



**HAL**  
open science

## **Semaine d'Etude Maths-Entreprises 8 : Evaluation du choix d'un maillage de terrassement sur l'impact des filières de gestion des sols pollués : aspect développement durable et économique.**

Papa-Abdulaye Faye, Vy-Thuy-Lynh Hoang, Hoang-Minh Le, Thomas Michel, Vinh Duc Nguyen, Xuan-Lan Phun, Philippe Ung

### **► To cite this version:**

Papa-Abdulaye Faye, Vy-Thuy-Lynh Hoang, Hoang-Minh Le, Thomas Michel, Vinh Duc Nguyen, et al.. Semaine d'Etude Maths-Entreprises 8 : Evaluation du choix d'un maillage de terrassement sur l'impact des filières de gestion des sols pollués : aspect développement durable et économique. : Rapport d'étude réalisé pour Géo-Hyd (Groupe ANTEA) lors de la SEME d'Orléans (Semaine d'Etude Maths Entreprise).. [Rapport de recherche] MAPMO, Université d'Orléans. 2014. hal-01083638

**HAL Id: hal-01083638**

**<https://hal.science/hal-01083638>**

Submitted on 13 Mar 2015

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

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

# Evaluation du choix d'un maillage de terrassement sur l'impact des filières de gestion des sols pollués : aspect “développement durable et économique”

FAYE Papa Abdoulaye <sup>a</sup>    HOANG Vy Thuy Lynh <sup>b</sup>  
LE Minh Hoang <sup>b</sup>            MICHEL Thomas <sup>c</sup>  
NGUYEN Vinh Duc <sup>d</sup>        PHUN Xuan Lan <sup>b</sup>  
UNG Philippe <sup>b e</sup>

- a.* Université Blaise Pascal, Clermont-Ferrand
- b.* MAPMO, Université d'Orléans
- c.* IMB, Université de Bordeaux
- d.* IRMAR, INSA Rennes
- e.* Equipe ANGE, Inria-Rocquencourt

Sujet proposé par :



Correspondant : Patrick SUIRE (Antea Group)  
Coordinateur universitaire : Stéphane CORDIER (MAPMO - Université d'Orléans)

## Résumé

Ce document s'inscrit dans un contexte d'évaluation des sources de pollution des sols et leur gestion. En particulier, on s'intéresse à la dépollution de sols suivant quelques sondages opérés et observations de terrain, pour la plupart, fortement hétérogènes. Cette étape de dépollution consiste à déterminer des volumes de matériaux par degré de pollution. L'inconvénient d'une telle démarche se situe essentiellement dans la responsabilité juridique du bureau d'étude. En effet, une erreur sur les volumes impliqués peut entraîner des préconisations disproportionnées en contradiction avec la notion de développement durable.

En pratique, le maillage de terrassement est carré et chaque maille n'est pas nécessairement pourvue d'un sondage réalisé préalablement.

En prenant en compte les contraintes, principalement géométriques, intervenant dans la mise en place d'un maillage de terrassement, on décrira dans un premier temps, une méthode permettant d'enrichir artificiellement les données de sondage disponibles et de proposer plusieurs cartographies de pollution de la zone étudiée. Puis dans un second temps, on cherchera à estimer la variabilité des coûts relatifs à différents maillages.

## Remerciements

Nous tenons à remercier en premier lieu l'ensemble du comité d'organisation de cette SEME orléanaise, Romain Abraham, Stéphane Cordier, Laurent Delsol et Cécile Louchet, pour leur accueil chaleureux au sein du MAPMO ainsi que pour leur soutien indéfectible durant toute cette semaine chargée en émotion mais également pour toute l'aide ô combien importante qu'ils ont pu nous fournir pour mener à bien ce projet court en durée mais grande en intensité.

Nous adressons également nos remerciements à Patrick Suire pour ce sujet intéressant ainsi que GeoHyd-Anteagroup à travers Daniel Pierre et Hervé Noel, sans qui nous n'aurions pas pu travailler ensemble dans une saine émulation.

Dans un deuxième temps, nous voulions remercier tout particulièrement Mounir Haddou pour tous ses conseils et son aide. Dans une autre mesure, un merci à nos camarades du soir, Ibrahim Cheddadi et Alexander Getmanenko, pour les discussions que l'on a pu avoir et les propositions qu'ils ont pu faire toute cette semaine.

Dans un troisième temps, mais non des moindres, un grand merci également à Marie-France Grespier pour la bonne organisation de cette SEME.

Enfin, nous remercions chaleureusement le GdR Math-Entreprises et l'AMIES à l'origine de la SEME.

# Table des matières

<b>1</b>	<b>Introduction</b>	<b>4</b>
<b>2</b>	<b>Cartographie de la zone de pollution</b>	<b>4</b>
2.1	Champs de Gibbs . . . . .	5
2.1.1	Description d'une carte . . . . .	5
2.1.2	Modèle probabiliste – Champs de Markov . . . . .	6
2.1.3	Champs de Gibbs . . . . .	6
2.1.4	Echantillonneur de Gibbs . . . . .	7
2.1.5	Modèle de Potts . . . . .	8
2.1.6	Cas d'étude . . . . .	8
<b>3</b>	<b>Evaluation des coûts</b>	<b>9</b>
3.1	Calcul des coûts . . . . .	9
3.2	Application numérique . . . . .	11
<b>4</b>	<b>Conclusion</b>	<b>15</b>

## 1 Introduction

Dans le cadre des études définissant les sources de pollution dans les sols et les moyens à mettre en œuvre pour les gérer, il est nécessaire d'établir des évaluations de volumes de matériaux par degré de pollution. Ces évaluations sont basées sur les quelques sondages disponibles et des observations de terrain souvent très hétérogènes. Les volumes alors indiqués vont impliquer une responsabilité juridique très importante du bureau d'étude. En cas d'erreurs sur ces volumes impliqués, les préconisations peuvent être disproportionnées (nombre de camions important, centres de stockage sollicités pour rien) ce qui devient très contradictoire avec la notion de développement durable. Alors que les dépollutions de sol seront en général réalisées sur la base d'un maillage carré ( $10\text{ m} \times 10\text{ m}$ ) en considérant chaque maille unitaire par tranche de  $1\text{ m}$  de hauteur, on ne dispose pas, la plupart du temps, d'un sondage pour chaque maille pendant les études préalables. Pendant les travaux, l'entreprise de terrassement peut être amenée à réaliser des sondages complémentaires mais la plupart du temps, elle se fixe un maillage par défaut et ne fait les analyses qu'en interne avec quelques contrôles ce qui implique de très nombreuses réclamations tant sur l'approche technique que sur les intérêts défendus.

Pour faire face à ces problématiques, Antea Group recherche pour ses sites pollués un modèle d'estimation d'une configuration (ou d'un intervalle de configurations) de mailles optimales intégrant différentes contraintes de son métier comme le choix du maillage de dépollution, la stratégie d'attaque et surtout les contraintes de terrassement d'un engin dont le rayon de braquage, l'épaisseur pelletable à chaque passe de godet, la piste d'accès à la fosse et le choix de la zone de démarrage du terrassement en fonction des concentrations les plus fortes connues.

Le travail décrit par la suite s'inscrit dans ce cadre d'étude. On met en place une méthodologie permettant d'évaluer la variabilité des coûts liés à un maillage. Dans un premier temps, on présente la façon de générer un échantillon basé sur l'échantillonneur de Gibbs et le modèle de Potts, pour dans un second temps, caractériser la loi de distribution des coûts précédemment évoqués.

## 2 Cartographie de la zone de pollution

Un travail préalable effectué par Antea-GeoHyd [4] a permis de mettre en évidence quelques paramètres influant sur la validité et la fiabilité d'une proposition d'un maillage de terrassement ; celui-ci a fait l'objet d'un rapport. De cette étude, il en ressort que les principaux paramètres sont :

- la taille des mailles,
- leur forme,
- leur orientation,
- la position de l'origine du repère,
- ... .

Cependant, il paraît évident que l'élaboration d'un maillage sans prendre connaissance du terrain devient rapidement inenvisageable et que le manque de données issues des sondages de terrain freine considérablement la démarche. De plus, aucun modèle physique ne permet de connaître la migration qui pourrait s'opérer sur les déchets dont il est question.

Compte tenu de ces limitations, on est donc rapidement amené à proposer une méthode pour enrichir les données disponibles. Bien qu'il semble impossible de déterminer avec exactitude la

carte de pollution du terrain étudié, on se propose d'évaluer la variabilité d'un coût en lieu et place de l'évaluation du coût réel qui semble difficilement réalisable.

Ainsi, on met en place une méthode rigoureuse permettant de générer aléatoirement des cartes afin d'évaluer la variabilité des coûts engendrés par un type de maillage. L'approche consiste à construire une carte selon le modèle de Potts à l'aide de l'échantillonneur de Gibbs.

## 2.1 Champs de Gibbs

Les champs de Gibbs, et plus généralement les champs markoviens, constituent une des techniques de base en traitement d'image. Dans l'optique du travail proposé, le modèle de Potts est intéressant en ce qu'il permet de varier l'homogénéité de la distribution spatiale de degré de pollution. L'idée est simple : il est envisageable qu'une cellule partage le même degré de pollution que son proche voisinage. Cette méthode nécessite donc une description locale des interactions entre cellules.

En premier lieu, on se propose d'introduire la notion préliminaire de champs de Gibbs ; l'ensemble de ce qui suit reprend les grandes lignes de [3].

### 2.1.1 Description d'une carte

Les champs de Gibbs nécessitent au préalable de définir les champs de Markov. En général, les dépollutions de sol sont réalisées sur la base d'un maillage carré en considérant une maille unitaire de  $1 m$  de hauteur. L'ensemble des mailles unitaires formant la carte est noté  $\mathcal{M}$  et chaque maille sera désignée par un site  $s_i$ . Chaque site porte l'information du degré de pollution de la maille. On note  $E$  l'ensemble des degrés de pollution possibles ; dans le cas présent,

$$E = \{1, 2, 3, 4\}.$$

Afin d'introduire une notion d'interactions locales, on définit un système de voisinage  $\mathcal{V}$  associé à  $\mathcal{M}$  tel que

$$\mathcal{V}_s = \{t\} \text{ tels que } \begin{cases} s \notin \mathcal{V}_s, \\ t \in \mathcal{V}_s \Rightarrow s \in \mathcal{V}_t. \end{cases}$$

**Définition 1.** Une clique notée  $c$ , est

- soit un singleton de  $\mathcal{M}$ ,
- soit un ensemble de sites tous voisins entre eux.

On renvoie à [3] pour plus de détails. On note  $\mathcal{C}$  l'ensemble des cliques associé à  $\mathcal{V}$  et  $\mathcal{C}_k$  l'ensemble des cliques de cardinal  $k$ . On exprime les interactions locales comme un potentiel de clique noté  $U_c$ . Par conséquent, l'énergie globale de  $\mathcal{M}$  s'obtient par la somme des potentiels de toutes les cliques :

$$U = \sum_{c \in \mathcal{C}} U_c,$$

et l'énergie locale en un site correspond à la somme des potentiels de toutes les cliques auxquelles il appartient

$$U_s = \sum_{c \in \mathcal{C} | s \in c} U_c.$$

### 2.1.2 Modèle probabiliste – Champs de Markov

Une carte est vue comme une réalisation d'un champ aléatoire. En d'autres termes, on introduit  $X_s$  la variable aléatoire à valeur dans  $E$  indiquant le degré de pollution du site  $s$ , et on note  $X^s \in E^{|\mathcal{M}|-1}$  le vecteur aléatoire contenant les degrés de pollution des sites extérieurs au site  $s$ . Une carte n'est autre qu'une réalisation du champ aléatoire  $X = (X_s, X_t, \dots)$  à valeurs dans  $E^{|\mathcal{M}|}$ . Par la suite, on notera  $\Omega = E^{|\mathcal{M}|}$ .

La probabilité globale de  $x \in E^{|\mathcal{M}|}$ ,  $P(X = x)$ , donne la vraisemblance d'une carte, et les probabilités conditionnelles locales d'une valeur en un site établissent le lien statistique entre un degré de pollution et le reste de la carte. L'hypothèse markovienne permet d'évaluer ces quantités [3].

**Définition 2.**  $X$  est un champ de Markov si et seulement si la probabilité conditionnelle locale en un site  $s$  ne dépend que de la configuration du voisinage de  $s$ , soit

$$P(X_s = x_s | X^s = x^s) = P(X_s = x_s | \{X_t = x_t, t \in \mathcal{V}_s\}),$$

avec  $x_s$  le degré de pollution au site  $s$  et  $x^s = (x_t)_{t \neq s}$  la configuration de la carte excepté le site  $s$ .

Ainsi, le degré de pollution d'un site ne dépend que de celui de ses voisins. Bien que dans le cas actuel, il n'existe aucun moyen de prédire le degré de pollution d'un site en se limitant à ce principe, on verra par la suite que le champ de Potts permet de contrôler l'influence d'un site sur ses voisins et inversement.

### 2.1.3 Champs de Gibbs

Au préalable, on définit la mesure de Gibbs [3]

**Définition 3.** *Mesure de Gibbs*

La mesure de Gibbs de fonction d'énergie (ou d'Hamiltonien)  $U : \Omega \rightarrow \mathbb{R}$  est la probabilité  $P$  définie sur  $\Omega$  par :

$$P(X = x) = \frac{1}{Z} \exp(-U(x)),$$

avec

$$U(x) = \sum_{c \in \mathcal{C}} U_c(x),$$

où  $\mathcal{C}$  est le système de cliques associé au système de voisinage  $\mathcal{V}$  de  $U$ .

$$Z = \sum_{x \in \Omega} \exp(-U(x))$$

est une constante de normalisation appelée fonction de partition de Gibbs (en pratique,  $Z$  est incalculable en raison du grand nombre de configurations possibles). Par abus de notation, on a en réalité  $U_c(x) = U_c(x_j, j \in c)$ .

**Définition 4.** *Champ de Gibbs de potentiel associé au système de voisinage  $\mathcal{V}$*

Il s'agit d'un champ aléatoire  $X$  dont la probabilité est une mesure de Gibbs associée au système de voisinage  $\mathcal{V}$ , ce qui implique :

$$P(X = x) = \frac{1}{Z} \exp(-U(x)) = \frac{1}{Z} \exp\left(-\sum_{c \in \mathcal{C}} U_c(x)\right).$$



En d'autres termes, l'énergie globale d'un champ de Gibbs se décompose sous forme d'une somme d'énergies locales qui vont permettre par la suite d'accéder aux probabilités conditionnelles locales. Il apparaît que plus une configuration d'un champ de Gibbs possède une énergie faible et plus elle est probable.

On introduit un théorème important établissant une relation d'équivalence entre un champ de Markov et un champ de Gibbs sous certaines hypothèses.

**Théorème 1.** *Théorème de Hammersley-Clifford*

Soit  $\mathcal{M}$  fini ou dénombrable muni d'un système de voisinage  $\mathcal{V}$  borné, et soit  $E$  un espace d'états discrets, alors on a équivalence entre les affirmations suivantes :

1.  $X$  est un champ de Markov relativement à  $\mathcal{V}$  et  $P(X = x) > 0, \quad \forall x \in \Omega,$
2.  $X$  est un champ de Gibbs de potentiel associé à  $\mathcal{V}$ .

L'intérêt du théorème de Hammersley-Clifford est qu'il permet de lier les probabilités globales et locales comme suit :

$$P(X_s = x_s | X^s = x^s) = \frac{P(X_s = x_s)}{P(X^s = x^s)} = \frac{\exp(-U(x_s, x^s))}{\sum_{x_t \in E} \exp(-U(x_t, x^s))},$$

qui peut s'exprimer

$$P(X_s = x_s | X^s = x^s) = \frac{\exp(-U_s(x_s | V_s))}{\sum_{x_t \in E} \exp(-U_t(x_t | V_t))},$$

où  $V_s = (x_t, t \in \mathcal{V}_s)$ .

**Remarque** Le théorème de Hammersley-Clifford n'est valable que lorsque toute configuration est possible :

$$P(X = x) > 0, \quad \forall x.$$

### 2.1.4 Echantillonneur de Gibbs

Au vu de ce qui a été défini jusqu'à présent, la question est de déterminer la façon de tirer une configuration, étant donné un champ de Markov, en suivant la loi de probabilité de Gibbs caractéristique de ce champ. De nombreux algorithmes réalisant ces tirages existent, cependant, ici, on ne s'en tiendra qu'à l'échantillonneur de Gibbs.

Cet algorithme (Geman & Geman [2]) est une méthode itérative de construction de cartes. Partant d'une carte quelconque et au terme d'un nombre suffisant d'itérations, la méthode converge et retourne une carte dont les réalisations sont tirées suivant la loi de Gibbs globale.

L'algorithme est le suivant [3] :

- à l'étape  $n$ , choix d'un site  $s$ ,
- au site  $s$ , selon la configuration des voisins  $V_s$  pour la carte  $x^{(n-1)}$ , calcul de la probabilité conditionnelle locale :

$$P(X_s = x_s | V_s) = \frac{\exp(-U_s(x_s | V_s))}{\sum_{x_t \in E} \exp(-U_t(x_t | V_t))},$$

► Mise à jour du site  $s$  par tirage aléatoire selon la loi  $P(X_s = x_s | V_s)$ .

On considère que l'algorithme a convergé au bout d'un grand nombre d'itérations ou lorsque le nombre de changement est faible. Le choix du site  $s$  peut se faire de n'importe quelle façon mais la contrainte est que chaque site doit pouvoir être balayé plusieurs fois. Dans le cas présent on va imposer un balayage classique ligne par ligne de la carte (une autre possibilité aurait été de tirer un site selon une loi uniforme par exemple).

Cet algorithme construit une suite de cartes  $x^{(n)}$  correspondant aux observations d'une suite  $X^{(n)}$  de champs aléatoires constituant une chaîne de Markov pour un certain noyau de transition. Il est possible de prouver le théorème suivant (voir [3] pour les détails) :

$$\forall x^{(0)}, \forall x \in \Omega, \quad \lim_{n \rightarrow \infty} P(X^{(n)} = x | X^{(0)} = x^{(0)}) = P(x),$$

où  $P$  correspond à la mesure de Gibbs associée au champ de Markov considéré. Pour un nombre suffisamment large d'itérations, les cartes  $x^{(n)}$  sont donc des réalisations de la loi globale  $P(x)$ , indépendamment de la configuration initiale  $x^{(0)}$ .

### 2.1.5 Modèle de Potts

On se propose de présenter un des champs de Markov les plus utilisés : le champ de Potts. Ce dernier s'obtient par tirage de réalisations selon l'échantillonneur de Gibbs.

Pour commencer, on peut préciser que le modèle de Potts peut être vu comme une extension du modèle d'Ising ([1], [3]) initialement développé pour l'étude du ferromagnétisme en physique statistique, pour lequel on se place dans un domaine binaire  $E = \{-1, 1\}$  à la différence du modèle de Potts où celui-ci est décrit pour un espace  $m$ -aire. On décrira le modèle en l'appliquant à notre cas. Ainsi, l'espace des descripteurs équivalent à l'ensemble des degrés de pollution  $E$  est donné par

$$E = \{1, 2, 3, 4\}.$$

Le voisinage est composé des 4 ou 8 plus proches voisins. On définit les potentiels en tout ou rien seulement pour les cliques d'ordre 2 :

$$U_s(x_s | V_s) = -B \times x_s - \beta \sum_{t \in V_s} 1_{x_s = x_t}, \quad (1)$$

où  $\beta$  est la constante de couplage entre voisins et  $B$  représente un champ magnétique externe ou bien caractérise les interactions énergétiques externes du système. Pour des valeurs de  $\beta$  faible, la configuration finale sera aléatoire. En d'autres termes, les cartes que l'on génère seront totalement aléatoires et présenteront une distribution des degrés de pollution hétérogène. A l'opposé, de grandes valeurs de  $\beta$  favoriseront l'apparition de régions homogènes à un degré de pollution dont la valeur ne peut être prédite à l'avance ; en ce sens, chaque degré pourra être majoritaire sur les autres et inversement.

### 2.1.6 Cas d'étude

Par la suite, le cas traité est basé sur la carte en Figure 1.

Partant d'un domaine carré de taille  $12 \times 12$  cases, on dispose de 9 données de sondage et l'ensemble des degrés de pollution  $E$  est donné par

$$E = \{1, 2, 3, 4\},$$

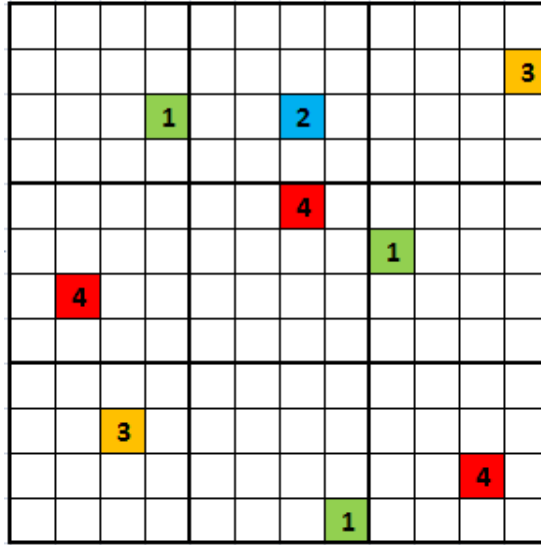


FIGURE 1 – Cas d'étude initial

D'un point de vue numérique, il est nécessaire d'imposer à l'échantillonneur de ne pas faire de tirages aléatoires pour les sites imposés (ceux où l'on dispose d'une donnée de sondage). Dans un premier temps, on pourra mettre en évidence l'influence du paramètre  $\beta$  dans le modèle de Potts. Pour ce faire, on génère des cartes pour différentes valeurs de  $\beta$

$$\beta \in \{0, 0.1, 0.3, 0.5\},$$

et on maintient  $B = 0$ .

Les réalisations en Figure 2 confirment que plus la valeur de  $\beta$  croît et plus les régions pour un degré de pollution sont homogènes.

On se propose de générer plusieurs cartes et d'afficher la carte de fréquence d'apparition de chaque degré de pollution pour tous les sites de la carte (Figure 3). Comme déjà observé, on constate que globalement,  $\beta$  influence la répartition des degrés de pollution ; les grandes valeurs de  $\beta$  privilégient l'apparition des valeurs extrêmes (1 et 4). De plus, ces figures permettent de mettre en évidence le fait que les sites influencent fortement leurs voisins, au travers des régions qui se forment autour des sites à valeurs imposées, dans lesquelles, les fréquences d'apparition de ces dernières sont importantes.

### 3 Evaluation des coûts

A partir des cartes générées suivant le modèle de Potts, il est possible d'évaluer, pour un maillage donné, le coût associé à chaque carte, ainsi que le coût moyen et la variabilité pour une valeur de  $\beta$  et un maillage fixés. Pour ces deux quantités, on fera varier la valeur de  $\beta$  ainsi que les maillages de terrassement.

#### 3.1 Calcul des coûts

On présente la méthode de calcul adoptée dans cette étude. On définit une matrice de coût  $A \in \mathbb{R}^{4 \times 4}$  telle que  $A_{xy}$  correspond au coût engendré lorsqu'un camion est envoyé initialement

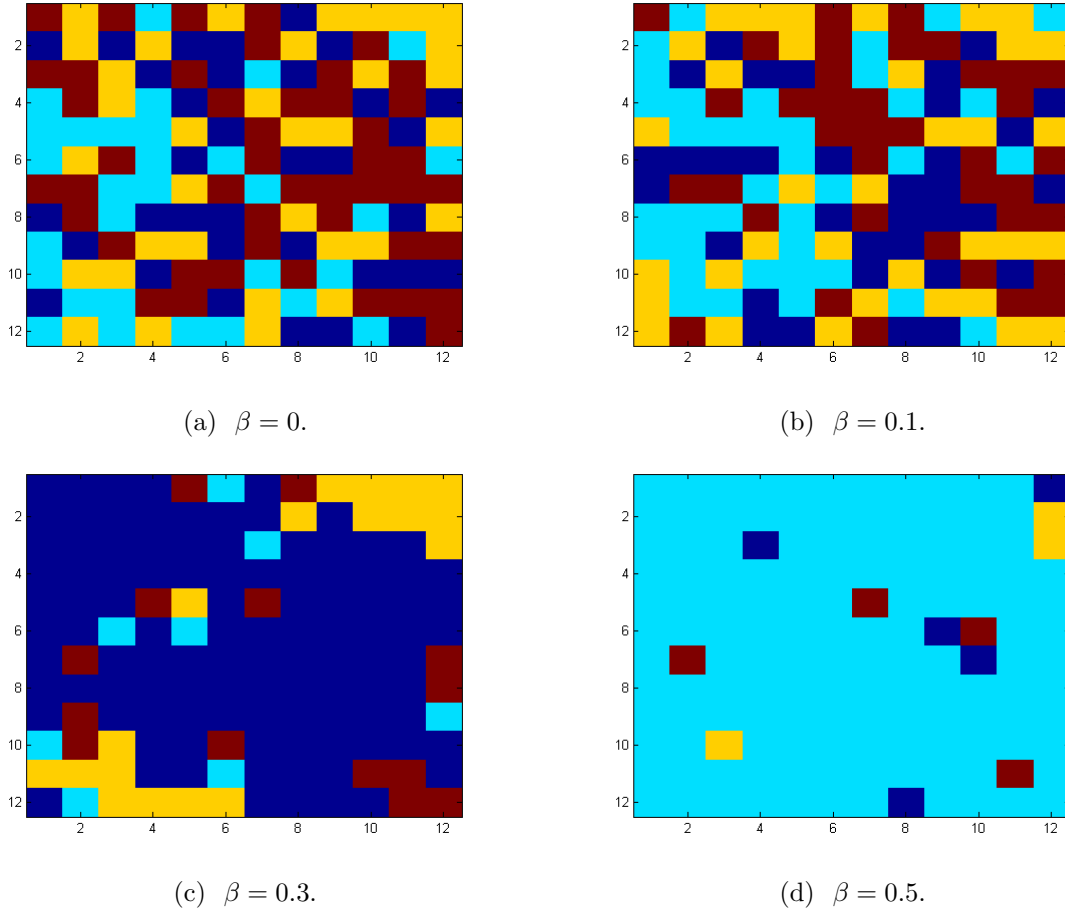


FIGURE 2 – Cartes de la répartition des degrés de pollution générées selon le modèle de Potts (bleu :1, cyan :2, jaune :3, rouge :4).

au centre de traitement  $x$  puis renvoyé au centre de traitement  $y$ . Il est évident que le renvoi ne se fait pas lorsque  $x = y$ . En d'autres termes, les coûts peuvent se résumer comme suit :

- si  $x \geq y$ , le coût correspond au coût de traitement des déchets de classe  $x$ ,  $A_{xy} = A_{xx}$ ,
- si  $x < y$ , le coût correspond au coût de traitement des déchets de classe  $y$  complété par les frais additionnels (amendes, frais juridiques, ...).

On rappelle que l'on travaille sur un domaine carré de taille  $N \times N$  cases. On note  $M$  un maillage de terrassement de taille variable,  $x \in \mathbb{N}^{N \times N}$  la matrice des classes estimées telle que  $x_{ij}$  correspond à la classe estimée par le maître d'ouvrage pour le site  $(i, j)$ , et  $y \in \mathbb{N}^{N \times N}$  la matrice des classes générées telle que  $y_{ij}$  correspond à la classe du site  $(i, j)$  pour une réalisation d'une carte suivant le modèle de Potts. Il vient que le coût estimé  $C_e$  est donné par

$$C_e = \sum_{(i,j)=(1,1)}^{(N,N)} A_{x_{ij}x_{ij}},$$

et le coût réel final  $C_f$  en se référant aux données de la carte aléatoire s'écrit

$$C_f = \sum_{(i,j)=(1,1)}^{(N,N)} A_{x_{ij}y_{ij}} \quad (\text{car } A_{x_{ij}y_{ij}} = A_{x_{ij}, \max(x_{ij}, y_{ij})})$$

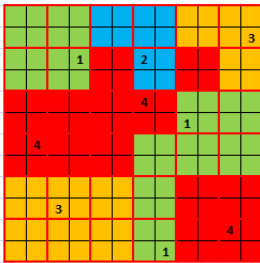
Afin d'évaluer la variabilité du coût final pour un maillage donné, on propose la méthode suivante :

- on génère un nombre fini de cartes  $N_c$ ,
- on calcule le coût final  $C_f$  associé aux différentes cartes générées pour le maillage considéré ; la variable aléatoire correspond au coût final et chaque réalisation est obtenue lorsqu'une carte est générée. Ce faisant, on obtient un échantillon de variables aléatoires de taille  $N_c$ ,
- on peut alors caractériser la loi de distribution du coût final pour un maillage donné en déterminant l'espérance  $\mathbb{E}$  et l'écart type  $\sigma$  (ou la variance) de ce coût. On caractérise également cette loi par la fonction de densité (PDF) pour le coût final.

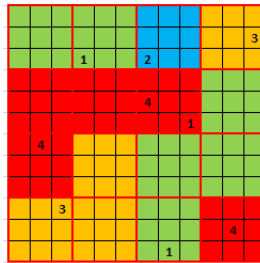
L'intérêt de cette démarche réside dans le fait que l'on va appliquer cette méthode à différents maillages ; celle-ci permettra de comparer les maillages entre eux.

### 3.2 Application numérique

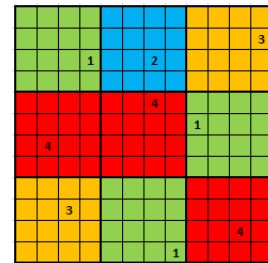
Certains maillages  $M$  considérés sont donnés en Figure 4 ; on traitera les trois maillages présentés en Figure 4 respectivement de taille de maille unitaire  $2 \times 2$ ,  $3 \times 3$  et  $4 \times 4$  cases ainsi qu'un maillage grossier de maille unitaire  $6 \times 6$  cases. On s'intéresse à la variabilité des coûts  $C_f$  liée au paramètre de taille des mailles.



(a) Maillage de maille unitaire  $2 \times 2$ .



(b) Maillage de maille unitaire  $3 \times 3$ .



(c) Maillage de maille unitaire  $4 \times 4$ .

FIGURE 4 – Maillages de taille différente.

Il est à noter que les coûts que l'on calcule ici ne sont pas contraints par le volume des camions. On fait l'hypothèse que le volume d'une case correspond exactement au volume d'un camion. De même, l'ordre de terrassement des mailles n'intervient pas dans le calcul. A titre indicatif, on travaille avec la matrice de coûts  $A$  suivante

$$A = \begin{pmatrix} 1 & 6 & 31 & 151 \\ 5 & 5 & 40 & 160 \\ 30 & 30 & 30 & 190 \\ 150 & 150 & 150 & 150 \end{pmatrix}.$$

Dans un premier temps, on s'intéressera à l'influence du paramètre  $\beta$  du modèle de Potts sur le coût final  $C_f$  relatif à un maillage. Dans un second temps, on évaluera la variabilité des coûts en fonction de la taille de maille pour une constante de couplage entre sites voisins  $\beta$  fixée.

Les résultats pour le premier cas sont présentés dans le Tableau 1. Dans ce cas test, on travaille avec un maillage de maille unitaire  $4 \times 4$  cases en imposant  $\beta = 0.1$  et on génère 100 cartes. Comme on ne traite qu'un seul maillage, le coût estimé  $C_e$  est indépendant de  $\beta$ . On remarque que la moyenne empirique de  $C_f$  pour différents  $\beta$  tend à augmenter à mesure que  $\beta$  croît et donc que les régions de classe de pollution s'homogénéisent. En l'état, il est important de préciser que le résultat est donné à titre indicatif ainsi que les fonctions de densité (PDF) que l'on présente en Figure 5. En effet, le fait de générer 100 cartes est équivalent à générer 100 réalisations de Monte-Carlo. Ce nombre n'est pas significatif compte tenu de la valeur des espérances obtenues et des écarts types, l'intervalle de confiance que l'on peut calculer montre que les espérances sont égales pour  $\beta = 0.1$  et  $\beta = 0.3$  d'une part, et pour  $\beta = 0.5$  et  $\beta = 1$  d'autre part.

	$\beta = 0.1$	$\beta = 0.3$	$\beta = 0.5$	$\beta = 1$
$C_e$	8288	8288	8288	8288
$\mathbb{E}(C_f)$	12292.2	12038	17178.95	18014.1
$\sigma$	640.7543	3651.679	6264.131	5571.262
$IC$ sans $\mathbb{E}(C_f)$	$[-64; +64]$	$[-365; +365]$	$[-626; +626]$	$[-557; +557]$
$\min(C_f)$	10783	8293	8288	8288
$\max(C_f)$	13983	20903	22163	22163

TABLE 1 – Variation du coût du maillage  $4 \times 4$  pour différents  $\beta$ .

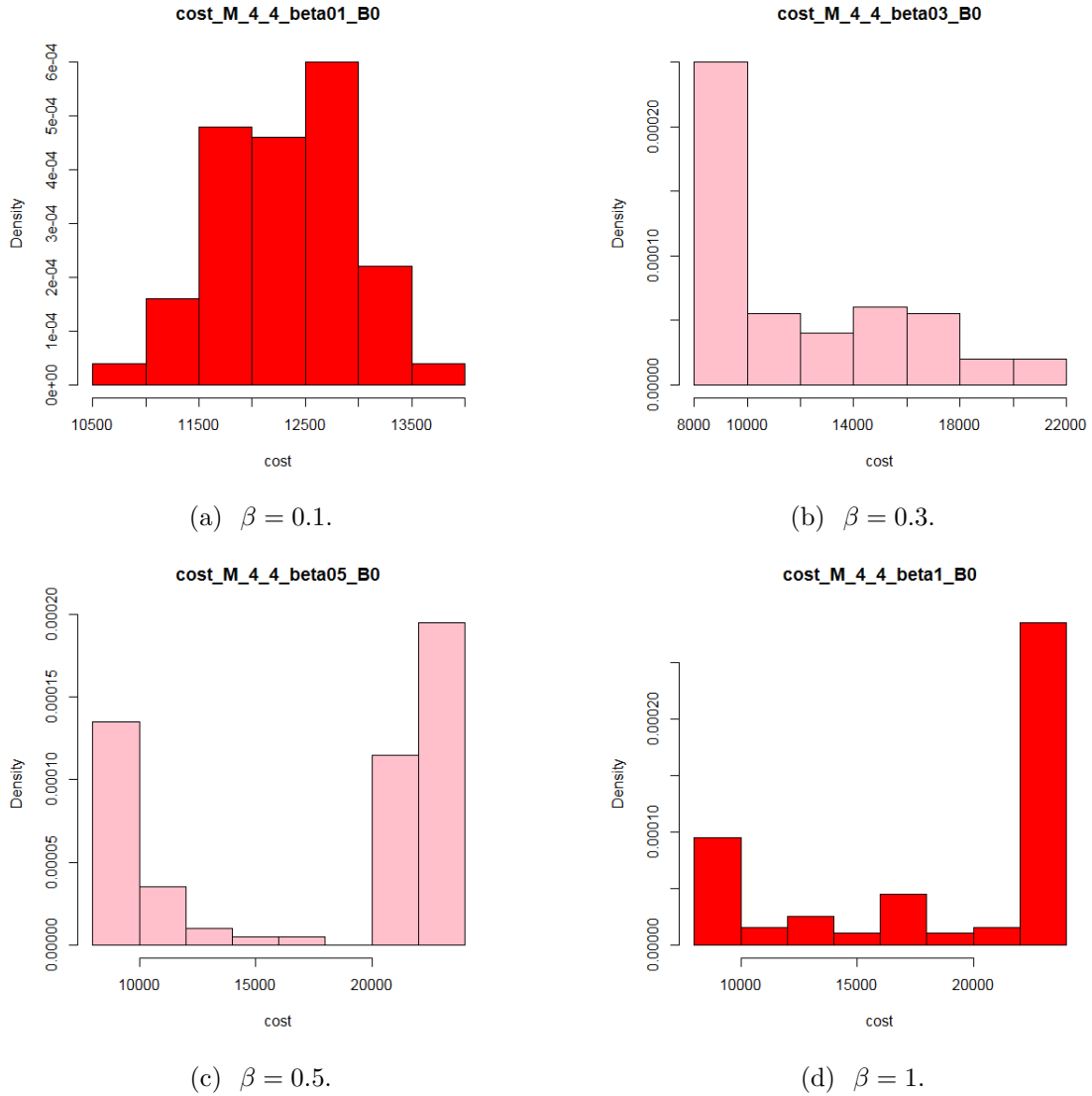


FIGURE 5 – Densité de la loi de  $C_f$  pour un maillage de maille unitaire  $4 \times 4$  cases pour différents  $\beta$ .

Concernant le deuxième cas test, les résultats sont donnés dans le Tableau 2. On génère également 100 cartes auxquelles on va opposer les quatre maillages de taille différente précédemment évoqués. Les coûts estimés  $C_e$  sont directement liés au maillage d'où la variation des valeurs. Il est intéressant de noter que plus la carte sera grossière et plus le coût sera élevé; le coût associé au maillage  $6 \times 6$  est pratiquement égal au double des coûts associés aux autres maillages. L'espérance des coûts est la plus faible pour le maillage  $3 \times 3$  auquel on pourrait adjoindre le maillage  $4 \times 4$  et dans une autre mesure, le maillage  $2 \times 2$  compte tenu des intervalles de confiance. Une nouvelle fois, les fonctions de densité présentées en Figure 6 sont données à titre indicatif. Néanmoins, il est à noter que l'on obtient globalement des lois centrées.

	$M 2 \times 2$	$M 3 \times 3$	$M 4 \times 4$	$M 6 \times 6$
$C_e$	8984	7929	8288	16236
$\mathbb{E}(C_f)$	12752.5	12037.3	12292.2	17913.7
$\sigma$	665.44	721.9444	640.7543	488.8734
$IC$ sans $\mathbb{E}(C_f)$	$[-66; +66]$	$[-72; +72]$	$[-64; +64]$	$[-48; +48]$
$\min(C_f)$	11259	10449	10783	16776
$\max(C_f)$	14369	13794	13983	19011

TABLE 2 – Caractérisation de la variation des coûts pour différentes tailles de maille.

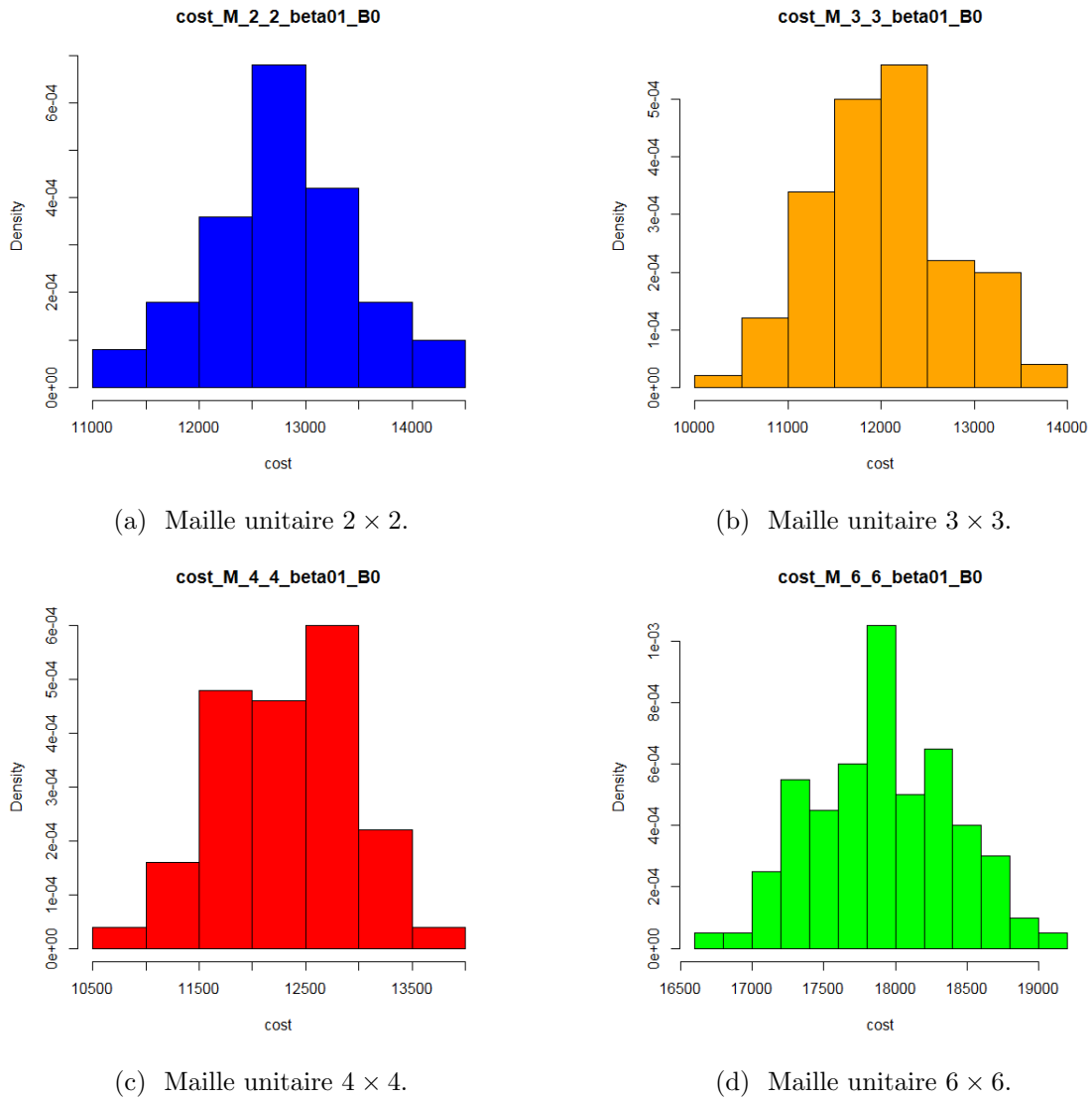


FIGURE 6 – Densité de la loi de  $C_f$  pour différents maillages.



## 4 Conclusion

Au cours de ce travail, on a pu proposer un générateur de cartes relativement paramétrable prenant en compte les données initiales liées au sondage de terrain. Celui-ci a permis de mettre en place une méthode permettant de comparer des maillages qui diffèrent au niveau de leur taille de maille. Il a également été possible de proposer une caractérisation de la distribution du coût final pour un maillage donné. Cependant, il peut être davantage rigoureux d'appliquer cette méthodologie pour un nombre de réalisations plus important en respectant les contraintes liées à l'intervalle de confiance.

Par la suite, il peut s'avérer intéressant d'envisager une étude statistique pour l'évaluation des paramètres liés au générateur de maillages possibles. Cela permet entre autres d'évaluer ces paramètres au cas par cas pour chaque site de terrassement voire même de faire varier ces paramètres d'une région d'un même site à l'autre. Il est également possible d'estimer ces paramètres en fonction des informations sur le site à terrasser (son histoire, d'éventuelles photographies,...).

Un autre aspect plus ambitieux consiste à prendre en compte les contraintes liées aux camions et à l'aménagement du terrain, mais également les paramètres géométriques directement liés à la construction du maillage tels que son orientation ou la forme des mailles. Afin de gérer ces derniers paramètres, il est possible de songer à la construction d'un échantillon de cartes aléatoires avec une résolution élevée. En effet, puisque les évaluations de coût entre différents maillages doivent être effectuées sur un même échantillon, la taille des cellules des cartes générées doit être suffisamment petite pour que le décalage induit par le changement d'orientation, de forme ou d'origine, n'entraîne pas d'erreurs trop importantes dans les calculs.

En outre, il reste à utiliser une méthode d'optimisation pour obtenir un maillage de coût minimal. La principale difficulté est que la fonctionnelle à minimiser (par exemple, le risque de se tromper pour un maillage donné) n'est pas une fonction continue de ses paramètres, soit les caractéristiques du maillage (la taille de maille par exemple). Il faudra donc envisager des méthodes de type optimisation aléatoire.

L'étude présentée a été réalisée en 2 dimensions d'espace sur maillage cartésien or la réalité du terrain imposerait une réflexion en 3 dimensions sur des maillages plus complexes. En conservant des hypothèses analogues à celles imposées en dimension 2, l'extension de la méthode est envisageable. Se poserait alors la question du nombre d'échantillons dont on disposerait mais également du paramétrage des potentiels qui constituerait la pierre angulaire du travail. Pour ce dernier, une étude des corrélations entre les données obtenues sur des précédents cas et les données physiques de terrain pourrait permettre de distinguer des premiers éléments d'étude et éventuellement de construire un nouveau modèle physique plus adapté à la réalité des terrains pollués. L'autre aspect plus géométrique concerne les irrégularités des mailles pour tenir compte des contraintes liées au terrain. Cette problématique nécessiterait de modifier la fonction de calcul de coût d'une part, mais également la manière de générer chaque carte aléatoirement en vertu des contraintes imposées.

## Références

- [1] E. Ising. *Beitrag zur theorie des ferromagnetisms*. Zeitschrift fur Physik, 31, 253–258, 1925.

- [2] S. Geman et D. Geman. *Stochastic Relaxation, Gibbs Distribution, and the Bayesian Restoration of Images*. IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-6(6), 721–741, 1984.
- [3] M. Sigelle et F. Tupin. *Champ de Markov en traitement d'image*. Telecom Paris, Module C3M, support de cours, 1999.
- [4] P. Suire, S. Belbèze et V. Vallon. *Evaluation du choix d'un maillage de terrassement sur l'impact des filières de gestion des sols pollués : aspect "développement durable et économique"*. AnteaGroup, Note méthodologique préliminaire, 2013.

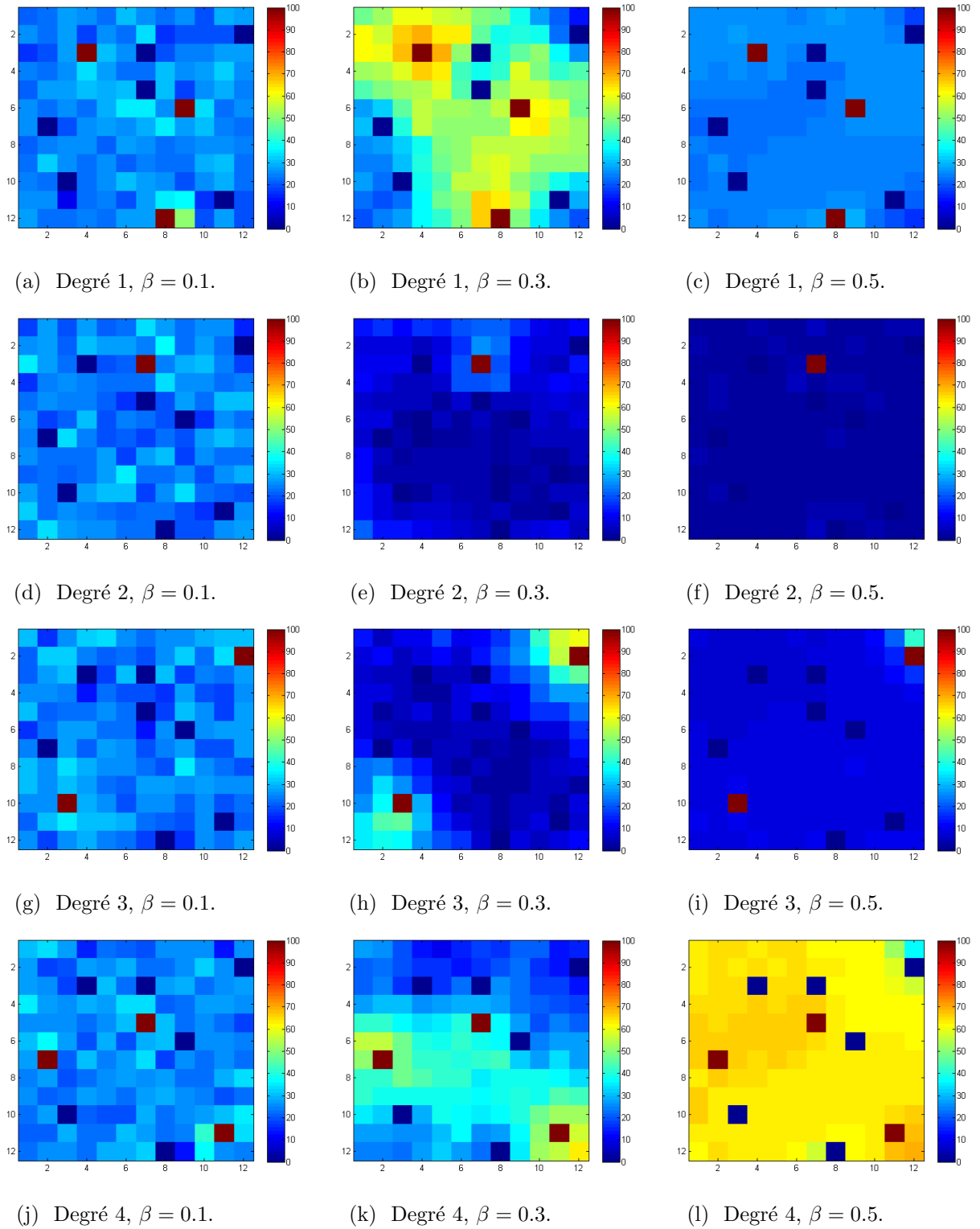


FIGURE 3 – Cartes des fréquences d'apparitions des degrés de pollution.