

# Formalisation de l'intégration de données géospatiales dans un modèle multi-agents : cas de LittoSIM-GEN

Cécilia Pignon-Mussaud<sup>1</sup>, Ahmed Laatabi<sup>1</sup>

<sup>1</sup>UMR 7266 LIENSs, CNRS, Université de La Rochelle, France.

## Introduction

LittoSIM [1] est un dispositif de simulation participative (figure 1) qui combine un jeu sérieux (modèle multi-agents distribué sous la plateforme GAMA) et un modèle hydrodynamique de submersion (LISFLOOD-FP). Co-construit en 2015 avec la Communauté de Communes de l'île d'Oléron, il vise essentiellement la sensibilisation aux risques de submersion marine auprès des acteurs locaux (élus et gestionnaires). Depuis 2019, le projet LittoSIM-GEN a pour objectif d'étendre ce dispositif à d'autres territoires métropolitains, afin d'évaluer sa généricité et son efficacité en matière d'apprentissage. La diversité, l'hétérogénéité et la multiplicité des données géospatiales nécessaires au modèle LittoSIM induisent une complexité supplémentaire à la modélisation et des traitements longs et chronophages lors de l'intégration de nouvelles données.

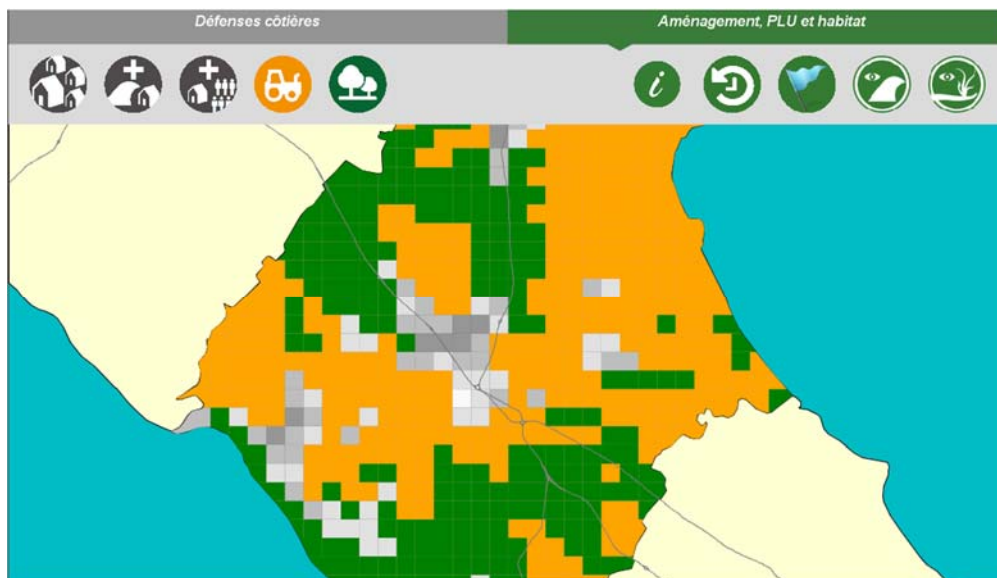


Figure 1 : l'interface joueur de la plateforme LittoSIM

## Les données géospatiales et les modèles multi-agents

La réutilisation des modèles est l'un des défis majeurs de la science aujourd'hui. Le manque de transparence dans la construction et la formalisation des relations données modèle complexifie la modélisation multi-agents. En effet, il est souvent difficile de savoir comment les données brutes sont utilisées dans le modèle, et comment elles sont transformées pour être intégrées dans sa structure. Cette situation est fréquente dans le cas des données géospatiales, qui sont par nature, hétérogènes, composites et incomplètes. L'application d'un modèle existant dans un autre contexte avec de nouvelles données nécessite une multitude d'opérations pour transformer les fichiers bruts

(shapefile, raster, grille...) à une structure adaptée au modèle. Il est donc nécessaire de définir une approche formalisée pour faciliter l'intégration des données dans un modèle géospatial. Une telle approche doit répondre principalement à deux questions :

- **Où** sont les données dans le modèle, et ceci à travers des liaisons de mapping.
- **Comment** ces données sont transformées, et ceci à travers une description détaillée du prétraitement des données.

### **Le mapping données-modèle**

Le mapping est la spécification d'un ensemble de mécanismes pour transformer une structure à une autre. Le mapping données-modèle consiste à définir les liens de correspondance entre la structure de données source, et la structure nécessaire au modèle. Le schéma de mapping permet donc de spécifier les opérations de transformation à faire lors de la réutilisation d'un modèle sur de nouvelles données. Pour cela, le processus de mapping nécessite, d'une part, d'identifier les structures source et destination, et d'autre part, de créer la liaison entre ces deux structures, en précisant les prétraitements à réaliser.

En termes de données géospatiales, les liens de mapping représentent l'ensemble des opérations sur les composantes spatiales (création, intersection, fusion, reprojection, contrôle topologique...) et textuelles/attributaires (création, conversion, agrégation, remplacement...). Le mapping constitue donc un protocole à suivre pour faciliter le prétraitement et l'intégration des données géospatiales dans la structure du modèle. Les schémas de mapping permettent également de guider l'automatisation des géo-traitements, en particulier s'ils sont documentés à travers des descriptions détaillées.

### **La description de l'intégration des données au modèle**

Les descriptions textuelles sont un moyen efficace pour favoriser la réutilisation des modèles multi-agents. Elles permettent de faciliter la compréhension, la dissémination, et notamment la réutilisation du modèle avec de nouvelles données pour dupliquer ou reproduire ses résultats. Ce genre de documentation peut contenir des éléments graphiques, notamment les schémas de mapping.

Le protocole ODD (*Overview, Design concepts, and Details*) s'est imposé comme l'un des formats les plus utilisés par la communauté multi-agents pour décrire la structure des modèles. ODD a été étendu pour répondre à des besoins spécifiques, notamment par ODD+2D (ODD + Décision + Données) [2] qui se focalise sur l'utilisation des données empiriques dans un modèle multi-agents. Ce protocole propose un schéma de mapping pour relier les données sources à la structure de données utilisée par le modèle.

### **Application : le projet LittoSIM-GEN**

Le tableau 1 montre un extrait de la structure de données requise par le modèle LittoSIM (fichier bâtiments). Chaque shapefile doit avoir une structure spatiale bien définie, ainsi qu'un certain nombre d'attributs prédéfinis (libellé, type et valeurs) pour permettre le bon fonctionnement du modèle. La construction de cette structure n'est pas évidente, et nécessite un ensemble

d'opérations pour transformer les données brutes qui proviennent de différentes sources (IGN, DDTM, collectivités territoriales, ...). Ce processus chronophage entrave la réutilisation de LittoSIM et son déploiement sur de nouveaux territoires.

Shapefile	Attribut	Type	Valeurs	Description
buildings	bld_type	Texte	{ <i>Residential, Other</i> }	Type du bâtiment
	bld_area	Double		La surface du bâtiment (m <sup>2</sup> )

Tableau 1 : un extrait de la structure de données du modèle LittoSIM : exemple des bâtiments

L'identification des données sources spécifiques à l'étude de cas (ici Oléron), permet de construire le schéma de mapping dont un extrait est représenté dans la figure 2 : l'entité *buildings* est construite à partir du fichier *bâtiments* en faisant plusieurs traitements (patterns de transformation) de génération, de remplacement et de contrôle. Le schéma est ensuite complété par une description détaillée de type ODD+2D pour expliciter chaque pattern de transformation, et permettre ainsi sa compréhension et son implémentation.

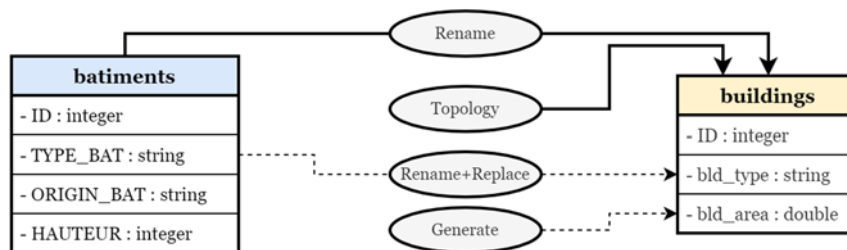


Figure 2 : le mapping du fichier buildings

Par exemple, le pattern « *Rename+Replace* » génère l'attribut *bld\_type* en renommant l'attribut original *TYPE\_BAT* et en remplaçant sa valeur « *résidentiel* » par « *Residential* » et toutes les autres valeurs « \* » par « *Other* ». En plus des transformations textuelles, des patterns de traitements spatiales sont également décrits. Le pattern « *Topology* » de *buildings* indique que les bâtiments qui intersectent avec plusieurs cellules de la grille LittoSIM doivent être découpés (figure 3). Un autre exemple de géo-traitement est celui des cellules limitrophes de la grille qui doivent être fusionnées pour avoir des cellules avec une surface minimale cliquable par le joueur (figure 4).

Le schéma de mapping et la description ODD+2D sont ensuite traduits à un fichier de mapping XML, puis à un script automatique de prétraitement de données sous R (tableau 2). Ce processus évite à l'utilisateur l'ensemble d'opérations coûteuses en termes de temps et d'effort pour transformer les données géospatiales spécifiques au territoire à la structure du modèle LittoSIM.

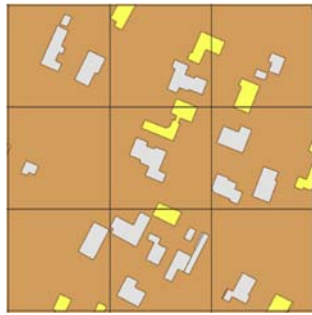


Figure 3 : découpage des bâtiments

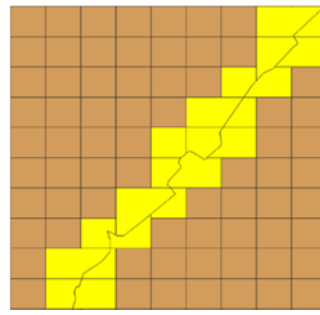


Figure 4 : fusion des cellules

Extrait du fichier de mapping XML (Oléron)	Extrait du script automatique R
<pre>&lt;file output-name="buildings.shp" input-name="batiments.shp"&gt;   &lt;attribute output-name="bld_type" type="string"&gt;     &lt;input-name&gt;TYPE_BAT&lt;/input-name&gt;     &lt;replace value="résidentiel"&gt;Residential&lt;/replace&gt;     &lt;replace value="*"&gt;Other&lt;/replace&gt;   &lt;/attribute&gt; &lt;/file&gt;</pre>	<pre>names(shape_file)[names(shape_file)   == attribute_input_name] &lt;-   attribute_output_name; shape_file@data[attribute_output_name] [[1]] &lt;- as.character(shape_file@data   [attribute_output_name][[1]]);</pre>

Tableau 2 : la transformation de l'attribut « bld\_type »

## Conclusion

La généralisation du modèle LittoSIM dans le cadre du projet LittoSIM-GEN est rendue possible par la formalisation et la description de la relation données-modèle. Elle a permis de faciliter le déploiement du dispositif sur plusieurs territoires littoraux (figure 5). L'adoption d'une telle approche facilite le traitement des données géospatiales et favorise ainsi la réutilisation des modèles complexes.

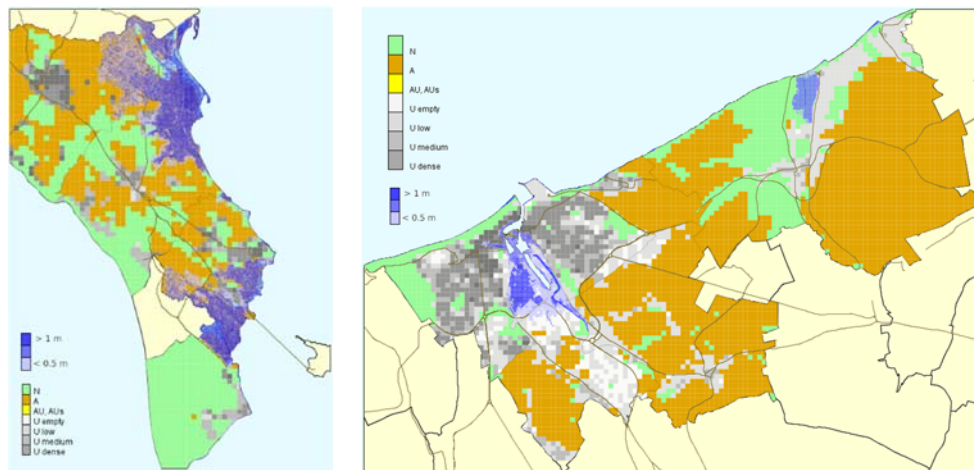


Figure 5 : l'application de LittoSIM-GEN à Oléron (gauche) et Normandie (droite)

## Références :

- [1] Becu, Nicolas, et al. "Participatory simulation to foster social learning on coastal flooding prevention." *Environmental modelling & software* 98 (2017): 1-11.
- [2] Laatabi, Ahmed, et al. "ODD+ 2D: an ODD based protocol for mapping data to empirical ABMs." *Journal of Artificial Societies and Social Simulation* 21.2 (2018).