



HAL
open science

Les indicateurs locaux d'association spatiale (LISA) comme méthode de régionalisation : Une application en Inde.

Sébastien Oliveau

► **To cite this version:**

Sébastien Oliveau. Les indicateurs locaux d'association spatiale (LISA) comme méthode de régionalisation : Une application en Inde.. 7ème rencontres ThéoQuant, Jan 2005, Besançon, France. pp.1-10. halshs-00519265

HAL Id: halshs-00519265

<https://shs.hal.science/halshs-00519265>

Submitted on 19 Sep 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Les indicateurs locaux d'association spatiale (LISA) comme méthode de régionalisation

Une application en Inde¹

Sébastien OLIVEAU : ESPACE² UMR 6012 CNRS - Université de Provence

sebastien.oliveau@univ-provence.fr

RÉSUMÉ :

En Inde, les données de recensement sont de suffisamment bonne qualité pour que l'on puisse se permettre une approche géographique fine des phénomènes socio-économiques, au niveau communal. Cependant, l'ampleur de la variabilité des données à cette échelle, ainsi que le nombre important d'unités à prendre en compte, entraînent des difficultés pour présenter des résultats synthétiques qui mettraient en évidence les grandes structures de l'espace étudié.

Notre objectif est de présenter une solution qui, pour reprendre l'expression de Gould, « laisse les données parler d'elles-mêmes ». Il s'agit de dévoiler les tendances régionales tout en conservant le détail des données communales. Pour cela, nous avons recours à des indices locaux d'autocorrélation spatiale qui permettent de mesurer les similarités (et les dissimilarités) entre voisins.

Les indicateurs locaux d'association spatiale (LISA), développés par Luc Anselin (1995) permettent en effet de mesurer le degré de ressemblance d'une unité spatiale avec ses voisins. On peut ainsi révéler les tendances régionales tout en conservant les valeurs locales, c'est-à-dire préserver l'information relative à l'hétérogénéité interne de ces zones.

Notre exposé présentera d'abord les LISA dans leur dimension statistique, étape nécessaire pour comprendre le fonctionnement de ces indicateurs et les limites inhérentes à leur construction. Nous proposerons ensuite une application où nous nous servons de ces indicateurs pour proposer un découpage spatial original du Tamil Nadu. Cet état, situé au sud-est de l'Inde, compte aujourd'hui plus de 62 millions d'habitants répartis sur plus de 16 000 communes, ce qui constitue un cadre remarquable pour tester l'efficacité de cette méthode.

ABSTRACT:

In India, quality of census data is good enough to allow geographic approach at micro-scale (village level) of the socio-economic phenomena. However, variability of data at this scale, as well as the huge number of units to take into consideration, make it difficult to present synthetic results, which would underline the major trends of the space structure.

Our aim is to present a solution, which, following Gould's expression, let "the data speak from themselves". The goal is to show the regional trend while keeping the local details visible. For this purpose, we use the local indicators of spatial association (LISA), which let us measure similarities (and dissimilarities) between neighbours. LISA have been developed by Luc Anselin (1995) and allow us to compute the level of likeness of one spatial unit with its neighbours (with the help of a weight matrix based on contiguity, distance or nearest neighbours).

We are then able to show regional trends without losing local values, i.e. to keep the information relative to the internal heterogeneity of the homogeneous zones.

Our paper will introduce the statistical dimension of LISA in order to understand how they run and their limits. We will then propose an application of LISA to create an original division of Tamil space, based on homogenous regions. This state, located in the South of India, has more than 60 millions inhabitants (in more than 16.000 villages), which constitutes a good case study to test the efficiency of this tool.

MOTS-CLÉS : autocorrélation spatiale, géostatistique, LISA, région, Inde

KEYWORDS : spatial autocorrelation, geostatistics, LISA, region, India

¹ Ce travail a bénéficié du soutien du groupe de travail EMIS (Espace et Mesure en Inde du Sud) dans le cadre de l'ACI « Géomatique, Espace, Territoires et Mobilité » (Programme Société de l'Information du CNRS). Je remercie les participants du colloque Théoquant pour leurs commentaires, et plus particulièrement François Bavaud.

² Cette présentation est le résultat d'un travail conduit au sein de l'UMR Géographie-cités.

Saisir la géographie des inégalités de développement

Le travail présenté ici constitue une tentative de réponse à un problème survenu lors d'une étude générale visant à comprendre les inégalités spatiales du développement dans le sud de l'Inde. L'étude se place dans le cadre de l'État du Tamil Nadu, et envisage le changement social à l'échelle des 16 085 villages qui le constituent. Il s'agit donc évidemment d'une approche qui exploite au maximum les outils de l'analyse spatiale quantitative.

La première étape fut la constitution d'un indice synthétique³, dit « indice de modernisation », qui reprend les différentes caractéristiques des villages tamouls, tant d'un point de vue sociologique (fécondité, alphabétisation, rapport de genre) qu'économique (orientation agricole, importances des secteurs secondaires et tertiaires).

La cartographie de cet indice (figure 1) laisse entrevoir différentes hypothèses explicatives. L'approche initialement suivie visait à introduire, comme élément explicatif majeur, le rôle de l'influence urbaine sur les inégalités spatiales observées en milieu rural. À titre d'abord comparatif (et ensuite complémentaire), la dynamique propre des villages a été envisagée. Nous avons donc pu proposer un modèle général des inégalités de développement au Tamil Nadu qui montrait d'une part, le rôle de la polarisation urbaine et d'autre part, la centralité propre des villages. La polarisation urbaine a été exprimée sous forme d'un indice d'enclavement des villages qui prenait en compte la distance à la ville la plus proche ainsi que la distance au réseau de communication (route et voie ferrée). La centralité propre des villages était exprimée par leur population, qui constitue classiquement un bon indicateur indirect.

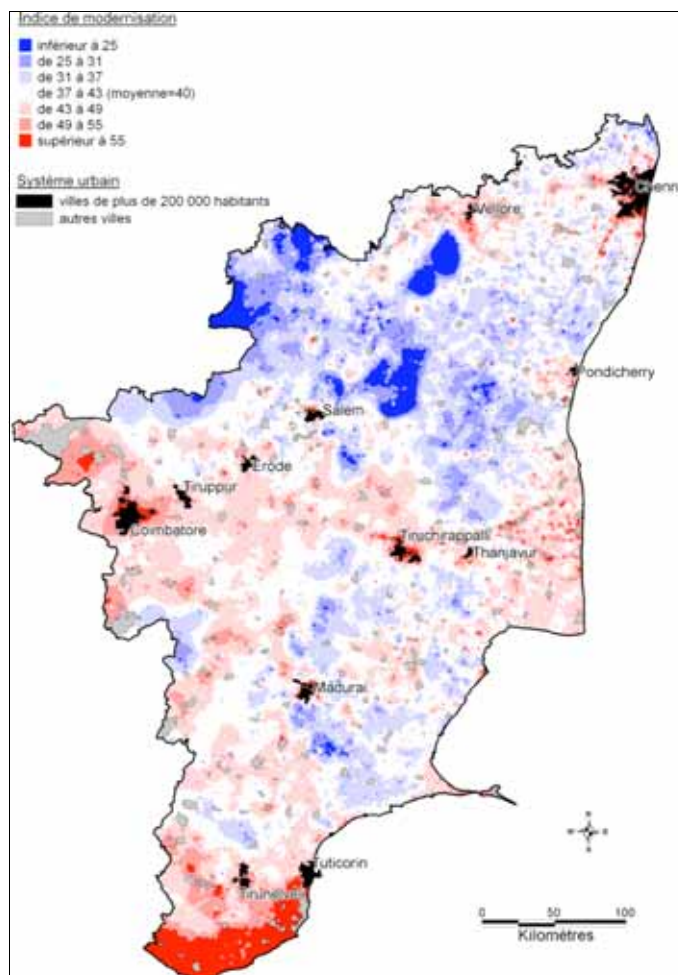


Figure 1 : les inégalités spatiales du développement au Tamil Nadu (cartographie au niveau villageois, lissé par krigeage)

³ La création de l'indice de modernisation repose sur une analyse en composante principale, dont le détail est exposé dans (Oliveau, 2004).

Un premier modèle géographique des différences spatiales de la modernisation au Tamil Nadu a donc été proposé, qui expliquait 35 % de la variance de l'indice :

$$I_m = 49,13 - 0,87I_e + 0,0014P + \theta$$

Avec I_m = l'indice de modernisation, I_e l'indice d'enclavement des villages (rôle de la polarisation urbaine corrigé par celui des axes de communication), P la population des villages (indicateurs indirects de leur niveau de centralité propre) et θ la part inexpliquée par le modèle.

Si le pourcentage de variance expliquée est important, la carte de l'indice de modernisation ajusté par ces deux éléments montre une structuration spatiale encore très nette (figure 2), qui n'a pas manqué de nous interroger. La trame administrative existante (district et taluks, équivalents aux NUTS⁴ 2 et 3 de l'Union Européenne) n'était pas à même d'expliquer les différences observées, et nous avons donc envisagé d'autres méthodes pour obtenir une régionalisation qui émerge de la structure de nos données. Nous avons par conséquent décidé de baser notre travail sur la ressemblance des villages entre eux, en « laissant les données parler d'elles-mêmes », pour reprendre l'expression de Gould (1981). C'est donc logiquement vers l'analyse exploratoire des données spatiales (ESDA – voir [Banos, 2001]) que nous nous sommes tournés pour mettre au point notre démarche de régionalisation.

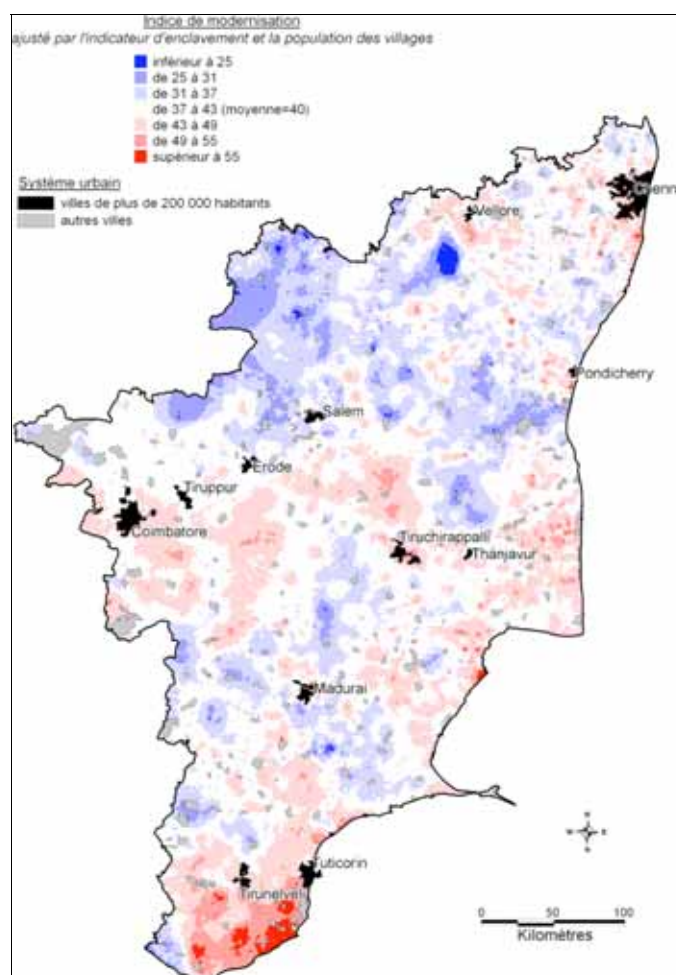


Figure 2 : les inégalités spatiales du développement au Tamil Nadu, après prise en compte de l'enclavement des villages et des différences de population (cartographie au niveau villageois, lissé par krigeage)

⁴ NUTS : Nomenclature des unités territoriales statistiques

Parmi les options proposées, l'utilisation des indicateurs locaux d'association spatiale (Local Indicators of Spatial Association — LISA) développés par Luc Anselin au milieu des années 1990 semblait prometteuse (Anselin, 1995). En effet, les LISA permettent le passage d'un indicateur global de la structure spatiale (indice d'autocorrélation spatiale) à une mesure locale des ressemblances (LISA)⁵. Nous allons présenter comment nous avons utilisé cet outil d'exploration spatiale pour en faire un outil de prise en compte de l'espace dans notre modèle.

1. La mesure de la structure spatiale

L'idée sous-jacente à notre entreprise est la capture, par les statistiques spatiales, de la structure visible sur la carte. Elle repose donc sur l'hypothèse que les valeurs en un lieu sont fonction des valeurs dans le voisinage, autrement dit que l'on est en présence d'autocorrélation spatiale⁶.

La mesure de l'autocorrélation spatiale des résidus d'une régression, et les biais qui y sont liés, sont aujourd'hui connus et les indices sont maîtrisés (Anselin et Griffith, 1988). Il reste néanmoins toujours à faire un choix qui concerne l'indice utilisé (parmi les plus connus : la semi-variance, l'indice de Moran, le coefficient de Geary) et la définition du voisinage pour le calcul (contiguïté, distance, plus proche voisins).

Notre choix s'est porté sur l'indice de Moran⁷ (I de Moran), qui semble globalement plus performant que les autres. D'abord, l'indice de Moran, comme le coefficient de Geary (c de Geary), se base sur la moyenne et est donc peu sensible aux valeurs aberrantes. De plus, il est moins fragile que le second car il mesure les écarts à la moyenne et non pas les écarts entre voisins. D'un point de vue statistique, les deux indices sont « raisonnablement » robustes, mais le I de Moran semble conservé un avantage sur le c de Geary (Cliff & Ord, 1981). Ensuite, et c'est un point important pour la vulgarisation des mesures d'autocorrélation spatiale, la similarité de son écriture et celle de ses résultats avec le coefficient de corrélation le rendent plus facile d'accès pour les néophytes. On ajoutera enfin un dernier argument, très pragmatique (parfois *ad hoc*), la mesure de l'indice de Moran est très répandue dans les logiciels géostatistiques, de statistiques, voire de systèmes d'information géographiques (SIG), ce qui a permis son développement et autorise le chercheur prudent à comparer ses résultats à partir de différents logiciels, qui produisent généralement leurs résultats au travers d'une « boîte noire », jamais très confortable.

Pour mémoire, l'indice de Moran se présente comme le rapport de la covariance sur la variance :

$$\frac{\sum_i \sum_j w_{ij} (z_i - \bar{z})(z_j - \bar{z})}{m} \div \frac{\sum_i (z_i - \bar{z})^2}{n}$$

Il s'écrit généralement sous la forme suivante :

$$\frac{n}{m} \times \frac{\sum_i \sum_j w_{ij} (z_i - \bar{z})(z_j - \bar{z})}{\sum_i (z_i - \bar{z})^2}$$

z_i = valeur de la variable au point "i" et de moyenne \bar{z}

i = unité de référence

j = unités voisines du point "i", définies par la matrice w_{ij}

n = nombre total d'individus dans l'échantillon (\sum_i)

m = nombre total de paires de voisins ($\sum_i \sum_j w_{ij}$)

w_{ij} = matrice de pondération

⁵ Le travail d'Anselin se place dans la continuité de celui de Getis et Ord (1992), qui avaient déjà cherché à mettre en place des indicateurs locaux d'association spatiale, mais sans lien avec les indicateurs globaux existants.

⁶ On se rappellera que l'autocorrélation spatiale est « la corrélation d'une variable avec elle-même (autocorrélation) attribuable à l'organisation géographique des données (spatiale) » (Griffith, 1992) (traduction libre)

⁷ L'indice de Moran est exposé dans un article de 1950 (Moran, 1950), mais c'est sous la forme présentée par Cliff et Ord (1981) qu'il a été popularisé. Il en est de même pour le coefficient de Geary (1954).

Le résultat du calcul du I de Moran est d'interprétation facile puisqu'il s'interprète comme un coefficient de corrélation classique. Il varie entre -1 (autocorrélation spatiale négative : les voisins ont des valeurs opposées) et +1 (autocorrélation spatiale positive : les voisins ont des valeurs semblables). On notera cependant que la valeur du I de Moran peut parfois être supérieure à +1 ou inférieure à -1 (Cliff & Ord, 1981). Il n'est donc pas strictement borné entre -1 et +1. La valeur $-1/(n-1)$, qui tend très rapidement vers 0, marque l'absence d'autocorrélation spatiale. En pratique on considère que l'autocorrélation spatiale est nulle lorsque l'indice se rapproche de 0.

Après avoir choisi un indice de mesure de l'autocorrélation spatiale, il faut définir une matrice de voisinage. Sans s'étendre, on rappellera que nous avons utilisé une matrice de voisinage basée sur la distance, pour trois raisons. D'abord la distance est une notion géographique, et se justifie donc dans notre approche, d'autant plus que la distance (à la ville, aux voies de communication) était déjà présente dans notre travail. Ensuite, la distance est cohérente avec l'idée sous-tendant notre approche, à savoir l'interaction spatiale, qui est généralement prise en compte par des modèles reposant sur la distance (Pini, 1992). Enfin, le voisinage, en fonction d'une distance, est de représentation aisée, puisqu'il s'agit d'un disque placé autour de l'unité considérée (principe des tampons dans les SIG).

Dans un voisinage de 2 km (15 721 paires de points), l'indice de Moran de l'indice de modernisation ajusté est de 0,37 et reste encore de 0,18 dans un voisinage de 20 km (avec 1 605 297 paires de points, l'autocorrélation spatiale reste significative), ce qui confirme la forte structuration spatiale observée.

2. Du global au local : les LISA

De cette mesure globale de la structuration, Luc Anselin a eu l'idée de développer des indicateurs locaux, les LISA⁸. Pour cela il s'appuie sur la description d'une statistique gamma (en fait, une statistique de produits croisés, voir [Getis, 1991]) qui s'écrit de la sorte :

$$\Gamma = \sum_i \sum_j w_{ij} v_{ij}$$

$$\text{Si } v_{ij} = \frac{(z_i - \bar{z})(z_j - \bar{z})}{\sum_i (z_i - \bar{z})^2} \times \frac{n}{m} \text{ on obtient alors le } I \text{ de Moran.}$$

En reprenant la formule du gamma, on voit que l'on peut facilement obtenir une mesure d'un gamma local :

$$\Gamma_i = \sum_j w_{ij} v_{ij} \text{ puisque } \Gamma = \sum_i \Gamma_i = \sum_i \sum_j w_{ij} v_{ij} .$$

Il décompose ainsi l'indice de Moran en isolant la part de chaque individu dans la mesure globale. Dans le cas du I de Moran, on doit d'abord effectuer une standardisation en ligne de la matrice de voisinage (opération qui consiste à rendre les poids des voisins égaux à 1 pour chaque unité). Le résultat de la standardisation en ligne est de rendre m égal à n , et donc à simplifier l'écriture de l'indice ($n/m=1$). Par ailleurs, la standardisation en ligne ne constitue pas un artifice statistique mais une correction nécessaire afin que chaque unité ait le même poids, quel que soit son nombre de voisin⁹. On obtient ainsi un I de Moran local (I_i) :

$$I_i = \frac{\sum_j w_{ij} (z_i - \bar{z})(z_j - \bar{z})}{\sum_i (z_i - \bar{z})^2} \text{ (d'après Anselin, 1995)}$$

⁸ Pour Anselin (1995, p.94), « un indicateur local d'association spatiale est une statistique qui répond à deux critères :

a) Le LISA de chaque observation donne une indication sur l'étendue d'un regroupement spatial significatif de valeurs similaires autour de cette observation.

b) La somme des LISA pour toutes les observations est proportionnelle à un indicateur global d'association spatiale. » (traduction libre)

⁹ À défaut, les unités ayant le plus de voisins (unités situées au centre et/ou dans des régions denses) auront plus de poids dans le calcul de l'autocorrélation spatiale. Ceci reste d'ailleurs valable aussi bien à l'échelle globale que locale.

L'intérêt de cette démarche est de proposer un lien direct de proportionnalité entre les indices locaux et l'indice global (Anselin, 1996, 1995 ; Oliveau, 2004) qui est de la forme :

$$\sum_i \Gamma_i = \gamma \Gamma \text{ avec } \gamma = n$$

Ce facteur de proportionnalité retenu n'est valable que si les variables sont préalablement standardisées et si la matrice de pondération est standardisée en ligne (ce qui est le cas). Son principal intérêt est de proposer alors une relation simple de la proportionnalité entre indice global et indice local. En effet, la moyenne des indices locaux est alors égale à l'indice global. Il sera donc possible de proposer pour des sous-espaces des indices de Moran globaux qui seront la moyenne pour ces sous-espaces des indices locaux, et qui pourront être comparés à l'indice global pour l'ensemble du Tamil Nadu, mais aussi entre eux.

C'est aussi grâce à ce lien de proportionnalité que l'on peut obtenir pour chaque village une estimation de sa ressemblance avec les villages voisins par rapport à sa ressemblance à l'ensemble des villages. On distingue alors quatre cas de figures :

- Les villages avec un indice fort dans un voisinage qui lui ressemble (autocorrélation spatiale positive et valeur de l'indice élevé). Situation résumée en « plus-plus » (high-high).
- Les villages avec un indice fort dans un voisinage qui ne lui ressemble pas (autocorrélation spatiale négative et valeur de l'indice élevé). Situation résumée en « plus-moins » (high-low).
- Les villages avec un indice faible dans un voisinage qui lui ressemble (autocorrélation spatiale positive et valeur de l'indice faible). Situation résumée en « moins-moins » (low-low).
- Les villages avec un indice faible dans un voisinage qui ne lui ressemble pas (autocorrélation spatiale négative et valeur de l'indice faible). Situation résumée en « moins-plus » (low-high).

On peut de plus tester le niveau de significativité des résultats par le biais d'un tirage aléatoire de type Monte-Carlo (Anselin, 1995). On obtient ainsi trois résultats complémentaires, qui vont nous permettre de réaliser une régionalisation (la démarche générale est illustrée par la figure 3). Tout d'abord, nous disposons des valeurs de l'indice ajusté (1^{ère} carte), puis de la valeur du LISA pour chaque unité (3^{ème} carte) et enfin d'une carte de la significativité (2^{ème} carte). L'addition de ces trois informations nous permet d'obtenir la 4^{ème} carte, à savoir celle des quatre cas de figures évoqués précédemment.

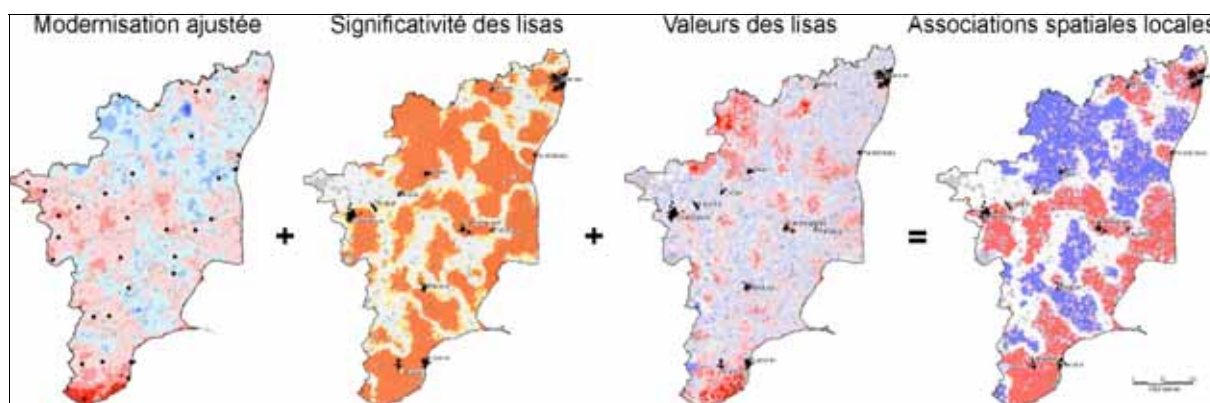


Figure 3 : une synthèse régionale issue du croisement de trois informations.

En isolant les groupes de villages par appartenance à une région dont la tendance est positive (plus-plus et moins-plus), négative (moins-moins et plus-moins) ou indifférenciée (LISA non significatifs), on obtient une carte régionale du développement au Tamil Nadu (figure 4) qui s'appuie sur la structure spatiale des résidus du modèle initialement proposé. On met ainsi en évidence des sous-espaces homogènes issus d'une sectorisation qui se base sur des niveaux locaux d'autocorrélation spatiale et des valeurs locales de l'indice. Il est ensuite possible de réintégrer cette information dans notre modèle de régression :

$$I_m = 45,58 - 0,73I_e + 0,0015P + \beta AR + \theta$$

avec I_e l'indicateur d'enclavement des villages, P la population des villages, βAR l'appartenance régionale¹⁰ et θ la part inexpliquée par le modèle.

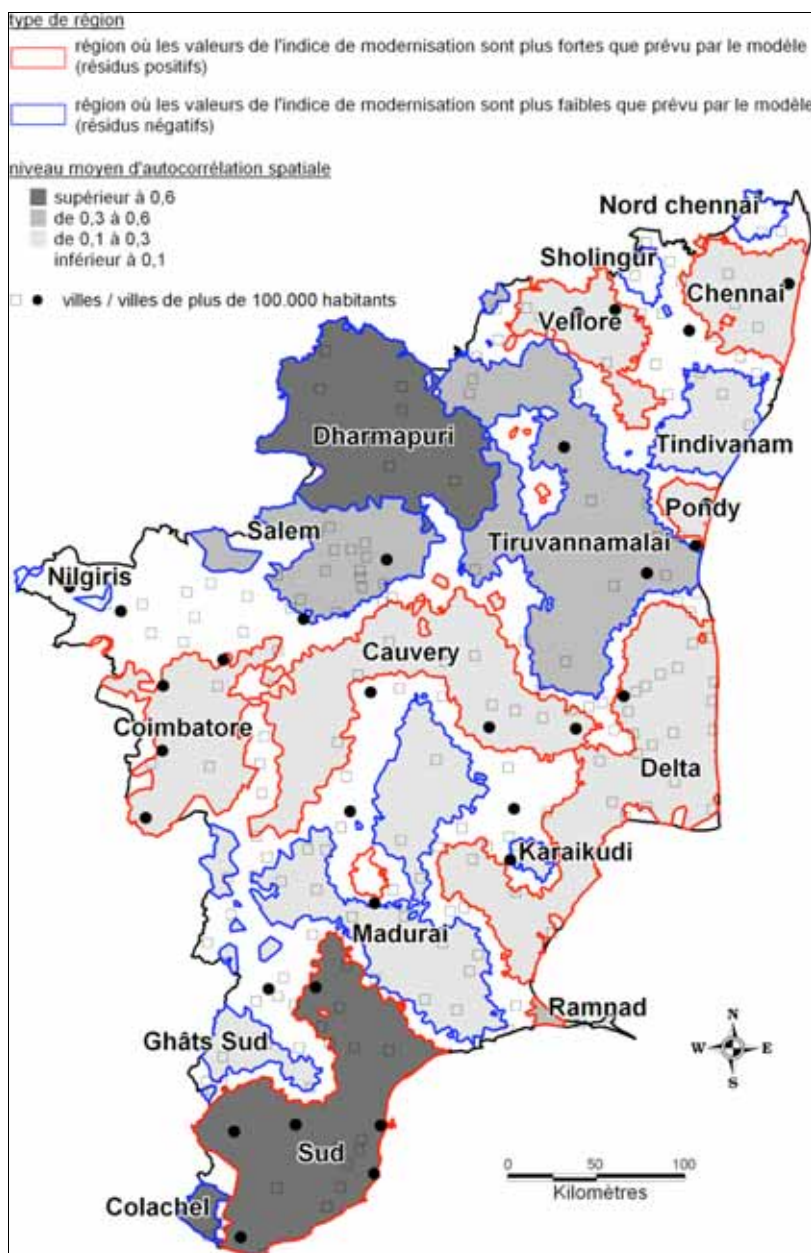


Figure 4 : la régionalisation issue de l'analyse des niveaux d'autocorrélation spatiale locale

Ce modèle explique 50 % de la variance de l'indice de modernisation, et surtout la quasi-totalité de sa variance spatiale. L'autocorrélation spatiale des résidus de ce dernier ajustement est quasi inexistante, puisqu'à 5 km, elle est déjà à peine supérieure à 0,1 et reste proche de 0 à partir du 15^{ème} kilomètre. Cela confirme le bien fondé de notre démarche de régionalisation.

On obtient ainsi un modèle explicitement spatial à deux dimensions que la figure 5 synthétise. La première dimension est verticale et concerne l'accumulation des effets, la seconde est horizontale et souligne

¹⁰ βAR est en fait de la forme « $\beta_1 AR_1 + \beta_2 AR_2 + \dots + \beta_n AR_n$ », avec « n » la région

l'indépendance des effets à l'intérieur de chaque niveau géographique. Ainsi, les trois niveaux géographiques distincts sont superposés et leurs effets, agissant à des échelles différentes, sont indépendants mais s'accumulent : toute chose égale par ailleurs, un village peuplé, près d'une ville et dans une région positive sera plus moderne que les autres. Mais si les effets s'additionnent, ils agissent aussi indépendamment les uns des autres : dans une même région, les villages les plus enclavés et/ou les plus petits seront moins modernes. De même, à degré d'enclavement égal et/ou dans une même région, les villages les plus gros seront plus modernes. Si deux villages sont de même taille et à même distance d'une ville, celui situé dans une région plus moderne le sera aussi. Les combinaisons possibles sont donc nombreuses, et ce sont elles qui créent la diversité observée à l'échelle du Tamil Nadu.

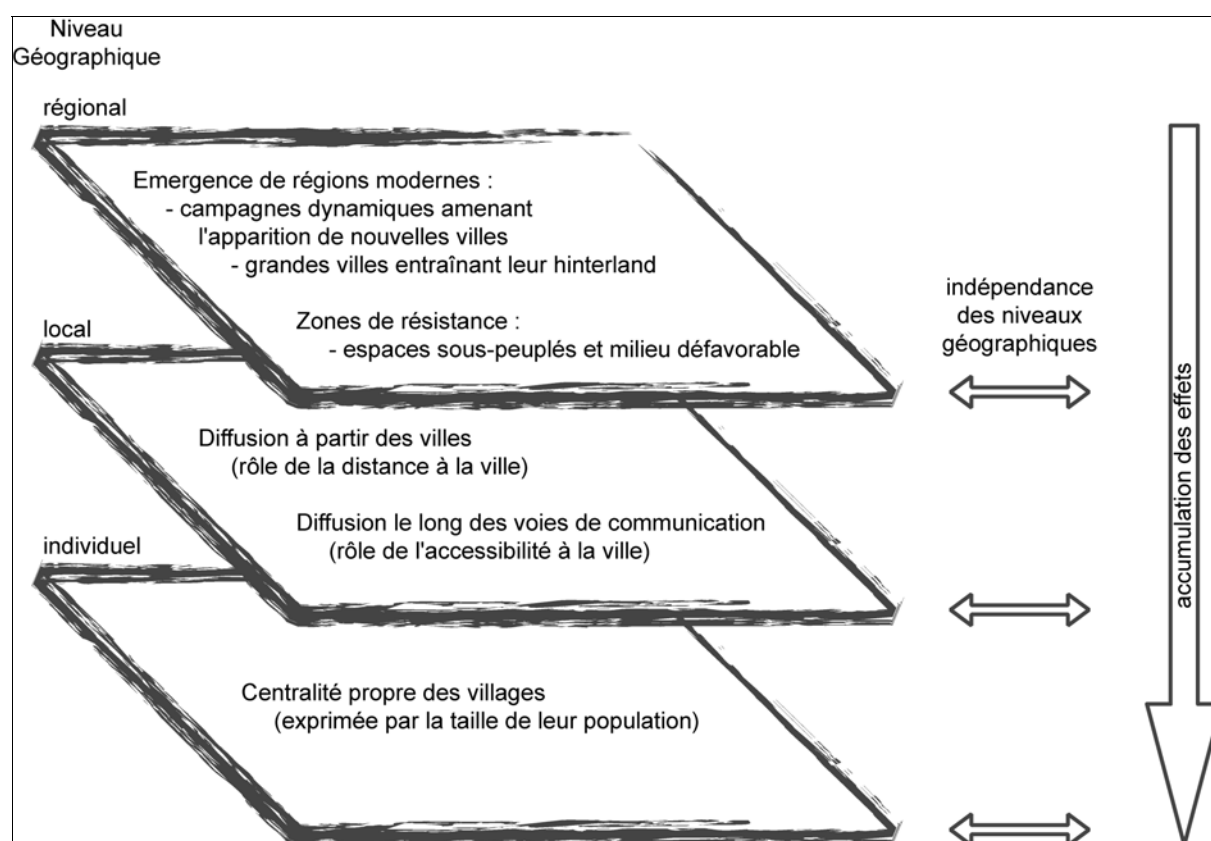


Figure 5 : un modèle explicitement spatial des inégalités de développement au Tamil Nadu

Quelques limites à évoquer

La méthode de régionalisation proposée ici fonctionne bien, et son application à un espace aussi vaste (1/3 de la surface de la France métropolitaine) et peuplé (60 millions d'habitants répartis dans 16 085 villages et 225 villes) le démontre bien. En utilisant les résultats de techniques issues de l'ESDA, donc prévues initialement pour explorer les données, on peut mettre en place des modèles descriptifs qui rendent bien compte des inégalités spatiales observées et nous permettent de proposer des modèles explicatifs de cette réalité.

Néanmoins, la proposition de régionalisation proposée en s'appuyant sur les LISA souffre encore de deux limites qu'il faut évoquer, en précisant qu'elles sont plus techniques que théoriques, et ne doivent pas rebuter les potentiels utilisateurs.

La première est inhérente aux outils géostatistiques actuels, à savoir non-intégration dans les SIG. Devoir exporter sa base de données pour lui appliquer les traitements géostatistiques et ensuite devoir la réimporter pour pouvoir continuer son utilisation est une opération toujours fastidieuse que l'on souhaiterait voir disparaître, d'autant plus que la base de données traitée ici est importante. Dans le même ordre d'idée, la lourdeur des traitements géostatistiques, directement proportionnel au nombre de voisins devient vite énorme (436 000 paires pour un voisinage de 10 km, 1,6 million pour un voisinage de 20 km). Ceci est d'autant plus vrai qu'il est difficile de prévoir à l'avance le voisinage adapté à notre travail, ce qui demande généralement une exploration spécifique des niveaux d'autocorrélation spatiale à différents pas de distance.

Mais c'est une limite plus fondamentale qui retiendra notre attention, liée à la construction même du I de Moran. L'indice de Moran mesure l'écart de la valeur d'un point et celle de ses voisins à celle de la moyenne. Par conséquent, il met en exergue les valeurs les plus éloignées de la moyenne, les valeurs trop proches se retrouvant alors avec des niveaux d'autocorrélation spatiale non-significatifs et n'influençant pas le niveau global d'autocorrélation spatiale. En d'autres termes, la mesure de l'autocorrélation spatiale avec le I de Moran prend d'abord en compte les valeurs éloignées de la moyenne, et non pas de manière indifférenciée toutes les valeurs similaires (proches les unes des autres). Il en résulte directement que les espaces où les LISA ne sont pas significatifs ne sont pas forcément des zones où la valeur des unités n'est pas liée à celle de ses voisines, mais plutôt des espaces où les valeurs sont toutes proches de la moyenne, ce que nous a confirmé une analyse détaillée des résultats obtenus (Oliveau, 2004).

Sans remettre en cause les résultats obtenus, cette limite est même un atout, elle permet de justifier notre régionalisation, qui isole des régions homogènes aux caractéristiques fortes. Cela nous rappelle, s'il en était besoin, que les limites construites des régions sont toujours arbitraires et ne sont que des artefacts nécessaires pour se saisir des territoires, dont les frontières sont toujours plus floues que ce que nous montrent les statistiques.

Bibliographie :

- ANSELIN, L., (1995), "Local indicators of spatial association – LISA", *Geographical Analysis*, Vol. 27, n°2, pp. 93-115.
- ANSELIN, L., (1996), "The Moran scatterplot as an ESDA tool to assess local instability in spatial association", in FISCHER, Manfred, SCHOLTEN, Henk K., UNWIN, David, *Spatial analytical perspectives on GIS*, Taylor & Francis, London, pp. 111-125.
- ANSELIN, L., (2003), *GeoDa 0.9 User's Guide*, Spatial Analysis Laboratory, Department of Agricultural and Consumer Economics and CSISS, University of Illinois, 125 p.
- ANSELIN, L., GRIFFITH, D.A., (1988), "Do spatial effect really matter in regression analysis ? ", *Papers of the regional science association*, Vol. 65, pp. 11-34.
- BANOS, A., (2001), « À propos de l'analyse spatiale exploratoire des données », *Cybergéo*, n°197.
- CLIFF, A.D., ORD, K.J., (1981), *Spatial processes. Models and applications*, Pion, Londres, 266 p.
- GEARY, R.C., (1954), "The contiguity ratio and statistical mapping". *The incorporated Statistician*, vol.5, n°3, pp.115–145.
- GETIS, A., (1991), "Spatial Interaction and Spatial Autocorrelation: A Cross-Product Approach", *Environment and Planning A*, Vol. 23, pp. 1269-77.
- GETIS, A., ORD, K.J., (1992), "The analysis of spatial association by use of distance statistics", *Geographical analysis*, Vol. 24, No. 3, pp. 189-206.
- GOULD, P., (1981), "Letting the Data Speak for Themselves", *Annals of the Association of American Geographers*, Vol. 71, n°2, pp. 166-176.
- GRIFFITH, D.A., (1992), "Spatial regression analysis on the PC: spatial statistics using SAS", department of geography and Interdisciplinary statistics Program, Syracuse University.
- MORAN, P.A.P., (1950), "Notes on continuous stochastic phenomena", *Biometrika*, Vol. 37, n°1/2, pp. 17-23.
- OLIVEAU, S., (2004) *Modernisation villageoise et distance à la ville en Inde du Sud*, Thèse de doctorat, UMR 8504 Géographie-cités, Université Paris 1 Panthéon-Sorbonne.
http://tel.ccsd.cnrs.fr/documents/archives0/00/00/76/81/index_fr.html.
- PINI, G., (1992), « L'interaction spatiale », in BAILLY, antoine, FERRAS, Robert, PUMAIN, Denise, *Encyclopédie de géographie*, Economica, Paris, pp. 557-576.