



Modèle individu-centré de transmission de la tremblante dans un troupeau ovin

Najat Ziyadi, Suzanne Touzeau, Caroline C. Bidot, Jean-Pierre Treuil, My Lhassan Hbid

► To cite this version:

Najat Ziyadi, Suzanne Touzeau, Caroline C. Bidot, Jean-Pierre Treuil, My Lhassan Hbid. Modèle individu-centré de transmission de la tremblante dans un troupeau ovin. [Rapport Technique] 2008-4, 2008. hal-02822910

HAL Id: hal-02822910

<https://hal.inrae.fr/hal-02822910>

Submitted on 6 Jun 2020

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.

Modèle individu-centré de transmission de la tremblante dans un troupeau ovin

Najat ZIYADI¹

Suzanne TOUZEAU²

Caroline BIDOT

Jean-Pierre TREUIL

My Lhassan HBID

Rapport technique 2008-4, 50 (2+48) pp.

Unité Mathématiques et Informatique Appliquées
INRA
Domaine de Vilvert
78352 Jouy-en-Josas Cedex
France

¹ najat_ziyadi@yahoo.fr

² Suzanne.Touzeau@jouy.inra.fr

Modèle individu-centré de transmission de la tremblante dans un troupeau ovin¹

Najat ZIYADI*, Suzanne TOUZEAU[†], Caroline BIDOT[‡],
Jean-Pierre TREUIL[‡], My Lhassan HBID*

*[LMDP](#), Faculté des Sciences Semlalia, Université Cadi Ayyad, Marrakech, Maroc

[†][UR341 MIA](#), INRA, Jouy-en-Josas, France

[‡][UR079 GEODES](#), IRD, Bondy, France

Décembre 2008

¹Ce travail a été réalisé dans le cadre de la thèse de doctorat de Najat ZIYADI, dirigée par M. L. HBID. Le modèle a été développé et implémenté lors des séjours de N. ZIYADI dans l'UR341 MIA de l'INRA Jouy-en-Josas, en collaboration avec S. TOUZEAU (co-encadrante – modélisation), J.-P. TREUIL (co-encadrant – simulateur) et C. BIDOT (estimation des paramètres).

*Ce travail a été partiellement soutenu par **NeuroPrion**, Réseau d'Excellence européen sur les maladies à prions.*

Table des matières

Table des matières	3
Liste des figures	5
Liste des tableaux	6
I Introduction	7
I.1 Contexte biologique : la tremblante ovine	7
I.1.1 Transmission et détection	7
I.1.2 Impact de la tremblante	8
I.2 Motivations	8
I.3 Approche retenue : modélisation individu-centrée	8
I.4 Objectif et plan du rapport	9
II Présentation du modèle individu-centré	11
II.1 Choix de modélisation	11
II.2 Les individus	12
II.3 Les processus	14
II.3.1 Hypothèses	14
II.3.1.1 Hypothèses sur le troupeau	14
II.3.1.2 Hypothèses de contamination	14
II.3.1.3 Hypothèses d'incubation	14
II.3.1.4 Hypothèses sur la mortalité	15
II.3.2 Contamination	15
II.3.2.1 Susceptibilité par âge	15
II.3.2.2 Susceptibilité génétique	15
II.3.2.3 Force d'infection	17
II.3.2.4 Risque de contamination	17
II.3.2.5 Implémentation	18
II.3.3 Incubation	18
II.3.3.1 Charge initiale d'infection	19
II.3.3.2 Évolution de la charge	19
II.3.3.3 Mortalité tremblante	19
II.4 La condition initiale	20
II.5 Les sorties	20
II.5.1 Sorties démographiques	21
II.5.2 Sorties épidémiologiques	21
III Description du simulateur	23
III.1 Architecture des classes du programme	23
III.1.1 La classe Mouton	23

III.1.2 La classe Troupeau	23
III.1.3 La classe MyModele	25
III.1.4 La classe MyFrame	25
III.2 Interface graphique	25
III.2.1 Paramètres d'entrée	25
III.2.1.1 Paramètres fixes	25
III.2.1.2 Paramètres ajustables	27
III.2.2 Sorties	27
III.2.2.1 Indicateurs quantitatifs	27
III.2.2.2 Indicateurs graphiques	27
III.2.3 Boutons de contrôle	28
III.3 Fonctionnement du simulateur	28
IV Estimation des paramètres	31
IV.1 Les données	31
IV.2 Les paramètres à estimer	32
IV.3 Le critère	34
IV.3.1 Niveau individuel	35
IV.3.1.1 Statut tremblante uniquement	35
IV.3.1.2 Statut tremblante et date de mort	35
IV.3.2 Niveau groupe d'animaux	35
IV.3.2.1 Statut tremblante uniquement	36
IV.3.2.2 Statut tremblante et date de mort	36
IV.4 L'algorithme d'optimisation	36
IV.4.1 Initialisation	37
IV.4.2 Apprentissage	37
IV.4.3 Exploitation	39
IV.4.4 Paramètres de l'algorithme et critère d'arrêt	40
IV.5 Premiers résultats	40
IV.5.1 Mise en œuvre	40
IV.5.1.1 Calcul du critère	40
IV.5.1.2 Gestion de la stochasticité du modèle	41
IV.5.1.3 Espace des paramètres	41
IV.5.1.4 Réglages de l'algorithme de minimisation	41
IV.5.2 Résultats	42
V Conclusion et perspectives	43
Bibliographie	45
Résumé	48

Liste des figures

II.1 Cycle de vie d'un mouton	13
II.2 Détermination de la date de mort simulée d'un individu	16
II.3 Susceptibilité par âge	16
III.1 Schéma mettant en relation les classes <code>MyModel</code> et <code>MyFrame</code> du simulateur .	25
III.2 Interface graphique du simulateur	26
IV.1 Cas de tremblante naturelle dans le troupeau expérimental de Langlade . .	31
IV.2 Détection d'un cas de tremblante dans le modèle	34
IV.3 Schéma de l'algorithme de recherche aléatoire par écart-type adaptatif <code>ARS</code>	38

Liste des tableaux

III.1 Attributs et méthodes de la classe Mouton	24
III.2 Attributs et méthodes de la classe Troupeau	24
III.3 Enchaînement des opérations du simulateur	30
IV.1 Classification des génotypes selon leur résistance à la tremblante classique .	32
IV.2 Paramètres à estimer dans le modèle avant et après simplification	33
IV.3 Nombre de calculs du critère pendant l'apprentissage de l'algorithme ARS .	39
IV.4 Espace des paramètres utilisé par l'algorithme ARS	41
IV.5 Valeurs des paramètres estimés par l'algorithme ARS	42

Chapitre I

Introduction

Ce chapitre introductif présente tout d’abord le contexte biologique de cette étude, qui porte sur la transmission de la tremblante dans un troupeau ovin. Puis nos motivations sont exposées. Notre approche, la modélisation individu-centrée, est rapidement justifiée et présentée. Enfin, l’objectif de notre étude est formulé et le plan du rapport est présenté.

I.1 Contexte biologique : la tremblante ovine

La tremblante est une encéphalopathie spongiforme transmissible (EST), qui touche naturellement les ovins et les caprins. Les EST sont des maladies neuro-dégénératives à progression lente et issue fatale, qui sont caractérisées par l’accumulation dans le cerveau d’une conformation anormale de la protéine prion PrP. Les EST humaines sont rares, la maladie de Creutzfeld-Jacob est la plus connue. Les EST animales incluent, entre autres, l’encéphalopathie spongiforme bovine ou « vache folle » qui a débuté au Royaume Uni en 1985 et la maladie du dépérissement chronique qui touche les cervidés aux États-Unis. La tremblante est connue depuis le XVIIIe siècle et c’est une pathologie relativement courante dans les troupeaux européens, voire endémique dans certaines régions [4].

I.1.1 Transmission et détection

La tremblante se propage dans les troupeaux, mais ses mécanismes de transmission ne sont que partiellement connus [9]. La transmission se produit principalement par ingestion, même si d’autres voies d’entrées sont possibles [1]. Ainsi, la présence de l’agent pathogène de la tremblante dans les placentas suggère une transmission accrue pendant les mises bas, par l’ingestion de placentas contaminés [14]. La tremblante ovine est intègre un facteur génétique, lié au polymorphisme du gène PrP codant pour la protéine prion. Selon leur génotype PrP, les animaux peuvent être très sensibles à totalement résistants. Par ailleurs, les animaux jeunes sont a priori plus susceptibles face à la maladie.

La durée d’incubation de la tremblante est particulièrement longue par rapport à la durée de vie des ovins, de l’ordre de 2 ans. Les animaux sont infectieux pendant cette période, mais on ne peut pas les distinguer des animaux sains. À l’issue de cette période, ils développent des signes cliniques et meurent ou sont abattus.

Le diagnostic de la tremblante se fait à l'aide des tests *post-mortem*, qui permettent de détecter la maladie vers la fin de la période d'incubation. Il n'existe pas de tests pré-cliniques fiables et validés sur animaux vivants.

I.1.2 Impact de la tremblante

Contrairement à l'encéphalopathie spongiforme bovine (ESB), la tremblante ne semble pas constituer un risque sanitaire pour les humains. Cependant, des études expérimentales ont montré que l'agent de l'ESB peut se transmettre au mouton et génère des signes cliniques semblables à ceux de la tremblante. Il est ainsi possible que l'agent de l'ESB soit présent dans les troupeaux ovins et que les cas soient confondus avec des cas de tremblante. Aucune infection naturelle n'a encore été mise en évidence chez les ovins; cependant un cas de contamination naturelle par l'ESB chez une chèvre a été observé en France.

L'Union Européenne a mis en place un contexte réglementaire qui vise à éradiquer la tremblante. Il repose sur des campagnes destinées à améliorer la résistance génétique des troupeaux par sélection et de la surveillance active et passive. Ces mesures ont un coût élevé. À l'échelle de l'élevage, l'impact de la tremblante est aussi important. Outre les pertes économiques liées à la mortalité engendrée par la maladie, des mesures coûteuses doivent être mises en place dès la détection des premiers cas, comme l'abattage sélectif et des tests systématiques.

I.2 Motivations

De nombreux chercheurs à l'INRA s'intéressent aux EST et en particulier à la tremblante du mouton du point de vue de la génétique, de la pathologie, du diagnostic de la maladie. Grâce à une collaboration avec la Station d'Amélioration Génétique des Animaux (SAGA) de l'INRA Toulouse, à la ferme expérimentale de Langlade et à l'IHAP de l'ENV-Toulouse, nous disposons de connaissances en matière de tremblante et de données de suivi de troupeau pour quelques élevages atteints de tremblante [6, 3, 5, 14]. Plusieurs éléments nous incitent donc à nous intéresser à la tremblante :

1. l'existence d'un risque sanitaire potentiel ;
2. un impact économique important ;
3. les incertitudes subsistant quant aux mécanismes de transmissions ;
4. la disponibilité des données.

I.3 Approche retenue : modélisation individu-centrée

Nous avons choisi d'aborder l'étude de cette maladie par une approche de modélisation. Un modèle est un bon outil pour étudier un système épidémiologique, à la fois intégratif et flexible. En effet, il permet de combiner données et connaissances issues de plusieurs sources. Il permet aussi d'explorer le comportement du système selon différents scénarios. C'est donc un complément relativement peu coûteux à l'expérimentation.

Plusieurs modèles de transmission de la tremblante dans un troupeau ont déjà été développés [13, 8, 12]. Il s'agit de modèles de dynamique de population, dans lesquels le troupeau est structuré en génotypes, âge et charge ou âge d'infection ; ils intègrent des processus démographiques et épidémiologiques réalistes. En résultent des modèles mathématiques relativement complexes, dont l'étude analytique est limitée [18], mais que l'on peut néanmoins étudier et confronter à des données [11, 10, 14].

Les modèles individus-centrés (ou MIC) sont des modèles informatiques issus de l'approche multi-agents [15, 16]. Ils sont plus flexibles que ces modèles de population. Ils permettent de représenter de manière fine et intuitive la variabilité et le comportement des individus, au-delà de la structuration des modèles de population. Un MIC est constitué d'un ensemble d'individus caractérisés par un certain nombre de variables. Il intègre des règles d'interaction entre individus, qui peuvent s'appliquer différemment selon les individus et évoluer au cours du temps. On simule et on observe l'évolution des caractéristiques des individus au cours du temps. Le comportement de la population est ainsi déterminé à partir des comportements individuels. Dans cette approche, chaque individu doit être caractérisé, ce qui nécessite de bonnes connaissances sur la population représentée. Comme nous disposons de données de suivi assez précises sur des troupeaux atteints de tremblante, nous avons retenu cette approche.

I.4 Objectif et plan du rapport

L'objectif de cette étude est donc de développer un modèle individu-centré représentant la propagation de la tremblante au sein d'un troupeau. L'outil obtenu doit être flexible et convivial. Il doit permettre de tester différentes hypothèses de transmission de la maladie. Le travail s'articule donc en trois phases : la première consiste à élaborer le modèle, la seconde à l'implémenter et la troisième à calibrer ses paramètres à partir des données disponibles.

Dans ce rapport, nous présentons tout d'abord en détails le modèle individu-centré. Puis nous donnons une description du simulateur Java implémenté pour ce modèle et de son interface graphique. Une méthode d'estimation des paramètres épidémiologiques du modèle est ensuite exposée, accompagnée de premiers résultats. Le rapport se termine par des conclusion et perspectives.

Chapitre II

Présentation du modèle individu-centré

Dans ce modèle de troupeau ovin, un individu correspond naturellement à un mouton. Le modèle représente les interactions entre ces individus. Dans ce chapitre, après avoir explicité nos choix de modélisation, nous caractérisons les individus. Puis nous détaillons les processus faisant interagir ces individus et évoluer leurs caractéristiques. Enfin, nous présentons la condition initiale et les sorties du modèle. Le déroulement d’une simulation est décrit plus en détails dans le chapitre [III](#).

II.1 Choix de modélisation

Plusieurs options de modélisation sont possibles. La première consiste à modéliser l’ensemble des processus régissant la dynamique des individus du troupeau :

- la reproduction,
- l’abattage / la réforme / l’équarrissage (pour raison autre que tremblante),
- la contamination,
- l’incubation,
- la mortalité tremblante.

La représentation de chacun de ces processus repose sur des hypothèses de modélisation et fait intervenir un grand nombre de paramètres.

La deuxième option consiste à modéliser uniquement les processus épidémiologiques :

- la contamination,
- l’incubation,
- la mortalité tremblante.

Les processus de reproduction et de réforme, liés à la conduite du troupeau, ne sont pas modélisés explicitement. Ils sont représentés directement par des dates de naissance et de réforme individuelles issues des données. Cette deuxième option est plus parcimonieuse en nombre de paramètres. En outre, elle isole les processus épidémiologiques et ne fait aucune hypothèse sur les processus démographiques. Mais elle nécessite la « généalogie complète » du troupeau, *i.e.* les dates de naissance et de mort de tous les individus, ainsi que leur génotype et leur statut tremblante.

Ces deux options de modélisation sont représentées sur la FIG. II.1. Dans ce modèle général, un individu peut naître sain ou infecté. S'il est sain, il peut changer de statut et devenir infecté par transmission horizontale de la maladie. Il entre alors dans une phase d'incubation infectieuse non détectable. Cette phase prend fin à l'apparition de signes cliniques, suite auxquels l'animal est abattu et considéré comme un cas de tremblante. La réforme s'applique aux animaux sains et infectés, qui sont abattus indifféremment de leur statut épidémiologique.

Comme nous disposons de données de suivi de troupeaux de bonne qualité, nous avons choisi la deuxième option : **ne modéliser que les processus épidémiologiques**.

Le modèle individu-centré développé dans cette étude est un **modèle stochastique en temps discret**. La stochasticité permet de représenter la variabilité entre individus du troupeau au niveau de la contamination et de la durée d'incubation. Le pas de temps est choisi comme un compromis entre le coût en temps de calcul d'une simulation et les échelles de temps des processus modélisés : 1-2 mois pour la période de contamination correspondant aux mises bas ; environ 2 ans pour la période d'incubation. Un pas de temps d'1 à 2 semaines est un bon compromis (minimum 1 jour, maximum 1 mois).

II.2 Les individus

Le modèle est constitué d'un ensemble d'individus, chaque individu correspondant à un mouton caractérisé par des paramètres et des variables.

Nous distinguons deux types de paramètres :

1. les paramètres issus directement des données :
 - (a) la date de naissance d_N ,
 - (b) la date de mort d_M ,
 - (c) le génotype PrP $g \in \{1, \dots, n_g\}$,
 - (d) le diagnostic tremblante h (*post-mortem*) admettant deux résultats :
 - i. $h = 1$ positif si la tremblante est confirmée à la mort de l'animal ;
 - ii. $h = 0$ négatif si elle n'est pas détectée ;
2. les paramètres déterminés par le modèle :
 - (a) la date de contamination d_c ,
 - (b) la date d'apparition des signes cliniques appelée date de mort tremblante d_T ,
 - (c) la date de mort dans le modèle, appelée aussi la date de mort simulée $d_s(d_M, d_T)$ (déterminée dans §II.3.1.4).

Les variables sont calculées à chaque instant t :

1. l'âge $a(t)$ est calculé à partir de la date de naissance d_N et de la date de mort simulée d_s comme suit :
 - $a(t) = 0$ par convention si $t < d_N$,
 - $a(t) = t - d_N$ si $d_N \leq t \leq d_s$,
 - $a(t) = d_s - d_N$ par convention si $t > d_s$.
2. la charge d'infection $\theta(t) \in [0, 1]$ est estimée par le modèle et représente l'inféctiosité (§II.3.2.3) et le degré d'incubation (§II.3.3) de l'individu.

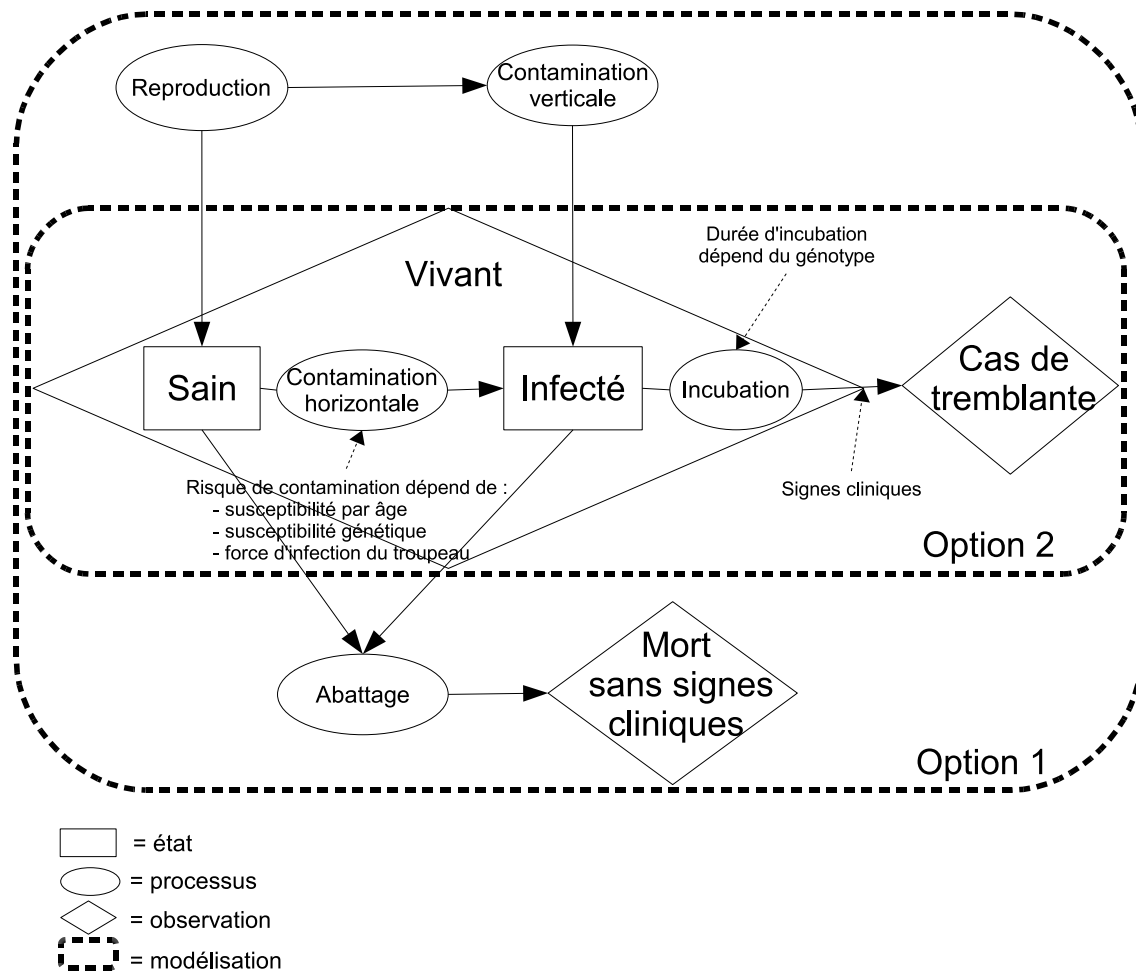


FIG. II.1 – Cycle de vie d'un mouton avec mise en évidence des états épidémiologiques (sain ou infecté), des états observés (vivant, mort sans signes cliniques, ou cas de tremblante) et des processus impliqués selon les deux options de modélisation :

- (1) modélisation de tous les processus, démographiques (reproduction et abattage) et épidémiologiques (transmission et incubation) ;
- (2) modélisation uniquement des processus épidémiologiques (transmission et incubation).

II.3 Les processus

Les seuls processus représentés dans le modèle sont les processus épidémiologiques, *i.e.* la contamination et l'incubation. La mortalité tremblante intervient à la fin de la période d'incubation. Ces processus reposent sur les hypothèses exprimées ci-dessous.

II.3.1 Hypothèses

II.3.1.1 Hypothèses sur le troupeau

1. Le troupeau est formé uniquement par les *brebis de renouvellement* (brebis utilisées pour la reproduction) ayant vécu tout ou partie de la période choisie pour la simulation.
2. Le troupeau est *fermé* : on suppose qu'il n'y a pas ni achat ni vente pendant la période de simulation.

II.3.1.2 Hypothèses de contamination

1. Le troupeau est *sans source externe de contamination*.
2. Seule la *transmission horizontale* est considérée, *i.e.* la contamination d'un animal par les animaux infectés du troupeau. La transmission verticale *in-utero* de la brebis à l'agneau est négligée. Nous supposons ainsi que tout animal naît sain, mais il peut s'infecter à la naissance dès l'âge 0 par transmission horizontale [2].
3. La transmission est accrue *pendant les périodes de mise bas* [14]. On fait l'hypothèse qu'elle est négligeable en dehors de ces périodes. Elle est donc saisonnière et est représentée par la fonction saison $s(t)$, définie comme un créneau périodique et telle que :

$$s(t) = \begin{cases} 1 & \text{si } t \in \text{ saison de transmission/mise bas,} \\ 0 & \text{sinon.} \end{cases} \quad (\text{II.1})$$

4. Au niveau individuel, la transmission se traduit par un *risque de contamination*, qui pour tout animal sain dépend de :
 - (a) son génotype (susceptibilité génétique),
 - (b) son âge (susceptibilité décroissant avec l'âge),
 - (c) le temps (saison de mise bas ou non),
 - (d) la force d'infection du troupeau au temps considéré.

II.3.1.3 Hypothèses d'incubation

1. La *durée d'incubation* est définie le temps écoulé entre la date de contamination et l'apparition des signes cliniques : $\tau = d_T - d_c$. Elle est longue par rapport à la durée de vie de l'animal. Elle dépend du *génotype* de l'individu infecté et est *variable* entre individus (même au sein du même génotype).
2. Les infectés sont *infectieux* pendant la période d'incubation.
3. La mort par tremblante survient dès l'apparition des signes cliniques à la fin de la période d'incubation.

II.3.1.4 Hypothèses sur la mortalité

La date de mortalité simulée d_s est déterminée par la date de mort des données d_M et la date de mort tremblante d_T calculée par le modèle. Deux choix simples en découlent :

1. $d_s = \max(d_M, d_T)$; dans ce cas, l'effectif du troupeau est surestimé par le modèle par rapport aux données ;
2. $d_s = \min(d_M, d_T)$; dans ce cas, l'effectif est sous-estimé.

Nous optons pour le deuxième choix :

$$d_s = \min(d_M, d_T). \quad (\text{II.2})$$

Plusieurs cas de figure sont possibles avec cette définition, comme le montre la FIG. II.2.

II.3.2 Contamination

A chaque instant, tout individu vivant et sain peut s'infecter. Sa probabilité d'infection, appelée *risque de contamination*, dépend du temps, de l'état du troupeau à cet instant et des caractéristiques de l'individu, âge et génotype (§II.3.1.2).

II.3.2.1 Susceptibilité par âge

On suppose que la susceptibilité d'un individu décroît avec son âge a et on modélise cet effet par une fonction seuil (FIG. II.3). La susceptibilité par âge est donc

$$\sigma_a = \begin{cases} 1 & \text{si } 0 \leq a < a_s, \\ \varepsilon_a \in [0, 1] & \text{sinon.} \end{cases} \quad (\text{II.3})$$

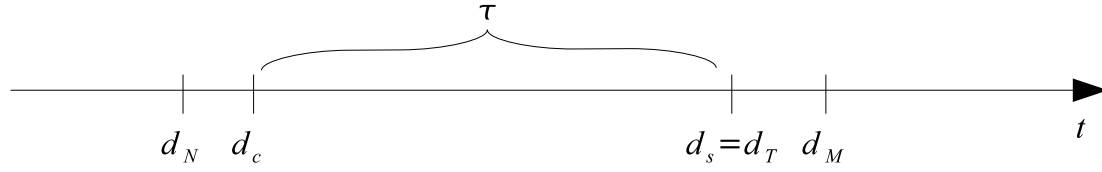
N.B. : Pour obtenir une susceptibilité par âge constante, il suffit de prendre $\varepsilon_a = 1$ (et a_s quelconque).

II.3.2.2 Susceptibilité génétique

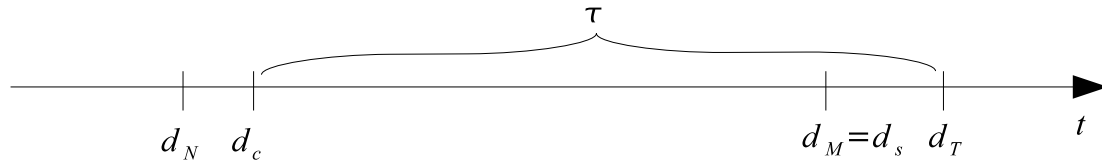
Le polymorphisme du gène PrP codant pour la protéine prion aux codons 136, 154 et 171 contrôle la susceptibilité génétique des individus à la tremblante. $n_a = 4$ allèles ont été détectés dans les données : VRQ, ARQ, AHQ et ARR. Ils engendrent $n_g = 10$ génotypes.

La susceptibilité génétique relative σ_g dépend du génotype g . Elle vérifie les relations suivantes :

$$\sigma_g = \begin{cases} 1 & \text{si } g : \text{VRQ-VRQ, le génotype le plus sensible;} \\ 0 & \text{si } g : \text{ARR-ARR, le génotype totalement résistant;} \\ \sigma_g \in [0, 1] & \text{sinon.} \end{cases} \quad (\text{II.4})$$



(a) Individu infecté mort de tremblante : $d_s = d_T$, avec contamination à $d_c < d_M$ et signes cliniques à $d_T \leq d_M$.



(b) Individu infecté abattu pendant la période d'incubation (réforme, autre maladie, etc.) : $d_s = d_M$, avec contamination à $d_c < d_M$ et sans signes cliniques ($d_T > d_M$).



(c) Individu sain : $d_s = d_M$ ($d_c, d_T > d_M$).

FIG. II.2 – Détermination de la date de mort simulée d_s d'un individu, à partir de sa date de mort provenant des données d_M et de sa date de mort tremblante d_T calculée par le modèle et correspondant à l'apparition des signes cliniques. d_N est sa date de naissance (données), d_c sa date de contamination, τ sa durée d'incubation.

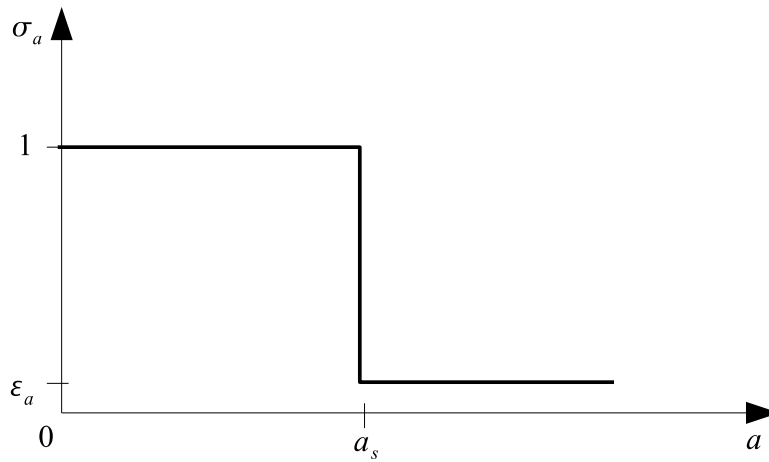


FIG. II.3 – Susceptibilité par âge représentée par une fonction seuil à deux variables : seuil a_s et valeur minimale ε_a .

II.3.2.3 Force d'infection

Notations

- \mathcal{J} représente l'ensemble des individus du troupeau.
- Soit j un individu de \mathcal{J} . Ses caractéristiques sont notées avec un exposant j .
EXEMPLES : sa date de naissance est notée d_N^j et son âge à l'instant t est $a^j(t)$.
- On définit une fonction indicatrice comme suit :

$$\mathbb{1}(x) = \begin{cases} 1 & \text{si la condition } x \text{ est vérifiée,} \\ 0 & \text{sinon.} \end{cases} \quad (\text{II.5})$$

EXEMPLES

- Avec $A^j = [t_N^j, t_s^j[$ l'intervalle de vie de l'individu j , $\sum_{j \in \mathcal{J}} \mathbb{1}(t \in A^j)$ représente la somme des individus vivants à l'instant t , soit l'effectif du troupeau à t .
- Soit g un génotype donné. $\sum_{j \in \mathcal{J}} \mathbb{1}(t \in A^j) \mathbb{1}(g^j = g)$ représente le nombre d'individus de génotype g dans le troupeau à t .

La force d'infection $f_I(t)$ représente le pouvoir de contamination du troupeau à l'instant t . Elle est définie comme une somme pondérée des animaux infectés présents dans le troupeau à l'instant t :

$$f_I(t) = \sum_{j \in \mathcal{J}} \mathbb{1}(t \in [d_c^j, d_s^j]) \varphi(\theta^j(t)),$$

où $\varphi(\theta^j)$ représente l'infectiosité de l'individu j à l'instant t , qui dépend de sa charge $\theta^j(t)$ (§II.3.3). Nous considérons deux fonctions d'infectiosité :

1. $\varphi = 1$, une infectiosité constante indépendante du degré d'incubation – la force d'infection est alors égale au nombre d'infectés présents dans le troupeau à t ;
2. $\varphi(\theta) = \theta$, une infectiosité croissante pendant la période d'incubation – la force d'infection est alors égale à la somme des charges des infectés présents dans le troupeau à t .

Comme $\theta \in [0, 1]$, l'infectiosité vérifie aussi $\varphi(\theta) \in [0, 1]$.

La force d'infection normalisée est obtenue en divisant la force d'infection ci-dessus par l'effectif du troupeau, soit

$$\overline{f_I}(t) = \frac{\sum_{j \in \mathcal{J}} \mathbb{1}(t \in [d_c^j, d_s^j]) \varphi(\theta^j(t))}{\sum_{j \in \mathcal{J}} \mathbb{1}(t \in [d_N^j, d_s^j])}. \quad (\text{II.6})$$

Comme $\varphi \in [0, 1]$, la force d'infection normalisée vérifie aussi $\overline{f_I}(t) \in [0, 1]$.

II.3.2.4 Risque de contamination

Le risque de contamination $r_c(t)$ d'un individu sain et vivant à l'instant t est supposé proportionnel à (i) sa susceptibilité génétique σ_g définie en (II.4) ; (ii) sa susceptibilité par âge $\sigma_{a(t)}$ définie en (II.3), dépendant de son âge $a(t)$; (iii) la force d'infection normalisée du troupeau (II.6). En outre, selon les hypothèses formulées dans §II.3.1.2, ce risque est supposé nul en dehors des saisons de mise bas, définies par (II.1). D'où

$$r_c(t) \propto \sigma_{a(t)} \sigma_g \overline{f_I}(t) s(t).$$

Grâce aux diverses normalisations, chacun des termes de droite est compris entre 0 et 1 ; leur produit est donc lui aussi compris entre 0 et 1.

Comme le risque de contamination est une probabilité, on pourrait transformer la proportionnalité ci-dessus en égalité. Néanmoins, le risque ainsi défini pourrait s'avérer trop faible, en particulier en début d'épizootie. C'est pourquoi nous introduisons un coefficient de transmission $\alpha > 0$. Pour garantir que le risque de contamination demeure inférieur à 1, on normalise ce coefficient par

$$\bar{\alpha}_t = \min\{\alpha, \sigma_{a(t)} \sigma_g \bar{f}_I(t) s(t)\}.$$

Le risque de contamination d'un individu à l'instant t est alors

$$r_c(t) = \bar{\alpha}_t \sigma_{a(t)} \sigma_g \bar{f}_I(t) s(t). \quad (\text{II.7})$$

On a bien $r_c(t) = \min\{1, \alpha \sigma_{a(t)} \sigma_g \bar{f}_I(t) s(t)\}$.

Le risque de contamination permet de déterminer si l'individu s'infecte ou non à t , par le biais d'un tirage aléatoire selon une loi de Bernoulli $\mathcal{B}(r_c(t))$ de paramètre ce risque.

II.3.2.5 Implémentation

À tout instant t , on calcule la force d'infection $\bar{f}_I(t)$ selon (II.6). Si elle est non nulle, pour tout individu sain et vivant à l'instant t , on réalise les opérations suivantes.

1. Calcul du risque d'infection $r_c(t)$ selon (II.7)
2. Tirage aléatoire selon une loi de Bernoulli de paramètre $r_c(t)$
 \Leftrightarrow tirage selon une loi uniforme $\mathcal{U}([0, 1])$ d'une réalisation η .
 - Si $\eta \leq r_c(t)$, contamination de l'individu à t :
 - (a) mise à jour de sa date de contamination $d_c = t$;
 - (b) attribution d'une charge initiale θ_0 selon (II.8) ;
 - (c) mise à jour de sa charge $\theta(t) = \theta_0$;
 - (d) calcul de sa durée d'incubation τ selon (II.11) et mise à jour de sa date de mort tremblante $d_T = t + \tau$;
 - (e) mise à jour de sa date de mort simulée $d_s = \min(d_M, d_T)$. selon (II.2).
 - Sinon, les paramètres de l'individu j restent inchangés.

II.3.3 Incubation

L'incubation est la période qui sépare la date de contamination de l'apparition des signes cliniques chez un animal infecté (FIG. II.2). C'est une longue période silencieuse pendant laquelle l'animal est infectieux, mais non détectable. Sa durée dépend du génotype de l'animal et est variable entre individus (§II.3.1.2). Pour représenter ce processus, on attribue à chaque individu infecté une charge d'infection qui croît au cours du temps.

II.3.3.1 Charge initiale d'infection

Chaque individu nouvellement infecté se voit attribuer une charge initiale d'infection $\theta_0 \in [0, 1]$ à sa date de contamination d_c . Afin d'obtenir des durées d'incubation variables entre individus, cette charge initiale est tirée aléatoirement selon une loi bêta $\mathcal{B}_0(p, q)$ de paramètres $p, q > 0$. La charge d'un individu lors de sa contamination est donc

$$\theta(d_c) = \theta_0 \sim \mathcal{B}_0(p, q). \quad (\text{II.8})$$

La densité de probabilité associée à cette loi bêta $\mathcal{B}_0(p, q)$ est

$$\Theta(\theta_0) = \begin{cases} \frac{1}{\beta(p, q)} \theta_0^{(p-1)} (1 - \theta_0)^{(q-1)} & \text{si } x \in [0, 1], \\ 0 & \text{sinon} \end{cases}$$

Sa moyenne m et sa variance v s'expriment en fonction de ses paramètres de la manière suivante :

$$\begin{cases} m = \frac{p}{p+q} \\ v = \frac{pq}{(p+q)^2(p+q+1)} \end{cases} \quad (\text{II.9})$$

On note que v doit vérifier la condition $v < m(1 - m)$.

II.3.3.2 Évolution de la charge

Durant la période d'incubation, la charge θ d'un individu infecté croît de manière exponentielle, le taux de croissance c_g dépendant de son génotype g :

$$\frac{d\theta}{dt} = c_g \theta \quad \text{avec } \theta(d_c) = \theta_0.$$

Pour tout $t \in [d_c, d_s]$, on a donc

$$\theta(t) = \theta_0 e^{c_g(t-d_c)}. \quad (\text{II.10})$$

II.3.3.3 Mortalité tremblante

La fin de la période d'incubation survient *par convention* lorsque la charge de l'individu infecté atteint la valeur 1. Cela correspond à l'apparition des signes cliniques, *i.e.* à sa date de mort tremblante d_T . D'où $\theta(d_T) = 1$. Sa durée d'incubation est donc

$$\tau = d_T - d_c = -\frac{\ln \theta_0}{c_g}. \quad (\text{II.11})$$

La durée d'incubation dépend bien du génotype g et varie entre individus, car la charge initiale (II.8) est issue d'un tirage aléatoire. Θ étant la densité de probabilité de la charge initiale, on en déduit la densité de probabilité Ψ de la durée de l'incubation :

$$\Psi(\tau) = c_g e^{-c_g \tau} \Theta(e^{-c_g \tau}).$$

N.B. : Il est possible qu'un individu infecté meure avant la fin de sa période d'incubation ($d_s = d_M < d_T$, cf. FIG. II.2(b)), auquel cas sa charge n'atteint pas la valeur 1.

II.4 La condition initiale

La simulation du modèle débute à l'instant initial noté t_0 . Il est nécessaire d'initialiser les caractéristiques de tous les individus à cet instant (§II.2). En particulier, il faut déterminer les infectés initiaux, qui seront responsables de la propagation de la tremblante dans le troupeau.

Les dates de naissance d_N , de mort d_M , le génotype g et le diagnostic tremblante h sont des paramètres fixés par les données. Pour que l'âge soit toujours une variable positive, il est initialisé à $a(t_0) = \max(0, t_0 - d_N)$.

Les animaux du troupeau sont considérés comme sains par défaut, donc leurs caractéristiques épidémiologiques sont initialisées comme suit :

- date de contamination $d_c = +\infty$,
- date de mort tremblante $d_T = +\infty$,
- charge d'infection $\theta(t_0) = 0$.

Les animaux infectés ne peuvent être distingués des animaux sains dans le troupeau, on doit donc déterminer les infectés initiaux de manière indirecte à partir des données. Les critères de sélection retenus pour qu'un animal soit supposé infecté à t_0 sont :

1. individu vivant à t_0 , soit $d_N \leq t_0 < d_M$;
2. individu mort de tremblante, soit $h = 1$ (d'où $d_T = d_M$) ;
3. individu déjà infecté à t_0 , que l'on peut formuler de manières équivalentes par :
 - $d_M - t_0 \leq \hat{\tau}$, durée d'incubation « raisonnable » ;
 - $\theta(t_0) = e^{c_g(t_0 - d_M)} \geq \hat{\theta}_0$, charge d'infection initiale « raisonnable » ;
 - $d_M + \log(\hat{\theta}_0)/c_g \leq t_0$.

On choisit $\hat{\theta}_0 = m$, la moyenne de la charge d'infection initiale donnée par (II.9). Les animaux vérifiant ces conditions sont les infectés initiaux du modèle. Leurs caractéristiques épidémiologiques sont initialisées comme suit :

- date de contamination $d_c = d_M + \log(m)/c_g$ ($d_c \leq t_0$),
- date de mort tremblante $d_T = d_M$,
- charge d'infection $\theta(t_0) = e^{c_g(t_0 - d_M)}$.

II.5 Les sorties

Les sorties du modèle sont calculées à chaque instant de simulation $t \in \{t_0, t_0 + \Delta t, \dots, t_f\}$. Les sorties associées à des flux prennent en t la valeur des flux intégrés sur l'intervalle de temps $[t, t + \Delta t[$, où Δt représente le pas de temps de la simulation.

Notations

- On note avec un indice d les sorties issues des données et sans indice les sorties simulées par le modèle.
- L'exposant j est utilisé pour les caractéristiques de l'individu $j \in \mathcal{J}$, \mathcal{J} représentant l'ensemble des individus du troupeau.
- On utilise ici aussi la fonction indicatrice $\mathbb{1}$ définie en (II.5).

II.5.1 Sorties démographiques

Effectif du troupeau Nombre d'animaux vivants à t :

$$N_d(t) = \sum_{j \in \mathcal{J}} \mathbb{1}(t \in [d_N^j, d_M^j]), \quad (\text{II.12})$$

$$N(t) = \sum_{j \in \mathcal{J}} \mathbb{1}(t \in [d_N^j, d_s^j]). \quad (\text{II.13})$$

Effectif à t par génotype g et classe d'âge Δ_a :

$$N(t, g, \Delta_a) = \sum_{j \in \mathcal{J}} \mathbb{1}(t \in [d_N^j, d_s^j]) \mathbb{1}(g^j = g) \mathbb{1}(t - d_N^j \in \Delta_a). \quad (\text{II.14})$$

Fréquences génotypiques Proportion d'animaux de génotype $g = 1, \dots, n_g$ à t :

$$f_{gd}(t) = \frac{\sum_{j \in \mathcal{J}} \mathbb{1}(g^j = g) \mathbb{1}(t \in [d_N^j, d_M^j])}{N_d(t)}, \quad (\text{II.15})$$

$$f_{gs}(t) = \frac{\sum_{j \in \mathcal{J}} \mathbb{1}(g^j = g) \mathbb{1}(t \in [d_N^j, d_s^j])}{N(t)}. \quad (\text{II.16})$$

II.5.2 Sorties épidémiologiques

Dans le troupeau, les seules observations en lien avec la maladie sont les cas de tremblantes. Donc peu de sorties épidémiologiques peuvent être estimées à partir des données.

Incidence Nombre de cas de tremblante déclarés sur l'intervalle de temps $[t, t + \Delta t[$:

$$C_d(t) = \sum_{j \in \mathcal{J}} \mathbb{1}(h^j = 1) \mathbb{1}(d_M^j \in [t, t + \Delta t]), \quad (\text{II.17})$$

$$C(t) = \sum_{j \in \mathcal{J}} \mathbb{1}(d_T^j = d_s^j) \mathbb{1}(d_T^j \in [t, t + \Delta t]) \quad (\text{II.18})$$

Incidence cumulée Idem sur $[t_0, t + \Delta t[$:

$$\tilde{C}_d(t) = \sum_{j \in \mathcal{J}} \mathbb{1}(h^j = 1) \mathbb{1}(d_M^j \in [t_0, t + \Delta t]), \quad (\text{II.19})$$

$$\tilde{C}(t) = \sum_{j \in \mathcal{J}} \mathbb{1}(d_T^j = d_s^j) \mathbb{1}(d_T^j \in [t_0, t + \Delta t]) \quad (\text{II.20})$$

Incidence cumulée par génotype g et classe d'âge Δ_a :

$$\tilde{C}(t, g, \Delta_a) = \sum_{j \in \mathcal{J}} \mathbb{1}(d_T^j = d_s^j) \mathbb{1}(d_T^j \in [t_0, t + \Delta t]) \mathbb{1}(g^j = g) \mathbb{1}(t - d_N^j \in \Delta_a). \quad (\text{II.21})$$

Les sorties suivantes ne peuvent pas être estimées à partir des données.

Infectés Nombre d'infectés à t :

$$I(t) = \sum_{j \in \mathcal{J}} \mathbb{1}(t \in [d_c^j, d_s^j]). \quad (\text{II.22})$$

Prévalence Proportion d'individus infectés à t :

$$P(t) = \frac{I_s(t)}{N_s(t)}. \quad (\text{II.23})$$

Nouveaux infectés Nombre d'individus contaminés sur $[t, t + \Delta t[$:

$$NI(t) = \sum_{j \in \mathcal{J}} \mathbb{1}(d_c^j \in [t, t + \Delta t[). \quad (\text{II.24})$$

Nouveaux infectés par génotype g et classe d'âge Δ_a :

$$NI(t, g, \Delta_a) = \sum_{j \in \mathcal{J}} \mathbb{1}(d_c^j \in [t, t + \Delta t[) \mathbb{1}(g^j = g) \mathbb{1}(t - d_N^j \in \Delta_a). \quad (\text{II.25})$$

Nouveaux infectés cumulés Nombre d'individus contaminés avant $t + \Delta t$:

$$\tilde{NI}(t) = \sum_{j \in \mathcal{J}} \mathbb{1}(d_c^j < t + \Delta t). \quad (\text{II.26})$$

Nouveaux infectés cumulés par génotype g et classe d'âge Δ_a :

$$\tilde{NI}(t, g, \Delta_a) = \sum_{j \in \mathcal{J}} \mathbb{1}(d_c^j < t + \Delta t) \mathbb{1}(g^j = g) \mathbb{1}(t - d_N^j \in \Delta_a). \quad (\text{II.27})$$

Âge de contamination moyen Calculé pour les individus contaminés avant $t + \Delta t$:

$$\tilde{A}(t) = \frac{\sum_{j \in \mathcal{J}} (d_c^j - d_N^j) \mathbb{1}(d_c^j < t + \Delta t)}{\sum_{j \in \mathcal{J}} \mathbb{1}(d_c^j < t + \Delta t)} \quad (\text{II.28})$$

Distribution de l'âge de contamination Proportion d'individus contaminés à l'âge $d_c - d_N \in \Delta_a$ (classe d'âge) avant $t + \Delta t$:

$$A(t, a_i) = \frac{\sum_{j \in \mathcal{J}} \mathbb{1}(d_c^j - d_N^j \in \Delta_a) \mathbb{1}(d_c^j < t + \Delta t)}{\sum_{j \in \mathcal{J}} \mathbb{1}(d_c^j < t + \Delta t)} \quad (\text{II.29})$$

Durée d'incubation moyenne Calculée pour les cas de tremblante sur $[t_0, t + \Delta t[$:

$$\tilde{B}(t) = \frac{\sum_{j \in \mathcal{J}} (d_T^j - d_c^j) \mathbb{1}(d_s^j = d_T^j) \mathbb{1}(d_T^j \in [t_0, t + \Delta t[)}{\sum_{j \in \mathcal{J}} \mathbb{1}(d_s^j = d_T^j) \mathbb{1}(d_T^j \in [t_0, t + \Delta t[)} \quad (\text{II.30})$$

Distribution de la durée d'incubation Proportion de cas de tremblante ayant incubé pendant $\tau = d_T - d_c \in \Delta_i$ (classe de durée d'incubation) sur $[t_0, t + \Delta t[$:

$$B(t) = \frac{\sum_{j \in \mathcal{J}} \mathbb{1}(d_T^j - d_c^j \in \Delta_i) \mathbb{1}(d_s^j = d_T^j) \mathbb{1}(d_T^j \in [t_0, t + \Delta t[)}{\sum_{j \in \mathcal{J}} \mathbb{1}(d_s^j = d_T^j) \mathbb{1}(d_T^j \in [t_0, t + \Delta t[)} \quad (\text{II.31})$$

Chapitre III

Description du simulateur

Dans ce chapitre, nous décrivons l'outil de simulation, qui a été développé dans le langage orienté-objet Java [7] sous la plateforme de programmation Eclipse. Il est fondé sur un simulateur développé par Jean-Pierre TREUIL [16, 15]. Nous exposons successivement l'architecture du programme, son interface graphique, et le fonctionnement du simulateur.

III.1 Architecture des classes du programme

Les langages orientés-objet reposent sur la notion de classe. Une classe est constituée d'un ensemble d'attributs et d'un ensemble de méthodes ou de règles. L'objet est une instance de classe. Pour simplifier la programmation, nous avons découpé le programme en un nombre relativement important de méthodes élémentaires, simples à comprendre et à résoudre. Ce programme contient deux classes fondamentales relatives à la modélisation, **Troupeau** et **Mouton**. Il comporte aussi plusieurs classes utilitaires ou spécifiques à l'interface, en particulier la classe **Myframe** qui s'occupe de l'affichage et la classe **MyModele** qui effectue l'enchaînement des opérations. Nous allons décrire plus en détails ces quatre classes principales.

III.1.1 La classe Mouton

Cette classe est consacrée à la description d'un individu : un mouton. Elle lui attribue des caractéristiques fournies par les données ou modélisées par le simulateur (§II.2). Elle décrit l'état d'un individu à travers différentes méthodes. TAB. III.1 décrit les attributs et méthodes de la classe **Mouton**.

III.1.2 La classe Troupeau

Cette classe décrit le troupeau, défini comme une liste de moutons. Elle contient aussi la date courante et a pour caractéristiques les sorties du modèle détaillées dans §II.5. Elle implémente les processus épidémiologiques au niveau du troupeau, fait évoluer les caractéristiques temporelles des individus, calcule les sorties et le critère d'estimation présenté en §IV.3. TAB. III.2 décrit les attributs et méthodes de la classe **Troupeau**.

Classe Mouton
Attributs : <ul style="list-style-type: none"> – identifiant, – sexe, – génotype g, – date de naissance d_N, – date de mort d_M, – diagnostic tremblante h, – âge a, – charge d’infection θ, – date de contamination d_c, – date de mort tremblante d_T.
Méthodes : <ul style="list-style-type: none"> – incrémentation de l’âge, – test si l’animal est vivant dans les données, – test si l’animal est vivant dans le modèle, – test si l’animal est sain, – incrémentation de la charge d’infection θ, – calcul de la charge c_g, – calcul de la susceptibilité par âge σ_a, – calcul de la susceptibilité génétique σ_g.

TAB. III.1 – Attributs et méthodes de la classe **Mouton** du simulateur, une des deux classes fondamentales avec **Troupeau**, en lien avec la modélisation.

Classe Troupeau
Attributs : <ul style="list-style-type: none"> – temps t, – liste des moutons, – sorties du modèle.
Méthodes : <ul style="list-style-type: none"> – progression de l’âge a, – progression de la charge θ, – implémentation de la contamination, – calcul des sorties, – calcul du critère J.

TAB. III.2 – Attributs et méthodes de la classe **Troupeau** du simulateur, une des deux classes fondamentales avec **Mouton**, en lien avec la modélisation.

III.1.3 La classe MyModele

Cette classe a pour fonction de rassembler les différents objets du modèle créés par les classes Mouton et Troupeau et d'enchaîner les différentes opérations du modèle. Ainsi, les paramètres d'entrée et les sorties sont aussi déclarés dans cette classe. Les opérations concernant l'initialisation et l'exécution du modèle apparaissent dans cette classe.

III.1.4 La classe MyFrame

Cette classe est en charge de l'affichage au niveau de l'interface. Elle permet de fixer certains paramètres d'entrée du modèle, modifiables ou non en cours de simulation. Elle représente de manière graphique les différentes sorties du modèle. Les interactions entre les classes MyModele et MyFrame sont représentées dans la FIG. III.1

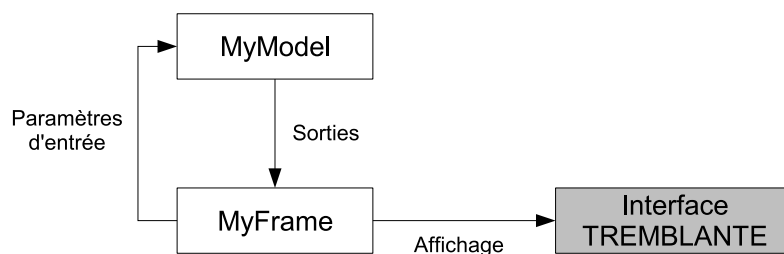


FIG. III.1 – Schéma mettant en relation les classes MyModel et MyFrame du simulateur : MyModel est en charge de l'initialisation et l'exécution du modèle ; MyFrame gère l'affichage sur l'interface graphique.

III.2 Interface graphique

L'interface graphique du simulateur permet de régler certains paramètres d'entrée du modèle et d'afficher ses sorties. Elle comporte aussi des boutons de contrôle. Une capture d'écran est montrée sur la FIG. III.2.

III.2.1 Paramètres d'entrée

Les entrées comportent deux types de paramètres : les paramètres fixes dont les valeurs sont fixées en début de simulation et les paramètres ajustables qui peuvent être modifiés au niveau de l'interface pendant la simulation.

III.2.1.1 Paramètres fixes

Ils apparaissent sur des boutons de réglage situés en haut à gauche de l'interface sur la première colonne (FIG. III.2, fenêtre 2).

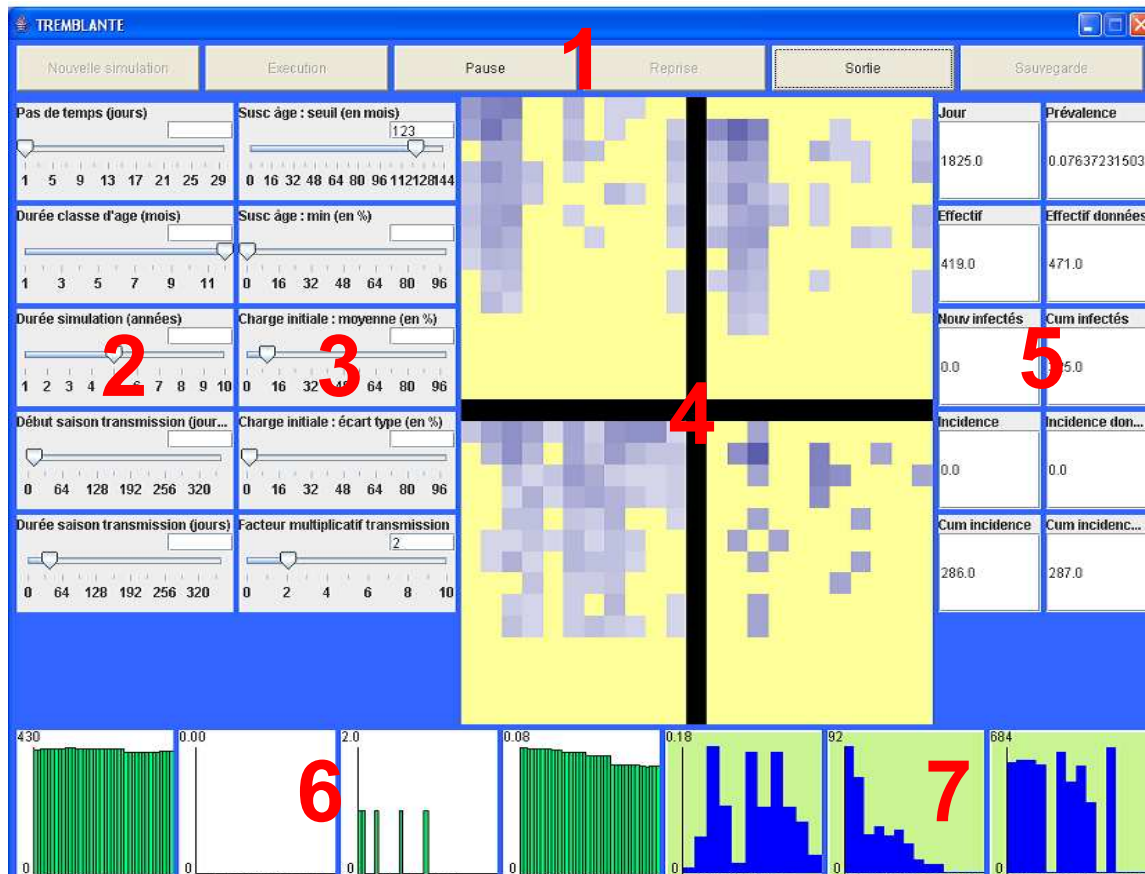


FIG. III.2 – Interface graphique du simulateur. Sept fenêtres sont identifiées (gros chiffres rouges), correspondant

- aux boutons de contrôle (1) ;
- aux paramètres d'entrée fixes (2) ou ajustables (3) ;
- aux différentes sorties : indicateurs 2D (4), indicateurs quantitatifs (5), indicateurs dynamiques (6) et histogrammes (7).

1. Pas de temps Δt .
2. Durée d'une classe d'âge Δ_a pour les sorties graphiques 2D (II.14,II.25,II.27,II.21).
3. Durée de la simulation $D = t_f - t_0$.
- 4-5. Paramètres (2) de la fonction saison $s(t)$ (II.1).

III.2.1.2 Paramètres ajustables

Ils apparaissent sur des boutons de réglage situés en haut à gauche de l'interface, sur la deuxième colonne (FIG. III.2, fenêtre 3).

- 1-2. Paramètres a_s et ε_a de la susceptibilité par âge (II.3).
- 3-4. Paramètres m et \sqrt{v} de la charge initiale d'infection (II.8,II.9).
5. Coefficient de transmission α (II.7).

III.2.2 Sorties

Il existe deux types de sorties, qui apparaissent sur l'interface du simulateur : les indicateurs quantitatifs et les indicateurs graphiques. Sauf mention contraire, les sorties sont des sorties de simulation et non des données.

III.2.2.1 Indicateurs quantitatifs

Ils sont situés en haut à droite de l'interface, sur deux colonnes (FIG. III.2, fenêtre 5). Un indicateur quantitatif donne simplement la valeur numérique de la sortie correspondante.

1. Temps t .
2. Prévalence (II.23).
3. Effectif (II.13).
4. Effectif données (II.12).
5. Nouveaux infectés (II.24).
6. Nouveaux infectés cumulés (II.26).
7. Incidence (II.18).
8. Incidence données (II.17).
9. Incidence cumulée (II.20).
10. Incidence cumulée données (II.19).

III.2.2.2 Indicateurs graphiques

On dispose de trois types d'indicateurs graphiques.

Indicateurs dynamiques Ils sont placés sur une ligne en bas à gauche de l'interface (FIG. III.2, fenêtre 6). Un indicateur dynamique représente, dans une fenêtre glissante avec le temps en abscisse, les valeurs instantanées de la sortie en ordonnée.

1. Effectif (II.13).
2. Nouveaux infectés (II.24).
3. Incidence (II.18).
4. Prévalence (II.23).

Histogrammes Ils sont placés sur une ligne en bas à droite de l'interface (FIG. III.2, fenêtre 7). Les histogrammes sont actualisés à chaque pas de temps.

1. Fréquence génotypiques (II.16).
2. Distribution de l'âge de contamination (II.29).
3. Distribution de la durée d'incubation (II.31).

Indicateurs 2D Il sont placés en carré au centre de l'interface (FIG. III.2, fenêtre 4). Un indicateur 2D représente par un gradient de couleur les valeurs instantanées de la sortie, selon les génotypes en abscisse et les classes d'âge en ordonnée. Plus le bleu d'un petit carré élémentaire 1 génotype \times 1 classe d'âge est foncé, plus la variable de sortie correspondante est élevée.

1. Effectif (II.14) (carré en bas à gauche).
2. Nouveaux infectés (II.25) (carré en bas à droite).
3. Nouveaux infectés cumulés (II.27) (carré en haut à gauche).
4. Incidence cumulée (II.21) (carré en haut à droite).

III.2.3 Boutons de contrôle

Ils sont situés en ligne en haut de l'interface graphique (FIG. III.2, fenêtre 1).

1. Nouvelle simulation : pour initialiser une simulation à partir des paramètres inscrits au tableau de commandes et du fichier de données.
2. Exécution : pour lancer l'exécution d'une simulation préalablement initialisée.
3. Pause : pour interrompre l'exécution du programme (la pause permet de figer l'interface).
4. Reprise : pour reprendre l'exécution momentanément interrompue.
5. Sortie : pour sortir du programme.

III.3 Fonctionnement du simulateur

Nous décrivons tout d'abord le lancement d'une simulation. Après compilation du programme, l'interface graphique du simulateur (FIG. III.2) est générée.

1. La première étape consiste à lancer une **Nouvelle simulation** depuis la fenêtre 1.
2. On fixe ensuite les paramètres d'entrée fixes et ajustables dans les fenêtres 2 et 3.
3. On lance ensuite l'**Exécution** du programme depuis la fenêtre 1. En arrivant à cette étape, le simulateur enchaîne des opérations, comme cela est décrit dans le TAB. **III.3**. Les fenêtres des sorties sont mises à jour.
4. À tout moment, on peut changer les valeurs des paramètres d'entrée ajustables. On peut aussi faire une **Pause** dans l'exécution du programme.

Pour une durée de simulation $D = N\Delta t$

I Initialisation

1. initialisation à partir des données de
 - (a) caractéristiques d'un mouton
 - (b) condition initiale
2. calcul de sorties
3. calcul du critère
4. affichage

II Opérations pendant un pas de temps Δt de la vie du troupeau

1. incrémentation du temps $t = t + \Delta t$
2. pour tous les animaux vivants à t
 - (a) incrémentation de l'âge $a = a + \Delta t$
3. pour tous les animaux vivants et infectés à t , outre l'incrément de l'âge
 - (a) incrémentation de la charge $\theta = \theta e^{c_g \Delta t}$
4. pour tous les animaux vivants et sains à t , en plus de l'incrément de l'âge
 - (a) incrémentation de l'âge $a = a + \Delta t$
 - (b) implémentation de la contamination (§II.3.2)
 - i. calcul de la force d'infection f_I
 - ii. calcul du risque de contamination r_c
 - iii. tirage de la charge initiale θ_0
 - iv. mise à jour de
 - A. date de contamination d_c
 - B. date de mort tremblante d_T
5. mise à jour des sorties
6. mise à jour du critère
7. mise à jour des paramètres ajustables
8. affichage

III répéter les opérations du paragraphe II ($N - 1$) fois

TAB. III.3 – Enchaînement des opérations du simulateur.

Chapitre IV

Estimation des paramètres

Pour simuler le modèle, on a besoin des données de suivi de troupeau afin d’initialiser les caractéristiques des individus (§II.4). Il faut aussi estimer les paramètres des processus épidémiologiques décrits dans les sections contamination §II.3.2 et incubation §II.3.3. Pour cela, on minimise un critère d’ajustement du modèle aux données. Dans ce chapitre, après avoir décrit les données et les paramètres à estimer, on présente plusieurs critères d’estimation. Puis on introduit la méthode d’optimisation retenue, un algorithme de recherche aléatoire. Enfin, on termine avec quelques résultats d’estimation.

IV.1 Les données

Les données que nous utilisons dans cette étude sont issues du troupeau expérimental de Langlade (SAGA, INRA Toulouse, France), dans lequel une épizootie de tremblante naturelle s’est déclarée en 1993 [6] (FIG. IV.1). Le troupeau est constitué de 600 à 800 brebis de renouvellement de race Romanov et est fermé (pas d’entrées d’animaux extérieurs) sur la période d’étude considérée [1992,1999].

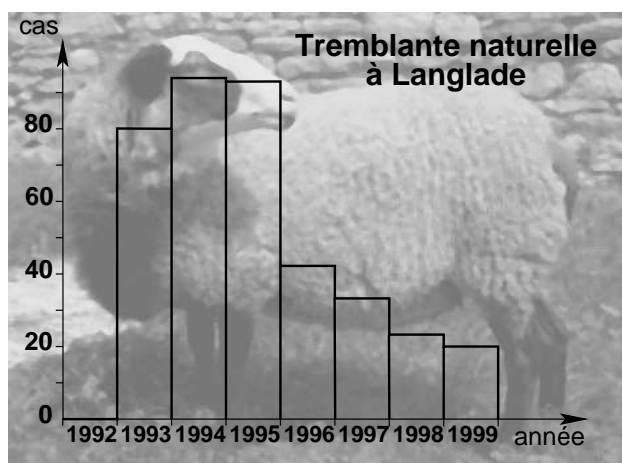


FIG. IV.1 – Cas de tremblante naturelle dans le troupeau expérimental de Langlade (SAGA, INRA Toulouse, France).

Pour tous les animaux des cohortes 1983 à 1999, on dispose des données suivantes :

- identifiant,
- sexe,
- génotype PrP,
- date de naissance,
- date de mort,
- diagnostic tremblante *post-mortem*.

On observe aussi les saisons de mise bas correspondant à la fonction saison (II.1).

Génotype PrP Le génotype PrP est défini à partir des codons 136, 154 et 171 du gène codant pour la protéine prion. 4 allèles ont été détectés dans les données : VRQ, ARQ, AHQ et ARR. VRQ est l'allèle de plus grande sensibilité à la tremblante, ARR de plus grande résistance. ARQ est l'allèle le plus fréquent, AHQ est assez rare. Ces 4 allèles engendrent 10 génotypes, dont les résistances à la tremblante sont données par le TAB. IV.1.

Ultra sensible	Très sensibles	Sensibles	Résistants	Vrai résistant
VRQ-VRQ (1)	VRQ-ARQ (2) ARQ-ARQ (3)	VRQ-AHQ (4) VRQ-ARR (5) ARQ-AHQ (6)	ARQ-ARR (7) AHQ-AHQ (8) AHQ-ARR (9)	ARR-ARR (10)

TAB. IV.1 – Classification des génotypes selon leur résistance à la tremblante classique. Correspondance entre génotype défini par ses allèles XXX-XXX et numéro de génotype (g).

La majorité des animaux nés après 1993 sont génotypés. Quand la donnée est manquante ou que le génotypage est incomplet pour un animal, on lui attribue le génotype « manquant » $g = 11$.

Diagnostic tremblante Les animaux ne sont pas tous testés et différents tests *post-mortem* ont été utilisés. Par souci de simplification, on suppose que le diagnostic est positif si l'animal est testé positif à la tremblante. Dans tous les autres cas, que le test soit négatif, incertain, ou qu'il n'ait pas été réalisé, le diagnostic est dit négatif.

IV.2 Les paramètres à estimer

Les paramètres à estimer sont ceux intervenant dans les processus épidémiologiques décrits dans les sections contamination §II.3.2 et incubation §II.3.3. Leur liste est donnée ci-dessous :

- 2 paramètres ε_a et a_s pour la susceptibilité par âge σ_a définie par (II.3) ;
- n_g susceptibilités génétiques relatives σ_g définies par (II.4) ;
- 2 paramètres v et m pour la charge initiale θ_0 définie par (II.8,II.9) ;
- n_g taux de croissance de la charge c_g définis par (II.10) ;
- 1 coefficient de transmission α pour le risque de contamination défini par (II.7).

On obtient donc $(2n_g + 5)$ paramètres à estimer. Avec 10 génotypes identifiés + 1 génotype inconnu dans les données de Langlade, cela correspond à 27 paramètres. Ce nombre est important et on cherche à le réduire, en agrégeant ou en fixant certains paramètres.

Susceptibilités génétiques relatives On a vu dans §II.3.2.2 que par définition, la susceptibilité génétique relative du génotype le plus sensible VRQ-VRQ est fixée à 1. Celle du génotype vrai résistant ARR-ARR est nulle. En outre, on utilise le TAB. IV.1 pour réduire encore ce nombre, en regroupant les génotypes intermédiaires « sensibles » et « résistants », ainsi que le génotype inconnu. On conserve ainsi 5 niveaux de susceptibilité génétique triés par ordre de résistance croissante :

- VRQ-VRQ : $\sigma_1 = 1$,
- VRQ-ARQ : σ_2 ,
- ARQ-ARQ : σ_3 ,
- $g = 4, \dots, 9, 11$: $\sigma_g = \sigma_4$,
- ARR-ARR : $\sigma_{10} = 0$.

On obtient ainsi 3 paramètres à estimer : σ_2 , σ_3 et σ_4 .

Taux de croissance de la charge Les taux de croissance de la charge déterminent les durées d'incubation en fonction du génotype (II.11). Comme le génotype ARR-ARR ne s'infecte jamais, on fixe $c_{10} = 0$. Comme précédemment, on regroupe les génotypes intermédiaires avec le génotype inconnu. On obtient ainsi 4 paramètres à estimer : c_1 , c_2 , c_3 et c_4 , avec $c_g = c_4$ pour $g = 4, \dots, 9, 11$.

Pour faciliter l'interprétation des paramètres, on peut reparamétriser le modèle en remplaçant les c_g par des durées d'incubation

$$\tau_g = -\frac{\ln m}{c_g},$$

où m est la moyenne de la charge initiale d'infection.

Charge initiale d'infection On peut aussi choisir de fixer arbitrairement m dans (II.9). La durée d'incubation dépend alors de 2 paramètres, v et c_g , la charge initiale du seul paramètre v .

Le nombre total des paramètres à estimer a diminué, par ces simplifications, de 27 à 11 paramètres. Les paramètres à estimer, leur nombre et les simplifications choisies sont récapitulés dans le TAB. IV.2.

<i>Paramètres</i>	Nombre initial	<i>Simplification</i>	Nombre simplif.
susceptibilité génétique σ_g	11	$\sigma_1 = 1, \sigma_{10} = 0, \sigma_{4,\dots,9,11} = \sigma_4$	3
susceptibilité par âge ε_a, a_s	2	–	2
charge initiale m, v	2	m fixé	1
durée d'incubation τ_g	11	$\tau_{4,\dots,9,11} = \tau_4$	4
coeff. de transmission α	1	–	1
Total	27	Total	11

TAB. IV.2 – Paramètres à estimer dans le modèle avant et après simplification.

IV.3 Le critère

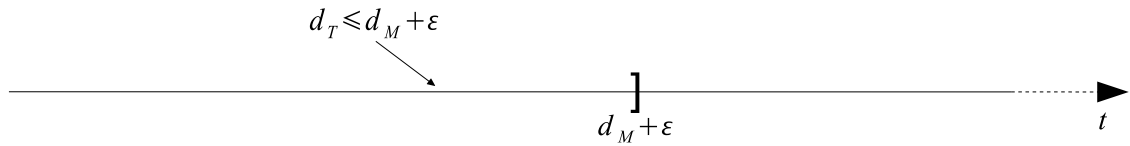
Le critère représente une distance entre les données et les simulations. En minimisant ce critère par rapport aux paramètres épidémiologiques du modèle, on peut estimer leur valeur.

Rappelons que les infectés ne peuvent être distingués des sains. Les seules grandeurs épidémiologiques observables dans un troupeau sont les cas de tremblante cliniques, ou plus précisément les infectés détectables par un test *post-mortem*. La sensibilité des tests permet généralement aussi de détecter la tremblante chez les infectés en fin d'incubation, quelques mois avant l'apparition des signes cliniques. On note ε la durée de la période pendant laquelle les tests sont sensibles ; ε est de l'ordre de 6 mois. On suppose que *dans les données, ces cas cliniques ou en fin d'incubation sont bien détectés* et correspondent aux diagnostics positifs.

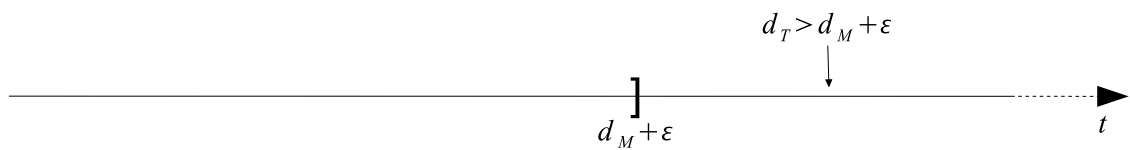
Dans le modèle, un animal infecté est dit *détectable* s'il développe des signes cliniques, *i.e.* $d_s = d_T$, ou s'il est abattu vers la fin de sa période d'incubation, soit $d_T - d_s \leq \varepsilon$. Comme $d_s = \min(d_M, d_T)$, ces deux conditions peuvent se reformuler ainsi :

$$d_T \leq d_M + \varepsilon.$$

Cette condition est représentée dans la FIG. IV.2(a).



(a) Cas de tremblante détectable par le modèle : mort moins de ε avant la date l'apparition des signes cliniques, *i.e.* la fin de la période d'incubation.



(b) Cas de tremblante non détectable dans le modèle : mort plus de ε après la date l'apparition des signes cliniques, *i.e.* la fin de la période d'incubation.

FIG. IV.2 – Détection d'un cas de tremblante dans le modèle en fonction de sa date de mort d_M issue des données et de sa date de mort tremblante d_T calculée par le modèle. ε correspond à la durée de la période pendant laquelle les tests sont sensibles avant l'apparition des signes cliniques.

On va donc comparer les cas détectés dans les données aux cas détectables dans le modèle. Le critère peut être construit soit en effectuant cette comparaison au niveau individuel, soit au niveau de groupes d'animaux de caractéristiques similaires. Par ailleurs, la comparaison peut porter uniquement sur le « statut tremblante » (détectée ou non dans les données *vs* détectable ou non dans le modèle), ou bien elle peut aussi intégrer les dates de mort (date de mort données *vs* date de mort tremblante calculée par le modèle).

IV.3.1 Niveau individuel

La comparaison données *vs* modèle se fait au niveau des individus. Le statut tremblante d'un individu est, dans les données :

$$stat_d = \begin{cases} 1 & \text{si le cas est détecté, i.e. si } h = 1, \\ 0 & \text{sinon ;} \end{cases}$$

dans le modèle :

$$stat_s = \begin{cases} 1 & \text{si le cas est détectable, i.e. si } d_T \leq d_M + \varepsilon, \\ 0 & \text{sinon.} \end{cases}$$

IV.3.1.1 Statut tremblante uniquement

Le critère correspond au nombre d'animaux pour qui les statuts tremblante données et modèle sont en désaccord :

$$J_i = \sum_{j \in \mathcal{J}} \mathbb{1}(stat_d^j \neq stat_s^j).$$

IV.3.1.2 Statut tremblante et date de mort

Dans ce critère, on pénalise aussi les écarts entre la date de mort données et la date de mort tremblante calculée par le modèle, lorsqu'ils sont supérieurs à un seuil ϵ' donné :

$$J_{id} = (1 - \kappa)J_i + \kappa \sum_{j \in \mathcal{J}} \mathbb{1}(|d_T^j - d_M^j| > \epsilon') \mathbb{1}(stat_d^j \neq stat_s^j).$$

Le coefficient κ sert à donner plus ou moins de poids à la contribution dans le critère aux dates de mort données et modèle non similaires.

IV.3.2 Niveau groupe d'animaux

La comparaison données *vs* modèle se fait à présent au niveau de groupes d'individus similaires. En effet, l'intervention de statuts tremblante entre deux individus de caractéristiques proches ne devrait pas être pénalisée, comme c'est le cas pour les critères de comparaison au niveau individuel. On choisit de regrouper les individus de même génotype et de même cohorte (animaux nés pendant la même période de mise bas). Ces animaux ont les mêmes génotypes et des âges proches. Ils ont donc quasiment le même risque de contamination. En outre, leur charge d'infection croît à la même vitesse, ils ont donc des durées d'incubation similaires.

On va donc comparer les cas détectables dans ces groupes. Il faut encore définir les périodes d'observation sur lesquelles effectuer les comparaisons. Pour cela, on introduit un intervalle de temps d'observation Δu multiple du pas de temps Δt . On note u les temps d'observation définis par $u = t_0, t_0 + \Delta u, \dots, t_f$.

Soit une cohorte k définie comme l'ensemble des animaux étant nés dans l'intervalle $[k_1, k_2[$. Le nombre de cas de tremblante détectés dans les données au temps d'observation u , *i.e.* sur l'intervalle de temps $[u, u + \Delta u[$, pour le groupe de génotype g et de cohorte k est

$$C_d(u, g, k) = \sum_{j \in \mathcal{J}} \mathbb{1}(g^j = g) \mathbb{1}(d_N^j \in [k_1, k_2]) \mathbb{1}(h^j = 1) \mathbb{1}(d_M \in [u, u + \Delta u]).$$

Le nombre de cas de tremblante détectables par le modèle pendant ce même intervalle de temps et dans ce même groupe est

$$C_s(u, g, k) = \sum_{j \in \mathcal{J}} \mathbb{1}(g^j = g) \mathbb{1}(d_N^j \in [k_1, k_2]) \mathbb{1}(d_T^j \in \{[d_s^j, d_M^j + \varepsilon]\}) \mathbb{1}(d_s^j \in [u, u + \Delta u]).$$

On peut définir les quantités ci-dessus de manière cumulée sur tout l'intervalle de simulation en remplaçant l'intervalle d'observation $[u, u + \Delta u[$ par $[t_0, t_f[$. On note respectivement $\tilde{C}_d(g, k)$ et $\tilde{C}_s(g, k)$ le nombre de cas de tremblante cumulés détectés dans les données et détectables par le modèle pour ce même groupe.

IV.3.2.1 Statut tremblante uniquement

On construit un critère des moindres carrés en comparant pour chaque groupe (g, k) le nombre de cas de tremblante cumulés dans les données et dans le modèle :

$$J_g = \sum_g \sum_k (\tilde{C}_d(g, k) - \tilde{C}_s(g, k))^2.$$

IV.3.2.2 Statut tremblante et date de mort

On construit un critère des moindres carrés en comparant pour chaque groupe (g, k) et à chaque instant d'observation u le nombre de cas de tremblante dans les données et dans le modèle :

$$J_{gd} = \sum_u \sum_g \sum_k (C_d(u, g, k) - C_s(u, g, k))^2. \quad (\text{IV.1})$$

Ce critère est retenu pour l'estimation des paramètres. En effet, l'échelle groupe d'animaux est préférable (*cf.* ci-dessus). En outre, ce dernier critère intègre des pénalités sur écarts entre les dates de mort données *vs* modèle, par le biais du temps d'observation u . De ce fait, il est sans doute mieux adapté pour l'estimation des paramètres en lien avec la durée d'incubation.

IV.4 L'algorithme d'optimisation

L'approche individu-centrée retenue pour le modèle ne permet pas d'établir une formulation exploitable du critère en relation avec les paramètres à estimer. En outre, elle nous place dans un cadre stochastique. Par ailleurs, le nombre de paramètres à estimer reste conséquent malgré les simplifications apportées. Dans ce contexte, l'optimisation par recherche aléatoire est une méthode exploratoire bien adaptée à notre problème, en particulier l'algorithme de recherche aléatoire par écart-type adaptatif, ou ARS (Adaptive

Random Search) [17]. Nous avons donc retenu cet algorithme, dont nous présentons le principe ci-dessous et que nous détaillons dans les sections suivantes.

L'algorithme ARS permet d'explorer un espace des paramètres borné

$$\mathbb{P} = \left\{ p = (p_1, \dots, p_{n_p})^\perp \mid p_{i_{\min}} \leq p_i \leq p_{i_{\max}}, i = 1, \dots, n_p \right\}.$$

Soit J le critère à minimiser, l'objectif est donc de trouver le vecteur de paramètres optimal

$$p^* \in \mathbb{P} \text{ tel que } J(p^*) = \min_{p \in \mathbb{P}} J(p).$$

Principe de l'algorithme ARS L'algorithme alterne des phases d'apprentissage et d'exploitation. Durant l'apprentissage, il cherche la meilleure précision avec laquelle se déplacer dans l'espace des paramètres (via un choix d'écart-type). La phase d'exploitation sert à tester plusieurs valeurs de paramètres en utilisant la précision choisie lors de l'apprentissage.

Dans chacune de ces deux phases, l'algorithme teste différentes valeurs de l'espace des paramètres \mathbb{P} , de la façon suivante : à partir du dernier p^* optimal trouvé, les paramètres testés seront $p^* + r$ où r est un vecteur tiré aléatoirement suivant une loi normale $\mathcal{N}(0, \Sigma(\sigma))$. $\Sigma(\sigma) = \text{diag}\{(\sigma_1)^2, \dots, (\sigma_{n_p})^2\}$ est la matrice de variance/covariance.

L'algorithme débute par une phase d'initialisation. Un critère d'arrêt doit être défini pour terminer l'optimisation.

Lest différentes phases de l'algorithme sont schématisées sur la FIG. IV.3.

IV.4.1 Initialisation

1. On part du milieu de l'espace des paramètres \mathbb{P}

$$p^* = (p_1^*, \dots, p_{n_p}^*)^\perp \text{ tel que } p_i^* = \frac{p_{i_{\min}} + p_{i_{\max}}}{2} \text{ pour tout } i = 1, \dots, n_p.$$

2. On sélectionne l'écart-type initial

$$\sigma^0 = (\sigma_1^0, \dots, \sigma_{n_p}^0)^\perp \text{ tel que } \sigma_i^0 = p_{i_{\max}} - p_{i_{\min}} \text{ pour tout } i = 1, \dots, n_p.$$

IV.4.2 Apprentissage

But Obtenir l'écart-type σ^* optimal pour la recherche du minimum.

Méthode Tester différentes valeurs pour l'écart-type σ et garder celui pour lequel on aura obtenu la plus petite valeur du critère.

En pratique On va tester successivement cinq écarts-types différents, tels que $\sigma^e = \sigma^{(e-1)}/10$, pour $e = 1, \dots, 5$. Pour chacun des écarts-types, il s'agit de calculer le critère pour plusieurs valeurs des paramètres p .

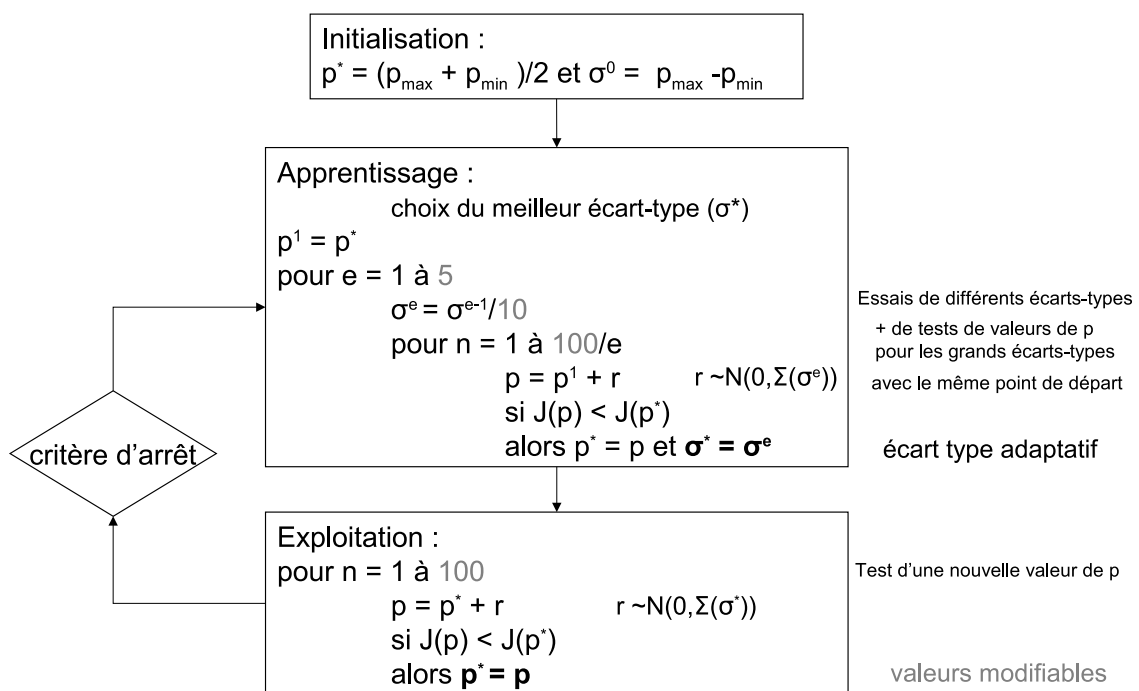


FIG. IV.3 – Schéma de l'algorithme de recherche aléatoire par écart-type adaptatif ARS, mettant en évidence les phases d'initialisation, d'apprentissage et d'exploitation.

1. On stocke la valeur optimale actuelle des paramètres : $p^1 = p^*$.
2. Pour chaque $e = 1, \dots, 5$:
 - (a) On calcule $\sigma^e = \frac{\sigma^{(e-1)}}{10}$.
 - (b) Pour $n = 1, \dots, \frac{100}{e}$ (plus l'écart-type est petit, moins on teste de nouvelles valeurs de p) :
 - i. On fait un tirage aléatoire pour obtenir $r \sim \mathcal{N}(0, \sum(\sigma^e))$.
 - ii. On calcule alors la valeur du critère pour les paramètres $p = p^1 + r$.
 - iii. Si $J(p) < J(p^*)$, alors $\sigma^* = \sigma^e$ et $p^* = p$; *i.e.* si la valeur du critère est meilleure, alors le $e^{\text{ème}}$ écart-type testé actuellement devient l'écart-type optimal pour la recherche du minimum.

N.B. : Les paramètres p résultant des déplacements doivent appartenir à \mathbb{P} .

Dans le TAB. IV.3, nous donnons le nombre de calculs du critère pendant la phase d'apprentissage avec 5 écarts-types testés.

Écart-type σ	σ^1	σ^2	σ^3	σ^4	σ^5
Nombre de calculs pour chaque écart-type	100	50	33	25	20
Nombre total de calculs du critère	228				

TAB. IV.3 – Nombre d'évaluations du critère d'optimisation pendant la phase d'apprentissage de l'algorithme ARS avec 5 écarts-types testés.

IV.4.3 Exploitation

But Obtenir les paramètres p^* optimaux pour notre critère.

Méthode Tester différentes valeurs du vecteur de paramètres p en utilisant la précision de l'écart-type optimal σ^* obtenu lors de la dernière phase d'apprentissage pour se « déplacer » autour du dernier p^* obtenu.

En pratique

1. Pour $n = 1, \dots, 100$
 - (a) On fait un tirage aléatoire pour obtenir $r \sim \mathcal{N}(0, \sum(\sigma^*))$.
 - (b) On calcule alors la valeur du critère pour les paramètres $p = p^* + r$.
 - (c) Si $J(p) < J(p^*)$, alors $p^* = p$ – si la valeur du critère est meilleure, alors les paramètres p utilisés sont actuellement deviennent les paramètres optimaux pour la recherche du minimum.

N.B. : Les paramètres p résultant des déplacements doivent appartenir à \mathbb{P} .

IV.4.4 Paramètres de l'algorithme et critère d'arrêt

Le nombre d'écarts-types testés (ici 5), le nombre d'itérations dans les phases d'apprentissage (ici $228 = \sum_{e=1}^5 \frac{100}{e}$) et d'exploitation (ici 100), ainsi que le rapport entre deux écarts-types successifs (ici 10) sont des paramètres de l'algorithme, et sont donc modifiables en fonction des besoins.

Dans le but de ne pas alterner les phases apprentissage/exploitation indéfiniment, on doit introduire un critère d'arrêt. Pour cela, on fixe :

1. un nombre maximum d'évaluations du critère J en phases d'apprentissage et d'exploitation ;
2. un nombre maximum d'utilisations consécutives du plus petit écart-type σ^5 en phase d'exploitation.

Ainsi, on arrête l'algorithme quand le maximum de calculs du critère est atteint ou si le plus petit écart-type a été utilisé un trop grand nombre de fois consécutivement (par exemple trois fois).

Le dernier vecteur de paramètres p^* obtenu est un *optimum local*. Si le nombre d'évaluations du critère autorisé tend vers l'infini, p^* va tendre vers le minimum global du critère J dans \mathbb{P} [17].

IV.5 Premiers résultats

Les paramètres à estimer sont détaillés dans §IV.2. Nous cherchons à minimiser le critère (IV.1), en nous appuyant sur les données de Langlade (§IV.1). La méthode d'optimisation choisie pour cette étude est un algorithme de recherche aléatoire décrit dans §IV.4. La mise en œuvre de ce problème d'optimisation nécessite quelques réglages.

Plusieurs tests ont été effectués, suivant différents cas de figure (modification des bornes des paramètres du modèle, ou des paramètres de l'algorithme). Le meilleur résultat obtenu est présenté ci-dessous.

IV.5.1 Mise en œuvre

IV.5.1.1 Calcul du critère

Le critère (IV.1) est une somme d'écarts entre incidences simulées et incidences observées dans les données par groupes d'animaux de mêmes génotype et cohorte, sur des intervalles de temps d'observation réguliers.

Pour cette application numérique, nous avons choisi un pas de temps Δt très court d'1 jour et un **pas de temps d'observation de 3 mois**. Le critère est donc calculé pour :

- 11 génotypes (10 + 1 pour les génotypes inconnus),
- des cohortes annuelles (1 mise bas par an),
- des intervalles de temps d'observation de 90 pas de temps.

IV.5.1.2 Gestion de la stochasticité du modèle

Le modèle étant stochastique, nous avons choisi de répéter plusieurs fois le calcul du critère pour un même jeu de paramètres et de minimiser la moyenne de ces différents calculs. Le choix du nombre de répétitions est un juste compromis entre la précision du critère et le nombre d'opérations effectuées, qui affecte le temps de calcul.

Nous avons choisi de faire **7 répétitions** : pour chaque jeu de paramètres, le modèle et le critère sont calculés 7 fois, la moyenne de ces 7 calculs étant la valeur à minimiser par l'algorithme de recherche aléatoire (critère J).

IV.5.1.3 Espace des paramètres

Les bornes des paramètres à estimer sont choisies suffisamment larges pour ne pas trop contraindre le problème d'optimisation, mais elles tiennent compte des connaissances biologiques disponibles, en particulier pour la susceptibilité génétique. Elles sont données dans le TAB. IV.4.

<i>Paramètres</i>	<i>Bornes</i>
Susceptibilité génétique :	$0,7 \leq \sigma_2 \leq 1$ $0 \leq \sigma_3, \sigma_4 \leq 0,6$
Susceptibilité par âge :	$0 \leq a_s \leq 4$ (en années) $0 \leq \varepsilon_a \leq 0,1$
Variance de la charge initiale :	$0 \leq v \leq 0,03$
Durée d'incubation (en années) :	$0 \leq \tau_1, \tau_2, \tau_3, \tau_4 \leq 4$
Coefficient de transmission :	$5 \leq \alpha \leq 10$

TAB. IV.4 – Espace des paramètres utilisé pour la minimisation du critère d'estimation par l'algorithme ARS.

IV.5.1.4 Réglages de l'algorithme de minimisation

Nous avons décidé de modifier certains réglages de l'algorithme pour nous adapter au modèle et mieux explorer l'espace des paramètres.

- L'algorithme n'a testé que **4** écarts-types successifs au lieu de 5, de rapport **2** au lieu de 10, *i.e.* tels que $\sigma^e = \sigma^{(e-1)}/2$.
- Le nombre d'itérations par écart-type en phase d'apprentissage est passé de 100 à **150**. Cela correspond à $312 = \sum_{e=1}^4 \frac{150}{e}$ calculs du critère dans cette phase.
- Le nombre d'itérations en phase d'exploitation est passé de 100 à **300**.

Pour le critère d'arrêt de l'algorithme, nous avons fixé à 6000 le nombre maximum d'évaluations du critère et à 3 le nombre maximum d'utilisations successives du plus petit écart-type.

IV.5.2 Résultats

Dans les conditions de mise en œuvre décrites ci-dessus, l'algorithme de recherche aléatoire s'est arrêté après 3 utilisations successives du plus petit écart-type. Il a effectué 5 cycles apprentissage/exploitation, correspondant à 3060 évaluations du critère. Les résultats obtenus figurent dans le TAB. IV.5.

<i>Paramètres</i>	<i>Valeurs</i>
Susceptibilité génétique :	$\sigma_2 = 0,82$ $\sigma_3 = 0,50$ $\sigma_4 = 0,033$
Susceptibilité par âge :	$a_s = 0,72$ années $\varepsilon_a = 0,052$
Variance de la charge initiale :	$v = 8,71.10^{-4}$
Durée d'incubation (en années) :	$\tau_1 = 1,82$ $\tau_2 = 1,72$ $\tau_3 = 1,67$ $\tau_4 = 1,59$
Coefficient de transmission :	$\alpha = 8,85$

TAB. IV.5 – Valeurs des paramètres estimés par l'algorithme ARS.

Ces valeurs sont tout à fait réalistes. Les susceptibilités génétiques étaient relativement plus contraintes que les autres paramètres, car nous disposions de connaissances a priori. Les résultats obtenus sont *en accord avec les précédentes études* menées sur les données Langlade [6, 14].

Ces résultats mettent en avant une *forte diminution de la susceptibilité des animaux avant l'âge de 1 an*. La contamination est donc presque exclusivement périnatale ; les individus se contaminent lors de la période de mise bas correspondant à leur naissance.

Les durées d'incubation mises en évidence sont peu variables entre individus (faible v). L'ordre de ces durées par rapport au génotype est le même que celui des susceptibilités génétiques : un animal génétiquement plus sensible a une durée d'incubation plus longue qu'un animal plus résistant. Ce résultat n'est pas intuitif. Néanmoins, les différences entre génotypes sont de moins 3 mois. Cette durée correspond justement au pas de temps d'observation du critère, *i.e.* la durée pendant laquelle le nombre de cas de tremblante au sein de chaque groupe génotype×cohorte est comparé entre données et simulations. À partir de ces résultats, *l'influence du génotype sur la durée d'incubation ne semble donc pas significative*.

Chapitre V

Conclusion et perspectives

Nous avons développé un modèle individu-centré stochastique simulant la transmission de la tremblante au sein d'un troupeau ovin. Comme nous disposions de bonnes données de suivi de troupeau, nous avons choisi de ne représenter que les processus de contamination et d'incubation, et d'utiliser directement les données pour la démographie du troupeau. Cela nous a permis de nous concentrer sur les processus épidémiologiques, qui ne sont encore que partiellement connus.

Le simulateur implémenté en Java sur la plateforme Eclipse est doté d'une interface graphique conviviale, qui permet de contrôler et suivre dynamiquement l'évolution de la tremblante dans le troupeau.

L'estimation des paramètres liés aux processus épidémiologiques a été réalisée à partir des données du troupeau expérimental de Langlade (INRA Toulouse, France), grâce à un algorithme de recherche aléatoire par écart-type adaptatif. Les premiers résultats obtenus sont prometteurs. Les valeurs des paramètres obtenues sont réalistes et en accord avec de précédentes études. Elles soulignent une forte diminution de la susceptibilité des animaux avant l'âge de un an. Les durées d'incubation mises en évidence sont peu variables, entre individus et entre génotypes.

Il nous semble intéressant de poursuivre cette étude, en affinant dans un premier temps nos résultats, puis en explorant plus avant différentes hypothèses et scénarios biologiques. Une analyse de sensibilité serait pertinente dans ce contexte. Une autre perspective serait de développer le simulateur en introduisant les processus démographiques dans le modèle, ce qui le rendrait moins dépendant aux données. Nous obtiendrions ainsi un outil de simulation convivial et flexible.

Bibliographie

- [1] ANDRÉOLETTI, O., BERTHON, P., MARC, D., SARRADIN, P., GROSCLAUDE, J., VAN KEULEN, L., SCHELCHER, F., ELSÉN, J.-M., and LANTIER, F. Early accumulation of PrP^{Sc} in gut-associated lymphoid and nervous tissues of susceptible sheep from a Romanov flock with natural scrapie. *J. Gen. Virol.*, 81(12) : 3115–3126, 2000.
- [2] ANDRÉOLETTI, O., LACROUX, C., CHABERT, A., MONNEREAU, L., TABOURET, G., LANTIER, F., BERTHON, P., EYCHENNE, F., LAFOND-BENESTAD, S., ELSÉN, J.-M., and SCHELCHER, F. PrP^{Sc} accumulation in placentas of ewes exposed to natural scrapie : influence of foetal PrP genotype and effect on ewe-to-lamb transmission. *J. Gen. Virol.*, 83(10) : 2607–2616, 2002.
- [3] CHASE-TOPPING, M. E., KRUUK, L. E. B., LAJOURS, D., TOUZEAU, S., MATTHEWS, L., SIMM, G., FOSTER, J. D., RUPP, R., EYCHENNE, F., HUNTER, N., ELSÉN, J.-M., and WOOLHOUSE, M. E. J. Genotype-level variation in lifetime breeding success, litter size and survival of sheep in scrapie affected flocks. *J. Gen. Virol.*, 86(4) : 1229–1238, 2005.
- [4] DETWILER, L. A., and BAYLIS, M. The epidemiology of scrapie. *Rev. sci. tech. Off. int. Epiz.*, 22(1) : 121–143, 2003.
- [5] DÍAZ, C., VITEZICA, Z. G., RUPP, R., ANDRÉOLETTI, O., and ELSÉN, J.-M. Polygenic variation and transmission factors involved in the resistance/susceptibility to scrapie in a Romanov flock. *J. Gen. Virol.*, 86(3) : 849–857, 2005.
- [6] ELSÉN, J.-M., AMIGUES, Y., SCHELCHER, F., DUCROCQ, V., ANDRÉOLETTI, O., EYCHENNE, F., TIEN KHANG, J. V., POLVEY, J.-P., and LAPLANCHE, J.-L. Genetic susceptibility and transmission factors in scrapie : detailed analysis of an epidemic in a closed flock of Romanov. *Arch. Virol.*, 144(3) : 431–445, 1999.
- [7] FLANAGAN, D. *Java examples in a nutshell*. O'Reilly (Berlin-Heidelberg), 2004.
- [8] HAGENAARS, T. J., DONNELLY, C. A., FERGUSON, N. M., and ANDERSON, R. M. The transmission dynamics of the aetiological agent of scrapie in a sheep flock. *Math. Biosci.*, 168(2) : 117–135, 2000.
- [9] HOINVILLE, L. J. A review of the epidemiology of scrapie in sheep. *Rev. Sci. Tech. Off. Int. Epiz.*, 15(3) : 827, 1996.
- [10] MATTHEWS, L., COEN, P. G., FOSTER, J. D., HUNTER, N., and WOOLHOUSE, M. E. J. Population dynamics of a scrapie outbreak. *Arch. Virol.*, 146(6) : 1173–1186, 2001.
- [11] MATTHEWS, L., WOOLHOUSE, M. E. J., and HUNTER, N. The basic reproduction number for scrapie. *Proc. R. Soc. Lond. B*, 266(1423) : 1085–1090, 1999.

- [12] SABATIER, P., DURAND, B., DUBOIS, M. A., DUCROT, C., CALAVAS, D., and DE WIELLE, A. V. Multiscale modelling of scrapie epidemiology : I. Herd level : a discrete model of disease transmission in a sheep flock. *Ecological Modelling*, 180(2-3) : 233–433, 2004.
- [13] STRINGER, S. M., HUNTER, N., and WOOLHOUSE, M. E. J. A mathematical model of the dynamics of scrapie in a sheep flock. *Math. Biosci.*, 153(2) : 79–98, 1998.
- [14] TOUZEAU, S., CHASE-TOPPING, M. E., MATTHEWS, L., LAJOUS, D., EYCHENNE, F., HUNTER, N., FOSTER, J. D., SIMM, G., ELSEN, J.-M., and WOOLHOUSE, M. E. J. Modelling the spread of scrapie in a sheep flock : evidence for increased transmission during lambing seasons. *Arch. Virol.*, 151(4) : 735–751, 2006.
- [15] TREUIL, J.-P., MULLON, C., PERRIER, E., and PIRON, M. Multi-agent simulations of spatial dynamics. In *Models in spatial analysis*, L. Sanders, ed., Geographical Information Systems. ISTE (John Wiley& Sons), 2007, ch. 7, pp. 197–230.
- [16] TREUIL, J.-P., DROGOUL, A., and ZUCKER, J.-D. *Modélisation et simulation à base d’agents : Exemples commentés, outils informatiques et questions théoriques*. Sciences Sup. Dunod, 2008.
- [17] WALTER, E., and PRONZATO, L. *Identification of parametric models from experimental data*. Springer-Verlag (Berlin-Heidelberg), 1997.
- [18] ZIYADI, N., BOULITE, S., HBID, M. L., and TOUZEAU, S. Mathematical analysis of a pde epidemiological model applied to scrapie transmission. *Communications on Pure and Applied Analysis*, 7(3) : 659–675, 2008.

Résumé

La tremblante, une encéphalopathie spongiforme transmissible ou « maladie à prions », est d'un intérêt majeur dans les troupeaux ovins à cause de son fort impact économique et des campagnes réglementaires d'éradication auxquelles elle est soumise. Ses mécanismes de transmission ne sont encore que partiellement connus. Sa caractéristique principale est un facteur de susceptibilité génétique, lié au polymorphisme du gène PrP codant pour la protéine prion. La combinaison d'une longue période d'incubation infectieuse et l'absence de diagnostic *ante-mortem* font des cas cliniques les seules observations possibles de la maladie, très rapidement suivies de la mort de l'animal ou de son abattage.

La modélisation est donc un outil de choix pour étudier la propagation de la tremblante dans un troupeau ovin. Pour sa flexibilité et son caractère intuitif, nous avons opté pour un modèle individu-centré stochastique, un modèle informatique relevant de l'approche multi-agents. Comme nous disposions de bonnes données de suivi pour un troupeau infecté (Langlade, INRA Toulouse), nous avons choisi de ne modéliser que les processus épidémiologiques, contamination et incubation. La démographie est directement issue des données. Nous avons ainsi développé un simulateur sous Java doté d'une interface graphique conviviale.

Les données de suivi de troupeau nous ont permis d'estimer les paramètres épidémiologiques du modèle individu-centré, grâce à une méthode d'optimisation reposant sur un algorithme de recherche aléatoire par écart-type adaptatif. Les premiers résultats sont encourageants. Les valeurs des paramètres obtenues sont réalistes et en accord avec de précédentes études. Elles soulignent une forte diminution de la susceptibilité des animaux avant l'âge de un an. Les durées d'incubation mises en évidence sont peu variables, entre individus et entre génotypes.

Il nous semble intéressant de prolonger cette étude, en affinant dans un premier temps nos résultats, puis en explorant plus avant différentes hypothèses et scénarios biologiques. Une analyse de sensibilité serait pertinente dans ce contexte. Une autre perspective serait de développer le simulateur en introduisant les processus démographiques dans le modèle, ce qui le rendrait moins dépendant aux données. Nous obtiendrions ainsi un outil de simulation convivial et flexible.