

THEMAMAP

Un outil de cartographie thématique des séries statistiques

par Gilles Domalain, par Céline Rodriguez

IRD - UMR 212 EME (Ecosystèmes Marins Exploités)
Centre de Recherche Halieutique Méditerranéenne et Tropicale (CRH)
Avenue Jean Monnet-BP 171-34203 SETE cedex-France
Gilles.Domalain@ird.fr
Celine.Rodriguez@ird.fr

par Jacques Madelaine, par Christophe Turbout

GREYC - CNRS UMR 6072
Université de Caen Basse-Normandie
BP 5186, F 14032 CAEN CEDEX
Jacques.Madelaine@unicaen.fr
Christophe.Turbout@unicaen.fr

par Carlota Estrella

Esq. Gamarra y Gral Valle s/n
Instituto del MAR del Peru
Chucuito –Callao – Peru
cestrella@imarpe.pe

Résumé

ThemaMap est un logiciel libre dédié à la cartographie thématique et spécialement conçu pour permettre à des utilisateurs néophytes le traitement de séries statistiques de données spatio-temporelles. Il intègre des outils de manipulation de données avancés qui le rendent apte à une exploitation approfondie de séries de données dans le but de produire et diffuser des cartes. Après une exploration de ses différentes fonctionnalités, nous montrerons quelques applications et cartes produites à partir de données d'enquêtes sur la pêche artisanale au Pérou.

1 Introduction

Les systèmes d'enquêtes sur les pêcheries, notamment dans le domaine du suivi des ressources marines, prennent différentes formes (suivi des débarquements, recueil d'observations en mer, campagnes de pêche expérimentales, enquêtes cadre...) qui sont, de manière récurrente, productrices de volumes importants de données souvent géolocalisées et dont l'exploitation optimale requiert l'utilisation de logiciels de type SIG. Si de nombreux produits sont déjà disponibles actuellement (ArcGIS, MapInfo, Quantum GIS, GvSIG, OrbisGIS, Geoda, SAGA GIS etc.), ceux-ci n'offrent pas toujours les outils adaptés à la manipulation de séries temporelles. De plus, les

logiciels existants, tant libres que propriétaires, nécessitent souvent un bon niveau d'expertise pour une prise en main et une pleine exploitation de leurs fonctionnalités.

Les séries de données produites par ces systèmes de collecte présentent souvent des caractéristiques similaires; ce sont en général des séries temporelles décrivant des valeurs mesurées (comptages, quantités...) selon plusieurs critères d'ordre temporel, spatial, ou même thématique (tels que les dates, les sites de débarquements ou les engins de pêche). Ces données peuvent se présenter sous forme « fines » (raw data) ou « agrégées », et peuvent être appréhendées non pas sous la forme d'un tableau plat mais sous la

forme d'un cube de données¹. Cette approche, répandue dans l'univers des statistiques d'enquêtes, se rencontre également dans celui des SIG basés sur le modèle SOLAP (Bédard, 2003). Elle permet d'observer les séries de données spatio-temporelles à travers une ou plusieurs statistiques dans leurs différentes dimensions notamment spatiales.

ThemaMap, application dont les sources sont en accès libre (Open Source), est dédiée à la réalisation de visualisations thématiques de jeux de données géo-référencés. Ce logiciel permet d'accéder à différents types de représentations (symboles colorés, camemberts, flux, histogrammes, étiquettes...) et fournit également plusieurs outils permettant de synthétiser l'information géographique (ex: iso-lignes, grille d'agrégation spatiale, lissage par la méthode des noyaux *kernel*, etc.). Son développement est le fruit d'un travail collaboratif entre le GREYC² et l'UMR 212 EME³, dans un objectif de mutualisation des compétences et en réponse à des besoins communs concernant le traitement cartographique des données d'enquête.

Enfin, bien que ThemaMap puisse prétendre à d'autres champs d'utilisation, la plupart des fonctionnalités développées répondent à des besoins et à des contraintes exprimés en halieutique⁴. Ceci a entraîné une orientation sensible du logiciel vers le traitement et la représentation des types d'informations utilisées dans ce domaine.

2 Les concepts de base de ThemaMap

2.1 Présentation générale de l'application

2.1.1 Généralités

ThemaMap, basé sur la librairie OpenMap™ et développé en Java, est multiplateforme. Il gère les informations en couches indépendantes, prend en compte la gestion de différents systèmes de projection géographique et intègre le rendu cartographique.

Les développements spécifiques réalisés pour ThemaMap concernent :

- **l'architecture** : les modifications apportées ont permis de dissocier la manipulation des données des représentations cartographiques ;

- **l'accès aux données** : divers modules d'accès aux données sont disponibles pour différents formats (csv, dbf, mif/mid, bases postgres, etc.) ; les données sont ensuite accessibles sous la forme d'une table attributaire bidimensionnelle ;

- **l'aspect multidimensionnel** : les données sont appréhendées comme un cube multidimensionnel de données, l'utilisateur peut naviguer au sein de son jeu de données selon les strates d'intérêt (spatiales, temporelles,...) ;

- **Le fonctionnement à travers le Web** : en applet ou en mode Java Web Start.

Outre un positionnement fort sur les analyses thématiques, la démarche de développement du logiciel a privilégié certains aspects :

- ergonomie et facilité de prise en main ;
- panel de fonctionnalités propres aux manipulations des bases de données statistiques ;
- large éventail des modes de visualisation ;
- disjonction entre données sources, chaîne de traitement et mode de représentation (ainsi une procédure conduisant à la construction d'une carte peut-être ré-appliquée sur d'autres jeux de données).

2.1.2 De la donnée à la carte

L'architecture globale du logiciel reprend celle que l'on retrouve dans de nombreux outils de cartographie se basant sur un empilement des couches géographiques. Une couche (*layer*) correspond à un ensemble autonome de données qui donne lieu à des visualisations. Les couches sont indépendantes les

1 Les cubes ou hypercubes (informatique décisionnelle) sont souvent utilisés dans les métiers du contrôle de gestion. Un cube de données correspond à l'union de toutes les combinaisons possibles de « Group By » sur une relation (Jim Gray et al. [ICDE 96]). Ils permettent de naviguer dans l'espace des données, à différents niveaux d'agrégats (zooms, rotation d'axes, etc.). Les tableaux croisés dynamiques de MS Excel® en sont un exemple classique.

2 Le Groupe de Recherche en Informatique, Image, Automatique et Instrumentation de Caen est un laboratoire de recherche associé au CNRS sous la forme d'une UMR rattachée à titre principal à l'INS2I (Institut des Sciences Informatiques et de leurs Interactions) et à titre secondaire à l'INSIS (Institut des Sciences de l'Ingénierie et des Systèmes). Ses thématiques couvrent les principales disciplines du domaine, à savoir les différentes branches de l'informatique et de l'intelligence artificielle, le traitement de signal et d'image, la sécurité informatique, l'automatique, l'électronique et les dispositifs.

3 L'UMR EME « écosystèmes marins exploités », associe l'Ifremer, l'IRD et l'Université de Montpellier II. Elle vise à améliorer la connaissance des impacts des changements globaux (changement climatique, croissance de la demande de produits de la mer et globalisation des marchés, pressions anthropiques de natures variées) sur les écosystèmes marins.

4 Notamment dans le cadre du développement d'un atlas électronique sur la pêche artisanale péruvienne

unes des autres et ne partagent entre elles que la zone d'affichage cartographique de la carte⁵. La construction d'une couche (fig.1) commence par l'importation du jeu de données dans une couche spécifique dédiée à la cartographie thématique⁶ ; de multiples transformations peuvent y être appliquées au moyen d'« adaptateurs » ; enfin l'utilisateur choisit les visualisations qu'il souhaite appliquer.

2.1.3 L'interface utilisateur

La figure 2 présente l'interface utilisateur de l'application. Elle se compose d'un menu contextuel, d'une barre de boutons et de quatre panneaux détachables et ré-organisables. L'ergonomie et la simplicité de l'interface ont été optimisées lors du développement de l'application; cela se traduit par l'absence de hiérarchie complexe de menus et sous-menus : tout ou presque est accessible directement et simplement sans tableau de bord surchargé.

Le **menu contextuel** a été simplifié au maximum pour permettre à l'utilisateur un accès rapide aux fonctionnalités classiques de gestion de couches (ouverture, fermeture, sauvegarde, impression etc.), de navigation globale dans la carte (zoom, système de projection, zone de visualisation), de choix de « vues » par défaut (boîte englobante centrée sur le Pérou par exemple), de gestion de l'affichage des panneaux.

La **barre d'outils** permet un accès rapide aux outils de navigation sur la carte (zoom, déplacement) et de sélection du mode de la souris (sélection, navigation, cliquer-déplacer, calcul de distance ou d'aire). Elle comprend aussi l'affichage de l'échelle courante et des raccourcis vers la création des couches d'habillage de la carte (graticule, légende) facilement configurables ainsi que des couches pré-définies « par défaut » (ex: *couche WMS*, *Etopo*, « *le monde politique* »).

Le panneau « **Navigation** » présente un onglet réduit d'affichage du planisphère (il n'apparaît pas dans la configuration d'affichage de la figure 2); Il permet à l'utilisateur de repérer sa zone de travail courante sur l'ensemble du planisphère et de gérer le niveau de zoom ou le panoramique principal. Remarque : grâce à un mécanisme dit de « vue », il est possible de revenir à une vue antérieure (étendue, niveau de zoom).

Le panneau « **Gestion des calques** » liste les couches et sous-couches chargées. Des boutons

dédiés permettent de les ajouter/ supprimer/afficher/masquer ou encore d'en modifier l'ordre de superposition. Un clic-droit sur un calque sélectionné donne accès à d'autres fonctionnalités (centrage, suppression, renommage, exportation...).

Le **panneau de visualisation cartographique** occupe une place centrale. C'est ici que se synchronisera en temps réel (ou à la demande de rafraîchissement de l'utilisateur) la visualisation des couches. C'est aussi ce « cadre » qui sera sauvegardé en l'état lors d'un export sous forme d'image ou de projet.

Le panneau « **Configuration** » a trait aux manipulations réalisables sur les données, depuis le chargement du fichier source via le choix du format (csv, shp, dbf, etc.) dans une liste déroulante et la spécification du chemin d'accès jusqu'à l'export des données (en gml ou simple csv). Ceci en passant par le choix des différents adaptateurs (figure 3) permettant, entre autres de trier, sélectionner, agréger les données. Ce panneau se compose de 3 onglets :

- l'onglet « Source » permet à l'utilisateur de visualiser à quel(s) fichier(s) se réfère son projet et quels adaptateurs ont été appliqués. C'est aussi à cet endroit qu'il pourra supprimer ou modifier les paramètres du dernier adaptateur ajouté pour revenir à l'étape précédente ;
- l'onglet « Données » permet à l'utilisateur de visualiser la table attributaire actualisée ;
- l'onglet « Représentation » permet le choix et la configuration de chaque visualisation, notamment la sélection des variables (champs de la table attributaire), les méthodes de discrétisation ainsi que les options de mise en forme du rendu à utiliser.

2.2 Les relations données - représentations

ThemaMap a été conçu de manière à rendre indépendante la source de données de l'ensemble des visualisations qui lui sont associées. À un instant t, la visualisation appliquée (ex: symbole proportionnel coloré) ne sera effective que sur le jeu de données stocké dans le *datastorage*, i.e. le jeu de données tel qu'il existe après les manipulations qu'il a subies jusqu'à cet instant. Un changement du fichier de données (si les intitulés de variables sont identiques) peut donc se faire en conservant les paramètres définis aupara-

5 Elles sont d'autant plus indépendantes qu'elles utilisent chacune un thread de calcul différent, ce qui évite qu'une couche gourmande en calcul puisse bloquer l'affichage. Ceci permet ainsi également d'utiliser efficacement les architectures multi-cœur de nos machines actuelles.

6 D'autres types de couche de type raster peuvent également être construites (WMS, Etopo...) mais elles ne seront pas décrites ici.

vant pour parvenir à un rendu identique. Ce principe vaut tant pour de simples modifications (filtre, pivot...) que pour un remplacement du fichier source lui-même.

2.3 Les « adaptateurs » de source de données

Les adaptateurs, outils agissant sur les données, sont disponibles dans l'onglet « source » de chaque couche. Différentes opérations peuvent ainsi être successivement mises en place par empilement, chacune d'entre elles construisant une nouvelle vue du tableau de données en mémoire (*datastorage*), sans altérer le fichier source, dans le but de les trier, filtrer ou encore de les agréger ou de pivoter la table. En outre, ils peuvent être supprimés afin de revenir à l'état antérieur du *datastorage*⁷.

2.4 Les données vues comme un cube multidimensionnel.

Comme évoqué précédemment, une table de données peut être perçue comme un cube multidimensionnel, comportant plusieurs dimensions (généralement composées de modalités) et de statistiques. Ces concepts sont notamment employés dans les entrepôts de données, spatialisés ou non (BI, OLAP, SOLAP). ThemaMap s'inspire de cette approche et appréhende les données comme un hypercube de données comportant de multiples dimensions.

L'avantage de cette démarche est d'autoriser différentes opérations sur les variables (création de filtres, animation, navigation...). Ce principe de conception est également conservé dans toutes les opérations visant à agréger ou à synthétiser les données (agré-gations spatiales, par pivot, enveloppe convexe ...) et qui donnent la possibilité de supprimer certaines dimensions et d'agréger à un niveau supérieur.

2.5 La dissémination des cartes

Si l'un des objectifs du logiciel est de faciliter la réalisation de cartes thématiques par des utilisateurs néophytes en matière de cartographie, un autre objectif tout aussi primordial est d'apporter une aide quant à la dissémination et au partage des cartes réalisées par les utilisateurs « thématiciens » leur permettant ainsi de mener à bien toutes les tâches depuis la conception jusqu'à la mise à disposition de leur carte.

Les cartes peuvent être exportées de plusieurs manières. La façon la plus simple est sous forme

d'image jpeg ou gif ou en format vectoriel SVG. Il est également possible d'exporter les éléments de représentations thématiques dans le format KML directement exploitable par *google earth* ou *google maps*.

D'autre part, il est possible d'exporter les données transformées par les adaptateurs à un instant (t) indépendamment des représentations visuelles en GML, ou en CSV pour des semis de points. Cela permet de restreindre l'export au semis de points et d'éviter les calculs lourds (succession d'adaptateurs) au rappel du fichier.

Outre ces modes d'exportation, ThemaMap permet la sauvegarde des projets (ensembles de couches) ou de simples couches dans des fichiers dits « *properties* ». Ces fichiers stockent le chemin d'accès au(x) fichier(s) sources, la succession des adaptateurs mis en œuvre ainsi que leurs paramètres et enfin les différentes visualisations appliquées. Il est possible de charger ultérieurement ce fichier et de le relire via le logiciel afin de reconstituer de nouveau le rendu dans son intégralité. Le rechargement de ce fichier va répéter toutes les opérations de construction et obtenir un résultat identique ou actualisé si les fichiers de données utilisés ont été mis à jour entre temps. Ces fichiers étant de simples fichiers « texte », ils peuvent être édités par un utilisateur averti à l'aide d'un éditeur de texte classique (Notepad++, Gedit etc.).

Cette propriété est également mise à profit dans la construction de l'atlas électronique de la pêche péruvienne. En effet, il est possible de sauvegarder l'ensemble des fichiers *properties* dans une arborescence (répertoire « Maître » ou un de ses sous répertoires). En mode « **Atlas** », il suffit d'indiquer le chemin du répertoire principal à l'application. Celle-ci lira les fichiers présents et reconstituera l'ensemble des couches et visualisations correspondantes. Dans ce mode, si l'ensemble des fonctions reste accessible, le panneau de gestion des couches est légèrement modifié et donne accès à un explorateur.

Tout ceci confère une grande souplesse au logiciel en termes de format d'échange entre des utilisateurs distants.

Rappelons enfin que le logiciel peut fonctionner en mode java web start et donc que l'ensemble des fonctionnalités peuvent être déployées en mode local ou sur serveur.

7 Toutes les opérations effectuées sont répercutées directement sur les représentations associées.

3 La construction des cartes thématiques

3.1 Les données en entrée

La couche thématique accepte comme source de données géographiques des fichiers de type vecteur au format *shapefile* (ESRI), MIF/MID (MapInfo), GML (*Geographical Markup Language*, extension du langage XML), CSV⁸, DBF ou encore des extractions de base de données PostGIS par construction de requête ou lecture de fichiers SQL.

3.2 La manipulation des données

Les données peuvent subir diverses transformations par le biais d'adaptateurs qui vont agir directement sur le *datastorage* actif :

- la jointure permet d'ajouter des tables attributaires par le biais d'un identifiant commun aux deux séries. On peut ainsi par exemple utiliser des données extraites d'une base PostgreSQL et les combiner à d'autres sources de données (ex : *shapefile*) pour obtenir une table unique et plus complète.

- le filtrage permet de sélectionner une partie des données et donc de restreindre le nombre d'objets géographiques à afficher en ne conservant que les lignes du tableau correspondant à certaines valeurs fixées. ThemaMap fournit un second outil de filtrage plus générique qui conserve les lignes validant une expression logique saisie par l'utilisateur. Celle-ci peut être construite par combinaison d'opérateurs relationnels de comparaison, d'opérateurs arithmétiques et logiques ou de certaines fonctions de calcul sur les géométries.

- l'outil d'édition permet à la fois de modifier ou de supprimer les éléments du tableau, et donne également accès à l'édition des objets géométriques (ajout, suppression, altération des nœuds...)

Les visualisations thématiques

Les visualisations (figure 4) disponibles en standard apparaissent sous la forme de boutons cliquables dans l'onglet « Représentation » du panneau « Configuration » :

1. rendu simple de la géométrie, polygones par exemple pour les découpages administratifs, points pour des lieux de relevé, ou ligne pour un trait de côte ;
2. choroplèthes⁹ ou rendu par aplats de couleurs par discrétisation ;
3. adjonction d'étiquettes ou de textes dont la police, la taille et l'orientation sont paramétrables ;
4. symboles proportionnels à une variable (disques, pictogrammes, polygones ou étoiles réguliers) ;
5. symboles de taille proportionnelle à un champ sélectionné et colorés en fonction d'un second ;
6. camemberts ;
7. histogrammes ;
8. représentation d'une quantité par un groupe de symboles de même taille, dont la règle de proportionnalité dépend du nombre de symboles et non de leur taille¹⁰ ;
9. représentation des flux ou des trajets, les données de départ et d'arrivée pouvant figurer dans la même ligne ou dans la même colonne.

Pour tous ces modes, un certain nombre d'options sont communes comme la modification des figurés des entités géométriques : type de trait, épaisseur, couleur de remplissage, transparence. De plus, tous ces objets sont déplaçables de manière collective ou individuelle par rapport à leur ancrage géographique par défaut.

En ce qui concerne les rendus faisant appel à des classifications colorées (choroplèthes, symboles colorés...), plusieurs méthodes de discrétisation sont proposées : intervalles égaux, amplitudes égales, moyennes emboîtées, etc.

Les éléments de légende et d'habillage de la carte sont accessibles depuis un raccourci de la barre d'outils et gérés via des couches indépendantes de la couche thématique principale. Tous les éléments présents sont modifiables individuellement, on peut ainsi agir sur les positions, les polices, les titres, l'échelle, la boussole et personnaliser totalement la légende.

8 Dans le cas de données ponctuelles géo-localisées, les coordonnées géographiques des points doivent être clairement identifiées par deux colonnes dont les intitulés commencent par des préfixes de « longitude » et « latitude » et un suffixe commun

9 Une carte choroplèthe est une représentation de quantités (plethos) relatives à des espaces, ou aires géographiques (khorè), par le moyen d'une échelle de tons gradués (<http://www.hypergeo.eu/spip.php?article274>).

10 Un groupe de symboles occupe la même aire qu'occuperait un simple symbole proportionnel, tout en ayant l'avantage de pouvoir évaluer de façon fiable à l'œil le doublement de quantité.

4 Outils d'exploration et de synthèse de l'information

Divers outils permettent d'explorer et d'analyser les informations cartographiques de manière un peu plus approfondie suivant deux grands axes : navigation dans les différentes dimensions du jeu de données et construction de nouvelles entités dans le but de synthétiser l'information.

4.1 Les adaptateurs d'agrégations

Deux outils permettent de modifier les niveaux d'agrégation du jeu de données. Pour cela, il suffit d'éliminer une ou plusieurs dimension(s) (variables) du système et de choisir l'opérateur statistique qui sera appliqué. Si la dimension est spatiale, en particulier pour les réseaux de points, il est possible de remplacer ceux-ci par une grille spatiale de résolution variable.

4.1.1 Le pivot

L'outil « pivot » permet, tout d'abord, de transposer (pivoter) le tableau de données de façon à obtenir, sur une même ligne, plusieurs variables attachées à la même entité géographique. Ceci permet de construire ensuite des camemberts et des courbes ou histogrammes. De plus, il permet d'agréger des données en fonction d'un ou plusieurs critères, en appliquant une fonction d'agrégation telle que somme, moyenne....

4.1.2 Grille d'agrégation spatiale

Une autre catégorie d'adaptateur permet de transformer les données notamment ponctuelles en grille régulière. Cette fonctionnalité est très importante notamment dans le cas de données liées au domaine marin. Le fonctionnement est similaire à celui du pivot mais la cellule de réception de l'agrégation s'inscrit dans une grille dont on a défini le maillage « à la volée ». Sur ces nouvelles entités, tous les modes de représentation peuvent s'appliquer.

4.2 Constructions de variables par ajout de colonne

Bien que cet adaptateur soit aisé à mettre en œuvre, il permet d'aboutir à des constructions complexes dépassant le simple ajout de colonnes. Le principe est basé sur la mise à disposition de fonctions de nature diverses pouvant être combinées entre elles et opérant sur les variables afin d'aboutir à la création d'un nouveau champ. Parmi les fonctions élaborées on peut retrouver :

- des fonctions sur les chaînes de caractères : TRIM, CONCAT, ...

- des fonctions de comparaisons : >, =, <, ou logiques classiques : OR, AND, .IF, ...
- des fonctions mathématiques : ROUND, LOG, COS, ...
- des fonctions à valeurs numériques opérant sur les géométries : calcul d'aire, de distance, ...
- des fonctions statistiques prenant en compte toutes les valeurs d'une colonne et donnant accès aux fonctions classiques telles que moyenne, écart type, minimum ou maximum. A ce niveau, des classifications sur une variable par les mêmes méthodes que celles utilisées pour les discrétisations sont également accessibles.

De plus, il est possible de construire des géométries : point ou disques en fonction des valeurs de variables numériques ou par union, intersection ou construction de *buffers* sur des géométries existantes.

4.3 Outils de synthèse de l'information

Plusieurs outils d'analyse permettent la construction de nouvelles entités géographiques tout en respectant le principe de multidimensionnalité des données, ils sont essentiellement destinés aux réseaux de points ou aux grilles maillées.

On peut ainsi de façon simple construire :

- des enveloppes convexes (*convex hull*) permettant de visualiser l'aire de répartition maximale d'un ensemble de points ;
- des centroïdes de distribution ou des centroïdes pondérés par la valeur d'une autre variable numérique ;
- des objets issus de l'union ou de l'intersection d'entités ayant des valeurs de modalités communes ;
- des tracés de contours pouvant ensuite être sauvegardés sous forme de lignes ou de polygones ;
- un lissage par la méthode des noyaux (*kernel*). Cette opération nécessite plusieurs étapes : le choix du maillage de la grille, le choix de la fonction noyau (gaussien, quadratique, uniforme...) et le choix du rayon.

Pour chacun de ces outils, les panneaux d'options permettent de retenir les dimensions d'intérêt. Ainsi, si dans la construction d'un centroïde, le mois et l'année sont conservés, une nouvelle entité géométrique sera construite pour chaque strate année/mois.

4.4 L'adaptateur d'animation

Cet adaptateur est destiné à la production de cartes animées. Il fonctionne grâce aux filtres pré-

sentés auparavant. Dans un premier temps, une variable d'animation est sélectionnée ; il est ensuite possible de faire défiler les cartes en fonction des modalités successives, ce défilement pouvant être opéré de manière manuelle ou paramétré par un timer¹¹. On peut également synchroniser différentes variables entre elles dans une couche en fonctions de critères d'asservissement des modalités (synchronisation entre l'année et le mois par exemple). Plusieurs couches peuvent également se voir synchronisées, permettant ainsi de faire défiler des thématiques différentes en fonction de valeurs communes.

Un autre intérêt de cet outil réside dans la production de cartes en séries au format image en fonction des variables d'animation sélectionnées.

5 Application à la pêche artisanale péruvienne

5.1 Le système d'enquête « pêche artisanale » de l'IMarPe¹²

Le réseau d'enquête déployé sur les côtes péruviennes¹² pour le suivi des pêches artisanales constitue un exemple classique des modalités de collecte de l'information dans ce secteur. Opérationnel depuis 1997, ce système mobilise une quarantaine d'enquêteurs répartis sur autant de ports de l'ensemble du littoral. Divers protocoles de recueil d'information coexistent avec des niveaux de renseignements spatiaux et temporels plus ou moins détaillés :

- des enquêtes de recensement quinquennales (dites enquêtes cadre) qui visent en premier lieu à dénombrer les effectifs (embarcations, engins...) au niveau des ports ;
- des enquêtes de suivi des débarquements qui permettent le recueil d'informations sur les quantités débarquées par type d'espèces ainsi que sur les caractéristiques et les positions géographiques des actions de pêche associées ;
- des enquêtes de nature socio-économique qui visent à recueillir des informations sur les prix à la vente, la situation économique et sociale des patrons de pêche (niveau d'étude, structure familiale...).

Parmi les éléments constitutifs des statistiques de pêche collectées figurent des valeurs telles que les quantités de poissons capturées, le nombre d'heures passées en mer, etc... Les dimensions thématiques

seront les 14 engins de pêche présents, les 200 espèces pêchées, etc... Les strates temporelles seront les dates, les mois, les années. En ce qui concerne la dimension spatiale, on rencontre des objets géographiques souvent régis par des règles de hiérarchies emboîtées :

- des objets ponctuels dont les positions peuvent être exprimées par des paires de coordonnées latitude/ longitude (ex : villages, positions de pêche, sites de débarquement...);
- des entités de type polygone (régions, pays, zones de pêche, zones statistiques...);
- des grilles dont la taille des mailles est variable.

5.2 Exemple 1 : Exploitation des données de recensement

L'IMARPE a mené une enquête cadre en 2005 visant à dénombrer les engins de pêche ainsi que les embarcations en activité. Au cours de ces enquêtes, ont été relevés les ports, le type d'engin, le type d'embarcation. Nous allons nous servir de ces résultats pour illustrer des opérations simples de création de cartes thématiques (cf. figures 5 et 6).

Les données de départ sont contenues dans un simple fichier CSV comportant l'intitulé des ports et leurs références spatiales et, la dénomination de la région d'appartenance, les noms des engins de pêche, et enfin les effectifs dénombrés. En d'autres termes, les dimensions spatiales sont constituées par les ports et les régions, les engins sont de nature thématique, les effectifs constituent les statistiques. Nous utilisons également une couche de fond cartographique du Pérou, disponible au format *shape*.

La première étape consiste à importer les deux fichiers (ajout couche thématique->choix du format), les objets géographiques sont identifiés automatiquement (coordonnées en Latitude/Longitude) et sont affichés sur le panneau de cartes. Le fichier comportant les effectifs de tous les engins observés. Nous ne retiendrons que les plus importants, pour simplifier l'illustration, en appliquant un adaptateur de filtre sur les modalités d'engin (*Pinta, Cortina, Cerco, Espinel*).

Comme nous l'avons précisé auparavant, les variables à représenter doivent être en colonne. Il faut donc passer d'un modèle Région/Engin/Effectifs à une structure où les ports restent en ligne mais les modalités des engins deviennent des colonnes et au croisement lignes/colonnes, on retrouve les effectifs. Cette

¹¹ Une option permet de fixer l'échelle à un même niveau pour toutes les cartes produites.

¹² Instituto del Mar del Peru

transformation passe par la mise en œuvre de l'outil « pivot » qui donne accès au choix des variables faisant office de lignes et de colonnes.

Il nous faut associer cette série au fichier d'entités géographiques des différentes régions par l'application de l'adaptateur de jointure. Pour cela, il suffit de choisir la colonne commune aux deux fichiers (*datastorage* et fichier « *shape* » contenant les régions).

Pour réaliser la figure 5, il faut :

- choisir une représentation de type camembert, dans le panneau d'option de la représentation, sélectionner les variables participant à la construction du graphe (ici le diamètre est proportionnel à la somme des valeurs mais d'autres choix sont offerts),
- ajouter une représentation pour placer les labels de région,
- modifier individuellement la position des labels et des camemberts,
- ajouter une couche de légende et placer les titres et les légendes,
- exporter la carte sous forme d'image ou la conserver sous forme de fichier « *properties* ».

Le principe de construction de la figure 6 est similaire mais les effectifs des engins sont exprimés en pourcentage du total. L'étape supplémentaire consiste à calculer les pourcentages par engin. Dans un premier temps, nous pouvons calculer une colonne « total » grâce à l'outil d'ajout de colonne, puis une seconde variable correspondant à la division de l'effectif de l'engin par ce total. Une solution plus élégante est d'appliquer directement la formule :
$$\text{engin1} / (\text{engin1} + \text{engin2}) * 100.$$

5.3 Exemple 2 : Exploitation des données d'effort de pêche

Dans cette seconde illustration, nous allons importer un fichier contenant des données brutes d'action de pêche déjà agrégées partiellement et contenant pour une année : les informations sur le port d'attache, le mois, la position de pêche référencée en latitude et longitude, l'engin de pêche, le nombre de voyages.

Après importation du fichier, nous allons tout d'abord appliquer un filtre sur les années pour ne conserver que l'année 2003, puis un second filtre sur les engins de pêche (limitation à la pêche au calamar géant ou *potá*), un troisième filtre aura pour effet de restreindre le *datastorage* à la région de Piura. L'application d'un filtre élaboré basé sur des opérateurs logiques de type « ET » aurait eu des conséquences identiques.

Une représentation élémentaire consiste à poser un symbole centré sur le point donné par la latitude/longitude et d'aire proportionnelle au nombre de marées comme montré figure 7. Cependant, la multiplication des données mensuelles rend très peu lisible le rendu qui empile les symboles sur un même lieu.

Une possibilité de synthèse consiste à agréger le nombre de voyages sur une grille maillée en fonction d'une statistique d'agrégation (ici la somme). Dans l'exemple présenté figure 8, le maillage est de 1×1 degré. Deux représentations du nombre de marées y sont appliqués ; la première est du type symbole proportionnel, la seconde est une classification colorée (méthode de discrétisation : moyennes emboîtées).

Nous pouvons conjuguer la construction de cette grille à la localisation du port d'origine (cf. figure 10). Dans cette optique, lors de la définition de la grille, nous avons conservé les coordonnées des ports, ce qui revient à effectuer la somme des marées pour chaque croisement cellule/port. Il suffit ensuite de sélectionner la représentation « flux-trajets », d'indiquer les entités géographiques de départ et d'arrivée et de préciser que la dimension de la flèche est proportionnelle aux effectifs (un filtrage a permis de ne montrer que les valeurs supérieures à 100).

D'autres outils permettent à la fois de visualiser et de synthétiser ce type de données. La figure 9 est un exemple de *convex hull* montrant l'extension géographique des zones de pêche de chaque port. Ce type de représentation est accessible par l'outil d'agrégation spatiale, l'opérateur d'agrégation sur la géométrie étant l'enveloppe convexe. Seul le port a été conservé comme dimension (nous obtenons donc une enveloppe par port).

En modifiant les paramètres de cet adaptateur, c'est à dire avec un choix d'opérateur portant sur « centroïde pondéré » et le mois comme dimension, nous aboutissons à la figure 10. Les flèches qui joignent les différents points montrent les déplacements moyens mensuels.

5.3 Exemple 3 : Exploitation des données de capture

Les trois cartes présentées (Figures 11, 12 et 13) illustrent la mise en œuvre des lissages par la méthode des noyaux *kernel* et la réalisation d'iso-contours. Les données de départ sont des quantités pêchées agrégées annuellement par zone de pêche. Sur la figure 11, elles sont représentées sous forme de grille sur laquelle est appliquée une discrétisation. La carte suivante est obtenue par l'application de l'adaptateur de lissage sur les données de capture avec les paramètres suivants : grille de maillage de 0,1 degré, fonc-

tion quartique, rayon de 0,2. Le résultat est une nouvelle grille dont les cellules contiennent les données de capture interpolées. Sur ce nouveau *datastorage*, les valeurs sont ensuite traduites en cercles proportionnels et en aplats de couleur. Sur cette grille résultante, la représentation des captures par courbe de niveau, met en évidence les zones de concentration fortes (seuls deux seuils figurent dans l'exemple). Elle nécessite l'emploi de deux adaptateurs, le premier construit une triangulation de Delaunay ; le second réalise les contours en fonction du paramétrage des valeurs de niveau.

6 Conclusion

Le logiciel a pour vocation d'aider à la représentation cartographique des séries statistiques (d'autres usages ne sont bien entendu pas exclus). Dans cette optique, l'application présente des fonctionnalités

dédiées à la valorisation cartographique des données d'enquête (notamment halieutiques) et qui favorisent l'exploration et la synthèse de ces informations.

Le logiciel est conçu dans un souci de simplification des interfaces, afin de faciliter sa prise en main par les utilisateurs néophytes en matière de cartographie statistique. De plus, ses possibilités de transfert des réalisations sur le Web, en particulier par la création aisée d'atlas électronique en font un outil efficace tant dans le traitement des données géo-référencées que dans la dissémination des cartes produites.

Le logiciel ThemaMap, traduit en Espagnol et en Anglais, est disponible pour téléchargement sur le site : <http://themamap.info.unicaen.fr/>. De nombreux exemples y sont présentés ainsi que des aides à l'utilisation.

Bibliographie

Rivest, S. Y. Bédard, M.J. Proulx & M. Nadeau, 2003, SOLAP: *a new type of user interface to support spatio-temporal multidimensional data exploration and analysis*, Workshop International Society for Photogrammetry and Remote Sensing (ISPRS), Quebec, Canada, October 2-3.

Estrella Arellano, C., Swartzman, G., 2010. *The Peruvian artisanal fishery : changes in patterns and distribution over time*. Fish. Res. 101, 133–145.

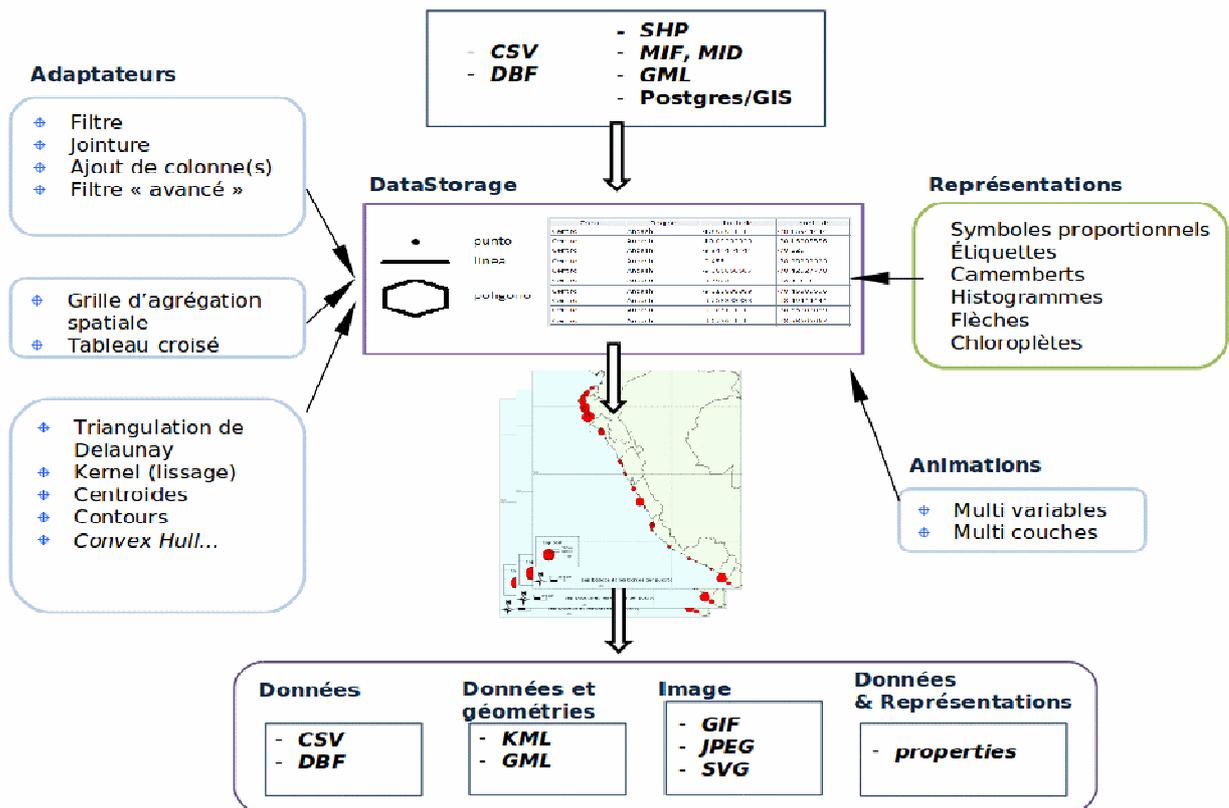


Figure 1 : Schéma conceptuel de ThemaMap

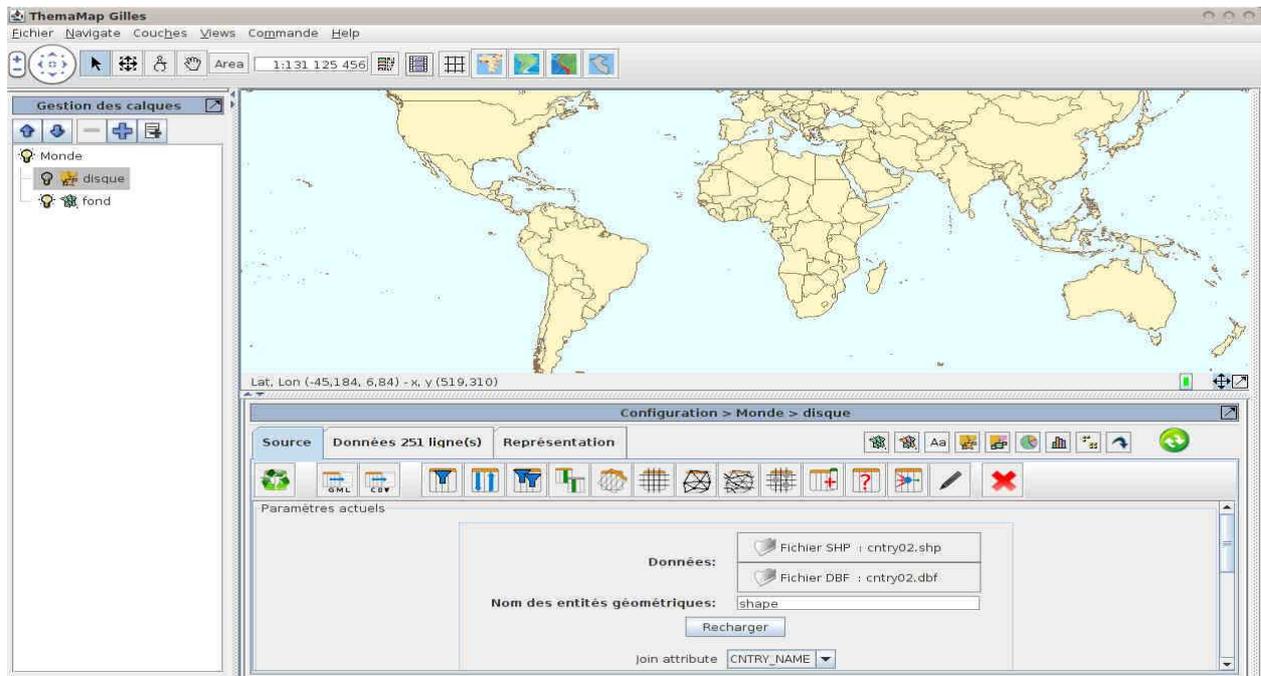


Figure 2 : L'application ThemaMap



Figure 3 : Liste des adaptateurs de ThemaMap



Figure 4 : Liste des boutons d'accès aux visualisations thématiques standards de ThemaMap

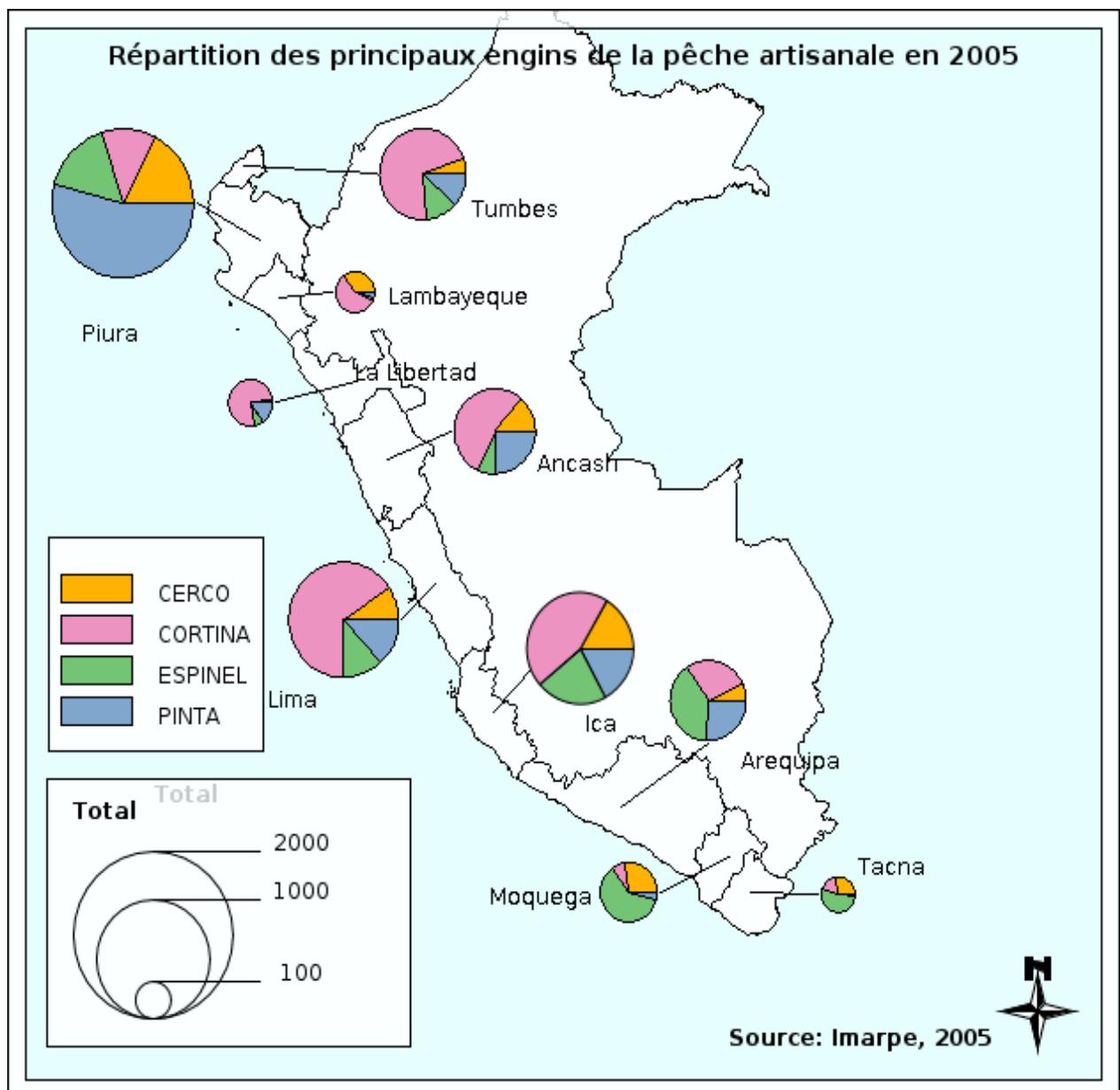


Figure 5 : Exemple de carte thématique (Camemberts et labels)

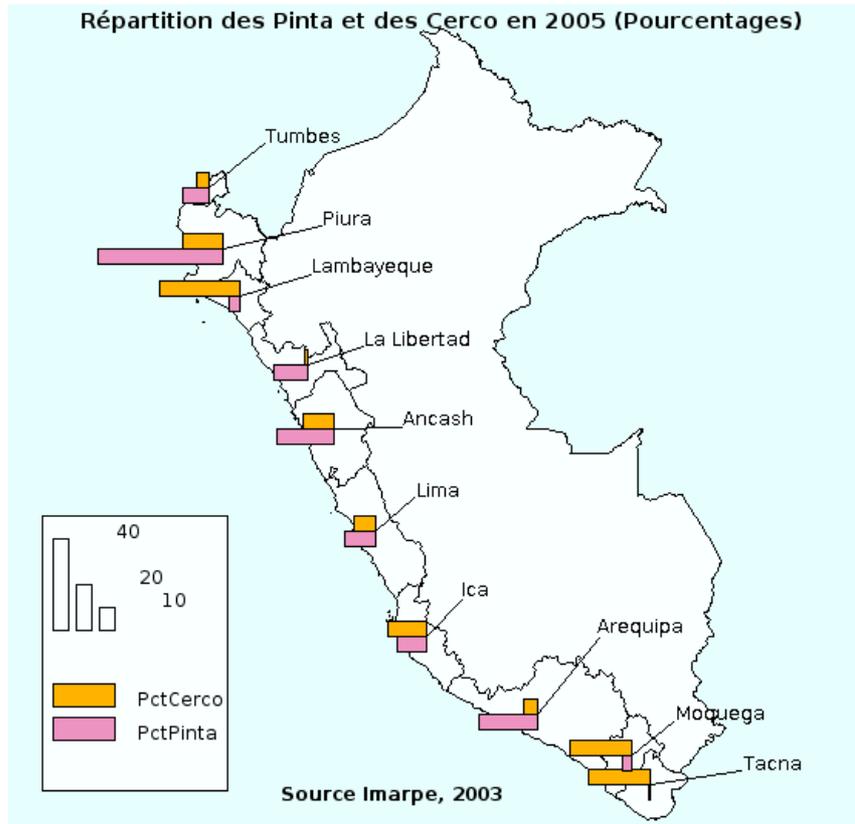


Figure 6 : Exemple de carte thématique (Histogrammes)

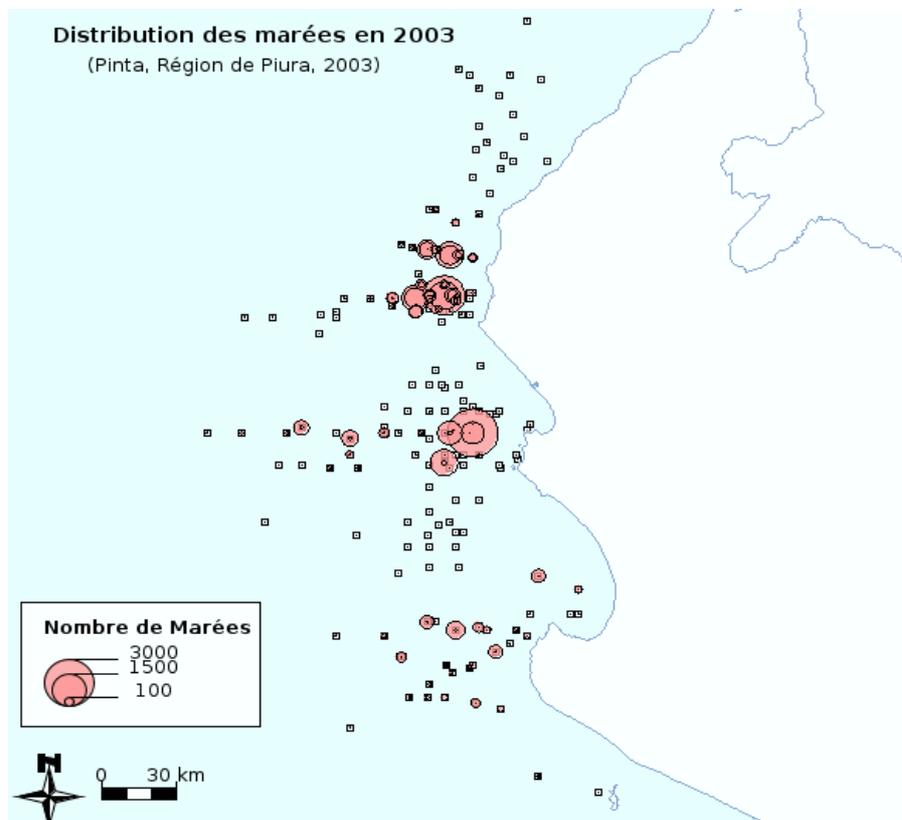


Figure 7 : Exemple de carte thématique (symboles proportionnels)

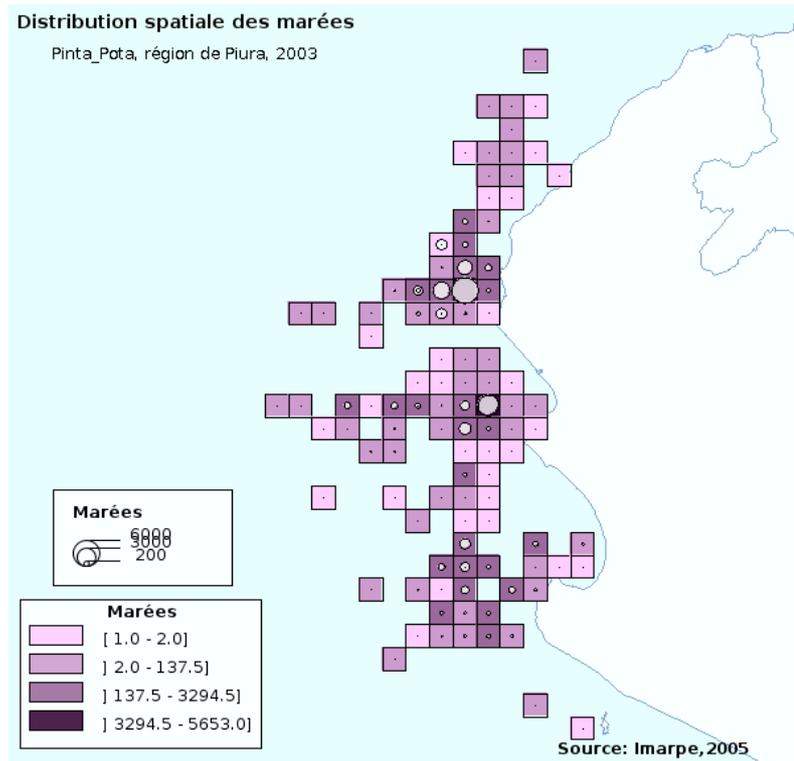


Figure 8 : Exemple de carte thématique (Grilles spatiales)

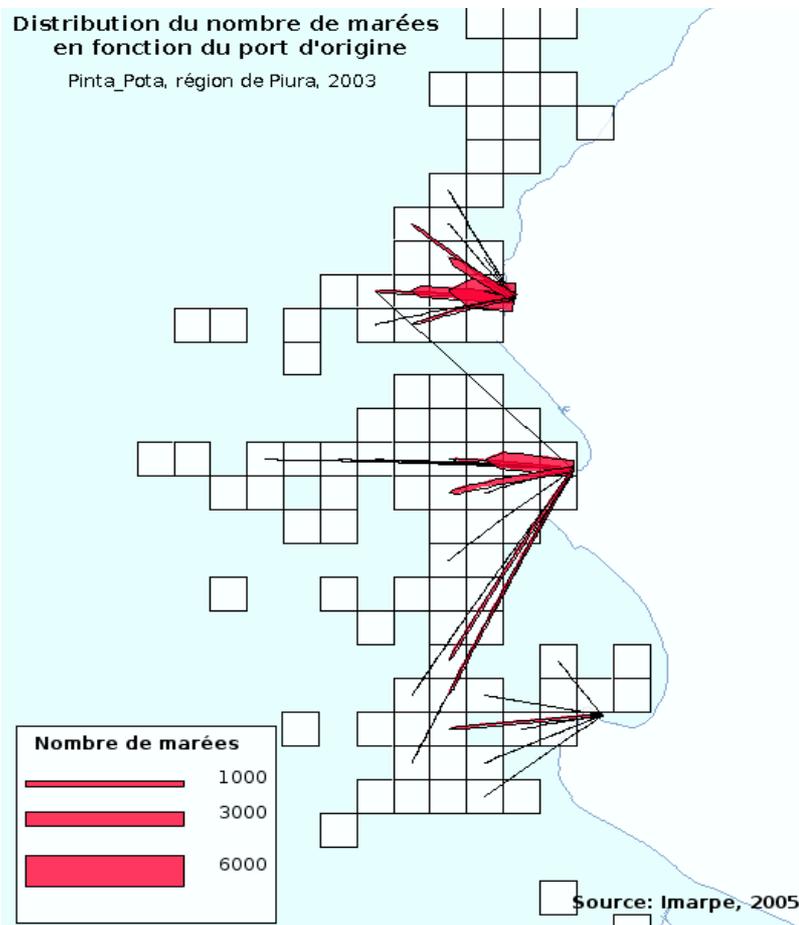


Figure 9 : Exemple de carte thématique (Flux)

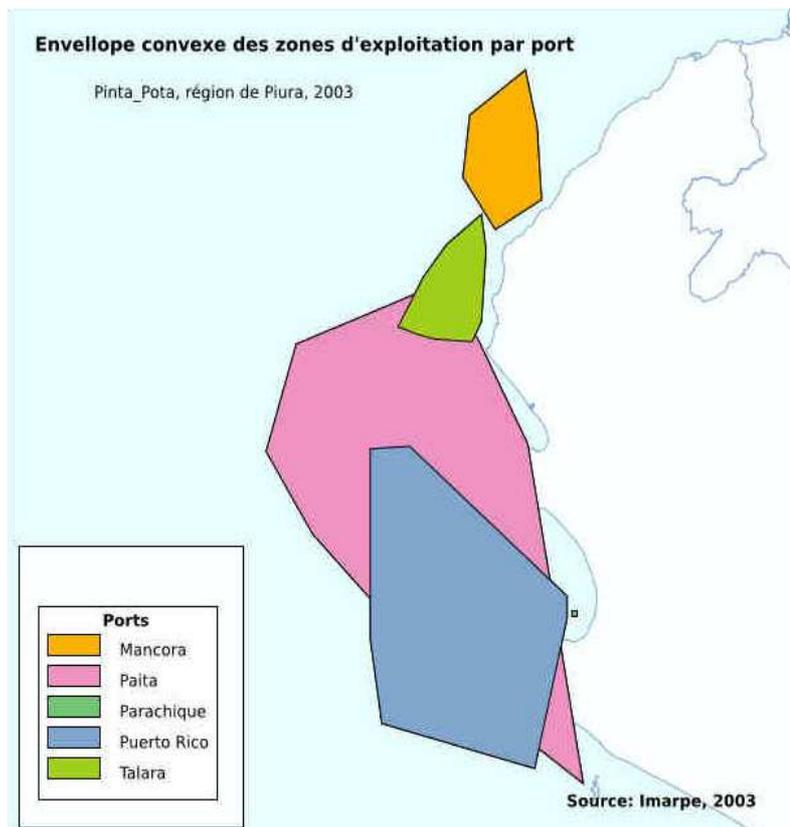


Figure 10 : Exemple de carte thématique (Convex Hull)

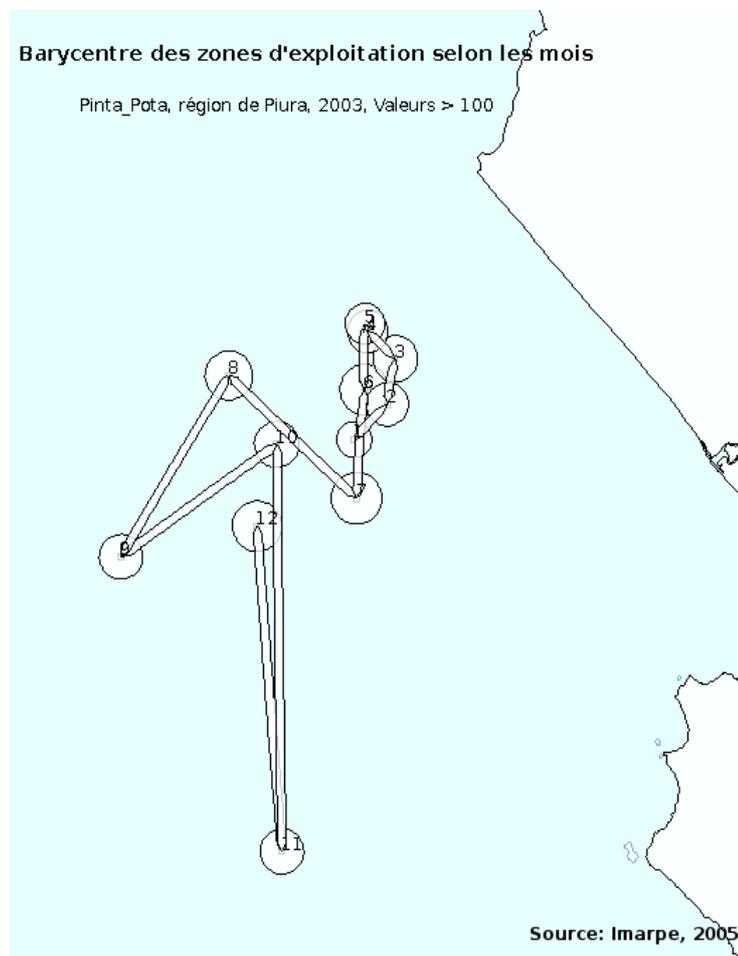


Figure 11 : Exemple de carte thématique (centroïdes pondérés)

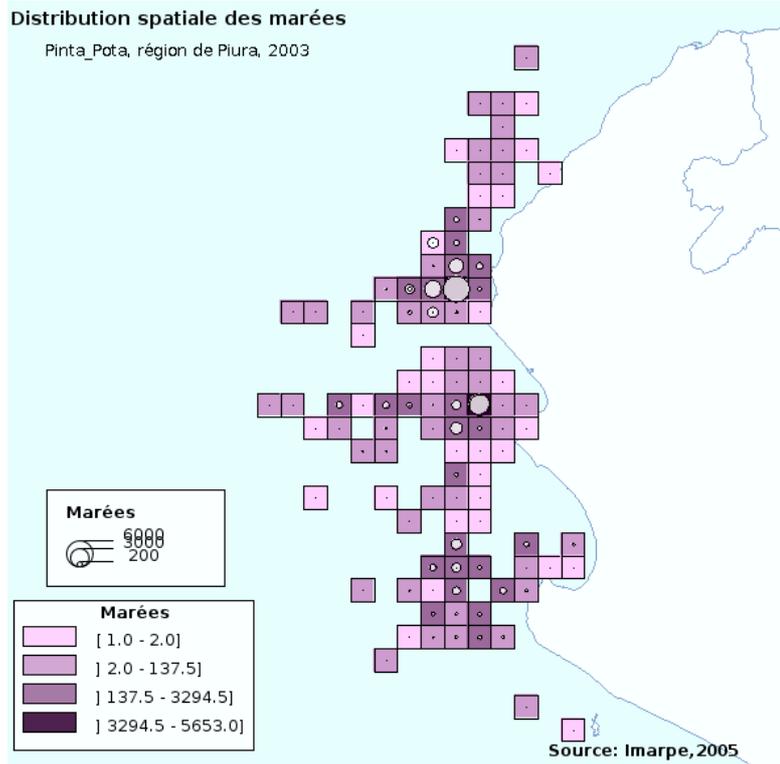


Figure 12 : Exemple de carte thématique (Grilles)

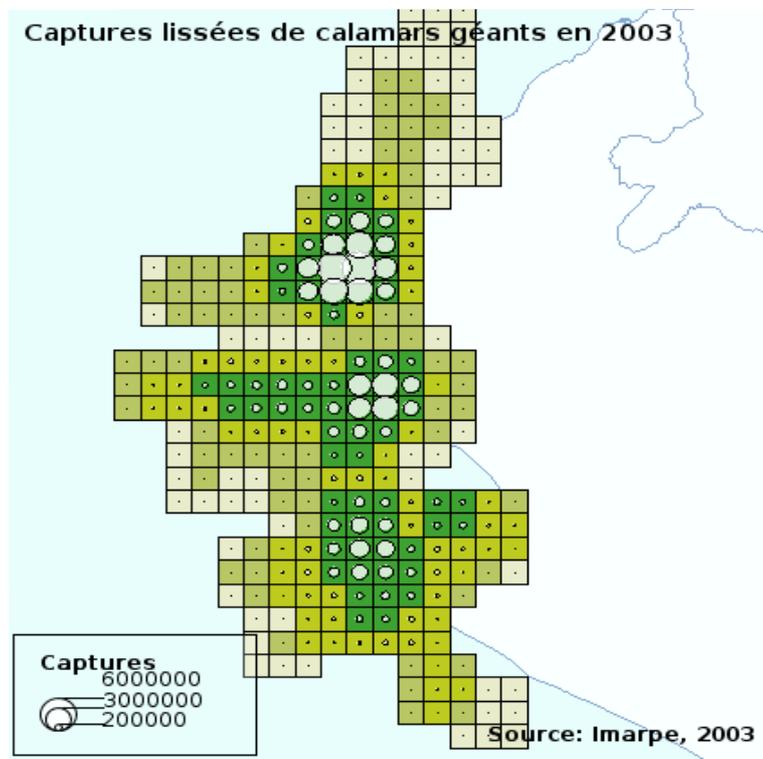


Figure 13 : Exemple de carte thématique (Kernel)

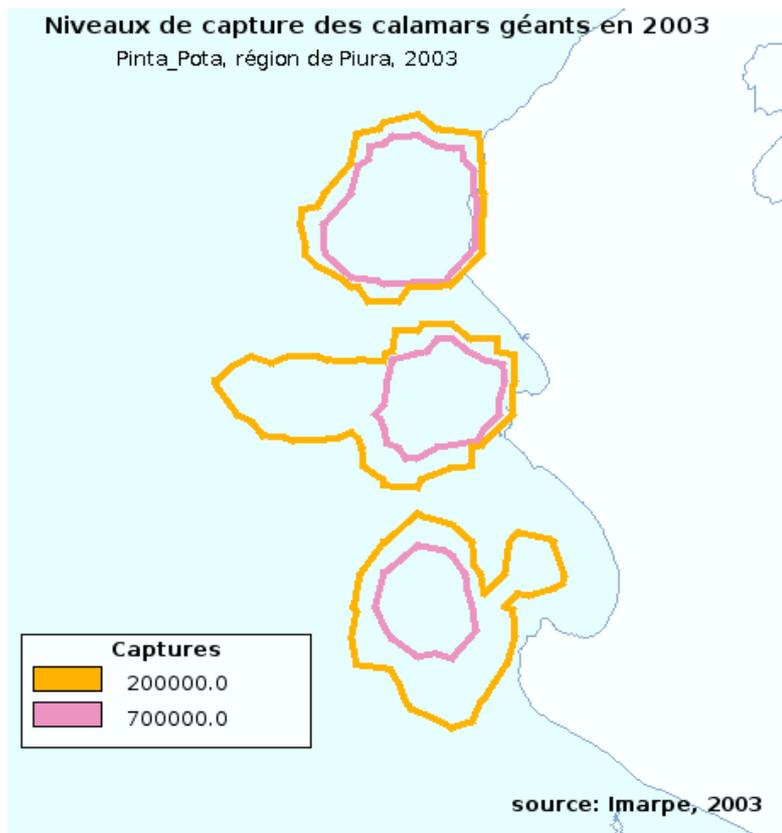


Figure 14 : Exemple de carte thématique (Contours)

Domalain Gilles, Rodriguez Céline, Madelaine J., Turbout C.,
Estrella C. (2010)

THEMAMAP : un outil de cartographie thématique des séries
statistiques

Le Monde des Cartes, (207), 33-48

Colloque : La Géomatique à l'IRD, Bondy (FRA), 2010/10/18-
19

ISSN 1634-3522