



**Modélisation, exploration de modèles, théorie de la  
viabilité pour l'aide à la gestion des espèces végétales  
invasives : application au cas des renouées asiatiques le  
long d'un cours d'eau**

François Lavallée

► **To cite this version:**

François Lavallée. Modélisation, exploration de modèles, théorie de la viabilité pour l'aide à la gestion des espèces végétales invasives : application au cas des renouées asiatiques le long d'un cours d'eau. Mathématiques [math]. Sorbonne Université, 2020. Français. NNT : . tel-03836187

**HAL Id: tel-03836187**

**<https://theses.hal.science/tel-03836187>**

Submitted on 1 Nov 2022

**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.

NNT :

# Thèse de doctorat de Sorbonne Université

École doctorale Sciences Mathématiques de Paris Centre

Présentée par

François Lavallée

## Modélisation, exploration de modèles, théorie de la viabilité pour l'aide à la gestion des espèces végétales invasives : application au cas des renouées asiatiques le long d'un cours d'eau

Thèse dirigée par Sophie Martin

préparée au sein du Laboratoire d'Ingénierie pour les Systèmes Complexes (LISC)  
INRAE

Soutenue publiquement à Paris le 13 novembre 2020

devant le jury composé de :

|      |                   |   |                             |                    |
|------|-------------------|---|-----------------------------|--------------------|
| M.   | Philippe Ciffroy  | Chercheur Sénior                                | Électricité de France (EDF) | Examineur          |
| Mme. | Fanny Dommanget   | Ingénieur des ponts, des eaux et forêts         | INRAE                       | Examineur          |
| Mme. | Hélène Frankowska | Directrice de recherche CNRS                    | Sorbonne Université         | Examineur          |
| Mme. | Sophie Martin     | Ingénieur en chef des ponts, des eaux et forêts | INRAE                       | Directeur de thèse |
| M.   | Alain Rapaport    | Directeur de recherche                          | INRAE                       | Rapporteur         |
| M.   | Romain Reuillon   | Chercheur                                       | CNRS                        | Examineur          |
| M.   | Georges Zaccour   | Professeur titulaire                            | HEC Montréal                | Rapporteur         |



Version du 25 octobre 2020





## Remerciements

Mes premiers remerciements s'adressent à Sophie. Un très grand merci pour la confiance que tu m'as accordée en acceptant d'encadrer ma thèse, merci pour ton soutien sans faille durant ces trois années, et merci de m'avoir transmis tant de choses.

Je remercie Alain Rapaport et Georges Zaccour pour avoir accepté de rapporter ma thèse. Je leur suis très reconnaissant pour la lecture attentive de mon manuscrit et leurs conseils.

Je suis très heureux que Philippe Ciffroy, Fanny Dommanget, Hélène Frankowska et Romain Reuillon aient accepté de participer à mon jury de thèse. Merci pour l'intérêt que vous portez à mon travail.

Je remercie la MIP pour m'avoir accompagné et guidé depuis le stage. Merci Sophie pour m'avoir fait découvrir la théorie de la viabilité, merci Fanny d'avoir répondu à mes nombreuses questions sur les renouées, merci Charline de m'avoir encadré sur les processus stochastiques, merci Isabelle pour tes nombreux conseils et merci Bjoern de m'avoir éclairé sur la modélisation. Je vous remercie, ainsi que Laure et Agnès, d'avoir participé à mes comités de thèse. Je remercie EDF et IRSTEA pour avoir co-financé ce travail.

Je remercie également mes co-auteurs et collaborateurs sur les différents projets liés à cette thèse : les membres de la MIP, mais également François Martin, Guillaume Chérel, Thibaud Monnin et Laurent Bolognini.

Mes sincères remerciements vont à Sophie, Charline, Fanny et Isabelle pour leur relecture attentive de ce manuscrit.

Je tiens à remercier l'Institut des Systèmes Complexes pour m'avoir accueilli durant ma thèse. Je remercie les personnes avec qui j'y ai eu le plaisir d'échanger, je garderai un très bon souvenir de ces trois années.

Je remercie les membres de la chaire MMB. J'ai beaucoup apprécié les journées organisées par celle-ci, pour leur qualité scientifique et leur convivialité.

Un grand merci à la communauté d'OpenMOLE pour les nombreux échanges que nous avons eu, notamment lors des coding camps et de la préparation d'eXModelo.

Merci à mes amis, Mélodie, Ilias, Yann, Thibaut et Anne pour leur soutien. Merci Romain pour tes conseils au moment de la rédaction et Amélie pour tes encouragements.

Enfin, merci à ma famille pour son affection et son soutien indéfectible et merci à tous ceux qui m'ont permis d'en arriver là.



## Avant-propos

Cette thèse s'inscrit dans le cadre d'une convention de collaboration d'encadrement entre INRAE et EDF. Plus précisément, cette collaboration regroupe des chercheurs du Laboratoire d'Ingénierie des Systèmes Complexes (LISC) du centre INRAE de Clermont-Ferrand, Isabelle Alvarez et Sophie Martin, du Laboratoire EcoSystèmes et Sociétés En Montagne (LESSEM) du centre INRAE de Grenoble, Fanny Dommanget, Björn Reineking et Charline Smadi, et du Laboratoire National d'Hydraulique et Environnement (LNHE) du centre EDF de Chatou, Laure Santoni.

Pour réaliser cette thèse, j'ai été accueilli à l'Institut des Systèmes Complexes de Paris Île de France (ISC-PIF), au sein duquel j'ai collaboré avec Mathieu Leclaire et Romain Reuillon sur la plateforme OpenMOLE et Guillaume Chérel sur l'exploration de modèles.

J'ai également collaboré avec Mireille Boyer (ingénieur responsable du pôle Ripisylves & Invasives, Concept.Cours.d'EAU SCOP Aquabio) sur les données de dispersion de la renouée le long de l'Hérault et François-Marie Martin (LESSEM) sur les données de croissance locale de la renouée.



## Résumé

La gestion efficace des espèces végétales exotiques envahissantes est un enjeu à différents niveaux : environnemental, économique et de santé publique. De nombreuses études portent sur la gestion optimale des espèces exotiques envahissantes, et certaines sont traitées par la modélisation mathématique. Les renouées asiatiques (*Fallopia sp.*) sont parmi les espèces exotiques envahissantes les plus menaçantes pour la biodiversité à l'échelle mondiale. Cependant, peu de modèles traitent de l'invasion par les renouées asiatiques et la plupart s'intéressent à leur croissance sans considérer de gestion.

L'objectif de cette thèse est de modéliser la croissance de renouées asiatiques soumises à de la fauche et dispersées le long d'un cours d'eau afin de comparer l'efficacité de différentes stratégies de gestion. La croissance locale de la plante peut être étudiée à l'aide de modèles stochastiques. Il est possible d'utiliser la théorie de la viabilité pour étudier sa croissance et sa dispersion à l'échelle d'un cours d'eau. La théorie de la viabilité s'intéresse à des systèmes dynamiques soumis à un ensemble de contraintes de viabilité. Une étude viabiliste cherche à déterminer les états à partir desquels il existe des évolutions au cours desquelles les contraintes sont satisfaites à chaque pas de temps. Le formalisme des inclusions différentielles permet d'introduire deux types d'incertitudes dans les variations de l'état du système : celle liée aux contrôles, qui permet de prendre en compte les différentes possibilités de gestion, et l'incertitude tychastique liée aux perturbations qui permet de prendre en compte l'effet des crues, dont la dispersion des renouées asiatiques dépend, et qui ne sont pas prévisibles.

Notre démarche s'appuie sur le développement et l'étude de deux modèles mathématiques appliqués à la gestion des renouées asiatiques. L'enjeu est de trouver des stratégies de gestion efficaces. Nous identifions d'abord l'influence des paramètres de gestion sur la croissance locale de la plante, puis nous présentons une méthode pour la résolution d'un problème de viabilité général dont celui de la dispersion des renouées le long d'un cours d'eau est une instance. Nous démontrons la faisabilité de cette démarche en l'appliquant au cas d'un cours d'eau fictif. Ces travaux ouvrent plusieurs pistes, en particulier en ce qui concerne l'application de cette méthode pour l'obtention d'indications de gestion dans des cas pratiques.



## Abstract

The effective management of invasive alien plant species is an issue at different levels: environmental, economic and public health. Many studies focus on the optimal management of invasive alien species, some of which are addressed through mathematical modeling. Asian knotweed (*Fallopia sp.*) is among the most threatening invasive alien species to biodiversity worldwide. However, few models address invasion by Asian knotweed and most focus on their growth without considering management.

The aim of this thesis is to model the growth of mown Asian knotweed and its dispersal along a river to compare the effectiveness of different management strategies. Local plant growth can be studied using stochastic models. Viability theory can be used to study plant growth and dispersal on a river scale. Viability theory deals with dynamic systems subject to a set of viability constraints. A viability study aims at determining the states from which there exists at least one an evolution that remains in the constraint set. The formalism of differential inclusions enables to consider two types of uncertainties into the variations of the system state: the uncertainty related to controls, which enables to take into account the different management strategies, and the tyochastic uncertainty related to disturbances, which enables to take into account the effect of floods. Indeed, floods impact the dispersal of Asian knotweeds and flood dynamics is not predictable.

Our approach is based on the development and study of two mathematical models applied to the management of Asian knotweed. The challenge is to determine effective management strategies. We first identify the influence of management parameters on local plant growth, and then present a method for solving a general viability problem of which the dispersal of knotweed along a river is one instance. We demonstrate the feasibility of this approach by applying it to a fictitious river. This work opens up several promising developments, in particular with regard to the application of this method to obtain management indications in practical cases.





# Table des matières

|   |             |
|---|-------------|
| <b>Remerciements</b>  | <b>i</b>    |
| <b>Avant-propos</b>   | <b>iii</b>  |
| <b>Résumé</b>   | <b>v</b>    |
| <b>Abstract</b>   | <b>vii</b>  |
| <b>Liste des figures</b>  | <b>xiii</b> |
| <b>Liste des tableaux</b>   | <b>xvii</b> |
| <b>Introduction</b>   | <b>1</b>    |
| <b>1 Les invasions végétales : modélisation de leur gestion</b>   | <b>7</b>    |
| 1.1 Enjeux . . . . .  | 9           |
| 1.2 Des modèles incluant la gestion . . . . .   | 9           |
| 1.2.1 Le modèle déterministe de Baker and Bode (2016) . . . . .   | 10          |
| 1.2.2 Le modèle individu-centré de Harris et al. (2009) . . . . .   | 11          |
| 1.2.3 Analyse de la sensibilité des paramètres d'un modèle déterministe et<br>d'un modèle individu centré . . . . . | 12          |
| 1.2.4 Le modèle statique de Lookingbill et al. (2014) pour la ré-invasion de<br>l'Ailante . . . . .                 | 13          |
| 1.3 Végétation riparienne . . . . .   | 13          |
| 1.3.1 Généralités . . . . .   | 14          |
| 1.3.2 Noyaux de dispersion par hydrochorie . . . . .  | 14          |
| 1.3.3 Deux modèles incluant une dispersion par l'eau . . . . .  | 15          |
| 1.4 Généralités sur les renouées asiatiques . . . . .   | 17          |
| 1.5 Croissance locale de la renouée du Japon . . . . .  | 19          |
| 1.5.1 Description des mécanismes de croissance locale . . . . .   | 19          |
| 1.5.2 Trois modèles pour la croissance locale de la renouée issus de la littérature                                 | 20          |
| 1.6 Étude de la dispersion . . . . .  | 24          |

|          |  |           |
|----------|--|-----------|
| 1.6.1    | Propagules pour la dispersion . . . . .  | 24        |
| 1.6.2    | Variables explicatives de la présence de renouées le long de cours d'eau . . . . .                                       | 26        |
| <b>2</b> | <b>La modélisation comme outil d'aide à la décision</b>  | <b>29</b> |
| 2.1      | Formalismes pour la modélisation . . . . .   | 31        |
| 2.1.1    | Les mesures pour décrire un modèle individu-centré . . . . .   | 31        |
| 2.1.2    | Les inclusions différentielles pour décrire un modèle global . . . . .   | 32        |
| 2.2      | Validation de modèles . . . . .  | 32        |
| 2.2.1    | Calibration . . . . .  | 32        |
| 2.2.2    | Recherche de diversité : PSE . . . . .   | 36        |
| 2.3      | Analyse de sensibilité d'un modèle . . . . .   | 37        |
| <b>3</b> | <b>La théorie de la viabilité - son intérêt pour l'aide à la gestion de systèmes contrôlés soumis à des incertitudes</b> | <b>41</b> |
| 3.1      | Concepts et outils fondamentaux de la théorie de la viabilité . . . . .  | 43        |
| 3.1.1    | Système contrôlé . . . . .   | 43        |
| 3.1.2    | Système tychastique . . . . .  | 46        |
| 3.1.3    | Système contrôlé et tychastique . . . . .  | 48        |
| 3.2      | Calculs de noyaux de viabilité . . . . .   | 50        |
| 3.2.1    | Outils et méthodes pour le calcul générique d'approximation de noyaux de viabilité en petite dimension . . . . .         | 50        |
| 3.2.2    | Calcul en grande dimension dans des cas particuliers . . . . .   | 52        |
| 3.3      | Un problème de viabilité garanti en grande dimension . . . . .   | 55        |
| 3.3.1    | Description du problème . . . . .  | 55        |
| 3.3.2    | Formalisation mathématique . . . . .   | 56        |
| 3.3.3    | Résolution du problème . . . . .   | 57        |
| <b>4</b> | <b>Modèle de croissance locale</b>   | <b>61</b> |
| 4.1      | Matériels et méthodes . . . . .  | 62        |
| 4.1.1    | Description des phénomènes modélisés . . . . .   | 62        |
| 4.1.2    | Synthèse des paramètres du modèle . . . . .  | 66        |
| 4.1.3    | Formalisme mathématique associé au modèle . . . . .  | 66        |
| 4.1.4    | Simulation du modèle . . . . .   | 68        |
| 4.1.5    | Données pour la calibration . . . . .  | 69        |
| 4.1.6    | Méthode pour la calibration . . . . .  | 70        |
| 4.1.7    | Analyse numérique . . . . .  | 71        |
| 4.2      | Résultats . . . . .  | 71        |
| 4.2.1    | Existence et Unicité de la solution de (4.10) . . . . .  | 71        |
| 4.2.2    | Calibration . . . . .  | 77        |
| 4.2.3    | Influence des paramètres de gestion et de la taille initiale . . . . .   | 79        |
| 4.3      | Discussion . . . . .   | 85        |
| <b>5</b> | <b>Modèle pour la dispersion des renouées le long d'un cours d'eau</b>   | <b>91</b> |
| 5.1      | Matériels et méthodes . . . . .  | 93        |
| 5.1.1    | Mécanismes modélisés . . . . .   | 93        |
| 5.1.2    | Pistes pour l'évaluation des paramètres . . . . .  | 98        |
| 5.1.3    | Objectif de gestion et réponse apportée au problème . . . . .  | 103       |
| 5.1.4    | Application numérique pour l'évaluation des paramètres dans le cas d'un cours d'eau "patchwork" . . . . .                | 109       |

|   |  |            |
|---|--|------------|
| 5.1.5   | Application numérique pour le calcul de la valeur de l'aire totale en dessous de laquelle on a la garantie que l'aire envahie par la renouée restera après $T_{ini}$ . . . . . | 117        |
| 5.2   | Résultats . . . . .  | 119        |
| 5.2.1   | Calibration . . . . .  | 119        |
| 5.2.2   | Cascade de noyaux de viabilité . . . . .   | 121        |
| 5.3   | Discussion . . . . .   | 125        |
| 5.3.1   | Données disponibles, choix de modélisation et calibration . . . . .  | 126        |
| 5.3.2   | L'implémentation de la méthode des cascades de noyaux de viabilité . .   | 126        |
| 5.3.3   | Pertinence des résultats des cascades pour la gestion . . . . .  | 127        |
| <b>Conclusion</b>   |  | <b>129</b> |
| <b>Annexes</b>  |  | <b>135</b> |
| <b>A Limite en grande population</b>                                    |  | <b>137</b> |
| A.1   | Résultats mathématiques . . . . .  | 137        |
| A.1.1   | Mesures Ponctuelles de Poisson . . . . .   | 137        |
| A.1.2   | Lemme de Gronwall . . . . .  | 141        |
| A.1.3   | Convergence vague, étroite . . . . .   | 141        |
| A.1.4   | Critères de tension . . . . .  | 141        |
| A.2   | Calcul du générateur . . . . .   | 143        |
| A.3   | Limite en grande population . . . . .  | 148        |
| A.3.1   | Propriété de moment . . . . .  | 150        |
| A.3.2   | Décomposition du processus avec une martingale . . . . .   | 154        |
| A.3.3   | Tension de la suite $(Z^n)$ . . . . .  | 155        |
| <b>B Code utilisé dans le logiciel VIABLAB</b>                          |  | <b>163</b> |
| <b>C Étude de la diversité des trajectoires produites par Electra 2</b> |  | <b>171</b> |
| C.1   | Matériel et méthodes . . . . .   | 173        |
| C.1.1   | Présentation d'Electra 2 et mécanismes modélisés . . . . .   | 173        |
| C.1.2   | Paramètres . . . . .   | 174        |
| C.1.3   | Formalisme mathématique . . . . .  | 174        |
| C.1.4   | Simulation . . . . .   | 175        |
| C.1.5   | Caractérisation des trajectoires . . . . .   | 175        |
| C.1.6   | Méthode pour la recherche de diversité dans les trajectoires . . . . .   | 177        |
| C.1.7   | Méthode de représentation sous forme de carte topologique . . . . .  | 179        |
| C.2   | Résultats . . . . .  | 180        |
| C.3   | Discussion . . . . .   | 181        |
| <b>D Étude d'un modèle de fission d'une colonie de fourmis</b>          |  | <b>183</b> |
| D.1   | Matériel et Méthodes . . . . .   | 185        |
| D.1.1   | Description des mécanismes . . . . .   | 185        |
| D.1.2   | Synthèse des paramètres . . . . .  | 186        |
| D.1.3   | Simulations . . . . .  | 186        |
| D.1.4   | Données pour la calibration . . . . .  | 187        |
| D.1.5   | Méthode pour la calibration . . . . .  | 187        |

|                      |  |            |
|----------------------|--|------------|
| D.1.6                | Analyse numérique . . . . .                      | 188        |
| D.1.7                | Méthode pour l'approximation du modèle . . . . . | 188        |
| D.2                  | Résultats . . . . .                              | 189        |
| D.2.1                | Calibration . . . . .                            | 189        |
| D.2.2                | Approximation du modèle . . . . .                | 190        |
| D.3                  | Conclusion . . . . .                             | 195        |
| <b>Bibliographie</b> |  | <b>197</b> |

## Liste des figures

|     |  |    |
|-----|--|----|
| 1.1 | Renouée du Japon, partie aérienne (photo personnelle) . . . . .  | 18 |
| 1.2 | Renouée du Japon, rhizome (photo personnelle) . . . . .  | 18 |
| 1.3 | Schéma du développement des tiges et des bourgeons le long du rhizome,<br>extrait de Adachi et al. (1996a) . . . . .   | 20 |
| 1.4 | Direction de branchement, extrait de Smith et al. (2007) . . . . .   | 22 |
| 1.5 | Une simulation du modèle, extrait de Smith et al. (2007) . . . . .   | 22 |
| 1.6 | Cycle de vie simplifié de la renouée utilisé dans le modèle IPM, extrait de Dauer<br>and Jongejans (2013) . . . . .  | 22 |
| 3.1 | En bleu, le noyau de viabilité pour la dynamique (3.21) avec l'ensemble des<br>contraintes $K = [0; +\infty] \times [0.1; 0.18]$ , calculé avec le logiciel KAVIAR, extrait<br>de Martin et al. (2011). . . . .              | 52 |
| 3.2 | En gris, le noyau de viabilité pour la dynamique (3.22) avec l'ensemble des<br>contraintes $K = [L_{min}, L_{max}] \times [0, P_{max}]$ , calculé avec le logiciel Viabilitree,<br>extrait de Alvarez et al. (2016). . . . . | 53 |
| 4.1 | Histogramme des longueurs d'extension du rhizome, extrait de Smith et al.<br>(2007). . . . .   | 63 |
| 4.2 | Schéma représentant un événement de naissance . . . . .  | 64 |
| 4.3 | Exemple illustratif : simulation d'une trajectoire du modèle . . . . .   | 70 |
| 4.4 | Loi Gamma pour la distance de dispersion . . . . .   | 78 |
| 4.5 | Taux de mortalité en fonction de la biomasse . . . . .   | 78 |
| 4.6 | Fonction décrivant l'effet de la fauche . . . . .  | 78 |
| 4.7 | Évolution de la biomasse d'un individu . . . . .   | 78 |
| 4.8 | Régression linéaire sur l'aire finale moyenne en fonction de la durée du projet de<br>gestion $T$ , avec $\tau = 0,5 \text{ an}^{-1}$ et $initialPopSize = 1500$ , extrait de Lavallée<br>et al. (2019b). . . . .            | 80 |
| 4.9 | Régression quadratique sur l'aire finale moyenne en fonction de la durée du<br>projet de gestion $T$ , avec $\tau = 8 \text{ an}^{-1}$ et $initialPopSize = 500$ . . . . .   | 81 |

|      |   |     |
|------|---|-----|
| 4.10 | Courbes de régression quadratique sur l'aire finale moyenne en fonction de $T$ , pour différentes valeurs de $\tau$ avec $initialPopSize = 1500$ et $proportionMowing = 0,9$ . . . . .  | 82  |
| 4.11 | Régression non linéaire sur la taille finale moyenne en fonction de la durée du projet de gestion $T$ avec $\tau = 8 \text{ an}^{-1}$ et $initialPopSize = 500$ . . . . .   | 83  |
| 4.12 | Régression linéaire sur l'aire finale moyenne en fonction du nombre moyen de fauches par année $\tau$ , avec $T = 8 \text{ ans}$ et $initialPopSize = 1500$ , extrait de Lavallée et al. (2019b). . . . .   | 84  |
| 4.13 | Régression non linéaire sur la taille finale moyenne en fonction du nombre moyen de fauches $\tau$ avec $T = 8 \text{ ans}$ et $initialPopSize = 1500$ , extrait de Lavallée et al. (2019b). . . . .  | 85  |
| 4.14 | Régression linéaire de la taille finale moyenne en fonction de la taille de population initiale avec $T = 8 \text{ ans}$ et $\tau = 4 \text{ an}^{-1}$ , extrait de Lavallée et al. (2019b). . . . .  | 86  |
| 4.15 | Régression linéaire de l'aire finale moyenne en fonction de la taille population initiale avec $T = 8 \text{ ans}$ et $\tau = 4 \text{ an}^{-1}$ , extrait de Lavallée et al. (2019b). . . . .  | 87  |
| 4.16 | Régression exponentielle sur la taille finale moyenne en fonction de la durée du projet de gestion ( $T$ ) . . . . .  | 88  |
| 4.17 | Paramètres ayant la plus grande influence sur la taille finale . . . . .  | 89  |
| 4.18 | Paramètres ayant la plus grande influence sur l'aire finale . . . . .   | 90  |
| 5.1  | Schéma synthétisant les processus écologiques intervenant pendant une année dans le modèle. La variable $a_i$ désigne l'aire de la zone $i$ , $b_{i,A}$ l'accumulateur d'akènes dans la zone $i$ , $b_{i,R}$ l'accumulateur de fragments de rhizomes dans la zone $i$ . $N_{i,A}$ (resp. $N_{i,R}$ ) désigne le nombre de nouvelles plantules créées dans la zone $i$ issues d'akènes (resp. de fragments de rhizomes). . . . . | 93  |
| 5.2  | Cartes de progression des renouées asiatiques sur le Haut-bassin de l'Hérault entre 2007 et 2015, figure extraite de Barthod and Boyer (2019) . . . . .   | 110 |
| 5.3  | Evolution de la hauteur d'eau à la station de Laroque entre 2007 et 2015 par jour. Les zones colorées correspondent aux périodes de crues importantes identifiées dans Barthod and Boyer (2019) . . . . .   | 111 |
| 5.4  | Points de confluence sur le Haut-bassin de l'Hérault dans la zone d'étude considérée. . . . .   | 112 |
| 5.5  | Découpage du Haut-bassin de l'Hérault en zones. . . . .   | 112 |
| 5.6  | Evolution de la hauteur d'eau à la station de Laroque entre 2007 et 2015 par an (valeur maximale de l'année). . . . .   | 112 |
| 5.7  | Aires envahies en 2015, simulées en noir avec les valeurs de paramètres issues de la calibration, et observées en rouge. Les points bleus montrent la valeur $A_i$ de l'aire totale de chaque zone $i$ . . . . .  | 121 |
| 5.8  | Aires envahies en 2015, simulées en noir avec les valeurs de paramètres issues de la calibration si on impose un seuil maximal sur la distance entre les nouvelles plantules et les nouvelles taches entre 2008 et 2015, et observées en rouge. Les points bleus montrent la valeur $A_i$ de l'aire totale de chaque zone $i$ . . . . .   | 122 |
| 5.9  | Coupe du noyau de viabilité du système (5.17) avec la contrainte (5.18) selon $T = 0$ et $r = r_{max}$ pour $i = 1$ , $\bar{a}_1 = 149 \text{ m}^2$ , $U_{1,max} = 8$ , $u_{1,max} = 4$ et $T_{ini} = 2$ . La situation initiale est indiquée par le point bleu et l'ensemble de contrainte est délimité par la droite rouge. La situation initiale appartient à la coupe. . . . .  | 123 |

|      |   |     |
|------|---|-----|
| 5.10 | Coupe du noyau de viabilité du système (5.17) avec la contrainte (5.18) selon $T = 0$ et $r = r_{max}$ pour $i = 1$ , $\bar{a}_1 = 145 \text{ m}^2$ , $U_{1,max} = 8$ , $u_{1,max} = 4$ et $T_{ini} = 2$ . La situation initiale est indiquée par le point bleu, et l'ensemble de contrainte est délimité par la droite rouge. La situation initiale n'appartient pas à la coupe. . . . .           | 123 |
| 5.11 | Coupe du noyau de viabilité du système (5.17) avec la contrainte (5.18) selon $T = T_{ini} + 1$ et $r = r_{max}$ pour $i = 1$ , $\bar{a}_1 = 149 \text{ m}^2$ , $U_{1,max} = 8$ , $u_{1,max} = 4$ et $T_{ini} = 2$ . La situation initiale est indiquée par le point bleu, et l'ensemble de contrainte est délimité par la droite rouge. La situation initiale n'appartient pas à la coupe. . . . . | 123 |
| 5.12 | Coupe du noyau de viabilité du système (5.17) avec la contrainte (5.18) selon $T = 0$ et $r = r_{max}$ pour $i = 1$ , $\bar{a}_1 = 145 \text{ m}^2$ , $U_{1,max} = 15$ , $u_{1,max} = 4$ et $T_{ini} = 2$ . La situation initiale est indiquée par le point bleu, et l'ensemble de contrainte est délimité par la droite rouge. La situation initiale appartient à la coupe. . . . .                | 123 |
| 5.13 | Coupe du noyau de viabilité du système (5.17) avec la contrainte (5.18) selon $T = 0$ et $r = r_{max}$ pour $i = 1$ , $\bar{a}_1 = 70 \text{ m}^2$ , $U_{1,max} = 8$ , $u_{1,max} = 4$ et $T_{ini} = 10$ . La situation initiale est indiquée par le point bleu, et l'ensemble de contrainte est délimité par la droite rouge. La situation initiale appartient à la coupe. . . . .                 | 124 |
| C.1  | Electra 2 . . . . .   | 173 |
| C.2  | Electra 2, tracé lumineux . . . . .   | 173 |
| C.3  | Schéma d'Electra 2 . . . . .  | 174 |
| C.4  | Trajectoire simulée avec $v_1 = 1, v_2 = -3, v_3 = 2$ , tous les points . . . . .   | 175 |
| C.5  | Trajectoire simulée avec $v_1 = 1, v_2 = -3.1, v_3 = 2.1$ , tous les points . . . . .   | 175 |
| C.6  | Illustration de la courbure, $c = 12,56$ . . . . .  | 176 |
| C.7  | Illustration de la courbure, $c = 13,19$ . . . . .  | 176 |
| C.8  | Illustration de la courbure, $c = 18,84$ . . . . .  | 176 |
| C.9  | Illustration de la courbure, $c = 19,47$ . . . . .  | 176 |
| C.10 | Illustration du nombre de points singuliers, $s = 30$ . . . . .   | 177 |
| C.11 | Illustration du nombre de points singuliers, $s = 0$ . . . . .  | 177 |
| C.12 | Illustration du nombre de points singuliers, $s = 31$ . . . . .   | 177 |
| C.13 | Illustration du nombre de points singuliers, $s = 22$ . . . . .   | 177 |
| C.14 | Illustration de la longueur d'une structure élémentaire, $l = 666$ . . . . .  | 178 |
| C.15 | Illustration de la longueur d'une structure élémentaire, $l = 4000$ . . . . .   | 178 |
| C.16 | Illustration de la longueur d'une structure élémentaire, $l = 1000$ . . . . .   | 178 |
| C.17 | Illustration de la longueur d'une structure élémentaire, $l = 905$ . . . . .  | 178 |
| C.18 | Extrait de Chérel et al. (2015), étapes de l'algorithme PSE . . . . .   | 179 |
| C.19 | Carte de Kohonen de 100 trajectoires issues de l'exploration par l'algorithme PSE . . . . .   | 180 |
| C.20 | Moving cells, une configuration . . . . .   | 181 |
| C.21 | Moving cells, une configuration . . . . .   | 181 |
| D.1  | Lois marginales de la loi <i>a posteriori</i> pour les paramètres soumis à la calibration, pour 3 colonies (lignes). . . . .  | 189 |
| D.2  | Comparaison des répliquations de sorties du modèle simulé avec les jeux de paramètres issus de la loi trouvée par la calibration et des données de terrain. . . . .   | 190 |
| D.3  | Nombre moyen de nids visités par une fourmi en fonction de la durée de la phase d'exploration, pour différentes valeurs de nombre de nids potentiels . . . . .  | 192 |



- D.4 Nombre moyen de nids après fission prédit et simulé en fonction de la durée de la phase d'exploration, pour différentes valeurs de nombre de nids potentiels et de nombre de fourrageuses . . . . . 193
- D.5 Nombre moyen de fourmis dans chaque nid après fission prédits et simulés, en fonction de la durée de la phase d'exploration, pour différentes valeurs du nombre de nids potentiels et du nombre de fourrageuses . . . . . 195

## Liste des tableaux

|     |  |     |
|-----|--|-----|
| 4.1 | Synthèse des paramètres pour la dynamique de la plante dans le modèle . . . .  | 66  |
| 4.2 | Résultat de la calibration obtenu avec OpenMOLE (Reuillon et al., 2013) . . .  | 77  |
| 4.3 | Synthèses des résultats des régressions effectuées dans la Section 4.2.3. . . . .  | 83  |
| 5.1 | Synthèse des paramètres pour la dynamique de l'invasion dans le modèle de dispersion . . . . .   | 98  |
| 5.2 | Synthèse des paramètres pour la dynamique de l'invasion dans le modèle de dispersion, les lignes grisées désignent les paramètres soumis à la calibration, les autres désignent les paramètres dont la valeur est fixée à partir de données et de la littérature. . . . .  | 103 |
| 5.3 | Longueurs, largeurs, aires et hauteurs de la crue maximale considérée pour chaque zone du Haut-bassin de l'Hérault. . . . .  | 113 |
| 5.4 | Valeurs des paramètres pour la dynamique de l'invasion dans le modèle de dispersion lorsqu'elles sont fixées à partir de données et de la littérature. Les lignes grisées désignent les paramètres soumis à la calibration. (*) signifie que ces valeurs de $a_{0,i}$ sont obtenues par calibration (voir Section 5.2.1). . . . .                                | 114 |
| 5.5 | Données d'abondance de renouées sur le Haut-bassin de l'Hérault utilisées pour la calibration. Les zones sont celles de la Figure 5.5. . . . .   | 115 |
| 5.6 | Valeurs des paramètres pour la dynamique de l'invasion dans le modèle de dispersion lorsqu'elles sont fixées à partir de données et de la littérature. Valeurs des plages de valeurs testées pour les paramètres soumis à la calibration (lignes grisées). (*) signifie que ces valeurs de $a_{0,i}$ sont obtenues par calibration (voir Section 5.2.1). . . . . | 116 |
| 5.7 | Valeurs des paramètres pour la dynamique de l'invasion dans le modèle de dispersion. Ces valeurs sont fixées à partir de données et de la littérature pour les lignes non grisées et issues de la calibration dans le cas des lignes grisées. .  | 120 |
| 5.8 | Comparaison des aires envahies en 2015, observées et simulées avec les valeurs de paramètres issues de la calibration. . . . .   | 120 |
| 5.9 | Comparaison des aires envahies en 2015, observées et simulées avec les valeurs de paramètres issues de la calibration si on impose un seuil maximal sur la distance entre les nouvelles plantules et les nouvelles taches entre 2008 et 2015.  | 121 |

|      |   |     |
|------|---|-----|
| 5.10 | Valeurs de l'aire en début de projet de gestion et de l'aire totale par zone. . . .   | 125 |
| 5.11 | Répartitions des efforts entre zones utilisées dans la cascade de viabilité pour le calcul de l'aire totale envahie minimale. . . . .   | 125 |
| 5.12 | Valeurs de $\tilde{a}_i$ dans chaque zone $i$ , obtenues par l'application de la méthode de cascade de noyaux de viabilité, pour les répartitions des efforts décrites dans le Tableau 5.11. Nous indiquons également l'aire totale minimale garantie par chaque stratégie. . . . . | 125 |
| C.1  | Bornes choisies pour les paramètres du modèle dans l'algorithme PSE pour le modèle Electra 2. . . . .   | 179 |
| C.2  | Discretisation de l'espace des motifs utilisé dans l'algorithme PSE pour le modèle Electra 2 . . . . .  | 179 |
| D.1  | Synthèse des paramètres pour le modèle de fission d'une colonie . . . . .   | 186 |

# Introduction

Le savoir mathématique, constitué de connaissances abstraites, est appliqué à des problèmes issus de différents domaines. Il peut être mobilisé pour formaliser des problématiques concrètes, qu'il permet de mieux comprendre. Ces connaissances issues des analyses mathématiques peuvent ainsi servir à appuyer et justifier des prises de décision.

L'écologie fait partie des domaines d'application des mathématiques. Dans cette thèse, nous nous intéressons plus particulièrement à l'écologie des plantes exotiques envahissantes, dites plantes invasives. Nous allons montrer que le formalisme mathématique est utile pour l'aide à la gestion des espèces végétales invasives. Plus précisément, nous avons choisi de nous intéresser au cas des renouées asiatiques le long d'un cours d'eau. Pour obtenir des recommandations de gestion, nous avons choisi de nous placer dans le cadre de la modélisation mathématique. Nous avons ainsi commencé par construire un modèle pour la croissance locale de ces plantes. Nous avons ensuite changé d'échelle pour nous intéresser à la dispersion de ces plantes le long de cours d'eau. Nous avons construit un modèle pour étudier l'influence des paramètres de gestion sur l'aire envahie par la renouée, ce qui a permis d'obtenir des recommandations pour sa gestion.

La gestion efficace des espèces végétales exotiques envahissantes est un enjeu à différents niveaux. Les espèces exotiques envahissantes ont d'abord un impact environnemental, économique et peuvent représenter un enjeu de santé publique à l'échelle mondiale. De nombreuses études portent sur la gestion optimale des espèces exotiques envahissantes. La littérature sur la gestion des espèces exotiques envahissantes traitée par la modélisation mathématique s'intéresse notamment à l'évaluation des bénéfices liés à la priorisation spatiale d'un projet de gestion ou à l'étude du compromis entre l'intensité et la durée de la stratégie de lutte pour maintenir l'invasion en dessous d'un seuil souhaitable. Il existe plusieurs modèles pour la croissance d'espèces exotiques envahissantes, qui reposent sur des formalismes mathématiques différents et apportent ainsi diverses réponses à ces problématiques. Les renouées asiatiques (*Fallopia* sp.) sont parmi les espèces exotiques envahissantes les plus menaçantes pour la biodiversité à l'échelle mondiale. Espèces clonales à forte capacité de reproduction et de dispersion, elles posent des problèmes de gestion importants. En effet, elles se dispersent rapidement le long de cours d'eau ou d'infrastructures linéaires de transport et colonisent durablement les milieux. Les gestionnaires confrontés à ces espèces ont peu de moyens d'intervention et sont contraints par des compromis coût/efficacité. Les modèles mathématiques peuvent aider à reproduire le

système biologique étudié et les actions de gestion afin d'en étudier les effets sur le développement des plantes ciblées. Ainsi, chez les renouées asiatiques, les modèles visent à reproduire sa croissance locale et sa dispersion sur la base des mécanismes biologiques et écologiques impliqués.

Comme nous l'avons indiqué, pour répondre aux problématiques de gestion, nous avons choisi de nous placer dans le cadre de la modélisation mathématique. L'invasion des renouées, que ce soit au niveau de leur croissance locale ou de leur dispersion, est un phénomène dynamique, nous faisons donc le choix d'utiliser des modèles dynamiques. Il existe différents formalismes correspondant à différentes échelles : par exemple, nous utilisons les mesures pour le modèle de croissance locale et les inclusions différentielles pour le modèle de dispersion. Cependant, si les mathématiques donnent un cadre rigoureux pour l'étude de modèles, il n'est pas toujours aisé, du fait de la complexité de certains modèles, d'obtenir des résultats analytiques sur les propriétés du modèle particulier considéré. Nous avons alors recours à des techniques d'exploration de modèles. Ces techniques sont valables quel que soit le formalisme mathématique utilisé. Deux étapes sont essentielles dans la modélisation : la calibration dont l'objectif est de trouver des valeurs de paramètres qui permettent de faire correspondre au mieux les sorties du modèle aux données et la validation qui consiste à vérifier qu'un modèle ne produit pas de valeurs non souhaitées. Les méthodes d'analyse de sensibilité d'un modèle permettent l'étude de la manière dont un modèle réagit à une modification de ses entrées et ainsi de répondre à des questions sur l'impact de différentes décisions sur l'évolution d'un système.

Le problème de gestion posé par la renouée est de contenir l'invasion le long d'un cours d'eau, c'est-à-dire de maintenir l'aire envahie en dessous d'un seuil. Il s'exprime naturellement par une contrainte à respecter plutôt que par une fonction objectif à optimiser. La théorie de la viabilité s'intéresse à des systèmes dynamiques soumis à un ensemble de contraintes de viabilité. Une étude viabiliste cherche à déterminer les états à partir desquels il existe des évolutions au cours desquelles les contraintes sont satisfaites à chaque pas de temps. Le formalisme ensembliste des inclusions différentielles permet d'introduire deux types d'incertitude dans les variations de l'état du système : celle liée aux contrôles, qui permet de prendre en compte les différentes possibilités de gestion, et l'incertitude tychastique liée aux perturbations qui permet de prendre en compte l'effet des crues, dont la dispersion des renouées asiatiques dépend, et qui ne sont pas prévisibles. C'est pourquoi nous avons choisi de réaliser l'étude du modèle de gestion de la renouée le long d'un cours d'eau dans le cadre de la théorie de la viabilité. La résolution analytique d'un problème de viabilité n'est possible que dans quelques cas particuliers. Des algorithmes d'approximation existent dans les autres cas, mais ils sont applicables en pratique seulement si la dimension de l'espace d'état du système n'est pas trop grande (ce sont des méthodes exponentielles en l'espace des états). Cependant, notre problème de viabilité issu de la gestion des renouées le long d'un cours d'eau ne peut être résolu par ces méthodes car il contient trop de variables. Nous avons donc développé la méthode de "cascade de noyaux de viabilité" pour la résolution d'une classe de problèmes de viabilité dont notre problème particulier de gestion fait partie.

Pour répondre à la question de l'influence de la fréquence de fauche (technique de gestion la plus couramment utilisée pour gérer des renouées), nous développons un modèle individu-centré qui décrit la croissance locale d'un massif ou "tache" de renouées et qui inclut la fauche comme stratégie de gestion (à notre connaissance aucun modèle ne permet de répondre à des problématiques de gestion concernant la fauche). Nous calibrons le modèle avec des données de terrain et étudions l'influence de la taille initiale de la population, du nombre moyen de fauches par année et de la durée du projet de gestion, sur l'aire et la densité moyennes de la tache. Notre étude renseigne ainsi sur la fréquence minimale de fauche pour obtenir une

diminution de l'aire de la tache, comme sur la durée minimale du projet de gestion nécessaire pour parvenir à l'éradication (si elle est possible, compte tenu d'une certaine fréquence de fauche). Nous pouvons aussi comparer l'efficacité des différentes stratégies de fauches.

Cependant, les décisions de gestion ne sont pas prises tache par tache mais à l'échelle d'unités de gestion, par exemple le long d'un cours d'eau, qui constitue en outre un vecteur de dispersion très efficace lors des crues. Pour traiter une telle échelle, nous développons cette fois un modèle global qui s'appuie sur les résultats moyens du modèle de croissance d'une tache. L'objectif de gestion considéré étant la maîtrise de l'aire envahie le long d'un cours d'eau, nous exprimons cet objectif de gestion sous la forme de la satisfaction d'une contrainte au cours de temps, et utilisons le formalisme des inclusions différentielles pour décrire notre modèle. Ainsi, nous obtenons grâce aux méthodes et outils de la théorie de la viabilité des résultats sur l'aire envahie minimale qu'il est possible de garantir étant donné un effort de gestion à répartir spatialement.

Les travaux sur la croissance locale des renouées asiatiques ont donné lieu à trois publications. La première s'intéresse aux apports de la modélisation pour l'aide à la gestion de la renouée du Japon, et a été publié dans Sciences Eaux & Territoires (Lavallée et al., 2019a). La seconde porte sur l'élaboration et l'étude d'un modèle de croissance locale pour cette plante, travail présenté dans le Chapitre 4 et publié dans Ecological Modelling (Lavallée et al., 2019b). La dernière traite d'une expérience menée pour étudier les dynamiques de la croissance clonale sous différentes conditions environnementales, et a été publié dans NeoBiota (Martin et al., 2020).

Concernant l'exploration de modèles, un article est en cours d'écriture sur le travail présenté dans l'Annexe D (Monnin et al., en cours d'écriture), et une présentation a été faite dans le cadre des journées de Rochebrune 2020 sur le travail présenté dans l'Annexe C. Concernant la méthode de cascade de noyaux développée pour résoudre un problème de viabilité, un article portant sur une technique similaire est en cours d'écriture (Martin et al., en cours d'écriture).

## Organisation du document

La première partie de la thèse présente quelques éléments écologiques sur le système biologique étudié et les méthodes et outils utilisés dans le manuscrit. Le Chapitre 1 s'intéresse aux espèces exotiques envahissantes et plus particulièrement aux renouées asiatiques. Nous présentons plusieurs modèles pour la croissance d'espèces exotiques envahissantes reposant sur des formalismes mathématiques différents. Les renouées asiatiques, modèle biologique étudié, affectionnant particulièrement les berges de cours d'eau, nous traitons aussi de quelques spécificités des végétations ripariennes. Nous présentons enfin des modèles pour la croissance locale de renouées ainsi que les mécanismes intervenant dans la dispersion, particulièrement le long de cours d'eau. Des outils méthodologiques issus de la modélisation pour l'aide à la décision sont introduits dans le Chapitre 2. Nous introduisons les formalismes mathématiques utilisés pour l'écriture de nos deux modèles, puis nous exposons des méthodes de validation de modèles pertinentes pour nos études. Enfin, nous présentons les méthodes d'analyse de sensibilité d'un modèle. Le Chapitre 3 vise à présenter des outils méthodologiques issus de la théorie de la viabilité utiles au problème de gestion de la renouée le long d'un cours d'eau. Nous énonçons d'abord des concepts essentiels de la théorie de viabilité. Nous présentons ensuite des méthodes de résolution de ces problèmes et d'approximation de leur solution. Nous détaillons enfin la méthode de "cascade de noyaux de viabilité" mise au point pour la résolution d'une classe de problèmes de viabilité.

La seconde partie est appliquée et concerne la gestion des renouées asiatiques. Le Chapitre

4 présente le modèle pour la croissance locale de la plante alors que le Chapitre 5 présente le modèle de l'aire envahie soumise à la dispersion le long d'un cours d'eau. Dans ces deux chapitres, nous présentons la construction du modèle ainsi que son étude et les implications pour la gestion qui en sont déduites.

Enfin, la dernière partie de la thèse est consacrée au bilan du travail réalisé. Nous revenons sur les indications de gestion qui ont été mises en lumière par les études des deux modèles, et les méthodes utilisées. Nous formulons également de possibles perspectives à ce travail.







## Chapitre

# 1

## Les invasions végétales : modélisation de leur gestion

Ce chapitre introduit les problématiques liées aux espèces exotiques envahissantes et à leur gestion. De nombreuses études portent sur la gestion optimale des espèces exotiques envahissantes (Baker and Bode, 2016; Harris et al., 2009; Travis et al., 2011). Une question centrale dans la littérature sur la gestion des espèces envahissantes traitée par la modélisation mathématique concerne l'évaluation des bénéfices liés à la priorisation spatiale d'un projet de gestion. Par exemple, pour les espèces dispersées par des graines, vaut-il mieux retirer les individus au cœur de l'invasion ou ceux qui se trouvent à la périphérie (Harris et al., 2009) ? D'autres questions importantes portent sur le compromis entre l'intensité et la durée de la stratégie de lutte pour maintenir l'invasion en dessous d'un seuil souhaitable. Par exemple, est-il préférable d'agir de manière significative au début du projet de gestion et de contrôler ensuite l'invasion avec un coût moindre (comme dans Meier et al. (2014)), ou d'avoir un projet de gestion plus long qui nécessite un effort annuel moins important ? Ce compromis entre l'intensité de l'effort de gestion et la durée du projet de gestion est une question cruciale que nous allons traiter dans le Chapitre 4 dans le cas des renouées asiatiques. En effet, les renouées asiatiques sont classées parmi les espèces envahissantes menaçant le plus la biodiversité, elles soulèvent des problématiques de gestion importantes et plusieurs études portent sur leur évolution en fonction de différentes conditions (parmi elles, Dommanget et al. (2013) et Martin et al. (2020) à laquelle nous avons participé).

Ce chapitre est organisé comme suit. Dans une première partie, nous allons présenter différents modèles pour la croissance d'espèces envahissantes reposant sur des formalismes mathématiques différents. Nous indiquerons les questions de gestion auxquelles ces modèles permettent de répondre, et les réponses qu'ils apportent. Cependant, dans ces modèles, la principale source de dispersion considérée est le vent. Or, l'objectif qui nous anime dans cette thèse concerne la gestion d'une plante envahissante le long de cours d'eau. C'est pourquoi nous traiterons de quelques spécificités des végétations ripariennes et nous détaillerons deux noyaux de dispersion par hydrochorie pour des graines et des tiges, ainsi que deux modèles de dispersion de plante

à l'échelle régionale qui incluent la dispersion par des cours d'eau. Dans une seconde partie, nous nous intéresserons plus particulièrement aux renouées asiatiques, espèces sur lesquelles nous avons choisi de nous concentrer plus particulièrement dans cette thèse, et pour lesquelles nous proposons différents modèles de croissance (voir Chapitres 4 et 5). Nous allons d'abord décrire les enjeux qui concernent ce complexe d'espèces. Puis nous nous intéresserons à leur croissance locale, en présentant les phénomènes impliqués et des modèles tachant de la reproduire. Enfin nous nous intéresserons à la dispersion de la plante, et particulièrement à sa dispersion le long de cours d'eau. Nous présenterons d'abord les propagules intervenant dans ce phénomène et leurs caractéristiques, ainsi que la manière dont elles sont impactées par le transport dans l'eau. Nous préciserons aussi les résultats d'études s'attachant à comprendre les variables explicatives de la présence de renouées le long de cours d'eau.

## 1.1 Enjeux

Les invasions biologiques sont un des changements globaux affectant la biodiversité (voir <https://uicn.fr/>). Processus naturel d'origine anthropique, elles impliquent le développement très dynamique et rapide de populations d'espèces introduites par l'homme en dehors de leur aire de répartition d'origine (Richardson et al., 2000). Les impacts sont multiples et peuvent être d'ordre économique, sanitaire et écologique. En particulier, certaines espèces végétales exotiques envahissantes (dites aussi espèces végétales invasives) peuvent affecter gravement la biodiversité ainsi que certaines fonctions écosystémiques. C'est le cas des renouées asiatiques dont nous parlerons plus amplement ci-après. La gestion efficace de ces espèces exotiques envahissantes présente donc des enjeux forts. Il est admis que plus le processus est traité en amont, plus la gestion est efficace et peu coûteuse (Simberloff et al., 2013). En revanche, il est souvent difficile de traiter des populations d'espèces invasives précocement pour des questions de détection et le problème de la gestion de populations déjà bien installées se pose alors. La modélisation mathématique permet, dans ce contexte, de mieux comprendre les effets de certains paramètres de gestion (priorisation spatiale ou temporelle des efforts de gestion) afin d'optimiser son efficacité.

## 1.2 Des modèles incluant la gestion

De nombreux articles, dont Baker and Bode (2016), Harris et al. (2009) et Travis et al. (2011), s'intéressent à la gestion optimale de l'espèce invasive. Dans les stratégies de contrôle, on souhaite amener la densité de l'espèce envahissante en dessous d'un seuil fixé. La question posée est de savoir s'il faut opérer une priorisation spatiale : retirer les individus au cœur de l'infection, ou ceux à la périphérie (Harris et al., 2009). La réponse à cette problématique dépend de manière essentielle des paramètres du modèle (et donc de l'espèce considérée). Ainsi, pour une espèce qui disperse ses graines à de grandes distances, celles-ci augmentant avec l'âge de la plante (et donc sa taille), il semble qu'il faille privilégier le retrait des plantes au cœur de l'infection. Par contre, si les plantes responsables de l'avancée du front de propagation de l'invasion se situent en périphérie (et donc n'ont pas une grande distance de dispersion), il est souhaitable de traiter en priorité les plantes en périphérie de l'infection (Harris et al., 2009). Une autre question se pose sur la répartition temporelle de l'effort : vaut-il mieux agir significativement en début de projet de gestion et ensuite contrôler l'invasion avec un effort plus faible (comme dans Meier et al. (2014)), ou augmenter l'effort au cours du temps (comme dans Baker and Bode (2016)) ?

Nous constatons qu'une hypothèse fréquemment émise dans les articles abordant les stratégies de contrôle, est de supposer que l'espèce envahissante est déjà installée depuis longtemps (par exemple dans Baker and Bode (2016)). Pour de nombreux modèles, la principale source de dispersion considérée est le vent. De plus, ces articles considèrent souvent que le terrain sur lequel se déroule l'invasion est homogène. Or, on constate que la présence de cours d'eau peut modifier la dispersion des graines. De plus, les modèles qui tiennent compte de la complexité du terrain soumis à l'invasion sont essentiellement des modèles statiques, qui ne considèrent donc pas l'invasion comme un phénomène dynamique (pas de composante temporelle). Ces modèles utilisent souvent des indices tels que la qualité de l'habitat (Lookingbill et al., 2014), construits à partir de données de terrain comme l'ensoleillement, l'humidité ou la nature des sols, pour établir des cartes. Ces cartes attribuent un score pour chaque zone, qui décrit sa probabilité d'être envahie Lookingbill et al. (2014).

Dans cette section, nous présentons plus précisément quelques uns des modèles pour la croissance d'espèces envahissantes cités précédemment, qui nous ont permis d'apprécier la variété des formalismes mathématiques employés et la diversité des problématiques de gestion traitées.

### 1.2.1 Le modèle déterministe de Baker and Bode (2016)

Baker and Bode (2016) proposent un modèle général pour la croissance et la dispersion d'une espèce exotique envahissante. En plus de la croissance de l'espèce et de sa dispersion spatiale, le modèle prend en compte la réponse à l'effort de contrôle. L'objectif de gestion est de réduire la densité de population en dessous d'un seuil. A cette fin, les auteurs cherchent les stratégies de contrôle optimales et étudient donc la répartition (en temps et en espace) optimale de l'effort à appliquer. Les problèmes de contrôle d'espèces exotiques envahissantes sont par nature spatiotemporels, puisque la taille et la densité de la population envahissante ainsi que la politique de gestion varient dans le temps et l'espace. Les auteurs présentent ici un modèle déterministe, spatiotemporel, et expliquent les hypothèses à ajouter pour le simplifier en un modèle purement spatial, ou purement temporel. Les deux simplifications supposent que l'espèce est établie depuis longtemps. Les auteurs appliquent ce modèle pour la gestion d'une espèce animale, mais ils soulignent que ce dernier peut aussi s'écrire pour une espèce végétale. Ainsi, ils estiment les paramètres du modèle pour une population de chats sauvages envahissant une île en Australie, et les résultats de leur étude concernent cette population.

Le modèle général repose sur l'équation de réaction-diffusion (1.1)

$$\frac{\partial N}{\partial t} = D\Delta N + rN\left(1 - \frac{N}{k}\right) - N(\mu E)^q, \quad (1.1)$$

où  $N$  est la taille de la population,  $r$  désigne le taux de croissance intrinsèque de la population et  $k$  la capacité de charge. Le laplacien modélise la dispersion spatiale. Le second terme du membre de droite de l'Équation (1.1) correspond à une croissance logistique de la population. Enfin, le terme  $N(\mu E)^q$  décrit les effets de la stratégie de contrôle : le contrôle agit comme une mortalité induite sur la population, proportionnelle (par  $\mu$ ) à l'effort de contrôle  $E$  sur la population. Le paramètre  $q \in [0, 1]$  rend compte du caractère décroissant du rendement des efforts de contrôle  $E$ . Le paramètre  $\mu$  et le contrôle  $E$  sont susceptibles de dépendre de l'espace et du temps. Du fait de la diffusion, les effets du contrôle à une position sont aussi ressentis aux environs. L'objectif de gestion est donc de trouver l'effort  $E$  qui permet de réduire la taille de la population sous un seuil pour un coût minimal.

Les auteurs s'intéressent d'abord au cas où l'invasion se produit dans une petite région fermée, pour lequel la densité moyenne (spatiale) rend bien compte de la densité. Ils ne considèrent donc plus la variable spatiale et se concentrent donc sur les aspects temporels du contrôle. Ils considèrent de plus que l'espèce envahissante est établie depuis longtemps. L'Équation (1.2) résulte de la simplification de l'Équation (1.1) dans ce cas. Elle est écrite pour la densité  $n = N/k$ , supposée initialement valoir 1.

$$\frac{\partial n}{\partial t} = rn(1 - n) - ne(t)^q \quad (1.2)$$

où  $e(t) := \mu E(t)$  est l'effort engagé pour l'éradication.

L'objectif est, à durée de la stratégie de gestion fixée, de trouver la répartition temporelle de l'effort qui réduit la densité de la population en dessous d'un seuil fixé pour un coût minimal. Les auteurs trouvent que le choix optimal est ici de commencer le contrôle par un faible effort qui s'intensifie à la fin de la période. Un effort faible initial réduit la taille de la population de manière optimale (le taux de croissance par individu est ici faible, l'individu retiré par le

contrôle n'est donc pas remplacé). Ils remarquent que lorsque le projet de gestion est plus court, l'allocation des ressources est plus uniforme. De plus, ils vérifient que le coût total du contrôle diminue bien (et admet une limite) lorsque la durée du projet de gestion augmente. Les auteurs étudient également l'influence de certains paramètres. Ils observent que lorsque le taux de croissance  $r$  est élevé, il est alors avantageux d'avoir un projet court (sinon de nombreux individus sont produits pendant la durée du projet). Si  $q$  est important, alors les efforts de contrôle peuvent être appliqués à une grande intensité sans perte significative de rendement, il est alors optimal d'avoir un contrôle élevé.

Les auteurs s'intéressent ensuite à la simplification du modèle général en un modèle spatial. Le but est ici d'éradiquer l'espèce sur une zone particulière (par exemple dans un parc national). L'objectif est de savoir si l'effort de gestion optimal doit être concentré uniquement sur la zone à conserver ou s'il doit également être appliqué à l'extérieur. Les auteurs se demandent aussi si l'effort est à appliquer uniformément sur la zone, et dans le cas contraire savoir s'il doit augmenter ou diminuer avec l'éloignement au centre de la zone de conservation. Pour cela, les auteurs supposent que le système est à l'état d'équilibre :  $\frac{\partial n}{\partial t} = 0$ , ce qui est obtenu après un temps long d'invasion. Le contrôle considéré dépend donc uniquement de l'espace. De plus, ils considèrent le problème comme étant radialement symétrique, ce qui restreint la dynamique à une variable spatiale réelle. Les auteurs testent notamment l'efficacité d'une zone tampon autour de la zone de conservation.

Les auteurs trouvent que le contrôle optimal accorde un effort plus important au centre de la zone de conservation, est décroissant avec l'éloignement du centre de cette zone, mais qu'il ne se limite pas à la zone de conservation. Ils remarquent que la densité de population augmente avec l'éloignement du centre de la zone de conservation. La stratégie qui consiste à répartir l'effort sur tout le domaine n'est pas optimale, tout comme une répartition uniforme (cas d'une zone tampon avec un effort constant). Cependant, comme sur le terrain les zones tampon avec effort uniforme sont pratiquées, les auteurs étudient la taille optimale que doit avoir cette zone. Ils constatent que la taille de la zone tampon optimale diminue lorsque la taille de la zone de conservation augmente.

### 1.2.2 Le modèle individu-centré de Harris et al. (2009)

Harris et al. (2009) ont construit un modèle stochastique pour l'invasion de Rhododendron, reposant sur la dispersion de graines par le vent (son principal mode de dispersion). Ils étudient l'efficacité de différentes stratégies de contrôle. Ils comparent des actions menées au cœur de la population (anciennes plantes) ou en périphérie (nouvelles plantes). Ils analysent également l'impact d'une stratégie opérant un retour sur les zones déjà traitées, mais qui ont été recolonisées depuis. Ces stratégies sont comparées à une stratégie basée sur l'âge, qui consiste à retirer les plus vieilles plantes (sous l'hypothèse qu'elles produisent plus de graines). Enfin, les auteurs analysent plus particulièrement l'efficacité de "corridors" dans lesquels on considère que l'espèce envahissante ne peut pas s'implanter. Pour ce qui concerne l'action de contrôle, les auteurs font l'hypothèse que le nombre de plantes retirées par année est constant, et ils supposent que l'effort pour retirer une plante ne dépend pas de son âge.

Le modèle est individu-centré, spatialement explicite, et rend compte de la dynamique de la population issue d'une seule plante. Chaque année, et pour tous les individus, l'âge est incrémenté, et si la plante est mature elle produit des graines, qui sont dispersées selon une loi donnée. L'ordre dans lequel les individus sont mis à jour est aléatoire et suit une loi uniforme. Le noyau donnant la distance de dispersion des graines est paramétré par la distribution quotidienne moyenne des vitesses de vent, la hauteur à laquelle la graine est lâchée, et le nombre

de graines estimé à partir de l'âge de la plante. Le modèle est paramétré avec des données issues de forêts en Écosse.

Le modèle étant stochastique, les résultats concernent des proportions d'événements sur 50 répliquations. Les auteurs étudient le succès d'éradication en fonction de la durée de l'invasion avant le début des actions de contrôle. Si l'invasion a pu se dérouler pendant 20 ans sans contrôle, ils constatent que l'éradication est possible avec toutes les stratégies de gestion (centre ou périphérie, avec ou sans retour, et selon l'âge). Néanmoins, ils notent que la stratégie d'action en périphérie avec retour est celle qui prend le plus de temps pour éradiquer l'espèce. Ils vérifient aussi que plus l'invasion a perduré sans contrôle, plus l'effort nécessaire pour l'éradication augmente, et constatent que les stratégies sans retour sur les zones recolonisées sont plus efficaces que celles qui opèrent un retour. Enfin, ils observent que la stratégie la plus efficace est celle basée sur l'âge (retirer les plus vieilles plantes).

L'étude s'intéresse ensuite à l'impact d'un corridor sur l'invasion (zone sur laquelle on considère que les rhododendron ne s'implante pas), les auteurs étudient en particulier l'influence de la longueur de ce dernier. Pour cela, ils observent le nombre de fois pour lesquelles au moins un rhododendron a réussi à franchir le corridor, en fonction du temps, pour différentes vitesses moyennes de vent et différentes longueurs de corridor. Ils constatent que pour une longueur de 50 mètres et une vitesse moyenne de vent de 9m/s, il faut attendre au moins 10 ans pour qu'une plante s'établisse de l'autre côté du corridor. Sur ce même cas, ils observent qu'après 40 ans, au moins une plante s'est systématiquement établie de l'autre côté du corridor. Les résultats sont similaires pour des vitesses moyennes du vent plus élevées.

Enfin, les auteurs s'intéressent à la provenance des graines sur le front d'invasion. Ils trouvent qu'au début de l'invasion (avant 80 ans), les vieilles plantes situées au cœur de la zone envahie sont responsables d'un grand nombre de graines qui arrivent au front d'invasion, mais qu'après 100 ans, la source principale s'est déplacée.

### 1.2.3 Analyse de la sensibilité des paramètres d'un modèle déterministe et d'un modèle individu centré

Travis et al. (2011) proposent deux modèles pour l'invasion par le rhododendron. L'un repose sur une approche individu-centrée proche du modèle présenté à la Section 1.2.2, tandis que l'autre est déterministe. L'article étudie l'importance de la paramétrisation sur le modèle. En particulier, les auteurs analysent l'influence du nombre de classes d'âge choisi pour le modèle sur la vitesse limite de propagation du front de l'invasion. Les auteurs utilisent le même noyau de dispersion des graines dans les deux modèles.

Le modèle déterministe est à temps discret et décrit l'évolution de la densité notée  $n$  d'individus en position  $x$  :

$$n(x, t + 1) = \int_{-\infty}^{+\infty} (K(x - y)B_n)n(y, t) dy. \quad (1.3)$$

Dans l'Équation (1.3),  $n$  est un vecteur de classes d'âge.  $B_n$  décrit la croissance de la population en position  $y$ , et est choisi indépendant de la densité, et est noté  $A$  de ce fait. La matrice  $A$  donne la fécondité des individus dans chaque classe, et décrit le passage des individus d'une classe à l'autre. Cette modélisation rend aussi compte des individus qui meurent ou restent dans la même classe. Enfin,  $K$  est le noyau de dispersion.

Le modèle individu-centré stochastique est identique à celui développé par Harris et al. (2009) (voir Section 1.2.2). On note toutefois l'ajout par les auteurs d'un taux de mortalité pour certaines classes d'âges.

Pour les deux modèles, les auteurs observent que lorsque le nombre de classes d'âge augmente, la vitesse de propagation de l'espèce diminue. Les auteurs constatent également que la vitesse de propagation dans le modèle déterministe est plus importante que celle dans le modèle stochastique. Enfin, les auteurs réalisent une analyse de sensibilité aux paramètres. Pour le modèle déterministe, ils constatent que la vitesse limite de propagation est plus sensible au paramètre de survie et au taux de transition lorsque qu'il y a peu de classes d'âge. Pour le modèle stochastique les résultats sont trop bruités et ne permettent pas de conclure.

### 1.2.4 Le modèle statique de Lookingbill et al. (2014) pour la ré-invasion de l'Ailante

Enfin, Lookingbill et al. (2014) proposent une approche statique pour déterminer les régions prioritaires à traiter afin d'éradiquer une espèce ré-envahissante. L'objectif est donc de fournir une carte qui donne le risque de ré-invasion de l'espèce si on l'éradique sur un territoire. Le cas d'étude concerne l'Ailante dans un parc national Américain.

Le modèle repose sur un découpage utilisé pour la gestion et tient compte de trois échelles pour estimer la réinvasion d'une zone. Au niveau global, il intègre le niveau d'envahissement qui permet de calculer la pression de propagule. A une échelle intermédiaire, il rend compte des vecteurs de propagation des graines (route, rivière) et de la densité envahie des aires voisines pour chaque zone. A l'échelle locale, il tient compte de la probabilité qu'a une graine de se développer, qui dépend de la qualité du sol de chaque zone.

La pression de propagule voisine est calculée pour chaque zone, elle dépend d'un indice de qualité de l'habitat des zones voisines ainsi que d'une probabilité de dispersion d'une graine par le vent entre deux zones. D'autres moyens de dispersion sont pris en compte, comme la présence de routes, de rivières, auxquels sont associés un poids d'importance. Chaque zone a donc une probabilité de recevoir des graines par ces autres moyens de dispersion. L'indicateur de l'échelle intermédiaire consiste en la moyenne pondérée des pressions de propagules pour les différents moyens de dispersion. La qualité de l'habitat d'une zone est calculée par une distance à un habitat idéal sur des critères tels que l'ensoleillement ou le pH du sol.

A partir des quantités introduites, les auteurs définissent et calculent le risque de réinvasion en chaque zone. La carte obtenue permet d'identifier celles prioritaires pour le traitement. Les auteurs étudient aussi l'effet d'un changement dans la valeur des poids pour les facteurs de dispersion. Ils trouvent alors que l'ordre des risques de réinvasion des différentes zones n'est pas trop altéré.

## 1.3 Végétation riparienne

Dans la plupart des articles de la Section 1.2, la dispersion considérée se fait par le vent. Par exemple, Harris et al. (2009) utilisent le noyau "Windissiper" pour connaître la distance de dispersion des graines de Rhododendron. Ce noyau requiert notamment la connaissance de la distribution quotidienne des vitesses moyennes de vent et la hauteur à laquelle la graine est lâchée. Travis et al. (2011) utilisent le noyau "Wald" pour la dispersion des graines. Le noyau est donné par :

$$\rho(x) = \left( \frac{\lambda}{2\pi x^3} \right)^{\frac{1}{2}} \exp \left( -\frac{\lambda(x - u')^2}{2u'x} \right),$$

où  $\lambda$  et  $u'$  sont des paramètres qui ne dépendent que de la vitesse du vent, la hauteur de chute de la graine et sa vitesse terminale.



Le cas qui nous intéresse dans la thèse est celui de la dispersion d'une plante envahissante le long d'un cours d'eau. Dans cette partie, nous nous concentrons donc plus particulièrement sur ce mode de dispersion.

### 1.3.1 Généralités

Les cours d'eau sont des vecteurs importants dans la dispersion des espèces envahissantes (Richardson et al., 2007). En particulier, les inondations, et plus généralement l'hydrologie des cours d'eau jouent un rôle important dans la dissémination des propagules et la distribution des espèces. Comme le soulignent Planty-Tabacchi et al. (1996), les niveaux d'invasions dans les milieux riverains sont positivement corrélés aux niveaux de perturbations de ces milieux. Ainsi, les perturbations anthropiques, du fait des installations humaines le long de cours d'eau, y favorisent les invasions végétales (références citées dans Richardson et al. (2007), paragraphe : "Establishment of new (alien) plant species"). De plus, les auteurs soulignent qu'il n'est pas toujours clair que la présence d'une espèce envahissante soit associée au déclin de la biodiversité. Ils en déduisent qu'il peut être préférable de porter l'effort sur le maintien d'un faible niveau de perturbation plutôt que sur le maintien de la diversité. Cependant, les auteurs affirment que la plantation d'espèces natives dans ces espaces devrait accélérer le processus de restauration, au moins dans les espaces très altérés.

### 1.3.2 Noyaux de dispersion par hydrochorie

Nous avons trouvé des noyaux de dispersion par hydrochorie pour des graines et pour des fragments de tiges, nous en présentons deux ici. Ces deux noyaux sont calibrés et testés avec des données de terrain. Les noyaux donnent pour chaque distance la probabilité qu'une propagule s'immobilise après avoir parcouru cette distance dans le cours d'eau. Les espèces considérées pour la calibration sont autres que la renouée, mais ces modèles ont une portée générale, et nous pouvons examiner dans quelle mesure les résultats présentés pourraient s'appliquer à la renouée.

Riis and Sand-Jensen (2006) cherchent à caractériser la dispersion de fragments de tiges pour des petits cours d'eau. Pour cela ils mesurent les distances parcourues par des tiges issues de deux espèces (une qui flotte et l'autre moins) qu'ils lâchent dans 4 rivières. Si on note  $N_0$  le nombre de fragments de tiges lâchés, ils trouvent que le nombre de tiges ayant parcouru une distance  $x$  vaut  $N_x = N_0 e^{-kx}$ , quantité qui dépend du paramètre  $k$  ( $m^{-1}$ ), appelé coefficient de rétention. Les auteurs s'intéressent à l'influence de caractéristiques des cours d'eau (profondeur, largeur, place occupée par la végétation dans le cours d'eau) sur ce coefficient de rétention  $k$ . Ils trouvent que le coefficient de rétention  $k$  varie significativement entre les rivières pour l'espèce qui flotte moins bien, mais n'a pas eu d'écart significatif pour les tiges provenant de l'autre espèce. Le rhizome de la renouée ne flottant pas bien (voir Section 1.6.1), nous précisons les résultats pour l'espèce de l'expérience correspondant à cette caractéristique. Les auteurs trouvent que (pour les deux espèces) la largeur du cours d'eau libre (non occupé par des plantes aquatiques), et la vitesse de l'eau sont les deux facteurs qui expliquent le mieux la variabilité entre les différents cours d'eau de la valeur du coefficient de rétention. Ils trouvent également que l'espèce flottant moins bien a été plus retenue en profondeur par des plantes submergées et des obstacles. Enfin, les auteurs se sont intéressés à la distance maximale parcourue par 90 % des tiges de *E.canadensis* (espèce flottant le moins bien). Ils trouvent 19 m (resp. 658 m) pour la plus petite (resp. grande) des valeurs sur les différents cours d'eau. Groves et al. (2009) proposent un noyau de dispersion pour les graines

dispersées par hydrochorie, inspiré de modèles existants pour la dispersion par le vent. Les deux paramètres clefs du noyau de dispersion sont la vitesse d'écoulement de la rivière et la largeur de la rivière (moyenne). Le noyau, moyennant la connaissance de la vitesse du cours d'eau et de sa largeur, donne donc la probabilité qu'une graine parcoure une certaine distance. Les auteurs proposent un noyau gaussien, mais non linéaire en la vitesse contrairement à la dispersion par le vent, qu'ils ont vérifié et calibré avec des akènes de tournesol. Les expériences de terrain montrent que les graines ont tendance à s'accumuler près de l'endroit où elles ont été déposées, mais avec une longue queue de distribution. De plus, la croissance de la vitesse augmente la distance moyenne de dispersion. Toutefois, les auteurs soulignent une limitation importante du modèle, qui n'est valable qu'en l'absence d'inondation (la dynamique de dispersion n'est pas la même).

Les deux modèles identifient la vitesse du cours d'eau comme étant explicative des différences de patron de dispersion entre les cours d'eau. Ces modèles renseignent donc sur la distance que va parcourir une propagule. Des connaissances supplémentaires sont nécessaires pour savoir ce que la propagule va devenir : si elle va germer et persister, ou ne pas donner naissance à une plantule.

### 1.3.3 Deux modèles incluant une dispersion par l'eau

Dans ce paragraphe, nous présentons deux modèles de dispersion d'espèces végétales à l'échelle régionale. Ces modèles prennent en compte plusieurs vecteurs de dispersion, dont celui par les cours d'eau. Les deux modèles sont spatiaux et stochastiques, et utilisent des informations sur la démographie (probabilité de survie, nombre de graines produites) de l'espèce considérée, ainsi que des informations sur les vecteurs de dispersion (eau, vent, route) et des données géographiques sur les types d'habitat. Les deux modèles divisent l'espace en cellules et donnent le nombre de plantes présentes dans chaque cellule. Les deux modèles ont été implémentés et testés avec des données de terrain. Le logiciel PSS (Plant Spread Simulator) est associé au modèle présenté dans Fennell et al. (2012), le cas d'étude qui y est développé porte sur la dispersion de graines de *Gunnera tinctoria* sur la côte ouest de l'Irlande (région de 50 x 30  $km^2$  couverte par 17000 cellules de 300 m de côté). Fox et al. (2009) ont créé une bibliothèque Python et ont testé leur modèle avec la dispersion du Chilean Needle Grass sur un territoire d'environ 100  $km^2$  en Australie, divisé en cellules de 100m de côté. Ainsi, Fox et al. (2009) utilisent leur modèle pour évaluer l'efficacité de stratégies de surveillance alors que Fennell et al. (2012) cherchent à identifier les paramètres du modèle qui expliquent le mieux la dispersion de la plante, ce dernier modèle a également une visée prédictive du devenir de l'invasion. Les modèles gèrent la dispersion à courte et longue distances, mais diffèrent sur la manière de déterminer la distance de dispersion : Fox et al. (2009) utilisent des noyaux exponentiels et gaussiens (similaires à ceux présentés dans la Section 1.3.2), alors que Fennell et al. (2012) utilisent le principe du "tournament selection algorithm". Cet algorithme repose sur le calcul du potentiel de dispersion pour un nombre déterminé de cellules se situant à une distance inférieure à la distance maximale de dispersion, par rapport à une cellule donnée. Le potentiel de dispersion est calculé à partir des probabilités de dispersion d'une cellule à une autre et des probabilités d'établissement de l'espèce dans les cellules concernées. Fennell et al. (2012) étudient plus particulièrement les variables qui expliquent le mieux (i.e. si on retire cette variable le modèle ne reproduit plus bien les données observées) la dispersion de la plante. Ils trouvent que ce sont le type d'habitat (terre agricole, marécage,...) et la présence de routes ou rivières qui expliquent le mieux la distribution des plantes. Quant à Fox et al. (2009), ils étudient différents types de surveillances qui correspondent à la visite

d'un nombre fixé de cellules. Ainsi, ils comparent l'efficacité de la surveillance systématique dans une région donnée, d'une surveillance aléatoire, de la technique "seek and destroy", d'une surveillance dont le choix des zones concorde avec les directions de propagation, et enfin de la surveillance des habitats préférentiels pour la reproduction de la plante. A chacune de ces techniques est associée une version adaptative, qui utilise les informations de présence lors de la détection au pas de temps (année) précédent, et consiste à chercher l'année suivante dans un rayon autour de ces cellules. Pour la comparaison des techniques, ils imposent un nombre de cellules inspectées identique entre les méthodes. Ils montrent que la version adaptative des surveillances a toujours de meilleures performances que la technique standard correspondante. Ils observent également des différences significatives dans l'efficacité des techniques selon que l'habitat préférentiel de l'espèce soit continu dans le paysage ou fragmenté. Enfin, notons que pour déterminer le nombre maximum d'individus qui peuvent exister au sein d'une même cellule, Fennell et al. (2012) divisent l'aire d'une cellule par l'aire d'une plante mature (le nombre obtenu est divisé par 2 en cas de présence dans la cellule de caractéristiques telles que des habitations qui réduisent l'aire disponible pour la plante). Nous utiliserons une démarche analogue pour estimer le nombre maximum de taches de renouée qui pourront être présentes le long d'un cours d'eau.

Cependant, ces modèles étudient la dispersion au niveau régional et ne modélisent pas de manière fine la dispersion le long d'un cours d'eau.

## 1.4 Généralités sur les renouées asiatiques

Les renouées asiatiques sont parmi les espèces envahissantes menaçant le plus la biodiversité et posent des problématiques de gestion importantes. Ce complexe est composé de trois espèces : la renouée du Japon, *Fallopia japonica* [Houtt.] Ronse Decraene, la renouée de Sakhaline, *Fallopia sachalinensis* [Schmidt Petrop.] Ronse Decraene (moins commune que la Renouée du Japon, elle possède des fleurs fertiles), et l'hybride entre les deux précédentes, la renouée de Bohême, *Fallopia* × *bohemica* Chrtek & Chrtkova (fertile). Les renouées sont originaires de l'Asie du Sud-Est et ont été introduites en Europe au 19<sup>ème</sup> siècle, mais sont désormais envahissantes en Amérique du Nord et en Europe du Nord et Centrale (Bailey and Wisskirchen, 2006; Barney et al., 2006; Beerling et al., 1994; De Waal, 2001). Elles sont aussi présentes en Australie, en Nouvelle Zélande et au Chili (Alberternst and Böhmer, 2006; Saldaña et al., 2009).

Les renouées asiatiques envahissent rapidement l'environnement dans lequel elles poussent (Gowton et al., 2016) et ont des impacts importants (Lavoie, 2017). Elles supplantent d'autres espèces végétales par la compétition lumineuse et l'allélopathie (Dommanget et al., 2014; Siemens and Blossey, 2007), affectent la diversité de la faune locale (Abgrall et al., 2018; Gerber et al., 2008; Maerz et al., 2005; Serniak et al., 2017) et modifient le fonctionnement des écosystèmes (Dassonville et al., 2011; Tharayil et al., 2013). Elles sont également capables de percer l'asphalte, ce qui est la cause de détériorations d'équipements routiers et s'ajoute à des problèmes liés à la visibilité, les tiges peuvent atteindre 3 mètres de hauteur (Dommanget et al., 2015). En outre, les coûts de contrôle sont très élevés, estimés à 250 millions de dollars par an en Grande-Bretagne (Colleran and Goodall, 2014) et à plus de 2 milliards d'euros par an en Europe (Kettunen et al., 2009).

Les renouées asiatiques poussent dans une grande variété de sols : sableux, marécageux et rocheux. Elles envahissent principalement les habitats modifiés par l'homme, tels que les bords de route, les décharges, mais aussi les berges des rivières. Ce sont des géophytes pérennes : leurs rhizomes leur permettent de passer l'hiver enfouis dans le sol (De Waal, 2001). Leurs rhizomes jouent également un rôle majeur dans leur propagation, grâce à leur forte capacité de régénération (Bailey et al., 2009; Brock et al., 1992) (ce point sera développé dans la Section 1.6.1). Une fois arrivé dans une nouvelle zone, le rhizome se développe et une nouvelle tache peut s'établir durablement en quelques semaines (Gowton et al., 2016; Smith et al., 2007). Les mécanismes de croissance locale (formation d'une tache) sont détaillés dans la Section 1.5. La Figure 1.1 montre la partie aérienne d'une renouée du Japon, après une année de croissance. Sur la Figure 1.2, on voit le réseau de rhizomes issu de la croissance d'un fragment de rhizome après une année.

Une fois établies, les renouées asiatiques sont extrêmement difficiles à éliminer en raison de leur vaste réseau de rhizomes. Les rhizomes représentent les deux tiers de leur biomasse (Barney et al., 2006) et peuvent s'étendre à plusieurs mètres du front d'invasion visible (Barney et al., 2006). Le rhizome peut atteindre jusqu'à 8 cm de diamètre, et peut s'étendre sur 15 à 20 mètres (Barney et al., 2006). Il est aussi capable de s'enfouir entre 2 et 3 mètres de profondeur. Les ressources stockées par les renouées peuvent être efficacement remobilisées après des événements de fauche (Rouifed et al., 2011a). Certains auteurs estiment que six fauches sont nécessaires pour réduire de manière significative la biomasse souterraine (Gerber et al., 2010). Il est donc crucial de comprendre le développement souterrain des renouées asiatiques pour avoir une idée de leur croissance locale et de leurs performances.

Dans Dommanget et al. (2013), les auteurs ont montré que l'intensité de lumière disponible



FIGURE 1.1 – Renouée du Japon, partie aérienne (photo personnelle)



FIGURE 1.2 – Renouée du Japon, rhizome

pour la croissance de la plante est le facteur le plus important pour expliquer sa performance (en terme de biomasse aérienne produite). Les auteurs ont ensuite testé une technique de gestion pour contrôler les renouées basée sur la compétition par la lumière par une espèce à croissance rapide. Il s'agit d'opérer une fauche sur les renouées et ensuite de faire du bouturage de Saules sur des taches de renouées asiatiques. L'expérience a montré des résultats positifs. En effet, on a observé une réduction significative de la vigueur des renouées asiatiques due à une revégétalisation par des boutures de saule et une action de fauche. Une technique de gestion couramment utilisée est la fauche. Les gestionnaires peuvent faire varier son intensité et sa fréquence. Par exemple, dans Gerber et al. (2010) les auteurs ont comparé la biomasse résultante après une seule coupe ou une coupe mensuelle pendant 6 mois. Les résultats obtenus après 4 années montrent que la coupe régulière des *Fallopia* pendant la période de végétation est la méthode la plus efficace. La biomasse des *Fallopia* est réduite de 98%. Ils ont également testé l'application d'un géotextile à base de fibres végétales (précédée par une coupe pendant la première année). Cette technique réduit la biomasse des *Fallopia*, mais les nouvelles pousses de renouées réussissent à percer ce matériel. Martin et al. (2020) montrent l'importance d'une fauche répétée sur l'entièreté de la tache car la plante semble nettement moins affectée par une fauche partielle. Ils mettent aussi en évidence une croissance différente de la plante lorsqu'elle pousse en pleine lumière (stratégie "phalange" d'occupation de l'espace) et lorsqu'elle est entièrement ombragée (stratégie "guérilla" de recherche de lumière). De plus, les herbicides (glyphosphate) n'ont qu'une efficacité à court terme, de plus ils présentent des risques écologiques. Il est également possible d'excaver les parties souterraines (rhizome), mais ces techniques sont très coûteuses et doivent être réalisées avec une grande attention pour ne pas disperser des morceaux de rhizome, et ainsi contribuer à l'invasion de la plante. Dans Gourley et al. (2016), les auteurs développent un modèle de lutte contre la renouée du Japon basée sur un insecte qui affaiblit la plante en se nourrissant de sa sève. Un paramètre déterminant de leur modèle est la durée que prend la larve pour consommer et digérer la biomasse du rhizome de la plante. Nous décrirons plus précisément ce modèle dans la Section 1.5. Il existe aussi d'autres méthodes de gestion comme le pâturage ou le brûlage.

## 1.5 Croissance locale de la renouée du Japon

Dans cette section, nous nous intéressons à la croissance locale (donc au niveau d'une tache) de la renouée du Japon. Nous décrivons d'abord les mécanismes intervenant dans sa croissance, puis nous présentons des modèles qui visent à reproduire ce comportement.

### 1.5.1 Description des mécanismes de croissance locale

Adachi et al. (1996a) décrivent la croissance et l'organisation du rhizome. Ces observations sont faites sur le mont Fuji, et si des similitudes existent avec les mécanismes de développement de la plante en Europe (Adachi et al., 1996a; Smith et al., 2007), les auteurs remarquent des différences au niveau des durées de vie des couronnes (plus courtes au Japon qu'en Europe), et de la qualité du sol (désertique au mont Fuji). De plus, le phénomène de dépérissement central (densité moins importante de tiges au centre) est observé au Japon mais pas en Europe (Smith et al., 2007). Adachi et al. (1996a) soulignent qu'en périphérie du massif, les rhizomes grandissent horizontalement, alors que dans la partie centrale, ils croissent de manière oblique, et s'enfoncent donc. Le diamètre du rhizome est plus important au centre de la tache (5 cm) alors qu'en périphérie les mesures varient entre 5 et 10 mm. Les rhizomes produisent des bourgeons latéraux, soumis à la dominance apicale (la définition de ce terme est donnée dans le glossaire à la fin de cette section) et qui ne débourent que lorsque la couronne mère est sectionnée ou morte. Les tiges aériennes produisent également chaque année, à la fin de leur croissance, des bourgeons hivernaux à leur base. Seules les tiges qui achèvent leur croissance produisent des bourgeons hivernaux, qui émergent presque à la même position que la tige précédente. C'est cet ensemble de tiges proches que l'on appelle couronne. Pour la partie souterraine, les rhizomes grandissent latéralement depuis une racine pivotante qui s'enfonce dans le sol. Les extrémités des rhizomes forment éventuellement une tige aérienne. C'est à la mort de la couronne que certains bourgeons latéraux poussent et forment de nouveaux rhizomes, à l'extrémité de ces rhizomes se trouve un bourgeon terminal, qui forme une nouvelle couronne.

La Figure 1.3, extraite de Adachi et al. (1996a), montre un schéma du développement des tiges et des bourgeons le long du rhizome.

On présente maintenant un glossaire qui reprend les définitions importantes sur les plantes clonales (extraites de Kroon and Groenendael (1997)) :

- Rhizome : tige souterraine horizontale.
- Bourgeon terminal : bourgeon situé à l'extrémité d'une tige aérienne.
- Bourgeon latéral : bourgeon situé le long du rhizome et qui peut potentiellement faire grandir la plante, par la croissance du rhizome.
- Dominance apicale : l'influence exercée par un bourgeon terminal (apical) inhibant la croissance des bourgeons latéraux.
- Reproduction clonale : formation d'unités génétiquement identiques depuis une unique cellule somatique.
- Ramet : la plus petite unité d'un clone qui peut être indépendante (une tige avec sa racine).
- Intégration : processus de partage de ressource entre des ramets interconnectés d'un clone.



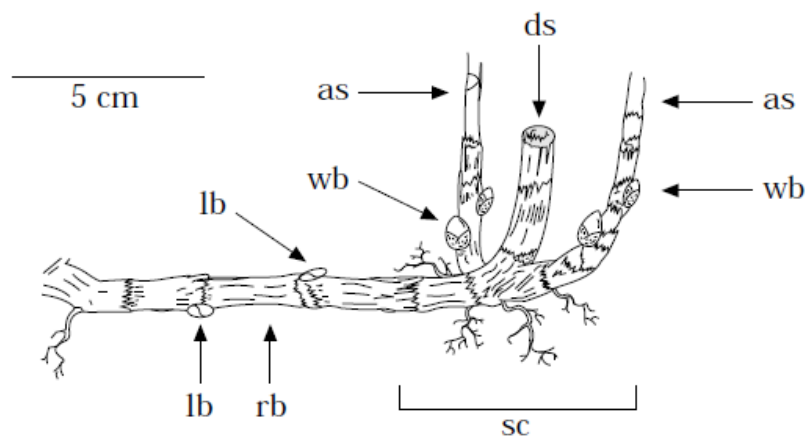


FIGURE 1.3 – Schéma du développement des tiges et des bourgeons le long du rhizome, extrait de Adachi et al. (1996a).

Légende : as: current aerial shoot, ds: dead aerial shoot of previous or earlier year, rb: rhizome branch, sc: shoot clump (couronne), lb: lateral bud along the rhizome, wb: winter bud formed at the bas of an aerial shoot.

### 1.5.2 Trois modèles pour la croissance locale de la renouée issus de la littérature

Comme les parties souterraines d'une tache de renouée sont en pratique inaccessibles aux observateurs, les observations directes sont rares et les modèles peuvent donc aider à aborder leur dynamique et à mieux comprendre comment les mesures de gestion peuvent affecter leur développement. À notre connaissance, il existe très peu de modèles dans la littérature qui décrivent la croissance d'une tache de renouées asiatiques, et parmi eux, rares sont ceux qui incluent une technique de gestion.

Dans cette partie nous en détaillons trois, que nous introduisons succinctement ici. Dans Smith et al. (2007), les auteurs ont construit un modèle de marche aléatoire corrélé en 3D pour le développement du réseau de rhizomes souterrains d'une tache de renouée du Japon. Leur modèle est construit à partir de connaissances sur la morphologie et la physiologie de la plante. Ils ont étudié le modèle par des simulations et ont observé une expansion quadratique de l'aire envahie. Dauer and Jongejans (2013) ont construit un "Integral Projection Model" pour la dynamique de la renouée du Japon au niveau de la tache. Ils s'intéressent en particulier à la hauteur ou la biomasse totale de la plante. Enfin, Gourley et al. (2016) ont développé un modèle mathématique pour le biocontrôle de *Fallopia japonica* reposant sur l'insecte *Aphalara itadori*. Ils décrivent la dynamique des insectes (larves et adultes), celle du poids total des tiges de renouée ainsi que l'évolution de la biomasse du rhizome.

#### Le modèle de Smith et al. (2007)

Smith et al. (2007) proposent un modèle pour le développement du réseau de rhizomes d'une tache. Ce modèle ne comprend pas de stratégie de gestion, il modélise le comportement d'une tache seule sans dispersion. Les auteurs se sont inspirés d'un modèle existant (Adachi et al., 1996b), qui modélise la croissance de la renouée sur le mont Fuji au Japon. Cependant, le modèle proposé par Adachi est en deux dimensions spatiales, et surtout, l'évolution de la plante sur sa terre native est différente de celle en Europe (phénomène de dépérissement

central observé sur le mont Fuji, voir Section 1.5.1). Comme dans Adachi et al. (1996b), des données sur la morphologie et la physiologie de la plante sont utilisées pour paramétrer le modèle. Il reste tout de même trois paramètres non estimables avec les données de terrain, comme par exemple la probabilité qu'une couronne se forme. Ces paramètres sont alors estimés de sorte qu'en moyenne la densité de couronne simulée soit comparable à celle de leurs données de terrain.

Le modèle est construit selon une marche aléatoire corrélée pour le développement du rhizome. La corrélation vient du fait que la direction de branchement du rhizome fils (donnée par l'angle entre un segment de rhizome père et son fils) est corrélée à la direction du rhizome père. Ainsi, le modèle décrit la croissance du réseau de rhizome depuis un fragment initial, et traite de la formation des couronnes. Nous décrivons ici les mécanismes de croissance pris en compte par les auteurs.

Tous les segments de rhizome grandissent à la même vitesse (paramètre du modèle), mais leur longueur est tirée selon une loi obtenue de manière empirique (en mesurant la longueur de rhizomes excavés). Si un segment a atteint sa longueur prescrite, il cesse de grandir, produit un bourgeon et un nouveau segment de rhizome qui forme un angle avec le rhizome précédent. La direction de branchement est choisie de manière uniforme dans un cône formé à partir de la direction précédente et d'un angle maximal de déviation (estimé avec des données de terrain). Un exemple de direction de branchement est donné par la Figure 1.4. La longueur du nouveau fragment de rhizome est tirée avec la même loi que l'étape précédente. De plus, après un an, un deuxième morceau de rhizome est produit depuis ce bourgeon. Au niveau du bourgeon, on a donc une séparation en deux du rhizome, un bout qui continue de grandir, et un bout "fils" qui doit attendre un an avant de grandir. Un fragment de rhizome meurt (devient inactif) à chaque itération avec une probabilité fixée (paramètre du modèle) et de manière certaine lorsqu'il atteint la longueur prescrite. Au niveau de chaque bourgeon, une couronne est formée avec une probabilité fixée (paramètre du modèle). Cependant, la couronne n'est réellement formée que si le bourgeon est à une profondeur inférieure à 30 cm (issue de données de terrain) et s'il n'est pas dans une zone de compétition intraspécifique. En effet, les auteurs considèrent une zone de compétition intraspécifique autour de chaque couronne, de sorte que d'autres couronnes ne peuvent y pousser. la distance minimale entre deux couronnes est un paramètre du modèle. Nous modélisons ce phénomène de manière analogue dans le modèle développé à la Section 4. A chaque pas de temps du modèle, les auteurs choisissent de faire la mise à jour (croissance des segments de rhizome,...) dans un ordre aléatoire et de manière uniforme pour éliminer les biais qui pourraient résulter d'un ordre fixe. Enfin, notons que si un fragment de rhizome atteint la limite de profondeur de 2 mètres imposée par les auteurs, il ne subit qu'une croissance horizontale pour ne pas sortir du domaine considéré.

Les caractéristiques des taches simulées étudiées sont la densité de couronnes et l'aire de la tache (enveloppe convexe formée par les couronnes). La Figure 1.5 montre la simulation d'une trajectoire de réseau de rhizomes pour 70 itérations (23 ans). Les points noirs représentent les couronnes, les lignes grises désignent le réseau de rhizomes.



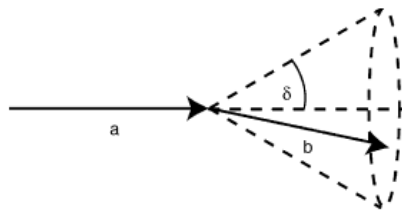


Fig. 7 – An example of the growth in one time-step. Here 'a' is the previous growth vector and 'b' is a potential choice of direction for this time-step. The dashed lines represent the maximum possible change of direction in this plane for this step and the continuation of the previous growth vector.

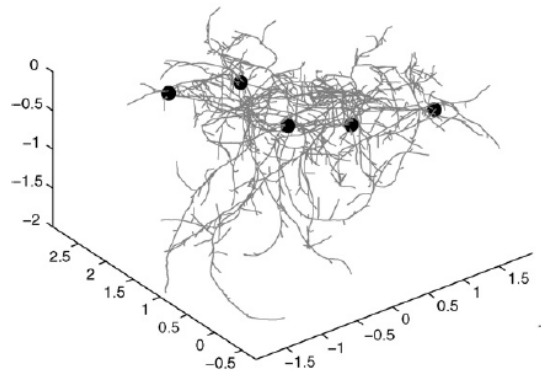


FIGURE 1.4 – Direction de branchement, extrait de Smith et al. (2007)

FIGURE 1.5 – Une simulation du modèle, extrait de Smith et al. (2007)

Les auteurs observent une expansion quadratique de l'aire de la zone colonisée ( $A$ ) par les rhizomes issue de la simulation de leur modèle :

$$A \approx \alpha t^2$$

Les auteurs étudient aussi l'influence des 3 paramètres pour lesquels ils n'avaient pas d'estimation (comme l'angle maximal de déviation ou la probabilité de former une couronne) sur la densité de tige et le paramètre  $\alpha$  (lié à la croissance quadratique).

### Le modèle de Dauer and Jongejans (2013)

Dauer and Jongejans (2013) utilisent un "Integral Projection Model" (IPM), paramétré en utilisant des données démographiques de terrain, pour décrire l'évolution de la biomasse totale de la plante et de sa taille. Pour leur modèle, les auteurs utilisent le cycle de vie simplifié de la plante représenté à la Figure 1.6. Il comprend deux états : nouvelles tiges et couronnes. Sur une année, les nouvelles tiges peuvent survivre et former une couronne, ou mourir. Les couronnes qui ne meurent pas restent des couronnes. Les nouvelles tiges et les couronnes peuvent former des nouvelles tiges par l'intermédiaire d'un rhizome.

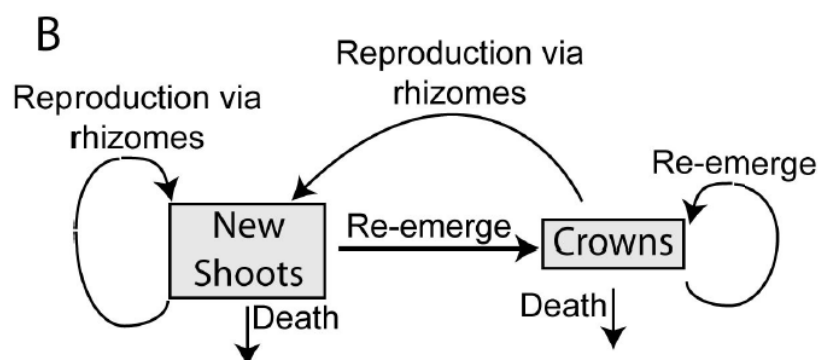


FIGURE 1.6 – Cycle de vie simplifié de la renouée utilisé dans le modèle IPM, extrait de Dauer and Jongejans (2013)

Les auteurs ont collecté et analysé les données de quatre sites (deux dans des parcs nationaux, un sur une parcelle privée, et le dernier sur un bord de route). Le modèle IPM est construit à partir de données sur la survie, la croissance et la fécondité des plantes observées. Pour chaque site, deux modèles IPM ont été réalisés, un avec la hauteur des tiges, et l'autre avec la biomasse des tiges, comme variable d'état.

Une analyse a été menée sur ces modèles pour identifier les plages de valeurs de paramètres qui ont le plus d'effet sur la variation de la plus grande valeur propre de la matrice (de grande taille) associée à l'IPM (appelée  $\lambda$ , "projected population growth rates"). Les auteurs mènent aussi une analyse pour déterminer la contribution des taux (fécondité, survie, croissance) sur le taux de croissance de la population. Ils montrent que ce sont les individus (tiges aériennes) ayant une taille entre 30 et 80 cm qui influent le plus sur le taux de croissance global de la population. De même, ils trouvent que la survie et la croissance impactent plus cette valeur que la fécondité. Enfin, c'est le ratio du nombre de nouvelles tiges, sur le nombre de rhizomes qui affecte le plus la valeur de  $\lambda$ .

### Le modèle de Gourley et al. (2016)

Gourley et al. (2016) proposent un modèle pour le biocontrôle de la renouée du Japon en utilisant l'insecte *Aphalara itadori*. Cet insecte se nourrit de la sève des tiges de la plante, l'affaiblissant ainsi. Les auteurs ne considèrent qu'une seule tache de renouée. La plante est décrite par deux variables temporelles : la biomasse totale des tiges ( $S$ ) et des rhizomes ( $R$ ). Ce sont les larves d' *A. itadori* qui endommagent le plus la plante. La population d' *A. itadori* est donc décrite par le nombre de larves ( $L$ ) et d'adultes ( $A$ ).

Les auteurs se placent dans un cadre déterministe et le formalisme adopté est celui des systèmes d'équations différentielles à retard. Ils déduisent l'évolution de ces quatre variables de la modélisation de l'évolution de densités de population dépendant de l'âge, aussi bien pour les larves que pour les tiges. Le modèle prend donc en compte la durée nécessaire au développement d'une larve *A. itadori* pour passer du stade d'œuf au stade d'adulte et la durée de vie d'une tige de renouée (respectivement les paramètres  $\tau_p$  et  $\tau_s$ ). Ainsi, les auteurs considèrent l'évolution d'une population où les individus sont différenciés par leur âge. La modélisation conduit à une Équation aux Dérivées Partielles structurée en âge. Précisément, ils utilisent des équations de transport linéaire.

Ils considèrent que les larves meurent à un taux qui dépend de la biomasse totale des tiges  $S(t)$ . Ils supposent ce taux décroissant, puisque la mortalité des larves augmente quand la taille totale des tiges diminue. Le taux de naissance de *A. itadori* dépend du nombre d'adultes  $A$  et de la biomasse totale des tiges  $S$ , et tient compte de la compétition intra-spécifique pour les larves issue du fait qu'une tige ne peut recevoir qu'un nombre fixé de larves. L'intégration de l'équation ainsi obtenue pour la densité de larves par rapport à l'âge sur  $[0, \tau_p]$ , puis sa différentiation par rapport au temps, permet d'obtenir l'équation régissant l'évolution de  $L$ . Le taux auquel les larves deviennent adulte est déduit de cette dernière étape, ce qui permet directement d'écrire l'évolution pour le nombre d'adultes ( $A$ ) en fonction du temps. Les auteurs y ajoutent un terme de mortalité intrinsèque.

Concernant l'évolution de la biomasse des tiges, la sève des tiges est consommée par les larves, ce qui est considéré comme étant le facteur le plus important dans l'endommagement des tiges (perte de biomasse). Les auteurs tiennent également compte d'une mortalité naturelle qui impacte la biomasse des tiges (taux constant). Comme pour les larves, les auteurs décrivent l'évolution de la densité (en âge et temps) de la quantité de biomasse des tiges par une équation de transport linéaire. Ils modélisent la perte de biomasse due à la larve par une réponse fonctionnelle d'Holling de type II (si la biomasse des tiges est élevée, le taux de consommation

de sève par larve atteint un plateau). C'est dans cette réponse fonctionnelle qu'intervient le paramètre  $h$  de durée de digestion par la larve d'une unité de biomasse consommée. C'est l'influence de ce paramètre sur la dynamique du système qui est ensuite particulièrement étudiée.

Le taux de production de biomasse de tige (les nouvelles tiges), dépend de la quantité de biomasse pour le rhizome souterrain. Puis, comme pour les larves, les auteurs intègrent la densité de biomasse de tiges par rapport à l'âge sur  $[0, \tau_s]$  (mortalité des tiges dès que l'âge devient  $\tau_s$ ).

Enfin, l'évolution de la biomasse du rhizome a une dynamique simple. La larve n'attaque pas le rhizome de la plante, et les auteurs ne considèrent qu'un taux de mortalité intrinsèque. En particulier, ils ne considèrent pas de contrepartie à la création de nouvelles tiges issues des ressources du rhizome. Par contre, le taux de production de nouveaux rhizomes dépend bien de la quantité de biomasse dans les tiges ( $S$ , en effet, la croissance du rhizome dépend de la photosynthèse des feuilles, qui est directement reliée à la quantité de tiges).

Les auteurs étudient d'abord le système dynamique sans la présence de *A. itadori* (conditions initiales nulles). Ils trouvent (Théorème 3.1) deux cas selon la valeur de paramètres (taux de naissance et de mortalité pour les tiges et rhizomes, et durée de vie d'une tige) : la croissance exponentielle de la tache de renouée ou sa décroissance exponentielle. En particulier, la tache croît si les taux de naissance des tiges et rhizomes sont suffisamment grands, et que les taux de mortalité sont suffisamment faibles (c'est en pratique le cas pour le rhizome). Les auteurs montrent ensuite (Théorème 3.2) que dans le cas de croissance exponentielle sans insecte, l'ajout de *A. itadori* ne permet pas d'éradiquer la tache (i.e. on n'a pas  $S(t) \rightarrow 0$  quand  $t \rightarrow \infty$ ). Après une analyse reposant sur la linéarisation des équations, les auteurs aboutissent à une conjecture sur une valeur critique de  $h$  en dessous de laquelle la solution du système serait bornée. Ils montrent par des simulations que la valeur seuil est en pratique plus élevée que celle conjecturée. Ils montrent (Théorème 3.4) également avec des hypothèses supplémentaires sur le taux de naissance de l'insecte que si  $h$  est suffisamment faible, alors le système admet un point équilibre, ce qui n'est plus le cas si  $h$  est trop grand.

Le contrôle de la renouée par biocontrôle avec l'insecte *A. itadori* est donc possible seulement si ce dernier consomme et digère la sève de la plante suffisamment rapidement.

## 1.6 Étude de la dispersion

Dans cette section, nous présentons des propriétés liées à la dispersion des renouées asiatiques.

### 1.6.1 Propagules pour la dispersion

Les renouées asiatiques se dispersent essentiellement de manière clonale par transport de tige ou rhizome, elles sont également capables de reproduction sexuée par les akènes (Barney et al., 2006). Le potentiel de reproduction sexuée des renouées asiatiques est important puisqu'une tige peut théoriquement produire plus de 100 000 akènes (fruits issus de la fertilisation de la fleur), observé aux États-Unis (Bram and McNair, 2004), alors qu'au Royaume-Uni le taux de fertilisation (entendu ici aussi comme le nombre d'akènes produits par tige) semble être 100 fois moindre (Bailey et al., 1994). Bien que 16 % des graines plantées dans des conditions contrôlées (en serres) ont germé et ont produit au moins un individu portant une fleur (Forman and Kesseli, 2003), la même étude identifie des conditions sur la croissance lors de la première année pour que la plantule créée survive l'année suivante. De plus, la germination d'une propagule issue d'un akène, puis la formation d'une plantule est beaucoup moins fréquente sur

le terrain que celle qui pourrait résulter de ces observations (Bailey et al., 2009; Martin et al., en préparation; Tiébré et al., 2007). Cela peut s'expliquer par la pression environnementale qui s'exerce sur ce type de propagule, entre autres la compétition par d'autres espèces et les conditions abiotiques (Martin et al., en préparation). En ce qui concerne la régénération par un fragment de rhizome, Bímová et al. (2003) montrent qu'elle est de 100 % dans certaines conditions (profondeur d'enfouissement), mais que le taux de régénération varie en fonction du sol et du taxon considéré. De plus, la régénération peut se faire à partir de très petits morceaux de rhizome. Pour ce qui est de la régénération à partir de morceaux de tige, Bímová et al. (2003) montrent que le taux de régénération des tiges enfouies dans des sols ou à la surface est de 20 %. Pour l'étude de la dispersion, les données dont on dispose quantifiant les populations "sources" étant bien souvent des surfaces colonisées, il est aussi intéressant de connaître l'ordre de grandeur du nombre de propagules produites par mètre carré de tache de renouée. De Waal (2001) estime à environ 250 propagules par mètre carré pour le réseau de rhizome. De plus, à partir de la donnée du poids minimal de rhizome qui peut permettre la création d'une plantule (moins de 1 g), Barney et al. (2006) trouvent un nombre plus important (environ 2000 propagules par mètre carré). L'étude de De Waal (2001) a aussi porté sur les fragments de tige, dont l'auteur conclut qu'entre 8 et 90 propagules peuvent être produites par tige.

Le contact des propagules avec l'eau lors d'un transport par un cours d'eau modifie leur taux de germination et de régénération.

Rouifed et al. (2011b) s'intéressent à l'hybride *Fallopia* × *bohemica*, dont les akènes sont fertiles. Ils veulent déterminer si ces akènes jouent un rôle dans la dispersion de la plante, et s'intéressent en particulier au vecteur de dispersion que représente l'eau pour les akènes de cette plante. Ils étudient en expérience contrôlée la flottabilité, la durée de germination et la probabilité de survie (une fois plantée) pour des akènes mis dans une cuve d'eau agitée robotiquement. Ils observent que la plupart des akènes flottent d'abord pendant une période de 2 à 5 jours, puis coulent. Ensuite, les akènes germent dans l'eau et reflottent. Les auteurs montrent que l'exposition à l'eau des akènes favorise significativement (par rapport aux graines plantées directement dans le sol) la germination (sans affecter la probabilité de survie une fois plantée). Ils concluent que ces propriétés des akènes dans l'eau augmentent considérablement les distances que ces akènes peuvent parcourir en comparaison aux distances en jeu lors de la dispersion par le vent. Lamberti-Raverot (2016) étudie les traits de dispersion et de colonisation précoce de la renouée et leur variabilité entre des taches distribuées sur différents cours d'eau. Dans la deuxième partie de sa thèse, l'auteur traite de la caractérisation écomorphologique des propagules sexuées et de leur variabilité à l'échelle régionale. L'auteur s'intéresse donc à la forme, taille, volume de la graine, sa surface en contact avec l'eau. L'auteur étudie le lien entre ces mesures et la capacité de flottaison des akènes. Ces observations très fines réalisées en laboratoire ont permis de montrer que la morphologie des akènes prélevés sur des taches de renouées poussant le long de rivières modifie leur capacité de flottaison. Cependant, en utilisant des données de terrain, ils n'ont pas trouvé de corrélation entre la capacité de flottaison des akènes provenant d'une tache, et leur capacité de dispersion (ni avec leur position sur la rivière). L'auteur a confirmé les observations de Rouifed et al. (2011b) et les a prolongées en comparant les durées de flottaison d'akènes issues de plusieurs taches. Il en résulte une grande variabilité entre ces dernières.

Il est important de noter que les rhizomes des renouées ne flottent pas. Néanmoins, ils peuvent être dispersés par hydrochorie. Bímová et al. (2003) ont montré que le taux de régénération de rhizomes est faible dans l'eau (moins de 20 %). Lamberti-Raverot (2016) a observé une grande variabilité sur la régénération de rhizomes issus de 10 taches sans contact avec de l'eau

(entre 25 et 95 %). De plus, les taux de régénération de rhizomes issus de taches différentes suite à une exposition à l'eau dans des conditions contrôlées sont plus faibles que dans le cas contrôle. Là encore, l'auteur a observé une grande variabilité entre les taches, mais après un mois d'exposition à l'eau, la moitié des rhizomes régénère encore. Pour ce qui est des fragments de tiges, Bímová et al. (2003) ont montré que la régénération était accrue pour des tiges qui flottent (>75 %). Lamberti-Raverot (2016) a de nouveau étudié la variabilité sur les mesures de flottaison et de viabilité entre taches pour des fragments de tige. Contrairement aux conclusions précédentes pour le rhizome et les akènes, la variabilité au sein d'une même tache était importante, et n'a pas permis de montrer de la variabilité entre tache. L'auteur constate une perte significative de viabilité et de capacité de flottaison des tiges après 3 semaines d'exposition à l'eau. L'expérience montre néanmoins que 10 % des tiges étaient encore en capacité de flotter après 2 mois. L'auteur souligne que dans ses expériences, plus de la moitié des fragments de tige avaient régénéré dans l'eau, ce qui leur donne un avantage lors de l'arrivée sur le site de dépôt.

La dispersion des propagules peut se faire à courtes et longues distances. Martin et al. (en préparation) soulignent que les rivières représentent le vecteur de dispersion naturel le plus observé pour les renouées. Cependant ils ajoutent qu'on ne connaît que peu de choses des dynamiques et des patrons de dispersion à courte distance par les rivières des renouées. Pour des petits cours d'eau, ils mentionnent les travaux de Barthod and Boyer (2019) qui ont trouvé que 85 % des nouvelles pousses de renouées se trouvent à moins 100 mètres de la plus proche tache de renouée voisine. Ces résultats ne devraient plus être valides pour des cours d'eau plus importants, Martin et al. (en préparation) énoncent donc l'importance d'améliorer notre connaissance de ces processus. Pour ce qui est de la dispersion à longue distance de propagules, la connaissance de leur capacité de flottaison ne donne pas d'indication sur la fréquence de tels événements. En effet, ces événements dépendent en grande partie de caractéristiques environnementales, qui évoluent avec le temps. Martin et al. (en préparation) attirent aussi notre attention sur le fait que nous ne connaissons pas l'allure des courbes de dispersion (densité en fonction de la distance), nous présenterons toutefois dans la suite de cette section deux noyaux de dispersion utilisés pour des graines et des tiges d'autres espèces.

### **1.6.2 Variables explicatives de la présence de renouées le long de cours d'eau**

Plusieurs études ont été menées pour déterminer les variables environnementales explicatives de la présence/absence et l'abondance de renouées le long de cours d'eau. Notons déjà que Rouifed et al. (2014) ont comparé au niveau régional les facteurs explicatifs de la présence de renouées en plaine versus montagne. Les auteurs montrent qu'en plaine, la présence de renouées est expliquée par des variables anthropogéniques (comme la présence d'une route) et naturelles (comme la présence d'une rivière), alors qu'en altitude, seules les variables anthropogéniques expliquent la présence de la plante. Les trois études citées dans la suite de ce paragraphe se placent au niveau de cours d'eau. Beerling (1991) étudie l'effet des différentes utilisations du sol sur le niveau de présence de renouées au bord de 5 rivières du sud du pays de Galles. L'auteur choisit 4 classes d'abondance pour les renouées (pourcentage de présence par secteur de 500 mètres), et 6 classes pour l'utilisation du sol : habitat naturel non perturbé avec des plantes natives bien établies, pâturages, forêt avec canopée continue, forêt avec canopée discontinue, terrain en friche, terrain modifié par des travaux en lien avec le cours d'eau. Il trouve que la plante est plus présente sur les terrains en friche, et moins présente dans les pâturages ou les habitats naturels. L'auteur constate aussi que là où la plante est déjà implantée, les taches

sont moins denses quand il y a une canopée continue. Il en déduit des implications sur une méthode de gestion qui pourrait consister à planter des arbres pour réduire l'accès à la lumière de la plante. Duquette et al. (2016) ont fait un travail similaire le long d'une rivière au Canada. Les auteurs ont cherché à identifier les facteurs expliquant la répartition spatiale des renouées. Les données pour les variables explicatives concernent l'abondance et le type de végétation sur la berge, le type d'occupation des sols voisins (industriel, résidentiel,...), la distance de la ville ou village, route, voie ferrée la plus proche,... Ils font ensuite une régression pour expliquer la présence et l'absence de renouées en fonction des variables explicatives mesurées. Ils trouvent que c'est la distance à la plus proche ville/village qui est la seule significative pour expliquer la présence de renouée. Ils trouvent aussi qu'un bon prédicteur de la présence d'une tache de renouée à un endroit est la présence d'une tache à proximité. Enfin, Descombes et al. (2016) ont mesuré sur un cours d'eau de 18 km en Suisse 11 prédicteurs environnementaux analogues à ceux de Duquette et al. (2016). Les auteurs disposent également de données moins précises mesurées 11 ans auparavant, et se sont intéressés à un ensemble d'espèces envahissantes incluant la renouée. Ils cherchent les prédicteurs qui expliquent le mieux la présence et l'abondance d'espèces par secteur à la seconde date. Ils observent d'abord une augmentation significative de 565 % du nombre de taches de renouée entre les deux dates. Pour la renouée, les auteurs trouvent que seules trois variables expliquent de manière significative la présence de la plante : la distance à la rivière, la densité de la canopée et la densité d'habitation. La présence de route et à nouveau la densité de la canopée sont les prédicteurs qui expliquent le mieux l'abondance de la renouée (essentiellement l'aire occupée par les taches). Les auteurs notent enfin que certaines populations d'une espèce présentes à la première date ont été remplacées par des populations de renouées, ce qui pourrait s'expliquer par de la compétition dans les habitats les mieux adaptés.

Ainsi, le long des cours d'eau, la renouée est plus présente sur des terrains en friche, mais est moins abondante dans les zones à forte canopée. De plus, la distance à la plus proche ville et la densité d'habitation semblent être des variables explicatives importantes de la présence de la renouée le long d'un cours d'eau.



## Chapitre

# 2

### **La modélisation comme outil d'aide à la décision**

Ce chapitre présente des outils méthodologiques issus de la modélisation pour l'aide à la décision étant définie comme "l'activité de celui qui s'appuyant sur des modèles clairement explicites mais non nécessairement complètement formalisés, aide à obtenir des éléments de réponses aux questions que se pose un acteur inclus dans le processus de décision." (Roy, 1985) Dans le cas de l'invasion par les renouées asiatiques, l'acteur est un gestionnaire en charge de gérer l'invasion. Le processus de décision concerne ses moyens d'actions tels que la fréquence et l'intensité des actions de gestion, leur répartition spatiale. L'invasion des renouées, que ce soit au niveau de sa croissance locale ou de sa dispersion, est un phénomène dynamique (voir Chapitre 1). Nous faisons donc le choix d'utiliser des modèles dynamiques. Ces modèles peuvent être individu-centrés, reposant sur une description fine du comportement de chaque individu (comme dans le cas des couronnes de renouées de notre modèle de croissance locale de la Section 4.1) ou globaux, à l'échelle d'une population, dans le formalisme des inclusions différentielles comme dans le cas de notre modèle de dispersion de la Section 5.1.

Quelle que soit leur échelle, les modèles ont des paramètres d'entrée, ils permettent de décrire la dynamique du modèle, et incluent ainsi, dans le cadre de l'aide à la décision, les moyens d'actions du gestionnaire. Les modèles produisent des résultats appelés sorties à partir desquelles sont déduites les réponses utiles pour l'aide à la décision.

Un aspect important de la modélisation est le choix de valeurs pour les paramètres d'entrée du modèle. La valeur de certains paramètres peut être estimée par des connaissances issues de la littérature, ou des données issues d'expériences. Pour d'autres paramètres, ces valeurs ne sont pas disponibles. Les paramètres pour lesquels une valeur n'est pas imposée sont dits libres. Néanmoins, le modélisateur peut avoir à sa disposition des données concernant les sorties du modèle et souhaiter que le modèle reproduise ces données. L'objectif est alors de trouver des valeurs de paramètres qui permettent de faire correspondre au mieux les sorties du modèle aux données. Ce processus fait partie de la validation de modèle. Mais, s'il est important que le modèle reproduise des motifs attendus, il faut aussi vérifier dans le cadre de sa validation, que le modèle ne produise pas de valeurs non souhaitées. Pour cela, il est possible d'employer



des algorithmes pour la recherche de diversité.

Pour apporter des réponses dans le cadre de l'aide à la décision, les relations entre paramètres d'entrée et les sorties doivent pouvoir être décrites. Du fait de la complexité des modèles, des outils généraux (indépendants du modèle) sont développés pour permettre cette étude. Ces méthodes font partie du domaine de l'analyse de sensibilité d'un modèle.

Ce chapitre est organisé comme suit. Nous présentons d'abord deux formalismes mathématiques : celui des mesures (dont nous nous servons dans le modèle de la Section 4.1) et celui des inclusions différentielles (dont nous nous servons dans le modèle de la Section 5.1). Nous présentons ensuite des techniques employées pour la validation de modèles. Nous commençons par la calibration et détaillons deux techniques qui reposent sur la minimisation d'un (ou plusieurs) critère(s) et valables pour des modèles généraux. La première utilise un algorithme évolutionnaire, la seconde regroupe des méthodes bayésiennes approchées. Nous décrivons ensuite un algorithme de recherche de diversité qui permet de vérifier que le modèle ne produit pas de valeurs non souhaitées. Enfin, nous présentons des méthodes d'analyse de sensibilité d'un modèle qui permettent l'étude de la manière dont un modèle réagit à une modification de ses entrées.

## 2.1 Formalismes pour la modélisation

En fonction de la question posée, différentes échelles de modélisation peuvent être pertinentes : les modèles à l'échelle de l'individu pour rendre compte de la variabilité individuelle, les modèles globaux à l'échelle d'une population lorsque l'approximation champ moyen est valide. Ces deux types de modèles peuvent être décrits dans différents formalismes, nous en présentons ici un pour chacun d'entre eux.

### 2.1.1 Les mesures pour décrire un modèle individu-centré

Un modèle individu-centré décrit le comportement de chacun des éléments (appelés individus) constituant le système (appelé population). Chaque individu est caractérisé par un ensemble de valeurs  $x$  qui évoluent au cours du temps, en fonction notamment des autres individus. Si on suppose que l'espace dans lequel évoluent les valeurs caractérisant un individu est le même pour tous les individus, et que c'est un sous-ensemble de  $\mathbb{R}^d$  noté  $X$ , on peut voir un individu avec l'ensemble de valeurs  $x$  comme une masse de Dirac en  $x$  notée  $\delta_x$ . Ainsi, la population appartient au sous-ensemble  $\mathcal{M}$  des mesures sur  $X$  composé des mesures ponctuelles finies :

$$\mathcal{M} := \left\{ \sum_{i=1}^N \delta_{x_i}, n \in \mathbb{N} \text{ et } x_i \in X \text{ pour tout } i \in [1, n] \right\}.$$

Remarquons que la taille de population n'est pas nécessairement fixée, il est donc possible de créer ou d'enlever des individus du système. L'étude mathématique porte sur le processus  $(Y_t, t \geq 0)$ , à valeurs dans  $\mathcal{M}$  et qui décrit les individus présents dans la population au temps  $t$ . Il est défini par

$$Y_t = \sum_{i=1}^{N_t} \delta_{X_t^i},$$

où  $N_t = \int 1dY_t$  est la taille de la population au temps  $t$ , et  $X_t^1, \dots, X_t^{N_t}$  sont les valeurs caractérisant chacun des individus au temps  $t$ .

Lorsqu'il y a un grand nombre d'individus, des résultats sur l'approximation de ces systèmes existent et concernent les limites d'échelles de ces modèles et l'étude de leur comportement en temps long. On suppose que les individus ont un poids  $\frac{1}{K}$  et on considère le processus de population renormalisé noté  $(Z_t^K)_{(t \in \mathbb{R}_+)}$  défini par

$$Z_t^K := \frac{1}{K} Y_t^K.$$

Il est possible d'étudier l'approximation de ce processus de population renormalisé lorsque  $K$  tend vers l'infini.

En physique statistique, les individus sont des particules. Les limites obtenues pour de grands systèmes s'appellent des limites en champ moyen, et permettent notamment d'obtenir des modèles classiques tels que ceux de Vlasov ou Schrödinger (Golse, 2003).

Ce cadre se prête donc à l'étude de l'évolution darwinienne de population. L'ensemble de valeurs caractérisant un individu est un trait qui se transmet à la descendance, et qui est sujet à la sélection et la mutation (Méléard and Bansaye, 2015). Dans ce cas, les processus sont des processus stochastiques à valeurs mesures. Leur étude probabiliste rigoureuse a d'abord été

menée par Fournier and Méléard (2004). Par la suite, ces modèles ont été également étudiés et généralisés notamment par Champagnat (2006), Champagnat et al. (2006), Costa et al. (2016), Coron et al. (2018) et Tran (2006, 2008). Une interprétation biologique de la renormalisation opérée pour l'étude en grande population est que les individus ont moins de poids dans une grande population, ce qui peut être la conséquence d'une quantité de ressources disponibles limitée à répartir entre les individus. Ainsi, les individus sont soumis à la compétition pour les ressources et le taux de mortalité individuel peut dépendre de la distribution des traits des individus dans la population, ce qui entraîne des non-linéarités dans la limite. Pour des grandes populations, l'équation limite obtenue est une équation intégro-différentielle non linéaire. Différents changements d'échelle des paramètres peuvent être étudiés, ce qui conduit à des scénarios évolutifs très différents. Les modèles classiques de dynamique des populations dans un cadre écologique ou mutationnel (Malthusien, logistique, diffusion de Feller logistique, Lotka-Volterra) peuvent ainsi être obtenus à partir de modèles individu-centrés construits initialement.

### 2.1.2 Les inclusions différentielles pour décrire un modèle global

Un modèle dynamique à l'échelle d'une population décrit les règles d'évolution de l'état de cette population. Le formalisme des inclusions différentielles permet de prendre en compte la dépendance des variations vis-à-vis de l'état de la population comme avec une équation différentielle ordinaire mais aussi des contrôles et des perturbations par l'appartenance à une correspondance (une application dont les valeurs sont des ensembles). Une inclusion différentielle s'écrit donc

$$x' \in F(x),$$

où  $F$  est la correspondance.

Ce formalisme est utilisé par exemple dans le cadre de la théorie de la viabilité, où la dynamique est associée à une contrainte de viabilité (voir Chapitre 3) ou en théorie du contrôle optimal, où l'objectif est de sélectionner une trajectoire qui optimise un critère.

## 2.2 Validation de modèles

Dans cette section, nous traitons de la validation de modèles. Dans le cas de modèles visant à reproduire des motifs observés, il est important de vérifier que le modèle proposé est effectivement en mesure de les reproduire. La calibration permet de rendre compte de la faisabilité de cet objectif. Nous présentons deux techniques principales utilisées pour réaliser une calibration. Cependant, la validation de modèles ne se limite pas à de la corroboration (accord avec les motifs attendus), elle doit aussi chercher à mettre en défaut le modèle, en analysant la capacité du modèle à produire des sorties non souhaitées (qui contrediraient les connaissances du domaine ou des données). Une méthode de recherche de diversité dans les sorties d'un modèle directement applicable à cette fin est présentée.

### 2.2.1 Calibration

La modélisation individu-centrée est beaucoup utilisée car elle permet de décrire des dynamiques individuelles avec des interactions non linéaires. Cette technique de modélisation introduit cependant de nombreux paramètres libres qui influencent de manière critique le comportement du modèle. Le choix de valeurs pour ces paramètres issues de la littérature est

un processus qui peut être long (Reuillon et al., 2015). La calibration de modèles est une procédure qui vise à trouver les valeurs de paramètres pour lesquelles le comportement simulé par le modèle est proche du comportement à reproduire (données de terrain ou connaissances issues de la littérature, dire d'experts). Une calibration effectuée manuellement en testant des valeurs pour les paramètres et en vérifiant visuellement que la sortie du modèle correspond aux résultats attendus soulève des problèmes mis en évidence par Stonedahl (2011), comme le fait que les relations entre les paramètres sont souvent non linéaires, que certains paramètres ne peuvent être comparés avec des valeurs empiriques ou que l'exploration manuelle est fastidieuse. La calibration automatique des modèles peut alors libérer le modélisateur de cette tâche et rendre le processus de modélisation plus rigoureux en réduisant le nombre de choix arbitraires. La méthode de calibration automatique qui prévaut actuellement consiste à considérer la calibration comme un problème d'optimisation dans lequel les différences entre les données connues et les prédictions du modèle doivent être minimisées (Reuillon et al., 2015). En raison du nombre de paramètres qui peut être élevé et de la plage continue de variation de la valeur des paramètres, toute stratégie de recherche exhaustive n'est pas efficace, car elle générerait un nombre d'expériences qui croît de manière exponentielle avec le nombre de paramètres étudiés. Cette optimisation peut être réalisée de manière plus efficace à l'aide de différents algorithmes d'optimisation comme les algorithmes évolutifs (en particulier l'algorithme Non-dominated Sorting Genetic Algorithm, NSGA2, (Deb et al., 2002; Schmitt et al., 2015)) et de méthodes issues du calcul bayésien approché (Approximate Bayesian Computation, ABC, (Beaumont, 2010)). Ces deux méthodes sont présentées dans la suite de cette section.

### Méthode issue d'algorithmes génétiques

Les algorithmes génétiques (aussi appelés algorithmes évolutionnaires) apportent une solution appropriée à la façon d'envisager la calibration comme une optimisation. Les algorithmes évolutionnaires sont des heuristiques qui explorent l'espace de recherche en utilisant des stratégies inspirées des processus naturels (sélection, crossover et mutation génétique) pour résoudre les problèmes d'optimisation. Afin d'utiliser les algorithmes évolutifs pour calibrer un modèle, la première étape consiste à formaliser le résultat attendu d'une "bonne" simulation par le choix d'une ou plusieurs fonctions objectifs (fitness functions). Ces fonctions quantifient ainsi la différence entre la sortie du modèle et la caractéristique (parfois appelée fait stylisé) qu'on cherche à reproduire. Ce sont ces fonctions qui doivent être minimisées ou maximisées. Un problème d'optimisation courant est l'ajustement du modèle à des données. Dans ce cas particulier, la fonction objectif calcule la distance entre les résultats de la simulation et les données (la norme euclidienne est fréquemment utilisée). Schmitt et al. (2015) proposent une méthode de calibration automatique reposant sur l'algorithme évolutionnaire NSGA2 (Deb et al., 2002) pour l'exploration de l'espace des paramètres. Cette exploration est guidée par les objectifs choisis. La qualité de chaque jeu de paramètres testé est mesurée en fonction des valeurs des fonctions objectifs pour les résultats de la simulation qu'il produit. Les jeux de paramètres qui reçoivent les meilleures évaluations sont ensuite utilisés comme base pour générer de nouveaux jeux de paramètres, qui seront ensuite testés. Cette sélection se fait selon l'appartenance à un front de Pareto de la valeur des fonctions objectifs associées à ces valeurs de paramètres. Un ensemble de valeurs appartient au front de Pareto si cet ensemble est optimal au sens de Pareto, c'est-à-dire qu'il ne peut être modifié de manière à améliorer la valeur d'un objectif sans en détériorer au moins un autre. En raison de leur nature heuristique, ce sont des procédures qui peuvent se dérouler indéfiniment. Dès lors, on considère qu'un algorithme évolutif a convergé lorsqu'il n'apporte plus d'améliorations dans la recherche de bonnes solutions. Pour

un algorithme d'optimisation multi-objectifs, une mesure de convergence est donnée par la stabilité de l'hypervolume de la partie dominée de la solution dans l'espace des objectifs (Schmitt et al., 2015). Le front de Pareto ne se serait probablement pas beaucoup plus amélioré avec du temps de calcul supplémentaire.

Cependant, la calibration des modèles stochastiques conduit à des fonctions objectifs bruitées, ce qui peut altérer la convergence de l'algorithme évolutionnaire. Une stratégie efficace pour traiter de telles fonctions objectifs est mise en œuvre dans le logiciel d'exploration de modèles OpenMOLE (Reuillon et al., 2013) (nous utilisons cette implémentation de l'algorithme pour la calibration des deux modèles de croissance de la renouée que nous proposons aux Sections 4 et 5). Cette stratégie gère la nécessité de réplication et celle découverte de nouvelles solutions. Ainsi, elle réplique plusieurs fois chaque jeu de paramètres et considère la médiane ou la moyenne de ses valeurs. De plus, l'utilisation des algorithmes évolutionnaires pour la calibration de modèles individu-centrés (et en particulier ceux qui sont stochastiques) entraîne un coût de calcul élevé. Le nombre d'opérations à effectuer est trop important pour être exécuté sur un seul ordinateur, c'est pourquoi la méthode développée par Schmitt et al. (2015) fonctionne sur des grilles de calcul, qui donnent accès à la puissance de calcul requise. Les algorithmes évolutifs sont donc des solutions efficaces lorsqu'aucune méthode d'exploration exhaustive n'est disponible. Remarquons cependant qu'ils ne garantissent pas l'optimalité de la solution calculée.

### Calcul Bayésien Approché (ABC)

La section précédente a présenté une méthode d'optimisation reposant sur l'algorithme évolutionnaire NSGA2. Comme nous l'avons vu, cette méthode gère la stochasticité en répliquant plusieurs fois le jeu de paramètres et en prenant la médiane ou la moyenne de ses valeurs. Cependant, ces quantités ne sont pas toujours représentatives des sorties d'un modèle (par exemple pour des distributions bi-modales). Les techniques de calcul bayésien approché (ABC) sont pertinentes pour calibrer les modèles stochastiques car elles sont faciles à mettre en œuvre et applicables à tout modèle. Les méthodes ABC conservent l'information aléatoire et la retranscrivent dans la solution proposée puisqu'elles génèrent un échantillon de valeurs des  $N$  paramètres d'entrée du modèle  $(\theta_i)_{i=1,\dots,N}$  à partir de la distribution *a priori*  $\pi(\theta)$  considérée et sélectionnent les valeurs  $\theta_i$  conduisant aux sorties du modèle  $x \sim p(x|\theta_i)$  satisfaisant un critère de proximité avec les données à reproduire  $y$  ( $\rho(x, y) \leq \epsilon$ ,  $\rho$  exprimant une distance,  $\epsilon$  étant un niveau de tolérance). L'échantillon sélectionné de valeurs de paramètres se rapproche de la loi *a posteriori* des paramètres, ce qui permet d'obtenir des sorties de modèle avec la qualité d'approximation souhaitée.

L'inférence bayésienne nécessite le calcul de la fonction de vraisemblance  $p(x|\theta)$ , c'est-à-dire la probabilité d'obtenir les observations compte tenu d'une certaine valeur des paramètres. Cependant, du fait de la complexité de nombreux modèles, le calcul de cette fonction de vraisemblance peut être difficile. Il est souvent plus aisé d'écrire un programme informatique pour simuler la dynamique du modèle. Le calcul bayésien approché permet ainsi de ne pas avoir recours au calcul exact de la vraisemblance. Comme cela est souligné par Beaumont (2010), ces méthodes ont été développées originellement pour des modèles de génétique des populations mais dépassent désormais ce cadre.

La densité conjointe des valeurs des paramètres  $\theta$  et des données  $x$  est donnée par  $p(x, \theta) = p(x|\theta)\pi(\theta)$ , où  $\pi(\theta)$  est la loi *a priori* sur l'espace des paramètres.

Il s'agit ensuite de calculer la loi *a posteriori* (loi de probabilité conditionnelle)

$$p(\theta|x) = \frac{p(x|\theta)\pi(\theta)}{p(x)}$$

ce qui est généralement difficile car la probabilité marginale

$$p(x) = \int p(x|\theta)\pi(\theta)d\theta,$$

requiert le calcul d'une intégrale sur un espace souvent de grande dimension. Néanmoins, des méthodes de Monte Carlo peuvent être utilisées pour approcher cette intégrale.

Si les sorties du modèle sont discrètes et de faible dimension, il est alors possible d'échantillonner la loi *a posteriori* des paramètres sans approximation, par l'algorithme de Rubin (1984) : Étant donnée l'observation  $y$ , répéter ce qui suit jusqu'à ce que  $N$  jeux de paramètres  $\theta_i$  soient acceptés :

1. Tirer  $\theta_i \sim \pi(\theta)$ .
2. Simuler  $x_i \sim p(x|\theta_i)$ .
3. Rejeter  $\theta_i$  si  $x_i \neq y$ .

Ainsi, l'échantillon  $(\theta_i)_{i=1,\dots,N}$  obtenu est tiré selon la loi de probabilité  $p(\theta|y)$  de la valeur des paramètres qui a donné lieu aux observations. L'utilisation de méthodes de noyaux peut permettre d'estimer la densité de la loi à partir de l'échantillon (Parzen, 1962). Pour la visualisation, il peut être utile de s'intéresser aux lois marginales. Il est aussi possible, à partir de l'estimation de la densité, d'estimer le mode de cette distribution, qui correspond à l'estimation du maximum de vraisemblance lorsque la loi *a priori* est uniforme. Il est également possible d'estimer la moyenne, la médiane de la loi *a posteriori*, ainsi que diverses quantités qui renseignent sur cette distribution.

Lorsque les sorties du modèles appartiennent à un espace continu, l'étape 3 de l'algorithme de Rubin (1984) peut être modifiée comme suit :

3'. Rejeter  $\theta_i$  si  $\rho(x_i, y) > \epsilon$ ,

où  $\rho$  est une fonction mesurant la distance entre les sorties du modèle simulé  $x_i$  et les données observées  $y$ . Si de plus les sorties du modèles appartiennent à un espace de grande dimension, il peut être nécessaire de réduire cette dimension avec l'utilisation d'une fonction résumant les sorties, notée  $S$  (summary statistic).

L'étape 3 de l'algorithme de Rubin (1984) peut être modifiée comme suit :

3''. Rejeter  $\theta_i$  si  $\rho(S(x_i), S(y)) > \epsilon$

En général, des informations sont perdues avec l'utilisation de la fonction  $S$ , mais l'objectif est que les fonctions  $S$  résumant les sorties satisfassent

$$p(\theta|x) = p(\theta|S(x)),$$

pour la loi *a priori*  $\pi(\theta)$ .

En pratique, l'utilisation de ces techniques est, comme pour les algorithmes évolutionnaires, très coûteuse en calcul car l'échantillonnage de l'ensemble de l'espace des paramètres nécessite un certain nombre de simulations qui croît exponentiellement avec le nombre de paramètres concernés par la méthode. C'est pourquoi certains auteurs, comme Lenormand et al. (2013), proposent des méthodes bayésiennes visant à minimiser le nombre d'exécutions du modèle nécessaire pour atteindre la qualité donnée d'approximation désirée. Lenormand et al. (2013) (voir les références incluses) énoncent les différentes améliorations du schéma original disponibles : utilisation de régressions locales pour améliorer l'inférence des paramètres, des techniques sélectionnant automatiquement des statistiques informatives, le couplage à des

Chaînes de Markov Monte Carlo (ou l'amélioration séquentielle des distributions *a posteriori* avec des distributions séquentielles de Monte Carlo. Nous utilisons la méthode de Lenormand et al. (2013), implémentée dans le logiciel OpenMOLE, pour la calibration du modèle étudié à l'Annexe D.

## Profil de calibration

Les techniques de calibration présentées dans les sections précédentes, malgré leur coût de calcul élevé, ne permettent pas d'obtenir une image globale de la relation entre l'espace des paramètres d'entrée et l'espace de sorties du modèle. Dans les deux cas considérés, le résultat de la calibration indique uniquement si le modèle peut reproduire les données avec une précision donnée, et si oui, il donne un jeu de paramètres, ou une loi sur l'ensemble des paramètres, qui permet d'y arriver. Comme le soulignent Reuillon et al. (2015), ces méthodes de calibration ne disent donc rien sur la fréquence à laquelle les ensembles de paramètres conduisent aux comportements qu'on vise à reproduire, ni sur la manière dont chaque paramètre modifie le comportement du modèle. Par exemple, il est souvent intéressant de savoir quand certaines valeurs de paramètres empêchent le système d'atteindre un comportement réaliste, plutôt que de connaître seulement un unique ensemble de valeurs de paramètres "optimales".

C'est à partir de ce constat que Reuillon et al. (2015) ont conçu la méthode de profil de calibration (Calibration Profile, CP) qui étend le concept de processus de calibration automatisé et rend compte de la sensibilité de chaque paramètre lors de la calibration d'un modèle indépendamment des autres paramètres. Elle calcule l'erreur de calibration la plus faible qui puisse être obtenue lorsque la valeur d'un paramètre est fixée et que les autres sont libres. La méthode calcule cette erreur minimale pour des valeurs du paramètre échantillonnées dans le domaine de définition considéré. Le profil peut ensuite être tracé sur un graphique en deux dimensions qui décrit l'influence du paramètre étudié sur la calibration du modèle.

### 2.2.2 Recherche de diversité : PSE

Une approche classique de validation d'un modèle consiste à le calibrer et comparer les résultats de simulations obtenus aux données à reproduire. Ces étapes amènent les modélisateurs à conclure si le modèle proposé peut expliquer les données ou non, et à le corriger si nécessaire. Or, comme énoncé dans Chérel et al. (2015) cette information ne suffit pas pour affirmer qu'il reflète avec précision la façon dont les données observées sont produites dans la nature ou la société. Un motif global peut en effet être expliqué par différents modèles, régis par des dynamiques locales différentes. Une approche de la validation basée uniquement sur la corroboration est donc insuffisante pour construire des modèles explicatifs et prédictifs solides. Chérel et al. (2015) appliquent alors le principe de falsifiabilité de Karl Popper à la modélisation. Ainsi, les simulations du modèle peuvent être divisées en deux ensembles : celles dont le résultat correspond à l'observation empirique et celles qui la contredisent. Dans le premier cas les simulations corroborent le modèle, dans le second elles le falsifient. S'il est bien sûr important de s'assurer qu'un modèle est capable de reproduire les données qu'il est censé expliquer le principe de falsifiabilité énonce qu'il est tout aussi important de rechercher des simulations qui peuvent falsifier le modèle, c'est-à-dire d'identifier les valeurs de paramètres pour lesquelles les sorties du modèle contredisent les données. En effet, les modèles ayant une complexité importante peuvent produire des résultats inattendus. Un tel résultat peut être soit un falsificateur du modèle s'il est contredit par une observation empirique, soit une prédiction du modèle s'il est validé par celui-ci.

Le problème de la recherche de motifs inattendus est distinct du problème de calibration pour lequel le motif à reproduire est connu. Dans Chérel et al. (2015), les auteurs proposent de l'aborder en recherchant plus généralement les divers motifs qu'un modèle peut produire et proposent pour cela une méthode évolutive PSE (Pattern Space Exploration) basée sur la "Novelty Search". Nous utilisons cet algorithme dans l'Annexe C, dans le but de la recherche de diversité de pattern produit en soi, et non pour la validation d'un modèle. Le motif obtenu dans une simulation dépend des entrées du modèle. Ainsi, bien que l'objectif de la recherche de la diversité concerne l'espace des sorties du modèle, il s'agit d'un problème d'exploration de l'espace des paramètres d'entrée. Une telle exploration rencontre les difficultés classiques que posent les modèles de simulation de systèmes complexes, déjà évoquées dans les méthodes précédentes : la haute dimensionnalité de l'espace des paramètres, la non-linéarité entre les entrées et les sorties, et la stochasticité. Pour le problème de l'échantillonnage dans l'espace des paramètres qui peut être de grande dimension, les auteurs appliquent une méthode adaptative, en choisissant itérativement le prochain paramètre échantillonné sur la base des informations disponibles dans la population à l'instant courant. La méthode PSE est un algorithme évolutif. Il est donc composé des blocs classiques des algorithmes évolutifs : une population d'individus dont les génomes codent chacun un élément particulier de l'espace des paramètres du modèle, une méthode de sélection pour choisir les individus à partir desquels une nouvelle population sera créée, un mécanisme de croisement et de mutation pour créer un nouvel individu à partir de ceux qui ont été sélectionnés précédemment, et un mécanisme d'élitisme pour filtrer tous les individus (nouvelle population créée et ancienne) à conserver et former la nouvelle population. Au fur et à mesure que l'algorithme progresse, les individus non désirés sont rejetés et oubliés à mesure que de nouveaux s'ajoutent à la population. L'algorithme commence par générer des individus avec des génomes aléatoires représentant des valeurs de paramètres. Pour chacun d'eux, une simulation est effectuée et son motif est obtenu par évaluation du modèle. Au stade de l'élitisme, les individus sont filtrés pour ne retenir dans la population que ceux qui sont significativement différents les uns des autres (motifs différents). Des paires d'individus sont sélectionnées en fonction de la rareté de leur motif par rapport à la population actuelle et aux motifs passés (la rareté est la mesure de la nouveauté). Le génome de chaque paire sélectionnée est recombinié pour former un nouveau génome, qui subit ensuite une nouvelle mutation. Une nouvelle génération d'individus est ainsi créée, et la boucle continue. L'efficacité de cette heuristique est fondée sur l'hypothèse qu'en sélectionnant des parents dont les motifs sont rarement observés aux étapes précédentes, les chances de trouver des motifs encore non découverts augmentent.

## 2.3 Analyse de sensibilité d'un modèle

L'analyse de sensibilité correspond à un ensemble de méthodes permettant de comprendre la manière dont un modèle réagit à une modification de ses entrées. L'objectif de ces méthodes statistiques est de mesurer comment la variation se propage des entrées aux sorties. Plus précisément, l'analyse de sensibilité est définie par Saltelli et al. (2008) comme la description de l'importance relative de chaque entrée dans la détermination de la variabilité des sorties. Dans le cas de modèles stochastiques (ce qui est souvent le cas des modèles individu-centrés), il est nécessaire de procéder à des répliques afin d'obtenir une robustesse statistique nécessaire pour tester les hypothèses et distinguer la stochasticité du modèle de l'effet de la variation de paramètres (Lee et al., 2015). Or, puisque la forme des distributions de sortie d'un modèle stochastique, à paramètres d'entrée fixée, est généralement inconnue *a priori* (et non nécessairement gaussienne), la taille de l'échantillon à considérer pour que la moyenne et la variance



des sorties soient stables doit être considérée.

L'analyse de sensibilité implique généralement un échantillonnage *a priori* de l'espace des paramètres d'entrée. La conception de plan d'expérience (Design of Experience, DOE) dans un contexte de simulation de modèle est l'art de mettre en place de telles expérimentations. Cela consiste donc à choisir les paramètres d'entrée étudiés et les valeurs qu'ils prendront, pour ensuite simuler le modèle sur ces valeurs dans le but de révéler des propriétés du modèle. L'échantillonnage le plus simple est défini sur une grille qui dépend d'un pas de discrétisation. Le modèle est ensuite évalué pour toutes les combinaisons possibles des valeurs d'entrée fournies (pour un nombre raisonnable de dimensions). L'échantillonnage par hypercube latin (LHS) est une méthode statistique pour l'échantillonnage quasi-aléatoire. En dimension  $d$ , un échantillon est un carré latin si chaque point de l'échantillon est le seul point dans chaque hyperplan de dimension  $d - 1$  aligné sur les coordonnées qui définissent sa position. Ainsi, un échantillon d'une grille carrée (2D) est un carré latin si (et seulement si) il n'y a qu'un seul point de l'échantillon dans chaque rangée et chaque colonne. Enfin, l'échantillonnage selon une suite de Sobol crée une suite quasi aléatoire à faible discrédance. La discrédance d'une suite est faible si la proportion des points de la suite appartenant à un ensemble  $B$  est proche de la valeur de la mesure de  $B$  (pour tout ensemble  $B$ ).

L'analyse de sensibilité peut être effectuée à un niveau global ou local. Les méthodes globales fournissent des informations sur les effets de la variation des entrées dans l'espace d'entrée complet, tandis que les méthodes locales concentrent l'effet de la variation des entrées autour d'un point donné de l'espace d'entrée. L'analyse de sensibilité pour les modèles individu-centrés fait appel à une ou plusieurs des méthodes suivantes, sélectionnées parmi celles exposées dans Lee et al. (2015) : one-parameter-at-a-time (OAT), elementary effects (EE), méta-modèles, décomposition de la variance (VBSA). Nous examinons ci-dessous chacune de ces méthodes. Dans l'OAT (one-parameter-at-a-time), chaque paramètre d'entrée est examiné indépendamment des autres, sur un ensemble de valeurs déterminées, en maintenant les autres paramètres à une valeur nominale fixée. La méthode OAT est donc une méthode de sensibilité locale. Les effets des modifications des paramètres sur les sorties du modèle sont ensuite étudiées. Il est possible de répéter cette opération pour différentes valeurs des autres paramètres, afin d'augmenter la robustesse de l'analyse. La simplicité de l'OAT est contrebalancée par des limites dans le cadre de l'analyse de sensibilité des modèles individu-centrés (Ligmann-Zielinska, 2013). D'une part, la méthode nécessite de fixer des valeurs nominales de chaque paramètre d'entrée important pour la dynamique, mais ces dernières peuvent être *a priori* inconnues. De plus, ce type d'échantillonnage ne tient pas compte des interactions potentielles entre les paramètres et n'explore qu'une très petite fraction de l'espace des paramètres. Enfin, la nature de cette analyse réduit considérablement l'espace de variation des paramètres d'entrée au sein de l'hypercube. Par exemple, avec seulement 10 paramètres d'entrée, l'OAT ne couvre que 0,25 % de l'espace d'entrée (Saltelli and Annoni, 2010).

En guise d'étude préliminaire du modèle du Chapitre 4, nous avons effectué plusieurs OAT pour chaque paramètre (c'est-à-dire pour différentes valeurs nominales des autres paramètres). Les effets élémentaires (Elementary effects, EE) étendent l'OAT, qui ne considérerait la variation d'un paramètre que sur une droite. Pour la méthode EE, une modification d'un paramètre d'entrée est maintenue lors de l'étude d'une modification de la valeur d'un autre paramètre, plutôt que de reprendre la valeur nominale (comme dans l'OAT). Proposée à l'origine par Morris (1991) et améliorée par Campolongo et al. (2000), l'EE convient aux modèles coûteux en termes de temps de calcul et ayant de larges ensembles pour les paramètres d'entrées. La méthode permet de sélectionner les paramètres non influents. La méthode commence par l'échantillonnage d'une valeur (initiale) pour chacun des paramètres (variable d'entrée), dans

des plages définies de valeurs possibles et calcule le résultat de l'évaluation du modèle sur ce jeu de paramètres. La deuxième étape consiste à modifier la valeur d'un paramètre (toutes les autres valeurs de paramètres restant à leurs valeurs de départ) en prenant la valeur d'un de ses voisins les plus proches sur une grille et à calculer la nouvelle valeur du modèle pour ce jeu de paramètres. Ensuite, la valeur d'un autre paramètre est modifiée (le paramètre dont la valeur a été modifiée à l'étape précédente conserve sa valeur modifiée et tous les autres paramètres conservent leur valeur initiale) et la valeur du modèle correspondant est à nouveau calculée. Cette opération se poursuit jusqu'à ce que toutes les variables d'entrée soient modifiées. Étant donnée une valeur initiale, le modèle est donc évalué une fois de plus que le nombre de paramètres. Cette procédure est répétée plusieurs fois avec des valeurs initiales différentes pour chacun des paramètres. Un effet élémentaire pour l'entrée  $i$  et la sortie  $y_j$  est obtenu lorsque la valeur  $x_0$  du paramètre  $i$  est modifiée par un pas  $\delta_i$  à partir d'un point  $x_0$  (vaut donc  $x_0 + \delta_i$ , et calculé comme  $\epsilon_{i,j} = (y_j(x_0) - y_j(x_0 + \delta_i))/\delta_i$ . Enfin, la méthode EE résume ces effets élémentaires en calculant trois indicateurs de sensibilité pour chaque entrée et chaque sortie du modèle, afin d'estimer la sensibilité globale dans l'espace de sortie. La première mesure de sensibilité globale  $\mu_{i,j}$  est la moyenne des effets élémentaires  $\epsilon_{i,j}$ , calculée sur les effets pour lesquels le facteur  $i$  a été modifiée. Une version plus robuste de la mesure de sensibilité finale,  $\mu_{i,j}^*$  est calculée comme la moyenne de la valeur absolue des effets élémentaires, ce qui garantit la robustesse par rapport aux modèles non monotones, mais il s'agit là encore d'une moyenne qui ne permet pas de rendre compte des effets non linéaires. Pour tenir compte des non-linéarités ou des interactions entre les facteurs, la mesure  $\sigma_{i,j}$  correspond à l'écart type des effets élémentaires. Un faible écart-type signifie que les effets sont constants, c'est-à-dire que la sortie est linéaire sur ce facteur. Une valeur élevée signifie soit que l'indicateur est non linéaire sur ce facteur (de faibles variations à certains endroits et de fortes variations à d'autres), soit que les variations changent beaucoup lorsqu'on change la valeur d'autres paramètres, c'est-à-dire que ce paramètre a des interactions fortes avec les autres paramètres.

L'analyse de sensibilité basée sur la variance (VBSA) est considérée comme étant l'approche la plus prudente pour évaluer la sensibilité des modèles car elle ne suppose pas de linéarité (Lee et al., 2015). Dans le VBSA, la variance de la sortie du modèle ou du système est décomposée en fractions qui peuvent être attribuées aux différents paramètres et à leurs interactions. Deux indices par entrée ( $i$ ) sont tirés des variances partielles : un indice du premier ordre (effets principaux)  $S_i$  et un indice des effets totaux  $ST_i$ .  $S_i$  est le rapport entre la variance partielle du paramètre  $i$  et la variance totale.  $ST_i$  est la somme des effets de tous les indices différents de  $i$  ( $\sum S_{-i}$ ) et rend compte des interactions entre  $i$  et les autres entrées. La définition des indices et la méthode sont présentées dans Saltelli et al. (2010). Le calcul des variances intervenant dans ces indices requiert le calcul d'intégrales. Ces intégrales sont estimées par des méthodes de Monte Carlo, s'appuyant sur un échantillon de points générés par une méthode de Sobol.

Les méta-modèles (aussi appelés émulateurs) sont le produit d'un ensemble de méthodes qui tentent d'expliquer les variations des sorties du modèle en réponse aux changements des entrées du modèle en ajustant les sorties et les entrées de ce modèle dans une fonction mathématique (plus simple) (Ligmann-Zielinska et al., 2020). On suppose donc que la relation entre les sorties et les entrées peut être décrite en termes mathématiques relativement faciles à interpréter. La forme la plus élémentaire d'un émulateur est une régression linéaire, mais celui-ci peut être plus complexe. Les coefficients de régression deviennent alors des mesures de sensibilité des entrées du modèle. Ils fournissent à la fois l'intensité et la direction de l'influence d'une entrée sur la sortie du modèle. L'utilisation de l'analyse de régression rend le résultat facile à comprendre, mais des émulateurs plus complexes (c'est-à-dire des méta-modèles qui utilisent des équations complexes, des méthodes d'intelligence artificielle - voir,

par exemple, (Lamperti et al., 2018)) permettent une évaluation détaillée des interactions des variables d'entrée, qui sont plus appropriées pour les modèles fortement non linéaires (Lamperti et al., 2018). Toutefois, dans ce cas de formulation plus complexe, les méta-modèles peuvent également être difficiles à interpréter, c'est le cas pour des modèles représentant des interactions à plusieurs échelles, courants dans les systèmes complexes. Les méta-modèles peuvent également ne pas saisir toutes les interactions telles que les causalités simultanées et les oscillations, ce qui conduit à des informations incomplètes concernant le comportement du système (Ligmann-Zielinska et al., 2020).

Par exemple, Happe et al. (2006) étudient le modèle individu-centré de politique agricoles AgriPoliS avec un méta-modèle. Au lieu de faire d'abord une analyse de sensibilité pour analyser les propriétés du modèle et d'examiner ensuite la politique introduite de manière isolée, les auteurs utilisent un méta-modèle construit à partir d'un plan d'expérience permettant d'analyser l'influence des paramètres qu'ils souhaitent étudier. Les auteurs ont défini deux valeurs pour chaque paramètre concerné, autour d'une valeur par défaut (qui représente une configuration standard du modèle). Le choix de ces valeurs est basé sur des dires d'experts et des données statistiques. Les paramètres non inclus dans le DOE sont supposés rester à leur valeur par défaut. À partir de ces deux valeurs pour chaque paramètre, ils ont donc construit le plan d'expérience complet comprenant toutes les  $2^5$  combinaisons de paramètres possibles. Ce plan d'expérience sert à ajuster le méta-modèle, bien que les auteurs reconnaissent qu'il ne fournit qu'une estimation très grossière et linéaire de la sortie du modèle. Les auteurs choisissent une forme polynomiale du second ordre pour leur méta-modèle, dont les coefficients sont à ajuster. La régression effectuée dans ce but cherche à minimiser les moindres carrés. Les coefficients du méta-modèle permettent d'identifier l'effet principal de chaque paramètre ainsi que la nature des interactions entre deux paramètres. La sortie du modèle considérée par les auteurs est moyennée. La procédure leur permet d'identifier les paramètres correspondant à des interactions négligeables. L'ajustement du méta-modèle aux données simulées par le modèle issues du plan d'expérience est évalué avec le coefficient de détermination  $R^2$ . Les auteurs remplacent ensuite l'étude du modèle par celle du méta-modèle.

Nous utilisons un méta-modèle sur le modèle du Chapitre 4, afin d'étudier l'influence de certains paramètres sur sa dynamique. Notre démarche est analogue à celle développée par Happe et al. (2006).

## Chapitre

# 3

### **La théorie de la viabilité - son intérêt pour l'aide à la gestion de systèmes contrôlés soumis à des incertitudes**

Ce chapitre présente des outils méthodologiques utiles au problème de gestion de la renouée le long d'un cours d'eau. Dans le Chapitre 5, nous modélisons la dynamique d'invasion des renouées le long d'un cours d'eau. Cette dynamique dépend des événements de crues que connaît le cours d'eau. Cependant, ces derniers ne sont pas prévisibles et nous considérons ainsi un ensemble de crues possibles contre les conséquences desquelles des gestionnaires peuvent vouloir se prémunir. De plus, on considère que ces gestionnaires peuvent agir sur la croissance des taches, notamment par la fauche, et plus généralement sur la dynamique de l'invasion. L'objectif de gestion est de maintenir l'invasion en dessous de certains seuils. Nous sommes donc en présence d'un système contrôlé soumis à des incertitudes et pour lequel on souhaite conserver des contraintes. Le modèle proposé prend en compte ces différentes caractéristiques. La théorie de la viabilité, introduite par Aubin (1991), s'intéresse à des systèmes dynamiques contrôlés et soumis à un ensemble de contraintes de viabilité. Pour chaque état du système, on dispose d'un ensemble de contrôles sur la dynamique du système. Une étude viabiliste cherche à déterminer les suites de contrôles qui permettent au système de satisfaire à chaque temps les contraintes de viabilité. La théorie de la viabilité intègre également des perturbations du système, cela permet de prendre en compte des incertitudes quant à son évolution. C'est pourquoi nous avons choisi de réaliser l'étude du modèle de gestion de la renouée le long d'un cours d'eau dans le cadre de la théorie de la viabilité. Dans ce formalisme, il y a des concepts essentiels : noyau de viabilité, bassin de capture, noyau de viabilité garanti, qui vont être très utiles pour formaliser la réponse à notre problème de gestion. Cependant, la résolution analytique d'un problème de viabilité est rarement possible, des méthodes d'approximation sont alors employées. Ces méthodes reposent sur des algorithmes qui permettent des calculs effectifs en petite dimension. Or, la dimension de notre problème dépend d'un découpage du cours d'eau en zones sur lesquelles la dynamique des taches est analogue et pour lesquelles la même gestion est pratiquée. La dimension de notre problème de viabilité est donc grande et

ne remplit pas le critère requis pour pouvoir appliquer un algorithme d'approximation. Nous proposons donc une méthode permettant de simplifier l'étude de ce problème de viabilité en grande dimension en le remplaçant par plusieurs problèmes de viabilité en dimension plus petite, et pour lesquels les algorithmes pour le calcul de noyau de viabilité sont applicables. Un article portant sur une technique similaire pour la résolution d'un problème de viabilité est en cours de rédaction (Martin et al., en cours d'écriture).

Ce chapitre est organisé comme suit. Nous énonçons d'abord des concepts essentiels de la théorie de viabilité que nous utilisons dans ce manuscrit. Nous présentons ensuite des méthodes pour le calcul d'ensembles de viabilité. Nous nous intéressons d'abord aux algorithmes permettant une résolution numérique. Nous en présentons le principe général, quelques exemples d'utilisation issus de la littérature ainsi que leurs limites d'utilisation. Toutefois, les algorithmes permettant la résolution approchée de ces problèmes sont effectivement fonctionnels pour des petites dimensions. C'est pourquoi nous présentons ensuite les solutions proposées dans la littérature pour résoudre un problème de viabilité de grande dimension. Cependant, pour la plupart d'entre elles, ces méthodes reposent sur le choix de fonctions particulières qui rendent ces calculs analytiques possibles. Notre problème de viabilité issu de la gestion de la renouée le long d'un cours d'eau ne se prête pas à de tels calculs et nous avons souligné qu'il ne respecte pas la contrainte de dimension nécessaire à une résolution numérique. Nous présentons donc un cadre général de problème de viabilité vérifié par notre problème particulier de gestion. Nous proposons une méthode, appelée "cascade de noyaux de viabilité" pour la résolution de ce problème de viabilité. La méthode consiste à remplacer le calcul d'un noyau de viabilité en grande dimension par le calcul de plusieurs noyaux de viabilité en dimension plus petite, et pour lesquels les méthodes numériques sont applicables. Nous énonçons enfin un résultat qui montre que le calcul de ces quantités permet bien de répondre à la question initiale.

## 3.1 Concepts et outils fondamentaux de la théorie de la viabilité

Le modèle que nous proposons au Chapitre 5 s'écrit sous la forme d'un système dynamique contrôlé. Nous imposons des contraintes sur la dynamique et autorisons des contrôles. La théorie de la viabilité, introduite par Aubin (1991), traite de telles dynamiques, et nous avons choisi ce formalisme pour l'étude du modèle. Les problèmes de viabilité sont présents dans la gestion des ressources renouvelables (Oubraham and Zaccour, 2018) et notamment dans la réglementation de la pêche (Béné et al. (2001), Chapel et al. (2008), De Lara et al. (2007), Martinet et al. (2007), Gajardo et al. (2018)), la population de pigeons urbains (Haidar et al., 2017), la préservation des forêts (Domenech et al. (2011), Bernard and Martin (2013)), la gestion de l'eutrophisation des lacs (Martin, 2004). Ils se posent également dans la dynamique des (éco)-systèmes au sens large (Cury et al. (2005), Doyen et al. (2007), Wei et al. (2013)), dans les problèmes purement économiques ou sociaux (Aubin et al. (2001), Bernard and Martin (2012)), ou dans divers problèmes d'ingénierie (Kalisiak and van de Panne (2004), Sicard et al. (2012), Tomlin et al. (2003)). Nous présentons dans cette partie les outils de cette théorie que nous utilisons pour la résolution du problème de viabilité lié à la gestion des renouées le long d'un cours d'eau.

Le modèle proposé au Chapitre 5 est écrit en temps discret. Les résultats de la théorie portent sur des systèmes en temps continu et en temps discret, nous présentons dans cette section les résultats dans les deux cadres. Nous traitons d'abord des systèmes dynamiques contrôlés, puis des systèmes dynamiques tychastiques et enfin des systèmes dynamiques contrôlés et tychastiques (ceux qui nous intéresseront pour notre application).

### 3.1.1 Système contrôlé

#### Évolution et système contrôlé

Soit  $X \subset \mathbb{R}^n$  l'espace des états du système.

##### Cas continu :

Une évolution  $t \in [0, +\infty[ \mapsto x(t) \in X$  décrit l'état du système pour tout  $t \in [0, +\infty[$ . Elle peut être influencée par des actions extérieures appelées contrôles. La dynamique dépend donc de l'état du système  $x(t)$  et des différentes actions de contrôle  $u(t)$  qui influencent l'évolution du système au temps  $t$ . Le système contrôlé continu est donné par :

$$S_c \begin{cases} x'(t) = f(x(t), u(t)) \text{ pour } t \geq 0 \\ u(t) \in U(x(t)) \end{cases} . \quad (3.1)$$

Les contrôles  $u(t)$  appartiennent à un ensemble  $U(x(t)) \subset \mathbb{R}^p$  qui dépend de l'état du système au temps  $t$ . La correspondance  $U : X \rightsquigarrow Y$  décrit l'ensemble des contrôles  $U(x) \subset Y$  admissibles lorsque le système est dans l'état  $x$ . A partir d'un état initial  $x_0$ , il peut exister plusieurs évolutions possibles, en fonction des différents choix de contrôles au cours du temps  $t \mapsto u(t)$ . Une solution du système (3.1) est une évolution  $t \mapsto x(t)$  telle qu'il existe une fonction de contrôle  $t \mapsto u(t)$  telle que les conditions de (3.1) soient vérifiées pour tout  $t \geq 0$ . On note  $S_c(x)$  l'ensemble des évolutions issues de  $x$  gouvernées par le système contrôlé continu (3.1) pour presque tout  $t$ , i.e. :

$$S_c(x) = \left\{ x(\cdot) \mid x(0) = x \text{ et } \begin{cases} x'(t) = f(x(t), u(t)) \\ u(t) \in U(x(t)) \end{cases} \text{ pour presque tout } t \geq 0 \right\}. \quad (3.2)$$

**Cas discret :** Une évolution  $(x^n)_{n \in \mathbb{N}}$  décrit l'état du système pour tout  $n \in \mathbb{N}$ . Le système contrôlé discret est donné par :

$$S_d \left\{ \begin{array}{l} x^{n+1} = f(x^n, u^n) \text{ pour } n \in \mathbb{N} \\ u^n \in U(x^n) \end{array} \right. . \quad (3.3)$$

Les contrôles  $u^n$  appartiennent à un ensemble  $U(x^n) \subset \mathbb{R}^p$  qui dépend de l'état du système à l'instant  $n$ . A partir d'un état initial  $x_0$ , il peut exister plusieurs évolutions possibles, en fonction des différents choix de contrôles aux différents instants  $(u^n)_{n \in \mathbb{N}}$ . Une solution du système (3.3) est une suite  $(x^n)_{n \in \mathbb{N}}$  telle qu'il existe une suite de contrôles  $(u^n)_{n \in \mathbb{N}}$  telle que les conditions de (3.3) soient vérifiées pour tout  $n \in \mathbb{N}$ . On note  $S_d(x)$  l'ensemble des évolutions issues de  $x$  gouvernées par le système contrôlé discret (3.3) pour tout  $n \in \mathbb{N}$ , i.e. :

$$S_d(x) = \left\{ (x^n)_{n \in \mathbb{N}} \mid x^0 = x \text{ et } \begin{cases} x^{n+1} = f(x^n, u^n) \\ u^n \in U(x^n) \end{cases} \text{ pour tout } n \in \mathbb{N} \right\}. \quad (3.4)$$

### Évolution et ensemble viable

La théorie de la viabilité s'intéresse à des systèmes contrôlés soumis à des contraintes de viabilité. Soit  $K \subset X$  un ensemble de contraintes de viabilité.

Une évolution est viable dans  $K$  si sa trajectoire reste toujours dans l'ensemble de contraintes  $K$  :

$$\forall t \geq 0, x(t) \in K \quad (3.5)$$

dans le cas continu, et

$$\forall n \in \mathbb{N}, x^n \in K \quad (3.6)$$

dans le cas discret.

**Définition 3.1.1.** *Un ensemble  $L$  est viable pour le système contrôlé continu (3.1) (resp. discret (3.3)), si*

$$\forall x_0 \in L, \exists x(\cdot) \in S_c(x_0) \text{ (resp. } S_d(x_0)) \mid x(\cdot) \text{ est viable dans } L.$$

### Noyau de viabilité

Le noyau de viabilité associé au système dynamique contrôlé continu (3.1) (resp. discret (3.3)) et soumis à l'ensemble de contraintes de viabilité  $K$ , noté  $Viab_{c,f,U}(K)$  (resp.  $Viab_{d,f,U}(K)$ ), est l'ensemble de tous les états initiaux  $x_0 \in K$  pour lesquels il existe une solution  $x(\cdot)$  au système (3.1) (resp.  $(x^n)_{n \in \mathbb{N}}$  au système (3.3)) partant de  $x_0$  et viable dans  $K$ .

Dans le cas continu, on a donc :

$$Viab_{c,f,U}(K) = \{x_0 \in K \mid \exists x(\cdot) \in S_c(x_0) \text{ telle que } \forall t \geq 0, x(t) \in K\}. \quad (3.7)$$

Dans le cas discret, on a donc :

$$Viab_{d,f,U}(K) = \{x_0 \in K \mid \exists (x^n)_{n \in \mathbb{N}} \in S_d(x_0) \text{ telle que } \forall n \in \mathbb{N}, x^n \in K\}. \quad (3.8)$$

**Cas continu :**

Aubin (1991) a trouvé un ensemble de conditions suffisantes sur le système dynamique contrôlé pour que le noyau de viabilité associé soit fermé. Pour énoncer ce résultat, nous rappelons quelques définitions sur les correspondances et les systèmes dynamiques.

Soit  $F$  une correspondance  $F : A \rightsquigarrow B$ , le domaine de  $F$  est :

$$Dom(F) := \{x \in A \mid F(x) \neq \emptyset\}.$$

Le graphe  $Graph(F)$  de  $F$  est l'ensemble  $\{(x, y) \in A \times B, y \in F(x)\}$ .

Soit  $F : X \rightsquigarrow X$  une correspondance.  $F$  a une croissance linéaire s'il existe une constante positive  $c$  telle que  $\forall x \in Dom(F), \|F(x)\| \leq c(\|x\| + 1)$ .

**Définition 3.1.2.** Soient  $U : X \rightsquigarrow Y$  et  $f : X \times Y \rightarrow X$ .  $(U, f)$  est un système contrôlé Marchaud si

- $Graph(U)$  est fermé
- $f$  est continue
- $f$  et  $U$  ont des croissances linéaires
- les ensembles  $\{f(x, u) \mid u \in U(x)\}$  sont convexes pour tout  $x \in Dom(U)$ .

**Théorème 3.1.3.** (Aubin, 1991) Si  $K$  est fermé et  $(f, U)$  est un système dynamique contrôlé Marchaud, alors  $Viab_{c,f,U}(K)$  est fermé.

**Cas discret :**

**Théorème 3.1.4.** Si  $K$  est fermé et  $(f, U)$  est un système dynamique contrôlé tel que  $Graph(U)$  est fermé,  $f$  est continue et  $U$  a une croissance linéaire, alors  $Viab_{d,f,U}(K)$  est fermé.

*Démonstration.* Montrons que  $\overline{Viab_{d,f,U}(K)} = Viab_{d,f,U}(K)$ . Soit  $x \in \overline{Viab_{d,f,U}(K)}$ , il suffit de montrer qu'il existe  $u \in U(x)$  tel que  $f(x, u) \in Viab_{d,f,U}(K)$ , ainsi puisque  $Viab_{d,f,U}(K)$  est le plus grand ensemble vérifiant cette propriété et que  $Viab_{d,f,U}(K) \subset \overline{Viab_{d,f,U}(K)}$ , on aura nécessairement  $Viab_{d,f,U}(K) = \overline{Viab_{d,f,U}(K)}$ . Soit  $(x_n)_{n \in \mathbb{N}} \in Viab_{d,f,U}(K)^{\mathbb{N}}$  une suite d'éléments du noyau de viabilité qui converge vers  $x$ . Puisque pour tout  $n \in \mathbb{N}$ ,  $x_n \in Viab_{d,f,U}(K)$ , il existe  $u_n \in U(x_n)$  tel que  $f(x_n, u_n) \in Viab_{d,f,U}(K)$ . Or  $U$  a une croissance linéaire,  $Graph(U)$  est fermé et  $(x_n)_n$  converge vers  $x$ , il existe donc une sous-suite de  $(u_n)_n$  (encore notée  $(u_n)_n$ ) qui converge vers  $u \in U(x)$ . De plus  $f$  est continue, et comme  $(x_n)_n$  et  $(u_n)_n$  convergent respectivement vers  $x$  et  $u \in U(x)$ , on a  $f(x, u) \in \overline{Viab_{d,f,U}(K)}$ .  $\square$

**Bassin de capture**

La théorie de la viabilité s'intéresse aussi à la recherche de l'ensemble des trajectoires respectant les contraintes de viabilité et atteignant une cible  $C \subset X$  en un temps fini.

Le bassin de capture est l'ensemble de tous les états initiaux  $x_0 \in K$  pour lesquels il existe une telle trajectoire. Cet ensemble est noté  $Capt_{c,f,U}(K, C)$  dans le cas d'un problème continu, avec :

$$Capt_{c,f,U}(K, C) = \left\{ x \in K \mid \exists x(\cdot) \in S_c(x), \exists t^* \geq 0 \left| \begin{array}{l} x(t^*) \in C \\ \forall t \in [0, t^*], x(t) \in K \end{array} \right. \right\} \quad (3.9)$$



et est noté  $Capt_{d,f,U}(K, C)$  dans le cas d'un problème discret, avec :

$$Capt_{d,f,U}(K, C) = \left\{ x \in K \mid \exists (x^n)_{n \in \mathbb{N}} \in S_d(x), \exists n^* \geq 0 \left| \begin{array}{l} x^{n^*} \in C \\ \forall n \in [0, \dots, n^*], x^n \in K \end{array} \right. \right\}. \quad (3.10)$$

Le bassin de capture à temps fixé  $T$  pour le cas continu (resp.  $N$  pour le cas discret) est l'ensemble de tous les états initiaux  $x_0 \in K$  pour lesquels il existe une trajectoire respectant les contraintes de viabilité et atteignant une cible  $C \subset X$  en un temps  $t^* \leq T$  (resp.  $n^* \leq N$ ). Cet ensemble est noté  $Capt_{c,f,U,T}(K, C)$  dans le cas d'un problème continu, avec :

$$Capt_{c,f,U,T}(K, C) = \left\{ x \in K \mid \exists x(\cdot) \in S_c(x), \exists t^* \in [0, T] \left| \begin{array}{l} x(t^*) \in C \\ \forall t \in [0, t^*], x(t) \in K \end{array} \right. \right\} \quad (3.11)$$

et est noté  $Capt_{d,f,U,N}(K, C)$  dans le cas d'un problème discret, avec :

$$Capt_{d,f,U,N}(K, C) = \left\{ x \in K \mid \exists (x^n)_{n \in \mathbb{N}} \in S_d(x), \exists n^* \in [0, N] \left| \begin{array}{l} x^{n^*} \in C \\ \forall n \in [0, \dots, n^*], x^n \in K \end{array} \right. \right\}. \quad (3.12)$$

### 3.1.2 Système tychastique

#### Évolution et système tychastique

**Cas continu :** Une évolution  $t \in [0, +\infty[ \mapsto x(t) \in X$  décrit l'état du système pour tout  $t \in [0, +\infty[$ . Elle peut être influencée par des actions extérieures appelées perturbations. La dynamique dépend donc de l'état du système  $x(t)$  et des différentes perturbations  $v(t)$  qui influencent l'évolution du système au temps  $t$ . Le système tychastique continu est donné par :

$$St_c \begin{cases} x'(t) = f(x(t), v(t)) \text{ pour } t \geq 0 \\ v(t) \in V(x(t)) \end{cases}. \quad (3.13)$$

Les perturbations  $v(t)$  appartiennent à un ensemble  $V(x(t)) \subset \mathbb{R}^q$  qui dépend de l'état du système au temps  $t$ . La correspondance  $V : X \rightsquigarrow Z$  décrit l'ensemble des perturbations  $V(x) \subset Z$  admissibles lorsque le système est dans l'état  $x$ . A partir d'un état initial  $x_0$ , il peut exister plusieurs évolutions possibles, en fonction des différentes perturbations au cours du temps  $t \mapsto v(t)$ . Une solution du système (3.13) est une évolution  $t \mapsto x(t)$  telle qu'il existe une fonction de perturbations  $t \mapsto v(t)$  telle que les conditions de (3.13) soient vérifiées pour tout  $t \geq 0$ . On note  $St_c(x)$  l'ensemble des évolutions issues de  $x$  gouvernées par le système tychastique continu (3.13) pour tout  $t$ , i.e. :

$$St_c(x) = \left\{ x(\cdot) \mid x(0) = x \text{ et } \left| \begin{array}{l} x'(t) = f(x(t), v(t)) \\ v(t) \in V(x(t)) \end{array} \right. \text{ pour presque tout } t \geq 0 \right\}. \quad (3.14)$$

**Remarque importante :** Bien que formellement la variable de perturbation  $v$  soit de même nature que la variable de contrôle  $u$  (on a d'ailleurs volontairement copié ici le texte associé

au système dynamique (3.1), en remplaçant  $u$  par  $v$  et "contrôle" par "perturbation" ou "tychastique", la question traitée dans cette section est différente de celle posée dans la Section 3.1.1 : on ne cherche plus à savoir si une propriété (celle d'appartenir à l'ensemble  $K$ ) est satisfaite pour au moins une trajectoire d'un système régulé (trajectoire viable d'un système contrôlé), mais on veut qu'elle soit satisfaite pour toutes les trajectoires d'un système tychastique.

Ainsi, le cas d'un système tychastique discret s'adapte de manière analogue à partir de la Partie 3.1.1.

**Cas discret :** Une évolution  $(x^n)_{n \in \mathbb{N}}$  décrit l'état du système pour tout  $n \in \mathbb{N}$ . Le système tychastique discret est donné par :

$$St_d \left\{ \begin{array}{l} x^{n+1} = f(x^n, v^n) \text{ pour } n \in \mathbb{N} \\ v^n \in V(x^n) \end{array} \right. . \quad (3.15)$$

On note  $St_d(x)$  l'ensemble des évolutions issues de  $x$  gouvernées par le système tychastique discret (3.3) pour tout  $n \in \mathbb{N}$ , i.e. :

$$St_d(x) = \left\{ (x^n)_{n \in \mathbb{N}} \mid x^0 = x \left\{ \begin{array}{l} x^{n+1} = f(x^n, v^n) \\ v^n \in V(x^n) \end{array} \right. \text{ pour tout } n \in \mathbb{N} \right\} . \quad (3.16)$$

### Évolution et ensemble invariant

On conserve la notion de trajectoires viables définies aux Équations (3.5) et (3.6).

**Définition 3.1.5.** *Un ensemble  $L$  est invariant pour le système tychastique continu (3.13) (resp. discret (3.15)), si*

$$\forall x \in L, \forall x(\cdot) \in St_c(x) \text{ (resp. } St_d(x)), x(\cdot) \text{ est viable dans } L.$$

### Noyau invariant

Le noyau invariant associé au système dynamique tychastique continu (3.13) (resp. discret (3.15)) et soumis à l'ensemble de contraintes de viabilité  $K$ , noté  $Inv_{c,f,V}(K)$  (resp.  $Inv_{d,f,V}(K)$ ), est l'ensemble de tous les états initiaux  $x_0 \in K$  pour lesquels toutes les solutions  $x(\cdot)$  du système (3.13) (resp.  $(x^n)_{n \in \mathbb{N}}$  du système (3.15)) partant de  $x_0$  sont viables dans  $K$ .

Dans le cas continu, on a donc :

$$Inv_{c,f,V}(K) = \{x_0 \in K \mid \forall x(\cdot) \in St_c(x_0), \forall t \geq 0, x(t) \in K\}. \quad (3.17)$$

Dans le cas discret, on a donc :

$$Inv_{d,f,V}(K) = \{x_0 \in K \mid \forall (x^n)_{n \in \mathbb{N}} \in St_d(x_0), \forall n \in \mathbb{N}, x^n \in K\}. \quad (3.18)$$

### Cas continu :

Aubin (1991) a trouvé un ensemble de conditions suffisantes sur le système dynamique tychastique pour que le noyau invariant associé soit fermé. Pour énoncer ce résultat, nous rappelons quelques définitions sur les correspondances et les systèmes dynamiques.

Soit  $F$  une correspondance  $F : X \rightsquigarrow Y$ . On note  $B$  la boule unité de l'espace d'état  $Y$ . Une correspondance  $F : X \rightsquigarrow Y$  est  $\lambda$ -Lipschitz (ou Lipschitz pour la constante  $\lambda > 0$ ) si

$$\forall x_1, x_2, F(x_1) \subset F(x_2) + \lambda \|x_1 - x_2\| B$$

**Définition 3.1.6.** Le système dynamique tychastique  $(f, V)$  est Lipschitz si

- $V : X \rightsquigarrow Z$  est Lipschitz à images compactes
- $f : X \times V \rightsquigarrow X$  est Lipschitz

**Théorème 3.1.7.** Aubin (1997) Si  $K$  est fermé et  $(f, V)$  est un système dynamique tychastique Lipschitz, alors  $\text{Inv}_{c,f,V}(K)$  est fermé.

**Cas discret :**

**Théorème 3.1.8.** Si  $K$  est fermé et  $(f, V)$  est un système dynamique tychastique avec  $f$  et  $V$  est continues, alors  $\text{Inv}_{d,f,V}(K)$  est fermé.

*Démonstration.* On adapte la preuve du Théorème 3.1.4. Montrons que  $\overline{\text{Inv}_{d,f,V}(K)} = \text{Inv}_{d,f,V}(K)$ . Soit  $x \in \overline{\text{Inv}_{d,f,V}(K)}$ , il suffit de montrer que pour tout  $v \in V(x)$   $f(x, v) \in \overline{\text{Inv}_{d,f,V}(K)}$ . Soit  $(x_n)_{n \in \mathbb{N}} \in \text{Inv}_{d,f,V}(K)^{\mathbb{N}}$  une suite d'éléments du noyau d'invariance qui converge vers  $x$ . Puisque  $V$  est continue et que  $x_n$  converge vers  $x$ , pour tout  $n \in \mathbb{N}$  il existe  $v_n \in V(x_n)$  tel que  $(v_n)_n$  converge vers  $v$ . On a pour tout  $n \in \mathbb{N}$ ,  $f(x_n, v_n) \in \text{Inv}_{d,f,V}(K)$  car  $x_n \in \text{Inv}_{d,f,V}(K)$ . De plus  $f$  est continue, donc  $f(x, v) \in \overline{\text{Inv}_{d,f,V}(K)}$ .  $\square$

### 3.1.3 Système contrôlé et tychastique

#### Évolution et système contrôlé et tychastique

**Cas continu :**

La théorie de la viabilité traite aussi le cas de systèmes contrôlés soumis à des perturbations, appelées tyches. On considère donc à la fois contrôle  $u(t) \in U(x(t))$  et perturbation  $v(t) \in V(x(t))$  issus respectivement des systèmes contrôlés (3.1) et des systèmes tychastiques (3.13), ce qui donne le système contrôlé et tychastique continu (3.19) donné par :

$$\begin{cases} x'(t) = f(x(t), u(t), v(t)) \text{ pour } t \geq 0 \\ u(t) \in U(x(t)) \\ v(t) \in V(x(t)) \end{cases}, \quad (3.19)$$

où les contrôles  $u(t)$  appartiennent encore à un ensemble  $U(x(t)) \subset \mathbb{R}^p$  qui dépend de l'état du système au temps  $t$ . La correspondance  $U : X \rightarrow Y$  décrit l'ensemble des contrôles  $U(x) \subset Y$  admissibles lorsque le système est dans l'état  $x$ . Les perturbations  $v(t)$  appartiennent à un ensemble  $V(x(t)) \subset \mathbb{R}^s$  qui dépend également de l'état du système au temps  $t$ . La correspondance  $V : X \rightarrow Z$  décrit l'ensemble des perturbations  $V(x) \subset Z$  admissibles lorsque le système est dans l'état  $x$ .

A partir d'un état initial  $x_0$ , il peut exister plusieurs évolutions possibles, en fonction des différents choix de contrôles  $t \mapsto u(t)$  et des différentes perturbations subies au cours du temps  $t \mapsto v(t)$ . Une solution du système (3.19) est une évolution  $t \mapsto x(t)$  telle qu'il existe une fonction de contrôle  $t \mapsto u(t)$  et une fonction de perturbation  $t \mapsto v(t)$  telles que les conditions de (3.19) soient vérifiées pour presque tout  $t \geq 0$ .

**Cas discret :** Par analogie, le système contrôlé et tychastique discret (3.20) est donné par :

$$\begin{cases} x^n = f(x^n, u^n, v^n) \text{ pour } n \in \mathbb{N} \\ u^n \in U(x^n) \\ v^n \in V(x^n) \end{cases} . \quad (3.20)$$

### Évolution et ensemble viable garanti

Un ensemble  $L$  a la propriété de viabilité garantie pour le système dynamique contrôlé et tychastique (3.19) dans le cas continu (resp. (3.20) dans le cas discret) s'il existe un contrôle  $\tilde{u}(\cdot) : x \in L \mapsto \tilde{u}(x) \in U(x)$  tel que pour toutes conditions initiales  $x \in L$ , toutes les solutions  $x(\cdot)$  (resp.  $(x^n)_{n \in \mathbb{N}}$ ) de (3.20) avec  $u(t) = \tilde{u}(x(t))$  (resp.  $u^n = \tilde{u}(x^n)$ ) et telles que  $x(0) = x$  (resp.  $x^0 = x$ ) sont viables dans  $L$ , i.e., if  $L$  est invariant sous  $f(\cdot, \tilde{u}(\cdot), V(\cdot))$ .

Dans le cas continu, quand les contrôles sélectionnés  $\tilde{u}$  sont  $\lambda$ -Lipschitz, on dit que  $L$  satisfait la propriété de  $\lambda$ -leading.

### Noyau de viabilité garanti

On définit le noyau de viabilité garanti associé à l'ensemble de contraintes de viabilité  $K$ , comme étant le plus grand ensemble  $L \subset K$  ayant la propriété de viabilité garantie, pour des contrôles sélectionnés  $\tilde{u}$  qui sont  $\lambda$ -Lipschitz dans le cas continu. Dans le cas continu, le noyau de viabilité garanti dépend donc du choix de  $\lambda$ , on le note  $Guar_{c,f,U,V,\lambda}(K)$ . Il est noté  $Guar_{d,f,U,V}(K)$  dans le cas discret.

#### Cas continu :

Aubin (1991) a trouvé un ensemble de conditions suffisantes sur le système dynamique contrôlé et tychastique pour que le noyau de viabilité garanti associé soit fermé. Pour énoncer ce résultat, nous rappelons quelques définitions sur les correspondances et les systèmes dynamiques.

**Définition 3.1.9.** *Un système contrôlé tychastique  $(U, V, f)$  est Lipschitz s'il satisfait les conditions suivantes :*

- $U$  et  $V$  sont Lipschitz à images compactes
- $f$  est Lipschitz

**Théorème 3.1.10.** *Aubin (1997) Si  $K$  est fermé et  $(f, U, V)$  est un système dynamique contrôlé et tychastique Lipschitz, alors  $Guar_{c,f,U,V,\lambda}(K)$  est fermé.*

#### Cas discret :

**Théorème 3.1.11.** *Si  $K$  est fermé et  $(f, U, V)$  est un système dynamique contrôlé tychastique avec  $Graph(U)$  fermé,  $U$  à croissance linéaire et  $V$  et  $f$  continues, alors  $Guar_{d,f,U,V}(K)$  est fermé.*

**Démonstration.** On adapte les preuves des Théorèmes 3.1.4 et 3.1.8. Montrons que  $\overline{Guar_{d,f,U,V}(K)} = Guar_{d,f,U,V}(K)$ . Soit  $x \in \overline{Guar_{d,f,U,V}(K)}$ , il suffit de montrer qu'il existe  $u \in U(x)$  tel que pour tout  $v \in V(x)$   $f(x, u, v) \in Guar_{d,f,U,V}(K)$ . Soit  $(x_n)_{n \in \mathbb{N}} \in Guar_{d,f,U,V}(K)^\mathbb{N}$  une suite d'éléments du noyau de viabilité qui converge vers  $x$ . Puisque pour tout  $n \in \mathbb{N}$ ,  $x_n \in Guar_{d,f,U,V}(K)$ , il existe  $u_n \in U(x_n)$  tel que  $f(x_n, u_n, V(x_n)) \in Guar_{d,f,U,V}(K)$ . Or  $U$  a une croissance linéaire,  $Graph(U)$  est fermé et  $(x_n)_n$  converge vers  $x$ , il existe donc une

sous-suite de  $(u_n)_n$  (encore notée  $(u_n)_n$ ) qui converge vers  $u \in U(x)$ . De plus, puisque  $V$  est continue et que  $x_n$  converge vers  $x$ , pour tout  $n \in \mathbb{N}$  il existe  $v_n \in V(x_n)$  tel que  $(v_n)_n$  converge vers  $v$ . On donc pour tout  $n \in \mathbb{N}$ ,  $f(x_n, u_n, v_n) \in \overline{Guar_{d,f,U,V}(K)}$ . De plus  $f$  est continue, donc  $f(x, u, v) \in \overline{Guar_{d,f,U,V}(K)}$ .  $\square$

## 3.2 Calculs de noyaux de viabilité

Dans la section précédente, nous avons introduit les concepts fondamentaux de la théorie de la viabilité. Parmi eux, le noyau de viabilité pour un système dynamique contrôlé, et le noyau de viabilité garanti pour un système dynamique contrôlé et tychastique. Dans cette section, nous présentons les moyens dont nous disposons pour calculer ces noyaux. Lorsque la dimension du système dynamique est faible, des algorithmes existent et permettent de calculer des approximations de noyaux de viabilité associés au problème général de viabilité que nous avons introduit à la Section 3.1. Nous les présentons dans une première partie. Par contre, dès lors qu'on dépasse ce cadre (dimension du système trop grande), ces algorithmes ne sont plus utilisables en pratique. Dans la littérature, les résolutions de problèmes de viabilité en grande dimension, formulés en dimension quelconque, se font par le calcul analytique de noyaux de viabilité sur des problèmes particuliers, comme dans le cas d'un système linéaire (Rapaport et al., 2006), ou par l'utilisation de méthodes permettant de ramener le problème de viabilité à l'étude de problèmes en dimension plus petite. Nous présentons les techniques employées pour la résolution de certains de ces problèmes dans une seconde partie.

### 3.2.1 Outils et méthodes pour le calcul générique d'approximation de noyaux de viabilité en petite dimension

Peu de calculs exacts de noyaux de viabilité par des méthodes analytiques ont été effectués. C'est le cas pour des systèmes contrôlés, dans Bernard and Martin (2012) sur un modèle de population où deux langues sont en compétition, dans Aubin and Saint-Pierre (2007) sur un modèle de dynamique de population du type Verhulst avec rétroaction malthusienne, et dans Aubin (1991) (chap. 6.2) sur un modèle économique simple. Dans ces trois exemples, l'espace des états de dimension deux et espace des contrôles de dimension un. La possibilité de détermination exacte de noyaux de viabilité est à étudier au cas par cas (Martin, 2015). Des algorithmes d'approximations des noyaux existent depuis les années 1990.

#### Les algorithmes

Le premier d'entre eux est celui de Saint-Pierre (1994). L'auteur a proposé une méthode de construction de noyaux de viabilité utilisant des approximations discrètes dans le cas Lipschitz. Cet algorithme se décompose en deux étapes :

- l'approximation du noyau de viabilité du système continu par des noyaux de systèmes discrets en temps,
- puis par des noyaux de systèmes discrets en temps et en espace.

Le calcul du noyau de viabilité du système dynamique discret fini est itératif. Il repose sur l'utilisation d'une grille régulière sur l'espace d'état. A chaque étape, le programme construit et mémorise une structure qui représente un sous-ensemble de cette grille, qu'il modifie en retirant certains points. Ainsi, le schéma numérique proposé approche le noyau de viabilité par une suite décroissante d'ensembles qui contiennent le noyau de viabilité. La convergence de

cette suite vers le noyau de viabilité a été prouvée par Saint-Pierre (1994). L'implémentation de l'algorithme est réalisée dans le logiciel VIABLAB (Bibliothèque d'algorithmes numériques de viabilité, le code est utilisé dans les expériences numériques de Durand et al. (2017)). Cependant, l'utilisation d'une grille fait que la mémoire nécessaire pour le calcul effectif du noyau est exponentielle avec la dimension de l'espace des états. Des travaux ont été réalisés depuis pour améliorer cet algorithme. Ainsi, une approche utilisant des méthodes de classification a été développée dans Deffuant et al. (2007). Elle vise à ne plus conserver en mémoire tous les points de la grille appartenant au noyau, mais uniquement une partie d'entre eux nécessaire pour construire la fonction de classification. Cette fonction permet d'approcher la frontière du noyau. Cet algorithme a été utilisé avec plusieurs fonctions de classification :

- Les Support Vector Machines qui permettent une définition parcimonieuse du noyau de viabilité (Deffuant et al., 2007). Cet algorithme est implémenté dans le logiciel KAVIAR (Kernel Approximation for VIability and Resilience, (Chapel and Deffuant, 2011)), disponible à l'adresse <http://www.kaviar.prd.fr>.
- Les kd-tree Alvarez et al. (2016) qui permettent de représenter un ensemble avec des divisions plus fines près de sa frontière. Cet algorithme est implémenté dans le logiciel Viabilitree, disponible à l'adresse <https://github.com/ISCPIF/viabilitree>.

Une autre amélioration de l'algorithme de Saint-Pierre (1994) concerne le temps de calcul. Pour cela, Brias et al. (2016) proposent une implémentation de l'algorithme permettant le calcul parallèle sur des processeurs graphiques (GPU). Sur un exemple, les auteurs montrent un gain de temps de l'ordre d'un facteur 20 comparé aux temps de calcul sur un processeur classique (CPU).

### Exemples de calculs de noyaux de viabilité

Dans cette section, nous présentons des approximations de noyaux de viabilité obtenues par des algorithmes présentés à la section précédente. Nous ne présentons ici que les équations des dynamiques utilisées pour ces approximations, se référer aux articles cités, et aux références incluses dans ces derniers pour des précisions concernant ces modèles (notamment concernant leurs paramètres).

Le premier exemple concerne une dynamique de gestion de pâturage (Martin et al., 2011). Le système est composé de deux variables : la biomasse aérienne  $s$  et la pression de pâturage  $\gamma_g$ . Des gestionnaires du territoire ont la possibilité de modifier la pression de pâturage, qui constitue la seule variable de contrôle du système (le contrôle porte sur les variations de  $\gamma_g$ , par sa dérivée). La dynamique du système est donnée par :

$$\begin{cases} s'(t) = \frac{r_c}{\gamma_c} s(t) (a_c + r_s s(t)) \left( 1 - \frac{s(t)}{s^*} - \alpha_{ws} \right) - \gamma_g(t) s(t) \\ \gamma_g'(t) = u \in [-0.02; 0.02] \end{cases} \quad (3.21)$$

Martin et al. (2011) utilisent le logiciel KAVIAR pour calculer une approximation du noyau de viabilité du système (3.21) associé à l'ensemble de contraintes  $K = [0; +\infty] \times [0.1; 0.18]$ .

Le second exemple concerne une dynamique d'eutrophisation des lacs (Alvarez et al., 2016). Le système est composé de deux variables : la concentration de phosphore  $P$  dans le lac et la quantité de phosphore entrant  $L$  (due par exemple aux fertilisants). Le contrôle porte sur les variations de la quantité de phosphore entrant (par sa dérivée). La dynamique du système est donnée par :

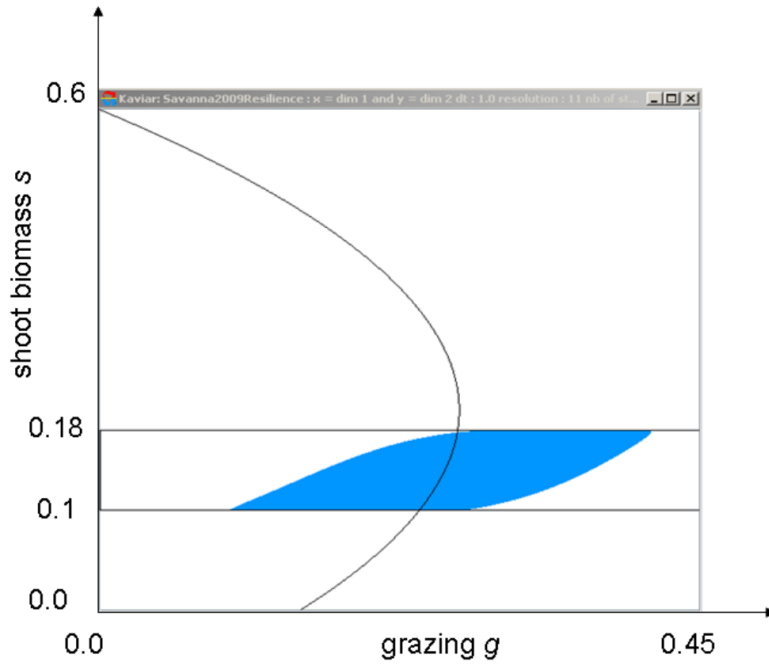


FIGURE 3.1 – En bleu, le noyau de viabilité pour la dynamique (3.21) avec l'ensemble des contraintes  $K = [0; +\infty] \times [0.1; 0.18]$ , calculé avec le logiciel KAVIAR, extrait de Martin et al. (2011).

$$\begin{cases} P'(t) = -bP(t) + r \frac{P^q(t)}{m + P^q(t)} + L(t) \\ L'(t) = u(t) \in [-V, V] \end{cases} \quad (3.22)$$

Alvarez et al. (2016) utilisent le logiciel Viabilitree pour calculer une approximation du noyau de viabilité du système (3.22) associé à l'ensemble de contraintes  $K = [L_{min}, L_{max}] \times [0, P_{max}]$ .

### Les limites d'utilisation

Une limite aux calculs d'approximations de noyaux de viabilité vient des ordinateurs standards actuels (Chapel and Deffuant, 2011). Par exemple, les auteurs soulignent que dans un espace d'état de dimension six, si on considère une grille régulière de 20 points par dimension, on obtient 64 millions de points au total (qui commence à être difficile à manipuler). De plus, la dimension de l'espace d'action (de contrôle) représente un autre défi. En effet, l'approche la plus élémentaire consiste à discrétiser l'espace des actions afin de l'utiliser pour chaque point de l'espace d'état. Là encore, les ressources informatiques nécessaires pour mener ces calculs augmentent de manière exponentielle avec la dimension de l'espace de contrôle.

### 3.2.2 Calcul en grande dimension dans des cas particuliers

Dans cette section, nous présentons des méthodes employées pour la résolution de problèmes de viabilité en grande dimension, formulés en dimension quelconque. Les systèmes dynamiques considérés ici sont composés d'une variable globale et de variables locales. Chaque problème repose sur des dynamiques particulières que nous préciserons. Les auteurs utilisent des méthodes



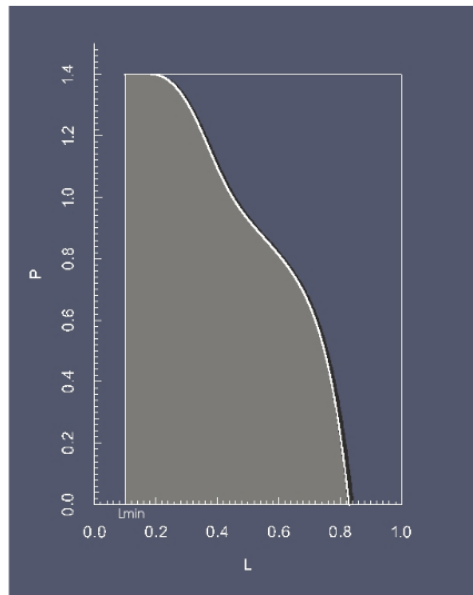


FIGURE 3.2 – En gris, le noyau de viabilité pour la dynamique (3.22) avec l'ensemble des contraintes  $K = [L_{min}, L_{max}] \times [0, P_{max}]$ , calculé avec le logiciel Viabilitree, extrait de Alvarez et al. (2016).

reposant sur ces spécificités pour soit ramener le problème de viabilité en grande dimension à l'étude de problèmes en dimension plus petite, soit calculer analytiquement le noyau.

Dans Aubin et al. (2012) et Aubin et al. (2013), la dynamique de la variable globale dépend de son état ainsi que de celui des variables locales et des contrôles associés à ces dernières. L'évolution de chaque variable locale est indépendante des autres variables du système, mais est soumise à des perturbations. Aubin et al. (2013) s'intéressent à la gestion de la récolte de  $n$  espèces alors que Aubin et al. (2012) étudient la répartition des droits d'émission de polluants pour  $n$  entités. Ces deux articles développent des modèles et des techniques très similaires, c'est pourquoi nous nous référons plus particulièrement à Aubin et al. (2013). Le problème posé est de garantir une récolte sur l'ensemble des espèces (contrainte globale) tout en assurant un seuil minimal pour la ressource de chaque espèce (contrainte locale). La contrainte imposée sur chaque variable est indépendante de celle imposée aux autres variables. L'évolution de la variable globale est soumise à des contrôles et des perturbations, le cadre considéré est donc tychastique. Les auteurs étudient d'abord le problème de viabilité sur cette variable. Ils calculent donc un noyau de viabilité garanti, qui donne l'existence de contraintes sur chaque variable locale telles que la variable globale respecte sa contrainte pour toutes les évolutions des variables locales. La deuxième étape consiste à assurer que la contrainte initiale de chaque variable locale, ainsi que les nouvelles contraintes déduites de l'étape précédente sur chaque variable locale sont respectées quelles que soient les perturbations sur les rendements, ils calculent alors le noyau de viabilité invariant pour chaque espèce. Ils trouvent ainsi la ressource minimale de chaque espèce à l'instant initial nécessaire pour qu'à chaque instant les contraintes sur la variable globale et les variables locales soient vérifiées. Cette méthode en deux étapes est appelée méthode de décentralisation de la contrainte globale. Son application repose ici sur le fait que la dynamique de chaque variable locale ne dépend pas de l'évolution des autres variables locales. En particulier, le problème initial de calcul d'un noyau de viabilité garanti en grande dimension est remplacé par le calcul d'un noyau de viabilité garanti pour la variable globale et par le calcul de noyaux invariants pour toutes les variables locales.



Dans Doyen and Péreau (2012) et Péreau et al. (2012), les auteurs considèrent que la dynamique de la variable globale dépend de son état ainsi que des contrôles associés à chaque variable locale. De plus, la dynamique de chaque variable locale ne dépend que du contrôle associé et de la variable globale. Dans ces deux articles, les auteurs considèrent des contraintes sur les variables locales et la variable globale, et cherchent à calculer le noyau de viabilité associé au système. Pour cela, ils transfèrent les contraintes associées aux variables locales sur les contrôles correspondants, ce qui permet de ne plus considérer les variables locales dans la dynamique du système qui ne comporte ainsi que la variable globale. Cela est rendu possible par le fait que chaque variable locale ne dépend pas des autres variables locales.

Ainsi, dans Doyen and Péreau (2012), les auteurs s'intéressent à la répartition de l'effort de pêche (contrôle) de  $n$  pêcheurs sur une espèce pour garantir un revenu minimal à chaque pêcheur. Plus précisément, ils étudient les caractéristiques que doit vérifier une coalition (groupe formé par des pêcheurs, dont les auteurs considèrent la recette commune) pour que cette dernière ait un gain strictement positif quelles que soient les actions des pêcheurs ne faisant pas partie de cette coalition (donc vues comme des perturbations). Le système est donc composé d'une variable globale, de contrôles et de perturbations. Le fait de garantir un gain minimum à la coalition implique (du fait des dynamiques choisies), de garantir un niveau minimal de ressource pour l'espèce. Les auteurs calculent analytiquement le noyau de viabilité garanti associé à ce système pour cette dernière contrainte. Ce calcul analytique est rendu possible par la dynamique simple du modèle et la contrainte ouverte considérée. Ce dernier point implique que le cadre de viabilité dans lequel se place cet article ne correspond pas au cadre dans lequel nous avons présenté les concepts généraux de la théorie de la viabilité (Section 3.1) et dans lequel s'exprime notre problème de viabilité (Section 3.3). Néanmoins, les ensembles de viabilité sont encore définis dans ce cas, seules leurs propriétés diffèrent. De plus, le calcul du noyau de viabilité repose sur le calcul d'une quantité à chaque pas de temps (définition "backward induction" du noyau), qui se trouve être indépendante du temps. Suite au calcul du noyau garanti, les auteurs montrent que le joueur ayant le meilleur rendement est un joueur clef : il est obligé d'appartenir à la coalition pour que la dynamique soit viable et cela encourage les grandes coalitions.

Dans Péreau et al. (2012), les auteurs étudient les conditions pour qu'un modèle bio-éco-socio soit viable. Le modèle concerne là encore la pêche, et ils s'intéressent au quota global à imposer sur la quantité de ressource à prélever pour que la ressource ne disparaisse pas (contrainte biologique), que tous les pêcheurs aient un minimum de travail à fournir (contrainte sociale), et qu'ils maximisent leur profit (contrainte économique, les pêcheurs ont la possibilité de vendre ou d'acheter des quotas à d'autres pêcheurs). Les contraintes sur les variables locales (profits) sont transférées sur les contraintes des contrôles associés, ce qui permet de ne plus considérer directement les profits dans la dynamique du système. De plus, les auteurs considèrent que chaque agent maximise son profit, ce qui correspond à un contrôle optimal pour chaque agent. Ces derniers sont ensuite agrégés dans une somme pour n'obtenir qu'une variable pour le contrôle. Ainsi, les auteurs remplacent ce système d'équations avec  $n$  variables locales, une variable globale et  $n$  contrôles par un système ne comprenant qu'une variable globale et un contrôle. Ils calculent ensuite analytiquement le noyau de viabilité du système. Ils trouvent donc les ressources initiales depuis lesquelles il existe un quota global qui, à chaque temps, permet de respecter les trois contraintes.

Dans Martin et al. (en cours d'écriture), qui étend le problème de la coopérative agricole de Martin et al. (2015), nous considérons que la dynamique de la variable globale dépend de son état ainsi que de celui des variables locales et des contrôles associés. De plus, nous supposons que la dynamique de chaque variable locale ne dépend que de son état, du contrôle associé à

cette variable, et de la variable globale. Nous nous intéressons à l'obtention de garanties sur la viabilité collective à partir de viabilités individuelles. Nous nous intéressons au cas d'une coopérative agricole et nous décrivons l'évolution du gain de chaque membre de la coopérative ainsi que celui du prix de vente de la coopérative. Chaque membre apporte chaque année une quantité fixée à la coopérative, et contrôle la qualité de sa production. Nous cherchons à savoir s'il existe des trajectoires de la coopérative qui respectent des contraintes sur la variable globale et sur chacune des variables locales. Ainsi, l'objectif de l'article est de déterminer le noyau de viabilité associé à ce système. Plus précisément, nous cherchons un sous-ensemble du noyau. Pour cela, nous transformons l'étude de ce problème de viabilité en l'étude d'un problème de viabilité pour chaque variable locale (associée à la variable globale). Ainsi, il y a un problème de viabilité à résoudre pour chaque membre de la coopérative. Pour cela, et pour chacun de ces problèmes, nous considérons l'ensemble des contrôles des autres participants comme relevant d'une perturbation. Ainsi, il s'agit de calculer dans chacun des cas un noyau de viabilité garanti. Nous explicitons le passage de l'ensemble des noyaux individuels à un ensemble garantissant la viabilité collective, c'est-à-dire que pour des conditions initiales dans cet ensemble, il existe une trajectoire pour chaque membre (au moins un choix de la qualité à chaque instant), tel que les gains de chaque membre et le prix de vente de la coopérative sont à tout temps dans l'ensemble de contraintes. Dans le cas de la coopérative, nous calculons de manière analytique le noyau de viabilité garanti pour chacun des problèmes de viabilité individuels (de petite dimension, donc les algorithmes de calcul de noyau peuvent s'appliquer). Cela est rendu possible notamment par le fait que l'influence des perturbations sur la dynamique est résumée par une variable unidimensionnelle et par des propriétés de monotonie de la dynamique considérée.

La technique que nous développons pour résoudre notre problème de viabilité (voir section suivante), est proche de celle développée par Martin et al. (en cours d'écriture). Certes le problème de départ n'est pas le même puisqu'il s'agit pour ces auteurs de déterminer un noyau de viabilité, alors que nous voulons ici déterminer un noyau de viabilité garanti. Nous reprenons l'idée de réduire la dimension de l'espace d'état pour transformer le problème initial en grande dimension en plusieurs problèmes de petite dimension. De plus, dans chacun de ces problèmes, l'action des autres entités est considérée comme une perturbation. Nous procédons également de la sorte en considérant l'état des autres sous-systèmes comme une perturbation. Néanmoins, dans Martin et al. (en cours d'écriture), comme dans Doyen and Péreau (2012), Péreau et al. (2012) et Aubin et al. (2012), la dynamique de chaque variable locale ne dépend que de son état et de la variable globale. Notre problème est différent en ce sens puisque la dynamique d'un sous-système peut dépendre de l'état d'autres sous-systèmes, mais nous utilisons aussi la structure liée à notre problème pour mener un raisonnement analogue.

### 3.3 Un problème de viabilité garanti en grande dimension

Dans cette section, nous commençons par une formalisation générale du problème dont la gestion de l'invasion des renouées le long d'un cours d'eau est une instance (voir Chapitre 5). Nous en proposons ensuite une résolution.

#### 3.3.1 Description du problème

On s'intéresse à un système composé de plusieurs sous-systèmes. Dans le cas de l'invasion des renouées le long d'un cours d'eau, les sous-systèmes décrivent la dynamique de l'aire envahie dans chaque zone proche du cours d'eau. Le système est aussi composé d'une variable globale

qui décrit l'environnement dans lequel évoluent les sous-systèmes. Dans notre exemple, la variable globale est un réservoir d'eau pour les crues. A la manière du cours d'eau qui ne peut transporter les propagules que de l'amont à l'aval, on suppose les sous-systèmes ordonnés et tels que l'évolution d'un sous-système ne dépend que des variables des sous-systèmes précédents. Ainsi, les sous-systèmes sont composés de variables qui leur sont propres (l'aire envahie dans notre exemple) ainsi que d'autres variables (le réservoir d'eau, l'aire des zones en amont) servant à modéliser l'environnement. On suppose que l'ensemble des contrôles admissibles dans chaque sous-système ne dépend que de l'état de ce sous-système. Dans notre exemple, cela revient à considérer que l'intensité maximale de gestion (dans notre cas la fauche) dans chaque zone ne dépend que de l'état de cette zone. On considère que ce système évolue dans un environnement qui provoque des perturbations sur sa dynamique. Cet environnement est décrit par des variables dont l'évolution est indépendante de celles du système mais qui influencent l'ensemble des perturbations possibles à chaque instant sur la dynamique du système. On considère que pour chaque sous-système, l'ensemble des perturbations possibles dépend également de son état.

On suppose que seules les variables globales interviennent dans les contraintes sur l'environnement et on suppose que dans les contraintes d'un sous-système, les états des autres sous-systèmes n'interviennent pas. Par exemple dans notre cas d'application, les contraintes sur les aires maximales envahies sont fixées pour chaque zone mais peuvent dépendre du temps écoulé depuis le début du projet de gestion. On considère un ensemble de contraintes pour l'environnement et chacun des sous-systèmes tel que l'ensemble des contraintes de chaque sous-système ne dépend que de la variable globale.

Le problème est donc de trouver les conditions initiales à partir desquelles la dynamique du système peut respecter les contraintes de viabilité, et ce pour toutes les perturbations envisagées. C'est-à-dire, dans notre cas, identifier les niveaux d'invasion dans chaque zone pour lesquels il existe une stratégie de gestion qui permet de garantir un niveau d'invasion dans chaque zone, à chaque temps, et pour tous les scénarios de crues considérés. Cela revient à calculer un noyau de viabilité garanti (voir Section 3.1.3).

### 3.3.2 Formalisation mathématique

Nous présentons ici la formalisation mathématique du système contrôlé tychastique particulier associé aux hypothèses de la section précédente.

Soit  $N \in \mathbb{N}$  le nombre de sous-systèmes (grand, correspond au nombre de zones) et pour tout  $i \in [1, N]$  on note  $k_i \in \mathbb{N}$  la dimension du sous-système (petite). On note  $X_i := \mathbb{R}^{k_i}$  l'espace d'état du sous-système  $i$ . La variable globale est de dimension  $q \in \mathbb{N}$  (petite). Ainsi, on note  $X$  l'espace d'état du système, constitué des  $N$  sous-systèmes et de la variable globale. On a  $X = \mathbb{R}^q \times X_1 \times \dots \times X_N$  et pour  $x \in X$ , on note  $x = (g, x_1, \dots, x_N)$ .

On considère le système dynamique contrôlé et tychastique à temps discret suivant :

$$\begin{cases} x(n+1) = f(x(n), u(n), v(n)) \text{ pour } n \in \mathbb{N} \\ u(n) \in U(x(n)) \\ v(n) \in V(x(n)) \end{cases} \quad (3.23)$$

Les contrôles et les perturbations sont donnés respectivement par  $U : X \rightsquigarrow Y := Y_1 \times \dots \times Y_N$  et  $V : X \rightsquigarrow Z := Z_g \times Z_1 \times \dots \times Z_N$ .

Pour  $x \in X$ ,  $U(x)$  est tel que  $U(x) = U(g, x_1, \dots, x_N) = (U_1(g, x_1), \dots, U_N(g, x_N))$ . On note  $u = (u_1, \dots, u_N) \in Y$ .

Pour  $x \in X$ ,  $V(x)$  est tel que  $V(x) = V(g, x_1, \dots, x_n) = (V_g(g), V_1(g, x_1), \dots, V_N(g, x_N))$ . On note  $v = (v_g, v_1, \dots, v_N) \in Z$ .

$f$  est définie pour tout  $x \in X$ ,  $u \in U(x)$  et  $v \in V(x)$  par :

$$f(x, u, v) = (f_g(g, v_g), f_1((g, x_1), u_1, v_1), f_2((g, x_1, x_2), (u_1, u_2), v_2), \dots, f_N((g, x_1, \dots, x_N), (u_1, \dots, u_N), v_N)),$$

ainsi l'évolution de l'état de l'environnement ne dépend que de son état et d'une perturbation, l'évolution de l'état du sous-système  $i$  ne dépend que de son état, des états des sous-systèmes précédents, d'un contrôle de ce sous-système et des contrôles des sous-systèmes précédents, et d'une perturbation. L'ensemble des contrôles admissibles qui influencent l'évolution du sous-système  $i$  ne dépend que de l'état de l'environnement et du sous-système  $i$ . L'ensemble des perturbations qui peuvent impacter l'état de l'environnement ne dépend que de cet état, celui de celles qui impactent le sous-système  $i$  ne dépend que de l'état de l'environnement et de celui du sous-système  $i$ .

Soit  $L_g \subset \mathbb{R}^q$  et soit pour tout  $g \in \mathbb{R}^q$ , pour tout  $i$ ,  $L_i(g) \subset X_i$ . On note

$$L := \{(g, x) \in X, g \in L_g \text{ et } \forall i, x_i \in L_i(g)\} \subset X, \quad (3.24)$$

l'ensemble de contraintes. Les contraintes sont bien indépendantes entre chacun des sous-systèmes mais dépendent de l'état de l'environnement.

Notre objectif est de calculer le noyau de viabilité garanti du système (3.23) associé à la contrainte  $L$ . La Section 3.1.3 identifie des conditions suffisantes pour qu'un tel ensemble soit fermé. Nous supposons que le système contrôlé tychastique  $(f, U, V)$  considéré ici remplit ces critères.

### 3.3.3 Résolution du problème

La dimension de l'espace d'état du système (3.23) est  $(q + \sum_{i=1}^N k_i)$ .  $N$  est considéré grand, il n'est donc pas possible d'utiliser un algorithme comme ceux présentés à la Section 3.2.1 pour calculer le noyau garanti associé à ce système dynamique.

Contrairement à la Section 3.2.2, l'évolution de chaque variable locale dépend des autres variables locales. Comme dans Martin et al. (en cours d'écriture), nous proposons une méthode pour transformer le problème de viabilité de ce système en  $N$  problèmes de viabilité en dimension  $q + n_i$  associés chacun à un sous-système. Pour cela, nous considérons les états et contrôles des autres sous-systèmes comme des perturbations.

#### Description de la méthode

Nous présentons ici une méthode appelée cascade de noyaux de viabilité, qui utilise les spécificités du problème de viabilité associé au système (3.23) pour trouver un domaine de viabilité inclus dans le noyau de viabilité garanti.

La méthode que nous proposons est itérative. Elle vise à construire à l'étape  $i$  un noyau de viabilité garanti, noté  $G_i$ , pour une dynamique correspondant à celle du sous-système  $i$ , avec le même ensemble de contrôles mais pour un ensemble de perturbations augmenté. Ces perturbations permettent de prendre en compte les états et contrôles des sous-systèmes précédents.

A l'étape  $i \in [1, N]$ , on suppose  $G_1, \dots, G_{i-1}$  construits. Ce sont des sous-ensembles respectivement de  $L_g \times X_1, \dots, L_g \times X_{i-1}$ . Et on s'intéresse au système dynamique contrôlé tychastique suivant dont l'espace d'état est  $\mathbb{R}^q \times X_i$  :

$$\begin{cases} (g(n+1), x_i(n+1)) = (f_g(g(n), v_g(n)), \tilde{f}_i((g(n), x_i(n)), u_i(n), (v_i(n), \bar{v}_i(n)))) \\ u_i(n) \in U_i(g(n), x_i(n)) \\ (v_g(n), v_i(n), \bar{v}_i) \in V_g(g(n)) \times V_i(g(n), x_i(n)) \times V'_{G_1, \dots, G_{i-1}}(g(n)) := V'_i(g(n), x_i(n)) \end{cases} \quad (3.25)$$

où pour tout  $g \in \mathbb{R}^q$ , on note

$$V'_{G_1, \dots, G_{i-1}}(g) = \overline{V_1}(g, G_1) \times \dots \times \overline{V_{i-1}}(g, G_{i-1})$$

avec

$$\overline{V_j}(g, G_j) := \{(x_j, u_j) \in X_j \times Y_j \mid x_j \in P_j(g, G_j) \text{ et } u_j \in U_j(g, x_j)\},$$

et tel que

$$P_j(g, G_j) := \{x_j \in X_j \mid (g, x_j) \in G_j\}.$$

On note  $v'_i := (v_g, v_i, \bar{v}_i) = (v_g, v_i, ((x_1, u_1), \dots, (x_{i-1}, u_{i-1}))) \in V'_i$

Ainsi, les perturbations de ce système sont celles qui touchent la variable globale et qui appartiennent à  $V_g(g)$ , celles de la zone  $i$  considérées précédemment et qui appartiennent à  $V_i(g, x_i)$ , toutes deux décrites dans le système initial (3.23), auxquelles on ajoute celles issues des variables des sous-systèmes précédents. Ainsi, les variables d'état  $x_j$  et les contrôles  $u_j$  des sous-systèmes précédents ( $j < i$ ) sont considérés comme des perturbations, et appartiennent à l'ensemble  $V'_{G_1, \dots, G_{i-1}}$  construit à partir des coupes  $P_j$ , selon la variable globale, des noyaux garantis calculés aux étapes précédentes.

De plus, on pose pour tout  $(g, x_i) \in \mathbb{R}^q \times X_i$ , pour tout  $u_i \in U_i(g, x_i)$  et pour tout  $(v_i, \bar{v}_i) \in V_i(g, x_i) \times V'_{G_1 \times \dots \times G_{i-1}}(g)$

$$\begin{aligned} \tilde{f}_i((g, x_i), u_i, (v_i, \bar{v}_i)) &= \tilde{f}_i((g, x_i), u_i, (v_i, ((x_1, u_1), \dots, (x_{i-1}, u_{i-1})))) \\ &:= f_i((g, x_1, \dots, x_i), (u_1, \dots, u_i), v_i), \end{aligned} \quad (3.26)$$

où  $f_i$ ,  $U_i$ ,  $V_i$  et  $V_g$  sont définis à la Section 3.3.2. Pour tout  $(g, x_i) \in \mathbb{R}^q \times X_i$  et pour tout  $v'_i = (v_g, v_i, \bar{v}_i) \in V'_i(g, x_i)$ , on définit aussi la fonction  $\hat{f}_i$  par  $\hat{f}_i((g, x_i), u_i, v'_i) = (f_g(g, v_g), \tilde{f}_i((g, x_i), u_i, (v_i, \bar{v}_i))))$ .

Finalement, on définit  $G_i$  comme étant le noyau de viabilité garanti associé au système dynamique (3.25) et avec la contrainte  $C_i := \{(g, x_i) \in \mathbb{R}^q \times X_i, g \in L_g \text{ et } x_i \in L_i(g)\}$ , soit

$$G_i := \text{Guar}_{(\hat{f}_i, U_i, V'_i)}(C_i).$$

De plus,  $\hat{f}_i$  est continue dès que  $f$  l'est.  $\text{Graph}(U_i)$  est fermé et  $U_i$  a une croissance linéaire dès que  $U$  a ces deux propriétés.  $V_g$  et  $V_i$  sont continues dès que  $V$  l'est. Or,  $V'_i = V_g(g) \times V_i \times V'_{G_1, \dots, G_{i-1}}$ , donc  $V'_i$  est continue dès lors que  $g \mapsto P_j(g, G_j)$  est continue pour tout  $j$ . Lorsque que ces conditions sont vérifiées,  $G_i$  est fermé (voir Théorème 3.1.11).

### Résultat obtenu

On note  $G := \{(g, x_1, \dots, x_N) \in X \mid \forall i (g, x_i) \in G_i\}$ .

**Proposition 3.3.1.**  $G \subset \text{Guar}_{(f,U,V)}(L)$  (le noyau garanti associé au système (3.23) pour l'ensemble de contraintes (3.24)).

*Démonstration.* Remarquons déjà que par construction  $G \subset L$ . En effet, soit  $x = (g, x_1, \dots, x_N) \in G$ , pour  $i \in [1, N]$  on a  $(g, x_i) \in G_i$  et comme la contrainte associée à  $G_i$  est  $C_i := \{(g, x_i) \in \mathbb{R}^q \times X_i, g \in L_g \text{ et } x_i \in L_i(g)\}$ , on a donc  $g \in L_g$  et pour tout  $i$ ,  $x_i \in L_i(g)$ .

Montrons maintenant que  $G$  est un domaine de viabilité garanti.

Soit  $x = (g, x_1, \dots, x_N) \in G$ . Montrons qu'il existe  $u = (u_1, \dots, u_N) \in U(x)$  tel que  $\forall v = (v_g, v_1, \dots, v_N) \in V(x)$ ,  $f(x, u, v) \in G$ .

Pour tout  $i \in [1, N]$ ,  $(g, x_i) \in G_i$ , donc il existe  $u_i^* \in U(g, x_i)$  tel que  $\forall v'_i \in V'_i(g, x_i)$ ,  $\hat{f}_i((g, x_i), u_i^*, v'_i) \in G_i$ .

On pose  $u^* = (u_1^*, \dots, u_N^*)$ . Soit  $v = (v_g, v_1, \dots, v_N) \in V(x)$ . On pose  $y = f(x, u^*, v)$ , on veut montrer que  $y \in G$ .

On note  $y = (y_g, y_1, \dots, y_N)$ , on veut donc montrer que pour tout  $i \in [1, N]$ ,  $(y_g, y_i) \in G_i$ .

Or,  $y_g = f(g, v_g)$  et

$y_i = f_i((g, x_1, \dots, x_i), (u_1^*, \dots, u_N^*), v_i) = \tilde{f}_i((g, x_i), u_i^*, (v_i, ((x_1, u_1^*), \dots, (x_{i-1}, u_{i-1}^*)))$ .

Donc  $(y_g, y_i) = \hat{f}_i((g, x_i), u_i^*, v'_i)$  avec  $v'_i = (v_g, v_i, (x_1, u_1^*), \dots, (x_{i-1}, u_{i-1}^*))$ , donc  $v'_i \in V'_i(g, x_i)$ .

Ainsi  $(y_g, y_i) \in G_i$  du fait du choix de  $u_i^*$ , et ceci est vrai pour tout  $i \in [1, N]$ .

Donc  $y \in G$ .

□

**Bilan :** L'ensemble  $G$  ainsi construit est donc inclus dans le noyau de viabilité garanti du système (3.23) pour l'ensemble de contraintes (3.24). De plus  $G$  est construit à partir de  $N$  noyaux de viabilité garantis en petite dimension  $q + k_i$ , ce qui les rend calculables par un des algorithmes cités dans la Section 3.2.1.



## Chapitre

# 4

### Modèle de croissance locale

Ce chapitre vise à l'étude de l'influence de la fauche sur la croissance d'une tache de renouée asiatique, il contient les résultats que nous avons publié dans Lavallée et al. (2019b). Plus précisément, nous avons étudié l'influence de la taille initiale de la population, du nombre moyen de fauches par année et de la durée du projet de gestion sur l'aire moyenne et le nombre de couronnes des taches. Nous soulignons les apports de la modélisation pour l'aide à la gestion de la renouée du Japon dans Lavallée et al. (2019a). Cependant, à notre connaissance, les modèles existants n'ont pas été conçus pour étudier ces questions. Nous adoptons un formalisme stochastique qui nous permet d'étudier le stade initial de l'invasion lorsque l'approximation du champ moyen n'est pas valable. En effet, l'approximation en champ moyen nécessite la présence de nombreux individus et donc la présence de l'espèce depuis une longue période. La description des phénomènes au niveau individuel nous a permis de prendre en compte la variabilité entre les couronnes, par exemple due à leur âges différents. Néanmoins, le formalisme adopté pour l'écriture du modèle permet également une étude en grande population. Nous avons mené une telle étude pour le modèle sans fauche. Cet aspect la rend moins utile pour les gestionnaires, elle est donc présentée en Annexe A.

Ce chapitre est organisé comme suit. La Section 4.1 est consacrée à la description des mécanismes écologiques considérés (dominance apicale, compétition intraspécifique, etc.), à la présentation du modèle mathématique, ainsi qu'aux méthodes utilisées pour son étude. Les résultats sont présentés dans la Section 4.2. En particulier, nous avons montré l'existence et l'unicité du processus qui décrit le modèle. Nous avons aussi calibré la dynamique des plantes avec des données de terrain collectées dans les Alpes. Des simulations numériques ont été réalisées avec le logiciel OpenMOLE (Reuillon et al., 2013) pour étudier l'influence des paramètres de gestion sur la croissance de la population. Nous avons exprimé et quantifié la dépendance de l'aire et la taille moyenne d'une tache de renouée en fonction de la taille initiale de la population, du nombre moyen de fauches par an et de la durée du projet de gestion. Enfin, nous résumons nos résultats et discutons leurs implications et des améliorations possibles dans la Section 4.3.



## 4.1 Matériels et méthodes

Dans cette section, nous décrivons le modèle pour la dynamique de croissance d'une tache de renouées asiatiques incluant la fauche comme technique de gestion. Nous décrivons également les méthodes utilisées pour l'étudier avec des simulations numériques. Le modèle est écrit dans le formalisme des modèles individu-centrés (voir Section 2.1.1). Dans ce chapitre, le terme "individu" fera référence aux couronnes. Nous rappelons qu'une couronne est l'emplacement d'un bourgeon d'où émergent les tiges. Les individus sont caractérisés par leur position  $x$  dans le plan et leur biomasse souterraine  $a$  (c'est-à-dire celle du rhizome qui est relié à la couronne, appelée simplement biomasse dans la suite). Les notations suivantes seront utilisées pour décrire le modèle :

- $\chi := \mathbb{R}^2 \times \mathbb{R}_+$ , est l'espace d'état des positions et des biomasses. Dans le modèle, une couronne est représentée par une masse de Dirac  $\delta_{(x,a)}$ , avec  $(x,a) \in \chi$ , où  $x$  représente la position de la couronne et  $a$  est la biomasse associée à la couronne.
- L'ensemble des couronnes présentes au temps  $t$  est décrit par la mesure  $Z_t \in \mathcal{M}(\chi)$ , où  $\mathcal{M}(\chi)$  est l'ensemble des mesures ponctuelles finies sur  $\chi$  dont les masses de points sont 0 ou 1.

$$\mathcal{M}(\chi) := \left\{ \sum_{i=1}^n \delta_{(x_i, a_i)}, n \geq 0, (x_1, a_1), \dots, (x_n, a_n) \in \chi \right\}.$$

L'utilisation de  $\mathcal{M}(\chi)$  nous permet de ne pas fixer *a priori* le nombre d'individus dans le modèle, puisqu'il contient toutes les tailles de population possibles.

### 4.1.1 Description des phénomènes modélisés

#### Naissance

Un individu de trait  $(x, a) \in \chi$  (c'est-à-dire qui se situe en position  $x$  et qui a une biomasse  $a$ ) donne naissance à un taux  $b(x, Z)$ , où  $Z \in \mathcal{M}(\chi)$  décrit l'état de la population (c'est-à-dire les positions et les biomasses de tous les individus). Nous supposons que ce taux de naissance ne dépend pas de la biomasse  $a$  d'un individu. Une couronne peut donner naissance à deux couronnes filles au maximum (Smith et al., 2007) à un taux dépendant des individus voisins présents à une distance inférieure au seuil noté *distanceParent*. Ainsi, si une couronne a déjà donné naissance à deux filles (en fait, s'il y a déjà trois individus à une distance inférieure à *distanceParent* de celle-ci puisque l'on compte son parent), elle a un taux de naissance nul (elle ne donne pas naissance). Nous verrons dans la section suivante qu'une couronne fille peut en principe se trouver à une distance supérieure à *distanceParent* de son parent. Ce phénomène peut être compensé par le fait que les couronnes d'un autre parent peuvent se trouver à une distance inférieure à *distanceParent*.

Cette modélisation permet de rendre compte des effets de la dominance apicale : si une couronne meurt, la dominance apicale qu'elle exerce sur les bourgeons latéraux voisins cesse, et ceux-ci peuvent se développer pour former des tiges et créer une nouvelle couronne. Bashtanova et al. (2009), Adachi et al. (1996a) et Dauer and Jongejans (2013) mentionnent ce phénomène de dominance apicale de manière générale, mais ils ne le quantifient pas. C'est pourquoi nous utiliserons une méthode de calibration pour fixer sa valeur (en fait, nous utilisons cette méthode pour tous les paramètres). Le taux auquel un individu en position  $x$  donne naissance peut donc être exprimé comme suit

$$b(x, Z) = \bar{b}_{\{\sum_{y \in V(Z)} 1_{\{\|x-y\| \leq \text{distanceParent}\}} \leq 3\}}, \quad (4.1)$$

où  $\|\cdot\|$  désigne la norme euclidienne usuelle sur  $\mathbb{R}^2$ ,  $\bar{b}$  est le taux de naissance maximal (dans des conditions idéales) et

$$V(Z) := \{x \in \mathbb{R}^2, Z(\{x\} \times \mathbb{R}_+) > 0\} \quad (4.2)$$

est l'ensemble des positions des couronnes présentes dans la population  $Z$ .

### Dispersion de l'individu créé

Un individu de trait  $(x, a)$  qui donne naissance génère un individu à la position  $x'$ . Nous choisissons ici une loi gamma pour la distribution de la longueur de dispersion. Sa densité sur  $\mathbb{R}^+$  est donnée par

$$x \mapsto 1/(\text{scale}^{\text{shape}} \Gamma(\text{shape})) x^{(\text{shape}-1)} e^{-(x/\text{scale})}. \quad (4.3)$$

Les paramètres  $(\text{shape}, \text{scale})$  de la distribution seront soumis à la calibration. Toutefois, Smith et al. (2007) indiquent l'ordre de grandeur des longueurs correspondant aux longueurs de dispersion (histogramme de la Figure 4.1), ainsi que l'ordre de grandeur de la longueur d'extension du rhizome par année (0.174m)

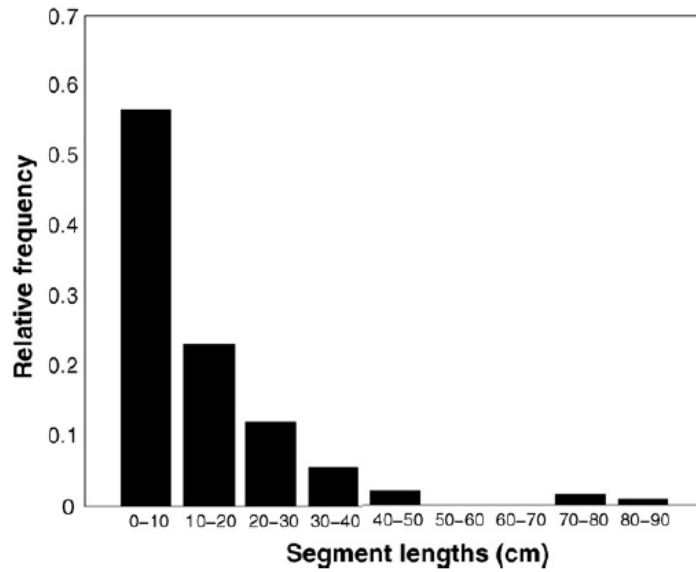


FIGURE 4.1 – Histogramme des longueurs d'extension du rhizome, extrait de Smith et al. (2007).

Nous supposons une distribution uniforme sur l'angle de la direction de  $x' - x$  par rapport à l'axe des abscisses, et on note  $D$  la loi sur  $\mathbb{R}^2$  pour la position du nouvel individu (relativement à celle de son parent). De plus, nous modélisons le phénomène de compétition intraspécifique en considérant qu'un individu ne naît réellement que s'il n'est pas trop proche d'une couronne déjà existante. Nous introduisons donc l'ensemble  $C$ , fonction de l'état de la population  $Z$  et de la position du parent potentiel  $x$  :

$$C_{x,Z} := \{z \in \mathbb{R}^2, \forall y \in V(Z) \setminus \{x\}, \|y - z\| > \text{distanceCompetition}\}. \quad (4.4)$$

L'individu créé doit donc se situer à une distance supérieure à  $distance_{Competition}$  de ses voisins pour ne pas tomber dans la zone de compétition intraspécifique. Le principe des zones exclues pour la naissance d'un individu est aussi utilisé dans Smith et al. (2007) : les zones entourant les couronnes sont soumises à la compétition pour la lumière et on n'y autorise pas l'apparition de nouvelles couronnes. Le schéma de la Figure 4.2 montre les distances jouant un rôle lors d'un événement de naissance.

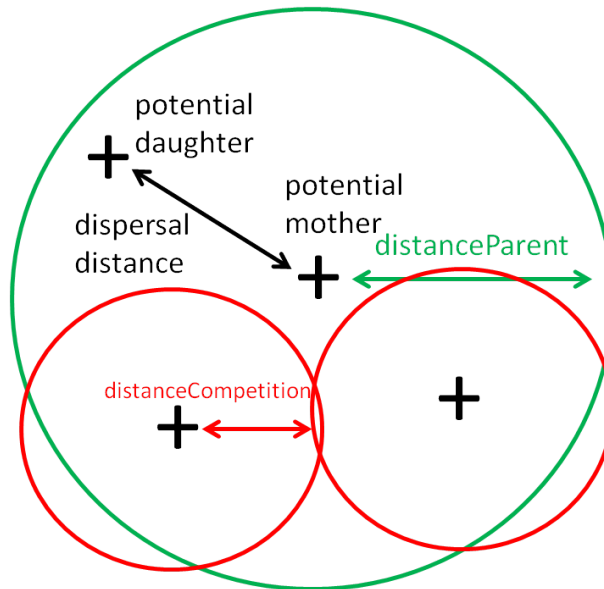


FIGURE 4.2 – Schéma représentant un événement de naissance. Une croix représente la position d'une couronne, extrait de Lavallée et al. (2019b). Nous indiquons également les différentes distances utilisées dans le modèle.

## Évolution de la biomasse

Dans Seiger et al. (1997), les auteurs présentent les effets de la fauche sur la croissance du rhizome. Ils constatent que la biomasse du rhizome croît significativement tout au long de la saison, contrairement à la biomasse aérienne dont la croissance n'est plus significative à la fin de la saison. Si on ne coupe pas les parties aériennes, on suppose une croissance de la biomasse souterraine (trait  $a$ ) selon une loi de Von Bertalanffy, explicitée ci-après.

Les effets de la fauche de parties aériennes sur le développement du rhizome ne sont pas bien connus. On suppose ici que la fauche entraîne une diminution de la biomasse souterraine. Cette hypothèse vient du fait que les ressources du rhizome sont utilisées pour la régénération des parties aériennes, ce qui est observé dans Gerber et al. (2010). Dans Rouifed et al. (2011a), les auteurs remarquent également que la fauche impacte la quantité de biomasse souterraine à la fin de la saison, et que la fauche induit une densité de rhizome décroissante avec la profondeur (alors que sans fauche elle est constante). Nous ne prenons cependant pas ce phénomène en compte, puisque le modèle est situé dans le plan.

Les fauches ont lieu au taux  $1/\tau$  (il y a donc en moyenne  $\tau$  fauches par an), et on fauche une proportion  $proportionMowing$  constante de la tache. Seiger et al. (1997) ont montré que la répartition des instants de fauches dans la saison n'a pas d'influence sur la quantité de biomasse souterraine à la fin de la saison (seul le nombre de fauches est significatif). Après une fauche, on suppose que la biomasse souterraine  $a$  est directement impactée et devient  $a.F(a)$ , où  $F$  prend ses valeurs dans  $[0, 1]$  et décrit l'effet de la fauche en fonction de la biomasse

de l'individu. Afin de rendre compte de la fragilité des jeunes couronnes, on suppose que la fonction  $F$  décrivant l'impact de la fauche sur la biomasse a des valeurs plus élevées pour les biomasses faibles (nous prenons donc  $F$  croissante, ce qui implique que pour deux biomasses  $a_1 < a_2$ , nous aurons  $a_1 * F(a_1) < a_2 * F(a_2)$ ).

Nous supposons que  $F$  s'écrit de la manière suivante :

$$\forall a \in \mathbb{R}_+, F(a) = 1 - \exp(-mowingParameter * a), \quad (4.5)$$

où  $mowingParameter$  est un paramètre qui agit sur le taux de décroissance de la fonction exponentielle.

Pour la croissance de la biomasse d'une couronne lorsqu'il n'y a pas de fauche, nous utilisons l'équation de Von Bertalanffy (4.6). Introduite dans von Bertalanffy (1934), cette équation a été beaucoup utilisée en foresterie (Zeide, 1993). Ici, elle décrit la dynamique de la biomasse en fonction du temps. Elle est basée sur des arguments physiologiques simples : le taux de croissance de l'organisme diminue avec la biomasse.

$$\frac{da(t)}{dt} = L(K - a(t)) =: v(a(t)), \quad (4.6)$$

où  $L$  est le taux de croissance à faible biomasse et  $K$  est la biomasse asymptotique maximale. On sait résoudre analytiquement l'Équation (4.6). Si on suppose que la biomasse à l'instant  $t_0$  vaut  $a_0$ , alors  $\forall t \geq t_0$ , on a :

$$a(t) = a_0 e^{-L(t-t_0)} + K(1 - e^{-L(t-t_0)}). \quad (4.7)$$

De plus, nous supposons que toutes les couronnes naissent avec la même biomasse  $a_0 \in \mathbb{R}_+$ .

## Mortalité

On suppose que le taux de mortalité est indépendant de la position  $x$  d'un individu. Un individu en vie au temps  $t$  et avec une biomasse  $a(t)$  meurt au taux  $m(a(t))$ . On suppose ici que la mortalité  $m$  est une fonction décroissante de la biomasse : un individu avec une faible biomasse, soit parce qu'il vient d'être créé, soit parce qu'il a subi des fauches, a un taux de mortalité plus important. Donc si  $T_0$  est la date à laquelle meurt un individu né à l'instant 0 et dont la biomasse jusqu'à l'instant  $t$  est donnée par la fonction  $a$ , on obtient :

$$\mathbb{P}(T_0 \geq t) = e^{-\int_0^t m(a(s)) ds}. \quad (4.8)$$

Dans Smith et al. (2007), les auteurs fixent la probabilité qu'un segment de rhizome meure sur une période de 4 mois (un pas de temps dans leur modèle) à 0,0083. Lorsqu'il n'y a pas de fauche, les événements de mortalité pour les couronnes se produisent rarement dans la nature, ce qui explique pourquoi la valeur proposée dans Smith et al. (2007) est faible. Pour obtenir une bonne estimation de cette valeur, il faut un nombre suffisamment important d'observations ; la calibration est donc très utile pour estimer une valeur pour ce type de paramètre. Nous supposons que  $m$ , la fonction qui décrit le taux de mortalité d'une couronne en fonction de sa biomasse, est exprimée par :

$$m(a) = deathParameterScaling e^{-deathParameterDecrease*a}. \quad (4.9)$$

L'Équation (4.9) fait intervenir deux paramètres :  $deathParameterDecrease$ , qui influe sur le taux de décroissance de la fonction et  $deathParameterScaling$  qui permet de choisir le taux de mortalité des individus ayant une faible biomasse.

## Fauche

Nous considérons deux possibilités pour le choix des couronnes à faucher. La première consiste à choisir les couronnes fauchées de manière aléatoire et uniforme (appelée ensuite technique de gestion aléatoire). La proportion peut ainsi représenter différentes qualités de fauche, si l'objectif est de faucher l'ensemble de la tache, en fonction des outils utilisés (à la main, débroussailleuse). Nous considérons cette technique lorsque nous écrivons le formalisme mathématique du modèle dans la Section 4.1.3 et pour l'analyse de l'influence des paramètres du modèle dans la Section 4.2.3. La seconde technique consiste à faucher un côté de la tache : elle consiste à déterminer une abscisse à droite de laquelle chaque couronne est fauchée, et à gauche de laquelle aucune couronne n'est fauchée. Elle reflète par exemple le cas d'une tache située sur deux parcelles, appartenant à des propriétaires différents, l'un gérant la tache tandis que l'autre ne le fait pas. Cette situation se produit fréquemment sur le bord des routes. Nous utilisons cette technique de gestion dans le modèle pour la calibration (Section 4.1.5) en raison des caractéristiques des données.

### 4.1.2 Synthèse des paramètres du modèle

Les paramètres du modèle sont de deux types : les paramètres de gestion et les paramètres pour la dynamique de la plante.

Le Tableau 4.1 synthétise les paramètres dynamiques de la plante. Nous précisons les unités des paramètres pour ceux qui ont une signification biologique.

| Paramètre                | Description   | Unité     |
|--------------------------|---|-----------|
| <b>Biomasse</b>          |   |           |
| $K$                      | biomasse maximale (FegrefsolVonBert)  | g         |
| $L$                      | taux de croissance de la biomasse à faible biomasse (Équation (4.7))          | $an^{-1}$ |
| $a_0$                    | biomasse initiale (d'une couronne à la naissance)                             | g         |
| <b>Fauche</b>            |   |           |
| $mowingParameter$        | dans la fonction pour l'effet de la fauche (Équation (4.5))                   | $g^{-1}$  |
| <b>Mortalité</b>         |   |           |
| $deathParameterScaling$  | taux de mortalité à faible biomasse (Équation (4.9))                          | $an^{-1}$ |
| $deathParameterDecrease$ | taux de décroissance dans la fonction pour la mortalité dans (Équation (4.9)) | $g^{-1}$  |
| <b>Naissance</b>         |   |           |
| $distanceParent$         | distance de dominance apicale (Équation (4.1))                                | $m$       |
| $distanceCompetition$    | distance de compétition intra-spécifique (Équation (4.4))                     | $m$       |
| $\bar{b}$                | taux de naissance (conditions idéales)  | $an^{-1}$ |
| (shape, scale)           | loi Gamma, dispersion du nouvel individu                                      |           |

TABLE 4.1 – Synthèse des paramètres pour la dynamique de la plante dans le modèle

Les paramètres de gestion sont les suivants :

- Nombre moyen de fauches par année :  $\tau$  ;
- Durée du projet de gestion :  $T$  ;
- Proportion de couronnes fauchées :  $proportionMowing$ .

Le paramètre pour la taille de la population initiale  $initialPopSize$  est aussi considéré.

### 4.1.3 Formalisme mathématique associé au modèle

Le formalisme utilisé est celui des processus stochastiques à valeurs mesures. Les modèles individus centrés avec une taille de population variable et de la compétition entre individus ont été introduits par Bolker and Pacala (1997), et par Dieckmann et al. (2000). Une étude probabiliste rigoureuse a ensuite été menée par Fournier and Méléard (2004). Par la suite, ces modèles ont été également étudiés dans Champagnat (2006), Champagnat et al. (2006),

Costa et al. (2016) et Coron et al. (2018). Le modèle que je vais présenter dans cette partie est inspiré des travaux de Sylvie Méléard et de Viet Chi Tran. En particulier, je me suis inspiré des notations et des techniques présentes dans Tran (2006, 2008) et Fournier and Méléard (2004). Nous rappelons que les individus, ici les couronnes sont caractérisés par un trait spatial (une position dans le plan) et un trait décrivant la biomasse souterraine (i.e. celle du rhizome qui est relié à la couronne).

L'équation différentielle stochastique (4.10)

$$\begin{aligned}
Z_t = & \sum_{i=1}^{N_0} \delta_{(X_i(Z_0), A_b(t, 0, A_i(Z_0)))} \\
& + \int_{[0, t] \times \mathbb{N}^* \times \mathbb{R}_+ \times \mathbb{R}^2} 1_{\{i \leq N_{s-}\}} \delta_{(X_i(Z_s) + z, A_b(t, s, a_0))} 1_{\{\theta \leq b(X_i(Z_s), Z_s)\}} \\
& \quad 1_{\{X_i(Z_s) + z \in C_{X_i(Z_s), Z_s}\}} M_1(ds, di, d\theta, dz) \\
& + \int_{[0, t] \times [0, 1]^{N^*}} \sum_{i=1}^{N_{s-}} 1_{\{y_i \leq \text{proportion Mowing}\}} (\delta_{(X_i(Z_s), A_b(t, s, A_i(Z_{s-})) * F(A_i(Z_{s-})))} \\
& \quad - \delta_{(X_i(Z_s), A_b(t, s, A_i(Z_s)))}) M_2(ds, dy) \\
& - \int_{[0, t] \times \mathbb{N}^* \times \mathbb{R}_+} 1_{\{i \leq N_{s-}\}} 1_{\{\theta \leq m(A_i(Z_s))\}} \delta_{(X_i(Z_s), A_b(t, s, A_i(Z_s)))} M_3(ds, di, d\theta)
\end{aligned} \tag{4.10}$$

décrit la dynamique de la tache (population de couronnes). Elle est régie par trois Mesures Ponctuelles de Poisson<sup>1</sup> indépendantes, définies comme suit :

- $M_1(ds, di, d\theta, dz)$  est une Mesure Ponctuelle de Poisson sur  $\mathbb{R}_+ \times \mathbb{N}^* \times \mathbb{R}_+ \times \mathbb{R}^2$  d'intensité  $ds \otimes n(di) \otimes d\theta \otimes D(dz)$ , où  $n(di)$  désigne la mesure de comptage sur  $\mathbb{N}^*$  et  $D$  est la densité de la loi pour la dispersion d'un individu. La mesure  $M_1$  décrit les événements de naissance.
- $M_2(ds, dy)$  est une Mesure Ponctuelle de Poisson sur  $\mathbb{R}_+ \times [0, 1]^{N^*}$  d'intensité  $1/\tau ds \otimes \mathcal{U}^{N^*}([0, 1])$ , où  $\mathcal{U}([0, 1])$  est la loi uniforme sur  $[0, 1]$ . On note  $y = (y_1, y_2, \dots)$  pour  $y \in [0, 1]^{N^*}$ . La mesure  $M_2$  décrit les événements de fauche.
- $M_3(ds, di, d\theta)$  est une Mesure Ponctuelle de Poisson sur  $\mathbb{R}_+ \times \mathbb{N}^* \times \mathbb{R}_+$  d'intensité  $ds \otimes n(di) \otimes d\theta$ , où  $n(di)$  désigne la mesure de comptage sur  $\mathbb{N}^*$ . La mesure  $M_3$  décrit les événements de mortalité.

Dans l'Équation (4.10),

$$Z_t = \sum_{i=1}^{N_t} \delta_{(X_i(Z_t), A_i(Z_t))}$$

est donc la mesure qui décrit la population au temps  $t \geq 0$ ,  $A_b$  est le flot de l'équation différentielle décrivant la dynamique de la biomasse d'une couronne (Équation (4.7)).  $X_i$  (resp.  $A_i$ ) indique la position (resp. la biomasse) du  $i$ -ième individu dans la population (selon l'ordre lexicographique).

On note  $b : (x, Z) \in \mathbb{R}^2 \times \mathcal{M}(\chi) \mapsto b(x, Z)$  et  $m : a \in \mathbb{R}_+ \mapsto m(a)$  respectivement les taux individuels de naissance et de mort. L'application

$$C : Z \in \mathcal{M}(\chi) \mapsto C_{X_i(Z), Z} \in \mathcal{P}(\mathbb{R}^2),$$

où  $\mathcal{P}(\mathbb{R}^2)$  est l'ensemble des parties de  $\mathbb{R}^2$ , donne la région admissible pour les naissances de nouveaux individus, qui est liée à la compétition intraspécifique. La fonction  $F : [0, K] \rightarrow [0, 1]$

1. La définition et des propriétés des Mesures Ponctuelles de Poisson sont rappelées en Annexe A

modélise l'effet de la fauche des couronnes et  $\tau$  est le nombre moyen d'événements de fauche par an.

Le premier terme de l'Équation (4.10) correspond à la croissance de la population initiale : la biomasse des individus augmente selon le flot  $A_b$ . Le second terme se réfère aux naissances. Un événement de naissance consiste à choisir un parent potentiel et à vérifier s'il remplit les conditions pour donner naissance : c'est le rôle des fonctions indicatrices. Si un tel événement se produit, on ajoute une masse de Dirac correspondant à un nouvel individu créé dans la population avec une biomasse initiale  $a_0$ . Le terme intégral suivant correspond à l'événement de fauche, après lequel les individus fauchés sont remplacés par un nouvel individu de même position et de biomasse réduite. Le dernier terme représente les événements de mortalité, pour lesquels nous retirons un individu de la population en soustrayant la masse de Dirac correspondante.

#### 4.1.4 Simulation du modèle

##### Algorithme utilisé pour simuler la dynamique des taches

Nous présentons une étape de l'algorithme de Gillespie utilisé pour simuler la dynamique des taches.

Trois types d'événements peuvent se produire : une naissance, une mort ou la fauche d'une proportion *proportionMowing* d'individus dans la population. Supposons que nous ayons  $N$  individus au temps  $t$ . Nous commençons par calculer le temps de l'événement suivant, ce qui nécessite la somme des taux de naissance, de mortalité et de fauche. La loi de ce temps ne dépend que de l'état actuel de la population puisque le processus est Markovien. S'il s'agit d'un événement de naissance, nous sélectionnons le parent de manière uniforme dans la population. Nous vérifions que l'individu sélectionné pour être le parent n'a pas déjà trop de voisins à une distance inférieure à *distanceParent*. Si c'est bien le cas, nous déterminons la position du nouvel individu (enfant) conformément à la loi décrite dans la Section 4.1.1. Si la position de l'enfant tombe dans l'ensemble  $C$  de l'Équation (4.4) (c'est-à-dire qu'elle ne tombe pas dans une zone de compétition intraspécifique), alors l'individu naît à cette position, et la population devient de taille  $N + 1$ . Dans le cas contraire, c'est-à-dire si le parent a déjà suffisamment de couronnes à proximité ou si le nouvel individu à naître est en dehors de l'ensemble  $C$ , il ne se produit pas de naissance et la taille de la population reste  $N$ . S'il s'agit d'un événement de fauche, alors chaque individu est fauché avec une probabilité *proportionMowing* : sa biomasse  $a$  est remplacée par  $a * F(a)$ . Enfin, si l'événement est un événement de mortalité, alors l'individu susceptible de mourir est choisi uniformément dans la population puis la réalisation d'une variable aléatoire détermine si cet individu meurt effectivement en fonction de son taux de mortalité, qui dépend de sa biomasse. S'il meurt, il est retiré de la population dont la nouvelle taille est  $N - 1$  ; sinon, il ne se passe rien. Entre deux événements, la biomasse de chaque individu croît de manière déterministe. L'algorithme s'arrête dès qu'il n'y a plus d'individus dans la population.

Pour les simulations, nous devons définir une population initiale (au temps 0, positions et biomasses) qui va évoluer. Pour créer une population initiale, nous utilisons l'algorithme ci-dessus avec initialement un individu jusqu'à ce que la population atteigne une taille prescrite, avec une stratégie de gestion faible ( $\tau = 1$  et *proportionMowing* = 0,9). En effet, si nous mettons une valeur plus élevée de  $\tau$ , nous ne pouvons pas souvent créer la population initiale car le premier individu meurt avant de produire un enfant. C'est pourquoi nous autorisons

également plusieurs tentatives dans l'algorithme pour créer la population initiale.

### Code utilisé pour la simulation et sa mise à disposition

Pour simuler la dynamique de la tache dans le modèle, nous utilisons le langage de programmation Scala (version 2.11.12). Nous utilisons le logiciel OpenMOLE (Reuillon et al. (2013), version 8.0) pour l'exploration du modèle. Enfin, nous utilisons le logiciel R (version 3.4.4) pour l'analyse statistique des sorties du modèle. Les simulations sont réalisées sur la grille de calcul européenne (EGI, <http://www.egi.eu/>). Le code scala pour la simulation du modèle est disponible sur GitHub (<https://github.com/francoismaths/renouee/>). Un billet de blog a également été rédigé (<https://blog.openmole.org/user-case-management-of-japanese-knotweed/>). Il s'adresse à des lecteurs déjà familiers des concepts d'OpenMOLE (workflow, task, hook, on peut aussi se référer à la documentation d'OpenMOLE) et permet, dans le cadre de l'étude présentée ici, de suivre en détail le déroulement d'une expérience numérique complète avec OpenMOLE.

### Exemple d'évolution de la dynamique

La Figure 4.3 montre un exemple d'évolution de la taille de la population (nombre de couronnes) pour une trajectoire du modèle, pour des paramètres de la dynamique de la plante et de gestion fixés. La taille de population initiale est fixée à 1000, le nombre moyen de fauches par an à  $\tau = 3$ , la durée de projet de gestion à  $T = 4$  ans et la proportion de couronnes fauchées est *proportionMowing* = 0.9. On a ainsi un nombre moyen d'événements de fauche égal à  $3 \times 4 = 12$  (il y en a eu 11 dans la simulation). La taille de population finale est de 619, cette stratégie de gestion a donc conduit à une réduction d'environ un tiers de la taille de la population.

#### 4.1.5 Données pour la calibration

Notre objectif est de comprendre l'influence des paramètres de gestion sur la dynamique d'une tache. Pour cela, nous devons fixer les paramètres de la dynamique de la plante. Pour certains de ces paramètres, nous ne pouvons pas trouver de valeur dans la littérature. Nous procédons donc à une calibration pour attribuer une valeur à ces paramètres pour lesquelles le modèle reproduit le mieux les données de terrain.

Les données de terrain sont celles utilisées par Martin et al. (2018). Les auteurs étudient le potentiel d'invasion des renouées asiatiques le long d'un gradient d'altitude (i.e. en montagne) en identifiant les déterminants de sa dynamique spatiale. L'expérience consiste en la collecte de données sur 19 taches de renouée asiatique à différentes altitudes dans les Alpes. Les mesures ont été effectuées en 2008 et 2015, sur les taches elles-mêmes (contour, densité de tiges), mais aussi sur des variables biotiques et abiotiques. Les taches étaient fauchées ou pas, et pour chaque tache, nous avons accès à la stratégie de gestion qui lui a été appliquée : la fréquence de fauche et une estimation de la proportion de la tache fauchée, lorsque la tache n'était fauchée que sur une partie seulement. Il existe une grande variabilité dans les taches observées, tant sur leur taille (de moins de 2  $m^2$  à 350  $m^2$ ) que sur les conditions dans lesquelles elles poussent telles que la qualité du sol, la proximité d'une rivière, d'une route ou d'un terrain en friche. A partir des sorties du modèle, nous calculons les tailles et aires des populations finales et initiales (les aires sont les aires de l'enveloppe convexe formée par l'ensemble des couronnes). Nous utilisons les données sur les aires et la densité de couronnes dans les taches pour déduire la taille de population, pour les taches de 2008 et 2015.



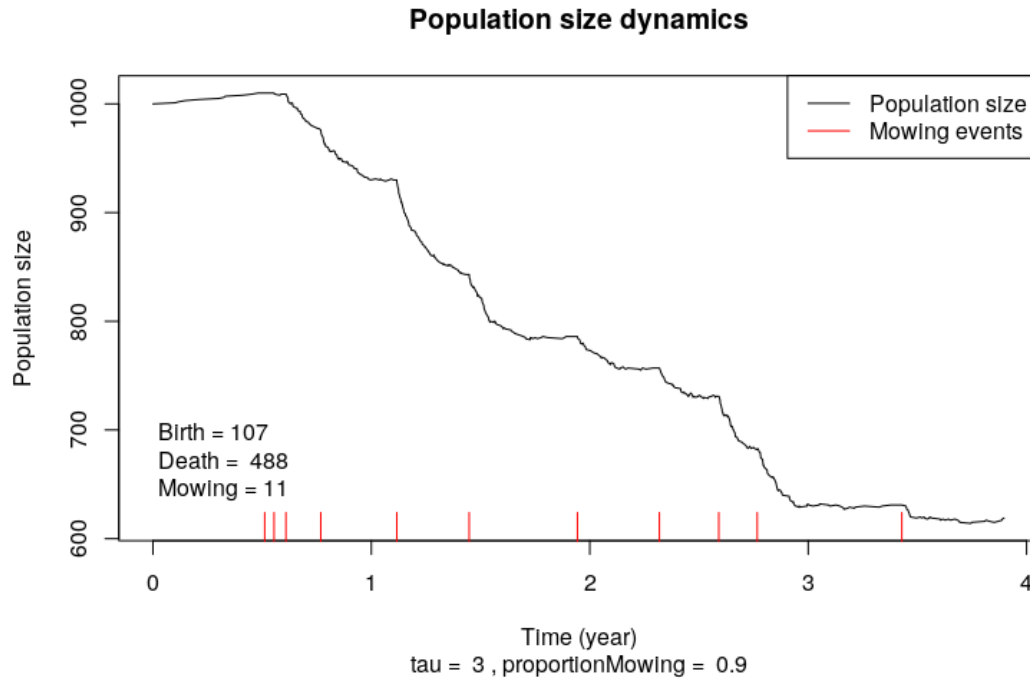


FIGURE 4.3 – Exemple illustratif : simulation d’une trajectoire du modèle (Équation (4.10)) avec  $\tau = 3$ ,  $initialPopSize = 1000$ ,  $T = 4$ ,  $proportionMowing = 0.9$  et pour les valeurs des paramètres de la dynamique de la plante celles indiquées dans le Tableau 4.2, extrait de Lavallée et al. (2019b). La ligne noire montre la taille de population, les traits rouges indiquent les instants de fauche.

Nous rappelons (voir Section 4.1.4) que l’opération pour simuler une tache se déroule en deux étapes : d’abord la création de la population initiale étant donnée la taille de population à atteindre (nous choisissons la taille de la tache en 2008), puis son évolution avec les paramètres de gestion issus des données (Martin et al. (2018)).

#### 4.1.6 Méthode pour la calibration

Pour chaque jeu de paramètres à tester, nous simulons les 19 taches : on obtient l’aire et la taille de populations initiales et finales. Nous pouvons donc comparer ces valeurs simulées aux  $19 \times 4$  valeurs mesurées dans Martin et al. (2018). Notre objectif est de trouver un jeu de paramètres de la plante commun à toutes les taches, qui fait le mieux correspondre les sorties du modèle (aire et taille) aux données de terrain. Cette procédure compare les tailles initiales des observations (2008) aux tailles initiales simulées. Cette comparaison est utilisée pour vérifier que le jeu de paramètre testé rend possible l’obtention de la taille de population souhaitée. En effet, certains jeux de paramètres peuvent conduire à un échec dans la création d’une population initiale (par exemple si la distance de compétition est trop importante, alors que la distance de dispersion est trop faible).

Nous devons donc choisir une distance pour comparer les résultats de simulations aux observations. Pour chaque tache et chaque sortie (aire et taille), nous prenons la distance :  $dist(simu, data) = |simu - data|/data$ . Nous optons pour une erreur relative (renormalisation par la donnée  $data$ ) car les tailles de population et les aires n’ont pas les mêmes ordres de grandeur, et puisqu’il y a également une différence importante dans les valeurs d’aires et tailles respectivement.

La distance globale à minimiser est la somme de ces distances sur les 19 taches et sur les quatre observations (taille et aire, en 2008 et 2015). Remarquons que si le jeu de paramètres testé ne permet pas la création d'une population initiale comme souhaitée, une population de taille nulle et d'aire nulle pour l'instant initial (2008) sont obtenues, les valeurs simulées pour l'année 2015 sont aussi nulles. Ainsi, la distance entre les données de terrain et une population simulée triviale (nulle) vaut  $76 = (19 \times 4)$ .

Pour minimiser la distance globale, nous utilisons la méthode NSGA2 du logiciel OpenMOLE, qui repose sur un algorithme génétique (voir Section 2.2.1). Le résultat obtenu est présenté dans la Section 4.2.2. L'algorithme de calibration est un algorithme itératif, qui donne un ensemble de solutions à chaque étape. La distance  $dist(simu, data)$  entre les données et le résultat de la simulation décroît lorsque l'algorithme progresse.

### 4.1.7 Analyse numérique

Les simulations sont effectuées avec le jeu de paramètres obtenu par calibration (Section 4.2.2, Table 4.2). Nous allons préciser la manière dont nous étudions l'influence des paramètres de gestion  $\tau$  et  $T$ , ainsi que la taille de population initiale. Nous rappelons que :

- $\tau$  désigne le nombre moyen de fauches par année,
- $T$  désigne la durée du projet de gestion.

Nous nous intéressons à l'influence de ces trois paramètres, mais nous n'étudions pas l'influence du paramètre *proportionMowing*. Sa valeur est fixée à 0.9 et nous utilisons la technique de fauche aléatoire. En effet, nous considérons que le gestionnaire souhaite faucher l'entièreté de la tache, mais nous ne fixons pas la valeur de *proportionMowing* à 1 pour considérer une fauche imparfaite due à l'outil utilisé (comme expliqué dans la Section 4.1.1).

Pour l'étude du modèle, nous utilisons les outils et techniques introduits dans la Section 2.3. Les échantillonnages des paramètres concernés sont effectués dans OpenMOLE, avec une réplication de taille  $n = 50$  pour chaque jeu de paramètres de gestion (ces échantillonnages sont détaillés dans la Section 4.2.3) et les sorties moyennes sont calculées sur ces  $n$  valeurs. Nous faisons d'abord varier un paramètre (méthode OAT, voir Section 2.3). Après un premier examen visuel des résultats de la simulation, nous ajustons trois relations avec la méthode des moindres carrés : une régression linéaire effectuée avec la fonction *lm* du logiciel *R*, une régression quadratique tronquée, et une régression exponentielle effectuée avec la fonction *nls* du logiciel *R*. Nous évaluons la performance du modèle en utilisant le coefficient de détermination ( $R^2$ ) et la racine carrée de la moyenne des erreurs (root mean squared errors, *RMSE*). Enfin, dans la Section 4.2.3, nous utilisons les mêmes outils statistiques (*lm* et *nls*) pour trouver des formules de régression générales pour les quantités moyennes de sortie en fonction des paramètres de gestion, et deux constantes que l'algorithme vise à trouver (méthode du méta-modèle, voir Section 2.3).

## 4.2 Résultats

### 4.2.1 Existence et Unicité de la solution de (4.10)

Dans cette section nous présentons un résultat d'existence et d'unicité pour le processus solution de l'équation différentielle stochastique (4.10).

Les techniques utilisées pour montrer les résultats de cette section sont adaptées de Tran (2006).

**Notations :**

Pour deux espaces métriques  $E$  et  $F$ ,  $\mathcal{B}_b(E, F)$  désigne l'ensemble des fonctions boréliennes bornées de  $E$  dans  $F$ . Lorsque  $F = \mathbb{R}$ , on note  $\mathcal{B}_b(E)$ .  $\mathcal{C}^{0,1}(E \times \mathbb{R})$  est l'ensemble des fonctions continues, bornées, avec des dérivées partielles bornées par rapport à la variable dans  $\mathbb{R}$ .

Pour  $f \in \mathcal{C}^{0,1}(E \times \mathbb{R})$ , le gradient de  $f : (x, a) \in E \times \mathbb{R} \rightarrow f(x, a) \in \mathbb{R}$  par rapport à la variable  $a \in \mathbb{R}_+$  est noté  $\nabla_a f$ . A chaque fonction  $f \in \mathcal{B}_b(E)$  et  $F \in \mathcal{B}(\mathbb{R})$ , on peut associer une fonction  $F_f \in \mathcal{B}(\mathcal{M})$ , définie par

$$\forall m \in \mathcal{M}, F_f(m) = F(\langle m, f \rangle).$$

On note

$$\begin{aligned} E_1 &= \mathbb{N}^* \times \mathbb{R}_+ \times \mathbb{R}^2 \\ E_2 &= [0; 1]^{N^*} \\ E_3 &= \mathbb{N}^* \times \mathbb{R}_+. \end{aligned} \tag{4.11}$$

**Hypothèse 4.2.1.** On a choisi une fonction pour le taux de naissance telle que :

$$\forall (x, Z) \in \mathbb{R}^2 \times \mathcal{M}, b(x, Z) \leq 1.$$

On note plus généralement  $\bar{b}$  une constante telle que

$$\forall (x, Z) \in \mathbb{R}^2 \times \mathcal{M}, b(x, Z) \leq \bar{b}.$$

On suppose une majoration analogue pour le taux de mortalité : il existe une constante  $\bar{m}$  telle que

$$\forall a \in \mathbb{R}^+, m(a) \leq \bar{m}.$$

Enfin, on généralise la définition de  $v$  donnée à l'Équation (4.6) pour que  $v$  vérifie

$$\frac{da}{dt} = v(a(t)), \tag{4.12}$$

$v$  décrit donc l'évolution de la biomasse, on suppose que cette fonction vérifie l'hypothèse :

$$\exists \bar{v} \forall a \in \mathbb{R}_+, v(a) \leq \bar{v}(1 + a).$$

**Proposition 4.2.2.** Sous l'Hypothèse 4.2.1, l'équation différentielle stochastique (4.10) admet une solution forte  $(Z_t)_{t \in \mathbb{R}_+} \in \mathbb{D}(\mathbb{R}_+, \mathcal{M})$  telle que pour tout  $T > 0$ ,  $N_t := \langle Z_t, 1 \rangle =$

$\int_{\mathbb{R}^2 \times \mathbb{R}} Z_t(dx, da)$  satisfait :

$$\mathbb{E}(\sup_{t \in [0, T]} N_t) < \mathbb{E}(N_0)e^{\bar{b}T} < \infty.$$

*Démonstration.* On va séparer la dynamique entre les temps de sauts et celle au moment des instants de sauts.

Si la taille de la population  $N_s$  au temps  $s$  est finie, alors le taux de saut au temps  $s$  est majoré par  $N_s(\bar{b} + \bar{m}) + \tau$ . On peut donc définir presque sûrement (p.s)  $(T_k)_{k \geq 1}$  la suite des temps de saut du processus  $Z$ . Comme cette suite est croissante, on peut également définir

$$T_\infty = \lim_{k \rightarrow \infty} T_k.$$

On va montrer que  $T_\infty = \infty$ , et on construira la solution de manière itérative.

En l'absence de saut dans le processus, la position des individus reste constante, seule la biomasse évolue. Le théorème de Cauchy-Lipschitz assure l'existence et l'unicité de la solution du problème de Cauchy sur  $a$  de l'Équation (4.12) avec la condition initiale  $a(t_0) = a_0$ .

On va maintenant montrer par l'absurde que  $T_\infty = \infty$  *p.s.*

$\forall n \in \mathbb{N}$ , on définit le temps d'arrêt

$$\tau_n := \inf\{t > 0, N_t > n\}.$$

On a  $\forall t \geq 0$  :

$$\begin{aligned} \langle Z_{t \wedge \tau_n}, 1 \rangle &= N_0 + \int_0^{t \wedge \tau_n} \int_{E_1} 1_{\{i \leq N_{s-}\}} 1_{\{\theta \leq b(X_i(Z_s), Z_s)\}} 1_{\{X_i(Z_s) + z \in C_{X_i(Z_s), Z_s}\}} M_1(ds, di, d\theta, dz) \\ &\quad - \int_0^{t \wedge \tau_n} \int_{E_3} 1_{\{i \leq N_{s-}\}} 1_{\{\theta \leq m(A_i(Z_s))\}} M_3(ds, di, d\theta) \end{aligned}$$

car

$$\left\langle \int_0^{t \wedge \tau_n} \int_{[0,1]^{N^*}} \sum_{i=1}^{N_{s-}} 1_{Y_i \leq p} (\delta_{(X_i(Z_s), A_b(t, s, A_i(Z_{s-})).F(A_i(Z_{s-})))}) - \delta_{(X_i(Z_s), A_b(t, s, A_i(Z_s)))}) M_2(ds, dy), 1 \right\rangle = 0,$$

où on note  $p := \text{proportionMowing}$ .

Donc

$$\langle Z_{t \wedge \tau_n}, 1 \rangle \leq N_0 + \int_0^{t \wedge \tau_n} \int_{E_1} 1_{\{i \leq N_{s-}\}} 1_{\{\theta \leq b(X_i(Z_s), Z_s)\}} 1_{\{X_i(Z_s) + z \in C_{X_i(Z_s), Z_s}\}} M_1(ds, di, d\theta, dz) \quad (4.13)$$

en prenant le sup, puis l'espérance et en appliquant le théorème de Fubini (car les quantités intégrées sont bornées, par définition de  $\tau_n$ ), on obtient :

$$\mathbb{E} \left( \sup_{s \in [0, t \wedge \tau_n]} N_s \right) \leq \mathbb{E}(N_0) + c \int_0^t \mathbb{E} \left( \sup_{u \in [0, s \wedge \tau_n]} N_u \right) ds.$$

En effet, on a

$$\begin{aligned} &\mathbb{E} \left( \int_0^{t \wedge \tau_n} \int_{E_1} 1_{\{i \leq N_{s-}\}} 1_{\{\theta \leq b(X_i(Z_s), Z_s)\}} 1_{\{X_i(Z_s) + z \in C_{X_i(Z_s), Z_s}\}} M_1(ds, di, d\theta, dz) \right) \\ &= \mathbb{E} \left( \int_0^{t \wedge \tau_n} \int_{E_1} 1_{\{i \leq N_{s-}\}} 1_{\{\theta \leq b(X_i(Z_s), Z_s)\}} 1_{\{X_i(Z_s) + z \in C_{X_i(Z_s), Z_s}\}} ds n(di) d\theta D(z) \right). \end{aligned}$$

De plus,  $\int_{\mathbb{N}^*} 1_{\{i \leq N_{s-}\}} n(di) = N_{s-}$  et il existe une constante  $c_1$  telle que

$$\int_{\mathbb{R}^2} 1_{\{z \in C_{x, Z}\}} D(z) < c_1, \text{ où } c_1 \text{ est indépendante de } Z \text{ et } x. \quad (4.14)$$

De même, il existe une constante  $\bar{b}$  telle que

$$\int_{\mathbb{R}_+} 1_{\{\theta \leq b(X_i(Z_s), Z_s)\}} d\theta \leq \bar{b}.$$

On pose donc  $c = \bar{b} + c_1$ .

Enfin, on applique le Lemme de Gronwall (voir Section A.1 de l'Annexe A) à l'Équation (4.13), ce qui donne une majoration indépendante de  $n$  :

$$\mathbb{E}\left(\sup_{s \in [0; t \wedge \tau_n]} N_s\right) \leq \mathbb{E}(N_0)e^{ct}. \quad (4.15)$$

Montrons maintenant que

$$\lim_{n \rightarrow \infty} \tau_n = \infty.$$

Nous allons procéder par l'absurde.

S'il existe  $M < \infty$  et  $A_M \subset \Omega$  tel que  $\mathbb{P}(A_M) > 0$  de sorte que :

$$\forall \omega \in A_M, \lim_{n \rightarrow \infty} \tau_n(\omega) < M.$$

On a alors,  $\forall t > M, \forall n \in \mathbb{N}$  :

$$\begin{aligned} \mathbb{E}\left(\sup_{s \in [0; t \wedge \tau_n]} N_s\right) &= \mathbb{E}\left(\sup_{s \in [0; t \wedge \tau_n]} N_s | A_M\right) \cdot \mathbb{P}(A_M) + \mathbb{E}\left(\sup_{s \in [0; t \wedge \tau_n]} N_s | A_M^c\right) \cdot (1 - \mathbb{P}(A_M)) \\ &\geq \mathbb{E}\left(\sup_{s \in [0; t \wedge \tau_n]} N_s | A_M\right) \cdot \mathbb{P}(A_M) \\ &\geq n \mathbb{P}(A_M). \end{aligned}$$

La dernière inégalité contredit l'inégalité (4.15) (majoration indépendante de  $n$ ).

On a donc montré :

$$\lim_{n \rightarrow \infty} \tau_n = \infty.$$

On en déduit par le lemme de Fatou, que  $\forall T > 0$  :

$$\begin{aligned} \mathbb{E}\left(\sup_{s \in [0; T]} N_s\right) &= \mathbb{E}\left(\liminf_{N \rightarrow \infty} \sup_{s \in [0; T \wedge \tau_n]} N_s\right) \\ &\leq \liminf_{N \rightarrow \infty} \mathbb{E}\left(\sup_{s \in [0; T \wedge \tau_n]} N_s\right) \\ &\leq \mathbb{E}(N_0)e^{cT}. \end{aligned}$$

Montrons maintenant par l'absurde que  $T_\infty = \infty$  p.s.

On suppose que :  $\exists M < \infty, A_M \subset \Omega$  tel que  $\mathbb{P}(A_M) > 0$  et  $\forall \omega \in A_M, T_\infty(\omega) < M$

Montrons que

$$\forall \omega \in A_M, \lim_{k \rightarrow \infty} N_{T_k(\omega)} = \infty, \quad (4.16)$$

ce qui entraînera  $\forall n > 0, \forall \omega \in A_M, \tau_n(\omega) < M$ , ce qui contredit  $\lim_{n \rightarrow \infty} \tau_n = \infty$ .

Une nouvelle fois, raisonnons par l'absurde pour montrer (4.16) :

$$\text{Si } \exists N' > 0, B \subset A_N \text{ tel que } \mathbb{P}(B) > 0 \text{ et } \forall \omega \in B, \forall k \in \mathbb{N}, N_{T_k(\omega)} < N',$$

alors,  $\forall \omega \in B$ , on peut construire (voir le paragraphe construction des instants de saut ci-dessous), les instants de saut  $T_k(\omega)$  comme une sous-suite de points d'un processus ponctuel de Poisson d'intensité  $\bar{b}N' + \bar{d}N' + \tau$ , qui a pour valeur d'adhérence  $+\infty$ . Cette dernière assertion contredit le fait que  $\forall \omega \in B, T_\infty(\omega) < M$ .

### Construction des instants de saut

Construisons la suite des instants de saut par récurrence. On pose  $T_0 = 0$  et on note  $Z_0 \in \mathcal{M}$  la condition initiale. Supposons qu'on ait déjà simulé  $k$  événements, et que le dernier ait lieu au temps  $T_k$ . On a  $N_{T_k} = \langle Z_{T_k}, 1 \rangle$ , la taille de la population à l'instant  $T_k$  qui est finie. On rappelle que  $\tau$  désigne le taux auquel ont lieu les événements de fauche. Le taux de saut est majoré par

$$S_k := \bar{b}N_{T_k} + \bar{d}N_{T_k} + \tau.$$

Nous allons déterminer l'instant  $T_{k+1}$  du  $(k+1)$ -ème événement, et la nature de cet événement par une méthode d'acceptation-rejet. A partir du temps  $T_k$ , on simule des temps  $(\mu_{k+l})_{l \in \mathbb{N}}$  suivant un processus de Poisson d'intensité  $S_k$ . Le premier instant  $\tau_{k+l}$  accepté par la procédure définit l'instant  $T_{k+1}$

Les étapes de la procédure sont :

- 1 : On pose  $\mu_k = T_k$ ,  $N_{\mu_k} = N_{T_k}$ , et  $l = 0$ .
- 2 : On simule  $\epsilon_{k,l}$  une variable aléatoire de loi de Poisson de paramètre  $S_k$ , indépendante des autres variables aléatoires. On pose  $\mu_{k+l+1} = \mu_{k+l} + \epsilon_l$
- 3 : Sur l'intervalle  $[\mu_{k+l}; \mu_{k+l+1}[$ , seul le phénomène de croissance de la biomasse se produit, l'individu  $i$  de trait  $X_i(Z_{\mu_{k+l}})$  devient au temps  $\mu_{k+l+1}$  :  $A_b(\mu_{k+l+1}, \mu_{k+l}, A_i(Z_{\mu_{k+l+1}}))$ .
- 4 : On simule une variable aléatoire  $I_{k,l}$  à valeur entière, uniformément dans  $\llbracket 1; N_{T_k} \rrbracket$  et on définit les quantités suivantes :

$$\begin{aligned} - \tilde{m}_1(\mu_{k+l+1}, Z_{\mu_{(k+l+1)-}}, I_{k,l}) &= \frac{b(X_{I_{k,l}}(Z_{\mu_{(k+l+1)-}}))}{S_k} \\ - \tilde{m}_2(\mu_{k+l+1}, Z_{\mu_{(k+l+1)-}}, I_{k,l}) &= \tilde{m}_1(\mu_{(k+l+1)-}, Z_{\mu_{(k+l+1)-}}, I_{k,l}) + \frac{m(A_{I_{k,l}}(Z_{\mu_{(k+l+1)-}}))}{S_k} \\ - \tilde{m}_3(\mu_{k+l+1}, Z_{\mu_{(k+l+1)-}}, I_{k,l}) &= \tilde{m}_2(\mu_{(k+l+1)-}, Z_{\mu_{(k+l+1)-}}, I_{k,l}) + \frac{\tau}{S_k}. \end{aligned}$$

- 5 : On tire une variable aléatoire  $\Theta_{k,l}$  de loi uniforme dans  $[0, 1]$ .

a) Si  $0 \leq \Theta_{k,l} < \tilde{m}_1$ , alors on réalise le tirage noté  $U$  dans  $\mathbb{R}^2$  suivant la loi  $D$  de dispersion. Si  $X_{I_{k,l}}(Z_{\mu_{(k+l+1)-}}) + U \in C_{X_{I_{k,l}}(Z_{\mu_{(k+l+1)-}}, Z_{\mu_{(k+l+1)-}}}$ , alors le  $(k+1)$ -ème événement a lieu : l'individu  $I_{k,l}$  donne naissance à une couronne de biomasse  $a_0$  en position  $X_{I_{k,l}}(Z_{\mu_{k+l+1}}) + U$ , et on définit  $T_{k+1} = \mu_{k+l+1}$ .

Sinon, on ne fait rien, il n'y a pas d'événement au temps  $\mu_{k+l+1}$ , et on reprend l'algorithme en 2 avec,  $N_{\mu_{k+l+1}} = N_{T_k}$ , et  $l = l + 1$  jusqu'à obtenir le  $(k+1)$ -ème événement.

b) Si  $\tilde{m}_1 \leq \Theta_{k,l} < \tilde{m}_2$ , le  $(k+1)$ -ème événement a lieu : mort de l'individu  $I_{k,l}$ . On définit  $T_{k+1} = \mu_{k+l+1}$ .

c) Si  $\tilde{m}_2 \leq \Theta_{k,l} < \tilde{m}_3$ , le  $(k+1)$ -ème événement a lieu : c'est une fauche. On réalise le tirage noté  $y$  selon une loi uniforme sur  $[0, 1]^{N_{T_k}}$  et on note

$$F_y := \{i \in \llbracket 1; N_{T_k} \rrbracket / y_i \leq p\}.$$

$\forall i \in F_y$ , on fait :

$$A_i(Z_{\mu_{k+l+1}}) = A_i(Z_{\mu_{(k+l+1)-}}) \cdot F(A_i(Z_{\mu_{(k+l+1)-}})).$$

On définit  $T_{k+1} = \mu_{k+l+1}$ .

d) si  $\Theta_{k,l} > \tilde{m}_3$ , rien ne se produit. On reprend l'algorithme en 2 avec,  $N_{\mu_{k+l+1}} = N_{T_k}$ , et  $l = l + 1$  jusqu'à obtenir le  $(k+1)$ -ème événement.

### Existence et unicité au niveau des temps de saut

Soit  $Z_0 \in \mathcal{M}$  la condition initiale et  $M_1, M_2, M_3$  3 mesures ponctuelles de Poisson définies dans l'EDS (4.10).

Pour  $k = 0$ ,  $Z_0$  est déterminé de façon unique par le quadruplet  $(Z_0, M_1, M_2, M_3)$ .

On va montrer par récurrence que les couples  $(T_k, Z_k)_{k \in \mathbb{N}}$  sont déterminés de façon unique par  $(Z_0, M_1, M_2, M_3)$ .

Supposons que ça soit le cas pour  $k \in \mathbb{N}$ , montrons que  $(T_{k+1}, Z_{k+1})$  est aussi déterminé de façon unique par  $(Z_0, M_1, M_2, M_3)$ .

On définit :  $\forall t \geq T_k$ ,

$$\tilde{Z}_t^k(dx, da) = \sum_{i=1}^{N_{T_k}} \delta_{(X_i(Z_t), A_i(t, T_k, A_i(Z_{T_k})))}$$

Par hypothèse,  $\tilde{Z}_t^k(dx, da)$  est déterminé de façon unique par  $(Z_0, M_1, M_2, M_3)$ .

On a de plus  $T_{k+1} = \min(\inf_1, \inf_2, \inf_3)$

où

$$\inf_1 = \inf\{t > T_k, \int_{T_k}^t \int_{E_1} 1_{\{i \leq N_{s-}\}} 1_{\{\theta \leq b(X_i(\tilde{Z}_s^k), \tilde{Z}_s^k)\}} 1_{\{X_i(\tilde{Z}_s^k) + z \in C_{X_i(\tilde{Z}_s^k), \tilde{Z}_s^k}\}} M_1(ds, di, d\theta, dz) > 0\}$$

$$\inf_2 = \inf\{t > T_k, \int_{T_k}^t \int_{E_2} M_2(ds, dy) > 0\}$$

$$\inf_3 = \inf\{t > T_k, \int_{T_k}^t \int_{E_3} 1_{\{i \leq N_{s-}\}} 1_{\{\theta \leq m(A_i(\tilde{Z}_s^k))\}} M_3(ds, di, d\theta, dz) > 0\}$$

Donc  $T_{k+1}$  est déterminé de façon unique par  $(T_k, Z_k)$ , donc par  $(Z_0, M_1, M_2, M_3)$ .

On note  $at$  l'atome correspondant au saut, c'est un atome d'une des trois mesures de Poisson :

$$\begin{aligned} at &= (T_{k+1}^1, I_{k+1}^1, \Theta_{k+1}^1, U_{k+1}^1) 1_{T_{k+1}=\inf_1} \\ &+ (T_{k+1}^2, Y_{k+1}^2) 1_{T_{k+1}=\inf_2} \\ &+ (T_{k+1}^3, I_{k+1}^3, \Theta_{k+1}^3) 1_{T_{k+1}=\inf_3} \end{aligned}$$

On a alors,

$$\begin{aligned} Z_{T_{k+1}}(dx, da) &= \tilde{Z}_{T_k}^k(dx, da) + \delta_{(X_{I_{k+1}^1}^1(Z_{T_k}) + U_{k+1}^1, a_0)} 1_{\{I_{k+1}^1 \leq N_{s-}\}} \\ &\times 1_{\{\Theta_{k+1}^1 \leq b(X_{I_{k+1}^1}^1(Z_{(T_{k+1})-}), Z_{T_k})\}} 1_{\{X_{I_{k+1}^1}^1(Z_{T_k}) + U_{k+1}^1 \in C_{X_{I_{k+1}^1}^1(Z_{T_k}), Z_{T_k}}\}} 1_{\{\min(\inf_1, \inf_2, \inf_3) = \inf_1\}} \\ &+ \sum_{i=1}^{N_{T_k}} 1_{Y_i^2 \leq p} (\delta_{(X_i(Z_{(T_{k+1})-}), A_i(Z_{(T_{k+1})-}))} \cdot F(A_i(Z_{(T_{k+1})-}))) \\ &\quad - \delta_{(X_i(Z_{(T_{k+1})-}), A_i(Z_{(T_{k+1})-}))} 1_{\{\min(\inf_1, \inf_2, \inf_3) = \inf_2\}} \\ &- \delta_{(X_{I_{k+1}^3}^3(Z_{T_k}), A_{I_{k+1}^3}^3(Z_{(T_{k+1})-}))} 1_{\{I_{k+1}^3 \leq N_{T_k}\}} 1_{\{\Theta_{k+1}^3 \leq m(A_{I_{k+1}^3}^3(Z_{(T_{k+1})-}))\}} 1_{\{\min(\inf_1, \inf_2, \inf_3) = \inf_3\}} \end{aligned} \tag{4.17}$$

Ainsi,  $Z_{T_{k+1}}$  dépend de manière unique de  $(Z_{T_k}^k, T_{k+1}, at)$  donc de  $(Z_0, M_1, M_2, M_3)$ , ce qui conclut la démonstration.  $\square$

### 4.2.2 Calibration

Les valeurs des paramètres calibrés sont présentées dans le Tableau 4.2. L'ensemble des solutions fournies par l'algorithme NSGA2 s'est stabilisé après 165000 itérations. Pour un jeu de paramètres, *evol.sample* fait référence au nombre de réplifications qui ont été effectuées par l'algorithme. Comme le modèle est stochastique, nous devons choisir une solution avec une valeur suffisamment grande pour *evol.sample* afin d'obtenir une solution fiable. Parmi l'ensemble des solutions fournies par l'algorithme, nous avons choisi la solution qui a été répliquée au moins 50 fois et qui minimise la distance  $dist(simu, data)$ .

| Paramètre              | Valeur issue de la calibration | unit      |
|------------------------|--------------------------------|-----------|
| K                      | 12.72                          | $g$       |
| L                      | 0.26                           | $an^{-1}$ |
| distanceCompetition    | 0.15                           | $m$       |
| distanceParent         | 0.20                           | $m$       |
| shape                  | 4.34                           |           |
| scale                  | 2.36                           |           |
| deathParameterDecrease | 2.32                           | $g^{-1}$  |
| deathParameterScaling  | 1.12                           | $an^{-1}$ |
| mowingParameter        | 0.11                           | $g^{-1}$  |
| bbar                   | 0.18                           | $an^{-1}$ |
| a0                     | 1.73                           | $g$       |
| score                  | 26.06                          |           |
| evol.sample            | 79                             |           |

TABLE 4.2 – Résultat de la calibration obtenu avec OpenMOLE (Reuillon et al., 2013)  
 Dans le Tableau 4.2, *score* est la médiane sur les 79 réplifications de la somme des distances  $dist(simu, data)$  sur les 19 taches et les quatre caractéristiques (initiale ou finale et aire ou taille).

Un score de 26,06 signifie que dans la moitié des cas, la distance relative pour une caractéristique entre la tache simulée et les données correspondantes est inférieure en moyenne à 0,3. La raison de cette différence est que les données ont été obtenues à partir d'expériences de terrain qui n'ont pas été effectuées pour calibrer le modèle, et peuvent donc contenir un biais dû à l'altitude ou au type de sol.

Bien que nous n'ayons pas trouvé de valeurs pour les paramètres de la dynamique des plantes dans la littérature, des experts peuvent fournir des bornes pour certains d'entre eux, ce qui permet d'évaluer la qualité du résultat donné par l'algorithme selon leur pertinence écologique. Tout d'abord, les paramètres *distanceCompetition* et *distanceParent* sont proches de ce qui est observé lors d'expériences de terrain. Ensuite, la distribution pour la dispersion des individus est proche de celle suggérée par les spécialistes (Figure 4.4). La Figure 4.5 représente le taux de mortalité d'une couronne en fonction de sa biomasse. On constate qu'une couronne non fauchée a un taux de mortalité très faible, ce qui est en accord avec les observations de terrain. En effet, pour comparer avec la valeur trouvée par Smith et al. (2007), en utilisant l'Équation (4.8) et les valeurs des paramètres données par la calibration, on calcule la probabilité qu'une couronne meure avant 4 mois. Cette quantité est égale à 0,0027, qui a le même ordre de grandeur que la valeur trouvée dans Smith et al. (2007) pour la probabilité qu'un segment de rhizome meure dans une période de 4 mois (0,0083). Enfin, le rapport entre la valeur de  $K$  (biomasse maximale, qui est susceptible d'être atteinte pour la couronne la plus ancienne, c'est-à-dire au centre de la tache lorsqu'il n'y a pas de fauche), et  $a_0$ , la biomasse d'une couronne à la naissance (plutôt qu'à la périphérie) est égale à 7,4 (le rapport devrait être



d'environ 10 dans Adachi et al. (1996a)). La Figure 4.7 représente l'évolution de la biomasse en fonction du temps, sans fauche, avec les valeurs de paramètres issues de la calibration. Quant à la Figure 4.6, elle représente la fonction décrivant l'effet de la fauche sur la biomasse également pour des valeurs de paramètres issues du Tableau 4.2.

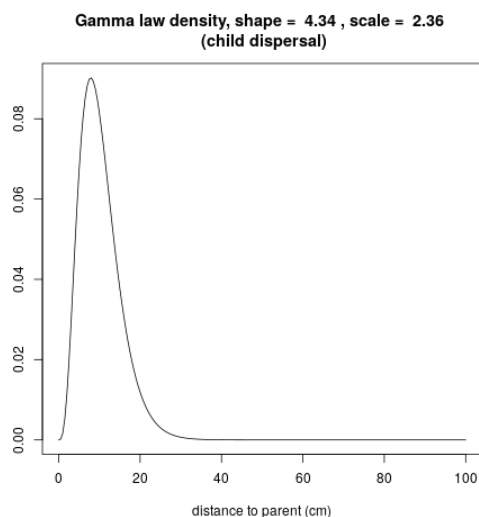


FIGURE 4.4 – Loi Gamma pour la distance de dispersion (Équation (4.3)) avec la valeur des paramètres shape scale issue de la calibration (Tableau 4.2), extrait de Lavallée et al. (2019b).

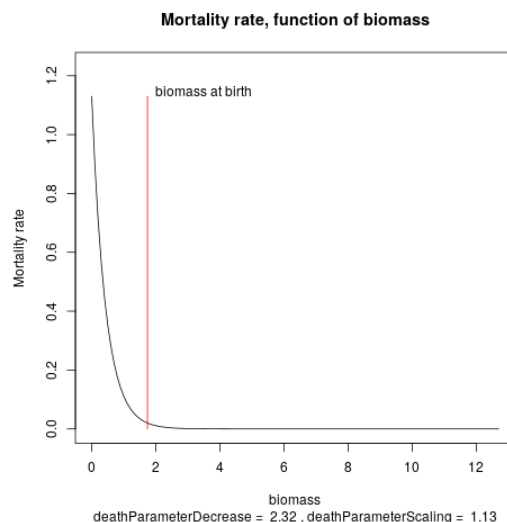


FIGURE 4.5 – Taux de mortalité en fonction de la biomasse (Équation (4.9)) avec la valeur des paramètres *deathParameterScaling* et *deathParameterDecrease* issue de la calibration (Tableau 4.2), extrait de Lavallée et al. (2019b).

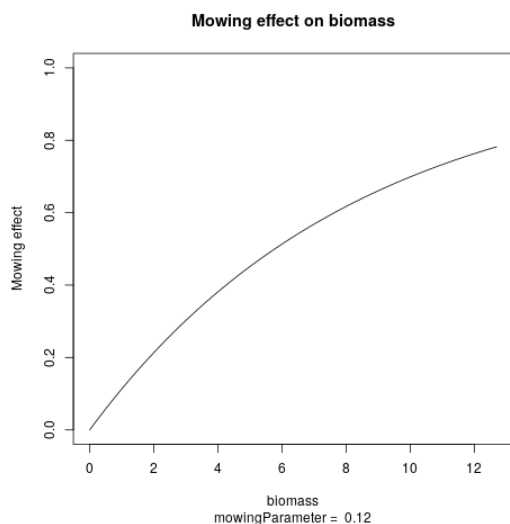


FIGURE 4.6 – Fonction décrivant l'effet de la fauche, Équation (4.5) avec la valeur du paramètre *mowingParameter* issue de la calibration (Tableau 4.2), extrait de Lavallée et al. (2019b).

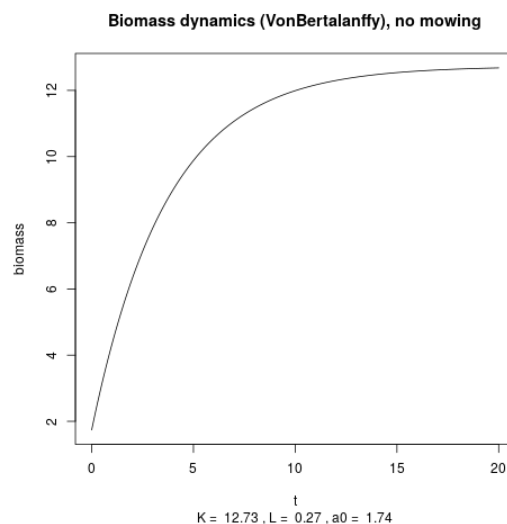


FIGURE 4.7 – Évolution de la biomasse d'un individu, sans fauche pendant 20 ans (Équation (4.7)) avec la valeur des paramètres  $K$  et  $L$  issue de la calibration (Tableau 4.2), extrait de Lavallée et al. (2019b).

### 4.2.3 Influence des paramètres de gestion et de la taille initiale

Dans cette section, l'objectif est de trouver des relations statistiques entre les variables explicatives (paramètres de gestion ou taille initiale de la population) et les sorties du modèle (aire moyenne et taille d'une tache).

Nous considérons les deux échantillonnages suivants :

- Dans *sampling1*, nous faisons varier  $\tau$  dans  $[0, 15]$  par pas de 0,5,  $T$  dans  $[0, 16]$  par pas de 1 et *initialPopSize* est fixé à 500 ou 1500. Nous effectuons 50 simulations du modèle stochastique pour chaque ensemble de valeurs. *sampling1* a un taux d'échantillonnage élevé sur  $\tau$  et  $T$ , avec des valeurs élevées pour la taille initiale de la population. Il est utilisé pour étudier plus précisément l'influence de  $\tau$  et  $T$ .
- Dans *sampling2*, nous faisons varier  $\tau$  dans  $[0, 14]$  par pas de 2,  $T$  dans  $[0, 16]$  par pas de 2 et *initialPopSize* dans  $[50, 1200]$  par pas de 50. Nous effectuons 50 simulations du modèle stochastique pour chaque ensemble de valeurs. *sampling2* a un taux d'échantillonnage élevé sur la taille initiale de la population, il est utilisé pour étudier plus précisément son influence.

#### Influence de la durée du projet de gestion $T$

Dans cette section, le premier échantillonnage (*sampling1*) est utilisé pour étudier l'influence de la durée du projet de gestion  $T$  sur les aires et tailles finales moyennes. Étant donné  $\tau$  et une valeur de la taille initiale de la population, nous effectuons :

- une régression quadratique tronquée pour l'aire finale moyenne ;
- une régression non linéaire sur des valeurs strictement positives pour la taille finale moyenne de la population, en utilisant la fonction  $f(T) = \text{initialPopSize} * \exp(-T/\text{rate})$ , avec *rate* étant une constante sur laquelle l'algorithme *nls* maximise  $R^2$ . Cette constante est différente pour chaque jeu de paramètres, puisqu'elle dépend des valeurs de  $\tau$  et de *initialPopSize*. Nous étudions cette dépendance dans un paragraphe suivant de cette section intitulé "Formules pour l'aire et la taille finales moyennes, en fonction de  $\tau$ ,  $T$  et de la taille de population initiale".

Pour des valeurs de  $\tau \leq 2,5$ , l'aire moyenne reste proche de sa valeur initiale à l'instant  $T = 0$  (une différence relative maximale de  $3 \text{ m}^2$ ), et la variation est plutôt linéaire du type

$$\text{aire}(T) = e(2,5 - \tau)T + a_{ini}, \quad (4.18)$$

où  $e$  et  $a_{ini}$  sont des constantes, mais les valeurs correspondantes de  $R^2$  sont inférieures à 0,9. La Figure 4.8 donne un exemple de régression linéaire sur l'aire moyenne en fonction de la durée de gestion  $T$ , pour des valeurs fixes données de  $\tau \leq 2,5$  et la taille initiale de la population.

Les valeurs de  $R^2$  et de *RMSE* nous permettent de conclure que l'aire finale moyenne dépend quadratiquement de la durée du projet de gestion  $T$  lorsque  $\tau > 2,5$ . En effet, 47 régressions sur les 50 de l'échantillonnage (variation de la taille initiale de la population et  $\tau > 2,5$ ) donnent une valeur de  $R^2$  supérieure à 0,95. La valeur maximale de *RMSE* sur ces 50 régressions est de  $1,31 \text{ m}^2$ , ce qui est faible par rapport à l'aire initiale moyenne, qui a pour valeurs  $20 \text{ m}^2$  ou  $60 \text{ m}^2$ . La Figure 4.9 donne un exemple de la régression quadratique sur l'aire moyenne en fonction de la durée du projet de gestion  $T$  pour des valeurs fixées de  $\tau$  et de la taille initiale de la population.

Ces résultats donnent des informations sur l'influence de la durée du projet de gestion sur la croissance des taches. Un premier fait, très important pour la gestion, est qu'il ne suffit pas de faucher pour diminuer la taille et l'aire de la tache : si le nombre de fauches par an est trop

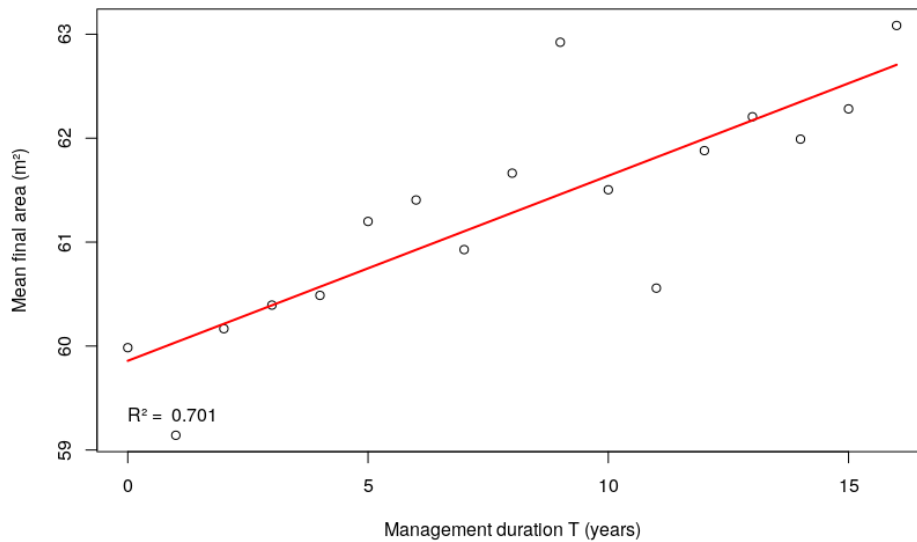


FIGURE 4.8 – Régression linéaire sur l'aire finale moyenne en fonction de la durée du projet de gestion  $T$ , avec  $\tau = 0,5 \text{ an}^{-1}$  et  $\text{initialPopSize} = 1500$ , extrait de Lavallée et al. (2019b).

faible (moins de 2,5 dans le cas présent), la taille de la population et l'aire augmentent pendant le projet de gestion. La Figure 4.10 présente les courbes de régression quadratique pour l'aire finale moyenne par rapport à la durée du projet de gestion ( $T$ , en abscisse), obtenues pour différentes valeurs de  $\tau$ . Un deuxième fait important pour la gestion est que l'éradication d'une tache de renouée s'étendant initialement sur  $60 \text{ m}^2$  ne peut pas être atteinte en moins de 11 ans (pour 15 fauches par an). La figure illustre également le gain en terme d'aire sur l'aire finale moyenne lors d'une fauche supplémentaire par an. Par exemple, en fauchant six fois par an au lieu de cinq, pendant 10 ans, l'aire finale moyenne de la tache est réduite en moyenne de  $4 \text{ m}^2$  (voir la verticale bleue à  $T = 10$  dans la Figure 4.10).

Comme pour l'aire, la taille varie de façon linéaire pour des valeurs de  $\tau \leq 2,5$  ( $R^2$  vaut environ 0,9). Lorsque  $\tau > 2,5$ , les valeurs de  $R^2$  et de  $RMSE$  nous permettent de conclure que la taille finale moyenne dépend de manière exponentielle de la durée du projet de gestion  $T$ . En effet, l'ensemble des 50 régressions de l'échantillonnage (variation de la taille initiale de la population et de  $\tau > 2,5$ ) conduisent à une valeur de  $R^2$  supérieure à 0,95. La valeur maximale de  $RMSE$  sur ces 50 régressions est de 39 couronnes, ce qui est faible par rapport à la taille initiale de la population, qui a pour valeurs 500 couronnes ou 1500 couronnes. La Figure 4.11 donne un exemple de régression quadratique sur l'aire moyenne en fonction de la durée de gestion  $T$  pour des valeurs fixées de  $\tau$  et de taille initiale de la population.

### Influence du nombre moyen de fauches par année $\tau$

Dans cette section, le premier échantillonnage (*sampling1*) est utilisé pour étudier l'influence du nombre moyen de fauches par année sur les aires et tailles finales moyennes. Étant donné  $T$  et une valeur de taille initiale de la population, nous effectuons :

- une régression linéaire sur les valeurs strictement positives de l'aire finale moyenne ;
- une régression non linéaire sur des valeurs strictement positives de la taille finale moyenne de la population, en utilisant la fonction  $f(\tau) = \text{initialPopSize} * \exp(-\tau/\text{rate})$ ,  $\text{rate}$  étant une constante sur laquelle l'algorithme *nls* minimise  $RMSE$ .

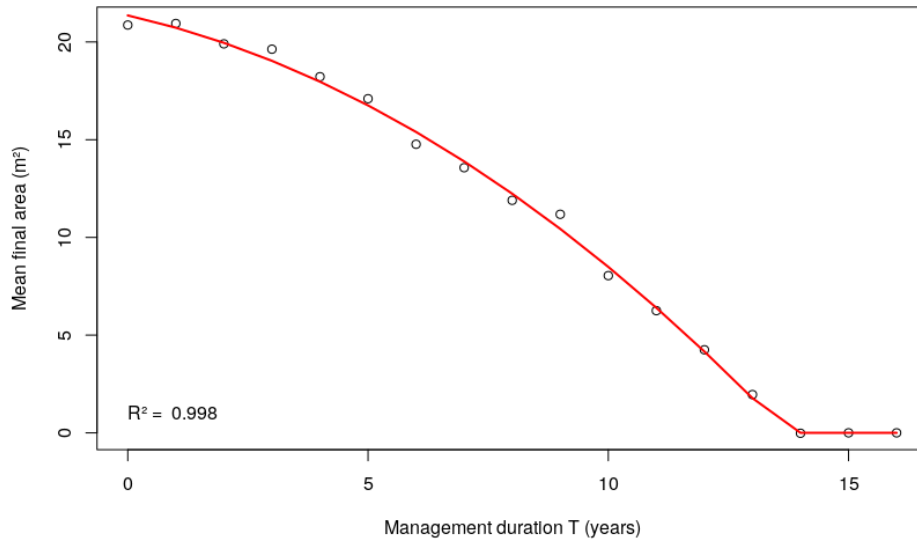


FIGURE 4.9 – Régression quadratique sur l'aire finale moyenne en fonction de la durée du projet de de gestion  $T$ , avec  $\tau = 8 \text{ an}^{-1}$  et  $\text{initialPopSize} = 500$ , extrait de Lavallée et al. (2019b). Pour les trois derniers points dans le coin inférieur droit, au moins la moitié des 50 simulations conduisent à l'extinction de la tache.

Cette constante *rate* diffère pour chaque jeu de paramètres car elle dépend des valeurs de  $T$  et de *initialPopSize*. Comme nous l'avons mentionné pour la constante similaire dans le paragraphe précédent, nous étudions cette dépendance dans un paragraphe suivant de cette section intitulé "Formules pour l'aire et la taille finales moyennes, en fonction de  $\tau$ ,  $T$  et de la taille de population initiale".

La régression linéaire présentée Figure 4.12 est valable pour  $\tau > 2,5$  (comme dans le paragraphe précédent) et  $T \geq 2$  (pour avoir une population décroissante). Les valeurs de  $R^2$  et de *RMSE* nous permettent de conclure que l'aire finale moyenne dépend linéairement du nombre moyen de fauches par année  $\tau$ . En effet, 24 régressions sur les 30 de l'échantillonnage (variation de la taille initiale de la population et de  $T \geq 2$ ) conduisent à une valeur de  $R^2$  supérieure à 0,95. La valeur maximale de *RMSE* sur ces 50 régressions est de  $2,09 \text{ m}^2$ , ce qui est faible par rapport à l'aire initiale moyenne, qui a pour valeurs  $20 \text{ m}^2$  ou  $60 \text{ m}^2$ . La Figure 4.12 donne un exemple de régression linéaire sur l'aire finale moyenne en fonction du nombre moyen de fauches pour des valeurs fixées de  $T$  et de la taille initiale de la population.

Nous concluons également que la taille finale moyenne dépend de manière exponentielle du nombre moyen de fauches par année  $\tau$ . En effet, les 30 régressions effectuées sur l'échantillonnage (variation de la taille initiale de la population et de  $T \geq 2$ ) conduisent à une valeur de  $R^2$  supérieure à 0,95. La valeur maximale de *RMSE* sur ces 50 régressions est de 64 couronnes, ce qui est faible par rapport à la taille initiale de la population, qui a pour valeurs 500 couronnes ou 1500 couronnes.

La Figure 4.13 donne un exemple de la régression non linéaire sur la taille finale moyenne en fonction du nombre moyen de fauches par an pour des valeurs fixées de  $T$  et de la taille de la population initiale.

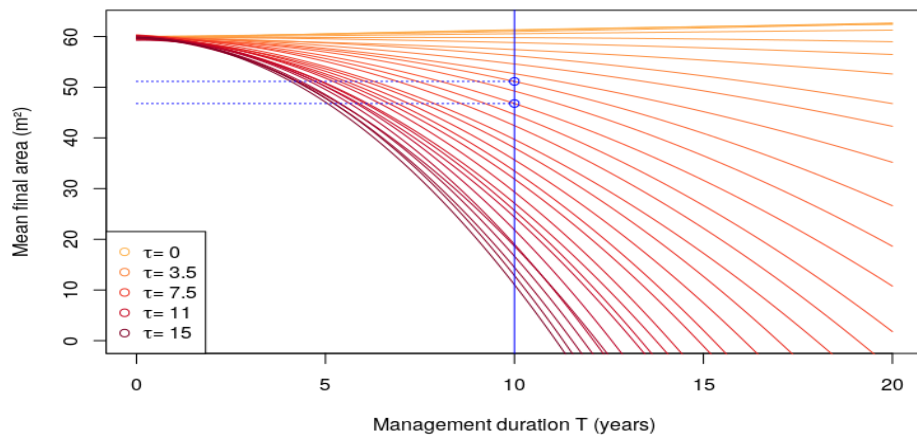


FIGURE 4.10 – Courbes de régression quadratique sur l'aire finale moyenne en fonction de  $T$ , pour différentes valeurs de  $\tau$  (faible en couleurs claires, jusqu'à 15 fauches par an en couleurs sombres), avec  $initialPopSize = 1500$  et  $proportionMowing = 0,9$ , extrait de Lavallée et al. (2019b).

### Influence de la taille de population initiale

Dans cette section, nous utilisons le second échantillonnage (*sampling2*) pour étudier l'influence de la taille initiale de la population sur les aires et tailles finales moyennes. Étant donnée une valeur pour  $\tau$  et  $T$ , nous effectuons une régression linéaire sur des valeurs de sortie strictement positives. En raison du large éventail de valeurs pour la taille initiale de population dans le *sampling2*, il peut se produire beaucoup d'extinctions pour un jeu de paramètres de gestion. Nous n'effectuons donc la régression que s'il y a au moins cinq valeurs de sorties strictement positives.

Les valeurs de  $R^2$  et  $RMSE$  nous permettent de conclure que l'aire et la taille finale moyenne dépendent toutes les deux linéairement de la taille de la population initiale. En effet, dans les deux cas, 60 régressions sur 63, parmi les 72 jeux de paramètres de l'échantillonnage, conduisent à une valeur de  $R^2$  supérieure à 0,95. La valeur maximale de  $RMSE$  sur ces 63 régressions est de  $1,1 \text{ m}^2$  (resp. 18 couronnes) pour l'aire finale moyenne (resp., taille), ce qui est faible par rapport à l'aire initiale moyenne (resp. taille de la population initiale), qui prend ses valeurs entre  $2 \text{ m}^2$  et  $48 \text{ m}^2$  (resp. entre 50 couronnes et 1200 couronnes).

Remarquons que l'influence de la taille de la population initiale sur l'aire initiale est également linéaire. En effet, *sampling2* contient le cas  $T = 0$ , et pour cette valeur spécifique de  $T$  l'aire finale est l'aire initiale.

Les Figures 4.14 et 4.15 donnent respectivement un exemple de la régression linéaire sur la taille moyenne et l'aire moyenne en fonction de la taille de la population initiale pour des valeurs fixées de  $T$  et  $\tau$ .

### Formules pour l'aire et la taille finales moyennes, en fonction de $\tau$ , $T$ et de la taille de population initiale

Nous synthétisons les résultats des régressions effectuées dans le Tableau 4.3.

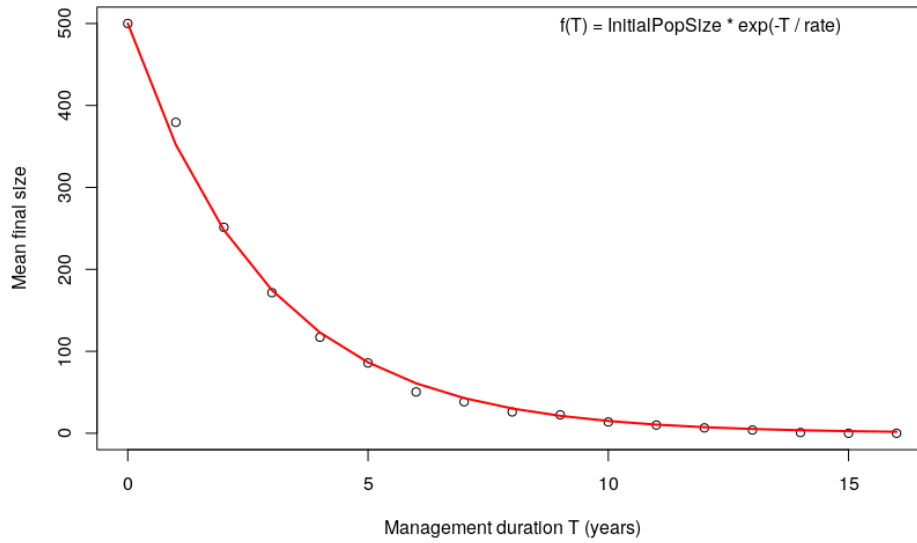


FIGURE 4.11 – Régression non linéaire sur la taille finale moyenne en fonction de la durée du projet de gestion  $T$  avec  $\tau = 8 \text{ an}^{-1}$  et  $initialPopSize = 500$ , extrait de Lavallée et al. (2019b). Pour les trois derniers points dans le coin inférieur droit, au moins la moitié des 50 simulations conduisent à l'extinction de la tache.

| Paramètre                   | Sortie moyenne | Variation       | $R^2 > 0.95$ | RMSE |
|-----------------------------|----------------|-----------------|--------------|------|
| $T$ pour $\tau \geq 2.5$    | aire finale    | quadratique ↘   | 47/50        | 1.31 |
| $T$ pour $\tau \geq 2.5$    | taille finale  | exponentielle ↘ | 50/50        | 39   |
| $\tau$ pour $\tau \geq 2.5$ | aire final     | linéaire ↘      | 24/30        | 2.09 |
| $\tau$ pour $\tau \geq 2.5$ | taille finale  | exponentielle ↘ | 30/30        | 64   |
| $initialPopSize$            | aire finale    | linéaire ↗      | 60/63        | 1.1  |
| $initialPopSize$            | taille finale  | linéaire ↗      | 60/63        | 18   |

TABLE 4.3 – Synthèses des résultats des régressions effectuées dans la Section 4.2.3.

Dans les sous-sections précédentes, nous avons étudié l'influence d'un seul paramètre à la fois, les deux autres étant fixés. Les deux échantillonnages utilisés dans les sections précédentes et introduits au début de la Section 4.2.3 ont été choisis pour contrôler la variation des paramètres de gestion et de la taille initiale de la population, afin d'étudier leur influence sur les sorties du modèle. Sur la base des résultats obtenus dans les sous-sections précédentes, nous sommes maintenant en mesure de proposer une formule pour les aires et tailles moyennes en fonction des deux paramètres de gestion (le nombre moyen de fauches par an ( $\tau$ ) et la durée du projet de gestion ( $T$ )) et de la taille initiale de la population. Nous utilisons un échantillonnage de Sobol (qui maximise la discrédance d'une suite de points, pour que l'espace soit couvert de manière uniforme) avec 5000 points pour  $\tau \in [0; 15.0]$ ,  $T \in [0; 20]$ , et  $initialPopSize \in [100; 1500]$ . Pour la même raison que celle évoquée précédemment, nous considérons le cas  $\tau \geq 2, 5$ . Les Équations (4.19) et (4.20) synthétisent les relations entre les sorties finales, les paramètres de gestion et la taille de population initiale :

$$\text{Taille Finale Moyenne} = initialPopSize * \exp(-T \cdot (\tau - a)/b), \quad (4.19)$$

avec  $a, b \in \mathbb{R}$  des constantes, et

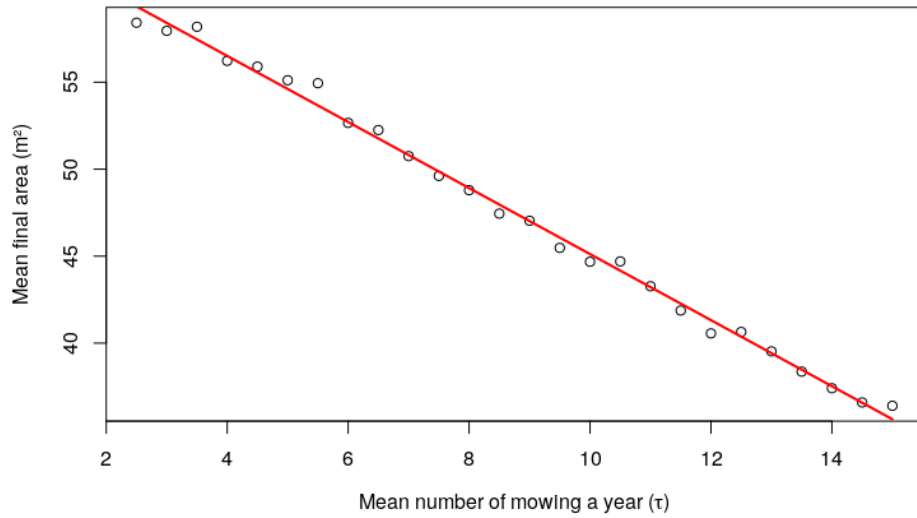


FIGURE 4.12 – Régression linéaire sur l'aire finale moyenne en fonction du nombre moyen de fauches par année  $\tau$ , avec  $T = 8$  ans et  $initialPopSize = 1500$ , extrait de Lavallée et al. (2019b).

$$Aire\ Finale\ Moyenne = \max((c \times \tau + d) * T^2 + 0.04 * initialPopSize, 0) \quad (4.20)$$

avec  $c, d \in \mathbb{R}$  des constantes.

Nous examinons maintenant les résultats de la régression non linéaire en fonction des deux paramètres de gestion ( $T$  et  $\tau \geq 2,5$ ) et de la taille de population initiale (l'échantillonnage contient 4332 valeurs pour le triplet  $(T, \tau, initialPopSize)$ ). Les valeurs de  $R^2$  et de  $RMSE$ , pour la comparaison entre les valeurs prédites et les données pour la taille moyenne valent respectivement 0,99 et 26,12 couronnes. Les intervalles de confiance à 95 % pour les constantes  $a$  et  $b$  obtenues avec  $R$  sont  $a \in [0,90; 0,94]$  et  $b \in [20,46; 20,77]$ . Les valeurs de  $R^2$  et de  $RMSE$ , pour la comparaison entre les valeurs prédites et les données pour l'aire moyenne valent respectivement 0,99 et  $2,23 m^2$ . Les intervalles de confiance à 95 % correspondants pour les constantes  $c$  et  $d$  sont respectivement  $c \in [-0,0342; -0,0336]$  et  $d \in [0,0960; 0,0998]$ .

En prenant  $T = 0$  dans le membre de droite des Équations (4.19) et (4.20), on obtient respectivement  $initialPopSize$  et  $initialPopSize * 0,04$ . La dernière quantité correspond donc à l'aire initiale moyenne. Il existe en effet une relation linéaire entre l'aire initiale moyenne et la taille de population initiale.

Nous soulignons le fait que les résultats obtenus pour les sorties moyennes sont encore pertinents pour les sorties directes (non moyennées). La Figure 4.16 illustre ce point, en traçant la Formule (4.19) avec  $T$  variant et pour des valeurs fixées de  $\tau$  et de taille de population initiale. Nous attirons aussi l'attention sur le fait que la courbe rouge sur la Figure 4.16 a été obtenue par une régression effectuée sur un ensemble de points beaucoup plus important que le sous-ensemble choisi pour tracer cet exemple.

Les Équations (4.19) et (4.20) nous permettent de trouver le paramètre qui influence le plus les sorties du modèle, c'est donc sur ce paramètre qu'il est préférable de concentrer les efforts de gestion. Pour ce faire, pour chaque valeur du triplet  $(T, \tau, initialPopSize)$ , la taille et l'aire finales ont été comparées en utilisant les Équations (4.19) et (4.20), pour les trois combinaisons

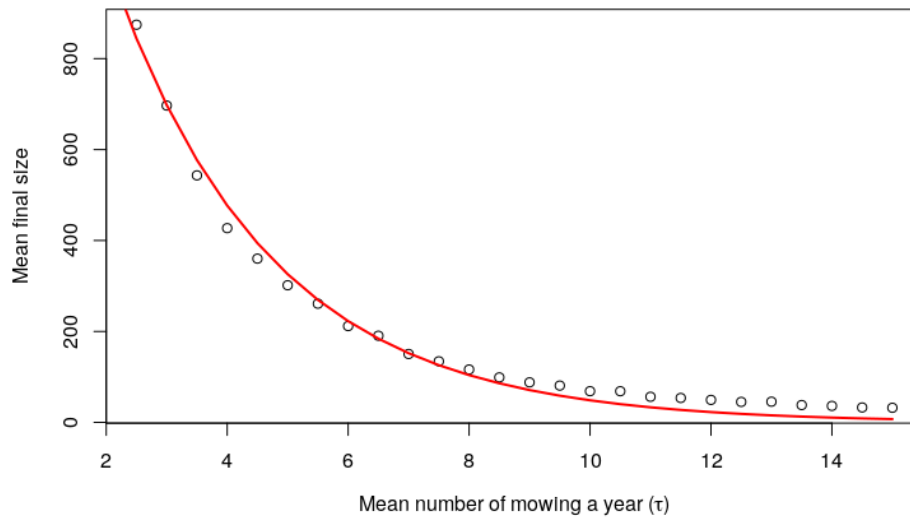


FIGURE 4.13 – Régression non linéaire sur la taille finale moyenne en fonction du nombre moyen de fauches  $\tau$  avec  $T = 8$  ans et  $initialPopSize = 1500$ , extrait de Lavallée et al. (2019b).

de valeurs de paramètres suivantes :  $(T + 1, \tau, initialPopSize)$ ,  $(T, \tau + 1, initialPopSize)$  et  $(T, \tau, initialPopSize \times 0,9)$ . Chaque graphique sur les Figures 4.17 et 4.18 correspond à une valeur fixe de  $initialPopSize$ , avec  $\tau$  variant en abscisse et  $T$  variant en ordonnée, et pour chaque triplet est associé le paramètre le plus important du point de vue de la gestion, c'est-à-dire le paramètre dont la modification produit la sortie (aire ou taille) la plus faible (il n'est pas nécessairement unique). Les zones en marron correspondent à l'ensemble des valeurs des paramètres qui conduisent directement à une éradication ; par conséquent, dans cette zone aucun gain supplémentaire ne peut être réalisé en considérant ces différentes gestions. La Figure 4.17 montre qu'en dehors des zones d'extinction, les paramètres  $T$  ou  $\tau$  ont la plus grande influence sur la taille finale, ce qui permet de déterminer la modification de gestion à adopter la plus efficace. Plus précisément, lorsque  $\tau$  est faible, il est plus efficace, pour obtenir une réduction de la taille, de faucher une fois de plus chaque année, et inversement, lorsque  $T$  est faible, il est plus efficace de continuer à faucher une année de plus. En ce qui concerne l'aire finale, nous observons sur la Figure 4.18 que les zones correspondant à la plus grande influence de  $T$  ou de  $\tau$  sont réduites par rapport à la Figure 4.17, en faveur de la zone de plus grande influence du paramètre  $initialPopSize$ . Dans ces régions, le fait de commencer le projet de gestion sur des taches plus petites (taille égale à 90% de la taille de référence) a un impact plus important sur l'aire finale que de faucher une fois de plus chaque année ou sur une période d'un an plus longue. Par conséquent, dans ces conditions, il convient d'encourager la détection précoce et la fauche de ces taches. Notez cependant que pour éradiquer une invasion, il faut toujours augmenter à la fois les paramètres  $\tau$  et  $T$ .

### 4.3 Discussion

Dans ce chapitre, nous avons proposé un modèle stochastique individu-centré pour la croissance d'une tache de renouées asiatiques incluant la fauche comme stratégie de gestion. Nous



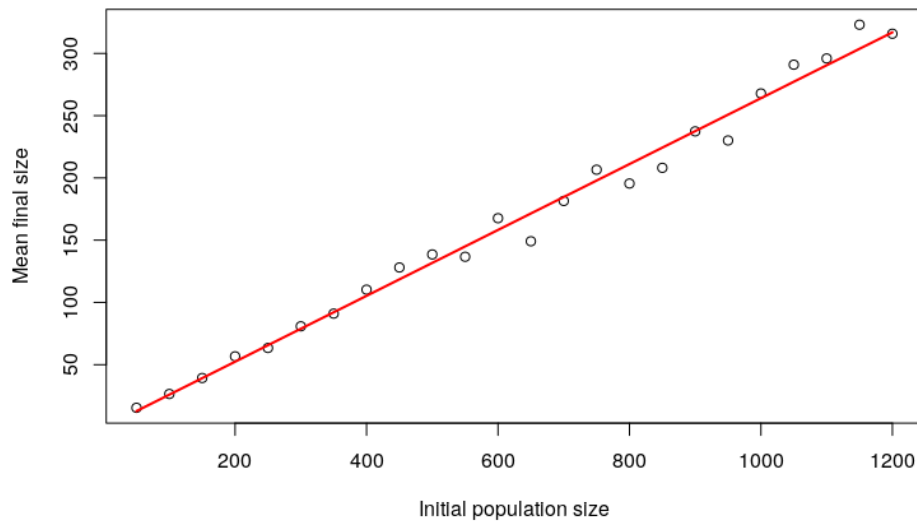


FIGURE 4.14 – Régression linéaire de la taille finale moyenne en fonction de la taille de population initiale avec  $T = 8 \text{ ans}$  et  $\tau = 4 \text{ an}^{-1}$ , extrait de Lavallée et al. (2019b).

avons ensuite calibré les paramètres de la dynamique de la plante avec des données de terrain dans la Section 4.2.2. Les valeurs de paramètres obtenues étaient en accord avec les valeurs des paramètres disponibles dans la littérature et avec notre connaissance du terrain. Dans la Section 4.2.3, nous avons étudié l'influence de la taille de la population initiale, du nombre moyen de fauches par an et de la durée du projet de gestion sur l'aire moyenne et la taille des taches. Nous avons également obtenu des formules pour l'aire et la taille moyennes d'une tache de renouée, en fonction de ces paramètres de gestion (pour  $\tau > 2,5$ ) et de la taille initiale de la population. Nous avons souligné l'intérêt des modèles mathématiques pour mieux comprendre les déterminants essentiels de la croissance et de la propagation des espèces exotiques envahissantes et donc d'aider à cibler les actions de gestion dans Lavallée et al. (2019a). En particulier, nous avons montré dans Lavallée et al. (2019b) que le fait de faucher une seule fois par an ne suffit pas pour réduire la taille et l'aire des taches. En effet, si le nombre de fauches par an est trop faible (inférieur à 2,5 ici), la taille de la population et l'aire augmentent au cours du projet de gestion. Nous avons également montré comment ces résultats pouvaient être utilisés par les gestionnaires. Les résultats issus des simulations indiquent la durée minimale du projet de gestion nécessaire pour parvenir à l'éradication (si elle est possible, compte tenu d'une certaine fréquence de fauche). La Figure 4.10 trace les courbes de régression quadratique pour l'aire finale moyenne en fonction de la durée du projet de gestion ( $T$ ), obtenues pour différentes valeurs de  $\tau$  et pour des valeurs fixées de *initialPopSize* et *proportionMowing*. La figure indique le bénéfice potentiel, en termes de réduction de l'aire envahie, à faucher la tache une fois de plus chaque année. Plus généralement, les Équations (4.19) et (4.20), donnent la relation entre les sorties finales et les paramètres de gestion ainsi que la taille initiale de la population, ce qui permet de répondre aux questions sur l'efficacité des différentes stratégies de fauches.

Comme dans Smith et al. (2007) et Dauer and Jongejans (2013), nous avons supposé que l'invasion se déroule sur un terrain homogène. Les modèles qui prennent en compte l'inhomogénéité de la zone envahie sont souvent des modèles statiques (sans composante temporelle (Buchadas et al., 2017; Hui and Richardson, 2017; Lookingbill et al., 2014)). Lookingbill et al.

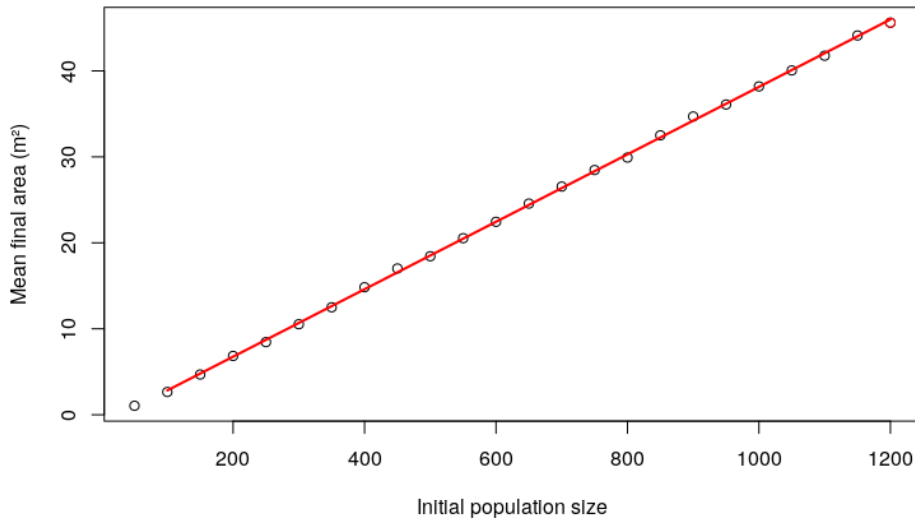


FIGURE 4.15 – Régression linéaire de l'aire finale moyenne en fonction de la taille population initiale avec  $T = 8 \text{ ans}$  et  $\tau = 4 \text{ an}^{-1}$ , extrait de Lavallée et al. (2019b).

(2014) a utilisé des indices tels que l'adéquation de l'habitat (*habitat suitability*), construits à partir de données de terrain telles que l'humidité ou le type de sol, pour produire des cartes d'invasibilité. Ces cartes attribuent à chaque zone un score qui décrit sa probabilité d'être envahie. L'invasibilité dans les réseaux écologiques peut également être liée aux caractéristiques des plantes (Hui et al., 2016). Une autre simplification dans le modèle a été de ne pas prendre en compte la dispersion des fragments de rhizome due à la fauche. Cela peut être un moyen important de propagation de la plante dans certaines conditions et cela contribue à son caractère envahissant (Sásik and Eliás, 2006). La dispersion doit être prise en compte si l'on veut modéliser l'invasion de renouées asiatiques à l'échelle d'une région composée de plusieurs taches, ce que nous faisons dans la Section 5. L'importance de l'intégration (transfert de biomasse entre les couronnes) est toujours en débat. Dans Price et al. (2002), les auteurs notent qu'il y a relativement peu d'intégration, alors que dans Suzuki (1994) les auteurs ont trouvé une plus grande importance de ce phénomène. Nous n'avons pas pris en compte ce processus dans le modèle. L'ajout de ce phénomène pourrait contribuer à produire des résultats de simulation plus proches de la réalité.

Les données de terrain que nous avons utilisées pour la calibration du modèle ont été extraites de Martin et al. (2018). Les mesures ont été prises en 2008 et 2015 sur des taches qui ont été fauchées ou pas. Pour chaque tache, les données fournissent des informations sur la stratégie de gestion utilisée par le propriétaire du terrain : la fréquence de la fauche et une estimation de la proportion de la tache fauchée. Les résultats de la calibration montrent que certains paramètres de la dynamique de la plante basés sur ces mesures sont en accord avec des données issues de la littérature.

Le modèle est écrit dans le formalisme des processus stochastiques à valeur mesure. Les outils que nous avons utilisés dans cette section dans le cas d'une tache de renouées asiatiques peuvent être utilisés par les praticiens pour tester différentes stratégies de gestion (comme le fauchage d'une partie seulement de la tache) et différents objectifs (confinement ou éradication). Ces outils peuvent de plus être utilisés dans un contexte plus général. En particulier, on pourrait utiliser le formalisme des processus stochastiques à valeur mesure pour modéliser la

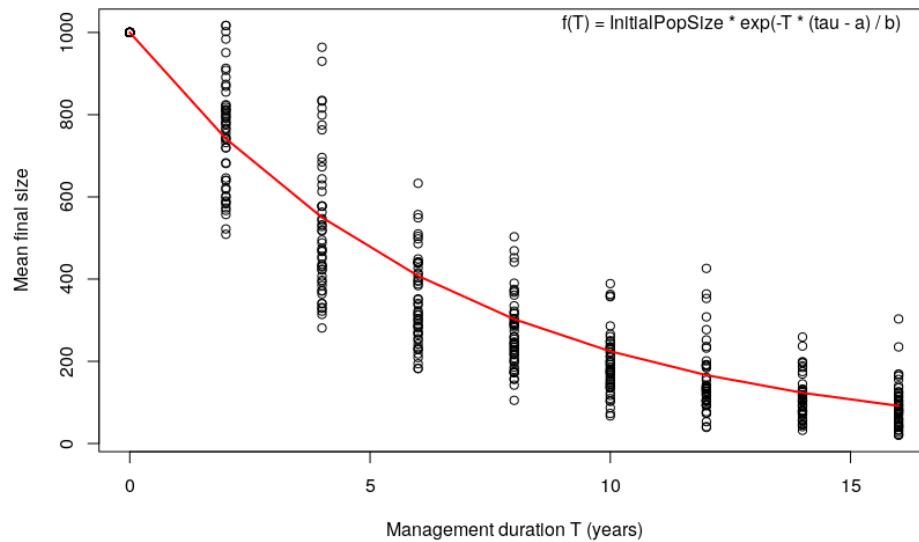


FIGURE 4.16 – Régression exponentielle sur la taille finale moyenne en fonction de la durée du projet de gestion ( $T$ ), extrait de Lavallée et al. (2019b). La courbe rouge est la fonction de prédiction qui dépend de  $T$  et définie par l'Équation (4.19). Les cercles noirs représentent la taille des taches résultant de 50 réplifications avec  $\tau = 4$ ,  $initialPopSize = 1000$ , en faisant varier  $T$ .

dynamique d'autres plantes envahissantes, comme des espèces qui produisent et dispersent des graines. On pourrait même considérer le déplacement d'individus dans de tels modèles. Dans Leman (2016), l'auteur a pris en compte le mouvement spatial d'individus dans un modèle dynamique de population stochastique et individu-centré. En outre, inclure dans le modèle la reproduction sexuelle des individus, comme dans Smadi et al. (2018), permettrait également de prendre en compte les espèces animales envahissantes, comme les moustiques (Juliano and Philip Lounibos, 2005) ou les chats sauvages (Baker and Bode, 2016).

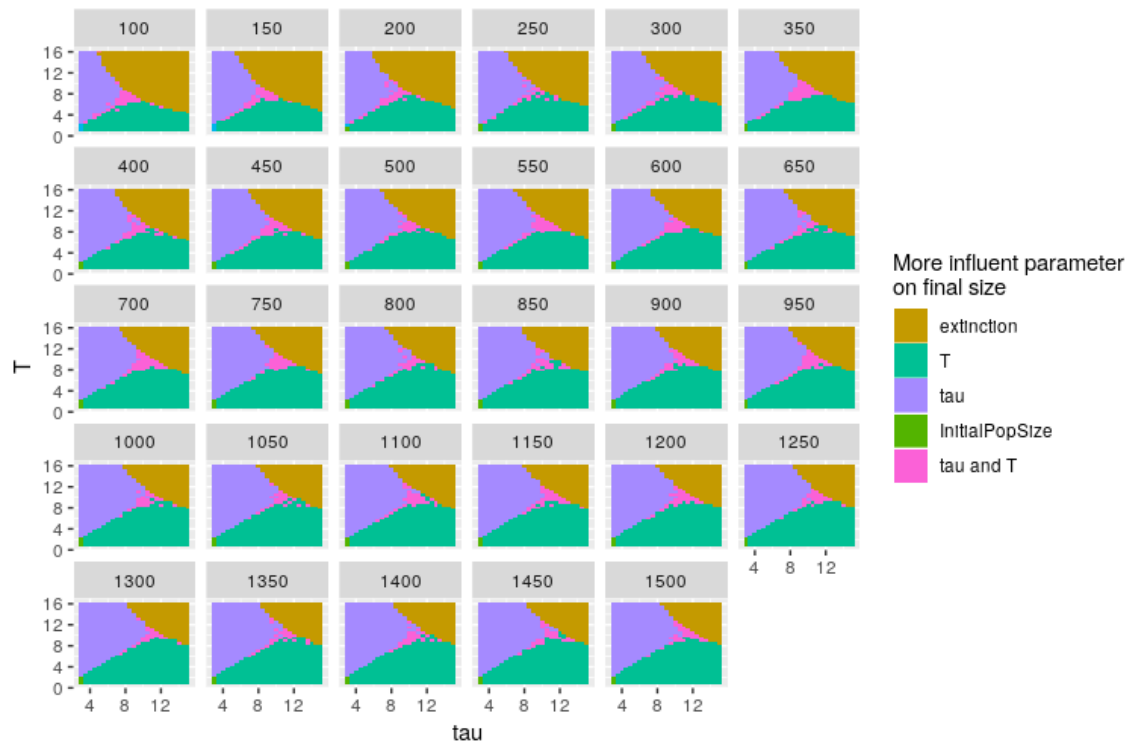


FIGURE 4.17 – Paramètres ayant la plus grande influence sur la taille finale, extrait de Lavallée et al. (2019b). Chaque graphe correspond à une valeur fixe de  $initialPopSize$  indiquée au-dessus du graphe (de 100 à 1500 couronnes),  $\tau$  varie sur l'axe des abscisses et  $T$  varie sur l'axe des ordonnées.

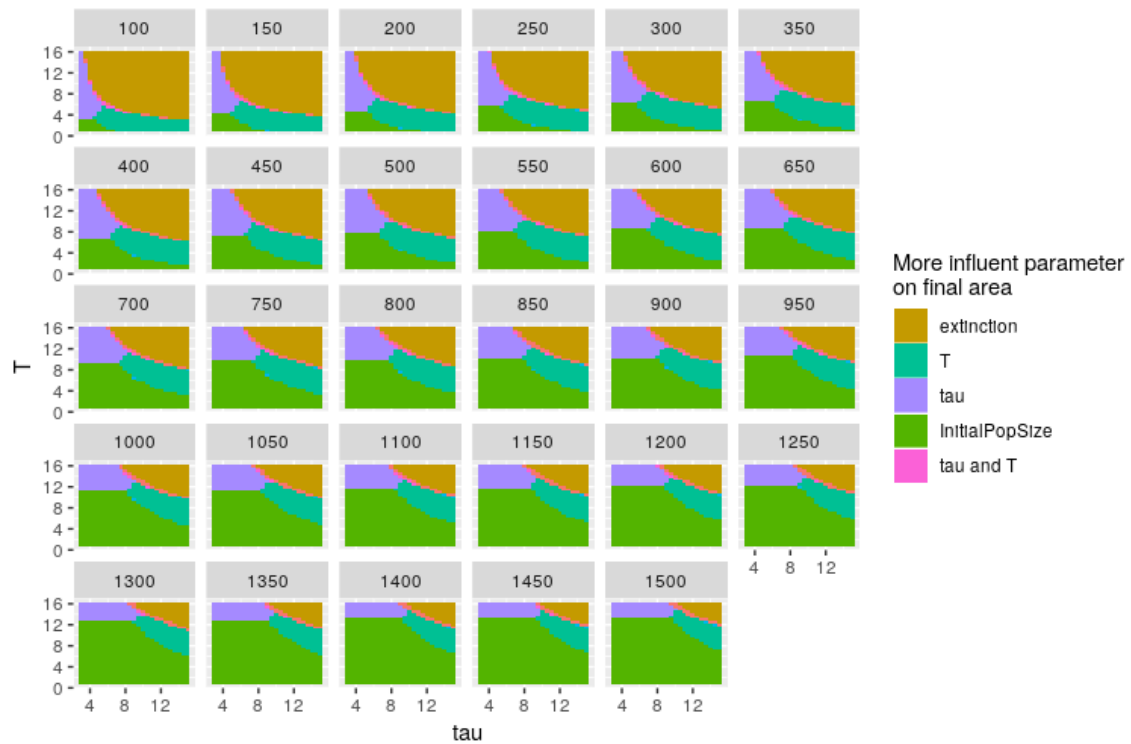


FIGURE 4.18 – Paramètres ayant la plus grande influence sur l'aire finale, extrait de Lavallée et al. (2019b). Chaque graphe correspond à une valeur fixe de *initialPopSize* indiquée au-dessus du graphe (de 100 à 1500 couronnes),  $\tau$  varie sur l'axe des abscisses et  $T$  varie sur l'axe des ordonnées.

## Chapitre

# 5

### **Modèle pour la dispersion des renouées le long d'un cours d'eau**

Ce chapitre vise à comparer l'efficacité de stratégies de gestion pour contrôler l'abondance de renouées le long d'un cours d'eau. Plus précisément, nous nous sommes intéressés à la répartition des efforts de gestion par zone qui permet de minimiser la borne supérieure de l'aire totale envahie par les renouées, au cours du temps et pour toutes les séquences de crues jugées possibles. Nous définissons les zones en lien avec la géodynamique fluviale et les conditions de terrain (facilité d'accès, même conditions environnementales pour la croissance de la plante). Nous considérons aussi qu'au sein d'une zone la même technique de gestion (fréquence de fauche) est appliquée à l'ensemble des taches de renouées présentes. Nous avons aussi analysé l'influence de la répartition spatiale de l'intensité des efforts sur l'aire totale envahie. Cependant, ce chapitre s'intéresse principalement à la méthode mise en place pour la comparaison de l'efficacité des stratégies de gestion, et moins aux résultats en eux-mêmes. En effet, nous ne disposons pas de toutes les données nécessaires pour une application numérique sur un cours d'eau particulier. Ces résultats ont donc été produits sur un cours d'eau "patchwork", qui a les caractéristiques de différentes études. En pratique, il faudrait reproduire ces études sur un même cours d'eau afin d'obtenir des informations sur sa gestion.

Pour étudier la comparaison de l'efficacité des stratégies de gestion, nous proposons un modèle de la dispersion des taches de renouées le long d'un cours d'eau qui inclut la possibilité de gestion par fauche des taches au sein de zones. Ce modèle prend en compte les propagules responsables de la dispersion de la plante, qui ont été présentées à la Section 1.6.1. Nous construisons aussi des noyaux de dispersion de ces propagules entre les zones, à l'image de ceux présentés dans la Section 1.3.2. Les crues impactent la dispersion des propagules le long du cours d'eau mais leur dynamique ne peut être prédite précisément. Toutefois, nous pouvons exprimer un ensemble de scénarios de crues considérés comme possibles et contre les conséquences desquels nous souhaitons nous prémunir dans le cadre de l'invasion par les renouées. Nous nous intéressons donc à une dynamique soumise à des contrôles (la fauche) et des incertitudes (les crues). Ce cadre correspond à celui de la théorie de la viabilité, dont nous avons introduit les concepts essentiels à la Section 3.1. Le problème de viabilité que nous

nous posons consiste à trouver des contrôles dans chaque zone qui permettent de garantir les seuils d'invasion minimaux dans chaque zone et ce pour toutes les crues considérées. La réponse à ce problème consiste à calculer un noyau de viabilité garanti. Or, la dimension du système considéré, qui dépend du nombre de zones, est trop grande pour pouvoir appliquer directement un algorithme de calcul de noyau (voir Section 3.2.1). Pour calculer le noyau garanti, nous utilisons donc la méthode de cascade de noyaux de viabilité, décrite à la Section 3.3. Cette méthode ramène l'étude du problème de viabilité initial à l'étude d'un problème de viabilité pour chaque zone, pour lequel on peut utiliser un des algorithmes de calcul du noyau présentés à la Section 3.2.1. Les bornes supérieures minimales sur les aires envahies de chaque zone (quels que soient les scénarios de crues jugés possibles et étant donné un effort total de gestion) ont été obtenues pour la dynamique du modèle calibrée sur un cours d'eau "patchwork". Ces seuils d'invasion dépendent de la répartition de cet effort de gestion parmi les différentes zones. Nous avons ainsi comparé l'aire totale envahie pour deux répartitions d'effort total constant.

Ce chapitre est organisé comme suit. La Section 5.1 est consacrée à la description des mécanismes écologiques considérés, à la présentation du modèle mathématique, ainsi qu'à l'évaluation des paramètres du modèle. Nous décrivons également le problème de gestion considéré et la réponse apportée, et présentons les méthodes utilisées pour l'étude du modèle. Les résultats sont présentés dans la Section 5.2 : d'abord les valeurs de paramètres issues de la calibration de la dynamique des plantes sur un cours d'eau "patchwork", puis les seuils d'invasion en fonction de la répartition des efforts de fauches donnés par le calcul des noyaux de viabilité garantis, ainsi que les indications de gestion mises en évidence. Enfin, nous résumons nos résultats et discutons leurs implications et les améliorations possibles dans la Section 5.3.

## 5.1 Matériels et méthodes

### 5.1.1 Mécanismes modélisés

Le modèle que nous proposons décrit la croissance de l'aire envahie par les renouées le long des berges d'un cours d'eau par lequel elles sont dispersées. On appellera berge dans la suite du document la zone sur le lit d'une rivière qui est sèche en absence de crue et inondée lors de la plus grande crue. On note  $A_i$  l'aire de la berge dans la zone  $i$  et  $L_i$  sa longueur. Ce zonage est lié à la géodynamique fluviale, mais également à des conditions de terrain (facilité d'accès, même conditions environnementales pour la croissance de la plante). Nous considérons aussi qu'au sein d'une zone la même technique de gestion est appliquée à l'ensemble des taches de renouées présentes. On suppose que les zones sont ordonnées de l'amont à l'aval, de sorte que la dispersion se fait d'une zone  $j$  vers une zone  $i \geq j$ . Le modèle a une dynamique temporelle discrète (le pas de temps est d'une année). Nous décrivons ici la dynamique de chaque variable du modèle durant un pas de temps. Les mécanismes intervenant dans le modèle sont identiques dans chaque zone, nous les décrivons pour une zone  $i$ .

La Figure 5.1 est un schéma synthétisant les processus écologiques et de gestion intervenant pendant une année dans le modèle. Les carrés bleus représentent les zones issues du découpage le long du cours d'eau. L'amont est situé à gauche, les flèches droites indiquent la dispersion des propagules. L'évolution temporelle se lit verticalement. On indique dans chaque zone certaines variables de la dynamique du modèle, on peut ainsi suivre leur évolution et les processus qui y sont impliqués durant un pas de temps. Nous ferons référence à la Figure 5.1 tout au long de la section.

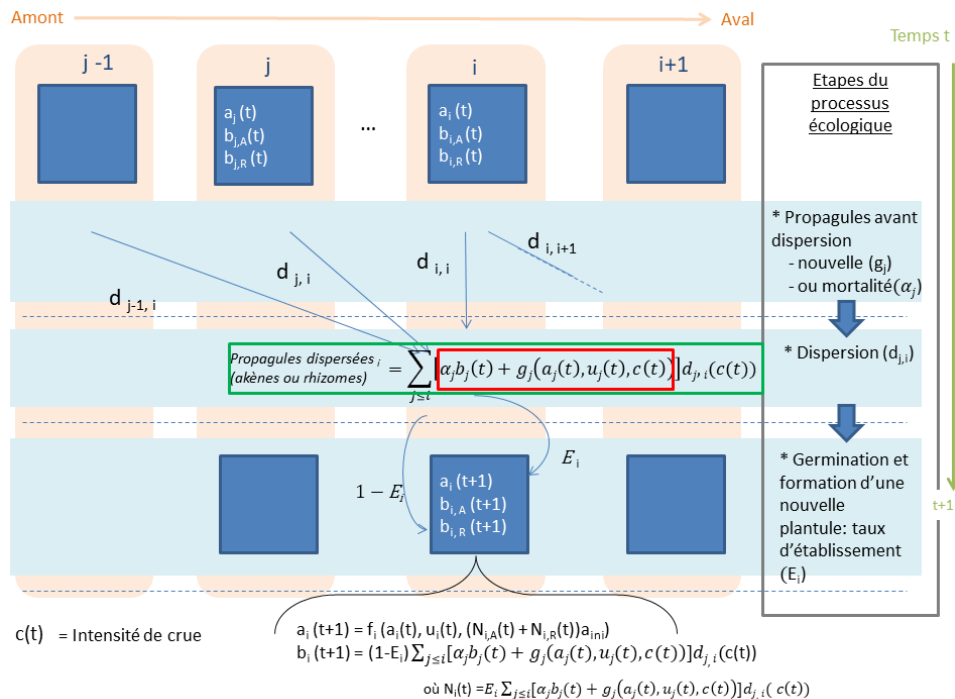


FIGURE 5.1 – Schéma synthétisant les processus écologiques intervenant pendant une année dans le modèle. La variable  $a_i$  désigne l'aire de la zone  $i$ ,  $b_{i,A}$  l'accumulateur d'akènes dans la zone  $i$ ,  $b_{i,R}$  l'accumulateur de fragments de rhizomes dans la zone  $i$ .  $N_{i,A}$  (resp.  $N_{i,R}$ ) désigne le nombre de nouvelles plantules créées dans la zone  $i$  issues d'akènes (resp. de fragments de rhizomes).



## Réservoir de crues

La dispersion des propagules dépend du régime hydraulique du cours d'eau considéré. Le modèle rend ainsi compte de l'intensité  $c(t)$  d'une crue à chaque pas de temps  $t$  et décrit l'enchaînement de crues d'intensités différentes qu'on considère possibles. Les crues sont gérées à partir d'une variable, notée  $r(t)$ , qui décrit la quantité d'eau disponible pour une crue et qui joue ainsi le rôle de "réservoir de crues" (ou d'accumulateur). Nous considérons que cette variable s'incrémente chaque année d'une quantité constante  $A$ , mais diminue lors d'une crue d'une quantité égale à l'intensité de la crue  $c(t)$  (nécessairement plus faible que le niveau du réservoir). Ainsi, pour une année donnée, plusieurs intensités de crues sont possibles. De plus, nous considérons que le réservoir de crues est borné par une quantité fixe, qui correspond à la crue la plus importante qu'on souhaite considérer, notée  $r_{max}$ .

On choisit de modéliser l'incertitude sur les crues dans le modèle en considérant la variable  $c$  comme tychastique (voir Section 3.1.2). Si à l'issue d'une crue le réservoir est peu rempli, il ne pourra pas y avoir une crue importante l'année suivante. Une crue d'intensité 0 correspond au régime du cours d'eau sans crue.

L'Équation (5.1) donne l'évolution des variables  $r$  et  $c$  :

$$\begin{cases} r(t+1) = \min(r(t) - c(t) + A, r_{max}) \\ c(t) \in [0, r(t)]. \end{cases} \quad (5.1)$$

Un scénario de crues considéré comme possible correspond ainsi à une suite  $(r(t), c(t))_{t \in \mathbb{N}}$  vérifiant l'Équation (5.1). Il est aussi possible de considérer une autre forme pour l'incrémentation du réservoir de crues, qui dépend de la valeur courante du réservoir.

## Accumulateur de propagules

Dans le modèle, on considère deux types de propagules (akènes et fragments de rhizomes (Rouifed et al. (2011b), Lamberti-Raverot (2016))) qui peuvent être dispersées d'une zone à une zone en aval (ou rester dans la même zone). Nous considérons pour chaque zone  $i$  une variable  $b_i$  (respectivement notée  $b_{i,A}$  et  $b_{i,R}$  pour les akènes et les rhizomes) qui compte chaque année les propagules dispersées par les taches et déposées sur les berges de cette zone  $i$ . Cette variable agit donc comme un accumulateur de propagules à l'échelle d'un tronçon de cours d'eau, qui s'incrémente avec les nouvelles propagules dispersées l'année courante. Les fragments de rhizomes considérés dans l'accumulateur de rhizomes ont été arrachés d'une tache en amont lors d'une crue. Pour les akènes, on compte ceux qui sont détachés de la plante et qui sont déposés sur une berge. Cette quantité peut dépendre de l'intensité d'une crue. Ainsi, une crue, selon son importance contribue à incrémenter les accumulateurs.

Nous considérons que les dynamiques des accumulateurs associés à ces deux types de propagules ont une forme analogue et ne diffèrent que par les valeurs de leurs paramètres. Nous décrivons ici l'évolution de la variable accumulateur de propagules sur un pas de temps.

Dans une zone  $j \leq i$ , avant la dispersion des propagules, on considère la mortalité des propagules présentes ainsi que la création de nouvelles propagules. On note  $\alpha_j \in [0, 1]$  ( $\alpha_{j,A}$  pour les akènes,  $\alpha_{j,R}$  pour les rhizomes) la proportion de propagules arrachées qui restent capables de produire une nouvelle tache après une année dans la zone  $j$  (ceux de la proportion  $1 - \alpha_j$  sont morts). Cette quantité est liée au taux de survie de la propagule. Nous considérons que la capacité régénérative des propagules n'est pas impactée par le temps (i.e. il n'y a pas de différence dans l'accumulateur entre une propagule qui vient d'arriver et une autre présente depuis plus de temps et encore en vie).

On note  $g_{j,A}(a, c)$  le nombre d'akènes dispersés par les taches de la zone  $j$  au cours d'une

année et susceptibles de donner une nouvelle plantule, si l'aire envahie de cette zone vaut  $a$  et qu'au cours du pas de temps la crue la plus importante qui s'est produite est d'intensité  $c$ . Cette fonction tient compte du taux de production d'akènes d'une tache, et on suppose que le nombre d'akènes dispersés sur le pas de temps considéré dépend de l'intensité de la plus grande crue sur cette période. On choisit d'exprimer cette fonction comme le produit du nombre d'akènes produits par unité d'aire ( $q_{c,A}$ ) avec l'aire de la zone envahie  $a$  et l'impact d'une crue  $c$  sur la quantité d'akènes dispersés  $k_{l_j,A,p_j}$  :

$$g_{j,A}(a, c) := a * q_{c,A} * k_{l_j,A,p_j}(c). \quad (5.2)$$

On suppose que les fonctions  $k_{l_j,A,p_j}$  sont de même type, et on les note  $k_{l,p}$  pour plus de lisibilité. On suppose de plus que  $k_{l,p}$  croît linéairement avec la proportion de surface inondée et vaut 1 lorsque cette proportion vaut 1,  $k_{l,p}$  s'exprime donc :

$$k_{l,p}(c) = l + (1 - l)p(c) \quad (5.3)$$

avec le paramètre  $l \in [0; 1]$  lié au nombre d'akènes dispersés en absence de crue et la fonction  $p$  désigne la proportion de la berge atteinte par une crue (elle vaut zéro en absence de crue et 1 pour la crue d'intensité maximale considérée). La valeur de  $l$  et la fonction  $p$  dépendent de la zone considérée.

Ainsi, le nombre d'akènes dispersés par une zone d'aire envahie  $a$ , en l'absence de crue ( $c = 0$ ) vaut  $q_{c,A} * l_{j,A} * a$ , et vaut  $q_{c,A} * a$  pour les crues dont l'impact est maximal sur la quantité d'akènes dispersés.

De même,  $g_{j,R}(a, u, c)$  désigne la quantité de fragments de rhizomes arrachés après une année (liée à l'érosion des berges ou à la gestion) dans la zone  $j$  si l'aire envahie de cette zone vaut  $a$ , que la zone est fauchée à l'intensité  $u$  (éventuellement nulle), et qu'au cours du pas de temps la crue la plus importante qui s'est produite est d'intensité  $c$ . Nous considérons que la quantité de fragments de rhizomes dispersés augmente avec l'intensité de gestion, car lors de la fauche des fragments de rhizomes peuvent être dispersés.

On choisit d'exprimer cette fonction comme la somme d'un terme dépendant de la crue et d'un autre dépendant de la gestion. Le premier est le produit de la longueur de la zone envahie en contact avec l'eau avec le nombre de fragments de rhizomes produits par unité de longueur ( $q_{c,R}$ ) et l'impact d'une crue  $c$  sur la quantité de fragments de rhizomes dispersés. On suppose cet impact de la même forme que celui sur les akènes avec la fonction  $k$  définie en (5.3) comprenant un nombre de fragments de rhizomes dispersés en l'absence de crues  $l_{j,R}$  et la proportion de berge inondée décrite par la fonction  $p_j$ . En assimilant la zone  $j$  à un rectangle et l'aire envahie à un carré tant que son côté est plus petit que la largeur de la berge et à un rectangle ensuite, on estime la longueur de la zone envahie en contact avec l'eau par

$$lg_j(a) := \begin{cases} \sqrt{a} & \text{si } \sqrt{a} \leq \frac{A_j}{L_j} \\ \frac{a}{A_j} L_j & \text{sinon} \end{cases}$$

où on rappelle que l'on note  $A_j$  l'aire de la berge dans la zone  $j$  et  $L_j$  sa longueur. On considère que des fragments de rhizomes peuvent être dispersés par la fauche et que cette quantité dépend linéairement du nombre de fauches. On a ainsi

$$g_{j,R}(a, u, c) := lg_j(a) * (q_{c,R} * k_{l_j,R,p_j}(c) + q_{u,R} * u). \quad (5.4)$$

Ainsi, la quantité de fragments de rhizomes dispersés par une zone d'aire envahie  $a$  dans le cas du régime sans crue et sans fauche ( $c = 0$  et  $u = 0$ ) vaut  $q_{c,R} * l_{j,R} * lg_j(a)$ , et vaut

$q_{c,R} * lg_j(a)$  pour les crues dont l'impact est maximal sur la quantité d'akènes dispersés. Le paramètre  $q_{u,R}$  donne la quantité de fragments de rhizomes dispersés par fauche et par unité de longueur.

La quantité d'akènes (resp. de fragments de rhizomes) de la zone  $j$  soumis à la dispersion est donc  $\alpha_{j,A}b_{j,A} + g_{j,A}(a_j, c)$  (resp.  $\alpha_{j,R}b_{j,R} + g_{j,R}(a_j, u_j, c)$ ).

Nous conservons la notation générique suivante pour évoquer l'une ou l'autre (voir case rouge du schéma de la Figure 5.1) :

$$\alpha_j b_j + g_j(a_j, u_j, c).$$

Les propagules d'une zone peuvent provenir d'autres zones situées en amont, c'est pourquoi nous considérons la matrice de dispersion de ces propagules. Cette dernière précise la proportion moyenne de propagules issues d'une tache appartenant à la zone  $j$  qui se trouvent après un pas de temps dans la zone  $i \geq j$ . En toute généralité, nous considérons que les matrices de dispersion dépendent de l'intensité  $c$  de la crue. On note  $D(c) = (d_{j,i}(c))_{(j,i) \in [1,N]^2}$  cette matrice de dispersion. Puisque la dispersion par hydrochorie se fait de l'amont vers l'aval,  $d_{j,i}(c) = 0$  si  $j > i$ . On suppose que des propagules peuvent sortir des zones de la rivière, par exemple au niveau d'un estuaire. Ainsi, la matrice  $D$  vérifie pour tout  $i$  et pour tout  $c$  :  $\sum_j d_{i,j}(c) \leq 1$ . De plus, on note  $D_A$  et  $D_R$  les matrices de dispersion respectivement pour les akènes et pour les rhizomes.

La quantité de propagules dispersées (akènes ou rhizomes) des zones  $j$  à la zone  $i$  est donc :

$$\sum_{j \leq i} [\alpha_j b_j + g_j(a_j, u_j, c)] d_{j,i}(c).$$

Cette quantité est indiquée dans la case verte sur le schéma de la Figure 5.1.

Enfin, une proportion  $E_i$  de ces propagules se trouvant dans la zone  $i$  va former de nouvelles plantules (voir la sous-section suivante), la proportion restante  $(1 - E_i)$  demeure sur place (zone de dépôt) et est à nouveau stockée dans l'accumulateur de propagules correspondant de la zone  $i$ .

Ainsi, l'évolution de la variable accumulateur  $b_i$  (une pour les akènes et une pour les fragments de rhizomes), qui indique le nombre de propagules présentes dans la zone  $i$  au temps suivant, est donnée par l'équation :

$$b_i(t+1) = (1 - E_i) \sum_{j \leq i} [\alpha_j b_j(t) + g_j(a_j(t), u_j(t), c(t))] d_{j,i}(c(t)). \quad (5.5)$$

## Aire envahie

L'aire envahie par les renouées après une année dans une zone résulte de deux phénomènes : l'apport des nouvelles plantules issues de la germination de propagules et l'évolution de l'aire des taches déjà présentes l'année précédente.

**Apport des nouvelles plantules issues de la germination de propagules :** Pour modéliser le fait que toutes les propagules présentes dans la zone  $i$  ne germent pas nécessairement (et ne forment ainsi pas toutes une plantule viable), nous considérons un paramètre qui indique le taux d'établissement d'une propagule. Ce paramètre est noté  $E_i \in \mathbb{R}_+$  ( $E_{i,A}$  pour les akènes,  $E_{i,R}$  pour les rhizomes) et sa valeur est différente pour les akènes et pour les rhizomes (pour les akènes, voir Bailey et al. (2009), Martin et al. (en préparation), Tiébré et al. (2007), Rouifed et al. (2011b), Lamberti-Raverot (2016) et pour les rhizomes, voir Bímová et al. (2003), Lamberti-Raverot (2016)). La valeur de ce paramètre est liée au taux de germination

de la propagule. Ainsi, la proportion de création peut être moins élevée si la zone est moins favorable à la création de nouvelles plantules issues de ce type de propagules.

Le nombre de nouvelles plantules créées dans la zone  $i$  vaut donc :

$$N_i := E_i \sum_{j \leq i} [\alpha_j b_j(t) + g_j(a_j(t), u_j(t), c(t))] d_{j,i}(c(t)). \quad (5.6)$$

On note  $N_{i,A}(t)$  et  $N_{i,R}(t)$  le nombre de nouvelles plantules créées dans la zone  $i$  issues respectivement d'akènes et de rhizomes. On note  $a_{ini} \in \mathbb{R}_+$  l'aire d'une nouvelle plantule qui vient d'être créée à partir de la germination d'une propagule (akène ou rhizome). Ainsi l'aire des nouvelles plantules est notée  $a_{new}$  et vaut  $(N_{i,A}(t) + N_{i,R}(t))a_{ini}$ .

**Évolution de l'aire des taches déjà présentes l'année précédente :** On utilise les formules trouvées à la Section 4.2.3 pour l'aire moyenne (le modèle était stochastique) d'une tache de renouée en fonction de la durée du projet de gestion  $T$ , de l'intensité de fauche  $u$ , et de l'aire initiale  $a_0$ . Dans le cas  $u > -d/h$  (croissance de la tache) où  $h$  est une constante négative et  $d$  une constante positive, on a trouvé (voir la Formule (4.20)) :

$$a(T) = \max((h \times u + d) * T^2 + a_0, 0).$$

Nous cherchons à exprimer cette aire à l'instant  $t + 1$  en fonction de l'aire à l'instant  $t$ . Avec une approximation au deuxième ordre, on obtient :

$$a(t+1) = \max \left( a(t) + (hu + d) * \left( 2\sqrt{\frac{\max(a_0 - a(t), 0)}{|hu + d|}} + 1 \right), 0 \right).$$

Dans le cas  $u \leq -d/h$ , on a trouvé (voir la Formule (4.18)) :

$$a(t) = e \times (-d/h - u) * t + a_0,$$

où  $e > 0$  est une constante. Donc dans ce cas,

$$a(t+1) = a(t) + e \times \left(-\frac{d}{h} - u\right).$$

Comme les formules de la section 4.2.3 on été validées sur des taches d'aire inférieure ou égale à  $60 \text{ m}^2$ , nous faisons l'hypothèse que l'aire envahie  $a$  par l'ensemble des taches de chaque zone  $i$  évolue de la même manière que  $n$  taches de moins de  $60 \text{ m}^2$  avec  $n$  qui vaut donc  $E(a/60)$ . Ainsi,  $a_{0,i}$  désigne l'aire envahie dans la zone  $i$  en début de projet de gestion. Dans l'évolution de l'aire envahie de la zone  $i$ , il faut également prendre en compte l'apport de l'aire envahie par les nouvelles plantules. De plus, l'aire envahie par les renouées est nécessairement inférieure à l'aire  $A_i$  de la zone. On note  $f_i(a, u, a_{new})$  l'aire envahie dans la zone  $i$  après une année, si l'aire envahie de cette zone vaut  $a$  l'année précédente, qu'elle est fauchée à l'intensité  $u$  (éventuellement nulle), et étant donnée que l'aire due aux nouvelles plantules est  $a_{new}$ .

Avec ce qui précède, nous obtenons l'expression de  $f_i$  :

$$f_i(a, u, a_{new}) = \begin{cases} \min(a + a_{new} + e \times \left(E\left(\frac{a}{60}\right) + 1\right) \times \left(-\frac{d}{h} - u\right), A_i) & \text{si } u \leq -\frac{d}{h} \\ \min \left( \max \left( a_{new} + a + (hu + d) \times \left(E\left(\frac{a}{60}\right) + 1\right) \times \left(2\sqrt{\frac{\max(a_{0,i} - (a + a_{new}), 0)}{|hu + d|}} + 1\right), 0 \right), A_i \right) & \text{si } u > -\frac{d}{h} \end{cases} \quad (5.7)$$

Ainsi, l'Équation (5.8) donne l'évolution de l'aire envahie dans la zone  $i$  :

$$a_i(t+1) = f_i(a_i(t), u_i(t), (N_{i,A}(t) + N_{i,R}(t))a_{ini}). \quad (5.8)$$

## Gestion

On considère qu'un gestionnaire dispose d'un effort de gestion annuel à répartir sur les différentes zones. On considère aussi qu'en début de projet de gestion, durant les  $T_{ini}$  premières années, le gestionnaire peut appliquer dans la zone  $i$  une intensité de fauche maximale supérieure ( $U_{i,max}$ ) à celle qu'il applique ensuite ( $u_{i,max}$ ). On note  $S_U$  (resp.  $S_u$ ) l'effort total de gestion mobilisé par le gestionnaire sur une année, et donc à répartir sur toutes les zones, avant  $T_{ini}$  (resp. après  $T_{ini}$ ). On a ainsi  $\sum_i U_{i,max} \leq S_U$ , et  $\sum_i u_{i,max} \leq S_u$ .

### 5.1.2 Pistes pour l'évaluation des paramètres

Dans cette section, nous commençons par résumer les paramètres du modèle présentés dans la section précédente, puis nous présentons une méthode générale pour leur donner une valeur.

#### Synthèse des paramètres du modèle

Les paramètres du modèle peuvent être soit des réels soit des fonctions, et sont de deux types : les paramètres de gestion et les paramètres pour la dynamique de l'invasion. Le Tableau 5.1 synthétise les paramètres de la dynamique de l'invasion.

| Paramètre                                     | Description   |
|---|---|
| $N$   | nombre de zones   |
| Crue  | Équation (5.1)  |
| $A$   | incrément annuel du réservoir de crues  |
| $r_{max}$                                     | intensité maximale d'une crue   |
| Propagules                                    | Équation (5.5)  |
| $\alpha_{i,A}$                                | taux de survie des akènes dans la zone $i$  |
| $\alpha_{i,R}$                                | taux de survie des fragments de rhizomes dans la zone $i$   |
| $q_{c,A}$                                     | nombre d'akènes produits par unité de surface (Équation (5.2))                                    |
| $q_{c,R}$                                     | nombre de fragments de rhizomes produits par unité de longueur (Équation (5.4))                   |
| $q_{u,R}$                                     | nombre de fragments de rhizomes dispersés par fauche et par unité de longueur (Équation (5.4))    |
| $l_{i,A}$                                     | proportion d'akènes dispersés en absence de crue par la zone $i$ (Équation (5.2))                 |
| $l_{i,R}$                                     | proportion de fragments de rhizomes dispersés en absence de crue par la zone $i$ (Équation (5.4)) |
| $p_i(c)$                                      | proportion de berge inondée dans la zone $i$ en fonction de la crue $c$ (Équation (5.3))          |
| $D_A(c) = (d_{i,j,A}(c))_{(i,j) \in [1,N]^2}$ | matrice de dispersion des akènes en fonction de la crue $c$                                       |
| $D_R(c) = (d_{i,j,R}(c))_{(i,j) \in [1,N]^2}$ | matrice de dispersion des fragments de rhizomes en fonction de la crue $c$                        |
| Aire envahie                                  | Équation (5.8)  |
| $E_{i,A}$                                     | taux d'établissement d'akènes dans la zone $i$  |
| $E_{i,R}$                                     | taux d'établissement de fragments de rhizomes dans la zone $i$                                    |
| $h$   | paramètre de la fonction $f_i$ , pour la décroissance de l'aire envahie (Équation (5.7))          |
| $d$   | paramètre de la fonction $f_i$ , pour la décroissance de l'aire envahie (Équation (5.7))          |
| $e$   | paramètre de la fonction $f_i$ , pour la croissance de l'aire envahie (Équation (5.7))            |
| $a_{ini}$                                     | aire d'une plantule nouvellement créée  |
| $a_{0,i}$                                     | aire envahie dans la zone $i$ en début de projet de gestion                                       |
| $A_i$   | aire de la zone $i$   |

TABLE 5.1 – Synthèse des paramètres pour la dynamique de l'invasion dans le modèle de dispersion

Les paramètres de gestion sont les suivants (voir Section 5.1.1) :

- $T_{ini} \in \mathbb{N}$  est la durée en début de gestion pendant laquelle on peut opérer un effort plus important que par la suite,
- $U_{i,max}$  l'intensité de fauche maximale autorisée dans la zone  $i$  avant  $T_{ini}$ ,
- $u_{i,max}$  l'intensité de fauche maximale autorisée dans la zone  $i$  après  $T_{ini}$ ,

- $S_U$  l'intensité de fauche globale (à répartir sur toutes les zones) avant  $T_{ini}$ ,
- $S_u$  l'intensité de fauche globale (à répartir sur toutes les zones) après  $T_{ini}$ .

### Valeurs des paramètres de la dynamique de l'invasion issues de la littérature ou d'autres études

**Découpage en zones ( $N$ ) :** Le modèle suppose que le bassin versant considéré soit découpé en  $N$  zones. Comme nous l'avons précisé au début de la Section 5.1.1, ces zones sont liées à la géodynamique fluviale, mais également à des conditions de terrain (facilité d'accès, même substrat pour la croissance de la plante). Nous considérons aussi qu'au sein d'une zone la même technique de gestion (fréquence) est appliquée à l'ensemble des taches de renouées présentes. De plus, nous avons supposé que les zones sont ordonnées de l'amont à l'aval, de sorte que la dispersion se fait d'une zone  $j$  vers une zone  $i \geq j$ .

Dans le modèle, la géodynamique fluviale intervient dans la dispersion des propagules au travers de noyaux de dispersion. Certains ont été présentés à la Section 1.3.2. Ils dépendent de la vitesse d'écoulement du cours d'eau, de sa largeur et de sa profondeur. Ces données peuvent être obtenues avec des mesures *in situ*, mais c'est rarement le cas. Ce type de données peut aussi être obtenu pour certains cours d'eau assez larges (plus de 50m) à partir d'informations spatiales, de divers satellites (le futur SWOT, Jason, Sentinel 1, 2 et 3, Envisat, Saral Altika, etc) et de bases de données globales (GWD-LR, GRWL, WBM, etc). Des travaux sont actuellement en cours sur des images issues de Sentinel 2 où la taille du pixel est plutôt d'environ 10 mètres. Les bases de données et logiciels tels que GWD-LR (Global Width Database - Large River), GRWL (Global River Width with Landsat), Google Earth Engine (données à diverses dates, à partir des images Landsat) permettent d'obtenir des informations sur la largeur de cours d'eau. Yamazaki et al. (2014) et Allen and Pavelsky (2018) utilisent respectivement les bases de données GWD-LR et GRWL pour avoir accès à la largeur de rivières. Des algorithmes permettent de calculer la profondeur moyenne sur des sections du cours d'eau. Par exemple, Tourian et al. (2017) estiment la profondeur de rivières à partir d'observations de même nature que celle de SWOT. Ces algorithmes peuvent nécessiter la connaissance d'une ligne d'eau et le débit associé (obtenus à partir de quelques stations in-situ). La vitesse du cours d'eau peut ensuite être déduite de données sur les débits et la géométrie du cours d'eau. Toutefois, ces informations sont essentiellement disponibles pour des cours d'eau suffisamment larges. En cas d'absence de telles données, on peut considérer le découpage basé sur les confluences des cours d'eau considérés. Celui-ci suppose que la dynamique est identique sur chaque section de cours d'eau entre deux confluences. Cependant, ce découpage peut ensuite être précisé en prenant en compte la sinuosité du cours d'eau, si cette dernière est disponible par exemple à partir de données GIS.

Les conditions de croissance de la plante peuvent être évaluées par des mesures sur le terrain. Duquette et al. (2016) attribuent le pourcentage de couvert végétal par les arbres dans des classes de valeurs. Cette étude utilise aussi des données au format GIS pour qualifier la nature du sol (anthropique, sableux,...). Descombes et al. (2016) trouvent quant à eux la nature du sol par une étude sur le terrain, ils mesurent également la couverture de la canopée et utilisent des données GIS pour les compléter. Martin et al. (2018) ont prélevé des échantillons de sol de différentes zones et analysent entre autres le pH et le ratio carbone/azote pour ensuite ne garder qu'une caractéristique du sol qui est la coordonnée sur le premier axe d'une analyse en composante principale sur ces mesures. Cette technique donne en particulier des valeurs très différentes aux sols sablonneux et rocailloux.

**Paramètres pour le réservoir de crues ( $A$  et  $r_{max}$ ) :** Les paramètres  $A$  et  $r_{max}$  peuvent être estimés à partir des données d'historiques sur les crues. Pour cela, on peut utiliser des mesures du niveau d'eau effectuées dans des stations de contrôle, comme celles de la banque Hydro (<http://www.hydro.eaufrance.fr/>).  $r_{max}$  désigne le niveau d'eau maximal atteint par une crue. L'historique des crues peut donner un ordre de grandeur de la valeur à considérer. La valeur de  $r_{max}$  est choisie en fonction du type de crue contre lequel on souhaite se prémunir (décennale, centennale). Une fois cette valeur choisie, nous décrivons la manière dont la valeur du paramètre  $A$  est fixée. Ce paramètre désigne l'incrément du niveau d'eau annuel pour former des crues. Étant donnée une valeur initiale pour le réservoir de crues  $r_0$  et une série de crues  $(c_i)_{i \in [0, n]}$ , on cherche une valeur minimale pour  $A$  telle que la valeur du réservoir de crues reste positive à chaque pas de temps (condition pour que le scénario de crues historique fasse partie de l'ensemble des scénarios envisagés comme possibles, cf Section 5.1.4) Il est possible de calculer numériquement cette valeur, qui dépend de  $r_0$  et notée  $\underline{A}(r_0)$ , et qui est nécessairement inférieure à  $\max_i(c_i)$  et décroissante. Ensuite, le choix de  $A$  dépend de l'ensemble des crues que l'on souhaite considérer, plus  $A$  est choisi grand, plus cet ensemble est grand : dès que  $A$  est pris supérieur à  $\min_{r_0}(\underline{A}(r_0)) = A(r_{max})$ , le scénario historique appartient à l'ensemble des crues considérées, si  $A$  est pris supérieur à  $\max_i(c_i)$ , le scénario avec une crue d'intensité  $\max_i(c_i)$  à tous les pas de temps est considéré comme envisageable.

**Proportion de la berge atteinte par une crue ( $p_i(c)$ ) et aire de chaque zone ( $A_i$ ) :** La fonction  $p_i(c)$  peut être estimée à partir d'images satellites, comme celles évoquées dans le paragraphe "Découpage en zones" ci-dessus. Néanmoins, si de telles données ne sont pas disponibles, il est possible d'utiliser des données sur le profil du lit d'un cours d'eau, desquelles on peut trouver la proportion de la berge atteinte par une crue. De tels profils de cours d'eau sont utilisés par exemple dans Doble et al. (2012) et Terrio and Nazimek (1997). En effet, la proportion de berge inondée peut être déduite du profil et de la hauteur d'eau. Dans le cas le plus simple où ce profil peut être résumé par une pente constante dans la zone  $i$ , notée  $\gamma_i$ , si  $c_i^{max}$  est la hauteur de la crue maximale considérée en zone  $i$ , on peut déduire la largeur de la berge :  $w_i = c_i^{max} / \gamma_i$  ainsi que son aire :  $A_i = L_i * w_i$ . On peut aussi déduire la proportion de berge inondée en zone  $i$  en fonction de la hauteur d'eau  $c_i$  :  $p_i(c_i) = c_i / c_i^{max}$ .

**Paramètres pour la matrice de dispersion ( $D$ ) :** Nous précisons l'obtention de la matrice de dispersion des propagules considérées dans le modèle. Dans la Section 1.3.2, nous avons présenté des noyaux de dispersion continus pour différentes propagules. Or, dans le modèle (Section 5.1.1), nous utilisons une matrice de dispersion pour décrire la proportion moyenne de propagules déplacées d'une zone à l'autre pendant un pas de temps. Nous précisons ici le passage de noyaux de dispersion spatialement continus à une matrice de dispersion (discrète) entre zones. Certains auteurs ont été confrontés à un problème similaire. Dans Mari et al. (2011), les auteurs présentent un modèle pour l'invasion de la moule zébrée *Dreissena polymorpha* dans le bassin du Mississippi-Missouri. Le modèle décrit le déplacement de la larve par une équation de transport (équation aux dérivées partielles avec une variable spatiale et une variable temporelle) qu'ils discrétisent sur le réseau hydraulique (nœuds et arêtes d'un graphe) pour obtenir un système d'équations différentielles ordinaires par la "method of lines", en gardant la dimension temporelle. Ainsi, en chaque point du réseau les auteurs connaissent la quantité de larves (l'équivalent des propagules dans le modèle que nous proposons). Cependant, nous ne pouvons pas appliquer cette technique dans notre cas car le modèle que nous proposons est à temps discret et la dispersion est décrite par un noyau continu et non une équation aux dérivées partielles.



Dans Chipperfield et al. (2011), les auteurs présentent et comparent l'efficacité de quatre méthodes pour approximer des noyaux continus dans des espaces discrets (centroid-to-centroid, centroid-to-area, area-to-centroid and area-to-area dispersal). Les auteurs attirent l'attention sur le fait que les erreurs commises en remplaçant le noyau continu par son approximation augmentent avec la taille des cellules. Les méthodes centroid-to-area et area-to-area ont des erreurs d'approximation plus faibles que les deux autres méthodes considérées. Cependant, dans Chipperfield et al. (2011), la dispersion est considérée homogène : le noyau de dispersion continu ne dépend pas de la zone de l'espace. Dans notre modèle, chaque zone a un noyau de dispersion qui lui est propre (qui dépend des caractéristiques de cette zone). Nous adaptons donc la méthode area-to-area à notre modèle.

Nous rappelons que  $N$  désigne le nombre de zones, et que nous notons  $L_i$  la longueur du cours d'eau dans la zone  $i$ . Les zones sont indicées par ordre croissant depuis l'amont vers l'aval. On rappelle qu'on note  $D = (d_{i,j})_{(i,j) \in [1,N]^2}$  la matrice de dispersion, qui donne la proportion moyenne de propagules dispersées entre deux zones.

On note  $n_i(y)$  le noyau de dispersion de propagules se situant dans la zone  $i$ , il donne la proportion de ces propagules qui ont parcouru une distance  $y$ .

On note  $A_{i,j}(x)$  la proportion des propagules parties de la position  $x$  dans la zone  $i$  et qui restent dans la zone  $j$  sachant qu'elles sont déjà arrivées dans la zone  $j$ .

Ainsi,

$$A_{i,i}(x) = \int_0^{L_i-x} n_i(y) dy.$$

Pour la dispersion dans les zones suivantes, on considère qu'il n'y a pas de diffuence le long du cours d'eau (division d'un cours d'eau en plusieurs branches). Par contre, il est possible que la dispersion ne se fasse pas entre toutes les zones (par exemple si la zone  $j > i$  est sur un affluent, voir le zonage réalisé à la Section 5.1.4). Pour rendre compte de cela, nous considérons une matrice  $(c_{i,j})_{(i,j) \in [1,N]^2}$ . Le terme  $c_{i,j}$  vaut 0 si la dispersion ne peut pas se faire entre les zones  $i$  et  $j$  et 1 si elle est possible. On s'intéresse au devenir de propagules parties de la zone  $i$ , nous notons  $\bar{L}_{i,j} := c_{i,j} L_j$ , de sorte que  $\bar{L}_{i,j}$  vaut 0 si les propagules issues de la zone  $i$  ne traversent pas la zone  $j$ .

Ainsi, en supposant que les propagules, parties de la position  $x$  en zone  $i$  et arrivées à l'entrée de la zone  $j$ , se dispersent dans la zone  $j$  suivant le noyau de dispersion  $n_j$  en gardant la mémoire de la distance déjà parcourue, nous obtenons pour  $j > i$  avec  $c_{i,j} = 1$  :

$$A_{i,j}(x) = \begin{cases} \frac{\int_{\bar{L}_{i,i}+\dots+\bar{L}_{i,j-1}-x}^{\bar{L}_{i,i}+\dots+\bar{L}_{i,j}-x} n_j(y) dy}{\int_{\bar{L}_{i,i}+\dots+\bar{L}_{i,j-1}-x}^{+\infty} n_j(y) dy} & \text{si } \int_{\bar{L}_{i,i}+\dots+\bar{L}_{i,j-1}-x}^{+\infty} n_j(y) dy \neq 0 \\ 1 & \text{sinon.} \end{cases} \quad (5.9)$$

On obtient alors :

$$d_{i,i}(x) = A_{i,i}(x)$$

et pour  $j > i$ ,  $d_{i,j} = 0$  si  $c_{i,j} = 0$  et si  $c_{i,j} = 1$ , de manière itérative :

$$d_{i,j}(x) = \left( 1 - \sum_{k=i}^{j-1} d_{i,k}(x) \right) A_{i,j}(x).$$

Enfin, pour tout  $j \geq i$ ,  $(i,j) \in [1,N]^2$ , l'expression du coefficient de la matrice de dispersion  $d_{i,j}$  s'en déduit par intégration :



$$d_{i,j} = \frac{1}{L_i} \int_0^{L_i} d_{i,j}(x) dx.$$

Nous pouvons remarquer que l'hypothèse sur laquelle nous avons basé notre méthode de discrétisation implique de retrouver les formules attendues en cas de noyaux de dispersion identiques dans toutes des zones. En effet, si pour tout  $i$ ,  $n_i = n$ , on obtient alors par récurrence

$$d_{i,j}(x) = \int_{\bar{L}_{i,i}+\dots+\bar{L}_{i,j-1}-x}^{\bar{L}_{i,i}+\dots+\bar{L}_{i,j}-x} n(y) dy.$$

Pour la forme de  $n_i$ , on peut se référer à Riis and Sand-Jensen (2006) pour un noyau de dispersion de fragments de tiges et à Groves et al. (2009) pour un noyau de dispersion de graines. Ces deux noyaux dépendent de paramètres, on note  $\eta_i$  ( $\eta_{i,A}$  pour les akènes,  $\eta_{i,R}$  pour les rhizomes) l'ensemble des paramètres pour le noyau de la zone  $i$ . Il est possible de tenir compte d'une dépendance de la valeur du paramètre  $\eta_i$  selon la nature du régime de crues ( $c$ ), si de telles données sont disponibles.

**Paramètres pour la croissance de la renouée avec fauche ( $h$ ,  $d$ ,  $a_{0,i}$ ) et pour l'aire initiale ( $a_{ini}$ ) :** Les paramètres  $h$  et  $d$  pour la croissance de la plante avec gestion peuvent être évalués comme dans le Chapitre 4. Cela suppose d'avoir des données de croissance (aire) de taches de renouées soumises à de la fauche à deux dates, et d'en connaître l'intensité de gestion. L'aire envahie dans la zone  $i$  en début de projet de gestion,  $a_{0,i}$ , est issue de mesures d'aires. L'aire  $a_{ini}$  d'une plantule nouvellement créée après une année peut être trouvée par des observations de terrain.

**Paramètres sur la quantité de fragments de rhizomes dispersés par la fauche ( $q_{u,R}$ ) :** Pour obtenir une borne supérieure de ce paramètre, on peut utiliser une étude analogue à De Waal (2001), qui trouvent le nombre maximal de fragments de tiges issus d'une tige et pouvant donner une plantule viable. Il faudrait mener une étude analogue avec des rhizomes.

### Valeurs des paramètres de la dynamique de l'invasion issues de la calibration

Pour certains paramètres, nous n'avons pas trouvé de protocole ou d'étude permettant d'estimer la valeur, nous leur attribuons alors une valeur par calibration. C'est le cas des paramètres :

- $\alpha_{i,A} \in [0, 1]$ , taux de survie des akènes dans la zone  $i$
- $\alpha_{i,R} \in [0, 1]$ , taux de survie des fragments de rhizomes dans la zone  $i$
- $q_{c,A} \in \mathbb{R}_+$ , nombre d'akènes produits par unité de surface
- $q_{c,R} \in \mathbb{R}_+$ , nombre de fragments de rhizomes produits par unité de longueur
- $l_{i,A} \in [0, 1]$ , proportion d'akènes dispersés en absence de crue par la zone  $i$
- $l_{i,R} \in [0, 1]$ , proportion de fragments de rhizomes dispersés en absence de crue par la zone  $i$
- $\eta_{i,A} \in \mathbb{R}_+$  paramètres intervenant dans la matrice de dispersion des akènes (voir le paragraphe y référant dans la Sous-Section 5.1)
- $\eta_{i,R} \in \mathbb{R}_+$  paramètres intervenant dans la matrice de dispersion des fragments de rhizomes (voir le paragraphe y référant dans la Sous-Section 5.1)
- $E_{i,A} \in [0, 1]$ , taux d'établissement d'akènes dans la zone  $i$
- $E_{i,R} \in [0, 1]$ , taux d'établissement de fragments de rhizomes dans la zone  $i$

—  $e \in \mathbb{R}_+$ , paramètre de la fonction  $f_i$  pour des faibles fréquences de fauche ( $u \leq -d/h$ ).

Ces paramètres sont évalués par calibration dans le logiciel d'exploration de modèles OpenMole (Reuillon et al., 2013), à partir de données d'évolution d'une invasion le long d'un cours d'eau et pour l'historique des crues associé, selon une méthode analogue à celle utilisée à la Section 4.1.6. Pour cela, des données sur les aires envahies par les renouées dans les zones considérées à deux dates sont nécessaires. L'inventaire à la première date sert de condition initiale dans le modèle, celui à la seconde date sert d'objectif dans la fonction à minimiser (et requiert la définition d'une distance entre les aires envahies simulées et observées, voir Section 5.1.4 pour celle choisie dans notre cas d'étude). De plus, la condition initiale sur les propagules (akènes et fragments de rhizomes) n'est pas accessible et devra être calibrée.

Le Tableau 5.2 indique pour chaque paramètre de la dynamique de l'invasion dans le modèle de dispersion s'il est soumis à la calibration (lignes grisées) ou si sa valeur est fixée à partir de données et de la littérature.

| Paramètre                     | Description   |
|-------------------------------|---|
| $N$                           | nombre de zones   |
| Crue                          | Équation (5.1)  |
| $A$                           | incrément annuel du réservoir de crues  |
| $r_{max}$                     | intensité maximale d'une crue   |
| Propagules                    | Équation (5.5)  |
| $\alpha_{i,A}$                | taux de survie des akènes dans la zone $i$  |
| $\alpha_{i,R}$                | taux de survie des fragments de rhizomes dans la zone $i$   |
| $q_{c,A}$                     | nombre d'akènes produits par unité de surface (Équation (5.2))                                    |
| $q_{c,R}$                     | nombre de fragments de rhizomes produits par unité de longueur (Équation (5.4))                   |
| $q_{u,R}$                     | nombre de fragments de rhizomes dispersés par fauche par unité de longueur (Équation (5.4))       |
| $l_{i,A}$                     | proportion d'akènes dispersés en absence de crue par la zone $i$ (Équation (5.2))                 |
| $l_{i,R}$                     | proportion de fragments de rhizomes dispersés en absence de crue par la zone $i$ (Équation (5.4)) |
| $p_i(c)$                      | proportion de berge inondée dans la zone $i$ en fonction de la crue $c$ (Équation (5.3))          |
| $\eta_{i,A} \in \mathbb{R}_+$ | paramètres intervenant dans la matrice de dispersion des akènes                                   |
| $\eta_{i,R} \in \mathbb{R}_+$ | paramètres intervenant dans la matrice de dispersion des fragments de rhizomes                    |
| Aire envahie                  | Équation (5.8)  |
| $E_{i,A}$                     | taux d'établissement d'akènes dans la zone $i$  |
| $E_{i,R}$                     | taux d'établissement de fragments de rhizomes dans la zone $i$                                    |
| $h$                           | paramètre de la fonction $f_i$ , pour la décroissance de l'aire envahie (Équation (5.7))          |
| $d$                           | paramètre de la fonction $f_i$ , pour la décroissance de l'aire envahie (Équation (5.7))          |
| $e$                           | paramètre de la fonction $f_i$ , pour la croissance de l'aire envahie (Équation (5.7))            |
| $a_{ini}$                     | aire d'une plante nouvellement créée  |
| $a_{0,i}$                     | aire envahie dans la zone $i$ en début de projet de gestion                                       |
| $A_i$                         | aire de la zone $i$   |

TABLE 5.2 – Synthèse des paramètres pour la dynamique de l'invasion dans le modèle de dispersion, les lignes grisées désignent les paramètres soumis à la calibration, les autres désignent les paramètres dont la valeur est fixée à partir de données et de la littérature.

Plusieurs choix de modélisation ont été faits de manière arbitraire par manque de données. Une étude de la sensibilité de la dynamique par rapport à ces choix permettrait d'évaluer leur impact sur les résultats obtenus.

### 5.1.3 Objectif de gestion et réponse apportée au problème

## Expression de la dynamique de l'invasion le long d'un cours d'eau dans le formalisme des inclusions différentielles et de la théorie de la viabilité

La dynamique de la variable globale réservoir de crues  $r(t)$  est donnée par l'Équation (5.1), et pour chaque zone  $i$  les dynamiques des 3 variables d'état locales,  $a_i$ ,  $b_{i,A}$  et  $b_{i,R}$  sont données par les Équations (5.5) et (5.8). De plus, comme les ensembles de contrôles admissibles dépendent du temps (voir Section 5.1.1), pour écrire le modèle dans le cadre de la Section 3.1.3, nous devons considérer le temps comme une variable globale du système, notée  $T$ . On note  $X$  l'espace d'états du système, on a  $X = \mathbb{R}^2 \times (\mathbb{R}^3)^N$ , et on note  $x = (T, r, a_1, b_{1,A}, b_{1,R}, \dots, a_N, b_{N,A}, b_{N,R}) \in X$ . Pour tout  $t \in \mathbb{N}$  et pour tout  $i \in [1, N]$ , nous obtenons donc le système (5.10) :

$$(S) \left\{ \begin{array}{ll} T(t+1) &= T(t) + 1 \\ r(t+1) &= \min(r(t) - c(t) + A, r_{max}) \\ a_i(t+1) &= f_i(a_i(t), u_i(t), (N_{i,A}(t) + N_{i,R}(t))a_{ini}) \\ b_{i,A}(t+1) &= (1 - E_{i,A}) \sum_{j \leq i} [\alpha_{j,A} b_{j,A}(t) + g_{j,A}(a_j(t), c(t))] d_{j,i,A}(c(t)) \\ b_{i,R}(t+1) &= (1 - E_{i,R}) \sum_{j \leq i} [\alpha_{j,R} b_{j,R}(t) + g_{j,R}(a_j(t), u_j(t), c(t))] d_{j,i,R}(c(t)) \\ u_i(t) &\in [0, U_{i,max}] \text{ si } T(t) \leq T_{ini}, u_i(t) \in [0, u_{i,max}] \text{ si } T(t) > T_{ini} \\ c(t) &\in [0, r(t)] \end{array} \right. \quad \text{pour } i \in [1, N] \quad (5.10)$$

où  $N_{i,A}$  et  $N_{i,R}$  ont été définis à l'Équation (5.6).

Les variables  $(u_i)_{i \in [1, N]}$  sont des contrôles, on choisit leur valeur dans un intervalle. Il s'agit d'une incertitude contingente (voir Aubin et al. (2013)). Contrairement aux contrôles  $(u_i)_{i \in [1, N]}$ , la valeur de la variable  $c(t)$  est imposée (on ne la choisit pas). Elle appartient à un intervalle et est considérée comme relevant de l'incertitude tychastique (voir Aubin et al. (2013)).

La section suivante s'intéresse à l'objectif de gestion que nous allons considérer et à sa traduction en terme de contraintes pour le système dynamique.

## Expression des objectifs de gestion sous forme de contraintes sur l'espace des états

On considère que le gestionnaire dispose d'un budget fixé par année (voir Section 5.1.1), qu'il connaît l'état initial de l'invasion et qu'il souhaite savoir comment allouer l'effort de fauche de manière à pouvoir garantir une aire totale envahie minimale après les  $T_{ini}$  premières années. Nous rappelons qu'il peut se permettre un effort plus important durant les  $T_{ini}$  premières années.

Nous proposons d'associer à une répartition de l'effort de fauche la plus petite valeur de l'aire totale envahie qui peut être garantie après les  $T_{ini}$  premières années. Ainsi, le gestionnaire pourra comparer les différentes répartitions en fonction de cette borne sur l'aire totale envahie. Pour cela, pour chaque choix de  $\{(U_{i,max}, u_{i,max})\}_{i \in [1, N]}$ , il faut déterminer pour quelles valeurs des bornes sur les aires envahies des différentes zones  $(\bar{a}_i)_{i \in [1, N]}$ , il existe des stratégies de fauche qui permettent, à partir de l'état actuel, de garantir le maintien de l'invasion en dessous de ces bornes malgré les crues après  $T_{ini}$ . Cela revient à calculer pour chaque répartition de l'effort de fauche les noyaux de viabilité garantis de la dynamique (5.10) pour toutes les contraintes sur les aires envahies des différentes zones et à retenir ceux qui contiennent l'état actuel.

Plus précisément, nous rappelons que l'espace d'états est  $X = \mathbb{R}^2 \times (\mathbb{R}^3)^N$ , et qu'on note  $x = (T, r, a_1, b_{1,A}, b_{1,R}, \dots, a_N, b_{N,A}, b_{N,R}) \in X$ . Nous considérons l'ensemble des contraintes  $L \subset X$  défini par

$$L((\bar{a}_i)_{i \in [1, N]}) := \left\{ x \in X, x \in \begin{cases} [0; T_{ini}] \times [0, r_{max}] \times \prod_{i=1}^N (\mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_+) \\ [T_{ini}; +\infty[ \times [0, r_{max}] \times \prod_{i=1}^N ([0, \bar{a}_i] \times \mathbb{R}_+ \times \mathbb{R}_+) \end{cases} \right\}, \quad (5.11)$$

où  $\bar{a}_i$  désigne la contrainte sur l'aire maximale envahie dans la zone  $i$ . Les contraintes sur l'aire de chaque zone sont imposées indépendamment et ne portent que sur l'aire envahie dans la zone à partir du temps  $T_{ini}$  (nous ne considérons pas de contrainte avant ce temps pendant lequel le gestionnaire dispose d'un moyen d'action plus important, puisque nous considérons que ce qui lui importe est la garantie sur l'aire envahie après  $T_{ini}$ ). Nous imposons à la variable temporelle et aux variables accumulateurs de propagules de rester positives, et à la variable réservoir de crues d'appartenir à l'intervalle  $[0, r_{max}]$ . Ces contraintes sont nécessairement vérifiées, du fait de la dynamique du système (5.10), pour toute trajectoire ayant une condition initiale qui les satisfait.

Le calcul d'un noyau de viabilité garanti de la dynamique (5.10) pour l'ensemble de contraintes (5.11) rassemble ainsi l'ensemble des états  $x \in X$  tels qu'il existe des contrôles (intensités de fauche dans chaque zone)  $u_i^*(x) \in [0, U_{i,max}]$  si  $T(t) \leq T_{ini}$ , et  $u_i^*(x) \in [0, u_{i,max}]$  si  $T(t) > T_{ini}$  qui vérifient les conditions d'efforts maximaux de chaque zone, tels que toute trajectoire ayant pour condition initiale un point de cet ensemble et des fonctions de contrôle égales aux  $u_i^*$ , respecte les contraintes (aire maximale envahie pour chaque zone) pour tout temps après  $T_{ini}$ , quelles que soient les perturbations (crues) jugées possibles (qui vérifient (5.1)).

Ainsi, on peut sélectionner parmi tous ces noyaux ceux auxquels la situation actuelle mesurée de l'invasion appartient. En effet, si la situation actuelle de l'invasion appartient à un tel noyau, cela signifie que les contraintes considérées pour calculer le noyau peuvent être garanties. Cela signifie qu'à partir de la situation actuelle, en suivant les fonctions de contrôles (les  $u_i^*$ ) déduites de ces noyaux, quel que soit le scénario de crues considéré comme possible (vérifiant (5.1)), l'aire envahie dans chaque zone n'excédera pas la contrainte de la zone correspondante. On peut ensuite calculer pour chacun de ces noyaux la somme des contraintes sur l'aire envahie de chaque zone. Cette opération donne pour chaque noyau l'aire totale qu'il permet de garantir. Il suffit enfin de sélectionner celui (ou ceux) dont l'aire totale est minimale. Nous disposons ainsi, en plus de la valeur minimale de l'aire qu'il est possible de garantir, la répartition de l'effort de gestion dans chaque zone qui permet de l'obtenir, c'est à dire des fonctions qui à  $\{(U_{i,max}, u_{i,max})\}_{i \in [1, N]}$  et à la situation initiale  $s_0 := (r^0, a_1^0, b_{1,A}^0, b_{1,R}^0, \dots, a_N^0, b_{N,A}^0, b_{N,R}^0)$  associent :

$$\min_{(\bar{a}_i)_{i \in [1, N]} \text{ tels que } (0, s_0) \in Guar_S(L((\bar{a}_i)_{i \in [1, N]}))} \sum_{i=1}^N \bar{a}_i \quad (5.12)$$

et

$$\operatorname{argmin}_{(\bar{a}_i)_{i \in [1, N]} \text{ tels que } (0, s_0) \in Guar_S(L((\bar{a}_i)_{i \in [1, N]}))} \sum_{i=1}^N \bar{a}_i \quad (5.13)$$

Remarque : calculer ce noyau de viabilité garanti revient à calculer le bassin de capture garanti

au temps  $T_{ini}$  pour la dynamique (5.10) privée de la variable  $T$  et autorisant un contrôle inférieur à  $U_{i,max}$ , sans contrainte et ayant pour cible le noyau de viabilité garanti calculé pour la dynamique (5.10) privée de  $T$  avec un contrôle inférieur à  $u_{i,max}$  et avec la contrainte  $[0, r_{max}] \times \prod_{i=1}^N ([0, \bar{a}_i] \times \mathbb{R}_+ \times \mathbb{R}_+)$ . Les bassins de capture de noyaux de viabilité sont par exemple utilisés dans les calculs de résilience (Martin, 2019).

### Méthode de calcul d'une borne supérieure minimale pour l'aire envahie après $T_{ini}$

La section précédente associe à la situation actuelle et une répartition de l'effort de fauche la plus petite valeur de l'aire totale envahie qui peut être garantie après les  $T_{ini}$  premières années. La méthode énoncée dans la section précédente consiste à calculer les noyaux de viabilité garantis de la dynamique (5.10) pour toutes les contraintes sur les aires envahies des différentes zones (5.11), puis de sélectionner ceux auxquels la condition initiale appartient. Or il est possible, en s'inspirant de la méthode de calcul de noyaux de viabilité garantis en cascade, de calculer les aires envahies garanties minimales par zone successivement.

Nous vérifions d'abord que le système (5.10) est un système tychastique contrôlé (voir Section 3.1.3), puis qu'il vérifie bien les conditions requises pour pouvoir appliquer la méthode de cascade.

**Système tychastique contrôlé et noyau de viabilité garanti fermé :** Nous rappelons que nous notons  $N$  le nombre de zones. Nous rappelons aussi que l'espace d'états du système (5.10) est  $X = \mathbb{R}^2 \times (\mathbb{R}^3)^N$ , et qu'on note  $x = (T, r, a_1, b_{1,A}, b_{1,R}, \dots, a_N, b_{N,A}, b_{N,R}) \in X$ .

D'après (5.10), les ensembles de contrôles admissibles sont donnés par la correspondance

$$\begin{aligned} X &\rightsquigarrow \mathbb{R}^N \\ x &\longmapsto \prod_{i=1}^N U_i(T) := \begin{cases} \prod_{i=1}^N [0; U_{i,max}] & \text{si } T \leq T_{ini} \\ \prod_{i=1}^N [0; u_{i,max}] & \text{sinon} \end{cases} \end{aligned} \quad (5.14)$$

et on note  $u = (u_1, \dots, u_N) \in \mathbb{R}^N$  un contrôle admissible.

D'après (5.10), les ensembles de perturbations possibles sont donnés par la correspondance

$$\begin{aligned} X &\rightsquigarrow \mathbb{R} \\ x &\longmapsto [0, r] \end{aligned} \quad (5.15)$$

et on note  $v \in \mathbb{R}$  une perturbation.

Enfin, d'après (5.10), la dynamique du système est donnée par la fonction

$$\begin{aligned} X \times \mathbb{R}^N \times \mathbb{R} &\longrightarrow \mathbb{R}^N \\ (x, u, v) &\longmapsto \begin{cases} T + 1 \\ r - v + A \\ f_i(a_i, u_i, v, (N_{i,A} + N_{i,R})a_{ini}) \\ (1 - E_{i,A}) \sum_{1 \leq j \leq i} [\alpha_{j,A} b_{j,A} + g_{j,A}(a_j, v)] d_{j,i,A}(v) \\ (1 - E_{i,R}) \sum_{1 \leq j \leq i} [\alpha_{j,R} b_{j,R} + g_{j,R}(a_j, u_j, v)] d_{j,i,R}(v) \end{cases} \quad \text{pour } i \in [1, N] \end{aligned} \quad (5.16)$$

où

$$N_{i,A} := E_{i,A} \sum_{1 \leq j \leq i} [\alpha_{j,A} b_{j,A} + g_{j,A}(a_j, v)] d_{j,i,A}(v),$$

et  $N_{i,R}$  a une expression analogue.

Le système d'Équation (5.10) est donc bien un système dynamique contrôlé tychastique. De plus, il vérifie les conditions pour que le noyau de viabilité garanti soit fermé (voir Théorème 3.1.11). En effet, la correspondance (5.14) qui donne les ensembles de contrôles admissibles a un graphe fermé (car  $U_{i,max} \geq u_{i,max}$ ) et une croissance linéaire (car bornée). La correspondance (5.15) qui donne les ensembles de perturbations possibles est continue. La fonction (5.16) qui donne la dynamique du système est continue dès que la fonction  $p$  introduite dans l'Équation (5.3) l'est également (ce qui sera le cas, voir l'Équation (5.20) Section 5.1.4), et les fonctions  $g_{i,A}$ ,  $g_{i,R}$ ,  $f_i$ , respectivement définies aux Équations (5.2) et (5.4) et (5.7), sont également continues.

**Système tychastique contrôlé vérifiant les hypothèses d'application de la méthode de cascade de noyaux de viabilité (voir Section 3.3.3) :**

Le système (5.10) est composé de deux variables que l'on peut qualifier de globales : le temps  $T$  et la variable du réservoir de crues  $r$ . Les variables locales de la zone  $i$  sont  $a_i$ ,  $b_{i,A}$  et  $b_{i,R}$ . Comme requis dans la Section 3.3, l'évolution des variables locales de la zone  $i$  et l'ensemble des contrôles admissibles qui influencent les variables locales de la zone  $i$  ne dépendent que de l'état des variables globales et de celui des variables locales des zones  $j \leq i$ . L'ensemble des perturbations possibles qui influencent les variables locales de la zone  $i$  ne dépend que des variables globales (ce qui est plus restrictif que la condition requise dans la Section 3.3). Enfin, l'ensemble de contraintes  $L((\bar{a}_i)_{i \in [1,N]})$  défini à l'Équation (5.11) est bien un ensemble produit comme ceux considérés à la Section 3.3, puisque chaque ensemble constituant l'ensemble produit ne dépend que des variables globales.

Le système (5.10) avec les contraintes (5.11) vérifie donc les conditions requises pour appliquer la méthode de cascade de noyaux de viabilité garantis.

**Calcul d'une borne supérieure minimale pour l'aire envahie par zone après  $T_{ini}$**  A l'étape  $i$ , on note

$$G_j := Guar_{S_j}(\mathbb{R}_+ \times [0, r_{max}] \times [0, \bar{a}_j] \times \mathbb{R}_+ \times \mathbb{R}_+)$$

pour  $j < i$  les noyaux de viabilité garantis construits aux étapes précédentes et on s'intéresse au noyau du viabilité garanti du système :

$$(S_i) \left\{ \begin{array}{ll} T(t+1) & = T(t) + 1 \\ r(t+1) & = \min(r(t) - c(t) + A, r_{max}) \\ a_i(t+1) & = f_i(a_i(t), u_i(t), (N_{i,A}(t) + N_{i,R}(t))a_{ini}) \\ b_{i,A}(t+1) & = (1 - E_{i,A}) \sum_{j \leq i} [\alpha_{j,A} b_{j,A}(t) + g_{j,A}(a_j(t), c(t))] d_{j,i,A}(c(t)) \\ b_{i,R}(t+1) & = (1 - E_{i,R}) \sum_{j \leq i} [\alpha_{j,R} b_{j,R}(t) + g_{j,R}(a_j(t), u_j(t), c(t))] d_{j,i,R}(c(t)) \\ u_i(t) & \in [0, U_{i,max}] \text{ si } T(t) \leq T_{ini}, u_i(t) \in [0, u_{i,max}] \text{ si } T(t) > T_{ini} \\ c(t) & \in [0, r(t)] \\ ((a_j(t), b_{j,A}(t), b_{j,R}(t)), u_j(t)) & \in \bar{V}_j((T(t), r(t)), G_j) \text{ pour tout } j < i \end{array} \right. \quad (5.17)$$

où pour  $(T, r) \in \mathbb{R} \times \mathbb{R}$

$$\overline{V}_j((T, r), G_j) := \{(x_j, u_j) \in \mathbb{R}^3 \times \mathbb{R} \mid x_j \in P_j((T, r), G_j) \text{ et } u_j \in U_j(T)\},$$

avec  $U_j$  la correspondance définie à l'Équation (5.14) et

$$P_j((T, r), G_j) := \{x_j \in \mathbb{R}^3 \mid ((T, r), x_j) \in G_j\},$$

avec l'ensemble de contraintes

$$\mathbb{R}_+ \times [0, r_{max}] \times [0, \overline{a}_i] \times \mathbb{R}_+ \times \mathbb{R}_+. \quad (5.18)$$

La connaissance de la globalité du sous-ensemble viable garanti construit par la méthode de cascade de noyaux nécessite les  $N$  itérations. Cependant, pour savoir si la condition initiale appartient à ce sous-ensemble viable, il faut qu'à chaque étape de la cascade la condition initiale correspondante appartienne au noyau de viabilité garanti du sous-système tychastique et contrôlé construit à cette étape (en considérant les états et contrôles des sous-systèmes précédents comme des perturbations). Nous pouvons donc arrêter le calcul à l'étape  $i$  si la condition initiale n'est pas dans le noyau garanti de l'étape  $i$ . Ainsi, pour le problème de dispersion, une condition nécessaire pour que la situation initiale  $(0, r^0, a_1^0, b_{1,A}^0, b_{1,R}^0, \dots, a_N^0, b_{N,A}^0, b_{N,R}^0)$  appartienne au sous-ensemble viable construit par la méthode de cascade de noyaux est qu'à chaque étape  $i$  la situation initiale  $(0, r^0, a_i^0, b_{i,A}^0, b_{i,R}^0)$  appartienne au noyau de viabilité garanti de la dynamique (5.17) avec les contraintes (5.18). Comme les ensembles de contraintes sont croissants (pour l'inclusion) avec  $\overline{a}_i$ , on peut déterminer à chaque étape la valeur seuil de  $\overline{a}_i$  au-dessus de laquelle la satisfaction de la contrainte est garantie après  $T_{ini}$  et en-dessous de laquelle nous ne pouvons apporter cette garantie :

$$\tilde{a}_i := \min_{\overline{a}_i \in \mathbb{R}} (\overline{a}_i \mid (0, r^0, a_i^0, b_{i,A}^0, b_{i,R}^0) \in Guar_{S_i}(\mathbb{R}_+ \times [0, r_{max}] \times [0, \overline{a}_i] \times \mathbb{R}_+ \times \mathbb{R}_+)) . \quad (5.19)$$

Finalement, après les  $N$  itérations de la méthode, on a construit  $\tilde{a}_1, \dots, \tilde{a}_N$  tels que l'ensemble

$$G := \{(T, r, a_1, b_{1,A}, b_{1,R}, \dots, a_N, b_{N,A}, b_{N,R}) \in X \mid \forall i \ (T, r, a_i, b_{i,A}, b_{i,R}) \in G_i\}$$

défini comme dans la Section 3.3.3 est un ensemble viable garanti pour la dynamique (5.10) avec la contrainte  $L((\tilde{a}_i)_{i \in [1, N]})$  (5.11) qui contient la situation initiale  $(0, r^0, a_1^0, b_{1,A}^0, b_{1,R}^0, \dots, a_N^0, b_{N,A}^0, b_{N,R}^0)$ . On a ainsi accès à un ensemble viable garanti, mais pas au noyau de viabilité garanti (voir Proposition 3.3.1), et ainsi  $\tilde{a}_i$  est une borne supérieure de l'aire envahie minimale que l'invasion n'excédera pas dans la zone  $i$ , après les  $T_{ini}$  premières années, pour toutes les crues considérées.

Le calcul de  $\sum_{i=1}^N \tilde{a}_i$  donne alors une borne supérieure de l'aire totale minimale qu'il est possible de garantir étant donnée la répartition des efforts maximaux de gestion par zone (5.12).

Une fois les valeurs de  $\tilde{a}_i$  calculées pour les répartitions qu'on souhaite comparer, il est possible de trouver la répartition qui minimise la borne de l'aire totale maximale envahie.

Le calcul de  $\tilde{a}_i$ , la valeur seuil pour l'aire du système  $i$  (définie à l'Équation (5.19)) nécessite le calcul de plusieurs noyaux de viabilité garantis, dont le nombre dépend de la précision souhaitée. La méthode retenue pour son calcul est présentée à la Section 5.1.5.

### 5.1.4 Application numérique pour l'évaluation des paramètres dans le cas d'un cours d'eau "patchwork"

Nous avons suivi la méthode décrite à la Section 5.1.2 pour évaluer les paramètres dans le cadre du Haut-bassin de l'Hérault. Cependant, pour certains paramètres, il n'y a pas d'études sur l'Hérault qui permettent d'estimer leur valeur. Dans le cadre de cette étude, nous utilisons donc les valeurs des paramètres estimés sur d'autres terrains. De ce fait, notre application numérique porte sur un cours d'eau "patchwork", qui a les caractéristiques de ces différentes études. Nous réalisons toutefois la calibration des paramètres restants à partir des données sur l'Hérault afin de montrer la faisabilité de la méthode proposée.

#### Données sur le Haut-bassin de l'Hérault

Nous disposons pour le Haut-bassin de l'Hérault des trois types de données nécessaires au découpage des berges en zones, à l'estimation des crues et à la calibration des paramètres restants (données d'abondance de renouées).

**Zones du Haut-bassin de l'Hérault :** Les données dont nous disposons pour la renouée, décrites précisément dans le paragraphe suivant, sont localisées dans le Haut-bassin de l'Hérault (terrains paléozoïques, ayant une forte pente de 5 à 2 %, [rapport final de "flux exportés de l'Hérault vers la Méditerranée et origine des masses d'eau", 2003]). Les cours d'eau prospectés sont l'Hérault, l'Arre, le Coudoulous, le Coularou, le Glepe, l'Alzon, le Rieutord et le Recodier. Cependant, certains ne comptent pas de présence de renouée (ainsi que sur leurs affluents), c'est pourquoi nous ne les incluons pas dans le modèle (car leur dynamique est telle que la zone ne peut être envahie, et ces derniers ne contribuent pas non plus à l'apport de propagules dans les autres zones).

**Données d'évolution de l'invasion :** Nous avons obtenu un jeu de données sur des renouées le long de cours d'eau récoltées par Mireille Boyer (ingénieur responsable du pôle Ripisylves & Invasives, CCEAU) et son équipe, lors d'une étude dont le maître d'œuvre est le SIVU Ganges-Le Vigan. Les données ont été mesurées au mètre, et ont donc un niveau de précision satisfaisant. Un inventaire initial a été réalisé en 2007, un second inventaire s'est déroulé en 2015. Ainsi, les rivières prospectées sont l'Hérault, l'Arre, le Coudoulous, le Coularou, le Glepe, l'Alzon, le Rieutord et le Recodier. Les données des deux inventaires comprennent les coordonnées GPS de chaque tache recensée, ainsi que d'autres informations dont certaines sont à dire d'expert (mais pas nécessairement les mêmes pour les deux inventaires). Nous mentionnons ici les plus pertinentes pour notre étude. La Figure 5.2, extraite de Barthod and Boyer (2019), présente sur une carte la progression de l'invasion des renouées asiatiques sur le Haut-bassin de l'Hérault entre les deux dates.

En 2007, pour l'inventaire initial, 498 taches de renouée ont été recensées, dont 91 sur l'Hérault. L'Arre (200), et le Coudoulous (207) étaient déjà très envahis en 2007, il n'y a donc pas eu de relevé sur ces deux cours d'eau en 2015. Le seul cours d'eau en commun entre les relevés de 2007 et 2015 est donc l'Hérault. Parmi les taches sur l'Hérault, toutes se situaient dans des zones érodables et la plupart (83) sont jugées provenir de substrat "naturel" (non issus de remblais). De plus, 89 taches de l'Hérault ont été répertoriées comme provenant d'une crue. Nous avons également accès, pour chaque tache, à son appartenance à une classe d'aire. Sur l'Hérault, on observait entre autres en 2007 : 52 taches de  $1 \text{ m}^2$ , 12 taches de  $2 \text{ m}^2$ , ainsi qu'une tache de  $60 \text{ m}^2$ , et une autre de  $100 \text{ m}^2$ .



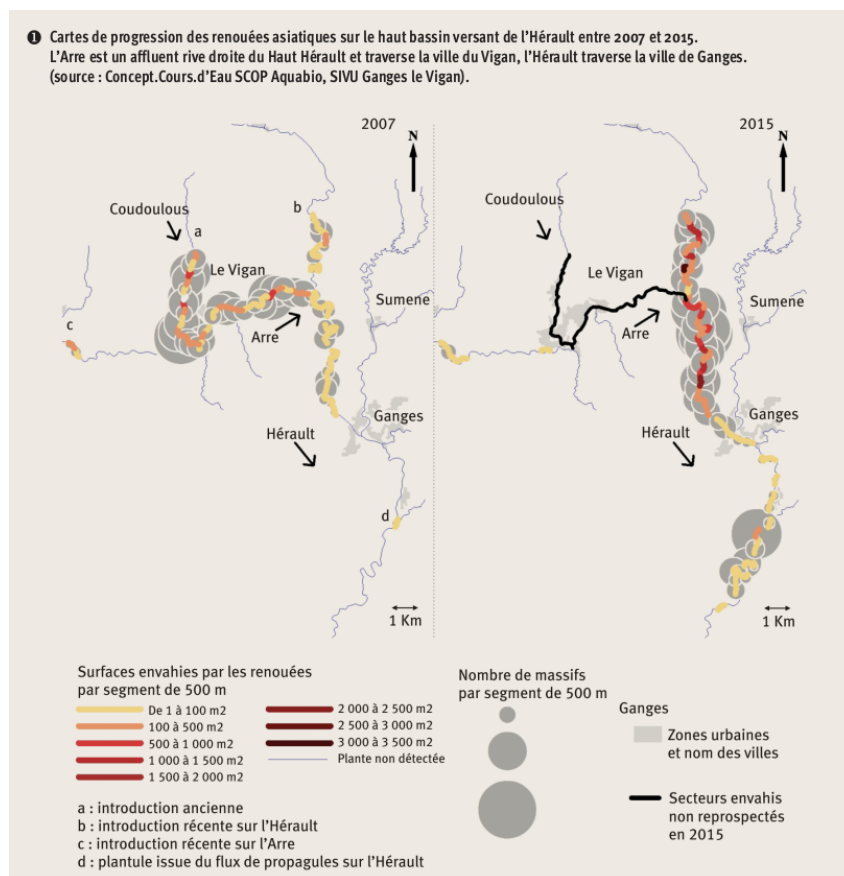


FIGURE 5.2 – Cartes de progression des renouées asiatiques sur le Haut-bassin de l'Hérault entre 2007 et 2015, figure extraite de Barthod and Boyer (2019)

En 2015, 672 taches ont été inventoriées, dont 660 sur l'Hérault (nous rappelons que l'Arre et le Coularou n'ont pas été inventoriés à nouveau, car déjà très envahis en 2007). Comme en 2007, la plupart des taches de l'Hérault se situent en zone érodable et la plupart sont encore issues de substrat "naturel".

La possibilité que la tache ait été gérée est également renseignée : la plupart des taches n'ont pas subi de gestion, certaines taches ont été recensées comme ayant subi une élimination mécanique. Cependant, Mireille Boyer nous a alerté sur la part importante d'incertitude que contient cette observation sur la gestion (de plus, aucune information sur la fréquence n'est renseignée).

**Historique des crues :** Pour connaître les crues de l'Hérault sur la période considérée, nous utilisons la station de mesure de débit de Laroque (cf <http://www.hydro.eaufrance.fr/>). En particulier, nous utilisons les données concernant la hauteur en fonction du temps. Les données sont présentées sur la Figure 5.3, il s'agit de mesures presque quotidiennes.

Barthod and Boyer (2019) signalent deux crues importantes sur la période considérée (l'une en novembre 2011 et l'autre en septembre 2014). Nous constatons que les deux niveaux d'eau maximum sur toute la période considérée sont effectivement atteints à ces dates (zones coloriées en rouge et bleu sur la Figure 5.3).

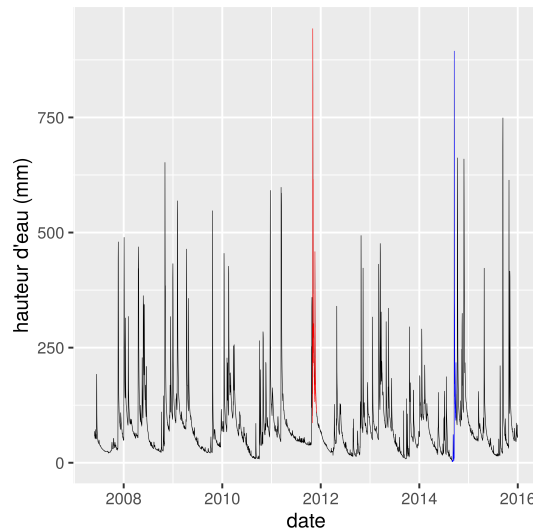


FIGURE 5.3 – Evolution de la hauteur d'eau à la station de Laroque entre 2007 et 2015 par jour. Les zones colorées correspondent aux périodes de crues importantes identifiées dans Barthod and Boyer (2019)

### Valeurs des paramètres de la dynamique d'invasion fixées à partir de données ou d'études

Dans cette section, nous précisons les valeurs des paramètres de la dynamique de l'invasion choisies à partir de données concernant le Haut-bassin de l'Hérault, et celles issues de la littérature que nous avons dû emprunter à d'autres cas d'études car non disponibles dans le cas de l'Hérault.

**Découpage en zones ( $N$ ) :** Nous ne disposons pas de données sur la géodynamique fluviale et les types de sol. Nous considérons de ce fait que dans notre zone d'étude la qualité du terrain et la dynamique fluviale ne varient pas au sein d'un même cours d'eau. Ainsi, le zonage du Haut-bassin de l'Hérault sur la partie qui nous intéresse est basé sur les confluences des cours d'eau. Le zonage ainsi réalisé donne  $N = 7$  zones. La Figure 5.4 montre le Haut-bassin de l'Hérault sur la zone prospectée, avec les points de confluence. La Figure 5.5 montre le graphe du découpage du Haut-bassin de l'Hérault en zones sur notre aire d'étude : les affluents qui ne font pas partie de zones sont des cours d'eau où les renouées ne sont pas présentes ni dans ces cours d'eau ni dans leurs affluents, si bien qu'ils ne peuvent recevoir de propagules du fait de la dynamique fluviale.



FIGURE 5.4 – Points de confluence sur le Haut-bassin de l'Hérault dans la zone d'étude considérée.

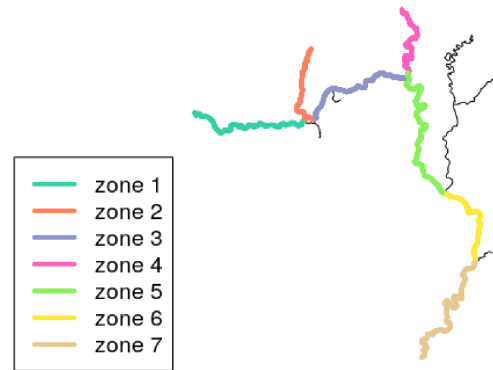


FIGURE 5.5 – Découpage du Haut-bassin de l'Hérault en zones.

**Paramètres pour le réservoir de crues ( $A$  et  $r_{max}$ ) :** Notre modèle a un pas de temps d'une année et l'intensité de crue retenue sur cette période est celle de la crue maximale (voir Section 5.1.1). Dans le Haut-bassin de l'Hérault, nous disposons des données de hauteurs d'eau de la station de Laroque entre 2008 et 2015 qui se trouve en zone 6. Cette zone notée  $i^*$  nous sert ainsi de référence et la quantité maximale du réservoir de crues  $r_{max}$  est prise égale à la hauteur d'eau maximale que nous souhaitons envisager comme possible à la station de Laroque  $c_{i^*}^{max}$ . La hauteur maximale observée est de  $9,43\text{ m}$ , nous souhaitons envisager la possibilité d'une crue plus importante encore, de 5% supérieure donc nous prenons  $r_{max} = c_{i^*}^{max} = 10\text{ m}$ .

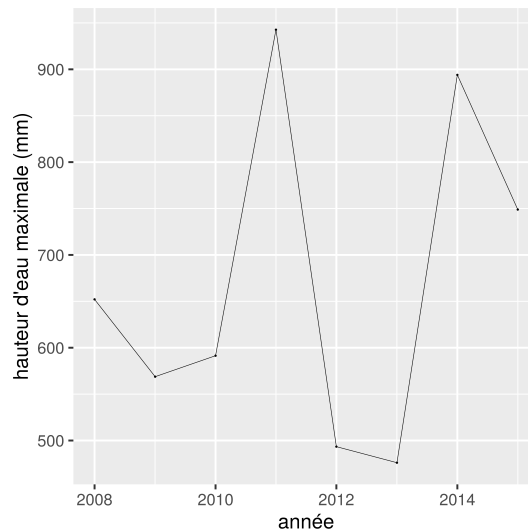


FIGURE 5.6 – Evolution de la hauteur d'eau à la station de Laroque entre 2007 et 2015 par an (valeur maximale de l'année).

Reprenant la dynamique (5.1) avec  $r_{max} = 10$  et avec  $c(t)$  valant les hauteurs d'eau de la Figure 5.6 et avec  $r(2007) = r_{max}$ , la valeur minimale de  $A$  pour que  $r(t)$  reste positive de 2007 à 2015 vaut  $\underline{A}_1(r_{max}) = 1,89\text{ m}$ . Pour que l'historique appartienne à l'ensemble des possibles, nous choisissons  $A = \underline{A}_1(r_{max}) = 1,89$ .

**Proportion de la berge atteinte par une crue ( $p_i(c)$ ) et aire  $A_i$  de chaque zone :**

Nous n'avons pas accès à des données satellites pour estimer en fonction de la hauteur d'eau la proportion de la berge inondée. Nous supposons donc que nous sommes dans le cas le plus simple où le profil de chaque zone  $i$  peut être décrit par une pente constante  $\gamma_i$ , que nous choisissons arbitrairement et qui vaut  $c_i^{max}/w_i$ , le quotient de la hauteur de la crue maximale considérée par la largeur moyenne de la berge. Nous fixons aussi arbitrairement les hauteurs d'eau maximales que l'on souhaite considérer  $c_i^{max}$  pour  $i \neq i^*$  ( $c_{i^*}^{max}$  est fixé en fonction de l'historique des crues de la station de Laroque qui appartient à la zone  $i^*$ , voir paragraphe précédent). Nous supposons également que lorsqu'une crue  $c \in [0; r_{max}]$  (avec  $r_{max} = c_{i^*}^{max}$ ) se produit dans la zone  $i^*$ , la hauteur d'eau observée en  $i$  est proportionnelle et vaut  $c/c_{i^*}^{max} \times c_i^{max}$ . Ainsi les proportions de berges inondées par une crue  $c$  valent pour toutes les zones :

$$p_i \left( \frac{c}{c_{i^*}^{max}} \times c_i^{max} \right) = \frac{c}{c_{i^*}^{max}} \times \frac{c_i^{max}}{c_{i^*}^{max}} = \frac{c}{c_{i^*}^{max}} := p(c). \quad (5.20)$$

La Tableau 5.3 indique les valeurs de longueurs, largeurs, aires et hauteurs de la crue maximale considérée pour chaque zone du Haut-bassin de l'Hérault.

| zone ( $i$ )            | 1      | 2     | 3     | 4     | 5      | 6     | 7      |
|-------------------------|--------|-------|-------|-------|--------|-------|--------|
| $L_i$ (km)              | 10.5   | 6     | 8.5   | 6.7   | 12.4   | 6.7   | 11.4   |
| $w_i$ (m)               | 10     | 10.5  | 10.5  | 11    | 11.5   | 14    | 15     |
| $A_i$ (m <sup>2</sup> ) | 105000 | 63000 | 89250 | 73700 | 142600 | 93800 | 171000 |
| $c_i^{max}$ (m)         | 7      | 7.2   | 7.5   | 8     | 8      | 10    | 10.5   |

TABLE 5.3 – Longueurs, largeurs, aires et hauteurs de la crue maximale considérée pour chaque zone du Haut-bassin de l'Hérault.

**Paramètres pour la croissance de la renouée ( $h, d$ ), pour l'aire initiale d'une nouvelle plantule ( $a_{ini}$ ) et pour l'aire envahie en début de projet de gestion ( $a_{0,i}$ ) :**

Nous considérons que la dynamique de croissance de la renouée est identique dans toutes les zones considérées. Pour la croissance de la plante avec fauche dans le cas  $u \geq -d/h$ , nous choisissons pour valeurs des paramètres  $h$  et  $d$  celles trouvées dans la Section 4, c'est à dire  $h = -0.033$  et  $d = 0.097$ . Nous notons que la croissance de la plante a été calibrée sur des données en contexte alpin, ce qui n'est pas le cas du Haut-bassin de l'Hérault. Nous avons utilisé ces valeurs faute de disposer d'autres données sur les croissances de taches avec fauche dans des terrains plus similaires à celui de l'Hérault.

Pour la valeur de l'aire initiale d'une plantule créée, nous choisissons  $a_{ini}$  arbitrairement également, faute de mesure, inférieure toutefois à la plus petite classe d'aire de tache répertoriée :  $a_{ini} = 0.5m^2$ .

Les variables  $a_{0,i}$  correspondent aux aires des zones envahies en début de projet de gestion. Dans notre application, les taches n'étant pas gérées entre 2008 et 2015, nous considérons que le projet de gestion débute en 2015. Ainsi, la valeur de l'aire envahie dans la zone  $i$  en début de projet de gestion  $a_{0,i}$  est fixée à partir des données de 2015 présentées à la Section 5.1.4 et regroupées par zone (voir Tableau 5.4 ci-dessous), et vaut pour la zone 1 puis les zones 4 à 7 : 105, 9065, 11655, 146, 701 m<sup>2</sup>. Ces données ne sont pas disponibles pour les zones 2 et 3, nous utilisons donc les aires simulées en 2015 avec le jeu de paramètres issus de la calibration.

**Paramètres sur la quantité de fragments de rhizomes dispersés par la fauche ( $q_{u,R}$ ) :**

Nous faisons l'hypothèse que la fauche n'est pas à l'origine de dispersion de fragments de

rhizomes, ainsi nous choisissons de fixer la valeur de  $q_{u,R} = 0$  pour les résultats de la section 5.2.2. Nous discuterons de l'impact du choix de la valeur de ce paramètre dans la section 5.3. Le Tableau 5.4 indique la valeur des paramètres de la dynamique de l'invasion dans le modèle de dispersion lorsqu'elle est fixée à partir de données. Les lignes grisées correspondent aux paramètres soumis à la calibration.

| Paramètre      | Valeur (par zone si c'est le cas) |       |       |       |        |       |        |       | Unité    |
|----------------|-----------------------------------|-------|-------|-------|--------|-------|--------|-------|----------|
| $N$            | 7                                 |       |       |       |        |       |        |       |          |
| $A$            | 1.89                              |       |       |       |        |       |        |       | $m$      |
| $r_{max}$      | 10                                |       |       |       |        |       |        |       | $m$      |
| $\alpha_{i,A}$ |                                   |       |       |       |        |       |        |       | $m^{-2}$ |
| $\alpha_{i,R}$ |                                   |       |       |       |        |       |        |       |          |
| $q_{c,A}$      |                                   |       |       |       |        |       |        |       | $m^{-1}$ |
| $q_{c,R}$      |                                   |       |       |       |        |       |        |       | $m^{-1}$ |
| $q_{u,R}$      | 0                                 |       |       |       |        |       |        |       | $m^{-1}$ |
| $l_{i,A}$      |                                   |       |       |       |        |       |        |       | $m^{-1}$ |
| $l_{i,R}$      |                                   |       |       |       |        |       |        |       |          |
| $p_i(c)$       | $c/10$                            |       |       |       |        |       |        |       |          |
| $\eta_{i,A}$   |                                   |       |       |       |        |       |        |       |          |
| $\eta_{i,R}$   |                                   |       |       |       |        |       |        |       |          |
| $E_{i,A}$      |                                   |       |       |       |        |       |        |       |          |
| $E_{i,R}$      |                                   |       |       |       |        |       |        |       |          |
| $h$            | -0.033                            |       |       |       |        |       |        |       |          |
| $d$            | 0.097                             |       |       |       |        |       |        |       |          |
| $e$            |                                   |       |       |       |        |       |        |       | $m^2$    |
| $a_{ini}$      | 0.5                               |       |       |       |        |       |        |       | $m^2$    |
| $a_{0,i}$      | 150                               | (*)   | (*)   | 9065  | 11655  | 146   | 701    | $m^2$ |          |
| $A_i$          | 105000                            | 63000 | 89250 | 73700 | 142600 | 93800 | 171000 | $m^2$ |          |

TABLE 5.4 – Valeurs des paramètres pour la dynamique de l'invasion dans le modèle de dispersion lorsqu'elles sont fixées à partir de données et de la littérature. Les lignes grisées désignent les paramètres soumis à la calibration. (\*) signifie que ces valeurs de  $a_{0,i}$  sont obtenues par calibration (voir Section 5.2.1).

### Valeurs des paramètres de la dynamique d'invasion fixées par calibration

Dans cette section, nous précisons d'abord des hypothèses supplémentaires sur les paramètres soumis à la calibration, puis la méthode de calibration utilisée.

**Hypothèses supplémentaires sur les paramètres soumis à la calibration** Nous supposons que les akènes et rhizomes ont des noyaux de dispersion du type  $n_i(x) = \eta_i e^{-\eta_i x}$  (voir Riis and Sand-Jensen (2006), que nous avons aussi présentés dans la Section 1.3.2), par manque des données comme la vitesse d'écoulement, la largeur et la profondeur du cours d'eau, nécessaires au paramétrage du noyau proposé par Groves et al. (2009).

Nous considérons que les rhizomes peuvent être arrachés seulement en cas de crue, cela implique que  $l_{i,R} = 0$ . De plus, nous supposons que la proportion d'akènes dispersés en absence de crue ( $l_{i,A}$ ) est identique dans chaque zone, sa valeur est notée  $l_A$ .

A dire d'expert, le taux de survie des akènes  $\alpha_{i,A}$  est très faible. Nous considérons dans cette première étude que sa valeur est nulle, ce qui permet de réduire la dimension du système dynamique. Cela revient à considérer que les akènes ne survivent qu'une seule année, et qu'au-delà de cette période ils meurent et ne sont plus en mesure de créer une nouvelle plantule. De plus, nous supposons que le taux de survie des fragments de rhizomes ( $\alpha_{i,R}$ ) est identique dans chaque zone, sa valeur est notée  $\alpha_R$ .

La calibration concerne donc les paramètres  $\alpha_R$ ,  $q_{c,A}$ ,  $q_{c,R}$ ,  $l_A$ ,  $\eta_{i,A}$ ,  $\eta_{i,R}$ ,  $E_{i,A}$ ,  $E_{i,R}$  et  $e$ .

**Méthode utilisée pour la calibration des paramètres  $\alpha_R$ ,  $q_{c,A}$ ,  $q_{c,R}$ ,  $l_A$ ,  $\eta_{i,A}$ ,  $\eta_{i,R}$ ,  $E_{i,A}$ ,  $E_{i,R}$  et  $e$  :** Les données de terrain utilisées pour la calibration du modèle sont celles du Tableau 5.5. Il présente les données d'abondance de renouées présentées à la Section 5.1.4 et regroupées par zone.

| zone | cours d'eau | longueur (km) | aire envahie 2007 ( $m^2$ ) | nb taches 2007 | aire envahie 2015 | nb taches 2015 |
|------|-------------|---------------|-----------------------------|----------------|-------------------|----------------|
| 1    | Arre        | 10.5          | 112                         | 3              | 150               | 8              |
| 2    | Coudoulous  | 6             | 6747                        | 207            | NA                | NA             |
| 3    | Arre        | 8.5           | 2800                        | 197            | NA                | NA             |
| 4    | Hérault     | 6.7           | 247                         | 28             | 9065              | 144            |
| 5    | Hérault     | 12.4          | 274                         | 62             | 11655             | 409            |
| 6    | Hérault     | 6.7           | 0                           | 0              | 146               | 36             |
| 7    | Hérault     | 11.4          | 1                           | 1              | 701               | 71             |

TABLE 5.5 – Données d'abondance de renouées sur le Haut-bassin de l'Hérault utilisées pour la calibration. Les zones sont celles de la Figure 5.5.

Elles donnent en particulier l'aire envahie par les renouées et le nombre de taches pour chaque zone du découpage du Haut-bassin de l'Hérault en 2008 et en 2015. Le modèle renseigne directement sur l'aire envahie dans chaque zone, et à partir de sa dynamique, nous pouvons calculer le nombre de nouvelles plantules créées sur la période considérée (voir Équation (5.6)). Nous ne disposons pas de données de terrain concernant le nombre de propagules sur les berges, pour comparer à la variable accumulateur de propagules, mais ce n'est pas un obstacle à la méthode de calibration employée : nous ajoutons la valeur initiale de l'accumulateur de propagules de rhizomes ( $b_{i,R}^0 \in \mathbb{R}_+$ ) dans chaque zone aux paramètres à calibrer. Pour chaque jeu de paramètres à tester, nous simulons l'aire et le nombre de nouvelles plantules créées après 7 années (en 2015), en prenant pour condition initiale dans la dynamique du modèle les aires envahies issues des données de 2008, avec une intensité de fauche nulle (nous pouvons considérer que les taches de renouées du Haut-bassin de l'Hérault n'ont pas été gérées durant cette période, voir Section 5.1.4). Nous pouvons donc comparer les aires simulées en 2015 aux données de terrain. Pour le nombre de taches, cela est moins direct puisque le modèle ne nous donne accès qu'aux nouvelles plantules et que ce qui est perçu comme une tache peut être en fait une association de plantules qui ont grandi à proximité. Le nombre de nouvelles plantules simulées doit seulement être supérieur au nombre de nouvelles taches observées sur le terrain. Nous supposons pour les données de terrain que la différence entre le nombre de taches aux deux années est uniquement due à la création de nouvelles taches (nous supposons qu'il n'y a pas eu d'éradication de taches car ces dernières n'ont pas été gérées). Notre objectif est de trouver un jeu de valeurs pour les paramètres de la dynamique de l'invasion concernés (voir Section 5.1) qui fait le mieux correspondre les sorties du modèle aux données de terrain. Ainsi, nous cherchons à minimiser la distance entre l'aire simulée et celle observée, puis celle entre le nombre de nouvelles plantules simulées et le nombre de nouvelles taches observées, sous la contrainte que cette dernière quantité soit inférieure à l'avant-dernière. Comme pour la Section 4.1.6, nous devons choisir une distance pour comparer les résultats de simulations aux observations. Pour chaque zone, nous prenons la distance :  $dist(simu, data) = |simu - data|/data$  pour comparer les aires. Nous optons pour une erreur relative (renormalisation par la donnée  $data$ ) car il y a une différence importante dans les valeurs d'aires et de nombre de taches au sein des différentes zones. Nous prenons une distance analogue entre le nombre de nouvelles plantules simulées et le nombre de taches observées, en la multipliant par une constante très importante lorsque le premier est inférieur au second afin de pénaliser énormément plus (l'algorithme ne connaît pas la valeur infini) un



TABLE 5.6 – Valeurs des paramètres pour la dynamique de l'invasion dans le modèle de dispersion lorsqu'elles sont fixées à partir de données et de la littérature. Valeurs des plages de valeurs testées pour les paramètres soumis à la calibration (lignes grisées). (\*) signifie que ces valeurs de  $a_{0,i}$  sont obtenues par calibration (voir Section 5.2.1).

Le résultat de la calibration est présenté dans la Section 5.2.1.

### 5.1.5 Application numérique pour le calcul de la valeur de l'aire totale en dessous de laquelle on a la garantie que l'aire envahie par la renouée restera après $T_{ini}$

#### Algorithme de calcul des $(\tilde{a}_i)_{i \in [1, N]}$

Nous précisons la méthode et les aspects numériques liés au calcul de la valeur de l'aire totale en dessous de laquelle on a la garantie que l'aire envahie par la renouée restera après  $T_{ini}$ . Nous avons appliqué la méthode de cascade de noyaux présentée à la Section 5.1.3, qui requiert le calcul des valeurs des aires par zone en dessous desquelles on a la garantie que l'aire envahie par la renouée restera après  $T_{ini}$ .

Ainsi, le calcul de

$$\tilde{a}_i := \min_{\bar{a}_i \in \mathbb{R}} (\bar{a}_i \mid (0, r^{ini}, a_i^{ini}, b_{i,A}^{ini}, b_{i,R}^{ini}) \in Guar_{S_i}(\mathbb{R}_+ \times [0, r_{max}] \times [0, \bar{a}_i] \times \mathbb{R}_+ \times \mathbb{R}_+)) .$$

repose sur le calcul d'un noyau de viabilité garanti  $Guar_{S_i}$ .

Pour calculer ce noyau, nous devons indiquer

- la zone  $i$  pour la dynamique
- la contrainte  $\bar{a}_i$  (voir Équation (5.18))
- les intensités de contrôle maximales  $U_{i,max}$  et  $u_{i,max}$  (voir Équation (5.17))
- l'ensemble de perturbations, qui est issu des noyaux garantis et des contrôles des zones précédentes  $j < i$ , noté  $perturbation_i$  (voir Équation (5.17)).

Nous notons  $Guar(i, \bar{a}_i, U_{i,max}, u_{i,max}, perturbation_i)$  ce noyau garanti.

L'Algorithme 1 décrit la méthode utilisée pour le calcul de  $(\tilde{a}_i)_{i \in [1, N]}$ . Il utilise la méthode de la dichotomie, dont le paramètre de précision est noté  $eps$ . On suppose connue la situation initiale dans chaque zone  $(situation\_initiale_i)_{i \in [1, N]}$ , où pour tout  $i \in [1, N]$ ,  $situation\_initiale_i = (0, r^{ini}, a_i^{ini}, b_{i,A}^{ini}, b_{i,R}^{ini})$ .

Nous utilisons le logiciel VIABLAB (voir Section 3.2.1) pour le calcul des noyaux garantis  $Guar(i, \bar{a}_i, U_{i,max}, u_{i,max}, perturbation_i)$ . Le logiciel prend en argument des paramètres et produit un fichier .dat. Il faut préciser les discrétisations sur les ensembles auxquels appartiennent les variables d'état, de contrôles et de perturbations et projeter la dynamique du modèle sur cette grille. Le fichier .dat produit donne pour chaque point de la grille issue de la discrétisation spatiale son appartenance ou non au noyau garanti. De plus, le logiciel requiert d'exprimer les ensembles auxquels appartiennent les contrôles, les contraintes et les perturbations par des expressions analytiques. Or, les ensembles de perturbations sont issus du calcul des noyaux précédents, donc sous forme de fichiers .dat. Nous choisissons de ce fait de considérer un ensemble de perturbations sous forme d'un hyper-rectangle, dont les bornes sont données par les valeurs extrémales des points appartenant au noyau pour chaque variable.

#### Expériences numériques réalisées dans le cas du cours d'eau "patchwork"

Dans notre cas d'étude de notre cours d'eau "patchwork" :



**Algorithme 1** : Calcule  $(\tilde{a}_i)_{i \in [1, N]}$ **Entrées** :  $N$ ,  $eps$ ,  $(situation\_initiale_i)_{i \in [1, N]}$ ,  $(U_{i, max})_{i \in [1, N]}$ ,  $(u_{i, max})_{i \in [1, N]}$ **Résultat** :  $(\tilde{a}_i)_{i \in [1, N]}$ **début**perturbation  $\leftarrow NULL$ **pour tous les**  $i \in [1, N]$  **faire** $m \leftarrow 0$  $M \leftarrow A_i$  $milieu \leftarrow \frac{m + M}{2}$ **tant que**  $M - m > eps$  **faire** $noyau \leftarrow \text{Guar}(i, milieu, U_{i, max}, u_{i, max}, \text{perturbation})$ **si**  $situation\_initiale_i \in noyau$  **alors** $m \leftarrow m$  $M \leftarrow milieu$ **sinon** $m \leftarrow milieu$  $M \leftarrow M$ **fin** $milieu \leftarrow \frac{m + M}{2}$ **fin**perturbation  $\leftarrow$  ajout du nouvel ensemble de perturbations issu de  $noyau$  et des intensités de contrôle maximales $\tilde{a}_i \leftarrow M$ **fin****fin**

- $N = 7$
- $\alpha_R = 0$ , de ce fait la variable accumulateur d'akènes  $b_{i, A}$  n'est plus une variable de la dynamique du système  $S_i$  (Équation (5.17)). Ainsi, le système  $S_i$  est de dimension 4 en variables d'état  $(T, r, a_i, b_{i, R})$ , 1 pour la variable de contrôle et  $\max(3 * (i - 1), 0) + 1$  pour les variables de perturbation ( $\max(2 * (i - 1), 0)$  pour les variables d'états des autres zones  $(a_1, b_{1, R}, \dots, a_{i-1}, b_{i-1, R})$ ,  $\max(i - 1, 0)$  pour le contrôle des autres zones  $(u_1, \dots, u_{i-1})$  et 1 pour les crues  $c$ )
- les valeurs des paramètres de la dynamique de l'invasion intervenant dans les  $S_i$  (5.17) sont celles du Tableau 5.6 et les valeurs des paramètres restants sont celles issues de la calibration (voir Section 5.2.1)
- les valeurs de la situation initiale sont celles de la situation en 2015 et on se place dans le pire des cas pour l'ensemble des crues à venir :  $r^{ini} = r_{max}$ ,  $a_i^{ini}$  est l'aire envahie observée en 2015 (et simulée en 2015 dans les zones 2 et 3) et  $b_{i, R}^{ini}$  est la quantité de fragments de rhizomes simulés en 2015
- nous choisissons  $eps = 1$ .

Pour tester l'influence des paramètres  $(U_{i, max}, u_{i, max})_{i \in [1, N]}$  et  $T_{ini}$  et afin de donner des indications de gestion, nous comparons différentes stratégies de gestion. Nous utilisons ainsi plusieurs valeurs de  $(U_{i, max}, u_{i, max})_{i \in [1, N]}$  et  $T_{ini}$  qui seront précisées dans les parties résultats correspondantes à la section 5.2.2.

Quant au calcul des noyaux garantis, le code VIABLAB pour le calcul du noyau garanti de la cascade de dispersion  $Guar_{S_i}(\mathbb{R}_+ \times [0, r_{max}] \times [0, \bar{a}_i] \times \mathbb{R}_+ \times \mathbb{R}_+)$  (avec  $S_i$  défini au système (5.17)) pour  $i = 1$  à 7 est reproduit à l'Annexe B et les principaux paramètres sont :

- la grille de l'espace des états : de dimension 4, le temps  $T$  varie de 0 à  $T_{ini} + 1$  avec  $T_{ini} + 2$  points, le réservoir de crues  $r$  varie de 0 à  $r_{max}$  avec 11 points, l'aire  $a_i$  varie de 0 à  $\min(2 * \max(a_i^{ini}, \bar{a}_i), A_i)$  avec 1001 points et l'accumulateur de fragments de rhizomes  $b_{i,R}$  varie de 0 à  $10 * b_{i,R}^{ini}$  avec 101 points
- la grille de l'espace des contrôles : de dimension 1, le contrôle varie de 0 à  $U_{max}$  avec 11 points si  $T \leq T_{ini}$ , et de 0 à  $u_{max}$  avec 11 points si  $T > T_{ini}$
- la grille de l'espace des perturbations : de dimension  $\max(3 * (i - 1), 0) + 1$ , la perturbation  $c$  varie de 0 à  $r_{max}$  avec 3 points, la perturbation  $u_j$  pour  $j < i$  vaut  $U_{j,max}$ , la perturbation  $a_j$  pour  $j < i$  vaut  $\max_{(T,r,a,b) \in G_j}(a)$ , la perturbation  $b_j$  pour  $j < i$  vaut  $\max_{(T,r,a,b) \in G_j}(b)$ .

Suivant la Section 5.1.3, il faudrait a priori considérer  $a_j \in [0, \max_{(T,r,a,b) \in G_j}(a)]$  et  $b_j \in [0, \max_{(T,r,a,b) \in G_j}(b)]$ . Toutefois dans le cas particulier de notre dynamique avec  $a_i(t+1)$  croissant avec  $b_{i,R}(t)$  et  $b_{i,R}(t+1)$  croissant avec  $a_j(t)$  et  $b_{j,R}(t)$ , et de notre contrainte sous forme de borne supérieure sur  $a(t)$ , si la trajectoire reste viable avec des perturbations  $a_j, b_{j,R}$  de valeurs maximales, la trajectoire avec les mêmes contrôles mais des perturbations  $a_j, b_{j,R}$ , d'intensité moindre le sera également.

Le calcul de  $(\tilde{a}_i)_{i \in [1,N]}$  pour chaque répartition de l'intensité de contrôle maximale dans chaque zone peut se faire de manière parallèle. Pour cela, nous utilisons le logiciel d'exploration de modèles OpenMOLE. L'implémentation de l'Algorithme 1 est réalisée dans ce logiciel.

## 5.2 Résultats

### 5.2.1 Calibration

Nous avons réalisé la calibration des paramètres  $\alpha_R, q_{c,A}, q_{c,R}, l_A, \eta_{i,A}, \eta_{i,R}, E_{i,A}, E_{i,R}$  et  $e$  parmi les plages de paramètres décrites dans le Tableau 5.6 à l'aide de l'algorithme NSGA2 implémenté dans le logiciel OpenMOLE (Reuillon et al., 2013). Cet algorithme fournit le jeu de paramètres donné dans le Tableau 5.7 qui a la distance minimale (aires et nombre de plantules) entre les simulations du modèle et les données observées parmi les jeux de paramètres testés par l'algorithme. Il faut ainsi noter que s'il existe plusieurs minima locaux de distance plus petite que celle trouvée par l'algorithme, il peut y avoir d'autres jeux de paramètres qui donnent de meilleurs résultats sans nécessairement être proches de ceux trouvés dans le Tableau 5.7.

Contrairement à la Section 4.2.2 pour la croissance locale, nous ne discutons pas la pertinence écologique des valeurs de paramètres obtenues, notre objectif est ici seulement de vérifier que le modèle est capable de reproduire les données. Toutefois, cet aspect sera discuté dans la Section 5.3.

La Figure 5.7 et le Tableau 5.8 comparent les aires envahies en 2015 simulées avec les valeurs de paramètres issues de la calibration et celles qui sont observées. Le modèle est donc capable de reproduire les données observées pour les aires. Cependant, on note que le nombre de plantules simulées par le modèle est très supérieur au nombre de taches observées.

La distance entre les aires vaut 0.0000206 alors que la distance entre le nombre de nouvelles plantules simulées et le nombre de nouvelles taches observées entre 2008 et 2015 vaut 1.99. Pour prendre davantage en compte cet écart dans la calibration, il faudrait avoir des connaissances supplémentaires sur le lien entre le nombre de nouvelles plantules et le nombre de

| Paramètre    | Valeur (par zone si c'est le cas) |        |       |        |        |        |        |          | Unité    |
|--------------|-----------------------------------|--------|-------|--------|--------|--------|--------|----------|----------|
| $N$          | 7                                 |        |       |        |        |        |        |          |          |
| $A$          | 1.89                              |        |       |        |        |        |        |          | $m$      |
| $r_{max}$    | 10                                |        |       |        |        |        |        |          | $m$      |
| $\alpha_A$   | 0                                 |        |       |        |        |        |        |          | $m^{-2}$ |
| $\alpha_R$   | 0.0014                            |        |       |        |        |        |        |          |          |
| $q_{c,A}$    | 5,2 e-4                           |        |       |        |        |        |        |          |          |
| $q_{c,R}$    | 18,8                              |        |       |        |        |        |        |          |          |
| $q_{u,R}$    | 0                                 |        |       |        |        |        |        |          | $m^{-1}$ |
| $l_A$        | 0.037                             |        |       |        |        |        |        |          |          |
| $l_R$        | 0                                 |        |       |        |        |        |        |          |          |
| $p_i(c)$     | $c/10$                            |        |       |        |        |        |        |          |          |
| $\eta_{i,A}$ | 0.067                             | 0.069  | 0.051 | 0.0001 | 0.1    | 0.027  | 0.061  | $m^{-1}$ |          |
| $\eta_{i,R}$ | 0.059                             | 0.009  | 0.009 | 0.028  | 0.064  | 0.011  | 0.031  | $m^{-1}$ |          |
| $E_{i,A}$    | 0.509                             | 0.990  | 0.785 | 0.398  | 0.343  | 0.267  | 0.0133 |          |          |
| $E_{i,R}$    | 0.00405                           | 0.440  | 0.311 | 0.956  | 0.994  | 0.225  | 0.994  |          |          |
| $h$          | -0.033                            |        |       |        |        |        |        |          |          |
| $d$          | 0.097                             |        |       |        |        |        |        |          |          |
| $e$          | 1.49                              |        |       |        |        |        |        |          | $m^2$    |
| $a_{ini}$    | 0.5                               |        |       |        |        |        |        |          | $m^2$    |
| $a_{0,i}$    | 150                               | 44132  | 12007 | 9065   | 11655  | 146    | 701    | $m^2$    |          |
| $A_i$        | 105000                            | 63000  | 89250 | 73700  | 142600 | 93800  | 171000 | $m^2$    |          |
| $b_{i,R}^0$  | 58812                             | 108190 | 35732 | 34988  | 7516   | 147471 | 8.54   |          |          |

TABLE 5.7 – Valeurs des paramètres pour la dynamique de l'invasion dans le modèle de dispersion. Ces valeurs sont fixées à partir de données et de la littérature pour les lignes non grisées et issues de la calibration dans le cas des lignes grisées.

nouvelles taches. Si on veut un écart plus faible entre le nombre de nouvelles plantules et le nombre de nouvelles taches, par exemple si on souhaite réduire de moitié cette distance, nous obtenons une distance pour les aires de 0.816. La Figure 5.8 et le Tableau 5.9 montrent les écarts d'aires obtenus. La différence concerne principalement la zone 4, pour laquelle une aire plus petite a été obtenue, ce qui a aussi fait diminuer le nombre de nouvelles plantules et ainsi la distance entre le nombre de nouvelles plantules et le nombre de nouvelles taches. Nous soulignons aussi que l'aire et le nombre de nouvelles plantules des zones 2 et 3 ont beaucoup changé, mais ces zones n'ayant pas été réinventoriées en 2015, ces différences n'ont pas d'influence sur les valeurs des distances.

| zone  | 1   | 2     | 3     | 4     | 5     | 6   | 7    |
|---|-----|-------|-------|-------|-------|-----|------|
| aire observée en 2015 ( $m^2$ )                           | 150 | NA    | NA    | 9065  | 11655 | 146 | 701  |
| aire simulée en 2015 ( $m^2$ )                            | 150 | 44132 | 12007 | 9065  | 11655 | 146 | 701  |
| nombre de nouvelles taches observées entre 2008 et 2015   | 5   | NA    | NA    | 144   | 409   | 36  | 71   |
| nombre de nouvelles plantules simulées entre 2008 et 2015 | 6   | 74700 | 18344 | 17566 | 22692 | 222 | 1330 |

TABLE 5.8 – Comparaison des aires envahies en 2015, observées et simulées avec les valeurs de paramètres issues de la calibration.

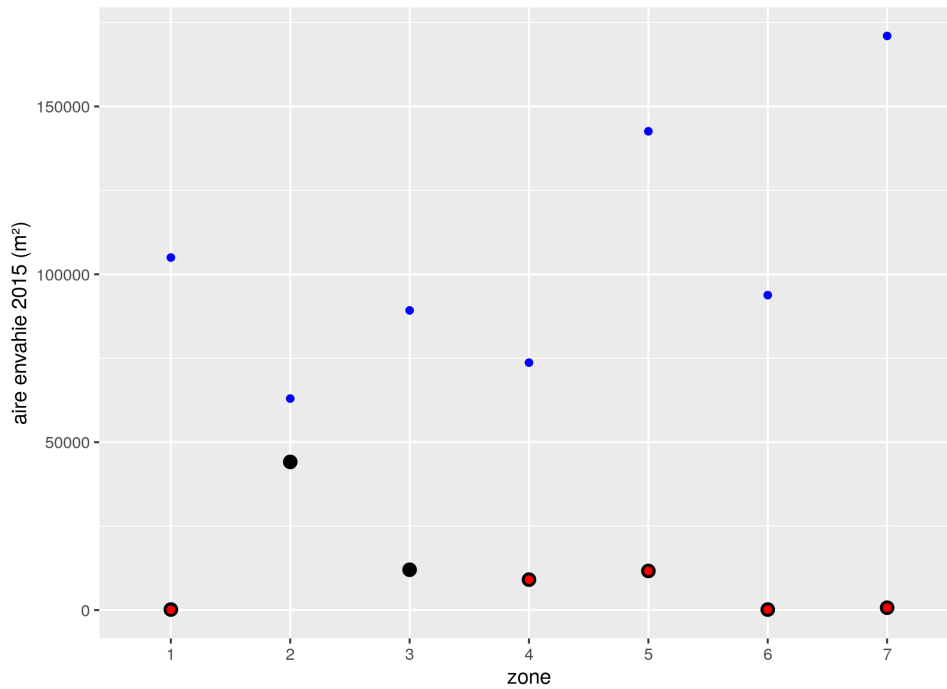


FIGURE 5.7 – Aires envahies en 2015, simulées en noir avec les valeurs de paramètres issues de la calibration, et observées en rouge. Les points bleus montrent la valeur  $A_i$  de l'aire totale de chaque zone  $i$ .

| zone  | 1   | 2       | 3    | 4    | 5     | 6   | 7    |
|---|-----|---------|------|------|-------|-----|------|
| aire observée en 2015 ( $m^2$ )                           | 150 | NA      | NA   | 9065 | 11655 | 146 | 701  |
| aire simulée en 2015 ( $m^2$ )                            | 150 | 63000   | 2878 | 1701 | 11657 | 146 | 699  |
| nombre de nouvelles taches observées entre 2008 et 2015   | 5   | NA      | NA   | 116  | 347   | 36  | 70   |
| nombre de nouvelles plantules simulées entre 2008 et 2015 | 6   | 1134154 | 86   | 2837 | 22696 | 222 | 1326 |

TABLE 5.9 – Comparaison des aires envahies en 2015, observées et simulées avec les valeurs de paramètres issues de la calibration si on impose un seuil maximal sur la distance entre les nouvelles plantules et les nouvelles taches entre 2008 et 2015.

## 5.2.2 Cascade de noyaux de viabilité

### La première étape de la cascade

Nous commençons par illustrer la première étape de la méthode de cascade de noyaux de viabilité présentée à la Section 5.1.5. Elle s'intéresse donc à la zone 1. Comme indiqué à la Section 5.1.5, nous prenons pour situation initiale  $r^{ini} = r_{max} = 10$  m,  $a_1^{ini} = 150$   $m^2$  l'aire envahie observée en 2015 dans la zone 1 et  $b_{1,R}^{ini} = 207$  la quantité de fragments de rhizomes simulés en 2015. Nous calculons dans le cas  $i = 1$  le noyau de viabilité garanti du système (5.17) avec la contrainte (5.18) et avec la valeur de  $\bar{a}_1$  qui vaut  $\bar{a}_1 = 149$   $m^2$  puis  $\bar{a}_1 = 145$   $m^2$ , avec dans les deux cas  $U_{1,max} = 8$ ,  $u_{1,max} = 4$  et  $T_{ini} = 2$ . Nous rappelons qu'en effet nous considérons des modes de gestion qui autorisent un effort plus important les  $T_{ini}$  premières années pour obtenir une garantie sur l'aire maximale envahie après  $T_{ini}$  pour un effort de fauche plus faible.

Les résultats numériques obtenus sont représentés aux Figures 5.9 - 5.13. Pour toutes ces figures, la situation initiale est indiquée par le point bleu, l'ensemble de contrainte est délimité

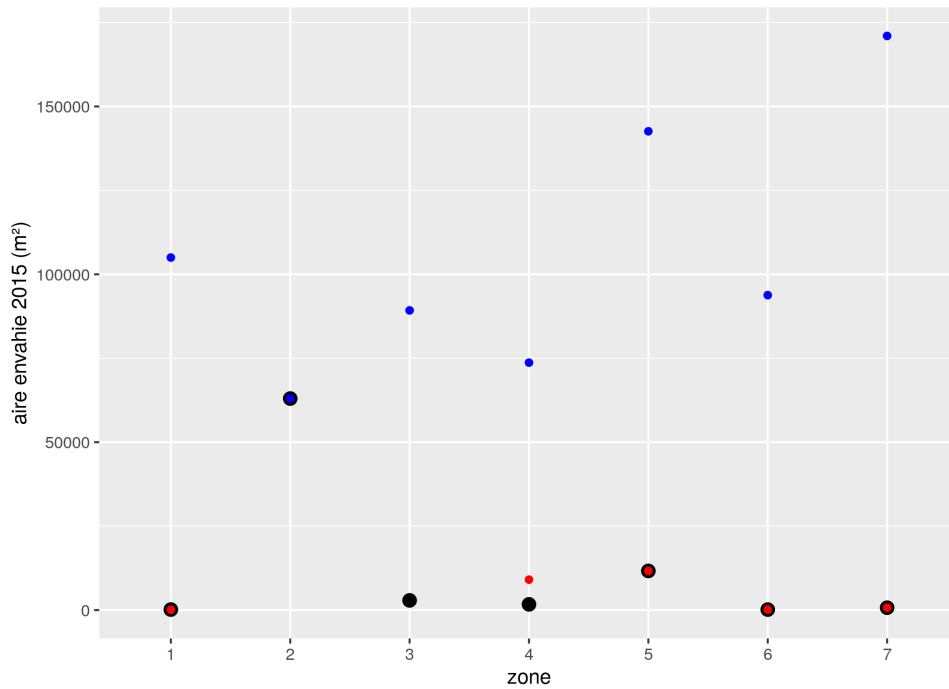


FIGURE 5.8 – Aires envahies en 2015, simulées en noir avec les valeurs de paramètres issues de la calibration si on impose un seuil maximal sur la distance entre les nouvelles plantules et les nouvelles taches entre 2008 et 2015, et observées en rouge. Les points bleus montrent la valeur  $A_i$  de l'aire totale de chaque zone  $i$ .

par la droite rouge et le pas de la discrétisation est de  $0.3 \text{ m}$  pour l'aire, et de  $20.6$  pour l'accumulateur de rhizomes.

Les Figures 5.9 et 5.10 montrent la coupe de ces noyaux selon  $T = 0$  et  $r = r_{max}$  dans les deux cas. Nous indiquons sur ces deux figures la situation initiale  $a_1^0, b_{1,R}^0$  par un point bleu. Nous constatons dans le premier cas qu'elle appartient au noyau, mais que ce n'est pas le cas pour le second noyau. Cela signifie qu'en s'autorisant un effort de fauche de 8 les trois premières années et un effort de fauche de 4 ensuite, il est possible de garantir que l'aire envahie n'excédera pas  $149 \text{ m}^2$  après  $T_{ini}$  quel que soit le scénario de crues jugé possible, mais pas de garantir que cette aire n'excédera pas  $145 \text{ m}^2$ . La valeur de  $\tilde{a}_1$  de l'aire de la zone 1 en dessous de laquelle on a la garantie que l'aire envahie par la renouée dans cette zone restera après  $T_{ini}$  (définie à l'Équation (5.19)) est donc comprise entre ces deux valeurs de  $\overline{a}_1$ .

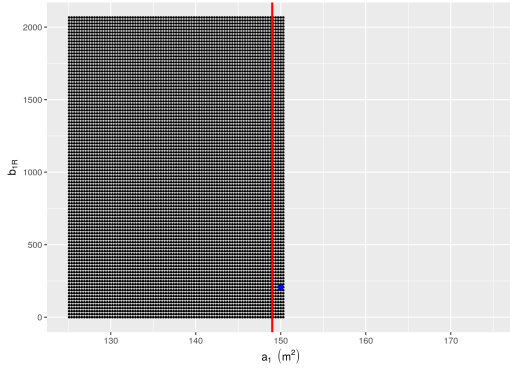


FIGURE 5.9 – Coupe du noyau de viabilité du système (5.17) avec la contrainte (5.18) selon  $T = 0$  et  $r = r_{max}$  pour  $i = 1$ ,  $\bar{a}_1 = 149 \text{ m}^2$ ,  $U_{1,max} = 8$ ,  $u_{1,max} = 4$  et  $T_{ini} = 2$ . La situation initiale est indiquée par le point bleu et l'ensemble de contrainte est délimité par la droite rouge. La situation initiale appartient à la coupe.

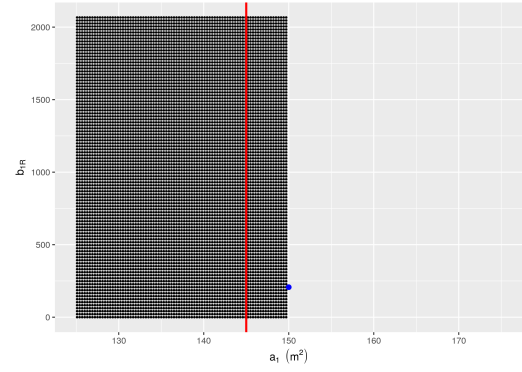


FIGURE 5.10 – Coupe du noyau de viabilité du système (5.17) avec la contrainte (5.18) selon  $T = 0$  et  $r = r_{max}$  pour  $i = 1$ ,  $\bar{a}_1 = 145 \text{ m}^2$ ,  $U_{1,max} = 8$ ,  $u_{1,max} = 4$  et  $T_{ini} = 2$ . La situation initiale est indiquée par le point bleu, et l'ensemble de contrainte est délimité par la droite rouge. La situation initiale n'appartient pas à la coupe.

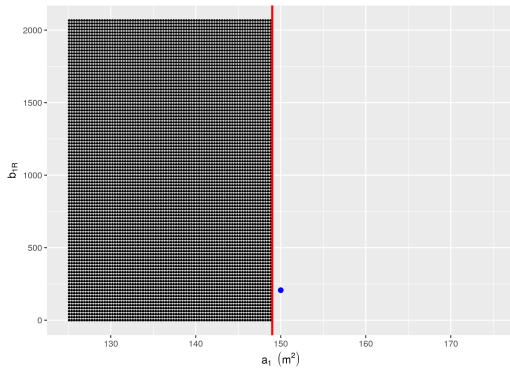


FIGURE 5.11 – Coupe du noyau de viabilité du système (5.17) avec la contrainte (5.18) selon  $T = T_{ini} + 1$  et  $r = r_{max}$  pour  $i = 1$ ,  $\bar{a}_1 = 149 \text{ m}^2$ ,  $U_{1,max} = 8$ ,  $u_{1,max} = 4$  et  $T_{ini} = 2$ . La situation initiale est indiquée par le point bleu, et l'ensemble de contrainte est délimité par la droite rouge. La situation initiale n'appartient pas à la coupe.

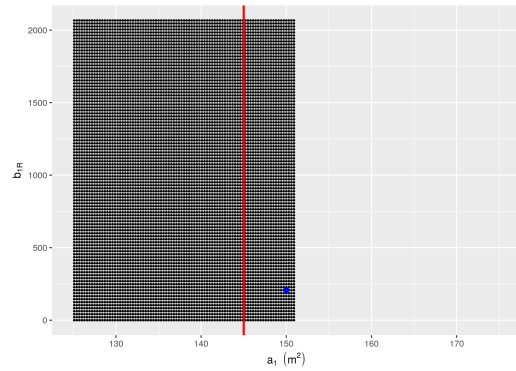


FIGURE 5.12 – Coupe du noyau de viabilité du système (5.17) avec la contrainte (5.18) selon  $T = 0$  et  $r = r_{max}$  pour  $i = 1$ ,  $\bar{a}_1 = 145 \text{ m}^2$ ,  $U_{1,max} = 15$ ,  $u_{1,max} = 4$  et  $T_{ini} = 2$ . La situation initiale est indiquée par le point bleu, et l'ensemble de contrainte est délimité par la droite rouge. La situation initiale appartient à la coupe.

La Figure 5.12 montre la coupe selon  $T = 0$  et  $r = r_{max}$  du noyau calculé avec la contrainte  $\bar{a}_1 = 145 \text{ m}^2$ , c'est-à-dire la même contrainte sur l'aire que celle considérée dans la figure 5.10, mais avec un effort de fauche  $(U_{1,max}, u_{1,max}) = (15, 4)$  plus important. Nous constatons que cette coupe contient effectivement celle représentée à la Figure 5.10 (avec  $U_{1,max} = 8$  et  $u_{1,max} = 4$ ). Nous constatons de plus que la condition initiale appartient à la coupe selon  $T = 0$  et  $r = r_{max}$  du noyau, ce qui signifie que ces plus grands efforts de fauche permettent de garantir que l'aire maximale envahie ne dépassera pas  $a_1 = 145$  après  $T_{ini}$  (alors que

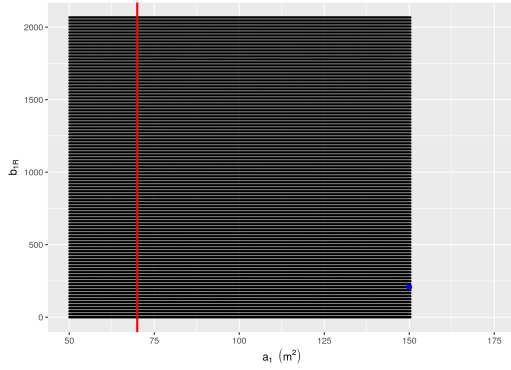


FIGURE 5.13 – Coupe du noyau de viabilité du système (5.17) avec la contrainte (5.18) selon  $T = 0$  et  $r = r_{max}$  pour  $i = 1$ ,  $\bar{a}_1 = 70 \text{ m}^2$ ,  $U_{1,max} = 8$ ,  $u_{1,max} = 4$  et  $T_{ini} = 10$ . La situation initiale est indiquée par le point bleu, et l'ensemble de contrainte est délimité par la droite rouge. La situation initiale appartient à la coupe.

$U_{1,max} = 8$  et  $u_{1,max} = 4$  ne le permettait pas, voir Figure 5.10).

De plus, la Figure 5.11 montre la coupe du noyau calculé avec la contrainte  $\bar{a}_1 = 149 \text{ m}^2$ ,  $U_{1,max} = 8$ ,  $u_{1,max} = 4$  et  $T_{ini} = 2$  selon  $T = T_{ini}$  et  $r = r_{max}$ . Nous constatons que cette coupe contient effectivement celle représentée à la Figure 5.9 (avec la coupe à  $T = 0$ ). En effet, ces coupes sont décroissantes avec  $T$  car la contrainte  $\bar{a}_1$  sur l'aire n'intervient qu'après les  $T_{ini} + 1 - T$  années, où  $T_0$  désigne la valeur pour la coupe selon  $T$ . De plus cette coupe en  $T_{ini} + 1$  représente l'ensemble des triplets  $(r_{max}, a_1, b_{1,R})$  à partir desquels l'aire envahie peut être conservée en dessous de  $\bar{a}_1$  avec un effort de fauche borné par  $u_{1,max}$  quels que soient les scénarios de crues jugés possibles. On constate ainsi qu'il faut avoir atteint certains états après  $T_{ini}$  années pour avoir la garantie de pouvoir conserver ensuite une aire envahie inférieure à  $\bar{a}_1 = 149$  avec un effort de fauche majoré par  $u_{1,max}$  quels que soient les scénarios de crues jugés possibles.

La forme rectangulaire du noyau de la Figure 5.11 vient du fait que la valeur faible de  $\alpha_R$  maintient un niveau de propagule faible et que le taux d'établissement de fragments de rhizomes dans la zone 1  $E_{1,R}$ , également faible, est à l'origine d'une contribution négligeable comparée au gain opéré par l'intensité de fauche considérée.

La Figure 5.13 montre la coupe du noyau calculé avec la contrainte  $\bar{a}_1 = 70 \text{ m}^2$ ,  $U_{1,max} = 8$ ,  $u_{1,max} = 4$  et  $T_{ini} = 10$  selon  $T = 0$  et  $r = r_{max}$ . Les efforts maximaux de fauche considérés sont identiques à ceux utilisés pour les Figures 5.9 et 5.10, la différence avec celles-ci concerne la durée  $T_{ini}$  supérieure ici. Nous constatons que la situation initiale appartient à cette coupe, ce qui signifie qu'en augmentant la durée de la période transitoire  $T_{ini}$ , une garantie bien plus stricte puisqu'elle est nécessairement inférieure à  $70 \text{ m}^2$  peut être obtenue ensuite après  $T_{ini}$  pour un même effort de fauche. Ainsi, la valeur de  $\tilde{a}_1$  est inférieure à la contrainte utilisée ( $\tilde{a}_1 \leq 70$ ) alors qu'avec  $T_{ini} = 2$  on avait  $\tilde{a}_1 \geq 145$ .

Les résultats énoncés précédemment étaient attendus : pour un même effort de fauche  $u_{1,max}$  après  $T_{ini}$ , l'aire envahie garantie minimale quel que soit le scénario de crues jugé possible décroît avec la valeur de  $U$  et avec celle  $T_{ini}$ . Cependant, ils permettent d'évaluer et de comparer l'impact de ces deux facteurs.

### Cascades complètes

Nous rappelons les valeurs des aires en début de projet de gestion et des aires totales des zones dans le Tableau 5.10.

| zone   | 1      | 2     | 3     | 4     | 5      | 6     | 7      |
|--|--------|-------|-------|-------|--------|-------|--------|
| Aire en début de projet de gestion ( $m^2$ ) | 150    | 44132 | 12007 | 9065  | 11655  | 146   | 701    |
| Aire totale de la zone ( $m^2$ )             | 105000 | 63000 | 89250 | 73700 | 142600 | 93800 | 171000 |

TABLE 5.10 – Valeurs de l'aire en début de projet de gestion et de l'aire totale par zone.

Nous décrivons maintenant les résultats obtenus par les cascades complètes de noyaux pour 2 répartitions des efforts de fauches : la première répartition concentre les efforts de fauche sur les zones faiblement envahies (voir Tableau 5.10), la seconde est homogène sur les 7 zones. La somme des efforts est identique pour les deux répartitions ( $S_U = 45$  et  $S_u = 28$ ).

| zone          |             | 1  | 2 | 3 | 4 | 5 | 6  | 7  |            |
|---------------|-------------|----|---|---|---|---|----|----|------------|
| Répartition 1 | $U_{1,max}$ | 15 | 0 | 0 | 0 | 0 | 15 | 15 | $S_U = 45$ |
|               | $u_{1,max}$ | 10 | 0 | 0 | 0 | 0 | 9  | 9  | $S_u = 28$ |
| Répartition 2 | $U_{1,max}$ | 6  | 7 | 7 | 7 | 6 | 6  | 6  | $S_U = 45$ |
|               | $u_{1,max}$ | 4  | 4 | 4 | 4 | 4 | 4  | 4  | $S_u = 28$ |

TABLE 5.11 – Répartitions des efforts entre zones utilisées dans la cascade de viabilité pour le calcul de l'aire totale envahie minimale.

En utilisant les logiciels Viablab et OpenMOLE avec les paramètres décrits à la Section 5.1.5, nous obtenons les résultats suivants :

| zone ( $i$ )  | 1   | 2     | 3     | 4     | 5      | 6    | 7     | aire garantie ( $m^2$ ) |
|---------------|-----|-------|-------|-------|--------|------|-------|-------------------------|
| Répartition 1 | 136 | 63000 | 89250 | 73700 | 142600 | 1106 | 26962 | 396754                  |
| Répartition 2 | 148 | 63000 | 89250 | 73700 | 142600 | 1280 | 26962 | 396940                  |

TABLE 5.12 – Valeurs de  $\tilde{a}_i$  dans chaque zone  $i$ , obtenues par l'application de la méthode de cascade de noyaux de viabilité, pour les répartitions des efforts décrites dans le Tableau 5.11. Nous indiquons également l'aire totale minimale garantie par chaque stratégie.

La première stratégie de gestion a permis de garantir une aire totale envahie de  $396754 m^2$ . Les zones non fauchées sont nécessairement totalement envahies (voir Tableaux 5.10 et 5.12), ce qui explique la valeur importante de cette quantité. Seule la zone 1 a une garantie inférieure à sa situation initiale. La seconde stratégie de gestion permet de garantir une aire totale envahie de  $396940 m^2$ , soit une valeur légèrement supérieure à celle de la première stratégie. Le calcul de la cascade de noyaux de viabilité nous permet ainsi de conclure qu'entre la répartition 1 et 2, c'est la 1 qui permet de garantir une borne supérieure minimale sur l'aire totale envahie, mais que la différence d'impact de ces deux stratégies de gestion en terme d'aire totale envahie minimale garantie est très faible, inférieure à 0.05%.

## 5.3 Discussion

Dans ce chapitre, nous avons proposé un modèle dans le formalisme des inclusions différentielles pour la croissance et la dispersion de renouées asiatiques le long d'un cours d'eau, incluant la fauche comme stratégie de gestion.



Nous avons ensuite calibré les paramètres de la dynamique de dispersion sur un cours d'eau "patchwork" dans la Section 5.2.1. Nous avons aussi identifié les seuils d'invasion en fonction de la répartition des efforts de fauches donnés par le calcul de noyaux de viabilité garantis. Pour illustrer notre démarche, nous avons calculé le noyau de viabilité associé à la première zone dans la Section 5.2.2. Dans la Section 5.2.2, nous avons déduit de l'application de la méthode de cascade de noyaux de viabilité des indications de gestion pour ce cours d'eau fictif. Nous avons illustré l'impact de la fréquence de fauche ( $U_{1,max}, u_{1,max}$ ) et de la durée de la première phase ( $T_{ini}$ ) sur l'aire envahie minimale garantie dans la zone 1. Nous avons ensuite comparé l'efficacité de deux répartitions d'effort à somme constante dont la différence s'avère faible.

Nous avons démontré avec ces premiers résultats sur l'aire d'invasion garantie minimale la faisabilité de la méthode de cascade de noyaux pour obtenir des informations sur les performances de stratégies de gestion spatialisées. Il y a de nombreuses perspectives à ce travail.

### 5.3.1 Données disponibles, choix de modélisation et calibration

Comme nous l'avons déjà énoncé dans les sections précédentes de ce chapitre, nous n'avons pas trouvé toutes les données nécessaires pour la calibration du modèle sur un même cours d'eau. Afin de démontrer la faisabilité de la démarche que nous proposons, nous avons opéré une application numérique sur un cours d'eau "patchwork" que nous avons défini et calibré suivant plusieurs sources. Nous disposons de données sur l'abondance de renouées le long de l'Hérault, que nous avons complété avec des données issues d'autres cours d'eau. Les indications de gestion obtenues dans ce cas ne sont donc pas applicables à l'Hérault. Nous soulignons que l'obtention d'indications de gestion pertinentes nécessite l'ensemble des données présentées à la Section 5.1.2 pour la calibration du modèle.

De plus, l'historique de crue utilisé pour l'estimation de  $r_{max}$  et  $A$  concernait la période de 2008 à 2015 (la période utilisée pour la calibration). Des données sont disponibles sur une plus grande période et permettraient ainsi de prendre en compte des scénarios de crues plus pertinents, comme ceux incluant une crue centennale.

Le modèle ne rend pas compte du nombre de taches par zone, il serait intéressant d'ajouter une telle variable dans la dynamique du modèle. Cela nécessiterait de décrire la manière dont des plantules s'entrecroisent en une tache (diminution du nombre de taches).

Cette amélioration pourrait contribuer à rendre plus satisfaisant le résultat de la calibration. En effet, bien que ce dernier permette au modèle de reproduire les données observées pour les aires, le nombre de nouvelles plantules créées avec est important dans certaines zones. Cela vient du fait que toutes les zones ont la même croissance annuelle, indépendante du nombre de taches, et ainsi beaucoup de plantules sont nécessaires pour compléter la croissance dans les zones ayant plus de taches.

### 5.3.2 L'implémentation de la méthode des cascades de noyaux de viabilité

Dans la Section 5.1.5, nous avons considéré, du fait de l'algorithme utilisé, un hyper-rectangle pour l'ensemble des perturbations, dont les bornes sont données par les valeurs extrémales des points appartenant au noyau pour chaque variable. Une perspective consisterait à développer une fonctionnalité afin que cet algorithme permette de considérer des ensembles de contrôles ou de perturbations issus de fichiers, par exemple ceux issus d'autres noyaux. Cela permettrait d'affiner la précision du calcul puisqu'on ne prendrait plus en compte les bornes maximales sur

les aires envahies et les quantités de propagules, mais les noyaux garantis des zones précédentes.

### 5.3.3 Pertinence des résultats des cascades pour la gestion

Dans la Section 5.1.3, on a considéré que l'objectif du gestionnaire est de garantir une aire totale envahie minimale après les  $T_{ini}$  premières années. La méthode que nous avons développée pour répondre à cet objectif à la Section 5.1.3 consiste à calculer les aires envahies garanties minimales par zone successivement. Ainsi, nous pouvons calculer l'aire totale envahie minimale, mais cette méthode permettrait aussi de considérer d'autres objectifs, comme celui d'indiquer un seuil dans des zones d'intérêts.

Nous n'avons comparé que deux répartitions, il est nécessaire d'en calculer bien davantage pour identifier celles qui produisent les meilleurs résultats. Il serait également intéressant d'étudier la dépendance de l'efficacité de la stratégie de gestion selon le niveau de prolifération. À intensité de contrôle fixée, cela correspondrait à indiquer l'aire envahie totale minimale qu'on peut garantir en fonction de l'aire totale initiale. Il faudrait alors définir les répartitions initiales à considérer, et étudier l'influence de ces dernières sur la répartition des efforts de gestion optimale par zone. L'analyse de l'influence du paramètre  $q_u$ , qui décrit le nombre de propagules dispersées lors de la gestion, sur les résultats semble aussi pertinente. Il reflète la qualité de la pratique de la fauche en termes de fragments de rhizomes dispersés. Une mauvaise qualité de fauche avec un grand nombre de fragments de rhizomes dispersés pourrait ruiner les efforts pour contenir l'invasion en conduisant à un trop grand nombre de propagules dispersées et donc à un trop grand nombre de nouvelles plantules. Il serait aussi intéressant d'étudier la dépendance de l'efficacité de la stratégie de gestion et de l'aire totale envahie selon l'intensité des efforts. Ces études ne demandent pas d'adaptation du code mais de modifier le plan d'expériences.

L'arrachage précoce de nouvelles plantules est également utilisé pour gérer la renouée le long de cours d'eau (Barthod and Boyer, 2019). Une étape de modélisation est nécessaire pour prendre en compte ce mécanisme dans le modèle. Pour des nouvelles invasions, caractérisées par un nombre faible de taches d'aires également très faibles (c'est-à-dire une zone faiblement envahie), l'arrachage précoce semble tout à fait approprié. Par contre, pour des invasions déjà présentes depuis plus de temps (caractérisées par un nombre de taches plus important et pour lesquelles l'arrachage manuel n'est pas possible, et donc pour des zones très envahies), la question se pose de la priorisation : vaut-il mieux gérer les zones déjà très envahies (fauche) ou les zones nouvellement colonisées (arrachage précoce), ou opérer une stratégie mixte.

Cependant, en pratique, le coût lié à l'effort de gestion n'est pas nécessairement le même pour gérer une superficie équivalente en amont et en aval (notamment du fait des conditions d'accès qui peuvent être plus compliquées en amont). Afin d'obtenir des indications de gestion plus précises, il serait nécessaire de prendre également ces informations en compte.



## Conclusion

Nous avons étudié la dynamique de la croissance de la renouée soumise à de la fauche à différentes échelles. D'abord au niveau local (tache) dans le chapitre 4, puis au niveau de zones pertinentes pour la gestion, dans le cadre de sa dispersion le long d'un cours d'eau dans le chapitre 5. Notre approche s'est inscrite dans le cadre de la modélisation. Nous avons commencé par présenter des mécanismes de croissance de la plante impliqués à ces différentes échelles, afin d'en tenir compte dans le modèle. Nous avons aussi présenté des modèles pour la gestion d'espèces envahissantes dans différents formalismes. Cependant, nous avons constaté que peu de modèles existent pour décrire la croissance de la renouée, et rares sont ceux qui incluent une possibilité de gestion. Afin d'étudier l'influence de la gestion sur la croissance des renouées, nous avons proposé deux modèles correspondant aux deux échelles évoquées. L'invasion des renouées, que ce soit au niveau de sa croissance locale, ou de sa dispersion est un phénomène dynamique. Nous avons donc fait le choix d'utiliser des modèles dynamiques. Nous avons ensuite introduit des outils pour la construction et l'étude de ces modèles. Nous avons d'abord présenté des outils méthodologiques issus de la modélisation pour l'aide à la décision. Ainsi, nous avons décrit les deux formalismes utilisés : celui des modèles individu-centrés et celui des inclusions différentielles. Nous avons mis en avant deux techniques pour la calibration de modèles, ainsi que des méthodes d'analyse de sensibilité. Nous avons également exposé des outils issus de la théorie de la viabilité utiles au problème de gestion de la renouée le long d'un cours d'eau. Plus précisément, le cadre introduit traite des systèmes contrôlés soumis à des incertitudes et pour lequel on souhaite conserver des contraintes. Nous avons présenté des méthodes pour la résolution analytique d'un problème de viabilité ainsi que des méthodes d'approximation. Cependant, notre problème de viabilité ne remplit pas les critères requis pour pouvoir appliquer ces outils, nous avons alors défini et étudié un cadre général de problème de viabilité vérifié par notre problème particulier de gestion. Nous avons proposé une méthode, appelée "cascade de noyaux de viabilité" pour la résolution de ce problème de viabilité. A l'échelle d'une tache, nous avons ensuite proposé le modèle de croissance locale de la renouée, dans le formalisme des modèles individu-centrés stochastiques. Nous avons aussi calibré la dynamique des plantes avec des données de terrain. Des simulations numériques ont été réalisées pour étudier l'influence des paramètres de gestion sur la croissance de la population. Nous avons exprimé et quantifié la dépendance de l'aire et la taille moyenne d'une tache de renouée en fonction de la taille initiale de la population, du nombre moyen de fauches

par an et de la durée du projet de gestion. Nous avons montré que le fait de faucher une seule fois par an ne suffit pas pour réduire la taille et l'aire des taches. En effet, si le nombre de fauches par an est trop faible, la taille de la population et l'aire augmentent au cours du projet de gestion. Nous avons également montré comment ces résultats pouvaient être utilisés par les gestionnaires. Les résultats issus des simulations indiquent la durée minimale du projet de gestion nécessaire pour parvenir à l'éradication (si elle est possible, compte tenu d'une certaine fréquence de fauche). Nous avons indiqué le bénéfice potentiel, en termes de réduction de l'aire envahie, à faucher la tache une fois de plus chaque année. Plus généralement, nous avons explicité la relation entre les sorties finales et les paramètres de gestion ainsi que la taille initiale de la population, ce qui permet de répondre aux questions sur l'efficacité des différentes stratégies de fauches.

A l'échelle d'un cours d'eau, nous avons proposé un modèle générique de croissance et de dispersion de la renouée, dans le formalisme des inclusions différentielles. Nous avons décrit comment certains paramètres peuvent être issus d'études préalables. Nous avons défini un cadre expérimental pour tester nos méthodes à l'échelle pertinente pour la gestion le long d'un cours d'eau, en définissant un cours d'eau "patchwork" à partir des données disponibles. Nous avons ensuite calibré la dynamique de l'invasion sur ce cours d'eau "patchwork". Nous avons aussi exprimé l'aire totale envahie à partir de noyaux de viabilité garantis et nous avons ainsi identifié les seuils d'invasion en fonction de la répartition des efforts de fauches. Nous avons comparé ces quantités pour deux stratégies de gestion a priori très différentes, l'une concentrant les efforts de fauche sur les zones faiblement envahies, la seconde étant homogène sur les zones. Nous avons constaté que la première donne l'aire totale minimale, mais que l'écart est très faible.

En résumé, à l'issue de ce travail de thèse, nos contributions concernant la dynamique de la renouée portent sur l'étude de l'influence de stratégies de gestion à différentes échelles. Nos contributions méthodologiques comportent la méthode de cascade de noyaux de viabilité pour résoudre certains problèmes de viabilité garantie de grande dimension sous forme de plusieurs problèmes de viabilité en dimension plus petite et donc rendre leur approximation numérique possible. Nous avons également proposé un cadre pour la modélisation de dispersion de renouées le long de cours d'eau.

## Perspectives

Concernant la dynamique de croissance d'une tache, une première perspective consiste à approfondir l'étude du modèle de croissance locale. Il serait intéressant d'obtenir un résultat de limite en grande population pour la dynamique avec terme de fauche. En effet, il serait alors possible d'étudier l'équation limite qui pourrait, selon la forme de la condition initiale, s'écrire comme une équation aux dérivées partielles sur la densité de couronnes en fonction de la biomasse. Une première piste pour réaliser cela pourrait être de commencer par considérer le cas où toutes les couronnes sont effectivement fauchées, ce qui simplifierait le modèle et pourrait faciliter le calcul du générateur, qui intervient dans l'équation limite. L'étude de l'équation obtenue permettrait de comparer les résultats déduits de l'étude de ce nouveau système (éventuellement effectuée de manière numérique à l'aide de schémas numériques) avec ceux obtenus par l'étude du modèle individu-centré. En effet, ce nouveau cadre permet également l'étude de l'influence des paramètres de gestion et en particulier, il permettrait d'analyser si l'on retrouve dans ce cas la présence d'un seuil minimal sur la fréquence de fauche pour observer une diminution de l'aire envahie.

Une autre piste de recherche concerne l'étude du cas sans fauche dans le modèle individu-

centré. En effet, les techniques utilisées n'ont pas permis d'obtenir de résultats significatifs. Cependant, la compréhension des déterminants de la croissance dans le cas sans fauche serait un atout majeur, en particulier pour déterminer les coûts issus d'une gestion différée. Une telle étude permettrait aussi d'identifier l'instant à partir duquel il serait illusoire (du fait de l'effort à appliquer par la suite) de débiter une gestion par la fauche, et que d'autres techniques de gestion devraient être employées.

Des perspectives peuvent également être envisagées pour le modèle de dispersion de la renouée le long d'un cours d'eau. D'un point de vue modélisation, d'abord, en incluant d'autres techniques de gestion tel que l'arrachage précoce. En effet, les crues sont à l'origine de la dispersion de fragments de rhizomes et d'akènes, qui créent de nouvelles plantules. Il serait alors intéressant de comparer les résultats portant sur l'aire minimale garantie pour des stratégies de gestion issues de la fauche et de l'arrachage précoce. Une solution efficace de gestion réside sûrement dans une stratégie mixte, pour laquelle il peut être utile de déterminer la répartition de l'effort.

Un effort de modélisation peut aussi être fait pour améliorer la dynamique de croissance de l'aire envahie dans une zone. En effet, la dynamique du nombre de taches n'est pas modélisée, alors que le modèle pour la croissance locale d'une tache est utilisé et qu'une zone est composée de plusieurs taches.

Une autre perspective concerne l'étude même du modèle. Nous avons considéré des efforts maximaux constants par zone dans le temps. Il pourrait être intéressant de considérer des ensembles de gestion plus riches pour ces répartitions. Par exemple, un aspect dynamique pourrait être envisagé dans les contrôles, qui pourrait en particulier dépendre de l'état du système et du temps, et permettrait de mettre en évidence les pertes liées à l'arrêt temporaire de la gestion.

De plus, nous avons considéré que le coût de l'effort de gestion est identique pour chaque zone. Or, cela n'est pas nécessairement le cas, par exemple le parcours d'une berge peut ne pas être aussi aisé tout le long de cours d'eau, ce qui rend le coût de gestion dépendant de la zone. Une perspective serait alors de prendre en compte les coûts liés à la gestion et de les intégrer dans le modèle pour proposer des comparaisons entre des stratégies plus conformes à celles que pourrait souhaiter un gestionnaire. Cela peut aussi prendre la forme d'enquête visant à déterminer la manière dont est effectuée en pratique la gestion.

Nous avons mené une étude sur les paramètres de gestion, il serait également intéressant d'analyser l'impact des caractéristiques du cours d'eau sur la dispersion. En effet, il semble que la renouée se propage plus rapidement le long de cours d'eau dynamiques plutôt qu'en plaine. En particulier, cela pourrait aboutir à des stratégies de gestion différentes (amont, uniforme, aval). De plus, dans le modèle, la quantité de propagules dispersées ne dépend que de la hauteur de crue, cela peut faire l'objet d'approfondissements visant à prendre en compte des caractéristiques du cours d'eau. Nous avons étudié le modèle pour répondre à un objectif précis de gestion. Pour cela, nous nous sommes placés dans le cadre de la théorie de la viabilité et nous avons mené une étude dans ce formalisme. Une étude du modèle pourrait donc être menée hors du cadre de la viabilité, en analysant directement la dynamique du modèle (sans l'associer à une contrainte de gestion), en utilisant par exemple des méthodes issues de l'analyse de sensibilité pour des modèles dynamiques.

Un autre enjeu consiste à réaliser une étude sur un cas pratique concret. Dans notre cas d'étude, nous avons considéré l'Hérault, pour lequel nous disposons de données sur la présence de renouées, mais pour lequel nous n'avons pas obtenu de données sur l'hydrodynamique. Une piste serait alors d'obtenir, en mettant par exemple en place des mesures sur le terrain, les données qui faisaient défaut. Les renouées sont aussi présentes le long d'autres cours d'eau. Si

des données de géodynamique fluviale sont plus facilement accessibles sur ceux-ci (en particulier s'ils sont plus larges), il peut être plus réaliste de faire des relevés sur la présence des renouées (évolution à entre deux dates) afin de calibrer le modèle puis d'obtenir des recommandations de gestion pour ce cours d'eau.

Enfin, les travaux réalisés pendant la thèse ont permis de produire des recommandations de gestion. Il est nécessaire pour évaluer la qualité des recommandations effectuées (trouvées par l'analyse du modèle) de confronter les prédictions du modèle à la réalisation de ces recommandations sur le terrain. Cela nécessiterait une étude reposant sur un plan d'expérience intéressant à mettre en place, puisque le nombre de tests à effectuer sur le terrain est limité. Si ces résultats ne correspondent pas aux prédictions des modèles, cela impliquerait de modifier les mécanismes de croissance, et éventuellement d'en considérer de nouveaux (comme l'intégration clonale). Ces allers-retours entre terrain et théorie sont essentiels dans le cadre de la modélisation, mais nécessitent du temps et des moyens pour les effectuer. Dans tous les cas, il pourrait être intéressant de rendre accessibles ces travaux de recherche à des gestionnaires. Le modèle nécessitant d'être calibré, cela encouragerait peut être les gestionnaires à effectuer des mesures pour quantifier les effets de leur gestion. Ces derniers pourraient ensuite les renseigner dans l'application, qui leur donnerait les indications de gestion issues de nos résultats. Le retour du gestionnaire après cette étape permettrait là encore de conforter ou d'invalidier les prédictions du modèle. Dans tous les cas, ce processus serait source de nouvelles données de croissance de renouées.

Concernant la méthode de cascade de noyaux, son utilisation ne se limite pas au cas de la dispersion de la renouée, et peut se révéler utile pour résoudre d'autres problèmes de jeux à plusieurs joueurs dès lors qu'une succession d'impacts peut être supposée : par exemple, en restant dans le domaine des cours d'eau, les problèmes d'irrigation ; dans d'autres domaines, les problèmes impliquant entre autres des filières de transformation de produits, par exemple.







---

## Annexes

---



## Annexe

# A

## Limite en grande population

Ce chapitre vise à montrer un résultat de limite en grande population relatif au processus défini à la Section 4.1.3. Le résultat principal est la convergence de la suite des processus renormalisés solutions de l'EDS (4.10) (renormalisée et simplifiée en ôtant le terme de fauche) vers la solution d'une équation intégral-différentielle (Théorèmes A.3.12 et A.3.13). Pour des grandes populations, il est donc possible de remplacer l'étude de l'EDS (4.10), sans le terme de fauche, par l'étude de ce processus.

Cette partie est organisée comme suit. Dans la Section A.1, nous énonçons les résultats mathématiques utilisés dans la Section 4.2.1 et dans la Section A.3. La preuve du résultat de limite en grande population nécessite notamment le calcul du générateur infinitésimal associé aux processus, ce qui fait l'objet de la Section A.2. Les techniques utilisées pour montrer les résultats de ces deux dernières sections sont adaptées de la thèse de Tran (Tran, 2006).

### A.1 Résultats mathématiques

#### A.1.1 Mesures Ponctuelles de Poisson

Dans le modèle pour la croissance d'une tache de renouée (voir Section 4.1.3), les événements (naissance, mort et fauche) sont gérés par mesures ponctuelles de Poisson. Nous en rappelons ici la définition et quelques premières propriétés.

Dans cette partie, les résultats cités sont issus de Çinlar (2011).

Soit  $(\Omega, \mathcal{F}, \mathbb{P})$  l'espace des probabilités et soit  $(E, \mathcal{E})$  un espace mesurable.

Une application  $M : \Omega \times E \rightarrow \mathbb{R}_+$  est une mesure aléatoire si  $\omega \rightarrow M(\omega, A)$  est une variable aléatoire pour chaque  $A \in E$  et si  $A \rightarrow M(\omega, A)$  est une mesure sur  $(E, \mathcal{E})$  pour chaque  $\omega \in \Omega$ . On note  $M(A)$  cette variable aléatoire. On peut alors voir  $M$  comme une collection de variables aléatoires  $M(A), A \in E$ . On note  $M_\omega$  la mesure  $A \rightarrow M(\omega, A)$ .

Le terme "mesure aléatoire" signifie que  $M$  est une variable aléatoire qui associe une mesure  $M_\omega$  à chaque événement  $\omega \in \Omega$ . Les termes finie,  $\sigma$ -finie, ... sont utilisés pour une mesure aléatoire  $M$  si les mesures  $M_\omega$  ont les propriétés correspondantes pour presque tout  $\omega \in \Omega$ .

En particulier, une mesure aléatoire est à valeurs entières si elle prend ses valeurs dans  $\overline{\mathbb{N}} = \{0, 1, \dots, \infty\}$ , et elle est dite mesure aléatoire de comptage si, pour presque tout  $\omega \in \Omega$ , la mesure  $M_\omega$  est atomique et que chaque atome a un poids égal à 1.

### Intégrales :

Soit  $M$  une mesure aléatoire sur  $(E, \mathcal{E})$ . Soit  $\mathcal{E}_+$  l'ensemble des fonctions positives  $\mathcal{E}$ -mesurables.

$$Mf(\omega) := \int_E M(\omega, dx) f(x), \quad \omega \in \Omega,$$

définit une variable aléatoire positive  $Mf$  pour tout  $f \in \mathcal{E}_+$ .

$$\mu(A) := \mathbb{E}M(A) = \int_\Omega \mathbb{P}(d\omega) M(\omega, A), \quad A \in \mathcal{E}$$

définit une mesure  $\mu$  sur  $(E, \mathcal{E})$ , appelée moyenne de  $M$ , et on a par le théorème de Fubini :

$$E(Mf) = \mu f, \quad f \in \mathcal{E}_+.$$

On étend la définition d'une variable aléatoire de Poisson  $X$  de paramètre  $c \in ]0, \infty[$  au cas  $c = 0$  et  $c = \infty$  par :

$$c = 0 \Leftrightarrow X = 0 \text{ ps et } c = \infty \Leftrightarrow X = \infty \text{ ps}$$

**Définition A.1.1.** Soit  $(E, \mathcal{E})$  un espace mesurable et soit  $\nu$  une mesure sur  $(E, \mathcal{E})$ . Une mesure aléatoire  $N$  sur  $(E, \mathcal{E})$  est une mesure de Poisson d'intensité  $\nu$  si :

- Pour tout  $A \in \mathcal{E}$ , la variable aléatoire  $N(A)$  a une loi de Poisson de paramètre  $\nu(A)$ .
- Pour tous  $A_1, \dots, A_n \in \mathcal{E}$  disjoints, les variables aléatoires  $(N(A_1), \dots, N(A_n))$  sont indépendantes, pour tout  $n \geq 2$ .

Soit  $N$  une mesure de Poisson d'intensité  $\nu$ . On a donc pour tout  $f \in \mathcal{E}_+$ ,  $E(Nf) = \nu f$  et  $\text{Var}(Nf) = \nu(f^2)$  si  $\nu f < \infty$ .

Le théorème suivant (théorème 2.9 du chapitre 6 de Çinlar (2011)) donne une caractérisation des mesures aléatoires de Poisson :

**Théorème A.1.2.** Soit  $N$  une mesure aléatoire sur un espace mesurable  $(E, \mathcal{E})$ . C'est une mesure aléatoire de Poisson d'intensité  $\nu$  si et seulement si :

$$\mathbb{E} \left( e^{\int_E N(dx) f(x)} \right) = e^{\int_E \nu(dx) (1 - e^{-f(x)})}, \quad \forall f \in \mathcal{E}_+$$

Le théorème suivant (théorème 2.15 du chapitre 6 de Çinlar (2011)) est un théorème d'existence de mesures de Poisson :

**Théorème A.1.3.** Soit  $\nu$  une mesure  $\sigma$ -finie sur  $(E, \mathcal{E})$ . Alors, il existe un espace de probabilité  $(W, \mathcal{G}, \mathbb{P})$  et une mesure  $N(w, \cdot)$  sur  $(E, \mathcal{E})$  pour chaque  $w \in W$  tels que  $N$  est une mesure de Poisson d'intensité  $\nu$ .

Soit  $(\mathcal{F}_t)$  une filtration. Les processus sont indexés sur  $\mathbb{R}_+$ , adaptés à  $(\mathcal{F}_t)$ , et à valeurs dans  $\mathbb{R}$ . Un processus  $F = (F_t)$  est une application  $(\omega, t) \rightarrow F_t(\omega)$  de  $\Omega \times \mathbb{R}_+$  dans  $\mathbb{R}$ . Soit  $(\Omega, \mathcal{F}, \mathbb{P})$  un espace de probabilité et  $(\mathcal{F}_t)_{t \geq 0}$  une filtration de  $\mathcal{F}$ . On considère la tribu  $\mathcal{F}^P$  sur  $\mathbb{R}_+ \times \Omega$  engendrée par les ensembles de la forme  $(s, t] \times F$ , où  $0 \leq s \leq t$ ,  $F \in \mathcal{F}_s$ . Le processus  $X : (t, \omega) \rightarrow X(t, \omega)$  est dit prévisible si il est  $\mathcal{F}^P$  mesurable. Le théorème suivant (théorème 6.2 du chapitre 6 de Çinlar (2011)) est utile dans les calculs faisant intervenir l'espérance d'une fonction intégrée contre une mesure ponctuelle de Poisson :

**Théorème A.1.4.** *Soit  $M$  une mesure ponctuelle de Poisson sur  $\mathbb{R}_+ \times E$  d'intensité  $\mu$  qui vérifie  $\mu(\{0\} \times E) = 0$ . Soit  $G = \{G(t, z) : t \in \mathbb{R}_+, z \in E\}$  un processus positif  $\mathcal{F}^P \otimes \mathcal{E}$ -mesurable.*

Alors,

$$E \int_{\mathbb{R}_+ \times E} M(dt, dz) G(t, z) = E \int_{\mathbb{R}_+ \times E} \mu(dt, dz) G(t, z).$$

Remarque : L'hypothèse sur  $G$  dit que l'application  $(\omega, t, z) \rightarrow G(\omega, t, z)$  de  $\Omega \times \mathbb{R}_+ \times E$  dans  $\overline{\mathbb{R}_+}$  est  $(\mathcal{F}^P \otimes \mathcal{E})$ -mesurable.

Avant de donner la définition de la mesure compensée d'une mesure ponctuelle de Poisson, on rappelle les définitions suivantes :

Si  $Z^i$  et  $Z^j$  sont deux martingales réelles, alors on a la formule d'intégration par parties suivante :

$$Z_t^i Z_t^j = Z_0^i Z_0^j + \int_0^t Z_{s-}^i dZ_s^j + \int_0^t Z_{s-}^j dZ_s^i + [Z^i, Z^j]_t$$

où  $[Z^i, Z^j]_t$  est un processus dont les chemins sont à variations finies. Si  $i = j$ , alors  $[Z^i, Z^i]_t$  est un processus croissant, appelé variation quadratique de  $Z^i$ .

On a  $\forall t$ ,

$$[Z^i, Z^j]_t = \lim_{\delta(\Pi) \rightarrow 0} \mathbb{P} \left( \sum_{t_n \in \Pi} (Z_{t_{n+1} \wedge t}^i - Z_{t_n \wedge t}^i)(Z_{t_{n+1} \wedge t}^j - Z_{t_n \wedge t}^j) \right)$$

où  $\Pi$  est une subdivision de  $0 = t_0 < t_1 < \dots < t_n < \dots$  de  $\mathbb{R}$  tel que  $\delta(\pi) = \sup_n (t_{n+1} - t_n)$  tend vers 0.

Si le processus  $([Z^i, Z^j])_{t \geq 0}$  est localement intégrable, alors il existe un unique processus prévisible, dont les chemins sont à variation finie, noté  $\langle Z_i, Z_j \rangle$  tel que

$$([Z^i, Z^j] - \langle Z_i, Z_j \rangle)_{t \geq 0}$$

soit une martingale locale.

On définit maintenant la mesure compensée d'une mesure ponctuelle de Poisson. Le théorème suivant est le théorème 10.3 de Méléard and Bansaye (2015).

**Théorème A.1.5.** *Soit  $N$  une mesure ponctuelle de Poisson d'intensité  $\mu(dh)dt$  sur  $\mathbb{R}_+ \times E$ . Soit la tribu prévisible  $\mathcal{F}^P$  sur  $\Omega \times \mathbb{R}_+$  (engendrée par les processus adaptés et continus à gauche) et soit le processus prévisible  $(G(s, h); s \in \mathbb{R}_+, h \in E)$  qui est  $\mathcal{F}^P \otimes \mathcal{E}$ -mesurable. On suppose que pour tout  $t \in \mathbb{R}_+$*

$$\mathbb{E} \left( \int_0^t \int_E |G(s, h)| N(ds, dh) \right) < \infty.$$

- Le processus  $M = (\int_0^t \int_E G(s, h) N(ds, dh) - \int_0^t \int_E G(s, h) \mu(dh) ds, t \leq T)$  est une martingale. La mesure aléatoire  $\tilde{N}(ds, dh) = N(ds, dh) - \mu(dh) ds$  est appelée la mesure compensée de  $N$ .

- Si on suppose de plus que  $\mathbb{E}(\int_0^t \int_E G^2(s, h) \mu(dh) ds) < \infty$ , alors la martingale  $M$  est de carré intégrable et sa variation quadratique est donnée par

$$\langle M \rangle_t = \int_0^t \int_E G^2(s, h) \mu(dh) ds$$

La formule d'Itô avec sauts est énoncé dans Ikeda and Watanabe (1989) (théorème 5.1 du chapitre 2).

On considère toujours l'espace  $(\Omega, \mathcal{F}, \mathbb{P})$  de probabilité muni de sa filtration naturelle continue à droite  $(\mathcal{F}_t)$ . On fait les hypothèses suivantes :

- $(M^i(t), 1 \leq \dots, \leq d)$  est une suite de martingales locales continues de carré intégrable.
- $(A^i(t), 1 \leq \dots, \leq d)$  est une suite de processus continus  $(\mathcal{F}_t)$ -adaptés, presque sûrement à variations bornées sur chaque intervalle de temps fini et tels que  $A^i(0) = 0$ .
- $\mu$  est une mesure ponctuelle qui admet un compensateur  $\mu^p$  par rapport à  $(\mathcal{F}_t)$  sur un espace d'état tel que  $f^i(t, x, \omega) g^j(t, x, \omega) = 0$ , pour  $1 \leq i, j \leq d$ . On suppose de plus que  $|g^i(t, x, \omega)|$  est bornée, i.e.  $\exists M > 0, |g^i(t, x, \omega)| \leq M, \forall i, t, x, \omega$ .
- $X^i(0), 1 \leq i \leq d$  sont des variables aléatoires  $\mathcal{F}_0$ -mesurables.

On définit une semi-martingale  $d$ -dimensionnelle  $(X^1(t), \dots, X^d(t))$  par

$$X^i(t) = X^i(0) + M^i(t) + A^i(t) + \int_0^t \int_X f^i(s, x, \cdot) \mu(ds dx) + \int_0^t \int_X g^i(s, x, \cdot) (\mu - \mu^p)(ds dx), \quad 1 \leq i \leq d$$

On note  $f = (f^1, \dots, f^d)$ , et  $g = (g^1, \dots, g^d)$ , on a alors le théorème suivant :

**Théorème A.1.6.** Soit  $F$  une fonction de classe  $\mathcal{C}^2$  sur  $\mathbb{R}^d$  et  $X(t)$  la semi-martingale  $d$ -dimensionnelle définie ci-dessus. Alors le processus stochastique  $F(X(t))$  est aussi une semi-martingale par rapport à  $(\mathcal{F}_t)$  et on a la formule suivante :

$$\begin{aligned} F(X(t)) - F(X(0)) &= \sum_{i=1}^d \int_0^t F'_i(X(s)) dM^i(s) + \sum_{i=1}^d \int_0^t F'_i(X(s)) dA^i(s) \\ &+ \frac{1}{2} \sum_{i,j=1}^d \int_0^t F'_{i,j}(X(s)) d\langle M^i, M^j \rangle(s) \\ &+ \int_0^t \int_X \{F(X(s^-)) + f(s, x, \cdot) - F(X(s^-))\} \mu(dx ds) \\ &+ \int_0^t \int_X \{F(X(s^-)) + g(s, x, \cdot) - F(X(s^-))\} (\mu - \mu^p)(dx ds) \\ &+ \int_0^t \int_X \{F(X(s^-)) + g(s, x, \cdot) - F(X(s^-)) - \sum_{i=1}^d g(s, x, \cdot) F'_i(X(s))\} \mu^p(dx ds) \end{aligned}$$

### A.1.2 Lemme de Gronwall

Le théorème suivant est le lemme I.1 du chapitre 10 de Queffélec and Zuily (2013). Nous l'utilisons pour obtenir une majoration dans la démonstration d'existence de la Proposition 4.2.2 (de la Section 4.2.1) et dans la preuve de l'unicité du Théorème A.3.13 (de la Section A.3.3).

**Lemme A.1.7.** Soit  $K > 0$  et  $t_0 \in \mathbb{R}$ . Soit  $\Psi \geq 0$  et  $\phi$  deux fonctions continues de  $[t_0, +\infty[$  dans  $\mathbb{R}_+$  qui vérifient :

$$\forall t \geq t_0, \quad \phi(t) \leq K + \int_{t_0}^t \Psi(s)\phi(s)ds.$$

Alors,

$$\forall t \geq t_0, \quad \phi(t) \leq Ke^{\int_{t_0}^t \Psi(s)ds}.$$

### A.1.3 Convergence vague, étroite

Soit  $E$  un espace métrique. On note  $\mathcal{C}_b(E)$  l'ensemble des fonctions continues et bornées de  $E$  dans  $\mathbb{R}$ ,  $\mathcal{C}_K(E)$  l'ensemble des fonctions continues à support compact de  $E$  dans  $\mathbb{R}$  et  $\mathcal{C}_0(E)$  l'ensemble des fonctions continues  $E$  dans  $\mathbb{R}$  qui ont pour limite 0 à l'infini.

Par défaut, on munit  $\mathcal{M}_F(E)$  de la topologie de la convergence étroite. Cependant, dans la preuve du Théorème A.3.12 (Section A.3.3), nous avons également besoin de la topologie de la convergence vague. Le Théorème A.1.10 utilise quant à lui la convergence faible. Nous rappelons donc les définitions suivantes :

La topologie de la convergence étroite est la plus petite topologie rendant continues les applications  $m \rightarrow \langle m, f \rangle$ , avec  $f \in \mathcal{C}_b(E)$ .

La topologie de la convergence faible est la plus petite topologie rendant continues les applications  $m \rightarrow \langle m, f \rangle$ , avec  $f \in \mathcal{C}_0(E)$ .

La topologie de la convergence vague est la plus petite topologie rendant continues les applications  $m \rightarrow \langle m, f \rangle$ , avec  $f \in \mathcal{C}_K(E)$ .

Nous pouvons définir la convergence vague, faible et étroite de suite dans  $\mathcal{M}_F(E)$ .

une suite  $\mu_n$  converge vaguement vers  $\mu$  pour la topologie vague [resp faible, resp étroite] si :

$$\lim_{n \rightarrow \infty} \mu_n(f) = \mu, \quad \forall f \in \mathcal{C}_K(E) \quad [\text{resp } \mathcal{C}_0(E), \text{ resp } \mathcal{C}_b(E)]$$

La topologie la plus fine (qui a le plus d'ouverts) est la topologie étroite, une suite qui converge étroitement converge faiblement et une suite qui converge faiblement converge vaguement.

### A.1.4 Critères de tension

Pour montrer la convergence de la suite des processus à valeurs mesure, on utilise le théorème de Prokhorov, on a donc besoin de critères de tension.

Soit  $H$  un espace métrique complet séparable, muni de la distance  $d$ .

**Définition A.1.8.** Une famille  $N \subset \mathcal{M}_1(H)$  de mesures de probabilité est dite tendue si, pour tout  $\epsilon > 0$ , il existe un compact  $K_\epsilon \subset H$  tel que,

$$\forall \mu \in N, \quad \mu(K_\epsilon) \geq 1 - \epsilon$$



Le théorème suivant (Chapitre 6, Section 3.a de Jacod and Shiryaev (2013)) permet de montrer la convergence faible d'une suite de mesure de probabilité.

**Théorème A.1.9** (Prokhorov). *Soient  $S$  un espace polonais et  $A$  une partie de l'ensemble  $\mathcal{M}_F(S)$  des mesures (positives) finies sur la tribu borélienne de  $S$ . Alors les deux propositions suivantes sont équivalentes :*

- l'ensemble  $\{\mu(S) \mid \mu \in A\}$  est borné et la partie  $A$  est tendue.
- l'ensemble  $A$  est relativement compact dans  $\mathcal{M}(S)$  pour la topologie de la convergence faible.

*Pour une suite de lois de probabilité associée à une suite de variables aléatoires à valeurs dans  $S$ , la convergence faible des lois correspond à la convergence en loi des variables aléatoires (c'est aussi la convergence étroite).*

On énonce maintenant un critère de tension de Roelly-Coppoletta. Ce critère est le Théorème 2.1 de Roelly-Coppoletta (1986). On considère la topologie de la convergence faible. Soit  $\Omega = \mathbb{D}([0, T]; \mathcal{M}_F)$ , l'espace des fonctions continues à droite sur  $\mathbb{R}_+$  et à valeurs dans l'espace des mesures positives finies.

**Théorème A.1.10** (Roelly-Coppoletta). *Soient  $(P^n)_{n \in \mathbb{N}}$  une suite de probabilités sur  $\Omega$ , et  $f_k$  une suite de fonctions dense dans  $\mathcal{C}_0$ . Si, pour tout  $k \in \mathbb{N}$ ,  $(\langle P_n, f_k \rangle)_{n \in \mathbb{N}}$  est une suite tendue de probabilités sur  $\mathbb{D}([0, T], \mathbb{R})$ , (fonctions càdlàg de  $[0, T]$  dans  $\mathbb{R}_+$ ), alors  $(P^n)_{n \in \mathbb{N}}$  est tendue sur  $\Omega$ .*

Enfin, énonçons un résultat de tension issu de Joffe and Métivier (1986).

Pour  $x \in \mathbb{D}(\mathbb{R}_+, H)$ , on définit,  $\forall N > 0, \delta > 0$ ,

$$w^N(x, \delta) = \inf_{\Pi_\delta} \max_{t_i \in \Pi} \sup_{t_i \leq s, t < t_{i+1}} d(x(t), x(s))$$

où  $\Pi_\delta$ , est l'ensemble des suites  $t = t_0 < t_1 < \dots < t_n < N$  dans  $\mathbb{R}_+$  qui vérifient  $\inf_{t_i \in [t_0, N]} |t_i - t_{i+1}| \geq \delta$ .

Soit  $(X^n)_{n \in \mathbb{N}}$  une suite de processus définis respectivement sur leurs espaces de probabilité  $(\Omega^n, \mathcal{A}^n, \mathbb{P}^n)$ , à valeurs dans un espace métrique complet séparable  $H$ . La suite  $(\tilde{P}^n)$  de lois du processus  $(X^n)$  est une suite tendue si et seulement si :

[T<sub>1</sub>]  $\forall t \in \mathcal{T}$ ,  $\mathcal{T}$  dense dans  $\mathbb{R}_+$ , les lois de  $(X_t^n)_{n \in \mathbb{N}}$ , sont tendues dans  $H$ .

[T<sub>2</sub>]  $\forall N > 0, \eta > 0, \epsilon > 0, \exists \delta > 0$  tel que

$$\limsup_n \mathbb{P}^n(\omega \in \Omega_n, w^N(X^n(\cdot, \omega), \delta) > \eta) \leq \epsilon$$

On énonce ensuite un critère dû à Aldous qui implique la condition [T<sub>2</sub>].

On dit que la suite  $(X^n)_{n \in \mathbb{N}}$  vérifie la condition [A] si :

$\forall N > 0, \epsilon > 0, \eta > 0, \exists \delta > 0$  et  $n_0$  tel que pour toute suite  $(\sigma_n, \tau_n)_{n \in \mathbb{N}}$ , de paires de temps d'arrêt  $(\sigma_n$  et  $\tau_n$  sont des temps d'arrêt par rapport à  $\mathcal{F}^n$  sur  $\Omega_n$ ) qui vérifient  $\sigma_n \leq \tau_n \leq N$ , on a

$$\sup_{n \geq n_0} \mathbb{P}^n(d(X_{\sigma_n}^n, X_{\tau_n}^n) > \eta, \tau_n < \sigma_n + \delta) \leq \epsilon.$$

**Théorème A.1.11** (Aldous). *La condition [A] implique le critère de tension [T<sub>2</sub>].*

**Théorème A.1.12** (Rebolledo). *Si  $X^n$  est une suite de semi martingales de la forme  $X^n = A^n + M^n$ , avec  $A^n$  un processus à variations finies et  $M^n$  une martingale localement intégrable (on suppose aussi que  $X_0^n = A_0^n$ ). Si les suites  $(A^n)$  et  $(\langle M^n \rangle)$  vérifient les conditions [T<sub>1</sub>] et [A], alors les lois de probabilité  $(\tilde{P}^n)$  de la suite  $(X^n)$  forment une famille tendue.*

## A.2 Calcul du générateur

Nous commençons par rappeler l'Équation (4.10), Section 4.1.3.

$$\begin{aligned}
Z_t = & \sum_{i=1}^{N_0} \delta_{(X_i(Z_0), A_b(t, 0, A_i(Z_0)))} \\
& + \int_{[0, t] \times \mathbb{N}^* \times \mathbb{R}_+ \times \mathbb{R}^2} \mathbf{1}_{\{i \leq N_{s-}\}} \delta_{(X_i(Z_s) + z, A_b(t, s, a_0))} \mathbf{1}_{\{\theta \leq b(X_i(Z_s), Z_s)\}} \\
& \quad \mathbf{1}_{\{X_i(Z_s) + z \in C_{X_i(Z_s), Z_s}\}} M_1(ds, di, d\theta, dz) \\
& + \int_{[0, t] \times [0, 1]^{N^*}} \sum_{i=1}^{N_{s-}} \mathbf{1}_{\{y_i \leq \text{proportionMowing}\}} (\delta_{(X_i(Z_s), A_b(t, s, A_i(Z_{s-})*F(A_i(Z_{s-}))))} \\
& \quad - \delta_{(X_i(Z_s), A_b(t, s, A_i(Z_s)))}) M_2(ds, dy) \\
& - \int_{[0, t] \times \mathbb{N}^* \times \mathbb{R}_+} \mathbf{1}_{\{i \leq N_{s-}\}} \mathbf{1}_{\{\theta \leq m(A_i(Z_s))\}} \delta_{(X_i(Z_s), A_b(t, s, A_i(Z_s)))} M_3(ds, di, d\theta)
\end{aligned} \tag{A.1}$$

Notons  $(Z_t)_{t \in \mathbb{R}_+}$  la solution de (A.1). Rappelons qu'à la Section 4.2.1 nous avons montré l'existence et l'unicité d'un tel processus, et nous utilisons ici les notations qui y ont été introduites en début de section. Dans cette section nous calculons le générateur infinitésimal du processus markovien  $(Z_t)_{t \in \mathbb{R}_+}$  défini par :

$$\forall f \in \mathcal{C}_b^{0,1}(\mathbb{R}^2, \mathbb{R}), \forall F \in \mathcal{C}_b^1(\mathbb{R}), LF_f(Z_0) = \frac{d}{dt} \mathbb{E}(F_f(Z_t))|_{t=0}$$

Le générateur apparaîtra dans l'équation limite de la Partie A.3. On rappelle que  $v$  est défini par  $\frac{dA(t)}{dt} = v(A(t))$ , et qu'on note  $p := \text{proportionMowing}$ . Dans le calcul du générateur, nous avons besoin d'une expression pour la quantité  $F(\langle Z_t, f_t \rangle)$  donnée par le lemme suivant :

**Lemme A.2.1.** *On se place dans le cadre de l'Hypothèse 4.2.1 (Section 4.2.1). Soit  $F \in \mathcal{C}^1(\mathbb{R})$ , et  $f : (x, a, s) \rightarrow f_s(x, a) \in \mathcal{B}(\mathbb{R}^2 \times \mathbb{R}_+ \times \mathbb{R}_+, \mathbb{R})$  telle que pour tout  $x \in \mathbb{R}^2$ ,  $(s, a) \rightarrow f_s(x, a) \in \mathcal{C}^{1,1}(\mathbb{R}_+ \times \mathbb{R}_+, \mathbb{R})$ , et  $f$  a des dérivées bornées uniformément en  $x$ . Alors,  $\forall t \in \mathbb{R}_+$ ,*

$$\begin{aligned}
F(\langle Z_t, f_t \rangle) = & F(\langle Z_0, f_0 \rangle) + \int_0^t \langle Z_u, \frac{df_u}{du} + v \cdot \nabla_a f_u \rangle F'(\langle Z_u, f_u \rangle) du \\
& + \int_0^t \int_{E_1} \mathbf{1}_{\{i \leq N_{s-}\}} \mathbf{1}_{\{\theta \leq b(X_i(Z_s), Z_s)\}} \mathbf{1}_{\{X_i(Z_s) + z \in C_{X_i(Z_s), Z_s}\}} \\
& [F(\langle Z_s, f_s \rangle + f_s(X_i(Z_s) + z, a_0)) - F(\langle Z_s, f_s \rangle)] \times M_1(ds, di, d\theta, dz) \\
& + \int_0^t \int_{E_2} F \left( \langle Z_s, f_s \rangle + \sum_{i=1}^{N_{s-}} \mathbf{1}_{Y_i \leq p} [f_s(X_i(Z_s), A_i(Z_{s-})) \cdot F(A_i(Z_{s-})) \right. \\
& \quad \left. - f_t(X_i(Z_s), A_i(Z_s)) \right] \Big) - F(\langle Z_s, f_s \rangle) M_2(ds, dy) \\
& - \int_0^t \int_{E_3} \mathbf{1}_{\{i \leq N_{s-}\}} \mathbf{1}_{\{\theta \leq m(A_i(Z_s))\}} \\
& [F(\langle Z_s, f_s \rangle - f_s(X_i(Z_s), A_i(Z_s))) - F(\langle Z_s, f_s \rangle)] \times M_3(ds, di, d\theta) \tag{A.2}
\end{aligned}$$

*Démonstration.* Soit  $t \in \mathbb{R}_+$ , en intégrant  $f_t(x, a)$  contre  $Z_t$ , on obtient :

$$\begin{aligned}
\langle Z_t, f_t \rangle &= \sum_{i=1}^{N_0} f_t(X_i(Z_0), A_b(t, 0, A_i(Z_0))) + \int_0^t \int_{E_1} 1_{\{i \leq N_{s-}\}} f_t(X_i(Z_s) + z, A_b(t, s, a_0)) \\
&\quad \times 1_{\{\theta \leq b(X_i(Z_s), Z_s)\}} 1_{\{X_i(Z_s) + z \in C_{X_i(Z_s), Z_s}\}} \times M_1(ds, di, d\theta, dz) \\
&\quad + \int_0^t \int_{E_2} \sum_{i=1}^{N_{s-}} 1_{Y_i \leq p} \left[ f_t(X_i(Z_s), A_b(t, s, A_i(Z_{s-})).F(A_i(Z_{s-}))) \right. \\
&\quad \left. - f_t(X_i(Z_s), A_b(t, s, A_i(Z_s))) \right] M_2(ds, dy) \\
&\quad - \int_0^t \int_{E_3} 1_{\{i \leq N_{s-}\}} f_t(X_i(Z_s), A_b(t, s, A_i(Z_s))) \times 1_{\{\theta \leq m(A_i(Z_s))\}} \times M_3(ds, di, d\theta).
\end{aligned} \tag{A.3}$$

Puis, pour tout  $(x, a) \in \mathbb{R}^2 \times \mathbb{R}_+$  et pour tout  $0 \leq s \leq t$ ,

$$f_t(x, A(t, s, a)) = f_s(x, a) + \int_s^t \frac{df_u}{du}(x, A(u, s, a)) + v(A(u, s, a)).\nabla_a f_u(x, A(u, s, a)) du$$

On obtient donc :

$$\begin{aligned}
\langle Z_t, f_t \rangle &= \sum_{i=1}^{N_0} f_t(X_i(Z_0), A_i(Z_0)) + C_1(t) \\
&\quad + \int_0^t \int_{E_1} 1_{\{i \leq N_{s-}\}} f_s(X_i(Z_s) + z, a_0) 1_{\{\theta \leq b(X_i(Z_s), Z_s)\}} 1_{\{X_i(Z_s) + z \in C_{X_i(Z_s), Z_s}\}} \\
&\quad \times M_1(ds, di, d\theta, dz) + C_2(t) \\
&\quad + \int_0^t \int_{E_2} \sum_{i=1}^{N_{s-}} 1_{Y_i \leq p} [f_s(X_i(Z_s), A_i(Z_{s-})).F(A_i(Z_{s-}))) - f_s(X_i(Z_s), A_i(Z_s))] \\
&\quad \times M_2(ds, dy) + C_3(t) \\
&\quad - \int_0^t \int_{E_3} 1_{\{i \leq N_{s-}\}} f_s(X_i(Z_s), A_i(Z_s)) \times 1_{\{\theta \leq m(A_i(Z_s))\}} \times M_3(ds, di, d\theta) + C_4(t)
\end{aligned} \tag{A.4}$$

où

$$C_1(t) = \int_0^t \sum_{i=1}^{N_0} \frac{df_u}{du}(X_i(Z_0), A_b(u, 0, A_i(Z_0))) + v(A_b(u, 0, A_i(Z_0))).\nabla_a f_u(X_i(Z_0), A_b(u, 0, A_i(Z_0))) du$$

$$\begin{aligned}
C_2(t) &= \int_0^t \int_{E_1} 1_{\{i \leq N_{s-}\}} 1_{\{\theta \leq b(X_i(Z_s), Z_s)\}} 1_{\{X_i(Z_s) + z \in C_{X_i(Z_s), Z_s}\}} \\
&\quad \times \int_s^t \left( \frac{df_u}{du}(X_i(Z_s) + z, A_b(u, s, a_0)) \right. \\
&\quad \left. + v(A_b(u, s, a_0)).\nabla_a f_u(X_i(Z_s) + z, A_b(u, s, a_0)) \right) du \times M_1(ds, di, d\theta, dz)
\end{aligned}$$

$$\begin{aligned}
C_3(t) = & \int_0^t \int_{E_2} \sum_{i=1}^{N_s-} 1_{Y_i \leq p} \int_s^t \frac{df_u}{du} (X_i(Z_s), A_b(u, s, A_i(Z_s)F(A_i(Z_s)))) - \frac{df_u}{du} (X_i(Z_s), A_b(u, s, A_i(Z_s))) \\
& + v \left( A_b(u, s, A_i(Z_s)F(A_i(Z_s))) \right) \cdot \nabla_a f_u \left( X_i(Z_s), A_b(u, s, A_i(Z_s)F(A_i(Z_s))) \right) \\
& - v \left( A_b(u, s, A_i(Z_s)) \right) \cdot \nabla_a f_u \left( X_i(Z_s), A_b(u, s, A_i(Z_s)) \right) du M_2(ds, dy)
\end{aligned}$$

$$\begin{aligned}
C_4(t) = & \int_0^t \int_{E_3} 1_{\{i \leq N_s-\}} 1_{\{\theta \leq m(A_i(Z_s))\}} \\
& \times \int_s^t \frac{df_u}{du} (X_i(Z_s), A_b(u, s, A_i(Z_s))) \\
& + v(A_b(u, s, A_i(Z_s))) \cdot \nabla_a f_u(X_i(Z_s), A_b(u, s, a_i(Z_s))) du \times M_3(ds, di, d\theta)
\end{aligned}$$

Dans  $C_2(t)$ ,  $C_3(t)$  et  $C_4(t)$  on applique le théorème de Fubini (par hypothèse de bornitude des dérivées de  $f_s$  par rapport à  $a$  et  $s$ ). On échange donc  $\int_0^t \int_{E_1} M_1 ds$  et  $\int_s^t du$ , afin de reconnaître une intégrale contre  $Z_u$ . Par exemple, pour  $C_4(t)$ , cela donne :

$$\begin{aligned}
C_4(t) = & \int_0^t \int_0^u \int_{E_3} 1_{\{i \leq N_s-\}} 1_{\{\theta \leq m(A_i(Z_s))\}} \times \left[ \frac{df_u}{du} (X_i(Z_s), A_b(u, s, A_i(Z_s))) \right. \\
& \left. + v(A_b(u, s, A_i(Z_s))) \cdot \nabla_a f_u(X_i(Z_s), A_b(u, s, a_i(Z_s))) \right] \times M_3(ds, di, d\theta) du
\end{aligned}$$

et on a ainsi, en utilisant l'écriture de l'Équation (A.1),

$$\begin{aligned}
C_1(t) + C_2(t) + C_3(t) + C_4(t) &= \int_0^t \int_{\mathbb{R}^2 \times \mathbb{R}_+} \frac{df_u}{du} (x, a) + v(a) \cdot \nabla_a f_u(x, a) Z_u(dx, da) du \\
&= \int_0^t \langle Z_u, \frac{df_u}{du} + v \cdot \nabla_a f_u \rangle du
\end{aligned}$$

Finalement, l'Équation (A.4) devient :

$$\begin{aligned}
\langle Z_t, f_t \rangle = & \langle Z_0, f_0 \rangle + \int_0^t \langle Z_u, \frac{df_u}{du} + v \cdot \nabla_a f_u \rangle du \\
& + \int_0^t \int_{E_1} 1_{\{i \leq N_s-\}} 1_{\{\theta \leq b(X_i(Z_s), Z_s)\}} 1_{\{X_i(Z_s) + z \in C_{X_i(Z_s), Z_s}\}} f_s(X_i(Z_s) + z, a_0) M_1(ds, di, d\theta, dz) \\
& + \int_0^t \int_{E_2} \sum_{i=1}^{N_s-} 1_{Y_i \leq p} \left[ f_s(X_i(Z_s), A_i(Z_s-)) \cdot F(A_i(Z_s-)) - f_s(X_i(Z_s), A_i(Z_s)) \right] M_2(ds, dy) \\
& - \int_0^t \int_{E_3} 1_{\{i \leq N_s-\}} f_s(X_i(Z_s), A_i(Z_s)) \times 1_{\{\theta \leq m(A_i(Z_s))\}} M_3(ds, di, d\theta) \tag{A.5}
\end{aligned}$$

On applique la formule d'Itô avec sauts (rappelée au Théorème A.1.6) à l'Équation (A.5) pour  $F \in \mathcal{C}^1(\mathbb{R}, \mathbb{R})$ , on obtient finalement l'Équation (A.2).

□

**Calcul du générateur :**

Pour cela, on introduit le temps d'arrêt  $\zeta_n$  défini par :

$$\zeta_n = \inf \left\{ t \geq 0, N_t \geq n \text{ ou } \int_{\mathcal{X}} |a| Z_t(dx, da) \geq n \right\} \quad (\text{A.6})$$

On va calculer la quantité  $\mathbb{E}(F_f(Z_{t \wedge \zeta_n}))$ . On rappelle que par hypothèse, il existe  $\bar{v}$  est tel que :

$$\forall a \in \mathbb{R}_+, v(a) \leq \bar{v}(1 + a).$$

Soit  $n > \max \left( N_0, \int_{\mathcal{X}} |a| Z_t(dx, da) \right)$

On commence par le terme de naissance de l'Équation (A.2). On a :

$$\begin{aligned} & \mathbb{E} \left( \int_0^{t \wedge \zeta_n} \int_{E_1} 1_{\{i \leq N_{s-}\}} 1_{\{\theta \leq b(X_i(Z_s), Z_s)\}} 1_{\{X_i(Z_s) + z \in C_{X_i(Z_s), Z_s}\}} \right. \\ & \quad \times \left[ F(\langle Z_s, f \rangle + f(X_i(Z_s) + z, a_0)) - F(\langle Z_s, f \rangle) \right] M_1(ds, di, d\theta, dz) \Big) \\ &= \mathbb{E} \left( \int_0^{t \wedge \zeta_n} \int_{E_1} 1_{\{i \leq N_{s-}\}} 1_{\{\theta \leq b(X_i(Z_s), Z_s)\}} 1_{\{X_i(Z_s) + z \in C_{X_i(Z_s), Z_s}\}} \right. \\ & \quad \times \left[ F(\langle Z_s, f \rangle + f(X_i(Z_s) + z, a_0)) - F(\langle Z_s, f \rangle) \right] ds n(di) d\theta D(dz) \Big) \\ &= \mathbb{E} \left( \int_0^{t \wedge \zeta_n} \int_{\mathbb{R}^2} \sum_{i=1}^{N_{s-}} b(X_i(Z_s), Z_s) 1_{\{X_i(Z_s) + z \in C_{X_i(Z_s), Z_s}\}} \right. \\ & \quad \times \left[ F(\langle Z_s, f \rangle + f(X_i(Z_s) + z, a_0)) - F(\langle Z_s, f \rangle) \right] D(dz) ds \Big) \\ &= \mathbb{E} \left( \int_0^{t \wedge \zeta_n} \int_{\mathbb{R}^2 \times \mathbb{R}_+} \int_{\mathbb{R}^2} b(x, Z_s) 1_{\{x+z \in C_{x, Z_s}\}} \right. \\ & \quad \times \left[ F(\langle Z_s, f \rangle + f(x + z, a_0)) - F(\langle Z_s, f \rangle) \right] D(dz) Z_s(dx, da) ds \Big). \end{aligned} \quad (\text{A.7})$$

On note  $\Psi(t, Z)$  la dernière quantité dont on prend l'espérance, c'est à dire,

$$\begin{aligned} \Psi(t, Z) &:= \int_0^{t \wedge \zeta_n} \int_{\mathbb{R}^2 \times \mathbb{R}_+} \int_{\mathbb{R}^2} b(x, Z_s) 1_{\{x+z \in C_{x, Z_s}\}} \\ & \quad \times \left[ F(\langle Z_s, f \rangle + f(x + z, a_0)) - F(\langle Z_s, f \rangle) \right] D(dz) Z_s(dx, da) ds. \end{aligned}$$

On fixe  $Z_0 \in \mathcal{M}$  déterministe, alors on a :

$$\begin{aligned} \frac{d\Psi}{dt}(0, Z_0) &= \int_{\mathbb{R}^2 \times \mathbb{R}_+} \int_{\mathbb{R}^2} [F(\langle Z_0, f \rangle + f(x + z, a_0)) - F(\langle Z_0, f \rangle)] \\ & \quad \times b(x, Z_0) 1_{\{x+z \in C_{x, Z_0}\}} D(dz) Z_0(dx, da) \end{aligned}$$

De plus,  $t \in [0; T] \rightarrow \psi(t \wedge \zeta_n, Z)$  est dominé p.s par  $2NT\bar{b}c_1 \|F\|_{\infty} < \infty$  où  $c_1$  a été introduit à l'Équation (4.14) dans la démonstration de l'existence et unicité du processus (Section 4.2.1,  $c_1$  est telle que  $\int_{\mathbb{R}^2} 1_{\{z \in C_{x, Z}\}} D(z) < c_1$ ). De même, on a une majoration pour  $\frac{d\Psi}{dt}(0, Z_0)$ . Ainsi,

le théorème de dérivation sous le signe espérance donne dans l'Équation (A.8) la partie du générateur associée aux naissances :

$$\int_{\mathbb{R}^2 \times \mathbb{R}_+} \int_{\mathbb{R}^2} [F_f(Z_0 + \delta_{(x+z, a_0)}) - F_f(Z_0)] b(x, Z_0) 1_{\{x+z \in C_{x, Z_0}\}} D(dz) Z_0(dx, da). \quad (\text{A.8})$$

De même, la partie associée au terme de mort dans le générateur est donnée dans l'Équation (A.9).

$$\int_{\mathbb{R}^2 \times \mathbb{R}_+} [F_f(Z_0 - \delta_{(x, a)}) - F_f(Z_0)] m(a(Z_s)) Z_0(dx, da). \quad (\text{A.9})$$

On a introduit le temps d'arrêt  $\zeta_n$ , pour majorer le terme issu de

$$\int_0^t \langle Z_u, \frac{df_u}{du} + v \cdot \nabla_a f_u \rangle F'(\langle Z_u, f_u \rangle) du$$

dans (A.2). On pose

$$\Psi_2(t, Z) := \int_0^t \int_{\chi} v(a) \cdot \nabla_a f(x, a) F'(\langle Z_s, f \rangle) Z_s(dx, da)$$

Alors  $t \in [0; T] \rightarrow \Psi_2(t \wedge \zeta_n, Z)$  est dominé ps par  $NT2\bar{v} \|\nabla_a f\|_{\infty} \|F'\|_{\infty} < \infty$ .

De plus,

$$\frac{\Psi_2(0, Z_0)}{dt} = \langle Z_0, v \cdot \nabla_a f \rangle F'(\langle Z_0, f \rangle)$$

est dominé ps par  $NT2\bar{v}(N_0 + \int_{\chi} Z_0(dx, da)) \|\nabla_a f\|_{\infty} \|F'\|_{\infty}$  ce qui permet à nouveau d'appliquer le théorème de dérivation sous le signe espérance à  $t \rightarrow \mathbb{E}(\Psi_2(t, Z_{t \wedge \zeta_n}))$ , en  $t = 0$ , ce qui donne le terme (A.10) pour le générateur

$$\int_{\chi} v(a) \cdot \nabla_a f(x, a) F'(\langle Z_0, f \rangle) Z_0(dx, da) \quad (\text{A.10})$$

Le terme de fauche de l'Équation (A.2) est plus compliqué et ne permet pas d'utiliser la même technique. On donne donc l'écriture du générateur sans le terme de fauche.

Avec les Équations (A.9), (A.8) et (A.10), le générateur s'écrit :

$\forall f \in \mathcal{C}_b^{0,1}(\chi), \forall F \in \mathcal{C}_b^1(\mathbb{R})$ ,

$$\begin{aligned} LF_f(Z_0) &= \frac{d}{dt} \mathbb{E}(F_f(Z_t))|_{t=0} = \int_{\chi} v(a) \cdot \nabla_a f(x, a) F'(\langle Z_0, f \rangle) Z_0(dx, da) \\ &+ \int_{\chi} \int_{\mathbb{R}^2} [F_f(Z_0 + \delta_{(x+z, a_0)}) - F_f(Z_0)] b(x, Z_0) 1_{\{x+z \in C_{x, Z_0}\}} D(dz) Z_0(dx, da) \\ &+ \int_{\chi} [F_f(Z_0 - \delta_{(x, a)}) - F_f(Z_0)] m(a(Z_s)) Z_0(dx, da) \end{aligned} \quad (\text{A.11})$$

### A.3 Limite en grande population

L'objectif de cette partie est d'étudier une renormalisation  $(Z^n)_{n \in \mathbb{N}}$  du processus dont on a calculé le générateur à la section précédente (issu de l'Équation (A.1) sans terme de fauche). Nous allons augmenter la taille de la population initiale, qui sera d'ordre  $n$ , sans modifier les interactions entre les individus. Nous allons montrer la convergence en loi des processus renormalisés vers un processus solution d'une équation différentielle. Ainsi, en grande population, on pourra remplacer l'étude de l'équation différentielle stochastique par l'étude de cette nouvelle équation. Les techniques et notations utilisées dans cette partie sont celles développées dans Tran (2006).

Nous allons également simplifier l'Équation (A.1) en retirant le terme de fauche.

Rappelons la notation  $\chi = \mathbb{R}^2 \times \mathbb{R}_+$ . Pour  $n \in \mathbb{N}^*$ , on note

$$\mathcal{M}^n := \left\{ \frac{1}{n}m, m \in \mathcal{M} \right\},$$

l'ensemble formé à partir des mesures ponctuelles finies. Nous allons généraliser les applications  $X$  et  $A$  introduits dans la Section 4.1.3, à des mesures à valeurs dans  $\cup_{n \in \mathbb{N}^*} \mathcal{M}^n$ . On considère donc les applications

$$X = (X_1, \dots, X_k, \dots) : \cup_{n \in \mathbb{N}^*} \mathcal{M}^n \rightarrow (\mathbb{R}^2)^{\mathbb{N}^*}$$

et

$$A = (A_1, \dots, A_k, \dots) : \cup_{n \in \mathbb{N}^*} \mathcal{M}^n \rightarrow (\mathbb{R}_+)^{\mathbb{N}^*}$$

respectivement définies par

$$X \left( \frac{1}{n} \sum_{i=1}^n \delta_{(x_i, a_i)} \right) = (x_1, \dots, x_n, 0, \dots, 0 \dots)$$

et

$$A \left( \frac{1}{n} \sum_{i=1}^n \delta_{(x_i, a_i)} \right) = (a_1, \dots, a_n, 0, \dots, 0 \dots)$$

où les individus (on conserve le terme de la Section 4.1.3) sont rangés suivant l'ordre lexicographique sur  $\mathbb{R}^2$ , et pour des individus de même position  $x \in \mathbb{R}^2$ , par ordre croissant de biomasse.

Introduisons maintenant le processus renormalisé. On considère l'espace probabilisé  $(\Omega, \mathcal{F}, \mathbb{P})$  et une suite  $(Z_0^n)_{n \in \mathbb{N}^*}$  de variables aléatoires indépendantes qui vérifient les Hypothèses A.3.1.

#### Hypothèse A.3.1.

$$\forall n \in \mathbb{N}^*, Z_0^n \in \mathcal{M}^n, \quad \text{et} \quad \sup_{n \in \mathbb{N}^*} \mathbb{E}(\langle Z_0^n, 1 \rangle) < \infty.$$

Dans cette partie, on se place encore dans le cadre de l'Hypothèse 4.2.1 (Section 4.2.1).

Pour tout  $n \in \mathbb{N}^*$ , l'Équation différentielle stochastique (A.12) décrit l'évolution du processus  $Z^n$  à valeurs dans  $\mathcal{M}^n$ .

$$\begin{aligned}
Z_t^n = & \frac{1}{n} \sum_{i=1}^{N_0} \delta_{(X_i(Z_0^n), A_b(t, 0, A_i(Z_0^n)))} + \frac{1}{n} \int_0^t \int_{\mathbb{N}^*} \int_{\mathbb{R}_+} \int_{\mathbb{R}^2} 1_{\{i \leq N_{s-}^n\}} \delta_{(X_i(Z_s^n) + z, A_b(t, s, a_0))} \\
& \times 1_{\{\theta \leq b(X_i(Z_s^n), Z_s^n)\}} 1_{\{X_i(Z_s^n) + z \in C_{X_i(Z_s^n), Z_s^n}\}} \times M_1(ds, di, d\theta, dz) \\
& - \frac{1}{n} \int_0^t \int_{\mathbb{N}^*} \int_{\mathbb{R}_+} 1_{\{i \leq N_{s-}^n\}} \delta_{(X_i(Z_s^n), A_b(t, s, A_i(Z_s^n)))} \times 1_{\{\theta \leq m(A_i(Z_s^n))\}} \times M_3(ds, di, d\theta).
\end{aligned} \tag{A.12}$$

Dans cette équation

$$— \forall t \in \mathbb{R}_+, N_t^n = n \langle Z_t^n, 1 \rangle.$$

—  $M_1$  et  $M_3$  sont définies de la même manière que pour l'EDS de l'Équation (A.1). On rappelle les définitions de  $E_1$  et  $E_3$  données à l'Équation (4.11) (Section 4.2.1).

On note  $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$  la filtration canonique engendrée par  $(Z_0^n)$ ,  $M_1$  et  $M_3$ . Enfin, pour tout  $n \in \mathbb{N}^*$ , on pose  $T_0^n = 0$ . La suite des temps de saut successifs de  $Z^n$  est donnée par  $(T_k^n)_{k \in \mathbb{N}}$ . On définit :

$$T_\infty^n = \lim_{k \rightarrow \infty} T_k^n.$$

Le paramètre  $n$  peut servir à modéliser une taille de ressource limitée, un individu aura moins de poids s'il est dans une grande population. Comme à la section précédente (Équation (A.6)), et à la Section 4.2.1, on introduit deux temps d'arrêt :

**Définition A.3.2.** Pour tout  $N > 0$  et  $n \in \mathbb{N}^*$ , on définit les temps d'arrêt :

$$\tau_N^n := \inf\{t > 0, \langle Z_t^n, 1 \rangle > N\}$$

$$\zeta_N^n = \inf\{t \geq 0, \langle Z_t^n, 1 \rangle \geq N \text{ et } \int_{\mathcal{X}} |a| Z_t^n(dx, da) \geq N\}.$$

On obtient alors l'existence et l'unicité de la solution de l'EDS (A.12) :

**Proposition A.3.3.** Sous les Hypothèses A.3.1, on a pour tout  $n \in \mathbb{N}^*$ ,

$$\lim_{N \rightarrow \infty} \tau_N^n = \infty, \quad \text{et} \quad T_\infty^n = \infty \text{ ps}$$

et l'Équation (A.12) admet une unique

*Démonstration.* Soit  $n \in \mathbb{N}^*$  fixé. La preuve est analogue à celle de la Section 4.2.1. On obtient ainsi :

$$\mathbb{E}\left(\sup_{s \in [0; t \wedge \tau_N^n]} \langle Z_s^n, 1 \rangle\right) \leq \mathbb{E}(\langle Z_s^n, 1 \rangle) e^{ct} \tag{A.13}$$

où  $c$  est une constante indépendante de  $n$  et  $N$ . Avec l'Hypothèse A.3.1 sur la suite  $Z_0^n$ , on en déduit que

$$\mathbb{E}\left(\sup_{s \in [0; t \wedge \tau_N^n]} \langle Z_s^n, 1 \rangle\right) < \infty.$$

L'existence de  $Z_t^n$  est immédiate car  $Z_t^n = \frac{1}{n} Z_t$ , où  $(Z_t)_{t \in \mathbb{R}_+}$  est le processus solution de l'Équation (A.1). On obtient alors l'unicité de la solution de l'Équation (A.12) de manière analogue au raisonnement mené à la Section 4.2.1.

□



Nous allons donc maintenant pouvoir montrer la convergence en loi dans  $\mathbb{D}(\mathbb{R}_+, \mathcal{M}_F(\chi))$  du processus  $(Z^n)_n$ . Nous allons d'abord montrer des propriétés de moment et de martingale pour  $Z^n$ , que nous utiliserons pour établir un résultat de tension, pour enfin prouver la convergence en loi de la suite  $Z^n$ .

Le lemme suivant adapte au nouveau processus le Lemme A.2 de la Section A.2 qui permettait le calcul du générateur, et la démonstration s'adapte directement.

**Lemme A.3.4.** *On rappelle qu'on se place dans le cadre de l'Hypothèse 4.2.1. Soit  $n \in \mathbb{N}^*$  et  $Z^n \in \mathbb{D}(\mathbb{R}_+, \mathcal{M}^n)$  solution de l'Équation (A.12). Soit  $F \in \mathcal{C}^1(\mathbb{R})$ , et  $f : (x, a, s) \rightarrow f_s(x, a) \in \mathcal{B}(\mathbb{R}^2 \times \mathbb{R}_+ \times \mathbb{R}_+, \mathbb{R})$  telle que  $\forall x \in \mathbb{R}^2, (s, a) \rightarrow f_s(x, a) \in \mathcal{C}^{1,1}(\mathbb{R}_+ \times \mathbb{R}_+, \mathbb{R})$ , avec des dérivées bornées uniformément en  $x$ .*

*Alors,  $\forall t \in \mathbb{R}_+$ ,*

$$\begin{aligned}
F(\langle Z_t^n, f_t \rangle) &= F(\langle Z_0^n, f_0 \rangle) + \int_0^t \langle Z_u^n, \frac{df_u}{du} + v \cdot \nabla_a f_u \rangle F'(\langle Z_u^n, f_u \rangle) du \\
&\quad + \int_0^t \int_{E_1} 1_{\{i \leq N_{s-n}\}} 1_{\{\theta \leq b(X_i(Z_s^n), Z_s)\}} 1_{\{X_i(Z_s^n) + z \in C_{X_i(Z_s^n), Z_s^n}\}} \\
&\quad \left[ F \left( \langle Z_s^n, f_s \rangle + \frac{1}{n} f_s(X_i(Z_s^n) + z, a_0) \right) - F(\langle Z_s^n, f_s \rangle) \right] \times M_1(ds, di, d\theta, dz) \\
&\quad - \int_0^t \int_{E_3} 1_{\{i \leq N_{s-n}\}} 1_{\{\theta \leq m(A_i(Z_s^n))\}} \\
&\quad \left[ F \left( \langle Z_s^n, f_s \rangle - \frac{1}{n} f_s(X_i(Z_s^n), A_i(Z_s^n)) \right) - F(\langle Z_s^n, f_s \rangle) \right] \times M_3(ds, di, d\theta).
\end{aligned} \tag{A.14}$$

### A.3.1 Propriété de moment

Dans cette section, on énonce deux résultats sur le contrôle des moments, qui vont être utiles pour montrer la tension de la suite  $(Z^n)$  à la section suivante.

**Théorème A.3.5.** *Soit  $(Z^n)_{n \in \mathbb{N}^*}$  la suite de solutions de l'Équation (A.12).*

*S'il existe  $q \geq 1$  tel que*

$$\sup_{n \in \mathbb{N}^*} \mathbb{E}(\langle Z_0^n, 1 \rangle^q) < \infty,$$

*alors,*

$$\forall T > 0, \sup_{n \in \mathbb{N}^*} \mathbb{E} \left( \sup_{t \in [0; T]} \langle Z_t^n, 1 \rangle^q \right) < \infty.$$

*Démonstration.* Comme pour le résultat analogue de la Section 4.2.1, on ne va tenir compte que des termes de naissance, qui font croître la population.

Soit  $N > 0, n \in \mathbb{N}^*$  et  $t \in [0; T]$ . On applique de Lemme A.3.4 avec  $\forall x \in \mathbb{R}, F(x) = x^q$  et  $f \equiv 1$ . Comme  $\nabla_a f = 0$  et  $f$  est indépendant du temps, on obtient :

$$\begin{aligned}
\sup_{u \in [0; t \wedge \tau_N^n]} \langle Z_u^n, 1 \rangle &\leq \langle Z_0^n, 1 \rangle + \int_0^{t \wedge \tau_N^n} \int_{E_1} 1_{\{i \leq N_{s-n}\}} 1_{\{\theta \leq b(X_i(Z_s^n), Z_s)\}} 1_{\{X_i(Z_s^n) + z \in C_{X_i(Z_s^n), Z_s^n}\}} \\
&\quad \times \left[ \left( \langle Z_s^n, 1 \rangle + \frac{1}{n} \right)^q - (\langle Z_s^n, f_s \rangle)^q \right] \times M_1(ds, di, d\theta, dz)
\end{aligned}$$

De plus, on utilise l'inégalité :

$$\left(x + \frac{1}{n}\right)^q - x^q \leq C(q)(1 + x^{q-1}), \quad \forall x \in \mathbb{R}_+,$$

où  $C(q)$  est une constante qui ne dépend que de  $q$ . Puis, en prenant l'espérance, on obtient la majoration suivante en utilisant la même technique de calcul qu'aux Équations (A.7) :

$$\begin{aligned} \mathbb{E}\left(\sup_{u \in [0; t \wedge \tau_N^n]} \langle Z_u^n, 1 \rangle\right) &\leq E(\langle Z_0^n, 1 \rangle) + E\left[\int_0^{t \wedge \tau_N^n} C(q) (1 + (\langle Z_s^n, 1 \rangle)^{q-1}) \right. \\ &\quad \left. \times \int_{\chi} \left( \int_{\mathbb{R}^2} b(x, Z_s^n) 1_{\{x+z \in C_{x, Z_s^n}\}} d(dz) \right) Z_s^n(dx, da) ds \right]. \end{aligned}$$

Or on a les inégalités :

$$\forall s \in [0; t \wedge \tau_N^n], \quad \mathbb{E}(\langle Z_s^n, 1 \rangle + (\langle Z_s^n, 1 \rangle)^q) \leq 2\mathbb{E}(1 + (\langle Z_s^n, 1 \rangle)^q) \leq 2(1 + \mathbb{E}(\sup_{u \in [0; s]} \langle Z_u^n, 1 \rangle^q))$$

et on applique le théorème de Fubini (les intégrands sont bornés par définition de  $\tau_N^n$ ), ce qui donne :

$$\mathbb{E}\left(\sup_{u \in [0; t \wedge \tau_N^n]} \langle Z_u^n, 1 \rangle^q\right) \leq E(\langle Z_0^n, 1 \rangle) + 2\bar{b}C(q)T + 2\bar{b}C(q) \int_0^t \mathbb{E}\left(\sup_{y \in [0; s \wedge \tau_N^n]} \langle Z_y^n, 1 \rangle^q\right) ds$$

On applique alors le lemme de Gronwall (Lemme A.1.7 de la Section A.1), ce qui donne :

$$\mathbb{E}\left(\sup_{u \in [0; T \wedge \tau_N^n]} \langle Z_t^n, 1 \rangle^q\right) \leq (E(\langle Z_0^n, 1 \rangle) + 2\bar{b}C(q)T) e^{2\bar{b}C(q)T}.$$

Finalement,  $\mathbb{E}\left(\sup_{u \in [0; T \wedge \tau_N^n]} \langle Z_t^n, 1 \rangle^q\right)$  est majoré par une constante  $C(q, T)$  indépendante de  $n \in \mathbb{N}^*$  et  $N$ . On obtient par le lemme de Fatou (en faisant  $N \rightarrow \infty$ ),

$$\sup_{n \in \mathbb{N}^*} \mathbb{E}\left(\sup_{u \in [0; T \wedge \tau_N^n]} \langle Z_t^n, 1 \rangle^q\right) < C(q, T) < \infty,$$

qui était la majoration recherchée. □

Nous allons maintenant établir une deuxième propriété de moment, en utilisant le temps d'arrêt  $\zeta_N^n$  introduit dans la Définition A.3.2.

**Théorème A.3.6.** *Soit  $(Z^n)_{n \in \mathbb{N}^*}$  la suite de solutions de l'Équation (A.12).*

*S'il existe  $q \geq 1$  et  $j \geq 1$  tel que*

$$\sup_{n \in \mathbb{N}^*} \mathbb{E}(\langle Z_0^n, 1 \rangle^j) < \infty, \quad \text{et} \quad \sup_{n \in \mathbb{N}^*} \mathbb{E}\left(\left[\int_{\chi} |a|^q Z_0^n(dx, da)\right]^j\right) < \infty,$$

*alors*

$$\forall T > 0, \quad \sup_{n \in \mathbb{N}^*} \mathbb{E}\left(\sup_{t \in [0; T]} \left[\int_{\chi} |a|^q Z_t^n(dx, da)\right]^j\right) < \infty.$$

*Démonstration.* Soit  $N > 0$ ,  $n \in \mathbb{N}^*$  et  $t \in [0; T]$ . Soit  $\zeta_N^n$  le temps d'arrêt de la Définition A.3.2. On applique le Lemme A.3.4 avec la fonction  $f_t(x, a) : (x, a) \in \mathbb{R}^2 \times \mathbb{R}_+^* \rightarrow |a|^q$ , indépendante du temps  $t$  (on la note  $f$ , et on a  $f(a) = |a|^q$ ), différentiable et on a  $\nabla_a f_u(a) = f'(a) = q|a|^{q-1}$  et  $\frac{df}{du} = 0$ , et avec la fonction  $F : x \in \mathbb{R}_+ \rightarrow x^j \in \mathbb{R}_+$  qui est aussi dérivable de dérivée  $F'(x) = jx^{j-1}$ .

On a alors

$$\begin{aligned} \mathbb{E} \left( \sup_{u \in [0; t \wedge \zeta_N^n]} \left[ \int_{\chi} |a|^q Z_u^n(dx, da) \right]^j \right) &\leq \mathbb{E} \left( \left[ \int_{\chi} |a|^q Z_0^n(dx, da) \right]^j \right) \\ &+ jq \mathbb{E} \left( \int_0^{t \wedge \zeta_N^n} \left[ \int_{\chi} |a|^{q-1} v(a) Z_s^n(dx, da) \right] \left[ \int_{\chi} |a|^q Z_s^n(dx, da) \right]^{j-1} ds \right) \\ &+ \mathbb{E} \left( \int_0^{t \wedge \zeta_N^n} \int_{E_1} \left[ \left[ \int_{\chi} |a|^q Z_s^n(dx, da) + \frac{1}{n} |a_0|^q \right]^j - \left[ \int_{\chi} |a|^q Z_s^n(dx, da) \right]^j \right] \right. \\ &\quad \left. \times 1_{\{i \leq N_{s-}^n\}} 1_{\{\theta \leq b(X_i(Z_s^n), Z_s)\}} 1_{\{X_i(Z_s^n) + z \in C_{X_i(Z_s^n), Z_s^n}\}} \times M_1(ds, di, d\theta, dz) \right). \end{aligned} \quad (\text{A.15})$$

Le dernier terme du membre de droite de (A.15) se traite de la manière que dans le théorème précédent, avec  $\left[ \int_{\chi} |a|^q Z_t^n(dx, da) \right]^j$  à la place de  $\langle Z_t^n, 1 \rangle^q$ . Il est donc majoré par

$$2\bar{b}C(q)T + 2\bar{b}C(q) \int_0^t \mathbb{E} \left( \sup_{y \in [0; s \wedge \tau_N^n]} \left[ \int_{\chi} |a|^q Z_s^n(dx, da) \right]^j \right) ds.$$

Pour le second terme du membre de droite de (A.15), on utilise l'hypothèse sur  $v$  :

$$\exists \bar{v} \forall a \in \mathbb{R}_+, \quad v(a) \leq \bar{v}(1 + a),$$

qui donne la majoration :

$$\begin{aligned} jq\bar{v} \mathbb{E} \left( \int_0^{t \wedge \zeta_N^n} \left[ \int_{\chi} |a|^{q-1} Z_s^n(dx, da) \right] \left[ \int_{\chi} |a|^q Z_s^n(dx, da) \right]^{j-1} \right) \\ + jq\bar{v} \mathbb{E} \left( \int_0^{t \wedge \zeta_N^n} \left[ \int_{\chi} |a|^q Z_s^n(dx, da) \right]^j \right) \end{aligned} \quad (\text{A.16})$$

On majore maintenant le premier terme de (A.16). On a :

$$\begin{aligned} \int_{\chi} |a|^{q-1} Z_s^n(dx, da) &\leq \int_{\chi} (|a| \vee 1)^{q-1} Z_s^n(dx, da) \\ &\leq \int_{\chi} (|a| \vee 1)^q Z_s^n(dx, da) \\ &\leq \langle Z_s^n, 1 \rangle + \int_{\chi} |a|^q Z_s^n(dx, da). \end{aligned}$$

On se sert de cette dernière inégalité pour majorer le premier terme de (A.16) :

$$\left[ \int_{\chi} |a|^{q-1} Z_s^n(dx, da) \right] \left[ \int_{\chi} |a|^q Z_s^n(dx, da) \right]^{j-1} \leq \langle Z_s^n, 1 \rangle \left[ \int_{\chi} |a|^q Z_s^n(dx, da) \right]^{j-1} + \left[ \int_{\chi} |a|^q Z_s^n(dx, da) \right]^j,$$

et comme

$$\langle Z_s^n, 1 \rangle \left[ \int_{\chi} |a|^q Z_s^n(dx, da) \right]^{j-1} \leq \langle Z_s^n, 1 \rangle^j + \left[ \int_{\chi} |a|^q Z_s^n(dx, da) \right]^j,$$

on a finalement, en utilisant le théorème de Fubini, car par définition de  $\zeta_N^n$ , les intégrands sont bornés :

$$\begin{aligned} \mathbb{E} \left( \sup_{u \in [0; t \wedge \zeta_N^n]} \left[ \int_{\chi} |a|^q Z_u^n(dx, da) \right]^j \right) &\leq \mathbb{E} \left( \left[ \int_{\chi} |a|^q Z_0^n(dx, da) \right]^j \right) \\ &+ jq\bar{v}T \mathbb{E} \left( \sup_{y \in [0; t \wedge \zeta_N^n]} \langle Z_y^n, 1 \rangle^j \right) ds + 3jq\bar{v} \int_0^t \mathbb{E} \left( \sup_{y \in [0; s \wedge \zeta_N^n]} \left[ \int_{\chi} |a|^q Z_y^n(dx, da) \right]^j \right) ds \\ &+ 2\bar{b}C(q)T + 2\bar{b}C(q)T \int_0^t \mathbb{E} \left( \sup_{y \in [0; s \wedge \zeta_N^n]} \left[ \int_{\chi} |a|^q Z_y^n(dx, da) \right]^j \right) ds. \end{aligned} \quad (\text{A.17})$$

Finalement, comme on a  $\mathbb{E} \left( \sup_{y \in [0; T \wedge \zeta_N^n]} \langle Z_y^n, 1 \rangle^j \right) \leq C(j, T)$ , la majoration de l'Équation (A.17) devient

$$\begin{aligned} \mathbb{E} \left( \sup_{u \in [0; t \wedge \zeta_N^n]} \left[ \int_{\chi} |a|^q Z_u^n(dx, da) \right]^j \right) &\leq \mathbb{E} \left( \left[ \int_{\chi} |a|^q Z_0^n(dx, da) \right]^j \right) + jq\bar{v}TC(j, T) + 2\bar{b}C(q)T \\ &+ (2\bar{b}C(q) + 3jq\bar{v}) \int_0^t \mathbb{E} \left( \sup_{y \in [0; s \wedge \zeta_N^n]} \left[ \int_{\chi} |a|^q Z_y^n(dx, da) \right]^j \right) ds \end{aligned}$$

Enfin, on applique à nouveau le lemme de Gronwall, ce qui donne :

$\forall t \in [0, T], \forall N > 0, \forall n \in \mathbb{N}^*$  :

$$\begin{aligned} &\mathbb{E} \left( \sup_{u \in [0; t \wedge \zeta_N^n]} \left[ \int_{\chi} |a|^q Z_u^n(dx, da) \right]^j \right) \\ &\leq \left( \left( \sup_{n \in \mathbb{N}^*} \mathbb{E} \left( \left[ \int_{\chi} |a|^q Z_0^n(dx, da) \right]^j \right) \right) + jq\bar{v}TC(j, T) + 2\bar{b}C(q) \right) \times e^{3jqT\bar{v} + 2\bar{b}C(q)} := D(q, j, T). \end{aligned}$$

La constante  $D(q, j, T)$  est indépendante de  $N$  et  $n$ , ce qui prouve le théorème.

En raisonnant par l'absurde, de manière analogue à la Proposition 4.2.2, Section 4.2.1 pour  $\tau_N^n$ , on montre que

$$\lim_{N \rightarrow \infty} \zeta_N^n = \infty \text{ ps.}$$

□

### A.3.2 Décomposition du processus avec une martingale

La décomposition obtenue dans le Théorème A.3.7 sera utilisée pour montrer la tension de la suite  $Z^n$  à la section suivante.

On note  $\mathcal{C}_b^{0,1,1}(\mathbb{R}^2 \times \mathbb{R}_+ \times \mathbb{R}_+)$  l'ensemble des fonctions continues bornées avec des dérivées partielles bornées par rapport aux variables  $a$  de  $\mathbb{R}_+$  et  $t$  de  $\mathbb{R}_+$ .

**Théorème A.3.7.** *Soit  $n \in \mathbb{N}^*$  fixé. Supposons que la condition initiale  $Z_0^n \in \mathcal{M}^n$  soit telle que  $\mathbb{E}(\langle Z_0^n, 1 \rangle^2) < \infty$ , et  $\mathbb{E} \left( \left[ \int_{\chi} |a| Z_0^n(dx, da) \right] \right) < \infty$*

*Pour tout  $f \in \mathcal{C}_b^{0,1,1}(\chi \times \mathbb{R}_+)$ ,*

$$\begin{aligned} M_t^{n,f} = & \langle Z_t^n, f_t \rangle - \langle Z_0^n, f_0 \rangle - \int_0^t \int_{\chi} [v(a) \nabla_a f_s(x, a) + \frac{df_s}{ds}(x, a) \\ & + \int_{\mathbb{R}^2} f_s(x+z, a_0) b(x, Z_s^n) 1_{\{x+z \in C_{x, Z_s^n}\}} d(dz) - f_s(x, a) m(a)] Z_s^n(dx, da) ds \end{aligned} \quad (\text{A.18})$$

*est une martingale de carré intégrable, et de variation quadratique prévisible :*

$$\langle M_t^{n,f} \rangle = \frac{1}{n} \int_0^t \int_{\chi} \left[ \int_{\mathbb{R}^2} f_s^2(x+z, a_0) b(x, Z_s^n) 1_{\{x+z \in C_{x, Z_s^n}\}} d(dz) + f_s^2(x, a) m(a) \right] Z_s^n(dx, da) ds. \quad (\text{A.19})$$

*Démonstration.* Le Lemme A.3.4 appliqué avec  $F = \text{Identité}$  donne :

$$\begin{aligned} \langle Z_t^n, f_t \rangle = & \langle Z_0^n, f_0 \rangle + \int_0^t \langle Z_u^n, \frac{df_u}{du} + v \cdot \nabla_a f_u \rangle du \\ & + \frac{1}{n} \int_0^t \int_{E_1} 1_{\{i \leq N_{s-}^n\}} 1_{\{\theta \leq b(X_i(Z_s^n), Z_s^n)\}} 1_{\{X_i(Z_s^n) + z \in C_{X_i(Z_s^n), Z_s^n}\}} f_s(X_i(Z_s^n) + z, a_0) M_1(ds, di, d\theta, dz) \\ & - \frac{1}{n} \int_0^t \int_{E_3} 1_{\{i \leq N_{s-}^n\}} f_s(X_i(Z_s^n), A_i(Z_s^n)) \times 1_{\{\theta \leq m(A_i(Z_s^n))\}} M_3(ds, di, d\theta). \end{aligned} \quad (\text{A.20})$$

On a donc par définition de  $M_t^{n,f}$  en (A.18)

$$\begin{aligned} M_t^{n,f} = & \frac{1}{n} \int_0^t \int_{E_1} 1_{\{i \leq N_{s-}^n\}} 1_{\{\theta \leq b(X_i(Z_s^n), Z_s^n)\}} 1_{\{X_i(Z_s^n) + z \in C_{X_i(Z_s^n), Z_s^n}\}} f_s(X_i(Z_s^n) + z, a_0) M_1(ds, di, d\theta, dz) \\ & - \frac{1}{n} \int_0^t \int_{E_3} 1_{\{i \leq N_{s-}^n\}} f_s(X_i(Z_s^n), A_i(Z_s^n)) \times 1_{\{\theta \leq m(A_i(Z_s^n))\}} M_3(ds, di, d\theta) \\ & + \int_0^t \int_{\chi} [f_s(x, a) m(a) - \int_{\mathbb{R}^2} f_s(x+z, a_0) b(x, Z_s^n) 1_{\{x+z \in C_{x, Z_s^n}\}} d(dz)] Z_s^n(dx, da) ds. \end{aligned}$$

Donc,

$$\begin{aligned} M_t^{n,f} = & \frac{1}{n} \int_0^t \int_{E_1} 1_{\{i \leq N_{s-}^n\}} 1_{\{\theta \leq b(X_i(Z_s^n), Z_s^n)\}} 1_{\{X_i(Z_s^n) + z \in C_{X_i(Z_s^n), Z_s^n}\}} f_s(X_i(Z_s^n) + z, a_0) \tilde{M}_1(ds, di, d\theta, dz) \\ & - \frac{1}{n} \int_0^t \int_{E_3} 1_{\{i \leq N_{s-}^n\}} f_s(X_i(Z_s^n), A_i(Z_s^n)) \times 1_{\{\theta \leq m(A_i(Z_s^n))\}} \tilde{M}_3(ds, di, d\theta). \end{aligned}$$

où  $\tilde{M}_1$  et  $\tilde{M}_3$  sont respectivement les mesures de Poisson compensées des mesures  $M_1$  et  $M_3$ . Ainsi,  $(M_t^{n,f})_{t \in \mathbb{R}_+}$  est une martingale locale associée à la suite de temps d'arrêt  $(\tau_N^n)_N$  (introduits à la Définition A.3.2).

Nous calculons maintenant sa variation quadratique. Pour cela, nous utilisons le Théorème A.1.5 qui donne directement le résultat souhaité. Nous avons donc montré que la martingale  $M^{n,f}$  est de carré intégrable et de variation quadratique donnée par l'Équation (A.19).  $\square$

### A.3.3 Tension de la suite $(Z^n)$

Cette partie montre la tension de la suite  $(Z^n)$ , qui permet de montrer la convergence en loi de la suite  $Z^n$  vers la solution d'une Équation différentielle.

Dans cette section, on considère les hypothèses suivantes :

**Hypothèse A.3.8.** *On suppose que la suite de variable aléatoire  $(Z_0^n)_{n \in \mathbb{N}^*} \in \mathcal{M}^n$  converge en loi dans  $(\mathcal{M}_F, e)$  (muni de la topologie étroite), vers une mesure  $\zeta_0 \in \mathcal{M}_F$ . On suppose que les conditions de moment suivantes sont satisfaites :*

$$\sup_{n \in \mathbb{N}^*} \mathbb{E}(\langle Z_0^n, 1 \rangle) < \infty, \quad \text{et} \quad \sup_{n \in \mathbb{N}^*} \mathbb{E} \left( \left[ \int_{\chi} |a| Z_0^n(dx, da) \right] \right) < \infty$$

Pour la fin de cette partie, afin d'établir les deux théorèmes suivants, nous devons ajouter des hypothèses de continuité. Elles vont concerner le terme

$$1_{\{\theta \leq b(X_i(Z_s^n), Z_s^n)\}} 1_{\{X_i(Z_s^n) + z \in C_{X_i(Z_s), Z_s}\}}$$

de l'Équation (A.12). Jusqu'à présent la fonction  $(x, Z) \in \mathbb{R}^2 \times \mathcal{M}_F(\chi) \rightarrow b(x, Z)$  correspondant au taux de naissance était simplement supposée bornée par l'Hypothèse 4.2.1. La fonction  $b$  du modèle définie à l'Équation (4.1) (Section 4.1.1) n'est pas continue. Dans la suite nous noterons encore  $b$  le taux de naissance de l'Équation (A.12) mais il ne désignera plus celui défini par l'indicatrice de l'Équation (4.1). On considère l'hypothèse suivante pour le taux de naissance  $b$  :

**Hypothèse A.3.9.** *La fonction  $(x, Z) \in \mathbb{R}^2 \times \mathcal{M}_F(\chi) \rightarrow b(x, Z)$  est continue. On suppose de plus que  $b$  est bornée :*

$$\exists C_b, \forall x \in \mathbb{R}^2, \forall Z \in \mathcal{M}_F(\chi), b(x, Z) \leq C_b.$$

Ensuite, on pose

$$\forall z \in \mathbb{R}^2, \forall x \in \mathbb{R}^2, \forall Z \in \mathcal{M}_F(\chi), h_1(x, z, Z) := 1_{\{x+z \in C_{x,Z}\}}.$$

Soit  $h$  une fonction de  $\mathbb{R}^2 \times \mathbb{R}^2 \times \mathcal{M}_F(\chi)$  dans  $\mathbb{R}$ , on considère l'hypothèse suivante :

**Hypothèse A.3.10.** *La fonction  $(z, x, Z) \in \mathbb{R}^2 \times \mathbb{R}^2 \times \mathcal{M}_F(\chi) \rightarrow h(x, z, Z)$  est continue. On suppose de plus  $h$  bornée :*

$$\exists C_h, \forall z \in \mathbb{R}^2, \forall x \in \mathbb{R}^2, \forall Z \in \mathcal{M}_F(\chi), h(x, z, Z) \leq C_h.$$

La fonction  $h_1$  issue du modèle (terme de compétition intra-spécifique) et utilisée jusqu'ici ne vérifie pas l'Hypothèse A.3.10. Nous modifions donc désormais l'Équation (A.12) en remplaçant le terme

$$1_{\{\theta \leq b(X_i(Z_s^n), Z_s^n)\}} 1_{\{X_i(Z_s^n) + z \in C_{X_i(Z_s), Z_s}\}},$$

provenant de la partie de naissance, par

$$1_{\{\theta \leq b(X_i(Z_s^n), Z_s^n)\}} h(x, z, Z),$$

où  $b$  et  $h$  vérifient respectivement les Hypothèses A.3.9 et A.3.10

Finalement, la fonction

$$(x, z, Z) \in \mathbb{R}^2 \times \mathbb{R}^2 \times \mathcal{M}_F(\chi) \rightarrow b(x, Z).h(x, z, Z) \text{ est continue}$$

et on conserve la notation  $\bar{b}$  pour désigner la constante telle que :

$$\forall z \in \mathbb{R}^2, \forall x \in \mathbb{R}^2, \forall Z \in \mathcal{M}_F(\chi), b(x, Z).h(x, z, Z) \leq \bar{b}.$$

Enfin, on ajoute une hypothèse de continuité à la fonction  $m$ , et on rappelle l'hypothèse sur  $v$  :

**Hypothèse A.3.11.** *On suppose que  $m$  est continue sur  $\mathbb{R}_+$  et qu'il existe une constante  $\bar{m}$  telle que*

$$\forall a \in \mathbb{R}^+, m(a) \leq \bar{m}.$$

*Enfin, on rappelle l'hypothèse sur l'évolution de la biomasse :*

$$\exists \bar{v} \forall a \in \mathbb{R}_+, v(a) \leq \bar{v}(1 + a).$$

**Théorème A.3.12.** *Sous les Hypothèses A.3.8 à A.3.11 énoncées ci-dessus, la suite  $(Z^n)_{n \in \mathbb{N}^*}$  solution de (A.12) est tendue sur  $\mathbb{D}(\mathbb{E}_+, (\mathcal{M}_F(\chi), e))$ .*

*Les valeurs d'adhérence de  $(Z^n)_{n \in \mathbb{N}^*}$  sont des processus à valeurs mesures continus  $\zeta$ , satisfaisant l'équation :*

$$\begin{aligned} \langle \xi_t, f_t \rangle = & \langle \xi_0, f_0 \rangle + \int_0^t \int_{\chi} [v(a) \nabla_a f_s(x, a) + \frac{df_s}{ds}(x, a) \\ & + \int_{\mathbb{R}^2} f_s(x + z, a_0) b(x, \xi_s).h(x, z, \xi_s) d(dz) - f_s(x, a) m(a)] \xi_s(dx, da) ds \end{aligned} \quad (\text{A.21})$$

Les techniques utilisées dans cette preuve sont tirées de Tran (2006) et Fournier and Méléard (2004).

*Démonstration.* Nous commençons par montrer que la suite  $(Z^n)_{n \in \mathbb{N}^*}$  solution de (A.12) est tendue sur  $\mathbb{D}(\mathbb{E}_+, (\mathcal{M}_F, v))$  (on considère la topologie de la convergence vague sur  $\mathcal{M}_F$ ). D'après le critère de Roelly-Coppoletta (Théorème A.1.10), il suffit de montrer que pour toute fonction  $f : (x, a) \rightarrow f(x, a)$  appartenant à un ensemble de fonctions denses dans  $\mathcal{C}_0(\chi, \mathbb{R})$  des applications continues de limite 0 lorsque  $|(x, a)| \rightarrow \infty$  (où  $|\cdot|$  est une norme sur  $\mathbb{R}^2 \times \mathbb{R}_+$ ), la suite  $(Z^{n,f})_{n \in \mathbb{N}^*} = \langle Z^n, f \rangle$  est tendue dans  $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$ . On prend donc  $f \in \mathcal{C}_b^{0,1}$  qui est dense dans  $\mathcal{C}_0(\chi, \mathbb{R})$ .

Or, avec le Théorème A.3.7, nous avons la décomposition suivante pour  $Z^{n,f}$  :

$$Z_t^{n,f} = M_t^{n,f} + V_t^{n,f},$$

où

$$\begin{aligned} M_t^{n,f} = & \langle Z_t^n, f_t \rangle - \langle Z_0^n, f_0 \rangle - \int_0^t \int_{\chi} [v(a) \nabla_a f_s(x, a) + \frac{df_s}{ds}(x, a) \\ & + \int_{\mathbb{R}^2} f_s(x + z, a_0) b(x, Z_s^n).h(x, z, Z_s^n) d(dz) - f_s(x, a) m(a)] Z_s^n(dx, da) ds \end{aligned} \quad (\text{A.22})$$

et

$$\begin{aligned} V_t^{n,f} = Z_t^{n,f} - M_t^{n,f} &= \langle Z_0^n, f_0 \rangle + \int_0^t \int_{\mathcal{X}} [v(a) \nabla_a f(x, a) - f_s(x, a) m(a) \\ &\quad + \int_{\mathbb{R}^2} f_s(x+z, a_0) b(x, Z_s^n) h(x, z, Z_s^n) d(dz)] Z_s^n(dx, da) ds \end{aligned} \quad (\text{A.23})$$

Donc, par les critères de Aldous et Rebolledo énoncés en Théorème A.1.11 et A.1.12, il suffit de montrer que :

- 1 -  $\forall t \in \mathcal{T}$ ,  $\mathcal{T}$  dense dans  $\mathbb{R}_+$ , les lois de  $(\langle M^{n,f} \rangle_t)_{n \in \mathbb{N}^*}$  et  $(V_t^{n,f})_{n \in \mathbb{N}^*}$  sont tendues sur  $\mathbb{R}$ .
- 2-  $\forall T \geq 0$ ,  $\forall \epsilon > 0$ ,  $\forall \eta > 0$ ,  $\exists \delta > 0$ ,  $n_0 \in \mathbb{N}$ , tels que :

$$\sup_{n \geq n_0} \mathbb{P}(|\langle M^{n,f} \rangle_{T_n} - \langle M^{n,f} \rangle_{S_n}| \geq \eta, T_n < S_n + \delta) \leq \epsilon \quad (\text{A.24})$$

Et

$$\sup_{n \geq n_0} \mathbb{P}(|V_{T_n}^{n,f} - V_{S_n}^{n,f}| \geq \eta, T_n < S_n + \delta) \leq \epsilon. \quad (\text{A.25})$$

Pour toute suite de paire de temps d'arrêt  $(S_n, T_n)_{n \in \mathbb{N}^*}$  de la filtration  $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$ , tels que  $S_n \leq T_n \leq T$ .

On commence par montrer 1 :

Il suffit de montrer que

$$\forall T > 0, \sup_{n \in \mathbb{N}^*} \mathbb{E}(\sup_{t \in [0, T]} |\langle M^{n,f} \rangle_t|) < \infty \quad \text{et} \quad \sup_{n \in \mathbb{N}^*} \mathbb{E}(\sup_{t \in [0, T]} |V_t^{n,f}|) < \infty$$

Pour la partie martingale, on utilise l'Équation (A.19) :

$$\sup_{n \in \mathbb{N}^*} \mathbb{E}(\sup_{t \in [0, T]} |\langle M^{n,f} \rangle_t|) \leq T \|f\|_{\infty}^2 (\bar{b} + \bar{m}) \sup_{n \in \mathbb{N}^*} \left( \frac{1}{n} \mathbb{E} \left( \sup_{t \in [0, T]} \langle Z_t^n, 1 \rangle \right) \right)$$

Cette dernière quantité est finie d'après le Théorème A.3.5.

De même, avec (A.23), on a la majoration suivante pour le terme à variations finies :

$$\begin{aligned} \sup_{n \in \mathbb{N}^*} \mathbb{E}(\sup_{t \in [0, T]} |V_t^{n,f}|) &\leq T (\|f\|_{\infty} (1 + \bar{b} + \bar{m}) + \bar{v} \|\nabla_a f\|_{\infty}) \sup_{n \in \mathbb{N}^*} \mathbb{E} \left( \sup_{t \in [0, T]} \langle Z_t^n, 1 \rangle \right) \\ &\quad + T \bar{v} \|\nabla_a f\|_{\infty} \sup_{n \in \mathbb{N}^*} \mathbb{E} \left( \sup_{t \in [0, T]} \int_{\mathcal{X}} |a| Z_t^n(dx, da) \right) \end{aligned}$$

Donc, d'après le Théorème A.3.6,  $\sup_{n \in \mathbb{N}^*} \mathbb{E}(\sup_{t \in [0, T]} |V_t^{n,f}|)$  est finie.

Nous allons maintenant prouver le deuxième point.

Soit donc  $T > 0$ ,  $\epsilon > 0$  et  $\eta > 0$ . Soit une paire de temps d'arrêt  $(S_n, T_n)_{n \in \mathbb{N}^*}$ , et  $\delta > 0$  tels que  $0 \leq S_n \leq T_n \leq T$  et  $T_n \leq S_n + \delta$ .

On déduit de (A.19) et du Théorème A.3.5, la majoration suivante :



$$\begin{aligned}
\mathbb{E}(|\langle M^{n,f} \rangle_{T_n} - \langle M^{n,f} \rangle_{S_n}|) &\leq \mathbb{E} \left( \frac{1}{n} \int_{S_n}^{T_n} \|f\|_\infty^2 (\bar{b} + \bar{m}) \langle Z_s^n, 1 \rangle ds \right) \\
&\leq \frac{\delta \|f\|_\infty^2 (\bar{b} + \bar{m})}{n} \mathbb{E} \left( \sup_{t \in [0, T]} \langle Z_t^n, 1 \rangle \right) \\
&\leq \frac{C(T) \delta}{n}
\end{aligned}$$

On peut donc trouver  $\delta$  et  $n_0$  qui satisfont (A.24).

De même, par (A.23), on obtient la majoration :

$$\begin{aligned}
\mathbb{E}(|V_{T_n}^{n,f} - V_{S_n}^{n,f}|) &\leq \delta (\|f\|_\infty (\bar{b} + \bar{m}) + \bar{v} \|\nabla_a f\|_\infty) \sup_{n \in \mathbb{N}^*} \mathbb{E} \left( \sup_{t \in [0, T]} \langle Z_t^n, 1 \rangle \right) \\
&\quad + \delta \bar{v} \|\nabla_a f\|_\infty \sup_{n \in \mathbb{N}^*} \mathbb{E} \left( \sup_{t \in [0, T]} \int_\chi |a| Z_t^n(dx, da) \right)
\end{aligned}$$

Donc par le Théorème A.3.6,  $\mathbb{E}(|V_{T_n}^{n,f} - M_{S_n}^{n,f}|)$  est majoré par une constante qui ne dépend pas de  $n$ . On peut donc trouver  $\delta$  et  $n_0$  qui satisfont (A.25).

Par le Théorème de Prokhorov A.1.9, on peut extraire de  $(Z^n)_{n \in \mathbb{N}^*}$  une sous-suite convergente en loi dans  $\mathbb{D}(\mathbb{R}_+, (\mathcal{M}_F, v))$ . On note alors  $(Z^{\phi(n)})_{n \in \mathbb{N}^*}$  cette sous suite, et  $Z$  un processus de la loi limite.

En raisonnant comme précédemment, on peut montrer la tension de la suite  $(\langle Z^n, 1 \rangle)_{n \in \mathbb{N}^*}$  dans  $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$ . On choisit donc une suite extraite de  $(Z^n)_{n \in \mathbb{N}^*}$  de sorte que  $\langle Z_{n \in \mathbb{N}^*}^{\phi(n)}, 1 \rangle$  converge en loi vers  $\langle Z, 1 \rangle$  dans  $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$ .

Or, comme on a

$$\sup_{t \in \mathbb{R}_+} \sup_{f \in \mathcal{C}_b^{0,1}(\chi)} \left| \langle Z_t^{\phi(n)}, f \rangle - \langle Z_{t-}^{\phi(n)}, f \rangle \right| \leq \frac{C \|f\|_\infty}{\phi(n)}, \quad \mathbb{P} - ps$$

où  $C$  est une constante indépendante de  $n$ , le processus limite  $Z$  est presque sûrement continu. Donc, par le théorème issu de Méléard and Roelly (1993), la suite  $(Z^{\phi(n)})_{n \in \mathbb{N}^*}$  converge en loi dans  $\mathbb{D}(\mathbb{R}_+, (\mathcal{M}_F, ))$  vers un processus continu. Enfin, le théorème de Prokhorov assure la tension de  $(Z^n)_{n \in \mathbb{N}^*}$  dans  $\mathbb{D}(\mathbb{R}_+, (\mathcal{M}_F, e))$ .

Nous pouvons maintenant identifier la limite. Pour montrer que le processus limite  $Z$  vérifie presque sûrement l'Équation (A.21), nous allons montrer que presque sûrement, pour tout  $f : (x, a, t) \rightarrow f_t(x, a)$  dans  $\mathcal{C}_b^{0,1,1}(\chi \times \mathbb{R}_+)$ , pour tout  $t \in \mathbb{R}_+$ ,  $\Psi_t(Z)$  défini à l'Équation (A.26) est nulle.

$$\begin{aligned}
\Psi_t(Z) &= \langle Z_t, f_t \rangle - \langle Z_0, f_0 \rangle - \int_0^t \int_\chi [v(a) \nabla_a f_s(x, a) + \frac{df_s}{ds}(x, a) \\
&\quad + \int_{\mathbb{R}^2} f_s(x + z, a_0) b(x, Z_s) h(x, z, Z_s) d(dz) - f_s(x, a) m(a)] Z_s(dx, da) ds \quad (\text{A.26})
\end{aligned}$$

Or, par le Théorème A.3.7, on sait que pour tout  $n \in \mathbb{N}^*$ , le processus  $M_t^{f, \phi(n)} = \Psi_t(Z^{\phi(n)})$  est une martingale de variation quadratique donnée par l'Équation (A.19). Nous avons donc :

$$\forall t \in \mathbb{R}_+, \quad \mathbb{E}(|M_t^{f, \phi(n)}|)^2 \leq \mathbb{E}(|M_t^{f, \phi(n)}|^2)$$

Or  $\mathbb{E}(|M_t^{f,\phi(n)}|^2) = \mathbb{E}(\langle M^{f,\phi(n)} \rangle_t)$ , donc

$$\mathbb{E}(|M_t^{f,\phi(n)}|)^2 \leq \frac{tC}{\phi(n)} \mathbb{E}(\sup_{s \in [0;t]} \langle Z_s^{\phi(n)}, 1 \rangle).$$

Et, par le Théorème A.3.5,  $\mathbb{E}(\sup_{s \in [0;t]} \langle Z_s^{\phi(n)}, 1 \rangle)$  est majoré par une constante indépendante de  $n$ . D'où,

$$\forall t \in \mathbb{R}_+, \lim_{n \rightarrow \infty} \mathbb{E}(|M_t^{f,\phi(n)}|) = 0.$$

Nous allons donc montrer que

$$\lim_{n \rightarrow \infty} \mathbb{E}(|M_t^{f,\phi(n)}|) = \mathbb{E}(|\Psi_t(Z)|),$$

ce qui nous permettra de conclure que  $\forall t \in \mathbb{R}_+, \Psi_t(Z) = 0$   $\mathbb{P}$ -p.s. Or, par ce qui précède, nous savons que  $Z^{\phi(n)}$  converge en loi vers  $Z$  et que  $Z$  est presque sûrement continu. Avec les Hypothèses A.3.9, A.3.10 et A.3.11,  $\Psi_y$  est continue en  $Z$ , et on a alors la convergence en loi :

$$\lim_{n \rightarrow \infty} \Psi_t(Z^{\phi(n)}) = \Psi_t(Z).$$

De plus, on a  $\forall t \in \mathbb{R}_+, \forall z \in \mathbb{D}(\mathbb{R}_+, \mathcal{M}_F)$ ,

$$|\Psi_t(z)| \leq C(t, f) \sup_{s \in [0,t]} (\langle z_s, 1 \rangle + \int_{\chi} |a| z_s(dx, da))$$

Par les Théorèmes A.3.5 et A.3.6, on a l'uniforme intégrabilité de la suite  $(|\Psi_t(Z^{\phi(n)})|)_{n \in \mathbb{N}^*}$ . On peut donc inverser la limite et l'espérance :

$$\mathbb{E}(\lim_{n \rightarrow \infty} \Psi_t(Z^{\phi(n)})) = \lim_{n \rightarrow \infty} \mathbb{E}(\Psi_t(Z^{\phi(n)})),$$

ce qui conclut la preuve. □

Nous allons maintenant montrer l'unicité de la solution de (A.21).

On considère la norme de variation totale définie pour  $\mu_1$  et  $\mu_2 \in \mathcal{M}_F$  par :

$$\|\mu_1 - \mu_2\| = \sup_{f \in L^\infty(\chi), \|f\|_\infty \leq 1} |\langle \mu_1 - \mu_2, f \rangle|$$

**Théorème A.3.13.** *Sous les hypothèses du théorème précédent. On suppose de plus que  $b$  vérifie la condition suivante : il existe  $L_b > 0$  telle que*

$$\forall x \in \mathbb{R}^2, \forall y \in \mathbb{R}^2, \forall (z_1, z_2) \in \mathcal{M}_F, |b(x, z_1)h(x, y, z_1) - b(x, z_2)h(x, y, z_2)| \leq L_b \|z_1 - z_2\|.$$

*Soit une condition initiale  $\xi_0 \in \mathcal{M}_F$  et deux solutions  $(\xi_t^1)_{t \in \mathbb{R}_+}$  et  $(\xi_t^2)_{t \in \mathbb{R}_+}$  de l'Équation (A.21), telles que  $\xi_0^1 = \xi_0^2 = \xi_0$ . Alors,  $\forall t \in \mathbb{R}_+, \xi_t^1(dx, da) = \xi_t^2(dx, da)$   $\mathbb{P}$ -p.s.*

*Démonstration.* Soit  $T > 0$ . Les solutions de L'Équation (A.21) restent presque sûrement de masse finie car  $\forall t \in \mathbb{R}_+$ ,

$$\begin{aligned} \langle \xi_t, 1 \rangle &\leq \langle \xi_0, 1 \rangle + \int_0^t \int_{\chi} b(x, a) \xi_s(dx, da) ds \\ &\leq \langle \xi_0, 1 \rangle + \int_0^t \bar{b} \langle \xi_s(dx, da), 1 \rangle ds \\ &\leq \langle \xi_0, 1 \rangle e^{\bar{b}t} < \infty. \end{aligned}$$

On définit donc

$$A_T(\omega) := \sup_{t \in [0, T]} \langle \xi_t^1(\omega) + \xi_t^2(\omega), 1 \rangle < \infty \quad \mathbb{P}p.s.$$

Soit  $\phi \in \mathcal{C}_b^{0,1}(\chi)$  telle que  $\|\phi\|_{\infty} \leq 1$ .  
 $\forall t \in [0, T]$ , on définit :

$$\forall (x, a) \in \chi, \quad \forall s \in [0, t], \quad f(x, a, s) := \phi(x, A(t, s, a))$$

On a alors (Proposition 2.1.8 de Tran (2006)),

$$\forall t \in \mathbb{R}_+, \forall a \in \mathbb{R}_+, \forall x \in \mathbb{R}^2, \quad \frac{df}{dt}(x, a, t) + v(x, a) \nabla_a f(x, a, t) = 0$$

Nous avons donc, pour  $i \in \{1, 2\}$  :

$$\begin{aligned} \langle \xi_t^i, \phi \rangle &= \int_{\chi} \phi(x, A(t, s, a)) \xi_0^i ds + \int_0^t \int_{\chi} \left[ -\phi(x, A(t, s, a)) m(a) \right. \\ &\quad \left. + \int_{\mathbb{R}^2} \phi(x+z, A(t, s, a_0)) b(x, \xi_s^i) 1_{\{x+z \in C_{x, \xi_s^i}\}} d(dz) \right] \xi_s^i(dx, da) ds \end{aligned}$$

On a donc,  $\forall t \in [0, T]$ ,

$$\begin{aligned} |\langle \xi_t^1 - \xi_t^2, \phi \rangle| &\leq \int_0^t \left| \int_{\chi} \phi(x+z, A(t, s, a_0)) m(a) (\xi_s^1(dx, da) - \xi_s^2(dx, da)) \right| ds \\ &\quad + \int_0^t \left| \int_{\chi} \int_{\mathbb{R}^2} \phi(x+z, A(t, s, a_0)) b(x, \xi_s^1) 1_{\{x+z \in C_{x, \xi_s^1}\}} (\xi_s^1(dx, da) - \xi_s^2(dx, da)) \right| ds \\ &\quad + \int_0^t \left| \int_{\chi} \int_{\mathbb{R}^2} \phi(x+z, A(t, s, a_0)) \left( b(x, \xi_s^1) 1_{\{x+z \in C_{x, \xi_s^1}\}} - b(x, \xi_s^2) 1_{\{x+z \in C_{x, \xi_s^2}\}} \right) \xi_s^2(dx, da) \right| ds \end{aligned}$$

Par hypothèse, il existe  $L_b > 0$  telle que

$$\forall x \in \mathbb{R}^2, \forall y \in \mathbb{R}^2, \forall (z_1, z_2) \in \mathcal{M}_F, \quad |b(x, z_1) 1_{\{x+y \in C_{x, z_1}\}} - b(x, z_2) 1_{\{x+y \in C_{x, z_2}\}}| \leq L_b \|z_1 - z_2\|$$

Donc

$$|\langle \xi_t^1 - \xi_t^2, \phi \rangle| \leq [\bar{m} + \bar{b} + L_b A_T] \int_0^t \|\xi_s^1 - \xi_s^2\| ds$$

Or toute fonction  $\phi \in \mathcal{C}_b(\chi, \mathbb{R})$  est limite d'une suite de fonction de  $\mathcal{C}_b^{0,1}(\chi, \mathbb{R})$ , on a :

$$\forall t \in [0, T], \|\xi_t^1 - \xi_t^2\| \leq C(T) \int_0^t \|\xi_s^1 - \xi_s^2\| ds \quad \mathbb{P} - ps$$

Donc, en appliquant le lemme de Gronwall, on a :

$$\forall t \in [0, T], \|\xi_t^1 - \xi_t^2\| = 0 \quad \mathbb{P} - ps.$$

Comme  $T$  est arbitraire, on en déduit que  $\xi_t^1 = \xi_t^2$   $\mathbb{P}$ -ps. □

En conclusion, sous les hypothèses du Théorème A.3.13, le processus renormalisé défini à l'Équation (A.12) converge en loi vers la solution de l'équation différentielle (A.21) (au sens faible).



## Annexe

# B

## Code utilisé dans le logiciel VIABLAB

Cette annexe présente le code utilisé pour le calcul du noyau de viabilité garanti du modèle de dispersion de la Section 5, réalisé avec le logiciel VIABLAB.

```
/*
 *
 * \author: A. Desilles, LASTRE
 * \brief Fichier contenant la definition d'un modele de viabilite pour les tests
 *
 */

#ifndef TESTDATA_H_
#define TESTDATA_H_

/*****
 * Some model specific parameters can be defined here
 *****/

int i = 0; // entre 0 et N-1
int N = 7;

// propagules
double vect_alpha_R[7] = {0.0014, 0.0014, 0.0014, 0.0014, 0.0014, 0.0014, 0.0014};
double vect_E_A[7] = {0.509, 0.990, 0.785, 0.398, 0.343, 0.267, 0.0133};
double vect_E_R[7] = {0.00405, 0.440, 0.311, 0.956, 0.994, 0.225, 0.994};
double q_c_A = 518* 0.000001; // m^2
double q_c_R = 18838 * 0.001; //m
double q_u_R = 0.0;

// zones
double vect_L[7] = {10500 , 6000 , 8500 , 6700 , 12400 , 6700 , 11400}; //m
double vect_A[7] = {105000 , 63000 , 89250 , 73700 , 142600 , 93899 , 171000}; //m^2
double c_i_star_max = 10.0 ;

// impact crue
double vect_l_R[7] = {0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0};
double vect_l_A[7] = {0.037, 0.037, 0.037, 0.037, 0.037, 0.037, 0.037};
double l_R = 0.0;
double l_A = 0.037;
```

```

// croissance plante
double h = -0.033 ;
double d = 0.097 ;
double e = 1.5 ; // m^2
double a_ini = 0.5 ; // m^2
double aire_envahie_0[7] = {150, 63000, 10994, 9065, 11655, 146, 701} ; //m^2 aires de
    2015 observees (et simulees pour zone 2 et 3)

// reservoir crue
double r_max = 10.0;
double A = 1.89 ;

// dispersion
double mat_dispersion_A[7][7] = {
    {0.9990304913927504, 0.0, 0.001969508607249681, 0.0, 0.0, 0.0, 0.0},
    {0.0, 0.9980638628807595, 0.0029361371192405017, 0.0, 0.0, 0.0, 0.0},
    {0.0, 0.0, 0.9981704621476356, 0.0, 0.0028295378523648116, 0.0, 0.0},
    {0.0, 0.0, 0.0, 0.30911781480141237, 0.6918821851985868, 0.0, 0.0},
    {0.0, 0.0, 0.0, 0.0, 0.999592769784679, 0.0014072302153208226, 0.0},
    {0.0, 0.0, 0.0, 0.0, 0.0, 0.9949570172400305, 0.006042982759969763},
    {0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.9989930698113254} };

double mat_dispersion_R[7][7] = {
    {0.9988476328043903, 0.0, 0.0021523671956100864, 0.0, 0.0, 0.0, 0.0},
    {0.0, 0.9819563876799532, 0.01904361232004846, 0.0, 0.0, 0.0, 0.0},
    {0.0, 0.0, 0.987124783290633, 0.0, 0.013875216709367515, 0.0, 0.0},
    {0.0, 0.0, 0.0, 0.9952287398181922, 0.005771260181807943, 0.0, 0.0},
    {0.0, 0.0, 0.0, 0.0, 0.9991855727105343, 0.0018144272894655064, 0.0},
    {0.0, 0.0, 0.0, 0.0, 0.0, 0.9875149944676882, 0.013485005532312053},
    {0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.9976853994915361} };

int T_chgt_u_U = 2; // = T_ini

// contrainte dynamique (a faire varier)
double contrainte_a_i = vect_A[i]/2 ; //m^2

// contrainte controle (a faire varier)
double small_u_max = 4.0;
double big_U_max =9.0;

// nb propagule (R) max par zone:
// Array( 206.661, 25519.43, 9693.359, 320.526, 50.318, 133.589, 2.686 ) // b_i 2015
//    simules dans OpenMOLE
double vect_propagR_max[7] = {2066.61, 255194.3, 96933.59, 3205.26, 503.18, 1335.89,
    26.86 }; // l'array precedent multiplie par 10

////////////////////////////////////////
/// variable d'etats
////////////////////////////////////////
/*! \var dim
 * \brief State dimension
 */
const int dim=4;

/*! \var discret_type
 * \brief the type of discretization scheme used for the dynamics
 * EE or 1 = Euler Explicit scheme
 * EI or 2 = Euler Implicit
 * RK2I or 3 = RK2 Implicit (RK2 for -F)
 * RK2E or 4 = RK2 Explicit (RK2 for F)
 */
const int discret_type=0;

/*!
 * \var dynType defines the type of dynamics in the model
 * 1 or CC : continuous in time and space

```

```

*      2 or DC : discrete time continuous space
*      2 or DD : discrete time discrete space
*      4 or HD : hybrid TODO
*/
const int dynType=DD;

/*! \var STATE_MIN[dim]
* \brief min values for state vector components
* \see gridParams
*/
double STATE_MIN[dim]={0.0,0.0,0.0,0.0};

/*! \var STATE_MAX[dim]
* \brief max values for state vector components
* \see gridParams
*/
double STATE_MAX[dim]={T_chgt_u_U+1,r_max,1.0,1.0};
// Dependance en i geree dans updateParams

/*!
* \var nbPointsState[dim]
* number of discretization points for the state variable
* \see gridParams
*/
unsigned long long int nbPointsState[dim] = {T_chgt_u_U+2,11,1001,101};

double pas_0 = double( (STATE_MAX[0] - STATE_MIN[0]) / (nbPointsState[0]-1)) ;
double pas_1 = double( (STATE_MAX[1] - STATE_MIN[1]) / (nbPointsState[1]-1)) ;
double pas_2 = double( (STATE_MAX[2] - STATE_MIN[2]) / (nbPointsState[2]-1)) ;
double pas_3 = double( (STATE_MAX[3] - STATE_MIN[3]) / (nbPointsState[3]-1)) ;

/*
*****
* Definition of constraints and target
***** */

/*!
* \brief Function defining the state constraints, corresponds to k(x)
*
* This function defines the set K for admissible states
* @param x state variable
* @return value that caraterise the constraints set
*/

inline double constraintsX_fd( unsigned long long int * x )
{
    double xReel0_t=STATE_MIN[0]+x[0]* pas_0;
    double xReel1_r=STATE_MIN[1]+x[1]* pas_1;
    double xReel2_a=STATE_MIN[2]+x[2]* pas_2;
    double xReel3_b=STATE_MIN[3]+x[3]* pas_3;

    double res = 1.0;
    if( (xReel0_t <= T_chgt_u_U) ) {res = 1.0;} else if ( (xReel2_a <= contrainte_a_i) & (
        xReel0_t > T_chgt_u_U) ) {res = 1.0;} else {res = PLUS_INF;};

    return res;
}

////////////////////////////////////////
///      Contrôles
////////////////////////////////////////

/*! \var dimc
* \brief Control variable dimension
*/
const int dimc=1;

```



```

/*! \var CONTROL_MIN[dimc]
 * \brief minimum values for control vector components
 * \see controlParams
 */
double CONTROL_MIN[dimc]={0};

/*! \var CONTROL_MAX[dimc]
 * \brief maximum values for control vector components
 * \see controlParams
 */
double CONTROL_MAX[dimc]={big_U_max};

/*!
 * \var nbPointsControl[dimc]
 * number of discretization points for the control variable
 * \see controlParams
 */
unsigned long long int nbPointsControl[dimc] = {11};

double pasC_0 = double( (CONTROL_MAX[0] - CONTROL_MIN[0]) / (nbPointsControl[0]-1)) ;

/*!
 * \brief Function defining the mixed constraints
 *
 * This function defines the set U(x) for admissible controls as function of the state
 * @param x state variable
 * @param u control variable
 * @return value that caraterise the constraints set
 */

inline double constraintsXU_fd( unsigned long long int * x, unsigned long long int * u )
{
    // Les controles sont geres dans la fonction dynamique: discretisation differente pour
    // u et U
    double res=1.0;
    return res;
}

////////////////////////////////////
/// Tychastique
////////////////////////////////////
// voir la fonction loadModelData pour la dependance en i
// vecteur tychastique: v = (u_0,a_0,b_0,...,u_i-1,a_i-1,b_i-1,c)

int dimc_ty=2;

float tab [30]; /

double *CONTROL_MIN_ty;
/*! \var CONTROL_MAX_ty[dimc]
 * \brief maximum values for tyches vector components
 * \see controlParams
 */
double *CONTROL_MAX_ty;

unsigned long long *nbPointsControl_ty;

double *pas_ty;

////////////////////////////////////
/// Parametres viablab
////////////////////////////////////

```

```

/*
 * Direction le long de laquelle l'ensemble sera represente par des suites de bits
 */
unsigned long long int dirTramage =1;

/*
 * Selection de la methode de representation de l'ensemble
 * Ce parametre determine quelle classe sera utilisee pour les calculs
 *
 * - BS = BitSet, representation par fonction caracteristique , pour
 *      tout type de calculs
 * - MM = MicroMacro, representation par valeurs reelles, pour les calculs
 *      d'ensembles epigraphiques , associes aux systeme micro-macro
 */
int gridMethod=BS;

/*
 * Selection de l'ensemble a calculer
 */
int setType=VIAB;

/*
 * Nombre de raffinements successifs dans le calcul
 */
int refine= 0; //2;

/*!
 * \var periodic[dim]
 * \brief indicator of periodicity
 * 1= periodic variable
 * 0=non periodic variable
 * \see gridParams
 */
int periodic[dim]={0};

/*!
 * \var dbFileName
 * the name for the main data base file for the roject
 */

string prefix="F_dispersion" ;

/*
 * Definition of the dynamics and associated functions and constants
 */

/*
 * Parametre qui indique si l'ensemble doit etre calcule
 * dans le cas contrainte il devra pouvoir etre charge depuis un fichier pour une etude
 * de
 * trajectoires
 */
int computeSet=1;

int saveBoundary=0;

const int nbTrajs=0;
double initPoints[dim*nbTrajs]={};

// important pour le temps ?
int sortieOKinf[dim]={0,0,0,0};
int sortieOKsup[dim]={0,0,0,0};

////////////////////
// Fonctions utiles pour la dynamique
////////////////////

double dynamic_r(double r, double c) { return min(r - c + A, r_max ) ; };

double p(double c, double c_i_star_max ) { return c/c_i_star_max ; };

```

```

double k_l_p( double c, double l, double c_i_star_max ) { return 1 + (1 - l) * p(c,
    c_i_star_max) ; };

double g_j_A(double a, double u, double c ) { return a * q_c_A * k_l_p(c,l_A,
    c_i_star_max) ; };

double lg_j(double a, double L, double Amax ) {
    double res = 1.0;
    if ( a <= L/Amax ) { res = sqrt(a); } else {res = L * a / Amax;};
    return res ; };

double g_j_R(double a, double u, double c, double L, double Amax ) { return lg_j(a,L,
    Amax) * ( q_c_R * k_l_p(c,l_R,c_i_star_max) + q_u_R * u ) ; };

double f_i(double a, double u, double a_new, double Amax, double aire_envahie_0 ) {
    double res = 1.0;
    if ( u <= (-d/h) ) { res = min(a + a_new + e * ((-d/h) - u), Amax ) ; } else {
        res = min( max( a + a_new + (floor(a/60)+1) * (h * u + d)* (2* sqrt( max(
            aire_envahie_0 -(a + a_new), 0.0) / ( abs( (h * u + d)) ) ) +1.0 ) ,
            0.0 ) , Amax) );};
    return res ; };

////////////////////////////////////
/// Fonction Dynamique
////////////////////////////////////
void dynamics_tych_fd(unsigned long long int * x, unsigned long long int *u,unsigned
    long long int *v, unsigned long long int * image) {

// etat
double xReel0_t=STATE_MIN[0]+x[0]* pas_0;
double xReel1_r=STATE_MIN[1]+x[1]* pas_1;
double xReel2_a=STATE_MIN[2]+x[2]* pas_2;
double xReel3_b=STATE_MIN[3]+x[3]* pas_3;

// controle
double uReel0=CONTROL_MIN[0]+u[0]* pasC_0;

// discretisation differente selon T pour u
double pasC_1 = double( (small_u_max - CONTROL_MIN[0]) / (nbPointsControl[0]-1)) ;
if (xReel0_t > T_chgt_u_U) { uReel0 = CONTROL_MIN[0]+u[0]* pasC_1;};

// perturbation (vecteur)
double vReel_vect[dimc_ty];
for(int k=0;k<dimc_ty; k++)
{
    vReel_vect[k] = CONTROL_MIN_ty[k]+v[k]* pas_ty[k];
}

// perturbation
double perturbation_u[7];
double perturbation_a[7];
double perturbation_b[7];

double c = 0.0;
c = vReel_vect[3*i];

// initialize elements of arrays with v
for ( int j = 0; j <= (i-1); j++ ) {
    perturbation_u[ j ] = vReel_vect[ 3*j ];
    perturbation_a[ j ] = vReel_vect[ 3*j +1 ];
    perturbation_b[ j ] = vReel_vect[ 3*j +2 ];
}

// la borne sur c depend de l'etat x
if (c > xReel1_r) {c=0.0;}; // si la perturbation c n'est pas dans l'ensemble
    admissible, on met la perturbation a 0

// temps
double temp_im0_t = xReel0_t +1 ;

```

```

if ( xReel0_t > T_chgt_u_U) {temp_im0_t = T_chgt_u_U+1;} else {temp_im0_t = xReel0_t
+1;};

// reservoir crue r
double temp_im1_r = dynamic_r(xReel1_r, c) ;

// apport akenes en i
double temp_apport_akene = 0.0;
//int j;

for ( int j = 0 ; j < i ; j++)
{
    temp_apport_akene = temp_apport_akene + g_j_A (perturbation_a[j], perturbation_u[j] ,
    c) * mat_dispersion_A[j][i];
}
temp_apport_akene = temp_apport_akene + g_j_A (xReel2_a, uReel0 , c) * mat_dispersion_A
[i][i] ;

// apport rhizomes en i
double temp_apport_rhizome = 0.0;

for ( int j = 0 ; j < i ; j++)
{
    temp_apport_rhizome = temp_apport_rhizome + ( vect_alpha_R[j] * perturbation_b[j] +
    g_j_R (perturbation_a[j], perturbation_u[j] , c, vect_L[j], vect_A[j]) ) *
    mat_dispersion_R[j][i];
}
temp_apport_rhizome = temp_apport_rhizome + ( vect_alpha_R[i] * xReel3_b + g_j_R (
    xReel2_a, uReel0 , c, vect_L[i], vect_A[i]) ) * mat_dispersion_R[i][i] ;

// aire
double a_new = (temp_apport_rhizome * vect_E_R[i] + temp_apport_akene * vect_E_A[i] ) *
a_ini;
double temp_im2_a = f_i( xReel2_a, uReel0, a_new, vect_A[i], aire_envahie_0[i] ) ;

// accumulateur rhizome
double temp_im3_b = temp_apport_rhizome * (1-vect_E_R[i]) ;

///////// projections
// 0
image[0] = (unsigned long long int)floor( (temp_im0_t - STATE_MIN[0]) / pas_0 );

// 1
image[1] = (unsigned long long int)floor( (temp_im1_r - STATE_MIN[1]) / pas_1 );

// 2
image[2] = (unsigned long long int)floor( (temp_im2_a - STATE_MIN[2]) / pas_2 );

// 3
image[3] = (unsigned long long int)floor( (temp_im3_b - STATE_MIN[3]) / pas_3 );
}

void loadModelData();
void loadModelData(){
/// lecture du fichier avec les bornes superieures pour l'ensemble de perturbations (
    calcule A partir des noyaux precedents)
int TAILLE_MAX = 1000 ;
FILE* fichier = NULL;
char chaine[TAILLE_MAX] = "";
int j;
fichier = fopen("test_perturbation.csv", "r");
j = 0;
if (fichier != NULL)
{
    while (fgets(chaine, TAILLE_MAX, fichier) != NULL)
    {
        tab[j] = atof(chaine);
        j =j+1;
    }
}
}

```

```

    }
    fclose(fichier);
};

dimc_ty=3*i+1;

CONTROL_MIN_ty= new double [dimc_ty];
for(int k=0;k<(dimc_ty-1); k++)
{
    CONTROL_MIN_ty[k]= tab[k];
}
CONTROL_MIN_ty[dimc_ty-1] = 0;

CONTROL_MAX_ty= new double [dimc_ty];
for(int k=0;k<(dimc_ty-1); k++)
{
    CONTROL_MAX_ty[k]= tab[k] ;
}
CONTROL_MAX_ty[dimc_ty-1]= r_max; // perturbation c (pour r)

nbPointsControl_ty = new unsigned long long [dimc_ty];
for(int k=0;k<(dimc_ty-1); k++)
{
    nbPointsControl_ty[k]=1; // remplir dans cette boucle les nombres de valeurs
    possibles des tyches
}
nbPointsControl_ty[dimc_ty-1]=3;

pas_ty = new double [dimc_ty];
for(int k=0;k<dimc_ty; k++)
{
    if (nbPointsControl_ty[k] != 1) { pas_ty[k]= ( (CONTROL_MAX_ty[k] - CONTROL_MIN_ty[k]
    ) / (nbPointsControl_ty[k]-1)) ; } else { pas_ty[k] = 0.0 ; };
}
}

void static updateParams(int local_i, double local_big_U_max, double local_small_u_max,
    double local_contrainte_a_i){
    i = local_i ;
    big_U_max = local_big_U_max ;
    small_u_max = local_small_u_max ;
    contrainte_a_i = local_contrainte_a_i;

    // MAJ Parametres
    // etat

    STATE_MAX[2]= min( 2*max( aire_envahie_0[i], contrainte_a_i ), vect_A[i] ) ;
    STATE_MAX[3]=vect_propagR_max[i];

    pas_2 = double( (STATE_MAX[2] - STATE_MIN[2]) ) / double( (nbPointsState[2]-1)) ;
    pas_3 = double( (STATE_MAX[3] - STATE_MIN[3]) ) / double( (nbPointsState[3]-1)) ;

    // controle
    CONTROL_MAX[0]={big_U_max};
    pasC_0 = double( (CONTROL_MAX[0] - CONTROL_MIN[0]) / (nbPointsControl[0]-1)) ;
}

#include "dispersion_renouee_unused.h"
#endif /* TESTDATA_H_ */

```

## Annexe

# C

### Étude de la diversité des trajectoires produites par Electra 2

Cette annexe vise décrire la mise en oeuvre de l'algorithme de recherche de diversité, présenté à la Section 2.2.2 du Chapitre 2. Ce travail a été réalisé avec Guillaume Chérel et Laurent Bolognini et concerne Électra 2, une structure cinétique lumineuse créée par Laurent Bolognini. La structure est mise en mouvement par des moteurs, et, dans l'obscurité nous voyons par le phénomène de persistance rétinienne des formes produites par la trajectoire des ampoules. Notre objectif est la production d'une cartographie permettant de représenter les formes qu'il est possible de créer à partir d'une sculpture cinétique lumineuse. Ce travail d'exploration sur les formes produites a été rendu accessible par des outils dédiés aux systèmes complexes. Afin de chercher la diversité dans les formes produites par la sculpture, nous utilisons un algorithme d'exploration de modèle. Notre choix s'est porté sur l'algorithme Pattern Space Exploration (PSE Chérel et al. (2015), voir la Section 2.2.2). Pour employer cet algorithme, nous avons modélisé Electra 2. Nous avons choisi de nous intéresser aux phases durant lesquelles les vitesses des moteurs sont constantes. Le modèle a donc comme paramètres les vitesses des moteurs et calcule les trajectoires des ampoules. Pour pouvoir distinguer les formes rares des plus fréquentes, l'algorithme d'exploration a besoin d'une description succincte, en quelques quantités, de chaque forme dessinée. Une partie importante du travail réalisé a été de trouver une telle description permettant de distinguer les formes qui nous semblent différentes à l'oeil nu. Nous avons par exemple choisi comme mesure la courbure moyenne de la trajectoire d'une ampoule, ou la longueur d'un motif (structure élémentaire) qui se répète. Avec ces mesures, l'algorithme nous a permis de trouver les formes qui combinaient ces différentes caractéristiques. Afin de visualiser les résultats obtenus, nous avons construit une carte auto-adaptative (algorithme de Kohonen Kohonen (1982, 2000)) à partir de l'ensemble des figures découvertes par l'algorithme. La carte produite respecte la topologie de l'espace des données (de dimension égale au nombre de mesures de caractéristiques choisies), et rend possible la visualisation des chemins entre les formes. Nous avons ainsi proposé une cartographie originale des figures produites par Electra 2. Cette carte a été présentée sous forme de poster conjointement avec la

sculpture lumineuse (et son alter-ego numérique) lors de l'événement Artex organisé en 2019 à l'ISC-PIF. Ce travail a été également présenté dans le cadre des journées de Rochebrune 2020, dont le thème était "Créativité et systèmes complexes".

Cette annexe est organisée comme suit. La Section C.1 est consacrée à la description de la structure cinétique lumineuse Electra 2 et aux mécanismes considérés, à la présentation du modèle mathématique ainsi qu'aux méthodes utilisées pour le recherche de diversité des formes. Ainsi, nous présentons les mesures sélectionnées pour caractériser les formes produites par la structure, la mise en œuvre de l'algorithme PSE et celle de la méthode de Kohonen utilisée pour cartographier les résultats. Les résultats sont présentés dans la Section C.2. En particulier, nous présentons la carte obtenue. Enfin, nous résumons notre démarche et discutons de perspectives dans la Section C.3.

## C.1 Matériel et méthodes

### C.1.1 Présentation d'Electra 2 et mécanismes modélisés

Electra 2 est une structure cinétique lumineuse créée par Laurent Bolognini. Éclairagiste de formation, Laurent Bolognini est un artiste plasticien qui s'intéresse depuis plus de vingt ans à ces dispositifs d'impression de tracés lumineux. Electra 2 est le résultat de ces années de recherche perfectionnement technologique. La structure est composée d'un grand bras métallique de 2 mètres de long, fixé à un support en son centre, tourne entraîné par un moteur (voir Figure C.1). Chaque moitié de ce bras porte un autre bras motorisé, plus court. Tous les trois tiennent, à chaque extrémité, une petite ampoule. L'artiste pilote la sculpture en actionnant les commandes électriques qui contrôlent la vitesse des trois moteurs, leur sens de rotation et l'intensité des lumières. Dans l'obscurité, quand les ampoules s'allument et que les moteurs s'animent, des traces lumineuses s'impriment dans l'espace. L'œil humain perçoit ces formes par le phénomène de persistance rétinienne. Ce sont d'abord des cercles, mais bien vite des boucles, des pointes, voire des lignes droites qui s'assemblent, se succèdent, se répètent et composent des rythmes et des formes surprenantes et captivantes. La Figure C.2 montre un exemple de tracé lumineux issu d'une photographie. La surprise devant cette diversité des formes générées par des compositions de mouvement circulaires a éveillé notre curiosité et nous avons voulu chercher les formes que la sculpture pouvait dessiner.



FIGURE C.1 – Electra 2

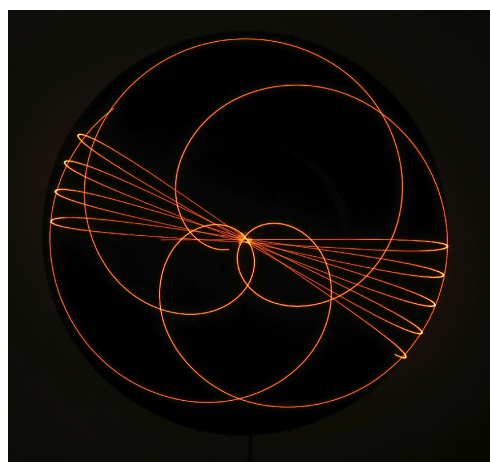


FIGURE C.2 – Electra 2, tracé lumineux

Notre objectif est de trouver les formes qu'il est possible de créer, dans un certain sens, que nous précisons à la Section C.1.6. Pour cela, il est nécessaire de caractériser les formes produites par Electra 2. Nous avons choisi de faire cette étude dans le cadre de la modélisation et de l'exploration de modèles. Pour cela, nous allons commencer par présenter le modèle d'Electra 2 sur lequel se base cette étude. Le modèle donne la position de chaque lumière en fonction du temps. Lors de représentation avec Electra 2, Laurent Bolognini fait varier les vitesses des moteurs au cours du temps. Nous choisissons pour notre étude de nous limiter aux figures obtenues avec des vitesses constantes au cours du temps et une intensité lumineuse fixe.



### C.1.2 Paramètres

La Figure C.3 montre une représentation schématique d'Electra 2, et présente les notations utilisées dans la suite de ce chapitre. Ainsi, les points  $B, C, D, E, F, G$  représentent les ampoules et les rayons des différents axes sont notés  $r_B, r_C, \dots, r_G$ . Les rayons sont fixés par la géométrie de la structure cinétique. La position initiale est donnée par angles :  $\phi_B, \phi_D, \phi_F$ . Nous considérons la position initiale fixée dans toute cette étude et nous nous intéressons uniquement aux vitesses des moteurs, notées  $v_1, v_2, v_3$  (nombre de tours par seconde). Cette vitesse peut être positive ou négative (sens de rotation).

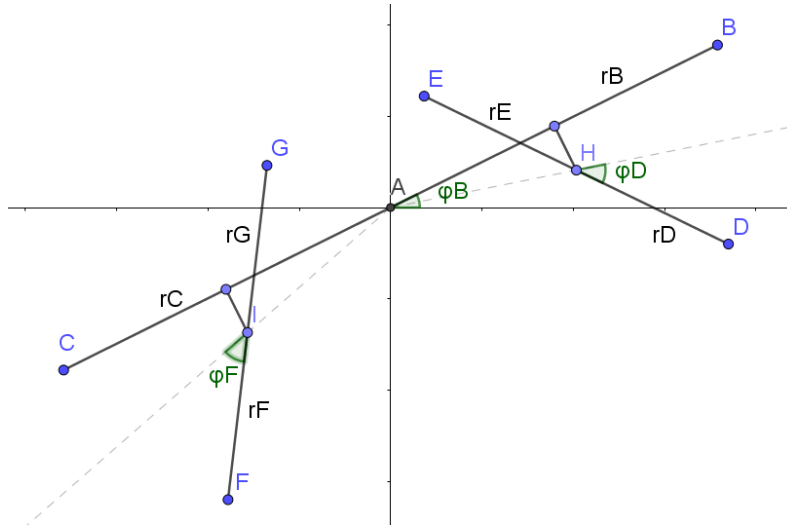


FIGURE C.3 – Schéma d'Electra 2, avec les notations utilisées dans le modèle (réalisé avec le logiciel GeoGebra)

### C.1.3 Formalisme mathématique

Le modèle donne la position des points (ampoules) en fonction du temps et des positions initiales. Cette dernière est déterminée de manière explicite. Par exemple, pour le point  $B$ , nous obtenons :

$$B(t) = (x_B(t), y_B(t)) = (r_B \cos(2\pi v_1 t + \phi_B), r_B \sin(2\pi v_1 t + \phi_B)) \quad (C.1)$$

Ainsi, à  $t = 0$ ,  $B(0) = (r_B \cos(\phi_B), r_B \sin(\phi_B))$ . Nous obtenons une équation analogue pour la position du point  $C, H$  et  $I$ . Pour déterminer la position des points  $D, E, F$  et  $G$ , nous utilisons les points intermédiaires  $H$  et  $I$  ainsi qu'une formule de changement de base ou de la trigonométrie. Par exemple, pour le point  $D$ , cela donne :

$$D(t) = \begin{pmatrix} x_D(t) \\ y_D(t) \end{pmatrix} = \begin{pmatrix} r_H \cos(2\pi v_1 t + \phi_B - \alpha_H) + r_D \cos(2\pi(v_2 + v_1)t + (\phi_D + \phi_B - \alpha_H)) \\ r_H \sin(2\pi v_1 t + \phi_B - \alpha_H) + r_D \sin(2\pi(v_2 + v_1)t + (\phi_D + \phi_B - \alpha_H)) \end{pmatrix},$$

où  $\alpha_H$  est une mesure de l'angle  $(\vec{AH}, \vec{AB})$ . Nous obtenons des formules analogues pour les autres points  $E, F$  et  $G$ . A partir de ces formules, nous pouvons facilement calculer la vitesse et l'accélération de chaque point à chaque instant. Ces quantités seront utilisées à la Section C.1.5 pour le calcul d'une mesure.

### C.1.4 Simulation

Nous avons implémenté le modèle dans le langage Scala. Les entrées du modèle sont les vitesses des moteurs (3 paramètres), ainsi que le fait d'allumer ou éteindre chaque lumière. Nous avons ainsi choisi de fixer la position initiale de la structure. Une sortie du modèle est la trajectoire de chaque lumière allumée. Pour toutes les simulations présentées dans cette partie, la durée de la simulation est fixée (indépendamment de la vitesse) à  $T = 10s$ . La trajectoire est la position des points à chaque pas de temps (fixé) et la sortie du modèle est l'ensemble des trajectoires des points allumés.

Les Figures C.4 et C.5 montrent des exemples de trajectoires simulées (toutes les lumières sont allumées). La première trajectoire a un rapport entier entre les vitesses, la seconde trajectoire est obtenue avec une légère variation de vitesses par rapport à celles de la première.

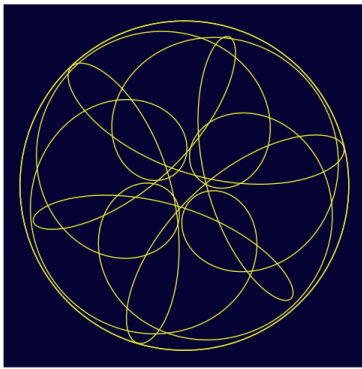


FIGURE C.4 – Trajectoire simulée avec  $v_1 = 1, v_2 = -3, v_3 = 2$ , tous les points

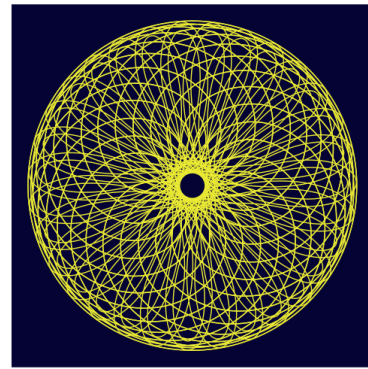


FIGURE C.5 – Trajectoire simulée avec  $v_1 = 1, v_2 = -3.1, v_3 = 2.1$ , tous les points

### C.1.5 Caractérisation des trajectoires

Pour utiliser l'algorithme PSE dans le cadre de recherche de diversité dans les trajectoires produites par Electra 2, nous devons résumer l'information de la trajectoire par des scalaires, appelés mesures (voir Section C.1.6). Nous avons choisi des mesures invariantes par rotations, que nous avons implémentées dans le langage Scala. Nous présentons ici trois mesures choisies pour cette étude : la courbure moyenne (Section C.1.5), le nombre de points singuliers (Section C.1.5) et la longueur d'une structure élémentaire (Section C.1.5). Nous avons aussi implémenté d'autres mesures (densité de la trajectoire, indice de Moran, points doubles), mais nous ne les détaillons pas ici. Dans cette section, nous précisons l'expression analytique de ces mesures, mais nous illustrons également les valeurs qu'elles prennent sur des trajectoires.

#### Courbure moyenne

La courbure mesure la manière dont une courbe s'éloigne localement d'une ligne droite. Pour une courbe paramétrée par  $(x(t), y(t))$ , la courbure est obtenue à partir de la vitesse (là où elle est non nulle) et de l'accélération par la formule :

$$\gamma(t) = \frac{\det \left( \frac{d\vec{M}}{dt}(t), \frac{d^2\vec{M}}{dt^2}(t) \right)}{\left\| \frac{d\vec{M}}{dt}(t) \right\|^3} = \frac{x'(t)y''(t) - y'(t)x''(t)}{(x'(t)^2 + y'(t)^2)^{3/2}}$$

Ainsi, la courbure est définie en chaque point de la trajectoire (de vitesse non nulle). Par exemple, la courbure d'une ligne droite est nulle, et la courbure d'un cercle est constante. Nous choisissons la courbure moyenne (temporelle) absolue, notée  $c$ , comme mesure de diversité. Les Figures C.6, C.7, C.8 et C.9 montrent la courbure moyenne absolue calculée sur 4 trajectoires. Les zones rouges (resp. bleues) correspondent à des valeurs de courbure maximales (resp. minimales).

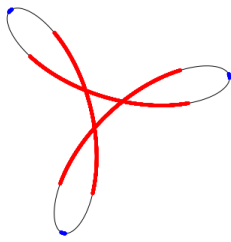


FIGURE C.6 – Illustration de la courbure,  $c = 12,56$

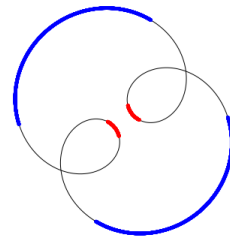


FIGURE C.8 – Illustration de la courbure,  $c = 18,84$

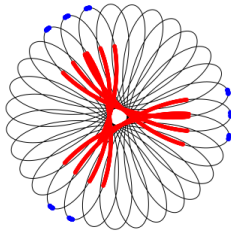


FIGURE C.7 – Illustration de la courbure,  $c = 13,19$

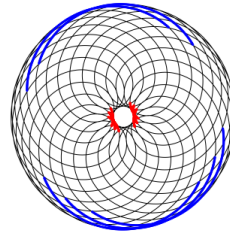


FIGURE C.9 – Illustration de la courbure,  $c = 19,47$

### Nombre de points singuliers

$M(t)$  est dit singulier (ou stationnaire) si sa vitesse est nulle ( $M'(t) = 0$ ).

On obtient par calcul une condition nécessaire sur les vitesses (et rayons) pour qu'un point soit singulier (pour  $D, E, F, G$ ) :

Soit  $t \in \mathbb{R}$ ,  $D'(t) = 0 \implies |v_1|r_H = |v_1 + v_2|r_D$

Cette condition limite l'obtention du nombre de points fixes différents pour une trajectoire, puisque nous considérons des vitesses constantes durant une trajectoire. Ainsi, on fixe un seuil  $c_{speed}$  et on s'intéresse aux points  $M(t)$  qui vérifient  $\|M'(t)\| \leq c_{speed}$ . Nous conservons

l'appellation de points singuliers pour ces points. La valeur de  $c_{speed}$  est fixée empiriquement. La mesure retenue correspond au nombre de points singuliers dans une trajectoire et est notée  $s$ .

Les Figures C.10, C.11, C.12 et C.13 montrent les valeurs de cette mesure du nombre de points singuliers pour 4 trajectoires.

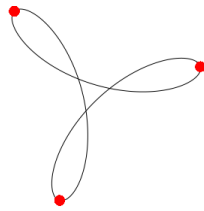


FIGURE C.10 – Illustration du nombre de points singuliers,  $s = 30$

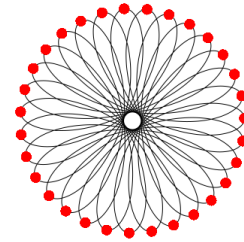


FIGURE C.12 – Illustration du nombre de points singuliers,  $s = 31$

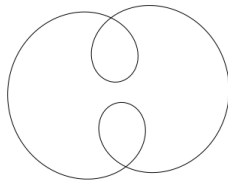


FIGURE C.11 – Illustration du nombre de points singuliers,  $s = 0$

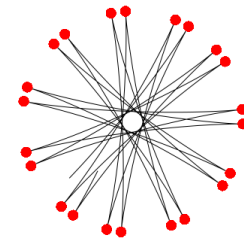


FIGURE C.13 – Illustration du nombre de points singuliers,  $s = 22$

### Longueur d'une structure élémentaire

On appelle structure élémentaire associée à une trajectoire une partie de la courbe (de taille minimale fixée) dont la trajectoire est l'union, en considérant des rotations du motif. La mesure retenue est la plus petite longueur (nombre de pas de temps) d'une structure élémentaire, notée  $m$ , qui peut être égal au nombre de pas de la trajectoire.

Les Figures C.14, C.15, C.16 et C.17 montrent les valeurs de cette mesure de la longueur d'une structure élémentaire pour 4 trajectoires.

## C.1.6 Méthode pour la recherche de diversité dans les trajectoires

Nous utilisons l'algorithme Pattern Space Exploration (PSE, Chérel et al. (2015), voir la Section 2.2.2) pour la recherche de diversité dans l'ensemble des trajectoires produites par Electra 2.

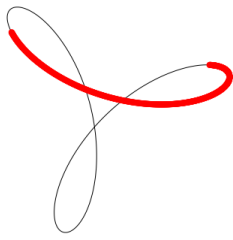


FIGURE C.14 – Illustration de la longueur d'une structure élémentaire,  $l = 666$

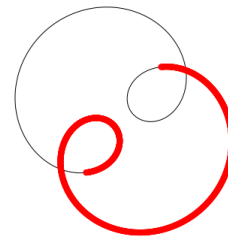


FIGURE C.16 – Illustration de la longueur d'une structure élémentaire,  $l = 1000$

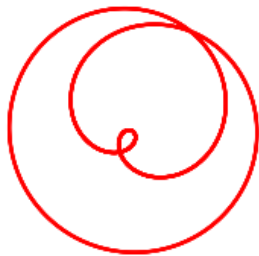


FIGURE C.15 – Illustration de la longueur d'une structure élémentaire,  $l = 4000$

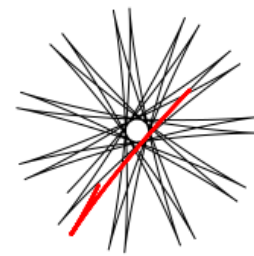


FIGURE C.17 – Illustration de la longueur d'une structure élémentaire,  $l = 905$

L'algorithme utilise une discrétisation de l'espace des sorties en une hit-map, constituée de cellules régulières (au lieu d'utiliser une distance entre les individus). Un motif correspond à une cellule de cette grille. La hitmap est inchangée pendant la phase d'exploration. Pour chaque motif produit par une simulation su modèle, un compteur de la cellule correspondante est incrémenté. La nouveauté d'un motif est de ce fait basée sur l'inverse de la valeur du compteur de la cellule. La Figure C.18, extraite de Chérel et al. (2015), montre les étapes de l'algorithme en tant qu'algorithme évolutionnaire.

Pour utiliser l'algorithme, nous devons choisir

- des paramètres du modèle qui constituent l'espace du génome
- des bornes de l'espace du génome
- un ensemble de mesures qui décrivent le résultat d'une simulation et constituent l'espace des motifs (espace phénotypique)
- une discrétisation de l'espace phénotypique

Pour l'espace du génome, nous choisissons les paramètres de vitesses  $v_2$  et  $v_3$ , et le paramètre qui indique d'allumer la lumière des points  $D$  et  $F$ . Le Tableau C.1 montre les bornes choisies sur ces paramètres. On choisit de fixer  $v_1 = 1.0$ , et les autres lumières sont éteintes.

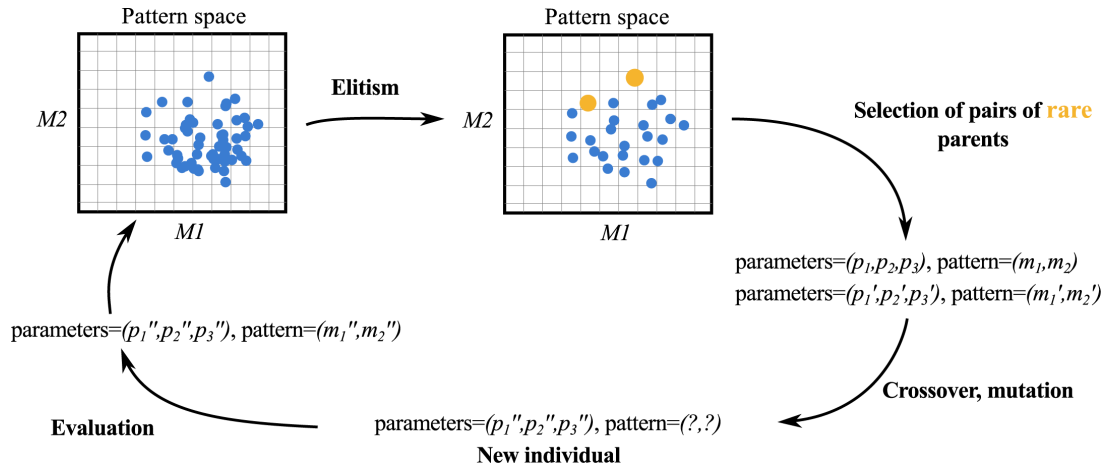


FIGURE C.18 – Extrait de Chérel et al. (2015), étapes de l'algorithme PSE

|     | v2   | v3   | lightD | lightF |
|-----|------|------|--------|--------|
| min | -5.0 | -5.0 | 0.0    | 0.0    |
| max | 5.0  | 5.0  | 1.0    | 1.0    |

TABLE C.1 – Bornes choisies pour les paramètres du modèle dans l'algorithme PSE pour le modèle Electra 2.

L'espace des motifs (espace phénotypique) est constitué des mesures présentées à la Section C.1.5 pour les deux points considérés ( $D$  et  $F$ ). Le Tableau C.2 montre la discrétisation de l'espace des motifs pour le modèle Electra 2. Dans ce tableau,  $s$  désigne le nombre de points singuliers,  $c$  la courbure moyenne et  $m$  la longueur de la structure élémentaire. Les deux premières quantités sont calculées sur une structure élémentaire. L'indice 1 (resp. 2) fait référence au point  $D$  (resp.  $F$ ).

|      | s1  | s2  | c1   | c2   | m1     | m2     |
|------|-----|-----|------|------|--------|--------|
| step | 1.0 | 1.0 | 5.0  | 5.0  | 125.0  | 125.0  |
| min  | 0.0 | 0.0 | 0.0  | 0.0  | 0.0    | 0.0    |
| max  | 5.0 | 5.0 | 50.0 | 50.0 | 2000.0 | 2000.0 |

TABLE C.2 – Discrétisation de l'espace des motifs utilisé dans l'algorithme PSE pour le modèle Electra 2

Nous avons utilisé l'implémentation de PSE dans le logiciel OpenMOLE (logiciel d'exploration de modèles, (Reuillon et al., 2013)) ainsi que du calcul distribué sur la grille de calcul européenne (EGI). L'algorithme requiert de décider de son arrêt. Nous choisissons de l'arrêter quand il ne trouve plus de motif, cela correspond à une asymptote sur la courbe du nombre de motifs découverts en fonction du temps.

### C.1.7 Méthode de représentation sous forme de carte topologique

L'application de l'algorithme PSE retourne dans notre cas des points dans un espace à 6 dimensions (les 6 mesures retenues). De plus, l'algorithme donne également les paramètres (vitesses de moteurs) qui permettent d'obtenir ces points. Pour résumer cette information, nous construisons une carte topologique (aussi appelée carte auto-organisatrice (Kohonen, 2000)) de taille  $10 \times 10$ . La carte produite regroupe les valeurs de motifs pour des trajectoires

"similaires" en un nœud de la carte, de sorte que les trajectoires semblables ont la même projection, et que par contre, plus des trajectoires sont dissemblables, plus elles seront associées à des groupes spatialement éloignés sur la carte. Enfin, nous choisissons de représenter une trajectoire par nœud, en conservant la topologie de la carte de Kohonen (voir Figure C.19).

## C.2 Résultats

L'algorithme PSE a découvert 1813 trajectoires différentes pour Electra 2. La Figure C.19 est la visualisation de la carte de Kohonen produite à partir de ces trajectoires.

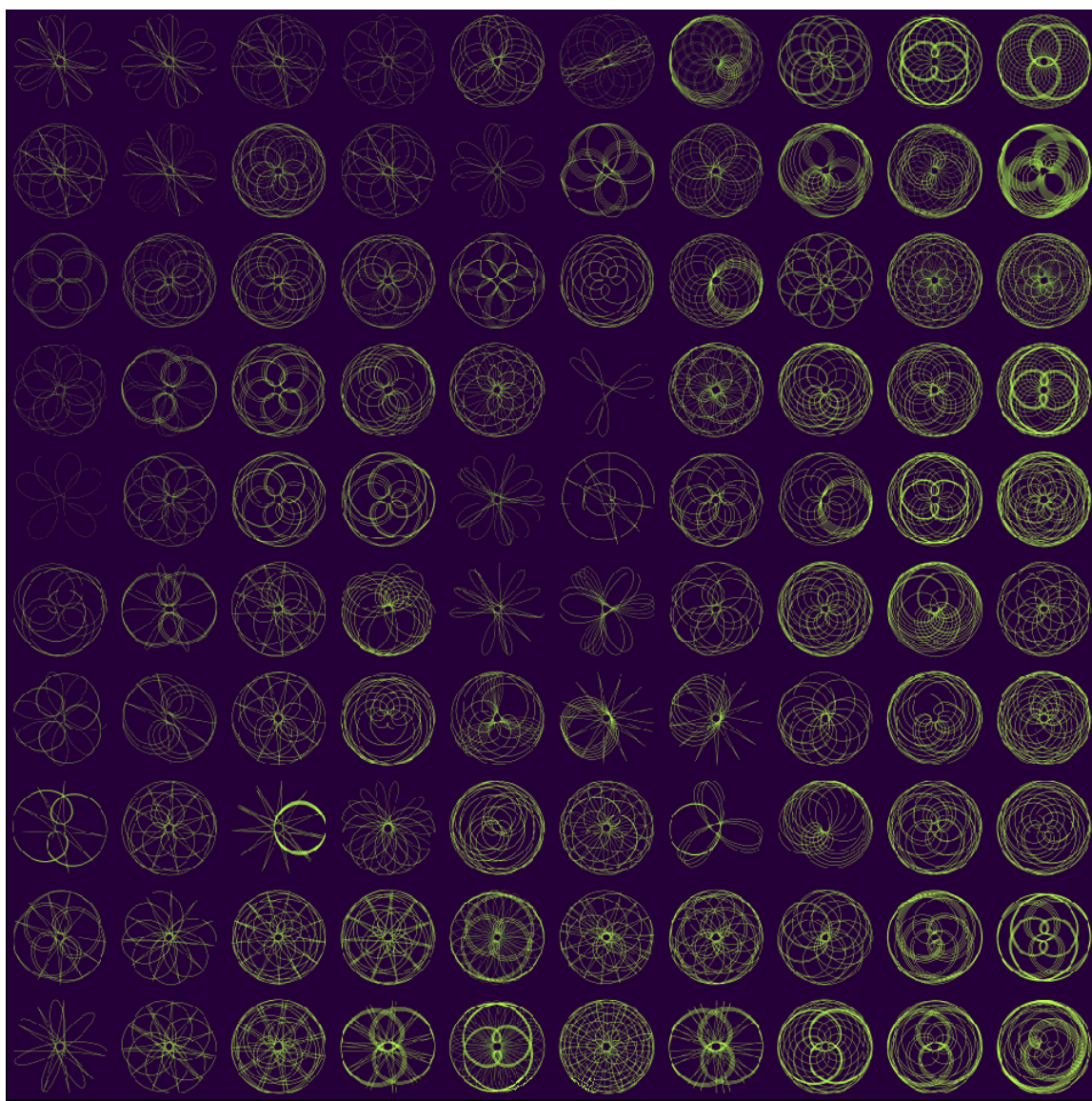


FIGURE C.19 – Carte de Kohonen de 100 trajectoires issues de l'exploration par l'algorithme PSE. L'intensité lumineuse d'une trajectoire est fonction de sa vitesse.



### C.3 Discussion

Nous avons construit une cartographie permettant de représenter les formes qu'il est possible de créer à partir d'Electra 2, une sculpture cinétique lumineuse. Cette cartographie originale des figures a été obtenue par une exploration de modèle. Pour cela nous avons commencé par modéliser la structure cinétique. Nous avons ensuite caractérisé les formes produites à l'aide de mesures. Nous avons utilisé ces dernières pour appliquer un algorithme de recherche de diversité, et à partir du résultat duquel nous avons construit une carte topologique.

Une première perspective à ce travail serait d'obtenir une visualisation de la découverte des motifs issus de l'algorithme PSE par Electra 2, qui prendrait ainsi la forme d'une balade dans les figures d'Electra 2 guidée par l'algorithme PSE. Laurent Bolognini commande souvent Electra 2 dans le cadre de concerts improvisés, où la musique et les formes produites par la structure cinétique dialoguent. Dans la continuité de ce lien entre le travail de l'artiste et la musique, nous pouvons envisager de trouver des mesures pour caractériser une musique, et produire une partition de figures correspondant pour Electra 2.

La méthode utilisée pour cartographier la diversité de figures produites par Electra 2 peut être utilisée dans d'autres contextes. Nous avons débuté une analyse similaire avec Moving Cells (Robin Lamarche-Perrin et Bruno Pace, 2015), une installation numérique proposant aux spectateurs d'interagir via le mouvement avec un nuage de particules virtuelles, ces dernières ayant une dynamique inspirée de la gravitation, le spectateur jouant le rôle du centre de masse. Comme pour Electra 2, nous pouvons remplacer le mouvement d'une personne par une combinaison de paramètres qui décrit le centre de masse, et étudier les figures produites par le programme informatique. Les Figures C.20 et C.21 montrent des exemples de figures produites par le programme.

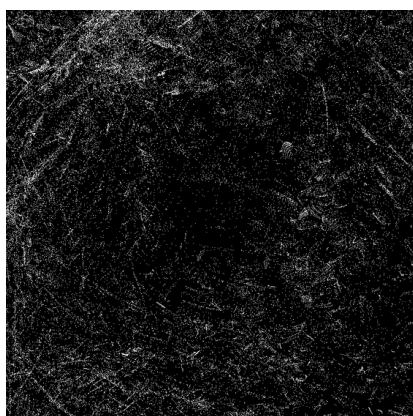


FIGURE C.20 – Moving cells, une configuration

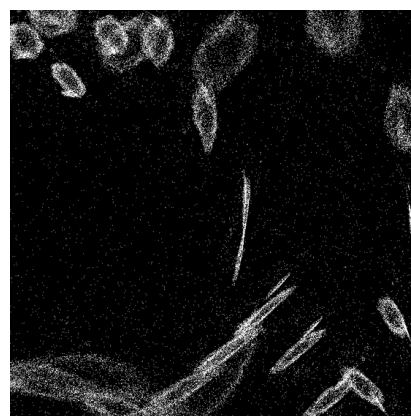


FIGURE C.21 – Moving cells, une configuration





## Annexe

# D

### Étude d'un modèle de fission d'une colonie de fourmis

Cette annexe vise à décrire la mise en œuvre d'une calibration avec l'algorithme ABC, présenté à la Section 2.2.1 du Chapitre 2, et l'obtention d'une formule analytique pour exprimer les sorties du modèle (moyennes) en fonction de ses paramètres d'entrée. Ce travail est en cours de réalisation en collaboration avec Thibaud Monnin et Guillaume Chérel et concerne la fission de colonies de fourmis à partir d'un modèle individu-centré développé par Thibaud Monnin. Nous étudions la manière dont les comportements individuels déterminent le comportement de la colonie. Le comportement de la colonie considéré est la fondation de nouvelles colonies chez une espèce qui se reproduit par fission de la colonie, c'est-à-dire lorsqu'une colonie mature se divise en plusieurs nouvelles colonies indépendantes. Plus précisément, nous examinons la capacité d'un modèle individu-centré d'allocation des ressources (les fourmis) à reproduire les motifs d'allocation observés dans la nature. Nous nous intéressons en particulier au nombre de nouvelles colonies et à leur taille (nombre d'individus) après fission. L'allocation des ressources est effectuée par les fourrageuses, qui transportent les autres fourmis de l'ancienne colonie vers les nouvelles. Le modèle s'inspire de la biologie de la fourmi *Cataglyphis piliscapa*, cependant nous ne présentons que brièvement les mécanismes considérés, l'article faisant le lien avec les interprétations biologiques est en cours de rédaction (Monnin et al., en cours d'écriture). Il est très simple et n'inclut aucune communication entre les fourmis. Chaque fourrageuse décide seule, sans échanger d'information avec d'autres fourrageuses, dans quel nid allouer les ressources de la colonie. Notre objectif est d'étudier si un tel modèle est capable de reproduire les patrons d'allocation de ressources après fission observés dans la nature. Il se place dans le cadre de la calibration automatique de modèles. Nous avons choisi pour cela une méthode bayésienne approchée (voir Section 2.2.1). Si effectivement le modèle reproduit les données, cela montrerait que l'allocation des ressources peut découler de nombreux choix faits indépendamment au niveau individuel. Cependant, cela ne signifierait pas nécessairement que les fourrageuses ne communiquent pas, mais plutôt que le processus d'allocation des ressources peut être résilient aux perturbations.

De plus, afin de mieux comprendre la dynamique du modèle, nous obtenons une formule analytique pour les sorties du modèle (moyennées) en fonction des paramètres d'entrée. En ce sens, cet objectif est commun avec les méthodes de méta-modèle, introduites dans la Section 2.3, qui cherchent à trouver les valeurs des coefficients du méta-modèle (défini par une fonction analytique qui dépend de paramètres d'entrée du modèle initial) qui font le mieux correspondre les sorties du modèle initial avec celles du méta-modèle. Cette étape vise ensuite à remplacer l'étude du modèle par celle du méta-modèle sélectionné. Ici, nous présentons l'obtention d'une formule analytique à partir de calculs sur la dynamique du modèle reposant sur des hypothèses simplificatrices. Cette approche est donc différente de celle menée à la Section 4.2.3, où nous proposons un méta-modèle (formule analytique) basé sur une étude préliminaire du modèle utilisant des OAT (voir Section 2.3). Cette approche est aussi différente de celle de Happe et al. (2006) (voir Section 2.3) où les auteurs choisissent une forme polynomiale du second ordre pour leur méta-modèle, dont les coefficients sont à ajuster. Nous n'avons donc pas de coefficients à ajuster, mais nous pouvons comparer avec un plan d'expériences les sorties du modèle individu-centré avec celles du méta-modèle.

Cette annexe est organisée comme suit. La Section D.1 est consacrée à la description du modèle de fission ainsi qu'aux méthodes utilisées pour son étude. Les résultats sont présentés dans la Section D.2. En particulier, nous avons montré que le modèle reproduit bien des données de terrain et nous avons obtenu une formule analytique qui approche le modèle.

## D.1 Matériel et Méthodes

### D.1.1 Description des mécanismes

Nous décrivons les mécanismes pris en compte dans le modèle d'allocation de ressources après fission d'une colonie de fourmis sans communication. Le modèle est individu-centré, les individus sont les fourmis. Le modèle considère :

- une colonie initiale : on considère une colonie composée de *initialSize* fourmis. Elle est composée de  $K$  ouvrières fourrageuses (on utilise aussi le pourcentage de fourrageuses dans la colonie noté *percentageForagers*). On suppose que toutes les fourrageuses ont les mêmes caractéristiques. On suppose qu'elles ne communiquent pas entre elles, et en particulier qu'elles ne laissent pas de trace de phéromone.
- un environnement : on suppose que les fourrageuses évoluent dans un terrain carré d'aire  $A_{tot}$ . Le nid de la colonie initiale est placé en son centre. Des nids potentiels, au nombre de  $N$ , sont placés de manière aléatoire et uniforme dans l'environnement (à l'exception d'un disque centré sur la colonie initiale et de rayon *exclusionRadius* fixé). Les nids sont caractérisés par une qualité  $Q_k \in \mathbb{R}$ .

Le modèle décrit le déplacement de fourmis, d'abord dans une phase d'exploration qui imite la connaissance de l'environnement acquise par les fourrageuses lorsqu'elles l'explorent pour collecter de la nourriture, au cours des semaines précédant le début de la fission de la colonie, puis dans une phase de fission pour allouer les fourmis à un nouveau nid. Le modèle est à temps discret.

**Phase d'exploration** : Les fourrageuses explorent l'environnement pendant une durée *exploringPhase* (nombre de pas de temps), les autres fourmis restent dans le nid. Afin de décrire leur connaissance de l'environnement au moment de la fission, c'est à dire la position et la qualité de nids potentiels, on suppose un mouvement très simplifié des fourmis. En un pas de temps, une fourmi parcourt une distance  $d$  et se déplace dans une direction formant un angle aléatoire (tiré de manière uniforme dans  $[-\pi/4, +\pi/4]$ ) avec sa direction actuelle. Le mouvement des fourmis est donc aléatoire. Il très simple et est utilisé pour mimer leur connaissance de l'environnement. Durant la phase d'exploration, une fourrageuse voit un nid et le visite s'il est à une distance inférieure à  $r$ . Chaque fourmi estime la qualité du nid visité durant cette phase, et conserve la connaissance de sa position et de sa qualité estimée. On suppose que l'estimation de la qualité du nid par une fourmi se fait avec une erreur. L'erreur dans cette estimation dépend du paramètre *NestQualityAssessmentError*  $\in [0, 100]$  (identique pour chaque fourmi). On suppose qu'un nid de qualité  $Q_k$  est estimé par  $Q_k + \mathcal{N}(0, Q_k \times \text{NestQualityAssessmentError}/100)$ . Ainsi, l'erreur d'estimation dépend également de la qualité du nid et la qualité peut être surestimée, sous-estimée, ou estimée sans erreur (si *NestQualityAssessmentError* = 0, dans ce cas la qualité estimée du nid égale la qualité réelle du nid). A l'issue de la phase d'exploration, chaque fourrageuse sélectionne le meilleur nid parmi ceux qu'elle a visités.

**Phase de fission** : A l'issue de la phase d'exploration, les fourrageuses retournent au nid de la colonie initiale. Dès lors, chaque fourrageuse fait des allers-retours entre le nid principal et le nid sélectionné durant la phase d'exploration. Lors de chaque déplacement, elle emporte une fourmi non fourrageuse du nid principal jusqu'au nid sélectionné. Ce processus se déroule jusqu'à ce que toutes les fourmis du nid initial soient transportées. Enfin, chaque fourrageuse rejoint le nid qu'elle a choisi.

A la fin de la fission, on s'intéresse au nombre de colonies formées (aussi appelé le nombre de nids après fission), ainsi qu'au nombre de fourmis composant chacune de ses colonies (ou

de manière équivalente à sa proportion par rapport à la taille de la colonie initiale). Nous avons aussi accès au nombre de nids visités par chaque fourrageuse, mais nous utilisons cette information uniquement à la Section D.2.2.

### D.1.2 Synthèse des paramètres

Certains paramètres de l'environnement ou liés au déplacement des fourmis sont fixés de manière arbitraire. Ainsi, le niveau de connaissance de l'environnement est caractérisé par un unique paramètre qui fixe la durée du déplacement des fourmis (*exploringPhase*). Le Tableau D.1 synthétise les paramètres du modèle, nous précisons les paramètres dont la valeur est fixée dans toute cette étude.

| Paramètres                                  | Description  | Valeur*       |
|---|--|---------------|
| Colonie initiale                            |  |               |
| <i>initialSize</i>                          | nombre de fourmis dans la colonie initiale                                   |               |
| <i>percentageForagers</i> (resp. <i>K</i> ) | pourcentage (resp. nombre) d'ouvrières fourrageuses dans la colonie initiale |               |
| Environnement                               |  |               |
| <i>A<sub>tot</sub></i>                      | aire de la zone (carrée)   | $61^2 = 3721$ |
| <i>N</i>                                    | nombre de nids potentiels  |               |
| <i>Q<sub>k</sub></i>                        | qualité du nid <i>k</i>  |               |
| <i>exclusionRadius</i>                      | rayon du disque d'exclusion des nids potentiels                              | 10            |
| Fourmis                                     |  |               |
| <i>d</i>                                    | distance parcourue en un pas de temps  | 3/4           |
| <i>r</i>                                    | distance à laquelle une fourmi voit un nid                                   | 1             |
| <i>NestQualityAssessmentError</i>           | intervient dans l'erreur d'estimation de la qualité d'un nid                 |               |
| Phase d'exploration                         |  |               |
| <i>exploringPhase</i>                       | durée de la phase d'exploration (en pas de temps)                            |               |

TABLE D.1 – Synthèse des paramètres pour le modèle de fission d'une colonie. \* pour les paramètres dont la valeur est fixée dans toute cette étude.

### D.1.3 Simulations

Pour simuler la dynamique du modèle, nous utilisons l'implémentation du modèle en Netlogo (version 6.1.1, (Wilensky et al., 1999)) faite par Thibaud Monnin. Nous utilisons le logiciel OpenMOLE (version 10.0, (Reuillon et al., 2013)) pour l'exploration du modèle. Enfin, nous utilisons le logiciel R (version 3.4.4) pour l'analyse des résultats des simulations. Les simulations sont réalisées sur la grille de calcul européenne (EGI, <http://www.egi.eu/>).

Le code Netlogo du modèle ainsi que les codes OpenMOLE et R utilisés pour l'étude du modèle sont disponibles sur GitLab (<https://gitlab.iscpif.fr/gcherel/antsdispersalstrategy>).

Pour réaliser une simulation du modèle, nous devons préciser la valeur de qualité des nids potentiels. Cela est fait de manière aléatoire selon une loi normale, ce qui introduit deux paramètres supplémentaires (moyenne et variance), fixés pour toutes les simulations dans ce document (respectivement à 70.0 et 50.0). Ainsi, étant donnée la manière dont est initialisé l'environnement pour une simulation (position et qualité des nids aléatoires), deux simulations effectuées avec le même jeu de paramètres n'auront pas le même environnement initial. De ce fait, lors de répliques de simulations effectuées avec le même jeu de paramètres, nous observons de la stochasticité issue de la manière dont est initialisé l'environnement et celle issue du comportement des fourmis. De plus, concernant les sorties du modèle, nous ne conservons que les valeurs des 10 nids après fission ayant le plus de fourmis (un vecteur de tailles inférieur à 10 est complété de 0). Dans un souci de répétabilité de nos travaux, nous conservons la valeur de la graine aléatoire généré pour les simulations réalisées dans la Section D.2.2.

### D.1.4 Données pour la calibration

Les données de terrain sont celles présentées dans Chéron et al. (2011). Des mesures ont été effectuées sur 19 colonies de fourmis. La taille des colonies varie entre 252 et 1808, le nombre de nids après fission entre 2 et 7 (la moyenne est de 4 nids). Pour chaque colonie, nous connaissons la taille de la colonie initiale, ainsi que le nombre de nids après fission et le nombre de fourmis dans chaque nid.

### D.1.5 Méthode pour la calibration

Notre objectif est de vérifier que le modèle est capable de reproduire les données observées sur le terrain, pour chaque colonie. Nous procédons donc à une calibration. Nous procédons différemment que pour la calibration effectuée à la Section 4.1.6 dans le cadre de la dynamique de croissance locale de tache renouée. Dans la Section 4.1.6, notre objectif était de trouver un jeu de paramètres de la plante commun à toutes les taches, qui fait le mieux correspondre les sorties du modèle aux données de terrain. Ici, nous cherchons un jeu de paramètres du modèle pour chaque colonie, qui fait le mieux correspondre les sorties du modèle aux données pour la colonie concernée. Nous effectuons ainsi autant de calibrations qu'il y a de colonies. La réalisation d'une inférence par colonie permet d'interpréter les différentes marginales comme différents mécanismes biologiques sous-jacents et ainsi d'étudier si la taille de la colonie et l'environnement (tel qu'il est pris en compte dans le modèle) peuvent expliquer à eux seuls la répartition des nids après fission, ou si d'autres mécanismes interviennent. Cela signifie par exemple que nous ne supposons pas que les colonies ont le même pourcentage de fourrageuses, ni la même durée pour la phase d'exploration. La colonie initiale est caractérisée par sa taille *initialSize*. Nous procédons à la calibration pour les paramètres suivants : durée de la phase d'exploration *exploringPhase*, erreur de la qualité d'estimation du nid *NestQualityAssessmentError*, pourcentage de fourrageuses *percentageForagers* et nombre de nids potentiels *N*. Les calibrations sont indépendantes entre les colonies, et nous simulons la fission de chaque colonie avec le jeu de paramètres correspondant à tester : nous obtenons le nombre de colonies formées après fission, ainsi que le nombre de fourmis composant chacune d'elles. Nous devons donc comparer ces valeurs simulées à celles mesurées sur le terrain, cela nécessite de choisir une distance. Pour chaque colonie, nous optons pour la distance euclidienne afin de comparer le vecteur simulé et observé donnant le nombre de fourmis dans chaque nid après fission (les colonies observées ont produit moins de 10 nids, les vecteurs de taille inférieure à 10 sont complétés de 0, voir Section D.1.3).

Pour minimiser la distance dans chaque colonie, nous utilisons une méthode de calcul bayésien approché (ABC, voir Section 2.2.1) implémenté dans le logiciel OpenMOLE. Une méthode bayésienne nécessite de choisir une loi *a priori* sur l'espace des paramètres. Nous choisissons une loi uniforme pour chaque paramètre, sur  $[0, 50]$  pour *NestQualityAssessmentError*, sur  $[1000, 10000]$  pour *exploringPhase*, sur  $[3, 100]$  pour *percentageForagers* et sur  $[15, 80]$  pour *N*. La méthode calcule une approximation de la loi *a posteriori*.

Le résultat est un échantillon, de taille 500 dans notre cas, de la loi *a posteriori* pour chaque colonie. Or cette distribution est à support dans un espace de dimension 4, nous choisissons donc d'illustrer le résultat obtenu en traçant les lois marginales (une estimation de leur densité) pour chaque paramètre, et trois colonies parmi les 19 (une de grande taille (1808), une de taille intermédiaire (720) et une de petite taille (278)). Le graphique correspondant est celui de la Figure D.1, dans la Section D.2.1.

### D.1.6 Analyse numérique

Notre objectif est d'évaluer la capacité du modèle à reproduire les données. Pour mesurer la qualité de la solution donnée par l'algorithme ABC pour chaque colonie, nous disposons de la distance calculée pour chaque jeu de paramètres composant l'échantillon de la loi *a posteriori*. Nous pouvons par exemple considérer la valeur maximale par colonie. Nous comparons aussi les sorties du modèle obtenues avec les jeux de paramètres issus de la loi *a posteriori* pour chaque colonie, avec les données de terrain correspondantes. Pour cela, pour chaque colonie nous ré-échantillonnons 2000 jeux de paramètres selon la loi *a posteriori* (à partir de l'échantillon) et nous traçons le graphe qui donne pour chaque nid après fission le nombre de fourmis qu'il contient. Le résultat est présenté sur la Figure D.1, dans la Section D.2.1.

### D.1.7 Méthode pour l'approximation du modèle

Nous présentons la méthode suivie pour l'obtention d'une approximation du modèle dans le cas particulier où les fourrageuses ne font pas d'erreur dans l'estimation de la qualité d'un nid (*NestQualityAssessmentError* = 0), réalisée à la Section D.2.2. Cette approximation est une formule analytique pour le nombre moyen de nids après fission et le nombre moyen de fourmis dans chaque nid après fission en fonction des paramètres du modèle. L'adjectif moyen réfère à la fois à l'environnement (position des nids) et au comportement des fourrageuses pendant la phase d'exploration. Cette approximation repose sur le fait que toutes les fourmis ont le même comportement, et qu'il n'y a pas d'interaction entre elles (absence de communication). Nous nous intéressons donc au comportement d'une fourmi. Nous estimons d'abord sa connaissance de l'environnement durant la phase d'exploration par le nombre moyen de nids visités. Pour cela, nous trouvons une formule qui décrit l'aire moyenne visitée par une fourmi en fonction de la durée d'exploration (et de l'aire de la zone et des paramètres de son déplacement). A partir de cette information, nous déduisons le nombre moyen de nids visités par une fourrageuse durant la phase d'exploration (*exploringPhase*). Nous notons  $n$  sa valeur (la dépendance en les paramètres cités n'est pas spécifiée pour alléger la notation). Nous exprimons ensuite le nombre moyen de nids après fission et le nombre moyen de fourmis dans chaque nid après fission en fonction de cette quantité. Pour cela, nous numérotions les  $N$  nids potentiels de qualité  $Q_1 < \dots < Q_N$  par ordre croissant de qualité. On suppose que ces nids sont visités par une fourrageuse de manière équiprobable. Sachant qu'une fourrageuse a visité  $n$  nids, on note  $X$  la variable aléatoire qui donne l'indice du nid visité qui a la plus grande qualité (cet indice est le plus grand des indices des nids visités). Nous notons  $N_k$  le nombre moyen de fourmis dans le nid  $k$  après fission. Nous rappelons que les nids sont répartis spatialement uniformément dans la zone, et que la distance d'un nid au nid initial est indépendante de la qualité du nid. Les fourrageuses transportent les *initialSize* –  $K$  fourmis dans le nid qu'elles ont choisi (de meilleure qualité). En moyenne, les fourrageuses parcourent la même distance. Le nombre moyen de fourmis dans chaque nid après fission est donc proportionnel au nombre de fourrageuses qui ont choisi ce nid, connu par la loi de  $X$ . Dans le cas où les fourrageuses ne font pas d'erreur d'estimation de la qualité du nid lors de la phase d'exploration, c'est à dire *NestQualityAssessmentError* = 0, nous calculons la loi de  $X$ . Elle ne dépend que de  $n$  (le nombre de nids visités par une fourmi pendant la phase d'exploration) et  $N$  le nombre de nids potentiels, mais pas des qualités des nids. La formule dans le cas général dépend de la valeur des qualités  $Q_k$ .

Enfin, nous comparons la sortie du modèle à celle de la formule analytique trouvée, pour différents jeux de valeurs pour les paramètres d'entrée. Le plan d'expériences utilisé à cette fin est détaillé dans la Section D.2.2. Toutes les simulations présentées dans la Section D.2.2

sont réalisées avec la taille de colonie initiale fixée à  $initialSize = 1000$  fourmis (nous faisons éventuellement varier la valeur de  $percentage\_foragers$ , donc de  $K$ ).

## D.2 Résultats

Dans cette section nous présentons les résultats, d'abord ceux qui concernent la calibration, puis ceux relatifs à l'approximation du modèle.

### D.2.1 Calibration

La Figure D.1 montre les lois marginales issues de la solution de l'algorithme ABC pour chaque paramètre et pour trois colonies de tailles différentes (une colonie par ligne). Afin de pouvoir comparer les marginales de différentes colonies, les échelles sont les mêmes pour les graphes d'une même colonne. Nous ne présentons pas ici les interprétations biologiques des valeurs obtenues.

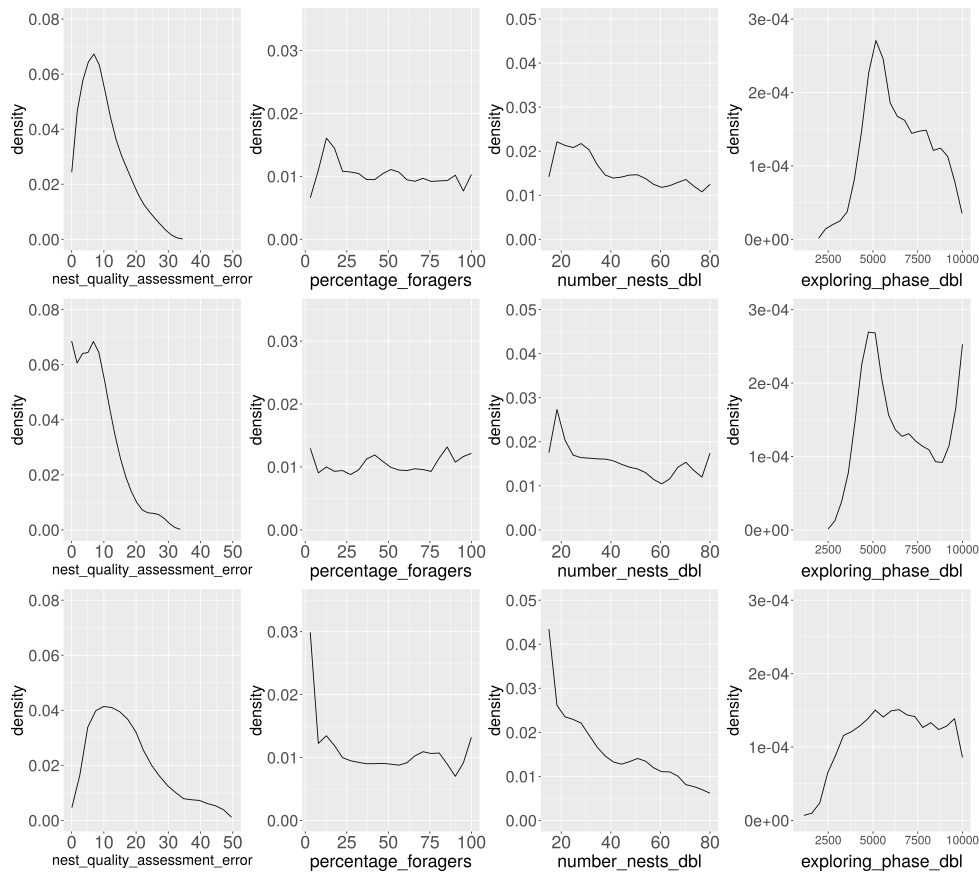


FIGURE D.1 – Lois marginales de la loi *a posteriori* pour les paramètres soumis à la calibration, pour 3 colonies (lignes).

La Figure D.2 montre pour chaque colonie la comparaison du nombre de fourmis dans chaque nid après fission entre les données (points rouges) et les sorties du modèle simulé avec les paramètres ré-échantillonnés dans la loi *a posteriori* (synthétisés dans une boîte à moustaches dont les bords du rectangle correspondent au premier et au troisième quartiles, coupée par la médiane et prolongée par des segments qui incluent 95 % des valeurs; les points noirs sont des valeurs extrêmes). Nous observons une bonne adéquation entre les sorties du modèle et les



données (les données sont souvent dans la boîte à moustaches), pour chaque colonie. Ainsi, le modèle est capable de reproduire des données.

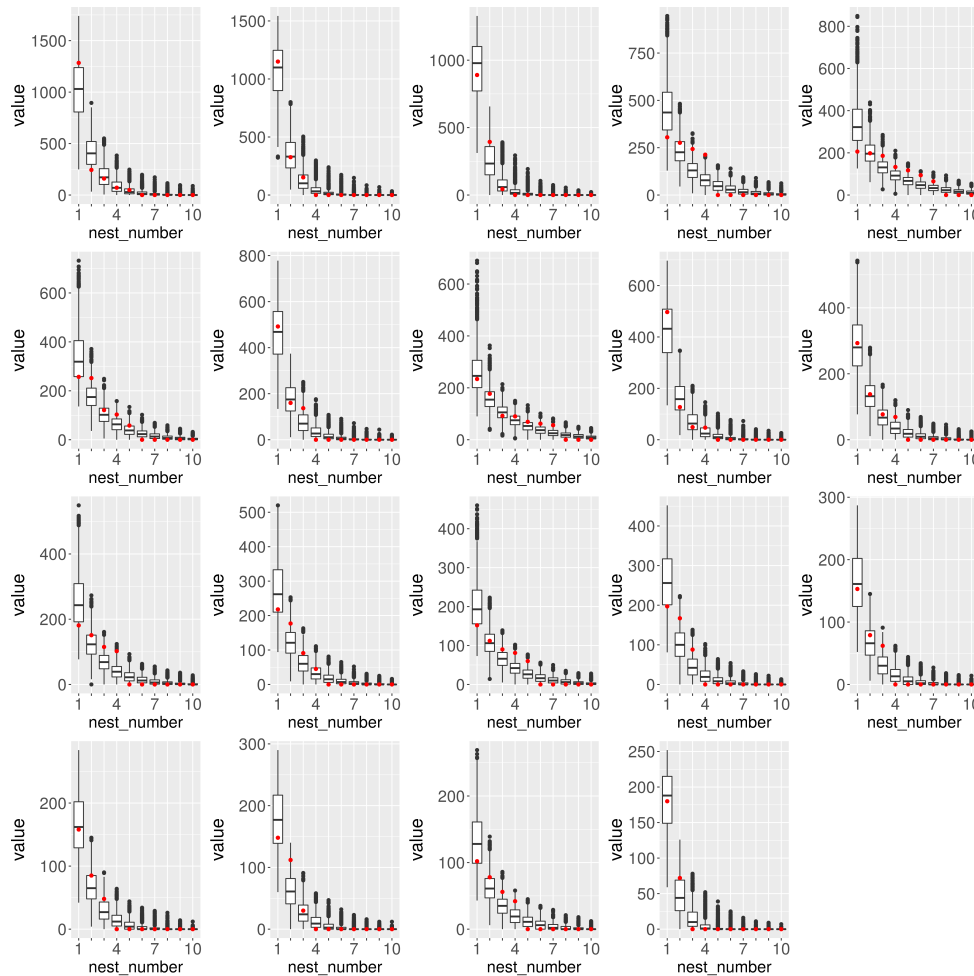


FIGURE D.2 – Comparaison des réplifications de sorties du modèle simulé avec les jeux de paramètres issus de la loi trouvée par la calibration (boîtes à moustaches et points noirs), et des données de terrain (points rouges). Les colonies sont rangées par taille décroissante (l'échelle n'est pas la même sur tous les graphiques).

### D.2.2 Approximation du modèle

Cette section présente les calculs pour l'obtention d'une approximation du modèle, suivant la méthode décrite à la Section D.1.7. Cette approximation est une formule analytique pour le nombre moyen de nids après fission et le nombre moyen de fourmis dans chaque nid après fission en fonction des paramètres du modèle. Nous exprimons d'abord le nombre moyen de nids visités par une fourrageuse durant la phase d'exploration, noté  $n$ . Puis, nous exprimons le nombre moyen de nids après fission et le nombre moyen de fourmis dans chaque nid après fission en fonction des paramètres du modèle dans le cas particulier où les fourrageuses ne font pas d'erreur dans l'estimation de la qualité d'un nid ( $NestQualityAssessmentError = 0$ ), avant de traiter du cas général.

### Nombre moyen de nids visités par une fourmi

Nous commençons par approcher l'aire de la zone visitée par une fourmi lors de la phase d'exploration. Nous rappelons que les fourrageuses évoluent dans un terrain carré d'aire  $A_{tot}$ . Nous rappelons aussi qu'à chaque pas de temps, une fourmi a une direction, et elle avance d'une distance  $d$  dans cette direction. Le choix de la nouvelle direction est fait par un angle tiré aléatoirement uniformément dans  $[-\pi/4, \pi/4]$ . Une fourmi voit les nids situés à une distance  $r$  (disque). On note  $a_i$  l'aire moyenne visitée par une fourmi après  $i$  pas de temps, c'est à dire l'aire moyenne de la zone qu'une fourmi a vue pendant sa trajectoire. Dans ce qui suit, on ne tient pas explicitement compte de l'intervalle  $[-\pi/4, \pi/4]$ . Au premier pas de temps, la fourmi voit un disque d'aire  $\pi r^2$ . Au pas de temps suivant, elle avance dans une direction, mais reste à l'intérieur de ce disque ( $d < r$ ), il y a donc une part du nouveau disque qu'elle voit actuellement qu'elle a déjà vue au pas de temps précédent. L'aire du domaine constitué du disque auquel on a retiré cette part est notée  $A$ , et vaut :  $A = \pi r^2 - 2r^2 \arccos\left(\frac{d}{2r}\right) + d\sqrt{r^2 - \frac{d}{2}}$ .

Or, lors de son déplacement, une fourmi peut revisiter des zones de l'espace qu'elle a déjà visitées, mais cela ne doit pas incrémenter l'aire visitée. Ainsi, à chaque pas de temps, nous considérons que l'aire moyenne visitée ne s'incrémente pas de  $A$ , mais d'une proportion de  $A$  égale à la proportion de l'aire moyenne pas encore visitée dans le domaine. Dans un premier temps, nous considérons donc que pour tout  $i \in \mathbb{N} \setminus \{0; 1\}$ ,  $a_{i+1} = a_i + A * \left(1 - \frac{a_i}{A_{tot}}\right)$ .

On a  $a_0 = 0$ ,  $a_1 = \pi r^2$ ,  $a_2 = a_1 + A$ . Dans un premier temps, on en déduit donc que pour tout  $i \in \mathbb{N} \setminus \{0; 1\}$ ,  $a_i = \left(1 - \frac{A}{A_{tot}}\right)^{i-2} (a_2 - A_{tot}) + A_{tot}$ .

Par simulation, on constate cependant que l'aire visitée par les fourmis croît moins vite que ce qui est prédit par l'équation précédente, mais que l'allure de la trajectoire est correcte. On introduit la possibilité de faire un décalage temporel avec le paramètre  $\theta \in \mathbb{R}$  (en pratique  $\theta \in [0, 1]$ ). Pour tenir compte de ce décalage, on a finalement pour  $i \geq 2$ ,  $a_i = \left(1 - \frac{A}{A_{tot}}\right)^{\theta(i-2)} (a_1 + A - A_{tot}) + A_{tot}$ .

Or, les nids étant répartis uniformément dans l'espace, le nombre moyen de nids visités pendant la phase d'exploration (de durée *exploringPhase*), est donc proportionnel à l'aire moyenne visitée  $a_{exploringPhase}$ , et vaut :

$$n = \frac{a_{exploringPhase}}{A_{tot}} * N \quad (D.1)$$

On calcule la moyenne du nombre de nids visités par chaque fourmi au sein d'une simulation (100 forageuses) et sur 50 réplifications (donc associées à des dispositions de nids différentes), pour différentes valeurs de nids potentiels  $N$  et de durée d'exploration *exploringPhase*. Ces valeurs ont été prises sur une grille régulière avec *exploringPhase* variant de 200.0 à 6200.0 par pas de 200.0 et  $N$  variant de 10.0 à 50.0 par pas de 20.0. La Figure D.3 montre le nombre moyen de nids visités par une fourmi en fonction de la durée de la phase d'exploration *exploringPhase* (en noir le résultat simulé par le modèle, en vert celui issu de la formule donnée par l'Équation D.1) pour ces différentes valeurs de nids potentiels  $N$ . Pour chaque graphique, la droite horizontale rouge correspond à la valeur maximale du nombre de nids visités (c'est à dire la valeur de  $N$ ). La valeur de  $\theta$  a été fixée à 0.576 afin de minimiser la distance entre les données simulées et la valeur prédite (pour le nombre moyen de nids, voir sous-section suivante).

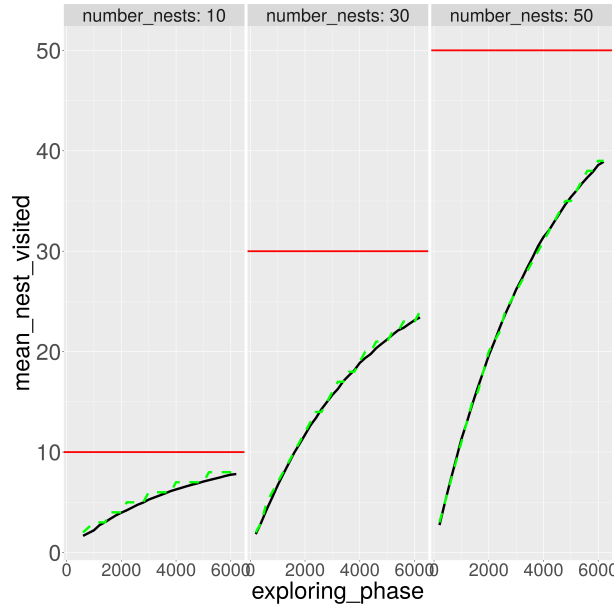


FIGURE D.3 – Nombre moyen de nids visités par une fourmi en fonction de la durée de la phase d'exploration, *exploringPhase*, pour différentes valeurs de nombre de nids  $N$  potentiels. En noir, les résultats issus de la simulation (réplications), en vert la courbe théorique donnée par l'Équation D.1. Ces deux courbes sont presque identiques (donc superposées) sur chaque graphique. Les droites horizontales rouges correspondent à la valeur maximale du nombre de nid visités (c'est à dire la valeur de  $N$ ).

### Cas particulier : les fourrageuses ne font pas d'erreur dans l'estimation de la qualité d'un nid

Nous commençons par examiner le cas particulier où les fourrageuses ne font pas d'erreur d'estimation de la qualité du nid, c'est-à-dire lorsque  $NestQualityAssessmentError = 0$ .

**Calcul du nombre moyen de nids après fission** On s'intéresse encore au comportement d'une seule fourmi. Sachant qu'une fourrageuse a visité  $n$  nids, nous rappelons que nous notons  $X$  la variable aléatoire qui donne l'indice du nid visité qui a la plus grande qualité. La loi de  $X$  est donnée par : soit  $k \in [1, N]$ , si  $k < n$ ,  $\mathbb{P}(X = k) = 0$ , si  $k = n$ ,  $\mathbb{P}(X = k) = \frac{1}{\binom{N}{n}}$ , si

$$k > n, \mathbb{P}(X = k) = \frac{\binom{k-1}{n-1}}{\binom{N}{n}}.$$

Or, il y a  $K$  fourrageuses, on considère donc  $K$  variables i.i.d de loi  $X : (X_1, \dots, X_K)$ .

Pour  $i \in [1, N]$ , on note  $Y_i$  la v.a qui vaut 1 si au moins une des v.a  $(X_k)_{k \in [1, K]}$  vaut  $i$  et 0 sinon. On a donc  $Y_i \sim \mathcal{B}(1 - (1 - \mathbb{P}(X = i))^K)$  pour  $i \geq n$  ( $Y_i$  nulle sinon).

Le nombre moyen de nids après fission correspond donc à  $\mathbb{E} \left[ \sum_{i=1}^N Y_i \right]$  et vaut

$$\mathbb{E} \left[ \sum_{i=1}^N Y_i \right] = (N - n + 1) - \sum_{i=n}^N \left( 1 - \frac{\binom{i-1}{n-1}}{\binom{N}{n}} \right)^K. \quad (\text{D.2})$$

Dans l'Équation (D.2), on constate que le nombre moyen de nids après fission est croissant avec  $K$  le nombre de fourrageuses. Pour  $K = 1$ , (1 fourrageuse), la formule donne bien 1 nid après fission (cette fourrageuse transporte toutes les fourmis dans le meilleur nid qu'elle a visité), et

$\lim_{K \rightarrow +\infty} \mathbb{E} \left[ \sum_{i=1}^N Y_i \right] = N - n + 1$ . Donc, lorsque le nombre de fourrageuses devient grand (et nécessairement le nombre de fourmis également), le nombre de nid moyen après fission tend vers le nombre maximum de nids possibles  $N - n + 1$ . En effet, la loi de  $X$  est à support dans  $[n, N]$  : si une fourmi visite  $n$  nids, celui de qualité maximale sera nécessairement d'indice  $\geq n$  (mais ce ne sera plus le cas lorsque  $NestQualityAssessmentError \neq 0$ ). De plus, la Formule (D.2) ne dépend pas de la valeur des qualités des nids. Nous notons que la valeur de  $n$  n'est pas nécessairement entière, nous prenons donc sa partie entière dans la somme de l'Équation (D.2).

La Figure D.4 compare le nombre moyen de nids après fission en fonction de la durée de la phase d'exploration *exploringPhase*, pour différentes valeurs de nombre de nids  $N$  et de nombre de fourrageuses  $K$ , donné par la Formule (D.2) (en vert) et simulé (en rouge). Nous constatons donc que la Formule (D.2) approxime bien le nombre moyen de nids après fission donné par le modèle.

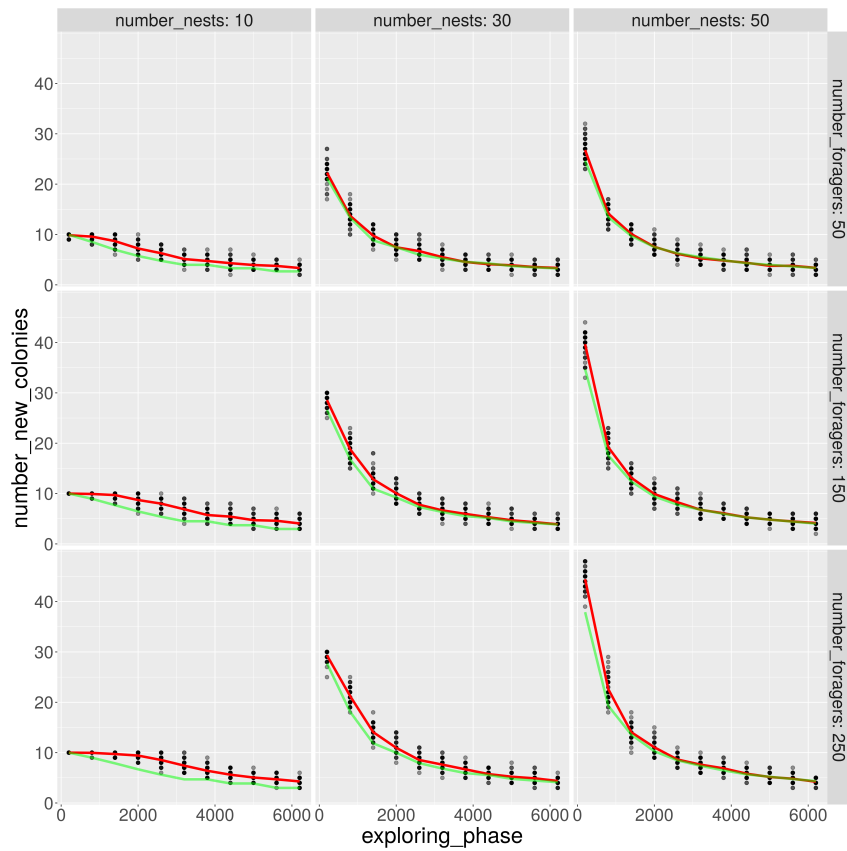


FIGURE D.4 – Nombre moyen de nids après fission en fonction de la durée de la phase d'exploration *exploringPhase*, pour différentes valeurs de nombre de nids  $N$  et de nombre de fourrageuses  $K$ . Les points noirs sont les résultats issus de la simulation (réplication), la courbe rouge en donne la moyenne. La courbe verte indique les valeurs données par la Formule (D.2).

**Calculs du nombre moyen de fourmis dans chaque nid après fission** Soit  $k \in [1, N]$ . On rappelle qu'on note  $N_k$  le nombre de fourmis dans le nid  $k$  après fission. On rappelle aussi que les nids sont répartis spatialement uniformément dans la zone, et que la distance d'un nid au nid initial est indépendante de la qualité du nid. Les fourrageuses transportent

les  $initialSize - K$  fourmis non fourrageuses dans le nid qu'elles ont choisi, puis chaque fourrageuse rejoint le nid qu'elle a choisi. En moyenne, elles parcourent la même distance. Le nombre moyen de fourmis dans chaque nid après fission est donc proportionnel au nombre de fourrageuses qui ont choisi ce nid. On a  $N_k = \frac{initialSize}{K} \sum_{i=1}^K 1_{X_i=k}$ , où  $X_i$  est défini à la sous-section précédente.

Le nombre moyen de fourmis dans le  $k$ -ième nid après fission vaut donc :

$$\mathbb{E}[N_k] = initialSize * \mathbb{P}(X = k). \quad (D.3)$$

La Figure D.5 compare le nombre moyen de fourmis dans chaque nid après fission (pour les 10 nids ayant le plus de ressources), en fonction de la durée de la phase d'exploration *exploringPhase*, pour différentes valeurs du nombre de nids  $N$  et du nombre de fourrageuses  $K$  (il s'agit des mêmes valeurs pour ces paramètres que celles utilisées pour la Figure D.4), donnés par la Formule (D.3) (prédits) et simulés à partir du modèle. Pour chaque durée d'exploration et pour chaque graphique, on représente l'histogramme des valeurs issues de la simulation du modèle. On indique les valeurs correspondantes issues de la Formule (D.3) par des points de différentes couleurs. On rappelle que la taille de la colonie initiale est fixée à 1000. Nous constatons là aussi que la Formule (D.3) approxime bien le nombre moyen de fourmis dans chaque nid après fission donné par le modèle (dès que la durée d'exploration est supérieure à 1000 pas de temps).

### Cas général, sans hypothèse sur l'erreur d'estimation de la qualité d'un nid

Nous examinons le cas général, pour lequel les fourrageuses font une erreur d'estimation de la qualité du nid.

**Calcul du nombre moyen de nids après fission** Comme précédemment, une fourmi choisit  $n$  nids de manière équiprobable. Elle évalue leur qualité avec une erreur. Nous rappelons que les  $N$  nids sont numérotés par ordre croissant de qualité :  $Q_1 < \dots < Q_N$ . Nous rappelons aussi que les erreurs d'estimation sont indépendantes et que leur loi dépend de la qualité du nid. Pour  $i \in [1, N]$ , on note  $A_i$  l'erreur commise en évaluant la qualité du nid  $i$  (qui est donc perçue comme étant  $Q_i + A_i$ ). On a donc, pour tout  $i$ ,  $A_i \sim \mathcal{N}(0, Q_i \times NestQualityAssessmentError/100)$ , et les variables aléatoires  $A_i$  sont indépendantes. Comme précédemment, on note  $X$  la variable aléatoire qui donne l'indice du nid visité de plus grande qualité (perçue ici).

Soit  $k \in [1, N]$ . On s'intéresse à l'événement  $X = k$ . On note, pour  $i, j \in [1, N]^2$ ,  $K_{i,j} = \{p \subset \{1, \dots, N\} \setminus \{i\}, |p| = j\}$ . Comme tous les tirages sont équiprobables, on a

$$\mathbb{P}(X = k) = \frac{1}{\binom{N}{n}} \sum_{p \in K_{k,n-1}} \mathbb{P}(\cap_{i \in p} Q_i + A_i < Q_k + A_k).$$

Pour  $(i_1, \dots, i_{n-1}) \in K_{k,n-1}$ ,  $(A_{i_1}, \dots, A_{i_{n-1}})$  est un vecteur gaussien (matrice de covariance diagonale), mais la matrice de covariance de  $(A_{i_1} - A_k, \dots, A_{i_{n-1}} - A_k)$  est  $d_k J + Diad(d_{i_1}, \dots, d_{i_{n-1}})$ , où  $J$  est la matrice carrée de taille  $n - 1$  remplie de 1 et où  $d_{i_j} = Var(A_{i_j})$ , en particulier les variables aléatoires qui composent ce vecteur gaussien ne sont pas indépendantes. Cela rend plus difficile le calcul de  $\mathbb{P}(X = k)$  par rapport au cas sans erreur d'estimation de la qualité d'un nid ( $NestQualityAssessmentError = 0$ ) pour lequel ces variables n'interviennent pas. Lorsqu'on intègre sur la loi de  $A_k$ , il est possible d'obtenir une relation de récurrence (sur  $n$ , considéré entier, à  $k$  fixé) pour les termes dans l'intégrale, à partir

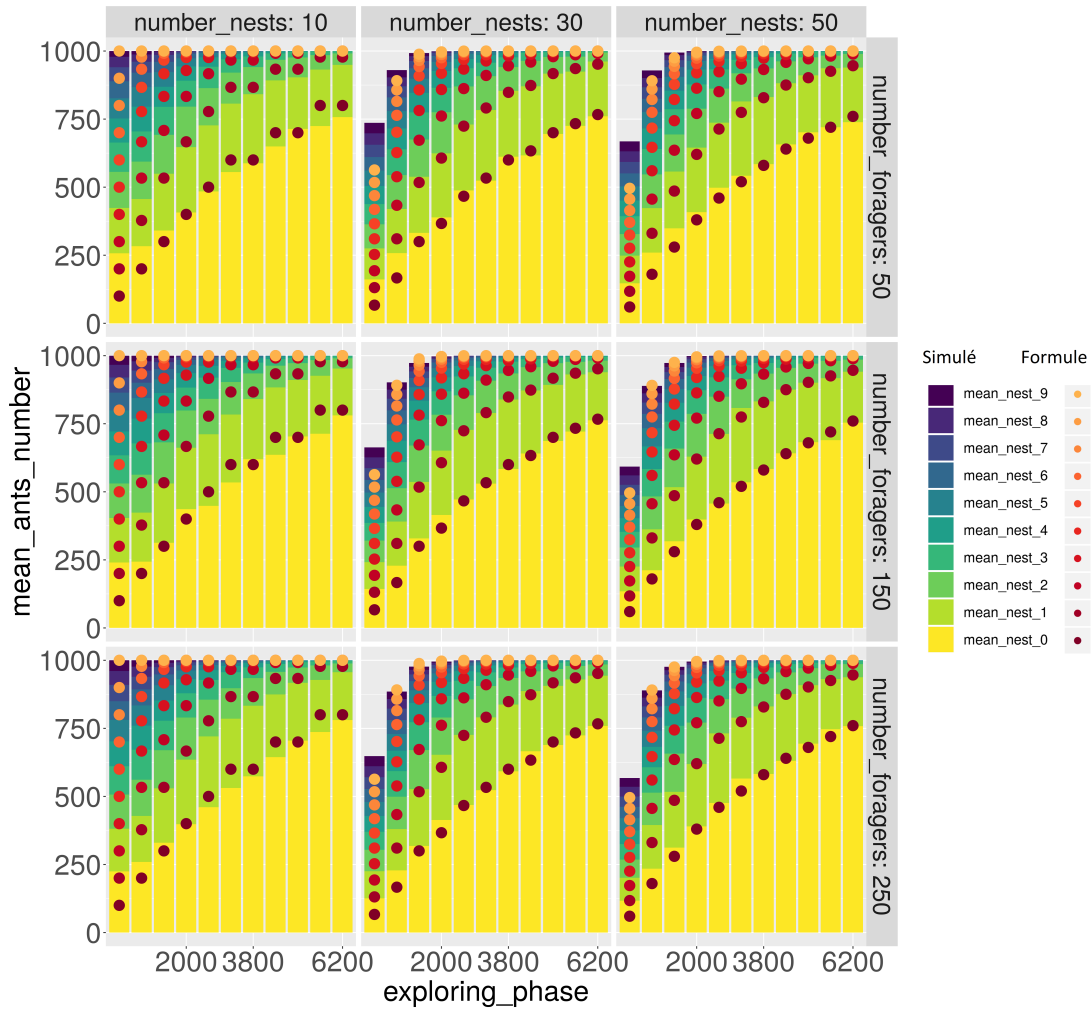


FIGURE D.5 – Nombre moyen de fourmis dans chaque nid après fission (pour les 10 nids ayant le plus de ressources), en fonction de la durée de la phase d'exploration, pour différentes valeurs du nombre de nids potentiels  $N$  et du nombre de fourrageuses  $K$ , prédits par la Formule (D.3) (points) et simulés (histogrammes).

de l'identité de Newton (qui utilise les polynômes symétriques élémentaires). Il n'est cependant pas aisé de déduire des propriétés du modèle à partir de cette formule. Toutefois, lorsque  $NestQualityAssessmentError = 0$  nous retrouvons la formule de la Section D.2.2 pour  $\mathbb{P}(X = k)$ . En effet, les variables  $A_i$  sont toutes nulles presque sûrement, et  $\mathbb{P}(\cap_{i \in p} Q_i < Q_k)$  vaut 0 ou 1 et est facile à calculer.

Enfin, à la Section D.2.2, nous avons exprimé le nombre moyen de fourmis dans chaque nid après fission en fonction de la loi de  $X$  (Équation (D.3)). Nous pouvons encore appliquer cette formule dans le cas où les fourrageuses font une erreur d'estimation sur la qualité du nid, mais l'écriture de la loi de  $X$  dans ce cas ne rend pas aisé l'obtention de propriétés du modèle à partir de cette formule.

### D.3 Conclusion

Nous avons d'abord vérifié que le modèle était en mesure de reproduire des données de terrain. Pour cela, nous avons opéré une calibration avec une méthode de calcul bayésien approché.

Nous avons ensuite ré-échantillonné des valeurs de jeux de paramètres selon la loi obtenue, pour simuler la dynamique du modèle, et comparer le nombre de fourmis dans chaque nid après fission avec les données de terrain. Cependant, nous n'avons pas utilisé l'information sur le nombre de nids après fission dans la distance pour la calibration. Or, nous trouvons en moyenne 7 nids après fission par simulation, contre 4 pour les données de terrain (et pour chaque colonie, le nombre de nids moyen simulé, calculé sur un échantillon de la loi *a posteriori* issue de la calibration, est supérieur à la valeur mesurée sur le terrain). Cette différence pourrait être expliquée par le phénomène de résorption de nids trop petits, observé sur le terrain mais dont le modèle ne tient pas compte. Il pourrait être intéressant d'introduire pour chaque colonie une contrainte sur le nombre de nids à obtenir, et d'observer comment cela impacte la capacité du modèle à reproduire les données.

Nous avons aussi obtenu une formule analytique pour le nombre de nids moyen après fission et le nombre de fourmis dans chaque nid après fission, dans le cas particulier où les fourmis ne font pas d'erreur d'estimation. D'autres approximations seraient nécessaires pour obtenir une formule analytique de laquelle on pourrait déduire des propriétés du modèle, dans le cas général où les fourrageuses font une erreur d'estimation de la qualité d'un nid. En particulier, un frein à l'obtention d'une telle formule vient de la non-indépendance de variables aléatoires. Une piste pourrait être de les considérer indépendantes et de quantifier l'erreur commise en remplaçant les sorties du modèle individu-centré par la formule obtenue, comme cela est fait pour les méta-modèles.

## Bibliographie

- Abgrall, C., Forey, E., Mignot, L., Chauvat, M., 2018. Invasion by *Fallopia japonica* alters soil food webs through secondary metabolites. *Soil Biology and Biochemistry* 127, 100–109.
- Adachi, I., Naoki, T., Terashima, M., 1996a. Central die-back of monoclonal stands of *Reynoutria japonica* in an early stage of primary succession on Mount Fuji. *Annals of Botany* 77, 477–486.
- Adachi, I., Naoki, T., Terashima, M., 1996b. Mechanisms of central die-back of *reynoutria japonica* in the volcanic desert on mt. fuji. a stochastic model analysis of rhizome growth. *Annals of botany* 78, 169–179.
- Alberternst, B., Böhmer, H., 2006. Invasive alien species fact sheet- *Fallopia japonica*-. online database of the North European and Baltic network on invasive alien species-nobanis. [www.nobanis.org](http://www.nobanis.org).
- Allen, G.H., Pavelsky, T.M., 2018. Global extent of rivers and streams. *Science* 361, 585–588.
- Alvarez, I., Reuillon, R., De Aldama, R., 2016. Viabilitree : A kd-tree Framework for Viability-based Decision. Technical Report. Irstea. URL : <https://hal.archives-ouvertes.fr/hal-01319738>.
- Aubin, J.P., 1991. Viability theory. Springer Science & Business Media.
- Aubin, J.P., 1997. Dynamic economic theory : a viability approach. volume 5. Springer Verlag.
- Aubin, J.P., Bonneuil, N., Maurin, F., Saint-Pierre, P., 2001. Viability of pay-as-you-go systems. *Journal of Evolutionary Economics* 11, 555–571.
- Aubin, J.P., Chen, L., Durand, M.H., 2012. Dynamical allocation method of emission rights of pollutants by viability constraints under tyochastic uncertainty. *Environmental Modeling & Assessment* 17, 7–18.
- Aubin, J.P., Chen, L., Durand, M.H., 2013. Dynamic decentralization of harvesting constraints in the management of tyochastic evolution of renewable resources. *Computational Management Science* 10, 281–298.



- Aubin, J.P., Saint-Pierre, P., 2007. An introduction to viability theory and management of renewable resources. *Advanced Methods for Decision Making and Risk Management in Sustainability Science*, 43–80.
- Bailey, J., Wisskirchen, R., 2006. The distribution and origins of *Fallopia* × *bohemica* (Polygonaceae) in Europe. *Nordic Journal of Botany* 24, 173–199.
- Bailey, J.P., Bímová, K., Mandák, B., 2009. Asexual spread versus sexual reproduction and evolution in Japanese knotweed *s.l.* sets the stage for the "battle of the clones". *Biological Invasions* 11, 1189–1203.
- Bailey, J.P., et al., 1994. Reproductive biology and fertility of *Fallopia Japonica* (Japanese knotweed) and its hybrids in the British Isles. *Ecology and management of invasive riverside plants*, 141–158.
- Baker, C.M., Bode, M., 2016. Placing invasive species management in a spatiotemporal context. *Ecological Applications* 26, 712–725.
- Barney, J.N., Tharayil, N., DiTommaso, A., Bhowmik, P.C., 2006. The biology of invasive alien plants in Canada. 5. *Polygonum cuspidatum* Sieb. & Zucc.[= *Fallopia Japonica* (Houtt.) Ronse Decr.]. *Canadian Journal of Plant Science* 86, 887–906.
- Barthod, L., Boyer, M., 2019. Un sac, des gants, un croc de jardin : le déterrage précoce, une technique douce contre l'envahissement des rivières par les renouées asiatiques. *Sciences Eaux & Territoires*.
- Bashtanova, U.B., Beckett, K.P., Flowers, T.J., 2009. Physiological approaches to the improvement of chemical control of Japanese knotweed (*Fallopia japonica*). *Weed science* 57, 584–592.
- Beaumont, M.A., 2010. Approximate bayesian computation in evolution and ecology. *Annual review of ecology, evolution, and systematics* 41, 379–406.
- Beerling, D., 1991. The effect of riparian land use on the occurrence and abundance of Japanese knotweed *Reynoutria Japonica* on selected rivers in south Wales. *Biological Conservation* 55, 329–337.
- Beerling, D.J., Bailey, J.P., Conolly, A.P., 1994. *Fallopia japonica* (Houtt.) Ronse Decraene. *Journal of Ecology* 82, 959–979.
- Béné, C., Doyen, L., Gabay, D., 2001. A viability analysis for a bio-economic model. *Ecological economics* 36, 385–396.
- Bernard, C., Martin, S., 2012. Building strategies to ensure language coexistence in presence of bilingualism. *Applied Mathematics and Computation* 218, 8825–8841.
- Bernard, C., Martin, S., 2013. Comparing the sustainability of different action policy possibilities : Application to the issue of both household survival and forest preservation in the corridor of Fianarantsoa. *Mathematical biosciences* 245, 322–330.
- von Bertalanffy, L., 1934. Untersuchungen über die Gesetzmäßigkeit des Wachstums. *Wilhelm Roux'Archiv für Entwicklungsmechanik der Organismen* 131, 613–652.

- Bímová, K., Mandák, B., Pyšek, P., 2003. Experimental study of vegetative regeneration in four invasive *Reynoutria* taxa (*Polygonaceae*). *Plant Ecology* 166, 1–11.
- Bolker, B., Pacala, S.W., 1997. Using moment equations to understand stochastically driven spatial pattern formation in ecological systems. *Theoretical Population Biology* 52, 179–197.
- Bram, M.R., McNair, J.N., 2004. Seed germinability and its seasonal onset of Japanese knotweed (*Polygonum Cuspidatum*). *Weed Science* 52, 759–767.
- Brias, A., Mathias, J.D., Deffuant, G., 2016. Accelerating viability kernel computation with cuda architecture : application to bycatch fishery management. *Computational Management Science* 13, 371–391.
- Brock, J., Wade, M., et al., 1992. Regeneration of Japanese knotweed (*Fallopia japonica*) from rhizomes and stems : observation from greenhouse trials., in : IXe Colloque international sur la biologie des mauvaises herbes, 16-18 September 1992, Dijon, France., ANPP. pp. 85–94.
- Buchadas, A., Vaz, A.S., Honrado, J.P., Alagador, D., Bastos, R., Cabral, J.A., Santos, M., Vicente, J.R., 2017. Dynamic models in research and management of biological invasions. *Journal of Environmental Management* 196, 594–606.
- Campolongo, F., Kleijnen, J., Andres, T., 2000. Screening methods, chapter 4 of sensitivity analysis, a. saltelli, k. chan, and em scott eds.
- Champagnat, N., 2006. A microscopic interpretation for adaptive dynamics trait substitution sequence models. *Stochastic Processes and their Applications* 116, 1127–1160.
- Champagnat, N., Ferrière, R., Méléard, S., 2006. Unifying evolutionary dynamics : from individual stochastic processes to macroscopic models. *Theoretical Population Biology* 69, 297–321.
- Chapel, L., Deffuant, G., 2011. Approximating Viability Kernels and Resilience Values : Algorithms and Practical Issues Illustrated with KAVIAR Software. volume 2011. pp. 161–192. doi :10.1007/978-3-642-20423-4\_7.
- Chapel, L., Deffuant, G., Martin, S., Mullan, C., 2008. Defining yield policies in a viability approach. *Ecological Modelling* 212, 10–15.
- Chéron, B., Cronin, A.L., Doums, C., Fédérici, P., Haussy, C., Tirard, C., Monnin, T., 2011. Unequal resource allocation among colonies produced by fission in the ant *cataglyphis cursor*. *Ecology* 92, 1448–1458.
- Chipperfield, J.D., Holland, E.P., Dytham, C., Thomas, C.D., Hovestadt, T., 2011. On the approximation of continuous dispersal kernels in discrete-space models. *Methods in Ecology and Evolution* 2, 668–681.
- Chérel, G., Cottineau, C., Reuillon, R., 2015. Beyond corroboration : Strengthening model validation by looking for unexpected patterns. *PLoS ONE* 10, 1–28. URL : <http://dx.doi.org/10.1371/journal.pone.0138212>, doi :10.1371/journal.pone.0138212.
- Çınlar, E., 2011. Probability and stochastics. volume 261. Springer Science & Business Media.

- Colleran, B.P., Goodall, K.E., 2014. In situ growth and rapid response management of flood-dispersed Japanese knotweed (*Fallopia japonica*). *Invasive Plant Science and Management* 7, 84–92.
- Coron, C., Costa, M., Leman, H., Smadi, C., 2018. A stochastic model for speciation by mating preferences. *Journal of Mathematical Biology* 76, 1421–1463.
- Costa, M., Hauzy, C., Loeuille, N., Méléard, S., 2016. Stochastic eco-evolutionary model of a prey-predator community. *Journal of Mathematical Biology* 72, 573–622.
- Cury, P.M., Mullon, C., Garcia, S.M., Shannon, L.J., 2005. Viability theory for an ecosystem approach to fisheries. *ICES journal of marine science* 62, 577–584.
- Dassonville, N., Guillaumaud, N., Piola, F., Meerts, P., Poly, F., 2011. Niche construction by the invasive Asian knotweeds (species complex *Fallopia*) : impact on activity, abundance and community structure of denitrifiers and nitrifiers. *Biological Invasions* 13, 1115–1133.
- Dauer, J.T., Jongejans, E., 2013. Elucidating the population dynamics of Japanese knotweed using integral projection models. *PloS one* 8, e75181.
- De Lara, M., Doyen, L., Guilbaud, T., Rochet, M.J., 2007. Is a management framework based on spawning-stock biomass indicators sustainable? a viability approach. *ICES Journal of Marine Science* 64, 761–767.
- De Waal, L., 2001. A viability study of *Fallopia japonica* stem tissue. *Weed Research* 41, 447–460.
- Deb, K., Pratap, A., Agarwal, S., Meyarivan, T., 2002. A fast and elitist multiobjective genetic algorithm : Nsga-ii. *IEEE transactions on evolutionary computation* 6, 182–197.
- Deffuant, G., Chapel, L., Martin, S., 2007. Approximating viability kernels with support vector machines. *IEEE transactions on automatic control* 52, 933–937.
- Descombes, P., Petitpierre, B., Morard, E., Berthoud, M., Guisan, A., Vittoz, P., 2016. Monitoring and distribution modelling of invasive species along riverine habitats at very high resolution. *Biological Invasions* 18, 3665–3679.
- Dieckmann, U., Law, R., Metz, J.A., 2000. The geometry of ecological interactions : simplifying spatial complexity. Cambridge University Press.
- Doble, R., Brunner, P., McCallum, J., Cook, P.G., 2012. An analysis of river bank slope and unsaturated flow effects on bank storage. *Groundwater* 50, 77–86.
- Domenech, P.A., Saint-Pierre, P., Zaccour, G., 2011. Forest conservation and co 2 emissions : a viable approach. *Environmental Modeling & Assessment* 16, 519.
- Dommanget, F., Breton, V., Forestier, O., Poupart, P., Daumergue, N., Evette, A., 2015. Contrôler des renouées invasives par les techniques de génie écologique : retours d'expérience sur la restauration de berges envahies. *Revue d'Ecologie (Terre et Vie)* 70, 215–228.
- Dommanget, F., Evette, A., Spiegelberger, T., Gallet, C., Pace, M., Imbert, M., Navas, M.L., 2014. Differential allelopathic effects of Japanese knotweed on willow and cottonwood cuttings used in riverbank restoration techniques. *Journal of Environmental Management* 132, 71–78.

- Dommanget, F., Spiegelberger, T., Cavaillé, P., Evette, A., 2013. Light availability prevails over soil fertility and structure in the performance of asian knotweeds on riverbanks : new management perspectives. *Environmental management* 52, 1453–1462.
- Doyen, L., De Lara, M., Ferraris, J., Pelletier, D., 2007. Sustainability of exploited marine ecosystems through protected areas : a viability model and a coral reef case study. *Ecological Modelling* 208, 353–366.
- Doyen, L., Péreau, J.C., 2012. Sustainable coalitions in the commons. *Mathematical Social Sciences* 63, 57–64.
- Duquette, M.C., Compérot, A., Hayes, L., Pagola, C., Belzile, F., Dubé, J., Lavoie, C., 2016. From the source to the outlet : understanding the distribution of invasive knotweeds along a north american river. *River Research and Applications* 32, 958–966.
- Durand, M.H., Désilles, A., Saint-Pierre, P., Angeon, V., Ozier-Lafontaine, H., 2017. Agroecological transition : A viability model to assess soil restoration. *Natural resource modeling* 30, e12134.
- Fennell, M., Murphy, J.E., Armstrong, C., Gallagher, T., Osborne, B., 2012. Plant spread simulator : a model for simulating large-scale directed dispersal processes across heterogeneous environments. *Ecological Modelling* 230, 1–10.
- Forman, J., Kesseli, R.V., 2003. Sexual reproduction in the invasive species *Fallopia Japonica* (*Polygonaceae*). *American Journal of Botany* 90, 586–592.
- Fournier, N., Méléard, S., 2004. A microscopic probabilistic description of a locally regulated population and macroscopic approximations. *Annals of applied probability* , 1880–1919.
- Fox, J.C., Buckley, Y.M., Panetta, F., Bourgoin, J., Pullar, D., 2009. Surveillance protocols for management of invasive plants : modelling chilean needle grass (*nassella neesiana*) in australia. *Diversity and Distributions* 15, 577–589.
- Gajardo, P., Olivares, M., Ramírez, H., 2018. Methods for the sustainable rebuilding of overexploited natural resources. *Environmental Modeling & Assessment* 23, 713–727.
- Gerber, E., Krebs, C., Murrell, C., Moretti, M., Rocklin, R., Schaffner, U., 2008. Exotic invasive knotweeds (*Fallopia spp.*) negatively affect native plant and invertebrate assemblages in European riparian habitats. *Biological Conservation* 141, 646–654.
- Gerber, E., Murrell, C., Krebs, C., Bilat, J., Schaffner, U., 2010. Evaluating non-chemical management methods against invasive exotic knotweeds, *Fallopia Spp.* CABI, Egham .
- Golse, F., 2003. The mean-field limit for the dynamics of large particle systems. *Journées équations aux dérivées partielles* , 1–47.
- Gourley, S.A., Li, J., Zou, X., 2016. A mathematical model for biocontrol of the invasive weed *Fallopia japonica*. *Bulletin of Mathematical Biology* 78, 1678–1702.
- Gowton, C., Budsock, A., Matlaga, D., 2016. Influence of disturbance on Japanese knotweed (*Fallopia japonica*) stem and rhizome fragment recruitment success within riparian forest understory. *Natural Areas Journal* 36, 259–267.

- Groves, J.H., Williams, D.G., Caley, P., Norris, R.H., Caitcheon, G., 2009. Modelling of floating seed dispersal in a fluvial environment. *River Research and Applications* 25, 582–592.
- Haidar, I., Alvarez, I., Prévot, A.C., 2017. Mathematical modeling of an urban pigeon population subject to local management strategies. *Mathematical biosciences* 288, 71–83.
- Happe, K., Kellermann, K., Balmann, A., 2006. Agent-based analysis of agricultural policies : an illustration of the agricultural policy simulator agropolis, its adaptation and behavior. *Ecology and Society* 11.
- Harris, C.M., Park, K.J., Atkinson, R., Edwards, C., Travis, J., 2009. Invasive species control : incorporating demographic data and seed dispersal into a management model for *Rhododendron ponticum*. *Ecological Informatics* 4, 226–233.
- Hui, C., Richardson, D.M., 2017. *Invasion dynamics*. Oxford University Press.
- Hui, C., Richardson, D.M., Landi, P., Minoarivelo, H.O., Garnas, J., Roy, H.E., 2016. Defining invasiveness and invasibility in ecological networks. *Biological Invasions* 18, 971–983.
- Ikeda, N., Watanabe, S., 1989. *Stochastic Differential Equations and Diffusion Processes*. volume Second Edition. North- Holland Mathematical Library.
- Jacod, J., Shiryaev, A.N., 2013. *Limit theorems for stochastic processes*. volume 288. Springer Science & Business Media.
- Joffe, A., Métivier, M., 1986. Weak convergence of sequences of semimartingales with applications to multitype branching processes. *Advances in Applied Probability* 18, 20–65.
- Juliano, S.A., Philip Lounibos, L., 2005. Ecology of invasive mosquitoes : effects on resident species and on human health. *Ecology Letters* 8, 558–574.
- Kalisiak, M., van de Panne, M., 2004. Approximate safety enforcement using computed viability envelopes, in : *IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA'04*. 2004, IEEE. pp. 4289–4294.
- Kettunen, M., Genovesi, P., Gollasch, S., Pagad, S., Starfinger, U., ten Brink, P., Shine, C., 2009. Technical support to EU strategy on invasive alien species (ias). London : Institut for European Environmental Policy (IEEP) .
- Kohonen, T., 1982. Self-organized formation of topologically correct feature maps. *Biological cybernetics* 43, 59–69.
- Kohonen, T., 2000. *Self-Organizing Maps*. Springer-Verlag Berlin Heidelberg.
- Kroon, H.d., Groenendaal, J.v., 1997. *The ecology and evolution of clonal plants*. Backhuys Publishers.
- Lamberti-Raverot, B., 2016. *Stratégies de propagation du complexe d'espèces Fallopiia par les cours d'eau : rôle des traits de dispersion et de colonisation*. Ph.D. thesis. Lyon.
- Lamperti, F., Roventini, A., Sani, A., 2018. Agent-based model calibration using machine learning surrogates. *Journal of Economic Dynamics and Control* 90, 366–389.

- Lavallée, F., Smadi, C., Alvarez, I., Reineking, B., Martin, F.M., Dommanget, F., Martin, S., 2019a. Quels apports de la modélisation pour l'aide à la gestion de la renouée du Japon. *Sciences Eaux et Territoires* 27, 108–111.
- Lavallée, F., Smadi, C., Alvarez, I., Reineking, B., Martin, F.M., Dommanget, F., Martin, S., 2019b. A stochastic individual based model for the growth of a stand of japanese knotweed including mowing as a management technique. *Ecological Modelling* 413, 108828.
- Lavoie, C., 2017. The impact of invasive knotweed species (*Reynoutria spp.*) on the environment : review and research perspectives. *Biological Invasions* 19, 2319–2337.
- Lee, J.S., Filatova, T., Ligmann-Zielinska, A., Hassani-Mahmoei, B., Stonedahl, F., Lorscheid, I., Voinov, A., Polhill, G., Sun, Z., Parker, D.C., 2015. The complexities of agent-based modeling output analysis. *The journal of artificial societies and social simulation* 18.
- Leman, H., 2016. Convergence of an infinite dimensional stochastic process to a spatially structured trait substitution sequence. *Stochastics and Partial Differential Equations : Analysis and Computations* 4, 791–826.
- Lenormand, M., Jabot, F., Deffuant, G., 2013. Adaptive approximate bayesian computation for complex models. *Computational Statistics* 28, 2777–2796.
- Ligmann-Zielinska, A., 2013. Spatially-explicit sensitivity analysis of an agent-based model of land use change. *International Journal of Geographical Information Science* 27, 1764–1781.
- Ligmann-Zielinska, A., Siebers, P.O., Magliocca, N., Parker, D.C., Grimm, V., Du, J., Cenek, M., Radchuk, V., Arbab, N.N., Li, S., et al., 2020. 'one size does not fit all' : A roadmap of purpose-driven mixed-method pathways for sensitivity analysis of agent-based models. *Journal of Artificial Societies and Social Simulation* 23, 1–6.
- Lookingbill, T.R., Minor, E.S., Bukach, N., Ferrari, J.R., Wainger, L.A., 2014. Incorporating risk of reinvasion to prioritize sites for invasive species management. *Natural Areas Journal* 34, 268–281.
- Maerz, J.C., Blossey, B., Nuzzo, V., 2005. Green frogs show reduced foraging success in habitats invaded by Japanese knotweed. *Biodiversity & Conservation* 14, 2901–2911.
- Mari, L., Bertuzzo, E., Casagrandi, R., Gatto, M., Levin, S.A., Rodriguez-Iturbe, I., Rinaldo, A., 2011. Hydrologic controls and anthropogenic drivers of the zebra mussel invasion of the mississippi-missouri river system. *Water Resources Research* 47.
- Martin, F.M., Dommanget, F., Janssen, P., Spiegelberger, T., Viguié, C., Evette, A., 2018. Could knotweeds invade mountains in their introduced range? An analysis of patches dynamics along an elevational gradient. *Alpine Botany* .
- Martin, F.M., Dommanget, F., Lavallée, F., Évette, A., 2020. Clonal growth strategies of *Reynoutria japonica* in response to light, shade, and mowing, and perspectives for management. *NeoBiota* , 89–110.
- Martin, F.M., Dommanget, F., Lavoie, C., Evette, A., en préparation. The spatiotemporal dynamics of a strong invader : a review on knotweeds (*Reynoutria spp.*) invasion patterns and processes across scales .

- Martin, S., 2004. The cost of restoration as a way of defining resilience : a viability approach applied to a model of lake eutrophication. *Ecology and Society* 9.
- Martin, S., 2015. D'oxymore en oxymore : du développement durable au contrôle complexe. Habilitation à diriger des recherches en mathématiques appliquées. Université Blaise Pascal Clermont-Ferrand.
- Martin, S., 2019. Viability analysis as an approach for assessing the resilience of agroecosystems. *Ecological Reviews*, Cambridge University Press. doi :10.1017/9781107705555.
- Martin, S., Alvarez, I., Kant, J.D., 2015. Micro/macro viability analysis of individual-based models : Investigation into the viability of a stylized agricultural cooperative. *Complexity* 21, 276–296.
- Martin, S., Alvarez, I., Lavallée, F., en cours d'écriture. Using individual strategies derived from guaranteed viability kernels to solve viability problems in  $N$ -player discrete-time dynamic games : application to an agricultural cooperative model .
- Martin, S., Deffuant, G., Calabrese, J.M., 2011. Defining resilience mathematically : from attractors to viability, in : *Viability and resilience of complex systems*. Springer, pp. 15–36.
- Martinet, V., Thébaud, O., Doyen, L., 2007. Defining viable recovery paths toward sustainable fisheries. *Ecological Economics* 64, 411–422.
- Meier, E.S., Dullinger, S., Zimmermann, N.E., Baumgartner, D., Gattringer, A., Hülber, K., 2014. Space matters when defining effective management for invasive plants. *Diversity and Distributions* 20, 1029–1043.
- Méléard, S., Bansaye, V., 2015. *Stochastic Models for Structured Populations : Scaling Limits and Long Time Behavior*. volume 1. Springer.
- Méléard, S., Roelly, S., 1993. Sur les convergences étroite ou vague de processus à valeurs mesures. *Comptes rendus de l'Académie des sciences. Série 1, Mathématique* 317, 785–788.
- Monnin, T., Lavallée, F., Chérel, G., en cours d'écriture. No communication required between workers for resources allocation during colony fission in a social insect : an individual-based model without communication reproduces empirical patterns .
- Morris, M.D., 1991. Factorial sampling plans for preliminary computational experiments. *Technometrics* 33, 161–174.
- Oubraham, A., Zaccour, G., 2018. A survey of applications of viability theory to the sustainable exploitation of renewable resources. *Ecological economics* 145, 346–367.
- Parzen, E., 1962. On estimation of a probability density function and mode. *The annals of mathematical statistics* 33, 1065–1076.
- Péreau, J.C., Doyen, L., Little, L., Thébaud, O., 2012. The triple bottom line : Meeting ecological, economic and social goals with individual transferable quotas. *Journal of Environmental Economics and Management* 63, 419–434.
- Planty-Tabacchi, A.M., Tabacchi, E., Naiman, R.J., Deferrari, C., Decamps, H., 1996. Invasibility of species-rich communities in riparian zones. *Conservation Biology* 10, 598–607.

- Price, E.A., Gamble, R., Williams, G.G., Marshall, C., 2002. Seasonal patterns of partitioning and remobilization of  $^{14}\text{C}$  in the invasive rhizomatous perennial Japanese knotweed (*Fallopia japonica* (Houtt.) Ronse Decraene). *Ecology and Evolutionary Biology of Clonal Plants*, 125–140.
- Queffelec, H., Zuily, C., 2013. *Analyse pour l'agrégation : cours et exercices corrigés*. Dunod.
- Rapaport, A., Terreaux, J.P., Doyen, L., 2006. Viability analysis for the sustainable management of renewable resources. *Mathematical and computer modelling* 43, 466–484.
- Reuillon, R., Leclaire, M., Rey-Coyrehourcq, S., 2013. Openmole, a workflow engine specifically tailored for the distributed exploration of simulation models. *Future Generation Computer Systems* 29, 1981 – 1990. URL : <http://www.openmole.org/files/FGCS2013.pdf>.
- Reuillon, R., Schmitt, C., De Aldama, R., Mouret, J.B., 2015. A new method to evaluate simulation models : The calibration profile (cp) algorithm. *Journal of Artificial Societies and Social Simulation* 18, 12. URL : <http://jasss.soc.surrey.ac.uk/18/1/12.html>, doi :10.18564/jasss.2675.
- Richardson, D.M., Holmes, P.M., Esler, K.J., Galatowitsch, S.M., Stromberg, J.C., Kirkman, S.P., Pyšek, P., Hobbs, R.J., 2007. Riparian vegetation : degradation, alien plant invasions, and restoration prospects. *Diversity and Distributions* 13, 126–139.
- Richardson, D.M., Pyšek, P., Rejmánek, M., Barbour, M.G., Panetta, F.D., West, C.J., 2000. Naturalization and invasion of alien plants : concepts and definitions. *Diversity and distributions* 6, 93–107.
- Riis, T., Sand-Jensen, K., 2006. Dispersal of plant fragments in small streams. *Freshwater Biology* 51, 274–286.
- Roelly-Coppoletta, S., 1986. A criterion of convergence of measure-valued processes : application to measure branching processes. *Stochastics : An International Journal of Probability and Stochastic Processes* 17, 43–65.
- Rouifed, S., Bornette, G., Mistler, L., Piola, F., 2011a. Contrasting response to clipping in the Asian knotweeds *Fallopia japonica* and *Fallopia*  $\times$  *bohemica*. *Ecoscience* 18, 110–114.
- Rouifed, S., Piola, F., Spiegelberger, T., 2014. Invasion by *Fallopia* spp. in a french upland region is related to anthropogenic disturbances. *Basic and Applied Ecology* 15, 435–443.
- Rouifed, S., Puijalon, S., Viricel, M.R., Piola, F., 2011b. Achene buoyancy and germinability of the terrestrial invasive *Fallopia*  $\times$  *Bohemica* in aquatic environment : A new vector of dispersion ? *Ecoscience* 18, 79–84.
- Roy, B., 1985. *Méthodologie multicritère d'aide à la décision*. BOOK, Economica.
- Rubin, D.B., 1984. Bayesianly justifiable and relevant frequency calculations for the applied statistician. *The Annals of Statistics*, 1151–1172.
- Saint-Pierre, P., 1994. Approximation of the viability kernel. *Applied Mathematics and Optimization* 29, 187–209.
- Saldaña, A., Fuentes, N., Pfanzelt, S., 2009. *Fallopia japonica* (Houtt.) Ronse Decraene. (Polygonaceae), : A new record for the alien flora of Chile. *Gayana. Botánica* 66, 283–285.



- Saltelli, A., Annoni, P., 2010. How to avoid a perfunctory sensitivity analysis. *Environmental Modelling & Software* 25, 1508–1517.
- Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., Tarantola, S., 2010. Variance based sensitivity analysis of model output. design and estimator for the total sensitivity index. *Computer Physics Communications* 181, 259–270.
- Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., Tarantola, S., 2008. *Global sensitivity analysis : the primer*. John Wiley & Sons.
- Sásik, R., Eliás, P., 2006. Rhizome regeneration of *Fallopia Japonica* (Japanese knotweed)(Houtt.) Ronse Decr. i. regeneration rate and size of regenerated plants. *Folia Oecologica* 33, 57.
- Schmitt, C., Rey-Coyrehourcq, S., Reuillon, R., Pumain, D., 2015. Half a billion simulations : Evolutionary algorithms and distributed computing for calibrating the simpoplocal geographical model. *Environment and Planning B : Planning and Design* 42, 300–315.
- Seiger, L.A., Merchant, H.C., et al., 1997. Mechanical control of Japanese knotweed (*Fallopia japonica* [Houtt.] Ronse Decraene) : Effects of cutting regime on rhizomatous reserves. *Natural Areas Journal* 17, 341–345.
- Serniak, L.T., Corbin, C.E., Pitt, A.L., Rier, S.T., 2017. Effects of Japanese knotweed on avian diversity and function in riparian habitats. *Journal of Ornithology* 158, 311–321.
- Sicard, M., Perrot, N., Reuillon, R., Mesmoudi, S., Alvarez, I., Martin, S., 2012. A viability approach to control food processes : Application to a camembert cheese ripening process. *Food Control* 23, 312–319.
- Siemens, T.J., Blossey, B., 2007. An evaluation of mechanisms preventing growth and survival of two native species in invasive bohemian knotweed (*fallopia* × *bohemica*, polygonaceae). *American Journal of Botany* 94, 776–783.
- Simberloff, D., Martin, J.L., Genovesi, P., Maris, V., Wardle, D.A., Aronson, J., Courchamp, F., Galil, B., García-Berthou, E., Pascal, M., et al., 2013. Impacts of biological invasions : what's what and the way forward. *Trends in ecology & evolution* 28, 58–66.
- Smadi, C., Leman, H., Llaurens, V., 2018. Looking for the right mate in diploid species : How does genetic dominance affect the spatial differentiation of a sexual trait ? *Journal of Theoretical Biology* 447, 154–170.
- Smith, J., Ward, J.P., Child, L.E., Owen, M., 2007. A simulation model of rhizome networks for *Fallopia japonica* (Japanese knotweed) in the United Kingdom. *Ecological Modelling* 200, 421–432.
- Stonedahl, F.J., 2011. Genetic algorithms for the exploration of parameter spaces in agent-based models. Ph.D. thesis. Northwestern University.
- Suzuki, J.I., 1994. Growth dynamics of shoot height and foliage structure of a rhizomatous perennial herb, *Polygonum cuspidatum*. *Annals of Botany* 73, 629–638.
- Terrio, P.J., Nazimek, J.E., 1997. Changes in cross-section geometry and channel volume in two reaches of the Kankakee River in Illinois, 1959-94. US Department of the Interior, US Geological Survey.

- Tharayil, N., Alpert, P., Bhowmik, P., Gerard, P., 2013. Phenolic inputs by invasive species could impart seasonal variations in nitrogen pools in the introduced soils : a case study with *Polygonum cuspidatum*. *Soil Biology and Biochemistry* 57, 858–867.
- Tiébré, M.S., Vanderhoeven, S., Saad, L., Mahy, G., 2007. Hybridization and sexual reproduction in the invasive alien *Fallopia* (*Polygonaceae*) complex in belgium. *Annals of Botany* 99, 193–203.
- Tomlin, C.J., Mitchell, I., Bayen, A.M., Oishi, M., 2003. Computational techniques for the verification of hybrid systems. *Proceedings of the IEEE* 91, 986–1001.
- Tourian, M.J., Elmi, O., Mohammadnejad, A., Sneeuw, N., 2017. Estimating river depth from swot-type observables obtained by satellite altimetry and imagery. *Water* 9, 753.
- Tran, V.C., 2006. Modèles particuliers stochastiques pour des problèmes d'évolution adaptative et pour l'approximation de solutions statistiques. Ph.D. thesis. Université de Nanterre-Paris X.
- Tran, V.C., 2008. Large population limit and time behaviour of a stochastic particle model describing an age-structured population. *ESAIM : Probability and Statistics* 12, 345–386.
- Travis, J.M., Harris, C.M., Park, K.J., Bullock, J.M., 2011. Improving prediction and management of range expansions by combining analytical and individual-based modelling approaches. *Methods in Ecology and Evolution* 2, 477–488.
- Wei, W., Alvarez, I., Martin, S., 2013. Sustainability analysis : Viability concepts to consider transient and asymptotical dynamics in socio-ecological tourism-based systems. *Ecological Modelling* 251, 103–113.
- Wilensky, U., et al., 1999. Center for connected learning and computer-based modeling, in : NetLogo. Northwestern University.
- Yamazaki, D., O'Loughlin, F., Trigg, M., Miller, Z., Pavelsky, T., Bates, P., 2014. Development of the global width database for large rivers. *Water Resources Research* 50. doi :10.1002/2013WR014664.
- Zeide, B., 1993. Analysis of growth equations. *Forest science* 39, 594–616.