



Master 2 Informatique
Ingénierie et Gestion Territoriales
Spécialité Géomatique

Développement d'une chaîne de traitements pour créer des cartes d'aléas en relation avec les différents facteurs de risque du paludisme

Gilles ENTRINGER

Stage réalisé du 06.02.2012 au 06.08.2012 à l'IRD sous la direction de Bertrand GUERRERO.

Tuteur pédagogique : Thérèse LIBOUREL

Date de soutenance : 25.09.2012

Membres du Jury : Thérèse LIBOUREL,

Remerciements

Je tiens tout d'abord à remercier mon maître de stage Bertrand Guerrero qui m'a permis de réaliser mon stage dans les meilleures conditions possibles. Je lui remercie pour ses conseils, le temps qu'il m'a consacré et pour ses commentaires sur mon travail.

Je remercie également mes co-tuteurs Nadine Dessay, Maud Loireau et Vincent Herbreteau pour leur disponibilité, leurs conseils et leur aide tout au long de mon stage.

Merci aussi à Christelle Pierkot, Jean-Christophe Desconnets, Samuel Andres et tous les membres de l'équipe SIC de l'UMR Espace-Dev pour leurs conseils techniques et pour leur patience.

Je tiens aussi à remercier Mireille Fargette et tous les membres de l'équipe AIMS de l'UMR Espace-Dev qui sont intervenus tout au long de l'élaboration de mon travail de stage.

Je voudrais remercier tout particulièrement mon tuteur pédagogique Thérèse Libourel de m'avoir aidé à trouver ce stage et de découvrir le monde de la recherche et pour ses conseils et pour le temps qu'elle m'a consacré.

Enfin, je tiens à remercier le personnel administratif de l'IRD, de l'UM2 et du LIRMM pour le temps qu'ils m'ont consacré.

Sommaire

1	Introduction	1
2	Contexte et problématique	3
2.1	Principes généraux	3
2.2	Le logiciel SIEL	3
2.2.1	Le "modèle" SIEL	4
2.2.2	Architecture	4
2.3	Problématiques du stage	4
3	Etat de l'art	6
3.1	Environnement-santé	6
3.1.1	Définitions...	6
3.1.2	Facteurs de risque Paludisme	7
3.2	Approches logicielles existantes	7
3.2.1	Architecture / Langages informatiques utilisés	7
4	Méthodologie	8
4.1	Modèle du cycle du paludisme	8
4.2	Chaîne de traitements	8
4.2.1	Modèle des traitements	8
4.2.2	Choix architecture informatique	8
4.3	Conceptualisation de la base de données	8
4.3.1	Modèle conceptuel	8
4.3.2	Modèle relationnel	8
4.3.3	Modèle physique / technique	8
5	Résultats	9
5.1	Les chaînes de traitements	9
5.2	Discussion	9
5.3	Perspectives	9
6	Conclusion	10
7	Glossaire et définitions	11
8	Bibliographie	12

Chapitre 1

Introduction

Le stage s'est déroulé au sein de deux équipes de l'UMR Espace-Dev¹, l'équipe SIC et l'équipe AIMS.

Les objectifs de l'UMR Espace-Dev sont multiples. L'UMR s'inscrit dans une perspective de développement durable des territoires et propose des méthodologies de spatialisation des dynamiques de l'environnement. L'UMR développe également et exploite également un réseau de stations de réception de satellites d'observation de la terre.

L'UMR se regroupe en trois équipes de recherche :

- Equipe OSE (Observation spatiale de l'environnement), spécialisée dans la télédétection et les images satellitaires
- Equipe AIMS (Approche intégrée des milieux et des sociétés) qui est spécialisée dans le domaine de l'environnement, de la télédétection et dans l'élaboration d'indicateurs et de modèles des interactions milieux/sociétés ou des dynamiques socio-environnementales
- Equipe SIC (Systèmes d'information et de connaissances) qui a pour objectif l'acquisition, la gestion, la représentation et le partage des données et des connaissances. D'autres objectifs de cette équipe sont la modélisation de dynamiques spatio-temporelles, la visualisation, la cartographie sémantique et l'aide à la décision

Le présent stage s'inscrit dans un contexte de coopération entre l'équipe SIC et l'équipe AIMS.

Depuis plusieurs années au sein de l'UMR Espace-Dev est développé le SIEL (Système d'Information sur l'Environnement à l'Echelle Locale). Le SIEL est un logiciel d'aide à la décision dans la gestion de l'environnement. Le SIEL permet d'évaluer le risque de dégradation de la végétation en milieu aride et permet de produire des indices environnementaux spatialisés sous la forme de cartes.

L'objectif du présent stage est d'élargir les potentialités en construisant dans un premier temps des indicateurs dans le contexte environnement-santé et plus précisément du paludisme. Le développement d'une chaîne de traitements indépendante devrait permettre de rendre le fonctionnement du SIEL plus souple, plus ouvert et plus facile face à des problématiques diverses comme par exemple l'environnement-santé.

Le mémoire est organisé de la façon suivante :

- Une première partie sera dédiée à la présentation du contexte du stage. Je présenterai les problématiques liées à la construction des indicateurs et au développement d'une chaîne de traitements. Dans cette partie je présenterai également plus en détail le SIEL.

1. <http://www.espace.ird.fr/index.php>

- Dans la deuxième partie, un état de l’art présentera à partir d’une recherche bibliographique les principes généraux liés à la problématique environnement-santé ainsi que des approches logicielles existantes.
- Une troisième partie présentera la méthodologie de mon travail. Une modèle du cycle du paludisme sera présenté. Les différentes étapes de l’élaboration de la chaîne de traitements seront également proposées : un modèle des traitements, l’architecture informatique choisie ainsi que la conceptualisation de la base de données.
- Dans la dernière partie du mémoire seront présentés les résultats de mon travail.

Par la suite, les termes en **gras** seront définis dans le glossaire en fin du mémoire.

Chapitre 2

Contexte et problématique

2.1 Principes généraux

A l'heure actuelle, le logiciel SIEL permet d'effectuer un certain nombre de traitements bien définis. Dans le domaine de l'environnement-santé il y a une forte demande en terme d'automatisation des traitements / automatisation de la création d'indicateurs.

Il est donc particulièrement intéressant d'intégrer la construction automatisée d'indicateurs dans le contexte environnement-santé au sein du SIEL. A l'heure actuelle, le SIEL est un plugin ArcGIS qui fonctionne avec des séries de traitements / opérations ciblés sur le type d'indices à produire. L'objectif à long terme est de retravailler ce logiciel, de préférence dans un contexte "OpenSource" et plus facilement réutilisable.

Dans ce contexte, le développement d'une chaîne de traitements indépendante pourrait servir comme base pour ce développement.

La définition des facteurs de risque est une autre étape indispensable de la conceptualisation et de l'élaboration de la chaîne de traitements. A partir d'un modèle du cycle de la malaria, des modèles de données et de traitements permettront de dégager les étapes nécessaires pour aboutir au résultat souhaité : des cartes de risques ...

2.2 Le logiciel SIEL

Le SIEL est un outil conçu pour répondre aux besoins des scientifiques et décideurs concernés par la lutte contre la dégradation et la gestion des ressources naturelles. Le SIEL est capable d'anticiper les risques et de les suivre, ce qui est important notamment dans le cadre de la mise en place d'observatoires. Cet outil donne la capacité aux scientifiques et aux décideurs de mesurer l'impact des opérations déjà effectuées et d'optimiser les actions futures notamment en terme de gestion des ressources.

Le logiciel a été conceptualisé à partir de 1993 par Maud Loireau dans le cadre de sa thèse. En même temps un premier prototype a été développé. En 2000, dans le cadre de la lutte contre la désertification, la démarche scientifique et le modèle général du SIEL ont été adoptés par le programme ROSELT (Réseau d'observatoires de Surveillance Ecologique à Long Terme) de l'OSS (Observatoire du Sahara et du Sahel). Ceci a permis de développer un premier logiciel SIEL qui couvrait l'ensemble de la chaîne de traitement proposée. Une première version du logiciel traitant la ressource « végétation » a pu être déployée dans le réseau ROSELT/OSS dès la fin de 2003. A partir de 2006, les différents acteurs ont poursuivi les développements du logiciel. Ceci a débouché sur un logiciel opérationnel et prêt à être diffusé, une documentation

(guide d'utilisateur, aide en ligne), une documentation technique (guide développeur, dictionnaire des données), une base de données exemple et des supports de formations pour objectif d'optimiser le logiciel. Actuellement, l'objectif principal reste le développement informatique du logiciel. Le responsable du projet informatique est depuis 2010 Bertrand Guerrero.

2.2.1 Le "modèle" SIEL

Le SIEL est un modèle environnemental instrumenté (Loireau et al. 2005^{6a}) qui spatialise les pratiques d'exploitation des ressources sur un territoire, modélise les paysages et évalue les risques de dégradation des ressources. (référence guide utilisateur)

Le SIEL se base sur le fonctionnement interactif milieux/sociétés sur un territoire dans lequel des acteurs, selon leurs représentations de leur territoire et leurs usages (agricole, pastoral, forestier), ont des pratiques qui permettent d'exploiter des ressources naturelles (physiques, biologiques) sur un territoire donné (Guide utilisateur). Les interactions entre l'espace, les acteurs, les usages et les ressources produisent des paysages qui vont composer les territoires. Ces paysages évoluent en fonction des interactions entre facteurs naturels et humains. Pour mettre en oeuvre ce phénomène, **le modèle SIEL décompose le territoire en Unités Spatiales dites « de Référence » (USR). Ces USR correspondent à une délimitation spatiale des différents paysages du territoire par voie de modélisation.**

Un minimum de données est nécessaire pour alimenter la chaîne de traitements :

- données socio-économiques (acteurs, stratégies, pratiques d'exploitation...)
- données biophysiques (ressources, végétation...)

Enfin, le modèle SIEL met en œuvre une méthode spatiale interdisciplinaire qui permet une cartographie dynamique des paysages et le calcul d'indices environnementaux associés (Loireau et al, 2007⁷).

2.2.2 Architecture

Le SIEL se présente sous forme d'un plugin ArcGIS et se base sur deux modules propriétaires :

- Module données (Logiciel Microsoft Access, langage de programmation : Visual Basic for Application (VBA)) pour la gestion et le stockage des tables attributaires. Gestion des bases de données (format .mdb).
- Module SIG (Logiciel : ArcGIS Desktop avec l'extension Spatial Analyst, langage de programmation : Visual Basic 6.0 et langage ESRI ArcObjects) pour la gestion et le stockage des couches géographiques et les géotraitements. Même format que pour le module données pour la gestion de la base de données.

Le système d'exploitation supporté est Windows XP, les logiciels supportés sont "Microsoft Office 2003" et "Microsoft Office 2007" pour le module données et "ArcGIS 9.1" et "ArcGIS 9.3" pour le module SIG.

Les deux modules, en faisant différents traitements, créent des indices environnementaux spatialisés qui permettent la cartographie des risques de dégradation des ressources.

2.3 Problématiques du stage

Étant données la pluridisciplinarité du sujet de stage, les problématiques se regroupent en 2 catégories :

- Problématiques liées à la maladie (paludisme) et aux facteurs de risques
- Problématiques au niveau informatique

Concernant problématiques liées à la maladie, la problématiques principale est la définition des facteurs de risque et la construction d'un modèle du cycle u paludisme.

Plusieurs problématiques au niveau informatique peuvent être soulevées : Choix de l'architecture informatique, choix du langage informatique principal, conceptualisation d'une base de données à partir du modèle du cycle de la malaria, implémentation de la chaîne.

En résumé une problématique générale se dégage : Conceptualisation et développement d'une chaîne de traitements à partir d'un modèle du cycle du paludisme.

Chapitre 3

Etat de l'art

3.1 Environnement-santé

On s'est focalisé sur les indicateurs pertinents dans le contexte environnement-santé et plus précisément du paludisme

Définitions...

3.1.1 Définitions...

Environnement-santé

Paludisme

Le paludisme, appelé également malaria est une parasitose due à un protozoaire transmis par la piqûre de la femelle d'un moustique (anophèle), provoquant des fièvres intermittentes. Il set cause d'environ deux millions de décès chaque année dans le monde, principalement dans les régions tropicales et en Afrique sub-saharienne. (5).

Un médecin français, Alphonse Laveran, a découvert la cause de la maladie en 1880 à Constantine (Algérie). La moustique anophèle se reproduit dans les zones marécageuses. Le parasite qui possède plusieurs hôtes intermédiaires, dans l'état endémique, infecte les cellules hépatiques de la victime puis circule dans le sang. Au cours de son cycle de vie, le parasite à l'intérieur de l'organisme humain, fait un certain nombre de transformation qui lui permettent d'échapper au système immunitaire humain. Au final, lorsqu'une moustique non-infectée pique une personne contaminée, le parasite est également transmis de l'homme à la moustique.

En France, le paludisme a disparu depuis les années 1960 mais n'a pas complètement de nos jours due aux nombreux voyages effectués dans les pays de risque, notamment dans les régions tropicales et subtropicales.

Actuellement, le paludisme est responsable de plus de 300 millions de cas de maladie aiguë et d'au moins un million de décès dans le monde. Quatre-vingt-dix pour cent des décès dus au paludisme surviennent en Afrique, au Sud du Sahara, principalement chez les jeunes enfants. De nos jours, aucun vaccin efficace n'a encore pu être développé et les scientifiques doutent de plus en plus qu'ils trouveront un jour la solution miracle contre cette maladie. (6)

vulnérabilité

Facteur de risque

Dans le cadre de cette étude, les facteurs de risques sont l'ensemble des éléments qui augmentent la probabilité que le paludisme se développe sur un territoire.

3.1.2 Facteurs de risque Paludisme

En me basant sur une recherche bibliographique, j'ai pu créer une liste des différents facteurs causant le développement du paludisme. Les facteurs ont été regroupés en 3 catégories : Les facteurs liés à l'environnement, le facteur biologiques et les facteurs humains et jouent un rôle plus ou moins important dans le développement du paludisme, ce qui sera expliqué dans les trois paragraphes qui suivent. Pour chaque catégorie un tableau présentera les différents facteurs ainsi que, si possible, les valeurs minimum et maximum nécessaires pour ces facteurs.

Facteurs liés à l'environnement

Numéro	Facteur	Valeur min	Valeur max
1	Précipitation	80mm	720 mm
2	Température	16°C	32°C
3	Distance au point d'eau le plus proche	0 km	7 km
4	Surface du point d'eau		
5	Altitude	0 m	2000 m
6	Normalized Difference Vegetation Index	-1	1
7	Humidité		
8	Humidité du sol		
9	Température mois précédant les précipitations	5	40
10	Qualité de l'eau		
11	Profondeur de l'eau		
12	turbidité de l'eau		
13	PH du sol		

Facteurs biologiques

Facteurs humains

3.2 Approches logicielles existantes

On s'est simultanément intéressé aux logiciels qui traitent ce genre de problématiques...

3.2.1 Architecture / Langages informatiques utilisés

Tableau récapitulatif avec les critères qui permettent de dégager ce qui manque et ce qu'on va proposer.

Chapitre 4

Méthodologie

L'objectif est de proposer une architecture logicielle permettant de calculer des indicateurs dans le contexte du paludisme (environnement-santé). Calculer les indicateurs revient à définir des chaînes de traitements. Celles-ci doivent être conceptualisées à partir de l'expertise du domaine. On a donc commencé par s'intéresser au modèle du cycle de paludisme, pour connaître les territoires vulnérables...

4.1 Modèle du cycle du paludisme

Les modèles / évolution

Introduire le fait que les observations (terrain / satellite) ont été faites sur le territoire concerné, qu'elles concernent le moustique, l'habitat humain et la végétation/surface aquatique, éventuellement les données climatiques. Ces observations vont être le point de départ de la chaîne de traitements à effectuer.

4.2 Chaîne de traitements

4.2.1 Modèle des traitements

S'appuyer sur la thèse de Yuan, je sais quels sont mes données, je sais ce que je veux fabriquer comme indicateur et je demande à l'expert les traitements qu'il faut faire. (A partir des indicateurs qu'ils veulent faire). Comment ensuite réaliser la chaîne réelle, il faut connaître les données d'entrées, d'où elles sont issues, les traitements (les algos).

Création de base de données et chargement de données

4.2.2 Choix architecture informatique

4.3 Conceptualisation de la base de données

4.3.1 Modèle conceptuel

4.3.2 Modèle relationnel

4.3.3 Modèle physique / technique

Chapitre 5

Résultats

5.1 Les chaînes de traitements

5.2 Discussion

5.3 Perspectives

Chapitre 6

Conclusion

Chapitre 7

Glossaire et définitions

Chapitre 8

Bibliographie