing the quality of the maps. Technical choices of two types were made. First, the programming of the site was done in the least demanding manner possible, without using the undoubtedly comfortable, but user contingent, techniques of resources such as the latest cartographic servers. The HTML language, simply programmed, was therefore retained. This makes it possible to limit to the maximum extent the size of the files generated. Then, and without abandoning the quality of the maps, we opted for a graphic format reducing the size of the images generated. The format of the images adopted is the GIF, one of the characteristics of which is the encoding in 256 colours. This lightens the weight of the maps, but entailed long phases of experimentation and the image visualized on the GIS software underwent numerous transformations before being published on line.

We shall conclude by recalling that there are no perfect methods in the matter of diffusion of spatial databases. Nevertheless, and according to the objective, solutions do exist and they are efficient. Finally, the utilization of two complementary supports, CD-ROM and online accessibility, made it possible to obtain an interesting result: making available to a large public information that had until then been reserved for well-informed professionals.

From villages to spatial clusters

The phases described up to now would almost cause one to forget that the initial project was not conceived to constitute and make available an abundance of documentary material. On the contrary, this development is only a secondary product, for the endeavour of the participants in the project aimed above all at enriching the issues proceeding from the spatialization of social change in India. As soon as our geomatic tool had been put in place, the volume of information suddenly proved to be too large to allow for efficient utilization in a research objective. We had in effect a spatialized base numbering more than 75,000 villages and provided with some 100 indicators. Such a collection of localities can hardly be mapped, if not on a micro-regional scale, for the quantity of information does not fit on a map of standard size. From a strictly visual point of view, several thousand localities are not really discernable. The maps would be unreadable and the statistical or geostatistical (see below) analyses would, moreover, often be impracticable for sheer computational reasons.

Table 1 enables a rapid examination of the database for South India. Let us first note that the number of villages in our base is slightly less than that of the figures published by the Census of India because we had to aggregate more than 100 villages (137) in Andhra Pradesh. This difference in the number of villages, less than 0.2% of the total, is due to our reclassifications and diverse corrections undertaken when the consolidated statistical base was created and during its spatialization in the GIS. It should be observed that these villages, in general with a very low population, are in the main located in the tribal parts of Andhra Pradesh where their location has not been identified. These are often not permanent settlements. Moreover, some are absent on other published maps because they are sometimes situated in regions more frequented by Naxalite groups than by agents of the state.

Table 1: Villages in the SIFP database

	Census of India	SIFP dataset			
State	*Inhabited Census Villages*	Number of Inhabited Villages	Population		Number of Uninhabited villages
			Average population	Standard deviation	
Andhra Pradesh	*26586*	26449	1838	2202	1411
Karnataka	*27066*	27066	1147	1402	2131
Kerala	*1384*	1384	15475	8182	0
Pondicherry	*263*	263	1105	1240	1
Tamil Nadu	*15822*	15822	2325	2588	958
South India	*71121*	70984	1946	3034	4501

Note: Means and standard deviations computed on inhabited villages

Sources: Official data from the Census of India are from Office of the Registrar General (1997)

The primary units in the census correspond to a very disparate assembly of "villages". The table indicates in this connection the number of villages by state and their average population. One will note the considerable differences, as for example between Karnataka (mean population of the villages = 1402 inhabitants) and Kerala (8182 inhabitants). In terms of the variation in demographic sizes, South India appears to be very heterogeneous as the standard deviation is 50% greater than the average population, which expresses the considerable differences prevailing among the regions of the south. These variations remain large within the states, notably in Andhra Pradesh and in Karnataka, where the villages are generally of very small size. In the two latter states, more than 10% of them contain less than 100 inhabitants, corresponding therefore more to hamlets than to villages strictly speaking. Conversely, several hundred "villages", located above all in Kerala, exceed 10,000 inhabitants. One will also observe that in this case, the administrative grid (that is, the density of census villages) is far from reflecting the density of the population.

The administrative, historical or ecological variables that govern these regional differences could themselves be the object of a special study employing the SIFP base. But for our purposes, the size of the villages poses other problems. The first concerns the quality of the data that can be derived from these units. The second relates to the interregional comparability of these data. We will examine in the first place the question pertaining to the reliability of the data, and to this end we will examine the variability of certain indicators according to the size of the villages. We have grouped the villages in 100 classes of increasing population, beginning with the villages of smallest demographic size. Each class represents one hundredth of the sample, which means around 710 villages. For these hundred classes of villages, we have examined a given variable by calculating first its mean value, its standard deviation, and its coefficient of variation (computed as (standard deviation)/ mean). The coefficient of variation reflects the variability of the variable across its size class: if classes were random, variability would remain almost the same across classes.

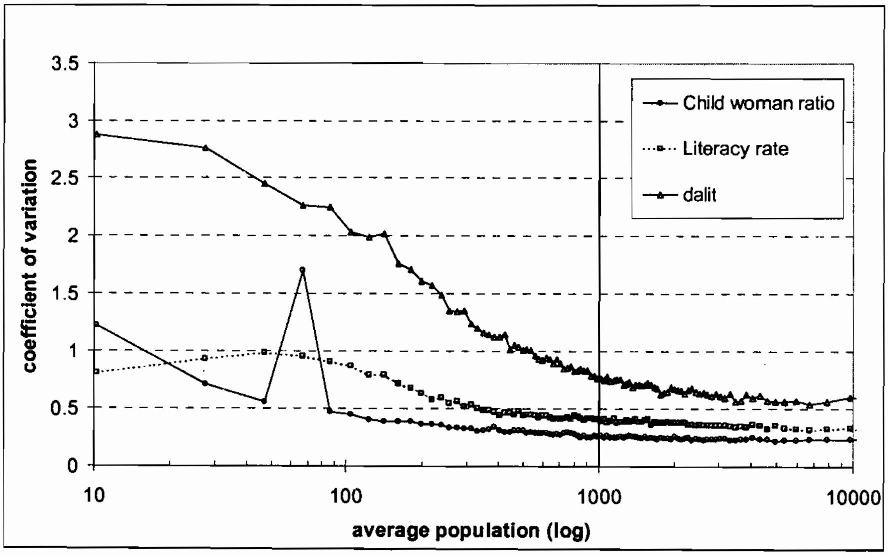


Figure 9: Coefficients of variation of three variables classified by average population size of villages

The results presented in Figure 9 refer to three variables: the child-woman ratio (CWR), used as proxy for fertility, the literacy rate (population aged 7 years or more) and the proportion of Dalits (Scheduled Castes). The values shown on the graph correspond to the variability (i.e. coefficient of variation) of these indicators according to the average size of the 100 classes of villages (omitting villages of more than 10,000 inhabitants). One will first note that the variability tends to diminish with the increase in the size of the villages, which confirms that the quality of the indicators improves for villages with the largest population. For villages with populations of less than 1000 inhabitants, the coefficient of variation is relatively high and this is still more striking for villages with less than 200 inhabitants. One will even note, for the CWR, a truly aberrant value for the fourth class of villages: this class includes, in fact, a village counting 22 children per woman, which results in a totally extravagant CWR value. This single outlier suffices to push the coefficient of variation of the CWR to a level of 1.7, which is thrice the mean value.

It can be reasonably assumed that the greater part of the variance observed in the small villages is above all due to disruptive effects of the small numbers rather than to real variations of the phenomena considered. This is perhaps less true concerning the Dalits: according to the severe spatial segregation which characterizes the village settlement, the minority Dalits are sometimes in an absolute majority in certain hamlets counted separately from the principal villages and this in part explains the strong variations observed in the small villages. But as regards literacy or the CWR, one much more suspects the effect of statistical variability characteristic of low population: the calculation of these indices on small numbers frequently leads to absurd values. Above 1000 inhabitants, this variability progresses much more slowly and tends to a minimal level that no doubt closely approaches the real variability of the considered variables. Without being able to draw an absolute limit, it thus seems reasonable to avoid treating of demographic units of less than 1000 inhabitants, because the smaller localities are greatly affected by statistical imprecision. The study of statistical robustness shows more generally that the indicators employed (rate, ratio or percentage) are dangerously unstable, or even devoid of significance, when the reference population was around 200 persons or less.

It therefore became necessary to adopt a procedure for aggregation so as to constitute larger populations. Different methods were considered, but it was soon evident that one could not make use of the administrative grid for reasons similar to those seen in the case of villages. The smallest units, namely the taluks, were not at all comparable in area or population from one state to another, as indicated in Table 1. In particular, the taluks had been eliminated in Andhra Pradesh before the 1991 census and replaced by much more scaled-down units (the *mandals*), which are much smaller. While the *mandals* are utilizable for comparative purposes in Andhra Pradesh (as Virginie Chasles does below), it is not possible to compare them to the taluks of the neighbouring states, the average population of which is from ten to twenty times larger. Thus, 1513 sub-district units are counted in South India, but among them more than two thirds (1099) are in fact *mandals* in Andhra Pradesh, whereas Kerala contains only 61 (taluks).

It was also necessary to avoid ascribing an identical weight to each unit (i.e. each village) independently of its population, as is commonly done in the smoothing programmes contained in commercial GIS software. It would have been a serious error, of course, to estimate the mean of an area by simply averaging the values in the villages belonging to it. The highly variable size of the villages (in terms of population or of area, etc.) would have made the direct smoothing of the observed values aberrant, for it would have accorded an undue weight to the small villages compared to large villages.

We therefore decided to proceed to a group of villages based exclusively on a criterion of spatial proximity, independent of administrative borders, including those of the states. The usual method of rectangular grids seemed hardly appropriate because it would have had the effect of placing at a disadvantage the border zones, along the sea or southern borders, which had a size smaller than the standard rectangular grid. Introducing units necessarily smaller than the others in the administrative or coastal bounds would have created a spatial bias. We sought to find a grid constituted of comparable units of area by employing Voronoi polygons.

The agglomeration procedure was as follows. First, an aggregation point around the villages was automatically identified, after having fixed a maximal distance. The point obtained is the centre of gravity of different villages within the radius of fixed size. Starting from this first layer of aggregation points, a spatial partitioning was created in Voronoi (or Thiessen) polygons covering South India. These polygons are areas created around points, usually in the form of hexagons[14]. Without being identical,

[14] This corresponds to a partitioning of South India with *n* aggregation points into *n* Voronoi polygons such that: each polygon contains exactly one aggregation point as its centre and every location in a given Voronoi polygon is closer to its central point than to any other central point. About Voronoi polygons and aggregation procedures, see Pumain and Saint Julien (1997) or Martin (1996).

they have relatively comparable areas, also in the border regions of our maps of South India. By way of example, we reproduce the map of Thiessen polygons on the basis of aggregations of 20 km in Figure 10. South India is divided into 620 units of relatively homogeneous size (see also Table 2). The villages are grouped together within each polygon and the observed socio-demographic values are accumulated.

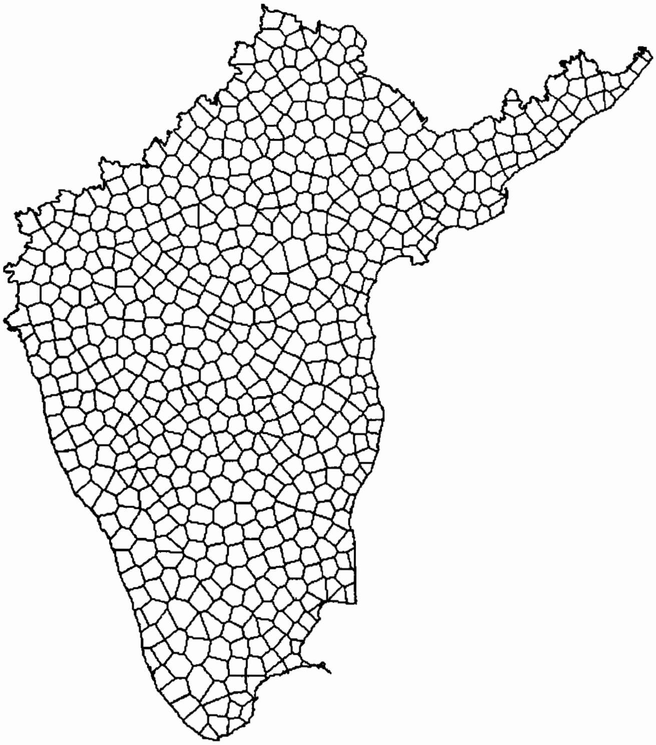


Figure 10: Division of South India into 20-km clusters

In this manner a series of "clusters" was defined, that is, groupings composed of a variable number of primary village units. The area of the clusters obtained is, as said, rather homogeneous. The socio-demographic indices are then recalculated within each cluster to obtain a global value. The same operation was carried out for four distinct sets of aggregation, determined by the distances of 2, 5, 10 and 20 kilometres respectively. The divisions into clusters obtained will be all the smaller given the fact that the distance of aggregation is reduced. This method, developed by

Christophe Z. Guilmoto, although in part empirical, is to our knowledge the best way to compose regular spatial divisions starting from a territory such as South India.

These aggregations considerably improve the quality of the statistics as the number of under-populated units is moderated, or becomes negligible. Thus, in the aggregation with a distance of two kilometres, the number of statistically vulnerable units of less than 200 inhabitants decreases from 15.8% to 2.1% of the sample. This proportion becomes truly negligible at higher levels of aggregation. As shown in Table 2, the number of units decreases, from 24,000 (for a distance of 2 km) to 7000 (for 5 km) to 2000 (for 10 km) and then to 600 (for 20 km) respectively, which makes all the foreseen mapping possible. It suffices to vary the level of aggregation utilized according to the scale retained, choosing notably the largest aggregations for the maps of the whole of South India. In the same way, thanks to this smaller sample of better quality, statistical and geographic modelling (see below) are now realizable in better conditions for calculation.

Table 2: Characteristics of villages and clusters

	SIFP villages	Cluster types			
		2 km	5km	10 km	20 km
Number of units	70984	23848	6974	2151	620
Average area	784.8	2403	8309	26991	93643
Standard deviation	1185.9	2046	4153	8854	26123
Variation (%)	151.1%	85.1%	49.9%	32.8%	27.8%
Average population	1946.6	5794.3	19814	64241	222877
Standard deviation	3035.0	6368	17961	53086	166218
Variation (%)	155.9%	109.9%	90.6%	82.6%	74.5%

Table 2 summarizes the characteristics of these families of clusters, eliminating the few remaining clusters inhabited in mountainous zones. Here, the number of units has been indicated, the average size and its coefficient of variation, as well as the average area and its coefficient of variation. As the diameter of the aggregation increases, the number of units decreases, while the population and the average area of each cluster continues to increase. Also important in our view is the rapid reduction of the variability of the units, as the line of coefficients of variation shows.

Also crucial is the variability of the areas. It is more than 150% for the villages of the census, which means that the database comprises village territories with areas that are difficult to compare. Thanks to clustering, these variations in area between the units diminish very quickly: the variability (coefficient of variation) thus amounts to 50% for the 6974 clusters of 5 km. This indicates that the units are henceforth of a much more harmonious size. Turning to the administrative units of sub-districts (i.e. 1513 taluks and mandals), one observes a variability in area of 100% between these units: this figure is thrice that of clusters of 10 km, which are comparable in number.

Of course, the population remains more variable than the area despite clustering because the effect of the differences in regional densities subsists. But the variability of the population of the villages in the first clusters of 2 km nevertheless decreases by one third. The sparsely populated clusters, or with a population of less than 1000 inhabitants, have not entirely disappeared because there are still large forest or mountainous regions that are sparsely populated. Their part in the whole, however, becomes very small and the effect of variations characteristic of the administrative grouping, illustrated by the "micro-villages" of the region of Mysore and the "macro-villages" of Kerala in Figure 6, tend to become indistinct.

In order to be mapped, the data will often then be smoothed and contoured. Figure 11 summarizes the different stages in the study of a phenomenon, starting from the primary data of our GIS to obtain a systematic cartography. We are using here the values of the child-woman ratio (computed as the ratio of children below 7 years to women aged 7 or more), which we will utilize as a proxy for fertility. The zone represented covers the north-east point of Andhra Pradesh, centred on the district of Visakhapatnam. All the 6700 villages of this region are represented on the first map. It will be noted that the highest densities of units of population do not necessarily correspond to a greater demographic density. Thus, the northern limits of the State (along the border with Orissa) encompasses several hundred census villages which, in fact, are only isolated tribal hamlets of a demographic size seldom greater than 100 inhabitants. Conversely, the villages on the coast oriented toward the south-east are much more populated, but appear very scattered.

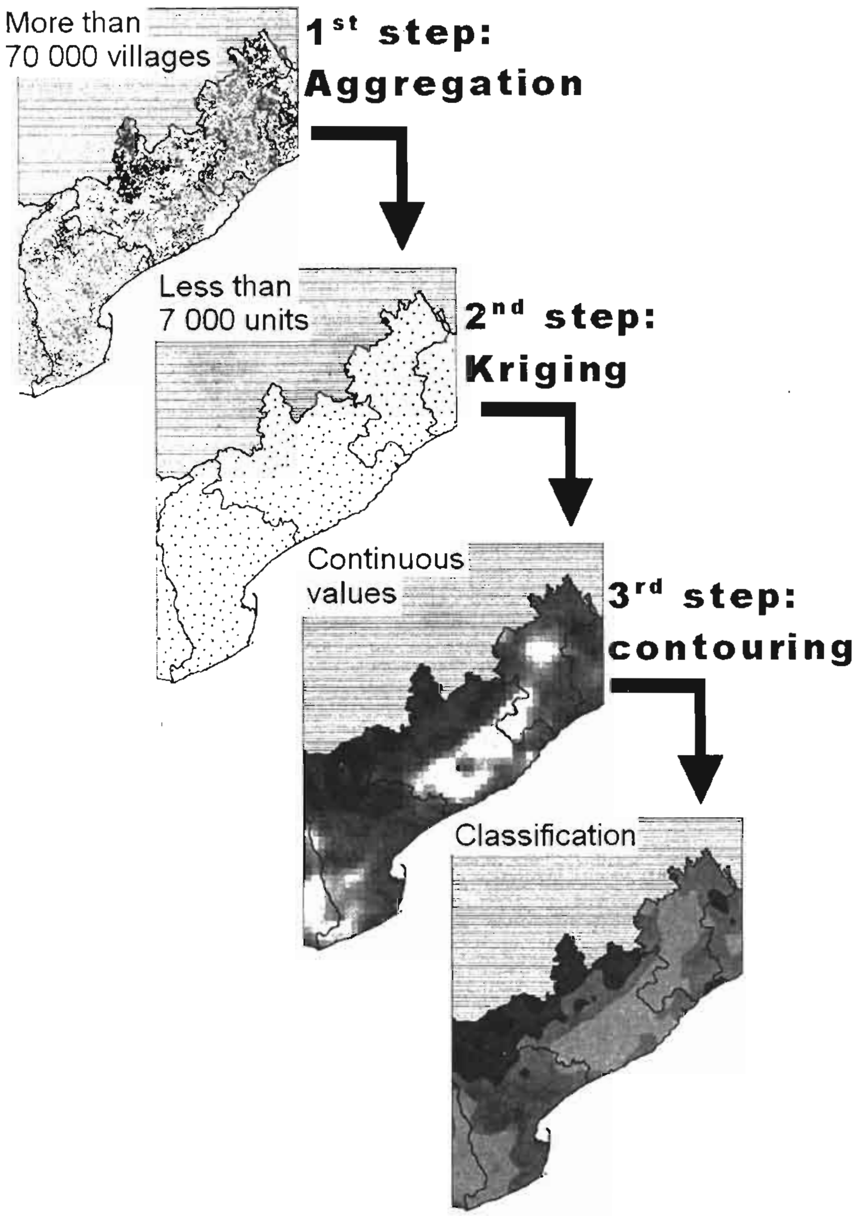


Figure 11: Mapping fertility data in North-East Andhra Pradesh: original villages, cluster aggregation, kriged estimates and final contouring

The second map in Figure 11 reduces this sample to a set of some 500 clusters in this rural pocket of Andhra Pradesh. The new units composed are very evenly spread out over the area and their average demographic size remains large. Only a negligible number of them count now less than 1000 inhabitants. In addition, the values of the fertility index are clearly less uneven, calculated now on aggregates of several thousand persons. They are then subjected to kriging[15] on the following map, in order to proceed from a vector representation to a representation by area (raster), which we have chosen to represent here with the help of cells (size of 5 km x 5 km). This is a spatial interpolation allowing of an estimation of fertility values over the entire area under consideration, and not only at a point (the centre of the clusters). Kriging, based on the observed values and the spatial autocorrelation measured in the sample, is a better method than direct linear interpolation.

The final map is a contour map prepared from interpolated values by classifying fertility values in four classes. Rather than keeping a raster map that usually proves difficult to analyze, we opt for isopleth maps by drawing isolines (continuous lines joining all points of identical value).

The last map clearly shows that here fertility values almost double from the central point of the region to the north-west frontier of the state, where fertility is the highest. We have chosen to represent this region precisely because the fertility differentials are the highest observed in South India: the central part is a rice-growing region with a high industrial development around the port of Visakhapatnam (a city of more than one million inhabitants) that is entering into the final phase of its demographic transition and where fertility rates sometimes approach 2.5 children per woman. At the same time, the forest regions in the north bordering Madhya Pradesh and Orissa are among the poorest in Andhra Pradesh, with a population whose majority is often tribal and with very low literacy levels: fertility remains very high and the impact of the proximity of rich and urbanized regions in the coastal plain seems to be imperceptible, so impregnable is the social barrier between the cultural groups.

One could, of course, proceed to a finer analysis of the cartography of fertility, for example, underscoring the local impact of urban agglomerations, as well as that of the major trunk roads in north-east Andhra

[15] The technique of kriging (here, ordinary kriging) constitutes from a statistical point of view the best smoothing technique. It is described, for example, in Fotheringam *et al.* (2000).

Pradesh. But that would mean departing from the objectives we have set in this section, for we want only to summarize the stages in the processing of spatialized information without dwelling at length on a particular example. The use of aggregations by cluster thus proves to be indispensable, as the detail of the maps obtained for all the villages would be too fine to perceive, sometimes too heavy for the machine processing the base information, and most often precarious because of the imperfection of the data when the reference numbers are reduced. The following section will abandon the strictly speaking technical discussions to illustrate the applications of the new database to different themes of interest to social scientists.

2

APPLIED SPATIAL ANALYSIS

The diversity of available information makes it possible to examine the South Indian database in terms of very disparate questions, ranging from demography in the strict sense to economy or sociology. The relative flexibility of the software tools used to manage the base has allowed numerous non-specialized researchers to have access to it without particular technical difficulty and has given rise to projects exploring numerous aspects of contemporary reality in South India. In this second section, we shall give the word to researchers who have chosen to concern themselves with only certain aspects of the base. It would, in fact, be illusory to want to deal at the same time with all the spatialized data of South India, if only because of the calculations involved in the number of spatial units or available variables.

We shall limit ourselves here to the programme Population and Space in South India, carried out by the IRD and the Department of Social Sciences of the French Institute in Pondicherry. We will cover different themes linked to the doctoral projects of four young researchers who have worked in Pondicherry, as well as two research topics explored by the principle investigator (Christophe Z. Guilmoto). Very disparate themes are to be found therein: irrigation, pilgrimage, town-country relationship, sexual discrimination, health provision, geostatistical analysis.

Other works have used this base as a support; however, they are too numerous to be presented in detail in this paper. We can nevertheless briefly summarize them, beginning with the programme that initially justified our entire enterprise: the South India Fertility Project, bringing together French and Indian demographers. The participating teams assumed responsibility for the different states of South India. Thus, Andhra Pradesh was studied by P. Ramachandran from the Sri Venkateshwara University in Tirupati, with the assistance of P. Ramesh, while K.S.R. Raju and T.V. Sekher from the Institute for Social and Economic Change (Bangalore) were responsible for the state of Karnataka. Kerala was studied by S. Irudaya Rajan and his collaborators, while the team of the Department of Population

Studies of Bharathiar University in Coimbatore -which included successively P.M. Kulkarni, S. Krishnamoorthy, N. Audinarayana and various collaborators- was entrusted with the analysis of Tamil Nadu. On the regional scale of South India, the teams first undertook a systematic analysis of the temporal and spatial profile of fertility; the works relied in particular on the village databases, which have a robust fertility indicator (child-woman ratio) and more than one hundred other descriptive variables that were subjected to detailed statistical analysis (modelling of variations in fertility according to locality). The regional rural fertility maps also made it possible to distinguish within the states, but also within the districts or taluks, the magnitude of the differentials which had in the past been ignored due to lack of finer data. It is today possible to examine fertility within the districts and to find the specific configurations of these variations, which are always very homogeneous from a micro-spatial point of view, but disparate on a more global scale[16].

The now completed second phase of the SIFP project enabled of the identification of representative villages in different typical zones of fertility on the basis of detailed cartographic analyses. Fieldwork was then conducted in a sample of ten villages, relating their fertility position and their social, economic and geographic profile. The basic idea was to complement the statistical and spatial analysis with a intensive qualitative survey using individual interview and focus group discussions: these investigations followed a common research protocol designed by the joint teams and were carried in 2002 and 2003 in a sample of representative rural localities in the four South Indian states. Findings from this phase will be available soon.

A different study relying on the database was also devoted to a forest milieu in the Ghats of Karnataka endangered by the pressure of plantation economy (Guilmoto, 2000). In this project, we tried to confront the social and demographic evolution in the district of Kodagu (Karnataka) to the local ecological dynamics, and above all to the progressive reduction of the evergreen forest to the advantage of the recent extension of areas planted with coffee. This work depended on a twofold cartography, human and forest, with the objective of showing the relationship between the changes in land use and the changes in the socio-economic profile of the district.

[16] This work was the subject of a collective publication (Guilmoto and Irudaya Rajan, forthcoming). More materials from the SIFP may be found on its web site: www.demographie.net/sifp

We may also mention other works, such as that of Corinne Pruvost-Giron on the Palakkad region in Kerala (Pruvost-Giron, 2002), based in particular on the exploitation of spatial data pertaining to several border taluks between Kerala and Tamil Nadu, or that by Olivia Aubriot on the system of tanks in Tamil Nadu, including detailed cartographic analyses of tank irrigation. A work under completion devoted to agricultural dynamics in the Palar region (Xavier Amelot) also makes use of a regional portion of the GIS of South India. To conclude, let it be mentioned that the data will be subjected to a more systematic treatment in a theoretical programme of spatial analysis and geostatistics that has just begun with a large number of the partners from the programme Population and Space[17].

The following section consists of essays of varying length written by those responsible for the different projects derived from our database. We have chosen to present them in order of increasing technical complexity, but the extreme variety of topics taken up necessarily makes this order somewhat arbitrary.

Graphic representations of the Sabarimala pilgrimage (Kerala)[18]

The pilgrimage to Sabarimala in the inland mountainous south of Kerala is organized around a cult of universalistic design – that of Ayyappa, ambivalent deity born of the marriage of Siva and Vishnu. It represents in the contemporary period one of the most significant phenomena in the reconstruction of the religious in India. As Radhika Sekar (1992) noted, the pilgrimage to Sabarimala presents anthropological and sociological characteristics that clearly distinguish it from other South Indian pilgrimages: a period of ritual initiation (*mala puja*), a basically male

[17] We are referring to EMIS project (Space and Measurement in South India), a programme supported in France by the CNRS and the IGN since 2003 and bringing together O. Aubriot, V. Chasles, C. Pruvost Giron, C.Z. Guilmoto, S. Oliveau, K. Radja, A. Vaguet and S. Vella. See http://www.demographie.net/emis.

[18] Section written by Rémy Delage. The author wishes to thank most particularly the team of a few professors and some dozen students for having participated in the survey (Kamaraj University in Madurai and Tamil Nadu, Sri Venkateshwara University in Tirupati, Andhra Pradesh, and colleges in Erumely, Palakkad and Pondicherry). A survey of such scope could not have taken place without the decisive assistance of an interpreter, P.V. Ashokan (Palakkad).

pilgrimage in order to preserve the ritual purity of the site (pubescent women are excluded from it), a period of austerity (vegetarianism, piety and chastity) or *vratam*, the promotion of a message of equality and an opening to all social strata as well as to all religious adherences[19]. An observer makes the following comments: "A silent and unnoticed religious fervour has been sweeping the southern states in the last two or three years. In numerous towns and villages, hundreds of boys are joining thousands of their elder menfolk in donning the black garb indicative of the Ayyappa *bhakta*. From street hawkers to auto drivers, bus conductors to salaried employees, skilled artisans to professionals [...] irrespective of wealth or riches, the Ayyappa fever seems to have infected everyone" (V. Suresh, 1992: 178). The Ayyappa cult also integrates, although in negligible proportions, Muslims and Christians.

To obtain the *darshan* ("vision" or divine benediction by the sight of the idol) of Ayyappa, which is the main goal of the pilgrimage to Sabarimala, the pilgrims must fulfil a series of austerities during a period of 41-61 days preceding the departure, calling to mind the figure of the ascetic renouncer. Paradoxically, this is hardly apposite as regards the Sabarimala pilgrimage, in view of the mass of pilgrims who make their way to receive a *darshan* responding above all to personal and very concrete preoccupations. The desire for liberation or deliverance (*mukti*) that characterizes the spiritual quest of the ascetic appears here to be secondary, that is, quite different from the actual social reasons of the pilgrimage expressed in the vows made and the thanks given. This pilgrimage thus bears the stamp of devotion (*bhakti*) and for this reason does not diverge from the general movement of uniformization of practices that characterizes modern Hinduism. "Structural factors partly account for the development of reformist ideals, for the standardization of practices and iconography, and for the progressive construction of a common 'Hinduism' " (Tarabout, 1997: 144).

---

[19] Among the many processions which punctuate the course of the pilgrimage, those for which Erumely is well-known (point of departure of the traditional route through the forest) involve a processional activity quite uncommon in India. The pilgrims go for the *darshan* of Ayyappa in a temple, then for that of his Muslim companion Vavar in a mosque, before ending with the *darshan* of Ayyappa in the largest of the two temples. Once the procession and the ritual ablutions have been completed, the pilgrims can begin their journey by foot through the forest, or even by vehicle via road to Pampa.

However, without dispensing with the anthropological material necessary for the interpretation of the phenomenon in its totality, we instead stress here the spatial dimensions of the pilgrimage analyzed on different scales, starting from the local or micro-regional level and proceeding to South India as a whole. While in the 1950s-1960s, Sabarimala only attracted groups originating from the Malabar Coast and from the Tamil districts neighbouring Kerala (Madurai, Virudunagar, etc.), the pilgrimage today draws several million devotees coming mostly from the four corners of South India (Kerala, Tamil Nadu, Karnataka and Andhra Pradesh). It is interesting to note in passing that the popular literature in Malayalam and Tamil on the Sabarimala pilgrimage often makes reference to an ideology known as "Dravidian" to (re)construct a discourse of identity around a phenomenon that is above all of social and religious nature.

From these elements a first hypothesis follows, that of the opposition between the Sabarimala pilgrimage, confined to the limits of South India, and the pan-Indian expression of pilgrimage, illustrated for example by Tirupati in Andhra Pradesh, an opposition that has refuelled the debate on the traditional North-South divide of India. The second hypothesis, which we shall partially verify in this contribution, is the geographic hypothesis of the temporary distortion of the map of South India by the itineraries of the pilgrimage that sustain here and there certain historical major routes.

Taking the map as a common thread of this study, but also as an expedient for describing certain structures of the pilgrimage, we shall begin by citing a number of original graphic descriptions of the event from the popular literature. The next stage consists in retracing the major methodological steps in the compilation of a database on the pilgrimage that we have been able to establish by combining the results of a survey by questionnaire with the georeferenced SIFP database. We will then be in a better position to present the initial results of a new cartography of the major routes of religious movement between Kerala and Tamil Nadu. This should, in conclusion, lead us to a redefinition, on the one hand, of certain migratory logics structuring the mobility of pilgrims through an array of preferential itineraries and, on the other hand, to the integration of this study in a historical geography of the flow and the networks of circulation in Tamil Nadu, as well as in South India.

Popular mappings of the Sabarimala pilgrimage

The micro-territory centred on Sabarimala is located at the intersection of four districts and at the confluence of spaces of differing natures. To understand the links existing between this cultural territory and the numerous networks established there, we must take into account the temple villages of Pampa (located 4 or 5 km from the Sabari Hill) and of Sabarimala (Pathanamthitta district), as well as the localities of Pandalam (Pathanamthitta district) and Erumely (Kottayam district). While the village of Erumely constitutes a major junction in the network through which the majority of the pilgrims must pass, it appears that some processions involving the localities of Pandalam[20] (Pathanamthitta district), Alangad[21] (Ernakulam district) and Ambalappuzha (Alappuzha district) go beyond the supposed frontiers of the territory under study.

Even though one never finds very accurate maps, the popular literature on the pilgrimage to Sabarimala (magazines, booklets, reviews and seasonal publications) nevertheless offers a collection of publications in the form of sketches or maps on different scales (see Figure 12), according to the nature of the booklet and the geographic origin of the publisher. They can be obtained in the bookshops of Erumely, Pampa and Sabarimala, but also more or less everywhere in South India if one explores the literature in vernacular language (sometimes in English) on the Ayyappa cult and the Sabarimala pilgrimage. These maps represent valuable research material to the extent that they indicate the main routes to reach the temple, situated in a sparsely populated mountainous region with a high forest density, in the centre of the Periyar Tiger Reserve (PTR). They also make it possible to identify the cult sites marking out the itineraries and structuring the pilgrimage to Sabarimala.

[20] Until the proclamation of the Travancore-Cochin Hindu Religious Institution XVth Act in 1950 and the transfer of authority from the royal family to the Travancore Devaswom Board (TDB), the royal family of Pandalam was in full possession of the temple and its lands, but also in charge of the ritual programme.

[21] Today, the two groups belonging to the neighbouring localities of Alangad and Manjapra today claim the legitimacy of the processional group of Alangad (*Alangad Yogam*).

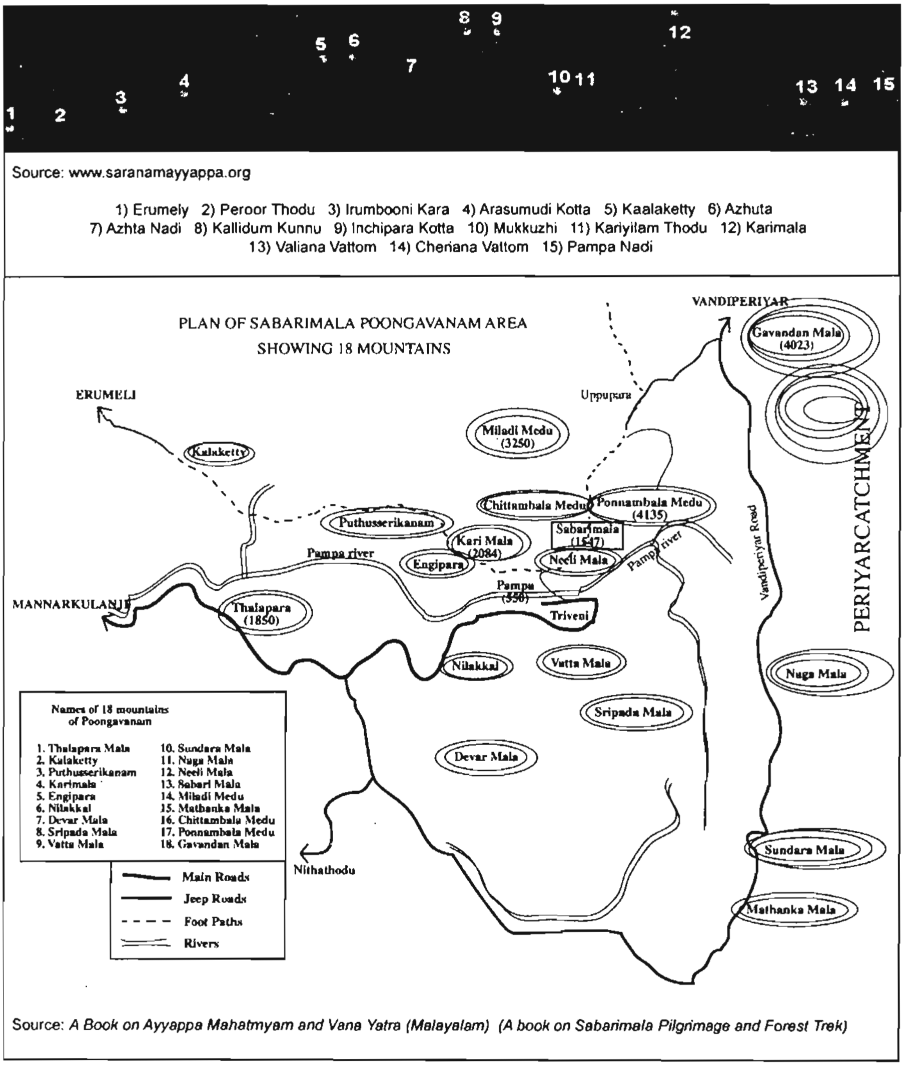


Figure 12: Topographical cross-section of the forest trek and « Ayyappa's garden » (Poonkavanam)

On the basis of two distinct sources of information (the Internet site of Akhila Bharatha Ayyappa Seva Sangham[22] and a Malayalam booklet entitled

[22] The address of the site is: www.saranamayyappa.org. Dating from 1946, the Akhila Bharatha Ayyappa Seva Sangham (ABASS) association was created under the regime of

"Ayyappa Mahatmyavum Vana Yatrayum"[23]), we have brought together in Figure 12 an approximate topographical section of the forest route between Erumely and Pampa (about 75 km) and a flat representation of the "Ayyappa garden" (Poonkavanam), a symbolic territory in the form of an envelope crowning the Sabarimala temple or immediate sphere of influence of the deity (Sabarimala Kshetram). The first allows one to discern the difficulties connected with the physical effort of the march, while the second drawing illustrates the role played by the mountain in the patterning of a geographic imagination of the pilgrimage (the forest route is represented by dotted lines on the sketch, between Erumely and Pampa via Kalaketty and Kari Mala).

It should be added that, on the regional scale, the Western Ghats have long been integrated in the collective imagination as a place of divine sojourn and as playing the role of a major geographic barrier, curbing or preventing interactions of all kinds. Numerous passes (for example, that of Shenkottai), and in particular the Pallakad Gap, have enabled commercial exchanges, tracing itineraries anchored in the cultural representations of space.

In Figure 13, we can identify on the basis of a map in Tamil and its English translation (figure 14), derived from a booklet that appeared in Madurai, the major gateways to Kerala: by way of the Pallakad (Palghat) Gap to the north, the most recent route from Kumily at the level of Madurai, the Shenkottai passes to the south or the long historical route by way of Kaniyakumari at the southernmost point of the Deccan. One will note that the geometrical centre of the map is none other than Sabarimala, representing thus an area demarcated in the north by the town of Coimbatore, in the east by the temple town of Madurai and in the south by the pilgrimage centre of Kaniyakumari, and overshadowing the rest of Tamil Nadu. This obviously does not correspond to the real zone of attraction of the pilgrimage, which merges with the administrative boundaries of South India (north of Karnataka and Andhra Pradesh).

---

the Charitable Trusts Act. It carries out the promotion of the Sabarimala pilgrimage and of the figure of Ayyappa, as well as the diffusion of the values of equality and fraternity.

[23] See Bhaskaran Nair (1998).

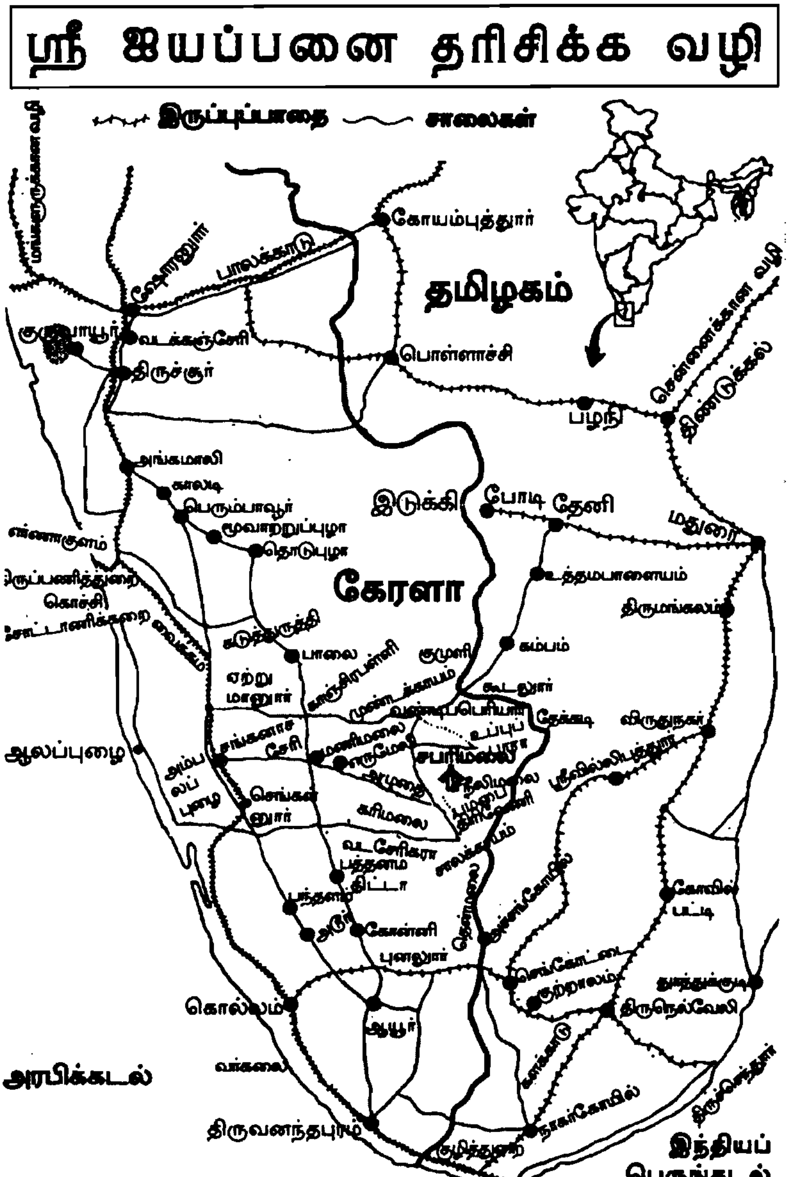


Figure 13: The road network leading to Sabarimala from Madurai, Tamil Nadu (sources: *Dinamalar*, 29/10/2000)

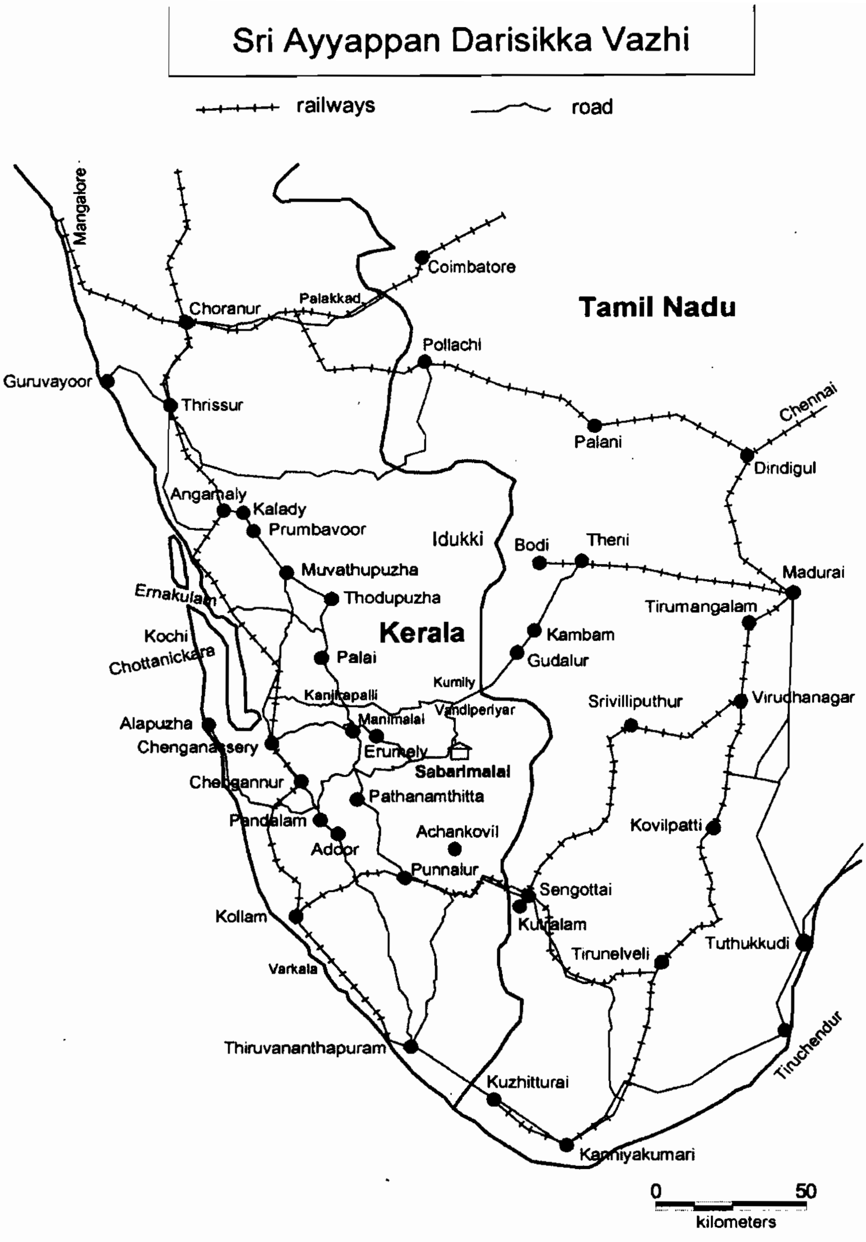


Figure 14: The road network leading to Sabarimala from Madurai, Tamil Nadu (previous map translated in English)

By matching the geographic information from sources of this type, it would have been possible to draw a map of the infrastructure of both geographic and ritual networks structuring a territory centred on Sabarimala. This will allow us to examine the more complex hypothesis of the regional opening of Kerala through the contradictory play of the increasing concentration of transport networks on its margins, conflicting movements between processional groups and the "deterritorializing" effects of the management of the pilgrimage by the Travancore Devaswom Board[24], a semi-governmental institution responsible for the temple affairs.

The analysis of these popular maps of the pilgrimage makes it possible to formulate the hypothesis of a discrepancy between the geographic imagination, the representation and the reality of the phenomenon. To respond to this hypothesis, it is necessary to compare the first results with those linked to the treatment of a complementary type of geographic information, namely, data recorded directly from the pilgrim population during the main season. The use of the SIFP database is here clearly distinct from its employment in other contributions in the sense that we do not make use of the census data to work on the geographic magnitude of a phenomenon. Rather, the georeferencing of this base is used to locate the places that constitute an itinerary of the pilgrimage, starting from a field survey of a phenomenon having a South Indian purview.

Methodological challenges of mapping pilgrimage data

After the exploration of a territory in the form of networks centred on Sabarimala, the cartography of the itineraries to Sabarimala between Tamil Nadu and Kerala will enable us to discover the gaps between the geographic imagination and the reality of the Sabarimala phenomenon in South India. In fact, the maps presented earlier showed above all the itineraries of the networks on the Kerala scale, or the internal ritual routes between Erumely and Sabarimala. The data collected during a survey by questionnaire now serve to represent in a complementary manner the routes outside this cultural micro-region, that is to say, itineraries on the scale of South India between

[24] Projects include the laying of a railway line between Kottayam and Erumely, of a cable car between Pampa and Sabarimala (a project abandoned because it was contrary to the spirit of the pilgrimage), new forest roads between Pampa and Sabarimala, and the invention of new ritual connections along the processional route by the Alangad group and in numerous temples situated in southern Kerala, in part destructuring the original geographic networks of the pilgrimage.

the place of residence of the pilgrims (point of departure) and the villages of Erumely and Pampa (places of arrival). We have confined ourselves in the framework of this contribution to the routes followed by the population of Tamil origin, as the latter is strongly represented during the main season and during the monthly openings throughout the year.

All stages of the methodological trajectory will be examined, from the conception of the questionnaire and the conditions of the survey to the data entry and processing in the georeferenced database, which makes it possible to accurately map the major pilgrimage routes leading to Sabarimala.

In order to draw up a very precise sociological and geographic profile of the Sabarimala pilgrimage, a vast survey by questionnaire was carried out in the village of Erumely among a sample of around 900 pilgrims during the main season, which extends from mid-November to the end of December in 2001 (*Mandalam* season) and to the end of January in 2002 (*Makara Vilakku* season). The challenge in this undertaking resides more in the size and the form of the sample than in the structure of the questionnaire itself. In fact, we shall not present here the socio-economic data of the pilgrim population derived from the collected information (age, family, level of education, profession, standard of living, etc.), but we shall instead focus our attention on the geographic origin of the pilgrims (state, district, taluk/mandal, village/town) and the trajectories they temporarily delineated. Also, for about three months, Malayalam and Tamil, but also Telugu, Kannada, Hindi and English are spoken in Sabarimala. This very broad linguistic gamut lends the pilgrimage an undeniable touch of originality, but at the same time represents an obvious constraint for the course of research. The original questionnaire was therefore prepared in English and then translated and distributed in six languages (English, the four main languages of the South and Hindi).

A range of possibilities for graphic representation emerges from this operation, resting on a cartography of variable scale centred on the most significant itineraries at the levels of South India and the four states represented, the main places of passage and the practices connected to them, as well as the other pilgrimage centres annually visited by the swamis[25].

[25] The pilgrims setting off for Sabarimala call themselves and call each other swami or Ayyappa, out of concern for equality among themselves and before the god. In effect, they elude the rigidity of the social order during the time of the pilgrimage, but nevertheless reinstate their social identity and their status once this geographic ritual has ended.

Continuing a bit further with the spatial imagination of the pilgrims, we can also map the places they dream of visiting at least once in their lives. Although difficult to realize for technical reasons, a flow-line map with proportionally sized arrows would represent the ultimate possibility of graphic representation and modelling of the Sabarimala pilgrimage by making it possible to visualize the most frequented nodes connecting different flows and perhaps revealing the competing catchment areas of a place.

Beyond the entry of the questionnaire data in a database, the stage that directly follows the empirical process of the field survey (questionnaire design and survey operations) consists of encoding, separating and classifying these chains of geographic places or trajectories.

From survey information to geographic and statistical data

Having collected information on the itineraries of the pilgrimage, which appear in the form of chains of places between the point of departure and Sabarimala (Table 3), two methods are available to us for sorting the data. The first consists in positioning the places manually so as to draw the itineraries, but this considerably limits the subsequent possibilities for working on the same type of thematic applications as the database would fail to be georeferenced.

Table 3: Coding of localities and geographic coordinates

Map code	Village/Town	X	Y
516100	Madurai U.A	78,102130000	9,887601000

The second method links the raw Sabarimala database and the SIFP database following an empirical process of location and identification of the places one-to-one. We allocate to each of them a set of data consisting of a general identification code, a name, as well as the *x* and *y* coordinates, which are indispensable for a precise plotting of the pilgrims in their migratory movement (Table 3). This method makes it possible to draw complex itineraries in a quasi-automated manner (Table 4).

4A

Table 4: Succession of places visited by a pilgrim on the way to Sabarimala

Chennai U.A
Villupuram
Tiruchirappalli U.A.
Tiruparangundram
Courtlam Slopes
Achankovil
Kulathupuzha
Erumeli South
Nilackal
Pampa
Sabarimala

While some of the places collected on paper are at times difficult to situate because of the bad spelling or writing of the surveyors, a still larger number of these places raise problems for a different reason. This is the case of the geographic origin of the Kerala pilgrims who often identify themselves with hamlets belonging to a much larger geographic entity, the *desam*. As the latter are not administrative entities, they are not listed in the SIFP database. In the case of places of passage such as the temples visited (which do not always correspond to administrative units), it was necessary to refer to tourist maps of the states concerned and then return to the original census maps or search the atlases, popular literature or any other source of information to locate these holy places[26]. When a village cannot be located, we replace it with the coordinates of the chief town of the taluk or, failing that, with those of the district capital. Once the missing links have been located, we create new points (code, name, *x*, *y*) so as to complete the database (Table 4).

The logic of the questionnaire would have it that a single journey there and back corresponds to several individuals from the same group of pilgrims. At this stage in the preparation of the final database, we identify the similar routes and then sort the database in such a way that only the outward journeys remain. We can then classify the itineraries according to

[26] See the maps indicated in the bibliography.

the number of places cited or the number of stopovers, from the most to the least informed in terms of geographic information, according to a process of triangulation of the table of data. Starting from a sample of 263 Tamil pilgrims and approximately 70 distinct itineraries, we first aggregate the individuals belonging to the same administrative entity for a map of departure representing the geographic distribution of pilgrims in Tamil Nadu (map with a pattern of points) and the number of individuals according to the district of origin. Then, we choose from among all the itineraries those which appear to us to be the most significant in order to best respect the variety of the routes according to the point of departure and the number of stopovers (selection of some dozen itineraries). Let us note that this type of cartography of pilgrimage routes is not statistically significant because of the unevenness of the geographic information (the number of stopovers mentioned in the questionnaires) collected from the pilgrims.

Between Tamil Nadu and Kerala: the pilgrim's routes to Sabarimala

We have chosen to break down the graphic method of plotting the itineraries of the pilgrimage between Tamil Nadu and Kerala into four stages: a table of geographic origins of the pilgrims by district (Table 5); the geographic distribution of the pilgrims in Tamil Nadu (Figure 15); two conventional forms of graphic representation of the pilgrimage (Figure 16); a map of the pilgrimage routes of Tamils visiting Sabarimala (Figure 17).

Table 5 constitutes a necessary stage in the mapping of the geographic distribution of the Tamil pilgrims according to the district of origin, but also town and village (see Figure 15). At first sight, one observes the relative concentration of the individuals in the centre Tamil Nadu, with a majority of pilgrims originating from the Madurai region[27]. Although it is true that the pilgrims coming from the districts neighbouring Kerala maintain privileged links with Sabarimala, this differential is perhaps to be ascribed to the place where the survey was conducted (Erumely), which is located on the Kumily-Kottayam axis and which is of easier access for them. In addition, we could

[27] Ayyappa is supposed to have spent time in Madurai, the seat of the Pandyan Kingdom. This would partly explain the traditional attachment of people from the Madurai region to the Sabarimala pilgrimage. "The folk songs concerning his [Ayyappa's] trip to the Pandyan Kingdom describe his journey [from Pandalam to Madurai] through dense forests and mountains" (Vaidyanathan, 1978).

Table 5: District-wise geographical origin of Tamil pilgrims

District Village/Town*	No.	District Village/Town*	No.	District Village/Town*	No.
Ariyalur (0)		*Kanniyakumari (0)*		*Sivaganga (10)*	
Chennai (11)		*Karaikal (1)*		Karaikudi*	9
Chennai*	11	Karaikal*	1	Tiruppattur*	1
Coimbatore (17)		*Karur (18)*		*Thanjavur (0)*	
Coimbatore*	10	Karur*	18	*The Nilgiris (0)*	
Panappatti	1	*Madurai (58)*		*Theni (21)*	
Sivanmalai	1	Kodangipatti	2	Bodhi Hill North	5
Tiruppur*	5	Madurai*	21	Kambam*	5
Cuddalore (33)		Melur*	4	Myaladumparai	8
Alappakkam	1	Peraiyur*	9	Silamalai	1
Avatti	1	Sholavandan	22	Thevaram	1
Cuddalore*	22	*Nagappattinam (3)*		Uttamapalaiyam	1
Keerapalayam	3	Palakurichi	3	*Tiruchirapalli (0)*	
Padirikuppam	5	*Namakkal (2)*		*Tirunelveli (6)*	
Tanur	1	Neikkarapatti	2	Sankarankovil*	6
Dharmapuri (3)		*Perumbaloor (1)*		*Tiruvallar (1)*	
Gobinathampatti	3	Tittagudi	1	Tiruvallur*	1
Dindigul (22)		*Pondicherry (34)*		*Tiruvanamalai (0)*	
Dindigul*	10	Pondicherry*	34	*Toothukudi (1)*	
Kurumbapatti	9	*Pudukkottai (4)*		Sindalakkarai	1
Palani*	3	Arantangi*	1	*Vellore (4)*	
Erode (1)		Kallakurichi	3	Guduivattam*	4
Kondalam	1	Ramanathapuram (0)		*Villupuram (1)*	
Kancheepuram (1)		*Salem (3)*		Nainarpalayam	1
Salur	1	Mahudanchavadi	3	*Virudhunagar (4)*	
Kanniyakumari (0)				Arupukkottai*	2
				Virudhunagar*	2

not test the gravitational constraint according to which the number of pilgrims would be inversely proportional to their distance from Sabarimala or from Kerala. There are certainly effects of distance that play on the motivation of the pilgrim movement, but the sample of the population retained for the survey may not representative from a statistical point of view. It can be representative within each linguistic group (150-250 individuals questioned), but the proportions of each in terms of participation or of number should not be confused with the size of the sample and its representativeness.

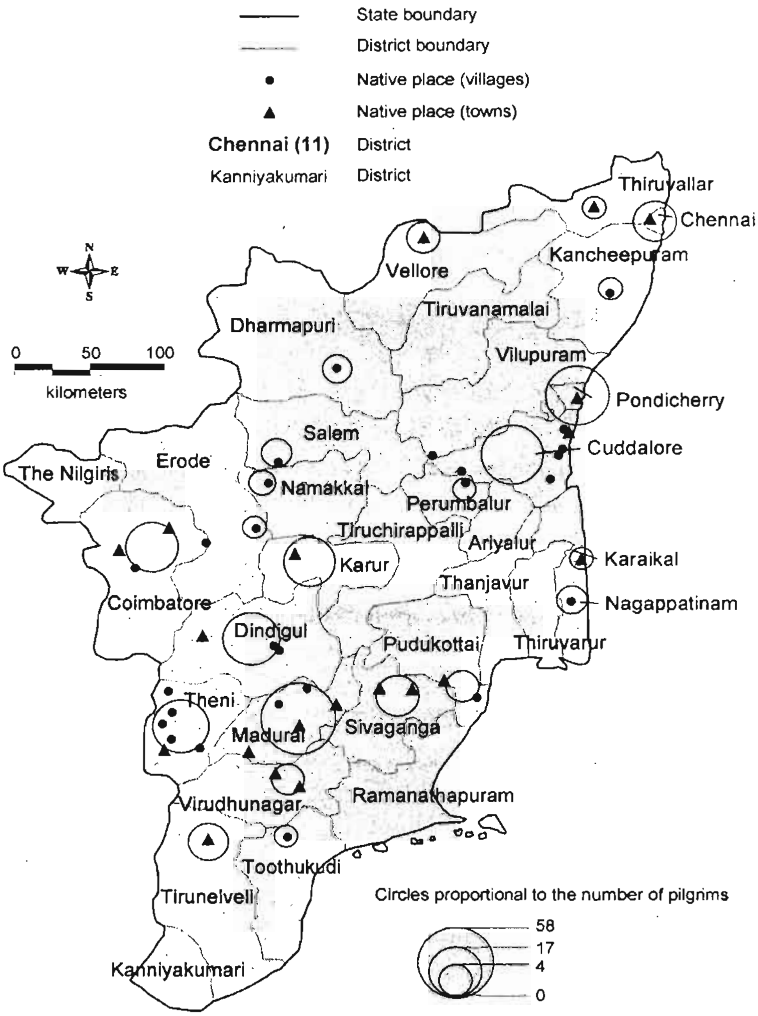


Figure 15: Geographical origin of Sabarimala pilgrims from Tamil Nadu

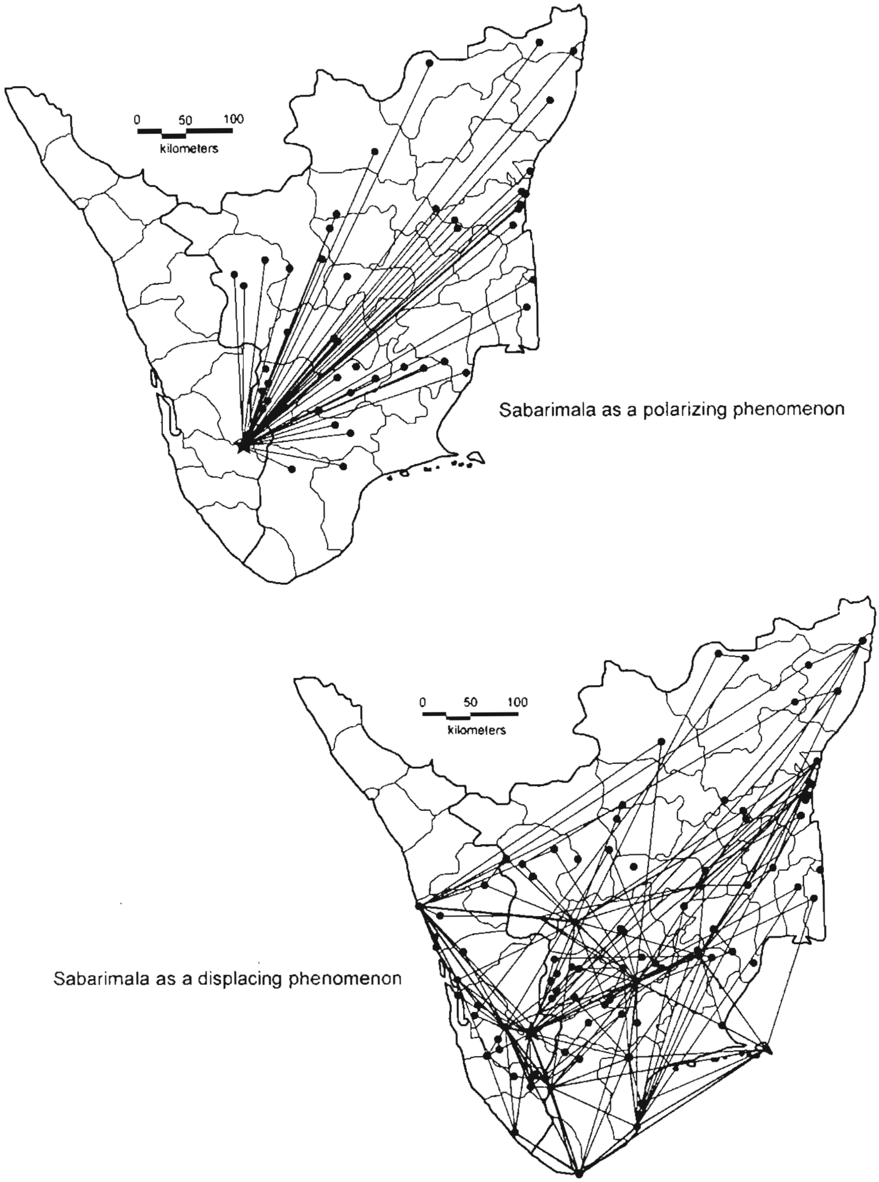


Figure 16: Two graphical representations of pilgrim flows to Sabarimala

For Figure 16, we have opted for a representation in segments or straight lines between two points so as to visualize, in the first case, the geographic audience of the pilgrimage and, in the second, the diversity and

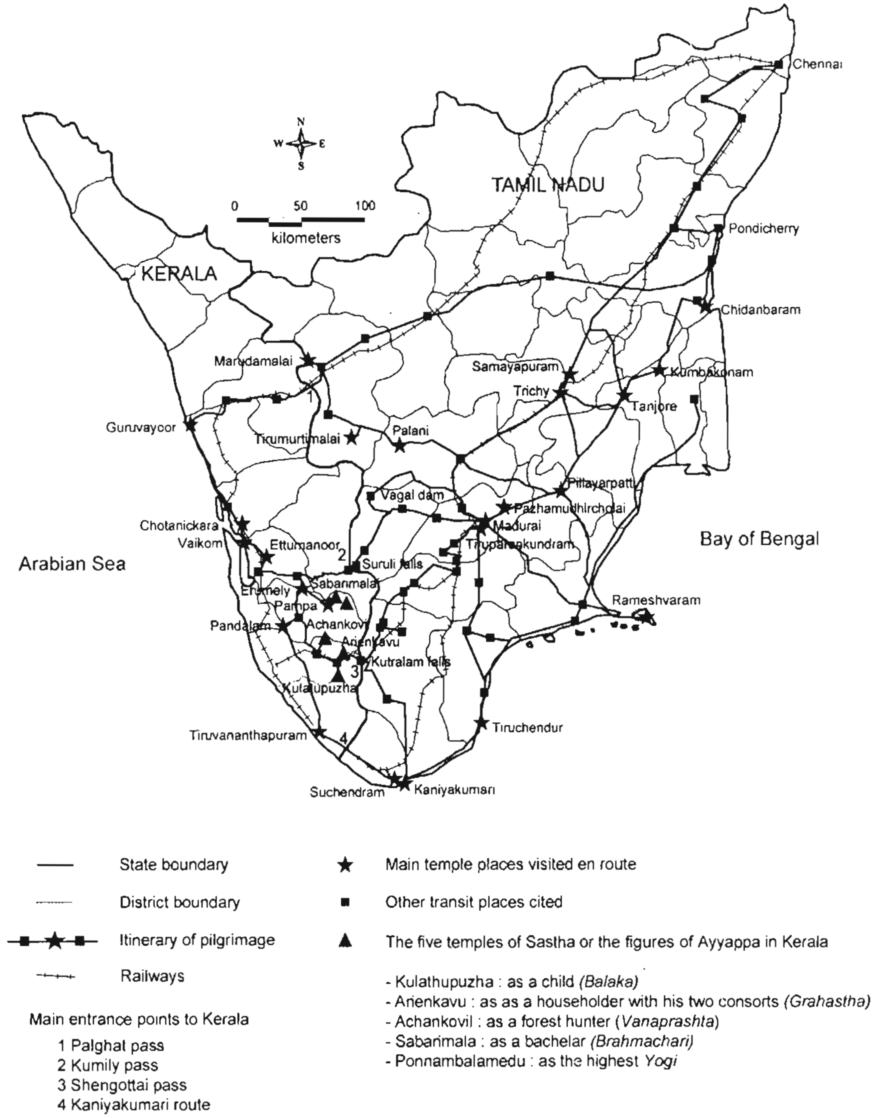


Figure 17: Main routes leading to Sabarimala

overlapping of the itineraries between Kerala and Tamil Nadu. While the first of the two maps exhibits the polarizing nature of Sabarimala, the second reveals the places frequented by the Tamil pilgrims beyond their borders (cross-border pilgrimage). However, this "aerial" (crow's flight) representation does not explain the high historical and symbolic charge of the pilgrimage routes in South India. For this reason, we have relied in Figure 17 on the physical networks of communications and transport to draw the pilgrimage itineraries and to elicit meaning from them.

Conclusions and discussion: pilgrimage and circulation in South India

On the basis of this attempt to graphically represent the pilgrimage itineraries between Tamil Nadu and Kerala, we can draw initial conclusions and put forward new working hypotheses to measure the magnitude of the Sabarimala phenomenon in South India.

First of all, it appears that the reticular geographic formation of pilgrimage obeys logics of original mobility beyond the most common logics governing regional migratory exchanges with Kerala[28]. On the other hand, the transport infrastructure plays a crucial role, as is known in the case of the train: "The first method of mass land transportation to be introduced in South Asia began operation in 1853 and quickly affected pilgrims and the temporal and spatial flows of pilgrim traffic" (Kerr, 2001: 326). The increased rapidity of movement, itself linked to the general improvement in the material transport conditions (notably the railways), certainly explains the standardization of the pilgrimage and the opening of Kerala in the contemporary period. It is, on the other hand, only an indirect cause of the renewal of attendance in the pilgrimage today; the grounds of the pilgrim logic are still, in our opinion, to be sought in a complex tangle of reasons and social and religious determinations.

Turning now to the final map of the pilgrimage itineraries (Figure 17), one notes first a historical continuity of certain routes leading to Sabarimala. The inland north-south axis of Tirupati-Tiruvanamalai-Madurai-Rameshwaram is greatly frequented by Tamil pilgrims before ascending to Sabarimala, either by way of Kaniyakumari in the south, or via the Shenkottai Pass to enter Kerala. The inland Tirupati-Rameshwaram axis belongs to an old pilgrimage route, that between Kashi and Rameshwaram, a historical continuity that is reinforced by a dense and solid network of rest-houses for pilgrims in Tamil Nadu (Deloche, 1993: 165). Other transversal

[28] For example, the logic of pilgrim mobility does not at all depend on the international emigration for work in the Gulf countries or the flows of tourists to Kerala, and the migratory networks of Tamil workers in Kerala (cf. tea plantations) remain confined to the adjoining districts.

axes of circulation, crossing the Western Ghats in the north by way of Palakkad (Madras-Coimbatore-Calicut), in the centre via Kumily (formerly through the forest) and in the south via the Shenkottai and Aramboli Passes, continue to structure the system of pilgrimage routes to Sabarimala.

Moreover, these routes or chains of places are significant from the moment certain of the links define the zones of attraction of the Tamil pilgrim population. These major nodes are for the most part temple towns such as Tiruvannamalai (Siva), Madurai (Meenakshi), Palani (Murugan), Rameshwaram (Siva), or Kaniyakumari (Devi Bhagavathi). On the outward as well as the return journey, the preferred pilgrimage itineraries of the Tamil population very often include the segments Madurai- Rameshwaram- Tiruchendur- Kaniyakumari- Suchendram- Kutralam Falls (before entering Kerala), or by way of the north, the Madurai-Palani-Marudamalai axis. In Kerala, the Tamil pilgrims visit almost exclusively the temples of Guruvayoor (Krishna), Chottanickara (Rajarajeshwari), Ettumanoor (Siva), as well as two temples of Sastha-Ayyappa, namely, Kulatupuzha and Arienkavu, in which the deity is represented in the former in his youthful form and in the latter as a householder.

In conclusion, the comparison of a local, regional (popular literature) and "scientific" production of maps on the pilgrimage is relevant in redefining the major axes of religious circulation between Kerala and Tamil Nadu. The analysis of popular maps reveals a clear gap between the representation and the reality of the Sabarimala phenomenon. For example, Figure 13 depicts a space oriented toward Kerala, the centre of which is Sabari Hill, while Figure 16 clearly shows a zone of attraction of the Tamil population that has shifted to the east. Finally, from Figure 17 emerges still more complete information, revealing, on the one hand, the historical axes of circulation (previously identified by Jean Deloche, *ibid*) and, on the other, more original itineraries traversing new places with a recreational role, for example, the Tekkady tourist park and the waterfalls at Suruli not far from Kumily, or again the dam on the Vaigai river near Theni in the district of the same name.

Moreover, a historical perspective shows us the importance of these routes. As was observed in a recent work, "the totality of circulations occurring in a given society and their outcomes could be viewed as defining a 'circulatory regime' (...) [which] in its turn tends to shape society" (Markovits *et al.*, 2003: 3). The circulation occasioned by the pilgrimage and its intensification in the contemporary period without question exhibits effects of the uniformization of Hinduism, revising the religious field through social change. Gilles Tarabout has clearly pointed out the importance of these changes by referring to recent modes of mass communication that play the role of diffusion channels of a "general attitude of piety" (*ibid*: 141-144).

Situating in this way the analysis at the intersection of an ethnocentric mapping and a more Cartesian cartography, with its geometric forms and representations of space, our approach is meant to contribute to the discussion on the mutual effects of pilgrim circulation and social change in India, while promoting the production of geographic discourse on South India.

The geography of irrigation in South India[29]

The SIFP database also provides non-demographic information on village infrastructures and land use. We shall give an initial illustration of their utility by presenting a general cartographic analysis of irrigation in South India. It should be pointed out that the data concerning irrigation are rarely adequate from either a statistical or geographic point of view. In the first place, they are not the object of a systematic collection or publication in India. If they exist at the district level according to statistics by state, they are not at the same time available for all of India. The ICRISAT (International Crops Research Institute for the Semi-Arid Tropics) collects agricultural information district-wise, but does not avail of data for all the states. The CMIE (Centre for Monitoring the Indian Economy) also occasionally publishes irrigation data according to district, but series available are usually incomplete even at the district level.

On a smaller scale, the regional statistics offices often have information for taluks/tahsils or for development blocks (see, for example, Rawal, 2001). But the collection is seldom systematic and it is not possible to compile data for several states. For this reason, the data gathered on the occasion of the census prove to be of great value, for they provide not only

[29] This section was written by Christophe Z. Guilmoto. For a more systematic treatment, see notably Guilmoto (2002).

exhaustive information, which in itself is very rare, but moreover at a maximum level of disaggregation. The village irrigation data are, in fact, nowhere else collected, and for this reason the map presented in Figure 18 has an entirely original character. We have plotted here the percentage of irrigated lands in the cultivated area of the villages[30]. This analysis rests on the 10-km clusters for South India, using the same procedure as described before: clustering, kriging interpolation, and final contouring.

We are not going to analyze here the determinants of irrigation in South India. A systematic interpretation would require in particular a detailed comparison with the maps of hydrography, relief and soils, which is beyond the scope of this brief presentation. But the map that is shown here makes it possible to describe the contours of irrigation in South India with an uncommon precision. One recognizes the very high concentration by sub-regions and the geostatistical analysis (see below) shows, furthermore, the strong geographic homogeneity of irrigation over a short distance. The first dimension that shows through is the west-east gradient of irrigation, closely linked to the slope running from the southern Ghats to the coast: irrigation, notably in the large deltas, is oriented toward the eastern littoral. One will also observe, in the more arid regions of the Deccan Plateau (in the central zones of the states of Karnataka and Andhra Pradesh), the effect of large hydro-agricultural developments that have enabled of the establishment of large irrigated pockets, as for example around the Tungabhadra in central Karnataka. One also finds, in the middle of a plateau characterized by a semi-arid climate and a primarily rain-fed agriculture, small zones where irrigation can involve more than half the lands under cultivation, which represents an extremely high figure.

The highest values of degree of irrigation, at levels close to fully irrigated (more than 80%), are observed mainly in the rice-growing deltas of the Bay of Bengal, that is, at the mouths of the Kaveri, Krishna and Godavari Rivers, the sources of which are in the Ghats. One can also recognize several inland river basins, such as those of the Palar to the west of Chennai, or the mid-valleys of the Vaigai around Madurai in Tamil Nadu, or again, of the Godavari on the border between Andhra Pradesh and Maharashtra, owing to the Nizam Sagar.

---

[30] Olivia Aubriot, who drew the maps of the modes of irrigation in Tamil Nadu, made in 2000 the first analyses on the basis of the village database.

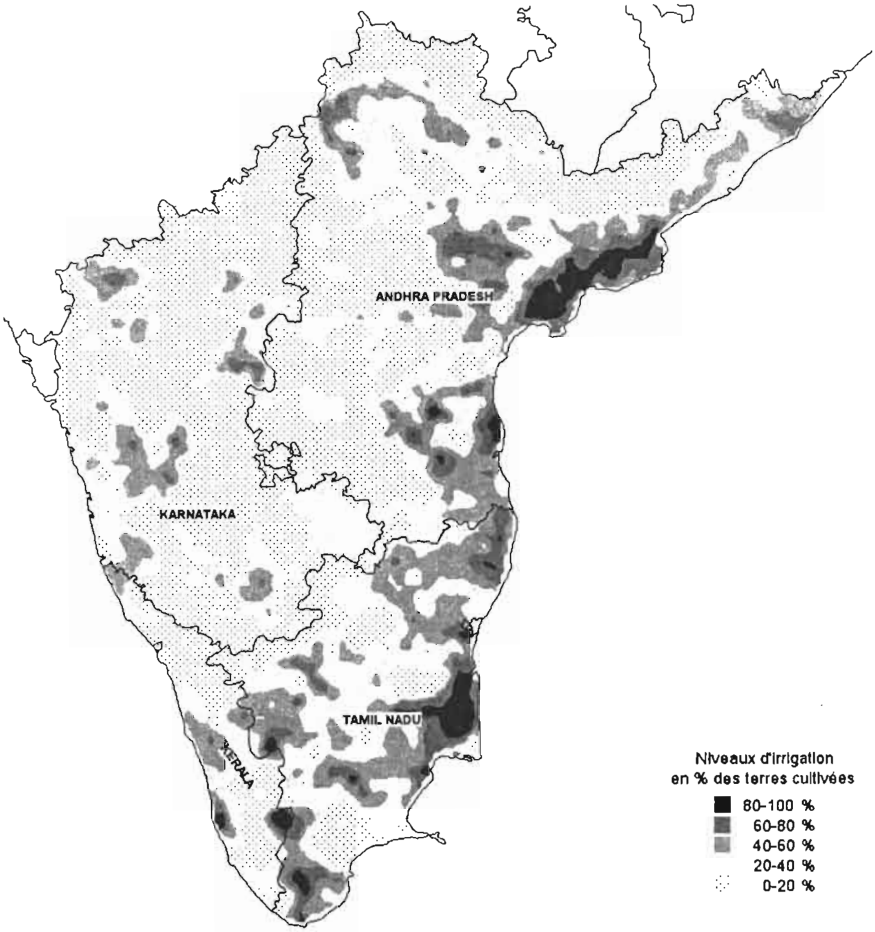


Figure 18: Irrigated land as percentage of cultivated areas in South India, 1991

This cartographic exercise is, however, also meaningful on a local scale, as for example in a given district: it highlights the fine differences to appear that exist between micro-regions, which are at times extremely pronounced because of physical constraints. The village data are not necessarily of exemplary quality, and this is all the more the case as we only avail here of the extent of irrigated lands, without knowing the real farming intensity enabled by the access to water, that is to say, the number of harvests per year. It is therefore often useful to aggregate the data to obtain a more reliable picture of the differentials between villages or groups of villages. This preliminary work would facilitate a subsequent analysis of irrigation according to each technique employed, for example, distinguishing the recourse to groundwater from that to river water, or showing the singular geography of the system of tanks that are scattered over South India and that are often the first source of irrigation in the driest regions deprived of accessible groundwater tables. Irrigation can also be related either statistically or cartographically to other data, in particular of a socio-economic nature, to study the potential impact of irrigation on the rural world (development of infrastructures, human development, composition of the population, etc.).

There are thus numerous paths to approach at the micro-level a phenomenon that is not known except through local monographic studies or aggregated analyses at the state level or for the country as a whole. Furthermore, one will have observed that the zones of irrigation appearing on the map of South India do not match the administrative boundaries, including those of districts that represent the smallest scale for which these statistics are sometimes published in India. This once again indicates that a spatial analysis of this kind judiciously reveals geographic units, at times sheer watersheds, which the official administrative division has a tendency to gloss over.

Measuring sexual discrimination in rural South India[31]

We wish to briefly describe how a study devoted to sexual discrimination can rely on a GIS. Discrimination against girls is seen to be more or less pronounced in India according to region or caste and religion, and the geographic dimension of the phenomenon is far from random. This discrimination can be broken up into numerous elements: educational, with clearly lower literacy rates for girls; residential, when forms of reclusion of women are encountered in certain regions of the North; economic or demographic in the case of negligence of girls in comparison to boys as regards allocated resources (health care, food, child development), etc. The aspect with which we are concerned in this study

[31] This section, written by Stéphanie Vella, lies in the framework of a doctoral work entitled "Spatialization of sexual discrimination in Tamil Nadu" conducted at Bordeaux University under the supervision of Professor Singaravélou.

represents two of the extreme components of this discrimination: the selective abortion of female embryos and infanticide of girls. These forms of discrimination are currently expressed in India by the lower number of girls compared to boys (today 927 girls for 1000 boys of less than 7 years), with pronounced variations between regions and social groups.

The study of these forms of discrimination was based in particular on the resources of the SIFP database. In effect, to approach all the aspects of this problematic, the fieldwork[32] was combined with a statistical and cartographic approach, for which work on the spatialized database was seen to be indispensable. From this database, numerous variables were employed, but most particularly sex ratio calculations: the general sex ratio of the population and, above all, the child sex ratio as defined below.

In this section, we shall show how a database derived from the census, having been spatialized in a GIS, makes it possible to analyze the phenomena of discrimination on several levels. We shall first examine the variable employed, the child sex ratio, and its significance in population geography and shall then interpret a number of maps that describe the spatial dimensions of sexual discrimination in South India and Tamil Nadu.

From population geography to child sex ratio

Population geography attempts to provide a spatial view of the facts of population; it is concerned with the spatial distribution of populations and that of their structures, general characteristics and evolutions. Thus, should one be concerned with the structure of a population according to a strictly demographic criterion, the distribution by sex is most often studied in association with the distribution by age (age pyramids). Although the sex distribution is of interest in itself, it is an aspect that is generally little studied by population geography in comparison to studies devoted to population numbers, fertility, migrations or mortality. Among the data currently employed to quantify sexual discrimination, the sex ratio is a good reflection of the status of the woman in a society. The sex ratio is, in fact, a crucial variable in the determination of a possible inequality between men and women according to the groups considered (population by age, urban,

[32] This study is based in part on a field work of six months in which surveys were conducted in a rural area in a group of hamlets comprising several castes in the state of Tamil Nadu. The objective of the questionnaires addressed to the villagers was to make a general appraisal of sexual discrimination on a day-to-day basis, from conception to adult age.

literacy, etc.). An imbalance in this ratio can express recent or past events (war, epidemic, etc.), as we shall see in the case of our study of sexual discrimination. Following normal practice, the sex ratio will be calculated for the Indian data as the number of women for 1000 men, and this choice facilitates the analysis of the deficits of women observed in India. On a smaller spatial scale, the sex ratio of the population can reveal certain socio-spatial practices. Distribution by sex is, moreover, one of the factors that in its turn influences demographic behaviours, for the imbalance of the sexes is nearly always expressed by a lower birth rate.

Composition by sex depends on the sex ratio at birth and on the sex differentials in mortality and migration. The migration factor, above all of an economic character, is responsible for a significant part of the observed variations in the sex ratio: for example, the preponderance of males among migrants to urban areas results n skewed sex ratios being observed in most Indian cities. On the other hand, the main factor used in this presentation, the child sex ratio (CSR)[33], is not susceptible to migratory mechanisms, which are very limited in these age classes, and consequently represents a valuable indicator of the situation of girls. It directly influences the total sex ratio, which has been extensively studied in India because of its imbalances (Visaria, 1999).

In 2001, the total sex ratio in India reached 107 men for 100 women. In most developing countries, one observes a slight numerical dominance of the male sex: the ratio is 109 men for 100 women in Melanesia, 107 in South Asia, 104 in China, 103 in the Middle East and 101 in North Africa and in Central America. But it is less than 100, and thus in favour of women, in South-East Asia and in Central and East Africa. In the developed countries, the women are more numerous. In Europe, the total sex ratio ranges from between 95 and 98 men for 100 women (Noin, 1996).

Indirect data for a sensitive topic

The SIFP database makes it possible to multiply the points of view for approaching a "sensitive topic" (Lee, 1993), so that the survey work would not be the only support and so as to be able to draw up a spatial model on different scales. These scales correspond to India, its states, South India, but also to districts, taluks and villages. Two types of geographic population

[33] The CSR is here computed as the number of girls for 1000 boys aged 0-6 years.

targets were therefore utilized, the second being included in the first: the total population of which a census was taken and the sample population in the villages studied. In fact, a difficulty in approaching a sensitive topic, concerning illegal practices such as foeticide or infanticide, is to elaborate methods for assessing the frequency of events that are hidden and go unreported. The official statistics under-estimate or ignore them. The problem is therefore to have an idea of the volume of concealed phenomena. Thus, to come to know even approximate figures pertaining to infanticides and abortions, as well as the utilization of new technologies of reproduction, proves to be nearly impossible in the field. Even the attempt to obtain governmental hospital data poses problems, for these figures are taboo. In addition, the cases of infanticide remain practically unreported by the police because very few complaints are registered; judicial statistics alone are therefore of little help. When the raw figures exist at the civil registration, they are very much lower than in reality[34]. In fact, village nurses can record infanticides as deaths due to "social causes", but their superiors frequently prefer to report them as "natural death". To conclude, let us point out that the number of many abortions, practised in private clinics or by traditional practitioners, is also unknown.

The computation of sex ratios at birth or the antenatal or post-natal mortality figures provide indirect indications as to the intensity of phenomena of discrimination. Similarly, in our case, the child sex ratio will serve as reference to indirectly quantify these two forms of discrimination. The figures from the census are therefore invaluable; they are calculated on the basis of distribution by sex available for all villages (and larger units) since the 1991 census[35]. The SIFP database has the additional advantage of being computerized and spatialized and thus enables a systematic utilization. We can thus search the correlations between the degree of development of the villages and sexual discrimination. Moreover, we can draw maps and establish a model of this spatialization, first statistically and then cartographically, alternating between a global perspective that brings out the regional contrasts and village variations.

---

[34] The study of the statistics of the civil registration would warrant a separate examination that will not be developed here.

[35] The census figures for 2001 are not yet available for the village level. Other figures exist for districts or taluks since 1871.

Variations in the child sex ratio

The imbalance of the child sex ratio, as well as the mortality sex ratio[36] and the sex ratio at birth, have in fact revealed for numerous years specific discriminatory socio-cultural practices anchored in the context of the Indian patriarchy, namely, the infanticide of girls and the selective abortion of female embryos[37]. The sex ratio of the Indian population has been diminishing with near regularity since 1901, as well as the ratio of the child population, notably in certain regions (the Punjab, Haryana, Rajasthan, and Tamil Nadu). These behaviours were statistically registered, but were difficult to show at the time of the first censuses; the excess in the number of men was often initially explained by the under-registration of women as

Table 6: Sex ratio below 7 in several States, 1991-2001

States	Sex ratio below 7 1991	Sex ratio below 7 2001	Change in %
India	945	927	-1,9
Punjab	875	793	-9,4
Haryana	879	820	-6,7
Rajasthan	916	909	-0,8
Uttar Pradesh	927	916	-1,2
Gujarat	928	878	-5,4
Madhya Pradesh	941	929	-1,3
Maharashtra	946	917	-3,1
Tamil Nadu	948	939	-0,9
Himachal Pradesh	951	897	-5,4
Kerala	958	963	+0,5
Karnataka	960	949	-1,1
Andhra Pradesh	975	964	-1,1

Source: Census of India (1991, 2001)

[36] The index is computed as the ratio of female to male mortality. It can be calculated on child mortality (number of deaths of children below one year on the number of births in the same year).

[37] Through the bias of the female child ratio, on the other hand, it is difficult to differentiate between infanticide and sex-selective abortion. One can say that until the years 1975 in the North and 1985 in the South, the child sex ratio did not depend of sex-selective abortion, but after these dates, both have an influence, with selective abortion gradually taking precedence over the infanticide of girls and excess female mortality during childhood.

compared to men. The government and a number of researchers continue to put forward the under-registration of girls at the time of censuses to explain their lower number, but these cases of under-declaration cannot explain the magnitude of such deficits. They can, on the other hand, entail differing results between surveys such as the NFHS and the census (IIPS, 1995; Guillot, 2002).

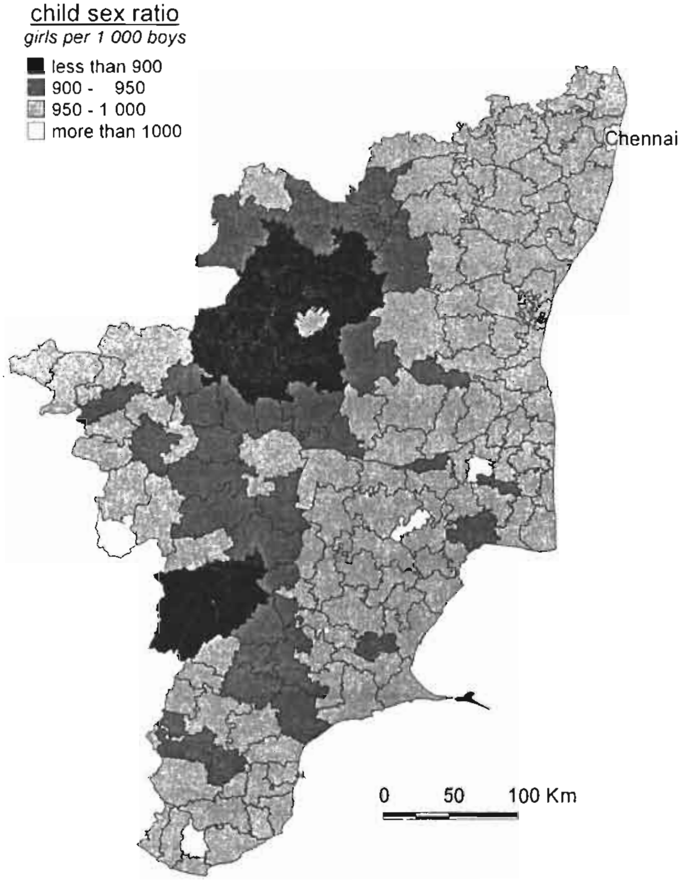


Figure 19: Child sex ratio (below 7), Tamil Nadu taluks (rural data), 1991

The child sex ratio has been regularly decreasing for several decades and the provisional results of the last census of 2001 show (Table 6) an almost general decline of ratios in the major states, except in Kerala. The Punjab and Haryana, in the north-west, are the two states having the lowest sex ratios and the most pronounced decline. Concerning this considerable fall in the CSR in 2001, the census office observed for the first time that it could be attributed to the recent use of sex determination tests, infanticide being in fact seldom mentioned by the government.

This research is conducted in South India for various reasons: sex discrimination is there less studied than in the North and the relative homogeneity of the states in the South seems to confirm the traditional North-South dichotomy. However, one can now show that heterogeneities, on smaller scales, exist between the four states and that Tamil Nadu presents specificities as regards sexual discrimination. The objective of this research is to provide preliminary elements to explain the mechanisms of sexual discrimination in specific regions of South India.

To analyze the demographic data, spatial analysis is thus seen to be relevant for presenting the results, making possible both an easy inter-state comparison and a study of variations in the child sex ratio on a more reduced scale within the states. The maps drawn by the Indian government are limited to the district scale[38], while the GIS of the SIFP allows the very necessary mapping of the data at the village level. The maps thus prepared are on a much more reduced scale and more precise than what a map by district or taluk can offer, which can be verified by comparing a map of taluks (Figure 19) with that of villages (Figure 21).

Child sex ratio in South India, 1991

The map in Figure 20 represents the child sex ratio in rural South India, in 1991. Before entering into an interpretation, we will point out that it was prepared according to the method described earlier: the villages are aggregated by clusters of 10 km so as to offer a regular grid with units that are sufficiently populated for a significant calculation of the child sex ratio. A spatial interpolation was done with the CSR data brought together by means of ordinary kriging (Chou, 1997). Then, the contouring of the homogeneous statistical regions was done. The value 900 was retained as threshold of abnormality of the CSR, so as to favour the clarity of the map, whereas some authors place it at 930 or 950 girls for 1000 boys.

[38] It should however be noted that the serious worsening of the child sex ratio in 2001 in the Punjab and in Haryana recently led the Census of India to publish separately a very unusual serie of maps of the sex ratio: maps of Indian districts, of the taluks in the Punjab, and even a map of villages in the district of Firozpur (maps published in mid-2003).

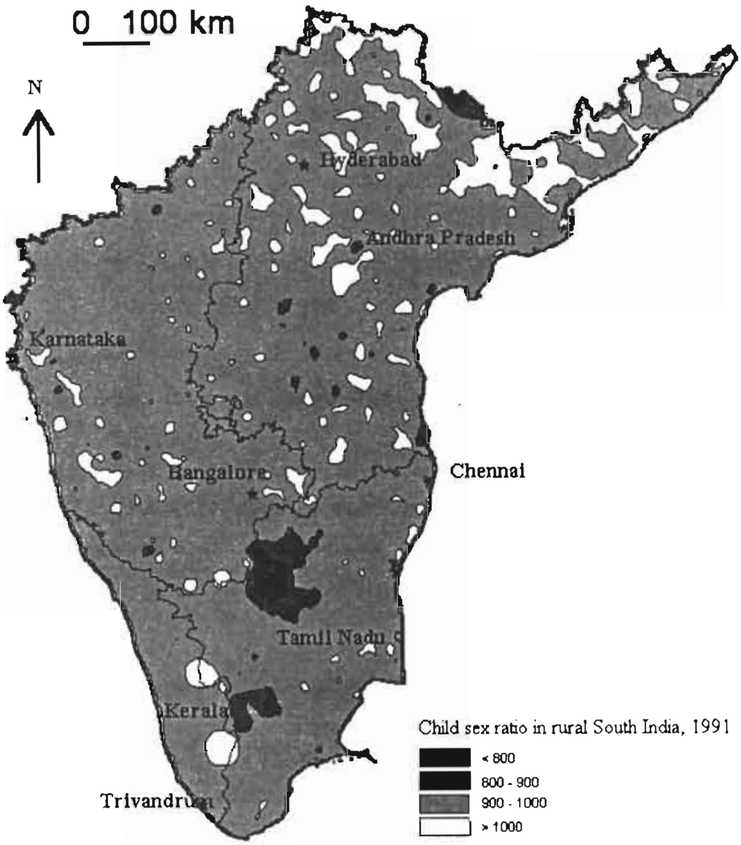


Figure 20: Child sex ratio (below 7) in rural South India, 1991

The map makes it possible to disaggregate large regional units and to reveal with greater precision the micro-regional contours. Here, it measures the intensity of sexual discrimination in rural South India. Globally, the rural CSR lies between 900 and 1000, but appreciable geographic variations appear upon reading this map. One will note first that few regions have a CSR above 1000. Among these regions, we may identify several tribal tracts bordering northern Andhra Pradesh, but also of isolated pockets along the western and eastern Ghats. In Andhra Pradesh, a few micro-regions with ratios disadvantageous for girls stand out, one of which is a zone in the north. In Karnataka, apart from the frontier zone adjacent to the region of

Salem in Tamil Nadu, the zones with low CSR are very isolated. In Kerala, it appears that the rural CSR values follow an average level, with two zones very advantageous to girls in two districts along Tamil Nadu.

This cartography brings out above all the particularism of Tamil Nadu, for it is there that two regions of sexual discrimination, of significant size, are identified in the most pronounced manner. Apparently, the gender imbalance is locally significant in Tamil Nadu and follows a precise geographic contour, which we shall now examine in greater detail.

Child sex ratio in Tamil Nadu, 1991

The spatial modelling of the child sex ratio at the level of Tamil Nadu in 1991 (Figure 21) reveals the large disparities within the state. In most of the state, the CSR is above 950 and thus corresponds to the normal values of mortality according to sex. But the situation presents itself quite differently in certain seemingly isolated pockets, where the CSR is very much lower than elsewhere. This is the case to the west of Madurai and in the pocket of Salem-Dharmapuri. One again notes the high spatial compactness of the phenomenon, which is not the result of the geostatistical smoothing carried out on the basis of spatial data on a small scale. The proportion of girls is the lowest in the north-west and the highest in the south and north-east of Tamil Nadu. Thus, the thirteen taluks having extremely low CSR values in 1991 belong to the then districts of Salem, Dharmapuri and Madurai.

Finally, one will notice that the magnitude of girl-boy imbalance is in several areas considerable, for values beneath 660 are observed in the Salem area. These are regions where there appears to be a deficit of one girl in three in 1991. At a more local level in this zone, there are numerous villages with more than 2000 inhabitants where the CSR is two boys for one girl. We have here without doubt the absolute peak in India of discrimination against young girls. Furthermore, the block that records the lowest CSR in Tamil Nadu (614) is a rural block in the district of Salem. The as yet incomplete results from 2001 do not contradict our observations, for one observes in the rural parts of the taluk of Omalur (Salem district) a CSR of 549 girls for 1000 boys. Even though very isolated, this value is much lower than those observed in all the tahsils in the Punjab or in Haryana.

The variations are very pronounced in the Kongu Nadu region that encompasses the districts of Coimbatore, Erode and Salem, whereas this area presents a certain economic and cultural homogeneity: agriculture is closely linked to industry and it is a rather prosperous region. The birth rates can be very low there, the town-country links are close in relation to the rest of the

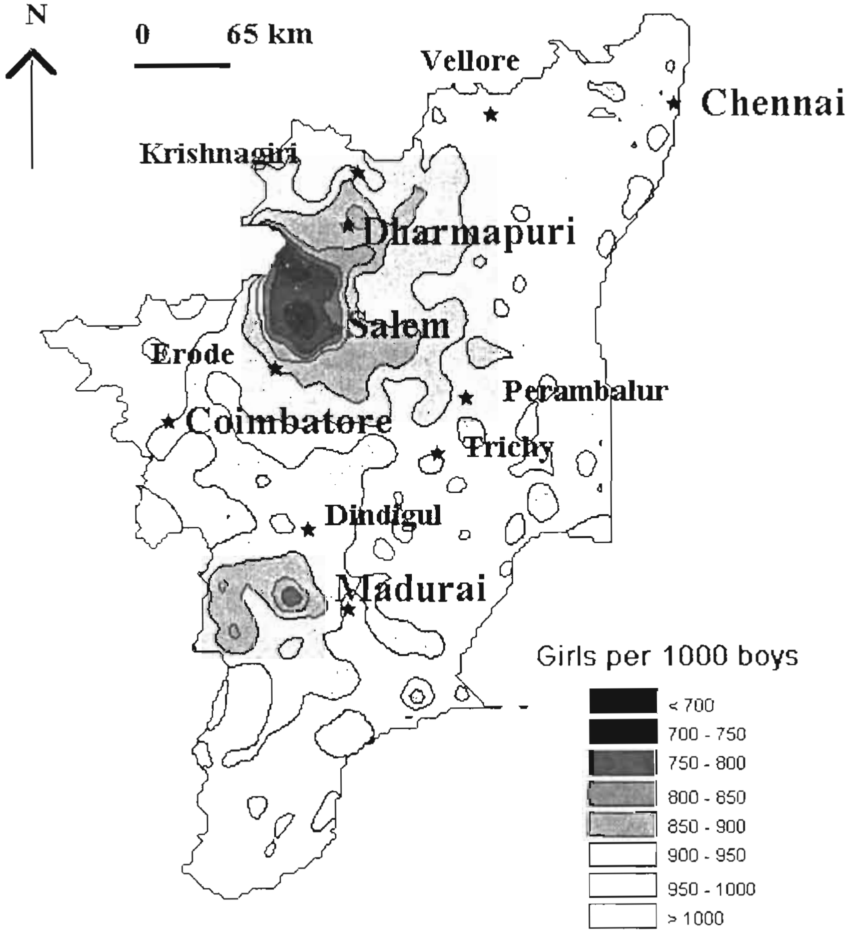


Figure 21: Child sex ratio (below 7) in rural Tamil Nadu, 1991

state and the intensity of irrigation there is also very high. Thus, following the major Coimbatore-Bangalore road that traverses this zone from south to north, one begins in Coimbatore, a very prosperous region where the CSR is at a normal level (950-1000). The road crosses the plateau of Kongu Nadu and the ratio remains within the average values (900-1000). But before crossing the Kaveri, a sudden change intervenes and the CSR declines strongly, moving from 950 to less than 830. Very quickly the absolute minimum of 614 is reached: in less than 50 km, the CSR has decreased by 50%. After the city of Salem, the CSR increases noticeably to return to 800 in Dharmapuri, then 950 in Krishnagiri more to the north and again attains to median values approaching Karnataka. It is difficult to think that this is actually a homogeneous zone, in particular as regards population, in view of the appearance of extreme variations in behaviour concerning sexual discrimination.

While the Salem zone is the most highly pronounced in terms of the deficit of girls, reinforced by that in Dharmapuri, this deficit appears to be much less extensive and of much lower intensity in the more southern region centred on Usilampatti (Madurai district). It seems to be connected to the preceding by zones in which the CSR lies between 900 and 950. Since the 1991 census, the phenomenon has progressed, an expansion to additional taluks has taken place and some authors spoke, in 1996, of a "belt" to spatialize the discrimination of girls (Athreya and Chunkath, 1998).

Thus the zoning of infanticide[39] begins in the western half of the district of Madurai and extends across the districts of Dindigul, Karur, Salem and Dharmapuri to the west of Vellore district. Salem and Usilampatti seem to function as secondary independent loci, after Madurai. In fact, even though Salem, Dharmapuri and Madurai are always cited as regards infanticide, on a smaller scale, the geographic distribution becomes thinner; one observes that at least twice as many of the districts are implicated in infanticides. On the other hand, the districts in the south, east and in the Kaveri delta seem to be spared from this phenomenon and no anomaly was discerned in these zones, if not, in 1999, in such districts as Perambalur and Tiruchirappalli that are somewhat peripheral to the Madurai-Salem axis. Another major exception concerns Coimbatore district; although it is in the west on the edge of the Usilampatti-Salem corridor, it is not at all affected by infanticide, according to the presently available data.

These maps therefore make it possible to bring to light very precise zones of sexual discrimination. We have presented only the maps of 1991,

[39] We speak of infanticide because the practice is old and goes back some fifty years in Tamil Nadu. In 1991, selective abortion was still infrequent in this state.

but the exploitation of the aggregated results from 2001 will enable us to conduct a more diachronic analysis and to study the spatial progression of child sex ratios from 1991 to 2001. These maps on the scale of South India and Tamil Nadu should confirm the socio-spatial diffusion of sexual discrimination. A parallel analysis is also being done on the scale of the whole of India, where the propagation of techniques for the detection of the sex of the foetus has caused an increase in sex-selective abortions.

Spatial distribution of medical infrastructure-Andhra Pradesh[40]

Devoted to maternal care in rural India, this research work began with the analysis of health care structures in the region of Rayalaseema, the southern part of Andhra Pradesh, through the information contained in the SIFP database and preceded an intensive survey in a specific subregion along the limits between Andhra Pradesh and Karnataka (Chasles 2001). The intended objective of this section is however more global: we will try to give here a regional representation of rural health facilities in Andhra Pradesh. To this purpose, cartography and geomatics (geographic information system) as well as statistics (principal component analysis, PCA) were called upon. As we shall see, the complementarity of these methods will allow us to shed light upon a health care gradient within the state, in addition to, and this is certainly the most important, providing elements for an understanding of the spatial logic of the distribution of health services.

Data and scale of analysis

The processing of data was carried out on the scale of the *mandal*. This administrative unit corresponds by and large to that of the *taluks* or *tahsils* of the other states, even though, as has been previously observed, the mandals are comparatively of a much smaller size. The 1099 mandals encompass the 26,686 villages of Andhra Pradesh and offer the advantage of being of relatively homogeneous size, in distinction to villages, whose population may vary in a ratio from 1 to 100. The use of clusters is thus not indispensable here.

[40] This section was written by Virginie Chasles in the framework of a doctoral research under the supervision of Alain Vaguet of the University of Rouen. See also Chasles (2002).

By means of a principal component analysis, we have examined the near-totality of variables relating to the available medical infrastructures in the census of 1991. We examine here the proportion of villages in each mandal possessing the following health facilities: hospital, maternal and child welfare centre, maternity home, child welfare centre, primary health centre, health centre, primary health sub-centre, dispensary, family planning centre, nursing home, community health workers, registered private practitioner and other medical practitioner. Should no infrastructure be present, the distance to the closest infrastructure has been included. Only two variables (other medical centres, tuberculosis clinic) had to be removed because of their rather dubious quality.

Principal component analysis is a mathematical procedure that transforms a number of correlated variables into a smaller number of uncorrelated variables (called principal components). This technique "allows [one] to extract from a dataset the largest amount of information in a simple and coherent format and helps to identify the interrelationships between variables and the gaps between geographical units under study" (Sanders, 1990: 17). In our analysis, the objective is to reduce the number of variables related to medical infrastructure that are found in the SIFP dataset. The first principal component accounts for as much of the variability in the data as possible: it serves therefore as a global indicator of the quality (quantity and diversity) of medical facilities found in villages of each taluk[41]. Hence, it is positively correlated with all variables listed above. This indicator mainly concerns the structures relating to the public sector. Different levels of the health hierarchy are here coincident, for one finds both centres having a range of relatively wide and elaborate services and, conversely, local centres delivering more preventive than curative care. From this plurality of listed centres there obviously result zones with diverse services, each of which is globally proportional to its "technical plateau". Thus, by way of example, a community health centre (CHC) corresponds theoretically to a population of one lakh, while the coverage for a sub-centre (which depends on a primary health centre) involves no more than 5000 persons (and less in tribal zones).

[41] Sébastien Oliveau uses a similar procedure to compute a "modernization index " for Tamil villages (see below).

Medical infrastructure in Andhra Pradesh

The principal component analysis (PCA) makes it possible to measure the medical service on the mandal level and to show certain characteristics of the location of medical structures. The variables have not been retained in a random manner, but because of their correlation to the first principal component. The result, and notably the first factor resulting of the PCA, is assessed by its eigenvalue that measures the quality of this factorization. In our analysis, the first factor that accounts for the quality of the medical infrastructure is endowed with a quite high eigenvalue as it amounts to 3.3 (the value of other eigenvalues is less than 1.4).

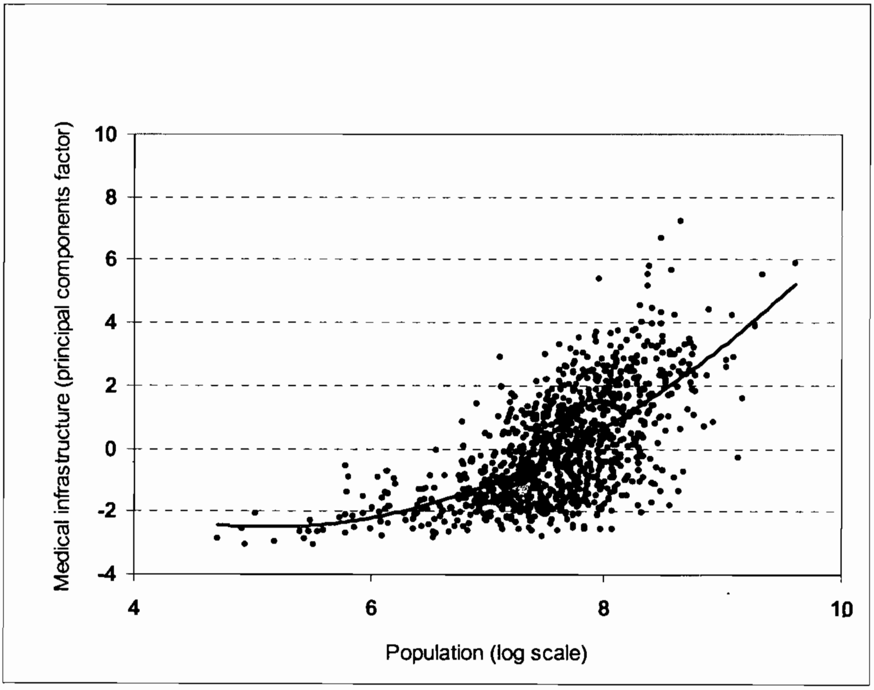


Figure 22: Medical infrastructure and population size
in Andhra Pradesh, 1991

This result was then compared with the average population of the villages. One observed a strong relation between the population size of the villages and the distribution of the health care infrastructures. This result is expressed by the scatter of points represented in Figure 22. The presence of the provision of health care facilities tends to slowly increase when the villages have less than 1000 inhabitants, to then rapidly increase and reach a new plateau near 10,000 inhabitants. It can be said generally that the number of health care centres is proportional to the average size of the population of the villages. In greater detail, one recognizes a log-linear curve that moves between the minimum level (no infrastructure) and the maximum level (all infrastructures), characteristic of mechanisms of diffusion.

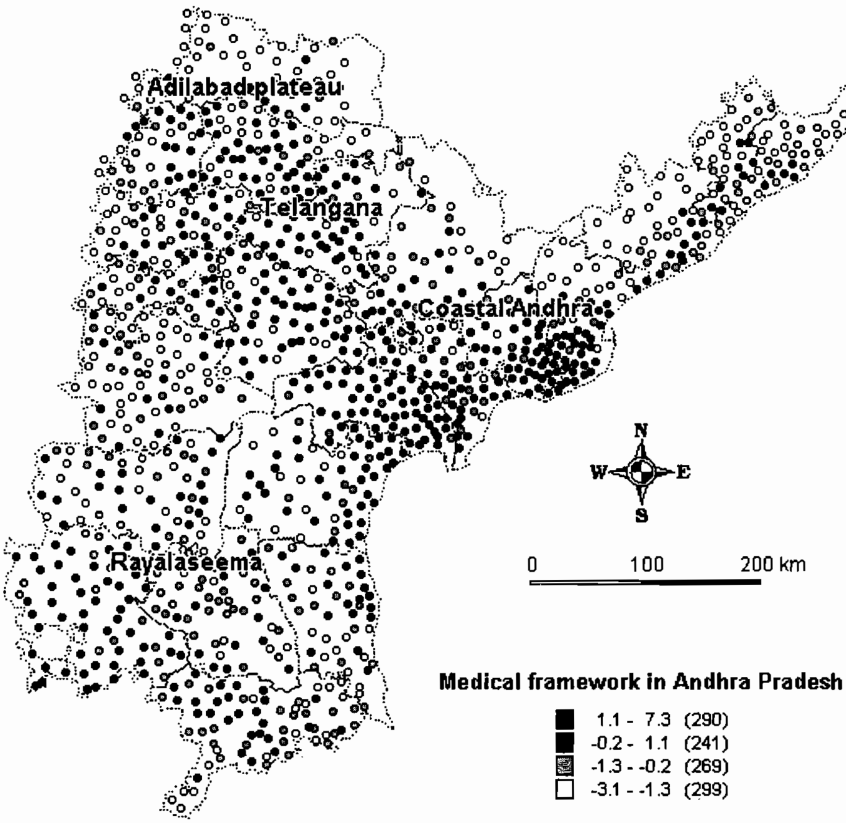


Figure 23: Medical infrastructure in Andhra Pradesh mandals, 1991

Mapping the first factor resulting of this PCA, one obtains in Figure 23 the map of the mandals classified according to the quality of their health infrastructure. This correlation sheds light on the different entities that make up Andhra Pradesh. More precisely, it can be noted that the health framework follows quite exactly the economic landscape of this state. In other words, the provision of health care is distributed according to a gradient proportional to the level of development and, in this case, to the population densities.

The most populated state in South India, Andhra Pradesh is characterized by real regional contrasts. Broadly speaking, the state includes two semi-arid climatic entities and a coastal area. Telengana extends to the north-east, while Rayalaseema includes the four districts in the south of this region. Finally, coastal Andhra corresponds to the deltaic region composed in particular of the Krishna and Godavari Rivers. On this map, two regions acquire distinctly individual characteristics through a high density of infrastructures compared to the rest of the state. The first region concerned is organized along the seafront and is characterized mainly by an exacerbated polarity around the main rivers mentioned above and corresponds overall to coastal Andhra and to the richest deltas of Andhra Pradesh, such as that of the Godavari. This sub-region took advantage of its geographic location through an optimal utilization of its territory. In particular, due to the efforts made to extend the percentage of irrigated lands, this region has ensured itself a relatively prosperous development in both the agricultural and industrial sectors. This has naturally entailed higher population densities than elsewhere and a real demand for medical care facilities. We are not examining whether these needs are completely met, but it appears that in this region the health care provision is the densest and probably the best.

To this first sub-set can be compared the region of Telengana, in which the state capital, Hyderabad, is located. In spite of a health system that seems less dense, the region is also looked on as being fortunate and distinguishes itself in this state. It behaves somewhat as a coastal extension, and this is above all attributable to the underlying dynamics of the two rivers flowing through it. Located on the Deccan Plateau, Telengana has an old urban and economic tradition. This explains why it is characterized by a relatively large provision of health facilities in comparison to the neighbouring regions. However, there are marked contrasts within Telengana itself. Thus, two somewhat marginalized zones are demarcated by their lack of infrastructures: the Aliabad Plateau in the north-west that is characterized by a severe economic backwardness, and south-west Telengana that suffers for its part from the aridity of its physical environment.

Rayalaseema is in a position that could be qualified as residual or even as intermediary. It should be said that it is characterized by the conjunction of several major handicaps: marked above all by a rainfall deficit, with an annual average of 691 mm, as opposed to 929 in Telengana and 1024 in coastal Andhra, this region is also characterized by very limited natural resources and a slow economic development. Hence, the weakness of its infrastructures, notably in the domain of health, would appear to be inevitable. But the regions that appear on this map as being the most underprivileged in Andhra Pradesh coincide with several mountainous and forest tribal zones in the north and centre of the country, where the population densities are very low and which avail of only minimal infrastructures.

Conclusion

The results obtained here unquestionably conform to the numerous theories pertaining to the location of services, in particular to those concerning local services, to which belong health care centres. Globally, since the 1960s (Merenne-Schoumaker, 1996: 39), a general dependence between the number and the nature of services and the size of the population has been shown. However, as regards this assertion, it has been pointed out that this relation is, more precisely, not continuous and that it thereby underscores the population thresholds: above, the infrastructure exists, below, it is absent. The infrastructures are organized along hierarchical lines, according to the size of the population necessary for their existence, and hence are established, in theory, according to the correspondence supply/demand, or supply/need. Let us also add that their distribution is also influenced by the socio-economic characteristics of the prospective users, in other words, by their solvency, which for us lies within the wider question of the economic dynamics of the regions studied. It must be observed that the health care framework in Andhra Pradesh is in line with these models, for, as we have been able to see, it is organized according to the size of the population that is itself, generally speaking, dependent on the level of economic development of the different spatial units studied. To conclude, let us note that one should not misunderstand the results presented here. They have revealed the main factors accounting for the distribution of health care facilities. They do not signify that the health care needs of the populations are really satisfied.

Urban spread and rural continuity. A quantitative approach[42]

The problem of modernization in a country like India constitutes a preferred domain of research in the social sciences. Our section intends to view modernization in its geographic dimension in order to go beyond the approach that consists in viewing development in distinct manners for town and country, and to establish a continuous framework of the space of development. We take here the example of Tamil Nadu.

A new approach to rural-urban linkages

The study of town-country relations is an old one in geography and represents a frequently used entry for the study of a region. This dual approach is characteristic of geography, and while it is found in other disciplines (in particular demography and economy), it is not utilized in such a systematic manner.

This approach has evolved and the simplistic dichotomy that it suggests has been emphatically called into question. Supplanting studies that would oppose towns to the surrounding countryside are studies that take the different rural spaces into account in a more precise manner. More than the milieu (rural or urban), it is now the space that is here considered in a more global dimension. The evolution of the vocabulary describing the different types of space according to their more or less urban or rural characteristics and their position in relation to the urban centre clearly reflects the general change of geographic approach. One now systematically distinguishes the town, its suburb and its peri-urban fringe from rural spaces. Countries such as France have refined the description of rural space by distinguishing different categories of spaces within the countryside: rural under weak urban influence, rural poles, periphery of the rural poles and finally the isolated rurality (Hilal and Schmitt, 1997). One can see in this evolution the necessity of adapting analysis to a space that has become more complex. Thus, the proportion of farmers, a traditional characteristic in the definition of rural space[43], has regressed to such an extent in Western societies that it no longer allows the distinction between town and country.

[42] This section was written by Sébastien Oliveau. This work is part of a doctoral research under the supervision of Prof. D. Pumain: "Village modernisation and distance to town in South India".

[43] This is still the case in India, as the census employs three characteristics to divide administrative units: the population, the density and the proportion of working males employed in agriculture.

This increasingly precise division reminds us of an essential element in the study of the relations between town and country: the dichotomy is artificial, the demarcation is never clear, and there exists a continuum from the rural to the urban. For this reason, some have decided to abandon this classification for a continuum analysis of rural space according to urban impact. Thus Kundu *et al.* (2002) have published a study showing the role played by the distance to town on different socio-economic variables. Their analysis, based on the statistics of a survey undertaken by the NCAER (National Council for Applied Economic Research), demonstrates quite clearly the interest of such an approach.

The authors show there how different economic (such as the per capita income or the size of landholding) and social (morbidity, literacy) variables decrease as they become more distant from urban centres. The method used consists in ordering the villages according to their distance from town, on the basis of the data provided by the census, then in examining how the value of the variables considered evolves.

The method is interesting, but the data employed bear a strong bias. The database of the NCAER is of good quality on the aggregated level of India, but using this base of 33,230 households spread over 1765 villages poses a considerable geographic problem: how to include in the same analysis Tamil villages, which are less than 30 km away from the closest town and which are well served by transport and communication, with villages in Rajasthan, which may be sometimes lost in the midst of the desert? The complete disregard of the regional dimension in this type of study is disturbing and creates a very significant bias. Thus, villages that are more than 30 km from a town will be less developed than the others, not because they are isolated, but because they belong to a less developed state. It should also be noted that the quality of the census concerning the distance to town, as we shall see, varies greatly.

To mitigate the limitations of this approach by sampling, the simplest method is the use of an exhaustive database and the systematic study of the different urban centres. This is what we have done by relying on the SIFP database, as presented in the first section of this paper[44].

---

[44] The database for Tamil Nadu is available in the form of a CD-ROM, see Guilmoto *et al.* (2000).

We thus wanted to study the urban influence on the countryside in Tamil Nadu, using the entire rural database for the state. To this purpose, a series of calculations was made using a GIS software in order to obtain the distance as the crow flies of each village from the nearest town. This enabled us to calculate the coefficients of correlation between the distance to the town and different variables, as Kundu *et al.* have done. The major results of these correlations are presented in Table 7. The role of the distance to town on the level is seen. Our findings broadly confirm the results of Kundu *et al.*

Table 7: Correlation between city size (logarithm) and various sociodemographic indicators

Density	Household size	Literacy	Workers in the agricultural sector	Workers in the service sector	Workers in household industry	Workers in non-household industry
0.4616	0.164	0.2805	-0.4486	0.4391	-0.1806	0.1327

Sources: S. Oliveau, based on SIFP database

Nevertheless, a comparison of the correlations between the distance to town and one of the variables common to our work shows slightly different results (see Table 8). The relation between literacy and distance is negative in all cases: the villages closest to towns have average levels of education that are considerably higher than the other villages.

Table 8: Coefficients of correlation between distance to town and literacy and household size. Comparison of different findings

Distance	Census estimate	Census estimate	GIS-computed
Correlation between literacy size and distance to the nearest town.	-0,1789	-0,1545	-0,2883
Correlation between household size and distance to the nearest town.	-0,1141	Not significant	0,0762
Sources	*Kundu et al. (2002)*	*Oliveau*	

The differences in the two series of results are no doubt connected, on the one hand, with the differences between the sources and the regions concerned, but also with the use of the distance calculated with our GIS. One will also note that our results for Tamil Nadu appear clearly more significant when the distance geometrically computed by the software is used[45], rather than the information from the census. This suggests that the census figures could be of rather poor quality.

Large-scale analysis

The following procedure, which consists in studying separately the effect of the distance to town on different indices, had, however, a detrimental effect: the synthesizing view provided by the study of distance is dissipated in the study of multiple indices. It is therefore preferable to construct a global index collecting the census data, and we have taken recourse to a factor analysis (principal component analysis) to construct an index of modernization, as Virginie Chasles has done regarding the medical infrastructure in Andhra Pradesh. The calculations carried out (not reproduced here) show that one obtains a composite index that is correlated to very numerous socio-economic characteristics. This modernization index corresponds to villages in which literacy is good and fertility is low. It is also correlated with a pronounced presence of industrial activities, but still more with tertiary activities, at the expense of the primary sector. In addition, the proportion of working population, characteristic of an agricultural setting, is low. The agricultural sector is there characterized by a weak presence of cultivators, which is connected with a more intensive irrigated agriculture and with a high demand for labourers[46]. These villages are rather well provided with educational, medical and transport infrastructures.

We then carried out a regression of our index according to the distance to town. The results are, as one can expect, significant, for more than 11% of the variance is thereby explained. The connection is highly negative between the level of modernization of the village and the distance to the nearest town.

---

[45] It is generally assumed that the distance as the crow flies usually provides a good approximation of the real distance on the scale of our study (see Berroir, 1998).

[46] In Tamil Nadu, intensive crops such as rice require a large labour force and are generally characterized by a high rate of agricultural labourers in comparison to that of cultivators. Conversely, the strong presence of cultivators is usually a sign of less intensive and thus generally poorer agriculture.

This underscores the pertinence of the distinction between different rural sub-spaces according to their distance to town.

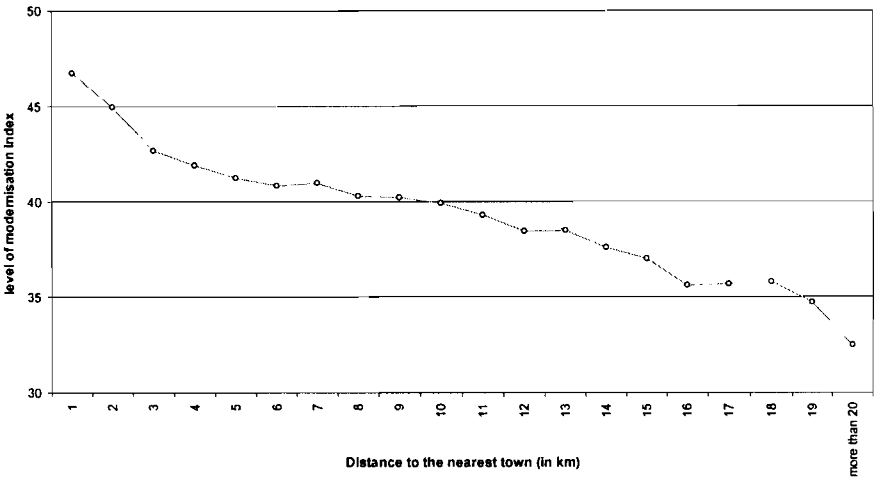


Figure 24: Modernization index and distance to the nearest town

To complete this result, we calculated the value of our index for each distance (measured here in kilometres) and these values are plotted on Figure 24. They show a decrease in the level of modernization of the villages according to their distance from towns. The first conclusion is that already proffered by Kundu *et al*: "the spatial distribution of the indicators [...] does not decline smoothly, as we move from the city/town periphery to distant areas". Thus, the decline of the index is first rapid, following an exponential curve, before stabilizing, following a more linear decline.

One can also measure the spatial impact of towns. For this, the point where the curve breaks off is considered as marking the end of the direct urban influence. According to our graph, it is situated at around 4 to 5 km (Figure 24). Another model consists in defining the limits of urban influence at the point where the average value of the index goes below the mean of the index for the entire set of villages (which is 0.014). The urban impact is thus of the order of 9 km. But the absence of a plateau at the end of the curve or of an upward turn of the latter leads us to consider that the urban influence, although diminishing with distance, has an impact on all villages, even the most distant.

Table 9: Characteristics of towns and their hinterland according to their size class

	Class 1	Class 2	Class 3	Class 4-6
Number of towns in this class	31	39	64	91
Average number of inhabitants	441 800	70 000	31 100	12 400
Number of villages, by size class of their nearest town	2839	3413	4910	4660
Average level of modernization of villages, by size class of their nearest town	0.52	-0.18	-0.22	0

Sources: S. Oliveau, based on SIFP database

This global approach can then be completed by global analyses that distinguish the towns among themselves. The first general criterion retained to classify the towns among themselves is their demographic weight. In fact, the size of the population of a town generally summarizes very well its socio-economic characteristics (Moriconi-Ebrard, 1994). The calculation of the correlation between certain indicators and the logarithms of the population of towns in Tamil Nadu confirms this (see Table 7). We therefore decided to separate the towns according to their size by following the usual Indian classification (see Ramachandran, 1989). One will distinguish here towns of class I (more than 100,000 inhabitants), class II (50,000 to 100,000 inhabitants), class III (20,000 to 50,000 inhabitants) and the set of classes IV, V and VI (less than 20,000 inhabitants). For each class of town (see their characteristics on table 9), we have recomputed the effect of the distance on the degree of modernization of the villages. Different curves are obtained, each corresponding to the effect of the towns classified according to their population; the results are shown in Figure 25.

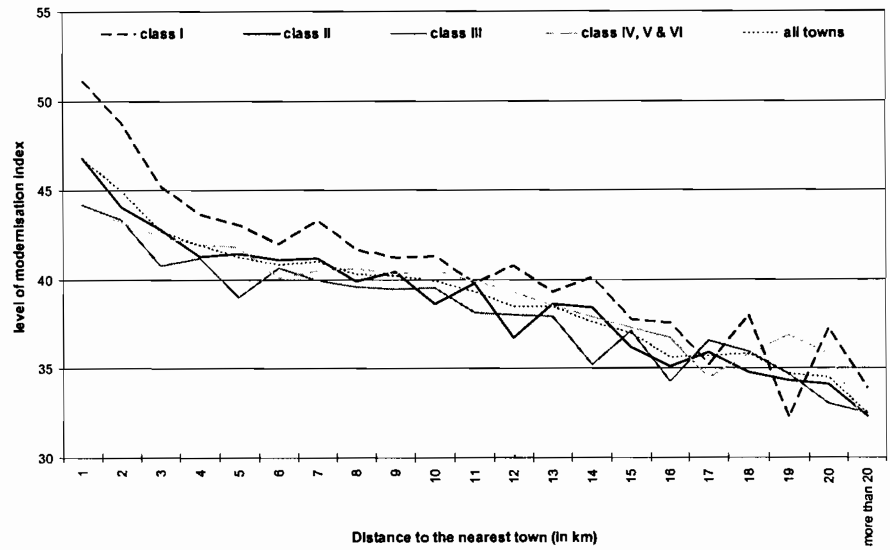


Figure 25: Modernisation index and distance to the nearest town classified by town class

The results are particularly clear as concerns villages near towns of class I: the effect of urban proximity is much greater than for the other towns and this is perceptible up to approximately a distance of 15 km. This means that the impact of the largest towns goes far beyond that of smaller towns. The interpretation of the curve relating to villages close to towns of classes II, III and above all IV, V and VI is more ambiguous. The effect of urban hierarchy is, in fact, hardly discernable among these classes, the respective curves of which distinguish themselves but little. In our hypothesis, the town would have an influence proportional to its size on the level of modernization of the villages. It appears that the smallest towns actually have as much influence as the medium-sized towns. This can result from our mode of calculation that takes into account the nearest town, even if a (larger) town is also close. It is also possible that the differential effect of the size of the towns is only perceptible as of a certain demographic threshold, which in our case study on Tamil Nadu would be around 100,000 inhabitants, confirming what Kundu observed in the 1980s (Kundu, 1992).

We shall now test the effect of the administrative status of these towns in order to see if the same differences are found. Three large categories of towns are distinguished in India: urban agglomerations, municipalities and town panchayats. This classification is interesting in that it takes into account the size of the towns, but also their local importance. Urban agglomerations (UA) are defined by the census as a continuum of several towns or a town of class I and its, or their, outgrowths. Municipalities are urban spaces having their own management system, whereas town panchayats (TP) are urban units which, according to the census, are without municipal status. There is quite logically a gradation of size from the UA (which may include municipalities and TPs) to the TP. Thus, in Tamil Nadu, the UAs have on the average 387,000 inhabitants, as opposed to 64,000 inhabitants for municipalities and 17,000 for town panchayats.

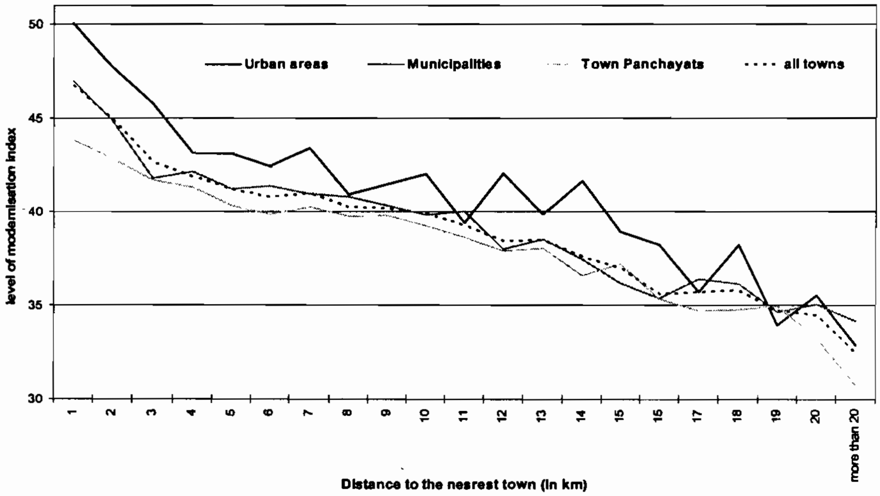


Figure 26: Modernisation index and distance to the nearest town classified by town status

The influence of towns on villages according to their status is much clearer (Figure 26). The UAs have a greater and more extensive influence and are in this respect very comparable to towns with more than one lakh examined previously. The municipalities function as do the medium-sized towns. As for town panchayats, their impact is much less strong and much less extensive. If one takes into consideration the widespread idea according to which the administrative status would be arbitrary, this result can be surprising. But if, on the contrary, one considers that the arbitrariness of these divisions hinges in fact on a less mechanical evaluation in which the human assessments of the role of each town serve to relativize the socio-demographic statistics, one then sees that this division corresponds better to reality.

In conclusion, we will call to mind that the use of Euclidean distance does not fail to raise problems, already profusely observed in geographic literature, as this measurement is but an approximation of the actual distance between localities. Thus, an approach based on the distance by road, as Sharma observes (1980: 226), would certainly be very enriching for the analysis. Better still, to view the time involved in going from one place to another, rather than the distance, would without doubt be a great step forward. Unfortunately, the distance given by the census is not reliable as the errors are too frequent and difficult to rectify. In addition, the automatic measurements of distance by GIS are not yet practicable for Tamil Nadu. However, this work has made it possible to establish a solid framework for the spatial analysis of the countryside in India and for the role played by urban proximity. A further step will consist in making this a viable decision support tool.

The geostatistical analysis of fertility in South India[47]

To conclude this paper, we shall provide another example of the application of our database to spatial analysis, this time not concerned with cartography or ordinary statistics. We shall return to one of the points of departure of the study conducted by the South India Fertility Project on fertility decline: What are the reasons for the very high degree of spatial compactness of fertility such as is observed in India at the district level, but also for South India? Is it a mere cartographic idiosyncrasy, or must one read therein the trace of structuring forces that consistently determine the fertility variations among the Indian regions? It would be difficult to briefly respond to the latter question, which prompted somehow our research programme (the SIFP), but it can be observed that, in the first place, the spatial configuration

[47] This section was written by Christophe Z. Guilmoto. This work benefited from the support of the EMIS project.

of fertility is far from being a matter of chance. It proves to be, furthermore, relatively invariant, discernible from one census to the next since the beginning of the fertility decline on a large scale in the 1960s[48]. But how far can we trust the maps about this? And furthermore, is cartographical inspection enough when we want to compare the way different social features appear on the map of South India?

Why geostatistics?

The analysis of maps rests on visual scrutiny by the reader and for this reason depends closely on their quality, but also on the initial cartographic choices. Maps are seldom unequivocal and the data can lend themselves to numerous manipulations (Monmonier, 1991). And, although data that are mapped can be compared with each other more efficiently with basic statistical analyses (such as regressions), the analysis of the polarization of a given phenomenon on one and the same map is much more empirical, if not approximate. We shall now make use of more powerful geostatistical tools to measure this spatial correlation, starting with fertility measurements.

We have used for this a common measurement, namely, spatial autocorrelation calculated with the Moran index[49]. Moran's I index, the definition of which is given below, makes it possible to measure the statistical correlation $I(d)$ between n pairs of localized observations (the values z_i or z_j), classified according to the distance d separating them.

$$I(d) = \frac{m \sum_{i,j} (z_i - \bar{z})(z_j - \bar{z})}{n \sum_{i} (z_i - \bar{z})^2}$$

The numerator of the index is the covariance between the observations separated by a given distance and the denominator is the total variance of the sample with m size: one compares the variations between distant observations with variations among all observations. By analogy with the

48 Various maps published in Guilmoto (2000) and Guilmoto and Rajan (2001, 2002) easily confirm this strong geographic patterning.

49 Regarding the Moran coefficient, see for example Bailey and Gatrell (1995). A simpler measurement of Moran's I consists in computing the index only for localities that are adjacent, and not by distance class as is done here.

measurement of the usual correlation between two variables (r^2), this index takes a value near to 1 when the spatial autocorrelation is maximal between pairs of observations, which may be the case when the localities are very close. It takes a value of zero when there is no statistical relation between the observations, and more rarely a value of -1 when the observed values are, on the contrary, opposed to each other.

The spatial correlation of social characteristics is seldom zero for close observations, but its intensity is highly variable: there can be strong variations on the micro-scale or, on the contrary, a great homogeneity and the visual examination of the maps scarcely make it possible to discern them. The Moran coefficient allows a systematic analysis of the effects of proximity and consequently of the spatial cohesion of the phenomena. In the case of spatial autocorrelation, the correlation between the observed values, here of fertility, is supposed to be very strong when the localities are close to each other. When the localities are distant, the correlation between the values will be zero, for the observed fertility values will have no relation to each other.

The results presented come from 2151 clusters of 10 km in rural South India. Having a similar area and a uniform distribution, these clusters are thus devoid of spatial or administrative bias. In addition, the average population of the clusters (approximately 65,000 inhabitants) is large enough to provide very robust values of child-woman ratios. Finally, their number makes it possible to carry out the necessary geostatistical computations. We must in fact examine the clusters, not individually, but by groups of two by measuring for each pair of clusters the distance separating them. With clusters of 10 km, one counts not less than 2.3 million pairs of localities, which still lies inside the limits of computation. A similar analysis based on 10,000 localities would be probably much more complex in computational terms.

The comparison of autocorrelation coefficients

The distribution of distances, measured between centres of clusters, is plotted on Figure 27. Distances have been classified by 20-km lags. Thus, the first lag includes localities (here clusters) that are distant by less than 20 km, while the second lag includes those that are distant by 20 to 40 km, and so on. As the figure shows, the number of pairs of observations rapidly increases with distance. For the first lag, we have less than 4000 pairs, but this is more than enough to compute Moran's I statistic. The maximum number of pairs (37,000) is reached for localities that are distant by about 400 km. The number of pairs then regularly decreases and the largest distance between South Indian localities (1450 km) is observed between Kaniyakumari and north-east Andhra Pradesh.

We have chosen in our examination to limit ourselves to distances of less than 400 km, representing around one third of the maximum distance between clusters and somewhat less than the global average distance between clusters. The autocorrelation measurement for distances of more than 400 km is generally zero, or negative, for the observations no longer bear any relationship to each other, and it is for this reason that we are more interested in short distances. We will show here a selection of results corresponding to certain variables of the database, paralleling presentations provided earlier in this book.

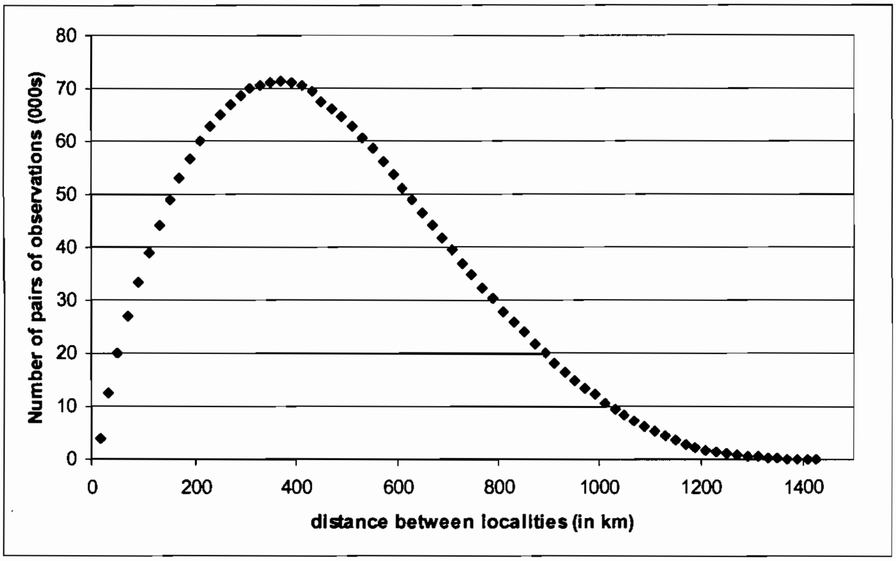


Figure 27: Distribution of distances between localities,
10-km clusters, South India

We will naturally begin with fertility (measured by the child-woman ratio) in Figure 28, which we compare with a common demographic index, that of rural density (in inhabitants per kilometre). The spatial autocorrelation for each of these two variables is manifest, because for each of them the correlation between the observations is very pronounced for short distances and then proceeds to diminish when the distance increases.

The nearest localities thus have very similar values. Fertility is, however, characterized by an extremely high Moran index for the shortest distances and reaches, in fact, 0.85 (85% of the total covariance). This suggests that the fertility measurements between adjacent clusters are practically identical. This similarity appears also for clusters that are less close. The fertility levels in the localities situated at 100 km distance remain in effect much correlated, with coefficients of around 0.5: approximately half the covariance between these localities would thus be explainable solely with the spatial modelling according to the distance separating them.

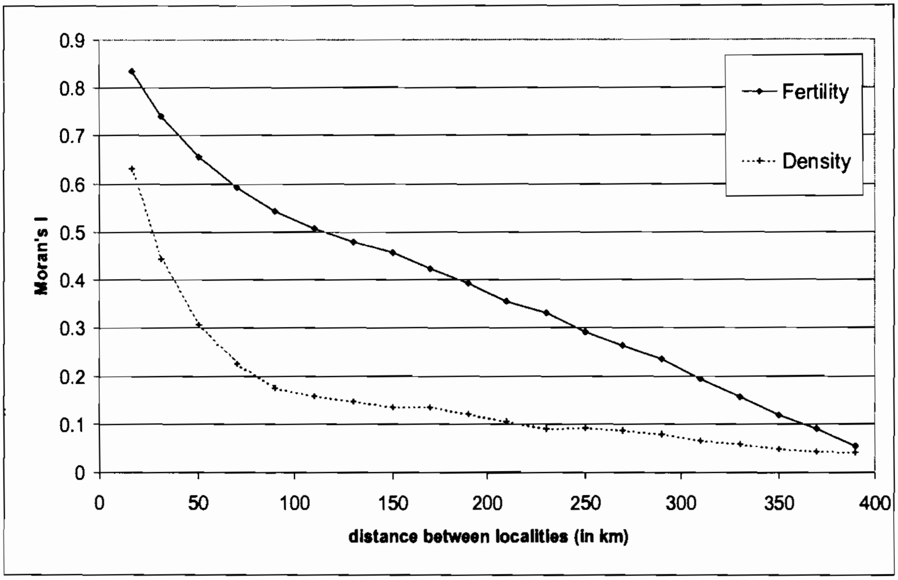


Figure 28: Moran's I for fertility and density, 10-km clusters, South India, 1991

A more detailed analysis according to orientation would show that the variations in fertility are stronger along the north-south or northwest-southwest axis than in the other directions. This pattern is broadly compatible with the pattern of the social geography of India as a whole, such as was conceived of some twenty years ago (Dyson and More, 1983).

The difference between density and fertility in Figure 28 is striking, for the degree of spatial autocorrelation for the former is much lower despite the well-known strong geographic groupings of the population: very high density in Kerala and in the delta regions, low density on the whole of the Deccan Plateau, etc. One thus notes that the correlation between the density values observed is indeed strong at short distance, but then abruptly decreases to become almost negligible when the localities are at a distance of 100 km. The local variability of the density is therefore high. One can assume that the rural density is linked to highly varied phenomena (such as hydrography, altitude, the presence of a city, etc.) that only have a very limited spatial impact and that, consequently, the spatial homogeneity of the density is only discernable over quite reduced spaces.

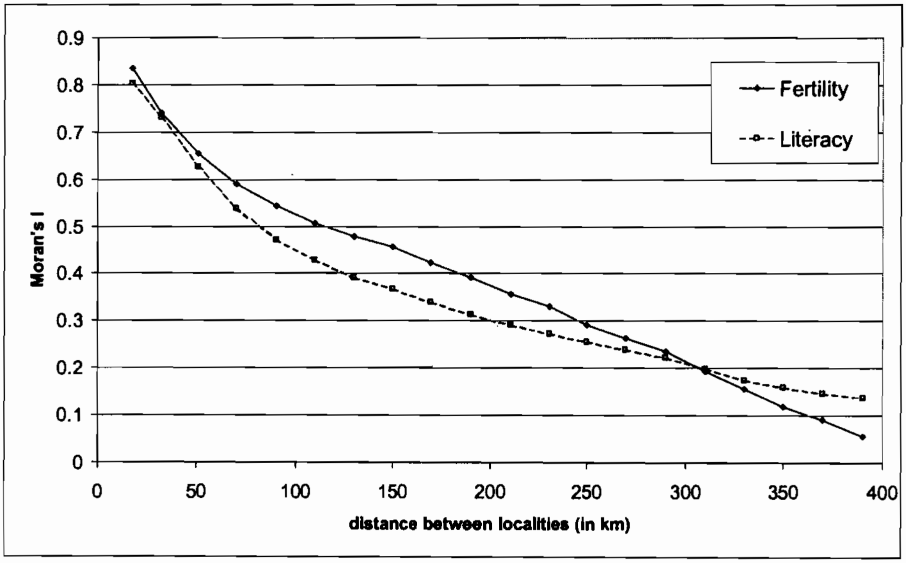


Figure 29: Moran's I for fertility and literacy, 10-km clusters, South India, 1991

Figure 29 takes up the spatial autocorrelation measurement of fertility, comparing it to that of female literacy. As is seen, the spatial correlation of female literacy (population of 7 years and above) is also at a very high level, comparable at first reading to that of fertility. A more detailed examination shows that the spatial correlation of fertility values is somewhat better than in the case of literacy values for the set of distances at less than 300 km. Beyond that, the curves appear to reverse, but the Moran index is then low and without real significance.

The comparison of education and fertility is far from being purely formal. It is in fact known that of all the fertility correlates in the developing world, female literacy usually appears first (Jejeebhoy, 1995). The education of women, along with urbanization, is the variable that best explains the differences in fertility observed between countries, or within countries between the regions or between social groups. It has been therefore suggested that the strong spatial autocorrelation of fertility would thus only be the by-product of that of literacy[50]. Our comparison shows that this reasoning is not entirely valid, for the spatial correlation for fertility values is here also strong, and even slightly stronger than that for education. Not only does the spatial concentration of fertility appear to be higher, but there are obvious differences between the distribution of values that could be detected without difficulty on maps of the South Indian atlas: fertility may for instance be very low in areas where literacy is not especially high as in Kongu Nadu. This means that spatial pattern of fertility has a spatial signature of its own, characterized by both very high concentration and specific geographic contours. It is also possible that fertility decline by itself tends to reinforce the spatial patterns of fertility, especially if diffusion mechanisms are at play. This question, of course, requires a more detailed examination, one that would, however, exceed the framework of this short section (see Guilmoto, forthcoming).

Figure 30, the last, is devoted to a general comparison of several variables. As is seen, fertility is the most correlated variable of all. As a matter of fact, we have not as yet identified any variable that would display a higher degree of autocorrelation than fertility. As minimum value, we have calculated the Moran index for the building sector, measured here by the percentage of the labour force working in this sector. The correlation is one of the lowest to be observed among the SIFP database variables and becomes practically zero for distances of 50 km or more. One no doubt observes the polarizations of this activity on the South Indian scale, for example, in Kerala because of the expenditures of migrants returning from the Gulf, in rapidly developing peri-urban zones, or even in regions where quarrying is done. But the intensity of this sector of activity in a given locality generally has but little relation to that noted in nearby communities, for the observed variations, which are quite minor, are very weakly

[50] I thank Patrice Vimard for drawing my attention to this issue.

determined by location. Beyond a certain distance, there is absolutely no link between the localities and the geographic distribution of the building sector seems to be nearly random. In fact, the needs in this sector are quite well distributed in the rural milieu and the concentrations that may exist are of a very localized character.

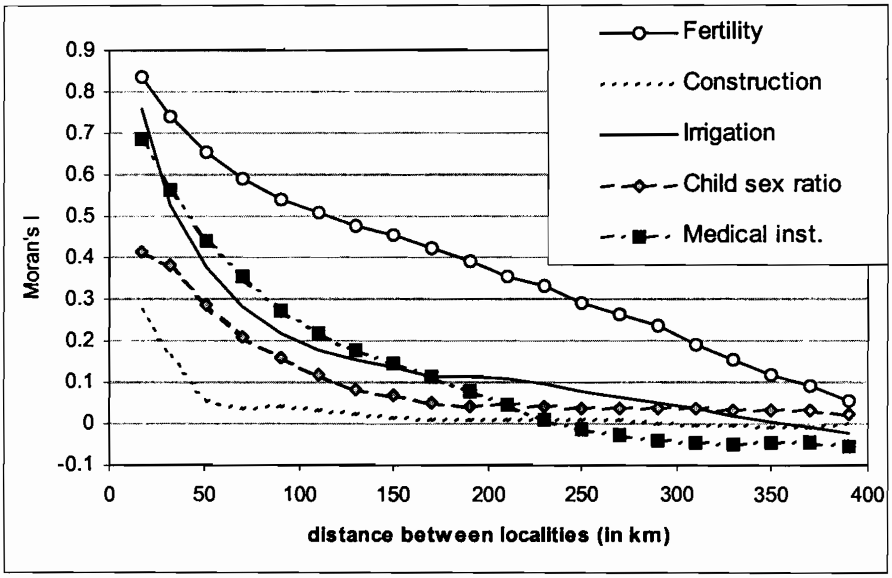


Figure 30: Moran's I for fertility and other indicators,
10-km clusters, South India, 1991

The other variables indicate clearly higher levels of autocorrelation that are located at between 40% and 75% of the maximum possible correlation. In particular, irrigation is spatially very strongly correlated, but this feature tends to fade away very rapidly with distance: the homogeneity observed among irrigation levels may be true firstly within small geographic units such as watersheds or irrigation systems (for example, by canals). At a distance of more than 100 km, the regional tendency becomes very weak. One can imagine that rainfall, which in large part determines the success of non-irrigated agriculture, will, on the contrary, be strongly correlated and on more extensive spaces. The density of reserved forests (not represented on the graph) is still less spatially autocorrelated, as forested areas usually correspond to isolated zones, the only concentrations of which are found to the mountainous parts of the Ghats.

The state of health care infrastructures, measured by the proportion of villages in the cluster endowed with a medical institution (census definition), displays a very strong spatial coherence and the favoured regions tend to be contiguous. One recognizes here the features illustrated in the study on the scale of Andhra Pradesh conducted by Virginie Chasles.

On the other hand, the child sex ratio on the scale of South India is seen to be very weakly correlated spatially, which requires additional explanation. Stéphanie Vella previously shows that sexual discrimination, far from being randomly distributed, is particularly acute in certain pockets of Tamil Nadu, especially concentrated in the regions of Salem and Usilampatti. According to the calculation of the Moran index, the spatial homogeneity of this phenomenon appears to be moderate (around 40% at short distance) and vanishes beyond 100 km. What this results indicates is, in fact, that the spatial distribution of the child sex ratio in most South India is nearly random: the variations are small from one locality to the next and the observed values are unconnected statistically. It is only in regions where the differences are more pronounced, as in Tamil Nadu, that the regional clustering is more significant. In fact, if the analysis were limited to the 466 clusters of Tamil Nadu, the results would be different: at short distance, the Moran index amounts to 60% and thus approaches that of the other indices. One will note, however, that the autocorrelation on this scale is still below that of fertility. This suggests that a significant part of the heterogeneity between the child sex ratio values cannot be attributable only to strictly speaking socio-spatial phenomena, but is a matter of non-spatialized variables or of an almost random distribution, as is often the case for the sex ratio levels at birth.

An interim conclusion

As an interim conclusion to an exploration that is in its inception, we will note that geostatistical formalization, while common in archaeology or in the environmental sciences, is new and very seldom undertaken in the social domain[51]. The results obtained here are therefore exploratory and must be interpreted above all as comparative: the differences in the calculated levels of spatial autocorrelation should be analyzed, rather than the raw values

[51] See however the efforts deployed by the Center for Spatially Integrated Social Sciences (www.csiss.org).

themselves, which often depend on the levels of scale of analysis as well as the spatial characteristics of the phenomena studied.

The level of clusters (10 km) retained for this exercise is seen to be quite efficient, for it makes it possible to avail of a sample that is neither too large (problem of time involved in calculation) nor too small (problem of non-representative small samples). Furthermore, the scale of aggregation has also made it possible to do away with clusters that are too small, such as in aggregates smaller than 2 and 5 kilometres, the inevitable effect of which is to impoverish the statistical calculations. This problem is frequently a source of difficulty in the calculation of the child-woman ratio because the numbers of children and women are sometimes very imbalanced in the small populations and the geostatistical calculations, which ascribe the same weight to all clusters independently of their demographic size, are particularly vulnerable.

As concerns fertility, there is no doubt that the estimated levels of spatial autocorrelation are extremely high and greater than most of the other variables tested, also on other scales of analysis (Guilmoto and Rajan, 2001). Geostatistical modelling thus goes far beyond the perceptions provided by cartography to point to a dimension that is crucial in the understanding of demographic change in India. It now remains necessary to link this spatial homogeneity of reproductive behaviours to their own logic of evolution within the social institutions that govern them (the couple, the family, the social group, the state, etc.). We are now far from geostatistical formalization, but we hope to have shown that the insights derived from a formal analysis, such as that based on the indices of spatial autocorrelation, compel us to directly approach the spatial question in the analysis of the fertility decline, in a manner parallel to the previously developed examination of the imbalances in the distribution of the sexes.

CONCLUSION

The intention of this collective volume was deliberately more descriptive than analytical because we have wanted to reconstitute the integrality of a geomatic approach. It is rare today that geographers would be led to cover the whole of this itinerary, extending from the measurements of georeferencing on the basis of public maps of uncertain quality to attempts at spatial or geostatistical modelling. The tasks are often disconnected and the technical parts ignored, as being superfluous, or are indeed abandoned to engineers. This path of research, although necessarily much longer than this volume would perhaps lead one to believe, has several virtues for disciplinary study.

In the first place, it calls to mind that numerous regions suffer from a glaring lack of modern cartography, for an entire range of reasons stemming by and large from the technical under-equipping of the local administrative or scientific apparatus, which is paradoxically held back by the rapidity of, often geographic, changes. In the India that we study here, the frequency and magnitude of the redrafting of borders, whether of urban territories or of political units such as districts or states, very certainly tend to discourage large cartographic investments that a few years of reform threaten to render obsolete. The studies are most often conducted on the basis of satellite imagery, which has the advantage of being indifferent to the continuous redrawing of administrative borders. The task of local digitization has, of course, begun in India, but it is concentrated more in the advanced regions than in the rural world: the large cities thus benefit from the ambitious "Urban Mapping Scheme" project (combining aerial photographs, satellite imagery and GIS) conducted by the Ministry in charge of urban affairs. But the coverage of the Indian countryside is particularly inadequate.

The volume of information collected in India is also prohibitive for an undertaking that would want to be exhaustive: we have devoted several years to the cartography of villages in southern India, but there remain four times as many villages in the rest of the country. However, the work to be undertaken cannot be ignored and investments made towards this end by researchers must quite naturally contribute to the enrichment of the public domain. The most significant venture thus far has been embarked upon under the aegis of the National Spatial Data Infrastructure (NSDI), which intends to facilitate and coordinate the compiling and sharing of spatial data.

It concerns both the dispersed suppliers of spatial information (hydrological, geographic, geological, agricultural, etc.) and the potential users by offering new formats of exchange[52]. At this time, these different projects have not yet been given tangible form, but without doubt a new era of geographic information in India has been announced.

In the second place, the results briefly presented in our article for the most part confirm the logic of the project. We assumed that the spatial approach would shed new light on the knowledge pertaining to numerous phenomena and this seems, in fact, to be confirmed. On the one hand, small-scale mapping sometimes reveals the magnitude of local or sub-regional phenomena that the usual format of available data often conceals. One now observes the very fine geographic differentiations that make it possible to more clearly discern the phenomena concerned, their constraints and their actors. It can even be a matter of phenomena that are by their nature dissimulated, as in the case of infanticide. On the other hand, the GIS enables one to cross information from diverse sources and thus to make good use of the large choice of available data. It is similar to a statistical approach, to the extent that one can systematically compare the phenomena to each other by comparing the layers of cartographic information. But it goes further by making it possible to create new information by gathering together different sources of information.

Finally, new research themes find here sustenance, for the very strong spatial patterning of the phenomena of social change conveys numerous elements of response to the study of their origins and, above all, to their mode of propagation in space and time across different layers of society. More theoretical questions, such as the origin of the high spatial compactness of the fertility decline in India or the geography of pilgrimage, thus find solid bases for analysis. The examples reviewed in this volume are revealing as they confirm that South India is built on both old and new spatial networks that the process of social change tends to reactivate and to reshape. It henceforth appears to be indispensable to take into account the spatial dimension both in the analysis of the conditions in which the

[52] The ambitious mission of the NSDI includes: "encouraging collection and distribution of spatial data on different themes in common defined standards and formats by different mapping agencies in India [and providing] metadata of all the data available with various participating agencies and facilitat[ing] users to obtain data with simple and smooth procedure". (www.idsiindia.org)

examined phenomena emerge and in the consideration of the future implications of social change.

In conclusion, our trajectory also indicates that technological development takes place not only to the advantage of formal tools of scientific treatment as illustrated by geostatistical calculations. Powerful tools for the reconstitution of information are now available to make more readable information that perhaps existed earlier, but was infrequently used outside circles of university specialists because of its chronic inaccessibility. The GIS does promote the use of spatialized information to a new range of casual users in the public. In the era of decentralized management (*Panchayati Raj*), the need for disaggregated information is becoming increasingly felt in India, in order to locally administer operations of resource development. Moreover, some more advanced regions, such as Kerala, have long ago integrated a local cartography in the first development plans of the *panchayats* (Chattopadhyay *et al.*, 1999). It is therefore important to have tools so as to be able to deliver information to highly varied users, from administrators to students, by way of NGOs and businesses which, moreover, appear to have been the only ones to have competed with us in this domain. Among these tools, CD-ROMs have the advantage of offering low-cost support, replacing atlases, the production and printing costs of which make them hardly publishable in developing countries. The possibility to use interactive modes of questioning with cartographic consultation software makes this tool even more attractive and the access to information easier. Online publications, via the web, represent additional assets, notably in terms of updating and delocalized accessibility. The non-scientific repercussions of a project, endowed with an active policy of development and diffusion, can thus be significant and silence the sometimes prevailing opinion regarding the relatively pointless character of fundamental research projects in developing countries.

Web sites

General

Cybergeo (European journal of geography): http://www.cybergeo.presse.fr

GIS tutorials: http://sunsite.berkeley.edu/GIS/gistuts

Spatial analysis in social sciences: http://www.csiss.org

Maps and GIS in India

Census of India: http://www.censusindia.net

GIS at development: http://www.gisdevelopment.net

Indian statistics: http://www.indiastat.com

Map store: http://www.mappls.com/mapstore

Maps of India: http://www.mapsofindia.com

Maps, Census of India: http://www.censusindiamaps.net

Maptell: http://www.maptell.com

National Atlas & Thematic Mapping Organisation: http://www.natmo.org

National Spatial Data Infrastructure: http://www.nsdiindia.org

South India Population Information System: http://www.demographie.net/sipis

Survey of India: http://www.surveyofindia.gov.in

Topic-related web sites

Andhra Pradesh: Health department: http://gist.ap.nic.in/health

Department of Women and Child Development: http://wcd.nic.in

Indian Girl Child and Women Welfare Society: http://www.girlchildwelfare.org

Ministry of Health and Family Welfare: http://mohfw.nic.in

Ministry of Urban Development and Poverty Alleviation: http://urbanindia.nic.in

Ministry of Water Resources: http://wrmin.nic.in

National Family and Health Survey: http://nfhsindia.org

Sabarimala web site: http://www.achichu.com/temples/sabarimala.htm

Tamil Nadu Ministry of Rural Development: http://www.rural.tn.gov.in

References

Maps and atlases

A Road Guide to (Tamil Nadu, Kerala, Karnataka, Andhra Pradesh), 1999, TTK Pharma Limited, Chennai.

Alphabetical List of Villages in the Taluks and Districts of the Madras Presidency, 1933, Re-printed by the Superintendent, Government Press, Madras.

Guilmoto, C. Z., and Oliveau S., with S. Vingadassamy and R. Amuda, 2000, South Indian Population Information System, Volume I: Tamil Nadu and Pondicherry, CD-ROM published by the French Institute, Pondicherry.

India and Bangladesh Road Atlas, 2001, Lonely Planet Publications, Victoria, Australia, 175 p.

Road Atlas and Distance Guide (Tamil Nadu, Kerala, Karnataka, Andhra Pradesh), 2001, Map India Service, Jodhpur.

Other publications

ATHREYA, V. and CHUNKATH, S. R., (1998), "Gender and Infant Survival in Rural Tamil Nadu. Situation and Strategy", *Economic and Political Weekly,* Vol. 33, No 40, 2588-2594.

BAILEY, T. C., and GATRELL, A. C., (1995), *Interactive Spatial Data Analysis*, Longman, Harlow.

BECK, B. E. F., (1979), *Perspectives on a Regional Culture. Essays About the Coimbatore Area of South India*, Vikas Publishing House, New Delhi.

BERROIR, S., (1998), *Concentration et polarisation : organisation des espaces urbanisés. Etude comparative des grandes villes françaises*, Ph.D. Thesis, Université Paris I.

BHASKARAN NAIR, P.G. alias VENUSWAMY, (1998), *A book on Ayyappa Mahatmyam and Vana Yatra (Malayalam) (A Book on Sabarimala Pilgrimage and Forest Trek)*, Kottayam.

CASTERLINE, J. B., (2001), *Diffusion Processes and Fertility Transition: Selected Perspectives.* Committee on Population, National Research Council, National Academy Press, Washington.

Census of India, (1998), *Availability of Infrastructural Facilities in Rural Areas of India: an Analysis of Village Directory Data*, Census of India 1991, Office of the Registrar General, New Delhi.

Census of India, (2001), *Series 1, Paper 1, India, 2001*, Office of the Registrar General, New Delhi.

CHASLES, V., (2001), "L'environnement culturel de la santé maternelle et périnatale en milieu rural en région bordière Andhra Pradesh/Karnataka", séminaire "les défis de l'Inde : enjeux et perspectives au XXIe Siècle", AJEI, Paris, 27 novembre. (http://www.ajei.org/downloads/sjc/2001/virginie.pdf)

CHASLES, V., (2002), "Entre sacralité et impureté, l'ambivalence de la maternité en Inde", *Espace Populations Sociétés*, 3, 387-396.

CHATTOPADHYAY, S., *et al.*, (1999), *Panchayat Resource Mapping to Panchayat-Level Planning in Kerala: An Analytical Study*, KRD Discussion Paper 14, Centre for Development Studies, Thiruvananthapuram.

CHOU, Y. H., (1997), *Exploring Spatial Analysis in Geographic Information Systems*, Onword Press, Santa Fe.

DELOCHE, J., (1993), *Transport and Communications in India. Prior to Steam Locomotion: Vol.1, Land Transport.* Oxford University Press, New Delhi.

DYSON, T., and MOORE, M., (1983), "Kinship Structure, Female Autonomy, and Demographic Behaviour in India", *Population and Development Review*, 9, 1, 35–60.

EDNEY, M. H., (1999), *Mapping an Empire*, Oxford University Press, New Delhi.

FOTHERINGHAM, S., BRUNDSON, C., and CHARLTON, M., (2000), *Quantitative Geography. Perspectives on Spatial Data Analysis,* Sage, London.

GOFFETTE-NAGOT, F., and SCHMITT, B., (1998), "Les proximités, la ville et le rural", *in* Huriot, J.M., *La ville ou la proximité organisée*, Anthropos, Coll. Villes, Paris, 173-183.

GUILLOT, M., (2002), "The Dynamics of the Population Sex Ratio in India, 1971-96", *Population Studies*, 56, 1, 56-63.

GUILMOTO, C. Z. and S. IRUDAYA RAJAN, (2001), "Spatial Patterns of Fertility Change in Indian Districts", *Population and Development Review*, 27, 4, 713-738.

GUILMOTO, C. Z. and S. IRUDAYA RAJAN, (2002), "District Level Estimates of Fertility from India's 2001 Census", *Economic and Political Weekly*, 37, 7, 665-672.

GUILMOTO, C. Z., (2000), "Demographic and Environmental Changes. Population Growth and Spatial Dynamics", *in* Ramakrishnan, P.S., *et al.*, *Mountain Biodiversity, Land Use Dynamics, and Traditional Ecological Knowledge*, Oxford and IBH, New Delhi, 54-70.

GUILMOTO, C. Z., (2002), "Irrigation and the great Indian rural database. Vignettes from south India", *Economic and Political Weekly*, March 30, 1223-1228.

GUILMOTO, C. Z., and A. VAGUET, (eds.) (2000), *Essays on Population and Space in India*, French Institute, Pondicherry.

GUILMOTO, C. Z., OLIVEAU, S., et VINGADASSAMY, S., (2002), "Une expérience de SIG en Inde du sud. Théorie, mise en oeuvre et application thématique", *Espace Populations Sociétés*, 1-2, 147-163.

GUILMOTO, C. Z., and S. IRUDAYA RAJAN, (eds.), forthcoming, *Fertility Transition in South India*.

GUILMOTO, C. Z., forthcoming, "A Comparative View of Fertility Decline: India vs. South India", *in* C. Z. Guilmoto and S. Irudaya Rajan, (eds.), *Fertility Transition in South India*.

HILAL, M., SCHMITT, B., (1997), "Les espaces ruraux : une nouvelle définition d'après les relations villes-campagnes", *INRA Sciences Sociales*, INRA, Ivry N°5, novembre, (www.inra.fr/Internet/Departements/ESR/publications/iss/pdf/iss97-5.pdf).

International Institute for Population Sciences, 1995, *National Family Health Survey, India 1992-93*, International Institute for Population Sciences, Bombay.

JEJEEBHOY, S. J. (1995), *Women's Education, Autonomy, and Reproductive Behaviour: Experience from Developing Countries*, Clarendon Press, Oxford.

KERR, I. J., (2001), "Reworking a Popular Religious Practice: The effects of Railways on Pilgrimage in 19th and 20th Century South Asia", *in* Kerr, I. J., Ed., *Railways in Modern India*, Oxford University Press, New Delhi, 304-327.

KRAAK, M.-J., ORMELING, F., (1996), *Cartography. Visualisation of spatial data*, Longman, Harlow.

KUNDU, A., (1992), *Urban development and urban research in India*, Khama Publishers, New Delhi.

KUNDU, A., PRADHAN, B.K., and SUBRAMANIAN, A., (2002), "Dichotomy or Continuum, Analysis of impact of urban centres on their periphery", *Economic and Political Weekly*, 14 December, 5039-5046.

LEE, M. RAYMOND, (1993), *Doing Research on Sensitive Topics*, Sage, New Delhi.

LONGLEY, P. and BATTY, M., (eds.) (1996), *Spatial Analysis: Modelling in a GIS Environment*, GeoInformation International, Cambridge.

MADAN, P.L., (1997), *Indian Cartography. A Historical Perspective*, Manohar, New Delhi.

MARKOVITS, C., POUCHEPADASS, J., and SUBRAHMANYAM, S., (2003), *Mobile People and Itinerant Cultures in South Asia (1750-1950)*, Permanent Black, New Delhi.

MARTIN, D., (1996), *Geographic Information Systems. Socio-economic Applications*, Routledge, London and New York.

MERENNE-SCHOUMAKER, B., (1996), *La localisation des services*, Editions Nathan, Coll. Géographie d'aujourd'hui, Paris.

MONMONIER, M., (1991), *How to Lie with Maps*, University of Chicago Press, Chicago.

MORICONI-EBRARD, F., (1994), *Geopolis. Pour comparer les villes du monde*, Anthropos, coll. Villes, Paris.

NOIN, D., (1996), *Atlas de la population mondiale*, Deuxième édition, Reclus, La Documentation Française, Paris.

OLIVEAU, S., (1998), *Atlas du district de Vellore : essai de typologie régionale*, mémoire de maîtrise de l'université de Paris X Nanterre.

PRUVOST-GIRÓN, C., (2002), *Palakkad Pass and the cattle trail. An example of a territory of reference in South India,* Pondy Paper in Social Sciences, 28, French Institute, Pondicherry.

PUMAIN, D., et SAINT-JULIEN, T., (1997), *L'analyse spatiale. 1. Localisations dans l'espace,* Armand Colin, Paris.

RAJ, K., (2003), "Circulation and Emergence of Modern Mapping. Great Britain and Early colonial India, 1764-1820" *in* Markovits, C., Pouchepadass, J., S. Subrahmanyam, eds, *Society and Circulation. Mobile People and Itinerant Culture in South Asia 1750-1950*, Permanent Black, Delhi, 23-54.

RAMACHANDRAN, R., (1989), *Urbanization and Urban System in India*, Oxford University Press, New Delhi.

RAWAL, V., (2001), "Irrigation Statistics in West-Bengal", *Economic and Political Weekly*, 36, 27, 2537-44.

SANDERS, L., (1990), *L'analyse statistique des données en géographie*, Reclus, Montpellier.

SEKAR, R., (1992), *The Sabarimala Pilgrimage and Ayyappan Cultus*, Motilal Banarsidas Publishers, New Delhi.

SHARMA, H.N., (1980), "City-hinterland relationships in regional context", *in* Sopher, D., *An Exploration of India, Geographical Perspective on Society and Culture*, Cornell University Press, New York, 213-233.

SURESH, V., (1992), "Sabarimala, Redefining Identities", *The Hindu*, January 19.

TARABOUT, G., (1997), "L'évolution des cultes dans les temples hindous. L'exemple du Kerala (Inde du sud)", *in* Clementin-Ojha, C., *Renouveaux religieux en Asie,* Presses de l'EFEO, Paris, 127-54.

VAIDYANATHAN, K.R., (1978), *Pilgrimage to Sabari*, Bhavan's Book University, Bombay.

VELLA, S., (2003), "Ethique et pratiques reproductives: les techniques de sélection sexuelle en Inde", *Autrepart*, 28, 147-163.

VELLA, S., forthcoming, "Trends and differentials of female discrimination in South India: a case study of Salem district" *in* Guilmoto CZ and S Irudaya Rajan, (eds.), *Fertility Transition in South India.*

VISARIA, L., (1999), "Deficit of Women in India. Magnitude, Trends, Regional Variations and Determinants" *in* Ray, B. and B., Aparna, (eds.), *From Independence Towards Freedom,* Oxford University Press, New Delhi.

LA CARTE DU CHANGEMENT SOCIAL EN INDE DU SUD

Un système d'information géographique et ses applications

L'Inde et son milliard d'habitants possèdent un riche recensement plus que centenaire, qui reste une source sous-exploitée pour l'étude des dynamiques socio-spatiales actuelles. L'article résume les étapes d'un projet de géomatique appliqué aux quelque 75.000 villages de l'Inde du sud.

L'apparition de la technologie des SIG dans un pays comme l'Inde bute sur de multiples problèmes (scientifiques, institutionnels et techniques). La localisation géoréférencée des villages a, par exemple, demandé des efforts considérables, en relation directe avec la rareté et la médiocre qualité des fonds cartographiques disponibles sur papier. De même, l'exploitation systématique de la statistique censitaire a mis en évidence ses faiblesses et ses lacunes, nécessitant des corrections de tous types. L'article fait le récit de ces problèmes pratiques et théoriques, ainsi que des solutions trouvées à la question de la restitution de l'information constituée grâce aux médias modernes.

En conclusion, nous présentons des exemples d'applications cartographiques de cette riche base de données et quelques outils géostatistiques qui permettent désormais d'aborder de manière renouvelée l'étude du changement socio-spatial. Les exemples choisis concernent des sujets aussi variés que la géographie du pèlerinage de Sabarimala, le système des soins en Andhra Pradesh, l'effet de l'urbanisation sur les villages tamouls, la discrimination sexuelle au Tamil Nadu, l'irrigation en Inde du sud ou encore les mesures d'autocorrélation spatiale.

Les auteurs de cet ouvrage sont membres du projet "Population et Espace en Inde du sud" co-ordonné par CZ. Guilmoto au département de Sciences Sociales, Institut Français de Pondichéry. **Christophe Z. Guilmoto** est démographe à l'Institut de Recherche pour le Développement (IRD, LPED) ; ses études portent sur la baisse de la natalité , la théorie de migration et la démographie spatiale. **Virginie Chasles** (Université de Rouen), **Rémy Delage** (Université de Bordeaux 3), **Sébastien Oliveau** (Université de Paris I) et **Stéphanie Vella** (Université de Bordeaux 3) sont sur le point de défendre une thèse en géographie et se sont spécialisés dans l'analyse spatiale des changements sociaux et économiques en Inde du sud.

CONTRIBUTORS

Virginie Chasles is doctoral candidate in geography and temporary lecturer at Rouen University (LEDRA/IDEES). She did her fieldwork in rural Andhra Pradesh in 2000 on health and maternity issues.

Rémy Delage is doctoral candidate in geography at Bordeaux 3 University (DYMSET). He has spent three years in India working on the Sabarimala pilgrimage in Kerala and organized an international workshop in Kolkata in 2003 while being affiliated to the French Institute and JNU.

Christophe Z Guilmoto is a demographer with the French *Institut de Recherche pour le Développement* (LPED research unit) and is coordinator of the programme "Population and Space in South India" at the Department of Social Sciences, French Institute of Pondicherry.

Sébastien Oliveau is doctoral candidate in geography at Paris I University (Géographie-Cités) and was fellow of the French Institute of Pondicherry in 2001-2003. He published a first CD-Rom (SIPIS) with the French Institute in 2000.

Stéphanie Vella is doctoral candidate in geography at Bordeaux 3 University (DYMSET). She is working on gender geography in India and was affiliated to the French Institute during her fieldwork in South India.

V. Chasles, CZ Guilmoto, S. Oliveau and S. Vella are members of the EMIS (Space and Measurement in South India) team funded by the GETM programme (CNRS-IGN-CEMAGREF) and have also enjoyed support from the LPED (Laboratoire-Population-Environnement-Développement), IRD-Université de Provence.

Pondy Papers in Social Sciences

GUÉRIN, I & SERVET, J.-M. (2003), 'Microfinance: From daily survival to social change', *Pondy Papers in Social Sciences*, 30, 153 p.

PORDIÉ, L. (2003), 'The Expression of Religion in Tibetan Medicine. Ideal Conceptions, Contemporary Practices and Political Use', *Pondy Papers in Social Sciences*, 29, 83 p.

PRUVOST-GIRON, C. (2002), 'Palaakkad Pass and the Cattle Trail: An Example of a Territory of Reference', *Pondy Papers in Social Sciences*, 28, 63 p.

SÉBASTIA, B. (2002), 'Māriyamman-Mariyamman, Catholic Practices and Image of Virgin in Velankanni (Tamil Nadu)', *Pondy Papers in Social Sciences*, 27, 74 p.

QUIEN, A. (2001), 'Friends of the Household: Study of a Catering Company in Mumbai', *Pondy Papers in Social Sciences*, 26, 67 p.

LANDY, F. (1997), 'Fertilizers, Structural Adjustment and Food Policy', *Pondy Papers in Social Sciences*, 25, 64 p.

ASSAYAG, J., LARDINOIS, R., VIDAL, D. (2001), 'Orientalism and Anthropology. From Max Müller to Louis Dumont', *Pondy Papers in Social Sciences*, 24 [1st edition 1997], 94 p.

BÉNÉÏ, V. & KENNEDY, L. (eds.) (1997), 'Industrial Decentralization and Urban Development', *Pondy Papers in Social Sciences*, 23, 165 p.

DE BRUIN, H.M. (1996), 'Leprosy in South India: Stigma and Strategies of Coping', *Pondy Papers in Social Sciences*, 22, 103 p.

KALAM, M.A. (1996), 'Sacred Groves in Kodagu District of Karnataka (South India). A Socio-Historical Survey', *Pondy Papers in Social Sciences*, 21 [1st edition 1996], 60 p.

BOQUÉRAT, G. and MUKHERJI, I.N., SAARC. (1996), 'Economic and Political Atlas, (41 maps, 21 tables and 2 charts)', *Pondy Papers in Social Sciences*, 20, 116 p.

BÉNÉÏ, V. (1996), 'To Give or not to Give. From Brideprice to Dowry in Maharashtra (Pune District)', *Pondy Papers in Social Sciences*, 19, 73 p.

ASSAYAG, J. (1995), 'The Making of Democratic Inequality. Caste, Class, Lobbies and Politics in Contemporary India (1880-1995)', *Pondy Papers in Social Sciences*, 18, 81 p.

BOURDIER, F. (1995), 'Sickness and Health in the Village. A Assessment of Health Conditions in South India (Coimbatore District)', *Pondy Papers in Social Sciences*, 17, 70 p.

BUFFETRILLE, K. (1994), 'The Halase-Martika Caves (Eastern Nepal). A Sacred Place Claimed by both Hindus and Buddhists', *Pondy Papers in Social Sciences*, 16, 70 p.

FREEMAN, R. (1994), Forests and the Folk: Perceptions of Nature in the Swidden Regimes of Highland Malabar', *Pondy Papers in Social Sciences*, 15, 36 p.

HOLMSTROMÖM, M. (1994), 'Bangalore as an Industrial District: Flexible Specialization in a Labour-surplus Economy?', *Pondy Papers in Social Sciences*, 14, 85 p.

BUCHY, M. (1993), 'British Forestry in the Western Ghats (India) and French Forestry in Indochina: a Comparison', *Pondy Papers in Social Sciences*, 13, 22 p.

GARRIGUES, J.P., DERAND, G.D. & HEDGE, R.(1993), 'Anthropic Action on the Vegetation in the Western Ghats of India (Shimoga District, Karnataka): a Study Using Agrarian Systems Analysis', *Pondy Papers in Social Sciences*, 12, 36 p.

MARIUS-GNANOU, K. (1993), 'Socio-economic Impact of the Green Revolution on Tamil Rural Society: The Example of the Pondicherry Area', *Pondy Papers in Social Sciences*, 11, 67 p.

LANDY, F. (1992), 'To Migrate or to Stay in the Maidan: a Survey in Two Rural Systems in South Karnataka', *Pondy Papers in Social Sciences*, 10, 70 p.

Coll. (1991), 'Urban Configurations and Merchant Networks in South India: a Workshop', *Pondy Papers in Social Sciences*, 9, 103 p.

ROGER, D. (1991), 'Food Practices and Traditional Medicine among the Muslims of Hyderabad', *Pondy Papers in Social Sciences*, 8, 32 p.

POUCHEPADASS, J. (1990), 'The Ecological History of the Central Western Ghats in the Modern Period: a Preliminary Survey', *Pondy Papers in Social Sciences*, 7, 32 p.

BUCHY, M. (1990), 'Colonial Forest Exploitation in the Western Ghats of India: a Case Study of North Kanara District', *Pondy Papers in Social Sciences*, 6, 56 p.

RACINE, J. (1990), 'To Migrate or to Stay? Mobility and Retention of Rural Populations in South India: a Field Survey of Villagers' Rationales in Mandya District, Karnataka', *Pondy Papers in Social Sciences*, 5, 34 p.

SCHAR, Ph. (1989), 'Well Irrigation and Socio-economic Changes. A Case Study in Pollachi Taluk, Coimbatore District, Tamil Nadu', *Pondy Papers in Social Sciences*, 4, 38 p.

AZIZ, A. (1989), 'Social and Economic Change in a Karnataka Village', *Pondy Papers in Social Sciences*, 3, 18 p.

RACINE, J. (ed.) (1989), 'Rural Change in Karnataka: a workshop', *Pondy Papers in Social Sciences*, 2, 80 p.

AZIZ, A. and RACINE ACINE, J. (1989), 'About Decentralization: Two Studies on Multi-level Planning in South India', *Pondy Papers in Social Sciences*, 1, 50 p.

The complete list of publications of the Department of Social Sciences is available on the French Institute Website: www.ifpindia.org

MAPPING OUT SOCIAL CHANGE IN SOUTH INDIA

A geographic information system and its applications

India, the population of which recently reached the one-billion mark, enjoys a rich and ancient census apparatus that remains an under-utilized source for the study of contemporary socio-spatial dynamics. Our paper summarizes the stages of a GIS project applied to the 75.000 localities of South India.

The emergence of GIS technology in a country like India comes up against many scientific, technical or institutional problems. The georeferencing of villages, for instance, required considerable effort because of the unavailability of reliable printed maps. Similarly, systematic examination of the census data has underlined their shortcomings and called for corrections of all kinds. This article chronicles some of these concrete and theoretical difficulties, as well as solutions found to make the collected information available to a large number of potential users.

By way of conclusion, we present some of the cartographic applications of this rich database, as well as some geostatistical tools that can be fruitfully applied in a renewed approach to the phenomenon of socio-spatial change. Illustrations given here come from various fields of interest: geography of Sabarimala pilgrimage, health care in Andhra Pradesh, impact of urbanization on Tamil villages, sexual discrimination in Tamil Nadu, irrigation in South India or spatial autocorrelation measurements.

The authors of this volume are members of the programme "Population and Space in South India" coordinated by CZ Guilmoto at the Department of Social Sciences, French Institute of Pondicherry. **Christophe Z Guilmoto** is a demographer at the French Research Institute for Development *(IRD, LPED)* and works on fertility decline, migration theory and spatial demography. **Virginie Chasles** (Rouen University), **Rémy Delage** (Bordeaux 3 University), **Sébastien Oliveau** (Paris I University) and **Stéphanie Vella** (Bordeaux 3 University) are doctoral candidates in geography and specialized in the spatial analysis of social and economic change affecting South India.