
Ségrégation de Schelling

Antonin Duroseau
Arthur Danjou
Valentin Grohens

Table des matières

Introduction	2
1 Explication du modèle	2
1.1 Description du modèle	2
1.2 Paramètres	2
2 Simulation et mesures	3
2.1 Explication de la mesure de ségrégation	3
2.2 Mesure de ségrégation pour plusieurs T_{max}	4
2.3 Mesure de ségrégation pour plusieurs L et plusieurs ρ et transition de phase	5
3 Variante	5
3.1 Présentation des changements	5
3.2 Energie du modèle et son évolution	5
Conclusion	6
Annexes	7
Simulation des valeurs significatives de L	7
Simulation de la modélisation	7
Simulation de l'algorithme de clusters	7
Simulation de la modélisation pour différentes valeurs de ρ et de L	7
Evolution du coefficient de ségrégation en fonction de L pour différente valeur de ρ	9

Introduction

Le modèle de ségrégation de Schelling est un modèle mathématique qui permet d'expliquer la répartition d'agents au sein d'une population selon leur préférence vis à vis des agents autour d'eux. Ses applications sont diverses, par exemple en économie pour justifier la séparation entre deux groupes de population (financière, ethnique, sociale, etc...); en physique dans le modèle de Ising décrivant les moments dipolaires magnétiques de spin atomique; ou encore en théorie du jeu qui est une variante du modèle de Schelling.

Nous avons décidé d'étudier ce sujet car nous souhaitons poursuivre nos études dans des masters en mathématiques appliquées aux statistiques. Ce projet est donc pour nous une ouverture au programme de master mais aussi un avant goût de sujet de recherche que nous pourrions entreprendre dans la suite de nos études.

Nous allons commencer par décrire le modèle et ses paramètres. Puis dans un second temps, nous allons expliquer la mesure de ségrégation, sa convergence et ce qu'on appelle les transition de phase. Enfin, nous analyserons une variante de ce modèle et introduirons une nouvelle donnée, l'énergie du modèle.

1 Explication du modèle

Dans cette première section, nous explorons en détail les fondements de notre modèle de ségrégation, en fournissant une description complète de ses mécanismes et de ses paramètres clés. Comprendre ces éléments est crucial pour appréhender la manière dont notre simulation fonctionne et comment les agents interagissent dans leur environnement.

1.1 Description du modèle

Nous commençons par une explication détaillée du modèle lui-même, en décrivant comment les agents sont disposés sur une grille spatiale, comment ils interagissent entre eux et comment leurs mouvements influencent la dynamique globale de la ségrégation. Notre modèle prend place sur un "échiquier" de taille $M \times M$, avec $M \geq 1$. On appelle agent, les individus sur les cases occupées. Il y a deux groupes d'agents représentés par deux couleurs, le bleu et le rouge. Les cases vides sont blanches et la proportion totale de case vide est ρ . La répartition des agents se fait selon une loi de probabilité que l'on définira dans la partie paramètres. On notera N_{tot} le nombre total de cases occupées. On appelle voisin une des huit cases adjacentes à notre agent, cf Figure 1. On définit un cluster comme un regroupement d'un ou plusieurs agents d'un même groupe, cf Figure 2. Pour qu'une case soit dite "satisfaisante" il faut que le ratio de ses cases voisines du même groupe soit supérieur à L , la tolérance que l'on impose au début de la modélisation. Si ce ratio est supérieur à L , on pose la satisfaction égale à 1, sinon la satisfaction est égale à 0. On arrête la modélisation soit après un temps max T_{max} soit lorsque tous les agents ont leur satisfaction égale à 1.

1.2 Paramètres

Nous examinons ensuite les différents paramètres qui influent sur le comportement des agents et sur la configuration spatiale résultante. Cela inclut notamment les seuils de tolérance et le taux de répartition de cases vides. Tout d'abord le paramètre ρ définit la proportion de cases vides (indexé par 0) de l'échiquier, on a alors $\mathbb{P}(A_{i,j} = 0) = \rho$ avec $A_{i,j}$ la case de la i -ème ligne et de la j -ème colonne de notre échiquier. On définit avec ce paramètre la probabilité d'être une case du groupe bleu (indexé par -1) ou du groupe rouge (indexé par 1) par :

$$\mathbb{P}(A_{i,j} = \pm 1) = \frac{1 - \rho}{2}$$

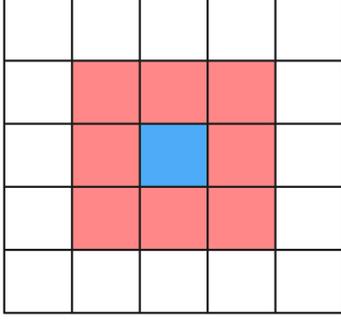


FIGURE 1 – Les voisins de la case bleue sont les huit cases rouges adjacentes

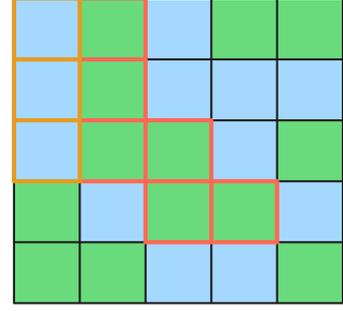


FIGURE 2 – Exemple de la définition d'un cluster. Ici, nous avons surligné deux clusters différents (orange et rose)

On constate que N_{tot} suit une loi de Binomiale $(M^2, 1 - \rho)$:

$$i.e. \mathbb{P}(N_{tot} = k) = \binom{M^2}{k} (1 - \rho)^k \rho^{M^2 - k}$$

On a donc

$$\mathbb{E}(N_{tot}) = M^2(1 - \rho) \text{ et } Var(N_{tot}) = M^2\rho(1 - \rho)$$

D'où,

$$\forall M \in \mathbb{R}, \frac{\mathbb{E}(N_{tot})}{Var(N_{tot})} = \frac{1}{\rho}$$

donc

$$\lim_{M \rightarrow +\infty} \frac{\mathbb{E}(N_{tot})}{Var(N_{tot})} = \frac{1}{\rho}$$

Le second paramètre que l'on détaille est la tolérance L . Ce paramètre nous sert à définir la satisfaction d'un agent. En effet comme dit plus tôt, si la fraction d'agent du même parmi les voisins d'un agent est supérieure à cette tolérance alors la case est dite satisfaisante et la satisfaction de l'agent sera égale à 1, sinon elle est dite insatisfaisante et la satisfaction de l'agent sera égale à 0. On a $L \in [0; 1]$, mais dans les faits si $L = 1$ ou $L = 0$ le modèle ne sera pas très parlant car dans le premier cas seules les cases entourées complètement d'agent du même groupe seront satisfaisantes et dans le second cas toutes les cases seront satisfaisantes; ce sont les valeurs significatives de L car ces deux valeurs changent grandement le comportement du modèle. Nous prendrons donc $L \in [0.2; 0.8]$. Nous avons simulé le modèle pour les deux valeurs significatives de L , soit $L \in \{0.05; 0.95\}$ (les valeurs 0 et 1 étant trop longues à simuler), cf Annexe 1.

2 Simulation et mesures

Dans cette section, nous plongeons dans les détails de notre processus de simulation ainsi que dans les différentes mesures utilisées pour évaluer les résultats obtenus. En explorant la dynamique de la ségrégation dans notre modèle, nous examinons comment les paramètres influent sur la répartition spatiale des agents. Nous pouvons illustrer nos simulations par l'Annexe 2.

2.1 Explication de la mesure de ségrégation

Nous commençons par définir en détail la mesure de ségrégation S utilisée pour évaluer la répartition des agents dans notre simulation. Cette mesure nous permet de quantifier le niveau de séparation entre les groupes homogènes au sein de notre population. On note C l'ensemble des clusters. Pour tout cluster

$c \in C$, on définit la masse n_c d'un cluster comme le nombre de points qu'il contient. On a bien sûr que $\sum_{c \in C} n_c = N_{tot}$ le nombre total de case occupées. On définit le coefficient de ségrégation $S \in [0; 1]$ comme la moyenne pondérée des poids des clusters, renormalisée. Pour une configuration donnée, avec $p_c = n_c/N_{tot}$, soit

$$S = \frac{2}{N_{tot}} \sum_{c \in C} n_c p_c = \frac{2}{N_{tot}^2} \sum_{c \in C} n_c^2$$

Nous avons codé un algorithme permettant de calculer tous les clusters de notre grille afin de déterminer le nombre total de clusters. Par exemple, l'Annexe 3 présente une configuration initiale aléatoire de notre grille et une configuration finale après $T_{max} = 50$, notre algorithme retourne bien 6 clusters.

2.2 Mesure de ségrégation pour plusieurs T_{max}

En analysant la mesure de ségrégation pour différentes valeurs de T_{max} , nous étudions comment la dynamique de ségrégation évolue au fil du temps et si des tendances significatives émergent. Pour représenter l'évolution du coefficient de ségrégation en fonction de $T = \{1; 10; 100; 400\}$ pour une grille de taille $M = 50$, $L = \frac{1}{2}$ et $\rho = 5\%$, nous avons généré la grille initiale puis avons lancé l'algorithme permettant de simuler le modèle. A chaque instant T , nous avons calculé le coefficient de ségrégation S . Nous avons donc obtenu le graphique suivant (Figure 3). Nous remarquons que plus T augmente, plus le coefficient de ségrégation S augmente et tend vers 1. En effet, on constate que

$$\lim_{T_{max} \rightarrow +\infty} C = 2 \quad \lim_{T_{max} \rightarrow +\infty} c = \frac{N_{tot}}{2}$$

D'où

$$\lim_{T_{max} \rightarrow +\infty} S = \frac{2}{N_{tot}} \sum_{c=1}^2 \left(\frac{N_{tot}}{2}\right)^2 = \frac{2}{N_{tot}} \times 2 \times \frac{N_{tot}^2}{4} = 1$$

On a donc bien $S \in [0, 1]$ ce qui justifie de la renormalisation de S

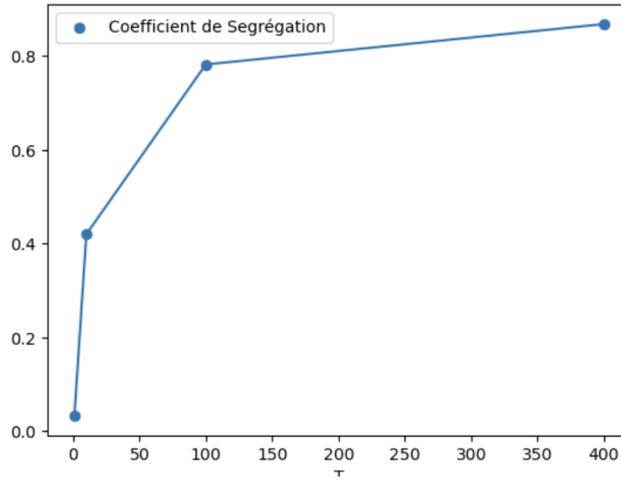


FIGURE 3 – Evolution du coefficient de ségrégation S en fonction de T_{max} pour une même configuration de grille donnée

2.3 Mesure de ségrégation pour plusieurs L et plusieurs ρ et transition de phase

En examinant la mesure de ségrégation pour différentes tolérance L et différentes densités de population ρ , nous explorons comment ces facteurs influent sur la formation de clusters et sur la ségrégation observée. Nous identifions également les valeurs critiques de L pour lesquelles la mesure de ségrégation S varie rapidement, indiquant ainsi une transition de phase dans la dynamique du modèle. Dans le modèle de ségrégation de Schelling, une transition de phase fait référence à un changement brusque ou à une transformation soudaine dans la répartition spatiale des agents. Plus précisément, cela se produit lorsque les agents se déplacent en réponse à des règles simples de préférence de voisinage. Une transition de phase se produit lorsque, à un certain point dans le temps ou après un certain nombre d'itérations, il y a un changement soudain dans la répartition spatiale des agents. Cela peut se traduire par une séparation nette des groupes homogènes dans différentes régions de la grille. La transition de phase est souvent associée à des concentrations croissantes d'agents similaires les uns aux autres, avec une diminution simultanée du nombre d'agents entourés de voisins différents.

Dans l'Annexe 4, nous avons généré différentes simulations pour différentes valeurs de L et de ρ . Globalement, nous voyons que le nombre final de clusters tend vers 2 ou 3 maximum, peu importe la répartition de cases vides ou le taux de tolérance. Dans les cas avec 3 clusters finaux, le temps final T_{max} n'est pas assez élevé pour voir l'apparition de deux clusters finaux, comme annoncé dans la partie précédente. Nous avons représenté l'évolution du coefficient de ségrégation S en fonction de L pour différentes valeurs de $\rho \in \{2\%; 6\%; 12\%; 18\%; 25\%\}$, cf Annexe 5. Les cinq graphiques confirment que L doit bien être compris entre 0.2 et 0.8 pour que la simulation soit intéressante à analyser. Globalement, la transition de phase se trouve au maximum de la fonction représentée, soit lorsque $S = 1$.

3 Variante

Dans cette section, nous abordons une variante de notre modèle de ségrégation, explorant les modifications apportées aux mécanismes fondamentaux ainsi que les conséquences de ces changements sur la dynamique globale. Cette variante nous permet d'approfondir notre compréhension du phénomène de ségrégation en examinant comment des ajustements spécifiques aux règles de déplacement des agents peuvent influencer les résultats.

3.1 Présentation des changements

Nous commençons par décrire en détail les modifications apportées au modèle original, en mettant en évidence les aspects spécifiques qui diffèrent de la version de base. Comprendre ces changements est essentiel pour évaluer leur impact sur la dynamique de ségrégation observée. Dans la variante que nous avons choisit, le mouvement des agents est déterminé exclusivement par leur niveau de satisfaction, créant ainsi une dynamique singulière. Seuls les agents qui sont insatisfaits sont actifs dans ce système. Pour être plus précis, nous considérons tous les agents du groupe k (où $k \in \{-1; 1\}$) qui expriment un sentiment de mécontentement, ainsi que l'ensemble des cases qui seraient satisfaisantes pour les agents du même groupe. À chaque instant t , un agent est sélectionné de manière uniforme parmi ceux qui sont insatisfaits, et il se déplace alors vers une case aléatoire parmi celles qui lui apporteraient satisfaction. Cependant, si aucune case de ce type n'est disponible, l'agent se déplace simplement vers une case libre au hasard. Ce mécanisme introduit une nouvelle dynamique dans le système, où les mouvements sont guidés par la recherche de satisfaction individuelle.

3.2 Energie du modèle et son évolution

Nous introduisons également le concept d'énergie dans le contexte de notre variante de modèle. Cette approche nous permet d'analyser le système du point de vue de la physique statistique, en évaluant comment l'énergie du système évolue au fil du temps en fonction des mouvements des agents. En étudiant l'évolution de l'énergie du système au cours de la simulation, nous explorons comment les agents interagissent pour minimiser leur énergie potentielle, et comment cela se traduit par des schémas de ségrégation spécifiques dans l'environnement. Nous constatons que l'énergie du schémas

évolue en décroissance avec le temps. Cela ce justifie si on regarde 2 instants très proches où juste une case se déplace vers une case vide satisfaisante, on a alors : pour $L = 1/8$

$$(t = 0) \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix} (t = 1)$$

Comme l'agent se déplace vers une case vide, on a besoin de calculer juste ce changement d'énergie, car le reste de la grille ne change pas, et donc seul l'énergie au niveau de l'agent qui se déplace importe. Nous ne calculeront pas l'énergie de tout l'échiquier mais nous pouvons simplement dire que si on soustrait l'ancienne énergie et la nouvelle seul l'énergie au niveau des anciens et nouveaux voisins de notre agent ne se simplifie pas.

$$ES(t = 0) = - \sum_{(k,l) \in V_{ij}} A_{ij} A_{kl} - K \sum_{(i,j) \in 3 \times 3} \sum_{(k,l) \in V_{ij}} A_{ij}^2 A_{kl}^2 = 2(1 - K)$$

$$ES(t = 1) = - \sum_{(k,l) \in V_{ij}} A_{ij} A_{kl} - K \sum_{(i,j) \in 3 \times 3} \sum_{(k,l) \in V_{ij}} A_{ij}^2 A_{kl}^2 = 2(-1 - K)$$

D'où

$$ES(t = 1) - ES(t = 0) = 2(-1 - K) - 2(1 - K) = -4 \leq 0$$

Donc bien décroissant.

Ce que nous donne cette décroissance est que pour un temps suffisamment long, l'énergie du modèle tend vers une situation stable où plus aucun agent ne se déplace, soit car ils sont tous satisfaits soit car il n'y a plus de case vers lesquels ils peuvent se déplacer.

Conclusion

bloblblo

Annexes

Annexe 1 : Simulation des valeurs significatives de L

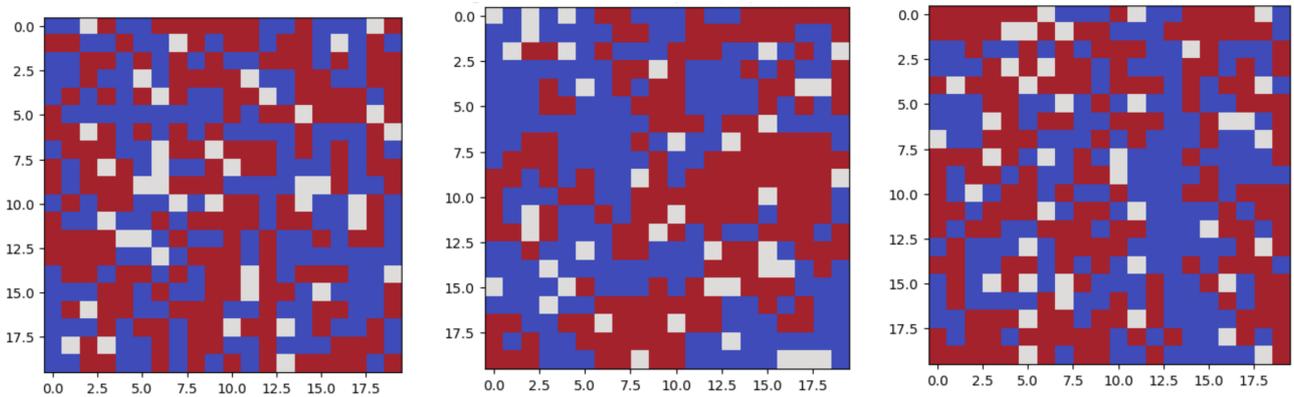


FIGURE 4 – Simulation des valeurs significatives de L pour $T_{max} = 50$ avec $(M, \rho) = (10, 105\%)$ et $L \in \{0.05; 0.95\}$

Annexe 2 : Simulation de la modélisation

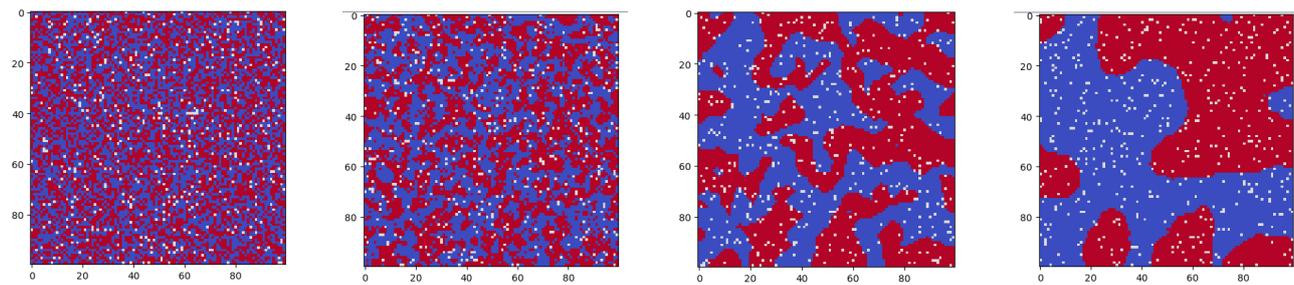


FIGURE 5 – Simulation de la modélisation pour $T_{max} = 1, 10, 100$ avec $(M, \rho, L) = (100, 5\%, \frac{1}{2})$

Annexe 3 : Simulation de l'algorithme de clusters

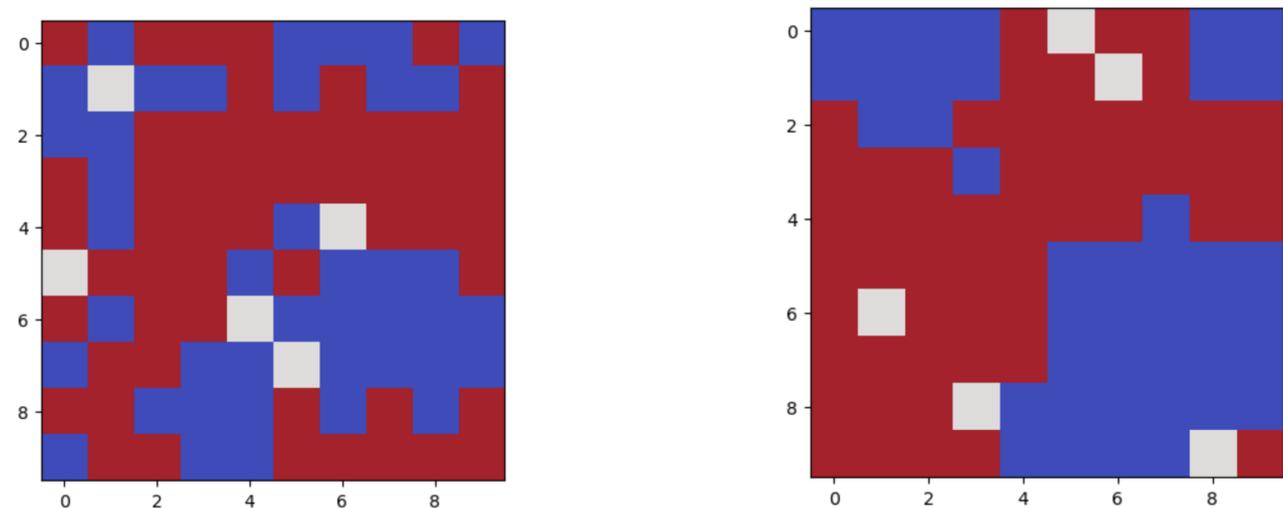
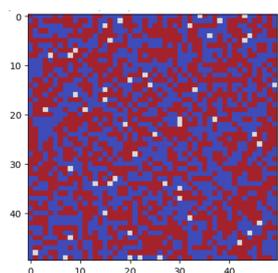
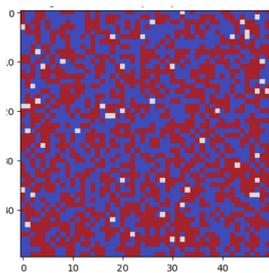


FIGURE 6 – Simulation de l'algorithme de clusters pour $T_{max} = 50$ avec $(M, \rho, L) = (10, 5\%, \frac{2}{7})$

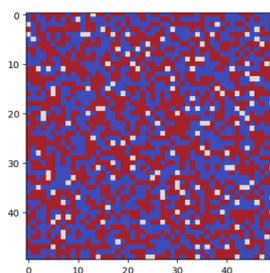
Annexe 4 : Simulation de la modélisation pour différentes valeurs de ρ et de L



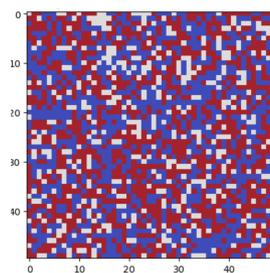
(a) Initial : $(M, \rho, L) = (50, 0.02, 0.33)$



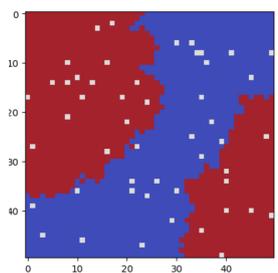
(b) Initial : $(M, \rho, L) = (50, 0.02, 0.75)$



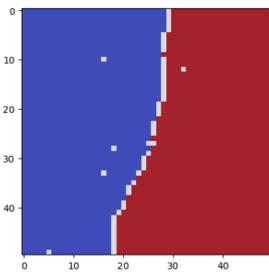
(c) Initial : $(M, \rho, L) = (50, 0.06, 0.71)$



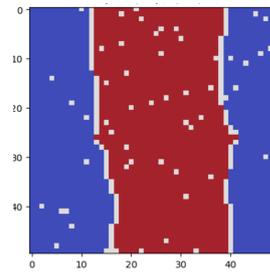
(d) Initial : $(M, \rho, L) = (50, 0.18, 0.75)$



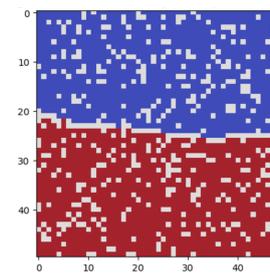
(e) Final : $(M, \rho, L) = (50, 0.02, 0.33)$



(f) Final : $(M, \rho, L) = (50, 0.02, 0.75)$



(g) Final : $(M, \rho, L) = (50, 0.06, 0.71)$



(h) Final : $(M, \rho, L) = (50, 0.18, 0.75)$

FIGURE 7 – Simulation de la modélisation pour $T_{max} = 400$ avec (M, ρ, L) avec $M = 50$, pour différente valeur de ρ et de L

Annexe 5 : Evolution du coefficient de ségrégation en fonction de L pour différente valeur de ρ

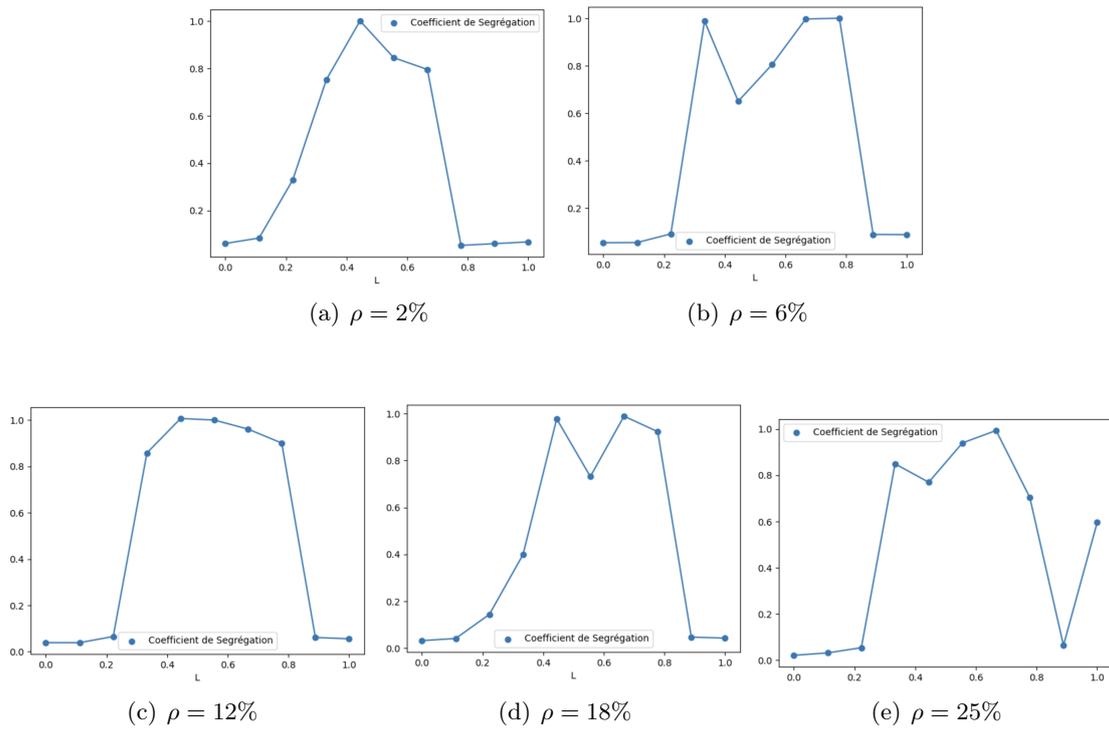


FIGURE 8 – Evolution du coefficient de ségrégation en fonction de L pour différente valeur de $\rho \in \{2\%; 6\%; 12\%; 18\%; 25\%\}$ pour $T_{max} = 100$ et $M = 30$