

**Mémoire de recherche**

=====

**CONCEPTION D'UN SYSTEME D'AIDE A LA  
DECISION EN LOCALISATION SPATIALE :  
CONCEPTS, FORMALISATION ET BASES  
METHODOLOGIQUES**

=====

Présenté et soutenu le 2 juillet 2004 par : **Igor AGBOSSOU**

Sous la direction de : **Professeur Christiane ROLLAND-MAY**

**Jury :**

**Christiane ROLLAND-MAY**  
**Patrice PAUL**  
**Michel PRUVOT**  
**Colette CAUVIN**

Professeur à l'ULP et directeur de recherche  
Professeur à l'ULP et responsable du DEA  
Maître de conférence à l'ULP  
Professeur à l'ULP et responsable de spécialité (invitée)

3, rue de l'Argonne  
F-67083 Strasbourg Cedex  
Tél : (33) 03 90 24 08 92  
Fax : (33) 03 90 24 08 80  
[www-ulp.u-strasbg.fr](http://www-ulp.u-strasbg.fr)

**Mention : Très bien**

**Année académique : 2003-2004**

Pour tout renseignement, prière me contacter à l'adresse suivante : [mous\\_igor@yahoo.fr](mailto:mous_igor@yahoo.fr)

## Tables des matières

<b>Remerciements</b> .....	4
<b>Résumé/Abstract</b> .....	5
<b>Introduction</b> .....	6
<b>Chapitre 1. Territoire et SIG</b> .....	8
1.1. Problématiques .....	8
1.1.1. Problématique générale .....	8
1.1.2. Problématique liée au DEA .....	10
1.2. Cadrage de la recherche .....	12
1.2.1. Notion de territoire .....	12
1.2.2. Approche systémique du territoire .....	13
1.2.3. Représentation orientée objet de la logique entité relation du territoire .....	14
1.3. Information géographique et base de données .....	16
1.3.1. Les données à référence spatiale .....	17
1.3.2. Les bases de données relationnelles .....	18
1.3.3. Les bases de données orientées objet .....	18
1.3.4. Nature des données à référence spatiale .....	19
1.4. Potentialités et limites du SIG dans le processus d'aide à la décision .....	20
1.4.1. Définition du SIG .....	22
1.4.2. Contributions du SIG .....	23
1.4.3. Limites du SIG .....	24
<b>Chapitre 2. La localisation spatiale : concepts, typologies et démarche d'analyse des besoins</b> .....	27
2.1. Concepts et définitions .....	27
2.2. Typologies .....	29
2.2.1. Typologie spatiale .....	29
2.2.2. Typologie thématique .....	30
2.3. Démarche d'analyse des .....	32
2.3.1. Analyse de la demande et de l'offre spatiale .....	32
2.3.2. Ajustement de l'offre à la demande .....	33
<b>Chapitre 3. L'aide à la décision en localisation spatiale : système expert et approche neuronale flous</b> .....	34
3.1. Concepts et définitions de l'aide à la décision .....	35
3.1.1. Aide à la décision .....	35
3.1.2. Décision .....	36
3.1.3. Décideur .....	36
3.1.4. Action, contrainte et critère .....	36
3.1.5. Le processus décisionnel .....	37
3.2. La chaîne de cohérence décisionnelle floue : application à la problématique de localisation spatiale .....	38
3.2.1. Contexte général de la chaîne de cohérence décisionnelle .....	38
3.2.2. La chaîne de cohérence : phase conceptuelle .....	40
3.2.3. La chaîne de cohérence : phase théorique et connaissances floues .....	42
3.2.4. La chaîne de cohérence : phase mathématique et modélisation des contraintes de localisation .....	47
3.2.5. La chaîne de cohérence : phase méthodologique .....	51

3.3. Méthodologie d'acquisition des connaissances .....	52
3.3.1. Problématique .....	52
3.3.2. Définition d'un neurone .....	53
3.3.3. Définition d'un réseau neuronal.....	54
3.3.4. Mise en œuvre du réseau neuronal flou .....	55
3.4. Le système expert .....	56
3.4.1. Définition du système expert .....	56
3.4.2. Définition de l'inférence .....	57
<b>Chapitre 4. Méthodologie de conception du système d'inférence pour l'aide en localisation :</b> .....	60
4.1. Extraction/Acquisition des connaissances .....	60
4.1.1. Généralités sur la base de données .....	60
4.1.2. Méthodologie d'extraction automatique des connaissances .....	62
4.1.3. Calibrage des connaissances par apprentissage du réseau .....	63
4.1.4. Synthèse .....	65
4.2. Le moteur d'inférence .....	66
4.2.1. Le chaînage arrière .....	66
4.2.2. Le chaînage avant .....	67
4.2.3. Synthèse : le chaînage mixte .....	68
4.3. La simulation : construction et évaluation des scénarii de localisation .....	69
4.3.1. La construction des modèles de simulation .....	70
4.3.2. Le générateur de scénarii .....	71
<b>Chapitre 5. Proposition de l'architecture du système d'aide à la décision en localisation spatiale : SADELOS</b> .....	73
5.1. Logique de conception du SADELOS .....	73
5.2. Architecture générale du SADELOS .....	74
5.3. Les flux de communication .....	78
<b>Conclusions et perspectives</b> .....	86
<b>Liste des figures</b> .....	87
<b>Références bibliographiques</b> .....	88

## REMERCIEMENTS

Notre reconnaissance va en premier lieu à notre directeur de recherche : **Madame Christiane ROLLAND-MAY**, Professeur à l'Université Louis Pasteur de Strasbourg I, pour les acquis scientifiques dont elle a su nous faire bénéficier en nous transmettant en permanence le goût de la recherche et la rigueur scientifique, à travers un encadrement aussi talentueux, malgré nos errances.

Nous souhaitons associer à ces remerciements le jury qui, convaincu de l'intérêt scientifique et opérationnel de la recherche entreprise, a accepté d'évaluer ce mémoire. Puissent ses commentaires et critiques, issus d'expériences et compétences diverses, nous être constructifs pour la suite de nos travaux de thèse.

## **RESUME**

La résolution des problématiques d'optimisation en terme de localisation spatiale constitue un axe de recherche à part entière en aménagement du territoire. Dans un contexte de réorganisation du territoire où la complexité est quasi permanente, la prise en compte du caractère flou du territoire et de l'information géographique est déterminante pour les décisions d'aménagement et le choix des actions de développement en matière de localisation spatiale. La conception d'un outil d'aide à la décision en localisation spatiale résultant de la combinaison d'un SIG et d'un système neuronal d'inférence flou s'avère donc très utile pour les acteurs ayant à charge la gestion du territoire d'une part, et contribue d'autre part à l'avancée de la recherche liée au développement des SIG "intelligents".

Ce mémoire de recherche qui constitue un préambule à nos travaux de thèse, pose les fondements théoriques et pratiques de la chaîne de cohérence épistémologique inhérente à la conception du Système d'Aide à la DÉcision en Localisation Spatiale (SADELOS).

**Mots clés** : Localisation spatiale, aide à la décision, SIG et système expert neuronal flous.

## **ABSTRACT**

Landsuitability evaluation constitutes a research orientation to whole share in land planning. In a context of land reorganization where complexity is quasi permanent, taking into account of fuzzy character of the land and geographical data associated is determining for planning decisions and the choice of development alternatives as regards spatial location. The design of a tool of decision-making in spatial location by combining GIS and reasoning module proves very useful for land planning actors on the one hand, and contributes on the other hand to advanced of research related to the development of "intelligent" GIS.

This research paper which constitutes a preamble for our thesis work, explains fluently the theoretical and practical bases of epistemological coherence chain to the design of the computerized decision-making system in spatial location (SADELOS).

**Key words**: Spatial location, decision making, fuzzy GIS and neural expert system.

## INTRODUCTION

Les problématiques de localisation spatiale constituent une préoccupation majeure de l'aménagement du territoire. Pour se soigner, les centres hospitaliers sont indispensables ; on fréquente les écoles et universités pour s'instruire ; pour se nourrir et se vêtir, on a constamment recours aux services et infrastructures du transport, de l'industrie et du commerce ; les réseaux d'assainissement interviennent dans la gestion de nos déchets etc. Un des défis des aménageurs est de parvenir à une adéquation optimale entre les besoins de nos sociétés et les réponses offertes par les travaux d'aménagement. Cette adéquation comporte de multiples facettes et difficultés : l'identification des besoins, le choix de ceux qui seront satisfaits, l'affectation de sites selon un certain nombre de contraintes et de critères. Toutes ces questions donnent lieu à de vastes débats de nature environnementale, économique, sociale et politique.

La question fondamentale qui nous intéresse, au travers de ces discussions, est l'adéquation spatiale entre l'offre et la demande, c'est-à-dire les choix de localisation optimale, au regard des contraintes environnementales et socio-économiques, des infrastructures et équipements en fonction de la nature des besoins qu'ils doivent satisfaire. Ce problème d'organisation spatiale des divers services devant répondre aux besoins des sociétés humaines est d'actualité aussi bien dans les pays développés que dans ceux en développement. Concrètement, les réflexions sur les localisations spatiales consistent à résoudre les problèmes suivants : faut-il modifier les localisations actuelles ? Quels équipements doit-on supprimer et où doit-on installer ceux qu'il serait nécessaire de créer ? Que gagne-t-on à redéfinir les zones desservies par les équipements existants ? A ces différentes questions il est maintenant possible d'apporter des éléments de réponses fiables, grâce aux nouvelles méthodes scientifiques d'aide à la décision.

En effet la disponibilité croissante de données, l'augmentation de la puissance de calcul de ordinateurs, la mise en application des acquis de l'intelligence artificielle et le développement des techniques d'implémentation des systèmes d'aide à la décision permettent d'aborder de manière efficace les problèmes décisionnels. Quant à la résolution des problèmes décisionnels à référence spatiale, en l'occurrence l'optimisation des localisations spatiales, l'intégration d'un système d'inférence, conçu par association de la logique floue et des réseaux de neurones, à un Système d'Information Géographique (SIG) présente beaucoup d'intérêts tant pour la recherche fondamentale que pour la recherche appliquée.

L'objectif de ce mémoire de DEA qui constitue un préambule à nos travaux de thèse, est d'exposer les principes conceptuels, la formalisation et de poser les bases méthodologiques de la conception d'un tel système.

Le document est organisé en cinq chapitres. Dans le chapitre premier, après avoir mis en évidence la nature complexe et floue du territoire, nous avons montré l'apport du SIG dans le processus de planification territoriale sans oublier de souligner ces limites en matière d'aide à la décision. Le second chapitre définit et explique les concepts relatifs à la localisation spatiale ainsi que la démarche analytique sous jacente. Le chapitre trois est consacré à la mise en application de la *chaîne de cohérence épistémologique* relative à l'aide à la décision en localisation spatiale. L'avant dernier chapitre définit et présente notre démarche méthodologique pour la conception du Système d'Inférence pour l'Aide en Localisation (SIAL). Le dernier chapitre fait état des aspects techniques et technologiques du couplage entre le SIG et le SIAL, les deux grandes composantes du Système d'Aide à la DÉcision en LOcalisation Spatiale (SADELOS).

## CHAPITRE I

# TERRITOIRE ET SIG

Ce premier chapitre a pour objectif d'exposer la problématique de la recherche entreprise, de la situer dans le contexte scientifique et opérationnel en regard des nouvelles méthodes et outils d'acquisition et de traitement de l'information géographique dans le cadre de l'aménagement du territoire. Après avoir énoncé et posé le cadrage de la recherche, nous définissons l'information géographique en mettant en exergue la nature *floue* des données à référence spatiale. Une attention particulière est ensuite accordée aux Systèmes d'Information Géographique (SIG) quant à leurs potentialités et limites dans le processus d'aide à la décision en matière d'aménagement du territoire.

### 1.1. Problématiques

#### 1.1.1. Problématique générale

La prise de conscience des effets néfastes de l'utilisation abusive de l'espace et l'émergence du concept du développement durable ont fait évoluer l'appréhension des problématiques sur le territoire. Désormais les aménageurs doivent reconsidérer ces problèmes dans toute leur globalité. Il est alors devenu évident, pour les acteurs de la gestion du territoire, que la maîtrise de cette complexité [LAA 00] nécessite l'utilisation d'une quantité d'informations plus grande sur le territoire ainsi que des méthodes et outils d'analyse, de modélisation et d'aide à la décision plus puissants et mieux adaptés.

Au nombre des préoccupations de l'aménagement du territoire, figurent les problématiques de localisation spatiale, objet de nos travaux de recherche. En effet, les questions de localisation se posent lors du choix du ou des meilleur(s) emplacement(s) pour l'implantation d'une infrastructure ou inversement, la ou le(s) meilleure(s) activités correspondante(s) à un site compte tenu de ses caractéristiques ; [JOE 95 ; BOR 02] notent à ce sujet que l'étude de la localisation amène à prendre une décision très importante par le fait qu'elle étendra ses effets

sur plusieurs années. Il est donc primordial que les méthodes choisies pour traiter cette question soient rigoureuses et fiables.

Sans être forcément exhaustif, on peut distinguer les problématiques de localisation spatiale suivantes :

- la recherche d'un site satisfaisant au mieux certaines conditions. Il s'agit par exemple de la recherche du meilleur site pour l'implantation d'un bâtiment administratif, d'une usine, d'une station d'épuration etc. ou inversement ;
- la recherche de plusieurs sites : [CHE 94] a abordé une problématique de ce type, où les surfaces devraient accueillir des ouvrages de lutte contre l'érosion. Dans ce cas, en plus de la recherche des meilleurs sites, il faudra tenir compte de leur nombre, leur ordre de réalisation ou leur combinaison, si toutes les surfaces ne sont pas destinées à la même utilisation ;
- la réalisation d'un tracé reliant deux points ou plus (dans le cas d'un réseau linéaire) : ce tronçon ou ce réseau peut être électrique, routier, d'approvisionnement en eau, ligne de tram et/ou train etc. Dans ce cas, doivent être pris en compte la forme du tronçon ou du réseau, sa localisation et le nombre de nœuds que doit comporter le réseau [MAT 03] ;
- la mise en place d'un réseau de polygones (ensemble de surfaces contiguës) : ce problème peut se poser lors de la conception de plans de zones, où l'ensemble d'une région est découpé en polygones définissant le type d'utilisation du sol choisi. L'objectif ne se restreint pas à attribuer à chaque surface les meilleures utilisations du sol, il faut aussi considérer le voisinage [ROL 00b ; MAT 03] des zones et l'organisation du plan proposés.

Parmi les différents types de localisation spatiale que nous venons de citer, quelques uns ont été partiellement traités avec l'analyse multicritère [EAS 93 ; PER 93 ; CHE 94 ; MAY 94 ; JOE 95 ; LAA 00]. Pour notre part, en raison de la nature floue [ROL 84 ; ROL 96 ; ROL 00a ; ROL 02 ; RUF 01] de certaines données à référence spatiale et compte tenu de la capacité des Systèmes d'Information Géographique (SIG) à stocker et à gérer les bases de données à référence spatiale [ARO 99 ; BIR 00 ; LAU 93 ; PAN 96 ; PRE 95], **nous proposons d'aborder ces problématiques de localisation spatiale par une approche hybride en couplant au SIG un système d'inférence que nous allons développer par association de la logique floue et des réseaux de neurones.**

En raison de l'ampleur de la recherche, et par souci de crédibilité scientifique et technique, notre mémoire de Diplôme d'Etudes Approfondies (DEA) en Systèmes Spatiaux et Environnement se consacre dès à présent au cadrage général de la recherche, à travers l'exploration des bases bibliographiques, méthodologiques, techniques et informatiques.

Dès à présent, pour introduire, cadrer et préparer au mieux le sujet de notre thèse, le DEA mené cette année aboutira donc à la conceptualisation, la formalisation et la mise en cohérence des fondements méthodologiques de notre travail de recherche.

### **1.1.2. Problématique du DEA**

La mise en application du concept de développement durable nécessite une approche globale du territoire et une démarche à la fois pluridisciplinaire et interdisciplinaire pour l'analyse et la résolution des problèmes de développement. Pour ce faire il est nécessaire de récolter, gérer, traiter et diffuser une masse considérable d'informations. Les responsables de l'aménagement et de la gestion du territoire sont ainsi confrontés à une colossale tâche de synthétisation pour organiser et transformer ces informations afin d'en extraire les données significatives et indispensables à une prise de décision éclairée et justifiée. Le redéploiement spatial des services et équipements au sein d'un territoire nécessite un effort de recherche d'une plus grande efficacité dans le nombre et la répartition. En général les bénéficiaires d'un service ne profitent d'un accès égalitaire [PEE 01] à ce service en raison de plusieurs facteurs tels que la distance, le coût de prestation, la satisfaction relative procurée par ce service. Ainsi les choix de localisation de services ou équipements font surgir deux problèmes fondamentaux. D'une part, on souhaite trouver un équilibre entre les coûts d'installation et de fonctionnement des équipements à mettre en place et les coûts d'accès des usagers à ces mêmes équipements. D'autre part, il faut parvenir à une organisation spatiale optimale des services tout en privilégiant la meilleure équité possible entre tous les usagers dans l'accès aux services. En d'autres termes, il s'agit de trouver "l'emplacement idéal pour l'équipement qu'il faut". Concrètement, la préparation et la réalisation d'un projet de localisation spatiale d'un équipement peuvent se résumer en trois interrogations :

- 1- Quels sont les besoins à satisfaire ? Cette question renvoie à la définition de la demande et de sa répartition géographique sur le territoire considéré.
- 2- Quelles sont les possibilités de réponses ? On se met alors dans une logique d'analyse de l'offre définie aussi bien par les équipements existants que par les modifications techniquement et économiquement réalisables que l'on peut y apporter ainsi que les nouveaux équipements qu'il conviendrait d'installer.
- 3- Comment ajuster l'offre à la demande ? En d'autres termes, il faut proposer les meilleures options qui tiennent compte des contraintes du développement durable.

L'évolution remarquable des technologies de l'informatique au cours de ces dernières décennies ouvre de nouveaux horizons et les outils qui en résultent ont pris une place importante dans le travail et la façon de faire des aménageurs. Ceux qui pensent l'espace en termes de développement s'appuient largement sur ces techniques [LAA 00]. Les collectivités territoriales, confrontées aux nouveaux défis que constituent ces mutations, introduisent

massivement les nouvelles technologies dans leurs activités. L'intérêt croissant pour la géomatique, ses méthodes et ses concepts, devient alors indéniable. Pourtant, l'apparente limpidité et le succès affichés par les projets géomatiques<sup>1</sup> [ROC 95 ; CAR 97] masquent la complexité des mécanismes de diffusion, d'appropriation et d'utilisation des techniques de l'information géographique. Ces technologies sont encore le plus souvent considérées par les géographes comme des outils au service de la réflexion sur l'espace. Elles ne sont pas de véritables objets d'étude. Malgré tout, certaines initiatives du NCGIA (National center for Geography Information and Analysis) aux Etats-Unis se concentrent sur les implications humaines et sociales des SIG, les aspects d'aide à la décision et à la négociation et sur les questions de représentations spatiales dans les SIG. Mais alors, dans quelle mesure l'étude du développement d'un système d'information géographique intelligent s'inscrit-elle dans une problématique de géographie sociale ? Comment aborder avec l'œil du géographe les aspects techniques sous-jacents ? Ces considérations nous ont naturellement conduit à construire notre réflexion en prenant davantage en compte la dimension spatiale dans sa globalité, trop souvent oubliée par les recherches menées dans le domaine de la technologie de l'information relative à l'aide à la décision.

Prendre des décisions en aménagement du territoire nécessite non seulement de posséder des informations nécessaires relatives au territoire dans sa globalité mais encore de disposer de méthodes de raisonnement efficaces. Une bonne décision en matière d'aménagement du territoire doit donc aboutir, à court, moyen et long terme, à maintenir ou améliorer globalement l'état du système que constitue le territoire dans la perspective d'un développement durable des sociétés humaines en rapport avec leur espace géographique. Pour essayer de répondre à un tel besoin, la plupart des décideurs ayant à charge la gestion du territoire ont de plus en plus recours à l'utilisation des SIG en raison de leur puissance et de leur excellence en tant qu'outils de stockage et gestion de base de données à référence spatiale. Mais trop souvent, pour des raisons de manque de fonctionnalités puissantes en matière d'aide à la décision, nombre d'utilisateurs se contentent d'utiliser leur SIG comme outil de cartographie automatique [LAU 93]. Actuellement les SIG répondent en grande partie aux besoins des usagers.

Mais la généralisation de leur utilisation et celle des nouvelles technologies informationnelles et méthodes scientifiques d'aide à la décision révèlent d'autres besoins en termes de fonctionnalités, beaucoup plus sophistiqués, afin de leur conférer des capacités de raisonnement intelligent. En ce qui concerne les SIG, il est tout à fait possible d'intégrer,

---

<sup>1</sup> La géomatique est la science et la technologie de l'information pour faciliter l'acquisition, le traitement et la gestion des données sur le territoire. Ses racines sont "Géo", qui veut dire Terre, et "matique" vient de informatique, soit le traitement automatique de l'information.

entièrement ou de façon biunivoque, de telles technologies et certaines sont déjà partiellement opérationnelles. En effet les problèmes de recherche et de localisation optimale de sites, de chemins, de zones, de découpage avec ou sans contraintes et/ou critères quantitatifs et/ou qualitatifs, de simulation de décisions d'aménagement du territoire, exigent l'utilisation d'outils de raisonnement spatial alliant non seulement les aspects déductifs, mais aussi intégrant des techniques de “*recherche opérationnelle soft*” [ROL 00a]. Ceci implique le développement d'un module d'inférence spatiale ou de raisonnement spatial basé sur des outils mathématiques et informatiques. A l'heure actuelle, aucun SIG commercialisé n'intègre toutes ces fonctionnalités dites “intelligentes” [COR 93, BIR 00 ; BER 02 ; ARE 03] à moins d'acquérir des modules optionnels qui coûtent chers et sont peu spécialisés [BOR 02 ; MAI 02]. Dans ce contexte, l'approche par l'utilisation de réseau neuronal flou pour faire évoluer les SIG vers de véritables systèmes d'aide à la décision en localisation spatiale nous paraît plus appropriée.

## **1.2. Cadrage de la recherche**

### **1.2.1. Notion de territoire**

On utilise souvent les termes *territoire* et *espace* de façon interchangeable sans leur donner une signification toujours très précise. Cela explique en partie le flou qui entoure ces différentes notions et aussi le fait qu'on trouve dans la littérature de nombreuses définitions qui ne sont pas toujours concordantes, quand elles ne sont pas franchement contradictoires. Nous ne voulons pas ici forger une définition définitive de ces concepts ; mais il nous semble nécessaire, à ce stade de notre réflexion, d'apporter quelques éléments de définitions basés sur quelques ouvrages de référence en géographie.

La notion la plus englobante est certainement celle d'espace. Brunet [BRU 92 ; ECK 96] la définit comme l'ensemble des dimensions dans lesquelles se déroulent nos actes, nos représentations, nos relations, nos sensations. Il distingue l'espace terrestre qui correspond à la surface du globe, l'espace de l'humanité (nommé *écoumène*) et l'espace géographique. Ce dernier correspond à l'étendue terrestre utilisée et aménagée par les sociétés en vue de leur reproduction au sens large, soit non seulement pour se nourrir et s'abriter, mais aussi dans toute la complexité des actes sociaux. Il comprend l'ensemble des lieux et des relations. Il est de ce fait à la fois un système de relations et un produit organisé résultant des interactions entre la nature et les sociétés.

Le territoire renvoie à la notion de milieu qui recouvre elle-même celle d'espace [ECK 96 op. cité] géographique en lui ajoutant l'idée de délimitation. En effet le territoire se distingue du milieu par le fait que les critères de délimitation retenus permettent de lui reconnaître des limites plus précises (d'un point de vue juridico administrative) et font référence à des notions

d'appartenance. Il résulte de la "projection", sur un espace donné, des structures spécifiques d'un groupe humain, qui incluent le mode de découpage et de gestion de cet espace, ainsi que son aménagement. Le territoire constitue de ce fait l'espace au sein desquels nous pouvons intervenir concrètement afin de préserver ou d'améliorer notre cadre de vie en remédiant aux problèmes de développement, qu'ils soient de nature sociale, politique, économique ou écologique. Lorsque les aspects juridico-administratifs sont prépondérants (ce qui correspond à la réalité), le territoire s'identifie à une portion de l'espace géographique circonscrite par des limites politico administratives [PRE 95]. Ce découpage juridico administratif ne respecte malheureusement pas les limites des sous systèmes qui le composent. Un modèle systémique du territoire doit toutefois s'affranchir de cette contrainte car il porte sur la définition des phénomènes et des processus au niveau conceptuel. Ce niveau de réflexion est nécessaire car il permet de comprendre le comment et le pourquoi de l'état de fonctionnement d'un territoire. Il est également le support du choix des phénomènes que l'on va représenter et des données que l'on doit intégrer dans un système d'information à référence spatiale pour aider à la gestion et à l'aménagement de ce territoire. Aussi dans la suite de notre développement nous utiliserons le territoire comme terme générique, sachant qu'il peut couvrir des réalités différentes. Et c'est ce territoire, réalité abstraite aux multiples facettes, que nous analyserons à l'aide de la systémique.

### **1.2.2. Approche systémique du territoire**

La *systémique* est fondée sur l'idée que, sous l'apparente diversité des phénomènes, il existe des traits communs, des régularités, des invariants que partage un grand nombre de **systèmes** qu'ils soient physiques, écologiques, sociaux ou cognitifs. En effet la promotion d'un développement durable a placé le territoire au croisement de plusieurs systèmes : écologique, social, économique et politique. Le rapport de la CMED<sup>2</sup> montre clairement que l'appréhension des problèmes qui se posent actuellement aux collectivités territoriales ne peut pas s'effectuer dans un seul de ces systèmes.

Dans "La théorie générale du système", Lemoigne [LEM 77] considère un système comme une unité non séparable constituée de composantes en interactions. Comme une unité indivisible ne peut être réduite à la somme de ses parties, et comme une relation ne peut être ramenée à des composantes, nous sommes amenés à considérer le territoire comme un système formé de trois catégories fondamentales : les composantes spatiales et les phénomènes sociaux (objets), les relations et le tout. Schwarz [SCH 94] propose une formalisation d'un système général dont nous nous inspirons. Il reprend ainsi ces trois catégories fondamentales et pose que tout système existant possède trois plans d'existence :

---

<sup>2</sup> Commission sur l'environnement et le développement

- un plan physique qui comprend les composantes, soit ce que nous désignons par “*objet*”<sup>3</sup> dans le cas du territoire ;
- un plan logique, parfois non visible par nos sens, qui est celui des relations entre les objets, qui détermine le fonctionnement et l'évolution du système ;
- un plan holistique qui est celui où le système est considéré comme un tout complexe, une unité indivisible sans laquelle les objets et les relations ne reproduiraient pas l'évolution observée. Autrement dit, le *tout*, soit en quelque sorte l'identité du système territorial qui émerge des interactions entre les objets, donne un sens à ces interactions.

Cette formalisation de l'approche systémique du territoire permet de mieux analyser les aspects [PIC 91 ; BER 92] touchant à l'évolution de la réalité. Ainsi agir sur le territoire nécessite de connaître et de comprendre les modes d'articulation et de complémentarité des différents sous-systèmes qui le composent ; ce qui permet, dans le cadre de l'aménagement et de la planification, de pouvoir analyser et évaluer voire prédire l'impact d'une action isolée ou d'un ensemble d'actions sur l'évolution du territoire. Nous pouvons alors concevoir et développer des modes d'interventions plus efficaces, car mieux adaptés aux caractéristiques des systèmes, qu'ils soient économiques, techniques, sociaux ou environnementaux.

Une action, aussi localisée soit-elle sur un territoire, a des répercussions à tous les niveaux identifiés. Pour trouver des solutions réellement efficaces, il convient de considérer ceux-ci dans leur globalité et leurs interactions pour en dégager une fusion adaptée aux exigences d'une approche intégrée et opérationnelle. De ce fait les acteurs territoriaux devraient s'appliquer plus à identifier et modéliser le fonctionnement et le comportement du territoire dans son ensemble.

### **1.2.3. Représentation orientée objet de la logique entité relation du territoire**

Vouloir donner une représentation du territoire, c'est se heurter au problème de l'objectivité et du rôle du modélisateur observateur. On peut donc admettre qu'il puisse, potentiellement, exister autant d'observateurs que de représentations de territoire. Pour notre part, la modélisation du territoire porte sur la représentation des phénomènes (objet) et des processus par le biais d'entités et de relations [PRE 95 ; PAN 96 ; FIS 01], en définissant et en décrivant leurs caractéristiques thématiques, spatiales et temporelles, et en mettant en exergue les liens de causalité attachés aux relations.

En conséquence nous considérons la représentation systémique de territoire comme un modèle interdisciplinaire du territoire fondé sur la détermination des phénomènes et des relations qui le composent et qui sont pertinents dans le cadre de la gestion, de l'aménagement

---

<sup>3</sup> Un objet correspond à la représentation informatique d'une entité au sein d'une base de données. Dans le cadre des bases de données à référence spatiale, l'objet se compose d'un *objet thématique* et d'un *objet graphique*. Dans un système qui intègre la gestion de la temporalité, l'objet peut présenter plusieurs *versions d'objet*.

et de la planification du territoire. Elle en donne une définition et une description adaptée aux exigences d'une approche interdisciplinaire (figure 1.1).

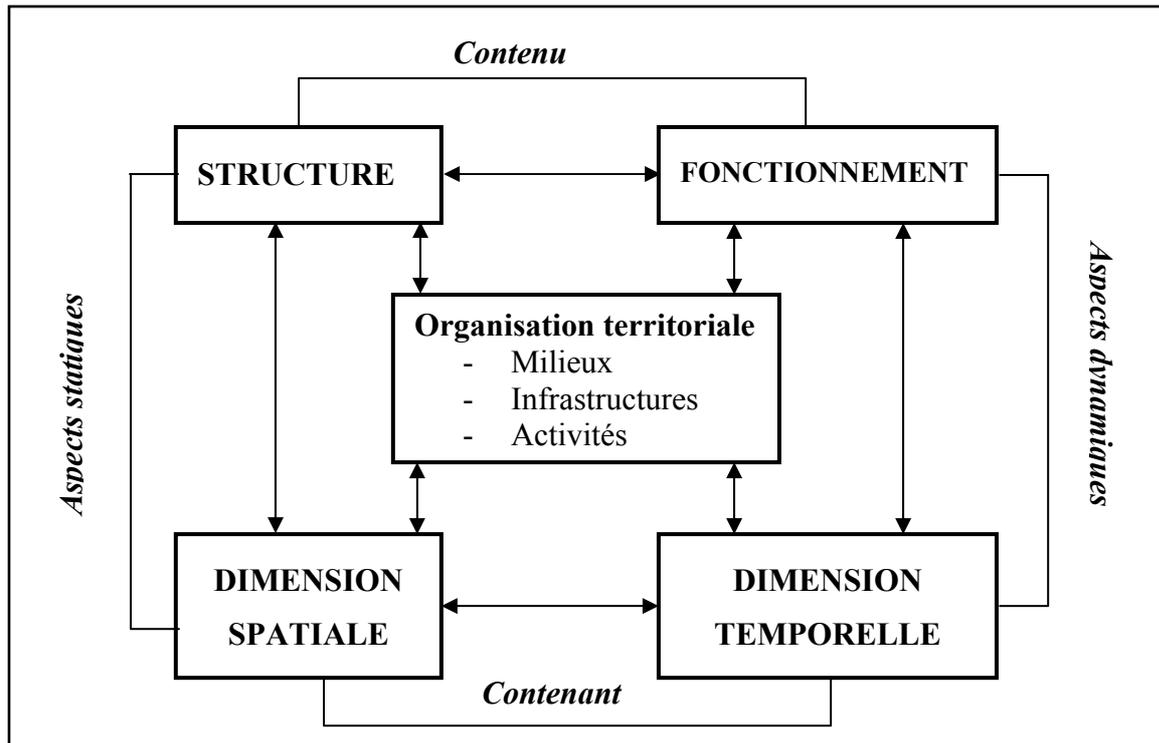
Cette représentation du territoire présente l'originalité de proposer une représentation des aspects statiques et dynamiques du territoire qui repose sur l'identification du *contenu* et de *contenant*, qui sont fortement interdépendants dans leur participation à l'organisation territoriale traduite par l'existence des activités, des infrastructures et des milieux (milieu naturel, le paysage et le milieu anthropique). Si le contenu exprime la structure et le fonctionnement du territoire, le contenant quant à lui prend en compte sa dimension spatiale et celle temporelle.

L'analyse des processus relationnels fait l'objet de travaux de recherche dans de nombreuses disciplines dans le but de comprendre et identifier les séquences logiques et de pouvoir reproduire leurs résultats, soit l'évolution du territoire. De manière générale, on remarque que les processus peuvent s'appréhender de façon simple au travers d'interactions entre les phénomènes [FRO 91]. Un objet peut tout d'abord avoir une action sur lui-même (autocontrôle). Il peut également interagir avec d'autres objets, de manière directe ou de manière indirecte. Ces interactions peuvent être à sens unique ou réciproque, une interaction pouvant être directe dans un sens et indirecte dans l'autre. La réalisation de ces interactions va modifier l'état des objets. Il en résulte des changements dans le système territorial qui provoquent son évolution vers un autre état. Dès peut se faire la modélisation spatio-temporelle des processus au sein du territoire.

En général, la modélisation des aspects statiques ou structurels et des aspects dynamiques ou fonctionnels est effectuée de façon séparée. Les modèles conceptuels de données<sup>4</sup> (MCD) développés dans le cadre des applications géomatiques ne prennent en compte généralement que les aspects statiques. La modélisation orientée objet ouvre des perspectives intéressantes pour la prise en compte du comportement du phénomène dans le territoire au niveau de l'entité qui le représente. La représentation conjointe des aspects statiques et dynamiques par une approche orientée objet mettant en exergue la logique relationnelle des objets du territoire est devenue indispensable [LAU 93 ; OPEN 97 ; OPEN 01 ; LAA 00] pour élaborer les systèmes d'information à référence spatiale (SIRS) capables de fournir les informations synthétiques nécessaires aux besoins de l'aménagement et de la planification [BRO 84 ; ROL 00b].

---

<sup>4</sup> Le modèle conceptuel des données (MCD) a pour but d'écrire et de décrire de façon formelle les données qui seront utilisées par le système d'information. Il s'agit donc d'une représentation des données, facilement compréhensible, permettant de décrire le système d'information à l'aide d'entités déclinées en objets et les relations.



**Figure 1.1** : Approche systémique orientée objet du territoire

On voit bien que l’approche modélisatrice du territoire correspond au niveau de définition et d’organisation des phénomènes en des ensembles dont les contenus et les frontières se révèlent parfois imprécis et incertains. Il se pose alors la question de crédibilité et fiabilité des données et même du processus décisionnel. Le raisonnement flou et les fondements de la logique floue ouvrent des perspectives [BOU 95 ; ROL 96 ; BOU 99 ; ROL 00a ; ROL 00b ; OPE 01] salvatrices sur la question. Nous y reviendrons en détail dans la section 2.3 du second chapitre.

### 1.3. Information géographique et base de données

#### *Définition de l’information géographique*

L’information géographique peut être définie d’une façon générale, en se basant sur la simple étymologie, comme un ensemble reliant une information relative à un objet ou un phénomène sur la terre, décrit plus ou moins complètement par sa nature, son aspect et ses attributs (cette description peut inclure des relations avec d’autres objets ou phénomènes) et sa localisation sur la surface de la terre dans un système de référence donnée. Se référant à l’aspect dynamique du territoire, on pourrait définir l’information géographique comme une “donnée quantitative ou qualitative décrivant un objet ou un phénomène réel, localisé dans l’espace à un moment précis”. Une meilleure connaissance de l’espace géographique des territoires pour

des fins d'aménagement et de gestion nécessite la collection d'une masse considérable d'informations relatives aux objets phénomènes qui constituent le système territorial. Cette masse d'information est ensuite organisée en une base de données.

### **1.3.1. Les bases de données à référence spatiale**

Selon Laurini et Thompson [LAU 92], une base de données est une collection de données non redondantes représentant les besoins d'un ensemble d'utilisateurs. C'est un ensemble d'informations organisé et structuré en fonction des besoins des utilisateurs et des phénomènes que ces informations traduisent. De manière générale une base de données présente les avantages suivants :

- gestion d'un grand volume de données ;
- organisation et structuration des données ;
- facilité de saisie, de consultation et de mise à jour des données ;
- rapidité d'extraction et de traitement des données ;
- sécurité et protection des données.

Lorsqu'une donnée se rapporte à un objet ou un phénomène localisé dans l'espace, elle est qualifiée de donnée à référence spatiale. Avec l'augmentation considérable des capacités de calcul des systèmes informatiques il devient de plus en plus relativement aisé de mettre en place des bases de données possédant une référence spatiale. Les bases de données à référence spatiale (BDRS) peuvent se définir comme un ensemble d'informations géoréférencées organisé en vue de son utilisation par des programmes, pouvant décrire plusieurs thèmes à différentes échelles sur un territoire. Elles se distinguent par les aspects suivants :

- les BDRS contiennent des données alphanumériques et des données graphiques. Ces dernières permettent de localiser et de définir la forme géométrique des phénomènes représentés ;
- les liens de voisinage, d'adjacence, d'intersection, d'appartenance et d'inclusion qui relient les phénomènes du territoire, nécessitent la prise en compte de la dimension topologique<sup>5</sup> de l'information spatiale. Celle-ci définit le niveau relationnel des données spatiales en décrivant la manière dont les géométries sont organisées pour reproduire l'organisation spatiale des phénomènes ;
- l'échelle de perception de la réalité n'est pas unique, ce qui nécessite de disposer de plusieurs niveaux de définition de l'information spatiale ;
- l'existence de deux modes de structuration de l'information géographique, à savoir les modes matriciel et vectoriel, nécessite de les faire cohabiter et de les combiner afin de

---

<sup>5</sup> La topologie est une théorie mathématique dont l'objectif général est d'étudier de manière qualitative les propriétés des espaces, comme les notions de limite ou de continuité.

pouvoir intégrer et exploiter l'ensemble des informations spatiales répondant aux besoins des utilisateurs ;

- le besoin d'accéder aux informations au travers des objets graphiques nécessite le développement de structures de représentation physiques appropriées.

Il existe actuellement deux grands types de base de données : les bases de données relationnelles et les bases de données orientées objet.

### **1.3.2. Les bases de données relationnelles**

Les bases de données relationnelles sont les plus répandues. Elles sont composées de fichiers alphanumériques reliés entre eux ; dans le cadre de la géomatique, ces fichiers font également référence aux objets graphiques. Leur construction se fait à l'aide de modèle relationnel lequel est un modèle logique de données. Il transcrit les entités et les relations définies au niveau conceptuel sous forme de relations logiques, soit d'un tableau à n colonnes qui correspondent aux attributs [LAU 93]. Le formalisme généralement utilisé représente la relation par un schéma qui est défini par le nom de la relation suivi de la liste des attributs, liste dans laquelle on en relief le ou les attributs qui servent d'identifiant. Ainsi, une base de données relationnelle contient un grand nombre de liens qui ne servent qu'à renforcer l'intégrité des données.

Il faut remarquer, que contrairement à une idée largement répandue, ce n'est pas la possibilité de mettre des tables en "relation" qui donne leur nom aux bases relationnelles. Une véritable base de données relationnelles doit posséder les propriétés suivantes :

- 1- toutes les informations sont organisées selon des structures mathématiques cohérentes appelées "relation" ;
- 2- tous les liens sont établis au travers de données et
- 3- toute opération globale appliquée à des relations doit donner de nouvelles relations (propriété appelée clôture).

Cette mauvaise compréhension de la technique relationnelle et de la signification de ce qu'est une relation fait que de nombreux produits commerciaux ne sont pas de véritables systèmes relationnels [COX 97].

### **1.3.3. Les bases de données orientées objet**

Les bases de données orientées objet (BDOO) sont relativement plus récentes. La technique orientée objet est issue des recherches portant sur les langages de programmation. Son application dans le domaine de la gestion des données à référence spatiale est très prometteuse [MIL 93 ; PRE 95 ; GRA 97 ; COX 97 ; MAI 02]. En effet l'orienté objet permet une représentation informatique plus proche de la perception que nous avons de la réalité et une

meilleure correspondance entre les entités définies au niveau conceptuel et les objets informatiques de la base de données. Il se fonde principalement sur le processus d'encapsulation, soit le regroupement de toutes les données (description graphique, sémantique, thématique et comportementale) concernant une entité au sein d'un seul et même objet informatique. Le comportement de l'objet au sein de la base de données, soit notamment les opérations qu'il y peut y déclencher, est décrit au travers de la notion de "méthode". D'une manière générale les techniques orientées objet s'appliquent à la conceptualisation, à la programmation, à la structuration des données et à la construction des systèmes. Elles présentent trois types d'avantages fondamentaux quant à l'utilisation des données territoriales.

♠ Les BDOO aident à combler le vide sémantique entre les objets et concepts du monde réel et leur représentation dans une base de données. Elles peuvent exprimer tout ce que le modèle entité relation est à même d'exprimer, formant un surensemble de données hiérarchiques et organisées en réseaux tout en enrichissant chacune d'elles du point de vue sémantique.

♠ Les BDOO donnent généralement de meilleures performances lorsqu'il s'agit de traiter des objets ou des relations complexes, ceci principalement du fait qu'il n'y a aucune nécessité à disposer les gros objets pour les stocker dans les tables normalisées, puis à les réassembler à l'exécution via des opérations de jointure lentes, c'est-à-dire la connexion informatique entre donnée. En outre, les quantités d'espace mémoire et d'espace disque occupées pourront être nettement inférieures à celles que demandent les systèmes relationnels du fait de leurs nombreux fichiers d'index.

♠ Les BDOO offrent un meilleur contrôle sur les requêtes (spatiales et thématiques) et ont une intégrité référentielle intégrée et une sémantique généralement plus riche que les autres systèmes. Elles partagent ces avantages avec les systèmes reposant sur les modèles sémantiques et les bases de données entité relation. L'identité d'objet résout le problème des n-uplets pendants<sup>6</sup> et supprime la nécessité de contrôle d'intégrité référentielle, puisque les objets sont référencés et non directement adressés. La sémantique plus riche repose sur la faculté de construire des types et d'encapsuler des contraintes au sein des objets à tous les niveaux du réseau.

#### **1.3.4. Nature des données à référence spatiale**

Les données à référence spatiale se rapportent à des objets ou des phénomènes par leur position et leur forme sur un territoire ; elles ont des propriétés uniques, leur conférant une

---

<sup>6</sup> Les n-uplets pendants sont des n-uplets non représentés dans toutes les relations (normalisées) qui contribuent à la représentation d'un objet, et qui risquent donc d'être perdus lorsqu'une opération de jointure est effectuée.

nature particulière qu'il convient de cerner. En effet ce qui distingue les données à référence spatiale, c'est qu'elles sont associées à des phénomènes dans un espace continu. Dans la pratique ces données sont multisources : certaines sont directement mesurées sur le terrain (levés topographiques) ou captées à distance (GPS, photos aériennes, images satellitales), ou obtenues par MNT, par digitalisation ou scannage de cartes et de plans existants. La majorité de ces données engendrent des formes géométriques, dont il faut s'assurer de l'intégrité spatiale après leur avoir associé des attributs spatio-temporels.

Un second élément distinctif des données à référence spatiale est ce que Anselin [ANS 89] appelle la dépendance spatiale, c'est-à-dire la propension des entités spatiales voisines à s'influencer, à posséder des attributs similaires ( qui ne sont pas forcément indentiques) ou des potentialités complémentaires. Ainsi la dépendance spatiale se rapporte à la taille, la forme et configuration des entités spatiales. Un autre élément caractéristique des données localisées est l'hétérogénéité spatiale. Celle-ci se produit quand il y a un manque d'uniformité spatiale des effets de la dépendance spatiale ou des relations entre les variables sous étude. L'hétérogénéité et la dépendance constituent en fait deux variantes de ce qu'on appelle impact spatial<sup>7</sup> [ANS 93].

Les données à référence spatiale sont aussi des informations exprimées en des termes linguistiques [ROL 00a ; ROL 00b] pour décrire, quantifier et/ou qualifier un phénomène. Ces données sont alors entachées d'imprécision, de subjectivité et d'incertitude ; elles relèvent ainsi des connaissances floues [ROL 84 ; ROL 96 ; ROL 00a ; ROL 02 ; RUF 01]. Ces différentes caractéristiques (appartenance, impact spatial, imprécision, subjectivité, ...) mettent en évidence la nature floue de l'espace géographique auquel elles font référence. De ce fait l'approche floue présenterait un intérêt indéniable [ROL 00b] pour ajuster au mieux ces données en cours de traitement dans un processus de décision.

#### **1.4. Potentialités et limites du SIG dans le processus d'aide à la décision**

Les SIG ont reçu des définitions variées allant de l'approche logicielle de la cartographie automatique et de l'analyse des données géographiques jusqu'à celle systémique, pour qualifier une composante du système d'information d'une institution [PAN 96, op cité] dans la mesure où il concerne des données de type géographique. Cet éventail de définitions est de nature à créer une confusion dans les concepts : aux Etats-Unis, l'acronyme GIS (Geographic Information System) est utilisé pour désigner indifféremment le système informatique avec un logiciel spécifique, le système d'information en tant que structure organisationnelle ou encore l'ensemble des disciplines géomatiques.

---

<sup>7</sup> On parle d'impact spatial du fait qu'en raison de la taille et de la configuration des unités spatiales, on puisse identifier des relations intra et inter variables, lesquelles relations sont dues aussi bien à la nature des unités spatiales qu'à la nature des variables étudiées.

Pour notre part, avant de parvenir à une définition et par souci de clarification des concepts, nous allons situer les SIG par rapport à la géomatique et aux systèmes d'information à référence spatiale (SIRS), qu'il convient de définir à leur tour.

### ***Géomatique***

La géomatique est la science et la technologie de la collecte, de l'analyse, de l'interprétation, de la distribution et de l'utilisation de l'information géographique. Elle englobe une foule de disciplines qui, dans un système à référence spatiale commune, concourent à créer une représentation à la fois détaillée et compréhensible du monde physique et de l'espace que nous y occupons. Parmi ces disciplines, on compte les SIG, le système de positionnement global (GPS), la télédétection, la cartographie numérique et le levé cadastral. La géomatique peut être définie comme étant "un champ d'activités scientifique et technique qui intègre, suivant une approche systémique, l'ensemble des moyens d'acquisition et de gestion des données à référence spatiale utilisés dans le processus de production et de gestion du territoire" [GAG 90]. Il s'agit donc d'une approche scientifique et moderne qui met à profit l'intégration de plusieurs moyens d'acquisition et de gestion de données sur le territoire. Ce néologisme, dont la paternité revient au scientifique français Bernard Debuissou<sup>8</sup> [LAA 00], a émergé au Canada et plus particulièrement au Québec; depuis, il est de plus en plus utilisé dans plusieurs pays.

### ***Système d'information à référence spatiale***

Le SIRS est une "organisation structurée de ressources humaines, d'équipements, de procédures et de données qui se coordonnent pour réaliser l'acquisition et la gestion des données et fournir l'information sur le territoire" [MAM 84]. Tel qu'il est défini ici, le SIRS constitue un moyen qui permet de fournir l'information sur le territoire sans avoir nécessairement recours à un système informatisé. Le SIRS informatisé gagnerait énormément, en terme d'efficacité, à se doter d'outils informatiques pour gérer un volume sans cesse croissant de données sur le territoire et faire face à la complexité de leur traitement. Son but est fournir l'information pertinente pour améliorer la compréhension des phénomènes et faciliter la prise de décision. Le terme SIG désigne ainsi, il système informatisé (en tant que logiciel) utilisé dans un SIRS pour effectuer de telles tâches.

---

<sup>8</sup> Bernard Debuissou a été le premier à utiliser le terme géomatique dans les années 1960.

### 1.4.1. Définition du SIG

La définition du concept de SIG est rendue difficile par les nombreuses applications voisines qui se substituent parfois au terme de SIG [BOR 02, op. cité] et les différentes significations qui lui sont attribuées. Il existe ainsi une large gamme de définitions [PAN 96, op. cité] allant de l'usage au contenu. Dans le cadre de notre travail de recherche nous nous intéressons à l'aspect logiciel du SIG plutôt qu'à sa version institutionnelle et organisationnelle. Ainsi nous définissons le SIG comme un système informatique de gestion de bases de données conçu pour saisir, stocker, manipuler, analyser, créer et visualiser des données à référence spatiale en vue de résoudre des problèmes complexes de gestion et de planification territoriale.

En définitive, le SIG devra inclure normalement, (figure 1.2), les composantes relatives à :

- l'acquisition des données d'entrée ;
- le stockage, à la récupération et à la gestion de base de données ;
- la manipulation et à l'analyse des données ;
- la visualisation et la génération d'autres données ;
- une interface utilisateur ;
- une ou plusieurs connexions à des modules d'inférence et autres SIG.

Cette approche du SIG traduit le souci majeur de sa promotion vers de véritables systèmes d'aide à la décision. D'autres auteurs l'appuient en plaidant pour que le SIG soit considéré comme une science de l'information spatiale plutôt qu'une technologie [FIS 93 ; OPE 97, op. cité ; ROL 96, op. cité ; GOO 02].

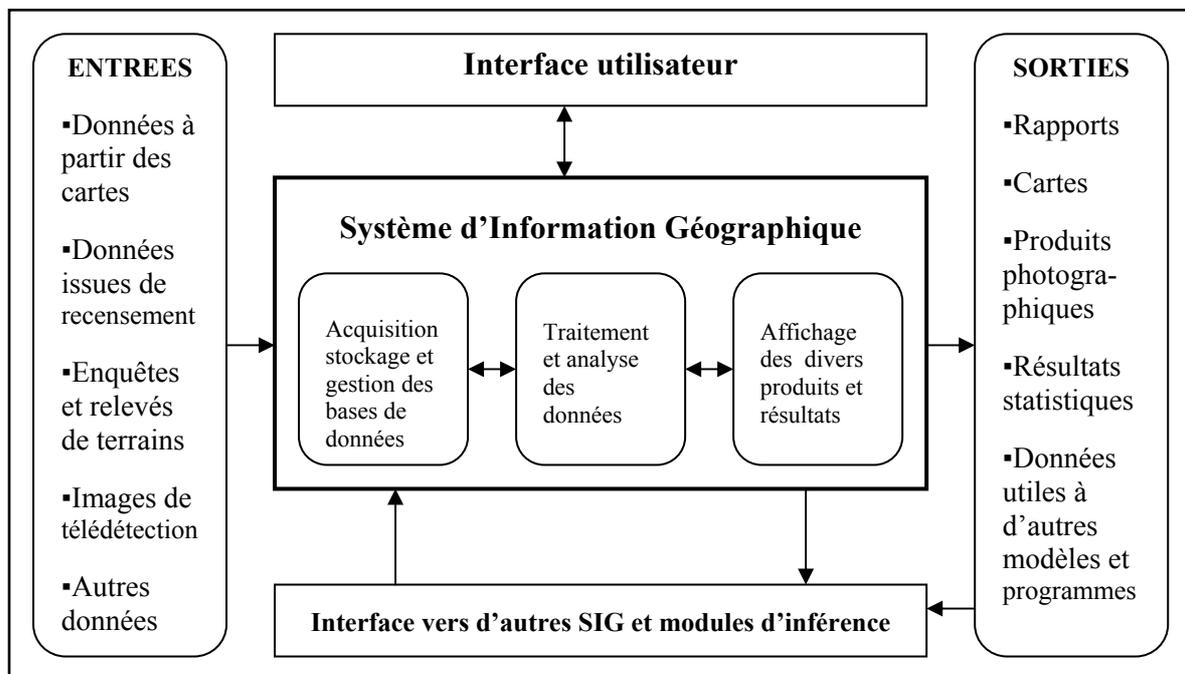


Figure 1.2 : Composantes d'un SIG [adapté de FIS 93]

### **1.4.2. Contribution du SIG**

Pour la problématique qui nous intéresse, le traitement des problèmes de localisation au moyen d'un SIG s'inscrit dans les contributions de type analyse spatiale et aide à la décision. Dans ce contexte, le recours aux outils informatiques a des avantages évidents : la manipulation d'un volume important de données aide à affronter le problème dans toute sa complexité et permet ainsi d'éviter des simplifications parfois hasardeuses. Le SIG servira donc à décrire et analyser le territoire puis participer à la simulation des scénarii plausibles.

#### ***Description du territoire***

La représentation du territoire se réalise par une base de données géographiques gérée par le SIG. Cette base de donnée fournit les informations la région concernée par un aménagement. Cette contribution est fondamentale. Par ailleurs, une représentation cohérente du territoire est nécessaire, aussi bien pour son analyse que pour l'exploitation des informations au travers de modèles de simulation. La représentation du territoire par une base de données géographiques est donc la condition de toute contribution du SIG en aménagement du territoire. Cette base de données doit sur l'identification des phénomènes et leurs interactions tout en se référant au *contenu* et au *contenant* du territoire dans sa représentation systémique telle que schématisée dans la section précédente.

#### ***Analyse du territoire***

L'usage des SIG comme outil d'analyse du territoire a pour objet de mettre en évidence les faits spatialisés. Ce type d'usage vise la validation d'une hypothèse, la confirmation d'une présomption ou encore la compréhension d'un phénomène localisé [BOR 02]. Cette analyse du territoire implique l'exploitation des informations géographiques dans le cadre de procédures ou de protocoles établis en vue d'une meilleure gestion des objets géographiques décrits. Les traitements sur les données ainsi que les données elles-mêmes sont alors en grande partie prédéfinis. Le potentiel du SIG pour l'analyse du territoire présente ainsi un grand intérêt dans les procédures d'aide à la décision concernant la localisation spatiale d'un ouvrage. Les fonctionnalités utiles sont les basées sur les opérateurs géométriques pour le calcul de distances, les opérateurs de voisinages ainsi que d'autres fonctions qui permettent de simuler l'interaction spatiale.

### ***Simulation des phénomènes spatiaux***

Le SIG permet de faire des simulations en mettant en exergue les conséquences ou les impacts des différents scénarii. Cette contribution des SIG est ainsi de façon privilégiée celle de l'analyse spatiale couplée à des modèles explicatifs au sens des géomaticiens et des géographes. L'utilisation de ces modèles de simulation nécessite généralement une grande quantité d'informations. [JEA 95] a mis en évidence l'intérêt de l'utilisation du SIG pour ce type d'application. Mais elle a aussi montré que la gestion, à l'aide d'un SIG, des données "input" et "output" d'un modèle de simulation pose de nombreuses difficultés. En effet pour des questions d'efficacité des calculs, la représentation du territoire associée au modèle de simulation est très spécifique et de ce fait, nécessite la combinaison du SIG avec un autre module.

#### **1.4.3. Limites du SIG**

Tous les auteurs semblent s'accorder sur la puissance et l'excellence des SIG en tant qu'outil de gestion des données à référence spatiale. Cependant, contrairement à sa puissante capacité à gérer les données, les fonctions d'aide à la décision du SIG sont relativement élémentaires [LAU 93, op. cité ; ROL 93 ; OPE 97, op. cité ; LAA 00, op. cité ; BOR 02 ; MAI 02] et limitées. Selon ces auteurs les méthodes d'analyse incluses dans les SIG standards se limitent aux opérations booléennes sur les attributs des entités géographiques et n'intègrent pas de véritables processus d'inférence spatiale. Ils notent particulièrement le manque du traitement de la notion du temps et la multidimensionnalité (3D, 4D) et le peu de performance dans le traitement des aspects multi-échelles et multisources. Au regard des critiques adressées aux SIG, nous pouvons distinguer deux grandes catégories de limites.

#### ***Limites au plan logique et technique***

Un problème pratique surgit quand on effectue des analyses de données spatiales avec les SIG : c'est l'absence d'une démarche systématique qui permettrait d'utiliser des séquences d'opérations pour obtenir les réponses escomptées. Pour répondre à une requête spatiale simple, chaque utilisateur se fie à son intuition et emprunte un chemin personnel, pour aboutir aux mêmes résultats [BUR 90]. Cette démarche intuitive qui fonctionne pour des opérations simples, serait peu opérationnelle pour des opérations plus complexes. Une telle situation témoigne, selon plusieurs chercheurs dont Goodchild [GOO 91 ; GOO 92], du manque de clarté logique qui entoure l'analyse spatiale, perçue comme un ensemble de techniques développées dans des domaines différents sans aucune codification claire ni un cadre conceptuel reposant sur une assise théorique robuste. Un espoir était né avec les SIG en ce

sens mais la faiblesse sous-jacente aux modèles conceptuels de données constituant encore un handicap pour la création de ce système de codification. On pourrait utiliser d'autres représentations spatiales que celle de l'approche "point-ligne-polygone" propre aux objets utilisés par les SIG vectoriels : nous pensons aux "*fonctions de probabilité spatiale*" et les méthodes des ensembles flous pour les entités qui ne se prêtent pas à une description par les attributs [ROL 84 ; ROL 200a, op. cité ; ROL 00b, op. cité] ou par une représentation spatiale dans les espaces métriques traditionnels [BUR 90, op. cité ; LOW 92].

Plusieurs études, qui se sont livrées à l'analyse des SIG existants, sont concordantes : on se trouve dans une situation caractérisée par un état riche des données et en progression croissante par rapport à un état pauvre de la théorie. En effet, il y a donc place pour une approche plus créative et pour les outils analytiques pouvant suggérer de nouvelles propositions théoriques et supporter des fonctions d'exploration des données.

### ***Limites au plan décisionnel***

Clarke [CLA 90] affirme que les praticiens des SIG ont besoin d'une meilleure compréhension des besoins des décideurs en rapport avec leurs environnements. Il défend l'idée qu'un besoin crucial des décideurs consiste à disposer de capacités analytiques de prédiction, de manière qu'ils puissent commencer à explorer les effets possibles de leurs décisions. Ce besoin est d'autant plus justifié en regard des applications très connues dans les domaines de l'aménagement du territoire qui nécessitent des statistiques spatiales ou des langages de modélisation. On constate que certains des SIG les plus sophistiqués incluent la fonction d'optimisation pour une classe de problèmes bien définie et échouent, en revanche, à représenter des modèles ou des capacités de simulation démontrés par des modèles mathématiques relatifs aux processus physiques et socio-économiques. Ce manque de capacités de modélisation et de simulation sophistiquées empêche les SIG actuels de pouvoir remplir leur potentiel de systèmes d'aide à la décision. La raison principale de cet "échec" est l'absence de *processus d'inférence spatiale* basé sur les outils mathématiques et informatique en la matière.

Dans un tel contexte, et afin d'aider le décideur dans sa tâche, il convient de trouver la meilleure combinaison possible des moyens (approches, méthodes et techniques) pour lui fournir des solutions pertinentes et appropriées à ses problèmes, de nature à lui permettre de prendre des décisions éclairées.

C'est ainsi que de plus en plus d'auteurs [LAU 93 ; ROL 93 ; ROL 00b] plaident pour que ***le SIG mute en un réel outil d'aide à la décision par couplage à des modules d'inférence.***

Une solution appropriée serait donc la combinaison du SIG avec un système d'inférence fondé sur les acquis et ressources de l'intelligence artificielle [OPE 97 ; ROL 02]. L'intérêt que nous portons à la logique floue et aux réseaux de neurones dans l'accomplissement de cette tâche se justifie par leur apport substantiel aux activités d'aide à la décision à référence spatiale.

Avant d'aborder cette préoccupation scientifique, il convient de définir ce que nous entendons par localisation spatiale. Ce sera l'objet du chapitre suivant.

## CHAPITRE II

# LA LOCALISATION SPATIALE : CONCEPTS, TYPOLOGIES ET DEMARCHE D'ANALYSE DES BESOINS

Dans un effort de recherche d'une signification cohérente et logique de l'expression "*localisation spatiale*", en accord avec les exigences des sciences géographiques, nous explicitons dans un premier temps les concepts et définitions sous-jacents. L'analyse typologique du concept de localisation spatiale nous permet en second lieu chapitre d'exposer la démarche pour l'identification et la satisfaction des besoins en localisation spatiale en termes d'aménagement du territoire.

### 2.1. Concepts et définitions

La localisation est une propriété spatiale essentielle en géographie. Si cette propriété désigne d'abord la position absolue exprimée en coordonnées géographiques (longitude et latitude) auxquelles on ajoute parfois l'altitude, c'est-à-dire la position par rapport au niveau de la mer (niveau 0 par convention) d'un lieu ou d'un objet, elle exprime surtout sa position relative c'est-à-dire la situation géographique vis-à-vis d'autres phénomènes. En ce domaine l'échelle d'analyse est très importante car elle détermine les facteurs de localisation c'est-à-dire tout paramètre susceptible d'influencer d'une manière ou d'une autre sur le choix d'un emplacement d'un objet. Ainsi il est possible de caractériser la localisation d'une activité ou d'un équipement de multiples façons en fonction des ordres de grandeur des espaces en considération : le monde, un continent, un pays, une région, un département, une commune, un quartier etc.

Une notion tout aussi fondamentale participant à la localisation est la *distance*, puisque qu'elle est directement utilisée pour caractériser la situation géographique. Cette notion est, par ailleurs, à la base de nombre de modèles mathématiques d'analyse spatiale (section 3.2.2). Elle permet en particulier de mesurer l'écart entre deux objets géographiques ou entre un objet et un repère. De ce fait, la distance engendre les notions de proximité et d'éloignement.

La proximité est la propriété d'objets qui sont voisins ; la notion est toutefois relative car la perception du voisinage peut varier d'un individu à un autre ; on lui préfère alors la notion de contiguïté qui traduit un réel voisinage puisqu'il s'agit d'objets qui se touchent d'une manière ou d'un autre. L'éloignement est bien entendu la propriété inverse qui reste aussi une notion relative.

La coexistence de ces différentes notions en terme de localisation spatiale renvoie à la distribution, qui traduit la localisation des objets dans l'espace, et à la répartition, qui indique la façon dont ces objets se partagent l'espace [BAV 98]. Distribution et répartition s'évaluent via un dénombrement d'objets et leur agencement spatial (nombre d'écoles primaires dans une commune, localisation de ces écoles les unes par rapport aux autres) ; elles peuvent aussi être appréhendées via des densités par la mise en rapport du nombre de ces objets et d'une surface donnée.

Dans le cadre de cette recherche, l'échelle d'analyse considérée est le territoire tel que défini plus haut. Les questions de localisation spatiale au sens de l'aménagement du territoire se posent lors du choix du ou des meilleur(s) emplacements pour l'implantation d'une infrastructure, un équipement ou inversement lors qu'il s'agit d'identifier le type d'activité à installer sur un site dont on connaît les caractéristiques : superficie, nature géologique, démographie, position géographique etc.

Qu'ils soient l'œuvre d'un organisme privé, semi-public ou d'une collectivité territoriale, ces équipements ont pour vocation d'améliorer le cadre de vie des populations du territoire considéré en leur offrant certains services ; ce sont donc des équipements collectifs. En conséquence les choix de localisation doivent intégrer plusieurs contraintes liées à l'accessibilité, la proximité, l'attractivité, la couverture spatiale etc. Au cœur de cette réflexion sur la localisation des équipement collectifs, on est alors confronté à deux types de problèmes. D'une part on souhaite trouver un équilibre entre les coûts d'accès des usagers aux services offerts par les équipements. D'autre part, il faudra parvenir à un consensus entre la recherche d'une plus grande efficacité dans l'affectation spatiale des équipements et une meilleure équité entre les usagers dans l'accès aux services. C'est donc une question d'arbitrage [PEE 01 ; SAN 01] entre efficacité du système et équité géographique et spatiale du point de vue des usagers mais aussi du point de vue de l'ensemble du territoire considéré.

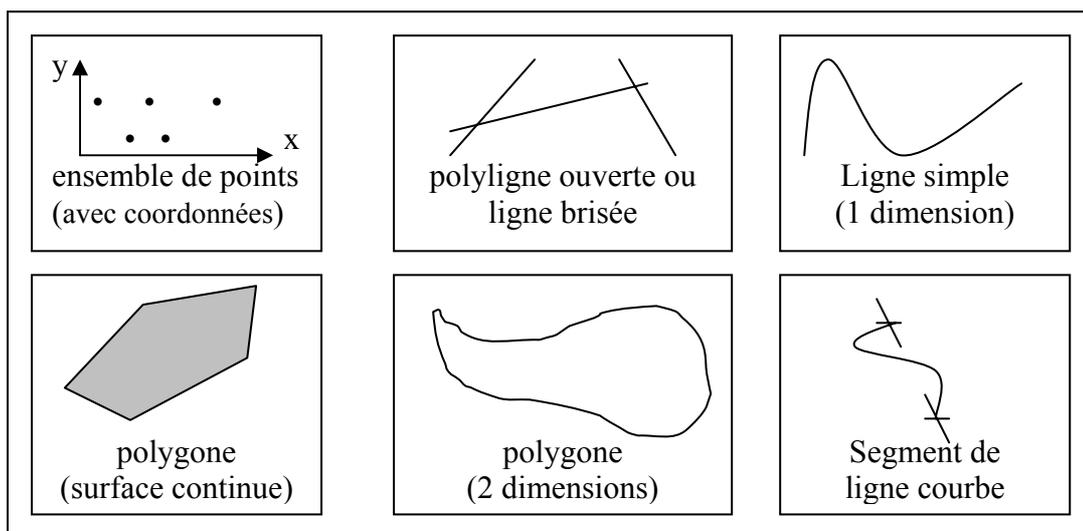
Ceci revient à optimiser les choix de localisation par la mise en œuvre de modèles de "*localisation-affectation*" efficaces qui tiennent compte des contraintes et des facteurs de localisation dont mention a été faite plus haut.

## 2.2. Typologies

Deux aspects fondamentaux coexistent dans l'étude des travaux à référence spatiale en l'occurrence la localisation spatiale : l'aspect spatial et l'aspect thématique. Dans cette section nous présentons la typologie spatiale des objets géographiques classiques selon le formalisme CONGOO (CONception Géographique Orientée Objet) [PAN 96, *op. cité*] s'appliquant à la modélisation conceptuelle des données géographiques c'est-à-dire les objets géoréférencés (points, lignes, polygones, surfaces élémentaires ou continues ou des ensembles de ces objets) et une typologie thématique (nature des services potentiels offerts par les équipements à localiser) associée.

### 2.2.1. Typologie spatiale

La représentation des objets spatiaux en particulier géographiques s'appuient sur un certain nombre de concepts issus de la géométrie euclidienne, de la théorie des graphes et de la topologie (figure 2.1).



**Figure 2.1** : Typologie des objets géographiques primitifs

En géométrie euclidienne, on distingue les objets à zéro dimension (*points*), à une dimension (*lignes*), à deux dimensions (*aires, surfaces ou polygones*) et enfin à trois dimensions (*volumes*). Dans la nature, tous les objets sont au moins tridimensionnels ; cependant, pour des raisons de représentation et de modélisation, on se contente de dimensions inférieures. Dans de nombreuses applications liées à l'occupation du sol, on ne s'intéresse qu'aux aspects ponctuel, ligne et surfacique des choses. Ainsi les concepts clés sont les **points**, les **segments** de ligne droite, les **lignes** (polylignes ou mixtilignes), les **polygones**.

De nombreux concepts géomatiques ont une modélisation empruntée à la théorie des graphes. Tel est le cas des réseaux de communication, les systèmes de distribution où les concepts dont le **nœud**, l'**arc** ou l'**arête**, ainsi que la **chaîne** d'arcs.

La troisième composante importante est la topologie qui permet de repérer les positions relatives des objets entre eux et d'exprimer des relations spatiales. Alliée à la théorie des graphes, elle offre des possibilités de modélisation efficace de la dynamique des objets spatiaux.

La combinaison de ces divers objets géographiques simples appartenant à une même classe ou à des classes différentes permet d'obtenir des objets géographiques composés et/ou complexes [PAN 96, *op. citée*] qui peuvent avoir ou non leur propre identité, selon le phénomène étudié, et être identifiés et traités de façon individuelle ou collective.

### 2.2.2. Typologie thématique

Les équipements ou infrastructures étant destinés à offrir un type de service aux bénéficiaires, nous allons les caractériser par la nature de ces services et la "forme" du déplacement entre le service et chaque usager pris séparément. Opérationnellement cette typologie s'avère indispensable pour comprendre la nature du problème de localisation afin d'en proposer des modèles adéquats.

Reprenant et élargissant les travaux de [PEE 01, *op. citée*; PEE 02; RUL 02], nous distinguons six principales catégories de services :

♣ **services ordinaires ou d'urgence** : les services peuvent être qualifiés d'ordinaires ou d'urgence, selon que le traitement de la demande d'un usager peut être différé ou non. Dans le cas où le service est d'urgence (ambulances, pompiers, police, etc.), la durée (en terme de temps) écoulée entre l'expression et la prise en charge du besoin peut être de performance du service ;

♣ **services fixes ou mobiles** : le service est fixe quand c'est l'utilisateur qui se déplace vers lui (universités, centres hospitaliers, grandes surfaces, etc.) et mobile lorsqu'il se déplace vers l'utilisateur (distributions de courriers et d'articles divers, soins médicaux à domicile, etc.). Cette distinction est nécessaire pour la modélisation du déplacement ;

♣ **services à demande élastique ou inélastique** : un service est élastique ou inélastique selon que la fréquentation de l'utilisateur est fonction du coût de déplacement et de la tarification de ce service (loisirs et services socioculturels) ou non (écoles primaires, services d'urgence). La modélisation de la demande de service des utilisateurs potentiels dépend de ces éléments ;

♣ **services à congestion** : un service est soumis à congestion lorsque son utilisation par un usager rend impossible son emploi par un autre ou est, en tout cas, susceptible d'altérer la qualité du même service offert aux autres usagers simultanément. Les infrastructures routières en sont un exemple. Dans ce cas on définira un seuil maximal, en ce qui concerne l'effectif des usagers utilisant de façon simultanée, au-delà duquel le service peut être inaccessible, en rapport avec les capacités de l'infrastructure ;

♣ **services obligatoires ou libres** : le caractère obligatoire de certains services comporte deux aspects : d'une part, une certaine catégorie de la population peut être "obligée" d'y avoir recours (écoles primaires), d'autre part, le service peut être contraint de répondre à toute demande lui parvenant (garde médicale du week-end). Pour certains services fixes, l'utilisateur choisit librement son lieu de consommation (services socioculturels), tandis que pour d'autre, il lui est "imposé" l'endroit où il doit se rendre (certains services administratifs).

♣ **services attractifs ou répulsifs** : toute infrastructure a un impact plus ou moins positif ou négatif sur l'environnement physique et humain. Aussi peut-on distinguer les services attractifs ou répulsifs selon la nature des externalités émises (en se référant aux normes existantes en la matière). Ainsi un hôpital sera-t-elle considéré comme "relativement attractif", tandis qu'un dépôt d'immondices ou un incinérateur sera qualifié de répulsif. Les choses peuvent se compliquer, en ce sens que les personnes qui subissent les inconvénients d'une situation ne sont pas nécessairement celles qui en bénéficient pleinement.

Bien entendu, cette typologie peut ne pas avoir un caractère universel en raison du contexte institutionnel dans lequel on se trouve. Cependant elle résume à peu près toutes les situations. Signalons tout de même qu'un service peut appartenir à plusieurs catégories à la fois. Par exemple la poste est à la fois un service fixe (guichet) et mobile (distribution de courriers par les facteurs), attractif (petit bureau de poste dans les zones isolées), répulsif (centre de tri), etc. Du fait de l'évolution des nouvelles technologies de l'information et de communication, il serait possible d'intégrer un facteur d'accélération quant à la modélisation de la demande de certains services compte de la culture et du niveau technologique du territoire considéré.

## **2.3. Démarche d'analyse**

### **2.3.1. Analyse de la demande et de l'offre spatiale**

#### *Analyse de la demande*

Après avoir subdivisé le territoire considéré en mailles (petites unités géographiques), on commence par localiser les besoins en fonction du type de service à partir d'une cartographie des usagers par observation ponctuelle (lorsque les usagers potentiels sont estimés peu nombreux) ou par groupes géographiques d'individus (îlots, quartiers, communes etc.). On évalue ensuite la quantité de demande de chacune des mailles. Ces informations qui sont généralement obtenues auprès des services statistiques et démographiques compétents seront éventuellement complétées par des enquêtes spécifiques auprès des usagers, dans une perspective d'ajustement. Ensuite on construit une fonction de demande pour chaque type de service en tenant compte des principales caractéristiques du service et des usagers (catégories socioprofessionnelles, tranches d'âges, etc.). Cette analyse économétrique doit être poussée à un niveau de finesse, en incorporant des caractéristiques du comportement de l'utilisateur [BAU 02 ; PEE 02, *op. citée*], comme les déplacements à objectifs multiples ou la possibilité pour un usager de fréquenter plusieurs unités de services.

La temporalité de la demande est également un élément essentiel dans la modélisation. En effet la prise en compte de la manière dont se répartissent les appels et /ou les visites dans l'horizon de la planification du service permettra de mieux traiter les aspects liés à la congestion. De plus, il sera possible de prévoir l'évolution probable de la demande pour chaque type de service.

#### *Analyse de l'offre*

Analyser l'offre tout d'abord à disposer de l'inventaire des équipements déjà existants (nombre, emplacement, taille ou capacité, nature des services offerts). Ensuite, il convient d'identifier les sites candidats c'est-à-dire les sites potentiels pouvant "théoriquement" recevoir d'éventuelles nouvelles installations d'équipements. Enfin, il convient d'établir le répertoire des normes qui doivent être respectées et que nous appellerons contraintes (terme qui est défini dans la section 3.1 du chapitre suivant), parce qu'elles seront imposées à toute proposition de redéploiement. Ces contraintes sont de plusieurs types. Les unes touchent au nombre d'unités à installer : généralement la collectivité territoriale dispose d'un budget à répartir entre un nombre d'unités à définir en fonction de la dispersion géographique des usagers (plus cette dispersion sera grande, plus on souhaitera un grand nombre de petites unités) et des économies d'échelles de la concentration (si elles sont élevées, on préférera un

petit nombre de grandes unités). D'autres contraintes concernent la capacité des équipements, les externalités (positives et négatives) émises par les équipements, la couverture territoriale qu'ils assurent. Nous reviendrons sur toutes ces différentes contraintes lors de la phase de modélisation mathématique des contraintes de décision (section 3.2.3).

### **2.3.2. Ajustement de l'offre à la demande**

Une fois l'offre et la demande définies, il convient de procéder à leur ajustement. Le recours aux modèles de localisation-affectation s'impose [DAS 95 ; HANS 96 ; LAB 95] ; et ces modèles seront définis par des fonctions d'objectif soumis à un ensemble de contraintes.

Les fonctions d'objectifs exprimeront de façon formelle les critères de configuration de localisation de chaque type de service. Ces critères devront résumer le bilan global de l'opération de redéploiement [PEE 01, *op. cité*]. En effet, pour chaque usager, l'avantage relatif d'une localisation réside, principalement, dans le coût (temps) de déplacement correspondant. Ces critères ne tiendront pas que des usagers, on intégrera aussi des fonctions de minimisation des coûts publics sous contraintes d'accessibilité au service. Le but de ce modèle de localisation-affectation est de fournir des propositions de redéploiement (les solutions) qui optimisent les critères retenus tout en respectant les contraintes explicitement formulées. Ces propositions indiqueront simultanément les localisations, le nombre et la taille des équipements. Bien évidemment, il peut s'agir d'équipements à supprimer, à ajouter ou à relocaliser. Les solutions préciseront également les affectations des usagers aux différents équipements de sorte qu'il soit possible d'établir la carte optimale des aires de service. Cette carte est particulièrement utile pour tous les services mobiles, puisque ceux-ci ont la possibilité de définir leurs zones d'interventions ; il n'en va de même pour les services fixes qui "imposent" leurs lieux de consommation.

En définitive, il ressort de ces considérations qu'il est essentiel d'asseoir notre recherche/développement sur deux fondements : la base conceptuelle, et les fondements typologiques (géométrique et thématique) des problématiques de localisation spatiale. Au plan décisionnel, la résolution de ces problèmes et le choix des "meilleures" solutions imposent une approche intégrative des SIG et des méthodes d'inférence. Ce sera l'objet du prochain chapitre.

## CHAPITRE III

# L'AIDE A LA DECISION EN LOCALISATION SPATIALE : SYSTEME EXPERT ET APPROCHE NEURONALE FLOUS

Le champ de l'aide à la décision a fait l'objet de recherches menées par plusieurs scientifiques de divers domaines ; il a donc constitué un pôle attractif pour différentes études et applications. Parmi ces applications, nous retrouvons d'ailleurs plusieurs cas relatifs à des problèmes à référence spatiale [KEE 76 ; ANC ; KEE 80 ; SIM 89 ; ROY 92 ; OPE 97, op. cité ; ROL 96, op. cité ; BIR 00, op. cité ; OPE 01] : planification urbaine et régionale, transport, gestion des ressources en eau, gestion environnementale, évaluation du territoire, localisation des activités industrielles etc. Afin de mieux cerner la portée de l'application des méthodes d'aide à la décision aux problématiques territoriales et la nécessité d'améliorer les méthodes classiques (comme la programmation linéaire), il convient de faire un bref rappel de l'évolution de ce champ d'études.

En effet les méthodes d'optimisation de la programmation linéaire ont en partie échoué à appréhender toutes les facettes, en particulier les dimensions cognitive, spatiale et socio-économique que présentent les problèmes sur le territoire. La raison principale est que ces méthodes s'appuient le plus souvent sur le jugement d'un critère unique, formalisé par une fonction-objectif linéaire et évaluée généralement sous la forme de coût, alors que la plupart de problèmes observés dans la réalité sont plus "difficiles", donc irréductibles à un seul critère ; l'approche de l'analyse multicritère constitue une alternative dans ce cas. Cependant, en dépit de son caractère multi objectif, cette approche se révèle très tôt insuffisante pour la résolution des problèmes sous plusieurs contraintes, car il faudra également se préoccuper de la nature floue et sémantique des connaissances, ainsi que les dimensions cognitive et spatio-temporelle.

C'est pourquoi, devant ce constat (insuffisance de la théorie statistique de la décision<sup>9</sup>), notre approche se place dans la perspective de l'aide à la décision gérée par un système d'inférence complexe. En premier lieu il sera fondé sur une base de connaissances extraites de divers experts traitée de façon automatique par une procédure neuronale et calibrée par *apprentissage* (en mode supervisé et/ou non supervisé). En second lieu cette base de connaissance alimentera un système expert qui sera apte à apporter l'aide à la décision en localisation dont nous avons besoin.

En conséquence ce chapitre représente le socle même de notre travail de recherche et le plan suivi est le suivant :

- concepts et définitions de l'aide à la décision,
- la chaîne de cohérence décisionnelle floue : application à la problématique de localisation spatiale,
- méthodologie d'acquisition des connaissances,
- le système expert.

### **3.1. Concepts et définitions de l'aide à la décision**

#### **3.1.1. Aide à la décision**

Bernard Roy [ROY 85] définit l'aide à la décision comme étant : *“L'activité de celui qui, prenant appui sur des modèles clairement explicités mais non nécessairement complètement formalisés, aide à obtenir des éléments de réponse aux questions que se pose un intervenant dans un processus de décision, éléments concourant à éclairer la décision et normalement à prescrire, ou simplement à favoriser un comportement de nature à accroître la cohérence entre l'évolution du processus d'une part, les objectifs et le système de valeurs au service desquels ce intervenant se trouve placé d'autre part”*.

Cette définition revêt deux aspects fondamentaux : d'une part, la modélisation totalement ou partiellement formalisée et, d'autre part l'analyse des solutions et l'évaluation de leurs conséquences. En tant que telle, l'expression “aide à la décision” ne fait l'unanimité au sein des scientifiques : l'école américaine utilise souvent “prise de décision” (*décision making*). La nuance est cependant importante car il ne s'agit pas d'une simple différence linguistique mais bel et bien deux paradigmes : d'un côté on adopte une posture plus modeste vis-à-vis du décideur en l'éclairant dans son choix de solution et donc en lui prêtant assistance dans sa décision ; de l'autre côté, on semble sûr de ses moyens et outils pour montrer au décideur comment prendre les bonnes décisions.

---

<sup>9</sup> Cette théorie était considérée comme étant *“la tentative la plus achevée de formalisation du comportement d'un individu confronté à un choix en avenir incertain”* [MAR 86], mais qui a connu relativement peu de succès.

### 3.1.2. Décision

La définition couramment admise du concept de la décision est la suivante : “ *un choix entre actions, solutions, la ou alternative(s)*”. Cette définition synthétique, quoiqu’en apparence claire, est peu satisfaisante car ce choix couronne l’aboutissement d’un processus dynamique qui intègre plusieurs facteurs et paramètres qui sont variables selon le problème considéré. Le dictionnaire Robert la définit comme “*la fin de la délibération dans l’acte volontaire de faire ou de ne pas faire*”.

### 3.1.3. Décideur

Outre l’aspect du résultat du processus que comporte la définition du concept de décision, il y a un deuxième aspect relatif à la délibération. En effet l’acte de délibération présuppose un débat au sein d’un groupe d’acteurs qui n’ont pas nécessairement des valeurs et intérêts convergents et qui n’exercent pas la même influence [ZOL 92]. La notion de décideur (decision-maker en anglais) “*désigne en dernier ressort l’entité qui apprécie le ‘possible’ et les finalités, exprime les préférences et est sensé de les faire prévaloir dans l’évolution du processus*” [ROY 85, *op. cité*].

Dans les problématiques relativement complexes comme la nôtre, cette entité (le décideur) est généralement guidée par *l’équipe d’études scientifiques (experts, chercheurs, ingénieurs de décision) dont le rôle consiste entre autres à concevoir, expliciter et exploiter un modèle en vue d’obtenir de élément de réponse, à éclairer le décideur sur les conséquences de tel ou tel comportement en les lui rendant intelligibles, en proposant des scénarii de solutions ou d’actions*. Dans un tel contexte, où la décision est souvent influencée par des “stratégies et intérêts politiques”, elle pourra être prise, au vu des scénarii, grâce à un travail complémentaire éventuellement basé sur des analyses multicritères multi-agents [LAA 00] où interviennent les notions de décision de groupe, participation, aide à la négociation, jeu de rôle, point de vue etc.

### 3.1.4. Action, contrainte et critère

Une **action** est une représentation d’un élément de solution contribuant à la décision. Lorsqu’elle est provisoirement jugée réaliste, on parlera d’action potentielle.

Une **contrainte** est un facteur permettant de circonscrire et de délimiter une action, un objet ou un phénomène territorial.

Un *critère* est facteur de jugement sur la base duquel on mesure et on évalue une action dans le domaine considéré.

### 3.1.5. Le processus décisionnel

Mintzberg [MIN 76] propose la définition suivante : “*Le processus décisionnel est considéré comme un ensemble d’activités déclenché par un stimulus, et aboutissant à un engagement spécifique à l’action*”. Cet ensemble d’activités se déroule en trois grandes phases (figure 3.1).

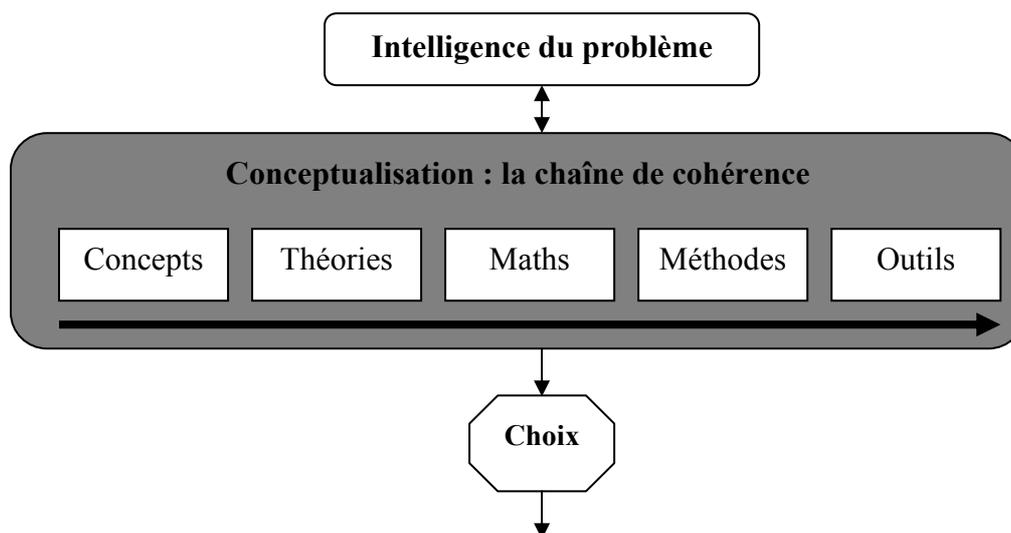


Figure 3.1 : La chaîne de cohérence intégrée au processus décisionnel

*L’intelligence<sup>10</sup> du problème* est la première étape du processus. Elle commence par la reconnaissance de la problématique générale qui suscite l’activité décisionnelle et s’achève par un diagnostic qui permet de comprendre le stimulus suscité afin de déterminer les relations de cause à effet se rapportant à la situation identifiée. Dans le cadre de notre recherche, l’exposé des chapitres précédents vaut cette phase.

*La conceptualisation* est le cœur du processus décisionnel. Elle se concrétise par la mise en œuvre de la *chaîne de cohérence* [ROL 03] (que nous développerons dans la section suivante). A cette phase, se réalise d’abord une étude détaillée des problèmes identifiés ; ce qui permet d’approcher la problématique dans toutes ses dimensions : les bases conceptuelles, théoriques et mathématiques sous-jacentes à la problématique générale.

<sup>10</sup> Utilisé dans le sens anglais du terme, signifie renseignement et collecte d’informations sur l’ensemble de la problématique posée.

**Le choix** constitue la dernière étape de ce processus. En d'autres termes, la sélection et le type de l'implantation informatique des différentes approches nécessaires à la résolution de la problématique, constituent également une démarche itérative impliquant progressivement une investigation profonde des variantes. Cette étape renseigne sur la fiabilité et l'applicabilité, au regard des différents tests et méthodes d'apprentissage, du modèle ou du système implémenté et permet ainsi son calibrage.

### 3.2. La chaîne de cohérence décisionnelle floue: Application à la problématique de localisation spatiale

On entend par chaîne de cohérence, “une démarche scientifique dont chaque étape est essentielle et qui doit être parcourue dans l'ordre sans **raccourcis**, ni **trahisons épistémologiques**” [ROL 03]. Cette chaîne se décline en quatre phases fondamentales : les **concepts**, les **théories**, les **mathématiques** et les **outils méthodologiques**. Du point de vue de la recherche scientifique en géographie, cette exigence s'impose non seulement au plan décisionnel, afin de permettre au chercheur de bien circonscrire les limites de l'objet de recherche (en l'occurrence, l'aide à la décision en localisation spatiale), mais aussi au plan cognitif et méthodologique dans le souci de mieux cerner et traiter les données qualitatives de façon efficace. Avant d'aborder les quatre différents maillons de la chaîne de cohérence, nous insistons dans un premier paragraphe sur le fait qu'elle s'inscrit dans un contexte général de logique floue.

#### 3.2.1. Contexte général de la chaîne de cohérence décisionnelle

La *logique classique* (logique propositionnelle, logique des prédicats, algèbre de Boole), caractérisée par ses deux valeurs de vérités (vrai et faux) et ses règles d'inférence telles que le *modus ponens*<sup>11</sup> et le *modus tollens*<sup>12</sup> [TON 95], est inappropriée à une description fine et souple des connaissances. En particulier elle est inadéquate à la représentation et au traitement des connaissances vagues, imprécises, ou incertaines et ne peut de ce fait être utilisée comme modèle de base du raisonnement humain.

---

<sup>11</sup> **Modus ponens** (en posant, je pose) veut dire que “L'affirmation de l'antécédent de la majeure dans la mineure permet l'affirmation du conséquent de la majeure dans la conclusion”.

Exemple

*Majeure* : Si il pleut, alors je ne sors pas.    *Mineure* : Et il pleut.    *Conclusion* : Donc, je ne sors pas.

<sup>12</sup> **Modus tollens** (en niant, je nie) : veut dire “La négation du conséquent de la majeure dans la mineure permet la négation de l'antécédent de la majeure dans la conclusion”.

Exemple

*Majeure* : Si il pleut, alors je ne sors pas.    *Mineure* : Et je sors.    *Conclusion* : Donc il ne pleut pas.

Certains systèmes logiques sont considérés comme extensions ou alternatives de la logique classique. Il faut différencier parmi ces systèmes, ceux qui malgré l'introduction d'éléments nouveaux (par exemple la notion de temps dans les valeurs de vérité en logique temporelle, ou le concept de modalité en logique modale) ne changent pas fondamentalement la logique classique, de ceux qui en constituent une véritable alternative comme la logique multivalente (ou multivaluée) et la logique floue. En effet, ces derniers systèmes, qui en particulier ne respectent plus le principe du tiers exclu, sont caractérisés par un nombre de valeurs de vérité supérieur à deux et une généralisation de l'implication et des règles classiques d'inférence.

### *La logique multivalente*

La logique multivalente ou logique *n*-aire est une généralisation de la logique bivalente ou binaire dans ce sens que ses valeurs de vérité dont le nombre *n* est supérieur à deux, prend des éléments de l'intervalle unité  $[0,1]$ . Ainsi on parle de logique trivalente ou ternaire (3-aire) lorsque  $n = 3$ , ou logique endécadaire pour  $n = 11$  (11 états ou valeurs de vérité). Contrairement à la logique binaire, il existe différentes logiques multivalentes proposées par divers auteurs. Leurs différences viennent essentiellement de la définition de l'implication dans chacun des systèmes. Ainsi il existe plusieurs logiques ternaires introduites par Kleene [KLE 71] et d'autres auteurs [TON 95, *op. cité*]. En toute généralité, la logique *n*-aire associe à chaque variable une valeur de vérité prise dans  $[0,1]$ .

### *La logique floue*

La logique multivalente généralise la logique binaire classique en augmentant le nombre des valeurs de vérité dont chacune est un nombre précis appartenant à l'intervalle  $[0,1]$ , ou à tout autre ensemble totalement ordonné. Le nombre de ces valeurs de vérité est fini en général. Cette généralisation s'avère obsolète lorsqu'on aborde certains domaines d'application et certaines classes de problèmes où les connaissances manipulées sont vagues, imprécises et mêmes incertaines [BOU 95, *op. cité* ; ROL 96, *op. cité* ; BOU 99 ; ROL 00a, *op. cité* ; ROL 00b, *op. cité*]. La logique floue également appelée la logique linguistique, au contraire, est une généralisation plus radicale de la logique classique. Le nombre de ses valeurs de vérité est bien entendu supérieur à deux. Mais une valeur de vérité dans ce cas n'est plus une valeur de l'intervalle  $[0,1]$ , mais un mot ou une expression du langage naturel : vrai, faux, oui, non, environ, un peu, beaucoup, pas du tout, passionnément, etc. On voit alors que les valeurs de vérité de la logique floue sont des notions vagues dont chacune est modélisée par un ensemble flou dans l'univers  $[0,1]$ .

En conséquence, la logique floue liée à la théorie des ensembles flous (support de base des notions vagues) est-elle la plus appropriée pour représenter et traiter des connaissances jugées

vagues, imprécises ou même incertaines en l'occurrence les connaissances territoriales [ROL 84, *op. cité* ; ROL 96, *op. cité* ; BOU 99 ; ROL 00a, *op. cité* ; ROL 00b, *op. cité* ; RUF 01].

### 3.2.2. La chaîne de cohérence : phase conceptuelle

Le terme *flou*, présente deux significations fondamentalement différentes d'imprécision et d'incertitude, proches et parfois confondues, mais qu'il convient de bien distinguer tant au niveau sémantique qu'opérationnel [ROL 96, *op. cité*] avant d'exposer les éléments principaux de la logique floue.

#### *Définition de l'imprécision*

L'imprécision d'un objet se définit (*Grand Larousse Encyclopédique*) par le manque de netteté, le vague de cet objet, son manque de précision. Cette quasi circulaire nous renvoie au terme "précision" que la même source définit comme la délimitation stricte, nette d'un objet [ROL 00b, *op. cité*].

Partant de cette proposition et l'élargissant, C. Rolland-May [ROL 00b, *op. cité*] définit un objet par trois points indissociables :

- l'imprécision est relative à un objet mal, incomplètement et/ou imparfaitement délimité, rendant ainsi impossible ou en tout cas difficile sa définition en tant qu'entité parfaitement individualisée par rapport à son environnement ;
- l'imprécision est inhérente à l'objet même. Elle n'est pas liée au fait d'une mauvaise connaissance ou d'une connaissance incomplète ou de la grossièreté des outils d'appréciation, de mesure et/ou d'étude, et de ce fait ne sera réduite ou résorbée si ces outils se perfectionnent ou si notre connaissance de l'objet progresse ;
- l'objet imprécis est donc, par nature, un objet à différenciation interne car : *grosso modo*, nous y distinguons un "cœur" qui appartient sans ambiguïté à l'objet, et une "frange" plus ou moins nette, plus ou moins continue, qui réalise la "transition" entre cet objet et son environnement [ROL 92, *op. cité*].

#### *Définition de l'incertitude*

Nous la définissons comme un état de doute, d'hésitation (*Grand Larousse Encyclopédique*) dans lequel peut se trouver un individu face à un objet ou une situation.

[ROL 00b, *op. cité*] va plus loin en démontrant que l'incertitude est donc liée, non à l'objet lui-même, mais à l'observateur. Elle est générée par des contraintes extérieures à l'objet que sont à la fois : l'importance et la qualité de l'information disponible sur l'objet, fortement dépendantes du niveau technologique et technique, du lieu et du moment d'observation ; les

capacités propres de l'observateur à traiter les informations, selon sa formation, sa perception, sa faculté de compréhension et d'interprétation de l'information. Ce " filtre cognitif" dépend largement de la personnalité même de l'observateur, de son expérience, de sa culture, et d'autre part son histoire ; il dépend enfin du niveau et de la qualité du contexte géo-socio-économique et culturel général dans lequel baigne l'observateur et dont il subit l'influence, de façon consciente ou non.

Une fois la lumière faite sur le terme "*flou*" nous allons dès à présent entrer dans le vif du sujet relatif aux éléments de logique floue.

### ***Valeurs de vérité floues***

En logique classique la valeur de vérité d'une proposition est l'une des valeurs *vrai* ou *faux* représentées en général par 1 et 0 ou V et F respectivement. La différence profonde est qu'en logique floue la valeur de vérité d'une proposition floue (la notion de proposition floue sera définie au paragraphe suivant) est une caractéristique floue d'un référentiel en l'occurrence [0,1]. Autrement dit, du fait que cette valeur de vérité dépend du prédicat flou c'est-à-dire de la notion vague (jeune, grand, beau, loin, proche etc.) énoncée dans la proposition, elle induit en conséquence la notion de degré (ou force) de vérité ou de fausseté. Ces valeurs de vérité sont obtenues par le biais des fonctions d'appartenance des ensembles flous qui les modélisent.

### ***Propositions floues***

Une proposition est un énoncé contenant un ou plusieurs prédicats flous. Sa valeur de vérité est elle-même floue. Les énoncés suivants sont des propositions floues : "les étudiants de cette année ont été peu brillants", "construire le super marché "proche" de la sortie d'autoroute", "l'hôpital doit se situer à "côté" d'un grand parking". On peut distinguer parmi ces exemples deux types de propositions floues : proposition floue simple et proposition floue composée. Nous n'allons pas entrer dans les détails ici.

### ***Implication floue***

L'opérateur d'implication floue : si (...) alors (...), ouvre le vaste domaine de la logique et du raisonnement flous, bases des systèmes à base de connaissance et d'aide à la décision floue [BOU 95, *op. cité* ; ROL 00b, *op. cité*].

Une implication floue associe deux propositions floues par l'opérateur "si (...) alors (...)".

Si  $V_1 = A \rightarrow V_2 = B$  [BOU 95, *op. cité* ; ROL 00b, *op. cité*],

où  $V_1$  est une variable linguistique,  $A$  un terme sémantique, élément de l'alphabet de  $A$ , et  $V_2$  est une variable linguistique définie également par un terme sémantique  $B$  élément de son propre alphabet (nous parlerons de l'alphabet dans la prochaine section).

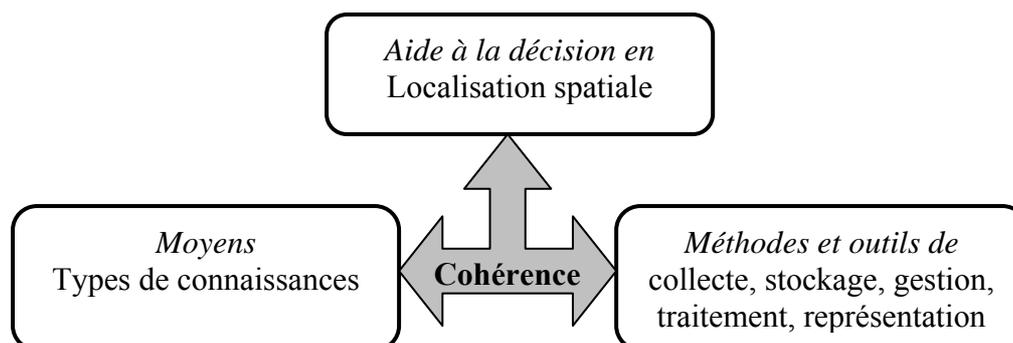
En toute généralité et au plan théorique, [BOU 95, *op. cité*] définit "l'implication floue entre deux propositions floues comme une proposition floue concernant le couple de variables associées ( $V_1$  et  $V_2$ ). La valeur de vérité de cette proposition peut être donnée par la fonction d'appartenance  $f_R$  d'une relation floue entre  $V_1$  et  $V_2$ , pour tout couple  $(x,y)$  de ( $V_1 \times V_2$ ) par :  
 $f_R(x,y) = \Phi((f_{V_1}(x), f_{V_2}(y)))$  ;  
avec  $\Phi$ , fonction de  $[0,1] \times [0,1]$  vers l'ensemble d'arrivée  $[0,1]$ ".

### 3.2.3. La chaîne de cohérence : phase théorique et connaissances floues

Très souvent les travaux d'études relatives au territoire privilégient les connaissances quantitatives, car elles s'avèrent plus faciles à traiter par des méthodes statistiques d'analyse uni- bi- ou multivariées. De même si des connaissances qualitatives sont collectées (enquêtes) elles sont le plus souvent exploitées de façon "ordinaire" (calcul de fréquences), sans tenir compte de leur caractère fondamentalement flou [ROL 00b, *op. cité*]. Ici, l'idée défendue par l'auteur revêt une triple facette :

- toute connaissance, même incertaine et imprécise, de quelque niveau qu'elle soit, est toujours bonne à prendre en matière d'aide à la décision à référence spatiale, car elle est susceptible d'apporter, ne serait-ce qu'un infime éclairage supplémentaire, et de contribuer ainsi à optimiser la démarche décisionnelle ;
- mais pour ce faire, il est nécessaire d'adapter les outils d'investigation, de stockage, de traitement de cette information à son caractère spécifique flou ;
- le questionnement sur le type d'information disponible, ou susceptible de l'être, doit être placé (et résolu) en amont de tout processus d'aide à la décision. En réponse à cette interrogation, il est indispensable de disposer, à la fois d'un vaste panel de connaissances quantitatives et qualitatives, mais aussi d'un arsenal d'approches méthodologiques et d'outils informatiques (en respect de la chaîne de cohérence) pour recueillir, stocker, traiter représenter et mettre à jour l'ensemble de ces données.

Il en résulte la nécessité de réunir ce "triplet gagnant" [ROL 00b, *op. cité*] qui associe, en un schéma de cohérence unique, l'objectif et la thématique (en l'occurrence il s'agit d'optimisation de la localisation des services collectifs) de l'aide à la décision, les types de connaissances, les méthodes et outils cohérents avec ces deux éléments (figure 3.2)



**Figure 3.2** : “Triplet gagnant” de l’aide à la décision [adapté de ROL 00b, op. cité]

### ***Information et donnée floues***

Dans la *section 1.3* nous avons explicité dans le premier chapitre les termes *information* et *donnée*. Reprenant [ROL 96, op. Cité; ROL 00b, op. Cité], notre approche des données et informations qui constituent l’ossature des connaissances floues, utilise le formalisme de la théorie des possibilités. Il nous paraît le plus adapté, à la fois à une démarche scientifique rigoureuse, claire et facilement explicitable, et à l’objectif qui est le nôtre c’est-à-dire la perspective de développement d’un système efficace d’aide à la décision et suffisamment simple pour être proposé aux gestionnaires du territoire.

Admettons que nous ayons une donnée telle que “*faible*”. Elle comprend bien une “granule” d’information flou : d’une part, sorti de son contexte, ce “granule” n’a aucun sens et n’apporte aucune information, d’autre part, cette donnée est linguistique. Exprimée en langage naturel, elle ne fait référence à aucune valeur numérique susceptible de nous renseigner sur une entité encore inconnue en l’état actuel de notre connaissance, dont nous savons seulement qu’elle est “faible”.

Ajoutons maintenant que : “*la couverture territoriale de tel service collectif est faible*”. Nous disposons alors une information inscrite dans un contexte bien déterminé : elle est relative à un objet (l’équipement qui assure le service collectif) défini par une variable (couverture territoriale) dont nous connaissons l’intensité (faible).

En l’état actuel de la connaissance énoncée, on constate que malgré son imprécision, “l’information passe bien”. Le terme linguistique utilisé est parfaitement compréhensible par les “acteurs territoriaux”, soit parce qu’ils n’ont pas mieux à proposer (quand toute information plus précise manque ou n’est pas disponible), soit parce que cette notation qualitative est justement suffisamment vague pour assouplir une information qui serait trop sèche si elle est quantifiée. Dans le premier cas l’imprécision est le seul moyen de disposer d’une information, même si elle est peu satisfaisante dans un contexte de rationalité absolue,

dans le second cas on recherche volontairement un contexte de rationalité relative, la souplesse sémantique d'un terme flou étant plus à même de traduire les situations du monde réel, où règnent naturellement l'entre-deux, l'imprécis et le vague.

Ainsi certaines informations habituellement non quantifiables sont à retenir, car elles apportent un éclairage supplémentaire. Les rejeter sous prétexte du flou revient à se priver d'une partie souvent stratégique de la connaissance du phénomène étudié.

### *Terminologie de base et formalisation*

On appelle :

- *variable linguistique*, une variable exprimée en langage naturel (ici la variable “couverture territoriale”), dont la qualité et/ou l'intensité sont exprimées, soit uniquement par des termes sémantiques. Dans ce cas elle est définie comme une variable linguistique sémantique, soit par des valeurs numériques et des termes linguistiques (par exemple la variable “âge d'une personne” peut accepter à la fois des valeurs numériques, et des termes linguistiques “une personne est âgée”). Dans ce cas elle est définie comme une variable linguistique numérisable ;
- *terme sémantique d'une variable* d'une variable linguistique, un terme permettant d'informer l'utilisateur sur l'intensité de la variable dans un cas précis ;
- *alphabet* de la *variable linguistique*, l'ensemble des termes sémantiques retenus pour évaluer (ou, plus exactement, pour “valuer”) la qualité, l'intensité de cette variable.

Soit  $V$ , la variable linguistique “couverture territoriale”. Il n'est pas à priori possible de l'exprimer par des valeurs numériques. Nous n'avons guère d'autres choix que d'utiliser, pour en exprimer la qualité et/ou l'intensité, des appréciations qualitatives. Nous parlons de variable linguistique sémantique. Son alphabet peut être :

$$A_V = \{\text{nulle, faible, moyenne, élevée, très élevée}\} \text{ [ROL 00b, op. cité]}$$

Soit  $A_g$  la variable “âge”. Si nous ne disposons pas d'éléments pour apprécier quantitativement l'âge d'une personne, nous pouvons éventuellement disposer d'informations qualitatives. Ainsi définie,  $A_g$  peut prendre ses valeurs à la fois dans l'intervalle  $[0,120]$  et dans son alphabet :

$$A_{A_g} = \{\text{très jeune, jeune, moyen, mûr, élevé, très élevé}\} \text{ [ROL 00b, op. cité]}$$

En toute généralité, toute variable utilisée peut être considérée comme une variable linguistique (sémantique ou numérisable) si, du fait de l'imprécision des connaissances, nous n'avons à notre disposition que des termes sémantiques la concernant. La connaissance que nous avons est incontestablement plus faible que si nous avons une expression quantifiée de l'information, mais, dans certains d'absence d'informations précises, elle a le mérite d'exister, à condition que nous sachions la formaliser le plus rigoureusement possible.

### **Formalisation d'une variable linguistique**

#### *Variables floues*

Dans l'analogie établie entre la théorie des ensembles classiques et la logique booléenne, une variable binaire  $a$  est associée à la fonction caractéristique  $f_A(x)$  d'un sous-ensemble  $A$  de l'univers de référence  $U$ , et  $x \in \{0,1\}$ . Cette analogie peut être étendue à la théorie des ensembles flous dans  $U$  et la logique floue. De cette manière, on peut immédiatement définir une variable linguistique. Ceci signifie que, pour tout  $x \in \{0,1\}$   $f_A(x) = 0$  ou  $f_A(x) = 1$  c'est-à-dire  $x$  appartient à  $A$ , ou (exclusif)  $x$  n'appartient pas à  $A$ .

Prenons par exemple un ensemble de communes  $E = \{c_1, c_2, \dots, c_{10}\}$  et définissons  $A$  un sous-ensemble de  $E$ , dont les éléments (les communes) bénéficient d'une subvention. On écrit :  $A = \{(c_1, 1), (c_5, 1), (c_9, 1)\}$  l'ensemble des communes candidates pour recevoir de nouvelles infrastructures.

Reprenant les résultats des travaux de Rolland-May [ROL 96, *op. cité*], nous définissons  $S$ , le support de la variable, comme l'intervalle sur la droite des réels dans lequel la variable linguistique se "situe" :

- si la variable est linguistique sémantique, nous fixons traditionnellement le support à  $[0,100]$ , [ZAD, 65]

- si la variable est linguistique numérisable, le support sera l'intervalle borné inférieurement et supérieurement, respectivement par les valeurs minimale et maximale (valeurs définies par dire d'expert ou par référence à des travaux antérieurs), puis généralement ramené par calcul à l'intervalle  $[0,100]$  (dans ce cas la variable est dite *normalisée*). L'objectif est de situer sur  $S$ , support de  $v$  chaque terme linguistique  $t \in A_V$ , alphabet de la variable  $V$ . Pour ce faire, on définit la *fonction de possibilité*  $\pi$  du *terme linguistique*  $t$  sur le support  $S$  de la variable  $V$  :

$$\begin{aligned} \pi : S &\longrightarrow U = [0,1] \\ \forall x \in S &\longrightarrow \pi_t(x) \in U \end{aligned}$$

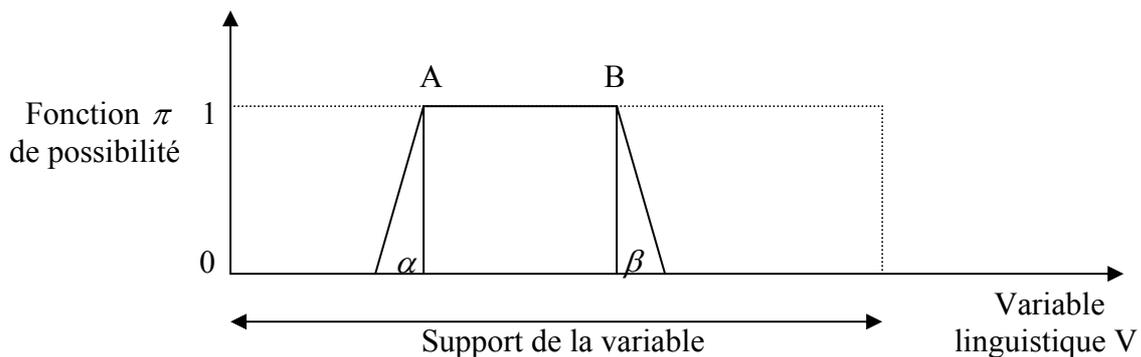
La fonction de possibilité  $\pi$  associe à chaque valeur  $x$  du support  $S$  de la variable  $V$ , la *possibilité* que cette valeur appartienne au terme sémantique  $t$  étudié.

La fonction de possibilité est formalisée de façon simple et programmable [DUB 87] comme suit :

- $\pi$  est normalisée : si et seulement si  $\forall x \in S, (\pi(x) \in [0,1])$ , avec  $\sup \pi(x) = 1$  ;
- elle est maximisée dans l'intervalle de  $S = [A, B]$ , avec  $A \leq B$  ; où  $\pi(x) = 1$  ;
- elle est linéaire et croissante à gauche de l'intervalle  $[A, B]$ , dans l'intervalle  $[(A-\alpha), A]$  ;
- elle décroît linéairement à droite de l'intervalle  $[A, B]$ , dans l'intervalle  $[B, (B+\beta)]$  ;
- elle est nulle dans les intervalles  $[0, (A-\alpha)]$  et  $[(B+\beta), 100]$ .

Il en résulte que le terme sémantique  $t$  est défini par :

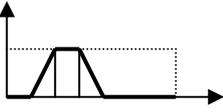
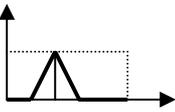
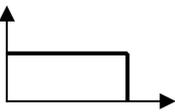
- la représentation graphique *trapézoïdale* de sa fonction de possibilité (figure 3.3) ;
- son profil sémantique  $Ps_t = (A, B, \alpha, \beta)$ .



L'intérêt de ces notations est considérable. En effet, une écriture unifiée de l'information précise ou "hybride" est désormais possible. Elle consiste à fuzzyfier "rendre flous" les variables utilisées. Ainsi, que les variables soient quantitatives mesurables, quantitatives intervalles, linguistiques mesurables et/ou linguistiques sémantiques, la "valeur" de la variable associée à un objet s'écrit toujours sous forme d'un profil sémantique de type  $(A, B, \alpha, \beta)$ , avec les contraintes  $(A-\alpha) \geq 0$  et  $(B+\beta) \leq 100$ , valables dans tous les cas et qui permettent à la base du trapèze représentatif du profil sémantique d'être contenu dans le support [ROL 96, *op. citée*].

Sans entrer dans les détails, nous présentons ici (tableau 2.1) les différentes représentations graphiques de la fonction de possibilité  $\pi$  selon la nature (modalité) de la variable d'étude.

Modalité de la variable	Valeur de la donnée de base	Ecriture unifiée par fuzzification	Représentation graphique de la fonction de possibilité $\pi$
Variable quantitative mesurable	$A = B$ $\alpha = \beta = 0$	$(A, A, 0, 0)$	
Variable quantitative intervalle	Valeur comprise entre A et B Avec $\alpha = \beta = 0$	$(A, B, 0, 0)$	

Variable qualitative	Valuation linguistique	$(A, B, \alpha, \beta)$	
Variable linguistique sémantique	Précision autour d'une valeur $A = B$ et $\alpha = \beta \neq 0$	$(A, A, \alpha, \alpha)$	
Variable de type quelconque. Absence d'information	Imprécision totale : valeur inconnue $A = 0$ ; $\alpha = \beta = 0$ et $B = 100$	$(0, 100, 0, 0)$	

**Tableau 2.1** : Modalités possibles et représentation de la fonction  $\pi$  des variables d'étude  
[Adapté de ROL 00b, op. cité]

### 3.2.4. La chaîne de cohérence : phase mathématique et modélisation des contraintes de localisation

Nous référant à l'approche systémique orientée objet du territoire où les objets (divers phénomènes socio-éco-géographiques) entretiennent des relations très dynamiques, la modélisation statique et comportementale de ces objets dans une perspective d'aide à la décision pose tout d'abord des problèmes proprement mathématiques et algorithmiques. Aussi doit-on disposer d'un modèle assez complet dans le cadre de la résolution des problèmes de localisation optimale des services collectifs.

La localisation optimale consiste, après avoir fait un maillage du territoire, à identifier :

- un ou plusieurs sites, soit dans un espace vierge, soit dans un espace déjà occupé (ce qui est le cas le plus fréquent) que l'on se propose de réorganiser ; ou
- un ou plusieurs services collectifs à implanter sur un ou des sites disponibles dont les caractéristiques sont bien connues afin d'optimiser non seulement l'accessibilité des populations à ces différents services mais aussi l'usage qu'elles en font.

Les approches les plus utilisées dérivent de l'économétrie et se fondent sur la recherche de la configuration optimisant une fonction d'utilité globale sous différentes contraintes. Dans le cadre de la recherche qui nous occupe, le recours au modèle de la p-médiane ou problème Loc-Alloc [DAS 95, op. cité ; HANS 96, op. cité ; LAB 95, op. cité] réputé pour la résolution des problèmes de localisation d'établissement et d'allocation de ressources, comme fonction d'utilité présente un grand intérêt [PEE 02, op. cité ; GRA 01 ; GOL 94 ; GOS 87 ; etc.]. Cependant, ce modèle demeure assez simpliste car faudra-t-il lui associer, en raison de l'ampleur de la dynamique socio-éco-territoriale et cognitive engendrée par le comportement

des usagers et la nature (type) des services à localiser, les conditions d'optimisation qui se traduisent par les contraintes relatives à :

- la couverture territoriale de chaque type d'équipement,
- la cohérence territoriale dans la répartition des équipements,
- les interactions spatiale et territoriale des usagers,
- la dynamique spatio-temporelle des usagers potentiels,
- la prise en compte des externalités positives et/ou négatives émises par chaque nouvel équipement.

Dans les relations mathématiques suivantes, modélisant les contraintes de localisation (à l'exception du modèle entropique), nous avons introduit des paramètres flous traduisant des facteurs qualitatifs. Nous n'explicitons pas ici la formalisation complète de ces paramètres. ce sera chose faite durant nos travaux de thèse.

**1- la couverture territoriale de chaque type de service** : cette couverture peut être mise en évidence à l'aide d'une méthode de mesure de la dispersion spatiale. Pour ce type de problématique, on aura recours, entre autres, à l'entropie [CLI 98a ; CLI 98b ; RUL 98]. Utilisée sous sa forme relative, l'entropie permet d'obtenir un indice compris entre 0 et 1, plus aisément comparable entre services. Cela permet d'éviter la nature arbitraire des comparaisons visuelles des cartes [SAN 90]. D'un point de vue opérationnel, son calcul exige peu de données, puisque seules la localisation des équipements et les mailles territoriales sont indispensables. Sa formalisation classique, est la suivante :

$$ER = E/\log k \ ; \text{ avec } E = \sum_{i=1}^k f_i \log 1/f_i \quad [1]$$

où :

ER = entropie relative et E = entropie ;

k = nombre total de mailles du découpage territorial ;

$f_i$  = fréquence de types de service dans la maille territoriale i ( $f_i = n_i/N$ ) ;

$n_i$  = nombre de service dans la maille i ;

N = nombre total de services sur l'ensemble du territoire considéré ;

**2- la cohérence territoriale dans la répartition** : établir la cohérence territoriale d'un réseau d'équipements offrant le même type de service collectif peut se faire à l'aide d'un coefficient d'autocorrélation spatiale [RUL 02, *op. Cité*]. L'optique d'un tel coefficient est différente de l'entropie relative. Ici, l'idée est orientée sur la différence entre mailles territoriales contiguës et non contiguës : les mailles frontalières (sachant que la frontière, bien que délimitée par une ligne, participe du flou du point de vue de

la dynamique du fonctionnement) sont-elles plus semblables entre elles que les mailles éloignées ? L'autocorrélation spatiale se révèle-t-elle un outil géographique mettant en évidence une structure spatiale. L'analyse de l'autocorrélation spatiale permettra de mettre en exergue la cohérence et la qualité de la logistique de chaque réseau d'équipements afin d'optimiser le choix des localisations potentielles. Nous inspirant des travaux de Cliff et Ord issus des résultats des tests d'autocorrélation de Durbin-Waston et de Von Neumann [CLI 73 ; CLI 75 ; RUL 02, *op. Cité*], nous posons la formalisation floue du coefficient d'autocorrélation C comme suit :

$$C = \frac{(k-1) \sum_{i=1}^k \delta_{ij} \sum_{j=1}^k \delta_{ij} (n_i - n_j)^2}{4F \sum_{i=1}^k (n_i - x)^2} \quad [2]$$

avec :

k = nombre total de mailles du découpage territorial ;

F = nombre de frontières total entre les mailles du territoire considéré ;

$n_i$  = nombre de services dans la maille i ;

$n_j$  = nombre de services dans la maille j ;

x = la moyenne arithmétique de l'ensemble des équipements selon chaque type de service ;

$\delta_{ij}$  = le *paramètre flou* (que nous aurons à estimer par des méthodes d'agrégation floue) exprimant la relation de voisinage entre les mailles i et j, sachant que  $0 \leq \delta_{ij} \leq 1$  ;

**3- les interactions spatiale et territoriale des usagers** : les modèles d'interaction spatiale partent de l'hypothèse que les relations qui se nouent entre les populations (flux de personnes entre) de deux lieux i et j dépendent de la distance (exprimée en termes de temps ou de kilométrage ou encore en termes de distance topologique). En poussant un peu plus loin l'analyse on se rend compte qu'en dehors de la distance, il existe d'autres effets de barrière<sup>13</sup> de type économique, sociologique ou cognitif [STO 60 ; GRA 01, *op. cité*]. En conséquence, partant de la définition de l'interaction spatiale, on définit l'interaction territoriale comme étant l'interaction spatiale conjuguée avec les effets de barrières (étant des facteurs qualitatifs d'interaction). Cette notion participe également du flou par l'introduction d'un paramètre flou  $\delta_{ij}$  défini dans la formalisation de la cohérence territoriale. Ainsi, nous nous appuyons sur le modèle additif plutôt que multiplicatif et à double contrainte d'interaction spatiale proposé par Tobler et Dorigo [TOB 83 ; GRA 97], modèle qui présente des propriétés

---

<sup>13</sup> Les effets de barrière, c'est-à-dire la réduction des flux induits par le franchissement d'une limite ne peuvent généralement pas être mesurés dans l'absolu [GRA 01, *op. cité*] car ils sont le plus souvent étroitement associés à des phénomènes d'*intégration territoriale*, c'est-à-dire de concentration des échanges à l'intérieur des limites des mailles territoriales.

remarquables [GRA 97, *op. cité* ; GRA 01, *op. cité*], notamment en ce qui concerne la question de la désagrégation des matrices en unités de tailles de plus en plus réduites et l'établissement des représentations cartographiques des champs de transfert (analyse de la composante dissymétrique des flux). Nous formalisons l'interaction territoriale comme suit :

$$M_{ij} = \lambda_{ij}(O_i + D_j) * f(d_{ij}) \quad [3]$$

avec :

$M_{ij}$  = matrice des flux des usagers entre les mailles i et j ;

$\lambda_{ij}$  = paramètre flou (à développer durant la thèse) ;

$O_i$  = la somme des flux d'origine de la maille i ;

$D_j$  = la somme des flux d'origines de la maille j ;

$f(d_{ij})$  = fonction (exponentielle ou parétienne) exprimant la distance (temporaire, topologique ou métrique) entre les mailles i et j ;

**4- la dynamique spatio-temporelle des usagers potentiels** : le territoire étant continu et hétérogène, la modélisation du processus ne peut se faire conformément aux règles de base d'un modèle discret [PUM 97 ; PUM 01 ; RAV 98 ; SAI 01] ; le recours à une formalisation d'un modèle continu du temps et de l'espace s'avère donc indispensable. De plus, faudra-t-il tenir compte de l'inégale répartition des usagers entre les mailles territoriales ainsi que leur évolution au cours du temps, afin sur afin de prévoir la capacité des infrastructures potentielles à implanter pour optimiser leur localisation. Ainsi, reprenant les résultats de Cliff et Ord [CLI 75, *op. cité* ; RUL 02, *op. cité* ; SAI 01, *op. cité*], nous formalisons la variation du nombre des usagers potentiels présents au cours du temps dans une maille territoriale quelconque i comme suit :

$$\frac{dX_i(t)}{dt} = [P_i - X_i(t)] * \left[ a_i + \sum_{j=1}^k \gamma_{ij} X_j(t) \right] \quad [\text{adapté de CLI 02}] \quad [4]$$

avec :

$X_i(t)$  = le nombre d'usagers de la maille i estimé pendant la période t ;

$X_j(t)$  = le nombre d'usagers de la maille j estimé pendant la période t ;

$P_i$  = la population totale de la maille i

$a_i$  = intensité du processus de communication interne à la maille i ;

$\gamma_{ij}$  = paramètre flou exprimant le processus de communication entre la maille i et la maille j ;

**5- la prise en compte des externalités positives et/ou négatives** : toute infrastructure a forcément à la fois des impacts positifs et négatifs sur l'environnement et plus globalement sur le cadre de vie des bénéficiaires et même des non bénéficiaires. L'objectif est donc de pouvoir prendre en comptes ces différentes externalités dans le

processus d'optimisation des choix de localisation affectation. En appelant  $E_p$ ,  $E_n$  et  $\varepsilon$  respectivement les externalités positives, négatives et le paramètre flou exprimant la qualité des externalités qui serait à développer, il sera possible d'établir une mesure d'externalité en fonction de  $E_p$  et  $E_n$ . Ainsi, on aura :

$$\text{Ext} = f(E_p, E_n) * \varepsilon \quad [5]$$

En définitive, techniquement, la complexité de notre problématique de recherche qui s'apparente en définitive à un problème d'optimisation, est mesurée par sa difficulté de résolution, elle-même mesurée par le nombre de fonctions mathématiques et le temps nécessaires pour que l'on puisse faire exécuter, par un système de traitement de l'information, un programme informatique qui en donne la solution. On a alors affaire à une explosion combinatoire, à priori, difficile à gérer.

En conséquence, dans ce contexte de modélisation très complexe, on conçoit d'emblée dans ces conditions, l'intérêt, sinon même la nécessité, de recourir à la mise en œuvre d'un système d'inférence neuronal [HOP 85 ; KOH 95 ; ROB 02 ; DRE 02] flous [COX 97, *op. cité* ; ROL 00b, *op. cité* ; ROL 02, *op. cité* ; RUT 01, *op. cité* ; FIS 01 ; GAO 04] qui constitue une méthode éprouvée d'intelligence artificielle très utile en recherche opérationnelle. On pourrait aussi tenter d'aborder notre problématique en ayant recours aux acquis de l'analyse multicritère [ROY 93 ; LAA 00, *op. cité* ; FUS 94 ; FUS 98] mais dans le cadre de notre recherche, le choix de l'approche système d'inférence (système neuronal et système expert) flou neuronal flou se justifie par sa puissance pour la modélisation non linéaire et ses avantages informatiques et mathématiques ; ce qui, du moins nous l'escomptons garantit de meilleurs résultats par rapport à l'approche multicritère.

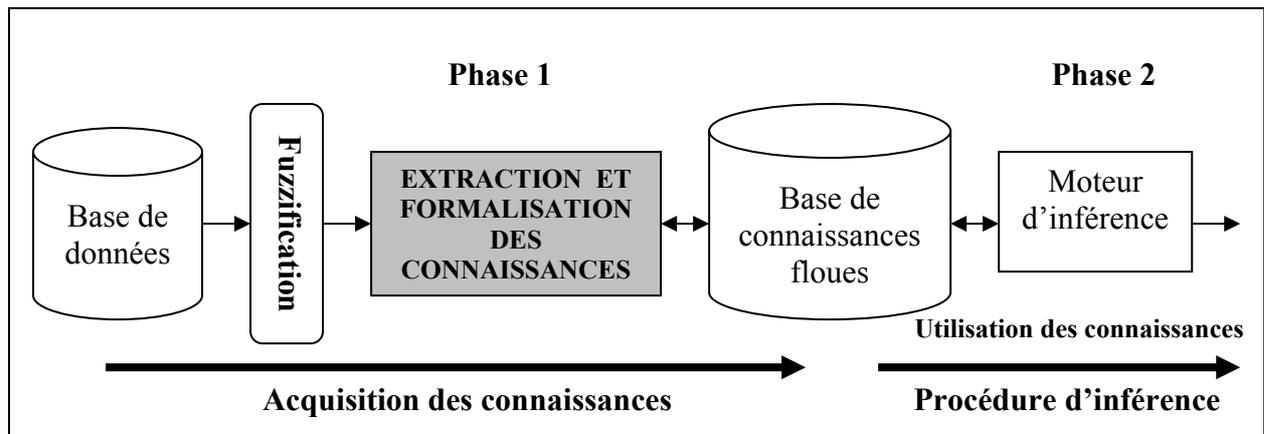
### **3.2.5. La chaîne de cohérence : phase méthodologique**

Notre démarche méthodologique consiste à construire un système d'inférence flou par association d'un réseau neuronal flou et d'un système expert flou. Concrètement, il s'agit d'une construction en deux phases.

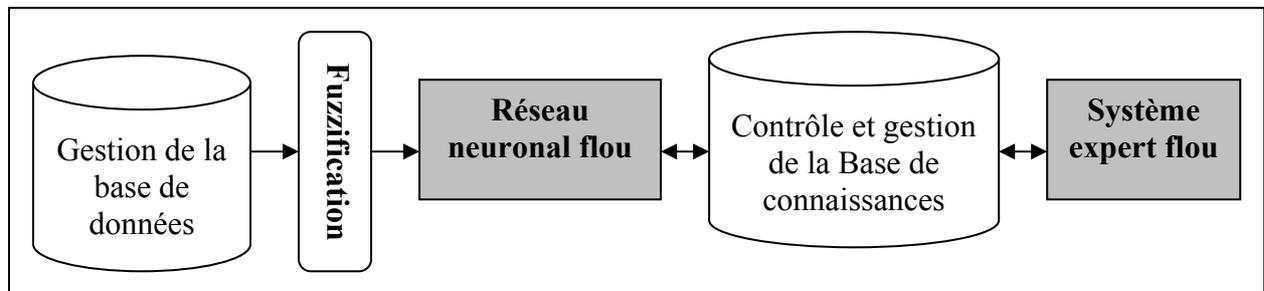
La première phase consiste à extraire et formaliser les connaissances (figure 3.3.a) relatives aux problèmes de localisation spatiale identifiés plus haut. Cette phase sera assurée par la mise en œuvre d'un réseau neuronal flou (figure 3.3.b). L'ensemble de ces connaissances sera ensuite stocké dans une base dite *base de connaissances floues* (figure 3.3.a).

La deuxième phase consiste à utiliser les informations contenues dans la base de connaissances floues pour produire l'aide à la décision au moyen d'un processus de

raisonnement intelligent supporté par le *moteur d'inférence* (figure 3.3.a), lequel faisant partie intégrante du système expert (figure 3.3.b).



**Figure 3.3.a :** Chaîne d'acquisition des connaissances au sein du système d'inférence



**Figure 3.3.b :** Système d'Inférence pour l'Aide en Localisation (SIAL)

Il ressort de ces considérations que le système d'inférence que nous entendons développer pendant la réalisation de notre thèse est un ensemble composé d'un module d'acquisition et de formalisation des connaissances relatives aux problèmes de localisation et d'un système expert dont le moteur d'inférence est chargé d'utiliser ces connaissances pour des fins d'aide à la décision.

### 3.3. Méthodologie d'acquisition des connaissances

#### 3.3.1. Problématique

L'objectif de notre travail de recherche étant la conception d'un système d'aide à la décision qui se caractérise par son extrême souplesse, autorisant à la fois l'acquisition et le traitement des connaissances floues relatives aux problématiques de localisation spatiale au sein du territoire ainsi que le mode de raisonnement associé, la démarche modélisatrice au plan logique et informatique impose une approche neuronale [DRE 02 ; FIS 01 ; ROL 02, *op. cité*]

qui participe aux nouvelles techniques efficaces d'acquisition de traitement de l'information [AMA 02].

En effet, cette phase qui consiste à recueillir la connaissance nécessaire n'est pas une tâche facile. Elle exige d'acquérir des connaissances auprès d'un expert, puis aidera à transcrire ces connaissances sous forme de règles en tentant d'être le plus exhaustif et le plus complet possible. Ce travail extrêmement ardu, représente un coût non négligeable en termes temporel et financier, car au-delà de la difficulté d'acquisition de la connaissance, il faudra surmonter d'autres obstacles d'ordre technique et tout particulièrement la transcription de la connaissance en symbole élémentaire. En résumé, on devra avoir à disposition un expert, soit d'énoncer et de formaliser les connaissances ; ce qui n'est la plupart du temps pas possible. La "solution de remplacement" est qu'actuellement les techniques de fouille de données (data mining) issus de l'intelligence artificielle permettent de surmonter ces difficultés, inhérentes à la disponibilité d'experts, par la mise en œuvre de réseaux de neurones fonctionnant en modes supervisé et/ou non supervisé [AMA 02, *op. citée*]. C'est cette voie que nous avons retenue dans notre projet.

### 3.3.2. Définition d'un neurone

Dans sa définition classique, un neurone est un automate élémentaire possédant une fonction algébrique non linéaire, paramétrée, à valeurs bornées, dont la tâche consiste, à partir de signaux<sup>14</sup> discontinus captés en entrée, d'élaborer un signal unique, par l'application d'une fonction prédéfinie (en général une fonction en forme de "S" telles les fonctions sigmoïde, seuillée, etc.) lors de la conception du neurone. La représentation d'un neurone (figure 3.4) est le reflet de l'inspiration biologique qui a été à l'origine des premières recherches dans le domaine pendant les années 1940 à 1970.

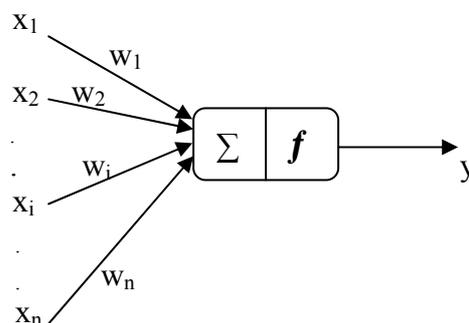


Figure 3.4 : Représentation d'un neurone artificiel

<sup>14</sup> On entend par signal, la variation d'une grandeur de nature quelconque (variables floues) servant de support à l'information.

Les variables sur lesquelles opère le neurone sont désignées par leurs signaux sous le terme d'entrées  $x_i$  du neurone, et la valeur  $y$  de la fonction sous celui de sortie.

Ce neurone artificiel est une simplification extrême du neurone biologique dont le comportement temporel, notamment, est fort complexe (importance dans le temps de propagation, importance des dispersions temporelles des entrées dans le comportement du neurone, évolution dans le temps lorsque le neurone a fonctionné, etc.). Dans ce modèle, les liaisons synaptiques sont remplacées par les paramètres (ou encore poids)  $w_i$  qui sont attachés aux entrées  $x_i$  du neurone. La fonction d'activation  $f$  est appliquée à la combinaison linéaire, des  $n$  entrées  $x_i$  pondérées par les poids  $w_i$ , générée par le sommateur  $\sum$ . Ainsi la sortie  $y$  donne  $y = f(\sum w_i x_i)$  ; ( $i = 1, \dots, n$ ).

### 3.3.3. Définition d'un réseau neuronal

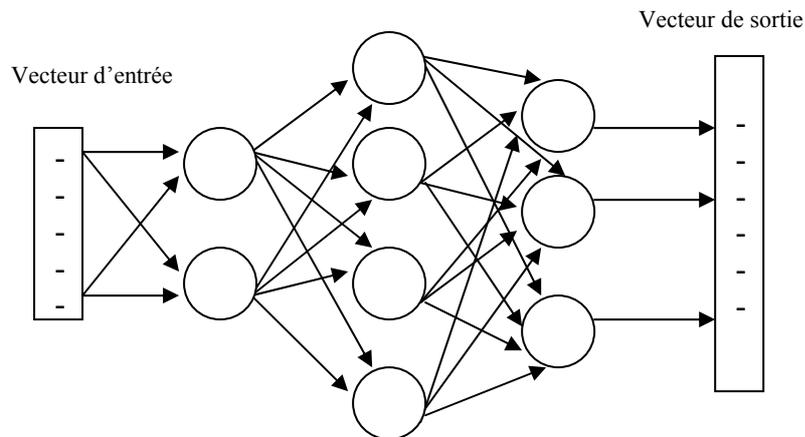
L'élément de base décrit plus haut, appelé neurone artificiel, est relativement simple et dénué d'intérêt si on le considère seul. Partant de la définition d'un neurone, le réseau neuronal se présente comme une association de plusieurs neurones interconnectés entre eux et organisés en plusieurs niveaux appelés "couches" selon une architecture de réseaux bien spécifique à chaque type de problème. Chaque couche du réseau comporte un ou plusieurs neurones entièrement connectés ou non entre eux. L'intérêt de l'organisation des neurones en réseaux réside dans le fait que le réseau favorise une meilleure *composition* des fonctions non linéaires issues des modèles traduisant la réalité.

De plus, les propriétés du réseau neuronal réellement remarquables sont leur capacité intrinsèque à la généralisation (aptitude à bien se comporter sur des données qu'ils n'ont pas appris). Nous y reviendrons dans la section 4.1 traitant de l'extraction de la base de connaissance du système d'inférence.

Il existe un certain nombre d'architectures spécifiques de réseaux possédant des propriétés particulières. Dans le cadre de notre recherche nous ne nous intéresserons qu'aux réseaux à couches (figure 3.5).

Le réseau à couches comprend :

- une couche d'entrée
- une ou plusieurs couches cachées dont le fonctionnement n'apparaît pas immédiatement en sortie. Les entrées d'une couche est assurée par les sorties de la couche précédente pondérées par des poids  $w_{i,j}$  (c'est-à-dire le poids du signal entre le neurone  $i$  de la couche  $c$  et le neurone  $j$  de la couche suivante) ; lorsqu'il s'agit de la couche d'entrée, les entrées ne sont rien d'autres que le  $x_i$  formant un vecteur; c'est-à-dire un ensemble de signaux
- la couche de sortie, dont le ou les neurones assure la ou les sorties du réseau.



**Figure 3.5 :** Configuration d'un réseau neuronal à couches

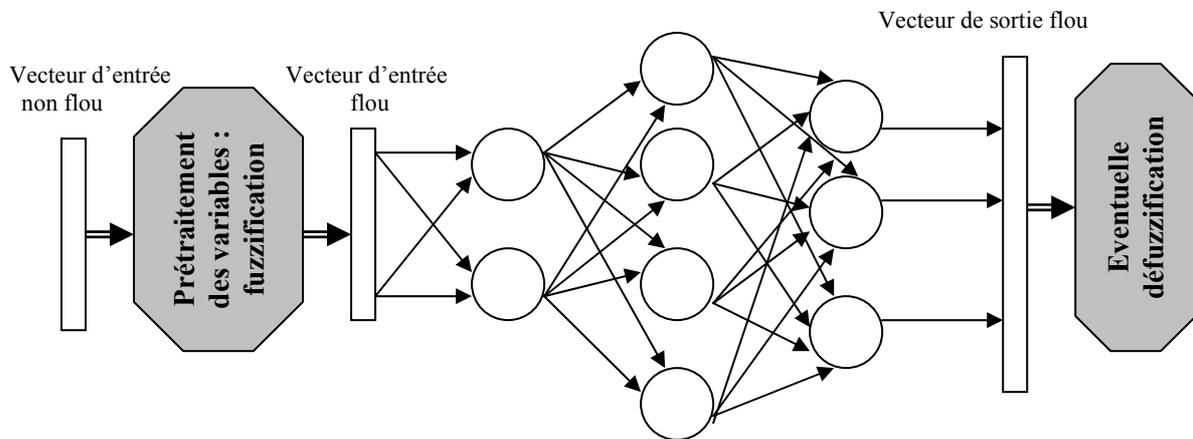
Les flèches sur les deux schémas représentent les poids synaptiques qui relient les neurones entre eux. Pour une architecture de réseau donnée, ces poids sont les seuls éléments qui peuvent varier au cours du fonctionnement du réseau (apprentissage du réseau).

### 3.3.4. Mise en œuvre du réseau neuronal flou

Le concept de réseau neuronal flou est largement développé tant au plan conceptuel qu'aux plans informatique et opérationnel [ROL 02, *op. cité*]. D'après la littérature spécialisée [BOU 95, *op. cité* ; BUC 99 ; ROL 02, *op. cité* ; RUT 01], on définit un réseau neuronal flou comme un réseau neuronal dont :

- les signaux d'entrées ne sont plus exclusivement numériques, mais peuvent également être qualitatifs. Dans ce cas l'entrée d'un neurone est libellée sous forme de termes sémantiques de la variable linguistique. Pour appréhender ces signaux d'entrée, nous utiliserons les profils sémantiques dans cette recherche;
- les poids des signaux sont soit, numériques, soit qualitatifs et sont exprimés selon les mêmes modalités ;
- les fonctions d'activation sont définies en cohérence avec les signaux d'entrée et leurs pondérations spécifiques en ayant recours aux opérateurs mathématiques flous ;
- les sorties du réseau neuronal sont exprimées, non pas par une valeur numérique, mais par une valuation floue. En d'autres termes les résultats produits par le réseau neuronal est exprimé selon un formalisme flou. Ce résultat pourrait être éventuellement "défuzzifié" grâce aux procédures de transformation d'une donnée floue en une donnée précise.

Il résulte de ces spécificités qu'un réseau neuronal flou est la généralisation d'un réseau neuronal classique, lequel n'est en réalité qu'un cas particulier de cette généralisation. En guise d'illustration (figure 3.6), nous représentons ici la formalisation floue du réseau de neurones à couches.



**Figure 3.6 :** Formalisation floue d'un réseau neuronal à couches

Soit une variable floue quelconque  $X$ , dont la valeur floue en entrée est représentée par un terme sémantique  $x_i$ , élément de son alphabet. Son écriture unifiée par fuzzification à l'entrée du réseau est représentée par son profil sémantique  $(A, B, \alpha, \beta)$ . Si la variable linguistique a un alphabet à  $n$  termes (cf section 3.2.3), chaque neurone devra être apte à capter l'un quelconque des  $n$  signaux pour l'entrée  $X$ . Le résultat (sortie du réseau) sera aussi flou. Il pourra être ensuite, éventuellement défuzzifié. Pour notre part, la sortie du réseau restera floue et ne sera défuzzifiée qu'après le processus d'inférence mené par le moteur d'inférence.

### 3.4. Le système expert

#### 3.4.1. Définition du système expert

Classiquement, on peut définir un système expert comme un programme informatique, possédant une masse de connaissances généralement fournies un spécialiste ou une équipe de spécialistes chevronnés (experts) dans le domaine d'application, implanté dans un ordinateur, et doté d'une faculté de raisonnement intelligent capable d'atteindre les performances de l'expert dans ce domaine, du moins dans les prestations les plus rudimentaires. Un tel programme informatique vise à simuler le raisonnement de l'expert humain. Les connaissances sont rangées dans une *base de connaissances* et exploitées par un *moteur d'inférence* en fonction d'une situation décrite dans une *base de faits*. A ce titre, le système

expert est composé d'une base de faits, d'une base de connaissances, d'un moteur d'inférence et éventuellement d'autres modules connexes comme dans le cas de cette recherche.

**La base de faits** comprend des *faits* permanents du domaine d'application (équivalents à ce que l'on trouve dans une base de données) et les faits propres à la problématique en cours de résolution. Ces faits sont généralement fournis par l'utilisateur du programme et/ou déduits par application d'une règle de production. C'est la mémoire de travail du système expert.

**La base de connaissances** est constituée de règles, dites règles de production qui permettent généralement de tirer des conclusions qui ajoutent éventuellement de nouveaux faits dans la base de faits à partir de ceux déjà connus. C'est le sous module contenant les connaissances exploitables par le système.

**Le moteur d'inférence** est chargé d'exploiter la base de connaissances de façon à conduire un raisonnement efficace sur le problème posé en fonction du contenu de la base de faits.

### 3.4.2. Définition de l'inférence

Par inférence, nous entendons le processus générale, qui consiste à utiliser les connaissances sur un objet, un système, au travers de *règles*, pour en tirer des conclusions. C'est une opération logique [SEA 89] qui consiste à affirmer q'une proposition est vraie en raison de ses liens avec une ou plusieurs autres propositions déjà vérifiées.

Une règle est un élément de connaissance mettant en relation la compatibilité de la prémisse (prédicat) d'une proposition floue avec celle d'un ou plusieurs espaces flous conséquents (conclusion) [COX 97, *op. cité*; ROL 00b, *op. cité*; BOU 03]. C'est le mode de représentation des connaissances dans un système expert. L'exécution de ces règles se fait à partir d'un gestionnaire de règles ou appelé *Gestionnaire de pile* et d'une *base de connaissances* c'est-à-dire une base de données formelle relative à la problématique étudiée, dans un processus de *chaînage arrière* et /ou *chaînage avant*.

L'inférence floue est celle basée sur des règles floues et qui utilise pour leur représentation et leur traitement des outils avancés (généralisation des outils classiques) tels que l'implication floue, le modus ponens généralisé ou les fonctions de possibilités. L'inférence floue est proche du raisonnement humain, car les règles sont celles énoncées constamment dans le langage naturel. En combinant l'implication floue et la fonction de possibilité, la logique inférentielle floue permet de prendre en compte le vague inhérent aux connaissances de base pour optimiser les paramètres d'aide à la décision. La structure courante des règles d'inférence est de type “*si-alors*” et “*si-alors-sinon*”.

En nous appuyant sur la figure 3.7 traduisant la chaîne de traitement des données, nous mettons en exergue l'illustration de l'inférence floue. D'après cette figure, nous distinguons trois étapes dans la chaîne de traitement :

- la *fuzzification* est la première opération du traitement flou proprement dit. Elle consiste à fournir des entrées floues au système d'inférence ;
- l'*inférence floue*, ou l'évaluation des règles, calcule les sorties floues à partir des entrées floues issues de la fuzzification et de l'ensemble des règles définies selon un mode d'agrégation<sup>15</sup> bien précis.
- La *défuzzification* opère la transformation des sorties floues en valeurs précises par des techniques particulières dont celle la plus usitée est connue sous le nom de "Centre de gravité" ou encore "Centroïde" (figure 3.7).

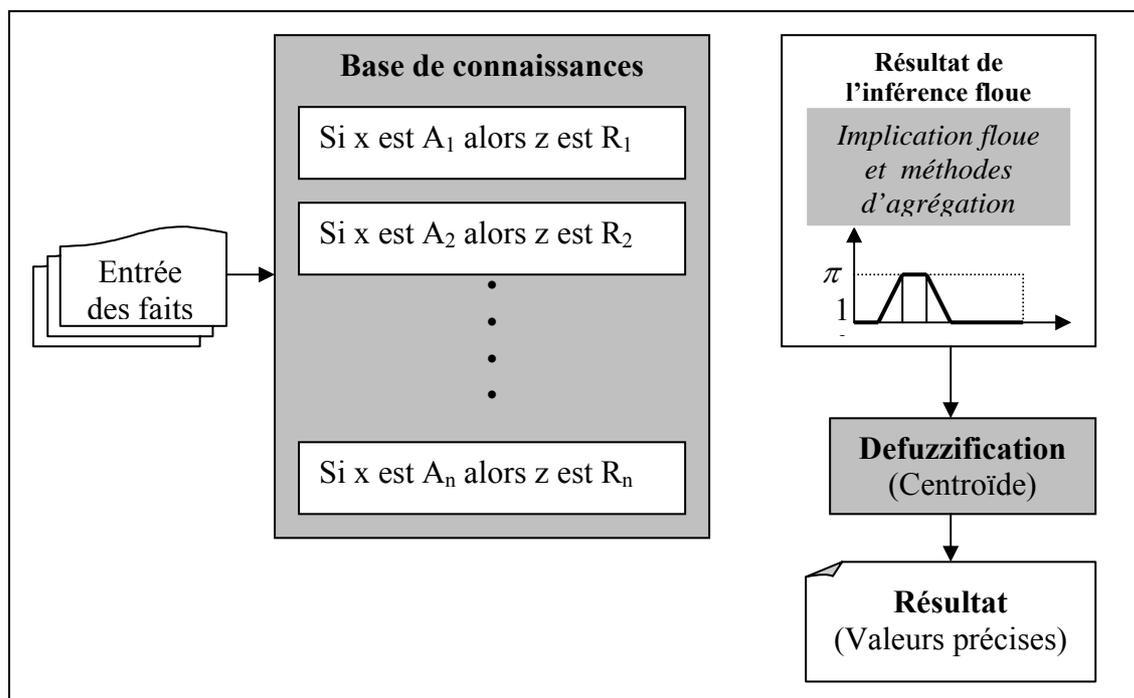


Figure 3.7 : Chaîne de traitement floue des données

En guise d'illustration, considérons l'exemple suivant où  $R_i$  ( $i = 1, 2, 3$ ) désigne une règle d'inférence :

$R_1$  : si une commune à plus de 5000 habitants et proche du centre-ville, alors construire une ligne de tram et installer des couloir de bus ;

$R_2$  : si une commune à moins de 5000 habitants et proche du centre-ville, alors installer un couloir de bus ;

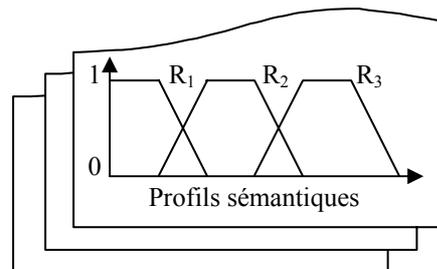
$R_3$  : si une commune est de taille moyenne (entre 1500 et 3000 habitants) et loin du centre-ville, alors construire une ligne de tram.

<sup>15</sup> L'agrégation est un processus de combinaison d'ensembles flous pour en construire un unique qui représente la variable solution. C'est la partie essentielle du raisonnement dans les systèmes flous. On distingue deux méthodes d'agrégation : la méthode additive et la méthodes Min/Max.

Problème à résoudre (fait) : une commune à plus de 5000 habitants et se situe loin du centre-ville que faire ?

Pour aider à prendre une décision face à ce problème de localisation (qui est en réalité un cas d'école), on pourrait procéder comme suit :

- 1- dresser le profil sémantique des termes “plus de 5000”, “moins de 5000”, “proche”, “loin” ;
- 2- évaluer les écarts sémantiques, en utilisant la méthode de Rolland-May [ROL 03, *op. cité*] ;



3 – une fois les écarts obtenus, on identifie sur le support des variables la valeur correspondant au terme résultant les écarts. Et par la même occasion on produit l'aide à la décision c'est-à-dire la partie droite (la conclusion) de la règle qui traduit le fait.

En conclusion, nous notons que, dans ce chapitre, nous avons exposé et appliqué les différentes phases de la chaîne de cohérence décisionnelle à la problématique qui nous vaut la recherche entreprise. Ceci nous a permis d'introduire et d'explicitier le réseau neuronal flou ainsi que le système expert flou tout en insistant sur le rôle qui leur incombe et leur partition dans le Système d'Inférence pour l'Aide en Localisation (SIAL). Ceci étant, nous allons procéder au développement de la méthodologie de conception du SIAL.

## CHAPITRE IV

# METHODOLOGIE DE CONCEPTION DU SYSTEME D'INFERENCE POUR L'AIDE EN LOCALISATION : S.I.A.L

Le SIAL est un ensemble associant deux composantes majeures : un module d'acquisition/extraction et de formalisation des connaissances relatives aux problèmes de localisation et un système expert dont le moteur d'inférence est chargé d'utiliser ces connaissances à des fins d'aide à la décision.

Afin d'accomplir avec efficacité le rôle qui est le sien dans un processus de décision, un véritable système d'aide à la décision en localisation spatiale associant les acquis de la logique floue et des réseaux de neurones doit se doter d'une capacité d'acquisition de connaissances (base de connaissances) nécessaires au traitement des problèmes qui lui seront confiés et être doté d'une faculté de raisonnement assez puissante. Dans le cadre de notre recherche, le SIAL, héritant des acquis des systèmes de raisonnement flou, se compose de sous modules tels la base de connaissances, le moteur d'inférence et le générateur de scénarii pour successivement acquérir et extraire les connaissances, raisonner sur les connaissances, simuler les situations inédites.

Pour répondre à ces trois objectifs, nous reprenons ici, en l'approfondissant les deux étapes de la figure 3.3.a et ajoutons dans le dernier paragraphe la fonction essentielle du SIAL : celle de construire et d'évaluer les scénarii de localisation.

### **4.1. Extraction/Acquisition des connaissances**

#### **4.1.1. Généralités sur la base de données**

En raison de la nature floue des données à référence spatiale et du volume d'information dont on a constamment besoin pour traiter les questions relatives à l'aménagement du territoire, la conception et la mise en œuvre d'outils efficaces permettant d'accéder aux données pertinentes lors du processus décisionnel devient une nécessité absolue [DRE 02, op. cité ;

EST 98 ; LON 02 ; ZEI 00a]. Recourant à la définition, des bases de données à référence spatiale, à laquelle nous souscrivons (*section 1.3*), nous articulons la gestion de la base de données, dans le cadre de la problématique générale qui nous occupe, sur les concepts, méthodes et outils de *Data Warehousing*, des *Métadonnées* et du *Data Mining*.

La définition classique du **Data Warehouse** donnée par Inmon [INM 94] est la suivante : “*Le Data Warehouse est une collection de données orientées sujet, intégrées, non volatiles et historisées, organisées pour le support d’un processus d’aide à la décision*”. Cette définition apporte deux considérations nouvelles qui s’ajoutent aux propriétés des SGBD traditionnels. La première considération est que la base de données doit être organisée autour des sujets majeurs du processus décisionnel. L’objectif est de ne plus se limiter à la vérification de la non redondance des données mais aussi procéder à leur organisation selon une approche thématique. L’intérêt de cette organisation est de disposer de l’ensemble des informations utiles sur chaque aspect, le plus souvent transversal, de la problématique étudiée. Cette orientation nous permettra également de développer le système d’aide à la décision via une approche incrémentale thématique après thématique.

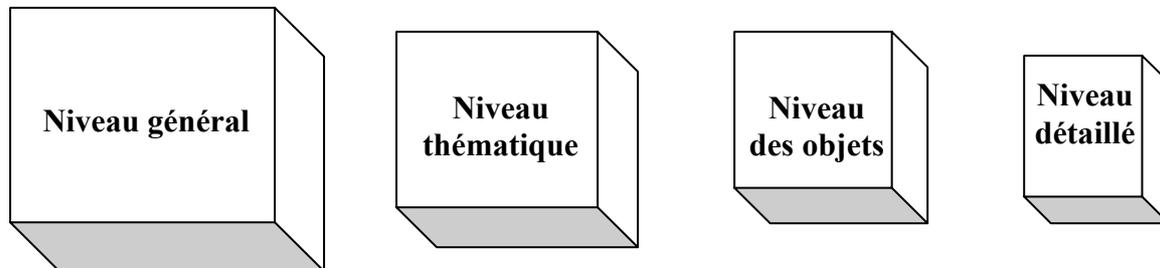
La deuxième considération est relative à l’historique des données. Généralement, à la mise à jour, l’ancienne valeur est perdue et la donnée reste constamment à jour. Le Data Warehouse permet de mettre à jour les données tout en conservant leurs anciennes valeurs. En réalité les données ne sont pas mises à jour mais elles présentent des valeurs insérées dans le système à un certain moment. L’historique des données est ainsi stocké, c’est-à-dire l’ensemble des valeurs que les données auront prises au cours du temps. Il est alors évident qu’un référentiel temporel soit associé à la base de données.

L’historisation des données favorisera non seulement le suivi et l’analyse dynamique des phénomènes mais servira aussi de support aux activités de simulation et de prédiction. La non volatilité est en quelque sorte une conséquence de l’historisation précédemment décrite et une opportunité très intéressante de “travail du SIG”.

Dans ce contexte, les **Métadonnées** ont pour but de fournir de l’information sur les données intégrées dans le système afin que l’on puisse mieux apprécier leur qualité et leur validité et en faire une utilisation pertinente. Elles portent généralement sur l’historique des données et leur mesure lors de la saisie informatique. Elles répondent aux besoins toujours plus grands d’identification, et de description des informations contenues dans la base (type de représentation, territoire couvert, échelles spatiale et thématique, temporalité, format, alphabet, sources etc.) ; elles contiennent enfin les éléments utiles aux requêtes spatiales et thématiques.

Les données de la base étant multisources, les métadonnées aident à gérer les structures, les contenues et les niveaux de qualité ; elles permettent ainsi de maîtriser l’exploration des

données. En effet, elles facilitent le filtrage et le contrôle des informations lors de leur intégration dans le processus d'aide à la décision. Nombreuses sont les informations utiles à la description des données et donc susceptibles d'être contenues dans les Métadonnées. En conséquence, nous distinguons plusieurs niveaux de métadonnées (figure 4.1)



**Figure 4.1** : *Différents niveaux de métadonnées*

Aussi le modèle conceptuel de données participe-t-il à la mise en œuvre des métadonnées qui constituent une aide non négligeable pour la concrétisation des opérations de Data Mining.

Les réservoirs de connaissances mises en évidence par le Data Warehouse et les métadonnées doivent être explorés afin de débusquer les relations entre les données et d'en déduire des modèles de comportement. Les outils de **Data Mining** basés sur les techniques des réseaux de neurones offrent d'excellents résultats [DRE 02 ; HER 00 ; MUL 00 ; TEG 01] en terme d'extraction de connaissances. Le terme data mining est souvent employé pour désigner l'ensemble des outils permettant d'accéder aux données d'une base. Nous le restreindrons dans le cadre de notre recherche aux outils ayant pour objet d'extraire des connaissances à partir d'une base de données à référence spatiale. Ainsi nous entendons par "fouille de données" (data mining) l'extraction des connaissances implicites, des relations spatiales ou d'autres propriétés non explicitement stockées dans la base de données (données spatiales et thématiques).

#### **4.1.2. Méthodologie d'extraction automatique des connaissances**

L'extraction automatique des connaissances relève du data mining. Comme nous l'avons explicité dans la section précédente, nous considérons le data mining (dont l'équivalent en français est la fouille de données) comme l'ensemble des techniques nous permettant d'explorer et de transformer les données géographiques et thématiques en connaissances nécessaires pour alimenter le moteur d'inférence du système.

Il s'agit en l'occurrence, à partir d'informations stockées dans la base de données, de générer des ensembles flous et de découvrir des règles traduisant le comportement des données. Dans

ce contexte, la mise en œuvre de réseaux de neurones fonctionnant en mode non supervisé, capables de constituer de façon autonome une classification complexe des données s'avère indispensable [AMA 02, *op. cité*].

En effet, la conception d'un système expert flou pour l'aide à la décision est complexe et difficile à développer selon les méthodes classiques (difficultés inhérentes à la disponibilité permanente d'une équipe d'experts) d'ingénierie et d'acquisition des connaissances [COX 97, *op. cité*]. Le recours aux à la mise en œuvre de réseaux de neurones permet de contourner de façon très efficace cette difficulté.

#### 4.1.3. Calibrage des connaissances par apprentissage du réseau

L'apprentissage artificiel s'intéresse à l'écriture de programmes d'ordinateur capables d'améliorer automatiquement au fil du temps, soit sur la base de leur propre expérience, soit à partir de données antérieures fournies par d'autres programmes [COR 02] leur performance. Dans le contexte informatique, l'apprentissage signifie la capacité d'un logiciel à améliorer sa performance. Au sens cognitif, l'apprentissage se définit comme "la capacité à améliorer les performances au fur et à mesure de l'exercice d'une activité".

En définitive, la notion d'apprentissage (artificiel ou automatique) englobe toute méthode permettant de construire un modèle de la réalité à partir de données, soit en améliorant partiellement ou totalement le modèle, soit en créant complètement le modèle. En ce sens participe de deux tendances principales : la tendance *symbolique* (issu de l'intelligence artificielle) fondée sur l'inférence inductive et la tendance *numérique* axée sur les lois statistiques. Dans le cadre de cette recherche, le sous module d'acquisition es connaissances est un système hybride neuronal flou vérifiant les propriétés de *consistance*<sup>16</sup>, de *complétude*<sup>17</sup> [ALC 03 ; JIN 00] et de *compacité* qui traduit ici le fait que la base de règles soit de taille raisonnable (en anglais, *compactness*).

L'une des caractéristiques importantes des réseaux de neurones est leur capacité d'apprentissage. Cet apprentissage, en mode supervisé ou non, est réalisé à partir de données numériques et se concrétise, du point de vue interne, par une adaptation automatique des seuls éléments susceptibles de varier dans un réseau de neurones : les poids synaptiques dont l'ensemble constitue la mémoire du réseau.

L'apprentissage supervisé consiste à apprendre une association particulière entre deux ensembles de couples (entrées et sorties escomptées) de vecteurs à partir des données d'exemples dont on connaît, a priori, les résultats (sorties escomptées) que devrait fournir le réseau. Durant une première phase appelée phase d'apprentissage, on présente successivement

---

<sup>16</sup> La consistance signifie qu'on ne peut trouver deux règles contradictoires dans une même base de règles.

<sup>17</sup> La complétude signifie que pour chaque variable d'entrée, il doit exister dans la base au moins une règle permettant de la traiter.

au réseau des vecteurs (ensemble de données brutes) en entrée. A chaque étape où un exemple est présenté au réseau en entrée, il calcule une sortie réelle. Ce calcul est effectué de proche en proche de la couche d'entrée à la couche de sortie. C'est la propagation avant. Ensuite l'erreur (somme quadratique des erreurs sur chaque sortie réelle) est calculée par rapport à la sortie escomptée. Cette erreur est ensuite rétro-propagée dans le réseau, donnant lieu à une modification des poids. Ce processus est répété en présentant successivement chaque exemple. Si pour tous les exemples, l'erreur est inférieure à un seuil choisi, on conclut que l'apprentissage est efficace. Ainsi l'apprentissage consiste à minimiser l'erreur quadratique commise par le réseau sur l'ensemble des exemples, considérée comme une fonction des poids. Pendant une seconde phase, appelée phase de test, on présente au réseau de nouveaux vecteurs qui n'ont pas servi à l'apprentissage et l'on observe ses réponses. Les performances du réseau sont alors évaluées par le pourcentage d'erreurs, la matrice de confusion.

La règle d'apprentissage dans ce cas est celle de la rétropropagation du gradient. Pour ce faire on appliquera, comme habituellement lorsqu'il s'agit d'un réseau neuronal multicouche, l'algorithme de la rétropropagation du gradient [COR 02, *op. cité* ; DRE 02, *op. cité* ; JOD 94, *op. cité* ; ROL 00b, *op. cité* ; etc.]. Il s'agit d'opérer une descente du gradient dans l'espace des poids et cela, pour tous les poids dans chaque couche. Le terme rétropropagation indique que les corrections à opérer sont calculées à partir de la sortie du réseau et sont propagées jusqu'à la première couche cachée. Le terme de décroissance des poids signifie une élimination automatique des connexions inutiles.

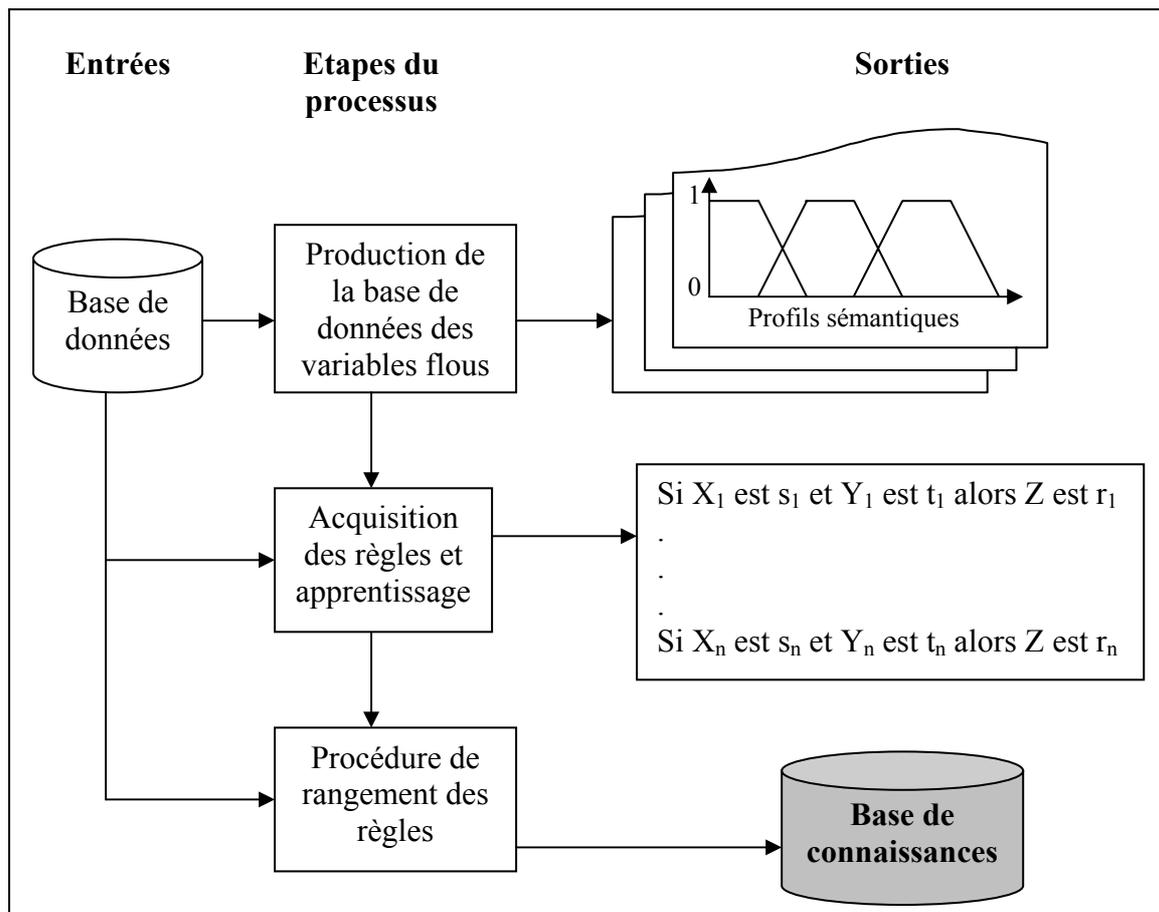
L'apprentissage non supervisé consiste à découper l'ensemble des vecteurs d'entrée en classes d'équivalence, sans qu'il soit nécessaire de donner les noms des classes pour chaque exemple. Cette séparation en classes s'opère par mesure de ressemblance entre les vecteurs proposés. La contrainte que l'on impose au réseau est nombre de classes d'équivalence. Pendant la phase d'apprentissage, le réseau construit une typologie de l'espace des vecteurs. L'intérêt de cette approche est que l'on n'a pas besoin de disposer d'exemples résolus pour servir de base d'apprentissage, c'est-à-dire des vecteurs de sortie correspondant à chaque vecteur d'entrée utilisé pour l'apprentissage.

La règle d'apprentissage non supervisé la plus connue et la plus utilisée, en l'occurrence lorsqu'on travaille sur des problématiques nécessitant l'usage de données géographiques et où l'étude de la proximité est un facteur important, est celle des cartes topologiques de Kohonen [AMA 02, *op. cité* ; DRE 02, *op. cité* ; KOH 88]. En effet le principe des cartes topologiques de Kohonen est très simple : il utilise le fait que le produit scalaire de deux vecteurs normés est une mesure de leur ressemblance. Les poids synaptiques étant initialisés au hasard, chaque neurone reçoit à l'entrée de son seuil une valeur égale au produit scalaire du vecteur d'entrée et du vecteur des poids synaptiques aboutissant sur ce neurone. Le neurone pour lequel ce produit scalaire est maximum devient prédominant (c'est celui dont les poids

synaptiques ressemblent le plus au vecteur d'entrée). En pratique la ressemblance entre deux vecteurs n'est pas nécessairement mesurée par le produit scalaire : nous pourrions utiliser un calcul de distance. Dans ce cas, les neurones ont une fonction d'activation non monotone et symétrique (la fonction gaussienne par exemple).

#### 4.1.4. Synthèse

En nous appuyant sur la figure 4.2, nous présentons ici la synthèse de la méthodologie d'extraction des connaissances.



**Figure 4.2** : *Processus d'extraction automatique des connaissances*

*Légende* :  $X_i, Y_i$  : variables linguistiques pour  $i = \{1, \dots, n\}$  ;  $Z$  : variable linguistique résultat ;  $s_i, t_i, r_n$  : termes sémantiques.

A partir des données, le mécanisme d'extraction génère les profils sémantiques nécessaires pour chaque variable, puis les règles qui décrivent leur comportement. Le processus d'extraction de règles trouve la meilleure combinaison de prémisses et de conséquents cohérente avec les données. Ce qui implique qu'il est nécessaire de vérifier et de corriger les anomalies éventuelles dans le fonctionnement du processus quant à la constitution de la base

de connaissances, détecter les variables absentes ou redondantes. Pour ce faire on devra procéder à l'apprentissage du réseau de neurones.

## 4.2. Le moteur d'inférence

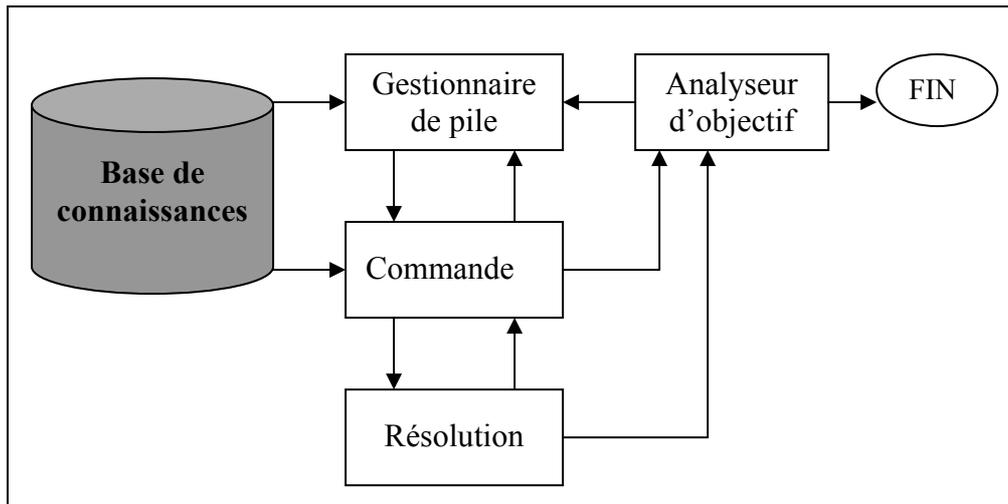
Le moteur d'inférence est chargé d'exploiter la base de connaissances de façon à mener un raisonnement sur le problème de localisation posé en fonction du contenu de la base de faits. Il contient donc un algorithme qui examine si la *prémisse* ou le *conséquent* d'une règle est vérifié par les faits contenu dans la base de faits, puis fait déclencher les règles nécessaires à la résolution du problème. Il existe deux modes principaux de déclenchement des règles. Le premier, appelé *chaînage avant*, consiste à examiner inlassablement les prémisses des règles et à appliquer celles-ci chaque que c'est possible. Le deuxième mode, appelé *chaînage arrière* consiste pour le moteur, étant donné un but qui a été fixé, à examiner les règles concluantes sur ce but, ce qui l'amène à vérifier les conséquents des règles.

Pour notre part, le moteur d'inférence sera doté d'un processus de chaînage mixte c'est-à-dire qu'il utilisera le chaînage avant et le chaînage arrière de façon complémentaire.

### 4.2.1. Le chaînage arrière

Le chaînage arrière est une technique de raisonnement automatique qui travaille en partant de l'objectif pour trouver toutes les propriétés permettant de l'atteindre. Il consiste à établir récursivement une série de sous-objectifs en travaillant à partir d'un ensemble de règles, pour chercher les valeurs concernant le sous-objectif courant (figure 4.3). Le chaînage arrière fonctionne ainsi selon une logique inductive. Pour cela, il recherche dans la base de connaissances une ou plusieurs règles qui permettraient d'atteindre ce but (c'est-à-dire les règles dont l'une des conclusions est identique à l'hypothèse fixée), puis il vérifie chacune de leurs prémisses jusqu'à ce que l'une des règles puisse être déclenchée.

*L'analyseur d'objectif* détermine l'état courant du but à atteindre. Cet objectif est choisi à partir du protocole d'exécution du système. Si nous cherchons par exemple le ou les types de service possible à implanter sur un site dont toutes les caractéristiques sont connues, ou inversement on peut être amené à identifier le site adéquat pour tel service dont tous les paramètres sont connus, cette variable sera retenue comme objectif courant (souvent appelé *sous-objectif*).



**Figure 4.3** : *Processus d'inférence en chaînage arrière* [adapté de COX 97]

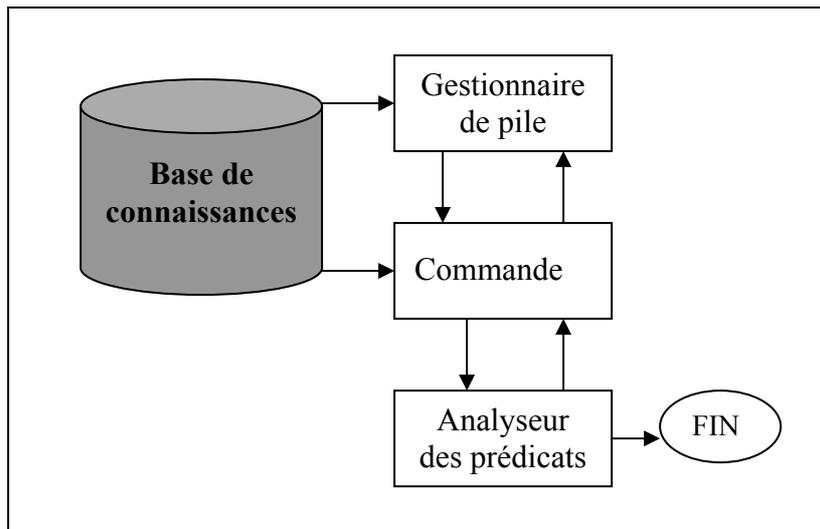
Le *gestionnaire de pile* établit la liste des règles (la pile) à déclencher pour atteindre l'objectif en fonction à partir de la base des règles et de la base des connaissances et passe la main au processeur *commande*.

Le processeur *commande* choisit et exécute les règles et met à jour la base de connaissance. A la sélection d'une règle, il fait appel au module *résolution* pour trouver les valeurs des variables non instanciées.

Le module *résolution*, se trouvant au cœur du processus, prend les variables non instanciées et recherche leur valeur. Pour cela, il définit la variable comme étant un sous-objectif et relance tout le chaînage arrière récursivement.

#### 4.2.2. Le chaînage avant

Le chaînage avant est une stratégie de raisonnement automatique qui, pilotée par les données, boucle continuellement sur une pile de règles en accumulant les résultats (figure 3.11). Au cours du processus, les règles non déclenchées au cycle  $k$  peuvent l'être au cycle  $k+1$  [COX 97, op. cité]. Le chaînage avant fonctionne selon une logique déductive. Pour cela, il compare les informations de la base des faits aux prémisses des règles de la base de connaissances afin d'en déduire de nouvelles informations en déclenchant certaines de ces règles.



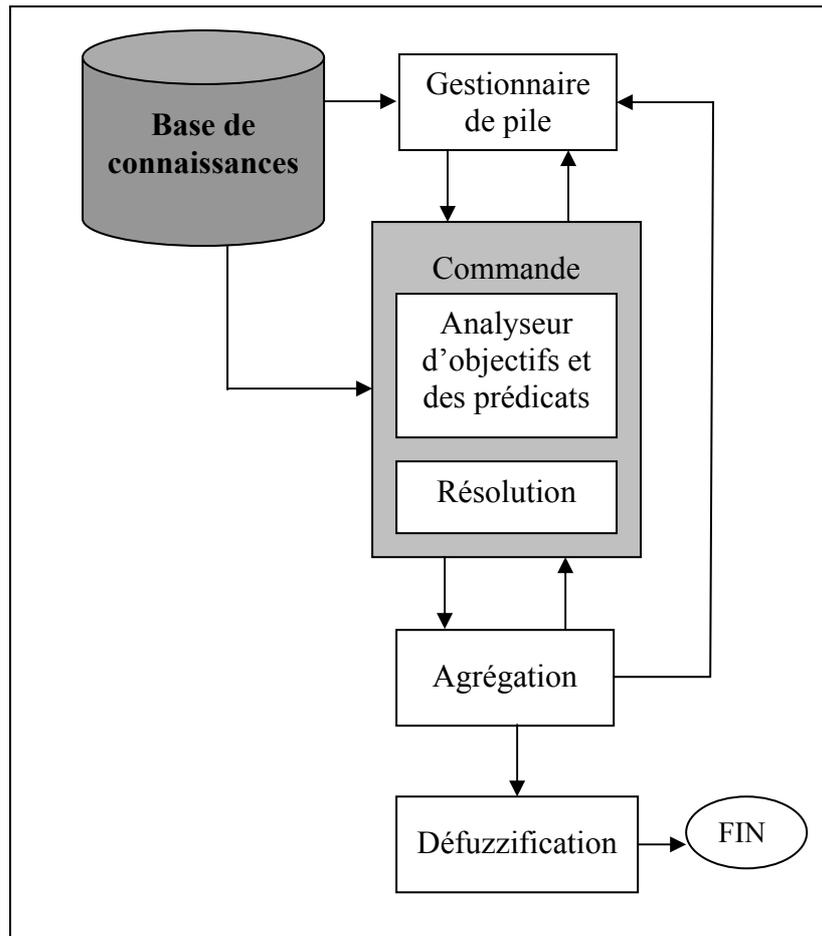
**Figure 4.4 :** *Processus d'inférence en chaînage avant* [adapté de COX 97]

Après réception des données du monde extérieur via l'interface utilisateur, le système recherche et trouve les règles qui leur correspondent. Ainsi, grâce au chaînage avant, le système collecte des informations dans son environnement et construit un espace de configuration qui corresponde à un ou plusieurs états du problème.

*L'analyseur des prédicats* examine l'état des règles et des variables de façon à en déduire la prochaine règle à exécuter. La base de règles est ainsi parcourue en boucle jusqu'à ce que plus aucune règle ne puisse être déclenchée. Le système compare les données aux prémisses de chaque règle afin de trouver celles qui sont déclenchables. Ainsi d'autres règles peuvent éventuellement être exécutées au cycle suivant. Si aucune règle n'a pu être déclenchée, le processus est terminé.

#### **4.2.3. Synthèse : le chaînage mixte**

La réalisation d'un chaînage mixte par combinaison du chaînage arrière et du chaînage avant permettra au système de posséder un moteur d'inférence véritablement opérationnel. Ceci favorisera l'exécution parallèle des règles (figure 4.5) afin d'optimiser le processus d'inférence.



**Figure 4.5 :** *Processus d'inférence en chaînage mixte*

Ici, le module *agrégation* s'occupe de la gestion et des mises en relation des solutions floues. Quand toutes les règles ont été déclenchées, le module *défuzzification* est appelé pour chaque chacune des variables solution. Il existe plusieurs méthodes de défuzzification (recherche du centroïde, du maximum, de la moyenne), mais l'idée générale est de trouver une valeur scalaire qui représente au mieux la variable solution. Dans tous les cas de figure, un indice d'évaluation du degré de validité de la solution, appelé *indice de compatibilité* [COX 97, *op. cité* ; FOU 03 ; MAR 03 ; ROL 00b, *op. cité*] est également calculé.

#### **4.3. La simulation : construction et évaluation des scénarii de localisation**

L'objectif ici est de doter le SIAL d'un simulateur capable de transformer les résultats numérique et/ou symbolique fournis par le moteur d'inférence en informations visuelles sous formes cartographique, statistique, etc.

Concrètement, le processus de simulation procédera en deux étapes : la construction du modèle et la génération des scénarii.

### 4.3.1. La construction des modèles de simulation

Au plan informatique, l'implémentation du système étant axée sur la programmation orientée objet pour des raisons de souplesse et d'efficacité [CLA 00 ; GAO 04 ; ROL 02, op. cité ; YEH 04], la modélisation des divers objets (fonctions, paramètres, variables, attributs, etc.) des modèles (traduits sous forme algorithmique) implantés dans le système, sera dirigée par la méthode de formalisme UML (Unified Modeling Language) [CLA 00, op. cité ; DUC 98 ; GRA 97, op. cité ; MUL 00, op. cité]. En conséquence, il est essentiel d'organiser les différentes composantes de chaque modèle en sous modèles conformément aux principes de la programmation objet avec la méthode UML. L'objectif ici est de mettre en exergue, à partir de la représentation conceptuelle d'un modèle, la logique séquentielle d'exécution du programme lors de la résolution des différentes tâches de traitement et les niveaux de dialogue (échelles d'articulation) entre les sous modèles. Nous distinguons ainsi, selon une hiérarchie descendante, quatre grandes échelles d'articulation (figure 4.6) à savoir le sous modèle *type*, le sous modèle *instanciation*, le sous modèle *composante* et le sous modèle *élémentaire*.

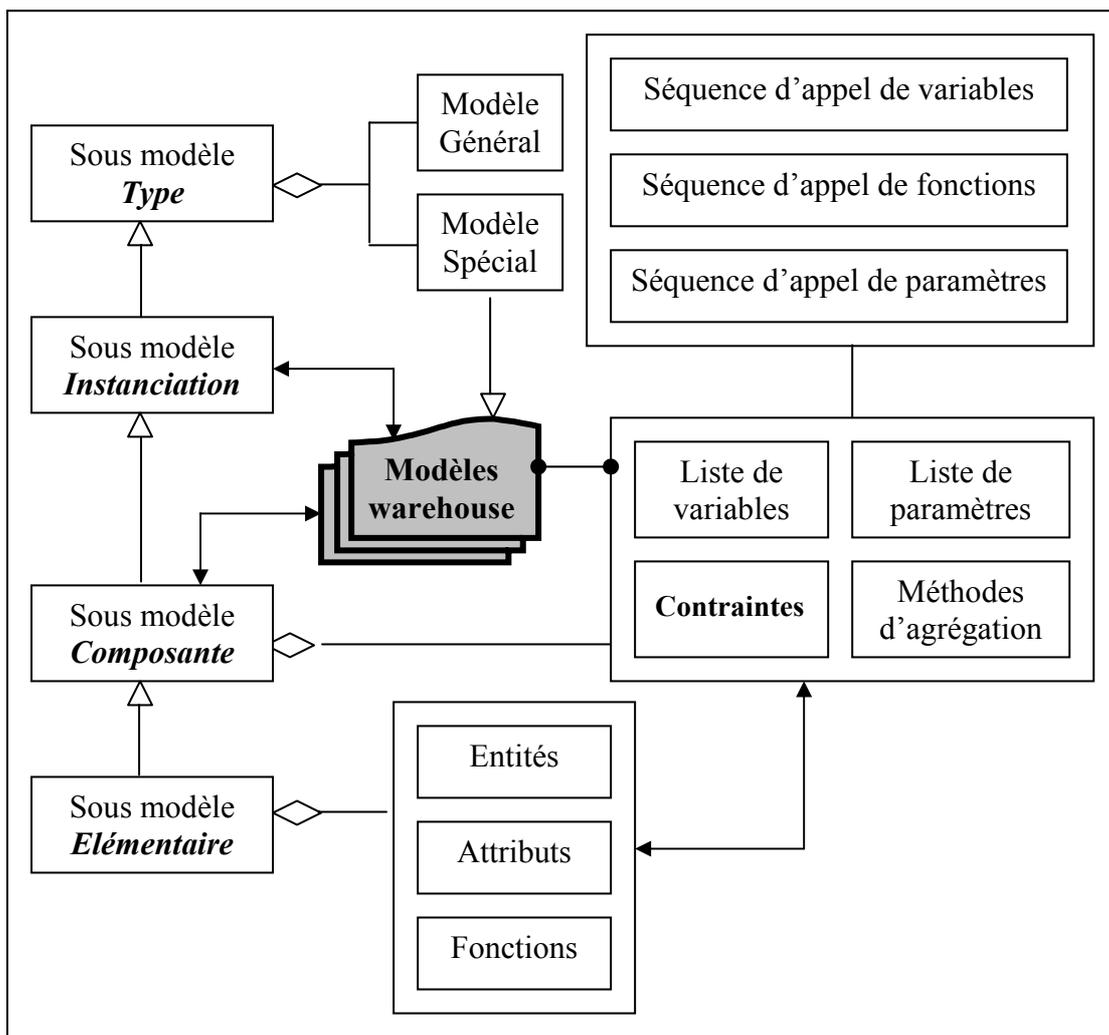


Figure 4.6 : Processus de construction d'un modèle de simulation

Le sous modèle *Type* représente l'architecture générale du modèle avec toutes ses capacités. Nous la divisons en deux formes abstraites : le modèle proprement dit (modèle général) et les spécificités (modèle spécial). Le modèle général définit la forme structurelle et les activités de modélisation. Quant aux spécificités, elles regroupent les applications possibles du modèle telles la prévision de population, la mesure de l'externalité, etc.

Le sous modèle *Instanciation* regroupe les blocs de fonctions mathématiques composant le modèle général. Ces blocs de fonctions instanciés, qui peuvent être complétés par d'autres via une procédure d'appel en direction de la base de modèles (model warehouse), sont utilisés pour analyser chaque aspect du problème auquel correspond le modèle. Par exemple dans le cas du modèle d'interaction territoriale, on aura un bloc de fonction analysant l'accessibilité, un autre pour les flux d'émissivité, un autre pour les flux d'attractivité et un dernier pour l'estimation du paramètre flou.

Le sous modèle *Composante* a pour objectif de regrouper les fonctions mathématiques élémentaires en unités fonctionnelles par combinaison avec les paramètres et les variables par de simples méthodes d'agrégation. Ce sont ces unités fonctionnelles qui composent les blocs de fonctions instanciés.

Le sous modèle *Elémentaire* inclut tous les objets élémentaires tels les variables, les constantes, les paramètres, etc. Ce sont les objets basiques nécessaires à la construction du modèle. Selon l'approche orientée objet, on peut les scinder en trois classes : entités, attributs et fonctions.

#### **4.3.2. Le générateur de scénarii**

Comme l'indique son nom, le générateur de scénarii est le sous module qui s'occupe des procédures de simulation. Grâce à ses différentes composantes, il permet de synthétiser et d'agréger les résultats spatiaux et thématiques en une couche d'informations cartographiques facilement visualisables en ayant recours au module de cartographie du SIG associé. Bien entendu, avant l'affichage des résultats cartographiques, il se passe en interne au système d'inférence des interactions analytiques entre le moteur d'inférence, la base de connaissances, la base de modèle. Le générateur de scénarii permettra également l'affichage des résultats statistiques sous formes graphique, tabulaire et/ou textuelle. L'intérêt du générateur de scénarii sera beaucoup plus mis en avant pendant la thèse.

Au cours de ce chapitre nous avons exposé la méthodologie complète de conception du SIAL. En définitive, il est composé de : la base de faits, la base de connaissances produite par le réseau neuronal flou, le moteur d'inférence flou fonctionnant en chaînage mixte, d'un constructeur de modèle et d'un générateur de scénarii.

## CHAPITRE V

# PROPOSITION DE L'ARCHITECTURE DU SYSTEME D'AIDE A LA DECISION EN LOCALISATION SPATIALE : SADELOS

Nous avons insisté plus haut que, si la plupart des chercheurs, dont les activités ont trait à la géomatique [BOR 02, *op. cité* ; GAO 04, *op. cité* ; LAA 00, *op. cité* ; LAU 93, *op. cité* ; MAI 02, *op. cité* ; ROL 02, *op. cité* ; etc.] s'accordent sur la puissance et l'excellence des SIG en tant qu'outils de gestion des données à référence spatiale, ils reconnaissent aussi que la pérennité de cet outil est tributaire d'un autre aspect fondamental : des capacités *intelligentes* d'inférence décisionnelle.

Rappelons aussi que, pour pallier ces insuffisances et contribuer activement à l'avancée de la recherche en ce domaine, notre choix s'est porté sur le couplage entre le SIG et le SIAL (Système d'Inférence pour l'Aide en Localisation). On propose de ce fait, dans ce dernier chapitre de présenter succinctement le Système d'Aide à la DECision en Localisation Spatiale (SADELOS) obtenu par encapsulation du SIAL.

Après la représentation logique du SADELOS, nous évoquerons son architecture ; ce qui nous permettra de mettre en évidence en troisième paragraphe, les flux qui parcourent le système : les flux de communication entre l'utilisateur et le SADELOS, les flux internes au SADELOS et les flux entre le SADELOS et l'expert (flux de mis à jour).

### 5.1. Logique de conception du SADELOS

D'après la revue de la littérature spécialisée [BIR 00, *op. cité* ; BUR 98 ; COR 93, *op. cité* ; LON 02, *op. cité*], on peut distinguer trois types d'intégration possibles des fonctionnalités analytiques et décisionnelles complémentaires à un SIG :

- l'ajout de certaines capacités analytiques directement au sein du SIG (*built in method*), à partir d'une mise à jour du logiciel de base. Ce type d'intégration a évolué depuis les années 1980, seule la mise à jour du logiciel de base ne suffit plus ;

- la combinaison de modules d'analyse avec les SIG : un module d'analyse indépendant (modules de simulation, d'optimisation, statistiques, etc.) s'appuie sur les potentialités de base du SIG pour la saisie des données d'entrée et sur certaines de ses fonctions pour l'affichage des résultats. Cette situation est encore à ses débuts et commence à connaître, toutefois, un certain progrès [LAA 00, op. cité] ;
- une intégration effective où les données peuvent naviguer entre les modules d'analyse spatiale ou d'aide à la décision et le SIG sans perte ou altération de qualités et/ou caractéristiques telles la topologie, l'identité des objets, les métadonnées ou les différentes relations. Ce type idéal de couplage n'existe pas encore à l'étape opérationnelle mais nombre de recherches [ARE 03 ; ARO 99 ; AUF 04, op. cité ; BIR 00, op. cité ; GAO 04, op. cité ; LAA 00, op. cité] comme celle qui nous occupe, en font déjà une préoccupation majeure.

Il résulte de ce constat que nous optons, dans le cadre de notre recherche, pour le dernier type d'intégration.

Ainsi, après le couplage du SIG et le SIAL par encapsulation, on aura plus qu'un seul logiciel avec son interface utilisateur unifié (interface offrant la possibilité de manipuler des données floues) et sa base de données ; celle-ci contient à la fois les données spatiales et thématiques et les méthodes d'agrégation neuronale floue (qui peuvent être gérées de façon interactive au sein du SGBD). Le système d'inférence localisé dans le SIAL sera programmé et compilé en langage C++ pour des raisons de performance et de souplesse puis intégré au SIG via des macrocommandes qui permettront également de relier tous les sous modules entre eux au sein de l'architecture générale du SADELLOS.

## 5.2. Architecture générale du SADELLOS

L'objectif de cette section est de présenter les diverses composantes du système d'aide en localisation spatiale et leur interconnexion.

Toujours guidé par l'approche objet, le système sera conçu dans un esprit de *justesse* (capacité du système à produire les résultats conformes à ses spécifications), *réutilisabilité* (possibilité de récupérer des parties du système pour de nouvelles applications) et d'*extensibilité* (facilité de modification du système pour d'éventuels élargissements de la problématique de base). Pour ce faire, il sera structuré en modules fonctionnels (*objet*). Ainsi, le système est considéré comme un ensemble d'objets "distincts", constitués eux-mêmes d'ensembles de structures de données et de routines (procédures associées aux données) intégrées (figure 5.1).

Structuré de façon flexible, le Système d'Aide à la DEcision en LOralisation Spatiale (ADELOS), comporte cinq composantes majeures : les interfaces (interface expert et interface utilisateur unifié), le "noyau" du SIG (les fonctionnalités de base du SIG), le SIAL (Système d'Inférence pour l'aide en Localisation), le système de gestion de la base de données (SGBD) et le méta-module de commande et de contrôle des interconnexions.

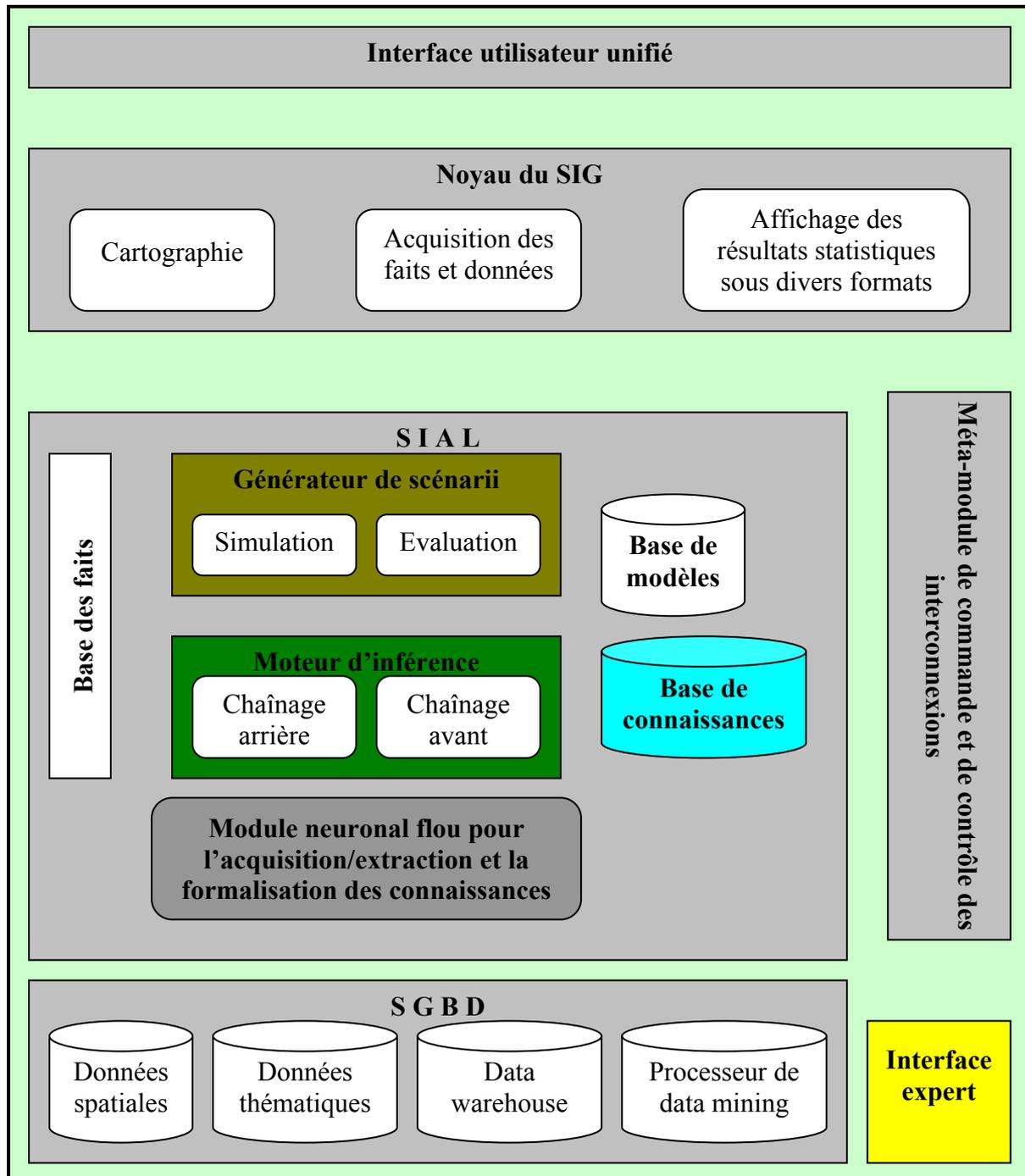


Figure 5.1 : Architecture générale du SADELOS

L'utilisateur se met en interactivité avec l'ensemble du système depuis l'interface qui est le module qui assure le lien entre le système et l'extérieur.

Les différentes manipulations sont opérées grâce au "noyau" du SIG, qui permet également l'acquisition de données externes, la cartographie et la visualisation des scénarii ainsi que l'affichage des résultats statistiques sous forme numérique, tabulaire et/ou graphique.

Pour accomplir sa tâche, le système d'inférence, par le biais du moteur d'inférence, exploite à la fois les informations fournies par l'utilisateur, rangées dans la base de faits, et les connaissances rangées dans la base correspondante, au moyen du chaînage avant et/ou arrière. Ensuite, générateur de scénarii analyse les objectifs de localisation et/ou d'affectation à atteindre et grâce au gestionnaire de dialogue (module de contrôle des interconnexion), il extrait les règles de décision à partir de la table de règles et les données du SGBD qui alimentent à leur tour les modèles nécessaires à la construction des scénarii.

Le tableau ci-après donne une vision synoptique des diverses composantes du SADELOS avec leur définition, leurs organes (sous composantes), leur fonction et les entrées/sorties.

<b>Composantes du SADELOS</b>	<b>Définition</b>	<b>Sous composantes</b>	<b>Fonction/Rôle</b>	<b>Entrées</b>	<b>Sorties</b>
<b>Interface utilisateur</b>	Partie graphique du système, il permet d'accéder aux différentes fonctionnalités destinées aux utilisateurs.	Il est composé de "menus", d'icônes et de zones de texte.	Assure la liaison entre le SADELOS et l'extérieur	Données et faits.	Cartes, résultats statistiques, résultats de simulation et de requêtes.
<b>Interface expert</b>	Partie non visible du système, il est destiné à l'administrateur du système.	Commandes et menus d'accès aux base de connaissances, base de modèles et au module neuronal.	Assure le lien entre le SADELOS et l'expert pour la mise à jour du système.	Modèles, connaissances sous forme d'algorithme compilé.	Connaissances et/ou modèles obsolètes.
<b>Noyau SIG</b>	Le maillon contenant les fonctionnalités de base du SIG.	Les modules de cartographie, d'acquisition des faits et données.	Transmet les données au SGBD, les faits au SIAL via la base de faits,	Données et faits transmis par l'utilisateur.	Résultats cartographiques et statistiques.

			enfin, il assure et gère l'affichage des divers résultats.		
<b>SIAL</b>	Système d'Inférence pour l'Aide en Localisation.	Base de faits, moteur d'inférence, base de modèles, base de connaissances, module neuronal d'acquisition des connaissances et le générateur de scénarii.	C'est le cœur du SADELOS. Il assure et gère le processus d'inférence permettant de produire l'aide à la décision. Il est aussi le garant de la construction et de l'évaluation des scénarii de localisation.	Faits et connaissances.	Scénarii de localisation.
<b>SGBD</b>	Comme l'indique son nom, il est le gestionnaire de la base de données.	Données et programmes de gestion de données.	Assure le stockage, la manipulation et la transmission de données au module d'extraction de connaissances.	Données et signaux de commande.	Données.
<b>Méta-module de commande et de contrôle</b>	Il représente les "système nerveux" du SADELOS	Petit logiciel de commande et de contrôle intermodulaire.	Envoie les signaux de commande entre les modules et assure le contrôle des flux de communication en temps réels.	signaux	Signaux.

**Tableau 5.1** : *Présentation synoptique des diverses composantes du SADELOS*

Dans l'architecture du SADELOS ainsi présentée, on ne perçoit pas de façon explicite les différents flux de communications qui susceptibles d'intervenir pendant le fonctionnement du SADELOS car ils sont faciles gérés grâce au module de contrôle des interconnexions. Dans la section qui suit, nous présentons ces flux de façon plus explicite sans faire intervenir le module de contrôle des interconnexions.

### 5.3. Les flux de communication

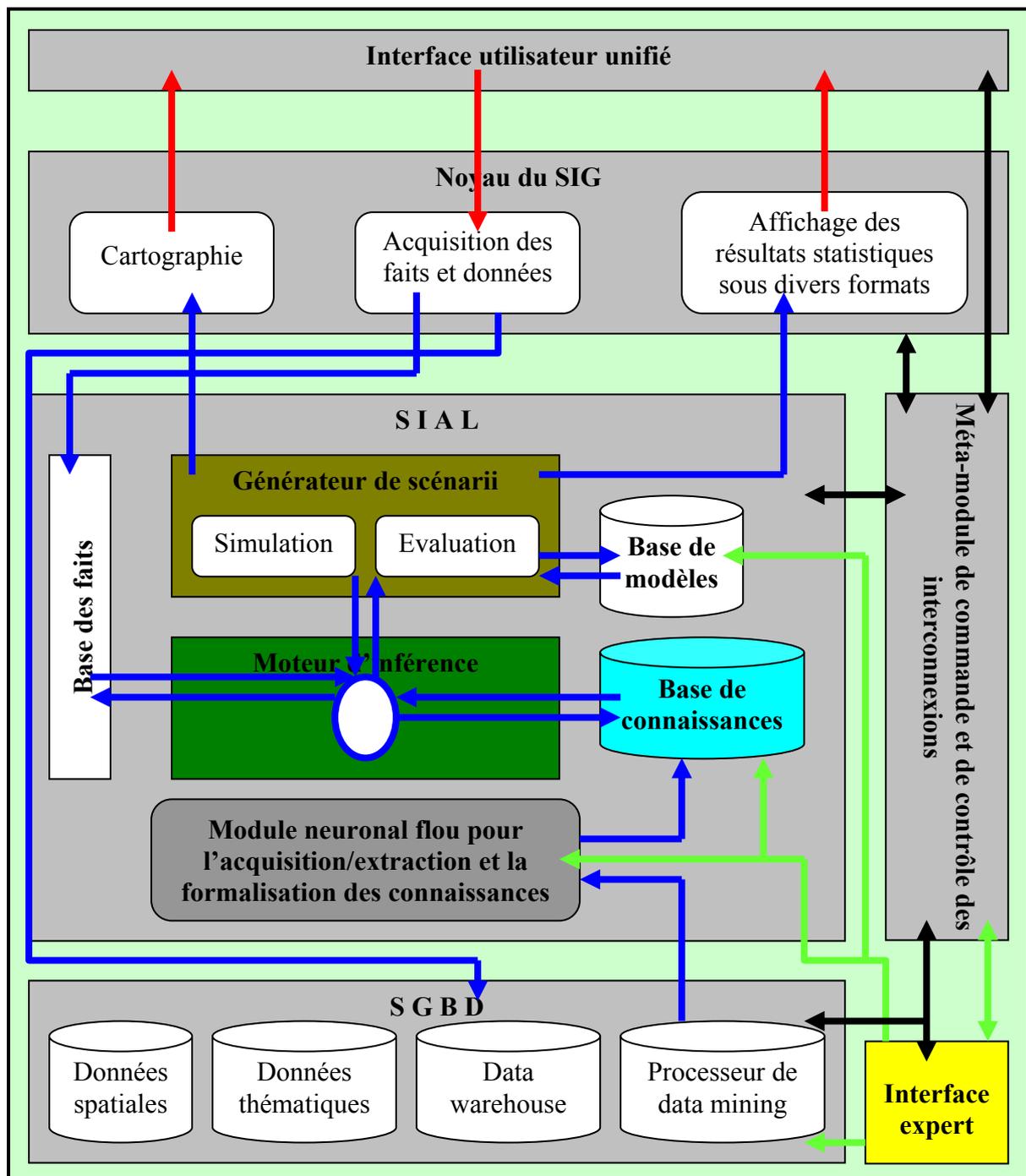


Figure 5.2 : Schéma des flux du SADELOS

Nous distinguons quatre sortes de flux (figure 5.2) :

- les flux d'information, représentés en rouge (figure 5.3) ;
- les flux de communication interne, représentés en bleu (figure 5.4) ;
- les flux de mise à jour, représentés en vert (figure 5.5);
- les flux de commande et de contrôle en noir (figure 5.6).

Chaque sorte de flux est mise en évidence sur un schéma distinct.

#### ♣ *Les flux d'information*

Il s'agit des flux d'entrées et de sortie entre l'utilisateur du système et le SADELOS lui-même. A partir de l'interface utilisateur unifié, l'utilisateur a la possibilité en entrée, de formuler les faits (qui sont stockés dans la base de faits) relatifs au problème auquel il souhaite obtenir des éléments des réponses, de fournir des données (qui seront stockées dans le SGBD) de toutes natures au système via les périphériques adéquats ou encore d'interroger le système sur l'état courant de la résolution d'un problème. En sortie le SADELOS fournit des résultats intermédiaires ou non sous formes cartographique, statistique, graphique, tabulaire et/ou numérique à l'utilisateur via le noyau du SIG.

#### ♣ *Les flux internes*

Il s'agit de la circulation des informations et des données à l'intérieur du SADELOS ; ces flux ne sont pas perceptibles par l'utilisateur du système. Une fois les données stockées dans le SGBD et les faits dans la base de faits, puis l'interrogation lancée par l'utilisateur, le moteur d'inférence est activé. Deux étapes se succèdent :

- en fonction des faits et du contenu de la base de connaissances, le moteur d'inférence produit l'aide à la décision ;
- la deuxième étape active le générateur de scénarii qui à son tour exploite les modèles pour construire et évaluer les scénarii. Les différentes solutions retenues pour un problème donné sont ensuite renvoyées au noyau du SIG pour être cartographiées et visualisées par l'utilisateur sous divers formats.

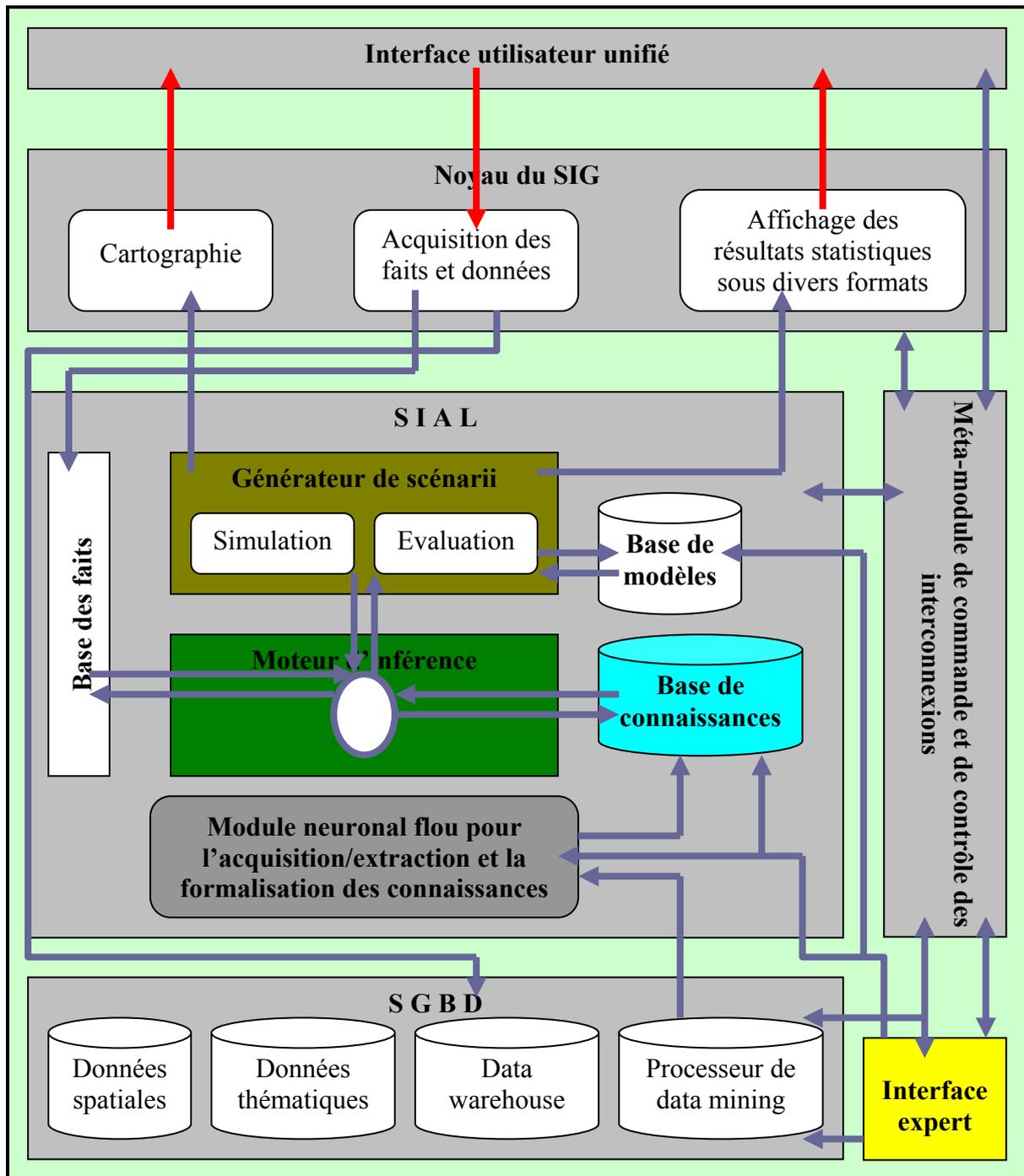


Figure 5.3 : Schéma des flux d'information entre le SADELOS et l'extérieur

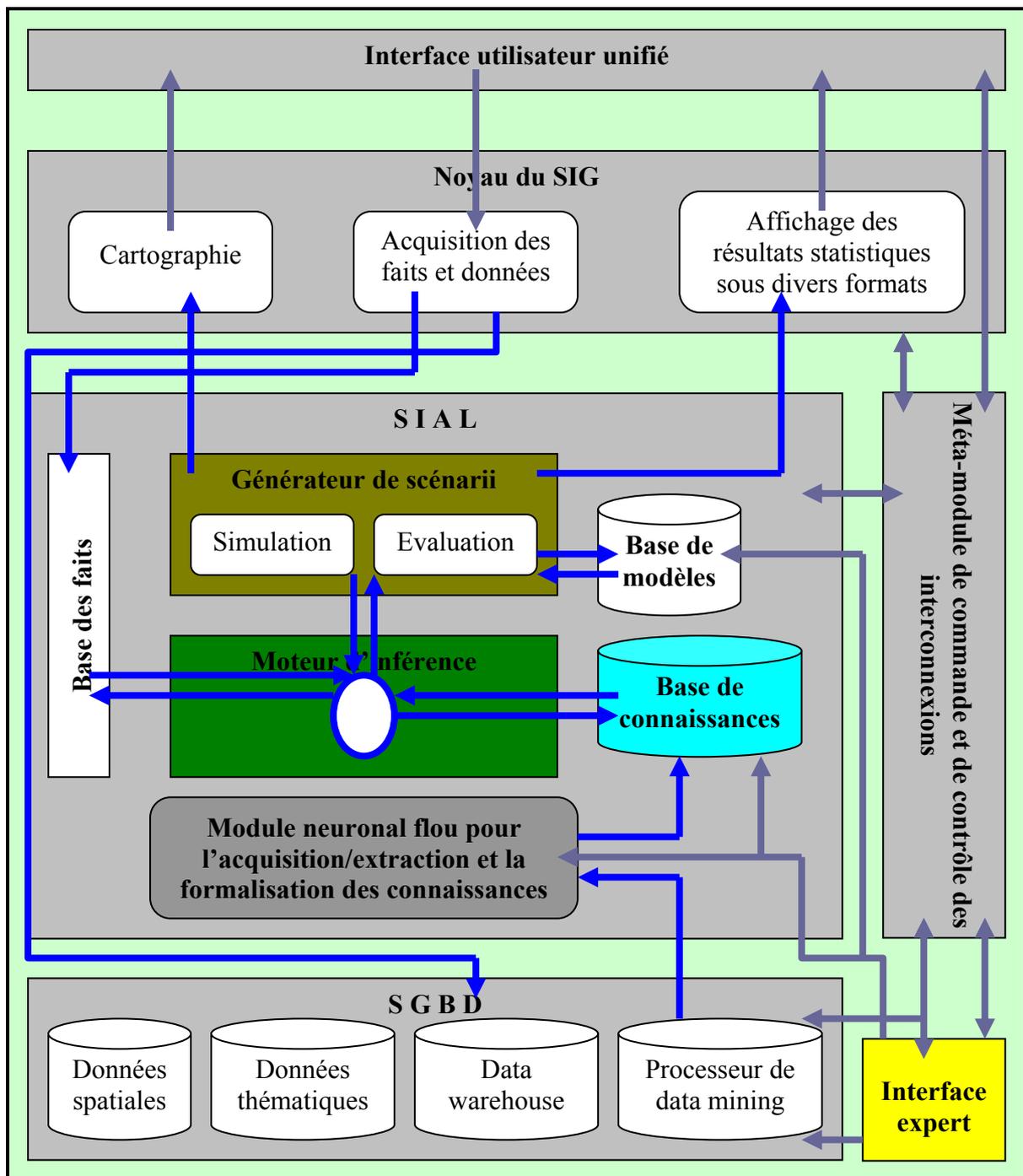
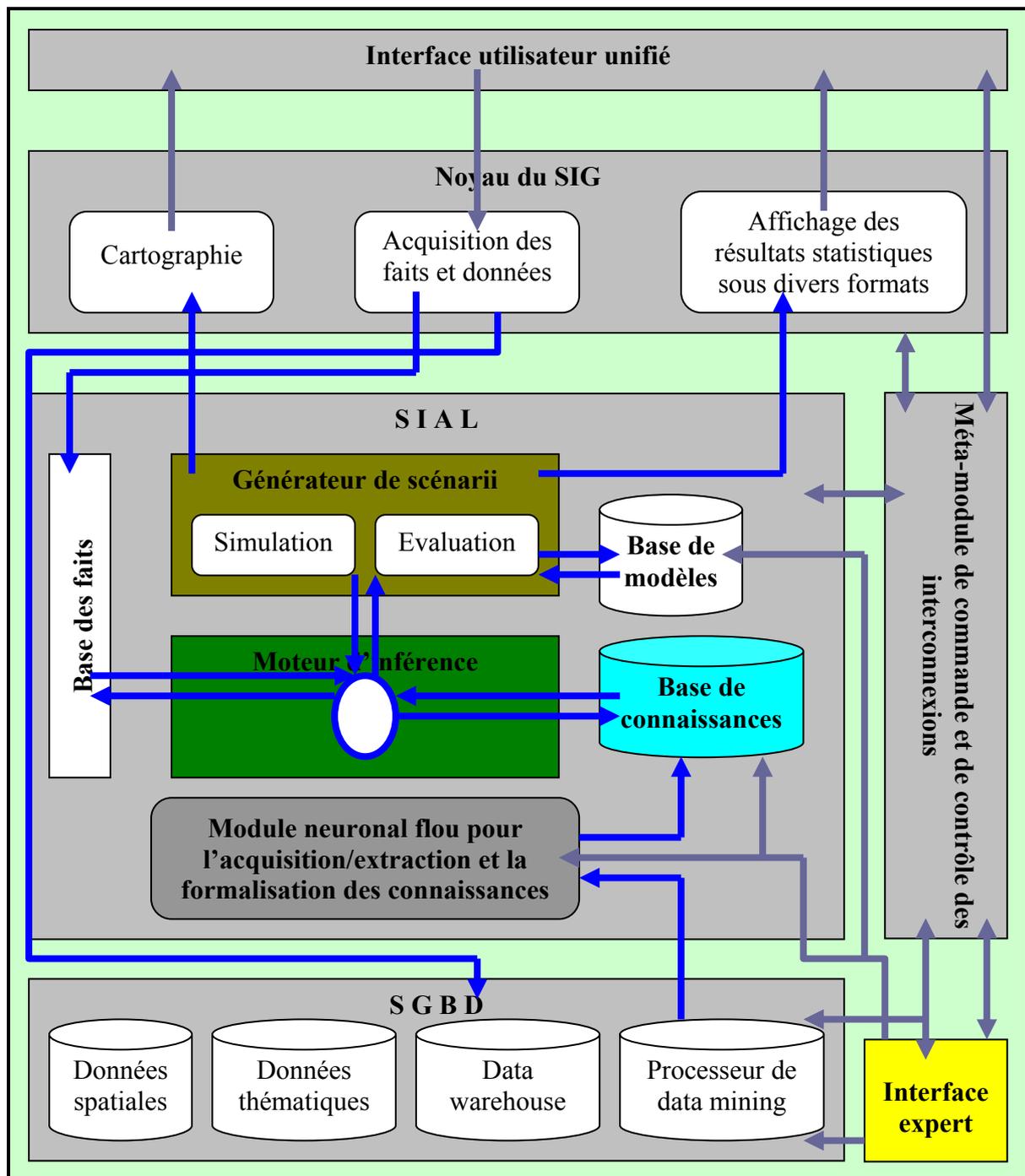


Figure 5.4 : Schéma des flux internes au SADEL0S

♣ *Les flux de mise à jour*



**Figure 5.5** : Schéma des flux de mise à jour du SADELOS

Le SADELOS étant composé d'un module neuronal flou, d'un système expert et d'un SIG, il convient de faire accroître ses capacités (inférence, stockage de données et extraction des connaissances) en rapport avec l'amélioration de l'état des connaissances aux plans scientifique et technique. Cette mise à jour permettra au SADELOS d'être toujours opérationnel même en cas d'éventuels élargissements des problématiques de base pour

lesquelles il est conçu à l'origine. Pour ce faire, l'expert a la possibilité d'intervenir par l'interface expert, directement sur le module neuronal donc la base de connaissances et sur les modèles afin d'opérer les mise à jour requises.

Trois types d'interventions sont possibles (cf. tableau 5.2).

Niveau d'intervention	Action	Type d'intervention	Module concerné	Contraintes
<b>Données</b>	Action 1	Ajout, suppression et/ou mise à jour des données.	SGBD	Vérification des cohérences avec le SGBD
<b>Règles</b>	Action 2	Création/Modification "directe" de règles.	Base de connaissances.	Vérification des cohérences et de non contradiction.
	Action 3	Suppression de règles.	Base de connaissances.	Création d'une échelle de droit d'accès et d'intervention.
<b>Production des connaissances</b>	Action 4	Création de règles par le SIAL	Activation du modèle neuronal	- Droit d'accès et d'activation - Vérification de la cohérence et de la pertinence avec l'existant.
<b>Modèles</b>	Action 5	Ajout, suppression et mis à jour des modèles.	Base de modèles.	- Droit d'accès et de mise à jour. - Vérification de la cohérence et de la pertinence avec l'existant.

**Tableau 5.2 :** *Les différents types d'interventions –expert.*

#### ♠ *Les flux de commande et de contrôle*

Il s'agit des flux de contrôle des interconnexions entre les composantes de l'ensemble du SADELLOS. Ces flux sont gérés par le méta-modèle de commande et de contrôle.

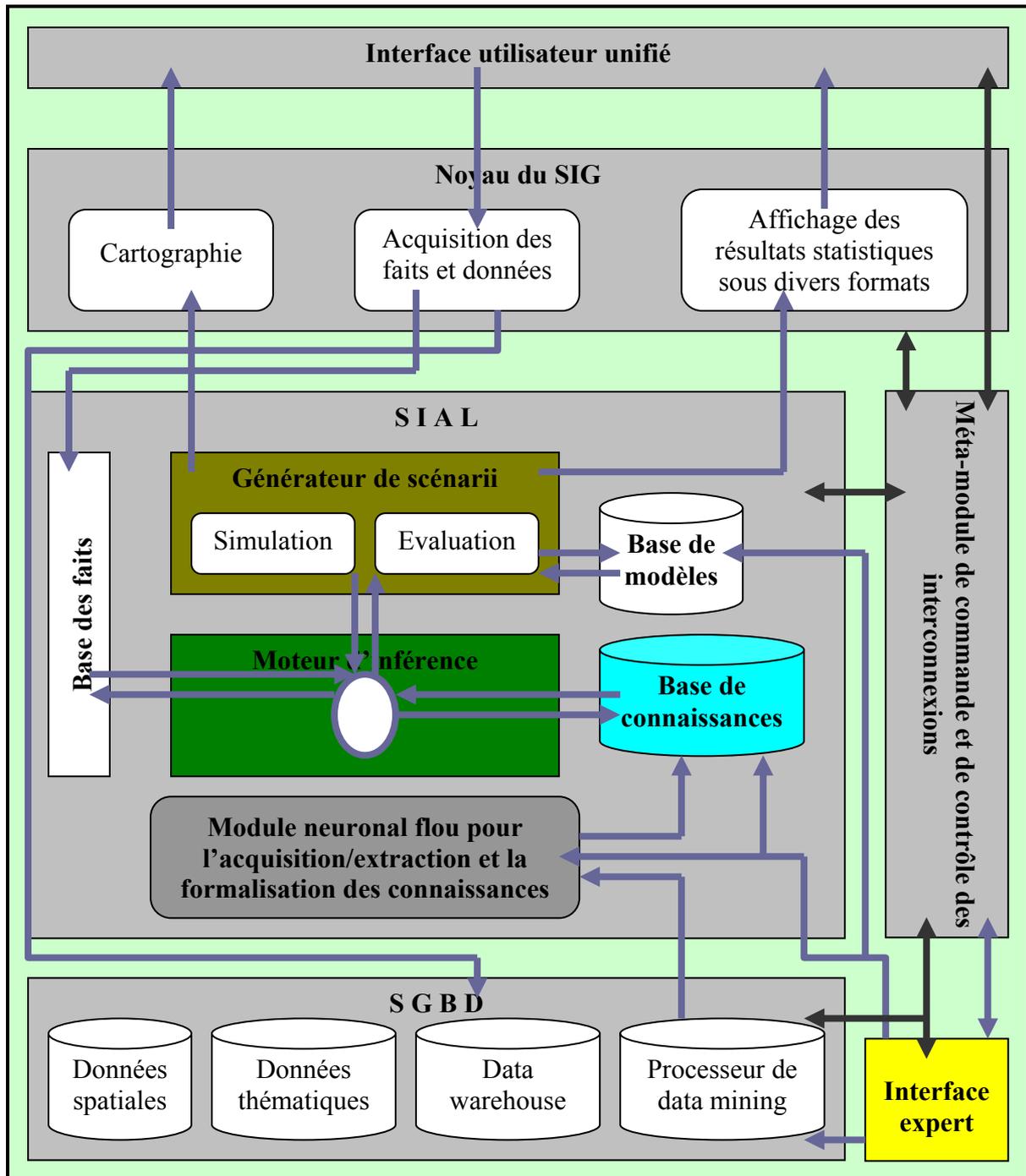


Figure 5.6 : Schéma des flux de commande et de contrôle du SADELOS

En général, l'intérêt de la Recherche/Développement dont nous amorçons ici la démarche tient donc dans sa caractéristique d'intégration. Elle se décline en cinq modalités :

- l'intégration logique : le SIAL (Système d'Inférence pour l'Aide en Localisation) est "encapsulé" dans un ensemble global : le SADELOS qui comprend l'ensemble de la chaîne de connaissances, depuis son extraction
- son utilisation et sa mise à jour ;
- l'intégration logicielle ;
- l'intégration cognitive des connaissances précises et certaines, des connaissances précises et incertaines, des connaissances imprécises certaines et les connaissances imprécises et incertaines ;
- l'intégration entre dimension scientifique et opérationnelle.

## CONCLUSIONS ET PERSPECTIVES

Notre thème de recherche se situe au cœur de trois domaines : l'aménagement du territoire, les systèmes d'informations géographiques et l'aide à la décision. Bien que cette démarche scientifique ne soit pas celle d'un spécialiste pointu, nous espérons qu'elle permettra de contribuer à l'avancée de la recherche appliquée en aménagement du territoire.

En effet, dans le contexte actuel de réorganisation territoriale où le développement et la recomposition des territoires prennent une ampleur cruciale, la nécessité de disposer d'un véritable outil d'aide à la décision en localisation spatiale devient de plus en plus impérieuse. Ce besoin est donc celui d'outils, qui soient scientifiquement fiables, adaptés à la complexité des objets territoriaux, capable de supporter et de gérer des procédures d'inférence et apte à traiter les données incertaines et imprécises ; en d'autres termes, il s'agit d'un outil qui dépasse les méthodes classiques en la matière.

L'utilisation combinée de SIG, de méthodes d'aide à la décision et la logique floue a déjà été explorée par plusieurs chercheurs. Ce triplet d'outils correspond aux deux principales phases de la décision. Le SIG permet au décideur de s'informer ; la méthode d'aide à la décision permet de synthétiser cette information pour en extraire la quintessence afin de produire les bonnes solutions qui se prêtent aux problèmes auxquels est confronté le décideur. L'originalité de recherche ne se situe donc pas, en absolu, dans la combinaison de ces outils. Elle se trouve dans le fait que l'utilisation de ces outils s'inscrit dans une démarche scientifique globale d'aide à la décision. Le SADELOS pourra ainsi satisfaire une partie des besoins de l'aménagement du territoire.

Le texte a abordé successivement le cadrage scientifique de la problématique posée, puis l'explication et la présentation conceptuelle et méthodologique du SIAL à travers la mise en application de la chaîne de cohérence épistémologique. En nous appuyant sur ces acquis, nous avons forgé la définition du SIAL et formalisé le SADELOS tout en abordant la méthodologie de sa conception. Les différentes interrelations internes au SADELOS et les interactions qu'il entretient avec le monde extérieur terminent l'exposé.

Cette recherche propose un mode d'utilisation combinée de SIG et d'un système d'inférence neuronal flou qui se tourne plus directement vers les gestionnaires du territoire. Il ne s'agit pas de concevoir seulement une interface conviviale, mais de véritablement se mettre à l'écoute, puis au service du décideur tout en restant fidèle aux exigences de la recherche scientifique. Nos travaux de recherche doivent évidemment se poursuivre dans cette direction.

## Liste des figures

Figure 1.1 : Approche systémique orientée objet du territoire

Figure 1.2 : Composantes d'un SIG

Figure 2.1 : Typologie des objets géographiques primitifs

Figure 3.1 : La chaîne de cohérence intégrée au processus décisionnel

Figure 3.2 : Triplet gagnant de l'aide à la décision

Figure 3.3.a : Chaîne d'acquisition des connaissances au sein du système d'inférence

Figure 3.3.b : Système d'Inférence pour l'aide en Localisation : SIAL

Figure 3.4 : Représentation d'un neurone artificiel

Figure 3.5 : Configuration d'un réseau neuronal à couches

Figure 3.6 : formalisation floue d'un réseau neuronal à couches

Figure 3.7 : Chaîne de traitement flou des données

Figure 4.1 : Différents niveaux de méta-données

Figure 4.2 : Processus d'extraction automatique des connaissances

Figure 4.3 : Processus d'inférence en chaînage arrière

Figure 4.4 : Processus d'inférence en chaînage avant

Figure 4.5 : Processus d'inférence en chaînage mixte

Figure 4.6 : Processus de construction d'un modèle de simulation

Figure 5.1 : Architecture générale du SADELOS

Figure 5.2 : Schéma des flux du SADELOS

Figure 5.3 : Schéma des flux d'information entre le SADELOS et l'extérieur

Figure 5.4 : Schéma des flux internes au SADELOS

Figure 5.5 : Schéma des flux de mise à jour du SADELOS

Figure 5.6 : Schéma des flux de commande et de contrôle du SADELOS.

## Références bibliographiques

- [ALC 03] Florence D'ALCHE-BUC "Association des systèmes d'inférence floue avec les méthodes connexionnistes et évolutionnistes" in Traitement des données complexes et commande en logique floue, Hermès-Lavoisier, 2003.
- [ARE 03] Arena (2003). Arena software. URL : <http://arenasimulation.com>
- [ARO 99] Stan ARONOFF, "Geographic Information Systems : a management perspective", WLD Publications, Ottawa, Canada 1999, 294 p.
- [AUF 04] Marie-Aude AUFAURE, Laurent YEH, Karine ZEITOUNI ; Fouille de données spatiales ; Laboratoire d'Ingénierie des Systèmes d'Information (2004). URL : <http://lisi.insa-lyon.fr>
- [BAU 02] Didier BAUDEWYN "La localisation intra-urbaine des firmes : une double approche empirique" ; in Stratégies de localisation des entreprises commerciales et industrielles ; De nouvelles perspectives ; Edition DE Boeck Université, Bruxelles, 2002.
- [BAV 98] J-J BAVOUX "Introduction à l'analyse spatiale" ; Editions A. Collin, Collection Synthèse, Série Géographie, Paris, 1998.
- [BIR 00] Mark BIRKIN, Graham CLARKE, Alan WILSON "Intelligent GIS : location decisions and strategic planning". Geoinformation international, 2000.
- [BER 02] F. BERNARDI "Conception des bibliothèques hiérarchisées de modèles réutilisables selon une approche orientée objet" Thèse de doctorat, Université de Corse, 2002.
- [BOR 02] Patricia BORDIN, "SIG concepts, outils et données" Editions Hermès Lavoisier, Paris 2002, 260 p.
- [BOU 95] Bernadette BOUCHON-MENIER "La logique floue et ses applications". Edition Addison-Wesley, Paris, New York Amsterdam, 1995.
- [BOU 99] Bernadette BOUCHON-MENIER "La logique floue" Que sais-je, troisième édition 1999, Presses Universitaires de France.
- [BUC 99] J.J BUCKLEY, T. FEURING, "Fuzzy and neural: Interactions and Applications", Edition Physica Verlag-Springer. Studies in Soft Computing, vol. 25, Heidelberg.
- [BUR 98] Petre A. BURROUGH, Rachel A. McDONNELL "Principles of Geographical Information Systems"; Oxford, University Press, 1998.
- [CAN 93] A. CAN "Residential quality assessment : alyernative approaches using GIS" in Fischer M.M., Nijkamp P. (eds), Geographic Information Systems, Spatial Modelling, and Policy Evaluation, 1993
- [CHE 94] J-J CHEVALIER "De l'information à l'action : vers des systèmes d'aide à la décision à référence spatiale (SADRS)" EGIS/MARI'94, Paris 1994.
- [CLA 00] Gilles CLAVEL, Nicolas FAGART, David GRENET, Jorge MIGUEIS "C ++ La Synthèse" : Concepts objet, standard ISO et modélisation UML ; Editions DUNOD, 2000.
- [COB 94] Claude COBICHON "L'information géographique : nouvelles technologies, nouvelles pratiques", Hermès, Paris 1994, 122 p.
- [COR 93] S. CORRADUCCI, PROMARC : prototype d'intégration de PROMETHE II dans Arc/Info, Travail pratique de Diplôme d'ingénieur, école Polytechnique Fédérale de Lausanne, Lausanne, Suisse, 1993.
- [COR 02] Antoine CORNUEJOLS, Laurent MICLET, Yves KODRATOFF "Apprentissage artificiel : Concepts et algorithmes" Editions Eyrolles 2002.
- [COX 97] Earl D. COX, "La logique floue" pour les affaires et l'industrie ; traduction de Maurice Clerc ; International Thomson Publishing France, Paris 1997, 572 p.
- [DAS 95] M. DASKIN, "Network and discrete Location, Models, algorithms and Applications", Wiley Interscience, New York, 1995
- [DRE 02] G. DREYFUS et al "Réseaux de neurones : Méthodologie et application" ; Editions Eyrolles 2002, 2<sup>ème</sup> tirage, 386 p.
- [DUC 98] B. M. DUC "Conception et modélisation objet des systèmes temps réel" Editions Eyrolles, 1998.
- [EAS 93] R. J. EASTMAN et al, "A procedure for Multiple-objective Decision Makin in GIS under conditions of conflicting Objectives", in Harts J. et al. (eds) EGIS'93, Genoa-Italy, vol 1(2), 1993.

- [FOU 03] Laurent FOULLOY, Sylvie GALICHET, Gilles MAURIS “Utilisation des techniques floues pour la mesure et la commande” in *Traitement des données complexes et commande en logique floue*, Hermès-Lavoisier, 2003.
- [GAO 04] Shan GAO, John PAYNTER, David SUNDARAM, “Flexible support for spatial decision-making” in *Proceedings of the 37<sup>th</sup> Hawaii International Conference on System Sciences*, 2004. URL: [www.jadeworld.com](http://www.jadeworld.com)
- [GRA 97] Ian GRAHAM “Méthodes orientées objet” traduction de Gilles Mathieu ; 2<sup>ème</sup> Edition International Thomson Publishing France, Paris 1997, 482 p.
- [GRA 01] Claude GRASLAND “L’analyse des discontinuités territoriales : l’exemple de la structure par âge dans les régions européennes”, *L’espace géographique* 4, 2001.
- [GRA 01] Claude GRASLAND “Analyse des phénomènes sociaux”, Université de Paris VII/ UMR Géographie cités ; Cours de modélisation ; Octobre 2001.
- [HANS 96] P. HANSEN, M. LABE, D. PEETERS, J-F. THISSE, “Facility Location Analysis” dans *Encyclopedia of Economics. Regional and Urban Economics*, Harwood Academic Publishers, Amsterdam, 1996.
- [INM 94] W. H INMON, Richard D. HACKATORN “Using the Data Warehouse”; Wiley-QED Publication, 1994.
- [HIL 00] D. HILL et al, “Multimodeling and object oriented design patterns, application to bio-control simulation”. Actes de conference Artificial Intelligence Simulation (AIS 2000) Towards Collaborative/distributed Modeling and simulation Environments, 2000.
- [HER 00] L. HERAULT “Rescaled Simulated Annealing-Acceleratin convergence of Simulated Annealing by rescaling the states energies”, *Journal of Heuristics*, p 215-252, vol. 6, Kluwer Academic Publishers, 2000.
- [HOR 03] Ivor HORTON “Microsoft Visual C++”, Solutions développeurs, Editions Eyrolles, 2003.
- [JEA 95] N. JEANDROZ “Conception d’un SIRS combiné à un modèle hydraulique pour la simulation des inondations” Rapport de travail pratique de diplôme, Institut d’Aménagement des Terres et des Eaux, Département de Génie rural, EPFL, 1015 Lausanne Suisse, 1995.
- [JIN 00] Y. JIN ; “Fuzzy Modeling of High Dimensional Systems : Complexity reduction and Interpretability Improvement”, *IEEE Trans. Neural Networks*, vol 8, n°2, avril 2000.
- [JOD 94] Jean-François JODOUIN, “Les réseaux de neurones : principes et définitions” ; Hermès 1994, 124 p.
- [JOE 95] F. JOERIN “Méthode multicritère d’aide à la décision et SIG pour la recherché d’un site”, *Revue internationale de géomatique*, 1995.
- [KEL 93] C.P. KELLER, J. D. STRAPP, “Design of Multi Criteria Decision Support System for Land Reform Using GIS and API”, *Proceeding of the second International Conference on Integrating Geographic Information System and Environmental Modelling*, Breckenridge, Colorado, September, 1993..
- [KLI 95] George J. KLIR, Bo YVAN, “Fuzzy sets and fuzzy logic, theory and applications”, 1995.
- [LAA 00] Amor LAARIBI “SIG et analyse multicritère”, Edition Hermès, Paris 2000.
- [LAB 95] M. LABE, D. PEETERS, J-F. THISSE, “Location on Network ” dans *Handbook of Operations Research and Management Science: Networks*, North-Holland, Amsterdam, 1995.
- [LAN 97] André LANGLOIS, Michel PHIPPS “Automates cellulaires, application à la simulation urbaine” ; Editions Hermès 1997, 197 p.
- [LAU 93] Robert LAURINI, Françoise MILLERET-RAFFORT, “Les bases de données en géomatique”, *Traité des nouvelles technologiques, série géomatique*, Hermès 1993, 340 p.
- [LEV 89] P. LEVINE, J-C. POMEROL ; *Systèmes interactifs d’aide à la décision et systèmes experts* ; Edition Hermès, Paris, 1989.
- [LON 02] A. Paul LONGLEY et al “Geographic Information Systems and Science”, Edition John Wiley & Sons, Ltd, October 2002.
- [MAI 02] Michel MAINGUENAUD “Langages pour les SIG” Conception, développement et IHM, Edition Hermès, 2002.

- [MAT 03] Philippe MATHIS “Graphes et réseaux” : modélisation mult niveau, Edition Hermès, 2003.
- [MAR 98] Pascal MARCHAND, “L’Analyse du Discours Assisté par Ordinateur”, Armand Colin 1998, 219 p.
- [MAR 03] Christophe MARSALA, B. BOUCHON-MEUNIER “Apprentissage et extraction de connaissances” in Traitement des données complexes et commande en logique floue, Hermès-Lavoisier, 2003.
- [MAT 02] Philippe MATHIEU, J-P MÜLLER, “Systèmes multi-agents et systèmes complexes” ingénierie, résolution de problèmes et simulation, Hermès, Lavoisier 2002, 264 p.
- [MUL 00] P-A MULLER, N. GAERTNER, Modélisation objet avec UML, 4<sup>ème</sup> tirage 2002, Editions Eyrolles, 2000.
- [OPE 97] S. OPENSHAW, C. OPENSHAW “Artificial Intelligence in Geography, 1997.
- [PAN 96] Dimos PANTAZIS, Jean-Paul DONNAY, “La conception de SIG ; méthode et formalisme” ; Collection géomatique, Hermès 1996, 344 p.
- [PEE 01] Dominique PEETERS, Isabelle THOMAS, “Localisation des services publics : de la théorie aux applications” in Modèles en analyse spatiale ; Edition Hermès, Paris, 2001.
- [PEE 02] Dominique PEETERS, Jacques-François THISSE, Isabelle THOMAS, “Modèles opérationnels de localisation des équipements collectifs” ; in Stratégies de localisation des entreprises commerciales et industrielles ; De nouvelles perspectives ; Edition DE Boeck Université, Bruxelles, 2002.
- [PRE 95] Roland PRELAZ-DROUX, “Système d’information et gestion du territoire, Approche systémique et procédure de réalisation” ; Presses Polytechniques et Universitaires Romandes 1995, 158 p.
- [PUM 97] Denise PUMAIN, Thérèse SAINT-JULIEN, L’analyse spatiale, Volume 1 : Localisation, A. Colin, Collection Cursus, Paris 1997.
- [PUM 01] Denise PUMAIN, Thérèse SAINT-JULIEN, Les interactions spatiales, Volume 2 : Interactions, A. Colin, Collection Cursus, Paris, 2001.
- [RAV 98] L. RAVENEL, La géographie du football de haut niveau en France, PUF, Paris, 1998.
- [RAV 01] V. RAVI, H-J ZIMMERMAN, “A neural network and fuzzy rule base hybrid for pattern classification”, Soft Computing, vol 5 n°2; 2001.
- [ROL 84] Christiane ROLLAND-MAY “Les espaces géographiques flous”, Thèse de doctorat d’Etat, Université de Metz, 1984
- [ROL 85] Christiane ROLLAND-MAY “Algorithm of fuzzy classification and application to fuzzy regionalization”, Sistemi Urbani, 1985.
- [ROL 91] Christiane ROLLAND-MAY, “ Fuzzyland, modèle de détermination et d’évaluation de territoire de cohérence. Application aux Pays et au département de la Moselle “, in Norois n° 181, 1991.
- [ROL 93] Christiane ROLLAND-MAY “Le concept de base déclarative de données spatialisées, cœur d’un système d’information et d’aide à la décision en géographie” in Systèmes d’information géographique et Systèmes experts. Sixième colloque européen de géographie théorique et quantitative, 1993.
- [ROL 96] Christiane ROLLAND-MAY “Méthodes de régionalisation d’un espace imprécis et incertain défini par des variables qualitatives”, Revue Géographique de l’Est, Nancy, n°3-4, 1996.
- [ROL 00a] Christiane ROLLAND-MAY Géostratégie de la recomposition de territoires. Cas particulier en espace fortement métropolisé : l’espace “médiolorrain” autour de Port-à-Mousson (Lorraine). Revue Géographique de l’Est tome XL, n°3-4, décembre 2000.
- [ROL 00b] Christiane ROLLAND-MAY, “Evaluation des territoires” concepts, modèles, méthodes. Edition Hermès, Paris 2000
- [ROL 02] Christiane ROLLAND-MAY “NEURO-RES, modèle d’évaluation de territoire et de projet” Approche géographique, logique et fonctionnelle. Revue du Centre d’Etudes géographiques de l’Université de Metz, MOSELLA tome XXVI n°3-4, juin 2002.
- [ROL 03] Christiane ROLLAND-MAY “Démarche de recherche et cadrage scientifique du système géographique”, Université Louis Pasteur de Strasbourg I ; notes de cours en DEA Systèmes Spatiaux et Environnement ; Octobre 2003.
- [ROY 85] Bernard ROY, “Méthodologie multicritère d’aide à la décision”, Economica 1985, 423 p.
- [RUF 01] S. DE RUFFRAY, C. ROLLAND-MAY. Vers une évaluation globale des espaces de marge :

l'exemple de la Moselle-Est et du Nord-Ouest du Bas-Rhin. In RITMA, Regards croisés sur les territoires de marges, Maison des sciences de l'Homme de Strasbourg, Presses Universitaires de Strasbourg, 2001, p. 207-220.

**[RUL 02]** David RULENCE, "Cohérence territoriale des stratégies spatiales des distributeurs français" ; in Stratégies de localisation des entreprises commerciales et industrielles ; De nouvelles perspectives ; Edition DE Boeck Université, Bruxelles, 2002.

**[RUT 01]** D. RUTKOWSKA, Neuro-fuzzy architecture and hybrid learning. Ed. Studies in Fuzzyness and Soft Computing. Heidelberg, 2001.

**[SAN 90]** L. SANDERS "L'analyse statistique des données en géographie", Montpellier : Alidade Reclus, 1990

**[SAI 01]** Thérèse SAINT-JULIEN "Processus de diffusion spatiale et modélisation du changement" in Modèles en analyse spatiale ; Edition Hermès, Paris, 2001.

**[SEA 89]** Michel LE SEAC'H " Développer un système expert" : Méthodes et exemples ; Editions TESTS, 1989.

**[TEG 01]** J.TEGHEM, M. PIRLOT; Métaheuristiques et outils nouveaux en recherche opérationnelle. Tome I : Méthodes. Tome II : Implémentations et Applications, Edition Hermès, 2001.