



HAL
open science

Construction et analyse d'outils d'apprentissage statistique biomimétiques basés sur des systèmes d'équation aux dérivées partielles pour l'assimilation de données d'élevage

Hélène Flourent

► To cite this version:

Hélène Flourent. Construction et analyse d'outils d'apprentissage statistique biomimétiques basés sur des systèmes d'équation aux dérivées partielles pour l'assimilation de données d'élevage. Statistiques [math.ST]. Université de Bretagne Sud, 2020. Français. NNT : 2020LORIS557 . tel-03324814

HAL Id: tel-03324814

<https://theses.hal.science/tel-03324814v1>

Submitted on 24 Aug 2021

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.

THESE DE DOCTORAT DE

L'UNIVERSITE BRETAGNE SUD

COMUE UNIVERSITE BRETAGNE LOIRE

ECOLE DOCTORALE N° 601

*Mathématiques et Sciences et Technologies
de l'Information et de la Communication*

Spécialité : Mathématiques et leurs Interactions

Par

Hélène FLOURENT

Construction et analyse d'outils d'apprentissage statistique biomimétiques basés sur des systèmes d'Equations aux Dérivées Partielles pour l'Assimilation de Données d'élevage.

Thèse présentée et soutenue à Vannes le 20 mars 2020

Unité de recherche : 6205

Thèse N° : 557

Rapporteurs avant soutenance :

Ernest Fokoué, Professeur, Rochester Institute

Jérôme Monnier, Professeur, INSA de Toulouse

Composition du Jury :

Président : Quansheng Liu, Professeur, Université de Bretagne Sud

Examineurs : Liliane Bel, Professeure, AgroParisTech
Catherine Choquet, Professeure, Université de La Rochelle
Quansheng Liu, Professeur, Université de Bretagne Sud
Vincent Sincholle, Directeur Data & Design Lab, Neovia

Dir. de thèse : Emmanuel Frénod, Professeur, Université de Bretagne Sud

Co-dir. de thèse : Gilles Durrieu, Professeur, Université de Bretagne Sud

À ceux qui m'ont fait grandir.

REMERCIEMENTS

Mes trois années de thèse correspondent à trois années d'investissement et de travail mais surtout d'épanouissement personnel. Cet épanouissement n'aurait pas été possible sans la présence de mon entourage professionnel et personnel.

Je tiens avant tout à remercier profondément mon directeur de thèse, Emmanuel Frénod, professeur à l'Université de Bretagne Sud, pour son soutien constant, pour la pertinence de ses conseils et pour tout ce qu'il m'a transmis. Il m'a permis de conserver pendant trois ans un état d'esprit serein, positif et combatif. Je tiens à le remercier pour son implication mais aussi pour la liberté qu'il m'a accordée. Il est pour moi une source d'admiration, autant sur le plan scientifique que sur le plan humain.

J'adresse également mes remerciements à mon co-directeur de thèse, Gilles Durrieu, professeur à l'Université de Bretagne Sud, pour les encouragements et les conseils qu'il m'a apportés malgré sa délégation à l'Université de Nouvelle-Calédonie durant la première moitié de ma thèse.

Je remercie également Vincent Sincholle, directeur de l'équipe Data et Design Lab chez Neovia, pour son encadrement pendant ces trois années. Il m'a donné la chance de travailler dans une équipe se situant au cur de l'innovation d'une entreprise internationale. J'ai énormément appris à ses côtés et je lui adresse ma gratitude pour tout cela.

Je souhaite également remercier Ernest Fokoué, professeur à l'Institut Technologique de Rochester, États Unis, de m'avoir fait l'honneur d'accepter de suivre mon travail et d'être rapporteur de cette thèse. Je le remercie pour sa sympathie, son enthousiasme, pour le temps qu'il m'a accordé et pour l'ensemble de ses remarques positives et constructives.

Je remercie Jérôme Monnier, professeur à l'INSA de Toulouse, d'avoir accepté

d'être rapporteur de cette thèse. Je le remercie pour sa lecture attentive de mon manuscrit et pour l'ensemble de ses remarques pertinentes et constructives qui m'ont permis d'améliorer la qualité du manuscrit.

Je remercie également Liliane Bel, professeure à AgroParisTech et Catherine Choquet, professeure à l'Université de La Rochelle de me faire l'honneur de leur présence dans mon jury de thèse.

Je remercie Quansheng Liu, professeur à l'Université de Bretagne Sud, d'avoir accepté de faire partie de mon jury et d'en être le président.

J'adresse des remerciements particuliers à David Causeur, professeur à Agrocampus-Ouest. Je le remercie de m'avoir fait l'honneur de continuer à participer à ma formation de scientifique après la fin de mes études à Agrocampus-Ouest, en acceptant de faire partie de mon Comité de Thèse et de mon Comité de Suivi Individuel. Je le remercie pour sa disponibilité, sa bienveillance et son implication à mes côtés au cours de mes trois années de thèse. Nos échanges et l'ensemble de ses conseils m'ont permis d'orienter et de structurer mes travaux.

Cette thèse doit beaucoup au cadre dans lequel elle s'est déroulée, et j'aimerais à ce titre remercier les personnes que j'ai côtoyées et qui m'ont soutenue pendant ces trois années.

Je remercie d'abord, Sandrine De Oliveira, la gestionnaire du laboratoire de mathématique pour son efficacité, ses conseils et sa gentillesse.

Je remercie l'ensemble des membres du LMBA, enseignants-chercheurs, doctorants et post-doctorants, pour leur sympathie, leur écoute et leur soutien.

Je remercie l'ensemble des collègues avec qui j'ai eu la chance de travailler chez Neovia, pour leur esprit de camaraderie, leurs blagues et leur bonne humeur. Nos parties de babyfoot acharnées m'ont permis de me défouler et de décompresser pendant ces trois ans.

Je remercie également les membres de l'équipe de See-d que j'ai eu le plaisir de côtoyer, pour nos échanges, pour leur gentillesse et tous les déjeuners et petits déjeuners partagés.

J'adresse des remerciements chaleureux et tout particuliers à Jamila et Cécile, doc-

torantes de l'Université de Bretagne Sud, avec lesquelles j'ai pu partager les joies et les difficultés de faire une thèse. Merci à elles pour les heures que nous avons passées à rire et discuter.

Je remercie également mes proches, mes amis et ma famille, qui m'ont soutenue, de près ou géographiquement de loin, mais toujours avec attention et affection.

Enfin, je voudrais terminer en remerciant mon petit "clan" familial d'avoir fait de moi ce que je suis aujourd'hui. Merci à ma mamie pour les valeurs et le goût du travail bien fait qu'elle s'est efforcée, non sans mal, de me transmettre quand j'étais enfant. Merci à ma "petite maman" pour sa confiance, sa patience, sa générosité (...) et l'incroyable bienveillance avec laquelle elle m'a élevée. Et enfin, merci à ma grande sur, pour la tendresse, la protection et l'attention qu'elle me porte depuis mes premiers jours. Merci à elles pour leurs conseils et leur soutien sans faille.

TABLE DES MATIÈRES

Introduction générale	23
1 Utilisation d'outils mathématiques et statistiques dans le domaine de la production animale : Etat de l'art et approche explorée	31
1.1 Les intérêts et les difficultés de la modélisation mathématique dans le domaine de la production animale.	32
1.2 Les différents types de modèles mathématiques et les différentes méthodes d'apprentissage	34
1.2.1 L'apprentissage supervisé et non-supervisé	34
1.2.2 Les modèles paramétriques et non-paramétriques	35
1.2.3 Des modèles statiques et des modèles dynamiques	36
1.2.4 Des modèles déterministes et des modèles stochastiques	37
1.2.5 Des modèles «boîtes-noires» et des modèles "boîtes-blanches"	37
1.3 Émergence d'approches intermédiaires	40
1.3.1 L'apprentissage statistique	40
1.3.2 Le Couplage Modèle-Données	42
1.4 Réapprentissage et Assimilation de Données	44
1.5 Résumé de l'état de l'art	45
1.6 L'approche explorée dans ces travaux de thèse	49
1.7 Description des données disponibles et de la méthode d'adimensionnement appliquée	53
1.7.1 Les données disponibles	53
1.7.2 Adimensionnement des données	55
1.8 Présentation de l'algorithme DIRECT	58
2 Construction et analyse d'outils d'apprentissage statistique biomimétiques	61
2.1 Construction d'un premier modèle mathématique biomimétique	62
2.1.1 Élaboration des équations constituant le modèle mathématique	63

TABLE DES MATIÈRES

2.1.2	Contextes d'application du modèle mathématique et les différentes versions développées	65
2.1.3	Les conditions initiales et les conditions aux bords	67
2.1.4	Les paramètres du modèle	68
2.1.5	Discrétisation du modèle mathématique	68
2.2	Etude des gammes de valeurs des paramètres du modèle mathématique.	68
2.2.1	Détermination des gammes de valeurs des vitesses de convection et de diffusion	69
2.2.2	Etude de l'effet de saturation de l'impact des paramètres sur le modèle	70
2.2.3	Détermination des gammes de valeurs de u_a , f_a et r_a	74
2.3	Tests par simulation de la capacité d'apprentissage de l'outil d'apprentissage statistique modélisant un phénomène de croissance	80
2.3.1	Génération d'une base de données fictive	81
2.3.2	Les relations existantes entre les paramètres	84
2.3.3	Ajustement des paramètres et calcul de la précision du modèle	90
2.3.4	Étude de la robustesse du modèle et de l'impact de la présence de bruit dans les données d'apprentissage	94
2.4	Tests par simulation de la capacité d'apprentissage des autres versions du modèle "A" : Le modèle "A d'augmentation" et de "diminution".	99
2.4.1	Génération de bases de données :	99
2.4.2	Ajustement des paramètres	101
2.4.3	Capacité d'extrapolation des différentes versions du modèle "A"	103
2.5	Application de l'approche développée sur des données réelles	104
2.5.1	Le modèle mathématique utilisé	105
2.5.2	Les données utilisées	105
2.5.3	Etude des relations existantes entre les paramètres du modèle	106
2.5.4	Ajustement des paramètres du modèle	106
2.5.5	Résultats	108
2.5.6	Discussion des résultats	109
2.6	Conclusion du chapitre 2	110
3	Existence, unicité et propriétés qualitatives du système d'EDP embarqué dans nos outils d'apprentissage statistique biomimétiques.	113

3.1	Introduction	113
3.2	Structure et analyse du modèle mathématique développé initialement .	122
3.2.1	Propriété de conservation du modèle mathématique développé initialement	122
3.2.2	Principe de la Formulation Variationnelle	123
3.2.3	Les premières estimations a priori	127
3.3	Identification des propriétés perdues du fait de la structure de "l'opérateur de diffusion" grâce à l'utilisation d'un premier modèle simplifié. . . .	133
3.3.1	Unicité de la solution	133
3.3.2	L'effet régularisant de "l'opérateur de diffusion"	136
3.3.3	La question de la positivité des fonctions $\{\Phi_f(d)\}$ et $\{\Phi_b(d)\}$	137
3.4	Démonstration de l'existence et de l'unicité de la solution d'un deuxième modèle simplifié : un modèle mathématique sans terme de convection .	142
3.4.1	Les expressions exactes de $\{\Phi_f(d)\}$ et $\{\Phi_b(d)\}$	143
3.4.2	Positivité de certaines composantes de la solution du système d'EDP	154
3.4.3	Propriété de conservation du second sous-modèle	160
3.4.4	Majoration de la norme infinie de $\{\Phi_f(d)\}$, $\{\Phi_b(d)\}$ et $\{\Psi(d)\}$	162
3.4.5	Existence de la solution des équations (2.1.4b) et (2.1.4)	165
3.4.6	Unicité de la solution du second modèle simplifié	168
3.5	Conclusion	172
4	Intégration de variables d'entrée dans des outils d'apprentissage statistique biomimétiques	175
4.1	Valorisation de l'information contenue dans la variable relative à l'âge des animaux	175
4.1.1	Construction d'une courbe de croissance moyenne	176
4.1.2	Modulation de la valeur du paramètre P_{max_i}	178
4.1.3	Ajustement des paramètres du modèle « A_{Age} »	179
4.2	Construction d'un modèle simulant la circulation et l'action de deux informations en interaction : le modèle «B»	182
4.2.1	Objectifs et intérêts de ce deuxième modèle	182
4.2.2	La structure du modèle "B"	182
4.2.3	Les données utilisées	186

TABLE DES MATIÈRES

4.2.4	Construction de l'équation d'utilisation	186
4.2.5	Ajustement des paramètres de la fonction d'utilisation, U :	188
4.2.6	Ajustement des paramètres du modèle "B"	189
4.3	Construction d'un modèle simulant la circulation et l'action de deux informations en interaction : le modèle «C»	192
4.3.1	Objectifs	192
4.3.2	La structure du modèle «C»	193
4.3.3	Perspective concernant l'utilisation du modèle «C»	196
4.4	Conclusion du chapitre 4	197
5	Positionnement et domaine d'application des outils d'apprentissage statistique biomimétiques	199
5.1	Comportement des modèles biomimétiques en fonction de leur structure et des caractéristiques des données d'apprentissage	199
5.1.1	Gain de temps d'exécution apporté par la parallélisation des calculs	200
5.1.2	Étude de l'évolution du temps d'apprentissage en fonction du nombre de flux d'information modélisés par le modèle biomimétique	202
5.1.3	Évolution du nombre d'itérations nécessaire avant d'atteindre la convergence des paramètres en fonction du nombre de paramètres à déterminer et du nombre d'individus présents dans la base d'apprentissage	208
5.1.4	Étude de l'impact de la prise en compte des relations existantes entre les paramètres sur le temps d'apprentissage des modèles	212
5.2	Positionnement de l'outil de prédiction de croissance biomimétique par rapport à des outils existants	216
5.2.1	Comparaison des performances du modèle de croissance biomimétique et de modèles "boîtes noires"	216
5.2.2	Comparaison des performances du modèle de croissance biomimétique et des modèles classiques de croissance logistique	223
5.2.3	Comparaison entre le modèle de croissance biomimétique et des modèles de croissance mécanistes	225
5.2.4	Conclusion du chapitre 5	226

6 Chapitre conclusif : Réflexions en cours et perspectives	229
6.1 Réflexion concernant une potentielle application mobile permettant de piloter un atelier d'engraissement.	229
6.2 Réflexion et premiers pas vers l'automatisation d'une procédure de sélection de modèles dans le cadre du Couplage Modèle-Données	231
6.2.1 Conception d'un outil de sélection de modèles	232
6.2.2 Sélection des modèles	234
6.2.3 Perspectives de ce projet	235
6.3 Premières réflexions autour de la construction et de l'évaluation d'un outil d'Assimilation de Données, basé sur un outil d'apprentissage statistique biomimétique.	236
6.3.1 Les notations utilisées dans cette section	236
6.3.2 Réflexion autour de l'élaboration d'une méthode permettant de tester la qualité d'un outil d'Assimilation de Données	237
Conclusion générale	241
Annexe I : L'article synthétisant les travaux produits lors du CEMRACS 2018	243

TABLE DES FIGURES

1	Schématisation de la capacité d'extrapolation des outils développés . . .	28
1.1	Résumé de l'état de l'art.	46
1.2	Synthèse des différents niveaux considérés dans notre approche, leurs composants et les processus qu'ils contiennent.	51
1.3	Articulation des différents éléments intervenant dans notre exploration .	52
2.1	Schématisation du modèle mathématique principal	62
2.2	Saturation de l'impact des paramètres f_a (à gauche), u_a (au centre) et r_a (à droite), sur le modèle	71
2.3	La valeur de l'indicateur I_{ijk} en fonction de la valeur de $u_{a,n+1}$	72
2.4	La valeur de u_{aUpp}^{ijk} en fonction des valeurs de ω_a , r_a et f_a	74
2.5	La valeur de l'indicateur f_{aUpp}^{ijk} en fonction des valeurs de u_a , r_a et ω_a . . .	75
2.6	La valeur de l'indicateur f_{aUpp}^{ijk} en fonction de la valeur de r_a , et colorée en fonction de la valeur de ω_a	75
2.7	La valeur de l'indicateur f_{aUpp}^{ijk} en fonction de la valeur de r_a	76
2.8	La relation reliant r_a et f_{aUpp}^{ijk} . Le modèle "ajusté" (Fitted) et le modèle "ajusté et majoré" (Majoré Fitted) correspondent à l'équation (2.2.8) pa- ramétrée respectivement avec les paramètres (A, B, C, D, E) et (A, B, C, D, E_M) . 77	
2.9	Distributions des paramètres ω_a , r_a , f_a et u_a	82
2.10	Exemple de variabilité pouvant être présente dans la base de données générée.	83
2.11	Exemple de courbes simulées avant et après l'ajout d'un bruit gaussien. 84	
2.12	La valeur de l'indicateur $RRSS$ en fonction de la valeur des paramètres ω_a et r_a (a : gauche), et le schéma des différentes séries de couples équivalents (EC) (b : droite)	86
2.13	La représentation 3D de la valeur de $RRSS$ en fonction de ω_a et r_a . . .	87
2.14	La régression de Nadaraya-Watson reliant les couples $(\omega_{dOpt}, r_{dOpt})$. . .	88
2.15	La valeur du $RRSS$ en fonction de f_a et u_a (a) et le schéma représentant les différentes séries de couples équivalents (EC) (b)	89

2.16	La représentation 3D de la valeur du $RRSS$ en fonction de f_a et u_a	89
2.17	La régression de Nadaraya-Watson reliant $f_{d_{Opt}}$ et $u_{d_{Opt}}$	91
2.18	Exemples de résultats obtenus concernant la concordance entre les courbes d'apprentissage (noire) et les prédictions fournies par le modèle (rouge).	92
2.19	Distributions des valeurs des indicateurs $RRSS$ et R^2 associées au modèle "générateur" et au modèle ajusté.	93
2.20	Exemples de courbes bruitées contenues dans les différentes bases de données fictives $n\check{r}1$, $n\check{r}2$, $n\check{r}3$, $n\check{r}4$, $n\check{r}5$ et $n\check{r}6$	96
2.21	Les relations non paramétriques reliant ω_a et r_a (Haut), et f_a et u_a (Bas) établies sur les nouvelles bases de données générées, numérotées de 2 à 6.	97
2.22	Différence entre les valeurs de R^2 obtenues pour les modèles ajustés et les modèles "générateurs" (10^{-3}).	98
2.23	Exemples de courbes de sortie générées par le modèle « de diminution » et induites par une seule (Gauche) ou deux injections d'information (Droite).	100
2.24	Exemples de courbes de sorties générées par le modèle « d'augmentation » et induites par une seule (Gauche) ou deux injections d'information (Droite).	101
2.25	Exemples de courbes simulées avec le modèle « d'augmentation » ajusté (–), en comparaison avec les courbes de la base de test (○).	102
2.26	Exemples de courbes simulées avec le modèle de « diminution » ajusté (–), en comparaison avec les courbes de la base de test (○).	102
2.27	Exemples de courbes simulées (–) avec le modèle « de diminution » (Gauche) et avec le modèle « d'augmentation » (Droite) ajustés, en comparaison avec les données de test modélisant la réponse à deux injections d'information (○).	103
2.28	Les régressions de Nadaraya-Watson reliant les couples $(\omega_{d_{Opt}}, r_{d_{Opt}})$ (gauche) et $(f_{d_{Opt}}, u_{d_{Opt}})$ (droite) sur les données d'apprentissage réelles.	107
2.29	Différences obtenues entre les valeurs de la variable de sortie mesurées (+) et celles prédites (×), à différents instants t pour les individus de la base de test.	109

TABLE DES FIGURES

4.1	Droite de poids moyen adimensionné et positionnement des individus étudiés sur cette droite.	178
4.2	Différences obtenues entre les poids mesurés (+) et ceux prédits (\times), à différents instants t , pour les individus de la base de test.	181
4.3	Schématisation de la structure globale du modèle «B».	183
4.4	Le Gain Quotidien Moyen de chaque groupe de l'élevage B.	187
4.5	Courbe de la fonction d'utilisation U (—) et la moyenne des mesures de GMQ (\circ).	189
4.6	Différences obtenues entre les valeurs mesurées (\times) et celles prédites (+) à $t = 1.52$ (Gauche) et à $t = 2.5$ (Droite) pour les individus de la base de test.	191
4.7	Structure du modèle «C».	193
5.1	Monitoring du fonctionnement des processus lors de l'ajustement d'un modèle biomimétique avant la parallélisation des calculs	201
5.2	Monitoring du fonctionnement des processeurs lors de l'ajustement d'un modèle biomimétique après la parallélisation des calculs	202
5.3	Courbe d'évolution du temps de calcul (min) en fonction du modèle biomimétique utilisé et du nombre d'individus présents dans la base d'apprentissage	203
5.4	Schématisation de la structure globale du modèle « $B_{SsInter}$ », sans interaction modélisée entre les informations	205
5.5	Schématisation de la structure globale du modèle « $C_{SsInter}$ », sans interaction modélisée entre les informations circulant	206
5.6	Courbe d'évolution du temps de calcul nécessaire pour réaliser une itération (min) en fonction du modèle biomimétique utilisé et du nombre d'individus présents dans la base d'apprentissage	207
5.7	Le nombre d'itérations effectuées en fonction de la valeur du critère d'arrêt, en utilisant une base d'apprentissage contenant 5 individus (Gauche) et 10 individus (Droite).	210
5.8	Le nombre d'itérations effectuées en fonction du nombre de paramètres à déterminer, en utilisant une base d'apprentissage contenant 5 individus (Gauche) et 10 individus (Droite).	210
5.9	Moyenne et variance de l'indicateur ARE en fonction du nombre d'individus dans la base de données d'apprentissage.	212

5.10	Le nombre d'itérations effectuées en fonction de la valeur du critère d'arrêt, en utilisant une base d'apprentissage contenant 5 individus (Gauche) et 10 individus (Droite).	213
5.11	Le nombre d'itérations effectuées en fonction du nombre de paramètres à déterminer, en utilisant une base d'apprentissage contenant 5 individus (Gauche) et 10 individus (Droite).	214
5.12	Précision obtenue ($ARE(1)$) en utilisant des Random Forest contenant différents nombres d'arbres (1 à 700), sur la base d'apprentissage (Gauche) et sur la base de test (Droite).	221
5.13	Précision obtenue ($ARE(1)$) en utilisant des Random Forest contenant différents nombres d'arbres (1 à 200), sur la base d'apprentissage (Gauche) et sur la base de test (Droite).	221
5.14	Représentation graphique des valeurs moyennes de l'indicateur ARA calculées sur la base de test à différents instants et associé aux différents modèles testés.	225
5.15	Exemples de courbes de croissance de la base de test simulées le modèle biomimétique, le modèle de Gompertz et celui de Verhulst.	225
6.1	Processus de comparaison de l'ensemble des modèles	233
6.2	Schématisation du processus d'ajustement et de test des différents modèles testés.	233
6.3	Schématisation des étapes permettant de tester la qualité d'un outil d'Assimilation de Données.	238

LISTE DES TABLEAUX

1.1	Description de la composition des données et des différents groupes étudiés (NI : Nombre d'Individus, × : Donnée disponible, † : Donnée estimée).	55
2.1	Le plan d'expérience permettant de déterminer la borne supérieure de la gamme de valeurs de p_E quelle que soit la valeur des autres paramètres, p_{F_1} , p_{F_2} et p_{F_3}	73
2.2	Coefficients de l'équation reliant f_{aUpp}^{ijk} et r_a	77
2.3	Les couples $(\omega_{a,H}; r_{a,H})$ testés et leurs couples équivalents $(\omega_{a,L}; r_{a,L})$	79
2.4	Les gammes de valeurs des différents paramètres.	80
2.5	Distributions suivies par les paramètres et les données d'entrée.	84
2.6	Moyenne et écart-type relatif associés aux paramètres et au coefficient de détermination calculé sur la base d'apprentissage	92
2.7	Comparaison des valeurs des indicateurs associées au modèle "générateur" et au modèle ajusté.	94
2.8	Les différentes quantités de bruit intégrée dans les 6 bases de données générées	95
2.9	La valeur des paramètres ajustés sur chaque base de données fictive	97
2.10	Les valeurs de R^2 associées aux modèles ajustés et aux modèles "générateurs", calculées sur les 6 bases de données fictives et les mesures de l'écart existant entre ces modèles.	98
2.11	Comparaison de la précision des modèles obtenus après ajustement (modèles "A" ajustés) et celle des modèles ayant été utilisés pour générer les données (modèles "A" "générateurs"), pour les modèles de "diminution" et "d'augmentation".	102

2.12	Comparaison de la précision des modèles obtenus après ajustement (modèles "A" ajustés) et celle des modèles ayant été utilisés pour générer les données (modèles "A" "générateurs"), pour les modèles de "diminution" et "d'augmentation", sur les données simulant la réponse de la variable de sortie à deux injections d'information.	104
2.13	Résumé des données utilisées.	106
2.14	La moyenne et l'écart type relatif (<i>RSD</i>) des paramètres ajustés, et l'indicateur <i>ARE</i> calculé à l'instant $t = 1$ sur la base d'apprentissage. . . .	108
2.15	Erreur relative moyenne (<i>ARE</i>) calculée à différents instants t pour les individus de la base de test.	109
4.1	La moyenne et l'écart type relatif (<i>RSD</i>) des paramètres ajustés, et l'indicateur <i>ARE</i> calculé l'instant $t = 1$ sur la base d'apprentissage.	180
4.2	Erreur relative moyenne (<i>ARE</i>) calculée à différents instants t sur les individus de la base de test.	181
4.3	La valeur des paramètres de la fonction d'utilisation, U	188
4.4	Valeurs des paramètres ajustés et <i>ARE</i> calculé à l'instant $t = 1.52$ et $t = 2.5$, sur la base d'apprentissage.	191
4.5	Erreur relative moyenne (<i>ARE</i>) calculée à l'instant $t = 1.52$ et $t = 2.5$, pour les individus de la base de test.	191
5.1	Évolution du temps de calcul nécessaire pour réaliser une itération (min) en fonction du modèle biomimétique utilisé et du nombre d'individus présents dans la base d'apprentissage	203
5.2	Évolution du temps de calcul nécessaire pour réaliser une itération (min) en fonction du modèle biomimétique utilisé et du nombre d'individus présents dans la base d'apprentissage	206
5.3	Évolution du nombre d'itérations nécessaire avant d'atteindre la convergence des paramètres en fonction du nombre de paramètres à déterminer (param.), de la valeur du critère d'arrêt (ϵ) et du nombre d'individus (NI) présents dans la base d'apprentissage.	209
5.4	Gain apporté en termes de nombre d'itérations nécessaire avant d'atteindre la convergence des paramètres lorsque le nombre d'individus dans la base d'apprentissage passe de 5 à 10 individus, en fonction de la valeur du critère d'arrêt et du nombre de paramètres appris.	211

5.5	Évolution du nombre d'itérations en fonction du nombre de paramètres à déterminer, de la valeur du critère d'arrêt et du nombre d'individus (NI) présents dans la base d'apprentissage en utilisant le modèle " A_{Rel} ". . .	213
5.6	Évolution de la durée (min) globale d'ajustement du modèle " A_{Rel} " en fonction du nombre de paramètres à déterminer, de la valeur du critère d'arrêt et du nombre d'individus (NI) présents dans la base d'apprentissage.	215
5.7	Gain Moyen (GM) et Gain Moyen Relatif (GMR) en terme de nombre d'itérations à réaliser avant d'atteindre la convergence des paramètres, lorsque les relations existantes entre les couples de paramètres (ω_a, r_a) et (f_a, u_a) sont prises en compte.	215
5.8	Gain Moyen (GM) et Gain Moyen Relatif (GMR) en terme de durée (minute) d'apprentissage, lorsque les relations existantes entre les couples de paramètres (ω_a, r_a) et (f_a, u_a) sont prises en compte.	216
5.9	Précision (ARA) obtenue en utilisant un modèle linéaire ou le modèle biomimétique sur la base d'apprentissage (BA) et sur la base de test (BT).218	
5.10	Les valeurs de l'indicateur ARA calculées sur la base d'apprentissage (ARA_{Train}), et sur la base de test (ARA_{Test}), à $t = 1$, en utilisant différents réseaux de neurones. Le réseau de neurone $(k_1, \dots, k_i, \dots, k_n)$ correspond à un réseau contenant n couches cachées et dans lequel la i^{me} couche contient k_i noeuds.	219
5.11	Gamme de précision, variance et écart-type de la précision relative ($ARE(1)$) pouvant être obtenue en utilisant des Random Forest contenant différents nombres d'arbres (1 à 200), sur la base d'apprentissage et sur la base de test.	222
5.12	La meilleure précision obtenue ($ARE(1)$) sur la base de test, en utilisant des Random Forest (BA : Base d'Apprentissage, BT : Base de Test). . .	222
5.13	Valeurs des paramètres obtenues après l'ajustement et valeurs moyennes de l'indicateur $ARA(1)$ calculées sur la base d'apprentissage pour les différents modèles étudiés.	224
5.14	Valeurs moyennes de l'indicateur ARA calculées sur la base de test à différents instants, pour le modèle biomimétique, celui de Verhulst et celui de Gompertz.	224

5.15 Tableau de comparaison de la précision (%) des différentes méthodes utilisées pour prédire le poids des individus à l'instant $t = 1$ ($P(1)$) et à l'instant $t = 2.5$ ($P(2.5)$) (BA : Base d'Apprentissage, BT : Base de Test, RL : Régression Linéaire, RN : Réseaux de Neurones, RF : Random Forest, Biomim. : Modèle Biomimétique).	227
--	-----

INTRODUCTION GÉNÉRALE

Dans leurs travaux WALTER *et al.* (2017) expliquent que l'agriculture a connu jusqu'à présent trois révolutions. La première correspond à la domestication des animaux et des plantes, il y a quelques milliers d'années. Plusieurs centaines d'années auparavant, il y eu la mise en place du système de rotation des cultures et l'amélioration des pratiques d'élevage. Enfin, il y a quelques dizaines d'années, une troisième révolution a eu lieu. Cette révolution, appelée "la révolution verte", correspond à l'augmentation significative de l'utilisation d'intrants et d'espèces animales et végétales à hauts potentiels. D'après WALTER *et al.* (2017), une quatrième révolution a démarré depuis quelques années.

L'agriculture doit aujourd'hui relever de nouveaux défis. Le premier est de nourrir une population dont le nombre d'individus ne cesse d'augmenter. D'après la FAO (Foundation and Agriculture Organisation), la production alimentaire doit augmenter de 70% d'ici 2050 pour espérer nourrir une population qui devrait alors atteindre les 9 milliard d'individus (FOOD et ORGANIZATION (2009)). Le deuxième défi est d'augmenter cette production agricole tout en préservant l'environnement.

La quatrième révolution agricole évoquée précédemment tente de faire face à ces nouvelles problématiques grâce au développement de l'agriculture de précision. Ce développement se fait via l'utilisation de plus en plus fréquente de technologies d'information et de communication (WALTER *et al.* (2017), JAYARAMAN *et al.* (2016) et WOLFERT *et al.* (2017)). La collecte d'informations diverses sur l'ensemble de l'exploitation agricole et ensuite leur mise en lien, semblent être aujourd'hui les leviers permettant de révolutionner l'agriculture (JAYARAMAN *et al.* (2016)). Ce que l'on appelle le "Smart Farming", et plus particulièrement l'agriculture de précision, consiste donc à utiliser de nouvelles technologies, telles que des capteurs et des machines connectées, dans le but de suivre, de contrôler et finalement de rendre plus efficaces les processus de production agricoles. L'objectif est de maximiser les rendements et la production à partir d'une quantité d'intrants utilisée (alimentation, fertilisation, *etc.*) (WALTER *et al.* (2017), O'GRADY et O'HARE (2017) et ESMEIJER *et al.* (2015)). Pour une même quantité produite, l'augmentation de l'efficacité des plantes et des animaux revient à

réduire les pertes et à optimiser la valorisation des intrants utilisés. L'agriculture de précision permet donc de limiter l'impact de l'agriculture sur l'environnement (SCHULZE *et al.* (2009)). En réduisant la quantité d'intrant nécessaire et les pertes, l'agriculture de précision permet aussi de réduire les pertes économiques et d'augmenter les revenus des agriculteurs (RYU *et al.* (2015) et WALTER *et al.* (2017)).

Augmenter la productivité agricole passe par la prédiction et l'anticipation des performances des différentes espèces animales et végétales en fonction des conditions d'élevage et de culture (O'GRADY et O'HARE (2017) et JAYARAMAN *et al.* (2016)). Il est également important de pouvoir suivre en temps réel l'évolution de différents indicateurs et l'état des cultures ou des troupeaux (JAYARAMAN *et al.* (2016) et WALTER *et al.* (2017)). Cela nécessite d'être capable de collecter des données fréquemment et ensuite de les assimiler. L'état réel des cultures ou des animaux peut alors être estimé en réalisant de l'Assimilation de Données. Cette méthode consiste à combiner l'information fournie par un outil de prédiction et celle contenue dans les données collectées pour estimer l'état le plus vraisemblable d'un système à un instant t (Voir AUROUX et BLUM (2005), GREGG *et al.* (2009), LGUENSAT *et al.* (2017) and LGUENSAT *et al.* (2019)). Pour cela il faut donc tout d'abord collecter des données (FAULKNER et CEBUL (2014)).

De nombreuses technologies ont vu le jour ces dernières années afin de permettre la collecte de données dans l'ensemble de l'exploitation agricole (JEMILA et PRIYADHAR-SINI (2018), MIEKLEY *et al.* (2012), TOL et KAMP (2010), BÜCHEL et SUNDRUM (2014) et HOLMAN *et al.* (2011)). Néanmoins, les données agricoles sont souvent très hétérogènes et d'assez mauvaise qualité (ISHII (2014), LI *et al.* (2014) et WOLFERT *et al.* (2014)). De plus l'accès aux nouvelles technologies représente encore aujourd'hui un investissement important pour les agriculteurs qui ne sont de ce fait pas encore tous équipés de machines connectées (WALTER *et al.* (2017) et KUTTER *et al.* (2011)). Un des grands enjeux de cette quatrième révolution agricole est donc le développement de modèles mathématiques et d'outils de prédiction performants, capables de gérer conjointement un manque de données fréquent, la complexité de celles-ci, mais aussi la complexité des phénomènes étudiés.

Concrètement, l'agriculture de précision repose donc sur la collecte de données mais aussi sur la construction d'outils mathématiques et statistiques capables d'assimiler ces données d'élevage, de les synthétiser et de leur donner du sens.

Ces travaux ont été réalisés dans le cadre d'une thèse en Convention Industrielle de Formation par la REcherche (CIFRE), qui a été effectuée dans une entreprise d'agroalimentaire. Cette entreprise produit de l'alimentation pour les animaux d'élevage. L'objectif de cette entreprise est de proposer aux éleveurs une ration alimentaire qui garantisse l'optimisation des performances des animaux selon l'espèce, la race et les conditions d'élevage. Cette entreprise souhaite également apporter du conseil aux éleveurs et les aider à piloter leur élevage. Pour cela cette entreprise a besoin d'outils capables de mettre en lien la composition de la ration distribuée et les performances des animaux, en termes de croissance ou de production laitière. L'objectif final est de déterminer la ration optimisant les performances des animaux, ou simplement de pouvoir indiquer à l'éleveur le poids que feront ses animaux à une certaine date selon la ration alimentaire qu'il souhaite utiliser.

Le premier objectif de cette thèse a donc été de construire des outils capables de mettre en lien des données d'entrée et des données de sortie associées à un animal ou à groupe d'animaux d'élevage, dans le but de prédire l'évolution de variables biologiques. Le second objectif a été de veiller à ce que les outils construits soient également capables de réaliser de l'Assimilation de Données, c'est-à-dire de synthétiser et d'interpréter un flot de données provenant de manière plus ou moins régulière d'un élevage connecté. Et enfin, l'objectif plus global de cette thèse a été de démontrer que les mathématiques et plus particulièrement les Équations aux Dérivées Partielles (EDP), peuvent être des outils intéressants pour développer des Intelligences Artificielles (IA) destinées à traiter des problèmes liés à l'agriculture de précision.

D'après VÁZQUEZ-CRUZ *et al.* (2014), il existe aujourd'hui deux principaux types d'approches permettant de construire des outils de prédiction capables de mettre en lien des données biologiques : Des approches de type mécanistes réalistes et des approches dites "boîtes-noires".

La modélisation réaliste d'un phénomène biologique est un processus complexe qui nécessite beaucoup de connaissances a priori et qui mène généralement à la construction de modèles mathématiques contenant un grand nombre d'équations et de paramètres à déterminer. L'implémentation de ces outils est donc complexe et limite leur adaptabilité surtout lorsqu'il s'agit de traiter des données brutes, hétérogènes et d'ajuster des paramètres. Néanmoins, ces modèles qui valorisent les connaissances et

l'expertise des biologistes, constituent des outils très informatifs et utiles pour pouvoir tester des hypothèses concernant le fonctionnement de systèmes biologiques (TOMLIN et AXELROD (2007)).

À l'inverse, les outils de d'apprentissage automatique, appelé couramment outils de "Machine Learning", et outils "boîtes-noires", ne nécessitent aucune connaissance a priori en ce qui concerne la dynamique du système étudié. En revanche, ces outils, tels que les Réseaux de Neurones, nécessitent de disposer d'une grande quantité de données collectées pour pouvoir les ajuster, surtout lorsqu'il s'agit de prédire des phénomènes complexes. Étant donné que ces outils n'intègrent pas de connaissances biologiques, ils sont très génériques et peuvent être utilisés pour traiter des problématiques très variées. Seulement, cette généricité limite leur adaptabilité et la possibilité de prendre en compte de l'existence de phénomènes biologiques sous-jacents, comme des phénomènes de diffusion, de retard, de saturation ou encore d'accumulation.

Leur manque d'adaptabilité ou leur incapacité à s'ajuster sur une faible quantité de données, rend les outils existants peu adaptés pour réaliser de l'Assimilation de Données d'élevage.

Au cours de ces travaux de thèse nous avons donc exploré une approche de modélisation se situant à mi-chemin entre une approche de modélisation réaliste et une approche dite "boîtes-noires". En effet, l'approche explorée, appelée "Couplage Modèle-Données", combine à la fois l'intégration de connaissances dans un modèle mathématique et l'utilisation de données pour ajuster la valeur des paramètres de ce modèle mathématique.

Nous avons fait l'hypothèse qu'une prise en compte globale et synthétique des processus biologiques pourrait permettre à des outils de prédiction d'être plus précis et d'avoir une capacité d'apprentissage plus grande que les Réseaux de Neurones qui n'intègrent aucune connaissance a priori. Nous avons également fait l'hypothèse que cette prise en compte synthétique permettrait cela tout en gardant un outil plus léger que les outils basés sur des modèles réalistes.

L'Assimilation de Données est un objectif à long terme des travaux de recherche présentés dans ce manuscrit. Néanmoins, pour aboutir à un outil particulièrement adapté à cet objectif, celui-ci a fortement orienté l'ensemble des choix que nous avons faits tout au long de l'élaboration de notre approche de modélisation.

Notre contexte de recherche a été marqué par deux autres contraintes : la com-

plexité des phénomènes biologiques et le peu de données disponibles pour ajuster nos outils. Ces deux problématiques ont également fortement influencé les choix que nous avons faits. C'est pourquoi, la recherche de parcimonie et à la fois d'un grand pouvoir d'extraction d'informations contenues dans les données, constituent deux quêtes centrales de ce travail de thèse.

Au vu de nos divers objectifs de précision, de parcimonie et de capacité d'apprentissage, nous avons décidé d'élaborer des systèmes d'EDP, contenant des paramètres associés à des facteurs biologiques. Ces paramètres pouvant être appris à partir de données collectées et d'un algorithme d'optimisation, apportent aux outils construits une capacité d'apprentissage. Ces équations et cette capacité d'apprentissage font de nos outils, des outils d'apprentissage statistique biomimétiques. Une fois les paramètres ajustés, le système d'EDP est embarqué dans un outil capable de prédire l'évolution de variables biologiques.

Plusieurs contributions ont été apportées par les travaux de thèse présentés dans ce manuscrit.

Tout d'abord, les applications de nos outils sur des données générées et des données collectées, relatives à la croissance d'animaux, ont mis en évidence leur capacité d'extrapolation, qui est un réel avantage par rapport aux outils existants. Comme illustré par la figure 1, nos outils ont été entraînés pour mettre en lien les données d'entrée et de sortie sur une période d'apprentissage assez courte. Malgré cela ils sont capables de prédire avec précision l'évolution de la variable de sortie en dehors de cette gamme de données ayant servi pour l'apprentissage. Dans notre approche, une des particularités est que les modèles apprennent une dynamique reliant de façon synthétique les données d'entrée et de sortie, en ajustant des équations d'évolution dépendantes de paramètres. Une fois les paramètres appris, cette dynamique peut ensuite être appliquée en dehors et même plus loin de la période d'apprentissage.

Cette capacité d'extrapolation permet de réduire la quantité de données à collecter et ainsi de réduire les coûts associés aux expérimentations et à la gestion des données. Elle permet également de rallonger la durée de validité des prédictions fournies par l'outil de prédiction. Ainsi, dans le cadre de l'Assimilation de Données, les corrections des prédictions via l'utilisation de données pourraient être moins fréquentes et les coûts de calculs pourraient être plus faibles.

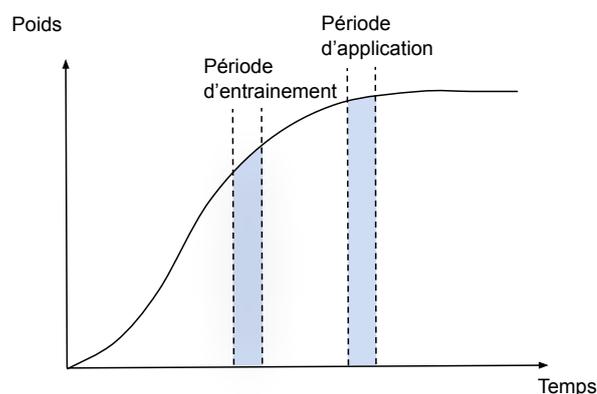


FIGURE 1 – Schématisation de la capacité d'extrapolation des outils développés

La seconde contribution de notre exploration est le développement d'un concept permettant de dissocier la réalité et le modèle mathématique. Dans la plupart des cas, l'objectif de la modélisation mathématique est de traduire la réalité en adoptant un niveau d'abstraction plus ou moins élevé. Dans notre approche, le support de réflexion utilisé pour construire le modèle mathématique n'est pas directement l'animal réel mais un Avatar, qui synthétise de façon conceptuelle la dynamique interne globale de l'animal. L'ensemble de cette dynamique est réduit à la convection, la diffusion et à l'action d'une information globale. Nous avons ensuite construit le système d'EDP traduisant mathématiquement ce flux d'information.

De plus, les applications de notre approche sur des données réelles ont montré que nos outils sont capables de mettre en lien avec précision des données d'élevage, même s'ils sont ajustés sur très peu de données.

Le Couplage Modèle-Données est aujourd'hui essentiellement utilisé dans les domaines de la météorologie (HOUTEKAMER et MITCHELL (1998), KALNAY (2003), MERTENS *et al.* (2009) et WANG *et al.* (2000)), de l'océanographie (CHAVE *et al.* (2009), DOWD (2007), GREGG *et al.* (2009), LGUENSAT *et al.* (2019) et ANDERSON *et al.* (2000)) et de la géologie (NIU *et al.* (2014), GAO *et al.* (2011) et WILLIAMS *et al.* (2005)). L'utilisation de cette approche émergente pour traiter des données biologiques peut également être considérée comme une autre contribution de nos travaux.

Le premier chapitre de ce manuscrit sera consacré à une présentation de l'état de l'art en ce qui concerne l'utilisation de la modélisation mathématique dans le domaine de la production animale. Ce chapitre sera l'occasion de rendre compte de l'intérêt

d'avoir recours à la modélisation mathématique pour traiter des problèmes liés à de la biologie animale mais également les difficultés rencontrées. Dans ce premier chapitre, les principales méthodes de modélisation seront présentées, ainsi que les notions de "Couplage Modèle-Données" et "d'Assimilation de Données". Au regard des méthodes classiques de modélisation mathématique, nous justifierons le choix de l'approche explorée. Ce premier chapitre contiendra également une description des données utilisées tout au long de ces travaux de thèse.

Le deuxième chapitre détaillera la construction et l'analyse d'outils d'apprentissage statistique Biomimétiques. Dans ce chapitre nous décrivons les EDP construites et l'ensemble des tests nous ayant permis de mettre en évidence le bon fonctionnement et la qualité de nos outils de prédiction. Nous commencerons par présenter les tests par simulation réalisés sur des données simulées. Ces premiers tests nous ont permis d'établir les relations existantes entre certains paramètres du modèle, de tester la capacité de nos outils à apprendre la valeur de ces paramètres sur des données bruitées et de mesurer l'impact du bruit présent dans les données sur l'ajustement du modèle. La capacité d'extrapolation de nos outils a également été mise en évidence grâce à ces tests par simulation. Dans ce second chapitre nous présenterons également une application de notre approche sur des données réelles. L'objectif de cette application a été de tester la capacité d'un outil d'apprentissage statistique biomimétique à prédire le poids d'un animal à un instant donné, à partir de son poids initial et de sa consommation alimentaire. Cette application a été l'occasion de valider les performances de prédiction et d'extrapolation de nos outils sur des données réelles.

Le troisième chapitre de ce manuscrit sera consacré à l'étude de l'existence et de l'unicité d'un des systèmes d'EDP construits. Cette étude a révélé certaines particularités du modèle mathématique construit, comme notamment l'écriture de l'opérateur de diffusion. Ces particularités nous ont empêchées d'appliquer les méthodes classiques utilisées dans l'analyse des équations paraboliques. De ce fait, deux versions simplifiées du système d'EDP initial ont été étudiées de façon à identifier les propriétés perdues et celles conservées, malgré la structure de l'opérateur de diffusion utilisé.

Un quatrième chapitre présentera une étude concernant la manière dont certaines variables biologiques, ou facteurs extérieurs, peuvent être pris en compte et intégrés dans des outils d'apprentissage statistique biomimétiques. Lors de cette étude nous avons cherché à valoriser l'information concernant l'âge de l'animal et la composition de la ration alimentaire. Pour cela nous avons construit trois autres systèmes d'EDP,

inspirés de celui présenté dans le second chapitre.

Dans le cinquième chapitre nous décrivons l'ensemble des études nous ayant permis de définir le domaine d'application de nos outils d'apprentissage statistique biomimétiques et de les positionner par rapport aux outils existants. Dans ce chapitre nous présenterons les études qui ont été réalisées afin d'établir l'évolution du temps d'ajustement des paramètres en fonction du nombre de flux d'information pris en compte dans le modèle et du nombre de paramètres à apprendre. Ces études ont permis de prendre conscience de l'impact de la structure du modèle mathématique sur le temps d'ajustement global de l'outil de prédiction embarquant ce modèle. Nous allons également comparer les performances d'un de nos outils avec celles de modèles de régressions linéaires, de modèles logistiques, de réseaux de neurones, de forêts aléatoires et de modèles mécanistes réalistes.

Enfin, un chapitre conclusif présentera les perspectives et les pistes de recherche amenées par ces travaux de thèse. Nous avons notamment entamé une réflexion concernant l'élaboration d'une procédure de sélection de modèles permettant de déterminer le niveau d'abstraction du modèle mathématique apportant la meilleure qualité de prédiction sur une certaine base de données. Nous avons également commencé à réfléchir à une méthode permettant de réaliser de l'Assimilation de Données basée sur des outils d'apprentissage statistique biomimétiques. Ces deux perspectives de recherche seront donc décrites dans ce dernier chapitre.

UTILISATION D'OUTILS MATHÉMATIQUES ET STATISTIQUES DANS LE DOMAINE DE LA PRODUCTION ANIMALE : ÉTAT DE L'ART ET APPROCHE EXPLORÉE

D'après WILSON et DAINTY (1999), la modélisation mathématique peut être décrite comme étant la construction d'une représentation simplifiée d'une partie du monde réel à partir d'un diagramme schématisant avec plus ou moins de réalisme l'objet réel étudié. Le diagramme imaginé est ensuite traduit mathématiquement.

Selon SOLTANI (2012) la procédure de modélisation mathématique consiste tout d'abord à formuler mathématiquement le problème étudié, à programmer ce modèle mathématique, à estimer ensuite la valeur des paramètres qu'il contient et enfin à évaluer la qualité du modèle paramétré.

Comme l'expliquent VÁZQUEZ-CRUZ *et al.* (2014), les mathématiques correspondent à un langage précis et concis permettant de formuler des idées que les ordinateurs peuvent ensuite utiliser pour réaliser des calculs numériques. Les mathématiques couplées à l'utilisation d'une puissance de calcul nous offre donc un cadre riche et bien défini pour pouvoir élaborer des modèles, puis des outils, issus à l'origine de la conceptualisation d'un objet réel.

Les organismes vivants et notamment les animaux, sont des organismes complexes dans lesquels se déroule un grand nombre de phénomènes biologiques. Construire des outils capables de prédire l'effet de facteurs extérieurs sur ces organismes est de ce fait une tâche difficile. Néanmoins, comme nous l'avons vu dans l'introduction cette tâche présente d'importants enjeux économiques et stratégiques pour les acteurs du secteur agroalimentaire.

1.1 Les intérêts et les difficultés de la modélisation mathématique dans le domaine de la production animale.

D'après TOMLIN et AXELROD (2007), depuis des décennies les biologistes utilisent des modèles mathématiques pour comprendre et extraire les mécanismes et les dynamiques se cachant derrière leurs observations expérimentales. Et aujourd'hui, d'après DUMAS *et al.* (2008), MCPHEE (2009a) et VÁZQUEZ-CRUZ *et al.* (2014), la modélisation mathématique reste un processus décisif pour décrire et simuler les mécanismes et les liens existants entre les facteurs régissant la dynamique d'un système biologique.

Les systèmes agricoles sont notamment des systèmes complexes et divers dans lesquels les problématiques biologiques, économiques et écologiques sont multiples. La modélisation mathématique de ces systèmes et de leurs sous-ensembles est donc un outil intéressant pour fournir des représentations simplifiées permettant de mettre en évidence les facteurs clés de ces systèmes et leurs impacts (MCPHEE (2009b) et COMTE (2017)).

Dans leur historique, DUMAS *et al.* (2008) expliquent que la construction de modèles mathématiques pour traiter des problèmes liés à la production animale, a commencé entre 1910 et 1925. Comme cela est expliqué dans les travaux de MCPHEE (2009a), PUILLET *et al.* (2011), MARTIN et SAUVANT (2010b), NKRUMAH *et al.* (2007), NESETRILOVA (2005) et BASARAB *et al.* (2003a), dans le domaine de l'agroalimentaire, simuler et prédire l'impact de l'alimentation sur les performances des animaux sont deux compétences décisives et stratégiques pour les éleveurs, comme pour les entreprises de ce secteur. Lenjeu pour ces acteurs de l'agronomie est d'être capable de comprendre et d'anticiper les mécanismes et les facteurs permettant d'optimiser l'efficacité alimentaire des animaux, c'est-à-dire le ratio entre la quantité d'aliments consommée et la production de viande ou de lait (BRAZDIL et CLARK (1990), PUILLET *et al.* (2011), CASTILLO *et al.* (2000), NAHM (2002), BASARAB *et al.* (2003b) et HERD *et al.* (2004)).

Cette capacité à utiliser efficacement les nutriments d'une ration semble être influencée par de multiples facteurs et processus biologiques (DINIZ *et al.* (2018)). Néanmoins le développement récent de nouvelles technologies permettant de monitorer avec précision la consommation alimentaire des animaux, ainsi que l'évolution de leur

performances a permis de faciliter l'étude de l'efficacité alimentaire des animaux d'élevage (ARCHER *et al.* (1999)).

En effet, les outils de mesure permettant de collecter des données sur des êtres vivants évoluent. JEMILA et PRIYADHARSINI (2018), MIEKLEY *et al.* (2012), TOL et KAMP (2010), BÜCHEL et SUNDRUM (2014) et HOLMAN *et al.* (2011) présentent de nouvelles technologies destinées à monitorer les animaux avec précision. Il existe notamment aujourd'hui des colliers connectés, essentiellement prévus aujourd'hui pour les bovins, qui permettent de collecter des données concernant le temps que les animaux passent à se déplacer et à ruminer. Ces colliers peuvent être couplés à d'autres objets connectés afin de collecter des données dans l'ensemble du bâtiment d'élevage. Par exemple, des outils tels que des auges peseuses, des balances, ou des robots de traite, permettent de collecter des données en lien avec la consommation alimentaire et la production des animaux. D'autres outils comme les bolus peuvent être ingérés par l'animal et permettre de collecter des données concernant l'évolution de variables biologiques internes telles que le pH ou encore la température du rumen. Mais ces technologies peuvent être considérées comme intrusives, leur durée de vie est assez courte et certaines d'entre elles restent chères. Leur utilisation lors des essais expérimentaux est donc encore peu fréquente et elle se fait sur peu d'individus.

De plus, comme cela est illustré dans les travaux de LOCKE *et al.* (2005) et QI *et al.* (2006), les données collectées à partir d'organismes vivants contiennent généralement une quantité importante de variabilité. Cette variabilité correspond en partie aux différences existantes entre les individus. Mais elle résulte également du bruit induit par les outils de mesure utilisés au cours de la collecte de données. En effet, les outils de mesure tels que des balances, des thermomètres ou encore des bolus, peuvent introduire des erreurs dans les données par manque de précision ou dû à un mauvais calibrage (BRAZDIL et CLARK (1990)).

Par ailleurs, les données biologiques utilisées pour ajuster ou corriger les modèles sont souvent collectées au cours de différents essais expérimentaux. Lors de ces différentes expérimentations le protocole appliqué pour collecter les données n'est pas toujours rigoureusement le même. Le manque de continuité qui existe parfois entre les essais expérimentaux introduit également de l'hétérogénéité dans les données.

À cette variabilité s'ajoute généralement une quantité plus ou moins importante de données aberrantes ou manquantes.

La biologie animale reste donc aujourd'hui un contexte dans lequel nous devons construire et ajuster, à partir de peu de données exploitables, des outils de simulation et de prédiction précis (PHUA *et al.* (2017) et MOURET (2016)).

Selon INGALLS (2013), tous les modèles sont par définition une abstraction d'un objet réel. Mais ces modèles peuvent être statiques ou dynamiques, déterministes ou stochastiques, paramétriques ou non et présenter un niveau d'abstraction plus ou moins élevé.

1.2 Les différents types de modèles mathématiques et les différentes méthodes d'apprentissage

L'apprentissage et l'ajustement de modèles mathématiques à partir de données est un concept très vaste comportant une large gamme de méthodes (HASTIE *et al.* (2009)). Deux grandes méthodes d'apprentissage peuvent être distinguées : L'apprentissage supervisé et non-supervisé (VAPNIK (1995)).

1.2.1 L'apprentissage supervisé et non-supervisé

L'apprentissage supervisé correspond à un problème d'apprentissage dans lequel l'utilisateur cherche à prédire une sortie spécifique, à partir de données d'entrée (HASTIE *et al.* (2009)). Par exemple, les méthodes de régression et de classification dont les classes attendues sont précisées sont des méthodes d'apprentissage supervisé.

Dans un problème d'apprentissage non-supervisé, il n'y a pas de sortie prédéfinie (GHAHRAMANI (2003)). L'objectif est alors de mettre en évidence la structure des données de manière exploratoire et sans faire d'hypothèse a priori. Les principales méthodes d'apprentissage non-supervisé correspondent aux méthodes de regroupement (clustering) et de réduction de dimensions (EVERITT *et al.* (2011) et LU *et al.* (2013)).

En biologie et notamment en biologie animale, le jeu est souvent de déterminer les liens qui existent entre des variables, à partir de données d'entrée et de sortie (INGALLS (2013), MARTIN et SAUVANT (2007) et NESETRILOVA (2005)). Nous tra-

vaillons donc dans un contexte dans lequel l'apprentissage est essentiellement supervisé. Nous allons de ce fait étudier ici les modèles principalement utilisés dans le cadre de l'apprentissage supervisé.

D'après KITANO (2002), établir clairement les objectifs de modélisation est une étape décisive. Avoir recours à la modélisation pour améliorer la compréhension des processus à l'origine de la dynamique d'un système, ne nécessitera pas les mêmes besoins en terme d'architecture du modèle que si l'objectif est de réaliser de l'Assimilation de Données ou de la prédiction en temps réel. Ainsi les modèles mathématiques construits peuvent être plus ou moins réalistes ou plus ou moins conceptuels selon la quantité de données disponibles, les objectifs de modélisation, mais aussi selon le niveau de connaissances acquises concernant la dynamique du système étudié.

Les méthodes d'apprentissage supervisé peuvent être réparties en deux catégories : les méthodes paramétriques et non-paramétriques (JAMES *et al.* (2013)).

1.2.2 Les modèles paramétriques et non-paramétriques

Dans les approches d'apprentissage paramétrique, l'apprentissage se base sur un modèle contenant des paramètres. Pendant la phase d'apprentissage l'objectif est donc dans ce cas, de déterminer la valeur des paramètres minimisant une fonction de coût (JAMES *et al.* (2013)). Cette fonction de coût, appelée aussi fonction "objectif", est généralement une différence calculée entre des données d'apprentissage empiriques et les prédictions fournies par le modèle (VAPNIK (1995) et HASTIE *et al.* (2009)).

Dans les approches d'apprentissage non-paramétrique aucune hypothèse n'est faite sur la fonction reliant les données d'entrée (X) et les données de sortie (Y). L'objectif de la phase d'apprentissage est donc de déterminer la fonction f telle que : $Y = f(X) + \epsilon$, où ϵ correspond aux résidus ou à l'erreur du modèle et f la fonction minimisant la fonction de coût (JAMES *et al.* (2013)).

D'après JAMES *et al.* (2013), les approches paramétriques simplifient les problèmes d'apprentissage. En effet, il est généralement plus facile d'estimer la valeur des paramètres d'une fonction donnée plutôt que d'estimer l'allure de la fonction toute entière. En revanche l'approche paramétrique nécessite de faire des hypothèses concernant

la fonction f et de construire un modèle. Si ce modèle est trop éloigné de la réalité ou qu'il est trop limité pour pouvoir simuler et résumer la dynamique du système étudié, l'apprentissage des paramètres ne pourra jamais aboutir à une prédiction satisfaisante. Le choix et l'élaboration du modèle est donc une étape décisive dans les approches d'apprentissage paramétrique.

1.2.3 Des modèles statiques et des modèles dynamiques

Les modèles mathématiques peuvent être statiques ou dynamiques, c'est-à-dire qu'ils peuvent être indépendants ou dépendants du temps (GUZMÁN-CRUZ *et al.* (2011))

Les modèles statiques sont utilisés pour mettre en lien des variables dont la relation ne dépend pas du temps. Ils peuvent par exemple être utilisés pour prédire l'impact de la génétique sur la qualité de la viande (LENSINK et GUERMONPREZ (2003)) ou sur les capacités de reproduction des animaux (LENSINK et GUERMONPREZ (2003)).

Néanmoins, les modèles construits pour traiter des problèmes liés à la biologie sont souvent des modèles dynamiques puisqu'ils modélisent la plupart du temps l'évolution dans le temps de variables biologiques (BASTIANELLI et SAUVANT (1997)).

Ces modèles dynamiques peuvent être basés sur des Équations Différentielles décrivant l'évolution de variables continues. Les Équations Différentielles Ordinaires correspondent à la forme d'Équations Différentielles la plus utilisée pour construire des modèles biologiques dynamiques (CHASSAGNOLE *et al.* (2002), TYSON *et al.* (2003), RIZZI *et al.* (1997), CHEN *et al.* (1999) et ZÚÑIGA *et al.* (2014)). Néanmoins des EDP peuvent également être utilisées pour décrire plus finement la cinétique d'un processus dans le temps et dans l'espace (GETTO et WAURICK (2016), FLEGG *et al.* (2012) et WAYNE et DAVID (2007)).

Dans les modèles mathématiques décrivant des processus biologiques, les Équations Différentielles servent essentiellement à modéliser des flux de matière, de molécules ou d'énergie, entre des compartiments schématisant des organes ou des tissus biologiques (BASTIANELLI et SAUVANT (1997)).

1.2.4 Des modèles déterministes et des modèles stochastiques

Les modèles mathématiques paramétriques peuvent être divisés en deux autres classes : Les modèles déterministes et les modèles stochastiques.

Par définition, les paramètres des modèles déterministes sont des constantes fixées une fois leur valeur déterminées ou optimisées. Ainsi, à partir des mêmes données, un modèle déterministe fournira les mêmes sorties.

En revanche, les modèles stochastiques sont des modèles probabilistes générant de l'aléa lors de chaque simulation, via l'intégration dans le modèle de variables aléatoires. D'après ALFONSI *et al.* (2005), les modèles stochastiques construits pour analyser l'évolution de phénomènes se basent sur le constat que ces phénomènes sont des processus aléatoires et que pour les simuler de façon réaliste il est nécessaire d'avoir des modèles reproduisant l'existence de cet aléa.

Ainsi, les outils de simulation réalistes sont souvent basés sur des modèles dynamiques et stochastiques (MARIA (1997)) alors que les outils prédictifs sont souvent basés sur des modèles dynamiques déterministes (VÁZQUEZ-CRUZ *et al.* (2014) et GONZE *et al.* (2003)).

1.2.5 Des modèles «boîtes-noires» et des modèles "boîtes-blanches"

D'après VÁZQUEZ-CRUZ *et al.* (2014), parmi les différentes méthodes existantes pour analyser des données biologiques, deux grandes approches peuvent être distinguées : Les modèles dits «boîtes-noires» et les modèles dits «boîtes-blanches».

Les modèles «boîtes-noires»

Les modèles d'apprentissage automatique, aussi appelés modèles de «Machine Learning», comme les réseaux de neurones, sont souvent appelés «boîtes-noires». Mais d'après VÁZQUEZ-CRUZ *et al.* (2014), les modèles «boîtes-noires» regroupent également les régressions linéaires simples et multiples, ainsi que les modèles logistiques tels que les modèles de GOMPERTZ (1825) et de VERHULST (1838).

Ces modèles sont plus ou moins complexes en termes d'architecture. Comme l'apprentissage automatique est considéré comme une technique d'Intelligence Artificielle, l'apprentissage profond, aussi appelé "Deep Learning" est souvent considéré comme une sous partie de l'apprentissage automatique (LECUN *et al.* (2015), BENGIO (2012) et AREL *et al.* (2010)). Globalement, lorsque la structure des algorithmes se complexifie et que le nombre de niveau de représentation augmente, on parle d'apprentissage profond. En effet, les algorithmes d'apprentissage profond se composent de plusieurs sous modules non-linéaires (BENGIO (2012), AREL *et al.* (2010) et CHO et SAUL (2009)). L'objectif de ces outils est d'explorer de manière plus approfondie les informations contenues dans les données (LECUN *et al.* (2015), BENGIO (2012) et GLOROT *et al.* (2011)).

Les algorithmes d'apprentissage automatique correspondent à une description directe des données. L'objectif de ces modèles est de mettre en lien des données d'entrée et de sortie sans avoir besoin de connaissances concernant les phénomènes biologiques sous-jacents.

Comme cela est expliqué dans les travaux de DOMINGOS (2012), depuis une dizaine d'années, l'utilisation d'algorithmes d'apprentissage automatique et plus particulièrement de réseaux de neurones a fortement augmenté. D'après GORCZYCA *et al.* (2018), VALLETTA *et al.* (2017), MA *et al.* (2014) et IP *et al.* (2018), la popularité de ces modèles peut s'expliquer par leur facilité d'utilisation, ainsi que par la diversité des problèmes pouvant être résolus grâce à de tels algorithmes.

Néanmoins ces modèles d'apprentissage automatique contiennent peu de connaissances a priori. Les outils basés sur ce genre de modèles sont donc peu informatifs d'un point de vue biologique (TAN et GILBERT (2003), SHAVLIK *et al.* (1995), HUBBARD et REINHARDT (1998)). De plus, DUMPALA *et al.* (2017) expliquent qu'il est nécessaire d'apprendre les paramètres de ces modèles à partir de beaucoup de données afin de compenser cette absence d'expertise biologique a priori.

De plus, la structure de ces modèles, décorrélée de la nature des processus biologiques reliant les variables étudiées, est par conséquent très générique et peut être appliquée pour traiter des problèmes pouvant être de natures très différentes (HASSOUN (1995)). Mais cette généricité limite l'adaptabilité de ces modèles à la dynamique de certains processus biologiques (LU *et al.* (1996)).

Ces outils "boîtes-noires" sont donc des outils très utiles lorsque l'objectif est de construire des outils de prédiction ou de simulation mettant en lien de nombreuses variables, sans volonté d'expliquer des phénomènes régissant l'évolution de ces variables. Ils sont également utiles lorsque l'utilisateur dispose de peu de connaissances concernant les processus étudiés mais qu'il dispose en revanche de beaucoup de données pour ajuster les paramètres de ces modèles (HASSOUN (1995)).

Les modèles "boîtes-blanches"

D'après VÁZQUEZ-CRUZ *et al.* (2014), les modèles "boîtes-blanches" correspondent à des modèles mécanistes décrivant les processus physiques qui se déroulent au sein du système étudié avec une volonté de réalisme et souvent dans un objectif bien précis. Dans le cadre de problèmes liés à la biologie animale, ils se placent de ce fait à un niveau de description plus fin et plus précis que l'animal complet.

Cette approche intègre beaucoup de connaissances et nécessite donc une bonne compréhension a priori des processus régissant la dynamique du système étudié.

À l'image des travaux réalisés par BASTIANELLI et SAUVANT (1997) et MARTIN et SAUVANT (2010a), la construction de modèles mathématiques mécanistes réalistes est une tâche complexe menant généralement à un modèle contenant beaucoup d'équations et de paramètres à déterminer.

Ces modèles complexes permettent de réaliser des expérimentations "In Sillico" à partir de données d'entrée bien précises. En effet, les modèles mécanistes réalistes décrivent des processus biologiques précis et dont l'expression mathématique requiert la valeur de paramètres spécifiques (concentrations sanguine de certaines molécules, coefficients de transfert de molécules entre deux organes, *etc.*). Ces modèles se basent donc sur de nombreuses variables d'entrée, spécifiques au problème étudié. Ce grand nombre d'informations à préciser en entrée rend l'implémentation de ces modèles compliquée, surtout lorsque ces informations sont difficiles à collecter (BASTIANELLI et SAUVANT (1997)).

L'architecture complexe des modèles «boîtes-blanches» les rend peu adaptés à l'optimisation des paramètres sur des données collectées et de ce fait à la réalisation de prédiction et d'Assimilation de Données d'élevage (BASTIANELLI et SAUVANT (1997)).

et MARTIN et SAUVANT (2010a)). La valeur des paramètres est donc généralement obtenue grâce à la littérature et à des expérimentations précises (prises de sang, dosages, *etc.*). Les valeurs manquantes sont alors souvent estimées ou extrapolées à partir d'autres valeurs connues.

Les modèles mécanistes réalistes sont donc des outils utiles lorsque l'utilisateur dispose déjà d'une bonne connaissance de la dynamique du système étudié mais qu'il souhaite visualiser les réactions du système lors de différentes simulations quantitatives et valider certaines hypothèses (BASTIANELLI et SAUVANT (1997) et BALDWIN (1995)).

Pour résumer, la construction des modèles «boîtes-blanches» nécessite d'avoir beaucoup de connaissances concernant le système étudié et donne naissance à des modèles complexes contenant beaucoup de paramètres à déterminer. De l'autre côté, l'ajustement des modèles «boîtes-noires» nécessite d'avoir beaucoup de données. Or lors de l'étude de problématiques liées à de la biologie animale, il arrive encore aujourd'hui que l'on ne dispose ni de beaucoup de données, ni de beaucoup de connaissances a priori ou que l'on ait besoin d'un modèle plus léger qu'un modèle réaliste. Il est donc nécessaire d'explorer des approches intermédiaires pour construire des outils parcimonieux capables de s'ajuster à partir de peu de données.

1.3 Émergence d'approches intermédiaires

1.3.1 L'apprentissage statistique

L'apprentissage statistique : une méthode de modélisation issue des statistiques

D'après JAMES *et al.* (2013) et DUNSON (2018), les méthodes d'apprentissage statistiques, appelées approches de «Statistical Learning», ont fortement évolué ces dernières années sous l'influence notamment de l'évolution en parallèle de l'apprentissage automatique. L'apprentissage statistique est donc une approche développée récemment (HASTIE *et al.* (2009) et JAMES *et al.* (2013)).

L'apprentissage automatique et l'apprentissage statistique sont deux approches

ayant pour objectif d'apprendre des coefficients de pondérations ou des paramètres à partir de données. Mais ces deux approches sont issues de deux visions différentes.

En effet, l'apprentissage automatique, issu de l'informatique, vise la rapidité d'exécution et la qualité de prédiction, en faisant le moins d'hypothèses possible et sans volonté d'expliquer les liens existants entre les variables (DUNSON (2018) et VALLETTA *et al.* (2017)).

D'autre part, l'apprentissage statistique, issu des statistiques, se base davantage sur l'intégration d'hypothèses ou de connaissances concernant la distribution des données et les relations existantes entre les variables (BOUSQUET *et al.* (2011) et VALLETTA *et al.* (2017)). Cette approche nécessite donc plus de connaissances a priori concernant la dynamique du système étudié (BARTOCCI *et al.* (2014)).

Certains auteurs comme JAMES *et al.* (2013) considèrent que les méthodes d'apprentissage statistique peuvent être réparties en deux catégories : les méthodes paramétriques et non-paramétriques.

Comme nous l'avons vu dans la section 1.2.2, l'apprentissage statistique non-paramétrique consiste alors à trouver la fonction f reliant les données d'entrée et de sortie et minimisant la fonction de coût (HASTIE *et al.* (2009), VAPNIK (1995), JAMES *et al.* (2013) et FOKOUÉ (2014)).

L'apprentissage statistique paramétrique nécessite la construction d'un modèle mathématique capable de mettre en lien des données d'entrée et des données de sortie (KOTSIANTIS *et al.* (2007) et BOUSQUET *et al.* (2011)). Dans ce cas le problème d'apprentissage revient alors à déterminer la valeur des paramètres du modèle minimisant la fonction de coût utilisée (JAMES *et al.* (2013)).

D'après KOTSIANTIS *et al.* (2007) et INIESTA *et al.* (2016), la construction d'un modèle intégrant des hypothèses ou des connaissances, fait de l'apprentissage statistique paramétrique un cadre adapté pour modéliser et comprendre des données complexes.

L'apprentissage statistique paramétrique : Intégrer des hypothèses pour étudier des phénomènes complexes et réduire la dépendance aux données

D'après MOURET (2016), PHUA *et al.* (2017) et MORIK *et al.* (1999), la prise en compte dans l'outil de prédiction d'hypothèses concernant la structure des données

permet de gagner en précision, tout en permettant aux outils d'apprendre plus rapidement et à partir d'une moins grande quantité de données .

Or comme nous l'avons vu précédemment dans la section 1.1, la biologie, et tout particulièrement la biologie animale, est un contexte dans lequel les scientifiques sont amenés à construire des outils pour prédire des phénomènes complexes à partir de peu de données exploitables. Ce contexte explique que l'apprentissage statistique paramétrique joue depuis quelques années un rôle important dans les domaines de la science, de l'agriculture et de l'industrie (HASTIE *et al.* (2009)).

MORIK *et al.* (1999), PHUA *et al.* (2017), LEE (2004), DOMINGOS et RICHARDSON (2007) et WANG *et al.* (2006), ont exploré des approches d'apprentissage statistique dirigées par les données («Data-Driven») mais en intégrant des connaissances concernant le domaine d'étude dans le modèle construit. L'intégration de connaissances concernant la distribution des données, mais aussi concernant la dynamique du système étudié semble être un levier intéressant pour réduire la dépendance aux données d'apprentissage et améliorer la prédiction de phénomènes complexes. D'après RENZULLO *et al.* (2008), ce qui peut aujourd'hui permettre d'améliorer la précision des outils de prédiction est une combinaison optimale entre l'utilisation de données expérimentales et de modèles mathématiques.

Or depuis quelques années on assiste à l'émergence d'une approche de modélisation qui consiste à intégrer des connaissances dans des outils basés sur des modèles mathématiques dynamiques et destinés à la réalisation de prédiction et d'Assimilation de Données : le «Couplage Modèle-Données».

1.3.2 Le Couplage Modèle-Données

Le Couplage Modèle-Données est une approche qui se situe à mi-chemin entre les approches "boîtes-noires" et "boîtes-blanches". En effet le Couplage Modèle-Données consiste à intégrer des connaissances concernant la dynamique du système étudié via la construction d'un modèle mathématique et à utiliser ensuite des données pour ajuster les paramètres du modèle construit.

Le Couplage Modèle-Données est aujourd'hui surtout utilisé dans le domaine de la météorologie (SIMMONS et HOLLINGSWORTH (2002)), de l'hydrologie (KIM et BAR-

ROS (2002), CROSSON *et al.* (2002) et MACKAY *et al.* (2003)), de la biogéochimie (BARRETT (2002), BARRETT *et al.* (2005), RAYNER *et al.* (2005) et SACKS *et al.* (2006)) et de l'océanographie (AILLIOT *et al.* (2006)). Tout comme la biologie, ces domaines sont des secteurs dans lesquels il est nécessaire de prendre en compte au moins certains aspects de la dynamique du système étudié pour réaliser des prédictions précises. Seulement, le système étudié est souvent très complexe et une modélisation réaliste de celui-ci prendrait du temps et aboutirait à la construction d'un modèle mathématique volumineux. Ainsi, comme cela est décrit dans les travaux de FRÉNOT (2017), ROUSSEAU et NODET (2013), SACKS *et al.* (2007) et WANG *et al.* (2010), l'approche du Couplage Modèle-Données consiste à construire un modèle mathématique parcimonieux, correspondant à une synthèse mathématique du système étudié. Le modèle mathématique obtenu contient ainsi un nombre limité de paramètres qui sont ensuite ajustés à partir de données collectées.

Dans cette approche le niveau d'abstraction du modèle mathématique est donc choisi de façon à extraire des données le plus d'informations possibles tout en gardant un outil souple et léger, capable d'être ajusté sur un jeu de données.

L'ajustement du modèle passe par l'optimisation de la valeur des paramètres du modèle qui se base sur la minimisation d'une fonction de coût (RENZULLO *et al.* (2008)).

Certaines études ont récemment exploré l'utilisation de modèles biomimétiques contenant des paramètres ajustés à partir de données expérimentales, pour traiter des problématiques liées à de la biologie. C'est le cas par exemple des études réalisées par ZÚÑIGA *et al.* (2014) et DA ROS *et al.* (2013). Mais ces modèles contiennent encore un nombre important d'équations et de paramètres à déterminer, ce qui peut rendre l'optimisation des paramètres complexe lorsqu'il y a peu de données disponibles.

Un des objectifs de notre approche a donc été de construire des modèles plus parcimonieux, c'est-à-dire un modèle capable de fournir des prédictions précises tout en contenant un faible nombre d'équations et de paramètres à déterminer.

1.4 Réapprentissage et Assimilation de Données

Parmi les méthodes permettant l'intégration et la prise en compte de nouvelles données collectées nous distinguons ici la notion de «réapprentissage» et d'«Assimilation de Données».

Le réapprentissage sera considéré ici comme l'actualisation, à partir d'un nouveau jeu de données, d'un modèle ajusté ultérieurement sur un historique de données plus ancien. Cette actualisation permet de corriger et d'améliorer les prédictions fournies par le modèle.

Comme l'ajustement, l'actualisation du modèle se fait généralement par le réapprentissage des paramètres de celui-ci, via la minimisation d'une fonction de coût (LEWIS *et al.* (2006), NIU *et al.* (2014), ANDERSON *et al.* (2000), NIU *et al.* (2014), LUO *et al.* (2011) et LADEAU *et al.* (2011)).

Nous restreindrons ici la notion d'Assimilation de Données à une méthode visant à estimer l'état d'un système à un instant t , en combinant l'information fournie par un outil de simulation et celle fournie par un flux de données collectées plus ou moins continu.

L'objectif de l'Assimilation de Données est d'utiliser les données pour corriger, de manière itérative et plus ou moins régulière, les simulations du modèle. Cette correction est réalisée en tenant compte du fait que les données contiennent du bruit et que l'outil de prédiction est associé à une erreur plus ou moins importante. Ainsi, dans le cadre de l'Assimilation de Données, plus les données sont faciles à collecter et plus les corrections peuvent être effectuées fréquemment (AUROUX et BLUM (2005), GREGG *et al.* (2009), LGUENSAT *et al.* (2017) et LGUENSAT *et al.* (2019)).

Les méthodes les plus utilisées actuellement pour estimer l'état actuel d'un système à partir des données collectées et de la simulation de l'outil utilisé sont le filtre de Kalman et la méthode de Monte Carlo (WANG *et al.* (2000), ANDERSON *et al.* (2000), LUO *et al.* (2011) et LGUENSAT *et al.* (2017)).

L'Assimilation de Données peut être basée sur un outil de simulation davantage guidé par les données («Data-Driven») ou par un modèle («Modèle-Driven») (LGUENSAT *et al.* (2017)). Dans le cas de l'assimilation dirigée par les données l'outil de simula-

tion utilisé est basé sur un modèle de type «boîtes-noires». Mais comme nous l'avons vu dans la section 1.2.5, l'ajustement de ce type de modèles nécessite de disposer de beaucoup de données. Lorsque l'on souhaite réaliser de l'Assimilation de Données dans des contextes où l'on dispose de peu de données il est nécessaire d'avoir un outil capable de fournir des prédictions précises malgré un ajustement effectué préalablement sur peu de données. Si nous sommes dans un contexte dans lequel la collecte de données de qualité est difficile, il faudra également que cet outil soit robuste et qu'il fournisse des prédictions ne nécessitant pas de corrections trop fréquentes. Dans ce genre de contextes l'Assimilation de Données est donc plutôt dirigée par un modèle mathématique dynamique contenant peu de paramètres à ajuster (NIU *et al.* (2014)).

L'Assimilation de Données dirigée par un modèle peut donc être vue comme une méthode couplant l'utilisation de données et de connaissances pour estimer l'état actuel d'un système. Basé sur le même principe, le Couplage Modèle-Données peut alors être utilisé pour construire l'outil fournissant les simulations (HOUTEKAMER et MITCHELL (1998), GREGG *et al.* (2009), LUO *et al.* (2011) et NIU *et al.* (2014)).

Comme le Couplage Modèle-Données, l'Assimilation de Données dirigée par un modèle est aujourd'hui essentiellement utilisé dans les domaines de la météorologie (HOUTEKAMER et MITCHELL (1998), KALNAY (2003), MERTENS *et al.* (2009) et WANG *et al.* (2000)), de l'océanographie (HAVE *et al.* (2009), DOWD (2007), GREGG *et al.* (2009), LGUENSAT *et al.* (2019) et ANDERSON *et al.* (2000)) et de la géologie (NIU *et al.* (2014), GAO *et al.* (2011) et WILLIAMS *et al.* (2005)).

L'enjeu majeur pour réaliser de l'Assimilation de Données biologiques est de disposer d'un modèle relativement léger, capable de fournir des prédictions précises et fiables malgré peu de données disponibles pour l'ajuster et le corriger (LUO *et al.* (2011) et NIU *et al.* (2014)).

1.5 Résumé de l'état de l'art

Différentes notions ont été abordées dans cet état de l'art. L'objectif de cette section est de résumer et de proposer un positionnement de ces différentes notions les unes par rapport aux autres. Ce positionnement proposé est synthétisé par la figure 1.1.

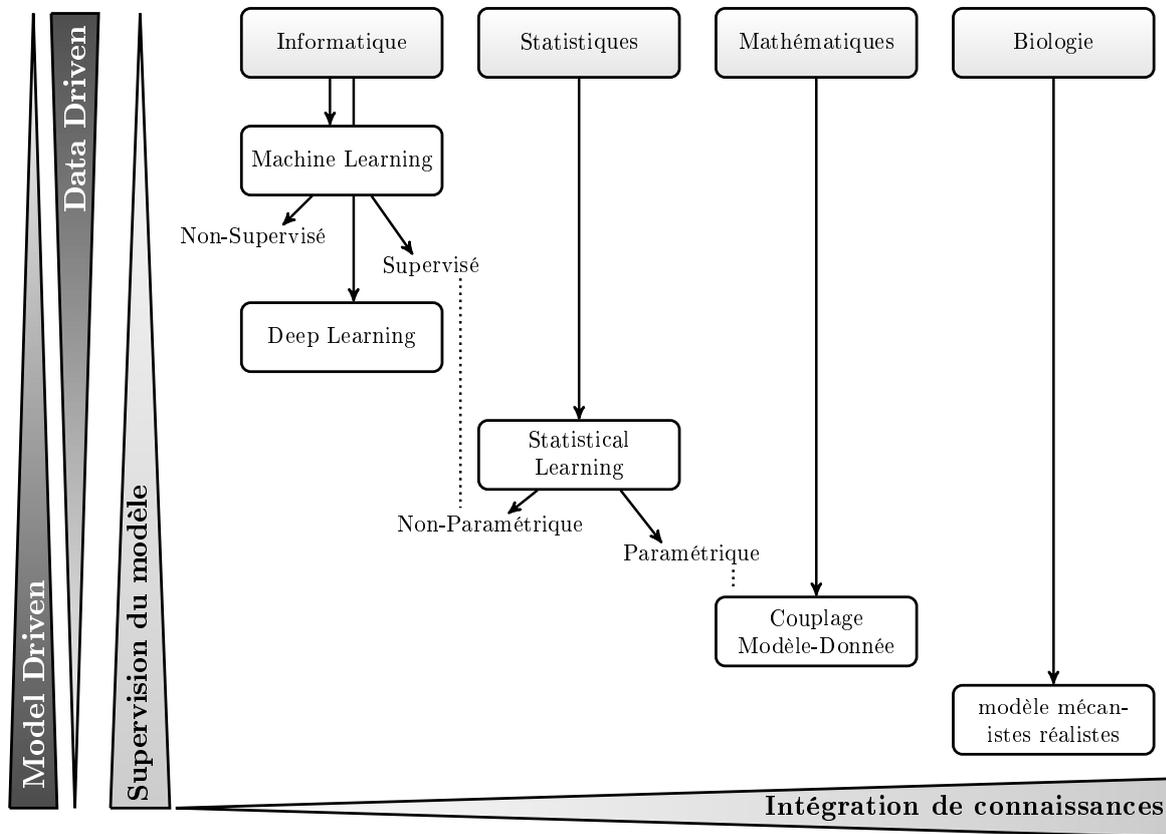


FIGURE 1.1 – Résumé de l'état de l'art.

Des auteurs comme KOTSIANTIS *et al.* (2007) considèrent que parmi les méthodes d'apprentissage automatique, deux catégories peuvent être distinguées : les méthodes supervisées et les méthodes non-supervisées. Comme nous l'avons vu dans la section 1.2.1, l'apprentissage automatique non-supervisé est une approche qui se base totalement sur l'information contenue dans les données («Data-Driven») et dans laquelle aucune hypothèse a priori n'est faite. Dans les approches d'apprentissage automatique supervisé, aucune hypothèse n'est faite concernant la distribution des données, mais les entrées et les sorties attendues sont spécifiées. La construction d'algorithmes d'apprentissage automatique et surtout d'apprentissage profond, nécessite de faire un minimum de choix a priori, notamment le nombre de couches cachées, le nombre de neurones présents dans chaque couche cachée ou encore le nombre d'arbres de dé-

cision utilisés (LECUN *et al.* (2015)). La supervision de la construction de ces modèles n'est donc pas nulle.

D'autre part, comme nous l'avons vu dans la section 1.2.2, l'apprentissage statistique non-paramétrique est une approche dans laquelle aucune hypothèse n'est faite concernant la fonction reliant les données d'entrée et de sortie. Ainsi, l'apprentissage statistique non-paramétrique semble s'approcher de l'apprentissage automatique supervisé en ce qui concerne l'importance accordée à l'information contenue dans les données et la volonté de garder une approche affranchie de toute hypothèse a priori.

Le «Couplage Modèle-Données» est une approche inspirée des mathématiques mais elle implique des études interdisciplinaires. En effet, dans la plupart des explorations basées sur du «Couplage Modèle-Données» il y a une volonté d'intégrer dans le modèle mathématique des connaissances concernant la dynamique globale du système étudié, via le langage mathématique. La construction de ces modèles est donc supervisée et nécessite quelques connaissances concernant le domaine d'étude.

En ce qui concerne la construction de l'architecture des modèles, deux approches peuvent être distinguées d'après TOMLIN et AXELROD (2007). La première, dite «Bottom-Up», consiste à construire un modèle en commençant par représenter les détails des différents composants du système étudié et les liens qui existent entre eux. À l'inverse, la seconde approche, dite «Top-Down», consiste à considérer dans un premier temps le système étudié dans son ensemble et à entrer ensuite plus ou moins dans les détails du fonctionnement du système.

La construction de l'architecture des modèles de type «boîtes-noires» est totalement décorrélée de la nature des processus biologiques étudiés et n'est donc pas concernée par ces deux approches.

La construction des modèles de type «boîtes-blanches» se base par définition sur une approche «Bottom-Up». En effet, ces modèles sont construits dans un objectif très précis et de manière extrêmement structurée et supervisée. Les approches "boîtes-blanches" contiennent beaucoup de connaissances traduites par un grand nombre d'équations. Peu de composants de ces modèles sont ajustés à partir de données et la quasi totalité de l'information est apportée par la construction du modèle.

En revanche, les modèles conçus dans le cadre du Couplage Modèle-Données

sont associés à une construction architecturale de type «Top-Down», afin de garder un modèle synthétisant la dynamique globale du système. L'objectif est d'avoir un modèle qui intègre des connaissances mais qui reste relativement souple et léger afin de pouvoir ajuster les paramètres qu'il contient sur des données. Cette approche peut donc être considérée comme une approche intermédiaire entre les approches «boîtes-noires» et «boîtes-blanches».

Enfin, l'Assimilation de Données et le réapprentissage du modèle correspondent à deux manières différentes de valoriser la collecte de nouvelles données. Mais ces deux pratiques se font à des échelles de temps différentes et dans deux optiques distinctes : La première est utilisée pour corriger de manière itérative, les simulations fournies par le modèle à partir de données collectées quasiment en temps réel et la deuxième est utilisée pour réajuster ponctuellement les paramètres du modèle sur un nouvel historique de données.

L'outil de simulation utilisé pour réaliser de l'Assimilation de Données peut être un outil d'apprentissage automatique, d'apprentissage statistique ou issu du Couplage Modèle-Données. Ainsi l'Assimilation de Données peut donc accorder plus d'importance aux données ou aux connaissances a priori selon le type d'outil de prédiction utilisé.

L'Assimilation de Données d'élevage est un contexte particulier qui peut être qualifié de "Biological Small Data Context". En effet, comme nous l'avons vu dans la section 1.1, les données relatives au fonctionnement interne des êtres vivants sont difficiles à collecter et contiennent généralement beaucoup de variabilité. De plus, les phénomènes à l'origine de l'évolution des variables biologiques sont complexes. Les modèles réalistes mécanistes ne sont pas construits dans l'optique d'être ajustés sur des données. D'autre part, les outils d'apprentissage automatique et d'apprentissage profond ont besoin de beaucoup de données pour pouvoir être ajustés. Dans ce contexte particulier, les outils de prédiction actuellement utilisés ne semblent donc pas adaptés pour gérer conjointement le manque de données, leur hétérogénéité et la complexité des phénomènes étudiés.

Il semble donc nécessaire d'explorer une nouvelle approche pour construire des outils de prédiction précis permettant de réaliser de l'Assimilation de Données d'élevage.

1.6 L'approche explorée dans ces travaux de thèse

Comme nous l'avons vu dans l'introduction, notre exploration se fait dans le cadre d'une thèse CIFRE dans une entreprise d'agroalimentaire. Cette entreprise désire disposer d'outils capables de mettre en lien l'alimentation et les performances d'animaux d'élevage. La prédiction des performances en fonction de la composition de la ration distribuée peut ensuite permettre de déterminer la ration optimisant les performances des animaux dans des conditions d'élevage données. Cette capacité de prédiction peut également permettre d'apporter du conseil aux éleveurs et de les accompagner dans le pilotage de leur élevage. Cependant, cette entreprise dispose de très peu de données pour ajuster et corriger des outils de prédiction et d'Assimilation de Données.

Le premier objectif de ces travaux de thèse a été de construire des outils capables de prédire avec précision l'évolution de variables biologiques malgré un ajustement réalisé sur peu de données.

Le second objectif était que ces outils soient capables de réaliser de l'Assimilation de Données d'élevage. Pour cela il nous a fallu veiller à conserver des outils relativement légers, capables de mettre en lien différents types de données et de fournir des prédictions ne nécessitant pas de corrections fréquentes.

De plus, l'entreprise dans laquelle cette thèse a été réalisée, produit de l'alimentation pour différentes espèces d'élevage (vaches, lapins, cochons.). Nous voulons donc des outils qui puissent s'adapter rapidement aux différentes espèces étudiées.

Pour atteindre ces objectifs, nous avons décidé d'explorer une approche se situant à mi-chemin entre une approche de modélisation réaliste et une approche de type "boîte-noire". Cette approche se base sur l'intégration dans le modèle mathématique de connaissances concernant la dynamique globale du système.

Nous avons fait l'hypothèse qu'un modèle dynamique contenant des connaissances a priori aura une capacité d'apprentissage plus forte et réalisera des prédictions plus précises que celles fournies par des outils de type "boîtes-noires". Nous avons aussi supposé qu'en veillant à conserver un modèle parcimonieux, c'est-à-dire un modèle ne contenant ni trop d'équations ni trop de paramètres à déterminer, cela permettra de conserver un outil léger par rapport à des outils de type mécanistes réalistes. Cette parcimonie permettra également de garder un outil capable de s'adapter relativement ra-

pidement et facilement aux différentes espèces animales étudiées par les entreprises du secteur agroalimentaire.

Comme dans les études explorant le Couplage Modèle-Données citées dans la section 1.3.2, la construction de notre outil est basée sur une combinaison optimale entre les connaissances que nous avons pour construire un modèle mathématique, et les données dont nous disposons pour ajuster les paramètres de ce modèle. L'approche explorée dans ce travail de thèse correspond donc à une approche de type "Couplage Modèle-Données".

Concrètement, nous avons construit un modèle mathématique principal, basé sur un système d'EDP embarquant l'expression mathématique de divers phénomènes, comme par exemple de la diffusion, de la convection et des effets de saturation. Ce modèle mathématique a pour but d'intégrer des connaissances biologiques, et non pas de décrire de façon réaliste et avec exactitude l'ensemble des processus biologiques qui se déroulent dans le corps de l'animal. Dans notre approche nous faisons une distinction entre la réalité et la dynamique traduite mathématiquement par notre modèle.

Ce modèle, même s'il peut être considéré comme biomimétique, il n'est pas "biologique" mais bien "mathématique". L'objectif est de construire un outil de prédiction embarquant une dynamique synthétique, dans l'espoir qu'il fournisse des prédictions précises malgré un apprentissage réalisé sur peu de données. Néanmoins, le modèle mathématique construit n'est pas directement assimilable à l'animal réel. En effet, l'animal réel est un organisme vivant très complexe dont la modélisation réaliste serait très longue. Le support de réflexion que nous avons utilisé pour construire notre modèle n'est donc pas directement l'animal réel, mais un Avatar de cet animal (figure 1.2). Cet Avatar correspond à un résumé conceptuel, réduisant à l'essentiel la dynamique globale des réactions du corps de l'animal. Un grand nombre de phénomènes physico-chimiques se déroulent dans le corps de l'animal en réponse à l'ingestion ou à l'injection de molécules. Ils mènent, quelques temps plus tard, à l'évolution de variables biologiques. Ces apports de molécules dans le corps de l'animal et ces variables biologiques peuvent être enregistrés et captés pour générer des données d'entrée et de sortie. Nous avons supposé que ces données d'entrée et de sortie peuvent être reliées par un modèle dynamique succinct.

Dans l'Avatar nous avons imaginé les flux de molécules et les organes de manière

globale, comme étant des flux d'informations et des compartiments. Nous avons ensuite construit les équations et les opérateurs mathématiques que nous avons jugés nécessaires et suffisants pour traduire mathématiquement cet Avatar et relier les données d'entrée et de sortie citées précédemment. Le principal modèle mathématique obtenu dans ces travaux est un système d'EDP décrivant la convection, la diffusion et l'action d'une information globale.

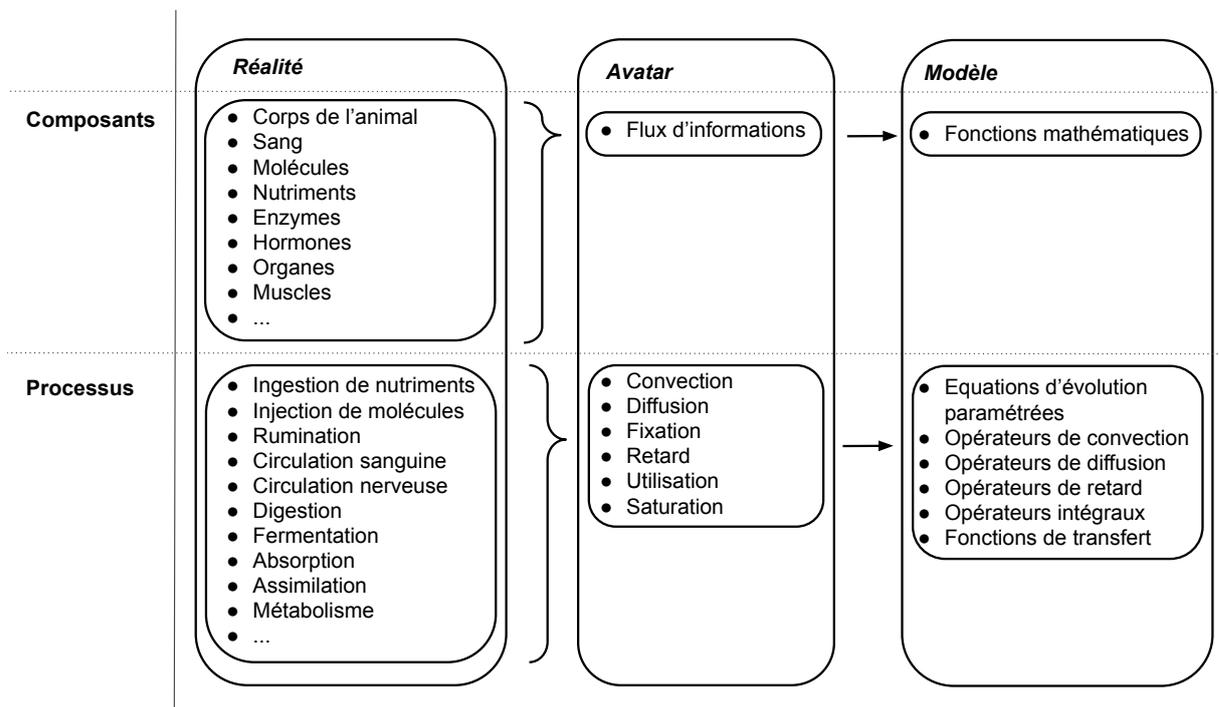


FIGURE 1.2 – Synthèse des différents niveaux considérés dans notre approche, leurs composants et les processus qu'ils contiennent.

Dans notre approche nous distinguons plusieurs éléments de différentes natures, dont les liens sont schématisés dans la figure 1.3. Nous considérons qu'il y a d'une part la réalité, dans laquelle des ingestions et des injections induisent des processus biologiques complexes dans le corps de l'animal. Différents types de capteurs remontent des données et permettent d'extraire de cette réalité des données d'entrée (Inputs) et des données de sortie (Outputs). Étant donné que le modèle n'est pas directement assimilable à une réalité physique, les entrées et les sorties du modèle ne sont pas non plus directement assimilables aux données d'entrée et de sortie collectées. Une fonction mathématique, appelée "fonction d'entrée", permet donc de traduire ces don-

nées et d'induire une réponse du modèle, elle-même reliée aux données de sortie par une fonction mathématique.

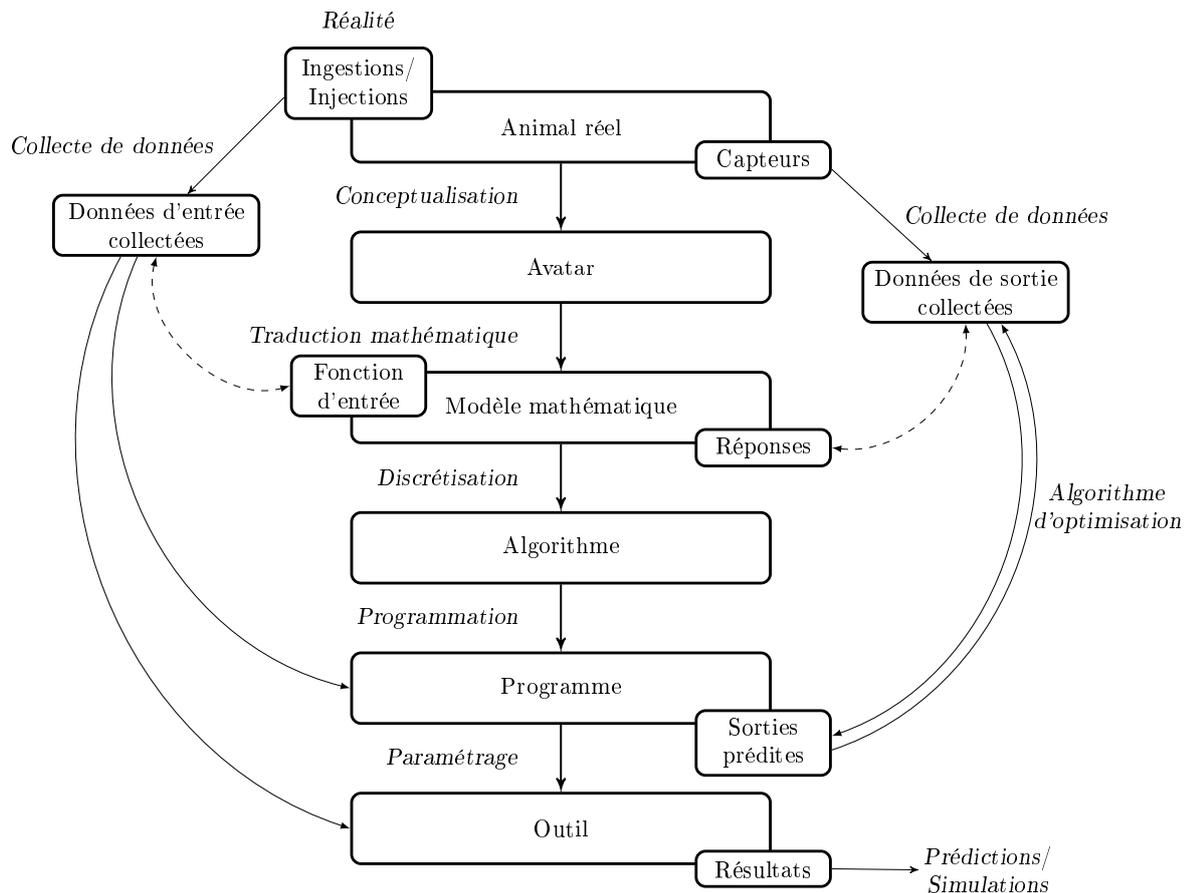


FIGURE 1.3 – Articulation des différents éléments intervenant dans notre exploration

L'algorithme de notre approche résulte de la discrétisation du modèle mathématique, c'est-à-dire du système d'EDP. Ces équations contiennent des paramètres associés à des facteurs biologiques. Elles contiennent par exemple une vitesse de convection, un coefficient de diffusion et des seuils de saturations. Les valeurs de ces paramètres sont apprises à partir de données et en utilisant un algorithme d'optimisation existant.

La présence dans le modèle mathématique de paramètres pouvant être appris à partir de données confère à nos outils une capacité d'apprentissage. En utilisant une base de données nous pouvons donc obtenir un outil capable de reconstituer une dyna-

mique reliant des données d'entrée et de sortie et de réaliser ensuite des simulations et des prédictions. Les outils construits au cours de ces travaux de thèse peuvent donc être considérés comme étant des outils d'apprentissage statistique.

Le programme correspond ensuite au code permettant de gérer cette étape d'apprentissage des paramètres, grâce à un processus itératif basé sur la comparaison des données de sortie issues de la réalité et des sorties prédites. Au cours de ce processus itératif, un algorithme d'optimisation permet de faire converger les paramètres vers des valeurs minimisant la différence entre les données de sortie collectées et les sorties générées par le modèle. Enfin, l'outil de simulation et de prédiction se base sur le modèle mathématique paramétré avec les valeurs des paramètres obtenues à la fin de l'apprentissage.

1.7 Description des données disponibles et de la méthode d'adimensionnement appliquée

1.7.1 Les données disponibles

Dans le cadre de ce projet, des données ont été mises à notre disposition pour ajuster nos outils et tester ensuite la qualité de leurs prédictions. Ces données appartiennent à l'entreprise dans laquelle cette thèse CIFRE a été réalisée et sont donc confidentielles. Dans cette section, ainsi que dans la suite du manuscrit, seules les données adimensionnées seront présentées, sans préciser l'espèce animale sur laquelle nous avons travaillé.

Les données utilisées ont été collectées dans deux élevages différents, au cours d'une période d'engraissement. Les individus, qui sont tous de la même espèce et de la même race, ont été monitorés pendant une certaine période afin de suivre leur croissance et leur consommation alimentaire.

Pour assurer la collecte des données, les animaux ont été équipés de colliers connectés contenant une puce électronique propre à chaque animal. Lorsqu'un animal mange, l'auge contenant l'aliment détecte la puce de ces colliers et collecte différentes informations associées à l'animal et à sa prise d'aliment. Ces auges connectées enregistrent le moment du repas, sa durée et la quantité d'aliment consommée. En parallèle

de cela, l'ensemble des individus ont été pesés grâce à des balances installées ponctuellement dans les élevages.

Dans le premier élevage, que l'on appellera l'élevage «A», 8 individus ont été monitorés pendant une période de référence allant de $t = 0$ jusqu'à $t = 1$. Ces individus ont été alimentés avec une ration de base, sans ajout d'additif. Ils constitueront le groupe que l'on appellera "Témoin A". Dans cet élevage, les animaux ont été pesés à l'instant $t = 0$ et à $t = 1$. Nous connaissons donc le poids de ces individus à $t = 0$ (P_0) et à $t = 1$ (P_1).

Dans le deuxième élevage, que l'on appellera l'élevage «B», 20 individus ont été monitorés pendant une période allant de $t = 0$ jusqu'à $t = 2.5$. Ces individus ont été alimentés avec une ration de base et avec une dose plus ou moins grande d'additif. Parmi ces 20 individus :

- 6 individus n'ont pas reçu d'additif en supplément de la ration de base. Ces animaux constitueront le groupe que l'on appellera «Témoin B»,
- 5 individus ont reçu une dose $Dose_1$ d'additif en supplément de la ration de base. Ces animaux constitueront le groupe que l'on appellera «Dose₁ B»,
- 5 individus ont reçu une dose $Dose_2$ d'additif en supplément de la ration de base. Ces animaux constitueront le groupe que l'on appellera «Dose₂ B»,
- et enfin, 4 individus ont reçu une dose $Dose_3$ d'additif en supplément de la ration de base. Ces animaux constitueront le groupe que l'on appellera «Dose₃ B».

Dans cet élevage les animaux ont été pesés à l'instant $t = 0$, à $t = 0.6$, à $t = 1.52$ et à $t = 2.5$. Nous connaissons donc le poids de ces animaux à ces différents instants (P_0 , $P_{0.6}$, $P_{1.52}$ et $P_{2.5}$).

Lors de chaque étude les données ont été divisées en deux parties. La première partie, appelée "base d'apprentissage", correspond aux données sur lesquelles les outils ont été ajustés. La seconde, appelée "base de test", contient les données sur lesquelles nous avons vérifié la qualité des outils de prédiction obtenus.

Dans plusieurs études nous avons utilisé les données du groupe «A» comme base d'apprentissage et celles associées au groupe «B» comme base de test. Dans plusieurs de nos études nous avons eu besoin de comparer le poids prédit et le poids mesuré des individus à l'instant $t = 1$. Dans ces cas là, pour avoir une information concernant le poids des individus à $t = 1$ dans les deux bases de données et pour pou-

1.7. Description des données disponibles et de la méthode d'adimensionnement appliquée

voir utiliser les individus du groupe «B» comme individus «Test» nous avons interpolé leur poids à $t = 1$ à partir de leur poids mesuré à $t = 0.6$ et à $t = 1.52$.

Dans les deux élevages, la consommation alimentaire de chaque animal a été collectée périodiquement. Nous avons de ce fait la consommation moyenne de chaque individu, sur la période allant de $t = 0$ jusqu'à $t = 0.67$ ($C_{[0-0.67]}$), sur la période allant de $t = 0.67$ jusqu'à $t = 1.5$ ($C_{[0.67-1.5]}$) et sur la période allant de $t = 1.5$ jusqu'à $t = 2.5$ ($C_{[1.5-2.5]}$).

Pour résumer le contenu de ces données nous avons construit la table 1.1.

TABLE 1.1 – Description de la composition des données et des différents groupes étudiés (NI : Nombre d'Individus, × : Donnée disponible, + : Donnée estimée).

	Témoin A	Témoin B	Dose ₁ B	Dose ₂ B	Dose ₃ B
NI	8	6	5	5	4
P₀	×	×	×	×	×
P_{0.6}		×	×	×	×
P₁	×	+	+	+	+
P_{1.52}		×	×	×	×
P_{2.5}		×	×	×	×
C_[0-0.67]	×	×	×	×	×
C_[0.67-1.5]	×	×	×	×	×
C_[1.5-2.5]	×	×	×	×	×

1.7.2 Adimensionnement des données

Les individus de l'élevage "A", qui ont constitué notre base d'apprentissage dans la plupart de nos études, ont été utilisés comme groupe de référence pour adimensionner les données.

Adimensionnement du temps et de l'espace

Les modèles que nous avons construits sont des modèles spatiaux temporels. Il nous faut donc adimensionner les données en temps et en espace.

La période d'engraissement du groupe "A" correspond à la période de référence. Cette période, allant de $t = 0$ jusqu'à $t = X_A$, a donc été ramenée à une période allant de $t = 0$ jusqu'à $t = 1$. Pour cela nous avons divisé le temps par X_A . Donc,

$$t_{Adim_i} = \frac{t_i}{X_A}, \quad (1.7.1)$$

où t_i et t_{Adim_i} correspondent respectivement à la i^{ime} valeur de la variable de temps, avec et sans dimension.

Les individus de l'élevage «B» ont été monitorés pendant une période 2.5 fois plus longue que la période de suivi des individus de l'élevage «A». En divisant cette période par X_A , celle-ci a donc été ramenée à une période allant de $t = 0$ jusqu'à $t = 2.5$.

Dans notre étude l'espace peut être assimilé à la longueur du corps de l'animal ou à celle de la zone étudiée. Nous avons décidé de réduire cet espace à un intervalle allant de 0 jusqu'à 1.

Adimensionnement des données de croissance

Dans notre base de données nous avons également des poids mesurés à différents instants. Comme pour les autres variables, nous voulons que ces poids, qui dépendent du temps, varient entre 0 et 1.

Comme précédemment, le groupe «Témoin A» constitue notre groupe de référence. Pour adimensionner les données de croissance des individus des élevages «A» et «B», nous divisons donc toutes les mesures de poids, aux différents instants t par le poids le plus élevé atteint à $t = 1$ dans le groupe «Témoin A» ($\max(P_1)_A$). Nous avons donc,

$$P_{Adim_{t,i}} = \frac{P_{t,i}}{\max((P_1)_A)} \quad (1.7.2)$$

où $P_{t,i}$ et $P_{Adim_{t,i}}$ correspondent respectivement au poids du i^{ime} individu à l'instant t , avec et sans unité.

Adimensionnement des données de consommation

Dans la base de données utilisée, la consommation alimentaire a été collectée périodiquement. Pour s'assurer que la sortie du modèle à $t = 1$ soit comprise entre 0 et 1, il est nécessaire que la quantité totale d'information injectée dans le modèle soit inférieure ou égale à 1.

Les individus du groupe «Témoin A» constituent notre groupe de référence. Nous avons donc commencé par calculer la quantité totale d'aliment consommée par chaque individu des groupes «Témoin A» et «Témoin B». Nous avons ensuite divisé cette quantité par la plus grande quantité totale d'aliment consommée dans le groupe "A" ($\max(\text{SumConsum})_A$). Nous avons donc :

$$\text{SumConsum}_{\text{Adim}_i} = \frac{\text{SumConsum}_i}{\max(\text{SumConsum})_A} \quad (1.7.3)$$

où SumConsum_i et $\text{SumConsum}_{\text{Adim}_i}$ correspondent respectivement à la quantité totale d'aliment consommée par le $i^{\text{ème}}$ individu, avec et sans dimension.

Dans les données utilisées, nous connaissons pour chaque individu, sa consommation moyenne sur 3 périodes consécutives. Pour adimensionner ces consommations moyennes, nous avons appliqué les opérations suivantes :

$$C_{[0-0.67]_{\text{Adim}_i}} = C_{[0-0.67]_i} \frac{1}{\max(\text{SumConsum})_A}, \quad (1.7.4)$$

$$C_{[0.67-1.5]_{\text{Adim}_i}} = C_{[0.67-1.5]_i} \frac{1}{\max(\text{SumConsum})_A}, \quad (1.7.5)$$

et,

$$C_{[1.5-2.5]_{\text{Adim}_i}} = C_{[1.5-2.5]_i} \frac{1}{\max(\text{SumConsum})_A}. \quad (1.7.6)$$

où $C_{[t_1-t_2]_i}$ et $C_{[t_1-t_2]_{\text{Adim}_i}}$ correspondent respectivement aux consommations moyennes sur la période allant de $t = t_1$ jusqu'à $t = t_2$, pour le $i^{\text{ème}}$ individu, avec et sans dimen-

sion.

1.8 Présentation de l'algorithme DIRECT

Les modèles mathématiques construits lors de ce travail de thèse contiennent plusieurs paramètres dont les valeurs sont à déterminer. Ces modèles peuvent contenir une dizaine de paramètres à ajuster. Néanmoins, étant donné le peu de données que nous avons pour apprendre la valeur de ces paramètres, nous avons fixé certains d'entre eux. Les autres doivent être appris à partir d'une base de données d'apprentissage et d'un algorithme d'optimisation. L'objectif de cet algorithme est de trouver les valeurs des paramètres minimisant l'erreur du modèle sur cette base de données.

La fonction à minimiser, appelée fonction «objectif», correspond donc à un indicateur d'erreur dépendant des valeurs des paramètres. Cette fonction calcule généralement l'écart relatif existant entre des valeurs de la variable de sortie mesurées et celles prédites par le modèle.

Dans notre approche, du fait de la complexité des modèles mathématiques, nous ne connaissons pas à priori l'allure de la fonction "objectif". Il nous faut donc un algorithme d'optimisation capable de trouver les valeurs des paramètres d'un modèle complexe à partir de données expérimentales pouvant contenir beaucoup de variabilité.

Parmi les algorithmes d'optimisation existant l'algorithme DIRECT, développé par JONES *et al.* (1993), est appliqué dans de nombreux problèmes industriels tels que ceux présentés par ZHU et BOGY (2002), SABER et SHAW (2008), NICHOLAS et ALDERSON JR (2014), CARTER *et al.* (2001) et HE *et al.* (2002).

Cet algorithme a été développé dans le but de résoudre des problèmes d'optimisation complexes (NICHOLAS (2014) et FINKEL et KELLEY (2004)). Son nom vient de la contraction de « DIViding RECTangles » et décrit la façon dont cet algorithme tend vers le minimum global de la fonction «objectif». En effet, à chaque itération le domaine, couvrant l'ensemble des valeurs pouvant être prises par les paramètres, est divisé en hyper-rectangles. L'algorithme calcule ensuite la valeur de la fonction «objectif» au centre de ces hyper-rectangles. En procédant ainsi l'algorithme parcourt l'intégralité du domaine en découpant de manière plus précise les zones du domaine dans lesquelles la fonction «objectif» prend de faibles valeurs (FINKEL (2003) et FINKEL et

KELLEY (2004)).

Le processus d'optimisation appliqué par cet algorithme est donc uniquement basé sur le calcul de la valeur de la fonction à minimiser à différents endroits du domaine. Cet algorithme ne nécessite donc pas de connaissance a priori concernant l'allure de la fonction «objectif» (FINKEL (2003)). Seules les gammes de valeurs des paramètres doivent être précisées. De plus, en parcourant l'ensemble du domaine, l'algorithme DIRECT assure une convergence de la fonction «objectif» vers le minimum global (FINKEL (2003) et FINKEL et KELLEY (2004)).

Cependant, cette convergence globale est assurée au prix d'une recherche exhaustive et de la réalisation d'un grand nombre d'itérations.

Dans notre exploration nous avons fait le choix de privilégier l'optimisation globale de nos outils plutôt que la rapidité d'apprentissage. Nous avons donc choisi d'utiliser l'algorithme DIRECT pour ajuster la valeur des paramètres de nos modèles.

Pour cela, comme nos outils sont codés sous R, nous utilisons la fonction *directL* développée par JOHNSON (2008).

CONSTRUCTION ET ANALYSE D'OUTILS D'APPRENTISSAGE STATISTIQUE BIOMIMÉTIQUES

L'objectif de ce chapitre est de présenter la construction d'outils d'apprentissage statistique biomimétiques, leur fonctionnement et leurs capacités de prédiction.

La première section de ce chapitre détaillera l'élaboration d'un premier modèle mathématique et de ses variantes. La construction de ce modèle mathématique, est une étape clé de notre approche. Une attention particulière a donc été portée au choix et à la construction des différents opérateurs mathématiques qui composent les équations de ce modèle.

Dans la seconde section nous exposerons la méthode ayant permis de déterminer les gammes de valeurs des paramètres du modèle mathématique.

Dans la troisième section nous présenterons les tests par simulation effectués. Ces tests ont été réalisés afin d'étudier le fonctionnement de l'outil de prédiction basé sur le modèle mathématique construit, les relations existantes entre certains paramètres et la capacité de l'outil à apprendre la valeur des paramètres sur des données générées artificiellement puis bruitées.

Ces tests par simulation ont mis en évidence les capacités d'apprentissage et de prédiction de l'outil construit. Nous avons donc ensuite réalisé une application de cet outil sur des données réelles. Cette application dont l'objectif était d'être capable de prédire la croissance d'individus en fonction de leur consommation alimentaire sera présentée dans la quatrième section.

2.1 Construction d'un premier modèle mathématique biomimétique

La conception de l'Avatar, nous a permis d'établir un résumé parcimonieux de l'ensemble des phénomènes qui peuvent se dérouler dans le corps d'un l'animal.

Nous avons fait l'hypothèse que, lorsqu'un ensemble de molécules entre dans le corps d'un animal, il circule à travers un réseau de vaisseaux contenant un fluide. Ces molécules utilisent alors ce fluide comme un vecteur pour évoluer via des phénomènes de convection et de diffusion. Dans ce réseau de vaisseaux, cet ensemble de molécules peut entrer en compétition avec d'autres éléments et d'autres processus pouvant retarder sa progression. Ensuite, ces molécules qui circulent peuvent être fixées, puis accumulées dans un organe ou un tissu spécifique. Au cours de leur stockage, elles peuvent être utilisées et induire la modification d'une variable biologique.

Nous avons construit un système d'EDP traduisant mathématiquement ce résumé. Ce système est illustré par la figure 2.1.

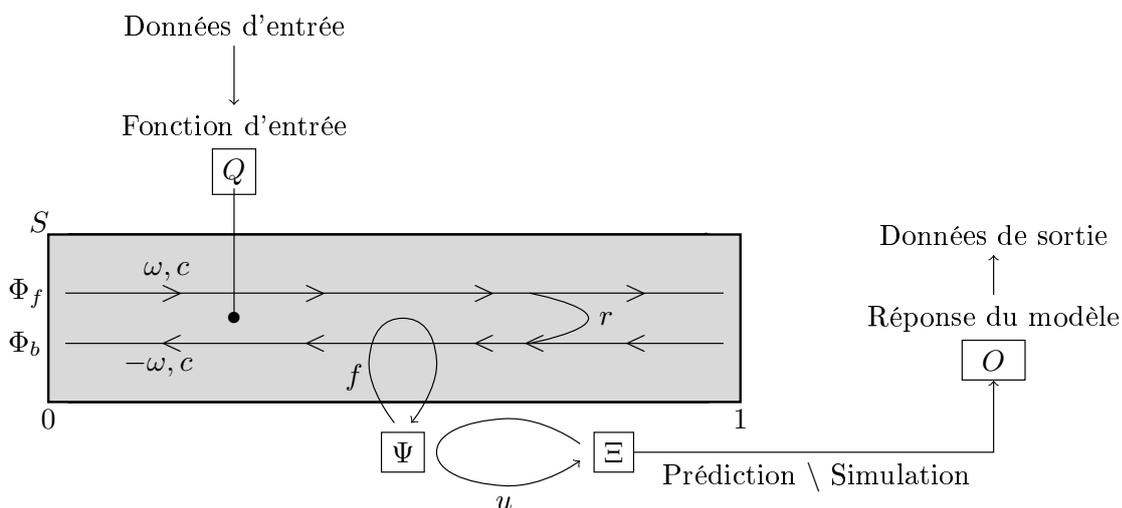


FIGURE 2.1 – Schématisation du modèle mathématique principal

2.1.1 Élaboration des équations constituant le modèle mathématique

Nous avons modélisé le fonctionnement de notre Avatar en utilisant des variables, des densités et des champs, qui sont tous sans dimension.

Nous considérons un flux avant Φ_f , et un flux arrière Φ_b , s'écoulant dans un espace géométrique à une dimension, que nous avons restreint à l'intervalle $(0, 1)$ (Voir la figure 2.1). Cet intervalle correspond à la schématisation d'une partie ou de l'ensemble du corps de l'animal, projetée sur une dimension.

Selon la problématique étudiée, les flux Φ_f et Φ_b peuvent être vus comme un résumé extrême d'une circulation sanguine, d'une circulation digestive ou encore d'un flux nerveux.

Les données d'entrée à intégrer dans ce modèle correspondent principalement à des données concernant la consommation alimentaire de l'animal, sa consommation d'eau et les injections de médicaments qui lui sont administrés. Ces données d'entrée peuvent être intégrées dans le modèle, via une fonction mathématique, sous la forme d'une information dont la quantité est relative à la quantité de molécules ingérées ou injectées.

Ainsi, plusieurs éléments sont associés à une information d , correspondant à une information susceptible d'être injectée dans le modèle. $\{\Phi_f(d)\}(t, x)$ et $\{\Phi_b(d)\}(t, x)$ sont à chaque instant t , deux densités spatiales associées respectivement à un flux avant avec une vitesse de convection ω_d et un flux arrière avec une vitesse de convection $-\omega_d$.

La densité spatiale $\{\Phi_f(d)\}(t, x)$ est supposée solution de :

$$\begin{aligned} \frac{\partial \{\Phi_f(d)\}}{\partial t}(t, x) + \omega_d \frac{\partial \{\Phi_f(d)\}}{\partial x}(t, x) - \frac{\partial \left[c_d \chi \frac{\partial [\{\Phi_f(d)\} + \{\Phi_b(d)\}]}{\partial x} \right]}{\partial x}(t, x) \\ = \frac{1}{2} \{Q(d)\}(t, x) - f_d \{F(d)\}(x) \{\Phi_f(d)\}(t, x) - r_d \{\Phi_f(d)\}(t, x). \end{aligned} \quad (2.1.1)$$

De la même manière, $\{\Phi_b(d)\}(t, x)$ est supposée solution de :

$$\begin{aligned} \frac{\partial \{\Phi_b(d)\}}{\partial t}(t, x) - \omega_a \frac{\partial \{\Phi_b(d)\}}{\partial x}(t, x) - \frac{\partial \left[c_a \chi \frac{\partial [\{\Phi_f(d)\} + \{\Phi_b(d)\}]}{\partial x} \right]}{\partial x}(t, x) \\ = \frac{1}{2} \{Q(d)\}(t, x) - f_a \{F(d)\}(x) \{\Phi_b(d)\}(t, x) + r_a \{\Phi_f(d)\}(t, x). \end{aligned} \quad (2.1.2)$$

Dans les équation (2.1.1) et (2.1.2), la fonction $\{Q(d)\}$ est ce que l'on appelle la "fonction d'entrée". Elle transforme les données collectées en flux d'information. La fonction $\{Q(d)\}$ est donc une source externe d'information.

La moitié de l'information injectée va circuler vers l'avant, via $\{\Phi_f(d)\}$ et l'autre moitié vers l'arrière, via $\{\Phi_b(d)\}$.

Pour modéliser de manière très synthétique la circulation et l'évolution de l'information dans un corps, nous avons intégré dans nos équations un terme de convection et un terme de diffusion, correspondant respectivement au deuxième et au troisième terme des deux premières équations du modèle, (2.1.1) et (2.1.2). Nous imaginons de manière conceptuelle, que $\{\Phi_f(d)\}$, $\{\Phi_b(d)\}$, la vitesse de convection ω_a et le coefficient de diffusion c_a , permettent de résumer très parcimonieusement l'ensemble du périphe que peuvent faire les molécules injectées dans le corps de l'animal.

La fonction χ , intégré dans le terme assimilable à un terme de diffusion des équations (2.1.1) et (2.1.2), est à support compact sur $(0, 1)$, principalement constante et de valeur 1 à l'intérieur de $(0, 1)$. Cette fonction permet de supprimer la diffusion de l'information aux bords du domaine.

Nous faisons ensuite l'hypothèse que l'information qui circule peut prendre du retard. Nous avons modélisé ce retard par le retrait d'une partie de l'information du flux avant $\{\Phi_f(d)\}$, vers le flux arrière $\{\Phi_b(d)\}$. Ce transfert permet de retarder la progression de l'information vers l'avant. Ce phénomène de retard, quantifié par le paramètre r_a , est modélisé par le dernier terme des deux premières équations du modèle. Le paramètre de retard r_a synthétise donc l'ensemble des phénomènes de compétition ou de freinage que peuvent subir les molécules injectées.

Nous avons également modélisé le fait que l'information qui circule peut être fixée, c'est à dire captée par une entité. Dans cette entité, qui peut correspondre à un tissu, à

un organe ou à un groupement d'organes, l'information peut être accumulée et stockée. Nous avons modélisé cela par le retrait de la circulation d'une partie de l'information, qui est alors transférée dans cette entité, notée $\{\Psi(d)\}$.

À chaque instant t , la densité spatiale $\{\Psi(d)\}(t, x)$, associée à l'information fixée, est solution de :

$$\frac{\partial\{\Psi(d)\}}{\partial t}(t, x) = f_a\{F(d)\}(x) \left[\{\Phi_b(d)\}(t, x) + \{\Phi_f(d)\}(t, x) \right] - u_a\{\Psi(d)\}(t, x). \quad (2.1.3)$$

Cette équation modélise le stockage et l'accumulation de l'information.

La localisation sur l'intervalle $(0, 1)$ de l'entité capable de capter l'information est indiquée par la fonction $\{F(d)\}$. Cette fonction vaut 0 à certains endroits de l'espace géométrique et 1 à d'autres. La zone où cette fonction vaut 1 correspond à la zone au niveau de laquelle l'information est captée. Le paramètre f_a détermine le taux d'information fixée.

Une fois mise en stock, l'information peut avoir un impact sur une variable. Nous avons modélisé cette utilisation de l'information par le fait qu'une partie de l'information prélevée du compartiment de stockage, induise une variation de la variable de sortie. Le paramètre u_a détermine le taux d'information utilisée. À chaque instant t , la densité spatiale $\{\Xi(d)\}(t, x)$, associée à l'information utilisée, est solution de :

$$\frac{\partial\{\Xi(d)\}}{\partial t}(t, x) = u_a\{\Psi(d)\}(t, x). \quad (2.1.4)$$

Le domaine $\Omega(d)$ correspond à la zone sur laquelle l'information utilisée agit sur la variable de sortie du modèle $\{O(d)\}(t)$, donnée par l'équation :

$$\{O(d)\}(t) = \int_{\Omega(d)} \{\Xi(d)\}(t, x) dx. \quad (2.1.5)$$

2.1.2 Contextes d'application du modèle mathématique et les différentes versions développées

Le modèle mathématique présenté précédemment, composé des équations (2.1.1), (2.1.2), (2.1.3), (2.1.4) et (2.1.5) peut simuler et prédire un phénomène d'accumulation. Il peut donc être utilisé pour mettre en lien des données relatives à une production

totale sur une période donnée.

La quatrième équation du modèle, correspondant à l'équation (2.1.4) dans le modèle présenté précédemment, est ce que l'on appelle l'équation «d'utilisation». Cette équation détermine l'action de l'information injectée sur la variable de sortie.

Les nutriments, l'eau ou les médicaments peuvent avoir des effets différents selon la variable biologique étudiée. En effet, ces molécules peuvent faire diminuer certaines variables et faire augmenter d'autres. Cette équation doit donc pouvoir s'adapter aux différentes façons dont une molécule peut impacter la variable à prédire.

Pour pouvoir modéliser une croissance logistique nous avons ajouté un limiteur dans cette équation. Dans ce cas, l'équation «d'utilisation» devient :

$$\frac{\partial \{\Xi(d)\}}{\partial t}(t, x) = u_a \{\Psi(d)\}(t, x) \left(\frac{L_a - \{O(d)\}(t)}{L_a} \right) \quad (2.1.4b)$$

Avec cette version de l'équation d'utilisation, des données relatives à l'évolution du poids d'un animal, borné par une constante positive L_a , peuvent être traitées. Cette équation peut être assimilée à l'équation Différentielle de VERHULST (1838),

$$\frac{\partial y}{\partial t}(t) = r y(t) \left(\frac{K - y(t)}{K} \right) \quad (2.1.6)$$

dont la structure est semblable. Nous constatons en effet que dans le cas où aucun terme ne dépend de x , où la valeur de u_a est très élevée et $\Omega(d)$ correspond à l'ensemble de l'intervalle $[0 ; 1]$, $\{\Xi(d)\}$, $\{\Psi(d)\}$ et $\{O(d)\}$ sont très proches les unes des autres. Ainsi, les équations (2.1.6) et (2.1.4b) sont sensiblement les mêmes.

Il peut également être nécessaire de modéliser des fluctuations afin de pouvoir, par exemple, utiliser notre outil pour traiter des données relatives à l'impact d'un médicament sur une variable biologique. Pour cela, nous devons être capables de modéliser l'augmentation ou la diminution de cette variable. Comme cela est le cas pour de nombreuses variables biologiques, nous avons fait en sorte que la variable de sortie du modèle puisse varier entre une borne supérieure et une borne inférieure. Nous avons donc construit l'équation,

$$\frac{\partial \{\Xi(d)\}}{\partial t}(t, x) = - \left(\{\Xi(d)\}(t, x) - U_{ppa} \right) - u_a \{\Psi(d)\}(t, x) \left(\{\Xi(d)\}(t, x) - Low_a \right), \quad (2.1.4c)$$

qui modélise le fait que l'information fixée $\{\Psi(d)\}$ attire la variable de sortie vers une valeur plus petite que celle atteinte durant l'état stationnaire, et l'équation,

$$\frac{\partial \{\Xi(d)\}}{\partial t}(t, x) = -u_a \{\Psi(d)\}(t, x) \left(\{\Xi(d)\}(t, x) - U_{ppa} \right) - \left(\{\Xi(d)\}(t, x) - Low_a \right), \quad (2.1.4d)$$

qui modélise le fait que l'information fixée $\{\Psi(d)\}$ attire la variable de sortie vers une valeur plus élevée que celle atteinte lors de l'état stationnaire.

Dans les deux cas présentés précédemment, la variable de sortie du modèle varie entre une valeur inférieure, Low_a et une valeur supérieure, U_{ppa} .

L'écriture de l'équation «d'utilisation» doit être définie a priori, en fonction de la problématique abordée et des connaissances acquises concernant l'impact du type de données d'entrée utilisées sur la variable à prédire.

Le modèle mathématique composé des équations (2.1.1), (2.1.2), (2.1.3), (2.1.5) et de l'équation d'utilisation (2.1.4), (2.1.4b), (2.1.4c) ou (2.1.4d), a été nommé "Modèle A". Le modèle mathématique composé de l'équation "d'utilisation" (2.1.4b) est appelé "modèle de croissance". Les modèles composés des équations "d'utilisation"(2.1.4c) et (2.1.4d) sont respectivement appelés "modèle de diminution" et "modèle d'augmentation".

2.1.3 Les conditions initiales et les conditions aux bords

Nous avons imposé :

$$\forall t \in (0, \infty), \{\Phi_f(d)\}(t, 0) = \{\Phi_b(d)\}(t, 0) \text{ et } \{\Phi_b(d)\}(t, 1) = \{\Phi_f(d)\}(t, 1). \quad (2.1.7)$$

Les conditions données par (2.1.7) permettent à l'information qui circule de réaliser des allers-retours entre les deux extrémités du domaine.

Les conditions initiales $\{\Phi_f(d)\}(0, x)$, $\{\Phi_b(d)\}(0, x)$, $\{\Psi(d)\}(0, x)$, $\{\Xi(d)\}(0, x)$ et $\{O(d)\}(0)$ sont données pour tout x appartenant à l'intervalle $(0, 1)$.

2.1.4 Les paramètres du modèle

Le système d'EDP développé dans cette section contient plusieurs paramètres dont les valeurs sont à déterminer. Il y a en particulier ω_a , c_a , r_a , f_a et u_a .

Étant donné que nous avons eu très peu de données pour ajuster ces différents paramètres, nous avons dû pour l'instant fixer certains d'entre eux. Nous avons fait notamment le choix de fixer le paramètre de diffusion c_a à une valeur assez faible de 0.001.

Selon les études réalisées, nous avons fixé un nombre plus ou moins grand de paramètres. Les autres sont appris en utilisant une base de données et un algorithme d'optimisation.

2.1.5 Discrétisation du modèle mathématique

Pour discrétiser les équations du modèle nous avons commencé dans un premier temps par discrétiser ces équations en espace. Pour cela nous avons utilisé la méthode classique des Différences Finies, en utilisant un pas d'espace donné. Nous obtenons des équations semi-discrétisées en espace. Ensuite, puisque le modèle mathématique est codé sous R (R CORE TEAM (2014)), nous avons utilisé la fonction *Ode.1D* existante dans ce logiciel et développée par SOETAERT *et al.* (2010), pour gérer la discrétisation en temps des équations semi-discrétisées. Cette fonction R fait appel à la méthode de Runge-Kutta à l'ordre 4 avec un pas de temps donné (voir ENRIGHT (1989)).

Dans les premières explorations, nous avons décidé de paramétrer le maillage avec un pas de temps de 0.001 et un pas d'espace de 0.025.

2.2 Etude des gammes de valeurs des paramètres du modèle mathématique.

Lors de l'ajustement du modèle sur des données d'apprentissage, l'algorithme DIRECT, présenté dans la section 1.8, va chercher la valeur des paramètres minimisant la valeur d'une fonction "objectif". Pour délimiter le domaine de recherche de l'algorithme il est nécessaire de préciser pour chaque paramètre, sa valeur minimale et sa valeur maximale. Nous devons donc déterminer a priori les gammes de valeurs des paramètres à ajuster.

L'objectif de cette étude est d'établir une méthodologie permettant de déterminer la gamme de valeurs de chaque paramètre.

2.2.1 Détermination des gammes de valeurs des vitesses de convection et de diffusion

Dans la plupart des cas, les conditions CFL servent à garantir l'utilisation d'un schéma numérique stable. En fonction de la valeur de la vitesse de convection et de celle du coefficient de diffusion, ces conditions nous permettent alors de connaître le rapport à respecter entre le pas de temps et le pas d'espace.

L'approche mise en oeuvre dans nos travaux implique successivement une définition d'un espace géométrique, puis une fixation du maillage de cet espace (fixation des pas de temps et d'espace), discrétisation par différences finies d'un système d'EDP/EDO sur ce maillage et ensuite un processus d'apprentissage des paramètres du système. Parmi ces paramètres, il y a la vitesse de convection et le coefficient de diffusion. Ces deux paramètres vont donc être déterminés via une optimisation dans des gammes de valeurs qu'il nous faut préciser et il faut assurer que pour toutes valeurs des paramètres prises dans ces gammes, les conditions CFL sont respectées. Pour cela nous avons considéré les conditions CFL comme une contrainte sur les gammes de valeurs de ω_a et de c_a . De plus, le fait de figer le maillage très en amont et pour toute la suite de la démarche, est nécessaire dû au besoin de pérennité des outils dans le cadre de leur déploiement industriel.

La vitesse de convection ω_a , et le coefficient de diffusion c_a , doivent donc respecter les conditions CFL suivantes (Voir COURANT *et al.* (1928) et WEISSTEIN (2014)) :

$$\omega_a \leq \frac{\Delta x}{\Delta t} \quad \text{et} \quad c_a \leq \frac{\Delta x^2}{\Delta t}, \quad (2.2.1)$$

où Δx correspond au pas d'espace et Δt correspond au pas de temps utilisés pour discrétiser le système d'EDP.

Nous pouvons donc conclure que, quelle que soit la valeur des autres paramètres :

$$\omega_a \in \left[0; \frac{\Delta x}{\Delta t} \right]$$

$$c_a \in \left[0; \frac{\Delta x^2}{\Delta t} \right]$$

Or, nous avons vu dans la section 2.1.5 que le pas de temps vaut 0.001 et que le pas d'espace vaut 0.025. Donc :

$$\omega_a \leq \frac{\Delta x}{\Delta t} = \frac{0.025}{0.001} = 25 \quad \text{et} \quad c_a \leq \frac{\Delta x^2}{\Delta t} = \frac{0.025^2}{0.001} = 0.625 \quad (2.2.2)$$

Ainsi nous avons :

$$\omega_a \in [0; 25] \quad (2.2.3)$$

$$c_a \in [0; 0.625] \quad (2.2.4)$$

2.2.2 Etude de l'effet de saturation de l'impact des paramètres sur le modèle

Nous avons mis en évidence un effet de saturation de l'impact du paramètre de retard (r_a), de fixation (f_a) et d'utilisation (u_a). La figure 2.2 illustre ce qui peut être observé lorsque plusieurs courbes de sorties $O(t)$, sont générées lorsque tous les paramètres sont fixés sauf 1. Le paramètre qui varie peut être f_a , u_a ou r_a .

Sur la figure 2.2, le gradient de couleur appliqué aux courbes est associé à la valeur du paramètre étudié : Plus la valeur du paramètre est élevée, plus la couleur de la courbe de sortie générée à partir du modèle paramétré avec cette valeur est foncée. La figure 2.2 montre que lorsque la valeur de f_a (à gauche), de u_a (au centre) ou de r_a (à droite) augmente, l'évolution du profil de la courbe de sortie ralentit, se stabilise et finit par ne plus évoluer. Ainsi, pour chaque paramètre, il existe un seuil de saturation au-delà duquel le paramètre n'influence plus le modèle.

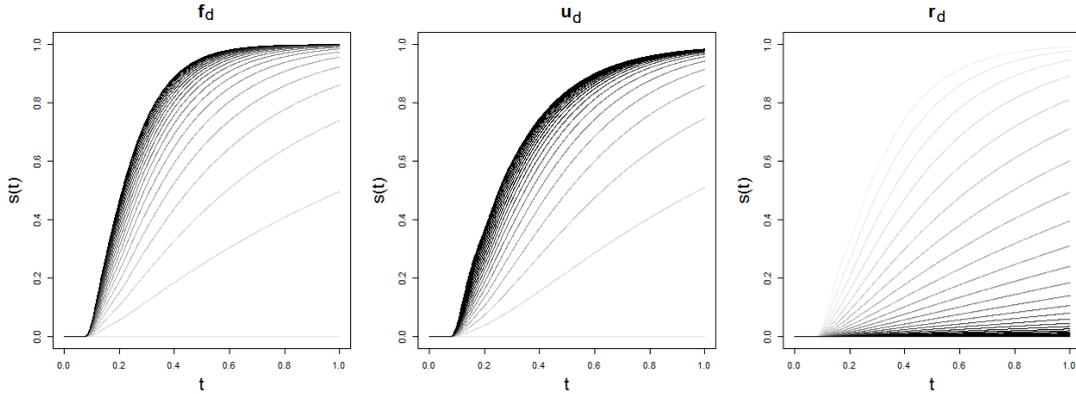


FIGURE 2.2 – Saturation de l’impact des paramètres f_d (à gauche), u_d (au centre) et r_d (à droite), sur le modèle

De ce fait, la gamme de valeurs de chaque paramètre correspond à un intervalle, allant de 0 jusqu’au seuil de saturation de l’impact de ce paramètre sur le modèle.

Pour connaître la valeur du seuil de saturation de chaque paramètre, nous avons calculé un indicateur $I_{ijk}(p_{E_n}, p_{E_{n+1}})$, correspondant à la vitesse d’évolution du profil des courbes de sortie, en fonction de la valeur du paramètre étudié, p_E (où p_E peut être r_d , f_d ou u_d).

Nous avons vu dans la section 2.1.4 que nous avons fixé le paramètre c_d à 0.001. Ainsi, pour étudier la saturation de l’impact d’un paramètre p_E donné, nous fixons les trois paramètres restant parmi ω_d , r_d , f_d et u_d , excepté p_E (p_{F_1} , p_{F_2} , p_{F_3}). Par exemple, pour étudier l’effet de saturation associé à u_d ($= p_E$), nous fixons ω_d ($= p_{F_1}$), r_d ($= p_{F_2}$) et f_d ($= p_{F_3}$).

Ensuite, nous générons N_E courbes de sortie, à partir de N_E valeurs de p_E prises dans une très large gamme de valeurs. Nous avons généré ces courbes pour N_1 valeurs de p_{F_1} , N_2 valeurs de p_{F_2} et N_3 valeurs de p_{F_3} , prises elles aussi dans des gammes de valeurs très larges. Nous avons ensuite calculé l’indicateur,

$$I_{ijk}(p_{E_n}, p_{E_{n+1}}) = \frac{\|O_{ijk}^{n+1} - O_{ijk}^n\|}{\|O_{ijk}^{n+1}\|}. \quad (2.2.5)$$

Dans 2.2.5,

$$\|f\| = \int_0^1 |f(t)| dt$$

- n correspond au rang du paramètre étudié p_E , $n \in \llbracket 1, N_E \rrbracket$
- i correspond au rang du premier paramètre p_{F_1} , $i \in \llbracket 1, N_1 \rrbracket$,
- j correspond au rang du deuxième paramètre p_{F_2} , $j \in \llbracket 1, N_2 \rrbracket$,
- k correspond au rang du troisième paramètre p_{F_3} , $k \in \llbracket 1, N_3 \rrbracket$,
- O_{ijk}^n correspond à la courbe de sortie générée à partir des valeurs i, j, k, n des paramètres p_E, p_{F_1}, p_{F_2} et p_{F_3} .

Et nous imposons,

$$p_E \neq p_{F_1}, p_E \neq p_{F_2}, p_E \neq p_{F_3}, p_{F_2} \neq p_{F_1}, p_{F_2} \neq p_{F_3} \text{ et } p_{F_3} \neq p_{F_1} \quad (2.2.6)$$

Par exemple, pour étudier l'évolution de l'impact du paramètre u_d , nous avons calculé $I_{ijk}(u_{d_n}, u_{d_{n+1}})$ pour plusieurs valeurs u_{d_n} , $n \in \llbracket 1, N_E \rrbracket$. Ainsi, $I_{ijk}(u_{d_n}, u_{d_{n+1}})$ correspond à la différence relative entre deux courbes générées à partir de deux valeurs successives de $u_d : u_{d_n}$ et $u_{d_{n+1}}$. Dans ce cas, i correspond à la valeur de ω_d , j correspond à la valeur de r_d et k correspond à la valeur de f_d .

Pour des valeurs données de r_d, ω_d et f_d , nous avons calculé $I_{ijk}(u_{d,n}, u_{d,n+1})$ pour 15 valeurs de u_d . Nous avons obtenu la courbe de la figure 2.3.

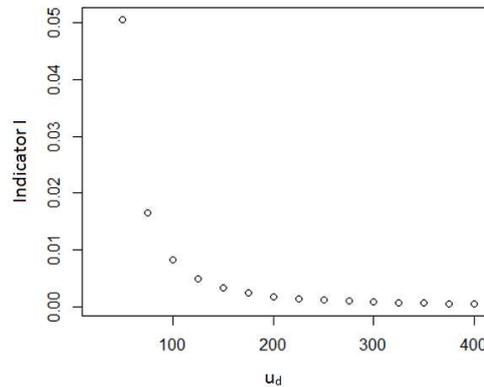


FIGURE 2.3 – La valeur de l'indicateur I_{ijk} en fonction de la valeur de $u_{d,n+1}$.

Dans ce cas, la borne supérieure de u_d ($u_{d_{\text{Upp}}}^{ijk}$), correspond à la valeur à partir de laquelle la différence relative entre deux courbes de sortie consécutives est plus petite que 0.005. Donc, $u_{d_{\text{Upp}}}^{ijk}$ correspond à la plus petite valeur de $u_{d,n+1}$ telle que,

$$I_{ijk}(u_{d,n}, u_{d,n+1}) \leq 0.005. \quad (2.2.7)$$

Nous avons calculé p_E^{ijk} , correspondant à la valeur de la borne supérieure du paramètre étudié p_E , pour des valeurs de p_{F_1} , p_{F_2} et p_{F_3} données. Nous avons ensuite construit une table équivalente à la table 2.1. Cette table correspond à un plan d'expérience permettant de déterminer la borne supérieure du paramètre p_E , quelle que soit la valeur des paramètres p_{F_1} , p_{F_2} et p_{F_3} .

TABLE 2.1 – Le plan d'expérience permettant de déterminer la borne supérieure de la gamme de valeurs de p_E quelle que soit la valeur des autres paramètres, p_{F_1} , p_{F_2} et p_{F_3} .

p_{F_1}	p_{F_2}	p_{F_3}	p_{Upp}^{ijk}
$p_{F_1,1}$	$p_{F_2,1}$	$p_{F_3,1}$.
$p_{F_1,1}$	$p_{F_2,1}$	$p_{F_3,2}$.
.	.	.	.
.	$p_{F_2,1}$	p_{F_3,N_3}	.
.	.	.	.
.	p_{F_2,N_2}	$p_{F_3,1}$.
.	p_{F_2,N_2}	$p_{F_3,2}$.
.	.	.	.
$p_{F_1,1}$	p_{F_2,N_2}	p_{F_3,N_3}	.
.	.	.	.
p_{F_1,N_1}	$p_{F_2,1}$	$p_{F_3,1}$.
p_{F_1,N_1}	$p_{F_2,1}$	$p_{F_3,2}$.
.	.	.	.
.	$p_{F_2,1}$	p_{F_3,N_3}	.
.	.	.	.
.	.	.	.
.	p_{F_2,N_2}	$p_{F_3,1}$.
.	p_{F_2,N_2}	$p_{F_3,2}$.
.	.	.	.
p_{F_1,N_1}	p_{F_2,N_2}	p_{F_3,N_3}	.

2.2.3 Détermination des gammes de valeurs de u_d , f_d et r_d

La gamme de valeurs de u_d

Afin de trouver la borne supérieure de u_d quelle que soit la valeur de f_d , ω_d et r_d , nous avons calculé u_{dUpp}^{ijk} pour plusieurs valeurs de f_d , ω_d et r_d . Ainsi, la borne supérieure de u_d , notée u_{dUpp} , correspond à la plus grande valeur obtenue de u_{dUpp}^{ijk} .

Nous avons sélectionné 10 valeurs de chaque paramètre, prise entre 0 et 25 pour ω_d et entre 0 et 400 pour f_d et r_d . Pour chaque triplet testé (f_d, ω_d, r_d) nous avons calculé la valeur de u_{dUpp}^{ijk} et nous avons obtenu la figure 2.4.

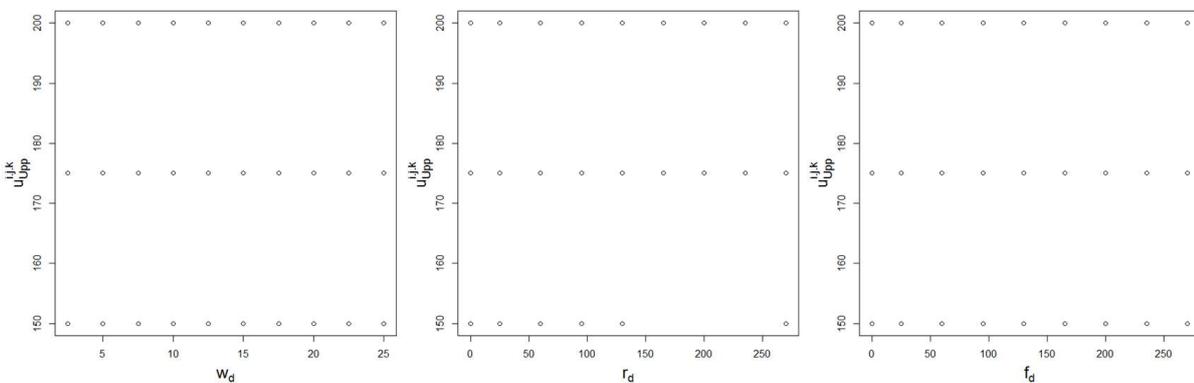


FIGURE 2.4 – La valeur de u_{dUpp}^{ijk} en fonction des valeurs de ω_d , r_d et f_d .

Sur cette figure 2.4 nous remarquons que u_{dUpp}^{ijk} ne dépasse pas 200, quelles que soient les valeurs de f_d , ω_d et r_d . Nous pouvons donc conclure que la borne supérieure de u_d vaut 200. Ainsi, u_d appartient à l'intervalle $[0; 200]$ quelle que soit la valeur des autres paramètres.

La gamme de valeurs de f_d

Comme précédemment, pour trouver la borne supérieure de f_d quelle que soit la valeur de u_d , ω_d et r_d , nous calculons la valeur de f_{dUpp}^{ijk} pour plusieurs valeurs de u_d , ω_d et r_d . Nous avons sélectionné 10 valeurs de chaque paramètre, prises entre 0 et 25 pour ω_d et entre 0 et 400 pour u_d et r_d et pour chaque triplet testé (u_d, ω_d, r_d) nous avons calculé la valeur de f_{dUpp}^{ijk} . Nous avons obtenu la figure 2.5.

À gauche de la figure 2.5, les valeurs prises par f_{dUpp}^{ijk} sont données pour chaque valeur testée de u_d . Au milieu et à droite de la figure les valeurs de f_{dUpp}^{ijk} sont données

pour chaque valeur testée de r_d et ω_d .

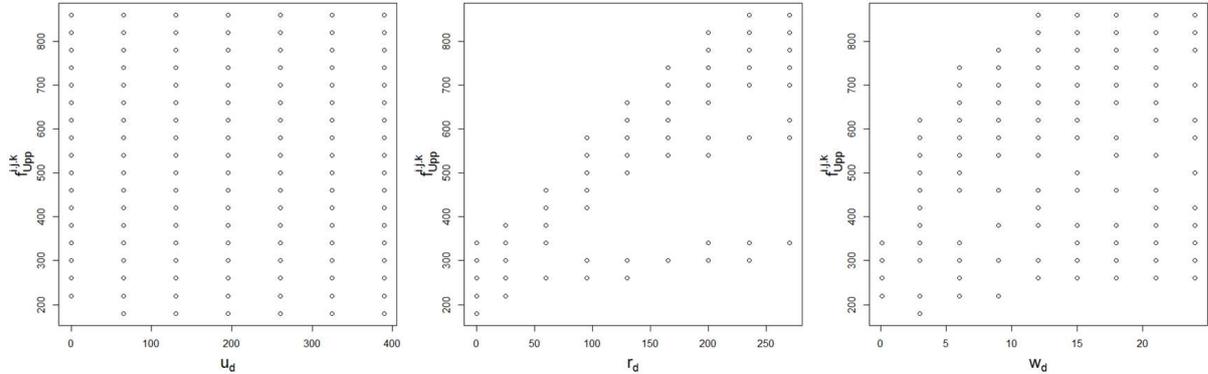


FIGURE 2.5 – La valeur de l'indicateur f_{aUpp}^{ijk} en fonction des valeurs de u_d , r_d et ω_d .

L'image de gauche de la figure 2.5 montre que f_{aUpp}^{ijk} ne dépend pas de la valeur de u_d . En revanche, Les deux autres images montrent qu'il existe une relation entre les valeurs prises par f_{aUpp}^{ijk} et celles de ω_d et celles de r_d . En effet, la valeur de f_{aUpp}^{ijk} semble être plus élevée lorsque les valeurs de ω_d et de r_d augmentent. La figure 2.6 montre plus précisément que l'augmentation de la valeur de f_{aUpp}^{ijk} est due à l'augmentation de la valeur de r_d . Nous avons donc étudié la relation existante entre f_{aUpp}^{ijk} et r_d .

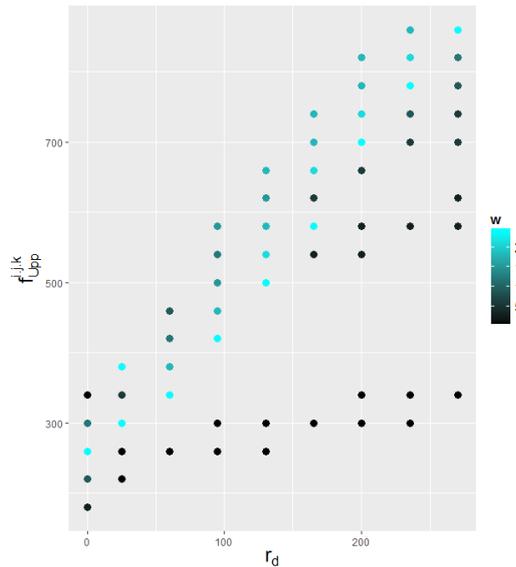


FIGURE 2.6 – La valeur de l'indicateur f_{aUpp}^{ijk} en fonction de la valeur de r_d , et colorée en fonction de la valeur de ω_d .

La plus grande valeur de f_{aUpp}^{ijk} a été obtenue pour $\omega_d = 15$ (Figure 2.5). Nous avons

donc fixé ω_d à cette valeur et u_d à une valeur moyenne de 100. Nous avons ensuite calculé f_{dUpp}^{ijk} pour plusieurs valeurs de r_d prises entre 0 et 900 (Figure 2.7).

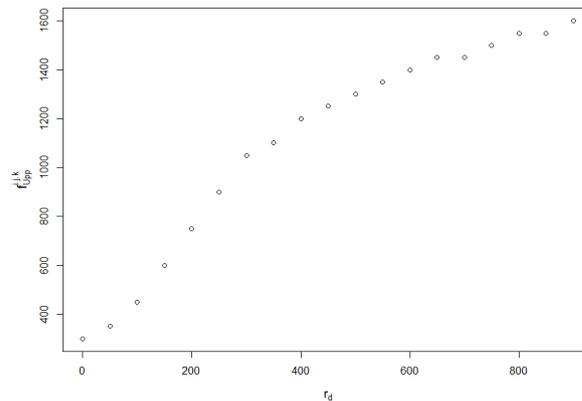


FIGURE 2.7 – La valeur de l'indicateur f_{dUpp}^{ijk} en fonction de la valeur de r_d .

Nous avons obtenu la courbe de la figure 2.7. Cette courbe peut être modélisée par une équation de la forme :

$$f_{dUpp}^{ijk} = \frac{A}{1 + B \cdot \exp(-C \cdot r_d^D)} + E \quad (2.2.8)$$

Pour ajuster la valeur des paramètres A , B , C , D et E de l'équation (2.2.8) sur la courbe de la figure 2.7, nous avons utilisé la fonction *directL* présentée dans la section 1.8.

Après l'ajustement des paramètres, dont les valeurs sont présentées dans la table 2.2, nous avons obtenu une équation reliant les valeurs de f_{dUpp}^{ijk} et celles du paramètre r_d .

L'objectif de cette méthode mise en place est de trouver la borne supérieure de f_a quelle que soit la valeur de r_d . De ce fait, nous avons majoré la valeur du coefficient E par une valeur E_M , de façon à être sûr que la gamme de valeur établie contient bien toutes les valeurs pouvant être prises par f_a . Ainsi, nous fixons $E_M = 345$ (Figure 2.8). Pour connaître la valeur de la borne supérieure de f_a il a donc fallu déterminer la borne supérieure de r_d et en déduire celle de f_a grâce à la relation donnée par

TABLE 2.2 – Coefficients de l'équation reliant f_{dUpp}^{ijk} et r_d .

Paramètre	Valeur
A	1394.7
B	1853.1
C	1.83
D	0.25
E	272.3

l'équation (2.2.8).

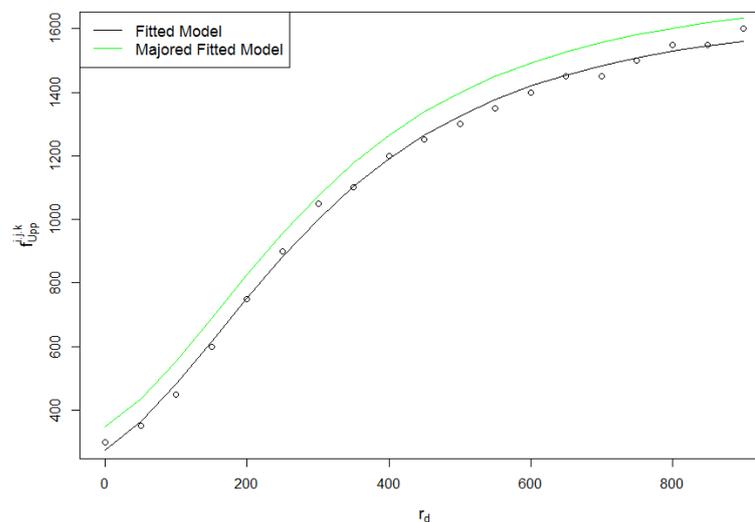


FIGURE 2.8 – La relation reliant r_d et f_{dUpp}^{ijk} . Le modèle "ajusté" (Fitted) et le modèle "ajusté et majoré" (Majored Fitted) correspondent à l'équation (2.2.8) paramétrée respectivement avec les paramètres (A, B, C, D, E) et (A, B, C, D, E_M) .

La gamme de valeurs de r_d

En réalisant les mêmes analyses que précédemment pour déterminer la borne supérieure de r_d , nous avons observé que ce paramètre peut prendre une large gamme de valeurs avant d'atteindre un seuil de saturation. Or le retard r_d , peut être compensé par la vitesse de convection ω_d . En effet, une faible vitesse de convection associée à un faible retard peut induire la même cinétique que celle générée par une vi-

tesse de convection élevée associée à un retard important. Ainsi, un couple $(\omega_{a,H}; r_{a,H})$ de valeurs importantes de ω_a et r_a peut être considéré comme équivalent à un couple $(\omega_{a,L}; r_{a,L})$ de plus faibles valeurs de ω_a et de r_a . Ainsi, cet effet de compensation permet de limiter la gamme de valeur de r_a à un intervalle restreint mais permettant malgré cela de simuler toutes les cinétiques possibles.

Nous avons fait l'hypothèse que la plus grande valeur atteinte par le paramètre r_a est 900 et nous l'avons fixé à cette valeur. Ensuite, pour plusieurs couples $(\omega_{a,H}; r_{a,H})$, où $r_{a,H} = 900$ et $\omega_{a,H} \in [2; 25]$, nous avons cherché le couple $(\omega_{a,te}; r_{a,te})$ permettant de générer la même courbe de sortie que le couple $(\omega_{a,H}; r_{a,H})$ mais à partir de valeurs plus faibles de ω_a et de r_a .

Pour cela nous avons tout d'abord fixé $u_a = 200$ et $f_a = 900$, qui sont d'importantes valeurs de ces paramètres. Nous avons sélectionné 20 valeurs de $\omega_{a,te}$ réparties de manière régulière entre 0.2 et 5. Ensuite, pour chaque couple $(\omega_{a,H}; r_{a,H})$ et pour chaque valeur de $\omega_{a,te}$, nous cherchons la valeur de r_a minimisant la somme des résidus relatifs au carré ($RRSS$). Ces résidus mesurent la différence entre la courbe générée à partir du couple $(\omega_{a,H}; r_{a,H})$ et celle générée à partir du couple $(\omega_{a,te}; r_{a,te})$:

$$RRSS((\omega_{a,H}; r_{a,H}); (\omega_{a,te}; r_{a,te})) = \sum_{i=1}^m \left(\sum_{l=1}^M \left(\frac{O_{[\omega_{a,H}; r_{a,H}]}i,l} - O_{[\omega_{a,te}; r_{a,te}]i,l}}{O_{[\omega_{a,H}; r_{a,H}]}i,l}} \right)^2 \right), \quad (2.2.9)$$

où $O_{[\omega_{a,k}; r_{a,k}]}$ correspond à la courbe de sortie générée à partir du couple $(\omega_{a,k}; r_{a,k})$, m est le nombre d'individus dans la base d'apprentissage et M est le nombre de points de la courbe de sortie.

Pour déterminer la valeur de $r_{a,te}$ nous avons utilisé l'algorithme d'optimisation DIRECT. Cet algorithme détermine la valeur de $r_{a,te}$ minimisant l'indicateur $RRSS$ donné par l'équation (2.2.9).

À la fin de cette étape, nous avons obtenu pour chaque couple $(\omega_{a,H}; r_{a,H})$, 20 couples $(\omega_{a,te}; r_{a,te})$ équivalents. Pour chaque couple $(\omega_{a,H}; r_{a,H})$ nous avons sélectionné parmi les 20 couples obtenus $(\omega_{a,te}; r_{a,te})$, le couple associé à la valeur de $RRSS$ la plus faible. Nous obtenons finalement la table 2.3, dans laquelle R^2 correspond au coefficient de détermination reliant les courbes de sortie $O_{[\omega_{a,H}; r_{a,H}]}$ et $O_{[\omega_{a,te}; r_{a,te}]}$.

Pour chaque couple $(\omega_{a,H}; r_{a,H})$ nous avons donc trouvé un couple équivalent $(\omega_{a,L}; r_{a,L})$ présentant une valeur plus faible de r_a .

TABLE 2.3 – Les couples $(\omega_{a,H}; r_{a,H})$ testés et leurs couples équivalents $(\omega_{a,L}; r_{a,L})$.

$\omega_{a,H}$	$r_{a,H}$	$\omega_{a,L}$	$r_{a,L}$	$RRSS$	R^2
2	900	0.7	284	35.2	0.98
4.5	900	1	194.4	53.9	1
7	900	1.5	185.1	61.2	1
9.5	900	2	177.5	76.8	1
12	900	2.5	171.4	76.6	1
14.5	900	3	166.6	66.2	1
17	900	3.5	162.9	55.9	1
19.5	900	4	160.1	48.7	1
22	900	4.5	157.9	44.2	1
24.5	900	5	156.1	42.1	1

Le couple $(\omega_{a,L}; r_{a,L})$ équivalent au couple $(\omega_{a,H} = 2; r_{a,H} = 900)$ présente la valeur la plus importante de r_a , qui est de 284. De ce fait, quelle que soit la valeur de f_a , u_a et ω_a nous avons fixé la borne supérieure de r_a à 284.

Déduction de la gamme de valeurs de f_a

Nous avons vu précédemment que la borne supérieure de f_a dépend de la borne supérieure de r_a . Nous avons démontré dans la section précédente que la borne supérieure de r_a vaut 284. Ainsi, à partir de l'équation (2.2.8), nous avons déduit que la borne supérieure de f_a vaut 1035.

Conclusion à propos des gammes de valeurs des différents paramètres

Pour conclure, nous pouvons dire que, étant donné que les paramètres ω_a et c_a sont ajustés via une procédure d'optimisation, les gammes de valeurs de ces paramètres sont imposées par le maillage que nous avons choisi. En effet cela permet d'assurer le respect des conditions CFL quelles que soient les valeurs prises par ces deux paramètres.

La borne supérieure du paramètre u_a a été déterminée en étudiant la saturation de l'impact de ce paramètre sur le modèle.

Un effet de saturation de l'impact du paramètre f_a sur le modèle a également été observé. Néanmoins, le seuil de saturation de ce paramètre dépend de la valeur du paramètre r_a . Nous avons donc établi une relation existante entre la valeur de r_a et le

seuil de saturation de f_a .

Le paramètre r_a peut prendre des valeurs importantes avant d'atteindre un seuil de saturation. Mais il existe un phénomène de compensation entre r_a et ω_a permettant de limiter la gamme de valeur de r_a à un intervalle restreint mais assurant au modèle la possibilité de simuler toutes les cinétiques possibles.

La détermination de la borne supérieure de r_a nous a ensuite permis de déduire la borne supérieure de f_a , en utilisant la relation établie précédemment.

Nous obtenons finalement les gammes de valeurs contenues dans la table 2.4.

TABLE 2.4 – Les gammes de valeurs des différents paramètres.

Paramètres	Gamme de valeurs
ω_a	[0; 25]
c_a	[0; 0.625]
u_a	[0; 200]
r_a	[0; 284]
f_a	[0; 1035]

2.3 Tests par simulation de la capacité d'apprentissage de l'outil d'apprentissage statistique modélisant un phénomène de croissance

L'objectif de cette section est de présenter les tests par simulation réalisés sur l'outil d'apprentissage statistique modélisant un phénomène de croissance. Cet outil est basé sur le modèle mathématique dont l'équation "d'utilisation" est l'équation (2.1.4b). Ce modèle "A", dit de "croissance", se compose donc des équations (2.1.1), (2.1.2), (2.1.3), (2.1.4b) et (2.1.5).

Dans ces tests par simulation les entrées de ce modèle correspondent à un volume d'information injecté ($VolQ$) et le moment de cette injection (ct). La sortie est une série temporelle correspondant à une courbe de croissance.

Ces tests ont eu pour but de vérifier la capacité de l'outil à apprendre la valeur des paramètres du modèle mathématique sur lequel il est basé, à partir de données

biologiques bruitées. Pour cela nous avons commencé par générer une base de données fictive à partir de notre modèle mathématique paramétré. Une méthodologie a été mise en place afin que notre base de données fictives puisse avoir les mêmes caractéristiques qu'une base de données collectées via des capteurs.

Nous avons ensuite utilisé cette base de données pour étudier les effets de compensation existants entre ω_d et r_d et entre f_d et u_d .

Nous avons finalement simulé l'apprentissage des paramètres ω_d et r_d et entre f_d et u_d sur ces mêmes données, puis vérifié si l'ajustement du modèle s'était fait correctement. La valeur du paramètre L_d de l'équation (2.1.4b) est fixée à 1.

Ces tests par simulation ont également été l'occasion de vérifier l'identifiabilité du modèle et de réaliser une étude de robustesse du modèle en fonction de la quantité de bruit intégrée dans les données.

2.3.1 Génération d'une base de données fictive

Afin de tester la capacité d'apprentissage de l'outil construit nous avons commencé par générer une base de données fictive contenant les données associées à 50 individus, c'est-à-dire 50 poids initiaux fictifs et les 50 courbes de sortie associées. L'objectif est d'obtenir une base de données générée artificiellement mais ayant les mêmes caractéristiques qu'une base de données expérimentales. Pour cela nous avons intégré dans ces données du bruit et de la variabilité inter-individuelle.

Intégration d'une variabilité inter-individuelle

Les paramètres du modèle sont des constantes à déterminer. Néanmoins, dans le but d'introduire des différences entre les individus générés, nous avons considéré ici, et uniquement dans cette section 2.3.1, que les paramètres étaient des facteurs biologiques suivant des lois Normales. En effet, pour simuler des différences inter-individuelles nous avons attribué à chaque paramètre une loi Normale. Nous avons centré ces lois sur des valeurs choisies arbitrairement dans les gammes de valeurs des paramètres et nous avons choisi la variance des paramètres de façon à introduire suffisamment de variabilité, mais sans excès. Pour cela nous avons commencé par fixer une variance relative de 0.005 et nous avons ensuite augmenté la variance de la loi associée à ω_d pour obtenir une variabilité inter-individuelle que nous avons jugé suffisante (Voir Table 5.2).

À partir de ces lois Normales nous avons généré 50 valeurs des paramètres ω_d , r_d , f_d et u_d . Leurs distributions respectives sont représentées dans la figure 2.9.

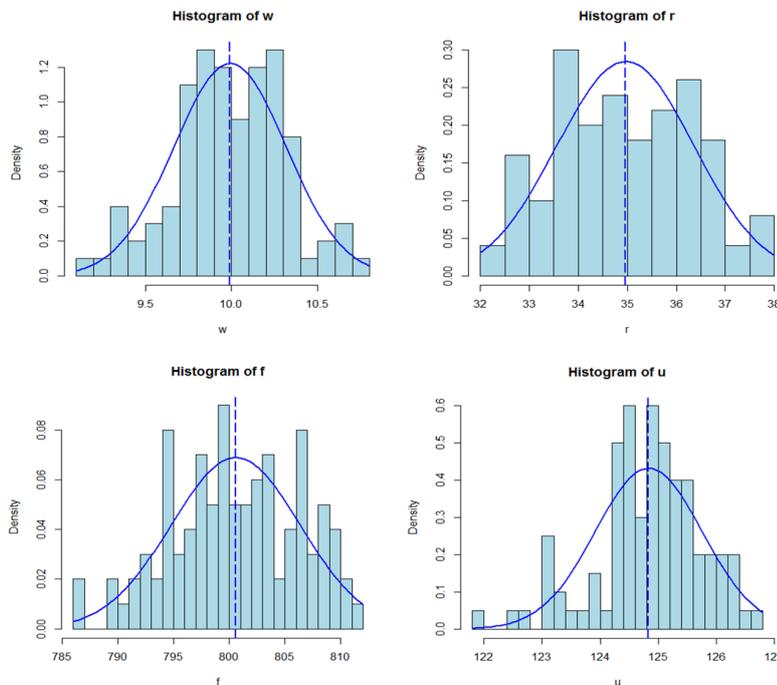


FIGURE 2.9 – Distributions des paramètres ω_d , r_d , f_d et u_d

Pour visualiser la variabilité intégrée dans les données nous avons tracé 50 courbes à partir de 50 quadruplets $(\omega_d, r_d, f_d, u_d)$ générés, et des mêmes données d'entrée, c'est-à-dire à partir d'un couple $(VolQ, ct)$. Nous avons obtenu la figure 2.10. Cette figure nous montre un exemple de variabilité pouvant être obtenue à partir des lois Normales que nous avons paramétrées.

Génération d'Inputs fictifs

Les données d'entrée intégrées dans le modèle utilisé pour réaliser ces tests par simulation correspondent donc à un volume d'information injecté ($VolQ$) et au moment de l'injection (ct). Nous avons considéré que chacun de ces paramètres pouvait prendre n'importe quelle valeur comprise entre 0 et 1. Nous avons donc appliqué à ces deux types de données d'entrée une loi Uniforme sur $[0; 1]$ (Table 5.2). Comme précédemment nous avons généré à partir de cette loi Uniforme, 50 valeurs de $VolQ$ et 50 valeurs de ct .

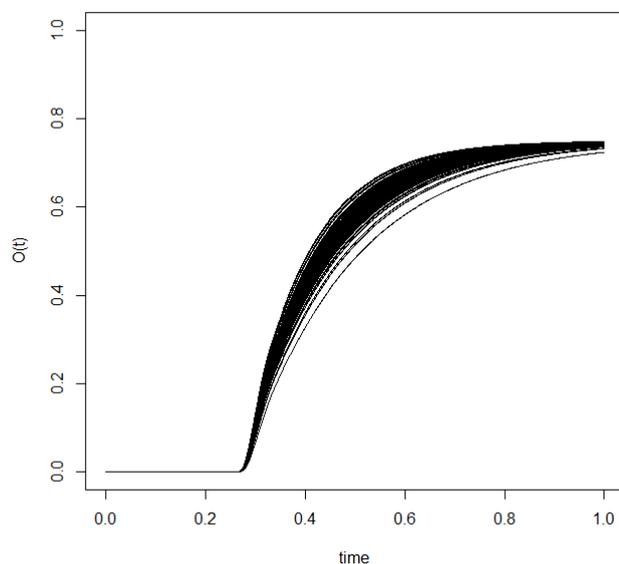


FIGURE 2.10 – Exemple de variabilité pouvant être présente dans la base de données générée.

À partir des 50 quadruplets de paramètres $(\omega_a, r_a, f_a, u_a)$ obtenus précédemment et des 50 couples d'entrées fictives générées $(VolQ, ct)$, nous avons simulé 50 courbes de sortie.

Ajout d'un bruit blanc

Toujours dans l'objectif d'obtenir une base de données fictive ressemblant à une base de données expérimentale, nous avons ajouté un bruit aléatoire aux courbes de sortie générées. Pour cela nous avons ajouté à chaque point de chaque courbe de sortie, une valeur aléatoire suivant une loi Normale centrée en 0 et présentant une variance de 0.05 (Table 5.2). La figure 2.11 montre des exemples de courbes générées avant et après l'ajout du bruit aléatoire.

Nous avons divisé la base de données obtenue en deux jeux de données distincts : nous avons constitué une base d'apprentissage composée de 30 courbes et une base de test constituée des 20 courbes restantes. L'objectif est de disposer d'une base de données d'apprentissage pour ajuster les valeurs des paramètres du modèle et une base de test pour tester la qualité du modèle obtenu et vérifier que le modèle ne sur-

ajuste pas les données d'apprentissage.

Dans la suite de cette section nous avons supposé que nous disposons d'une base de données fictives, ressemblant à une base de données expérimentale, et un modèle contenant des paramètres dont nous ne connaissons pas la valeur.

TABLE 2.5 – Distributions suivies par les paramètres et les données d'entrée.

Paramètre	Loi de probabilité
ω_d	$\mathcal{N}(10, 0.31)$
r_d	$\mathcal{N}(35, 1.42)$
f_d	$\mathcal{N}(800, 5.175)$
u_d	$\mathcal{N}(125, 1)$
$VolQ$	$\mathcal{U}(0, 1)$
c_t	$\mathcal{U}(0, 1)$
$Noise$	$\mathcal{N}(0, 0.05)$

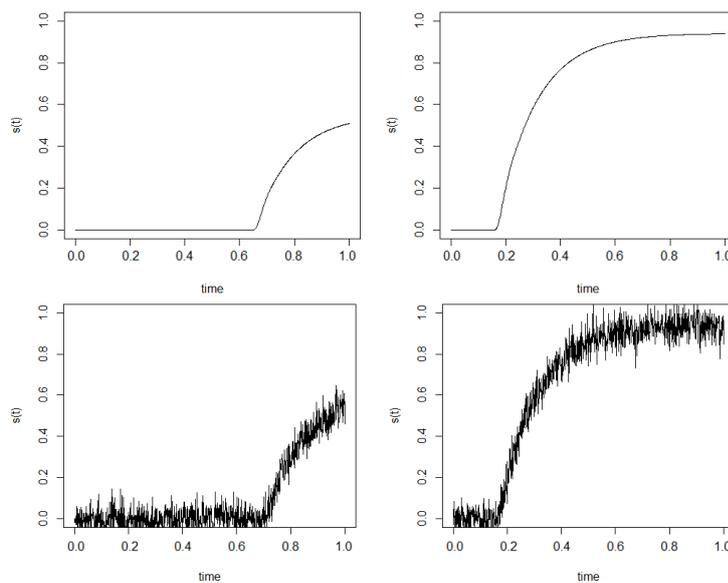


FIGURE 2.11 – Exemple de courbes simulées avant et après l'ajout d'un bruit gaussien.

2.3.2 Les relations existantes entre les paramètres

Parmi les paramètres ω_d , r_d , f_d et u_d , certains se compensent.

L'impact de la vitesse ω_d , peut être compensé par le retard r_d subi par l'information qui circule. En effet, une faible vitesse de convection associée à un faible retard de

l'information peut induire une cinétique équivalente que celle induite par une vitesse de convection élevée associée à un retard important.

La fixation et l'utilisation de l'information, contrôlées respectivement par les paramètres f_a et u_a , sont aussi deux processus qui se compensent. En effet, une importante fixation de l'information suivi d'une faible utilisation de celle-ci, peut induire le même impact sur la variable de sortie qu'une faible fixation de l'information suivi d'une importante utilisation.

Ainsi, des relations existent entre les couples de paramètres (ω_a, r_a) et (f_a, u_a) . L'objectif a donc été d'utiliser les données d'apprentissage fictives pour mettre en évidence ces relations et établir une équation reliant les paramètres ω_a et r_a et les paramètres f_a et u_a .

Etude des relations existantes entre ω_a et r_a

Dans un premier temps, nous avons décidé de mettre en évidence la relation existante entre ω_a et r_a en calculant l'erreur faite sur la base d'apprentissage par le modèle paramétré avec différents couples (ω_a, r_a) . Pour cela nous avons parcouru le domaine $\omega_a \times r_a$, fourni par les gammes de valeurs présentées dans la section 2.2. Nous avons ensuite calculé la somme des résidus relatifs au carré ($RRSS$) associée à chaque modèle paramétré avec les différents couples (ω_a, r_a) testés. Cet indicateur est donné par l'équation (2.3.1) :

$$RRSS(\omega_a, r_a) = \sum_{i=1}^n \left(\sum_{l=1}^m \left(\frac{(y_{obs_{i,l}} - O_{[\omega_a, r_a]_{i,l}})}{y_{obs_{i,l}}} \right)^2 \right). \quad (2.3.1)$$

Dans l'équation (2.3.1), y_{obs_i} et $O_{[\omega_a, r_a]_i}$ correspondent respectivement à la courbe générée initialement et à la courbe simulée à partir du couple (ω_a, r_a) , pour le i^{ime} individu. n correspond au nombre d'individus contenus dans la base d'apprentissage et m le nombre de points sur les courbes.

Les figures 2.12 et 2.13 représentent la valeur de l'indicateur $RRSS$ en fonction de ω_a et r_a . L'existence de séries de couples équivalents peut être identifiée sur la figure

2.12(a). En effet, sur la figure 2.12(a) le gradient de couleur indique qu'il existe des séries de couples $(\omega_{d_{Opt}}, r_{d_{Opt}})$ induisant la même valeur de $RRSS$. Ces séries de couples équivalents sont schématisées par la figure 2.12(b). On observe sur la figure 2.13 une zone dans laquelle l'indicateur $RRSS$ atteint un minimum. Cette zone est schématisée par la courbe $EC1$ de la figure 2.12(b). Nous avons supposé que le couple optimal $(\omega_{d_{Opt}}, r_{d_{Opt}})$ induisant la plus petite erreur de modèle, appartenait à cette courbe. Nous avons donc cherché l'équation de cette courbe $EC1$.

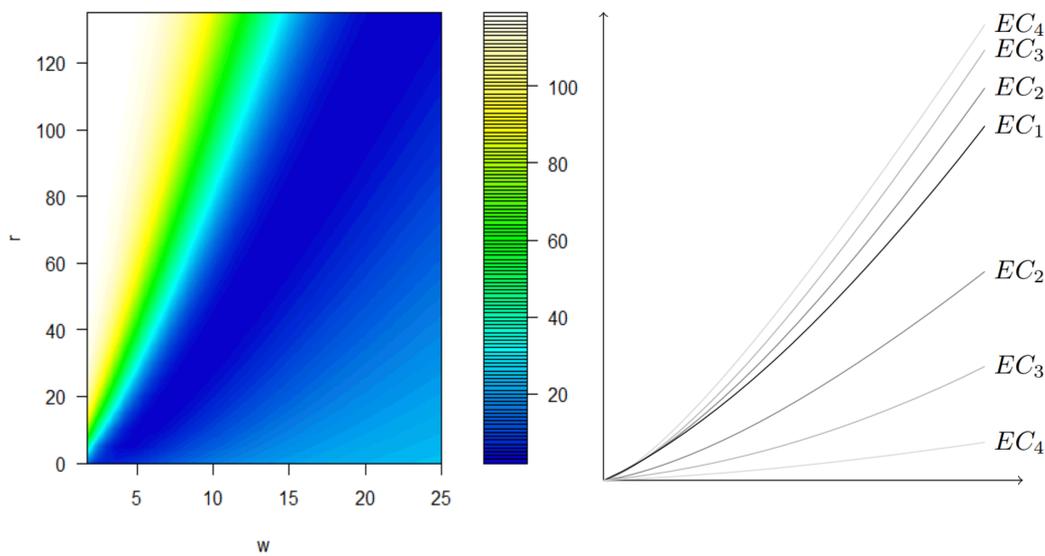


FIGURE 2.12 – La valeur de l'indicateur $RRSS$ en fonction de la valeur des paramètres ω_a et r_a (a : gauche), et le schéma des différentes séries de couples équivalents (EC) (b : droite)

Recherche des couples $(\omega_{d_{Opt}}, r_{d_{Opt}})$ minimisant la valeur de $RRSS$

Pour trouver l'équation de la courbe $EC1$ nous avons cherché, pour différentes valeurs de ω_a , la valeur de r_a minimisant la valeur du $RRSS$. Pour chaque valeur de ω_a testée, nous avons utilisé l'algorithme d'optimisation DIRECT pour trouver la valeur de r_a minimisant la fonction «objectif» correspondant ici à la valeur moyenne du $RRSS$. Cette fonction «objectif» est donnée par l'équation (2.3.2),

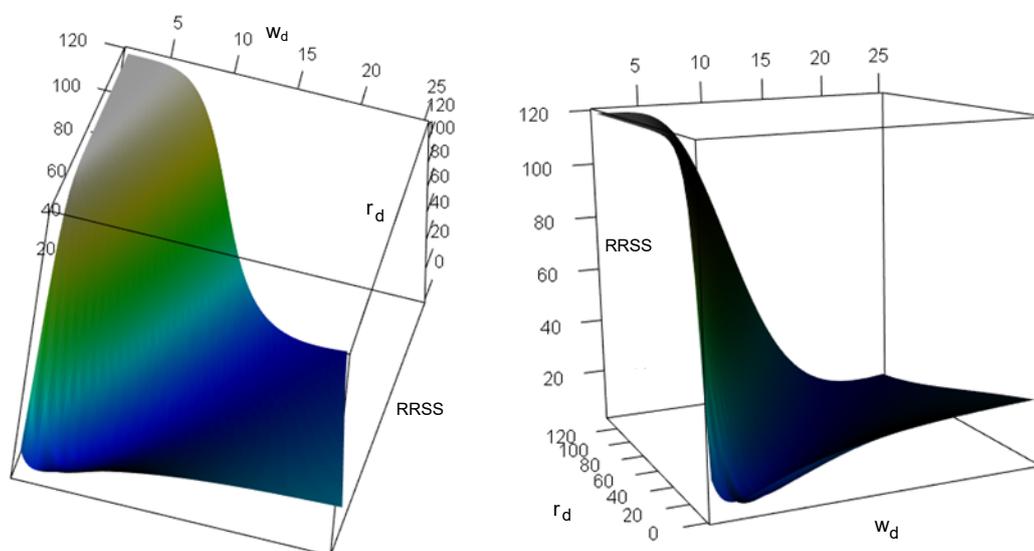


FIGURE 2.13 – La représentation 3D de la valeur de $RRSS$ en fonction de ω_d et r_d

$$f_{obj}(r_d) = \frac{RRSS(\omega_d, r_d)}{n}, \quad (2.3.2)$$

où n est le nombre d'individus sur lesquels nous avons ajusté la valeur des paramètres.

Afin d'avoir plusieurs valeurs de r_d pour chaque valeur de ω_d testée, nous avons réalisé un échantillonnage de la base d'apprentissage. En effet, avant chaque ajustement de r_d nous avons sélectionné 20 courbes parmi les 30 courbes de la base d'apprentissage et nous avons ajusté la valeur de r_d sur ces 20 courbes. À la fin de la procédure d'ajustement nous avons obtenu trois valeurs de r_d pour chaque valeur de ω_d testée (Figure 2.14). Grâce à une régression de Nadaraya-Watson (Voir NADARAYA (1964) et WATSON (1964)), nous obtenons une relation non-paramétrique reliant ω_{dOpt} et r_{dOpt} de la forme :

$$r_{dOpt} = \hat{m}_{\omega_d}(\omega_{opt}) + \epsilon_{\omega_d}, \quad (2.3.3)$$

où \hat{m}_{ω_d} correspond à l'estimateur de Nadaraya-Watson et ϵ_{ω_d} aux résidus.

Connaissant la relation reliant ω_{dOpt} et r_{dOpt} , il est désormais possible de déduire

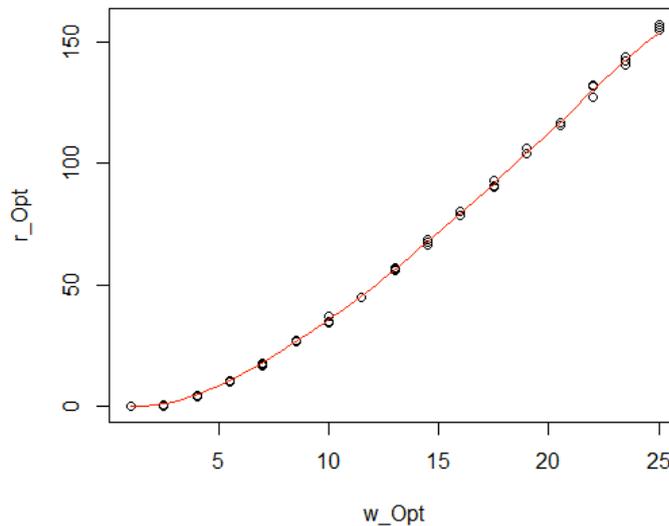


FIGURE 2.14 – La régression de Nadaraya-Watson reliant les couples $(\omega_{d_{Opt}}, r_{d_{Opt}})$.

la valeur de l'un des deux paramètres en fonction de la valeur de l'autre. Ainsi, cette relation permet de réduire le nombre de paramètres à apprendre simultanément.

Étude de la relation existante entre les paramètres f_a et u_a

Comme nous l'avons vu dans l'introduction de cette section 2.3.2, il existe également un effet de compensation entre f_a et u_a . Comme précédemment nous avons cherché la relation existante entre ces deux paramètres. L'objectif est de continuer à réduire le nombre de paramètres à ajuster simultanément.

Comme précédemment nous avons parcouru le domaine $f_a \times u_a$ établi dans la section 2.2 et nous avons calculé la valeur de l'indicateur $RRSS$ associé au modèle paramétré avec différents couples (f_a, u_a) (Figures 2.15 et 2.16). Comme précédemment, cette étude a mis en évidence des séries de couples (f_a, u_a) équivalents schématisées par la figure 2.15(a). Sur la figure 2.16 nous observons également une zone dans laquelle l'indicateur $RRSS$ est minimal. Cette zone est schématisée par la courbe $EC1$ de la figure 2.15(a). Nous avons supposé, comme précédemment, que le couple optimal $(f_{d_{Opt}}, u_{d_{Opt}})$ induisant la plus petite erreur de modèle, appartenait à cette courbe.

Nous avons donc cherché à établir l'équation de cette courbe EC_1 .

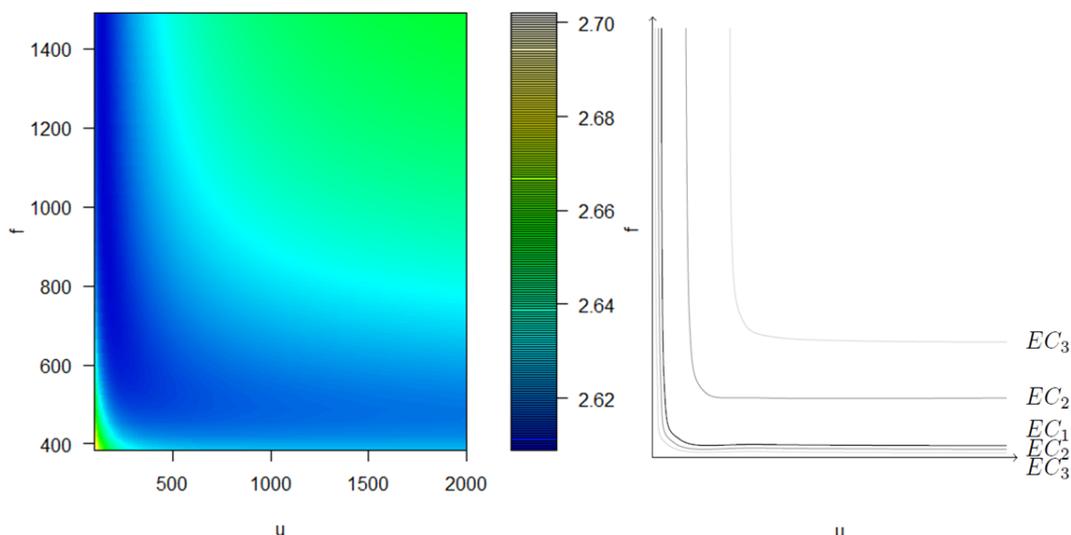


FIGURE 2.15 – La valeur du $RRSS$ en fonction de f_a et u_a (a) et le schéma représentant les différentes séries de couples équivalents (EC) (b)

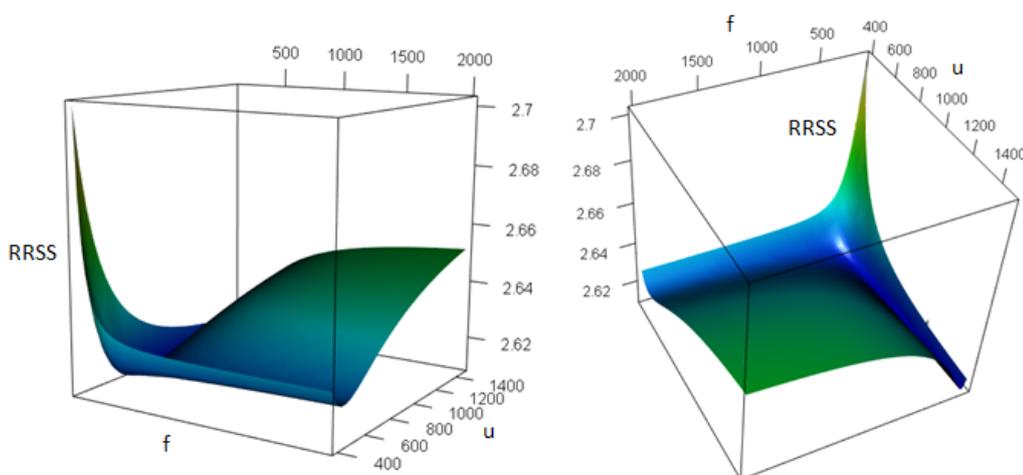


FIGURE 2.16 – La représentation 3D de la valeur du $RRSS$ en fonction de f_a et u_a

Recherche des couples $(f_{d_{Opt}}, u_{d_{Opt}})$ induisant la plus petite valeur du $RRSS$

Comme précédemment, pour trouver l'équation de la courbe EC_1 associée aux plus petites valeurs du $RRSS$, nous avons cherché pour plusieurs valeurs de f_a , la valeur de u_a minimisant l'indicateur $RRSS$. Pour chaque valeur de f_a testée, nous avons

utilisé l'algorithme d'optimisation DIRECT pour trouver la valeur de u_a minimisant la valeur de la fonction «objectif» correspondant à la moyenne du $RRSS$ donnée par l'équation (2.3.4),

$$f_{obj}(u_a) = \frac{RRSS(f_a, u_a)}{n}, \quad (2.3.4)$$

où n est le nombre d'individus sur lesquels nous avons ajusté la valeur des paramètres.

Comme précédemment, de façon à avoir plusieurs valeurs de u_a pour chaque valeur de f_a testée, nous avons réalisé un échantillonnage de la base d'apprentissage en sélectionnant avant chaque ajustement, 20 courbes parmi les 30 de la base d'apprentissage. Nous avons ensuite ajusté la valeur de u_a sur ces 20 courbes sélectionnées. À la fin de la procédure d'ajustement nous avons obtenu trois valeurs de u_a pour chaque valeur de f_a testée (Figure 2.17). Grâce à une régression de Nadaraya-Watson, nous avons obtenu une relation non-paramétrique reliant $f_{d_{Opt}}$ et $u_{d_{Opt}}$ de la forme :

$$u_{d_{Opt}} = \hat{m}_{f_a}(f_{d_{Opt}}) + \epsilon_{f_a}, \quad (2.3.5)$$

où \hat{m}_{f_a} correspond à l'estimateur de Nadaraya-Watson et ϵ_{f_a} les résidus.

Connaissant la relation existante entre les paramètres $f_{d_{Opt}}$ et $u_{d_{Opt}}$, il est désormais possible de déduire la valeur de l'un en fonction de la valeur de l'autre. Ainsi, cette relation permet de réduire encore le nombre de paramètres à ajuster simultanément.

2.3.3 Ajustement des paramètres et calcul de la précision du modèle

Nous avons ajusté les paramètres du modèle sur la base d'apprentissage et nous avons ensuite testé la qualité du modèle obtenu en calculant l'erreur associée à ce modèle sur la base de test.

Ajustement de ω_a et f_a

Afin de pouvoir réaliser plusieurs ajustements nous avons réalisé un échantillonnage de la base d'apprentissage. En effet, avant chaque ajustement nous avons sélectionné 20 courbes au hasard parmi les 30 de la base d'apprentissage et nous avons

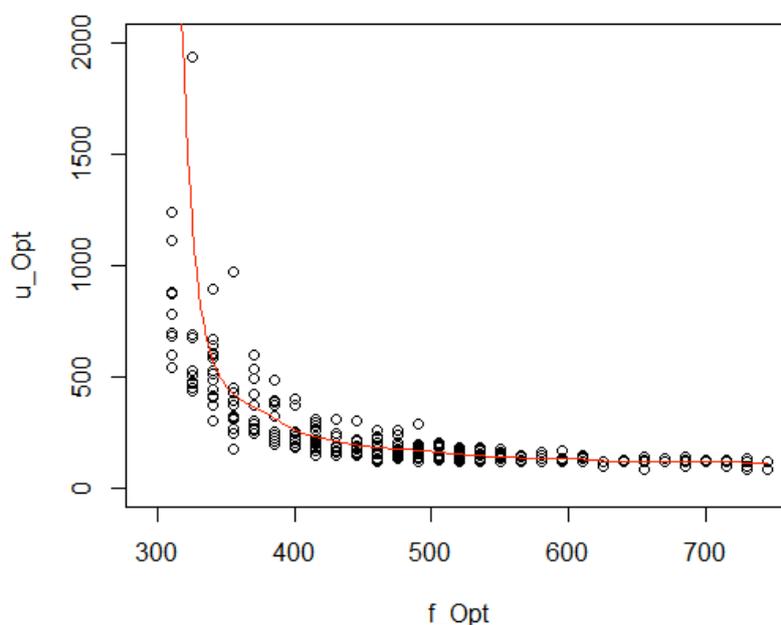


FIGURE 2.17 – La régression de Nadaraya-Watson reliant $f_{d_{Opt}}$ et $u_{d_{Opt}}$.

ajusté les paramètres sur ces 20 courbes sélectionnées. En procédant ainsi, nous avons réalisé 30 ajustements sur 20 courbes sélectionnées au hasard.

Pour déterminer les valeurs de ω_a , r_a , f_a et u_a , nous avons ajusté ω_a et f_a sur les courbes de la base d'apprentissage sélectionnées et nous en avons ensuite déduit les valeurs de r_a et u_a .

Pour ajuster la valeur des paramètres, nous avons utilisé l'algorithme d'optimisation DIRECT permettant de trouver le couple (ω_a, f_a) minimisant la fonction «objectif» donnée par l'équation (2.3.6).

$$f_{obj}(\omega_a, f_a) = \frac{RRSS(\omega_a, f_a)}{n}, \quad (2.3.6)$$

où n est le nombre d'individus sur lesquels nous avons ajusté les paramètres ω_a et f_a .

À la fin des 30 ajustements nous avons obtenu 30 valeurs des paramètres ω_a , r_a , f_a et u_a . Nous avons calculé la moyenne et l'écart-type relatif (RSD) de chaque paramètre

(Table 2.6). Nous avons paramétré le modèle mathématique avec la valeur moyenne des paramètres et nous avons testé la précision de l'outil de prédiction basé sur ce modèle en comparant les courbes de la base d'apprentissage et les celles prédites (Figure 2.18). Pour cela nous avons calculé la valeur du coefficient de détermination (R^2) du modèle obtenu sur la base d'apprentissage (Table 2.6). Nous avons remarqué que le coefficient R^2 était élevé, ce qui montre que le modèle s'ajuste correctement aux courbes de la base d'apprentissage. Ce résultat est confirmé par la figure 2.18.

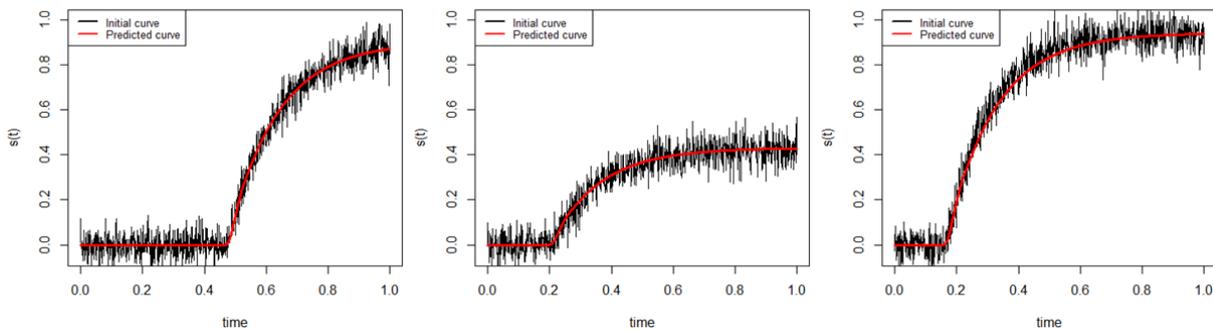


FIGURE 2.18 – Exemples de résultats obtenus concernant la concordance entre les courbes d'apprentissage (noire) et les prédictions fournies par le modèle (rouge).

TABLE 2.6 – Moyenne et écart-type relatif associés aux paramètres et au coefficient de détermination calculé sur la base d'apprentissage

<i>Paramtre</i>	<i>Moyenne</i>	<i>RSD</i>
ω_d	19.51	0.055
r_d	61.29	0.080
f_d	240.0	0.001
u_d	561.4	0.007
L_a	1.71	0.011
$ARE(1)$ (%)	1.83	0.013

Validation de la qualité de l'outil de prédiction obtenu

Pour vérifier la qualité de l'outil de prédiction obtenu nous avons calculé sa précision sur les données de la base de test. Pour cela nous avons calculé la valeur de l'indicateur $RRSS$ et celle du coefficient de détermination associé à chaque courbe

de la base de test. Nous avons ensuite représenté la distribution de ces indicateurs (Figure 2.19).

La valeur du $RRSS$ est faible et celle du coefficient de détermination est élevée (Table 2.7). Ces indicateurs confirment que le modèle est capable de simuler correctement les courbes de la base de test.

De plus, nous avons également comparé les valeurs du R^2 et du $RRSS$ associées au modèle ajusté (R_{Fit}^2 et $RRSS_{Fit}$) et celles associées au modèle "générateur" (R_{Gener}^2 et $RRSS_{Gener}$), c'est-à-dire le modèle qui a été initialement utilisé pour générer la base de données fictives (Figure 2.19 et Table 2.7). La valeur du $RRSS_{Fit}$ est faible et cette valeur est très proche de celle du $RRSS_{Gener}$. De même, les valeurs de R_{Fit}^2 et R_{Gener}^2 sont très proches.

Les valeurs de $RRSS_{Fit}$ et de R_{Fit}^2 montrent que l'ajustement des paramètres s'est fait correctement. En effet, le modèle "générateur" peut être considéré comme le meilleur modèle pouvant être obtenu puisqu'il s'agit du modèle ayant généré les courbes de la base d'apprentissage et de test. La valeur de l'indicateur R_{Gener}^2 quantifie donc le bruit et la variabilité inter-individuelle que nous avons intégré dans les données. Le fait que les valeurs de R_{Fit}^2 et R_{Gener}^2 et les valeurs de $RRSS_{Fit}$ et $RRSS_{Gener}$ soient très proches montre que l'erreur de l'outil de prédiction ajusté se limite à la quantité de bruit et aux différences inter-individuelles intégrées initialement dans la base de données générée. Les paramètres du modèle ont donc été ajustés correctement et l'outil obtenu présente une capacité de prédiction très satisfaisante.

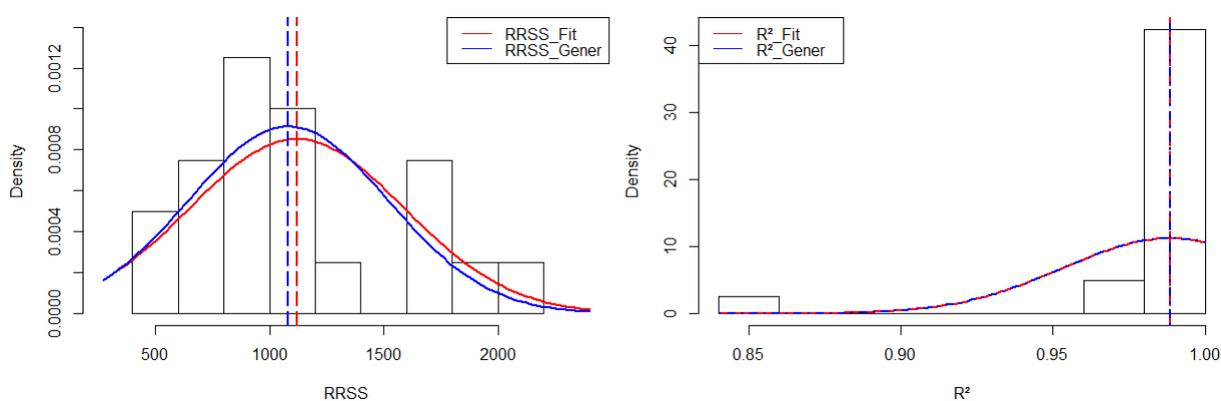


FIGURE 2.19 – Distributions des valeurs des indicateurs $RRSS$ et R^2 associées au modèle "générateur" et au modèle ajusté.

TABLE 2.7 – Comparaison des valeurs des indicateurs associées au modèle "générateur" et au modèle ajusté.

	$RRSS$	R^2
Modèle "générateur"	1082	0.9887
Modèle ajusté	1119	0.9886

Identifiabilité du modèle

Daprès PERASSO (2009) et WALTER (2014), l'identifiabilité d'un modèle paramétrique correspond à la capacité de ce modèle à apprendre un set de paramètres initialement inconnus, à partir de données. L'identifiabilité suppose donc que les valeurs de ces paramètres appris soient stables au cours de différents ajustements sur des données similaires.

La méthode d'échantillonnage nous a permis de réaliser 30 ajustements sur des données similaires. Nous avons donc obtenu 30 valeurs des paramètres ω_a , r_a , f_a et u_a . Pour chacun de ces paramètres, leur valeur moyenne et leur écart-type relatif ont été calculés (Table 2.6).

L'écart-type relatif associé à chaque paramètre est faible. Cela montre que les valeurs des paramètres ajustées sont stables au cours des différents ajustements. La méthode d'apprentissage développée assure donc l'identification d'un set de valeurs de paramètres que l'on peut considérer comme unique.

La structure du modèle construit et la méthode d'apprentissage utilisée assurent donc bien l'identifiabilité de notre outil d'apprentissage statistique biomimétique.

2.3.4 Étude de la robustesse du modèle et de l'impact de la présence de bruit dans les données d'apprentissage

Une étude de la robustesse du modèle et de l'impact de la présence de bruit dans les données a été réalisée. L'objectif de cette étude a été de vérifier la capacité de l'outil construit à apprendre la valeur des paramètres du modèle quelle que soit la quantité de bruit présent dans les données d'apprentissage.

Génération de six bases de données fictives

Pour réaliser cette étude de robustesse nous avons utilisé la base de données générée présentée dans la section 2.3.1, mais nous avons également généré 5 autres bases de données fictives dans lesquelles a été intégré une quantité de bruit plus ou moins importante. Ces 5 autres bases de données fictives ont été générées en appliquant les mêmes lois de probabilités aux paramètres ω_a , r_a , f_a et u_a et aux données d'entrée, $VolQ$ et ct (Voir la table 5.2). Seule la variance du bruit gaussien ajouté aux données varie selon les bases de données. Comme indiqué dans la table 5.2, selon les bases de données générées nous avons ajouté aux courbes un bruit dont la variance varie entre 0.05 et 1.

TABLE 2.8 – Les différentes quantités de bruit intégrée dans les 6 bases de données générées

Base de données générée	Loi associée au bruit gaussien ajouté
Base nř1	$\mathcal{N}(0, 0.05)$
Base nř2	$\mathcal{N}(0, 0.075)$
Base nř3	$\mathcal{N}(0, 0.1)$
Base nř4	$\mathcal{N}(0, 0.125)$
Base nř5	$\mathcal{N}(0, 0.15)$
Base nř6	$\mathcal{N}(0, 1)$

Comme dans la section 2.3.1, à partir des lois de probabilité appliquées nous générons pour chacune des 6 bases de données, 50 valeurs de ω_a , r_a , f_a , u_a , $VolQ$ et ct . Nous générons ensuite à partir de ces paramètres et de ces données d'entrée, 50 courbes de sortie. La figure 2.20 montre des exemples de courbes bruitées contenues dans les différentes bases de données fictives.

Comme précédemment, nous divisons ensuite chaque base de données en deux jeux de données : Nous avons constitué des bases d'apprentissage composées de 30 courbes et des bases de test constituées de 20 courbes.

Etude de la robustesse de l'outil construit en fonction de la variance du bruit présent dans les données

En utilisant les 6 bases de données générées, nous avons étudié la robustesse de notre outil d'apprentissage statistique en fonction de la variance du bruit ajouté.

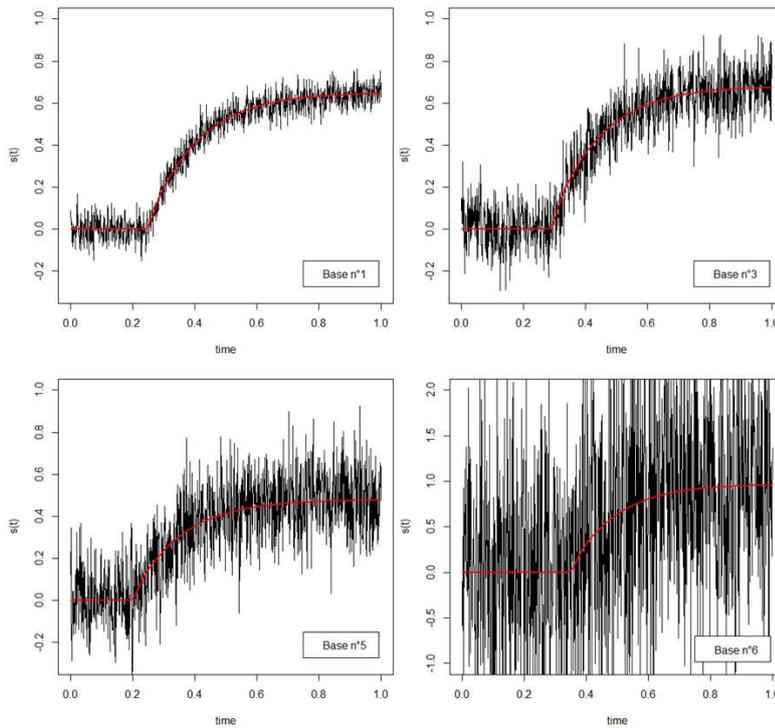


FIGURE 2.20 – Exemples de courbes bruitées contenues dans les différentes bases de données fictives n°1, n°2, n°3, n°4, n°5 et n°6.

Cette étude a consisté à ajuster les paramètres du modèle mathématique sur les 6 bases de données d'apprentissage. L'objectif a ensuite été de calculer le coefficient R^2 sur la base d'apprentissage et sur la base de test et ensuite de comparer les valeurs du R^2 associées aux 6 modèles ajustés et celles du R^2 associées aux modèles "générateurs".

Comme dans la section 2.3.2, avant d'ajuster les paramètres, il nous faut établir les relations existantes entre ω_a et r_a , et entre f_a et u_a . Ces deux relations doivent être établies sur les 6 bases de données générées.

En procédant comme dans la section 2.3.2, pour chaque base de données fictives nous obtenons des relations non paramétriques reliant ces deux couples de paramètres (Figure 2.21). Nous intégrons ces relations dans les modèles appliqués à chaque base de données.

En appliquant la même méthodologie que celle présentée dans la section 2.3.3, nous ajustons les paramètres du modèle sur chaque base de données fictives.

2.3. Tests par simulation de la capacité d'apprentissage de l'outil d'apprentissage statistique modélisant un phénomène de croissance

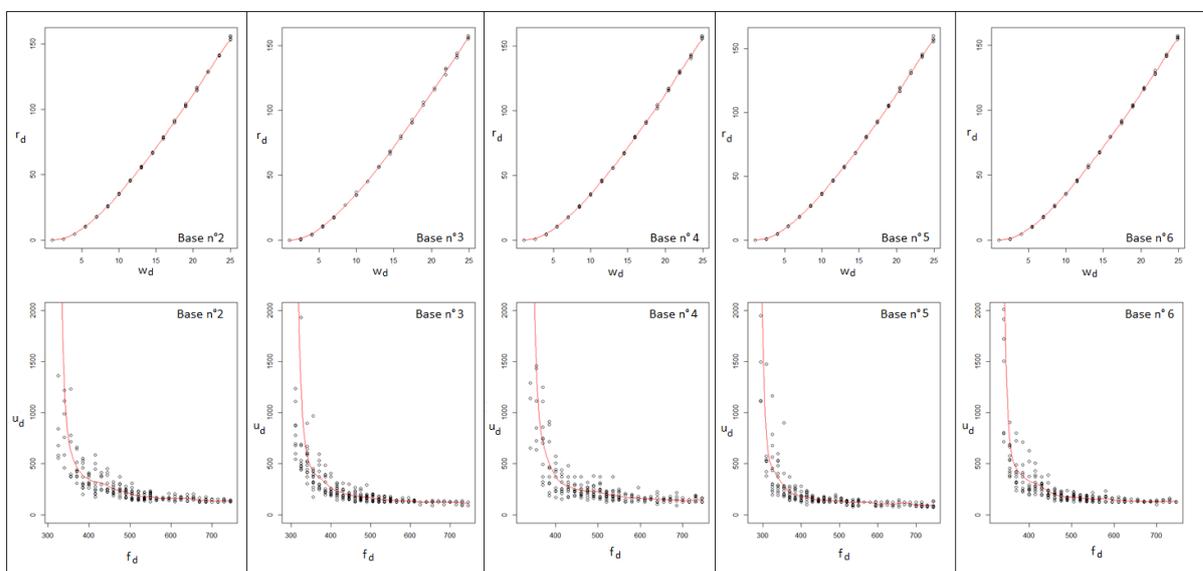


FIGURE 2.21 – Les relations non paramétriques reliant ω_d et r_d (Haut), et f_d et u_d (Bas) établies sur les nouvelles bases de données générées, numérotées de 2 à 6.

À la fin du processus d'apprentissage nous obtenons les valeurs moyennes des paramètres ω_a , r_a , f_a et u_a associées à chaque base de données. Nous paramétrons les 6 modèles avec ces valeurs moyennes et pour chaque modèle nous calculons la valeur du R^2 sur la base d'apprentissage. Nous avons obtenu la table 2.9.

TABLE 2.9 – La valeur des paramètres ajustés sur chaque base de données fictive

	Base n°1	Base n°2	Base n°3	Base n°4	Base n°5	Base n°6
ω_a	11.3	10.3	10.0	10.0	10.2	9.6
r_a	44.1	38.0	35.3	35.8	36.9	33.7
f_a	742.2	669.3	711.6	735.1	709.3	728.9
u_a	116.3	116.4	142.1	132.0	148.6	95.0
R^2	0.89	0.83	0.79	0.75	0.69	0.1

Afin d'estimer la robustesse de l'outil construit nous avons calculé pour chaque modèle ajusté sur chaque base de données générée, la valeur du R^2 sur la base de test. Sur les 6 bases de test, nous avons également calculé la valeur du R^2 associé aux six modèles "générateurs", c'est-à-dire les modèles ayant servi initialement à générer les six bases de données sur chaque base de test. Nous avons calculé la différence entre les valeurs de R^2 obtenues pour le modèle ajusté et le modèle "générateur" de

cette base. Nous avons obtenu la table 2.10.

TABLE 2.10 – Les valeurs de R^2 associées aux modèles ajustés et aux modèles "générateurs", calculées sur les 6 bases de données fictives et les mesures de l'écart existant entre ces modèles.

	Base n°1	Base n°2	Base n°3	Base n°4	Base n°5	Base n°6
Modèle ajusté	0.9050	0.8112	0.7736	0.7404	0.6713	0.0539
Modèle "générateur"	0.9051	0.8114	0.7736	0.7404	0.6715	0.0539
Différence	0.0001	0.0002	0.0000	0.0000	0.0001	0.0000

La table 2.10 et la figure 2.22 montrent que la différence entre la valeur du R^2 associée au modèle ajusté et celle du R^2 associée au modèle "générateur" est très faible quelle que soit la quantité de bruit ajouté. Cela montre que l'outil d'apprentissage statistique biomimétique est capable d'apprendre la valeur des paramètres du modèle malgré une quantité de bruit dans les données qui est parfois très élevée.

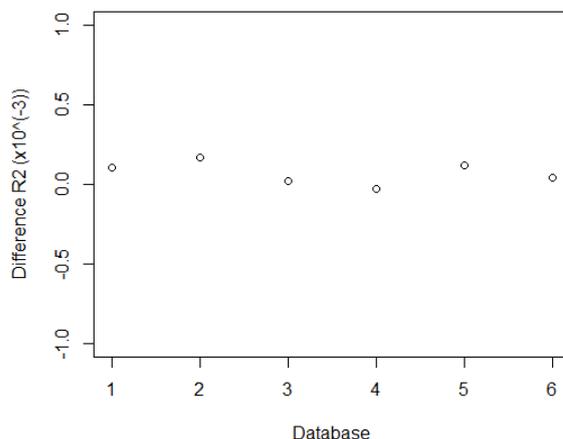


FIGURE 2.22 – Différence entre les valeurs de R^2 obtenues pour les modèles ajustés et les modèles "générateurs" (10^{-3}).

Cette étude a donc permis de mettre en évidence la robustesse du modèle en termes de capacité d'apprentissage, malgré la présence de bruit dans les données.

2.4 Tests par simulation de la capacité d'apprentissage des autres versions du modèle "A" : Le modèle "A d'augmentation" et de "diminution".

Dans la section 2.3 nous avons vérifié la capacité d'ajustement d'un modèle "A" intégrant l'équation d'utilisation (2.1.4b). Or deux autres versions de l'équation d'utilisation ont été présentées dans la Section 2.1.2. L'objectif de cette partie est de vérifier brièvement la capacité d'apprentissage des modèles "A" intégrant l'équation d'utilisation (2.1.4c) ("Modèle A d'augmentation") et l'équation d'utilisation (2.1.4d) ("Modèle A de diminution"). Nous avons également montré leur capacité d'extrapolation.

Pour cela nous avons commencé par construire d'autres bases de données fictives à partir de ces deux autres versions du modèle "A".

2.4.1 Génération de bases de données :

Nous avons commencé par générer des données à partir du "modèle A d'augmentation" et du "modèle A de diminution".

Pour chacun de ces modèles nous avons généré trois bases de données :

- Une base de données « d'apprentissage », correspondant à la simulation de la variation de la variable de sortie en réponse à une seule injection d'information. Cette base de données contient 10 courbes.
- Une base de données de « test », correspondant comme précédemment à la simulation d'une réaction induite par une seule injection d'information. Cette base de données contient 5 courbes.
- Une seconde base de données de « test », correspondant à la simulation de la variation de la variable de sortie en réponse à deux injections d'information. Cette base de données contient 5 courbes.

Pour chaque modèle, la base d'apprentissage nous a permis d'ajuster la valeur des paramètres sur des courbes de sortie induites par une seule injection d'information. La

première base de test nous a permis de vérifier la qualité du modèle obtenu et de s'assurer que le modèle ne surajuste pas les données d'apprentissage.

La seconde base de test modélise une réaction induite par deux injections d'information. Celle-ci nous a permis de voir si le modèle ajusté est capable de modéliser les variations de la variable de sortie induites par plusieurs injections d'information alors qu'il a été ajusté uniquement sur la réponse induite par une seule injection d'information. Cette seconde base de test nous a permis de savoir si les modèles de "diminution" et "d'augmentation" étaient capables de fournir des prédictions en dehors de la gamme des données ayant servi à les ajuster.

Pour générer ces différentes bases de données nous avons cette fois, fixé l'ensemble des paramètres du modèle. Chaque courbe des bases d'apprentissage et des premières base de test simule une réponse de la variable de sortie induite par une seule injection d'information. Ces courbes ont été générées à partir d'un couple $(VolQ, ct)$ donné. En revanche, les courbes des deuxièmes base de test simulent une réponse de la variable de sortie induite par deux injections d'information. chacune d'entre elles a donc été générée à partir de deux couples $(VolQ, ct)$ donnant le volume et l'instant des deux injections. Ainsi, en utilisant plusieurs couples $(VolQ, ct)$ nous avons généré plusieurs courbes de sortie. Nous avons ensuite bruité ces courbes en ajoutant un bruit gaussien suivant une loi Normale centrée en 0 et ayant une variance de 0.001.

Les figures 2.23 et 2.24 montrent des exemples de courbes contenues dans ces six nouvelles bases de données.

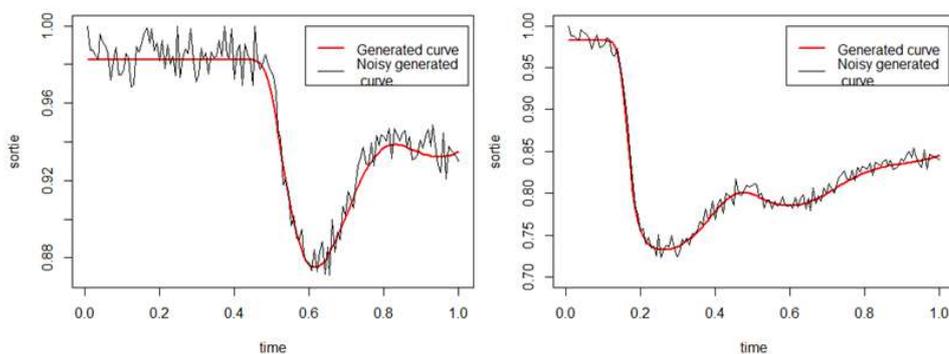


FIGURE 2.23 – Exemples de courbes de sortie générées par le modèle « de diminution » et induites par une seule (Gauche) ou deux injections d'information (Droite).

2.4. Tests par simulation de la capacité d'apprentissage des autres versions du modèle "A" :
Le modèle "A d'augmentation" et de "diminution".

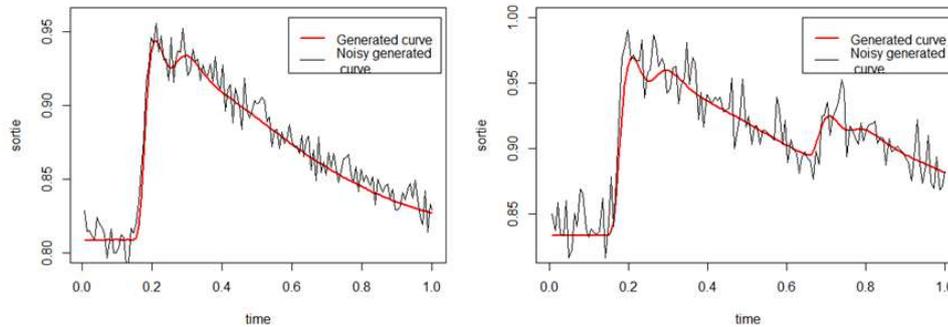


FIGURE 2.24 – Exemples de courbes de sorties générées par le modèle « d'augmentation » et induites par une seule (Gauche) ou deux injections d'information (Droite).

2.4.2 Ajustement des paramètres

Pour le "modèle A d'augmentation" et pour le "modèle A de diminution", nous avons choisi d'ajuster les paramètres ω_d et r_d sur les données de la base d'apprentissage. Les autres paramètres restent fixés.

Après une centaine d'itérations, nous avons obtenu des valeurs ajustées de ω_d et r_d pour ces deux modèles. Nous avons ensuite paramétré les modèles avec les valeurs obtenues de ω_d et r_d et nous avons testé leur qualité sur les données de la première base de test.

Comme dans la section 2.3.3, afin de voir si l'ajustement des paramètres s'est fait correctement nous avons comparé la précision des modèles obtenus après l'ajustement de ω_d et r_d (modèles "A" ajustés) et celle des modèles ayant été utilisés pour générer les données (modèles "A" "générateurs"). Les figures 2.25 et 2.26 nous montrent des exemples de courbes simulées à partir des modèles ajustés et comparées aux les courbes de la base de test. La table 2.11 contient les précisions relatives moyennes (*ARA*) calculées sur la première base de test et données par l'équation :

$$ARA = 1 - \frac{1}{n} \frac{1}{m} \sum_{j=1}^m \left(\sum_{i=1}^n \sqrt{\left(\frac{(O_{test,i,j} - O_{simul,i,j})}{O_{test,i,j}} \right)^2} \right). \quad (2.4.1)$$

où $O_{test,i,j}$ et $O_{simul,i,j}$ correspondent respectivement à la valeur de la courbe contenue dans la base de test et à celle de la courbe simulée pour l'individu i au point j . n est le nombre d'individus de la base de test et m est le nombre de points des courbes.

Les résultats contenus dans la table 2.11 montrent que l'ajustement des paramètres

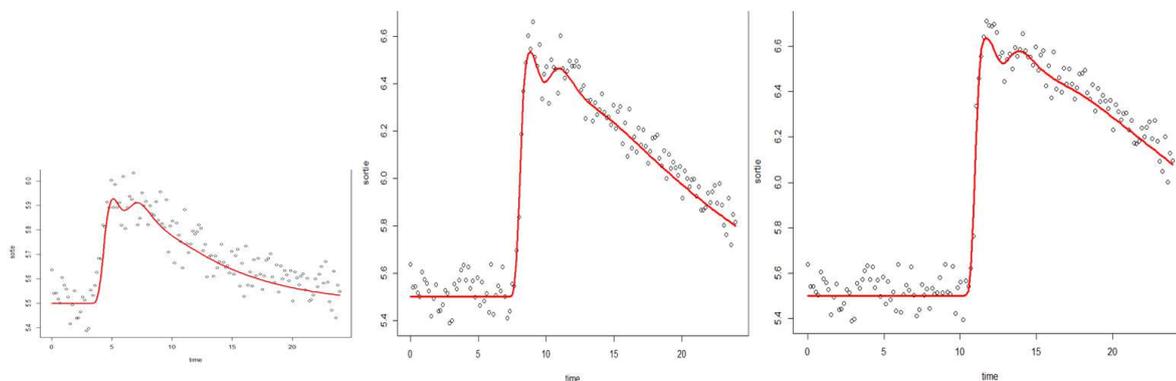


FIGURE 2.25 – Exemples de courbes simulées avec le modèle « d'augmentation » ajusté (–), en comparaison avec les courbes de la base de test (○).

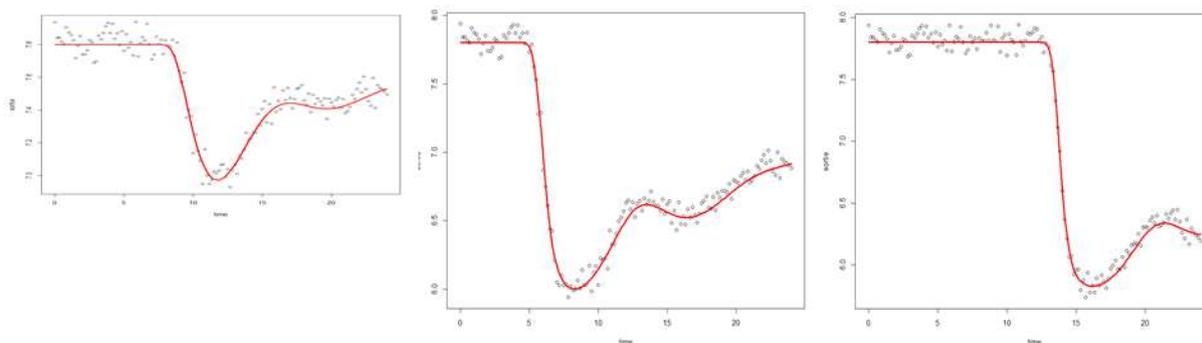


FIGURE 2.26 – Exemples de courbes simulées avec le modèle de « diminution » ajusté (–), en comparaison avec les courbes de la base de test (○).

TABLE 2.11 – Comparaison de la précision des modèles obtenus après ajustement (modèles "A" ajustés) et celle des modèles ayant été utilisés pour générer les données (modèles "A" "générateurs"), pour les modèles de "diminution" et "d'augmentation".

Modèles	Modèle "A" "d'augmentation"	Modèle "A" de "diminution"
Modèle "générateur"	97.99	95.42
Modèle ajusté	97.99	95.44

s'est fait correctement. En effet, que ça soit pour le modèle de "diminution" ou "d'augmentation", les différences de précision existantes entre les modèles "générateurs" et les modèles "ajustés" sont très faibles. La méthode d'ajustement mise en place permet donc de retrouver la valeur des paramètres minimisant l'erreur des modèles de "diminution" et "d'augmentation" à partir de données bruitées.

2.4.3 Capacité d'extrapolation des différentes versions du modèle "A"

Nous avons vu précédemment que nous avons ajusté les différentes versions du modèle "A" sur des données modélisant la variation de la variable de sortie induite par une seule injection d'information. L'objectif de cette étude est de voir si les modèles obtenus sont capables, sans réapprentissage, de simuler avec précision la variation de la variable de sortie induite cette fois par deux injections d'information.

Pour tester cette capacité d'extrapolation, nous avons appliqué les modèles ajustés précédemment sur les données de la seconde base de test, c'est-à-dire celle contenant des courbes modélisant des réponses induites par deux injections d'information. Pour cela nous avons utilisé les modèles ajustés précédemment pour simuler les courbes des secondes bases de test à partir des données d'entrée correspondant aux deux couples $(VolQ, ct)$ associés à chaque courbe.

Comme précédemment, nous avons calculé la précision des modèles ajustés (modèles "A" ajustés) et celle des modèles ayant été utilisés pour générer les données (modèles "A" "générateurs"). La figure 2.27 montrent des exemples de courbes simulées, en comparaison avec les courbes de la base de test. La table 2.12 contient les précisions relatives moyennes (ARA) données par l'équation (2.4.1) et calculées sur les secondes bases de test, pour les modèles "ajustés" et "générateurs".

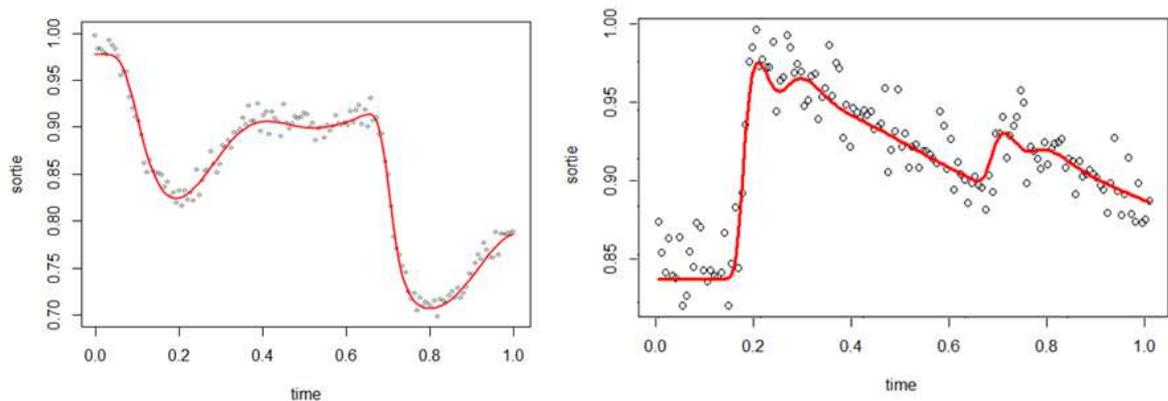


FIGURE 2.27 – Exemples de courbes simulées (–) avec le modèle « de diminution » (Gauche) et avec le modèle « d'augmentation » (Droite) ajustés, en comparaison avec les données de test modélisant la réponse à deux injections d'information (○).

TABLE 2.12 – Comparaison de la précision des modèles obtenus après ajustement (modèles "A" ajustés) et celle des modèles ayant été utilisés pour générer les données (modèles "A" "générateurs"), pour les modèles de "diminution" et "d'augmentation", sur les données simulant la réponse de la variable de sortie à deux injections d'information.

Modèles	Modèle "A" "d'augmentation"	Modèle "A" de "diminution"
Modèle "générateur"	98.83	95.79
Modèle ajusté	98.80	95.85

La différence de précision existante entre les modèles "générateurs" et les modèles "ajustés" est très faible, que ça soit pour le modèle de "diminution" ou "d'augmentation". Cela atteste que ces deux modèles "A" sont capables d'extrapoler et de modéliser la réponse de la variable de sortie à deux injections d'information, alors qu'ils n'ont été ajustés que sur des réponses induites par une seule injection d'information.

Il est donc possible d'appliquer ces modèles en dehors de la gamme de données ayant servi à les ajuster. En effet, il est possible grâce à ces modèles d'apprendre la dynamique reliant les données d'entrée et de sortie sur la réponse de la variable de sortie à une injection. Cette dynamique peut ensuite être appliquée pour simuler des réponses à plusieurs injections et donc sur des périodes plus longues que la période d'apprentissage.

2.5 Application de l'approche développée sur des données réelles

Dans cette section nous allons présenter une application de notre approche sur une base de données collectée dans une ferme expérimentale. Les données utilisées pour cette application correspondent aux données présentées dans la section 1.7.1.

L'objectif de cette application est de construire un outil capable de prédire la croissance logistique d'animaux en fonction de leur poids initial et de leur consommation alimentaire.

2.5.1 Le modèle mathématique utilisé

L'objectif de cette application est de prédire la croissance logistique d'animaux, nous choisissons donc comme équation d'«utilisation», l'équation (2.1.4b) contenant un coefficient limiteur. Le modèle mathématique utilisé correspond donc au modèle composé des équations (2.1.1), (2.1.2), (2.1.3), (2.1.4b) et (2.1.5).

Dans l'équation (2.1.4b) le paramètre L_a correspond au poids maximal pouvant être atteint par les individus de l'espèce et de la race sur lesquels les données ont été collectées. Les experts avec lesquels nous avons travaillé ont une idée assez précise de la valeur de ce paramètre. Au cours de l'ajustement, la valeur de L_a est donc chercher dans une gamme de valeurs très restreinte autour de la valeur attendue.

Au total il y a donc cinq paramètres à déterminer dans cette application : ω_a , r_a , f_a , u_a et L_a .

2.5.2 Les données utilisées

Dans cette application nous avons utilisé les données associées aux individus du groupe «Témoin A» pour ajuster le modèle et celles associées aux individus du groupe «Témoin B» pour tester la qualité du modèle obtenu (Voir la section 1.7.1).

Dans cette application nous ne connaissons pas la composition de la ration alimentaire distribuée aux animaux. Nous considérons donc la ration dans son ensemble. Dans cette application, l'information (d) évoquée dans la section 2.1.1 correspond à la ration distribuée dans sa globalité. Dans les deux groupes d'individus nous connaissons, la quantité moyenne d'aliment ingérée par chaque individu sur trois périodes ($C_{0-0.67}$, $C_{0.67-1.5}$ et $C_{1.5-2.5}$) (section 1.7.1). À partir de ces informations nous avons simulé une injection d'information plus régulière. C'est-à-dire que pour chaque individu, une information relative à sa consommation alimentaire est périodiquement injectée dans le modèle à chaque pas de temps $\Delta t = 0.16$. L'objectif est de répartir l'injection d'information mesurée sur des périodes plus courtes et ainsi lisser l'injection d'information.

TABLE 2.13 – Résumé des données utilisées.

	«Témoïn A»	«Témoïn B»
Nombre d'individus	8	7
Poids des individus mesurés à	$t = 0$ $t = 1$	$t = 0$ $t = 0.60$ $t = 1.52$ $t = 2.50$
Pas de temps d'injection de l'information	$\Delta t_{In} = 0.16$	$\Delta t_{In} = 0.16$

2.5.3 Etude des relations existantes entre les paramètres du modèle

Comme dans la section 2.3.2, nous avons analysé les relations existantes entre les paramètres du modèle sur les données du groupe «Témoïn A». Pour cela, nous avons appliqué la même méthodologie que celle appliquée sur les données simulées, mais cette fois sur les données réelles.

Grâce à une régression de Nadaraya-Watson, nous avons obtenu une relation non-paramétrique reliant $\omega_{d_{Opt}}$ et $r_{d_{Opt}}$ (Voir l'équation (2.5.1) et la figure 2.28). Connaissant la relation existante entre ces deux paramètres, il a été possible de déduire la valeur d'un de ces paramètres en fonction de la valeur du second.

De même, grâce à une régression de Nadaraya-Watson, nous avons obtenu une relation non-paramétrique reliant $f_{d_{Opt}}$ et $u_{d_{Opt}}$ (Voir l'équation (2.5.2) et la figure 2.28). Comme précédemment, connaissant cette relation reliant $f_{d_{Opt}}$ et $u_{d_{Opt}}$, il est possible de déduire la valeur de l'un en fonction de la valeur de l'autre.

$$r_{opt} = \hat{m}_{\omega_a Rel}(\omega_{opt}) + \epsilon, \quad (2.5.1)$$

$$u_{d_{Opt}} = \hat{m}_{f_a Rel}(f_{d_{Opt}}) + \epsilon, \quad (2.5.2)$$

2.5.4 Ajustement des paramètres du modèle

Grâce aux relations établies précédemment, nous n'avons ajusté que ω_a , f_a et L_a et nous en avons déduit les valeurs de r_a et u_a .

Pour ajuster les paramètres, nous avons utilisé l'algorithme DIRECT. Dans cette ap-

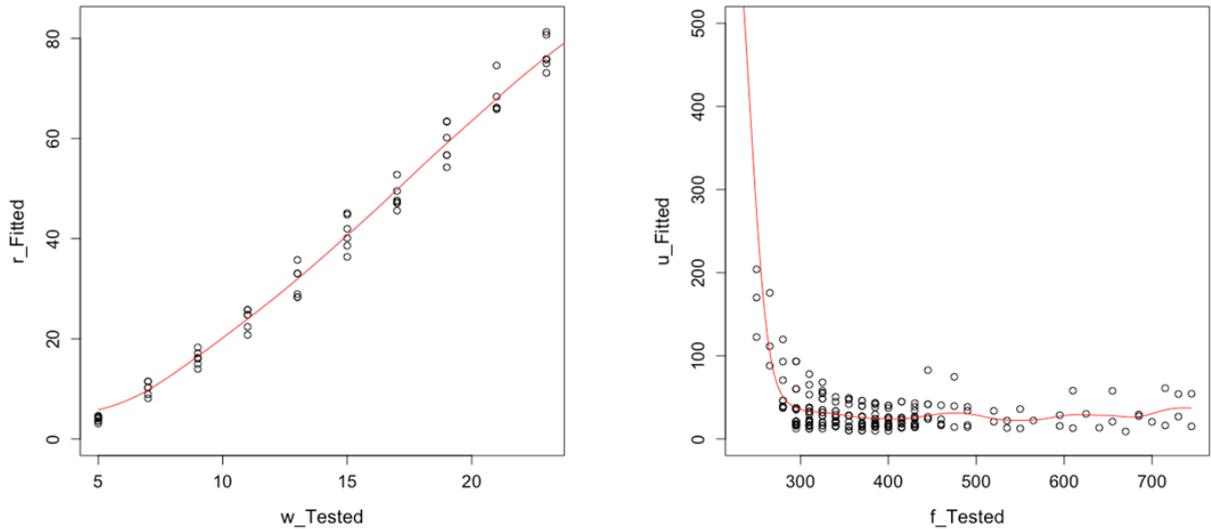


FIGURE 2.28 – Les régressions de Nadaraya-Watson reliant les couples $(\omega_{d_{Opt}}, r_{d_{Opt}})$ (gauche) et $(f_{d_{Opt}}, u_{d_{Opt}})$ (droite) sur les données d'apprentissage réelles.

plication, la fonction «objectif» à minimiser correspond à l'écart moyen au carré existant entre le poids des animaux prédit et celui mesuré l'instant $t = 1$. En effet, le poids à cet instant est la seule donnée renseignée pour les individus de la base d'apprentissage. Nous avons donc,

$$f_{obj}(\omega_a, r_a, f_a, u_a, L_a) = \frac{1}{n} \sum_{i=1}^n \left(\frac{(O_{i_{obs}}(1) - O_{i_{pred}}(1))}{O_{i_{obs}}(1)} \right)^2, \quad (2.5.3)$$

où n est le nombre d'individus de la base d'apprentissage et $O_{i_{obs}}(1)$ et $O_{i_{pred}}(1)$ correspondent respectivement au poids mesuré et au poids prédit à $t = 1$, pour le i^{ime} individu.

Pour tester la stabilité des valeurs du set de paramètres minimisant la fonction "objectif", nous avons réalisé plusieurs ajustements. Pour cela nous avons échantillonné la base d'apprentissage : Avant chaque ajustement nous avons sélectionné 7 individus sur les 8 de la base d'apprentissage et nous avons ajusté les paramètres sur les données associées à ces 7 individus sélectionnés. Nous avons donc réalisé 8 ajustements et nous avons obtenu 8 sets de paramètres ajustés $(\omega_a, r_a, f_a, u_a, L_a)$.

2.5.5 Résultats

À la fin de ces 8 ajustements nous avons calculé la moyenne et l'écart type relatif (*RSD*) associés à chaque paramètre (Table 2.14). Le *RSD* des paramètres est très faible. Cela signifie que notre méthode d'apprentissage permet d'identifier un unique set de paramètres minimisant l'erreur associée au modèle sur la base d'apprentissage. L'existence d'un unique set de paramètres optimaux $(\omega_a, r_a, f_a, u_a, L_a)$ atteste de l'identifiabilité du modèle.

Le modèle a été paramétré avec la valeur moyenne de chaque paramètre.

Dans un premier temps, nous avons calculé sur la base d'apprentissage l'écart relatif moyen (*ARE*) entre le poids mesuré et celui prédit à l'instant $t = 1$. Cet indicateur est donné par l'équation :

$$ARE(t) = \frac{1}{n} \sum_{i=1}^n \sqrt{\left(\frac{O_{i_{obs}}(t) - O_{i_{pred}}(t)}{O_{i_{obs}}(t)} \right)^2} \quad (2.5.4)$$

où n est le nombre d'individus dans la base de données et $O_{i_{obs}}(t)$ et $O_{i_{pred}}(t)$ correspondent respectivement aux poids mesuré et prédit à l'instant t pour l'individu i .

TABLE 2.14 – La moyenne et l'écart type relatif (*RSD*) des paramètres ajustés, et l'indicateur *ARE* calculé à l'instant $t = 1$ sur la base d'apprentissage.

Paramtre	Moyenne	RSD
ω	19.51	0.055
r	61.29	0.080
f	240.0	0.001
u	561.4	0.007
L_a	1.71	0.011
<i>ARE</i> (1) (%)	1.83	0.013

La valeur calculée de l'indicateur *ARE* à l'instant $t = 1$ sur la base d'apprentissage vaut 1.83%, ce qui est satisfaisant (Table 2.14). La précision du modèle a ensuite été calculée sur la base de test afin de tester la qualité du modèle et vérifier que le modèle ne surajuste pas les données d'apprentissage.

Pour cela nous avons calculé l'indicateur *ARE* sur la base de test aux instants

$t = 0.6$, $t = 1.52$ et $t = 2.5$ (Voir la table 2.15 et la figure 2.29).

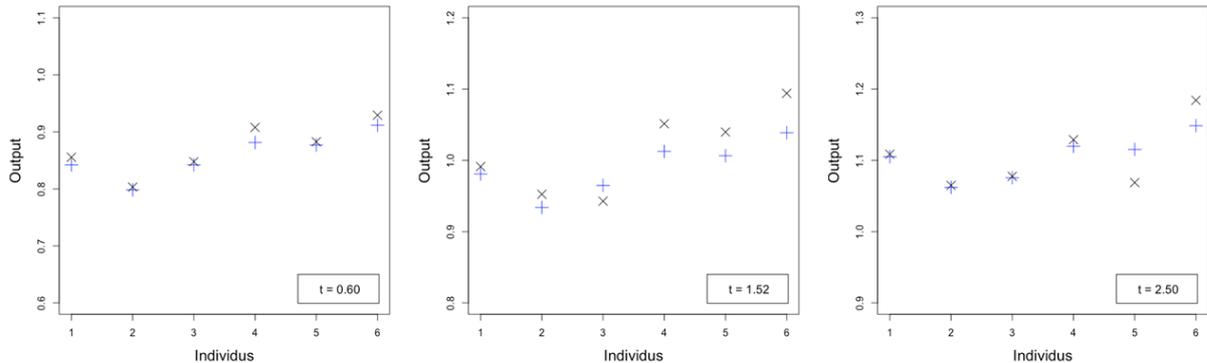


FIGURE 2.29 – Différences obtenues entre les valeurs de la variable de sortie mesurées (+) et celles prédites (×), à différents instants t pour les individus de la base de test.

TABLE 2.15 – Erreur relative moyenne (ARE) calculée à différents instants t pour les individus de la base de test.

t	0.6	1.52	2.5
$ARE(t)$ (%)	1.3	2.9	1.5

2.5.6 Discussion des résultats

L'erreur associée au modèle est faible sur la base d'apprentissage (Table 2.14) mais aussi sur la base de test (Table 2.15). Ces résultats montrent que le modèle s'ajuste correctement aux données et est capable de prédire avec précision le poids des individus à l'instant $t = 1$. Ces résultats montrent aussi que malgré un ajustement sur un très petit jeu de données, notre outil ne surajuste pas les données d'apprentissage.

Étant donné que nous avons très peu de données disponibles, les paramètres du modèle n'ont été ajustés que sur le poids des individus à l'instant $t = 1$, à partir de leur poids initial et de leur consommation alimentaire moyenne. Malgré cela nous avons remarqué que l'outil obtenu est capable de prédire avec précision le poids des animaux de la base de test avant et au-delà de l'instant $t = 1$ (Table 2.15). Ces résultats montrent que le modèle construit est capable d'apprendre une dynamique reliant les données d'entrées et les données de sorties sur une certaine période, et de rester valable sur une durée 2.5 fois plus longue que la période d'apprentissage. De ce fait, nous

pouvons conclure que l'intégration de connaissances dans un modèle parcimonieux confère à l'outil, en plus d'une capacité d'interpolation, une capacité d'extrapolation.

Cette capacité d'extrapolation, obtenue malgré un apprentissage sur très peu de données, illustre le fort potentiel d'extraction d'information de l'outil construit. Comme cela le sera confirmé dans le chapitre 5 consacré au positionnement de notre outil de prédiction, cette capacité représente un important avantage distinctif de cet outil par rapport aux outils existants.

Cette capacité d'extrapolation contribue également à réduire la dépendance de nos outils aux données d'apprentissage. En effet, celle-ci permet d'ajuster ces outils en utilisant des séries temporelles définies sur une période donnée et de les appliquer ensuite en dehors de cette période d'apprentissage. Ainsi, les outils peuvent être ajustés sur des bases d'apprentissage plus petites. Cette capacité d'extrapolation permet donc de réduire la durée de la collecte de données, la durée des essais expérimentaux et ainsi de réduire les coûts expérimentaux et computationnels liés aux données.

La capacité d'extrapolation de nos outils permet de rallonger la durée de validité des prédictions fournies par l'outil de prédiction. Dans le cadre de l'Assimilation de Données, comme nous l'avons vu dans la section 1.4, les prédictions fournies par l'outil sont régulièrement corrigées par des données collectées. Grâce à la capacité d'extrapolation de nos outils ces corrections pourraient être moins fréquentes.

2.6 Conclusion du chapitre 2

Pour conclure ce chapitre nous pouvons dire que nous avons construit des outils biomimétiques d'apprentissage statistique, basés sur un système d'EDP, embarquant l'expression mathématique de processus biologiques clés.

Plusieurs versions de ce système d'EDP ont été élaborées de façon à pouvoir traiter une gamme plus large de données d'entrée et de sortie.

Des études par simulation ont mis en évidence un phénomène de saturation de l'effet des paramètres sur le modèle mathématique. À partir de ces effets de saturation et des conditions CFL, nous avons établi la gamme de valeurs des différents paramètres du modèle.

Ces études par simulation ont également mis en évidence des relations existantes entre les paramètres. Un travail plus approfondi a ensuite permis d'établir ces relations

et de les quantifier. Ces relations a permis de renforcer la procédure d'ajustement du modèle. En effet, en connaissant la relation reliant les paramètres $\omega_{d_{Opt}}$ et $r_{d_{Opt}}$ et celle reliant $f_{d_{Opt}}$ et $u_{d_{Opt}}$, nous pouvons désormais apprendre les valeurs de $\omega_{d_{Opt}}$ et $f_{d_{Opt}}$ sur des données et d'en déduire les valeurs de $u_{d_{Opt}}$ et $f_{d_{Opt}}$. Cela réduit donc le nombre de paramètres à apprendre simultanément. Nous verrons également dans le chapitre 6 que la prise en compte de ces relations a permis de réduire le nombre d'itérations nécessaires pour ajuster le modèle. Cela a donc rendu le processus d'apprentissage plus rapide.

Les tests par simulation et l'application de notre approche sur des données réelles ont montré que nos outils possèdent une capacité de prédiction satisfaisante.

Néanmoins nous verrons dans le chapitre 5 que les outils biomimétiques d'apprentissage statistique que nous avons construits se distinguent réellement des outils existants par leur capacité d'interpolation et surtout d'extrapolation, par leur capacité d'apprentissage et plus globalement par leur capacité à réaliser de l'Assimilation de Données d'élevage.

La capacité d'extrapolation de nos outils couplée à une capacité d'apprentissage très forte, permet de réduire la dépendance des outils aux données d'apprentissage. Cela permet de réduire la quantité de données à collecter et d'espérer pouvoir réaliser de l'Assimilation de Données précise même si la fréquence de collecte de données est basse et la qualité des données collectées est faible.

EXISTENCE, UNICITÉ ET PROPRIÉTÉS QUALITATIVES DU SYSTÈME D'EDP EMBARQUÉ DANS NOS OUTILS D'APPRENTISSAGE STATISTIQUE BIOMIMÉTIQUES.

3.1 Introduction

Dans le chapitre 2, les applications réalisées ont montré que les outils basés sur le modèle "A" possèdent d'intéressantes capacités de prédiction et d'extrapolation. Nous avons donc voulu réaliser l'analyse mathématique de ce modèle composé des équations (2.1.1), (2.1.2), (2.1.3), (2.1.5), et des équations "d'utilisation" (2.1.4) ou (2.1.4b).

Cependant, dans les deux premières équations (2.1.1) et (2.1.2), le terme assimilable à un opérateur de diffusion dépend de deux densités inconnues, $\{\Phi_f(d)\}$ et $\{\Phi_b(d)\}$. Une étude approfondie du travail réalisé par LADYENSKAJA *et al.* (1968) nous a permis de mettre en évidence que cette dépendance à deux inconnues, couplée à la présence d'un terme de convection, rend impossible l'application des résultats classiques associés à l'analyse des équations paraboliques, apportés notamment par ces auteurs. De ce fait, l'analyse mathématique de notre modèle constitue un nouveau défi dans le domaine de l'analyse des EDP.

Au vu de cette complexité, ce chapitre présentera des résultats préliminaires concernant les propriétés du système d'EDP, composé des équations (2.1.1), (2.1.2), (2.1.3), (2.1.5), et des équations "d'utilisation" (2.1.4) ou (2.1.4b).

Pour rappel, χ est une fonction positive qui ne dépend que de x , qui est continuellement

dérivable et compacte sur l'intervalle $(0, 1)$:

$$\chi \geq 0, \chi \in C_0^1((0, 1)). \quad (3.1.1)$$

La fonction $\{F(d)\}$ est une "fonction de localisation" et la fonction $\{Q(d)\}$ est une "fonction d'entrée" (Section 2.1.1). Ces deux fonctions sont positives, carré intégrables sur l'intervalle $(0, 1)$ et leur support est inclus dans le support de la fonction χ :

$$\{F(d)\} \in L^2((0, 1)), \{Q(d)\} \in L^\infty((0, +\infty), L^2((0, 1))). \quad (3.1.2)$$

$$\{F(d)\} \geq 0, \{Q(d)\} \geq 0. \quad (3.1.3)$$

$$Supp(\{F(d)\}) \subset Supp(\chi), Supp(\{Q(d)\}) \subset Supp(\chi). \quad (3.1.4)$$

La première contribution de notre travail d'analyse correspond à l'établissement de la Formulation Variationnelle, aussi appelée Formulation Faible, des équations (2.1.1) et (2.1.2). Cette formulation consiste à chercher $\{\Phi_f(d)\}$ et $\{\Phi_b(d)\}$ dans l'espace \mathbf{H} que nous avons défini tel que,

$$\mathbf{H} = H^1((0, +\infty) \times (0, 1)) \cap C^0((0, +\infty), H^1(0, 1)) \cap C^1((0, +\infty), H^{-1}(0, 1)), \quad (3.1.5)$$

de façon à ce que, pour toutes fonctions $\varphi_f \in \mathbf{H}$ et $\varphi_b \in \mathbf{H}$, l'égalité suivante soit vraie

pour tout $T > 0$,

$$\begin{aligned}
& - \int_0^T \int_0^1 \{\Phi_f(\mathbf{d})\}(t, x) \frac{\partial \varphi_f}{\partial t}(t, x) dx dt - \int_0^T \int_0^1 \{\Phi_b(\mathbf{d})\}(t, x) \frac{\partial \varphi_b}{\partial t}(t, x) dx dt \\
& + \int_0^1 \{\Phi_f(\mathbf{d})\}(0, x) \varphi_f(0, x) dx - \int_0^1 \{\Phi_f(\mathbf{d})\}(T, x) \varphi_f(T, x) dx + \int_0^1 \{\Phi_b(\mathbf{d})\}(0, x) \varphi_b(0, x) dx \\
& - \int_0^1 \{\Phi_b(\mathbf{d})\}(T, x) \varphi_b(T, x) dx - \int_0^T \omega_a \{\Phi_f(\mathbf{d})\}(t, 1) \varphi_f(t, 1) dt + \int_0^T \omega_a \{\Phi_f(\mathbf{d})\}(t, 0) \varphi_f(t, 0) dt \\
& \quad + \int_0^T \omega_a \{\Phi_b(\mathbf{d})\}(t, 1) \varphi_b(t, 1) - \int_0^T \omega_a \{\Phi_b(\mathbf{d})\}(t, 0) \varphi_b(t, 0) \\
& = \int_0^T \int_0^1 \omega_a \{\Phi_f(\mathbf{d})\}(t, x) \frac{\partial \varphi_f}{\partial x}(t, x) dx dt - \int_0^T \int_0^1 \omega_a \{\Phi_b(\mathbf{d})\}(t, x) \frac{\partial \varphi_b}{\partial x}(t, x) dx dt \\
& \quad - 2 \int_0^T \int_0^1 c_a \chi \left(\frac{\partial(\{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\})}{\partial x}(t, x) \right) \left(\frac{\partial(\varphi_f + \varphi_b)}{\partial x}(t, x) \right) dx dt \\
& \quad + \frac{1}{2} \int_0^T \int_0^1 \{Q(\mathbf{d})\}(t, x) (\varphi_f(t, x) + \varphi_b(t, x)) dx dt \\
& \quad - \int_0^T \int_0^1 f_a \{F(\mathbf{d})\}(x) (\{\Phi_f(\mathbf{d})\}(t, x) \varphi_f(t, x) + \{\Phi_b(\mathbf{d})\}(t, x) \varphi_b(t, x)) dx dt \\
& \quad - \int_0^T \int_0^1 r_a \{\Phi_f(\mathbf{d})\}(t, x) \varphi_f(t, x) dx dt + \int_0^T \int_0^1 r_a \{\Phi_b(\mathbf{d})\}(t, x) \varphi_b(t, x) dx dt. \quad (3.1.6)
\end{aligned}$$

À partir de cette Formulation Variationnelle, nous avons établi les estimations a priori classiques qui restent vraies malgré les particularités des équations de notre modèle. Parmi elles, il y a l'estimation L^2 suivante :

Théorème 3.1.1. *Sous les hypothèses (2.1.7), (3.1.1) et (3.1.2), si $\{\Phi_f(\mathbf{d})\} \in \mathbf{H}$ et $\{\Phi_b(\mathbf{d})\} \in \mathbf{H}$ sont solutions de la Formulation Variationnelle donnée par l'équation (3.1.6), alors, pour tout $T > 0$, il existe trois constantes $L_{2a}(T)$, $L_{2b}(T)$ et $L_3(T)$ telles que pour tout $t \in (0, T)$, $T > 0$,*

$$\|\{\Phi_f(\mathbf{d})\}\|_2^2(t) \leq L_{2a}(T), \quad (3.1.7)$$

$$\|\{\Phi_b(\mathbf{d})\}\|_2^2(t) \leq L_{2b}(T), \quad (3.1.8)$$

et,

$$\left\| \left\| \sqrt{\chi} \frac{\partial (\{\Phi_f(d)\} + \{\Phi_b(d)\})}{\partial x} \right\| \right\|_2^2 \leq L_3(T), \quad (3.1.9)$$

où,

$$\|\Phi\|_2^2(t) = \int_0^1 \Phi^2(t, x) dx$$

et,

$$\left\| \left\| \frac{\partial \Phi}{\partial x} \right\| \right\|_2^2 = \int_0^T \int_0^1 \frac{\partial \Phi^2}{\partial x}(t, x) dx dt$$

La démonstration de ce théorème sera donnée dans la section 3.2.3.

Les équations (2.1.3), (2.1.5), et les équations "d'utilisation" (2.1.4) ou (2.1.4b) ne contiennent pas de particularités et ne posent donc pas de réels problèmes pour mener à bien l'analyse du modèle mathématique.

Le modèle mathématique composé des équations (2.1.1), (2.1.2), (2.1.3), (2.1.4) et (2.1.5), possède une propriété de conservation :

Propriété 3.1.1. *Sous les hypothèses (3.1.1) et (3.1.2), si $\{\Phi_f(d)\} \in \mathbf{H}$ et $\{\Phi_b(d)\} \in \mathbf{H}$ sont solutions de l'équation (3.1.6), $\{\Psi(d)\} \in C^1((0, +\infty); L^1((0, 1)))$ est solution de l'équation (2.1.3), $\{\Xi(d)\} \in C^1((0, +\infty); L^1((0, 1)))$ est solution de l'équation (2.1.4), et $\{O(d)\} \in C^1((0, +\infty))$ est solution de l'équation (2.1.5), alors pour tout $x \in (0, 1)$ et pour tout $t \in (0, T)$, $T > 0$, il existe $K_Q(T) > 0$ telle que,*

$$\frac{\partial \left[\int_0^1 \{\Phi_f(\mathbf{d})\}(t, x) dx + \int_0^1 \{\Phi_b(\mathbf{d})\}(t, x) dx + \int_0^1 \{\Psi(\mathbf{d})\}(t, x) dx + \int_0^1 \{\Xi(\mathbf{d})\}(t, x) dx \right]}{\partial t} \leq K_Q(T). \quad (3.1.10)$$

Remarque. L'estimation (3.1.10) fournirait une estimation L^1 dans le cas où nous serions capables de démontrer la positivité des fonctions $\{\Phi_f(\mathbf{d})\}$, $\{\Phi_b(\mathbf{d})\}$, $\{\Psi(\mathbf{d})\}$ et $\{\Xi(\mathbf{d})\}$.

Cependant, nous verrons dans la suite que cette question de la positivité est une réelle problématique dans le cadre de notre modèle. De plus, l'obtention de l'estimation (3.1.10) dans le cas où l'on remplace l'équation (2.1.4) par l'équation (2.1.4b) n'est pas évidente.

La démonstration de cette propriété sera donnée dans la section 3.2.1.

Au-delà de la question de la positivité, d'autres propriétés typiquement vérifiées dans le cadre des EDP paraboliques ne sont pas vérifiables pour notre système d'EDP. De ce fait, nous avons travaillé sur deux modèles simplifiés afin de mettre en évidence les propriétés perdues et celles que nous pouvons espérer conserver malgré la présence de ce terme de diffusion dépendant de la somme de deux densités inconnues.

Le premier modèle que nous avons étudié est composé de deux équations. Il considère $\{\Phi_f(\mathbf{d})\}$ et $\{\Phi_b(\mathbf{d})\}$, définies pour $x \in (0, 1)$ et $t \in (0, +\infty)$, sont respectivement solutions de,

$$\frac{\partial \{\Phi_f(\mathbf{d})\}}{\partial t}(t, x) - c_d \frac{\partial^2 [\{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\}]}{\partial x^2}(t, x) = 0, \quad (3.1.11)$$

et,

$$\frac{\partial \{\Phi_b(\mathbf{d})\}}{\partial t}(t, x) - c_d \frac{\partial^2 [\{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\}]}{\partial x^2}(t, x) = 0. \quad (3.1.12)$$

Dans ce modèle nous avons imposé pour tout $x \in (0, 1)$,

$$\{\Phi_f(\mathbf{d})\}(0, x) \geq 0 \text{ and } \{\Phi_b(\mathbf{d})\}(0, x) \geq 0, \quad (3.1.13)$$

et pour tout $t \in (0, +\infty)$ et pour toutes fonctions positives $\overline{\Phi}_f^0(t)$, $\overline{\Phi}_b^0(t)$, $\overline{\Phi}_f^1(t)$ et $\overline{\Phi}_b^1(t)$,

$$\begin{aligned} \{\Phi_f(\mathbf{d})\}(t, 0) &= \overline{\Phi}_f^0(t) \geq 0 \text{ et } \{\Phi_b(\mathbf{d})\}(t, 0) = \overline{\Phi}_b^0(t) \geq 0 \\ \{\Phi_f(\mathbf{d})\}(t, 1) &= \overline{\Phi}_f^1(t) \geq 0 \text{ et } \{\Phi_b(\mathbf{d})\}(t, 1) = \overline{\Phi}_b^1(t) \geq 0 \end{aligned} \quad (3.1.14)$$

La perte de l'effet régularisant, typiquement apporté par l'opérateur de diffusion, a été mise en évidence à travers l'étude de ce premier sous-modèle. Cette perte est illustrée par le fait que si $\overline{\Phi}_f^0(t) = 1$, $\overline{\Phi}_f^1(t) = 0$, $\overline{\Phi}_b^0(t) = 0$ et $\overline{\Phi}_b^1(t) = 1$, alors, pour tout $\tilde{x} \in (0, 1)$, le couple de fonctions

$$\{\Phi_f(\mathbf{d})\}(t, x) = \begin{cases} 1, & \text{si } x < \tilde{x} \\ 0, & \text{sinon} \end{cases} \quad (3.1.15)$$

et,

$$\{\Phi_b(\mathbf{d})\}(t, x) = \begin{cases} 0, & \text{si } x < \tilde{x} \\ 1, & \text{sinon} \end{cases} \quad (3.1.16)$$

est solution des équations (3.1.11), (3.1.12), (3.1.13) et (3.1.14). Or cette solution n'est clairement pas régulière.

Néanmoins, pour ce premier sous-modèle nous avons démontré que :

Propriété 3.1.2. Si $\overline{\Phi}_f^0(t) = \overline{\Phi}_f^1(t) = \overline{\Phi}_b^0(t) = \overline{\Phi}_b^1(t) = 0$, $\{\Phi_f(\mathbf{d})\}(0, \cdot) \in L^2((0, 1))$, $\{\Phi_b(\mathbf{d})\}(0, \cdot) \in L^2((0, 1))$ et $(\{\Phi_f(\mathbf{d})\}(0, \cdot) - \{\Phi_b(\mathbf{d})\}(0, \cdot)) \in H^1(0, 1)$, alors il existe une unique solution aux équations (3.1.11), (3.1.12), (3.1.13) et (3.1.14). Cette solution appartient à $C^1((0, +\infty), H^1(0, 1))$.

Cette propriété, démontrée dans la section 3.3.2, établit un cadre pour assurer la régularité des solutions.

L'analyse de ce premier modèle simplifié nous a aussi permis de montrer que la positivité des fonctions $\{\Phi_f(\mathbf{d})\}$ et $\{\Phi_b(\mathbf{d})\}$ n'est pas assurée même dans le cadre établi

par la propriété 3.1.2. Ceci sera illustré par un exemple donné dans la section 3.3.3.

Ensuite, en reformulant les équations (3.1.2), (3.1.3), (3.1.11), (3.1.12), (3.1.13) et (3.1.14) en utilisant les Transformations de Fourier, nous avons démontré certaines propriétés d'un deuxième modèle simplifié, plus proche du système d'EDP initial. En effet, ce deuxième modèle simplifié correspond au modèle initial mais sans terme de convection. De plus, les conditions aux bords (3.1.14) ont été remplacé par des conditions de périodicité pour $\{\Phi_f(d)\}$ et $\{\Phi_b(d)\}$ par rapport à x . Nous avons également introduit l'espace $\mathbf{H}_\#$, défini tel que :

$$\mathbf{H}_\# = \{\Phi \in \mathbf{H}, \forall t \in (0, +\infty), \Phi(t, 0) = \Phi(t, 1)\} \quad (3.1.17)$$

L'espace $\mathbf{H}_\#$ peut être vu comme l'ensemble des fonctions de \mathbf{H} qui sont périodiques et qui respectent la topologie de l'espace \mathbf{H} .

Ainsi, dans ce deuxième sous-modèle, $\{\Phi_f(d)\} \in \mathbf{H}_\#$ est une fonction périodique de période 1 par rapport à x , solution de,

$$\begin{aligned} \frac{\partial \{\Phi_f(d)\}}{\partial t}(t, x) - c_a \frac{\partial^2 [\{\Phi_f(d)\} + \{\Phi_b(d)\}]}{\partial x^2}(t, x) \\ = \frac{1}{2} \{Q(d)\}(t, x) - f_a \{\Phi_f(d)\}(t, x) - r_a \{\Phi_f(d)\}(t, x), \end{aligned} \quad (3.1.18)$$

et, $\{\Phi_b(d)\} \in \mathbf{H}_\#$ est une fonction périodique de période 1 par rapport à x , solution de,

$$\begin{aligned} \frac{\partial \{\Phi_b(d)\}}{\partial t}(t, x) - c_a \frac{\partial^2 [\{\Phi_f(d)\} + \{\Phi_b(d)\}]}{\partial x^2}(t, x) \\ = \frac{1}{2} \{Q(d)\}(t, x) - f_a \{\Phi_b(d)\}(t, x) + r_a \{\Phi_f(d)\}(t, x). \end{aligned} \quad (3.1.19)$$

Les trois autres équations du modèle restent les équations (2.1.3), (2.1.5) et l'équation (2.1.4b) ou (2.1.4).

Dans le cadre de l'étude de ce second sous-modèle nous supposons que,

$$\{Q(d)\} \in L^\infty((0, +\infty), L^2((0, 1))), \{Q(d)\} \geq 0, \quad (3.1.20)$$

$$\{\Phi_f(\mathbf{d})\}(0, \cdot) \geq 0 \text{ et } \{\Phi_b(\mathbf{d})\}(0, \cdot) \geq 0 \quad (3.1.21)$$

$$\{\Phi_f(\mathbf{d})\}(0, \cdot) + \{\Phi_b(\mathbf{d})\}(0, \cdot) \leq \frac{1}{f_d} \sup_{x \in (0,1)} (\{Q(\mathbf{d})\}(0, x)), \quad (3.1.22)$$

$$\{\Phi_f(\mathbf{d})\}(0, \cdot) \in L^\infty((0, 1)), \{\Phi_b(\mathbf{d})\}(0, \cdot) \in L^\infty((0, 1)), \quad (3.1.23)$$

$$\{\Phi_f(\mathbf{d})\}(0, \cdot) - \{\Phi_b(\mathbf{d})\}(0, \cdot) \in H^1((0, 1)), \quad (3.1.24)$$

$$\{\Phi_f(\mathbf{d})\}(0, 0) - \{\Phi_b(\mathbf{d})\}(0, 0) = \{\Phi_f(\mathbf{d})\}(0, 1) - \{\Phi_b(\mathbf{d})\}(0, 1), \quad (3.1.25)$$

$$\{\Psi(\mathbf{d})\}(0, \cdot) \in L^\infty((0, 1)), \{\Psi(\mathbf{d})\}(0, \cdot) \geq 0, \quad (3.1.26)$$

$$\{\Xi(\mathbf{d})\}(0, \cdot) \in L^\infty((0, 1)), \{\Xi(\mathbf{d})\}(0, \cdot) \geq 0, \quad (3.1.27)$$

$$0 \leq \{O(\mathbf{d})\}(0) \leq L_d, \quad (3.1.28)$$

et,

$$\exists T_F > 0, \forall t > T_F, \forall x \in (0, 1), \{Q(\mathbf{d})\}(t, x) = 0. \quad (3.1.29)$$

En utilisant ce deuxième modèle simplifié nous avons démontré que :

Théorème 3.1.2. *Sous les hypothèses allant de (3.1.20) jusqu'à (3.1.29), il existe un système unique composé de $\{\Phi_f(\mathbf{d})\} \in \mathbf{H}_\#, \{\Phi_b(\mathbf{d})\} \in \mathbf{H}_\#, \{\Psi(\mathbf{d})\} \in C^1((0, +\infty); L^1((0, 1)))$, $\{\Xi(\mathbf{d})\} \in C^1((0, +\infty); L^1((0, 1)))$ et $\{O(\mathbf{d})\} \in C^1((0, +\infty))$, solution du système d'EDP composé des équations (3.1.18), (3.1.19), (2.1.3), (2.1.5) et de l'équation (2.1.4b) ou (2.1.4). De plus, ce système d'EDP possède une propriété de conservation. Ainsi, pour*

tout $T > 0$, il existe une constante $K_{Q_2}(T) > 0$ telle que pour tout $t \in (0, T)$,

$$\frac{\partial \left[\int_0^1 \{\Phi_f(\mathbf{d})\}(t, x) dx + \int_0^1 \{\Phi_b(\mathbf{d})\}(t, x) dx + \int_0^1 \{\Psi(\mathbf{d})\}(t, x) dx + \int_0^1 \{\Xi(\mathbf{d})\}(t, x) dx \right]}{\partial t} \leq K_{Q_2}(T). \quad (3.1.30)$$

La démonstration du théorème 3.1.2 sera présentée dans la section 3.4.3.

Dans la section 3.2, nous présenterons la démonstration de la propriété de conservation du modèle initialement construit dans le chapitre 2. Dans cette section le principe de la Formulation Variationnelle appliqué à notre modèle sera expliqué. Après avoir donné du sens aux différentes équations du système d'EDP en définissant l'espace fonctionnel adapté dans lequel les solutions sont cherchées, les résultats des estimations a priori seront présentés. Ces estimations a priori ont permis de mettre en évidence les particularités du modèle initialement construit. Ces particularités empêchent l'application directe des résultats classiques concernant les équations paraboliques pour montrer l'existence et l'unicité de la solution de ce modèle.

Dans la section 3.3, nous présenterons les propriétés perdues en raison de la structure du terme assimilable à l'opérateur de diffusion. Pour cela nous avons analysé le premier modèle simplifié.

Dans la section 3.4 nous allons présenter l'analyse du deuxième modèle simplifié qui ne contient pas de terme de convection. Pour cela nous utiliserons des méthodes de changement de variables et de Transformation de Fourier. L'objectif de cette analyse a été de mettre en évidence les propriétés conservées malgré la structure de l'opérateur de diffusion. Dans cette section nous présenterons la démonstration de l'existence et de l'unicité de la solution de ce deuxième modèle simplifié. Nous avons également démontré la propriété de conservation de ce second sous-modèle et de la positivité de certaines composantes de la solution.

3.2 Structure et analyse du modèle mathématique développé initialement

3.2.1 Propriété de conservation du modèle mathématique développé initialement

Le modèle mathématique développé initialement est composé des équations (2.1.1), (2.1.2), (2.1.3), (2.1.4) et (2.1.5). Ce modèle possède une propriété de conservation.

Démonstration de la propriété 3.1.1 : Pour démontrer cette propriété nous avons supposé que $\{\Phi_f(\mathbf{d})\}$, $\{\Phi_b(\mathbf{d})\}$, $\{\Psi(\mathbf{d})\}$ et $\{\Xi(\mathbf{d})\}$ sont respectivement solutions de (2.1.1), (2.1.2), (2.1.3) et (2.1.4). Nous avons additionné et intégré ces quatre équations par rapport à x . Nous avons obtenu l'égalité suivante :

$$\int_0^1 \frac{\partial \left(\{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\} + \{\Psi(\mathbf{d})\} + \{\Xi(\mathbf{d})\} \right)}{\partial t} (t, x) dx$$

$$= \int_0^1 2 \frac{\partial \left[c_a \chi \frac{\partial \left[\{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\} \right]}{\partial x} \right]}{\partial x} (t, x) dx + \int_0^1 \{Q(\mathbf{d})\} (t, x) dx. \quad (3.2.1)$$

D'après la condition donnée par (3.1.1), la fonction χ est nulle en $x = 0$ et $x = 1$. En intégrant par rapport à x sur l'intervalle $(0, 1)$ nous obtenons donc,

$$\frac{\partial \left[\int_0^1 \{\Phi_f(\mathbf{d})\} (t, x) dx + \int_0^1 \{\Phi_b(\mathbf{d})\} (t, x) dx + \int_0^1 \{\Psi(\mathbf{d})\} (t, x) dx + \int_0^1 \{\Xi(\mathbf{d})\} (t, x) dx \right]}{\partial t}$$

$$= \int_0^1 \{Q(\mathbf{d})\} (t, x) dx \leq \sup_{t \in \mathbb{R}^+} \left(\int_0^1 \{Q(\mathbf{d})\} (t, x) dx \right). \quad (3.2.2)$$

Nous avons déduit de cette inégalité que pour tout x appartenant à $(0, 1)$ et pour tout t appartenant à $(0, T)$, où T appartient à $(0, +\infty)$, il existe une constante $K_Q(T) > 0$ satisfaisant l'inégalité (3.1.10). Nous avons donc démontré la propriété 3.1.1. ■

3.2.2 Principe de la Formulation Variationnelle

Dans le but de démontrer l'existence et l'unicité de la solution du système d'EDP développé initialement, il est essentiel de donner un sens aux différentes équations. Pour cela, il est nécessaire de définir l'espace mathématique dans lequel les solutions de ce système seront cherchées.

Au vu de la structure des équations (2.1.1) et (2.1.2), les deux premières composantes de la solution du système d'EDP pourraient être intuitivement cherchées dans des espaces tels que l'espace $C^1((0, +\infty), C^2(0, 1))$, puisque ces équations contiennent le Laplacien de $\{\Phi_f(d)\}$ et $\{\Phi_b(d)\}$. On parle alors de "solutions classiques". Cependant, les espaces de Sobolev sont plus adaptés pour analyser les propriétés des EDP et démontrer l'existence, la régularité et l'unicité de leur solution (ADAMS (1975), DI MENZA (2009), DAVID et GOSSELET (2015) et TAYLOR (1996)). En effet, nous cherchons les solutions dans des espaces plus grands que les espaces classiques C^m . Les espaces de Sobolev sont des outils très utiles pour l'analyse des EDP car dans ces espaces les dérivées sont prises dans un sens "faible", c'est-à-dire au sens des distributions, afin de rendre l'espace complet. Ces espaces fournissent donc un cadre adapté pour démontrer l'existence et l'unicité de la solution.

En travaillant dans ces espaces, la notion de "solution" peut être "affaiblie". Pour cela, il est nécessaire d'écrire une formulation faible des équations (2.1.1) et (2.1.2). En effet, l'objectif de la Formulation Variationnelle est d'affaiblir la notion de solution en multipliant les équations par une fonction choisie judicieusement et appelée "fonction test" (BRÉZIS *et al.* (1999) et ALLAIRE (2005)). L'intégration de ces équations au sens des distributions permet ensuite d'obtenir une équation qui implique des dérivées d'ordre plus faible qu'initialement.

Toujours dans cette optique de définir une solution au sens faible des équations (2.1.1) et (2.1.2), nous avons considéré ces deux équations comme étant des égalités de distributions et nous avons appliqué l'équation (2.1.1) à la fonction test $\varphi_f \in \mathcal{D}((0, +\infty) \times (0, 1))$ et l'équation (2.1.2) à la fonction test $\varphi_b \in \mathcal{D}((0, +\infty) \times (0, 1))$. L'espace $\mathcal{D}((0, +\infty) \times (0, 1))$ est l'espace topologique des fonction infiniment continuellement dérivables, à support compact dans $((0, +\infty) \times (0, 1))$. Ensuite, en utilisant la définition de l'égalité et de la dérivation au sens des distributions, nous avons obtenu :

$\forall \varphi_f \in \mathcal{D}((0, +\infty) \times (0, 1))$,

$$\begin{aligned} & -\langle \{\Phi_f(\mathbf{d})\}, \partial_t \varphi_f \rangle - \omega_a \langle \{\Phi_f(\mathbf{d})\}, \partial_x \varphi_f \rangle + c_a \langle \chi(\partial_x \{\Phi_f(\mathbf{d})\} + \partial_x \{\Phi_b(\mathbf{d})\}), \partial_x \varphi_f \rangle \\ & = \frac{1}{2} \langle \{Q(\mathbf{d})\}, \varphi_f \rangle - f_a \langle \{\Phi_f(\mathbf{d})\}, \varphi_f \rangle - r_a \langle \{\Phi_f(\mathbf{d})\}, \varphi_f \rangle, \end{aligned} \quad (3.2.3)$$

et $\forall \varphi_b \in \mathcal{D}((0, +\infty) \times (0, 1))$,

$$\begin{aligned} & -\langle \{\Phi_b(\mathbf{d})\}, \partial_t \varphi_b \rangle + \omega_a \langle \{\Phi_b(\mathbf{d})\}, \partial_x \varphi_b \rangle + c_a \langle \chi(\partial_x \{\Phi_f(\mathbf{d})\} + \partial_x \{\Phi_b(\mathbf{d})\}), \partial_x \varphi_b \rangle \\ & = \frac{1}{2} \langle \{Q(\mathbf{d})\}, \varphi_b \rangle - f_a \langle \{\Phi_b(\mathbf{d})\}, \varphi_b \rangle + r_a \langle \{\Phi_f(\mathbf{d})\}, \varphi_b \rangle. \end{aligned} \quad (3.2.4)$$

Au vu des Formulations Variationnelles données par les équations (3.2.3) et (3.2.4), il semble raisonnable de supposer que les solutions faibles des équations (2.1.1) et (2.1.2) sont des fonctions appartenant à l'espace $L^2((0, +\infty) \times (0, 1))$ et dont leur dérivées appartiennent également à $L^2((0, +\infty) \times (0, 1))$. Ainsi, nous supposons que les deux premières composantes de la solution du système d'EDP appartiennent à l'espace de Sobolev $H^1((0, +\infty) \times (0, 1))$, et respectent les conditions suivantes :

$$\forall t \in (0, +\infty), \quad \{\Phi_f(\mathbf{d})\}(t, 0) = \{\Phi_b(\mathbf{d})\}(t, 0), \quad (3.2.5)$$

et,

$$\forall t \in (0, +\infty), \quad \{\Phi_b(\mathbf{d})\}(t, 1) = \{\Phi_f(\mathbf{d})\}(t, 1). \quad (3.2.6)$$

Dans ce contexte, les équations (3.2.3) et (3.2.4) peuvent être réécrites telles que :

$\forall \varphi_f \in H_0^1((0, +\infty) \times (0, 1))$,

$$\begin{aligned} & - \int_0^{+\infty} \int_0^1 \{\Phi_f(\mathbf{d})\}(t, x) \frac{\partial \varphi_f}{\partial t}(t, x) dx dt - \int_0^{+\infty} \int_0^1 \omega_a \{\Phi_f(\mathbf{d})\}(t, x) \frac{\partial \varphi_f}{\partial x}(t, x) dx dt \\ & \quad + \int_0^{+\infty} \int_0^1 c_a \chi(x) \frac{\partial [\{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\}]}{\partial x}(t, x) \frac{\partial \varphi_f}{\partial x}(t, x) dx dt \\ & = \int_0^{+\infty} \int_0^1 \frac{1}{2} \{Q(\mathbf{d})\}(t, x) \varphi_f(t, x) dx dt - \int_0^{+\infty} \int_0^1 f_a \{F(\mathbf{d})\}(x) \{\Phi_f(\mathbf{d})\}(t, x) \varphi_f(t, x) dx dt \\ & \quad - \int_0^{+\infty} \int_0^1 r_a \{\Phi_f(\mathbf{d})\}(t, x) \varphi_f(t, x) dx dt, \end{aligned} \quad (3.2.7)$$

et $\forall \varphi_b \in H_0^1((0, +\infty) \times (0, 1))$,

$$\begin{aligned}
 & - \int_0^{+\infty} \int_0^1 \{\Phi_b(\mathbf{d})\}(t, x) \frac{\partial \varphi_b}{\partial t}(t, x) dx dt + \int_0^{+\infty} \int_0^1 \omega_a \{\Phi_b(\mathbf{d})\}(t, x) \frac{\partial \varphi_b}{\partial x}(t, x) dx dt \\
 & \quad + \int_0^{+\infty} \int_0^1 c_a \chi(x) \frac{\partial [\{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\}]}{\partial x}(t, x) \frac{\partial \varphi_b}{\partial x}(t, x) dx dt \\
 & = \int_0^{+\infty} \int_0^1 \frac{1}{2} \{Q(\mathbf{d})\}(t, x) \varphi_b(t, x) dx dt - \int_0^{+\infty} \int_0^1 f_a \{F(\mathbf{d})\}(x) \{\Phi_b(\mathbf{d})\}(t, x) \varphi_b(t, x) dx dt \\
 & \quad + \int_0^{+\infty} \int_0^1 r_a \{\Phi_f(\mathbf{d})\}(t, x) \varphi_b(t, x) dx dt, \quad (3.2.8)
 \end{aligned}$$

Dans le but de prendre en compte des conditions initiales et les conditions aux bords, nous avons enrichi les formulations (3.2.7) et (3.2.8) pour obtenir des formulations pour les fonctions test ayant des traces non-nulles en $\{(0, x), x \in (0, 1)\}$, $\{(t, 0), t \in (0, +\infty)\}$ et en $\{(t, 1), t \in (0, +\infty)\}$. Nous avons donc obtenu les formulations suivantes :

$\forall \varphi_f \in \mathbf{H}$,

$$\begin{aligned}
 & - \int_0^T \int_0^1 \{\Phi_f(\mathbf{d})\}(t, x) \frac{\partial \varphi_f}{\partial t}(t, x) dx dt + \int_0^1 \{\Phi_f(\mathbf{d})\}(0, x) \varphi_f(0, x) dx - \int_0^1 \{\Phi_f(\mathbf{d})\}(T, x) \varphi_f(T, x) dx \\
 & - \int_0^T \int_0^1 \omega_a \{\Phi_f(\mathbf{d})\}(t, x) \frac{\partial \varphi_f}{\partial x}(t, x) dx dt - \int_0^T \omega_a \{\Phi_f(\mathbf{d})\}(t, 1) \varphi_f(t, 1) dt + \int_0^T \omega_a \{\Phi_f(\mathbf{d})\}(t, 0) \varphi_f(t, 0) dt \\
 & \quad + \int_0^T \int_0^1 c_a \chi(x) \frac{\partial [\{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\}]}{\partial x}(t, x) \frac{\partial \varphi_f}{\partial x}(t, x) dx dt \\
 & = \int_0^T \int_0^1 \frac{1}{2} \{Q(\mathbf{d})\}(t, x) \varphi_f(t, x) dx dt - \int_0^T \int_0^1 f_a \{F(\mathbf{d})\}(x) \{\Phi_f(\mathbf{d})\}(t, x) \varphi_f(t, x) dx dt \\
 & \quad - \int_0^T \int_0^1 r_a \{\Phi_f(\mathbf{d})\}(t, x) \varphi_f(t, x) dx dt, \quad (3.2.9)
 \end{aligned}$$

et $\forall \varphi_b \in \mathbf{H}$,

$$\begin{aligned}
 & - \int_0^T \int_0^1 \{\Phi_b(\mathbf{d})\}(t, x) \frac{\partial \varphi_b}{\partial t}(t, x) dx dt + \int_0^1 \{\Phi_b(\mathbf{d})\}(0, x) \varphi_b(0, x) dx - \int_0^1 \{\Phi_b(\mathbf{d})\}(T, x) \varphi_b(T, x) dx \\
 & + \int_0^T \int_0^1 \omega_a \{\Phi_b(\mathbf{d})\}(t, x) \frac{\partial \varphi_b}{\partial x}(t, x) dx dt + \int_0^T \omega_a \{\Phi_b(\mathbf{d})\}(t, 1) \varphi_b(t, 1) - \int_0^T \omega_a \{\Phi_b(\mathbf{d})\}(t, 0) \varphi_b(t, 0) \\
 & \quad + \int_0^T \int_0^1 c_a \chi(x) \frac{\partial [\{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\}]}{\partial x}(t, x) \frac{\partial \varphi_b}{\partial x}(t, x) dx dt \\
 & = \int_0^T \int_0^1 \frac{1}{2} \{Q(\mathbf{d})\}(t, x) \varphi_b(t, x) dx dt - \int_0^T \int_0^1 f_a \{F(\mathbf{d})\}(x) \{\Phi_b(\mathbf{d})\}(t, x) \varphi_b(t, x) dx dt \\
 & \quad + \int_0^T \int_0^1 r_a \{\Phi_f(\mathbf{d})\}(t, x) \varphi_b(t, x) dx dt. \quad (3.2.10)
 \end{aligned}$$

Finalement, en additionnant (3.2.9) et (3.2.10) nous avons obtenu la Formulation Variationnelle globale donnée par l'équation (3.1.6).

Au vu de cette Formulation Variationnelle globale, $\{\Phi_f(\mathbf{d})\}$, $\{\Phi_b(\mathbf{d})\}$, $\frac{\partial \{\Phi_f(\mathbf{d})\}}{\partial x}$ et $\frac{\partial \{\Phi_b(\mathbf{d})\}}{\partial x}$ doivent appartenir à l'espace $L^2((0, +\infty) \times (0, 1))$. Ainsi, $\{\Phi_f(\mathbf{d})\}$ et $\{\Phi_b(\mathbf{d})\}$ doivent appartenir à $H^1((0, +\infty) \times (0, 1))$.

De plus, certaines conditions aux bords sont imposées pour $x = 0$ et $x = 1$. L'existence des traces en $(0, +\infty) \times \{0\}$ et $(0, +\infty) \times \{1\}$ pour les fonctions de cet espace amène le sens faible à donner à ces conditions aux bords.

Les formulations initiales (2.1.1) et (2.1.2) font intervenir le Laplacien de $\{\Phi_f(\mathbf{d})\}$ et $\{\Phi_b(\mathbf{d})\}$. Or, le Laplacien d'une fonction appartenant à $H^1((0, 1))$, appartient à $H^{-1}(0, 1)$. $\{\Phi_f(\mathbf{d})\}(t, \cdot)$ et $\{\Phi_b(\mathbf{d})\}(t, \cdot)$ doivent donc être des fonctions continues et dérivables à valeurs dans $H^{-1}((0, 1))$, pour tout $t \in (0, +\infty)$. Ainsi, $\{\Phi_f(\mathbf{d})\}$ et $\{\Phi_b(\mathbf{d})\}$ doivent appartenir à $C^1((0, +\infty), H^{-1}((0, 1)))$. Sachant qu'il semble raisonnable de chercher les fonctions $\{\Phi_f(\mathbf{d})\}$ et $\{\Phi_b(\mathbf{d})\}$ dans $H^1((0, +\infty) \times (0, 1))$ et dans $C^1((0, +\infty), H^{-1}((0, 1)))$, nous pouvons donc nous attendre à ce que les solutions que nous cherchons appartiennent également à l'espace $C^0((0, +\infty), H^1((0, 1)))$. En effet, l'appartenance des fonctions $\{\Phi_f(\mathbf{d})\}$ et $\{\Phi_b(\mathbf{d})\}$ à cet espace est important pour pouvoir mener à bien certaines estimations a priori nécessitant des calculs à t fixé et en intégrant les équations par rapport à x .

Il semble donc raisonnable de penser que les fonctions $\{\Phi_f(\mathbf{d})\}$ et $\{\Phi_b(\mathbf{d})\}$ appartiennent à l'espace \mathbf{H} , défini par (3.1.5).

Pour résumer, l'espace \mathbf{H} est l'espace des fonctions appartenant à $L^2((0, +\infty) \times (0, 1))$, telles que leur dérivées d'ordre 1 appartiennent aussi à $L^2((0, +\infty) \times (0, 1))$, et qui sont continues, dérivables et appartenant à $H^1(0, 1)$ et $H^{-1}(0, 1)$ à chaque instant $t \in (0, +\infty)$.

Finalement, nous avons considéré que $\{\Phi_f(\mathbf{d})\}$, $\{\Phi_b(\mathbf{d})\}$, φ_f et φ_b étaient cherchées dans \mathbf{H} telles que la formulation (3.1.6) soit respectée.

3.2.3 Les premières estimations a priori

Dans cette section, les premières estimations a priori obtenues seront présentées. Ces estimations nous ont permis de démontrer le théorème 3.1.1.

Démonstration du théorème 3.1.1 : Nous avons remarqué que si les fonctions $\{\Phi_f(\mathbf{d})\}$ et $\{\Phi_b(\mathbf{d})\}$, solutions de (2.1.1) et (2.1.2), existent et appartiennent à \mathbf{H} , alors elles peuvent être utilisées comme fonctions test dans les formulations (3.2.9) et (3.2.10). Dans un premier temps, nous avons donc choisi $\varphi_f = \{\Phi_f(\mathbf{d})\}$ et $\varphi_b = \{\Phi_b(\mathbf{d})\}$. Les équations (3.2.9) et (3.2.10) deviennent donc respectivement les équations :

$$\begin{aligned}
 & - \int_0^T \int_0^1 \{\Phi_f(\mathbf{d})\}(t, x) \frac{\partial \{\Phi_f(\mathbf{d})\}}{\partial t}(t, x) dx dt + \int_0^1 \{\Phi_f(\mathbf{d})\}^2(0, x) dx - \int_0^1 \{\Phi_f(\mathbf{d})\}^2(T, x) dx \\
 & - \int_0^T \int_0^1 \omega_a \{\Phi_f(\mathbf{d})\}(t, x) \frac{\partial \{\Phi_f(\mathbf{d})\}}{\partial x}(t, x) dx dt - \int_0^T \omega_a \{\Phi_f(\mathbf{d})\}^2(t, 1) dt + \int_0^T \omega_a \{\Phi_f(\mathbf{d})\}^2(t, 0) dt \\
 & \quad + \int_0^T \int_0^1 c_a \chi(x) \frac{\partial [\{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\}]}{\partial x}(t, x) \frac{\partial \{\Phi_f(\mathbf{d})\}}{\partial x}(t, x) dx dt \\
 & = \int_0^T \int_0^1 \frac{1}{2} \{Q(\mathbf{d})\}(t, x) \{\Phi_f(\mathbf{d})\}(t, x) dx dt - \int_0^T \int_0^1 f_a \{F(\mathbf{d})\}(x) \{\Phi_f(\mathbf{d})\}(t, x) \{\Phi_f(\mathbf{d})\}(t, x) dx dt \\
 & \quad - \int_0^T \int_0^1 r_a \{\Phi_f(\mathbf{d})\}(t, x) \{\Phi_f(\mathbf{d})\}(t, x) dx dt, \quad (3.2.11)
 \end{aligned}$$

et,

$$\begin{aligned}
 & - \int_0^T \int_0^1 \{\Phi_b(\mathbf{d})\}(t, x) \frac{\partial \{\Phi_b(\mathbf{d})\}}{\partial t}(t, x) dx dt + \int_0^1 \{\Phi_b(\mathbf{d})\}^2(0, x) dx - \int_0^1 \{\Phi_b(\mathbf{d})\}^2(T, x) dx \\
 & + \int_0^T \int_0^1 \omega_a \{\Phi_b(\mathbf{d})\}(t, x) \frac{\partial \{\Phi_b(\mathbf{d})\}}{\partial x}(t, x) dx dt + \int_0^T \omega_a \{\Phi_b(\mathbf{d})\}^2(t, 1) dt - \int_0^T \omega_a \{\Phi_b(\mathbf{d})\}^2(t, 0) dt \\
 & \quad + \int_0^T \int_0^1 c_a \chi(x) \frac{\partial [\{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\}]}{\partial x}(t, x) \frac{\partial \{\Phi_b(\mathbf{d})\}}{\partial x}(t, x) dx dt \\
 & = \int_0^T \int_0^1 \frac{1}{2} \{Q(\mathbf{d})\}(t, x) \{\Phi_b(\mathbf{d})\}(t, x) dx dt - \int_0^T \int_0^1 f_a \{F(\mathbf{d})\}(x) \{\Phi_b(\mathbf{d})\}(t, x) \{\Phi_b(\mathbf{d})\}(t, x) dx dt \\
 & \quad + \int_0^T \int_0^1 r_a \{\Phi_f(\mathbf{d})\}(t, x) \{\Phi_b(\mathbf{d})\}(t, x) dx dt. \quad (3.2.12)
 \end{aligned}$$

L'objectif est maintenant de vérifier à partir des équations (3.2.11) et (3.2.12) que les estimations a priori obtenues n'amènent pas de contradiction avec l'hypothèse que $\{\Phi_f(\mathbf{d})\}$ et $\{\Phi_b(\mathbf{d})\}$ appartiennent à \mathbf{H} .

Dans un premier temps nous allons présenter les étapes ayant permis d'obtenir des estimations dans l'espace $L^\infty((0, T), L^2((0, 1)))$, pour tout $T \in (0, +\infty)$.

Les équations (3.2.11) et (3.2.12) dépendent l'une de l'autre. De ce fait, il est nécessaire d'étudier les solutions de ces deux équations en même temps. Nous avons

donc additionné ces deux équations et nous avons obtenu l'égalité suivante :

$$\begin{aligned}
 & - \int_0^T \int_0^1 \{\Phi_f(\mathbf{d})\}(t, x) \frac{\partial \{\Phi_f(\mathbf{d})\}}{\partial t}(t, x) dx dt - \int_0^T \int_0^1 \{\Phi_b(\mathbf{d})\}(t, x) \frac{\partial \{\Phi_b(\mathbf{d})\}}{\partial t}(t, x) dx dt \\
 & + \int_0^1 \{\Phi_f(\mathbf{d})\}^2(0, x) dx - \int_0^1 \{\Phi_f(\mathbf{d})\}^2(T, x) dx + \int_0^1 \{\Phi_b(\mathbf{d})\}^2(0, x) dx - \int_0^1 \{\Phi_b(\mathbf{d})\}^2(T, x) dx \\
 & - \int_0^T \omega_a \{\Phi_f(\mathbf{d})\}^2(t, 1) dt + \int_0^T \omega_a \{\Phi_f(\mathbf{d})\}^2(t, 0) dt + \int_0^T \omega_a \{\Phi_b(\mathbf{d})\}^2(t, 1) dt - \int_0^T \omega_a \{\Phi_b(\mathbf{d})\}^2(t, 0) dt \\
 & = \int_0^T \int_0^1 \omega_a \{\Phi_f(\mathbf{d})\}(t, x) \frac{\partial \{\Phi_f(\mathbf{d})\}}{\partial x}(t, x) dx dt - \int_0^T \int_0^1 \omega_a \{\Phi_b(\mathbf{d})\}(t, x) \frac{\partial \{\Phi_b(\mathbf{d})\}}{\partial x}(t, x) dx dt \\
 & \quad - \int_0^T \int_0^1 c_a \chi(x) \left(\frac{\partial \{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\}}{\partial x}(t, x) \right)^2 dx dt \\
 & \quad + \int_0^T \int_0^1 \{Q(\mathbf{d})\}(t, x) \left(\{\Phi_f(\mathbf{d})\}(t, x) + \{\Phi_b(\mathbf{d})\}(t, x) \right) dx dt \\
 & \quad - \int_0^T \int_0^1 f_a \{F(\mathbf{d})\}(x) \left(\{\Phi_f(\mathbf{d})\}^2(t, x) + \{\Phi_b(\mathbf{d})\}^2(t, x) \right) dx dt \\
 & \quad - \int_0^T \int_0^1 r_a \{\Phi_f(\mathbf{d})\}^2(t, x) dx dt + \int_0^T \int_0^1 r_a \{\Phi_f(\mathbf{d})\}(t, x) \{\Phi_b(\mathbf{d})\}(t, x) dx dt. \quad (3.2.13)
 \end{aligned}$$

À partir de l'équation (3.2.13), lorsque nous calculons les intégrales nous obtenons :

$$\begin{aligned}
 & \int_0^T \left(\frac{\partial \|\{\Phi_f(\mathbf{d})\}\|_2^2}{\partial t}(t, \cdot) + \frac{\partial \|\{\Phi_b(\mathbf{d})\}\|_2^2}{\partial t}(t, \cdot) \right) dt = \\
 & \int_0^1 2 \left(\{\Phi_f(\mathbf{d})\}^2(0, x) dx - \{\Phi_f(\mathbf{d})\}^2(T, x) dx + \{\Phi_b(\mathbf{d})\}^2(0, x) dx - \{\Phi_b(\mathbf{d})\}^2(T, x) \right) dx \\
 & + \int_0^T 2\omega_a \left(- \{\Phi_b(\mathbf{d})\}^2(t, 1) + \{\Phi_b(\mathbf{d})\}^2(t, 0) + \{\Phi_f(\mathbf{d})\}^2(t, 1) - \{\Phi_f(\mathbf{d})\}^2(t, 0) \right) dt \\
 & \quad - 2 \int_0^T \int_0^1 c_a \chi(x) \left(\frac{\partial \left(\{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\} \right)}{\partial x}(t, x) \right)^2 dx dt \\
 & \quad + 2 \int_0^T \int_0^1 \{Q(\mathbf{d})\}(t, x) \left(\{\Phi_f(\mathbf{d})\}(t, x) + \{\Phi_b(\mathbf{d})\}(t, x) \right) dx dt \\
 & \quad - 2 \int_0^T \int_0^1 f_a \{F(\mathbf{d})\}(x) \left(\{\Phi_f(\mathbf{d})\}^2(t, x) + \{\Phi_b(\mathbf{d})\}^2(t, x) \right) dx dt \\
 & \quad - 2 \int_0^T \int_0^1 r_a \{\Phi_f(\mathbf{d})\}^2(t, x) dx dt + 2 \int_0^T \int_0^1 r_a \{\Phi_f(\mathbf{d})\}(t, x) \{\Phi_b(\mathbf{d})\}(t, x) dx dt, \quad (3.2.14)
 \end{aligned}$$

où $\|\cdot\|_2$ est la norme usuelle sur $L^2((0, 1))$.

En prenant en compte les conditions aux bords (2.1.7) et la positivité de la fonction χ , nous obtenons à partir de l'équation (3.2.14) l'inégalité suivante :

$$\begin{aligned} \int_0^T \left(\frac{\partial \|\{\Phi_f(\mathbf{d})\}\|_2^2}{\partial t}(t, \cdot) + \frac{\partial \|\{\Phi_b(\mathbf{d})\}\|_2^2}{\partial t}(t, \cdot) \right) dt &\leq -2 \int_0^T \int_0^1 c_a \chi \left(\frac{\partial (\{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\})}{\partial x}(t, x) \right)^2 dx dt \\ &+ 2 \int_0^T \int_0^1 \{Q(\mathbf{d})\}(t, x) (\{\Phi_f(\mathbf{d})\}(t, x) + \{\Phi_b(\mathbf{d})\}(t, x)) dx dt \\ &+ 2 \int_0^T \int_0^1 r_a \{\Phi_f(\mathbf{d})\}(t, x) \{\Phi_b(\mathbf{d})\}(t, x) dx dt + K_1, \quad (3.2.15) \end{aligned}$$

où K_1 est une constante positive.

En utilisant l'inégalité de Schwarz nous obtenons pour tout $t \in (0, T)$, $T \in (0, +\infty)$:

$$\left| \int_0^1 \{Q(\mathbf{d})\}(t, x) \{\Phi_f(\mathbf{d})\}(t, x) dx \right| \leq \|\{Q(\mathbf{d})\}(t, \cdot)\|_2 \|\{\Phi_f(\mathbf{d})\}(t, \cdot)\|_2, \quad (3.2.16)$$

$$\left| \int_0^1 \{Q(\mathbf{d})\}(t, x) \{\Phi_b(\mathbf{d})\}(t, x) dx \right| \leq \|\{Q(\mathbf{d})\}(t, \cdot)\|_2 \|\{\Phi_b(\mathbf{d})\}(t, \cdot)\|_2, \quad (3.2.17)$$

et,

$$\left| \int_0^1 \{\Phi_f(\mathbf{d})\}(t, x) \{\Phi_b(\mathbf{d})\}(t, x) dx \right| \leq \|\{\Phi_f(\mathbf{d})\}(t, \cdot)\|_2 \|\{\Phi_b(\mathbf{d})\}(t, \cdot)\|_2. \quad (3.2.18)$$

Ainsi, à partir de l'inégalité (3.2.15) et en utilisant les inégalités (3.2.16), (3.2.17) et (3.2.18), nous obtenons :

$$\int_0^T \left(\frac{\partial \|\{\Phi_f(\mathbf{d})\}\|_2^2}{\partial t}(t, \cdot) + \frac{\partial \|\{\Phi_b(\mathbf{d})\}\|_2^2}{\partial t}(t, \cdot) \right) dt \leq 2 \int_0^T \left(r_a \|\{\Phi_f(\mathbf{d})\}\|_2 \|\{\Phi_b(\mathbf{d})\}\|_2 + \|\{Q(\mathbf{d})\}\|_2 \left(\|\{\Phi_f(\mathbf{d})\}\|_2 + \|\{\Phi_b(\mathbf{d})\}\|_2 \right) \right) dt + K_1 \quad (3.2.19)$$

À partir de l'inégalité (3.2.19) et en utilisant les identités remarquables nous obtenons :

$$\int_0^T \left(\frac{\partial \|\{\Phi_f(\mathbf{d})\}\|_2^2}{\partial t}(t, \cdot) + \frac{\partial \|\{\Phi_b(\mathbf{d})\}\|_2^2}{\partial t}(t, \cdot) \right) dt \leq \int_0^T \left(r_a \left(\|\{\Phi_f(\mathbf{d})\}\|_2^2 + \|\{\Phi_b(\mathbf{d})\}\|_2^2 \right) + 2 \|\{Q(\mathbf{d})\}\|_2 \left(\sqrt{2 \left(\|\{\Phi_f(\mathbf{d})\}\|_2^2 + \|\{\Phi_b(\mathbf{d})\}\|_2^2 \right)} \right) \right) dt + K_1 \quad (3.2.20)$$

D'après le lemme de Gronwall, pour tout $T \in (0, +\infty)$, il existe une constante $L_1(T)$, qui peut être vu comme une fonction croissante de T telle que :

$$\|\{\Phi_f(\mathbf{d})\}\|_2^2(T, \cdot) + \|\{\Phi_b(\mathbf{d})\}\|_2^2(T, \cdot) \leq L_1(T). \quad (3.2.21)$$

Pour tout $T \in (0, +\infty)$ et pour tout $t \in (0, T)$, nous avons donc :

$$\|\{\Phi_f(\mathbf{d})\}\|_2^2(t, \cdot) + \|\{\Phi_b(\mathbf{d})\}\|_2^2(t, \cdot) \leq L_1(T). \quad (3.2.22)$$

De ce fait, pour tout $T \in (0, +\infty)$, il existe deux constantes positives $L_{2a}(T)$ et $L_{2b}(T)$ telles que les inégalités (3.1.7) et (3.1.8) soient vérifiées. Nous avons donc déduit de ces deux inégalités que $\{\Phi_f(\mathbf{d})\}$ et $\{\Phi_b(\mathbf{d})\}$ appartiennent à $L^\infty\left((0, T), L^2(\mathbb{R})\right)$, $T \in (0, +\infty)$. Ceci est cohérent avec le fait que $\{\Phi_f(\mathbf{d})\}$ et $\{\Phi_b(\mathbf{d})\}$ appartiennent à \mathbf{H} .

Dans le but d'obtenir des estimations dans $L^\infty\left((0, T), H^1((0, 1))\right)$, $T \in (0, +\infty)$, nous reprenons l'inégalité (3.2.15) et nous utilisons les inégalités (3.2.16), (3.2.17) et (3.2.18) pour obtenir l'inégalité suivante :

$$\begin{aligned} & \frac{\partial \|\{\Phi_f(\mathbf{d})\}\|_2^2}{\partial t}(t, \cdot) + \frac{\partial \|\{\Phi_b(\mathbf{d})\}\|_2^2}{\partial t}(t, \cdot) + 2 \int_0^1 c_a \chi(x) \left(\frac{\partial (\{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\})}{\partial x}(t, x) \right)^2 dx \\ & \leq 2 \|\{Q(\mathbf{d})\}(t, \cdot)\|_2 \left(\sqrt{2(\|\{\Phi_f(\mathbf{d})\}(t, \cdot)\|_2^2 + \|\{\Phi_b(\mathbf{d})\}(t, \cdot)\|_2^2)} \right) \\ & \quad + r_a \left(\|\{\Phi_f(\mathbf{d})\}(t, \cdot)\|_2^2 + \|\{\Phi_b(\mathbf{d})\}(t, \cdot)\|_2^2 \right) + K_1 \quad (3.2.23) \end{aligned}$$

Pour contrôler les deux premiers termes de l'inégalité (3.2.23) nous intégrons cette inégalité par rapport au temps. Ensuite, en utilisant (3.1.7) et (3.1.8) nous obtenons pour tout T appartenant à $(0, +\infty)$:

$$\begin{aligned} & \int_0^T \frac{\partial \|\{\Phi_f(\mathbf{d})\}\|_2^2}{\partial t}(t, \cdot) dt + \int_0^T \frac{\partial \|\{\Phi_b(\mathbf{d})\}\|_2^2}{\partial t}(t, \cdot) dt \\ & \quad + 2 \int_0^T \int_0^1 c_a \chi(x) \left(\frac{\partial (\{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\})}{\partial x}(t, x) \right)^2 dx dt \leq L_{2c}(T), \quad (3.2.24) \end{aligned}$$

où $L_{2c}(T)$ est une constante positive.

Dans l'inégalité (3.2.24) les deux premiers termes ont déjà été bornés. Nous avons donc l'inégalité (3.1.9).

Nous avons donc démontré le théorème (3.1.1) et l'inégalité (3.1.9) montre que,

$$\forall T \in (0, +\infty), \quad \sqrt{\chi} \frac{\partial (\{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\})}{\partial x} \in L^2((0, T) \times (0, 1)). \quad (3.2.25)$$

Ainsi, d'après (3.2.22) et (3.2.25), $\sqrt{\chi} (\{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\})$ appartient à $H^1((0, T) \times (0, 1))$, pour tout $T \in \mathbb{R}^+$. ■

Pour conclure cette section, nous pouvons dire que les différentes inégalités obtenues ne nous permettent pas de déduire que $\{\Phi_f(\mathbf{d})\}$ et $\{\Phi_b(\mathbf{d})\}$ appartiennent toutes les deux à l'espace $L^\infty((0, T), H^1((0, 1)))$, $T \in (0, +\infty)$. En effet, la dépendance existante entre les équations (2.1.1) et (2.1.2) nécessite d'étudier la somme $(\{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\})$

3.3. Identification des propriétés perdues du fait de la structure de "l'opérateur de diffusion" grâce à l'utilisation d'un premier modèle simplifié.

et nous n'avons pu borner le terme $\chi(\{\Phi_f(d)\} + \{\Phi_b(d)\})$ qu'en moyenne au cours du temps.

De plus, la fonction χ , qui est nulle aux bords de $\Omega(d)$, masque la régularité ou l'irrégularité des fonctions $\{\Phi_f(d)\}$ et $\{\Phi_b(d)\}$ aux bords du domaine. Or des conditions aux bords sont imposées pour ces deux fonctions. La fonction χ empêche donc la démonstration de la régularité des fonctions $\{\Phi_f(d)\}$ et $\{\Phi_b(d)\}$ sur $\partial\Omega(d)$.

Les estimations a priori nous ont montré que les différentes hypothèses que nous avons faites et les conditions aux bords que nous avons imposées compliquent fortement la démonstration de l'appartenance de $\{\Phi_f(d)\}$ et $\{\Phi_b(d)\}$ à l'espace $L^\infty\left((0, T), H^1((0, 1))\right)$, $T \in (0, +\infty)$.

Nous avons fait de nombreuses explorations en se basant sur les travaux réalisés par LADYENSKAJA *et al.* (1968) et nous avons conclu que les deux premières équations de notre modèles ne rentraient pas dans le cadre classique étudié par ces auteurs. Nous avons donc mis en évidence les propriétés perdues et celles conservées malgré les particularités de ces deux équations, sur des sous-modèles simplifiés. L'objectif est aussi d'essayer de trouver de nouvelles pistes pour achever la démonstration de l'existence et de l'unicité de la solution du modèle développé initialement.

3.3 Identification des propriétés perdues du fait de la structure de "l'opérateur de diffusion" grâce à l'utilisation d'un premier modèle simplifié.

Dans cette section nous allons présenter les propriétés perdues dû à l'écriture de "l'opérateur de diffusion" que nous avons choisie. Pour cela nous avons construit un premier modèle simplifié composé des équations (3.1.11) et (3.1.12). Les conditions initiales associées à ce modèle sont données par (3.1.13) et (3.1.14).

3.3.1 Unicité de la solution

L'objectif de l'analyse de ce premier modèle simplifié est de mettre en évidence les propriétés perdues du fait de l'écriture de "l'opérateur de diffusion". Mais nous avons tout de même commencé par démontrer l'unicité de la solution de ce sous-modèle :

Propriété 3.3.1. *Si la solution des équations (3.1.11) et (3.1.12) existe, alors elle est unique.*

Démonstration de la propriété 3.3.1 : Nous supposons que $\{\Phi_{f_1}(\mathbf{d})\} \in \mathbf{H}$ et $\{\Phi_{f_2}(\mathbf{d})\} \in \mathbf{H}$ sont solutions de (3.1.11), et $\{\Phi_{b_1}(\mathbf{d})\} \in \mathbf{H}$ et $\{\Phi_{b_2}(\mathbf{d})\} \in \mathbf{H}$ sont solutions de (3.1.12). L'objectif de cette démonstration est de montrer que :

$$\{\Phi_{f_1}(\mathbf{d})\} = \{\Phi_{f_2}(\mathbf{d})\} \quad (3.3.1)$$

et,

$$\{\Phi_{b_1}(\mathbf{d})\} = \{\Phi_{b_2}(\mathbf{d})\} \quad (3.3.2)$$

Puisque $\{\Phi_{f_1}(\mathbf{d})\} \in \mathbf{H}$ et $\{\Phi_{f_2}(\mathbf{d})\} \in \mathbf{H}$ sont solutions de (3.1.11), et que $\{\Phi_{b_1}(\mathbf{d})\} \in \mathbf{H}$ et $\{\Phi_{b_2}(\mathbf{d})\} \in \mathbf{H}$ sont solutions de (3.1.12), par soustraction nous obtenons,

$$\frac{\partial \{\Phi_{f_1}(\mathbf{d})\} - \{\Phi_{f_2}(\mathbf{d})\}}{\partial t}(t, x) - c_d \frac{\partial \left[\chi \frac{\partial \left[\{\Phi_{f_1}(\mathbf{d})\} - \{\Phi_{f_2}(\mathbf{d})\} + \{\Phi_{b_1}(\mathbf{d})\} - \{\Phi_{b_2}(\mathbf{d})\} \right]}{\partial x} \right]}{\partial x}(t, x) = 0 \quad (3.3.3)$$

et,

$$\frac{\partial \{\Phi_{b_1}(\mathbf{d})\} - \{\Phi_{b_2}(\mathbf{d})\}}{\partial t}(t, x) - c_d \frac{\partial \left[\chi \frac{\partial \left[\{\Phi_{f_1}(\mathbf{d})\} - \{\Phi_{f_2}(\mathbf{d})\} + \{\Phi_{b_1}(\mathbf{d})\} - \{\Phi_{b_2}(\mathbf{d})\} \right]}{\partial x} \right]}{\partial x}(t, x) = 0 \quad (3.3.4)$$

3.3. Identification des propriétés perdues du fait de la structure de "l'opérateur de diffusion" grâce à l'utilisation d'un premier modèle simplifié.

Nous multiplions l'équation (3.3.9) par la fonction test $\{\Phi_{f_1}(\mathbf{d})\} - \{\Phi_{f_2}(\mathbf{d})\}$:

$$\begin{aligned} & \frac{\partial \{\Phi_{f_1}(\mathbf{d})\} - \{\Phi_{f_2}(\mathbf{d})\}}{\partial t}(t, x) (\{\Phi_{f_1}(\mathbf{d})\} - \{\Phi_{f_2}(\mathbf{d})\})(t, x) \\ & - c_d \frac{\partial \left[\chi \frac{\partial [\{\Phi_{f_1}(\mathbf{d})\} - \{\Phi_{f_2}(\mathbf{d})\} + \{\Phi_{b_1}(\mathbf{d})\} - \{\Phi_{b_2}(\mathbf{d})\}]}{\partial x} \right]}{\partial x}(t, x) (\{\Phi_{f_1}(\mathbf{d})\} - \{\Phi_{f_2}(\mathbf{d})\})(t, x) = 0 \end{aligned} \quad (3.3.5)$$

Puis nous multiplions l'équation (3.3.4) par la fonction test $\{\Phi_{b_1}(\mathbf{d})\} - \{\Phi_{b_2}(\mathbf{d})\}$:

$$\begin{aligned} & \frac{\partial \{\Phi_{b_1}(\mathbf{d})\} - \{\Phi_{b_2}(\mathbf{d})\}}{\partial t}(t, x) (\{\Phi_{b_1}(\mathbf{d})\} - \{\Phi_{b_2}(\mathbf{d})\})(t, x) \\ & - c_d \frac{\partial \left[\chi \frac{\partial [\{\Phi_{f_1}(\mathbf{d})\} - \{\Phi_{f_2}(\mathbf{d})\} + \{\Phi_{b_1}(\mathbf{d})\} - \{\Phi_{b_2}(\mathbf{d})\}]}{\partial x} \right]}{\partial x}(t, x) (\{\Phi_{b_1}(\mathbf{d})\} - \{\Phi_{b_2}(\mathbf{d})\})(t, x) = 0 \end{aligned} \quad (3.3.6)$$

Nous additionnons les équations (3.3.5) et (3.3.6), nous intégrons la somme obtenue par rapport à x , puis en réalisant une intégration par parties du terme de diffusion nous obtenons :

$$\begin{aligned} & \frac{1}{2} \frac{\partial \left(\|\{\Phi_{f_1}(\mathbf{d})\} - \{\Phi_{f_2}(\mathbf{d})\}\|_2^2 + \|\{\Phi_{b_1}(\mathbf{d})\} - \{\Phi_{b_2}(\mathbf{d})\}\|_2^2 \right)}{\partial t}(t, \cdot) \\ & = - \int_0^1 c_d \chi(x) \left(\frac{\partial [\{\Phi_{f_1}(\mathbf{d})\} - \{\Phi_{f_2}(\mathbf{d})\} + \{\Phi_{b_1}(\mathbf{d})\} - \{\Phi_{b_2}(\mathbf{d})\}]}{\partial x} \right)^2(t, x) dx \end{aligned} \quad (3.3.7)$$

Donc,

$$\frac{1}{2} \frac{\partial \left(\|\{\Phi_{f_1}(\mathbf{d})\} - \{\Phi_{f_2}(\mathbf{d})\}\|_2^2 + \|\{\Phi_{b_1}(\mathbf{d})\} - \{\Phi_{b_2}(\mathbf{d})\}\|_2^2 \right)}{\partial t}(t, \cdot) \leq 0 \quad (3.3.8)$$

D'après le lemme de Gronwall nous obtenons :

$$\|\{\Phi_{f_1}(\mathbf{d})\} - \{\Phi_{f_2}(\mathbf{d})\}\|_2^2(t, \cdot) + \|\{\Phi_{b_1}(\mathbf{d})\} - \{\Phi_{b_2}(\mathbf{d})\}\|_2^2(t, \cdot) \leq 0 \quad (3.3.9)$$

Nous en déduisons que,

$$\|\{\Phi_{f_1}(\mathbf{d})\} - \{\Phi_{f_2}(\mathbf{d})\}\|_2^2(t, \cdot) = 0 \quad (3.3.10)$$

et,

$$\|\{\Phi_{b_1}(\mathbf{d})\} - \{\Phi_{b_2}(\mathbf{d})\}\|_2^2(t, \cdot) = 0 \quad (3.3.11)$$

Nous obtenons donc les égalités (3.3.1) et (3.3.2).

Nous avons donc démontré la propriété 3.3.1. ■

3.3.2 L'effet régularisant de "l'opérateur de diffusion"

Dans cette partie nous allons présenter le cadre assurant la régularité des solutions en démontrant la propriété 3.1.2.

Démonstration de la propriété 3.1.2 : Nous commençons par additionner les équations (3.1.11) et (3.1.12) et nous obtenons l'équation,

$$\frac{\partial[\{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\}]}{\partial t}(t, x) - 2c_a \frac{\partial^2[\{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\}]}{\partial x^2}(t, x) = 0, \quad (3.3.12)$$

Ensuite, nous soustrayons l'équation (3.1.12) et l'équation (3.1.11) et nous obtenons l'équation,

$$\frac{\partial[\{\Phi_f(\mathbf{d})\} - \{\Phi_b(\mathbf{d})\}]}{\partial t}(t, x) = 0, \quad (3.3.13)$$

L'équation (3.3.12) correspond à l'équation de la chaleur dans laquelle l'opérateur de diffusion apporte un effet régularisant. Donc, si $(\{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\})(0, \cdot) \in L^2((0, 1))$ et que $(\{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\})(\cdot, 1) = (\{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\})(\cdot, 0) = 0$, alors $(\{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\})$ est solution de cette équation, elle est unique et appartient à \mathbf{H} .

En revanche, dans l'équation (3.3.13) il n'y a pas d'effet régularisant. Néanmoins, cette équation montre que $(\{\Phi_f(d)\}(t, x) - \{\Phi_b(d)\}(t, x))$ est constante au cours du temps. Puisque $(\{\Phi_f(d)\} - \{\Phi_b(d)\})(., 1) = (\{\Phi_f(d)\} - \{\Phi_b(d)\})(., 0) = 0$ et que $\{\Phi_f(d)\}(0, x) - \{\Phi_b(d)\}(0, x)$ est dans $H^1(0, 1)$, alors il n'y a pas d'irrégularité générée dans l'équation (3.3.13). Donc $(\{\Phi_f(d)\} - \{\Phi_b(d)\})$ appartient à $C^1((0, +\infty); H^1(0, 1))$. Nous avons donc démontré l'existence de $\{\Phi_f(d)\}$ et $\{\Phi_b(d)\}$ et que ces fonctions appartiennent à $C^1((0, +\infty); H^1(0, 1))$. Grâce à ces calculs et à la propriété 3.3.1 nous avons donc démontré la propriété 3.1.2. ■

Ces résultats montrent que la structure du modèle construit et plus particulièrement l'écriture de "l'opérateur de diffusion" qu'il contient, nécessite d'imposer des conditions initiales régulières.

3.3.3 La question de la positivité des fonctions $\{\Phi_f(d)\}$ et $\{\Phi_b(d)\}$

Nous avons étudié la question de la positivité de $\{\Phi_f(d)\}$ et $\{\Phi_b(d)\}$, solutions des équations (3.1.11) et (3.1.12). Pour cela nous avons commencé par poser :

$$U(t, x) = \{\Phi_f(d)\}(t, x) + \{\Phi_b(d)\}(t, x), \quad (3.3.14)$$

et,

$$V(t, x) = \{\Phi_f(d)\}(t, x) - \{\Phi_b(d)\}(t, x). \quad (3.3.15)$$

À partir des équations (3.3.14) et (3.3.15), l'équation (3.3.12) devient :

$$\frac{\partial U}{\partial t}(t, x) - 2c_d \frac{\partial^2 U}{\partial x^2}(t, x) = 0, \quad (3.3.16)$$

et l'équation (3.3.13) devient :

$$\frac{\partial V}{\partial t}(t, x) = 0, \quad (3.3.17)$$

Donc,

$$V(t, x) = \{\Phi_f(\mathbf{d})\}(0, x) - \{\Phi_b(\mathbf{d})\}(0, x) = V_0(x). \quad (3.3.18)$$

Nous avons cherché la solution élémentaire de l'équation (3.3.16) en résolvant l'équation de la chaleur. Pour cela nous avons utilisé la méthode de la séparation de variable (EVANS (2010)). Nous avons donc cherché la solution élémentaire de (3.3.16) sous la forme,

$$U(t, x) = y(t)z(x) \quad (3.3.19)$$

L'équation (3.3.16) devient donc :

$$y'(t)z(x) = 2c_a y(t)z''(x) \quad (3.3.20)$$

Nous divisons l'équation (3.3.20) par $y(t)z(x)$:

$$\frac{y'(t)}{y(t)} = 2c_a \frac{z''(x)}{z(x)} \quad (3.3.21)$$

Le terme de gauche ne dépend que de t et le terme de droite ne dépend que de x . Nous posons donc une constante β et nous obtenons :

$$\frac{y'(t)}{y(t)} = 2c_a \beta \quad (3.3.22)$$

et

$$\frac{z''(x)}{z(x)} = \beta \quad (3.3.23)$$

L'équation (3.3.22) peut être écrite sous la forme,

$$y'(t) - 2c_a \beta y(t) = 0 \quad (3.3.24)$$

3.3. Identification des propriétés perdues du fait de la structure de "l'opérateur de diffusion" grâce à l'utilisation d'un premier modèle simplifié.

La fonction y peut donc s'écrire,

$$y(t) = \exp(2c_a\beta t) \quad (3.3.25)$$

Physiquement, nous savons que sans source de chaleur la température diminue au cours du temps. De ce fait, β est négative. Nous posons $\beta = -\sigma^2$.

L'équation (3.3.23) peut être écrite sous la forme,

$$z''(x) - \beta z(x) = 0 \quad (3.3.26)$$

La fonction z peut donc s'écrire,

$$z(x) = \exp(mx) \quad (3.3.27)$$

L'équation (3.3.26) donne donc,

$$m^2 \exp(mx) + \sigma^2 \exp(mx) = 0 \quad (3.3.28)$$

Nous en déduisons que,

$$m^2 = -\sigma^2. \quad (3.3.29)$$

Ainsi,

$$m = \pm\sigma \quad (3.3.30)$$

La fonction z peut donc s'écrire,

$$z(x) = A \cos(\sigma x) + B \sin(\sigma x) \quad (3.3.31)$$

Cependant,

$$z(0) = A = 0 \quad (3.3.32)$$

et

$$z(1) = B \sin(\sigma) = 0. \quad (3.3.33)$$

Donc,

$$\sin(\sigma) = 0. \quad (3.3.34)$$

Cela implique que σ est de la forme,

$$\sigma = n\pi \quad (3.3.35)$$

où $n \in \mathbb{Z}$.

Nous obtenons finalement,

$$z(x) = B \sin(n\pi x) \quad (3.3.36)$$

À partir des équations (3.3.19), (3.3.25) et (3.3.36), la fonction U devient,

$$U_n(t, x) = B \exp(-2c_d(n\pi)^2 t) \sin(n\pi x) \quad (3.3.37)$$

Pour toutes les valeurs de n nous obtenons donc une solution élémentaire de la forme,

$$U_n(t, x) = C_n \exp(-2c_d(n\pi)^2 t) \sin(n\pi x). \quad (3.3.38)$$

D'après le principe de superposition, toutes les combinaisons linéaires basées sur l'équation (3.3.38), sont solutions de l'équation (3.3.14). De ce fait, la solution générale de (3.3.14) est :

$$U(t, x) = \sum_{n=1}^{\infty} \left(C_n \exp(-2c_d(n\pi)^2 t) \sin(n\pi x) \right) \quad (3.3.39)$$

3.3. Identification des propriétés perdues du fait de la structure de "l'opérateur de diffusion" grâce à l'utilisation d'un premier modèle simplifié.

Nous fixons,

$$U(0, x) = C(x) = \sum_{n=1}^{\infty} (C_n \sin(n\pi x)) \quad (3.3.40)$$

Nous multiplions l'équation (3.3.40) par $\sin(k\pi x)$ et nous intégrons par rapport à x . Nous obtenons :

$$\int_0^1 C(x) \sin(k\pi x) dx = \int_0^1 \sum_{n=1}^{\infty} C_n \sin(n\pi x) \sin(k\pi x) dx \quad (3.3.41)$$

Étant donné que les fonctions sinus sont orthogonales lorsque $n \neq k$, il ne reste que le terme $n = k$, les autres termes de la somme sont nuls. Nous avons donc :

$$\int_0^1 C(x) \sin(k\pi x) dx = \int_0^1 C_k \sin^2(k\pi x) dx = \frac{1}{2} C_k + \int_0^1 C_k \cos(2k\pi x) dx = \frac{1}{2} C_k \quad (3.3.42)$$

Ainsi, la solution générale de (3.3.14) est donnée par l'équation (3.3.39) où,

$$C_n = 2 \int_0^1 U(0, y) \sin(n\pi y) dy \quad (3.3.43)$$

S'il existe un temps T_F tel que pour tout $t > T_F$, et pour tout $x \in (0, 1)$, $\{Q(d)\}(t, x) = 0$ dès que $t > T_F$, alors lorsque t continue d'augmenter, la fonction U décroît de façon exponentielle vers 0. Ainsi, si nous fixons $V_0(x) > 0$ pour tout $x \in (0, 1)$, alors il existe une constante $T_{U_1} > 0$ telle que pour tout $t > T_{U_1}$,

$$U(t, x) \leq V(t, x). \quad (3.3.44)$$

Cela implique que,

$$\{\Phi_b(d)\}(t, x) \leq 0. \quad (3.3.45)$$

Nous remarquons aussi que si nous fixons $V_0(x) < 0$ pour tout $x \in (0, 1)$, il existe une constante $T_{U_2} > 0$ telle que pour tout $t > T_{U_2}$,

$$U(t, x) \leq -V(t, x). \quad (3.3.46)$$

Cela implique que,

$$\{\Phi_f(\mathbf{d})\}(t, x) \leq 0. \quad (3.3.47)$$

Nous avons donc déduit de l'analyse de modèle très simplifié que la positivité de $\{\Phi_f(\mathbf{d})\}$ et $\{\Phi_b(\mathbf{d})\}$ pouvait ne pas être vérifiée.

En utilisant ce premier modèle simplifié nous avons donc mis en évidence que l'intégration d'un opérateur assimilable à un opérateur de diffusion dépendant de deux fonctions inconnues induit la perte de certaines propriétés comme l'effet régularisant typiquement apporté par l'opérateur de diffusion et la positivité de la solution.

Dans la section suivante nous allons utiliser un second modèle simplifié, plus proche du modèle initial construit dans le chapitre 2. L'objectif est de mettre en évidence les propriétés conservées malgré les particularités de notre modèle.

3.4 Démonstration de l'existence et de l'unicité de la solution d'un deuxième modèle simplifié : un modèle mathématique sans terme de convection

Dans cette section nous allons présenter les analyses réalisées sur un second modèle simplifié.

Ce deuxième modèle simplifié correspond au modèle développé initialement mais sans termes de convection. De plus, nous avons décidé de travailler dans un contexte périodique. Dans ce contexte, il n'y a donc pas de bords et la fonction χ n'est plus nécessaire.

Ainsi, les deux premières équations de ce second modèle simplifié sont les équations (3.1.18) et (3.1.19). Les trois autres équations restent les équations (2.1.3), (2.1.5) et (2.1.4b) ou (2.1.4). Dans ce modèle $\{\Phi_f(\mathbf{d})\}$ et $\{\Phi_b(\mathbf{d})\}$ sont des fonctions périodiques par rapport à x , appartenant à $\mathbf{H}_{\#}$ et nous faisons les hypothèses allant de (3.1.20) jusqu'à (3.1.29).

3.4.1 Les expressions exactes de $\{\Phi_f(\mathbf{d})\}$ et $\{\Phi_b(\mathbf{d})\}$

Dans cette section nous allons présenter les différentes étapes nous ayant permis de trouver les expressions exactes de $\{\Phi_f(\mathbf{d})\}$ et $\{\Phi_b(\mathbf{d})\}$. Pour cela nous avons utilisé les Transformations de Fourier (ALINHAC et GÉRARD (2012), TAYLOR (1996), COURANT et HILBERT (2008) et EVANS (2010)). Nous avons posé :

$$\{Q(\mathbf{d})\}(t, x) = \sum_{n \in \mathbb{Z}} \{\hat{Q}^n(\mathbf{d})\}(t) \exp(2i\pi n x), \quad (3.4.1)$$

$$\{\Phi_f(\mathbf{d})\}(t, x) = \sum_{n \in \mathbb{Z}} \{\widehat{\Phi_f(\mathbf{d})}\}^n(t) \exp(2i\pi n x), \quad (3.4.2)$$

et,

$$\{\Phi_b(\mathbf{d})\}(t, x) = \sum_{n \in \mathbb{Z}} \{\widehat{\Phi_b(\mathbf{d})}\}^n(t) \exp(2i\pi n x), \quad (3.4.3)$$

Nous avons défini,

$$K_u^n = 8\pi^2 n^2 c_a + f_a, \quad (3.4.4)$$

et nous avons obtenu la propriété suivante :

Propriété 3.4.1. *Sous les hypothèses (3.1.23), (3.1.24) et (3.1.25), les équations (3.1.18) et (3.1.19) admettent respectivement une solution $\{\Phi_f(\mathbf{d})\} \in \mathbf{H}_\#$ et $\{\Phi_b(\mathbf{d})\} \in \mathbf{H}_\#$, telle que pour tout $x \in (0, 1)$ et pour tout $t \in (0, +\infty)$:*

$$\begin{aligned} \{\Phi_f(\mathbf{d})\}(t, x) = & \frac{1}{2} \sum_{n \in \mathbb{Z}} \left(\int_0^t \{\hat{Q}^n(\mathbf{d})\}(\tau) \exp(-K_u^n(t-\tau)) d\tau \right. \\ & - \int_0^t \int_0^\tau r_a \{\hat{Q}^n(\mathbf{d})\}(\nu) \exp(-K_u^n(\tau-\nu)) \exp(-(r_a + f_a)(t-\tau)) d\nu d\tau \\ & - \int_0^t r_a \left(\{\widehat{\Phi_f(\mathbf{d})}\}^n(0) + \{\widehat{\Phi_b(\mathbf{d})}\}^n(0) \right) \exp(-K_u^n \tau) \exp(-(r_a + f_a)(t-\tau)) d\tau \\ & + \left(\{\widehat{\Phi_f(\mathbf{d})}\}^n(0) + \{\widehat{\Phi_b(\mathbf{d})}\}^n(0) \right) \exp(-K_u^n t) \\ & \left. + \left(\{\widehat{\Phi_f(\mathbf{d})}\}^n(0) - \{\widehat{\Phi_b(\mathbf{d})}\}^n(0) \right) \exp(-(r_a + f_a)t) \right) \exp(2i\pi n x), \quad (3.4.5) \end{aligned}$$

et,

$$\begin{aligned}
 \{\Phi_b(\mathbf{d})\}(t, x) = & \frac{1}{2} \sum_{n \in \mathbb{Z}} \left(\int_0^t \{\hat{Q}^n(\mathbf{d})\}(\tau) \exp\left(-K_u^n(t - \tau)\right) d\tau \right. \\
 & + \int_0^t \int_0^\tau r_a \{\hat{Q}^n(\mathbf{d})\}(\nu) \exp\left(-K_u^n(\tau - \nu)\right) \exp\left(-(r_a + f_a)(t - \tau)\right) d\nu d\tau \\
 & + \int_0^t r_a \left(\{\widehat{\Phi_f}(\mathbf{d})\}^n(0) + \{\widehat{\Phi_b}(\mathbf{d})\}^n(0) \right) \exp\left(-K_u^n \tau\right) \exp\left(-(r_a + f_a)(t - \tau)\right) d\tau \\
 & \quad + \left(\{\widehat{\Phi_f}(\mathbf{d})\}^n(0) + \{\widehat{\Phi_b}(\mathbf{d})\}^n(0) \right) \exp\left(-K_u^n t\right) \\
 & \quad \left. - \left(\{\widehat{\Phi_f}(\mathbf{d})\}^n(0) - \{\widehat{\Phi_b}(\mathbf{d})\}^n(0) \right) \exp\left(-(r_a + f_a)t\right) \right) \exp(2i\pi n x). \quad (3.4.6)
 \end{aligned}$$

Démonstration de la propriété 3.4.1 : Dans un premier temps nous avons additionné les équations (3.1.18) et (3.1.19) :

$$\begin{aligned}
 \frac{\partial \{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\}}{\partial t}(t, x) - 2c_a \frac{\partial^2 [\{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\}]}{\partial x^2}(t, x) \\
 = \{Q(\mathbf{d})\}(t, x) - f_a \left(\{\Phi_f(\mathbf{d})\}(t, x) + \{\Phi_b(\mathbf{d})\}(t, x) \right). \quad (3.4.7)
 \end{aligned}$$

Puis nous avons soustrait l'équation (3.1.19) et l'équation (3.1.18) :

$$\frac{\partial \{\Phi_f(\mathbf{d})\} - \{\Phi_b(\mathbf{d})\}}{\partial t}(t, x) = -f_a \left(\{\Phi_f(\mathbf{d})\}(t, x) - \{\Phi_b(\mathbf{d})\}(t, x) \right) - 2r_a \{\Phi_f(\mathbf{d})\}(t, x). \quad (3.4.8)$$

En faisant cela nous faisons apparaître dans les équations (3.4.7) et (3.4.8) plusieurs termes impliquant la différence ou la somme des fonctions $\{\Phi_f(\mathbf{d})\}$ et $\{\Phi_b(\mathbf{d})\}$. Nous avons donc fixé :

$$U(t, x) = \{\Phi_f(\mathbf{d})\}(t, x) + \{\Phi_b(\mathbf{d})\}(t, x), \quad (3.4.9)$$

3.4. Démonstration de l'existence et de l'unicité de la solution d'un deuxième modèle simplifié :
un modèle mathématique sans terme de convection

et,

$$V(t, x) = \{\Phi_f(\mathbf{d})\}(t, x) - \{\Phi_b(\mathbf{d})\}(t, x). \quad (3.4.10)$$

Les fonctions U et V respectent les conditions de périodicité par rapport à x et, $U(0, \cdot)$ et $V(0, \cdot)$ sont périodiques et appartiennent à $L^2((0, 1))$.

L'équation (3.4.7) devient :

$$\frac{\partial U}{\partial t}(t, x) - 2c_a \frac{\partial^2 U}{\partial x^2}(t, x) = \{Q(\mathbf{d})\}(t, x) - f_a U(t, x), \quad (3.4.11)$$

et l'équation (3.4.8) devient :

$$\frac{\partial V}{\partial t}(t, x) = -f_a V(t, x) - 2r_a \{\Phi_f(\mathbf{d})\}(t, x), \quad (3.4.12)$$

ou,

$$\frac{\partial V}{\partial t}(t, x) = -(r_a + f_a)V(t, x) - r_a U(t, x). \quad (3.4.13)$$

L'expression exacte de la fonction U

Nous appliquons les Transformations de Fourier à l'équation (3.4.11) en posant :

$$U(t, x) = \sum_{n \in \mathbb{Z}} \hat{U}^n(t) \exp(2i\pi n x), \quad (3.4.14)$$

puis à partir des équations (3.4.11) et (3.4.14), nous obtenons :

$$\begin{aligned} \sum_{n \in \mathbb{Z}} \left[\hat{U}^n(t) \exp(2i\pi n x) + 2c_a \hat{U}^n(t) 4n^2 \pi^2 \exp(2i\pi n x) \right] \\ = \sum_{n \in \mathbb{Z}} \left[\{\hat{Q}^n(\mathbf{d})\}(t) \exp(2i\pi n x) - f_a \hat{U}^n \exp(2i\pi n x) \right]. \end{aligned} \quad (3.4.15)$$

En regroupant certains termes nous obtenons :

$$\sum_{n \in \mathbb{Z}} \left[\hat{U}^n(t) \exp(2i\pi n x) + (8c_a n^2 \pi^2 + f_a) \hat{U}^n(t) \exp(2i\pi n x) \right] = \sum_{n \in \mathbb{Z}} \left[\{\hat{Q}^n(\mathbf{d})\}(t) \exp(2i\pi n x) \right]. \quad (3.4.16)$$

Les termes $(\exp(2i\pi n x))_{n \in \mathbb{Z}}$ forment une base de Hilbert. Donc, pour tout $n \in \mathbb{Z}$, nous obtenons :

$$\hat{U}^n(t) \exp(2i\pi n x) + (8c_a n^2 \pi^2 + f_a) \hat{U}^n(t) \exp(2i\pi n x) = \{\hat{Q}^n(\mathbf{d})\}(t) \exp(2i\pi n x). \quad (3.4.17)$$

Donc,

$$\hat{U}^n(t) + (8c_a n^2 \pi^2 + f_a) \hat{U}^n(t) = \{\hat{Q}^n(\mathbf{d})\}(t). \quad (3.4.18)$$

Pour tout $n \in \mathbb{Z}$ cette Équation Différentielle Ordinaire (EDO) est associée à une condition initiale $\hat{U}^n(0)$ relative aux coefficients de Fourier de $\{\Phi_f(\mathbf{d})\}(0, x)$ et $\{\Phi_b(\mathbf{d})\}(0, x)$.

Nous avons résolu cette équation (3.4.18) et nous avons obtenu :

$$\hat{U}^n(t) = \hat{U}^n(0) \exp(-K_u^n t) + \int_0^t \{\hat{Q}^n(\mathbf{d})\}(\tau) \exp(-K_u^n(t - \tau)) d\tau, \quad (3.4.19)$$

où K_u^n est donnée par l'égalité (3.4.4).

L'expression exacte de la fonction V

Nous avons également appliqué les Transformations de Fourier à l'équation (3.4.13) en posant :

$$V(t, x) = \sum_{n \in \mathbb{Z}} \hat{V}^n(t) \exp(2i\pi n x) \quad (3.4.20)$$

3.4. Démonstration de l'existence et de l'unicité de la solution d'un deuxième modèle simplifié :
un modèle mathématique sans terme de convection

Donc l'équation (3.4.13) devient :

$$\sum_{n \in \mathbb{Z}} \left[\hat{V}^{n'}(t) \exp(2i\pi n x) + (r_a + f_a) \hat{V}^n(t) \exp(2i\pi n x) \right] = \sum_{n \in \mathbb{Z}} \left[-r_a \hat{U}^n(t)(x) \exp(2i\pi n x) \right] \quad (3.4.21)$$

Comme précédemment, les termes $(\exp(2i\pi n x))_{n \in \mathbb{Z}}$ forme une base de Hilbert. Donc pour tout $n \in \mathbb{Z}$,

$$\hat{V}^{n'}(t) + (r_a + f_a) \hat{V}^n(t) = -r_a \hat{U}^n(t) \quad (3.4.22)$$

Pour tout $n \in \mathbb{Z}$ cette EDO est associée à une condition initiale $\hat{V}^n(0)$ relative aux coefficients de Fourier de $\{\Phi_f(d)\}(0, x)$ et $\{\Phi_b(d)\}(0, x)$.

Nous avons résolu l'équation (3.4.22) et nous avons obtenu :

$$\hat{V}^n(t) = \hat{V}^n(0) \exp\left(- (r_a + f_a)t\right) - \int_0^t r_a \hat{U}^n(\tau) \exp\left(- (r_a + f_a)(t - \tau)\right) d\tau \quad (3.4.23)$$

À partir de l'équation (3.4.19), nous obtenons :

$$\begin{aligned} \hat{V}^n(t) &= \hat{V}^n(0) \exp\left(- (r_a + f_a)t\right) \\ &\quad - \int_0^t \int_0^\tau r_a \{\hat{Q}^n(d)\}(\nu) \exp\left(- K_u^n(\tau - \nu)\right) \exp\left(- (r_a + f_a)(t - \tau)\right) d\nu d\tau \\ &\quad - \int_0^t r_a \hat{U}^n(0) \exp\left(- K_u^n \tau\right) \exp\left(- (r_a + f_a)(t - \tau)\right) d\tau. \end{aligned} \quad (3.4.24)$$

Les expressions exactes des fonctions $\{\Phi_f(d)\}$ et $\{\Phi_b(d)\}$

En additionnant les équations (3.4.9) et (3.4.10), nous obtenons :

$$\hat{U}^n(t) + \hat{V}^n(t) = 2\{\widehat{\Phi_f(d)}\}^n(t), \quad (3.4.25)$$

et en soustrayant l'équation (3.4.10) et l'équation (3.4.9), nous obtenons :

$$\hat{U}^n(t) - \hat{V}^n(t) = 2\{\widehat{\Phi_b(\mathbf{d})}\}^n(t), \quad (3.4.26)$$

où $\{\widehat{\Phi_f(\mathbf{d})}\}^n$ et $\{\widehat{\Phi_b(\mathbf{d})}\}^n$ sont données par (3.4.2) et (3.4.3).

Donc, à partir des équations (3.4.25), (3.4.26), (3.4.19) et (3.4.24), nous obtenons :

$$\begin{aligned} \{\widehat{\Phi_f(\mathbf{d})}\}^n(t) = & \frac{1}{2} \left(\int_0^t \{\hat{Q}^n(\mathbf{d})\}(\tau) \exp(-K_u^n(t-\tau)) d\tau \right. \\ & - \int_0^t \int_0^\tau r_a \{\hat{Q}^n(\mathbf{d})\}(\nu) \exp(-K_u^n(\tau-\nu)) \exp(-(r_a + f_a)(t-\tau)) d\nu d\tau \\ & - \int_0^t r_a \hat{U}^n(0) \exp(-K_u^n\tau) \exp(-(r_a + f_a)(t-\tau)) d\tau \\ & \left. + \hat{U}^n(0) \exp(-K_u^n t) + \hat{V}^n(0) \exp(-(r_a + f_a)t) \right), \quad (3.4.27) \end{aligned}$$

et,

$$\begin{aligned} \{\widehat{\Phi_b(\mathbf{d})}\}^n(t) = & \frac{1}{2} \left(\int_0^t \{\hat{Q}^n(\mathbf{d})\}(\tau) \exp(-K_u^n(t-\tau)) d\tau \right. \\ & + \int_0^t \int_0^\tau r_a \{\hat{Q}^n(\mathbf{d})\}(\nu) \exp(-K_u^n(\tau-\nu)) \exp(-(r_a + f_a)(t-\tau)) d\nu d\tau \\ & + \int_0^t r_a \hat{U}^n(0) \exp(-K_u^n\tau) \exp(-(r_a + f_a)(t-\tau)) d\tau \\ & \left. + \hat{U}^n(0) \exp(-K_u^n t) - \hat{V}^n(0) \exp(-(r_a + f_a)t) \right). \quad (3.4.28) \end{aligned}$$

En utilisant les équations (3.4.2) et (3.4.27), puis les équations (3.4.3) et (3.4.28), nous obtenons finalement les expressions exactes des fonctions $\{\Phi_f(\mathbf{d})\}$ et $\{\Phi_b(\mathbf{d})\}$.

Appartenance de $\{\Phi_f(d)\}$ et $\{\Phi_b(d)\}$ à l'espace $\mathbf{H}_\#$

Les expressions de $\{\Phi_f(d)\}$ et $\{\Phi_b(d)\}$ sont données par les équations (3.4.5) et (3.4.6). Nous devons désormais prouver que ces fonctions appartiennent à $\mathbf{H}_\#$. Pour cela nous avons démontré que les fonctions U et V , définies par les équations (3.4.9) et (3.4.10), appartiennent à $\mathbf{H}_\#$ et nous en avons déduit que $\{\Phi_f(d)\}$ et $\{\Phi_b(d)\}$ appartiennent aussi à $\mathbf{H}_\#$.

Dans un contexte périodique, l'espace $\mathbf{H}_\#$ peut être défini de la manière suivante,

$$\mathbf{H}_\# = H^1((0, +\infty) \times (0, 1)) \cap C^0((0, +\infty), H_\#^1(0, 1)) \cap C^1((0, +\infty), H_\#^{-1}(0, 1)). \quad (3.4.29)$$

De ce fait, par définition, une fonction Φ , définie telle que,

$$\Phi(t, x) = \sum_{n \in \mathbb{Z}} \hat{\Phi}^n(t) \exp(2i\pi nx), \quad (3.4.30)$$

appartient à $\mathbf{H}_\#$ si pour tout $t \in (0, T)$, $T \in (0, +\infty)$:

$$\sum_{n \in \mathbb{Z}} \left(n \hat{\Phi}^n(t) \right)^2 \leq L_I, \quad (3.4.31)$$

$$\sum_{n \in \mathbb{Z}} \frac{1}{|1 + n^2|} \hat{\Phi}^n(t)^2 \leq L_{II}, \quad (3.4.32)$$

et,

$$\sum_{n \in \mathbb{Z}} \frac{1}{|1 + n^2|} \hat{\Phi}^n(t)^2 \leq L_{III}, \quad (3.4.33)$$

où L_I , L_{II} et L_{III} sont trois constantes positives.

Dans cette section nous allons donc montrer que U et V vérifient les équations (3.4.31), (3.4.32) et (3.4.33).

À partir des égalités (3.4.14) et (3.4.19) et des identités remarquables nous obtenons :

$$\left(n\hat{U}^n(t)\right)^2 \leq C_1 \left(n^2\hat{U}^n(0)^2 \exp(-2K_u^n t) + n^2 \int_0^t \{\hat{Q}^n(\mathbf{d})\}(\tau)^2 \exp(-2K_u^n(t-\tau)) d\tau\right), \quad (3.4.34)$$

où K_u^n est donnée par (3.4.4) et C_1 est une constante positive.

Nous remarquons que :

$$\begin{aligned} n^2 \int_0^t \{\hat{Q}^n(\mathbf{d})\}(\tau)^2 \exp(-2K_u^n(t-\tau)) d\tau \\ = \int_0^t n^2 \{\hat{Q}^n(\mathbf{d})\}(\tau)^2 \exp(-16\pi^2 c_d n^2(t-\tau)) \exp(-2f_d(t-\tau)) d\tau \\ \leq C_2 \int_0^t n^2 \{\hat{Q}^n(\mathbf{d})\}(\tau)^2 \exp(-C_3 n^2(t-\tau)) d\tau, \end{aligned} \quad (3.4.35)$$

où C_2 et C_3 sont deux constantes positives. D'après l'hypothèse (3.1.20), $(\{\hat{Q}^n(\mathbf{d})\})_{n \in \mathbb{Z}}$ appartient à $L^\infty((0, T), l^2)$, pour tout $t \in (0, T)$, $T \in (0, +\infty)$. Donc, il existe \overline{Q}^n telle que pour tout $t \in (0, T)$, $T \in (0, +\infty)$:

$$\{\hat{Q}^n(\mathbf{d})\}(t)^2 \leq \{\overline{Q}^n(\mathbf{d})\}^2 \quad (3.4.36)$$

De plus,

$$\begin{aligned} \int_0^t n^2 \exp(-C_3 n^2(t-\tau)) d\tau &= \left[\frac{1}{C_3} \exp(-C_3 n^2(t-\tau)) \right]_0^t \\ &= \frac{1}{C_3} \left(1 - \exp(-C_3 n^2 t)\right) \end{aligned} \quad (3.4.37)$$

À partir des équations (3.4.35), (3.4.36) et (3.4.37), nous obtenons :

$$n^2 \int_0^t \{\hat{Q}^n(\mathbf{d})\}(\tau)^2 \exp(-2K_u^n(t-\tau)) d\tau \leq C_4 \{\overline{Q}^n(\mathbf{d})\}^2 \left(1 - \exp(-C_3 n^2 t)\right), \quad (3.4.38)$$

où C_4 est une constante positive. Puisque $0 \leq \exp(-C_3 n^2 t) \leq 1$, alors il existe $C_5 > 0$

3.4. Démonstration de l'existence et de l'unicité de la solution d'un deuxième modèle simplifié :
un modèle mathématique sans terme de convection

telle que pour tout $t \in (0, T)$, $T \in (0, +\infty)$,

$$n^2 \int_0^t \{\hat{Q}^n(\mathbf{d})\}(\tau)^2 \exp(-2K_u^n(t - \tau)) d\tau \leq C_5 \{\overline{Q}^n(\mathbf{d})\}^2. \quad (3.4.39)$$

À partir des équations (3.4.34) et (3.4.39) nous obtenons :

$$\sum_{n \in \mathbb{Z}} \left(n \hat{U}^n(t) \right)^2 \leq C_6 \left(\sum_{n \in \mathbb{Z}} n^2 \hat{U}^n(0)^2 \exp(-2K_u^n t) + \sum_{n \in \mathbb{Z}} \{\overline{Q}^n(\mathbf{d})\}^2 \right), \quad (3.4.40)$$

où C_6 est une constante positive.

D'après l'hypothèse (3.1.20), $(\{\hat{Q}^n(\mathbf{d})\}(t))_{n \in \mathbb{Z}}$ appartient à l^2 , pour tout $t \in (0, T)$, $T \in (0, +\infty)$. Donc, il existe $K_1 > 0$ telle que pour tout $t \in (0, T)$, $T \in (0, +\infty)$:

$$\sum_{n \in \mathbb{Z}} \{\hat{Q}^n(\mathbf{d})\}(t)^2 \leq K_1 \text{ et } \sum_{n \in \mathbb{Z}} \{\overline{Q}^n(\mathbf{d})\}^2 \leq K_1. \quad (3.4.41)$$

D'après l'hypothèse (3.1.23), $U(0, \cdot)$ appartient à $L^\infty((0, 1))$. cela signifie qu'il existe $K_2 > 0$ telle que :

$$\sum_{n \in \mathbb{Z}} \hat{U}^n(0)^2 \leq K_2 \quad (3.4.42)$$

Puisque la somme $\sum_{n \in \mathbb{Z}} n^2 \exp(-2K_u^n t)$ converge, il existe $K_3 > 0$ telle que :

$$\sum_{n \in \mathbb{Z}} n^2 \hat{U}^n(0)^2 \exp(-2K_u^n t) \leq K_3 \quad (3.4.43)$$

Finalement, à partir des équations (3.4.40), (3.4.41) et (3.4.43) nous avons prouvé que la fonction U vérifie (3.4.31).

La fonction U vérifie l'inégalité (3.4.31). Donc pour tout $t \in (0, T)$, $T \in (0, +\infty)$:

$$\sum_{n \in \mathbb{Z}} \hat{U}^n(t)^2 \leq K_2 \quad (3.4.44)$$

De plus, $0 \leq \frac{1}{|1+n^2|} \leq 1$, pour tout $n \in \mathbb{Z}$. Ainsi, la fonction U vérifie (3.4.32).

D'après les équations (3.4.18) et (3.4.41) nous avons,

$$\begin{aligned} \sum_{n \in \mathbb{Z}} \frac{1}{|1+n^2|} (\hat{U}^{n'}(t))^2 &\leq K_{1b} - \sum_{n \in \mathbb{Z}} C_7 \frac{n^4}{|1+n^2|} (\hat{U}^n(t))^2 \\ &\leq K_{1b} - \sum_{n \in \mathbb{Z}} C_7 \frac{n^2}{|1+n^2|} (n\hat{U}^n(t))^2 \end{aligned} \quad (3.4.45)$$

où K_{1b} et C_7 sont des constantes positives.

Nous avons démontré précédemment que la fonction \hat{U}^n vérifie l'équation (3.4.31) donc :

$$\sum_{n \in \mathbb{Z}} \frac{1}{|1+n^2|} (\hat{U}^{n'}(t))^2 \leq K_{1b} - K_{2b} \leq K_4 \quad (3.4.46)$$

où K_{2b} et K_4 sont des constantes positives.

Ainsi la fonction \hat{U}^n vérifie les inégalités (3.4.31), (3.4.32) et (3.4.33). Nous avons finalement prouvé que la fonction U appartient à $\mathbf{H}_{\#}$.

À partir des équations (3.4.20) et (3.4.24) et des identités remarquables nous obtenons :

$$\begin{aligned} \left(n\hat{V}^n(t) \right)^2 &\leq C_8 \left(n^2\hat{V}^n(0)^2 \exp(-2(r_a + f_a)t) \right. \\ &\quad \left. + \int_0^t r_a^2 n^2 \hat{U}^n(\tau)^2 \exp(-2(r_a + f_a)(t - \tau)) d\tau \right), \end{aligned} \quad (3.4.47)$$

où C_8 est une constante positive.

Pour tout $t \in (0, T)$, $T \in (0, +\infty)$, $0 \leq \exp(-2(r_a + f_a)t) \leq 1$ et $0 \leq \exp(-2(r_a + f_a)(t - \tau)) \leq 1$. À partir de l'inégalité (3.4.47), nous obtenons pour tout $t \in (0, T)$, $T \in (0, +\infty)$:

$$\left(n\hat{V}^n(t) \right)^2 \leq C_8 \left(n^2\hat{V}^n(0)^2 + \int_0^t r_a^2 n^2 \hat{U}^n(\tau)^2 d\tau \right). \quad (3.4.48)$$

3.4. Démonstration de l'existence et de l'unicité de la solution d'un deuxième modèle simplifié :
un modèle mathématique sans terme de convection

À partir de l'hypothèse (3.1.24) nous savons qu'il existe $K_5 > 0$ telle que :

$$\sum_{n \in \mathbb{Z}} n^2 \hat{V}^n(0)^2 \leq K_5 \quad (3.4.49)$$

Nous avons démontré précédemment que la fonction U vérifie l'inégalité (3.4.31), donc il existe $K_6 > 0$ telle que :

$$\sum_{n \in \mathbb{Z}} n^2 \hat{U}^n(t)^2 \leq K_6 \quad (3.4.50)$$

Finalement, à partir des inégalités (3.4.48), (3.4.49) et (3.4.50) nous avons démontré que pour tout $t \in (0, T)$, $T \in (0, +\infty)$ la fonction V vérifie (3.4.31).

La fonction V vérifie (3.4.31). Donc pour tout $t \in (0, T)$, $T \in (0, +\infty)$ il existe $K_7 > 0$ telle que :

$$\sum_{n \in \mathbb{Z}} \hat{V}^n(t)^2 \leq K_7 \quad (3.4.51)$$

De plus, $0 \leq \frac{1}{|1+n^2|} \leq 1$, pour tout $n \in \mathbb{Z}$. Donc, la fonction V vérifie (3.4.32).

D'après les équations (3.4.22), (3.4.51) et (3.4.50) nous avons :

$$\begin{aligned} \sum_{n \in \mathbb{Z}} \frac{1}{|1+n^2|} (\hat{V}^{n'}(t))^2 &\leq C_9 \left(\sum_{n \in \mathbb{Z}} \frac{1}{|1+n^2|} (\hat{V}^n(t))^2 + \sum_{n \in \mathbb{Z}} \frac{1}{|1+n^2|} (\hat{U}^n(t))^2 \right) \\ &\leq C_9 (K_{6b} + K_{5b}) \leq K_8 \end{aligned} \quad (3.4.52)$$

où C_9 , K_{6b} , K_{5b} et K_8 sont des constantes positives.

Ainsi la fonction \hat{V}^n vérifie les inégalités (3.4.31), (3.4.32) et (3.4.33). Nous avons finalement démontré que V appartient à $\mathbf{H}_\#$.

Dans cette section nous avons démontré que $\{\Phi_f(d)\} + \{\Phi_b(d)\}$ et $\{\Phi_f(d)\} - \{\Phi_b(d)\}$ appartiennent à $\mathbf{H}_\#$. Ainsi, $\{\Phi_f(d)\}$ et $\{\Phi_b(d)\}$ appartiennent à $\mathbf{H}_\#$.

Finalement nous avons démontré la propriété 3.4.1. ■

3.4.2 Positivité de certaines composantes de la solution du système d'EDP

Positivité de $(\{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\})$

Dans cette section nous allons démontrer que :

Propriété 3.4.2. *Sous les hypothèses allant de (3.1.20) jusqu'à (3.1.25) il existe une solution $(\{\Phi_f(\mathbf{d})\}, \{\Phi_b(\mathbf{d})\})$ de (3.1.18) et (3.1.19). Elle satisfait pour $\{\Phi_f(\mathbf{d})\} \in \mathbf{H}_\#$ et $\{\Phi_b(\mathbf{d})\} \in \mathbf{H}_\#$, pour tout $x \in (0, 1)$ et pour tout $t \in (0, +\infty)$,*

$$(\{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\})(t, x) \geq 0 \quad (3.4.53)$$

Démonstration de la propriété 3.4.2 :

Dans un premier temps, l'existence de la solution $(\{\Phi_f(\mathbf{d})\}, \{\Phi_b(\mathbf{d})\})$ de (3.1.18) et (3.1.19) est assurée par la propriété 3.4.1.

En ce qui concerne la démonstration de l'inégalité (3.4.53), nous avons commencé par poser,

$$m_\Phi(t, x) = \min \left(0, (\{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\})(t, x) \right), \quad (3.4.54)$$

nous avons ensuite multiplié l'équation (3.4.7) par m_Φ . nous avons obtenu :

$$\begin{aligned} & \frac{\partial [\{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\}]}{\partial t} (t, x) m_\Phi(t, x) - 2c_a \frac{\partial^2 [\{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\}]}{\partial x^2} (t, x) m_\Phi(t, x) \\ & = \{Q(\mathbf{d})\}(t, x) m_\Phi(t, x) - f_a \left(\{\Phi_f(\mathbf{d})\}(t, x) + \{\Phi_b(\mathbf{d})\}(t, x) \right) m_\Phi(t, x). \end{aligned} \quad (3.4.55)$$

En intégrant l'équation (3.4.55) par rapport à x , nous obtenons :

$$\begin{aligned} & \int_0^1 \frac{\partial [\{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\}]}{\partial t} (t, x) m_\Phi(t, x) dx + \int_0^1 2c_a \frac{\partial [\{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\}]}{\partial x} (t, x) \frac{\partial m_\Phi}{\partial x} (t, x) dx \\ & = \int_0^1 \left(\{Q(\mathbf{d})\}(t, x) m_\Phi(t, x) - f_a \left(\{\Phi_f(\mathbf{d})\}(t, x) + \{\Phi_b(\mathbf{d})\}(t, x) \right) m_\Phi(t, x) \right) dx. \end{aligned} \quad (3.4.56)$$

Nous avons,

3.4. Démonstration de l'existence et de l'unicité de la solution d'un deuxième modèle simplifié :
un modèle mathématique sans terme de convection

$$\frac{\partial[\{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\}]}{\partial t}(t, x)m_\Phi(t, x) = \frac{\partial m_\Phi}{\partial t}(t, x)m_\Phi(t, x) = \frac{1}{2} \frac{\partial m_\Phi^2}{\partial t}(t, x), \quad (3.4.57)$$

et,

$$\frac{\partial[\{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\}]}{\partial x}(t, x) \frac{\partial m_\Phi}{\partial x}(t, x) = \frac{\partial m_\Phi}{\partial x}(t, x) \frac{\partial m_\Phi}{\partial x}(t, x) = \left(\frac{\partial m_\Phi}{\partial x}(t, x) \right)^2, \quad (3.4.58)$$

Donc, à partir des équations (3.4.56), (3.4.57) et (3.4.58), nous obtenons :

$$\begin{aligned} \int_0^1 \frac{1}{2} \frac{\partial(m_\Phi^2)}{\partial t}(t, x) dx + \int_0^1 2c_a \left(\frac{\partial m_\Phi}{\partial x}(t, x) \right)^2 dx \\ = \int_0^1 \left(\{Q(\mathbf{d})\}(t, x)m_\Phi(t, x) - f_a(m_\Phi(t, x))^2 \right) dx. \end{aligned} \quad (3.4.59)$$

D'après l'hypothèse (3.1.20), $\{Q(\mathbf{d})\}$ est positive et par définition la m_Φ est négative pour tout $t \in (0, +\infty)$ et pour tout $x \in (0, 1)$. À partir de l'équation (3.4.59) nous obtenons donc :

$$\frac{1}{2} \frac{\partial \|m_\Phi\|_2^2}{\partial t}(t, \cdot) \leq \int_0^1 \{Q(\mathbf{d})\}(t, x)m_\Phi(t, x) dx \leq 0 \quad (3.4.60)$$

La fonction $\|m_\Phi\|_2^2$ est donc décroissante et d'après l'hypothèse (3.1.21), $m_\Phi^2(0, \cdot) = 0$. Donc, pour tout $t \in (0, +\infty)$,

$$\|m_\Phi(t, \cdot)\|_2^2 \leq 0 \quad (3.4.61)$$

Ainsi,

$$\|m_\Phi(t, \cdot)\|_2 = 0 \quad (3.4.62)$$

La propriété 3.4.1 et l'égalité (3.4.62) prouvent donc la propriété 3.4.2. ■

La positivité de $\{\Psi(d)\}$

Propriété 3.4.3. *Sous les hypothèses (3.1.20), (3.1.21) et (3.1.26) et si $\{\Phi_f(d)\}$ et $\{\Phi_b(d)\}$ appartiennent à $\mathbf{H}_\#$ sont solutions de (3.1.18) et (3.1.19), alors l'équation (2.1.3) admet une unique solution appartenant à $C^1((0, +\infty), L^1((0, 1)))$. De plus, cette solution est positive.*

Démonstration de la propriété 3.4.3 :

Les fonctions $\{\Phi_f(d)\} \in \mathbf{H}_\#$ et $\{\Phi_b(d)\} \in \mathbf{H}_\#$ sont solutions des équations (3.1.18) et (3.1.19) et $\{\Psi(d)\} \in C^1((0, +\infty), L^1((0, 1)))$ est solution de l'équation (2.1.3) qui est une EDO. Nous avons donc,

$$\begin{aligned} \{\Psi(d)\}(t, x) &= \{\Psi(d)\}(0, x) \exp(-u_a t) \\ &+ \int_0^t \exp(-u_a(t-s)) f_a(\{\Phi_b(d)\}(s, x) + \{\Phi_f(d)\}(s, x)) ds. \end{aligned} \quad (3.4.63)$$

D'après la propriété 3.4.2 ($\{\Phi_b(d)\}(t, x) + \{\Phi_f(d)\}(t, x)$) est positive pour tout x dans $(0, 1)$ et pour tout t dans $(0, T)$, T appartenant à $(0, +\infty)$. Donc, $\{\Psi(d)\}(t, x)$ est positive pour tout x dans $(0, 1)$ et pour tout t dans $(0, T)$, où T appartient à $(0, +\infty)$. ■

Positivité et majoration de la fonction $\{O(d)\}$

L'objectif de cette section est de prouver que :

Propriété 3.4.4. *Sous les hypothèses (3.1.20), (3.1.21), (3.1.26) et (3.1.28), si la fonction $\{\Xi(d)\} \in C^1((0, +\infty), L^1((0, 1)))$ est solution de l'équation (2.1.4b) et $\{O(d)\} \in C^1((0, +\infty))$ est solution de l'équation (2.1.5), alors pour tout $t \in (0, T)$, $T \in (0, +\infty)$,*

$$0 \leq \{O(d)\}(t) \leq L_a, \quad (3.4.64)$$

et

Propriété 3.4.5. *Sous les hypothèses (3.1.20), (3.1.21), (3.1.26) et (3.1.28), si la fonction $\{\Xi(d)\} \in C^1((0, +\infty), L^1((0, 1)))$ est positive et solution de l'équation (2.1.4) et si $\{O(d)\} \in C^1((0, +\infty))$ es solution de l'équation (2.1.5), alors pour tout $t \in (0, T)$, $T \in (0, +\infty)$, alors la fonction $\{O(d)\}$ est positive.*

Démonstration de la propriété 3.4.4 : Pour prouver la positivité de la fonction $\{O(d)\}$ dans le cas où la fonction $\{\Xi(d)\} \in C^1((0, +\infty), L^1((0, 1)))$ est solution de l'équation (2.1.4b) et $\{O(d)\} \in C^1((0, +\infty))$ est solution de l'équation (2.1.5) nous avons posé,

$$m_O(t) = \min \left(0, \{O(d)\}(t) \right). \quad (3.4.65)$$

À partir des équations (2.1.4b) et (2.1.5) nous avons obtenu,

$$\frac{\partial \{O(d)\}}{\partial t}(t) = \int_0^1 \frac{\partial \{\Xi(d)\}}{\partial t}(t, x) dx = \int_0^1 u_a \{\Psi(d)\}(t, x) \left(\frac{L_a - \{O(d)\}(t)}{L_a} \right) dx \quad (3.4.66)$$

Dans un premier temps nous multiplions l'équation (3.4.66) par m_O :

$$\begin{aligned} \frac{\partial \{O(d)\}}{\partial t}(t) m_O(t) &= \int_0^1 u_a \{\Psi(d)\}(t, x) \left(\frac{L_a - \{O(d)\}(t)}{L_a} \right) m_O(t) dx \\ &= u_a \left(\frac{L_a - \{O(d)\}(t)}{L_a} \right) m_O(t) \int_0^1 \{\Psi(d)\}(t, x) dx \\ &= u_a m_O(t) \int_0^1 \{\Psi(d)\}(t, x) dx - u_a \left(\frac{\{O(d)\}(t)}{L_a} \right) m_O(t) \int_0^1 \{\Psi(d)\}(t, x) dx \end{aligned} \quad (3.4.67)$$

Par définition, la fonction m_O est négative et d'après la propriété 3.4.3 $\{\Psi(d)\}(t, x)$ est positive pour tout $x \in (0, 1)$ et pour tout $t \in (0, T)$, $T \in (0, +\infty)$. Donc,

$$\frac{\partial \{O(d)\}}{\partial t}(t) m_O(t) \leq -u_a \left(\frac{\{O(d)\}(t)}{L_a} \right) m_O(t) \int_0^1 \{\Psi(d)\}(t, x) dx \quad (3.4.68)$$

Nous avons,

$$\frac{\partial \{O(d)\}}{\partial t}(t) m_O(t) = \frac{\partial m_O}{\partial t}(t) m_O(t) = \frac{1}{2} \frac{\partial (m_O)^2}{\partial t}(t) \quad (3.4.69)$$

et,

$$\{O(d)\}(t) m_O(t) = m_O^2(t) \quad (3.4.70)$$

Ainsi, en utilisant (3.4.69), (3.4.70) et l'équation (3.4.68) nous obtenons :

$$\frac{1}{2} \frac{\partial(m_O)^2}{\partial t}(t) \leq -\frac{u_a}{L_a} m_O^2(t) \int_0^1 \{\Psi(d)\}(t, x) dx \leq 0 \quad (3.4.71)$$

Donc la fonction m_O^2 est décroissante et d'après l'hypothèse (3.1.28), $m_O^2(0) = 0$. Pour tout $t \in (0, T)$, $T \in (0, +\infty)$, nous avons donc :

$$m_O^2(t) \leq 0 \quad (3.4.72)$$

De ce fait, la fonction m_O^2 est nulle pour tout $t \in (0, T)$, $T \in (0, +\infty)$. Cela signifie que la fonction $\{O(d)\}$ est positive.

Pour démontrer la majoration de la fonction $\{O(d)\}$ par la constante positive L_a nous posons,

$$M_O(t) = \max \left(0, (\{O(d)\}(t) - L_a) \right), \quad (3.4.73)$$

Nous multiplions l'équation (3.4.66) par M_O :

$$\frac{\partial\{O(d)\}}{\partial t}(t) M_O(t) = \int_0^1 u_a \{\Psi(d)\}(t, x) \left(\frac{L_a - \{O(d)\}(t)}{L_a} \right) M_O(t) dx \quad (3.4.74)$$

Nous avons,

$$\left(L_a - \{O(d)\}(t) \right) M_O(t) = -M_O^2(t), \quad (3.4.75)$$

et,

$$\frac{\partial\{O(d)\}}{\partial t}(t) M_O(t) = \frac{\partial M_O}{\partial t}(t) M_O(t) = \frac{1}{2} \frac{\partial M_O^2}{\partial t}(t) \quad (3.4.76)$$

Ainsi, d'après la propriété 3.4.3 à partir des équations (3.4.74), (3.4.75) et (3.4.76) nous obtenons :

3.4. Démonstration de l'existence et de l'unicité de la solution d'un deuxième modèle simplifié :
un modèle mathématique sans terme de convection

$$\frac{1}{2} \frac{\partial M_O^2}{\partial t}(t) = -\frac{u_d}{L_d} M_O^2(t) \int_0^1 \{\Psi(d)\}(t, x) dx \leq 0. \quad (3.4.77)$$

Donc, la fonction M_O^2 est décroissante et d'après l'hypothèse (3.1.28), $M_O^2(0) = 0$.
Donc,

$$\|M_O\|_2^2 \leq 0 \quad (3.4.78)$$

La fonction M_O^2 est nulle pour tout $t \in (0, T)$, $T \in (0, +\infty)$. Cela signifie que la fonction $\{O(d)\}$ est inférieure ou égale à L_d quel que soit $t \in (0, T)$, $T \in (0, +\infty)$.

Nous avons finalement démontré la propriété 3.4.4. ■

Démonstration de la propriété 3.4.5 :

Dans le cas où $\{\Xi(d)\} \in C^1((0, +\infty), L^1((0, 1)))$ est solution de l'équation (2.1.4) et $\{O(d)\} \in C^1((0, +\infty))$ est solution de l'équation (2.1.5), la fonction $\{O(d)\}$ dépend uniquement de $\{\Xi(d)\}$. Ainsi, si la fonction $\{\Xi(d)\}$ est positive, alors $\{O(d)\}$ est également positive. ■

La positivité de $\{\Xi(d)\}$

Dans le deuxième modèle simplifié, la fonction $\{\Xi(d)\}$ peut être solution de l'équation (2.1.4b) ou (2.1.4). Nous avons donc démontré que :

Propriété 3.4.6. *Sous les hypothèses (3.1.27) et (3.1.28) la fonction $\{\Xi(d)\} \in C^1((0, +\infty), L^1((0, 1)))$ est solution de l'équation (2.1.4b) ou (2.1.4) et si $\{O(d)\} \in C^1((0, +\infty))$ est solution de l'équation (2.1.5), alors pour tout $t \in (0, T)$, T appartenant à $(0, +\infty)$, la fonction $\{\Xi(d)\}$ est positive.*

Démonstration de la propriété 3.4.6 :

Dans le cas où la fonction $\{\Xi(d)\} \in C^1((0, +\infty), L^1((0, 1)))$ est solution de l'équation (2.1.4b) et $\{O(d)\} \in C^1((0, +\infty))$ est solution de l'équation (2.1.5), l'équation (2.1.4b) est une EDO non-linéaire. Pour tout $t \in (0, T)$ et pour tout $x \in (0, 1)$ nous avons,

$$\{\Xi(\mathbf{d})\}(t, x) = \{\Xi(\mathbf{d})\}(0, x) + \int_0^t u_a \{\Psi(\mathbf{d})\}(\tau, x) \left(\frac{L_a - \{O(\mathbf{d})\}(\tau)}{L_a} \right) d\tau. \quad (3.4.79)$$

D'après l'hypothèse (3.1.27) et les propriétés 3.4.3 et 3.4.4, les fonctions $\{\Psi(\mathbf{d})\}(t, x)$ et $\{O(\mathbf{d})\}(t)$ sont positives et $\{O(\mathbf{d})\}(t)$ est inférieure ou égale à L_a , pour tout $x \in (0, 1)$ et pour tout $t \in (0, T)$, $T \in (0, +\infty)$. Dans ce cas-là $\{\Xi(\mathbf{d})\}(t, x)$ est positive pour tout $x \in (0, 1)$ et pour tout $t \in (0, T)$, $T \in (0, +\infty)$.

Dans le cas où la fonction $\{\Xi(\mathbf{d})\} \in C^1((0, +\infty), L^1((0, 1)))$ est solution de l'équation (2.1.4) et la fonction $\{O(\mathbf{d})\} \in C^1((0, +\infty))$ est solution de l'équation (2.1.5), alors l'équation (2.1.4b) est une EDO linéaire. Donc, pour tout $t \in (0, T)$, $T \in (0, +\infty)$ et pour tout $x \in (0, 1)$ nous avons,

$$\{\Xi(\mathbf{d})\}(t, x) = \{\Xi(\mathbf{d})\}(0, x) + \int_0^t u_a \{\Psi(\mathbf{d})\}(\tau, x) d\tau. \quad (3.4.80)$$

D'après l'hypothèse (3.1.27) et la propriété 3.4.3, $\{\Psi(\mathbf{d})\}(t, x)$ est positive pour tout $x \in (0, 1)$ et pour tout $t \in (0, T)$, $T \in (0, +\infty)$. Dans ce cas-là $\{\Xi(\mathbf{d})\}(t, x)$ est également positive pour tout $x \in (0, 1)$ et pour tout $t \in (0, T)$, $T \in (0, +\infty)$.

Nous avons finalement démontré la propriété 3.4.6. ■

3.4.3 Propriété de conservation du second sous-modèle

Dans cette section nous allons démontrer que,

Propriété 3.4.7. *Sous les hypothèses allant de (3.1.20) jusqu'à (3.1.29), $\{\Phi_f(\mathbf{d})\} \in \mathbf{H}_\#$ et $\{\Phi_b(\mathbf{d})\} \in \mathbf{H}_\#$ solutions des équations (3.1.18) et (3.1.19), $\{\Psi(\mathbf{d})\} \in C^1((0, +\infty); L^1((0, 1)))$ solution de l'équation (2.1.3), $\{\Xi(\mathbf{d})\} \in C^1((0, +\infty); L^1((0, 1)))$ solution de l'équation (2.1.4b) ou (2.1.4), et $\{O(\mathbf{d})\} \in C^1((0, +\infty))$ solution de (2.1.5), alors pour tout $T > 0$, il existe $K_{Q_2}(T) > 0$ telle que pour tout $x \in (0, 1)$ et pour tout $t \in (0, T)$, l'inégalité (3.1.30) est satisfaite.*

Démonstration de la propriété 3.4.7 : Les fonctions $\{\Phi_f(\mathbf{d})\}$, $\{\Phi_b(\mathbf{d})\}$, $\{\Psi(\mathbf{d})\}$, $\{\Xi(\mathbf{d})\}$

3.4. Démonstration de l'existence et de l'unicité de la solution d'un deuxième modèle simplifié :
un modèle mathématique sans terme de convection

et $\{O(d)\}$ sont respectivement solutions de (3.1.18), (3.1.19), (2.1.3), (2.1.4) et (2.1.5). D'après l'hypothèse (3.1.20), la fonction $\{Q(d)\}$ sont aussi positives. En additionnant les équations (3.1.18), (3.1.19), (2.1.3), (2.1.4) et (2.1.5), nous obtenons donc,

$$\begin{aligned} & \frac{\partial}{\partial t} \left[\int_0^1 \{\Phi_f(d)\}(t, x) dx + \int_0^1 \{\Phi_b(d)\}(t, x) dx + \int_0^1 \{\Psi(d)\}(t, x) dx + \int_0^1 \{\Xi(d)\}(t, x) dx \right] \\ &= 2c_a \int_0^1 \frac{\partial^2 [\{\Phi_f(d)\} + \{\Phi_b(d)\}]}{\partial x^2} (t, x) dx + \int_0^1 \{Q(d)\}(t, x) dx \quad (3.4.81) \end{aligned}$$

Les fonctions $\{\Phi_f(d)\}$ et $\{\Phi_b(d)\}$ appartiennent à $\mathbf{H}_\#$, donc l'intégrale du premier terme du membre de droite est nulle. Donc,

$$\begin{aligned} & \frac{\partial}{\partial t} \left[\int_0^1 \{\Phi_f(d)\}(t, x) dx + \int_0^1 \{\Phi_b(d)\}(t, x) dx + \int_0^1 \{\Psi(d)\}(t, x) dx + \int_0^1 \{\Xi(d)\}(t, x) dx \right] \\ &= \int_0^1 \{Q(d)\}(t, x) dx \leq \sup_{t \in \mathbb{R}^+} \left(\int_0^1 \{Q(d)\}(t, x) dx \right) \quad (3.4.82) \end{aligned}$$

Nous considérons les fonctions $\{\Phi_f(d)\}$, $\{\Phi_b(d)\}$, $\{\Psi(d)\}$, $\{\Xi(d)\}$ et $\{O(d)\}$, respectivement solutions de (3.1.18), (3.1.19), (2.1.3), (2.1.4b) et (2.1.5). En additionnant les équations (3.1.18), (3.1.19), (2.1.3), (2.1.4b) et (2.1.5), nous obtenons,

$$\begin{aligned} & \frac{\partial}{\partial t} \left[\int_0^1 \{\Phi_f(d)\}(t, x) dx + \int_0^1 \{\Phi_b(d)\}(t, x) dx + \int_0^1 \{\Psi(d)\}(t, x) dx + \int_0^1 \{\Xi(d)\}(t, x) dx \right] \\ &= 2c_a \int_0^1 \frac{\partial^2 [\{\Phi_f(d)\} + \{\Phi_b(d)\}]}{\partial x^2} (t, x) dx \\ & \quad + \int_0^1 \{Q(d)\}(t, x) dx - \int_0^1 u_a \frac{\{O(d)\}(t)}{L_a} \{\Psi(d)\}(t, x) dx \quad (3.4.83) \end{aligned}$$

Comme précédemment, puisque les fonctions $\{\Phi_f(d)\}$ et $\{\Phi_b(d)\}$ appartiennent à $\mathbf{H}_\#$, l'intégrale du premier terme du membre de droite est nulle. De plus, d'après l'hypothèse (3.1.20) et les propriétés 3.4.4 et 3.4.3, les fonctions $\{O(d)\}$ et $\{\Psi(d)\}$ sont positives. Donc,

$$\frac{\partial}{\partial t} \left[\int_0^1 \{\Phi_f(\mathbf{d})\}(t, x) dx + \int_0^1 \{\Phi_b(\mathbf{d})\}(t, x) dx + \int_0^1 \{\Psi(\mathbf{d})\}(t, x) dx + \int_0^1 \{\Xi(\mathbf{d})\}(t, x) dx \right] \leq \sup_{t \in \mathbb{R}^+} \left(\int_0^1 \{Q(\mathbf{d})\}(t, x) dx \right) \quad (3.4.84)$$

Ainsi, avec ou sans coefficient limiteur, le modèle composé des équations (3.1.18), (3.1.19), (2.1.3), (2.1.5) et de l'équation (2.1.4b) et (2.1.4), possède une propriété de conservation.

Nous avons finalement démontré la propriété 3.4.7. ■

3.4.4 Majoration de la norme infinie de $\{\Phi_f(\mathbf{d})\}$, $\{\Phi_b(\mathbf{d})\}$ et $\{\Psi(\mathbf{d})\}$

L'objectif de cette section est de démontrer que :

Propriété 3.4.8. *Sous les hypothèses allant de (3.1.20) jusqu'à (3.1.29), si $\{\Psi(\mathbf{d})\}(0, \cdot) \in L^1((0, 1))$, alors il existe une unique solution à l'équation (2.1.3). De plus, s'il existe un temps T_F tel que pour tout $t > T_F$, et pour tout $x \in (0, 1)$, $\{Q(\mathbf{d})\}(t, x) = 0$, il existe une constante positive $K_\Psi > 0$ telle que,*

$$\sup_{\substack{t \in \mathbb{R}^+ \\ x \in (0,1)}} \{\Psi(\mathbf{d})\} = \|\{\Psi(\mathbf{d})\}\|_\infty \leq K_\Psi. \quad (3.4.85)$$

Avant cela nous devons démontrer que,

Propriété 3.4.9. *Sous les hypothèses allant de (3.1.20) jusqu'à (3.1.29) et d'après la propriété 3.4.2, s'il existe un temps T_F tel que pour tout $t > T_F$, et pour tout $x \in (0, 1)$, $\{Q(\mathbf{d})\}(t, x) = 0$, alors il existe une constante $K_\Phi > 0$ telle que $\{\Phi_f(\mathbf{d})\} \in \mathbf{H}$ et $\{\Phi_b(\mathbf{d})\} \in \mathbf{H}$, solutions de (3.1.18) et (3.1.19) satisfont,*

$$\sup_{\substack{t \in \mathbb{R}^+ \\ x \in (0,1)}} (\{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\}) = \|\{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\}\|_\infty \leq K_\Phi. \quad (3.4.86)$$

Démonstration de la propriété 3.4.9 : Nous posons,

3.4. Démonstration de l'existence et de l'unicité de la solution d'un deuxième modèle simplifié :
un modèle mathématique sans terme de convection

$$M_{\Phi}(t, x) = \max \left(0, f_a(\{\Phi_f(\mathbf{d})\}(t, x) + \{\Phi_b(\mathbf{d})\}(t, x)) - K \right), \quad (3.4.87)$$

où,

$$K = \sup_{\substack{t \in \mathbb{R}^+ \\ x \in (0,1)}} (\{Q(\mathbf{d})\}(t, x)). \quad (3.4.88)$$

Dans un premier temps nous additionnons les équations (3.1.18) et (3.1.19), nous multiplions l'égalité obtenue par M_{Φ} puis nous l'intégrons par rapport à x :

$$\begin{aligned} & \int_0^1 \frac{\partial \{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\}}{\partial t}(t, x) M_{\Phi}(t, x) dx - \int_0^1 2c_a \frac{\partial^2 [\{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\}]}{\partial x^2}(t, x) M_{\Phi}(t, x) dx \\ &= \int_0^1 \{Q(\mathbf{d})\}(t, x) M_{\Phi}(t, x) dx - \int_0^1 f_a(\{\Phi_f(\mathbf{d})\}(t, x) + \{\Phi_b(\mathbf{d})\}(t, x)) M_{\Phi}(t, x) dx. \end{aligned} \quad (3.4.89)$$

En réalisant une intégration par parties du second terme nous obtenons,

$$\begin{aligned} & \int_0^1 \frac{\partial \{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\}}{\partial t}(t, x) M_{\Phi}(t, x) dx + \int_0^1 2c_a \frac{\partial [\{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\}]}{\partial x}(t, x) \frac{\partial M_{\Phi}}{\partial x}(t, x) dx \\ &= \int_0^1 \{Q(\mathbf{d})\}(t, x) M_{\Phi}(t, x) dx - \int_0^1 f_a(\{\Phi_f(\mathbf{d})\}(t, x) + \{\Phi_b(\mathbf{d})\}(t, x)) M_{\Phi}(t, x) dx. \end{aligned} \quad (3.4.90)$$

Nous avons

$$\frac{\partial [\{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\}]}{\partial t}(t, x) M_{\Phi}(t, x) = \frac{1}{f_a} \frac{\partial M_{\Phi}}{\partial t}(t, x) M_{\Phi}(t, x) = \frac{1}{2f_a} \frac{\partial (M_{\Phi})^2}{\partial t}(t, x), \quad (3.4.91)$$

et,

$$\frac{\partial[\{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\}]}{\partial x}(t, x) \frac{\partial M_\Phi}{\partial x}(t, x) = \frac{1}{f_a} \frac{\partial M_\Phi}{\partial x}(t, x) \frac{\partial M_\Phi}{\partial x}(t, x) = \frac{1}{f_a} \left(\frac{\partial M_\Phi}{\partial x}(t, x) \right)^2. \quad (3.4.92)$$

Ainsi, en utilisant (3.4.91) et (3.4.92), et à partir de l'équation (3.4.90), nous obtenons :

$$\begin{aligned} & \frac{1}{2f_a} \int_0^1 \frac{\partial(M_\Phi)^2}{\partial t}(t, x) dx + \int_0^1 \frac{2c_a}{f_a} \left(\frac{\partial(M_\Phi)}{\partial x} \right)^2(t, x) dx \\ &= \int_0^1 \{Q(\mathbf{d})\}(t, x) M_\Phi(t, x) dx - \int_0^1 f_a \left(\{\Phi_f(\mathbf{d})\}(t, x) + \{\Phi_b(\mathbf{d})\}(t, x) \right) M_\Phi(t, x) dx. \end{aligned} \quad (3.4.93)$$

Par définition la fonction M_Φ est positive. Donc, d'après l'hypothèse (3.1.20) et l'égalité (3.4.88) nous avons,

$$\begin{aligned} \frac{1}{2f_a} \int_0^1 \frac{\partial(M_\Phi)^2}{\partial t}(t, x) dx &\leq \int_0^1 K M_\Phi(t, x) dx \\ &\quad - \int_0^1 f_a \left(\{\Phi_f(\mathbf{d})\}(t, x) + \{\Phi_b(\mathbf{d})\}(t, x) \right) M_\Phi(t, x) dx. \end{aligned} \quad (3.4.94)$$

Ainsi, nous avons,

$$\frac{1}{2f_a} \frac{\partial \|M_\Phi\|_2^2}{\partial t}(t) dx \leq - \int_0^1 (M_\Phi(t, x))^2 dx \leq 0. \quad (3.4.95)$$

La fonction $\|M_\Phi\|_2^2$ est donc décroissante et d'après l'hypothèse (3.1.22) $\|M_\Phi\|_2^2(0) = 0$. Donc,

$$\|M_\Phi\|_2^2(t, \cdot) \leq 0 \quad (3.4.96)$$

Pour tout $t \in (0, +\infty)$ et pour tout $x \in (0, 1)$, la fonction $M_\Phi(t, x)$ est donc nulle. Cela signifie que pour tout $t \in (0, +\infty)$ et pour tout $x \in (0, 1)$, $(\{\Phi_f(\mathbf{d})\} + \{\Phi_b(\mathbf{d})\})$ est bornée par une constante positive K_Φ . Donc, la propriété 3.4.9 est bien vérifiée. ■

Démonstration de la propriété 3.4.8 : L'expression exacte de la fonction $\{\Psi(d)\}$ est donnée par l'équation (3.4.63). À partir de cette équation et d'après l'hypothèse (3.1.26) et la propriété 3.4.9 il existe une constante positive K_Ψ telle que la propriété 3.4.8 soit vérifiée. ■

3.4.5 Existence de la solution des équations (2.1.4b) et (2.1.4)

Dans cette section nous allons démontrer que :

Propriété 3.4.10. *Sous les hypothèses allant de (3.1.20) jusqu'à (3.1.29), les équations (2.1.4) et (2.1.4b), associées à l'équation (2.1.5), admettent une unique solution appartenant à $C^1((0, T), L^1((0, 1)))$, $T \in (0, +\infty)$.*

Démonstration de la propriété 3.4.10 :

L'équation (2.1.4) est une EDO linéaire. Cette équation admet donc une unique solution.

L'objectif de cette section est de montrer que l'équation (2.1.4b) admet une unique solution. Pour cela nous sommes basés sur la démonstration du théorème de Cauchy-Lipschitz (TESCHL (2012) et O'REGAN (2013)). Nous posons :

$$G_{\xi_1}(t, x) = G_{\xi_1}(0, x) + \int_0^t g(\tau, x) \left(\frac{L_d - s_{\xi_1}(\tau)}{L_d} \right) d\tau, \quad s_{\xi_1}(t) = \int_0^1 \xi_1(t, x) dx, \quad (3.4.97)$$

et,

$$G_{\xi_2}(t, x) = G_{\xi_2}(0, x) + \int_0^t g(\tau, x) \left(\frac{L_d - s_{\xi_2}(\tau)}{L_d} \right) d\tau, \quad s_{\xi_2}(t) = \int_0^1 \xi_2(t, x) dx, \quad (3.4.98)$$

où,

$$g(t, x) = u_d \{\Psi(d)\}(t, x), \quad (3.4.99)$$

et,

$$G_{\xi_1}(0, x) = G_{\xi_2}(0, x), \quad (3.4.100)$$

Les fonctions ξ_1 et ξ_2 sont positives et appartiennent à $C^1((0, T), L^1((0, 1)))$ et les fonctions s_{ξ_1} et s_{ξ_2} sont positives appartiennent à $C^1((0, T))$, $T \in (0, +\infty)$.

Nous soustrayons l'équation (3.4.98) et l'équation (3.4.97) :

$$\begin{aligned} |G_{\xi_1} - G_{\xi_2}|(t, x) &= \left| \int_0^t g(\tau, x) \left(\frac{s_{\xi_2}(\tau) - s_{\xi_1}(\tau)}{L_d} \right) d\tau \right| \\ &= \left| \int_0^t \int_0^1 g(\tau, x) \left(\frac{\xi_1(\tau, \alpha) - \xi_2(\tau, \alpha)}{L_d} \right) d\tau d\alpha \right| \end{aligned} \quad (3.4.101)$$

et $\forall t \leq T_1$, $T_1 \in (0, T)$, nous avons :

$$\begin{aligned} |G_{\xi_1} - G_{\xi_2}|(t, x) &\leq \frac{T_1}{L_d} \int_0^1 \bar{g}(x) \|\xi_1(\alpha) - \xi_2(\alpha)\|_{L^\infty((0, T_1))} d\alpha, \\ &\leq \frac{T_1 \bar{g}}{L_d} \|\xi_1 - \xi_2\|_{L^\infty((0, T_1), L^1((0, 1)))}, \end{aligned} \quad (3.4.102)$$

où,

$$\bar{g}(x) = u_d \|\{\Psi(d)\}(x)\|_{L^\infty((0, T_1))} \quad \text{et} \quad \bar{g} = u_d \|\|\{\Psi(d)\}\|_{L^\infty((0, T_1) \times (0, 1))} \quad (3.4.103)$$

$$\|\{\Psi(d)\}(x)\|_{L^\infty((0, T_1))} = \sup_{t \in (0, T_1)} |\{\Psi(d)\}(t, x)| \quad (3.4.104)$$

et,

$$\|\|\{\Psi(d)\}\|_{L^\infty((0, T_1) \times (0, 1))} = \sup_{\substack{t \in (0, T_1) \\ x \in (0, 1)}} \|\{\Psi(d)\}(t, x)\| \quad (3.4.105)$$

D'après la propriété 3.4.8 \bar{g} est bornée, nous choisissons donc $T_1 < \frac{L_d}{\bar{g}}$. Pour tout $t \leq T_1$, nous avons donc :

3.4. Démonstration de l'existence et de l'unicité de la solution d'un deuxième modèle simplifié :
un modèle mathématique sans terme de convection

$$|G_{\xi_1} - G_{\xi_2}|(t, x) \leq k \left\| \xi_1 - \xi_2 \right\|_{L^\infty((0, T_1), L^1((0, 1)))}, \quad \text{où } k < 1 \quad (3.4.106)$$

Nous obtenons :

$$\left\| G_{\xi_1} - G_{\xi_2} \right\|_{L^\infty((0, T_1), L^1((0, 1)))} \leq k \left\| \xi_1 - \xi_2 \right\|_{L^\infty((0, T_1), L^1((0, 1)))}. \quad (3.4.107)$$

L'équation (3.4.107) démontre que l'opérateur $G : \xi \mapsto G_\xi$, est k -contractant de $C^0((0, T_1), L^1((0, 1)))$ dans $C^0((0, T_1), L^1((0, 1)))$. Or, $\{\Xi(d)\}$ qui est solution de l'équation (2.1.4b) correspond au point fixe de G . D'après le théorème du point fixe de Banach, ce point fixe existe et est unique. Donc l'équation (2.1.4b) admet une unique solution pour tout $t \in (0, T_1)$ et pour tout $x \in (0, 1)$.

La valeur de la fonction $\{\Xi(d)\}$ au point $t = T_1$ est donnée. Pour tout t tel que $T_1 \leq t \leq T_2$, $T_2 \in (0, T)$, nous avons donc :

$$\begin{aligned} |G_{\xi_1} - G_{\xi_2}|(t, x) &\leq \frac{T_2 - T_1}{L_d} \int_0^1 \bar{g}(x) \left\| \xi_1(\alpha) - \xi_2(\alpha) \right\|_{L^\infty((T_1, T_2))} d\alpha, \\ &\leq \frac{(T_2 - T_1) \bar{g}}{L_d} \left\| \xi_1 - \xi_2 \right\|_{L^\infty((T_1, T_2), L^1((0, 1)))}. \end{aligned} \quad (3.4.108)$$

Nous choisissons T_2 tel que $T_2 - T_1 < \frac{L_d}{\bar{g}}$. Donc, pour tout $t \in (T_1, T_2)$ et pour tout $x \in (0, 1)$, nous avons :

$$|G_{\xi_1} - G_{\xi_2}|(t, x) \leq k_2 \left\| \xi_1 - \xi_2 \right\|_{L^\infty((T_1, T_2), L^1((0, 1)))}, \quad \text{où } k_2 < 1 \quad (3.4.109)$$

Nous avons donc :

$$\left\| G_{\xi_1} - G_{\xi_2} \right\|_{L^\infty((T_1, T_2), L^1((0, 1)))} \leq k_2 \left\| \xi_1 - \xi_2 \right\|_{L^\infty((T_1, T_2), L^1((0, 1)))}. \quad (3.4.110)$$

L'équation (3.4.110) démontre que l'opérateur G est k -contractant de $C^0((T_1, T_2), L^1((0, 1)))$ dans $C^0((T_1, T_2), L^1((0, 1)))$. Or $\{\Xi(d)\}$ qui est solution de l'équation (2.1.4b), est le point fixe de l'application G . Or, d'après le théorème du point fixe de Banach, ce point fixe existe et est unique. Donc, l'équation (2.1.4b) admet une unique solution pour tout $t \in (T_1, T_2)$ et pour tout $x \in (0, 1)$.

Par itération et en procédant de la même manière nous pouvons parcourir ainsi l'ensemble de l'intervalle de temps $(0, T)$, $T \in (0, +\infty)$. Ainsi, par itération nous pouvons démontrer que G est k -contractante sur une série d'intervalles (T_n, T_{n+1}) couvrant l'intervalle $(0, T)$ et vérifiant $T_n - T_{n+1} \leq \frac{L_d}{g}$. D'après le théorème du point fixe de Banach, puisque $\{\Xi(d)\}$ est le point fixe de G et solution de l'équation (2.1.4b), cette équation admet une unique solution sur tous ces intervalles de temps et pour tout $x \in (0, 1)$.

Nous avons finalement démontré la propriété 3.4.10. ■

Puisque l'équation (2.1.5) est une EDO, l'existence et l'unicité de sa solution est évidente.

Nous avons finalement montré dans cette section l'existence de la solution du second sous-modèle composé des équations (3.1.18), (3.1.19), (2.1.3), (2.1.5) et de l'équation (2.1.4b) ou (2.1.4).

3.4.6 Unicité de la solution du second modèle simplifié

Dans cette section nous allons démontrer l'unicité de la solution du système d'EDP sur lequel est basé le second sous-modèle.

Nous supposons que $\{\Phi_{f_1}(d)\}$ et $\{\Phi_{f_2}(d)\}$ sont solutions de (3.1.18), et $\{\Phi_{b_1}(d)\}$ et $\{\Phi_{b_2}(d)\}$ sont solutions de (3.1.19). L'objectif de cette démonstration est de montrer que,

$$\{\Phi_{f_1}(d)\} = \{\Phi_{f_2}(d)\} \tag{3.4.111}$$

et,

$$\{\Phi_{b_1}(d)\} = \{\Phi_{b_2}(d)\} \tag{3.4.112}$$

3.4. Démonstration de l'existence et de l'unicité de la solution d'un deuxième modèle simplifié :
un modèle mathématique sans terme de convection

Par soustraction nous obtenons à partir de l'équation (3.1.18),

$$\begin{aligned} \frac{\partial \{\Phi_{f_1}(\mathbf{d})\} - \{\Phi_{f_2}(\mathbf{d})\}}{\partial t}(t, x) - c_a \frac{\partial \left[\frac{\partial \left[\{\Phi_{f_1}(\mathbf{d})\} - \{\Phi_{f_2}(\mathbf{d})\} + \{\Phi_{b_1}(\mathbf{d})\} - \{\Phi_{b_2}(\mathbf{d})\} \right]}{\partial x} \right]}{\partial x}(t, x) \\ = -f_a (\{\Phi_{f_1}(\mathbf{d})\} - \{\Phi_{f_2}(\mathbf{d})\})(t, x) - r_a (\{\Phi_{f_1}(\mathbf{d})\} - \{\Phi_{f_2}(\mathbf{d})\})(t, x). \quad (3.4.113) \end{aligned}$$

et à partir de l'équation (3.1.19),

$$\begin{aligned} \frac{\partial \{\Phi_{b_1}(\mathbf{d})\} - \{\Phi_{b_2}(\mathbf{d})\}}{\partial t}(t, x) - c_a \frac{\partial \left[\frac{\partial \left[\{\Phi_{f_1}(\mathbf{d})\} - \{\Phi_{f_2}(\mathbf{d})\} + \{\Phi_{b_1}(\mathbf{d})\} - \{\Phi_{b_2}(\mathbf{d})\} \right]}{\partial x} \right]}{\partial x}(t, x) \\ = -f_a (\{\Phi_{b_1}(\mathbf{d})\} - \{\Phi_{b_2}(\mathbf{d})\})(t, x) + r_a (\{\Phi_{f_1}(\mathbf{d})\} - \{\Phi_{f_2}(\mathbf{d})\})(t, x). \quad (3.4.114) \end{aligned}$$

Nous multiplions l'équation (3.4.113) par $(\{\Phi_{f_1}(\mathbf{d})\} - \{\Phi_{f_2}(\mathbf{d})\})$ et l'équation (3.4.114) par $(\{\Phi_{b_1}(\mathbf{d})\} - \{\Phi_{b_2}(\mathbf{d})\})$, puis nous additionnons les deux équations obtenues :

$$\begin{aligned}
 & \frac{\partial \{\Phi_{f_1}(\mathbf{d})\} - \{\Phi_{f_2}(\mathbf{d})\}}{\partial t}(t, x) (\{\Phi_{f_1}(\mathbf{d})\}(t, x) - \{\Phi_{f_2}(\mathbf{d})\}(t, x)) \\
 & \quad + \frac{\partial \{\Phi_{b_1}(\mathbf{d})\} - \{\Phi_{b_2}(\mathbf{d})\}}{\partial t}(t, x) (\{\Phi_{b_1}(\mathbf{d})\}(t, x) - \{\Phi_{b_2}(\mathbf{d})\}(t, x)) \\
 -c_a & \frac{\partial \left[\frac{\partial [\{\Phi_{f_1}(\mathbf{d})\} - \{\Phi_{f_2}(\mathbf{d})\} + \{\Phi_{b_1}(\mathbf{d})\} - \{\Phi_{b_2}(\mathbf{d})\}]}{\partial x} \right]}{\partial x}(t, x) (\{\Phi_{f_1}(\mathbf{d})\}(t, x) - \{\Phi_{f_2}(\mathbf{d})\}(t, x)) \\
 -c_a & \frac{\partial \left[\frac{\partial [\{\Phi_{f_1}(\mathbf{d})\} - \{\Phi_{f_2}(\mathbf{d})\} + \{\Phi_{b_1}(\mathbf{d})\} - \{\Phi_{b_2}(\mathbf{d})\}]}{\partial x} \right]}{\partial x}(t, x) (\{\Phi_{b_1}(\mathbf{d})\}(t, x) - \{\Phi_{b_2}(\mathbf{d})\}(t, x)) \\
 & = -f_a \left((\{\Phi_{f_1}(\mathbf{d})\}(t, x) - \{\Phi_{f_2}(\mathbf{d})\}(t, x))^2 + (\{\Phi_{b_1}(\mathbf{d})\}(t, x) - \{\Phi_{b_2}(\mathbf{d})\}(t, x))^2 \right) \\
 & \quad - r_a (\{\Phi_{f_1}(\mathbf{d})\}(t, x) - \{\Phi_{f_2}(\mathbf{d})\}(t, x))^2 \\
 & \quad + r_a (\{\Phi_{f_1}(\mathbf{d})\}(t, x) - \{\Phi_{f_2}(\mathbf{d})\}(t, x)) (\{\Phi_{b_1}(\mathbf{d})\}(t, x) - \{\Phi_{b_2}(\mathbf{d})\}(t, x)) \quad (3.4.115)
 \end{aligned}$$

Nous intégrons l'équation (3.4.115) par rapport à x et nous obtenons :

$$\begin{aligned}
 & \frac{1}{2} \frac{\partial \|\{\Phi_{f_1}(\mathbf{d})\} - \{\Phi_{f_2}(\mathbf{d})\}\|_2^2}{\partial t}(t) + \frac{1}{2} \frac{\partial \|\{\Phi_{b_1}(\mathbf{d})\} - \{\Phi_{b_2}(\mathbf{d})\}\|_2^2}{\partial t}(t) \\
 & \quad - \int_0^1 c_a \left(\frac{\partial \{\Phi_{f_1}(\mathbf{d})\} - \{\Phi_{f_2}(\mathbf{d})\} + \{\Phi_{b_1}(\mathbf{d})\} - \{\Phi_{b_2}(\mathbf{d})\}}{\partial x} \right)^2 (t, x) dx \\
 & = - \int_0^1 f_a \left((\{\Phi_{f_1}(\mathbf{d})\}(t, x) - \{\Phi_{f_2}(\mathbf{d})\}(t, x))^2 + (\{\Phi_{b_1}(\mathbf{d})\}(t, x) - \{\Phi_{b_2}(\mathbf{d})\}(t, x))^2 \right) dx \\
 & \quad + \int_0^1 r_a (\{\Phi_{f_1}(\mathbf{d})\}(t, x) - \{\Phi_{f_2}(\mathbf{d})\}(t, x))^2 dx \\
 & \quad + \int_0^1 r_a (\{\Phi_{f_1}(\mathbf{d})\}(t, x) - \{\Phi_{f_2}(\mathbf{d})\}(t, x)) (\{\Phi_{b_1}(\mathbf{d})\}(t, x) - \{\Phi_{b_2}(\mathbf{d})\}(t, x)) dx \quad (3.4.116)
 \end{aligned}$$

À partir de l'inégalité (3.2.18) nous obtenons donc l'inégalité suivante :

$$\begin{aligned} & \frac{\partial \|\{\Phi_{f_1}(\mathbf{d})\} - \{\Phi_{f_2}(\mathbf{d})\}\|_2^2}{\partial t}(t) + \frac{\partial \|\{\Phi_{b_1}(\mathbf{d})\} - \{\Phi_{b_2}(\mathbf{d})\}\|_2^2}{\partial t}(t) \\ & \leq 2r_a \|\{\Phi_{f_1}(\mathbf{d})\} - \{\Phi_{f_2}(\mathbf{d})\}\|_2(t, \cdot) \|\{\Phi_{b_1}(\mathbf{d})\} - \{\Phi_{b_2}(\mathbf{d})\}\|_2(t, \cdot) \end{aligned} \quad (3.4.117)$$

À partir de l'inégalité (3.4.117) et en utilisant les identités remarquables nous obtenons :

$$\begin{aligned} & \frac{\partial \|\{\Phi_{f_1}(\mathbf{d})\} - \{\Phi_{f_2}(\mathbf{d})\}\|_2^2}{\partial t}(t) + \frac{\partial \|\{\Phi_{b_1}(\mathbf{d})\} - \{\Phi_{b_2}(\mathbf{d})\}\|_2^2}{\partial t}(t) \\ & \leq r_a (\|\{\Phi_{f_1}(\mathbf{d})\} - \{\Phi_{f_2}(\mathbf{d})\}\|_2^2(t) + \|\{\Phi_{b_1}(\mathbf{d})\} - \{\Phi_{b_2}(\mathbf{d})\}\|_2^2(t)). \end{aligned} \quad (3.4.118)$$

D'après de lemme de Gronwall, pour tout $T \in (0, +\infty)$ et pour tout $t \in (0, T)$, nous avons :

$$\begin{aligned} & \|\{\Phi_{f_1}(\mathbf{d})\} - \{\Phi_{f_2}(\mathbf{d})\}\|_2^2(t) + \|\{\Phi_{b_1}(\mathbf{d})\} - \{\Phi_{b_2}(\mathbf{d})\}\|_2^2(t) \\ & \leq \|\{\Phi_{f_1}(\mathbf{d})\} - \{\Phi_{f_2}(\mathbf{d})\}\|_2^2(0) + \|\{\Phi_{b_1}(\mathbf{d})\} - \{\Phi_{b_2}(\mathbf{d})\}\|_2^2(0) = 0 \end{aligned} \quad (3.4.119)$$

Nous déduisons de cette équation (3.4.119) que :

$$\|\{\Phi_{f_1}(\mathbf{d})\} - \{\Phi_{f_2}(\mathbf{d})\}\|_2^2(t) + \|\{\Phi_{b_1}(\mathbf{d})\} - \{\Phi_{b_2}(\mathbf{d})\}\|_2^2(t) = 0 \quad (3.4.120)$$

Nous avons donc,

$$\|\{\Phi_{f_1}(\mathbf{d})\} - \{\Phi_{f_2}(\mathbf{d})\}\|_2^2(t, \cdot) = 0 \quad (3.4.121)$$

et,

$$\|\{\Phi_{b_1}(\mathbf{d})\} - \{\Phi_{b_2}(\mathbf{d})\}\|_2^2(t, \cdot) = 0 \quad (3.4.122)$$

Donc,

$$\{\Phi_{f_1}(\mathbf{d})\}(t, x) = \{\Phi_{f_2}(\mathbf{d})\}(t, x) \quad (3.4.123)$$

et,

$$\{\Phi_{b_1}(\mathbf{d})\}(t, x) = \{\Phi_{b_2}(\mathbf{d})\}(t, x) \quad (3.4.124)$$

Nous avons donc démontré l'unicité de la solution des équations (3.1.18) et (3.1.19).

L'équation (2.1.3) est une EDO linéaire. L'unicité de sa solution est donc évidente. Nous avons démontré dans la section 3.4.5 que dans le cas où la fonction $\{\Xi(\mathbf{d})\} \in C^1((0, +\infty); L^1((0, 1)))$ est solution de l'équation (2.1.4b) cette fonction admet une unique solution.

L'équation (2.1.4) est une EDO. Cette équation admet donc une unique solution.

La fonction $\{O(\mathbf{d})\} \in C^1((0, +\infty))$ solution de l'équation (2.1.5) ne dépend que de $\{\Xi(\mathbf{d})\}$. L'équation (2.1.5) admet donc également une unique solution. Finalement, nous avons démontré l'unicité de la solution du système d'EDP.

Dans cette section nous avons montré l'existence et l'unicité de la solution du deuxième modèle simplifié. De plus, nous avons montré la propriété de conservation de ce modèle et la positivité de certaines composantes. Finalement, nous avons donc démontré le théorème 3.1.2. ■

3.5 Conclusion

Les deux premières équations du modèle mathématique développé initialement et présenté dans le chapitre 2, contiennent un "opérateur de diffusion" dépendant de la somme de deux fonctions inconnues. Cette dépendance, couplée à la présence de termes de convection, rend impossible l'application des résultats classiques concernant les équations paraboliques. Nous avons donc travaillé sur deux modèles simplifiés.

La perte de l'effet de régularisation et de la positivité des solutions liée à la struc-

ture du terme de diffusion a été mise en évidence grâce à l'analyse du premier sous-modèle.

L'analyse du second modèle simplifié qui ne contient pas de termes de convection a mis en évidence les propriétés conservées malgré l'écriture du terme de diffusion, via l'utilisation de changements de variables et de Transformées de Fourier. Grâce à ce modèle nous avons démontré l'existence et l'unicité de la solution, ainsi que la propriété de conservation de ce modèle.

INTÉGRATION DE VARIABLES D'ENTRÉE DANS DES OUTILS D'APPRENTISSAGE STATISTIQUE BIOMIMÉTIQUES

Dans le chapitre 2, nous avons construit un modèle mathématique principal, appelé modèle "A", constitué des équations (2.1.1), (2.1.2), (2.1.3), (2.1.4) et (2.1.5). Trois autres versions de l'équation "d'utilisation", données par les équations (2.1.4b), (2.1.4c) ou (2.1.4d), permettent de simuler et prédire une large gamme de variables à partir de la consommation alimentaire et des injections de médicaments administrés aux animaux. Dans le modèle "A", l'alimentation est considérée comme une information globale et chaque ingestion est transformée de manière globale en une injection d'information. Un modèle "A" ajusté sur une certaine ration alimentaire ne peut donc prédire l'impact que de cette même ration sur une variable de sortie. Pour étudier l'impact de n rations il faudrait alors ajuster n fois le modèle "A" et nous obtiendrions n modèles ajustés.

L'objectif de ce chapitre est donc d'étudier la façon dont l'alimentation peut être prise en compte de manière plus précise. L'objectif est également d'étudier la façon dont d'autres variables d'entrée peuvent être prises en compte dans des outils biomimétiques d'apprentissage statistique.

4.1 Valorisation de l'information contenue dans la variable relative à l'âge des animaux

L'objectif de cette partie est d'étudier la manière dont le modèle mathématique biomimétique peut valoriser l'information apportée par la variable indiquant l'âge de chaque individu au démarrage de la collecte de données. Pour cela nous avons construit

un modèle " A_{Age} ". Ce modèle est basé sur le "modèle A de croissance", composé des équations (2.1.1), (2.1.2), (2.1.3) et (2.1.5). Une nouvelle équation "d'utilisation", inspirée de l'équation (2.1.4b) a été construite pour ce modèle et sera présentée dans cette section.

Comme pour le modèle "A de croissance", les données d'entrée de ce modèle correspondent au poids initial et aux consommations alimentaires de chaque animal (Voir les sections 2.5.2 et 1.7.1). Mais pour ce modèle " A_{Age} " l'âge des animaux constitue également une donnée d'entrée. Ce modèle fournit en sortie une prédiction du poids des animaux à un instant t .

4.1.1 Construction d'une courbe de croissance moyenne

Dans cette étude nous avons voulu moduler la valeur du poids maximum atteignable par chaque animal (P_{max_i}), en fonction de son potentiel de croissance. Pour cela nous avons construit un indicateur de ce potentiel de croissance en fonction du poids initial de chaque individu et de son âge.

Nous avons tout d'abord cherché à savoir si les individus étaient en retard ou en avance en termes de croissance par rapport aux autres. Nous avons fait l'hypothèse que cette information pouvait nous donner une indication concernant le potentiel de croissance de chaque individu. Nous avons imaginé qu'un animal qui est déjà en retard au démarrage de l'engraissement, n'atteindra pas le même poids à la fin de la période d'engraissement qu'un animal déjà en avance.

Pour savoir si un individu est en retard ou en avance il faut pouvoir comparer son poids de départ avec le poids d'un individu moyen du même âge. Nous avons donc construit une courbe de croissance moyenne. Les animaux de l'espèce étudiée commencent l'engraissement à un âge sans dimension variant entre 5.3 et 7.1. Nous devons donc construire une courbe de poids moyens définie sur cette gamme d'âge. Les animaux sont pesés lors de leur arrivée dans l'élevage puis conservés quelques jours tous ensemble dans un enclos, avant d'être répartis en groupe et de commencer réellement l'expérimentation et la collecte de données. Au cours de ces quelques jours les animaux ont reçu tous la même alimentation.

Nous connaissons donc le poids et l'âge des individus quelques jours avant le démarrage de l'expérimentation (P_{A_i} et Age_{A_i}) et au démarrage de l'expérimentation (P_{0_i}

et Age_{0_i}).

Nous avons construit pour chaque individu une courbe de croissance entre le moment son arrivé sur l'exploitation (t_{A_i}), jusqu'au démarrage de l'engraissement ($t_0 = 0$). La période adimensionnée qui sépare t_{A_i} et t_0 varie entre 0.64 et 0.88. Cette période est assez courte, nous faisons donc l'hypothèse que la courbe de croissance moyenne sur cette période est assimilable à une droite. L'équation de cette courbe pour l'individu i est donc une équation de la forme :

$$y_i = a_i x + b_i, \quad (4.1.1)$$

où,

$$a_i = \frac{P_{0_i} - P_{A_i}}{Age_{0_i} - Age_{A_i}}, \quad (4.1.2)$$

et,

$$b_i = P_{A_i} - \frac{P_{0_i} - P_{-1_i}}{Age_{0_i} - Age_{A_i}} Age_{A_i}. \quad (4.1.3)$$

Dans ces équations les variables Age_{A_i} , P_{A_i} , Age_{0_i} et P_{0_i} correspondent respectivement à l'âge et au poids de l'animal i , à l'instant $t = t_A$ et à $t = 0$.

Pour chaque individu i , nous avons calculé les coefficients a_i et b_i et nous avons ensuite calculé la pente moyenne (a_{Av}) et l'ordonnée à l'origine moyenne (b_{av}) pour l'ensemble des individus. Nous obtenons finalement une droite décrivant l'évolution de la croissance d'un individu moyen, dont l'âge sans dimension varie entre 5.3 et 7.1 et dont l'équation est donnée par,

$$y_{Av} = a_{Av}x + b_{Av}. \quad (4.1.4)$$

Nous avons ensuite placé les différents individus étudiés sur cette droite en fonction de leur âge et de leur poids au démarrage de l'expérimentation (t_0).

Les animaux dont le poids est placé au dessus de la droite sont donc considérés en avance en termes de croissance et ceux dont le poids est placé en dessous de cette

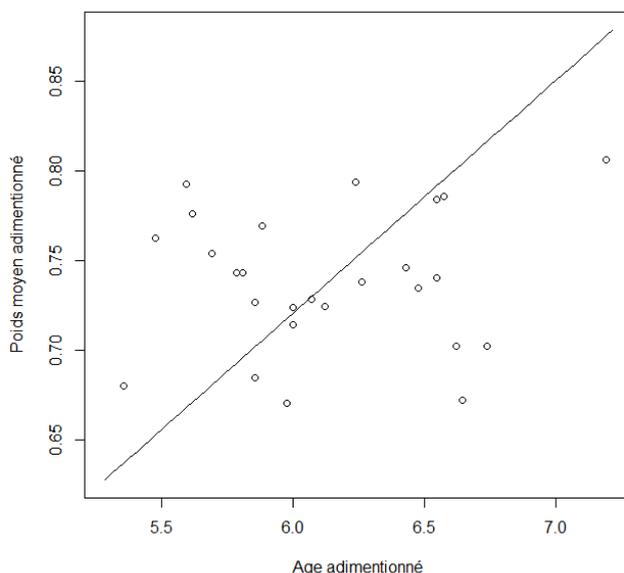


FIGURE 4.1 – Droite de poids moyen adimensionné et positionnement des individus étudiés sur cette droite.

droite sont considérés en retard.

4.1.2 Modulation de la valeur du paramètre P_{max_i}

Nous avons fait le choix d'utiliser l'information concernant l'avance ou le retard de croissance des individus pour moduler le poids qu'ils pourraient atteindre.

Plusieurs hypothèses peuvent être faites. Par exemple, pour les individus considérés en retard de croissance nous pouvons faire trois hypothèses :

- Le retard de croissance constaté au démarrage de l'engraissement s'estompe avec le temps et les animaux rattrapent leur retard ;
- Le retard de croissance constaté au démarrage de l'engraissement reste stable avec le temps et l'écart de poids existant par rapport au poids des individus moyens reste constant ;
- Le retard de croissance constaté au démarrage de l'engraissement s'accroît avec le temps et l'écart de poids augmente.

Afin de prendre ces trois hypothèses en compte nous avons décidé d'écrire l'équa-

tion "d'utilisation" telle que :

$$\frac{\partial \{\Xi(d)\}}{\partial t}(t, x) = u_d \{\Psi(d)\}(t, x) \cdot \left(\frac{P_{MaxCorr_i} - O(t)}{P_{MaxCorr_i}} \right) \quad (4.1.5)$$

où,

$$P_{MaxCorr_i} = \alpha(P_{0_i} - P_{0_{moy}}(Age_{0_i})) + P_{Max} \quad (4.1.6)$$

Les variables P_{0_i} et Age_{0_i} correspondent au poids et à l'âge de l'individu i au démarrage de l'expérimentation et la variable $P_{0_{moy}}(Age_{0_i})$ correspond au poids moyen donné à l'âge Age_{0_i} . Cette dernière variable peut être vue comme le poids que l'animal i devrait faire à l'âge Age_{0_i} . P_{Max} , correspond à un poids maximum de référence pour l'espèce étudiée.

En ajustant le paramètre α sur des données d'apprentissage, nous pouvons répondre aux hypothèses posées précédemment. En effet, si ce paramètre vaut 1 cela signifie que le retard de croissance constaté au démarrage de l'engraissement reste stable avec le temps. Si la valeur de ce paramètre est supérieure à 1 cela signifie que le retard de croissance constaté au démarrage de l'engraissement s'accroît avec le temps. Si la valeur de ce paramètre est comprise entre 0 et 1 cela signifie que le retard de croissance constaté au démarrage s'estompe au cours de l'engraissement. Enfin, si la valeur de α est négative cela signifie que la tendance s'inverse et qu'un animal en retard au début de l'expérimentation peut finalement rattraper les autres et devenir même plus gros que la moyenne.

4.1.3 Ajustement des paramètres du modèle « A_{Age} »

Comme dans la section 2.5.3, nous avons utilisé les données du groupe "Témoin A" comme base d'apprentissage et les données du groupe "Témoin B" pour tester la qualité de l'outil de prédiction obtenu (Voir la description des données dans la section 1.7.1).

Grâce aux relations établies dans la section 2.5.3, nous avons ajusté ω_d , f_d , P_{Max} et α et nous en avons déduit les valeurs de r_d et u_d .

Les paramètres ont été ajustés sur la base d'apprentissage en utilisant l'algorithme DIRECT et en minimisant la fonction «objectif» donnée par l'équation (2.5.3) donnée dans la section 2.5.3.

Comme dans la section 2.5.3, pour tester la stabilité des valeurs du set de paramètres minimisant l'erreur de l'outil de prédiction sur la base d'apprentissage, nous avons réalisé plusieurs ajustements. Pour cela nous avons sélectionné 7 individus sur 8 et nous avons ajusté les paramètres sur les données associées à ces 7 individus sélectionnés. Nous avons donc réalisé 8 ajustements et nous avons obtenu 8 sets de paramètres ajustés $(\omega_a, r_a, f_a, u_a, L_a, \alpha)$.

Nous avons calculé la moyenne et l'écart type relatif (*RSD*) associé à chaque paramètre (Table 4.1). Les valeurs du *RSD* sont très faibles. Cela signifie que notre méthode d'apprentissage permet d'identifier un unique set de paramètres minimisant l'erreur associée au modèle sur la base d'apprentissage. L'existence d'un unique set de paramètres optimaux $(\omega_a, r_a, f_a, u_a, P_{Max}, \alpha)$ atteste de l'identifiabilité du modèle.

Le modèle mathématique a été paramétré avec la valeur moyenne de chaque paramètre. Dans un premier temps, nous avons calculé l'erreur de l'outil de prédiction basé sur ce modèle, sur la base d'apprentissage. Pour cela nous avons calculé l'écart relatif moyen (*ARE*) entre la valeur mesurée et la valeur prédite de la variable de sortie à l'instant $t = 1$. Cet indicateur est donné par l'équation (2.5.4). Les résultats obtenus sont présentés dans la table 4.1.

TABLE 4.1 – La moyenne et l'écart type relatif (*RSD*) des paramètres ajustés, et l'indicateur *ARE* calculé l'instant $t = 1$ sur la base d'apprentissage.

<i>Paramtre</i>	<i>Moyenne</i>	<i>RSD</i>
ω	17.66	0.14
r	52.85	0.22
f	238.62	0.02
u	587.78	0.15
P_{Max}	1.71	0.01
α	-0.162	0.35
<i>ARE</i> (1) (%)	1.84	0.013

La valeur de l'*ARE* à l'instant $t = 1$ sur la base d'apprentissage vaut 1.84%, ce qui est satisfaisant. Nous avons ensuite calculé la valeur de l'indicateur *ARE* sur la base

de test aux instants $t = 0.6$, $t = 1.52$ et $t = 2.5$ (Voir la table 4.2 et la figure 4.2).

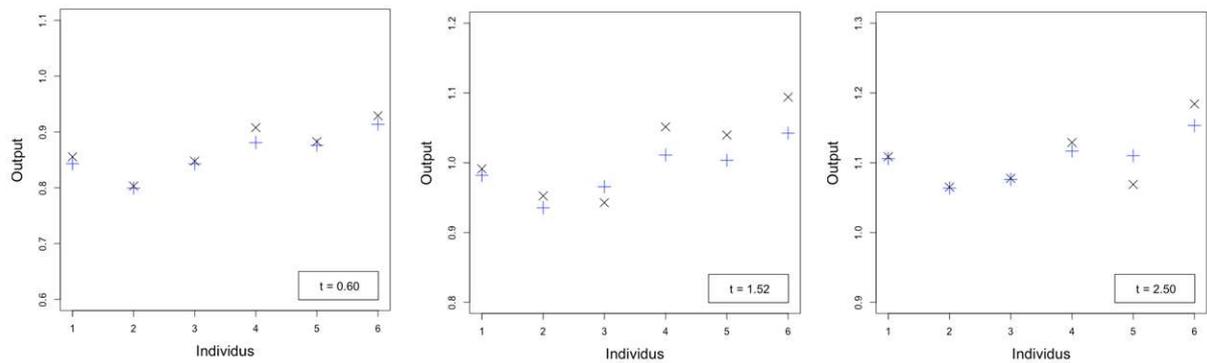


FIGURE 4.2 – Différences obtenues entre les poids mesurés (+) et ceux prédits (x), à différents instants t , pour les individus de la base de test.

TABLE 4.2 – Erreur relative moyenne (ARE) calculée à différents instants t sur les individus de la base de test.

t	0.6	1.52	2.5
$ARE(t)$ du modèle "A" (%)	1.3	2.9	1.5
$ARE(t)$ du modèle " A_{Age} " (%)	1.3	2.8	1.3

Les résultats contenus dans la table 4.2 montrent que la prise en compte de l'âge des animaux dans le modèle permet d'augmenter légèrement la précision du modèle sur les données de test par rapport aux résultats obtenus avec le modèle "A de croissance" (section 2.5.5). Nous constatons également que la valeur ajustée du paramètre α est négative (Table 4.1). Cela signifie que la tendance s'inverse et qu'un animal en retard au début de l'expérimentation a tendance à rattraper les autres et à devenir même plus gros que la moyenne. Néanmoins, comme ce paramètre est compris entre 0 et -1 , cela signifie que la différence observée une fois adulte n'est pas aussi importante que celle observée au démarrage de l'expérimentation.

4.2 Construction d'un modèle simulant la circulation et l'action de deux informations en interaction : le modèle «B»

4.2.1 Objectifs et intérêts de ce deuxième modèle

Un deuxième modèle plus complexe a été construit afin de modéliser plus finement l'effet de l'alimentation sur la croissance des animaux. En effet, l'alimentation des animaux d'élevage se compose généralement d'une ration de base et d'une quantité plus ou moins importante d'additif. L'additif étant une forme de complément alimentaire favorisant la digestion et l'assimilation des nutriments de la ration. L'objectif a donc été de modéliser séparément la circulation de l'information relative à l'ingestion de la ration de base et celle relative à l'ingestion de l'additif, d'apprendre ensuite leur dynamique d'évolution respective, leur interaction et leur action sur la variable de sortie qui modélise dans ce cas l'évolution du poids des animaux.

L'intérêt de ce modèle «B» est qu'une fois la dynamique et l'interaction existante entre la ration de base et de l'additif sont apprises, il est alors possible de faire varier la dose d'additif et de simuler l'effet de différents dosages sur la croissance des individus. L'objectif final est de trouver la dose d'additif optimisant les performances des animaux en termes de croissance, sans tester expérimentalement toutes les doses d'additif possibles.

L'outil de prédiction basé sur ce modèle "B" prend en données d'entrée le poids initial des individus, la part d'additif présente dans leur ration globale et leur consommation alimentaire ($C_{0-0.67}$, $C_{0.67-1.5}$ et $C_{1.5-2.5}$) (section 1.7.1). Ce modèle fournit en données de sortie une prédiction du poids des animaux à l'instant t .

4.2.2 La structure du modèle "B"

La structure globale du modèle «B» est schématisée par la figure 4.3.

Dans ce modèle "B" nous avons deux EDP décrivant le flux avant, $\{\Phi_f(A)\}$ et le flux arrière, $\{\Phi_b(A)\}$ de l'information relative à l'additif. Ainsi $\{\Phi_f(A)\}$ est solution de :

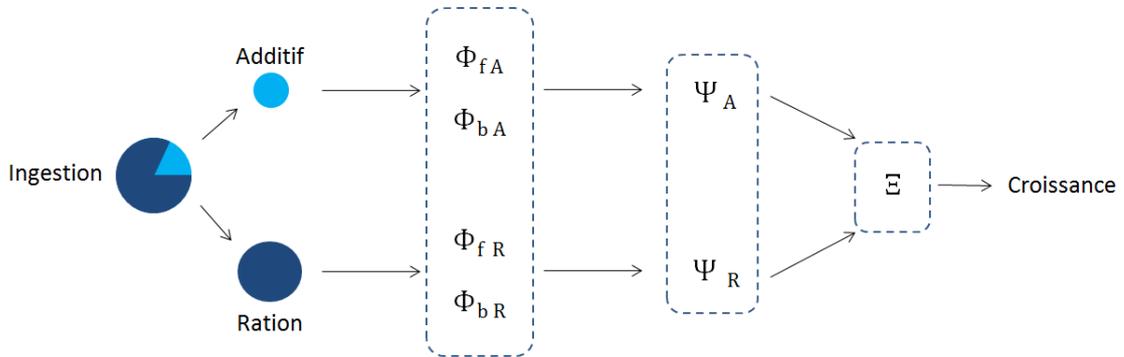


FIGURE 4.3 – Schématisation de la structure globale du modèle «B».

$$\frac{\partial \{\Phi_f(\mathbf{A})\}}{\partial t}(t, x) = \frac{1}{2}\{Q(\mathbf{A})\}(t, x) - \omega_{\mathbf{A}} \frac{\partial \{\Phi_f(\mathbf{A})\}}{\partial x}(t, x) + \frac{\partial \left[c_{\mathbf{A}} \chi \frac{\partial [\{\Phi_f(\mathbf{A})\} + \{\Phi_b(\mathbf{A})\}]}{\partial x} \right]}{\partial x}(t, x) - r_{\mathbf{A}} \{\Phi_f(\mathbf{A})\}(t, x) - f_{\mathbf{A}} \{F(\mathbf{A})\}(x) \{\Phi_f(\mathbf{A})\}(t, x), \quad (4.2.1)$$

et $\{\Phi_b(\mathbf{A})\}$ est solution de :

$$\frac{\partial \{\Phi_b(\mathbf{A})\}}{\partial t}(t, x) = \frac{1}{2}\{Q(\mathbf{A})\}(t, x) + \omega_{\mathbf{A}} \frac{\partial \{\Phi_b(\mathbf{A})\}}{\partial x}(t, x) + \frac{\partial \left[c_{\mathbf{A}} \chi \frac{\partial [\{\Phi_f(\mathbf{A})\} + \{\Phi_b(\mathbf{A})\}]}{\partial x} \right]}{\partial x}(t, x) + r_{\mathbf{A}} \{\Phi_f(\mathbf{A})\}(t, x) - f_{\mathbf{A}} \{F(\mathbf{A})\}(x) \{\Phi_b(\mathbf{A})\}(t, x). \quad (4.2.2)$$

Nous avons ensuite deux EDP décrivant le flux avant, $\{\Phi_f(\mathbf{R})\}$ et le flux arrière, $\{\Phi_b(\mathbf{R})\}$ de l'information relative à la ration de base. Ainsi $\{\Phi_f(\mathbf{R})\}$ est solution de :

$$\frac{\partial \{\Phi_f(\mathbf{R})\}}{\partial t}(t, x) = \frac{1}{2}\{Q(\mathbf{R})\}(t, x) - \omega_{\mathbf{R}} \frac{\partial \{\Phi_f(\mathbf{R})\}}{\partial x}(t, x) + \frac{\partial \left[c_{\mathbf{R}} \chi \frac{\partial [\{\Phi_f(\mathbf{R})\} + \{\Phi_b(\mathbf{R})\}]}{\partial x} \right]}{\partial x}(t, x) - r_{\mathbf{R}} \{\Phi_f(\mathbf{R})\}(t, x) - f_{\mathbf{R}} \{F(\mathbf{R})\}(x) \{\Phi_f(\mathbf{R})\}(t, x), \quad (4.2.3)$$

et $\{\Phi_b(\mathbf{R})\}$ est solution de :

$$\begin{aligned} \frac{\partial \{\Phi_b(\mathbf{R})\}}{\partial t}(t, x) = & \frac{1}{2} \{Q(\mathbf{R})\}(t, x) + \omega_{\mathbf{R}} \frac{\partial \{\Phi_f(\mathbf{R})\}}{\partial x}(t, x) + \frac{\partial \left[c_{\mathbf{R}} \chi \frac{\partial [\{\Phi_f(\mathbf{R})\} + \{\Phi_b(\mathbf{R})\}]}{\partial x} \right]}{\partial x}(t, x) \\ & + r_{\mathbf{R}} \{\Phi_f(\mathbf{R})\}(t, x) - f_{\mathbf{R}} \{F(\mathbf{R})\}(x) \{\Phi_b(\mathbf{R})\}(t, x) \end{aligned} \quad , \quad (4.2.4)$$

Dans les équations (4.2.1), (4.2.2), (4.2.3) et (4.2.4), les paramètres $\omega_{\mathbf{A}}$ et $\omega_{\mathbf{R}}$ correspondent à la vitesse de convection des informations relatives à l'ingestion de l'additif (A) et de la ration de base (R). Les paramètres $c_{\mathbf{A}}$ et $c_{\mathbf{R}}$ sont les coefficients de diffusion, $r_{\mathbf{A}}$ et $r_{\mathbf{R}}$, les paramètres quantifiant le retard qu'elles subissent et enfin $f_{\mathbf{A}}$ et $f_{\mathbf{R}}$ sont les paramètres qui déterminent la fixation de ces deux informations.

La fonction χ est la même fonction que celle présentée dans la section 2.1. Les fonctions d'injection, $\{Q(\mathbf{A})\}$ et $\{Q(\mathbf{R})\}$, et les fonctions de localisation, $\{F(\mathbf{A})\}$ et $\{F(\mathbf{R})\}$, sont également les mêmes que celles associées à l'information (d) et introduites dans la section 2.1 mais cette fois respectivement associée à l'additif et à la ration de base.

Comme expliqué dans la section 2.5.4, à chaque pas de temps $\Delta t = 0.16$, nous injectons dans le modèle une information relative à la quantité d'aliment consommée par les animaux. Dans ce modèle "B", à partir de la part d'additif indiquée en entrée, nous divisons cette information globale en deux afin d'avoir une quantité d'information allouée à l'additif et le reste alloué à la ration de base. La fonction $\{Q(\mathbf{A})\}$ permet d'injecter dans le modèle l'information relative à l'ingestion de l'additif et $\{Q(\mathbf{R})\}$ celle relative à l'ingestion de la ration de base. Nous avons donc :

$$\{Q(\mathbf{A})\} = p_{\mathbf{A}}(Q_{tot}) \quad (4.2.5)$$

et

$$\{Q(\mathbf{R})\} = (1 - p_{\mathbf{A}})(Q_{tot}) \quad (4.2.6)$$

où Q_{tot} est l'information relative à la consommation totale d'aliment et $p_{\mathbf{A}}$ correspond à la part d'additif dans la ration distribuée.

Deux équations décrivent ensuite la fixation de l'information relative à l'additif, $\{\Psi(A)\}$, et celle relative à la ration de base, $\{\Psi(R)\}$. $\{\Psi(A)\}$, est donc solution de :

$$\frac{\partial \{\Psi(A)\}}{\partial t}(t, x) = f_A \{F(A)\}(x) \left[\{\Phi_b(A)\}(t, x) + \{\Phi_f(A)\}(t, x) \right] - U \left(\{\Psi(A)\}, \{\Psi(R)\} \right) \{\Psi(A)\}(t, x), \quad (4.2.7)$$

et $\{\Psi(R)\}$ est solution de :

$$\frac{\partial \{\Psi(R)\}}{\partial t}(t, x) = f_R \{F(R)\}(x) \left[\{\Phi_b(R)\}(t, x) + \{\Phi_f(R)\}(t, x) \right] - U \left(\{\Psi(A)\}, \{\Psi(R)\} \right) \{\Psi(R)\}(t, x). \quad (4.2.8)$$

La fonction U est une fonction déterminant la quantité d'information utilisée en fonction de la quantité d'information relative à l'additif et à la ration de base, fixée. La construction de cette fonction sera présentée dans la section 4.2.4.

Ensuite une équation décrit l'utilisation de l'information fixée relative à la ration de base, en fonction de l'information fixée. Ainsi, Ξ est solution de :

$$\frac{\partial \Xi}{\partial t}(t, x) = U \left(\{\Psi(A)\}, \{\Psi(R)\} \right) \{\Psi(R)\}(t, x) \left(\frac{P_{max} - O(t)}{P_{max}} \right). \quad (4.2.9)$$

Le dernier terme de cette équation correspond à un limiteur permettant de modéliser une croissance logistique bornée par le paramètre P_{Max} .

Finalement, la variable de sortie est donnée par l'équation suivante :

$$O(t) = \int_{\Omega} \Xi(t, x) dx. \quad (4.2.10)$$

Le paramètre Ω correspond ici à l'intervalle $(0, 1)$.

4.2.3 Les données utilisées

Pour ajuster ce modèle "B" nous avons utilisé la base de données, associée au groupe "B" (voir Section 1.7.1). Cette base de données a été divisée en deux parties, de façon à avoir une base d'apprentissage et une base de test. La base d'apprentissage contient 13 individus : 4 individus du groupe "Témoin B", 3 individus du groupe "Dose₁ B", 3 individus du groupe "Dose₂ B" et 3 individus du groupe "Dose₃ B". Et la base de test contient 7 individus : 2 individus du groupe "Témoin B", 2 individus du groupe "Dose₁ B", 2 individus du groupe "Dose₂ B" et 1 individu du groupe "Dose₃ B".

4.2.4 Construction de l'équation d'utilisation

L'additif correspond à un ensemble de molécules permettant d'améliorer la digestibilité et l'assimilation des nutriments de la ration de base. L'additif n'agit donc pas directement sur le poids des animaux mais il influence l'impact de la ration sur celui-ci. Dans notre approche, ce phénomène peut être modélisé en faisant varier l'utilisation de l'information liée à l'ingestion de la ration de base, en fonction de la quantité d'information fixée liée à l'ingestion de l'additif. Nous avons donc construit une équation d'utilisation contenant une fonction U déterminant l'impact de la ration de base sur le poids des animaux en fonction de la dose d'additif ajoutée.

Le profil de la fonction d'utilisation, U

Nous avons déterminé l'expression mathématique de la fonction U en partant de la représentation graphique du Gain de poids Moyen Quotidien (GMQ) de chaque groupe de l'élevage "B" (Figure 4.4).

La figure 4.4 montre que jusqu'à la dose " $Dose_2$ ", le GMQ moyen augmente lorsque la dose d'additif augmente. Au-delà de cette dose, le GMQ moyen diminue lorsque la dose d'additif continue d'augmenter. L'additif utilisé améliore donc la croissance des animaux lorsqu'il est distribué jusqu'à une certaine dose (D_{Opt}) puis il devient néfaste pour celle-ci lorsqu'il est donné en trop grande quantité. À partir de ce constat nous avons fait l'hypothèse que l'utilisation de l'information relative à la ration de base devait elle aussi augmenter lorsque la dose d'additif augmente jusqu'à cette dose D_{Opt} , puis diminuer lorsque la dose d'additif distribuée continue d'augmenter. Nous avons donc fait l'hypothèse que la fonction, U , devait suivre la même tendance que le GMQ.

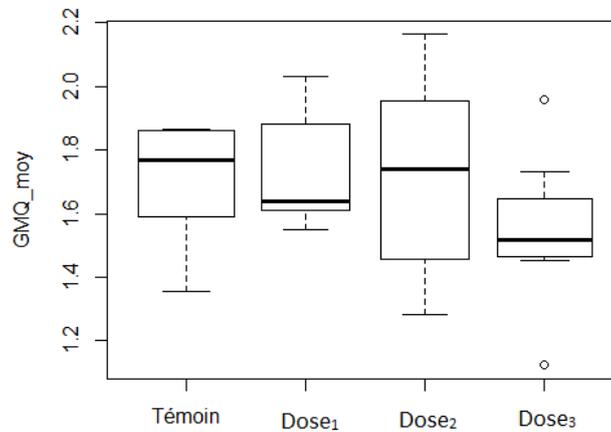


FIGURE 4.4 – Le Gain Quotidien Moyen de chaque groupe de l'élevage B.

L'équation mathématique de la fonction d'utilisation, U :

Nous avons déterminé l'équation mathématique capable de modéliser une gaussienne asymétrique dépendante de la part d'additif. En s'inspirant de la loi Log-Normale nous avons établi cette équation :

$$U(p_A) = Amp \exp(-k_1^2) [1 + (\alpha k_2 - \beta)] + B, \quad (4.2.11)$$

Où,

$$k_1 = \left(\frac{p_A - C}{\sigma_1} \right), \quad (4.2.12)$$

$$k_2 = \left(\frac{p_A - C}{\sigma_2} \right), \quad (4.2.13)$$

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2) \cdot dt, \quad (4.2.14)$$

et,

$$p_A = \frac{\Psi_A}{\Psi_A + \Psi_R}. \quad (4.2.15)$$

Cette équation de U contient 7 paramètres à ajuster : Amp , B , C , α , β , σ_1 et σ_2 . Le paramètre B est la valeur de la fonction U lorsque la part d'additif tend vers 1. Dans un premier temps nous avons fixé ce paramètre à 0.

4.2.5 Ajustement des paramètres de la fonction d'utilisation, U :

Nous sommes partis de l'hypothèse que la fonction U devait avoir la même allure que l'évolution du GMQ en fonction de la dose d'additif. Nous avons donc ajusté les paramètres de cette fonction sur les mesures de GMQ. Pour cela nous avons utilisé l'algorithme DIRECT qui nous a permis d'identifier la valeur des paramètres minimisant la fonction "objectif" donnée par l'équation :

$$f_{obj}(Amp, C, \alpha, \beta, \sigma_1, \sigma_2) = \frac{1}{4} \sum_{i=1}^4 \left| \frac{GMQ_m(p_a^i) - U(p_a^i)}{GMQ_m(p_a^i)} \right| \quad (4.2.16)$$

où p_a^i correspond à la part d'additif distribuée dans les différents groupes i , c'est-à-dire 0 pour $i = 1$, $Dose_1$ pour $i = 2$, $Dose_2$ pour $i = 3$ et $Dose_3$ pour $i = 4$. La variable $GMQ_m(p_a^i)$ correspond au GMQ moyen mesuré pour la dose d'additif p_a^i .

À la fin de l'ajustement nous avons obtenu la valeur des paramètres contenues dans la table 4.3. La figure 4.5 montre l'allure de la fonction U ajustée.

TABLE 4.3 – La valeur des paramètres de la fonction d'utilisation, U .

α	c	Amp	σ_1	σ_2	β	B
-415.53	0.0833	0.948	0.411	4.508	3.422	0

Nous avons donc obtenu une fonction U ayant l'allure globale attendue. En revanche, l'amplitude (Amp) et la valeur de U lors que la part d'additif tend vers 1 (B)

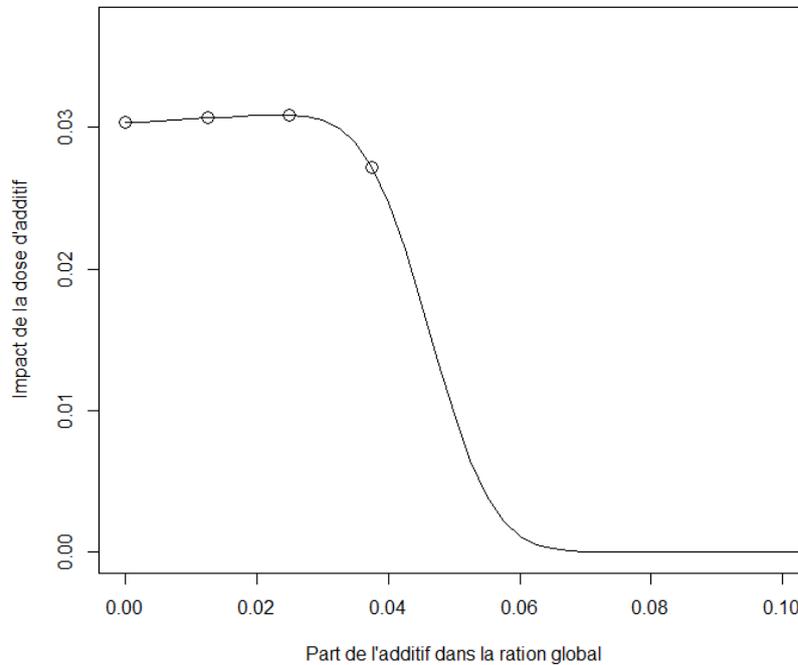


FIGURE 4.5 – Courbe de la fonction d'utilisation U (—) et la moyenne des mesures de GMQ (\circ).

doivent être ajustés sur les données de façon à avoir une fonction U permettant de relier nos données d'entrée et de sortie.

4.2.6 Ajustement des paramètres du modèle "B"

Méthode d'ajustement des paramètres

Dans la section 2.5 nous avons ajusté le modèle "A" sur les individus contenus dans le groupe "Témoin A", c'est-à-dire sur des individus qui ont été uniquement alimentés à partir de la ration de base, sans ajout d'additif. Ainsi, pour paramétrer les équations modélisant la circulation et l'évolution d'une information relative à la ration de base, nous avons repris les valeurs des paramètres obtenues lors de l'ajustement du modèle "A" sur les individus du groupe "Témoin A", soit : ω_R , r_R et f_R (Voir la table 2.5.4).

Afin de réduire le nombre de paramètres à ajuster nous avons décidé de fixer les valeurs des paramètres ω_A , c_A et r_A respectivement aux valeurs de ω_R , c_R et r_R . Nous avons ensuite ajusté les valeurs des paramètres f_A , Amp , B et P_{Max} sur la base d'apprentissage.

Pour cela nous avons utilisé l'algorithme DIRECT pour trouver les valeurs des paramètres minimisant la différence entre le poids des animaux mesuré et celui prédit, à l'instant $t = 1.52$. La fonction "objectif" est donnée par l'équation :

$$f_{obj}(f_A, Amp, B, P_{Max}) = \frac{1}{n} \sum_{i=1}^n \left| \frac{(O_{i_{obs}}(1.52) - O_{i_{pred}}(1.52))}{O_{i_{obs}}(1.52)} \right|, \quad (4.2.17)$$

où n est le nombre d'individus dans la base d'apprentissage et $O_{i_{obs}}(1.52)$ et $O_{i_{pred}}(1.52)$ correspondent respectivement à la valeur observée et à la valeur prédite de la variable de sortie pour le i^{me} individu, à l'instant $t = 1.52$.

Qualité de l'outil de prédiction basé sur le modèle "B" ajusté

A la fin de l'ajustement nous avons obtenu un quadruplet de paramètres (f_A, Amp, B, P_{Max}) (Table 4.4) et nous avons paramétré le modèle "B" avec ces valeurs. Nous avons ensuite calculé l'erreur de l'outil de prédiction basé sur le modèle "B" ajusté sur la base d'apprentissage. Pour cela nous avons calculé l'erreur relative moyenne (ARE) donnée par l'équation (2.5.4), à l'instant $t = 1.52$ (Table 4.4). Les résultats montrent que l'outil fournit des prédictions précises sur la base d'apprentissage.

Les paramètres du modèle «B» ont été ajustés à partir du poids des individus de la base d'apprentissage à $t = 1.52$, mais nous connaissons également le poids de ces individus à $t = 2.5$. Nous avons donc également calculé la valeur de l' ARE à l'instant $t = 2.5$ sur la base d'apprentissage (Table 4.4).

La valeur calculée de l'indicateur ARE à l'instant $t = 2.5$ sur la base d'apprentissage vaut 4.4%. Ce résultat est satisfaisant et met en évidence la capacité d'extrapolation du modèle "B". Néanmoins, la précision du modèle doit être calculée également sur la base de test afin de valider la qualité du modèle.

TABLE 4.4 – Valeurs des paramètres ajustés et ARE calculé à l'instant $t = 1.52$ et $t = 2.5$, sur la base d'apprentissage.

<i>Paramtre</i>	<i>Valeur</i>
Amp	0.11
B	0.26
f_A	58.44
P_{Max}	1.1
$ARE(1.52)$ (%)	2.9
$ARE(2.5)$ (%)	4.4

Nous calculons l'indicateur ARE sur la base de test aux instants $t = 1.52$ et $t = 2.5$ (Voir la table 4.5 et la figure 4.6).

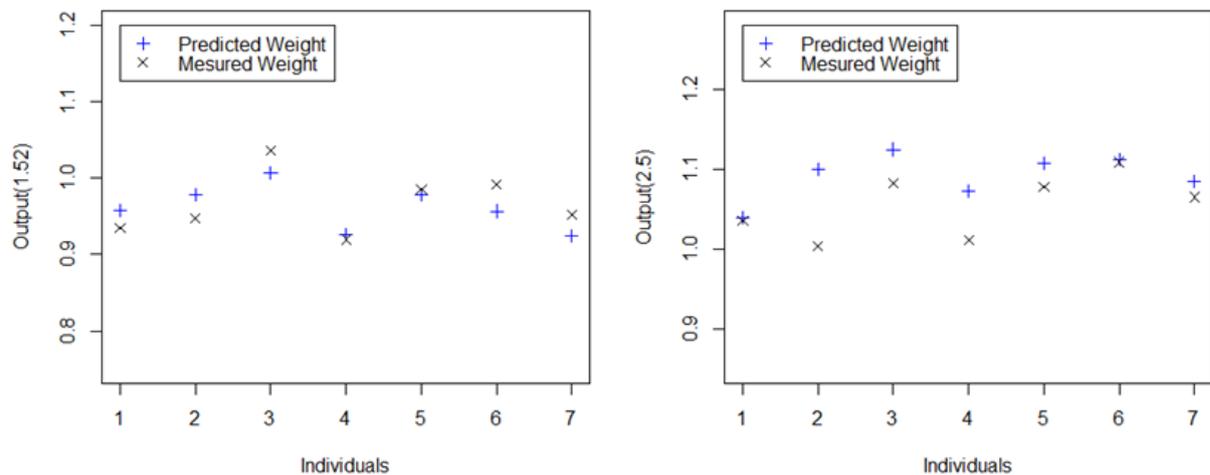


FIGURE 4.6 – Différences obtenues entre les valeurs mesurées (\times) et celles prédites ($+$) à $t = 1.52$ (Gauche) et à $t = 2.5$ (Droite) pour les individus de la base de test.

TABLE 4.5 – Erreur relative moyenne (ARE) calculée à l'instant $t = 1.52$ et $t = 2.5$, pour les individus de la base de test.

t	1.52	2.5
$ARE(t)$ (%)	2.3	3.5

L'erreur associée au modèle "B" est faible sur la base de test. Nous avons également remarqué que l'erreur faite au-delà de $t = 1.52$ est assez faible également. Ce résultat montre que, comme le modèle "A" (voir Section 2.5), le modèle "B" est capable d'apprendre une dynamique reliant les données d'entrées et de sorties sur une

certaine période, et de rester valable sur une durée plus longue. De ce fait, l'outil de prédiction basé sur le modèle "B" possède également une capacité d'extrapolation.

4.3 Construction d'un modèle simulant la circulation et l'action de deux informations en interaction : le modèle «C»

4.3.1 Objectifs

Un troisième modèle a été construit afin de modéliser encore plus finement l'effet de l'alimentation sur la croissance des individus. En effet, plusieurs additifs peuvent être ajoutés à la ration de base. De plus, l'ensemble de la ration peut être décomposée en trois parties : La quantité de protéines, de lipides et de glucides.

L'objectif a donc été de modéliser la circulation des informations relatives à l'ingestion de trois éléments de la ration globale séparément : la ration de base considérée comme un élément et deux additifs ou les protéines, les lipides et les glucides. L'objectif serait donc d'apprendre leur dynamique d'évolution respective, leurs interactions et leur action sur l'évolution du poids de l'animal.

L'intérêt de ce modèle «C» est qu'une fois la dynamique et les interactions existantes entre les différentes parties de la ration globale seront apprises, il sera possible de faire varier la part de chaque élément et de simuler l'effet de différents régimes alimentaires sur la croissance des animaux. Le but final étant de trouver le dosage optimisant les performances des animaux en termes de croissance, sans avoir à tester expérimentalement toutes les dosages possibles.

L'outil de prédiction basé sur ce modèle "C" prend en données d'entrée le poids initial des individus, la part d'additif présente dans leur ration globale et leur consommation alimentaire ($C_{0-0.67}$, $C_{0.67-1.5}$ et $C_{1.5-2.5}$) (section 1.7.1). Ce modèle fournit en données de sortie une prédiction du poids des animaux à l'instant t .

4.3.2 La structure du modèle «C»

La structure globale du modèle «C» est schématisée par la figure 4.7.

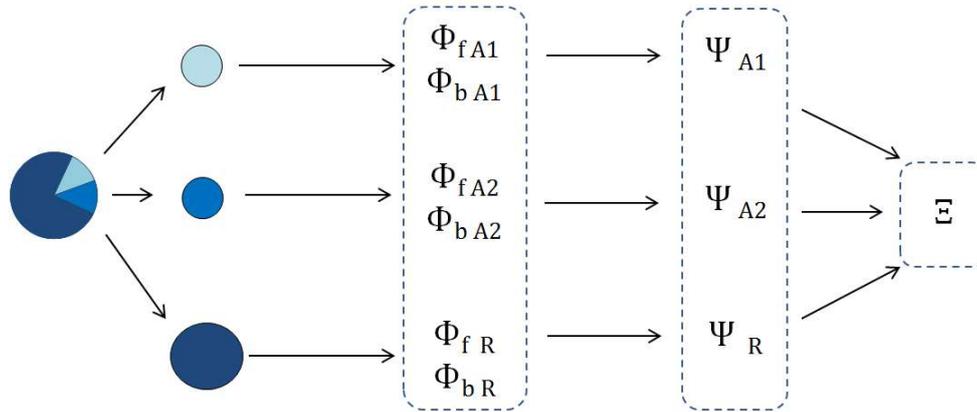


FIGURE 4.7 – Structure du modèle «C».

Par exemple, dans un modèle "C" modélisant les effets d'un régime alimentaire composé d'une ration de base et de deux additifs, nous avons deux équations modélisant le flux avant, $\{\Phi_f(A_1)\}$ et le flux arrière, $\{\Phi_b(A_1)\}$ de l'information relative au premier additif. Ainsi $\{\Phi_f(A_1)\}$ est solution de :

$$\frac{\partial \{\Phi_f(A_1)\}}{\partial t}(t, x) = \frac{1}{2} \{Q(A_1)\}(t, x) - \omega_{A_1} \frac{\partial \{\Phi_f(A_1)\}}{\partial x}(t, x) + \frac{\partial \left[c_{A_1} \chi \frac{\partial [\{\Phi_f(A_1)\} + \{\Phi_b(A_1)\}]}{\partial x} \right]}{\partial x}(t, x) - r_{A_1} \{\Phi_f(A_1)\}(t, x) - f_{A_1} \{F(A_1)\}(x) \{\Phi_f(A_1)\}(t, x), \quad (4.3.1)$$

et $\{\Phi_b(A_1)\}$ est solution de :

$$\frac{\partial \{\Phi_b(A_1)\}}{\partial t}(t, x) = \frac{1}{2} \{Q(A_1)\}(t, x) + \omega_{A_1} \frac{\partial \{\Phi_b(A_1)\}}{\partial x}(t, x) + \frac{\partial \left[c_{A_1} \chi \frac{\partial [\{\Phi_f(A_1)\} + \{\Phi_b(A_1)\}]}{\partial x} \right]}{\partial x}(t, x) + r_{A_1} \{\Phi_f(A_1)\}(t, x) - f_{A_1} \{F(A_1)\}(x) \{\Phi_b(A_1)\}(t, x), \quad (4.3.2)$$

Nous avons ensuite deux équations modélisant le flux avant, $\{\Phi_f(A_2)\}$ et le flux

arrière, $\{\Phi_b(\mathbf{A}_2)\}$ de l'information relative au deuxième additif. Ainsi $\{\Phi_f(\mathbf{A}_2)\}$ est solution de :

$$\begin{aligned} \frac{\partial \{\Phi_f(\mathbf{A}_2)\}}{\partial t}(t, x) = & \frac{1}{2}\{Q(\mathbf{A}_2)\}(t, x) - \omega_{\mathbf{A}_2} \frac{\partial \{\Phi_f(\mathbf{A}_2)\}}{\partial x}(t, x) + \frac{\partial \left[c_{\mathbf{A}_1} \chi \frac{\partial [\{\Phi_f(\mathbf{A}_2)\} + \{\Phi_b(\mathbf{A}_2)\}]}{\partial x} \right]}{\partial x}(t, x) \\ & - r_{\mathbf{A}_2} \{\Phi_f(\mathbf{A}_2)\}(t, x) - f_{\mathbf{A}_2} \{F(\mathbf{A}_2)\}(x) \{\Phi_f(\mathbf{A}_2)\}(t, x), \quad (4.3.3) \end{aligned}$$

et $\{\Phi_b(\mathbf{A}_2)\}$ est solution de :

$$\begin{aligned} \frac{\partial \{\Phi_b(\mathbf{A}_2)\}}{\partial t}(t, x) = & \frac{1}{2}\{Q(\mathbf{A}_2)\}(t, x) + \omega_{\mathbf{A}_2} \frac{\partial \{\Phi_b(\mathbf{A}_2)\}}{\partial x}(t, x) + \frac{\partial \left[c_{\mathbf{A}_1} \chi \frac{\partial [\{\Phi_f(\mathbf{A}_2)\} + \{\Phi_b(\mathbf{A}_2)\}]}{\partial x} \right]}{\partial x}(t, x) \\ & + r_{\mathbf{A}_2} \{\Phi_f(\mathbf{A}_2)\}(t, x) - f_{\mathbf{A}_2} \{F(\mathbf{A}_2)\}(x) \{\Phi_b(\mathbf{A}_2)\}(t, x), \quad (4.3.4) \end{aligned}$$

Nous avons enfin deux équations modélisant le flux avant, $\{\Phi_f(\mathbf{R})\}$ et le flux arrière, $\{\Phi_b(\mathbf{R})\}$ de l'information relative à la ration de base. Ainsi $\{\Phi_f(\mathbf{R})\}$ est solution de :

$$\begin{aligned} \frac{\partial \{\Phi_f(\mathbf{R})\}}{\partial t}(t, x) = & \frac{1}{2}\{Q(\mathbf{R})\}(t, x) - \omega_{\mathbf{R}} \frac{\partial \{\Phi_b(\mathbf{R})\}}{\partial x}(t, x) + \frac{\partial \left[c_{\mathbf{R}} \chi \frac{\partial [\{\Phi_b(\mathbf{R})\} + \{\Phi_f(\mathbf{R})\}]}{\partial x} \right]}{\partial x}(t, x) \\ & - r_{\mathbf{R}} \{\Phi_f(\mathbf{R})\}(t, x) - f_{\mathbf{R}} \{F(\mathbf{R})\}(x) \{\Phi_f(\mathbf{R})\}(t, x), \quad (4.3.5) \end{aligned}$$

et $\{\Phi_b(\mathbf{R})\}$ est solution de :

$$\begin{aligned} \frac{\partial \{\Phi_b(\mathbf{R})\}}{\partial t}(t, x) = & \frac{1}{2}\{Q(\mathbf{R})\}(t, x) + \omega_{\mathbf{R}} \frac{\partial \{\Phi_b(\mathbf{R})\}}{\partial x}(t, x) + \frac{\partial \left[c_{\mathbf{R}} \chi \frac{\partial [\{\Phi_b(\mathbf{R})\} + \{\Phi_f(\mathbf{R})\}]}{\partial x} \right]}{\partial x}(t, x) \\ & + r_{\mathbf{R}} \{\Phi_f(\mathbf{R})\}(t, x) - f_{\mathbf{R}} \{F(\mathbf{R})\}(x) \{\Phi_b(\mathbf{R})\}(t, x), \quad (4.3.6) \end{aligned}$$

Dans les équations (4.3.1), (4.3.2), (4.3.3), (4.3.4), (4.3.5) et (4.3.6), les paramètres $\omega_{\mathbf{R}}$, $\omega_{\mathbf{A}_2}$ et $\omega_{\mathbf{A}_1}$ correspondent à la vitesse de convection des informations relatives à

l'ingestion de la ration de base (R) et des deux additifs (A_1 et A_2). Les paramètres c_R , c_{A_1} et c_{A_2} , sont les coefficients de diffusion, r_R , r_{A_1} et r_{A_2} , les paramètres quantifiant le retard qu'elles subissent et enfin f_R , f_{A_1} et f_{A_2} sont les paramètres qui déterminent la fixation de ces trois informations.

La fonction χ est la même fonction que celle présentée dans la section 2.1. Les fonctions d'injection, $\{Q(A_1)\}$, $\{Q(A_2)\}$ et $\{Q(R)\}$, et les fonctions de localisation, $\{F(A_1)\}$, $\{F(A_2)\}$ et $\{F(R)\}$, sont également les mêmes que celles associées à l'information (d) et introduites dans la section 2.1 mais cette fois respectivement associée aux deux additifs et à la ration de base.

Trois équations modélisent ensuite la fixation de l'information relative au premier additif, $\{\Psi(A_1)\}$, au deuxième additif, $\{\Psi(A_2)\}$, et celle relative à la ration de base, $\{\Psi(R)\}$.

$\{\Psi(A_1)\}$, est solution de :

$$\frac{\partial \{\Psi(A_1)\}}{\partial t}(t, x) = f_{A_1} \{F(A_1)\}(x) \left[\{\Phi_b(A_1)\}(t, x) + \{\Phi_f(A_1)\}(t, x) \right] - u_{A_1} \{\Psi(A_1)\}(t, x), \quad (4.3.7)$$

$\{\Psi(A_2)\}$, est solution de :

$$\frac{\partial \{\Psi(A_2)\}}{\partial t}(t, x) = f_{A_2} \{F(A_2)\}(x) \left[\{\Phi_b(A_2)\}(t, x) + \{\Phi_f(A_2)\}(t, x) \right] - u_{A_2} \{\Psi(A_2)\}(t, x), \quad (4.3.8)$$

et $\{\Psi(R)\}$ est solution de :

$$\frac{\partial \{\Psi(R)\}}{\partial t}(t, x) = f_R \{F(R)\}(x) \left[\{\Phi_b(R)\}(t, x) + \{\Phi_f(R)\}(t, x) \right] - u_R \{\Psi(R)\}(t, x). \quad (4.3.9)$$

Dans les équations (4.3.7), (4.3.8) et (4.3.9) les paramètres u_{A_1} , u_{A_2} et u_R correspondent aux paramètres quantifiant l'utilisation des informations relatives à l'ingestion

des deux additifs et de la ration de base.

Ensuite une équation décrit l'utilisation de l'information fixée relative à la ration de base, en fonction de l'information fixée relative aux deux additifs. Ainsi, Ξ est solution de :

$$\frac{\partial \Xi}{\partial t}(t, x) = U_C\left(\{\Psi(A_1)\}, \{\Psi(A_2)\}, \{\Psi(R)\}\right) \{\Psi(R)\}(t, x) \left(\frac{P_{max} - s(t)}{P_{max}}\right), \quad (4.3.10)$$

Dans l'équation (4.3.10), la fonction U_C est une fonction d'utilisation dépendante de $\{\Psi(A_1)\}$, $\{\Psi(A_2)\}$ et $\{\Psi(R)\}$ et modélisant la potentielle interaction existante entre ces trois densités.

Finalement, la variable de sortie est donnée par l'équation suivante :

$$O(t) = \int_{\Omega} \Xi(t, x) dx. \quad (4.3.11)$$

4.3.3 Perspective concernant l'utilisation du modèle «C»

Le modèle "C" est donc un modèle permettant de simuler l'impact de trois éléments de la ration globale. Mais il permet surtout d'apprendre une dynamique reliant ces trois éléments et l'évolution d'une variable biologique qui peut être par exemple le poids des animaux.

Nous n'avons pas pour l'instant de données permettant de tester les qualités de ce modèle. Néanmoins les applications des modèles "A" et "B" sont encourageantes et nous laissent penser que ce modèle "C" pourra lui aussi être ajusté sur une certaine gamme de données et d'être appliqué en dehors de cette gamme. En effet, comme nous l'avons vu dans l'introduction de cette section, une fois que les paramètres associés aux trois éléments seront appris, nous pourrons simuler l'effet de différents dosages de ces trois constituants sur la variable de sortie. Ainsi, il sera possible de déterminer le dosage optimisant les performances des animaux sans tester expérimentalement tous les dosages possibles.

Comme dans la section 4.2, la fonction U_c déterminant l'utilisation de l'information pourra sans doute être établie en étudiant l'évolution du GMQ en fonction de la part des trois éléments étudiés.

4.4 Conclusion du chapitre 4

Dans ce chapitre nous avons étudié la façon dont certaines informations, comme l'âge des animaux et la composition de la ration alimentaire, peuvent être intégrées dans un modèle biomimétique d'apprentissage statistique. Pour cela nous avons construit un modèle " A_{Age} " pouvant être considéré comme une cinquième version du modèle "A" et deux autres modèles.

Davantage de données seraient nécessaires pour réellement éprouver ces outils, notamment le modèle "C". Toutefois, les applications des modèles " A_{Age} " et B sur des données réelles ont mis en évidence la capacité d'extrapolation des outils de prédiction basés sur ces modèles. En effet, tout comme le modèle "A" étudié dans le chapitre 2, ces modèles apprennent une dynamique reliant les données d'entrée et de sortie. Cette dynamique peut ensuite être appliquée en dehors de la gamme de données ayant servi pour ajuster les paramètres des modèles.

Cette capacité d'extrapolation pourrait permettre de simuler l'impact de différents régimes alimentaire en termes de dosage de deux ou trois éléments. Cela pourrait permettre de trouver ou en tout cas d'affiner la recherche du dosage optimisant les performances des animaux et les revenus des éleveurs, sans avoir à tester tous les dosages possibles. Cette capacité d'extrapolation des modèles "B" et "C" pourrait donc permettre de réduire fortement les coûts d'expérimentation et de gestion de données.

POSITIONNEMENT ET DOMAINE D'APPLICATION DES OUTILS D'APPRENTISSAGE STATISTIQUE BIOMIMÉTIQUES

5.1 Comportement des modèles biomimétiques en fonction de leur structure et des caractéristiques des données d'apprentissage

Dans cette section nous allons présenter plusieurs études nous ayant permis d'analyser le comportement des modèles biomimétiques, notamment en ce qui concerne le temps d'apprentissage de ces modèles.

Dans un premier temps, nous avons étudié le gain de temps d'exécution apporté par la parallélisation des calculs. En effet, pour pouvoir utiliser les 4 processeurs de l'ordinateur nous avons dû paralléliser les calculs réalisés au cours de l'ajustement des paramètres.

Dans le cadre d'une autre étude nous avons étudié l'évolution du temps d'apprentissage en fonction du nombre de flux d'information modélisés par le modèle biomimétique. Pour cela nous avons comparé le temps d'apprentissage des modèles "A", "B" et "C". Dans cette section nous avons également étudié l'évolution du temps de calcul en fonction du nombre d'individus contenus dans la base d'apprentissage.

Dans une troisième étude nous avons étudié l'évolution du nombre d'itérations nécessaire avant d'atteindre la convergence des paramètres en fonction du nombre de paramètres à déterminer et du nombre d'individus présents dans la base d'apprentis-

sage. Pour cela nous avons utilisé le modèle "A" dans lequel nous n'avons pas pris en compte les relations existantes entre les paramètres établies dans la section 2.3.2 (Modèle « A_{SsRel} »).

Enfin, nous avons étudié l'impact de la prise en compte des relations existantes entre les paramètres sur le temps d'apprentissage des modèles. Dans cette étude nous avons estimé le nombre d'itérations nécessaire avant d'atteindre la convergence des paramètres en utilisant toujours le modèle "A" mais cette fois en prenant en compte les relations établies entre les paramètres dans la section 2.3.2 (Modèle " A_{Rel} ").

5.1.1 Gain de temps d'exécution apporté par la parallélisation des calculs

Lors des premiers ajustements de modèles réalisés avec le logiciel R, nous avons monitoré les processeurs de l'ordinateur utilisé pour ajuster nos modèles. Nous avons obtenu les graphiques de la figure 5.1. Ces graphiques nous ont montré qu'un seul processeur était mobilisé pour effectuer les calculs alors que sur cet ordinateur, 3 autres processeurs restaient inutilisés. Pour pouvoir utiliser les 4 processeurs nous avons dû paralléliser les calculs réalisés au cours de l'ajustement des paramètres.

Initialement, les différents calculs étaient exécutés les uns après les autres. Or, lors de chaque itération réalisée par l'algorithme d'optimisation certains calculs peuvent être faits en parallèle. En effet, au cours d'un ajustement, l'objectif est de minimiser l'écart existant entre des valeurs mesurées et des valeurs prédites par le modèle, pour l'ensemble des individus de la base d'apprentissage. Dans un premier temps, à chaque itération, les courbes de croissance étaient simulées par le modèle pour tous les individus, les uns après les autres. Afin de réduire le temps de calcul, nous avons fait en sorte de simuler les courbes de croissance des différents individus de la base d'apprentissage en parallèle. Pour cela nous avons utilisé la fonction R *apply* (R CORE TEAM (2014)). Cette fonction applique en parallèle le modèle à ajuster à chaque individu de la base d'apprentissage à partir de ses données d'entrée.

Une fois cette parallélisation effectuée nous avons monitoré de nouveau les processeurs de l'ordinateur utilisé pour ajuster les modèles. Nous avons cette fois obtenu la figure 5.2.

Cette figure montre que désormais, l'ensemble des processeurs sont mobilisés lors

5.1. Comportement des modèles biomimétiques en fonction de leur structure et des caractéristiques des données d'apprentissage



FIGURE 5.1 – Monitoring du fonctionnement des processus lors de l’ajustement d’un modèle biomimétique avant la parallélisation des calculs

de l’apprentissage des paramètres.

Le calcul du temps nécessaire pour effectuer 10 itérations a montré que cette parallélisation des calculs a permis de diviser par 20 le temps d’apprentissage des paramètres du modèle.

Néanmoins, le temps d’apprentissage des modèles biomimétiques construits peut encore représenter une limite à leur utilisation. Nous avons donc réalisé une étude

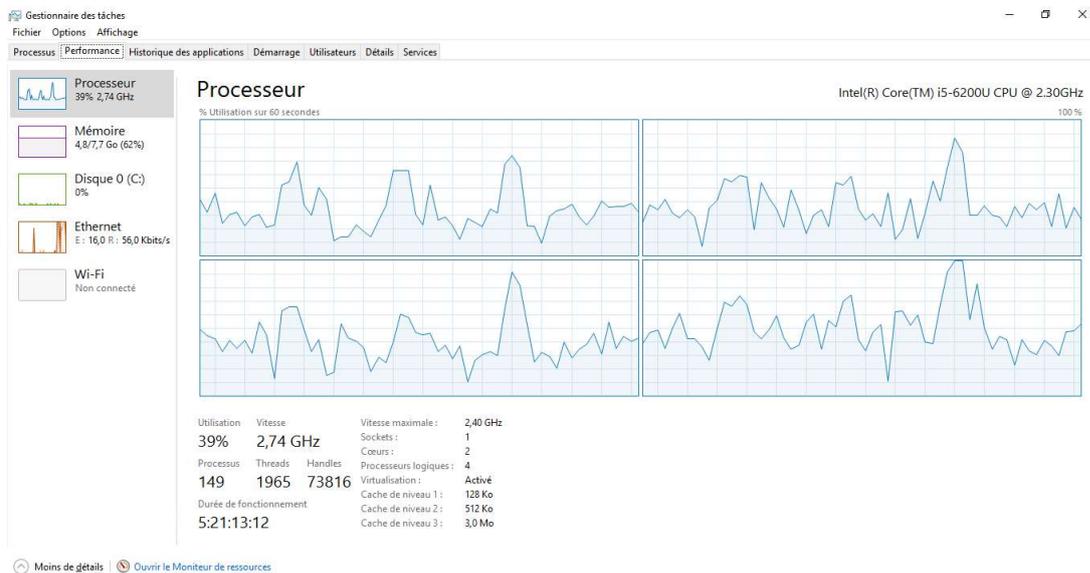


FIGURE 5.2 – Monitoring du fonctionnement des processeurs lors de l’ajustement d’un modèle biomimétique après la parallélisation des calculs

du temps d’apprentissage nécessaire en fonction de différentes variables comme le nombre de flux d’information modélisés, le nombre d’individus contenus dans la base d’apprentissage, ou encore le nombre de paramètres à ajuster.

5.1.2 Étude de l’évolution du temps d’apprentissage en fonction du nombre de flux d’information modélisés par le modèle biomimétique

Etude du temps d’apprentissage des modèles "A", "B" et "C"

Au cours d’un processus d’apprentissage, un certain nombre de valeurs de chaque paramètre doit être testé avant que l’algorithme DIRECT trouve les valeurs des paramètres minimisant la fonction "objectif" sur les données d’apprentissage.

Dans cette section nous avons donc étudié l’évolution du temps de calcul en fonction du nombre de flux d’information modélisés. Nous avons mesuré le temps nécessaire pour réaliser une itération avec le modèle "A" modélisant un flux d’information (Section 2.1), avec le modèle "B" modélisant deux flux d’informations (Section 4.2) et enfin avec le modèle "C" modélisant trois flux d’informations (Section 4.3).

5.1. Comportement des modèles biomimétiques en fonction de leur structure et des caractéristiques des données d'apprentissage

Dans cette section nous avons également étudié l'évolution du temps de calcul en fonction du nombre d'individus contenus dans la base d'apprentissage.

Pour cela nous avons mesuré le temps nécessaire pour que chaque modèle réalise 10 itérations et nous avons divisé le temps obtenu par 10. Ces itérations ont été réalisées en utilisant l'algorithme DIRECT, sur 4 bases de données contenant respectivement 2, 5, 10 et 15 individus. Nous avons obtenu la table 5.1.

TABLE 5.1 – Évolution du temps de calcul nécessaire pour réaliser une itération (min) en fonction du modèle biomimétique utilisé et du nombre d'individus présents dans la base d'apprentissage

Modèle A	Modèle B	Modèle C	Nombre d'individus
0.16	0.57	0.78	2
0.37	1.42	2.00	5
0.75	2.85	3.89	10
1.12	4.30	5.71	15

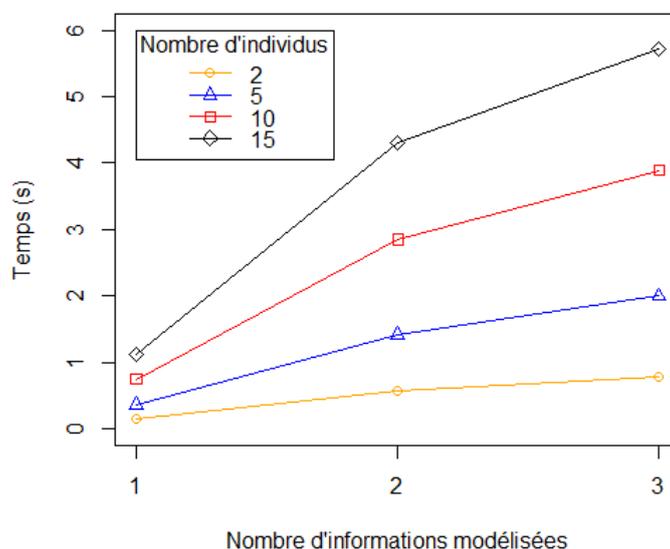


FIGURE 5.3 – Courbe d'évolution du temps de calcul (min) en fonction du modèle biomimétique utilisé et du nombre d'individus présents dans la base d'apprentissage

Le temps de calcul augmente de façon exponentielle lorsque le nombre d'individus

contenus dans la base d'apprentissage augmente. Le temps d'exécution augmente également lorsque le nombre de flux d'informations modélisés augmente mais pas de manière exponentielle (5.3). Cela peut s'expliquer par la structure des modèles «B» et «C». En effet, dans ces modèles nous avons modélisé une interaction entre les différentes informations en terme d'impact sur la variable de sortie (Figures 4.3 et 4.7). Ainsi, même si le nombre d'informations intégrées augmente, il n'y a toujours dans le modèle qu'une équation "d'utilisation". Le nombre d'équations contenues dans le modèle, n'est donc pas proportionnel au nombre de flux d'informations modélisés.

Étude du temps de calcul de modèles n'intégrant pas d'interactions entre les flux d'informations intégrés : «A», « $B_{SsInter}$ » et « $C_{SsInter}$ »

L'objectif de cette étude est d'analyser l'évolution du temps de calcul de modèles ne modélisant pas d'interactions entre les informations intégrées. Le but de cette étude est de mettre en évidence l'impact de ces interactions sur le temps de calcul.

Dans cette étude, le modèle, qui ne modélise qu'un flux d'information reste inchangé, mais nous avons construit un modèle « $B_{SsInter}$ » (figure 5.4), dans lequel $\{\Phi_f(A)\}$ et $\{\Phi_b(A)\}$ sont respectivement solutions des équations (4.2.1) et (4.2.2), $\{\Phi_f(R)\}$ et $\{\Phi_b(R)\}$ sont respectivement solutions des équations (4.2.3) et (4.2.4), les densités $\{\Psi(A)\}$ et $\{\Psi(R)\}$ sont solutions des équations (4.2.7) et (4.2.8), $\Xi(A)$ est solution de

$$\frac{\partial \Xi(A)}{\partial t}(t, x) = u_A \{\Psi(A)\}(t, x) \quad (5.1.1)$$

et $\Xi(R)$ est solution de :

$$\frac{\partial \Xi(R)}{\partial t}(t, x) = u_R \{\Psi(R)\}(t, x) \quad (5.1.2)$$

et finalement, la variable de sortie est donnée par l'équation suivante :

$$O(t) = \int_{\Omega} (\{\Xi(A)\}(t, x) + \{\Xi(R)\}(t, x)) dx. \quad (5.1.3)$$

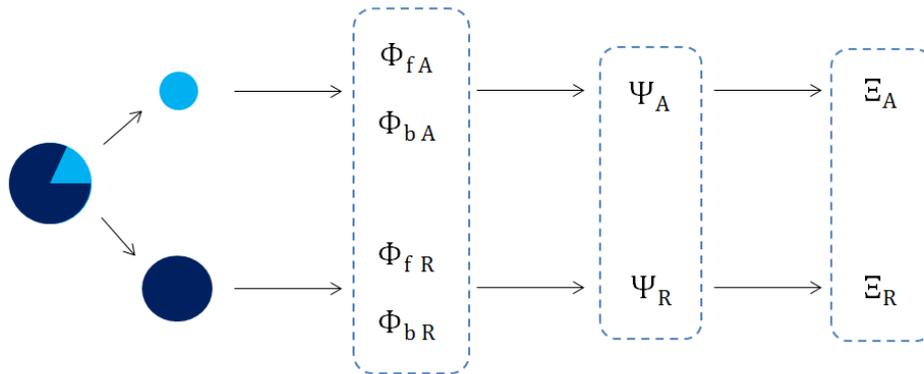


FIGURE 5.4 – Schématisation de la structure globale du modèle « $B_{SsInter}$ », sans interaction modélisée entre les informations

Pour cette étude nous avons également construit un modèle « $C_{SsInter}$ » (figure 5.5), dans lequel $\{\Phi_f(A_1)\}$ et $\{\Phi_b(A_1)\}$ sont respectivement solutions des équations (4.3.1) et (4.3.2), $\{\Phi_f(A_2)\}$ et $\{\Phi_b(A_2)\}$ sont respectivement solutions des équations (4.3.3) et (4.3.4), $\{\Phi_f(R)\}$ et $\{\Phi_b(R)\}$ sont respectivement solutions des équations (4.3.5) et (4.3.6), les densités $\{\Psi(A_1)\}$, $\{\Psi(A_2)\}$ et $\{\Psi(A)\}$ sont solutions de (4.3.7), (4.3.8) et (4.3.9). La fonction $\{\Xi(A_1)\}$ est solution de,

$$\frac{\partial \{\Xi(A_1)\}}{\partial t}(t, x) = u_{A_1} \{\Psi(A_1)\}(t, x), \quad (5.1.4)$$

la fonction $\{\Xi(A_2)\}$ est solution de,

$$\frac{\partial \{\Xi(A_2)\}}{\partial t}(t, x) = u_{A_2} \{\Psi(A_2)\}(t, x), \quad (5.1.5)$$

la fonction $\{\Xi(R)\}$ est solution de,

$$\frac{\partial \{\Xi(R)\}}{\partial t}(t, x) = u_R \{\Psi(R)\}(t, x), \quad (5.1.6)$$

et finalement, la variable de sortie est donnée par l'équation suivante :

$$O(t) = \int_{\Omega} (\{\Xi(A_1)\}(t, x) + \{\Xi(A_2)\}(t, x) + \{\Xi(R)\}(t, x)) dx. \quad (5.1.7)$$

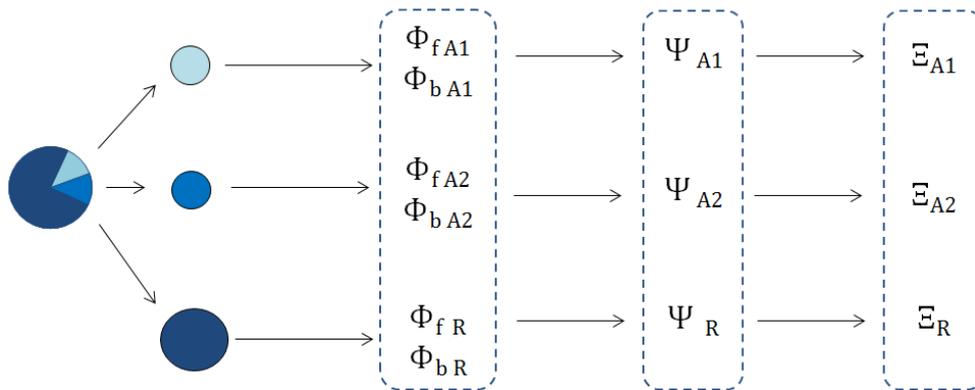


FIGURE 5.5 – Schématisation de la structure globale du modèle « $C_{SsInter}$ », sans interaction modélisée entre les informations circulant

Ainsi dans les modèles « A », « $B_{SsInter}$ » et « $C_{SsInter}$ » le nombre d'équations est proportionnel au nombre d'informations intégrées dans le modèle.

De la même manière que précédemment nous avons étudié l'évolution du temps de calcul de ces 3 modèles. Nous avons obtenu la table 5.2.

TABLE 5.2 – Évolution du temps de calcul nécessaire pour réaliser une itération (min) en fonction du modèle biomimétique utilisé et du nombre d'individus présents dans la base d'apprentissage

Modèle A	Modèle $B_{SsInter}$	Modèle $C_{SsInter}$	Nombre d'individus
0.16	0.34	0.80	2
0.37	0.85	1.94	5
0.75	1.55	3.81	10
1.12	2.43	5.53	15

Comme précédemment, le temps de calcul augmente de manière exponentielle lorsque le nombre d'individus contenus dans la base d'apprentissage augmente. Mais nous remarquons cette fois que lorsque le nombre de flux d'information augmente, le

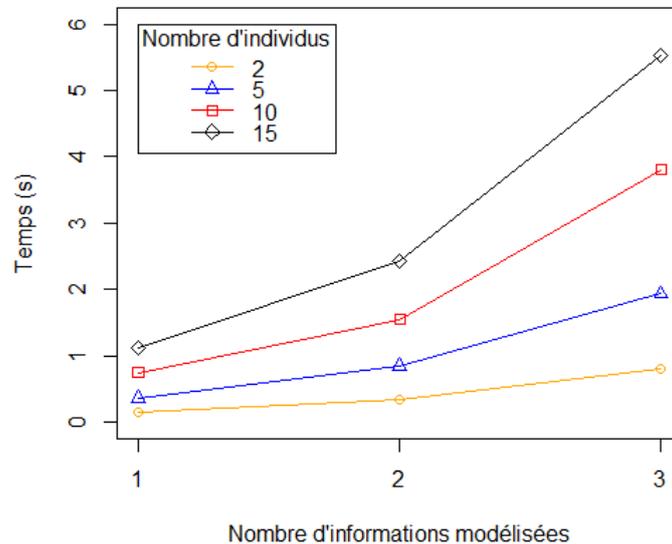


FIGURE 5.6 – Courbe d'évolution du temps de calcul nécessaire pour réaliser une itération (min) en fonction du modèle biomimétique utilisé et du nombre d'individus présents dans la base d'apprentissage

temps d'exécution augmente de manière exponentielle (figure 5.6).

Le nombre de paramètres à déterminer augmente lorsque le nombre de flux d'information modélisés augmente. Le nombre d'itérations nécessaire avant d'atteindre la convergence des paramètres augmente donc également. Une étude a donc été réalisée afin de déterminer l'évolution du nombre d'itérations nécessaire avant d'atteindre la convergence des paramètres en fonction du nombre de paramètres à déterminer et du nombre d'individus présents dans la base d'apprentissage.

5.1.3 Évolution du nombre d'itérations nécessaire avant d'atteindre la convergence des paramètres en fonction du nombre de paramètres à déterminer et du nombre d'individus présents dans la base d'apprentissage

Dans cette étude nous avons estimé le nombre d'itérations nécessaire avant d'atteindre la convergence des paramètres. Pour cela nous avons utilisé le modèle "A" dans lequel nous n'avons pas pris en compte les relations existantes entre les paramètres établies dans la section 2.3.2 (Modèle « A_{SsRel} »).

Nous avons ajusté ce modèle « A_{SsRel} » sur deux bases de données d'apprentissage contenant respectivement 5 et 10 individus.

L'algorithme DIRECT a été implémenté de façon à ce que le processus d'apprentissage s'arrête lorsque la fonction «objectif» atteint un seuil de convergence donné, et noté ϵ . Le processus d'apprentissage s'arrête donc lorsque,

$$\frac{|f_{objI_{k+1}} - f_{objI_k}|}{f_{objI_k}} \leq \epsilon, \quad (5.1.8)$$

où f_{objI_k} correspond à la valeur de la fonction «objectif» obtenue à la $k^{ième}$ itération et ϵ , le critère d'arrêt de l'algorithme. Le nombre d'itérations retenu est alors $k + 1$.

Dans cette étude la fonction "objectif" est donnée par l'équation,

$$f_{obj}(\omega_a, r_a, f_a, u_a, L_a) = \frac{1}{n} \sum_{i=1}^n \left(\frac{(O_{i_{obs}}(1) - O_{i_{pred}}(1))}{O_{i_{obs}}(1)} \right)^2, \quad (5.1.9)$$

où n est le nombre d'individus de la base d'apprentissage et $O_{i_{obs}}(1)$ et $O_{i_{pred}}(1)$ correspondent respectivement au poids mesuré et au poids prédit à $t = 1$, pour le $i^{ième}$ individu.

Nous avons testé différentes valeurs du critère d'arrêt ϵ : 10^{-4} , 10^{-5} , 10^{-6} , 10^{-7} et 10^{-8} .

Le modèle « A_{SsRel} » contient 5 paramètres pouvant être appris. Nous avons donc cherché la valeur de $(k + 1)$ lors de l'ajustement d'un seul paramètres (ω_a), de 2 (ω_a, r_a), de

5.1. Comportement des modèles biomimétiques en fonction de leur structure et des caractéristiques des données d'apprentissage

3 (ω_d, r_d, f_d), de 4 (ω_d, r_d, f_d, u_d) et de 5 paramètres ($\omega_d, r_d, f_d, u_d, Pmax$).

Nous avons réalisé au total 50 ajustements du modèle « A_{SsRel} » sur les données réelles présentée dans la section 1.7.1. Á la fin de ces 50 ajustements nous avons obtenu la table 5.3.

TABLE 5.3 – Évolution du nombre d'itérations nécessaire avant d'atteindre la convergence des paramètres en fonction du nombre de paramètres à déterminer (param.), de la valeur du critère d'arrêt (ϵ) et du nombre d'individus (NI) présents dans la base d'apprentissage.

1 param. (ω_d)	2 param. (ω_d, r_d)	3 param. (ω_d, r_d, f_d)	4 param. (ω_d, r_d, f_d, u_d)	5 param. ($\omega_d, r_d, f_d, u_d, L_d$)	ϵ	NI
187	267	5000	1999	1275	10^{-8}	5
155	177	271	429	717	10^{-7}	5
155	145	143	211	167	10^{-6}	5
51	105	71	211	167	10^{-5}	5
21	59	71	107	85	10^{-4}	5
163	225	205	385	499	10^{-8}	10
129	225	167	257	181	10^{-7}	10
57	225	143	115	63	10^{-6}	10
39	27	99	115	55	10^{-5}	10
15	27	99	45	55	10^{-4}	10

Analyse de l'évolution du nombre d'itérations effectuées en fonction de la valeur du critère d'arrêt et du nombre de paramètres à ajuster

Afin de visualiser les résultats obtenus dans la table 5.3 nous avons représenté le nombre d'itérations effectuées en fonction de la valeur du critère d'arrêt, ϵ (Figure 5.10). Sur cette figure, le nombre d'itérations effectuées pour ajuster le modèle « A_{SsRel} » augmente de façon exponentielle lorsque la valeur du critère d'arrêt, ϵ diminue.

Afin de mieux visualiser l'impact du nombre de paramètres sur ce nombre d'itérations nous avons représenté le nombre d'itérations effectuées en fonction du nombre de paramètres ajustés (Figure 5.11). Sur cette figure, le nombre d'itérations nécessaire avant d'atteindre la convergence des paramètres augmente lorsque le nombre de paramètres augmente, mais cette augmentation n'est pas exponentielle. Cela s'explique sans doute par l'existence de relations entre certains paramètres du modèle (Voir la section 2.3.2).

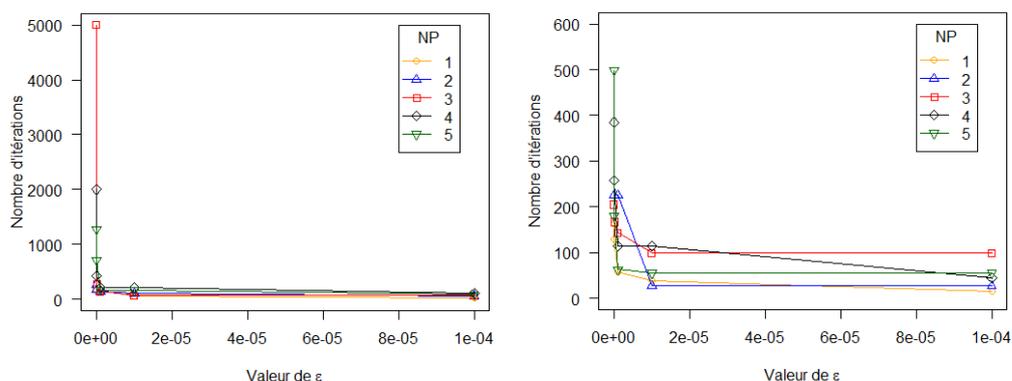


FIGURE 5.7 – Le nombre d'itérations effectuées en fonction de la valeur du critère d'arrêt, en utilisant une base d'apprentissage contenant 5 individus (Gauche) et 10 individus (Droite).

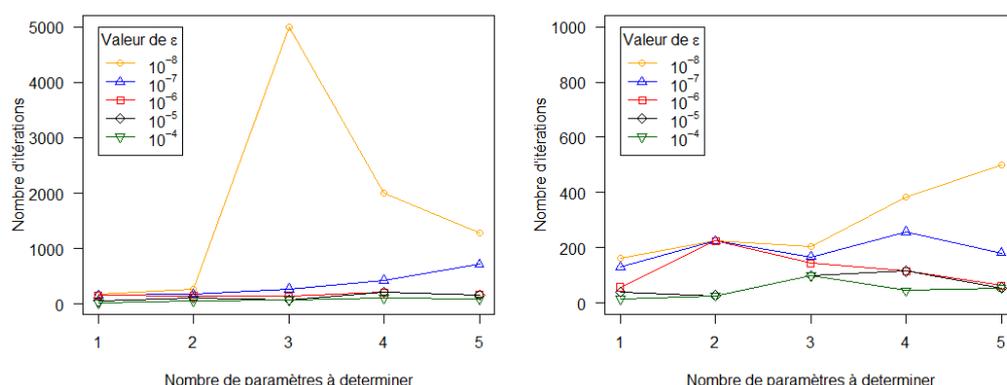


FIGURE 5.8 – Le nombre d'itérations effectuées en fonction du nombre de paramètres à déterminer, en utilisant une base d'apprentissage contenant 5 individus (Gauche) et 10 individus (Droite).

Analyse de l'évolution du nombre d'itérations effectuées en fonction du nombre d'individus présents dans la base d'apprentissage

Dans cette section nous allons étudier l'évolution du nombre d'itérations nécessaire avant d'atteindre la convergence des paramètres lorsque le nombre d'individus dans la base d'apprentissage passe de 5 à 10 individus, et ce en fonction de la valeur du critère d'arrêt. Pour cela nous avons construit la table 5.4.

La table 5.4 nous montre que lorsque le nombre d'individus présents dans la base d'apprentissage augmente, cela permet de réduire le nombre d'itérations nécessaire avant d'atteindre la convergence des paramètres. L'apprentissage de la valeur des

TABLE 5.4 – Gain apporté en termes de nombre d'itérations nécessaire avant d'atteindre la convergence des paramètres lorsque le nombre d'individus dans la base d'apprentissage passe de 5 à 10 individus, en fonction de la valeur du critère d'arrêt et du nombre de paramètres appris.

1 param. (ω_d)	2 param. (ω_d, r_d)	3 param. (ω_d, r_d, f_d)	4 param. (ω_d, r_d, f_d, u_d)	5 param. ($\omega_d, r_d, f_d, u_d, L_d$)	Gain Moyen	ϵ
24	42	4795	1614	776	1450.2	10^{-8}
26	-48	104	172	536	158	10^{-7}
98	-80	0	96	104	43.6	10^{-6}
12	78	-28	96	112	54	10^{-5}
6	32	-28	62	30	20.4	10^{-4}

paramètres se fait donc plus rapidement lorsque la taille de la base d'apprentissage augmente.

Pour tenter d'expliquer ce résultat nous avons utilisé les modèles obtenus à la fin de l'apprentissage pour calculer l'erreur relative associée à chaque individu des bases d'apprentissage contenant 5 et 10 individus. Pour cela nous avons calculé l'*ARE* à l'instant $t = 1$ pour chaque individu :

$$ARE(t) = \frac{1}{n} \sum_{i=1}^n \sqrt{\left(\frac{(O_{i_{obs}}(t) - O_{i_{pred}}(t))}{O_{i_{obs}}(t)} \right)^2} \quad (5.1.10)$$

où n est le nombre d'individus dans la base de données et $O_{i_{obs}}(t)$ et $O_{i_{pred}}(t)$ correspondent respectivement aux poids mesuré et prédit à l'instant t pour l'individu i .

Nous avons ensuite calculé la valeur moyenne et l'écart type de cet indicateur d'erreur. L'objectif est de mesurer l'hétérogénéité des individus de ces deux bases de données en termes de croissance. Nous avons obtenu la figure 5.9. Nous avons constaté sur cette figure que l'indicateur d'erreur *ARE* présente une moyenne et une variance plus élevées sur la base de données contenant 10 individus. L'augmentation du nombre d'individus a donc entraîné une augmentation de la variabilité. Cette plus grande variabilité semble permettre une convergence des paramètres plus rapide.

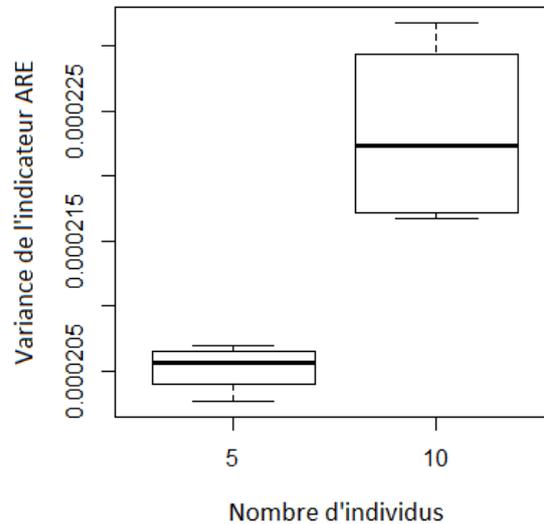


FIGURE 5.9 – Moyenne et variance de l'indicateur ARE en fonction du nombre d'individus dans la base de données d'apprentissage.

5.1.4 Étude de l'impact de la prise en compte des relations existantes entre les paramètres sur le temps d'apprentissage des modèles

Dans cette nouvelle étude nous avons estimé le nombre d'itérations nécessaire avant d'atteindre la convergence des paramètres en utilisant toujours le modèle "A" mais cette fois en prenant en compte les relations établies entre les paramètres dans la section 2.3.2 (Modèle " A_{Rel} ").

Comme dans la section 5.1.3, nous avons ajusté ce modèle " A_{Rel} " contenant jusqu'à 5 paramètres à ajuster, sur deux bases de données d'apprentissage, contenant respectivement 5 et 10 individus. L'algorithme DIRECT a été implémenté de façon à ce que le processus d'apprentissage s'arrête lorsque la fonction "objectif" a atteint un seuil de convergence, ϵ . Comme dans l'étude précédente nous avons étudié l'évolution du nombre d'itérations nécessaire avant d'atteindre la convergence des paramètres pour différentes valeurs du critère d'arrêt ϵ : 10^{-4} , 10^{-5} , 10^{-6} , 10^{-7} et 10^{-8} .

La fonction "objectif" est donnée par l'équation (5.1.10).

5.1. Comportement des modèles biomimétiques en fonction de leur structure et des caractéristiques des données d'apprentissage

Nous avons donc réalisé au total 50 ajustements du modèle " A_{Rel} " sur les données réelles. À la fin de ces 50 ajustements nous avons obtenu la table 5.5.

TABLE 5.5 – Évolution du nombre d'itérations en fonction du nombre de paramètres à déterminer, de la valeur du critère d'arrêt et du nombre d'individus (NI) présents dans la base d'apprentissage en utilisant le modèle " A_{Rel} ".

1 param. (ω_d)	2 param. (ω_d, f_d)	3 param. (ω_d, f_d, r_d)	4 param. (ω_d, f_d, r_d, u_d)	5 param. ($\omega_d, f_d, r_d, u_d, L_d$)	ϵ	NI
537	243	787	685	1000	10^{-8}	5
81	243	209	189	717	10^{-7}	5
51	93	181	101	167	10^{-6}	5
29	25	111	73	167	10^{-5}	5
5	23	49	47	85	10^{-4}	5
395	365	125	517	499	10^{-8}	10
29	129	83	271	181	10^{-7}	10
29	23	49	169	63	10^{-6}	10
9	15	25	65	55	10^{-5}	10
5	11	19	65	55	10^{-4}	10

Afin de visualiser les résultats obtenus dans ce tableau nous avons représenté le nombre d'itérations effectuées en fonction de la valeur du critère d'arrêt, ϵ et du nombre de paramètres à apprendre (figures 5.10 et 5.11).

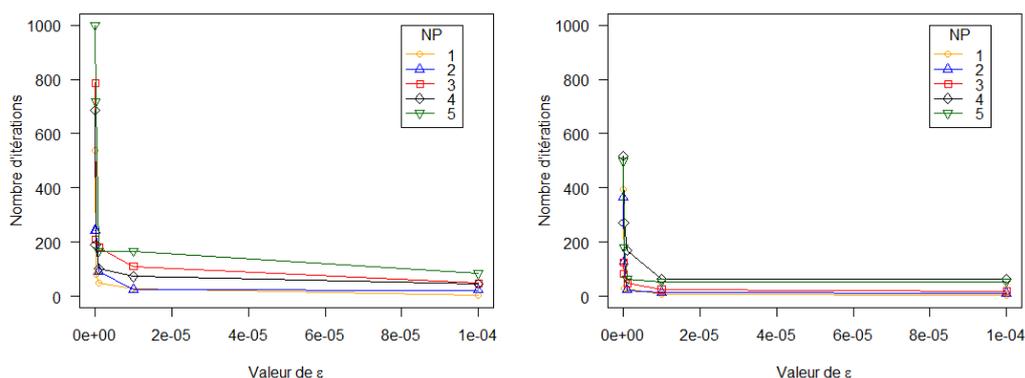


FIGURE 5.10 – Le nombre d'itérations effectuées en fonction de la valeur du critère d'arrêt, en utilisant une base d'apprentissage contenant 5 individus (Gauche) et 10 individus (Droite).

Comme dans l'étude précédente, le nombre d'itérations effectuées pour ajuster le modèle " A_{Rel} " augmente lorsque la valeur du critère d'arrêt, ϵ diminue et lorsque le

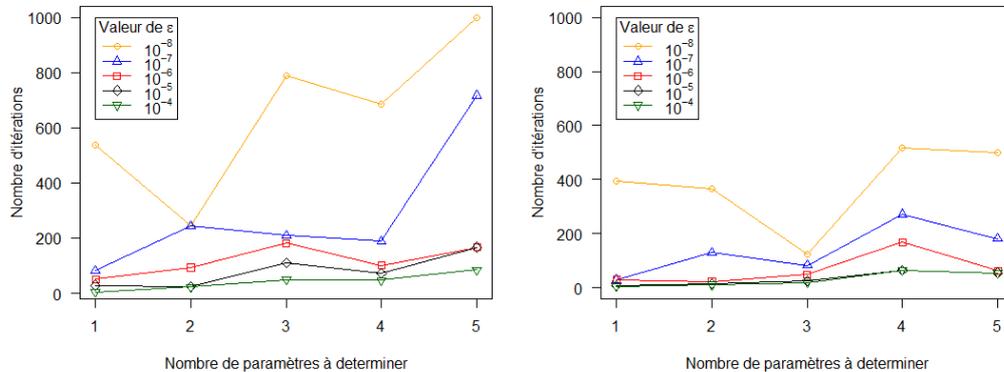


FIGURE 5.11 – Le nombre d'itérations effectuées en fonction du nombre de paramètres à déterminer, en utilisant une base d'apprentissage contenant 5 individus (Gauche) et 10 individus (Droite).

nombre de paramètres augmente.

Évolution du temps global d'apprentissage du modèle " A_{Rel} "

Grâce à l'étude réalisée dans la section 5.1.2 nous pouvons traduire l'évolution du nombre d'itérations effectuées en une évolution de la durée globale d'ajustement du modèle " A_{Rel} " en multipliant le nombre d'itérations effectuées par le temps nécessaire pour réaliser une itération. Nous avons obtenu la table 5.6.

Les résultats contenus dans cette table nous montrent les mêmes tendances que celles observées dans la section 5.1.4.

Etude du gain apporté par la prise en compte des relations existantes entre les paramètres en terme de vitesse d'apprentissage du modèle "A"

Afin de mesurer le gain apporté par la prise en compte des relations existantes entre les paramètres en terme de vitesse d'apprentissage, nous avons comparé l'évolution du nombre d'itérations à réaliser pour pouvoir apprendre les valeurs des paramètres, avec et sans la prise en compte de ces relations.

Lorsque l'on prend en compte les relations établies dans la section 2.3.2, l'apprentissage de la valeur de ω_d engendre également la détermination de la valeur de r_d . De la même façon, l'apprentissage de la valeur de f_d engendre également la détermination de la valeur de u_d . Avec le modèle " A_{Rel} " la détermination de la valeur de ω_d et

5.1. Comportement des modèles biomimétiques en fonction de leur structure et des caractéristiques des données d'apprentissage

TABLE 5.6 – Évolution de la durée (min) globale d'ajustement du modèle " A_{Rel} " en fonction du nombre de paramètres à déterminer, de la valeur du critère d'arrêt et du nombre d'individus (NI) présents dans la base d'apprentissage.

1 param. (ω_d)	2 param. (ω_d, f_d)	3 param. (ω_d, f_d, r_d)	4 param. (ω_d, f_d, r_d, u_d)	5 param. ($\omega_d, f_d, r_d, u_d, L_d$)	ϵ	NI
188	85	275	240	350	10^{-8}	5
28	85	73	66	251	10^{-7}	5
18	33	63	35	58	10^{-6}	5
10	9	39	26	58	10^{-5}	5
2	8	17	16	30	10^{-4}	5
296	274	94	388	374	10^{-8}	10
22	97	62	203	136	10^{-7}	10
22	17	37	127	47	10^{-6}	10
7	11	19	49	41	10^{-5}	10
4	8	14	49	41	10^{-4}	10

r_d ne nécessite donc que l'ajustement avec le modèle " A_{SsRel} " de ω_d . En revanche la détermination de la valeur de ω_d et r_d nécessite l'ajustement de ces deux paramètres. De même, avec le modèle " A_{Rel} ", la détermination des valeurs de ω_d, r_d, f_d et u_d ne nécessite l'ajustement que de ω_d et f_d alors qu'elle nécessite l'ajustement de ces quatre paramètres si le modèle " A_{SsRel} " est utilisé.

Dans cette étude nous avons donc comparé l'apprentissage de ω_d et r_d via l'utilisation du modèle " A_{SsRel} " et l'apprentissage de ω_d via l'utilisation du modèle " A_{Rel} ". Nous avons également comparé l'apprentissage de ω_d, r_d, f_d et u_d via l'utilisation du modèle " A_{SsRel} " et l'apprentissage de ω_d et f_d via l'utilisation du modèle " A_{Rel} ". Nous avons obtenu les tables 5.7 et 5.8.

TABLE 5.7 – Gain Moyen (GM) et Gain Moyen Relatif (GMR) en terme de nombre d'itérations à réaliser avant d'atteindre la convergence des paramètres, lorsque les relations existantes entre les couples de paramètres (ω_d, r_d) et (f_d, u_d) sont prises en compte.

	Détermination des valeurs de ω_d et r_d		Détermination des valeurs de ω_d, r_d, f_d et u_d	
	GM	GMR	GM	GMR
5 Individus	10	36.4%	466	70.8%
10 Individus	52	49.4%	74.8	59.5%

Les résultats contenus dans ces tables nous montrent que la prise en compte des

TABLE 5.8 – Gain Moyen (GM) et Gain Moyen Relatif (GMR) en terme de durée (minute) d'apprentissage, lorsque les relations existantes entre les couples de paramètres (ω_a , r_a) et (f_a , u_a) sont prises en compte.

	Détermination des valeurs de ω_a et r_a		Détermination des valeurs de ω_a , r_a , f_a et u_a	
	GM	GMR	GM	GMR
5 Individus	3,7	36,4%	172,4	70,8%
10 Individus	39,3	49,4%	27,7	59,5%

relations existantes entre les paramètres permet de réduire de manière assez conséquente le nombre d'itérations à réaliser. En effet, en moyenne cela permet un gain de vitesse d'apprentissage de 53.6% lorsque la base d'apprentissage contient 5 individus et de 54.4% lorsque la base d'apprentissage contient 10 individus (Tables 5.7 et 5.8).

Cette étude a également montré que le gain de vitesse d'apprentissage apporté par la prise en compte des relations existantes entre les paramètres est d'autant plus important lorsque le nombre d'individus dans la base d'apprentissage est faible. Comme dans la section 5.1.3 cela peut sans doute s'expliquer par le fait que plus la base de données d'apprentissage est petite, moins elle contient de variabilité (Voir figure 5.9).

5.2 Positionnement de l'outil de prédiction de croissance biomimétique par rapport à des outils existants

Comme nous l'avons vu dans l'état de l'art et plus précisément dans la section 1.2.5, les modèles permettant de simuler l'évolution de variables biologiques peuvent être répartis en deux catégories : les modèles dits «boîtes noires», et les modèles mécanistes. Dans cette section nous allons comparer le modèle "A de croissance" introduit dans la section 2.5, à des modèles appartenant à ces deux catégories.

5.2.1 Comparaison des performances du modèle de croissance biomimétique et de modèles "boîtes noires"

Parmi les modèles «boîtes noires», on distingue les régressions linéaire simple et multiples. Mais aussi les modèles logistiques, les réseaux de neurones et les forêts

aléatoires. Dans cette section nous avons donc choisi de comparer les capacités de notre outil biomimétique de croissance (modèle "A de croissance") à celles d'un modèle de régression linéaire, à celles de modèles logistiques, de réseaux de neurones et à celles de forêts aléatoires.

Pour cela nous avons ajusté ces modèles sur les mêmes données d'apprentissage que celles ayant servi à ajuster le modèle "A de croissance" (Section 2.5.2). De même, nous avons comparé les performances des différents outils en utilisant les mêmes données de test et les mêmes indicateurs d'erreur et de précision.

Comparaison entre le modèle biomimétique et une régression linéaire multiple

Nous avons appliqué une régression linéaire multiple sur notre base de données d'apprentissage, c'est-à-dire sur les individus du groupe «Témoin A». La base de test utilisée ici correspond aux individus du groupe «Témoin B» (voir Section 1.7.1).

Comme pour le modèle biomimétique, les données d'entrée correspondent au poids initial (P_0) de chaque individu et à leur consommation alimentaire ($C_{[0-0.67]}$ et $C_{[0.67-1.5]}$). La variable à prédire est le poids des individus à l'instant $t = 1$ (P_1). Le modèle linéaire testé est donc le suivant :

$$P_1 = \alpha_0 + \alpha_1 P_0 + \alpha_2 C_{[0-0.67]} + \alpha_3 C_{[0.67-1.5]} + \epsilon. \quad (5.2.1)$$

Nous avons ajusté les paramètres de ce modèle linéaire sur la base d'apprentissage et nous avons ensuite mesuré sa qualité en calculant la précision relative moyenne (ARA) sur les données d'apprentissage et sur les données de test. Cet indicateur ARA est donnée par l'équation (2.4.1) et rappelé ici :

$$ARA = 1 - \frac{1}{n} \frac{1}{m} \sum_{j=1}^m \left(\sqrt{\left(\sum_{i=1}^n \left(\frac{(O_{testi,j} - O_{simuli,j})}{O_{testi,j}} \right)^2 \right)} \right). \quad (5.2.2)$$

où $O_{testi,j}$ et $O_{simuli,j}$ correspondent respectivement à la valeur de la courbe contenue dans la base de test et à celle de la courbe simulée pour l'individu i au point j . n est le nombre d'individus de la base de test et m est le nombre de points des courbes.

En ajustant les paramètres α_0 , α_1 , α_2 et α_3 , uniquement sur le poids à l'instant $t = 1$ le modèle obtenu ne peut prédire le poids d'individus qu'à l'instant $t = 1$. Nous utilisons

donc le poids interpolé des individus de la base de test à $t = 1$ afin de comparer la qualité du modèle linéaire et celle du modèle biomimétique. Nous avons obtenu la table 5.9.

TABLE 5.9 – Précision (ARA) obtenue en utilisant un modèle linéaire ou le modèle biomimétique sur la base d'apprentissage (BA) et sur la base de test (BT).

Modèle	Précision _{BA} (%)	Précision _{BT} (%)
Modèle linéaire	98.4	97.0
Modèle biomimétique	98.1	98.5

Les résultats contenus dans cette table montrent que le modèle linéaire testé est associé à une plus grande précision sur la base d'apprentissage, mais à une précision plus faible sur la base de test que notre modèle biomimétique. Ainsi, dans le cadre de notre application, le modèle linéaire surajuste davantage les données d'apprentissage que notre modèle biomimétique.

Néanmoins la précision de ces outils est satisfaisante et le réel avantage du modèle biomimétique est sa capacité d'interpolation et d'extrapolation. En effet, en ajustant le modèle linéaire uniquement sur le poids des individus à l'instant $t = 1$, la notion de temps n'est pas intégrée dans le modèle. Dans le cadre de cette application, ce modèle linéaire ne peut donc pas être utilisé pour prédire le poids des individus avant et au-delà de l'instant $t = 1$.

De plus, le phénomène à prédire est une croissance logistique. Même si nous avons davantage de données d'apprentissage, le modèle linéaire ne pourrait donc pas extrapoler correctement le poids des animaux.

Comparaison entre le modèle biomimétique et des réseaux de neurones

Nous avons appliqué différents réseaux de neurones sur nos données de manière à comparer les capacités de ces outils d'apprentissage automatique à celles du modèle de croissance biomimétique que nous avons construit. Nous avons testé 6 réseaux de neurones ayant un nombre de couches cachées et un nombre de noeuds par couche différents (Table 5.10). Ces réseaux de neurones prennent en entrée le poids initial de chaque individu (P_0) et les informations concernant leur consommation alimentaire ($C_{[0-0.67]}$ et $C_{[0.67-1.5]}$).

Nous avons ajusté chaque réseau de neurones sur notre base de données d'apprentissage en utilisant la fonction *neuralnet* développée par FRITSCH *et al.* (2012) sur R, et nous avons calculé la précision de ces réseaux de neurones sur la base d'apprentissage dans un premier temps, puis sur la base de test (Table 5.10).

TABLE 5.10 – Les valeurs de l'indicateur ARA calculées sur la base d'apprentissage (ARA_{Train}), et sur la base de test (ARA_{Test}), à $t = 1$, en utilisant différents réseaux de neurones. Le réseau de neurone $(k_1, \dots, k_i, \dots, k_n)$ correspond à un réseau contenant n couches cachées et dans lequel la i^{me} couche contient k_i noeuds.

Structure	$ARA_{Train}(1)(\%)$	$ARA_{Test}(1)(\%)$
(4)	99.9	78.8
(4,3)	99.8	90.5
(6,5)	99.7	93.4
(4,6,6,3)	99.9	94.8
(5,7,7,7,4)	99.8	95.3
(5,9,9,9,5)	99.9	93
Modèle Biomimétique	98.1	98.5

Les résultats contenus dans la table 5.10 montrent que tous les réseaux de neurones testés sont associés à une très bonne précision sur la base d'apprentissage, mais à une précision plus faible sur la base de test. Les réseaux de neurones surajustent donc les données de la base d'apprentissage. Ce surapprentissage est d'autant plus important lorsque le réseaux contient trop ou trop peu de couches cachées et de noeuds. En effet, nous avons constaté que la précision des réseaux de neurones sur la base de test commence par augmenter lorsque le nombre de noeuds et de couches cachées augmente, puis diminue lorsque la complexité de la structure continue d'augmenter (Table 5.10). La précision la plus élevée sur la base de test (95.3%) est obtenue en utilisant un réseau de neurone contenant 5 couches cachées mais celle-ci reste plus faible que celle associée au modèle biomimétique (Table 5.10).

Ainsi, dans le cadre de notre application, étant donné qu'il y a très peu de données disponibles pour ajuster les outils, les réseaux de neurones surajustent les données d'apprentissage et restent moins précis que le modèle biomimétique. Ce constat confirme les problèmes de surajustement rencontrés lorsque les réseaux de neurones sont ajustés sur peu de données (Voir l'état de l'art, section 1.2.5).

Néanmoins la précision de ces outils d'apprentissage automatique est satisfaisante et le réel avantage du modèle biomimétique par rapport à ces outils est une fois de plus

sa capacité d'interpolation et d'extrapolation. En effet, comme le modèle biomimétique, les réseaux de neurones testés ont été ajustés uniquement sur la valeur de la variable de sortie à l'instant $t = 1$. Dans ce cas, les réseaux de neurones que nous avons testés ne prennent pas en compte la notion de temps et ne peuvent être utilisés que pour prédire la valeur de la variable de sortie à $t = 1$. Ainsi, contrairement au modèle biomimétique, les réseaux de neurones testés n'ont pas de capacité d'interpolation, ni d'extrapolation.

Comparaison entre le modèle biomimétique et des forêts aléatoires.

Nous avons appliqué une série de forêts aléatoires, appelées aussi «Random Forest», sur nos données. L'objectif est de comparer les capacités de ces outils d'apprentissage automatique et celles du modèle de croissance biomimétique.

Nous avons testé 700 forêts aléatoires contenant respectivement 1 à 700 arbres. Ces 700 forêts aléatoires prennent en entrée le poids initial de chaque individu (P_0) et les informations concernant leur consommation alimentaire ($C_{[0-0.67]}$ et $C_{[0.67-1.5]}$).

Nous avons ajusté ces 700 forêts aléatoires sur nos données d'apprentissage et leur qualité a ensuite été mesurée en calculant l'*ARE*, donné par l'équation (5.1.10) sur les données d'apprentissage et sur les données de test.

Nous constatons sur la figure 5.12 que la qualité des forêts aléatoires testées se stabilise lorsque le nombre d'arbres utilisés augmente. Nous déduisons donc de la figure 5.12, qu'il est inutile d'appliquer des forêts aléatoires contenant plus de 200 arbres.

Sur la figure 5.12 nous remarquons qu'il y a une certaine variabilité en ce qui concerne la précision des différentes forêts aléatoires. Cela est dû à l'initialisation aléatoire des différentes forêts testées. Nous donnerons donc une gamme de précision pouvant être atteinte en utilisant des forêts aléatoires.

Pour étudier cette gamme de précision, nous avons appliqué des forêts aléatoires contenant 1 à 200 arbres, en ajustant 5 fois chaque structure. Nous avons donc ajusté dans cette étude 1000 forêts aléatoires sur nos données d'apprentissage. Nous avons obtenu la figures 5.13.

Nous avons calculé sur la base d'apprentissage et sur la base de test la gamme de précision relative, la variance et l'écart-type de cette indicateur de qualité (Table 5.11). Nous avons également isolé la forêt aléatoire présentant la meilleure précision sur la

5.2. Positionnement de l'outil de prédiction de croissance biomimétique par rapport à des outils existants

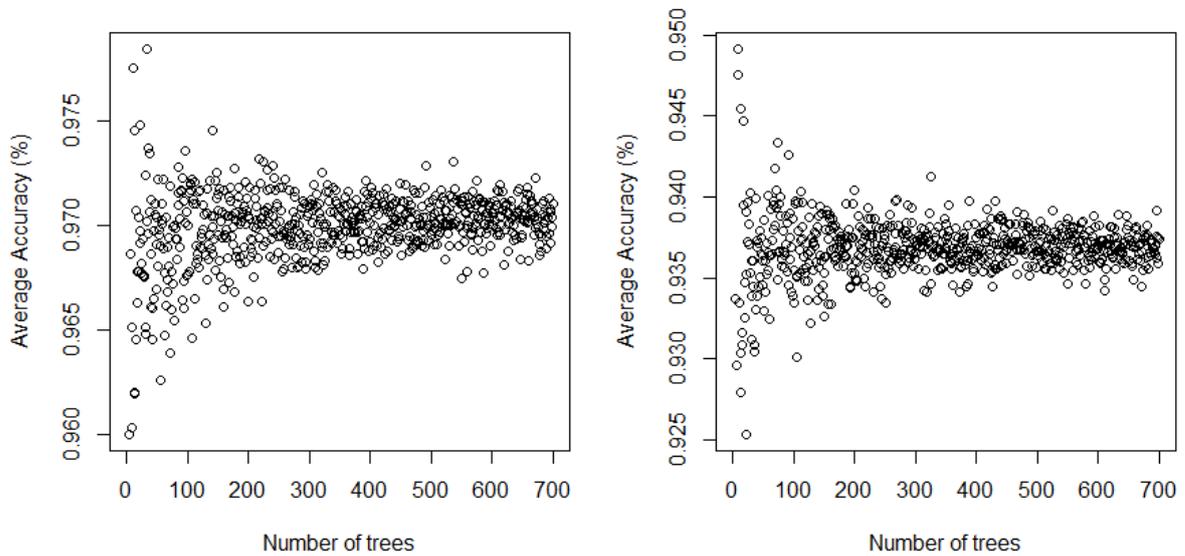


FIGURE 5.12 – Précision obtenue ($ARE(1)$) en utilisant des Random Forest contenant différents nombres d'arbres (1 à 700), sur la base d'apprentissage (Gauche) et sur la base de test (Droite).

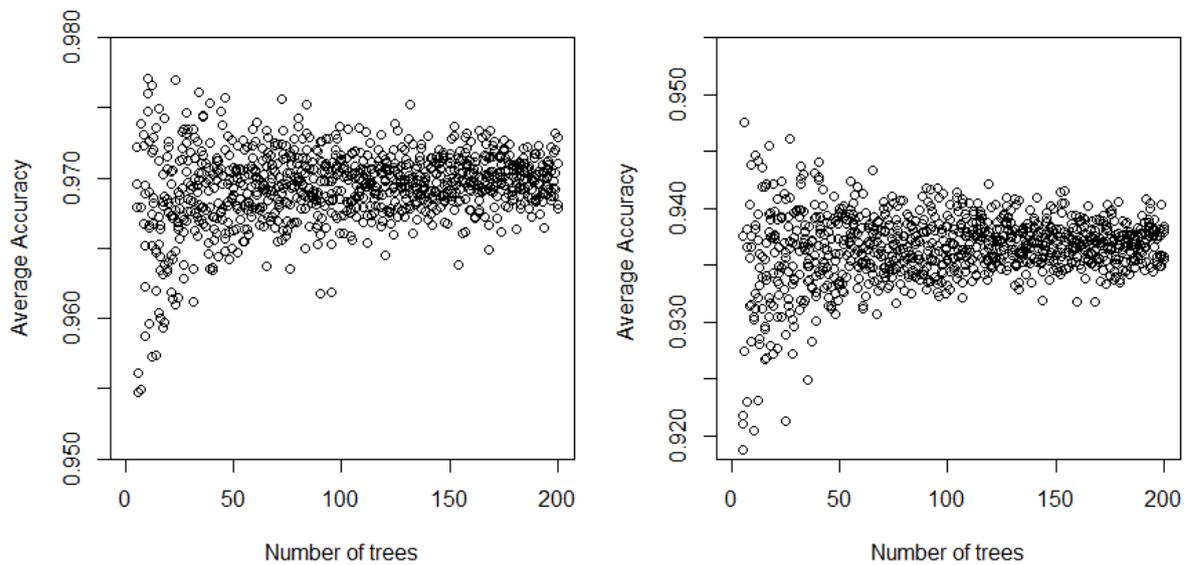


FIGURE 5.13 – Précision obtenue ($ARE(1)$) en utilisant des Random Forest contenant différents nombres d'arbres (1 à 200), sur la base d'apprentissage (Gauche) et sur la base de test (Droite).

base de test (Table 5.12).

Nous constatons plus précisément sur la figure 5.13 que la qualité des forêts aléa-

TABLE 5.11 – Gamme de précision, variance et écart-type de la précision relative ($ARE(1)$) pouvant être obtenue en utilisant des Random Forest contenant différents nombres d'arbres (1 à 200), sur la base d'apprentissage et sur la base de test.

	Gamme de précision	Variance	écart type
Base d'apprentissage	[0.955 ;0.987]	$7.2 \cdot 10^{-6}$	0.003
Base de test	[0.919 ;0.947]	$9.0 \cdot 10^{-6}$	0.003

TABLE 5.12 – La meilleure précision obtenue ($ARE(1)$) sur la base de test, en utilisant des Random Forest (BA : Base d'Apprentissage, BT : Base de Test).

Précision sur la BA	Précision sur la BT	Nombre d'arbres
0.955	0.947	6

toires testées se stabilise lorsque le nombre d'arbres est supérieur à 60.

La gamme de précision relative associées aux forêts aléatoires testées sur la base de test, varie entre 0.919 et 0.947 (Table 5.11). La précision la plus élevée a été obtenue en utilisant une forêt aléatoire contenant 6 arbres (Table 5.12) mais celle-ci reste néanmoins plus faible que celle associée au modèle biomimétique et qui était de 98.5% (Table 5.10).

Les résultats contenus dans la table 5.11 montrent que toutes les forêts aléatoires testées ont une grande précision sur la base d'apprentissage, mais une précision plus faible sur la base de test. Ainsi, dans le cadre de notre application, comme les réseaux de neurones, les forêts aléatoires surajustent les données d'apprentissage et restent moins précises que le modèle biomimétique. Néanmoins, la précision de ces outils d'apprentissage automatique est satisfaisante et le réel avantage du modèle biomimétique, reste sa capacité d'interpolation et d'extrapolation. En effet, les forêts aléatoires testées ont été ajustées uniquement sur la valeur de la variable de sortie à l'instant $t = 1$. Elles n'intègrent donc pas la notion de temps et ne peuvent être utilisées que pour prédire la variable de sortie à $t = 1$. Ces outils d'apprentissage automatique n'ont donc pas de capacité d'interpolation ni d'extrapolation.

5.2.2 Comparaison des performances du modèle de croissance biomimétique et des modèles classiques de croissance logistique

Le modèle de GOMPERTZ (1825),

$$\frac{dN(t)}{dt} = a_G \cdot N(t) \cdot \ln \left(\frac{K_G}{N(t)} \right), \quad (5.2.3)$$

et le modèle de VERHULST (1838),

$$\frac{dN(t)}{dt} = a_V \cdot N(t) \cdot \left(1 - \frac{N(t)}{K_V} \right), \quad (5.2.4)$$

sont deux modèles couramment utilisés pour modéliser des phénomènes de croissance (Voir par exemple les études réalisées par WINSOR (1932), SAKOMURA *et al.* (2005), BUYSE *et al.* (2004), ROBERTSON (1916), ROBERTSON (1923) et ROMÁN-ROMÁN et TORRES-RUIZ (2012)). Les modèles construits par Gompertz et Verhulst modélisent des processus de croissance bornée, respectivement par K_G et K_V .

Nous avons ajusté les paramètres de ces deux modèles sur les individus du groupe «Témoin A» en utilisant le même algorithme que celui utilisé pour ajuster le modèle biomimétique (Voir la section 2.5). Comme pour le modèle biomimétique, ces deux modèles classiques de croissance logistique sont ajustés en minimisant la valeur de l'indicateur d'erreur $ARE(1)$ donné par l'équation (5.1.10).

Une fois les paramètres ajustés nous avons calculé la valeur de l'indicateur $ARA(1)$ sur la base d'apprentissage, donné par l'équation (5.2.2) (Voir la table 5.13). Sur cette base de données, le modèle biomimétique est associé à la précision la plus élevée, néanmoins, la précision des autres modèles est également très satisfaisante.

Nous avons ensuite calculé l'indicateur ARA à l'instant $t = 0.6$, à $t = 1.52$ et à $t = 2.5$, sur la base de test. Les résultats contenus dans la table 5.14 et les courbes des figures 5.14 et 5.15 montrent que le modèle de Gompertz simule un ralentissement trop important et trop précoce de la croissance des individus. En revanche, le modèle de Verhulst présente une bonne précision sur l'ensemble de la période de suivi des individus de la base de test.

La similarité entre les résultats fournis par le modèle biomimétique et ceux fournis par le modèle de Verhulst était attendue. En effet, une équation assimilable à celle de Verhulst est intégrée dans notre modèle (Voir la section 2.1.2). Le réel avantage du modèle de croissance biomimétique est sa capacité à intégrer des données d'entrée et donc des informations supplémentaires concernant les individus étudiés. En effet le modèle de Verhulst ne prend en compte que le poids initial des individus, c'est-à-dire P_0 . Notre modèle prend lui aussi en compte ce poids initial mais il intègre également des données d'entrée concernant la consommation alimentaire des individus tout au long de la période de suivi. L'intégration d'informations supplémentaires semble permettre à l'outil d'affiner la prédiction du poids des individus sur l'ensemble de la période étudiée (Table 5.14). Cette capacité de nos outils biomimétiques à intégrer des informations d'entrée au cours du temps est aussi une capacité très intéressante pour réaliser de l'Assimilation de Données et qui les distingue des modèles de Gompertz et de Verhulst.

TABLE 5.13 – Valeurs des paramètres obtenues après l'ajustement et valeurs moyennes de l'indicateur $ARA(1)$ calculées sur la base d'apprentissage pour les différents modèles étudiés.

Modèle	a	K	ARA(1)
Gompertz	$a_G = 0.412$	$K_G = 0.563$	0.978
Verhulst	$a_V = 0.411$	$K_V = 1.563$	0.979
Biomimétique			0.981

TABLE 5.14 – Valeurs moyennes de l'indicateur ARA calculées sur la base de test à différents instants, pour le modèle biomimétique, celui de Verhulst et celui de Gompertz.

		Modèles		
		Verhulst	Gompertz	Biomimétique
t	0.6	0.985	0.980	0.986
	1.52	0.968	0.937	0.971
	2.5	0.979	0.923	0.985

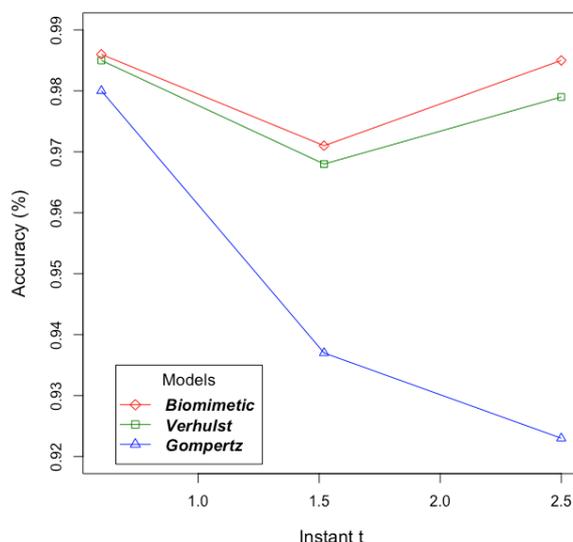


FIGURE 5.14 – Représentation graphique des valeurs moyennes de l'indicateur ARA calculées sur la base de test à différents instants et associé aux différents modèles testés.

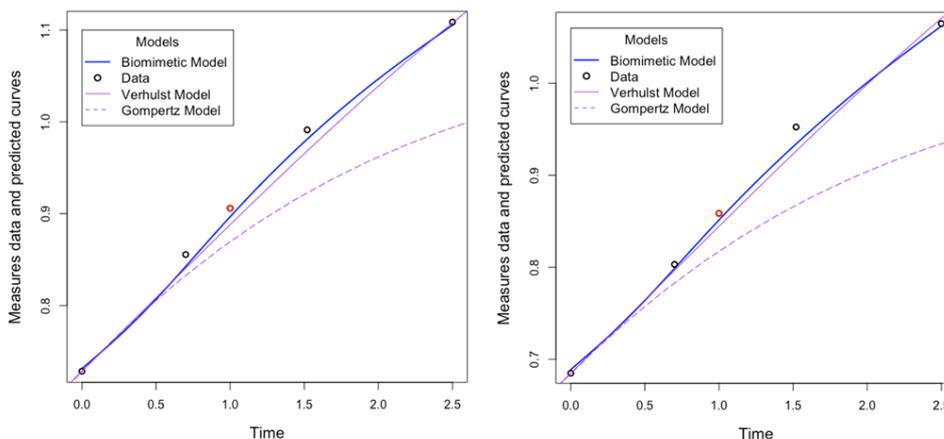


FIGURE 5.15 – Exemples de courbes de croissance de la base de test simulées le modèle biomimétique, le modèle de Gompertz et celui de Verhulst.

5.2.3 Comparaison entre le modèle de croissance biomimétique et des modèles de croissance mécanistes

Des modèles de croissance de type mécanistes réalistes ont été développés par BASTIANELLI *et al.* (1996), MACH et KRISTKOVA (2010), BRUN-LAFLEUR *et al.* (2013)

et ZÚÑIGA *et al.* (2014). Ces modèles intègrent de nombreuses données d'entrée et certaines d'entre elles ne sont pas disponibles dans nos données et ne peuvent donc pas être appliqués sur celles-ci. De ce fait, nous allons uniquement comparer ces modèles et notre modèle biomimétique en termes de structures, de fonctionnement et d'objectifs.

Les modèles mécanistes réalistes et notre modèle biomimétique ont en commun l'objectif d'embarquer des connaissances a priori, traduites par des équations mathématiques. En revanche, le niveau d'abstraction et de réalisme est très différent dans ces deux approches. Cette différence s'explique par des objectifs eux-mêmes très différents.

Contrairement au modèle biomimétique, les modèles mécanistes sont construits dans l'optique d'avoir un outil réaliste et explicatif, et non pas un outil pour réaliser de l'Assimilation de Données. En effet, comme nous l'avons vu dans l'état de l'art (Section 1.2.5), ces modèles de croissance mécanistes sont utilisés pour réaliser des analyses quantitatives et apporter des réponses à des hypothèses biologiques. Ces objectifs expliquent le besoin de prendre en compte la dynamique du système étudié avec un niveau de précision et de réalisme très important. Ces modèles contiennent de ce fait, un grand nombre de paramètres à déterminer (entre 10 et 30) et prennent en entrée un nombre important de facteurs. La structure de ces modèles les rend peu adaptés à l'optimisation de la valeur de paramètres et à l'Assimilation de Données.

5.2.4 Conclusion du chapitre 5

Les outils d'apprentissage statistique biomimétiques que nous avons construits ont pour vocation de prédire l'évolution de variables biologiques et de réaliser de l'Assimilation de Données d'élevage. Nous avons donc choisi de privilégier certaines qualités du modèle et de l'outil de prédiction basé sur ce modèle.

Les applications de nos outils sur des données générées et des données réelles ont montré que nos outils sont capables de prédire avec précision l'évolution d'une variable biologique, malgré une quantité de données d'apprentissage très faible. Les applications d'outils existants sur les mêmes données ont montré que nos outils d'apprentissage statistique biomimétiques se distinguent par plusieurs de leurs capacités.

TABLE 5.15 – Tableau de comparaison de la précision (%) des différentes méthodes utilisées pour prédire le poids des individus à l'instant $t = 1$ ($P(1)$) et à l'instant $t = 2.5$ ($P(2.5)$) (BA : Base d'Apprentissage, BT : Base de Test, RL : Régression Linéaire, RN : Réseaux de Neurones, RF : Random Forest, Biomim. : Modèle Biomimétique).

	Modèles					
	RL	Gompertz	Verhulst	RN	RF	Biomim.
$P(1)_{BA}$	98.4	97.8	97.9	99.8	98.7	98.1
$P(1)_{BT}$	97.0	97.1	98.3	95.3	94.7	98.5
$P(2.5)_{BT}$	×	92.3	97.9	×	×	98.5

Nos outils se distinguent des régressions linéaires et des outils d'apprentissage automatique par leur capacité à apprendre la dynamique reliant des données d'entrée et de sortie sur peu de données sans les surajuster, mais aussi par leur capacité d'extrapolation (Voir la table 5.15).

Nos outils se distinguent ensuite des outils mécanistes réalistes par la parcimonie de leurs modèles mathématiques. En effet, les modèles que nous avons construits contiennent beaucoup moins d'équations et de paramètres à déterminer que les modèles réalistes. Cela facilite leur ajustement sur des données et leur implémentation.

Enfin, notre modèle de croissance biomimétique se distingue des modèles de croissance logistique classiques par sa capacité à intégrer des données d'entrée au cours du temps.

Pour résumer, nos outils se distinguent par leur capacité à réaliser de l'Assimilation de Données complexes et potentiellement éparses. La capacité de nos outils à s'ajuster sur peu de données et à réaliser de l'extrapolation leur permet également de se distinguer par la réduction des coûts d'expérimentation et de gestion des données qu'ils permettent.

Néanmoins, la conception de nos outils nécessite d'avoir quelques connaissances a priori concernant la dynamique du système étudié et quelques notions de modélisation mathématique, notamment si l'utilisateur souhaite prendre en compte de nouvelles variables d'entrée ou d'autres processus.

Comme nous l'avons vu dans la partie 5.1, l'apprentissage des paramètres peut nécessiter du temps ou de disposer d'une importante puissance de calcul.

De plus, nos outils possèdent une capacité de prédiction très satisfaisante mais

ils n'ont pas pour vocation d'expliquer ou de quantifier les phénomènes biologiques sous-jacents reliant les données d'entrée et les données de sortie.

Nos outils d'apprentissage statistique biomimétiques ne seront donc pas les outils les plus adaptés si l'objectif est de mettre en lien un très grand nombre de variables différentes et si l'utilisateur a besoin que l'apprentissage de paramètres se fasse très rapidement. Nos outils ne seront pas non plus les plus adaptés si l'utilisateur attache beaucoup d'importance à l'interprétabilité du modèle et des paramètres obtenus après un apprentissage sur des données.

CHAPITRE CONCLUSIF : RÉFLEXIONS EN COURS ET PERSPECTIVES

Les travaux de thèse présentés dans ce manuscrit ont amené d'autres réflexions. Nous avons notamment réfléchi aux informations et aux conseils pouvant être apportés par à une potentielle application mobile qui serait basée sur notre modèle de croissance biomimétique.

Les résultats que nous avons obtenus dans les différentes études réalisées résultent d'un certain nombre d'hypothèses. En effet, nous avons défini la structure des équations constituant le modèle mathématique, nous avons fixé des paramètres et choisi une manière de relier certains d'entre eux. Une procédure de sélection de modèles pourrait permettre de déterminer la structure du modèle mathématique optimisant par exemple la précision de l'outil de prédiction final.

De plus, la finalité des outils d'apprentissage statistique que nous avons construits est de réaliser de l'Assimilation de Données. Nous avons donc entamé une réflexion sur la manière dont nos outils peuvent être embarqués pour réaliser de l'Assimilation de Données. Nous avons également réfléchi à la façon de tester leurs performances dans ce domaine.

6.1 Réflexion concernant une potentielle application mobile permettant de piloter un atelier d'engraissement.

Les capacités de prédiction et d'extrapolation de notre modèle de croissance ont été validées au cours de ces travaux de thèse. Une réflexion a donc été menée afin de savoir comment le modèle de croissance que nous avons construit pouvait être intégré dans une application mobile et quels services pouvaient être rendus par l'utilisation de cette application.

En fonction du prix de la viande et du prix de l'aliment, l'éleveur peut vouloir simuler et prévoir le scénario le plus intéressant économiquement. Nous avons donc pensé à une application mobile capable de comparer différents scénarii.

Le premier scénario pourrait consister à prédire le poids d'un animal ou d'un groupe d'animaux à la fin d'une période d'engraissement. En effet, nous avons vu que notre modèle de croissance est capable de prédire le poids d'un individu à un instant donné à partir de sa consommation alimentaire et de son poids initial. Le poids initial utilisé peut correspondre au poids des animaux à leur naissance ou lors de leur arrivée dans l'élevage.

Dans ce scénario l'éleveur peut aussi comparer plusieurs plans de nourrissage possibles, c'est-à-dire différents enchaînements de périodes plus ou moins longues pendant lesquelles une certaine quantité d'aliment, contenant une certaine dose d'additif est distribuée quotidiennement.

L'application pourrait fournir comme résultat une estimation du poids actuel des animaux et une prédiction du poids final. Certains indicateurs peuvent être calculés comme le gain de poids total, la durée d'engraissement prévue et la quantité totale d'aliment ingérée par les animaux. Ces indicateurs pourraient permettre à l'éleveur de prendre une décision en termes de plans de nourrissage, de prévoir le poids des animaux à la fin de la période d'engraissement et donc de prévoir le bénéfice qu'il pourrait faire. La quantité d'aliment consommée peut également permettre à l'éleveur de s'organiser et de gérer ses stocks d'aliments.

Le deuxième scénario pourrait consister à prédire la date correspondant à la fin de l'engraissement des animaux. Pour répondre à cette question nous avons besoin de connaître notamment l'objectif à atteindre en termes de poids (P_{Obj}). Comme précédemment, l'éleveur peut aussi comparer plusieurs plans de nourrissage.

Dans ce cas, l'application pourrait rappeler en sortie le poids initial des animaux. Elle pourrait également afficher une estimation du poids actuel et une prédiction de la durée d'engraissement nécessaire pour atteindre le poids objectif P_{Obj} . La quantité totale d'aliment ingérée par les animaux peut être calculée pour chaque plan de nourrissage proposé. Dans ce scénario, cela pourrait permettre à l'éleveur de prendre une décision en termes de plan de nourrissage, de prévoir la fin de la période d'engraissement de ces animaux en fonction de ses propres objectifs et de gérer ses stocks

d'aliments.

Enfin, un troisième scénario pourrait consister à prédire la quantité optimale d'aliment à distribuer pour atteindre un certain objectif d'engraissement. Un éleveur peut avoir pour ses animaux un objectif de poids à atteindre et un objectif de date avant laquelle il veut avoir terminé l'engraissement de ses animaux.

Pour répondre à cette question nous avons besoin de connaître notamment l'objectif à atteindre en termes de poids (P_{Obj}) et de date finale (D_{Obj}). Comme précédemment, l'éleveur peut aussi comparer plusieurs plans de nourrissage.

L'application pourrait rappeler en sortie le poids initial des animaux. Elle pourrait également afficher une estimation du poids actuel et une prédiction de la quantité d'aliment à distribuer pour atteindre le poids objectif P_{Obj} avant la date D_{Obj} . Dans ce scénario, cela pourrait permettre à l'éleveur de gérer ses stocks d'aliments et de prévoir ses coûts d'élevage sur l'ensemble de sa période d'engraissement.

Actuellement, notre modèle de croissance n'a été ajusté que sur un seul type de ration de base et un seul additif. Pour l'instant cette application mobile ne pourrait être utilisée que dans des élevages où les animaux reçoivent cette ration alimentaire de base et une quantité plus ou moins importante de l'additif étudié. L'ajustement du modèle sur d'autres rations et d'autres additifs pouvant être utilisées dans les élevages permettrait de pouvoir utiliser une application de ce type dans un nombre plus conséquent d'élevages.

Lorsque nous serons capables de prédire le poids des animaux en fonction de différentes rations nous serons également capables de comparer les courbes de croissance et les poids finaux obtenus en fonction de la composition de la ration. Un conseil pourra alors être apporté à l'éleveur pour qu'il puisse choisir un scénario optimisant les performances de ses animaux et son chiffre d'affaire final.

6.2 Réflexion et premiers pas vers l'automatisation d'une procédure de sélection de modèles dans le cadre du Couplage Modèle-Données

La définition du niveau d'abstraction du modèle est une étape décisive dans une démarche de type "Couplage Modèle-Données". En effet, un niveau d'abstraction élevé

peut permettre d'obtenir un outil très léger et facile à implémenter. Néanmoins cette simplicité peut être au détriment de la précision du modèle. À l'inverse, un modèle présentant un niveau de détails et de réalisme important, peut contenir un grand nombre de paramètres à déterminer, nécessiter un temps d'apprentissage important et aboutir à des problèmes de surapprentissage. L'objectif serait donc de déterminer, par des méthodes de sélection, le niveau d'abstraction le plus pertinent en termes de précision du modèle, de temps d'apprentissage et de nombre de paramètres à déterminer.

Dans cette optique nous avons initié quelques travaux en vue d'automatiser une procédure de sélection de modèles dans le cadre du Couplage Modèle-Données. Au cours de l'année 2018 – 2019, j'ai participé à l'encadrement d'un projet de recherche de 6 mois au sein de l'université. Lors de ce projet un outil a été construit afin d'ajuster un modèle sur des données et de collecter certaines informations associées à celui-ci.

Dans un premier temps nous avons imaginé la structure d'un outil gérant le processus globale de sélection de modèles.

6.2.1 Conception d'un outil de sélection de modèles

L'objectif est de lister une série de modèles que nous voulons comparer les uns aux autres et de sélectionner le meilleur d'entre eux au regard d'un certain nombre de critères. Pour cela il est nécessaire d'ajuster l'ensemble de ces modèles sur des données et de mesurer un certain nombre de métriques permettant de comparer ces modèles entre eux. La figure 6.1 schématise l'architecture de cet outil.

Dans la figure 6.1, le programme "Scoring" contient deux autres programmes, nommés "Fitting" et "Testing" dont le fonctionnement est schématisé par la figure 6.2.

L'un des objectifs de ce projet était d'imaginer un outil qui soit le plus générique possible. Pour cela, nous voulons que chaque modèle soit associé à une série d'informations. Parmi ces informations nous avons par exemple un identifiant unique (ID_M), les équations du modèle (E), le nombre de paramètres à ajuster ($NParam$), la gamme de valeurs de ces paramètres ($Range$) et la valeur de ceux qui sont fixés (VF). De même, chaque base de données disponible doit être associée à un identifiant unique (ID_B), et à une information précisant sa taille ($SizeB$) et la part de cette base destinée

6.2. Réflexion et premiers pas vers l'automatisation d'une procédure de sélection de modèles dans le cadre du Couplage Modèle-Données

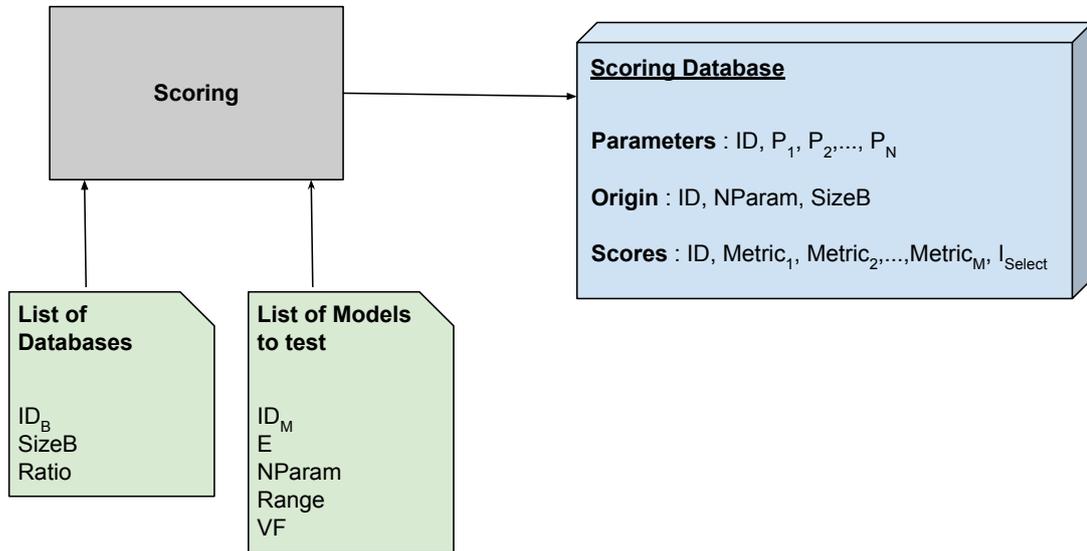


FIGURE 6.1 – Processus de comparaison de l'ensemble des modèles

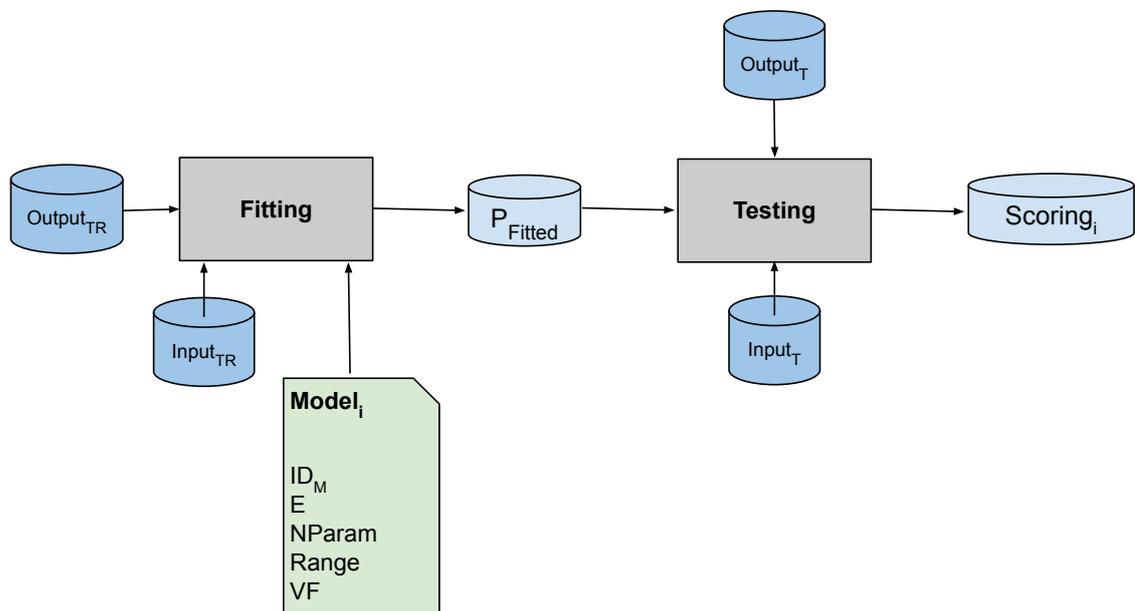


FIGURE 6.2 – Schématisation du processus d'ajustement et de test des différents modèles testés.

à l'apprentissage et celle destinée au test de qualité de prédiction des modèles (*Ratio*).

Dans un premier temps, nous avons besoin d'un historique de données, composé de données d'entrée (Inputs) et de données de sortie (Outputs). Cet historique de données sera ensuite divisé en deux parties de façon à avoir une base d'apprentissage ($Input_{TR}$ et $Output_{TR}$) et une base de test ($Input_T$ et $Output_T$).

Le programme nommé "Fitting", appelle tout d'abord le premier modèle ($Model_1$) parmi la listes de modèles que nous voulons comparer et utilise les données de la base d'apprentissage choisie, ainsi qu'un algorithme d'optimisation pour ajuster les paramètres de ce modèle. L'algorithme d'optimisation utilisé peut être DIRECT et la procédure d'apprentissage peut être la même que celle appliquée dans la section 2.3.3 et 2.5.4 de ce manuscrit.

Une fois les paramètres de ce premier modèle ajustés, nous le paramétrons et nous testons sa qualité sur les données de la base de test, grâce à un programme nommé "Testing". Ce programme calcule les premières métriques, comme par exemple l'erreur et la précision du modèle sur les données de test. L'ensemble des métriques pouvant être calculées sont ensuite stockées dans une base de données appelée "*Scoring*₁".

Ce processus menant à la construction de la base de données "Scoring" est ainsi répétée pour chaque modèle ($Model_i$) que nous souhaitons tester. Une fois tous les modèles ajustés et testés, la base de données "Scoring" contient alors les métriques associées à l'ensemble des modèles que nous voulons comparer.

6.2.2 Sélection des modèles

Les différents modèles testés peuvent être sélectionnés sur différents critères. Le critère qui nous semble primordial est la qualité des prédictions faites sur les données de la base de test. Mais d'autres critères peuvent être pris en compte pour sélectionner un modèle. En effet, le nombre de paramètres à ajuster, le temps d'apprentissage et l'interprétabilité des résultats peuvent influencer le choix d'un modèle plutôt qu'un autre.

L'objectif est donc de construire un indicateur, I_{Select} , représentant la qualité globale

d'un modèle en fonction des attentes de l'utilisateur. Cet indicateur peut correspondre à une somme dans laquelle l'utilisateur peut pondérer les différentes métriques en fonction de ses besoins. Par exemple, l'indicateur I_{Select} peut s'écrire :

$$I_{Select} = \alpha_1 ET + \alpha_2 NP + \alpha_3 TA + \alpha_4(10 - NI) \quad (6.2.1)$$

où,

- ET est l'erreur du modèle sur la base de test,
- NP est le nombre de paramètre à ajuster,
- TA est le temps d'apprentissage du modèle, et
- NI correspond à une note sur 10, d'interprétabilité des résultats

Ainsi, un utilisateur qui donnera par exemple plus d'importance à la précision du modèle qu'au temps d'apprentissage pourra attribuer une valeur plus élevée au poids α_1 qu'au poids α_2 .

Dans le cas présenté précédemment, le modèle sélectionné correspondra au modèle présentant la plus faible valeur de l'indicateur I_{Select} .

6.2.3 Perspectives de ce projet

L'outil construit est capable pour l'instant d'ajuster un modèle sur des données et de collecter certaines informations associées à celui-ci, comme notamment le nombre de paramètres et la précision de ce modèle sur des données de test. Dans le futur, un des objectifs sera donc d'éprouver cet outil, d'augmenter sa robustesse et son niveau de généralité et enfin d'automatiser la procédure complète de sélection de modèles.

6.3 Premières réflexions autour de la construction et de l'évaluation d'un outil d'Assimilation de Données, basé sur un outil d'apprentissage statistique biomimétique.

Comme il a été expliqué dans l'introduction générale de ce manuscrit, les outils d'apprentissage statistique développés dans ces travaux de thèse ont été construits dans l'optique d'être capable de réaliser de l'Assimilation de Données d'élevage. Dans le cadre de l'élevage de précision, l'objectif de l'Assimilation de Données est de synthétiser et d'interpréter un flot de données provenant de manière plus ou moins régulière d'un élevage connecté. Le but étant d'estimer l'état réel d'un système à un instant t , en combinant l'information générée par un modèle mathématique ajusté et celle contenue dans des données collectées (Voir l'état de l'art, section 1.4 et AUROUX et BLUM (2005), GREGG *et al.* (2009), LGUENSAT *et al.* (2017) et LGUENSAT *et al.* (2019)). L'outil d'Assimilation de Données doit donc être capable d'estimer l'erreur associée au modèle mathématique, notée η , et le bruit contenu dans les données, noté κ .

Dans cette section nous allons présenter les réflexions que nous avons entamées concernant la manière dont les outils d'apprentissage statistique biomimétiques peuvent être embarqués dans un outil d'Assimilation de Données. Un des aspects auxquels nous avons également réfléchi à la façon dont la qualité d'un outil d'Assimilation de Données peut être testée.

Dans l'ensemble de cette section, le modèle mathématique qui sera embarqué dans l'outil d'Assimilation de Données sera appelé M_k . Ce modèle mathématique correspond à un modèle de croissance semblable au modèle "A" présenté dans la section 2.1. Il contient L_T paramètres pouvant être appris mais nous ne faisons varier dans cette section que L paramètres et nous fixons les $(L_T - L)$ paramètres restants.

6.3.1 Les notations utilisées dans cette section

Dans cette section nous utilisons quelques notations :

- X , est le nom de l'élément,
- \tilde{X} , correspond aux valeurs numériques de l'élément X ,

- \hat{X} , correspond aux valeurs estimées de l'élément X ,
- \check{X} , contient les valeurs de l'élément X associées à la réalité, et
- \bar{X} , correspond à un élément dont les sous-éléments sont des distributions.

L'ensemble de ces notations seront reprises et développées dans les paragraphes suivants.

6.3.2 Réflexion autour de l'élaboration d'une méthode permettant de tester la qualité d'un outil d'Assimilation de Données

L'outil d'Assimilation de Données doit être capable d'estimer l'état réel, noté $\check{\epsilon}(t)$, d'un système ou d'un individu, \check{i} . Pour cela il doit combiner l'information fournie par un modèle M_k et celle contenue dans des données collectées, O_{TR} . Ces informations doivent être combinées sachant que le modèle est associé à une erreur η et que les données contiennent un bruit κ .

Pour vérifier la qualité de l'outil d'Assimilation, il faut comparer l'état estimé à l'instant t ($\hat{\epsilon}(t)$), à l'état considéré comme étant l'état réel du système ($\check{\epsilon}(t)$).

Dans la réalité, cet état $\check{\epsilon}(t)$ n'est jamais accessible. Mais pour ce projet nous voulons générer $\check{\epsilon}(t)$ à partir du modèle mathématique M_k paramétré avec des paramètres fixés \check{P}_k .

La figure 6.3 schématise l'enchaînement des différentes étapes permettant de tester la qualité d'un outil d'Assimilation de Données.

Génération de données fictives à partir d'un modèle mathématique paramétré

L'objectif de la première étape de cette méthode est de générer deux bases de données. Dans un premier temps, nous devons générer un historique de données bruitées, associées à une population d'individus ($O_H(t)$) contenant de la variabilité interindividuelle. Il faut également générer des données bruitées de suivi en temps réel, associées à un individu \check{i} . Cet individu \check{i} est un individu appartenant à la population citée précédemment, pour lequel nous voulons estimer l'état réel et dont les données seront assimilées.

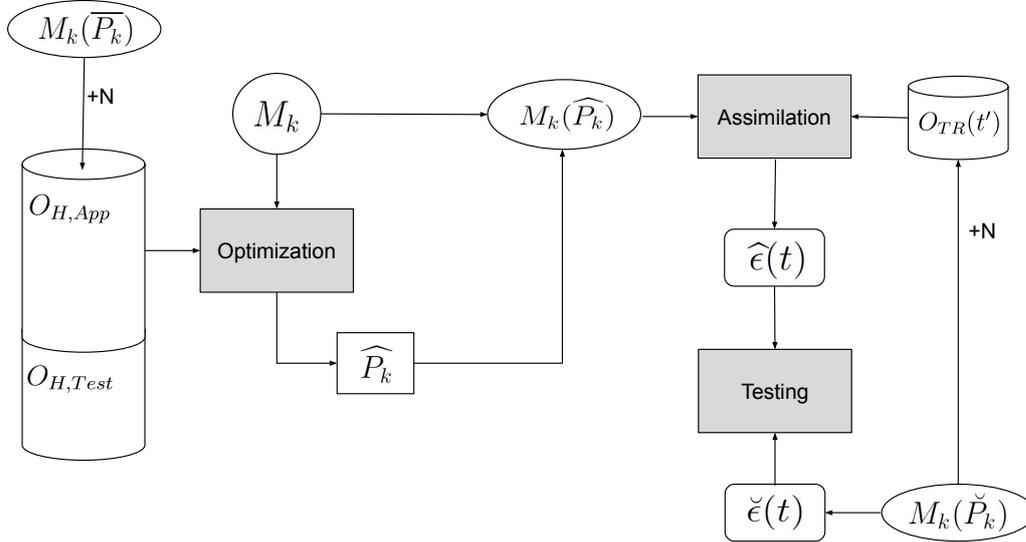


FIGURE 6.3 – Schématisation des étapes permettant de tester la qualité d'un outil d'Assimilation de Données.

Pour générer les données nous paramétrons le modèle M_k en utilisant la même méthode que celle élaborée dans la section 2.3.1.

Pour générer de la variabilité interindividuelle, le modèle M_k sera paramétré avec les paramètres contenus dans \overline{P}_k . Les sous-éléments de \overline{P}_k correspondent aux différents paramètres du modèle ($P_k^l, l \in \llbracket 0, L \rrbracket$). Chaque paramètre \overline{P}_k^l suit une loi Normales centrées sur les valeurs des paramètres de \check{P}_k perturbées. Nous avons donc,

$$\overline{P}_k^l \sim \mathcal{N}(\check{P}_k^l + \delta^l, \sigma^l) \quad (6.3.1)$$

L'objectif de la perturbation δ^l , est de générer artificiellement une erreur modèle, $\eta(t)$ et l'objectif de la variance σ^l est d'intégrer des différences inter-individuelles. Le comportement moyen de \check{i} peut être simulé par le modèle $M_k(\check{P}_k)$, alors que le comportement moyen de la population peut être simulé par le modèle $M_k(\check{P}_k + \delta)$. Le vecteur de perturbation δ ($\delta^1, \delta^2, \dots, \delta^L$), détermine le niveau de particularité de l'individu \check{i} par rapport aux autres individus de la population.

Comme dans la section 1.4, pour que l'historique de données O_H contienne Z courbes

de croissance, il faut générer Z set de paramètres $(P_k^1, P_k^2, \dots, P_k^L)$ à partir des lois Normales établies. Il faut ensuite paramétrer le modèle M_k avec ces différents set de paramètres et générer Z courbes. Une fois les Z courbes générées, nous leur ajoutons un bruit gaussien, N .

La base de données $O_H(t)$ est ensuite divisée en une base d'apprentissage ($O_{H,App}$) et une base de test ($O_{H,Test}$).

Les données de suivi en temps réel (O_{TR}), associées à l'individu \check{i} , sont générées à partir du modèle M_k paramétré avec les paramètres \check{P}_k^l . Nous ajoutons également aux données simulées un bruit gaussien, N . Ces données seront utilisées dans l'étape d'Assimilation (figure 6.3).

Ajustement du modèle mathématique sur l'historique de données

L'objectif de cette étape d'ajustement est d'utiliser l'historique de données O_H pour ajuster les L paramètres du modèle M_k . Pour cela, l'algorithme DIRECT et la procédure d'ajustement appliquée dans la section 2.3.3 et 2.5.4, peuvent être appliqués sur la base d'apprentissage $O_{H,App}$.

L'ensemble de la procédure d'apprentissage de la valeur des paramètres est gérée par le programme nommé "Optimization" (Figure 6.3). À la fin de l'apprentissage, nous obtenons les valeurs ajustées des paramètres, \widehat{P}_k . Le modèle M_k paramétré avec les valeurs de \widehat{P}_k est le modèle qui sera embarqué dans l'outil d'Assimilation.

Évaluation de la qualité de l'outil d'Assimilation de Données.

Le programme nommé "Assimilation" utilisera donc l'information fournie par le modèle $M_k(\widehat{P}_k)$ et celle contenue dans les données, O_{TR} . Il estimera l'erreur η et le bruit, κ , pour estimer l'état $\widehat{\epsilon}(t)$ (Figure 6.3).

Pour tester la qualité de l'estimation $\widehat{\epsilon}(t)$, nous la comparons à ce qui est considéré comme étant l'état réel, c'est-à-dire $\check{\epsilon}(t)$. Pour cela, un programme "Testing" calculera un indicateur d'erreur basé sur l'écart relatif existant entre $\widehat{\epsilon}(t)$ et $\check{\epsilon}(t)$.

CONCLUSION GÉNÉRALE

Nous avons construit des outils d'apprentissage statistique biomimétiques, basés sur des systèmes d'EDP, embarquant l'expression mathématique de processus biologiques clés. Dans notre approche la construction des EDP repose sur la conception d'un Avatar résumant, de manière conceptuel et très synthétique, la dynamique interne de l'animal. Le système d'EDP traduit la convection, la diffusion et l'action d'un flux d'information global évoluant dans l'Avatar. Les équations contiennent des paramètres correspondant à des facteurs biologiques. Ces paramètres sont appris sur des données et en utilisant un algorithme d'optimisation.

Lors de cette thèse, nous avons travaillé dans le contexte particulier du «Biological Small Data», c'est-à-dire un contexte dans lequel nous avons dû construire des outils capables de prédire l'évolution de phénomènes complexes, en utilisant très peu de données d'apprentissage. L'objectif de ces travaux de thèse a donc été de construire des outils de prédiction capables de réaliser de l'Assimilation de Données dans ce contexte particulier. Pour cela, il nous a fallu construire des outils intégrant des connaissances de manière parcimonieuse. Cela a permis d'obtenir des outils relativement légers, précis, ayant une capacité d'apprentissage et un potentiel d'extraction d'information très forts. De ce fait, nos outils ont une faible dépendance aux données d'apprentissage.

Dans le contexte du «Biological Small Data» nous avons montré que les outils d'apprentissage statistique que nous avons construits possèdent une précision plus élevée que celle des autres outils testés.

Cependant, ces outils biomimétiques se distinguent réellement des outils existants par leur capacité d'apprentissage et leur capacité d'extrapolation. Plus globalement, nos outils se distinguent par leur capacité à réaliser de l'Assimilation de Données d'élevage, même si la fréquence de collecte de données est basse et que la qualité des données collectées est faible.

La faible dépendance de nos outils aux données permet également de collecter moins de données pour les ajuster et donc de réduire les coûts associés aux expérimentations, à la collecte, au traitement et au stockage de données.

Les résultats issus des différentes études réalisées ont été obtenus à partir d'un

certain nombre d'hypothèses concernant la structure des modèles mathématiques. Il serait donc intéressant de poursuivre les travaux concernant l'élaboration d'une procédure de sélection qui permettrait de déterminer le meilleur modèle au regard de différents critères.

De plus, la capacité des outils d'apprentissage statistique biomimétiques à réaliser de l'Assimilation de Données n'a pas encore été validée. Les réflexions et les études à ce sujet sont donc également à poursuivre.

ANNEXE I : L'ARTICLE SYNTHÉTISANT LES TRAVAUX PRODUITS LORS DU CEMRACS 2018

CONSTRUCTION OF A STATISTICAL LEARNING TOOL BASED ON ORDINARY DIFFERENTIAL EQUATIONS TO MODEL THE DIGESTIVE BEHAVIOUR OF ROSS CHICKENS

NICOLAS BLOYET^{1,2}, HÉLÈNE FLOURENT², EMMANUEL FRÉNOT^{2,1}, MAROUAN HANDA³,
HAROLD MOUNDOYI⁴ AND THUY VO THI PHUONG⁵

Abstract. Being able to monitor and forecast farm animal performances is a strategic problem in the agronomy industry. We use a Data-Model Coupling approach to build a biomimetic Statistical Learning tool taking into account some aspects of the biological dynamics of the animal body. The objective is to build a tool which is able to assimilate data about daily feed consumption and measured performances.

The model encompasses several sub-models corresponding to compartments and permitting to mimic a kinetic process divided into several steps. Each sub-model contains parameters which can be learnt by using an optimization algorithm and data.

The goal of the first application of the model on field data was to simulate and predict the growth of chickens. An experiment was performed during 70 days to collect every day the feed consumption and the weight gain of a male and a female chickens. After the learning of the model parameters, the model shows a very good approximation of the chicken's weight evolution over time.

INTRODUCTION

Being able to monitor and forecast farm animals performances is a strategic problem in the agronomy industry: since several decades, endeavors have been made in this activity sector, with the aim to optimize the breeding-related production, as for example meat, eggs, milk, etc. (See Filipe et al. 2012).

Today, new technologies permit to monitor farm animals and collect a wide range of information as explained in Büchel and Sundrum 2014, Miekley et al. 2012, Suganthi Jemila and Suja Priyadharsini 2018, van der Tol and van der Kamp 2010 and Holman et al. 2011. Yet, those tools are still expensive and their use still infrequent. Furthermore, biological data contain high variability (Noise, heterogeneity, missing and aberrant values, etc.) (See Locke et al. 2005 Qi et al. 2006). Therefore, to treat most of biological issues we have to build precise predictive tools from few exploitable data.

Animals are complex living organisms in which intakes induce complex physico-chemical phenomena. Therefore to be able to link inputs and outputs concerning the evolution of some sensed biological factors, we need a mathematical model taking into account some aspects of the animal's body dynamics (See Renzullo et al. 2008).

¹ See-d, 6, rue Henri Becquerel - CP 101, 56038 Vannes Cedex, France

² Université Bretagne Sud, Laboratoire de Mathématiques de Bretagne Atlantique, UMR CNRS 6205, Campus de Tohannic, Vannes, France

³ Université de Picardie Jules Verne, LAMFA UMR 7352, 33 Rue St Leu, 80000 Amiens, France

⁴ Sorbonne Université, PR UPMC, Laboratoire de Biologie Intégrative des Modèles Marins, Place Georges Teissier 29680 Roscoff, France

⁵ Université Paris 13 - Galilée Institute - LAGA, 99 avenue Jean-Baptiste Clément - 93430 Villetaneuse, France

The main objective of this study is to build a tool which is able to predict outputs from inputs concerning farm animals and also able to manage the existence of underlying complex biological phenomena. In this paper we are going to present the construction of a Statistical Learning tool based on a dynamical mathematical model corresponding to a system of Ordinary Differential Equations (ODE). The first application of this Statistical Learning tool consisted to simulate and predict the growth of chickens.

Purpose

Growth simulation is a well-known problem, and some classical models - such as the Gompertz's and the Verhulst's models - are very well suited to fit some growth data as it is showed in Birch 1999 and Roush et al. 2006. Nevertheless those classical growth models do not permit to perform data assimilation. Indeed, almost no input data can be integrated in those kind of models.

On another hand, some works had already treated biological modeling issues by developing realistic and specific models as in Taghipoor 2012. But the construction of those realistic models is a costly task resulting in models containing a lot of equations and unknown parameters making them difficult to implement (Wallach et al. [2001], Bastianelli and Sauvant [1997] and Emmans [1995]).

A model-free approach is also explored, as in Guardabasso et al. 1987 for instance. Machine Learning tool based on Neural Networks were developed to simulate and predict the evolution of biological factors in Gorczyca et al. 2018. This approach, which are only based on data does not need knowledge about the link existing between the used inputs and outputs and permits to develop easier to handle models. But those models need a large amount of data to be fitted and compensate the lack of taken into account knowledge (See Valletta et al. 2017).

In the light of the existing methods for predicting biological responses, we decided to explore an approach which can be defined as being half-way between "model-free" and "full-model" approaches: we explored the Model-Data Coupling theory to construct a tool integrating biological knowledge in a mathematical model, and using data, to optimize the model parameters. Model-Data Coupling is an expanding approach but primarily developed to treat issues in meteorology (See Simmons and Hollingsworth 2002), hydrology (See Kim and Barros 2002, L. Crosson et al. 2002 and Mackay et al. 2003) and biogeochemistry (See Barrett et al. 2005, Barrett 2002, Rayner et al. 2005 and SACKS et al. 2006).

We built our model by splitting the whole complex kinetic into several parsimonious sub-biomimetic-processes performed by a combination of successive compartments. All those compartments compute their respective behaviors, and then exchange messages, assimilable to fluxes. To do that we built several Ordinary Differential Equations integrating the mathematical expressions of biological phenomena (storage, saturation, etc.).

Scope

In the scope of this study, we focus on meat poultry and specifically two biological strains :

- Ross 308 (Aviagen)
- Cobb 500 (Cobb-Vantress)

Some informations about those strains are publicly available (Aviagen and Cobb-Vantress). We know the average daily feed consumption and the average daily weight gain of an average female and an average male of each strain, during 70 days. Therefore, those available data correspond to smoothed experimental data. We decided to add a Gaussian noise to those data in order to obtain experimental-like data.

As a first step, we considered the growth (mass gained each day) and the loss (ratio of food unmetabolized). Therefore we have a database concerning four individuals (2 females and 2 males) and containing 70 lignes and 4 columns.

In this paper we are going to introduce in Section 1 the applied methodology. Then in Section 2 we will present the developed mathematical model and the results obtained by using this model to simulate the growth of chickens. In this section, the parameters of the model will be learnt by using the previously presented database.

1. METHODOLOGY

1.1. Biological-like function

The basis computational unit used in the following work is what we defined here as *organ-like compartment*. We define it as a tuple :

- $input \in \mathbb{R}^{in}, in \in \mathbb{N}$
- $output \in \mathbb{R}^{out}, out \in \mathbb{N}$
- $function : \mathbb{R}^{in} \rightarrow \mathbb{R}^{out}$
- $state \in \mathbb{R}^{states}, states \in \mathbb{N}$

where *input* and *output* can be treated as vectors of fixed dimensions *in* and *out*, which are each relative to a flux of a particular kind. There is a mapping between those vectors through *function*, which corresponds to a *biological-like function*. This function models a biological phenomenon in a synthesized way. Finally, *state* is a vector allowing these computational units to have a state, in other words a kind of memory, which can take part in the expression of the said function.

The dimensions of these input / output vectors therefore correspond to information exchanges in the form of messages between the organ-like compartments, the nature of which can be, for example, nutrients, drugs active principles, etc. These functions can for example try to mimic a phenomenon of fixing, convection, diffusion, etc. Figure 1 shows how such a biological-like function can be summarized as a computational unit.

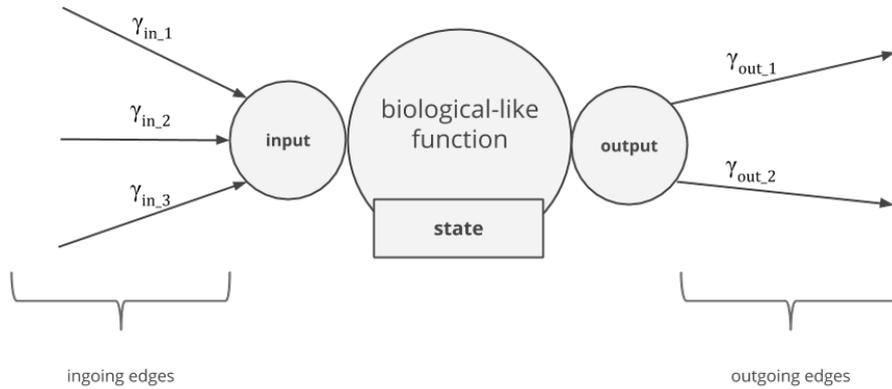


FIGURE 1. A biological-like function

1.2. Functions network

These deliberately simple functions are not complex enough to model an organism. However, by making the contact between them, it becomes possible to model a more complex structure. To do this, a communication is established from the output vector of a biomimetic function to the input vector of the following function (Figure 2). Since these functions have states, the network as a whole can be likened to a finite state machine.

1.3. Optimization steps

This network acting like a model containing several parameters to optimize. We can distinguish two groups of parameters.

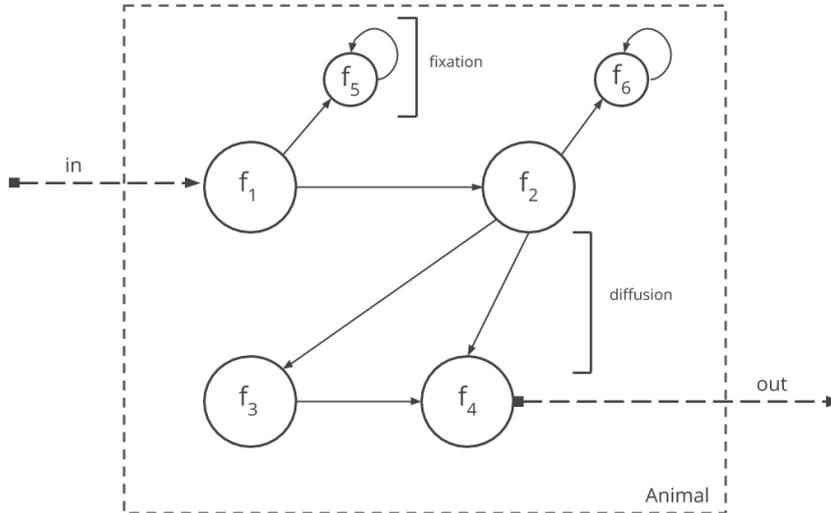


FIGURE 2. Example of a graph-model

1.3.1. *Biological-like functions (organ-level)*

The first step of optimization is at the level of biological-like functions. Most of the phenomena that we are trying to model involve parameters corresponding for example to the part of information exchanged between the compartments, saturation thresholds, etc. These are optimized in order to fit as much as possible the data.

To determine these parameters we use an non-linear optimization algorithm implemented in Matlab (fmincon). This algorithm allows us to find the set of parameters minimizing the model error towards data used.

1.3.2. *Network structure (organism level)*

The nature of the links between these different functions also involves parameters (weighting link, pre-processing of the vector), and is therefore also subject to optimization.

In our study, these parameters essentially correspond to the amount of information exchanged between the compartments. Given the small number of computational units involved at this stage, and so a reasonable dimensionality of the problem, it is possible to use a conventional optimization by just including them in the above set of parameters, but potentially on a complex structure, a heuristic search method (genetic algorithm for example) could be preferred.

We could also favor separated processes of optimization, where each organ would be adjusted independently, before adjusting the whole network.

1.4. **Main goals**

The goal of the following work is to build a Statistical Learning tool that can be used to simulate the food consumption and mass gain of poultry, to obtain results close to what one would obtain with a growth law fitting, with the singular difference that this association of function will not only be a function of time but also of the food consumed, and thus would have an increased biological correspondence.

It therefore seeks to replicate stem growth data, but adjusting a network of biomimetic functions instead of a Gompertz logistic function. The latter will, however, be used as a reference point for the metrics considered (coefficient of determination R^2 and root mean square error $RMSE$).

2. CONSTRUCTION OF THE MODEL

Growth requires substrate and complex physico-chemical phenomena to convert substrate into dry weight. Weight formation can be influenced by many factors such that substrate provision, assimilation, digestibility, storage, daily losses, environmental parameters, etc. Modeling of all those processes would lead to the construction of a huge model. Therefore we do not want to reproduce the digestion kinetic but just integrate into the mathematical model information exchanges, delay, fixation, accumulation and saturation effects.

In the model, we use the Gompertz growth equation as in M. C. Kathleen [2017] :

$$W'(t) = A \ln \left(\frac{W_f}{W(t)} \right) W(t), \quad (1)$$

where W is the weight which varies in time t , W_f is the maximum weight that can be reached with the available nutrients and A is a constant.

2.1. Assumptions

In our model, we assume that the organs participating in the food digestion are the stomach, the small intestine and the large intestine. We built our model under the following assumptions:

(A1) The ingested food is stored in the compartment named “crop” before moving to the compartment assimilable to the stomach.

(A2) The flux rates are constant between compartments.

(A3) There exist some transmission delays between some compartments.

(A4) The small intestine emptying dynamic starts after some time τ .

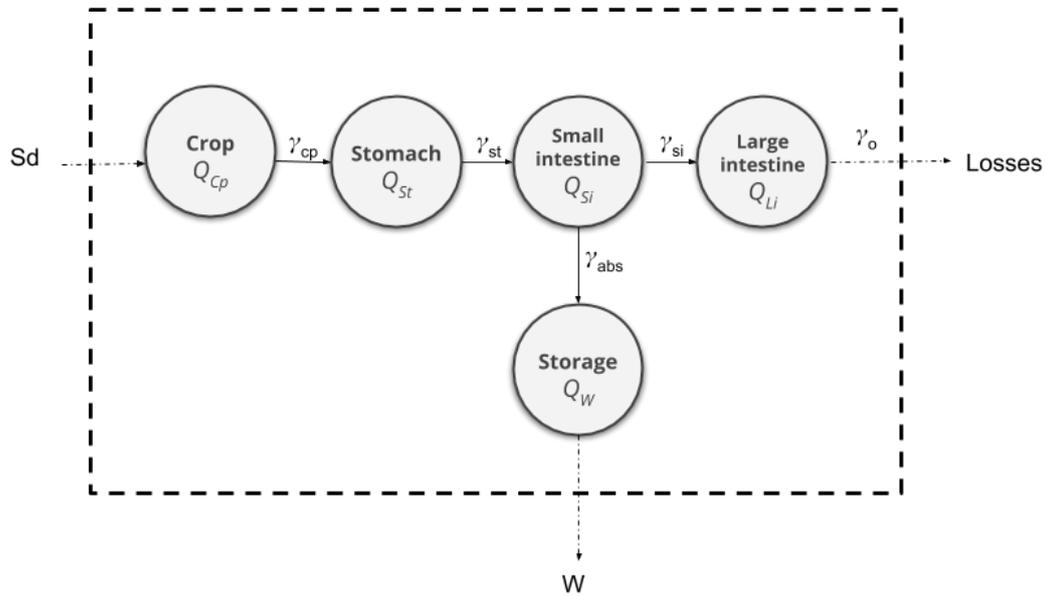


FIGURE 3. Digestive system graph

Following those assumptions, we can draw the graph associated to our model and synthesizing the digestive system of a chicken as in Figure 3, where

- S_d is the amount of food consumed each day;

- Q_{cp} (resp. $Q_{st}, Q_{si}, Q_{Li}, Q_w$) is the information inside the crop (resp. stomach, small intestine, large intestine, storage compartment);
- γ_{cp} , (resp. $\gamma_{st}, \gamma_{si}, \gamma_{abs}$) is the rate of nutrients transferred from crop to stomach (rep. from stomach to small intestine, from small intestine to large intestine, from small intestine to storage compartment).
- γ_o , corresponds to the non-metabolized part of the nutrients.

The dynamical system of the digestive functions at each day d are synthesized by the ODEs:

$$Q'_{cp,d} = -\gamma_{cp}Q_{cp,d} \quad (2)$$

$$Q'_{st,d} = \gamma_{cp}(1 - \exp(-\beta_1 t))Q_{cp,d} - \gamma_{st}Q_{st,d} \quad (3)$$

$$Q'_{si,d} = \begin{cases} \gamma_{st}Q_{st,d}, & t \leq \tau \\ \gamma_{st}Q_{st,d} - (\gamma_{abs} + \gamma_{si})Q_{si,d}, & t > \tau \end{cases} \quad (4)$$

$$Q'_{Li,d} = \gamma_{si}Q_{si,d} - \gamma_{Li}Q_{Li,d}, \quad (5)$$

with the initial conditions

$$Q_{cp,d}(0) = S_d; \quad (6)$$

$$Q_{st,d}(0) = 0; \quad (7)$$

$$Q_{si,d}(0) = 0; \quad (8)$$

$$Q_{Li,d}(0) = 0. \quad (9)$$

Note here that there is some delay in the procedure of information transmission from crop to stomach exponentially with rate β_1 . The storage compartment uses a proportion of nutrients going out from small intestine to gain weight and all the information left from the day before is kept to be used in the present day, that gives

$$Q'_{w,d} = \gamma_{abs}Q_{si,d} - \alpha Q_{w,d}; \quad (10)$$

$$Q_{w,d}(0) = Q_{w,d-1}. \quad (11)$$

The quantity of the growth machinery is proportional to the amount of information transmitted to the storage compartment. However, there are some losses during the information transmission corresponding to the part of the nutrient used via physical activities, reproduction needs and the non-metabolizable nutrients. Those losses are proportional to the weight of the animal and nutrients produced from metabolism,

$$\begin{cases} W'_d = \kappa Q_{w,d} - \left(W_{max} \frac{Q_{w,d}}{K + Q_{w,d}} \right) \frac{1}{K} W_d \\ W_d(0) = 0 \end{cases} .$$

The effectiveness of the machinery κ exponentially decays in time with the delay μ ,

$$\begin{cases} \kappa' = -\mu\kappa \\ \mu' = \beta_2(\mu_{max} - \mu) \end{cases} . \quad (12)$$

Equation (12) follows that

$$\kappa(t) = \kappa_0 \exp \left[-\mu_{max} \left(t - \frac{1 - \exp(-\beta_2 t)}{\beta_2} \right) \right]. \quad (13)$$

As a result, the final weight is the total gain up to considering day,

$$W_f = W_0 + \sum_{d=1}^{N_d} W_d(t_f), \quad (14)$$

where N_d is the number of days and t_f is the time at the end of a digestive cycle of a day.

2.2. Global optimization

The parameters contained into the different biological-like functions are initially unknown. Those parameters which can be learned by using data and optimization algorithms confer to the model a learning capability. Therefore we used the function *directL* (Johnson [2008]) existing in R (R Core Team [2018]), to learn the model parameters. To find the values of the parameters permitting the model to fit the data, this function minimize an objective function. In this application the objective function corresponds to the mean squared difference between the initial and the predicted curves (15),

$$f_{Obj} = \frac{1}{n} \sum_{i=1}^n \left(PredCurve(i) - InitCurve(i) \right)^2 \quad (15)$$

where $PredCurve(i)$ and $InitCurve(i)$ respectively correspond to the i^{th} point of the predicted curve and the i^{th} point of the initial curve, and n is the total number of point, that is 70 days.

In this application we fixed τ and we fitted β_1 , γ_{st} , γ_{abs} , γ_{si} , γ_o , μ_{max} , β_2 , V_{max} , K and γ_{cp} .

2.3. Simulation results

After the fitting we parametrized the model with the obtained values of the parameters (Table 1).

The dynamics of the model throughout one day and throughout two days (Figure 4) were simulated with the Female Model, that is the model simulated the growth of the female chicken. We also simulated the growth of the animals throughout the whole studied period (Figure 5).

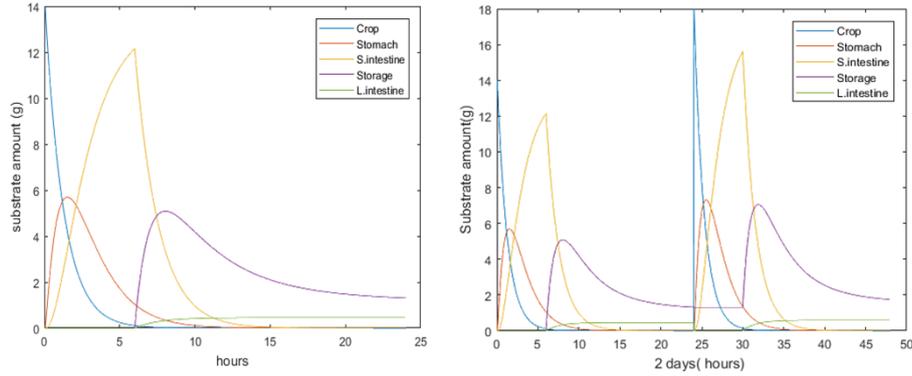


FIGURE 4. Dynamics of the model throughout one day (left) and throughout two days (right).

Parameters	Female Model	Male Model
β_1	0.405	0.402
γ_{st}	0.601	0.650
γ_{abs}	0.650	1.00
γ_{si}	0.020	0.014
γ_o	0.424	0.512
μ_{max}	0.008	0.008
β_2	0.016	0.048
V_{max}	0.051	0.050
K	24.70	24.90
γ_{cp}	1.654	0.971

TABLE 1. Parameter values of the Male and Female Models.

3. DISCUSSIONS

3.1. Knowledge feedback

First of all, considering for the same strain males on one side, and females on the other, it is interesting to note that some parameters seem to be stable, while only a part of them vary significantly among poultry's sex. In that way, the adjusted model give us back some information, that could be confronted to some business expertise (here zootechnicians).

Indeed, it appears that organism's difference of behavior is concentrated on parameters :

- β_2 , relative to the effectiveness of the organism ¹
- γ_{abs} , the link's strength between the small intestine and the storage compartment
- γ_{si} , the link's strength between the small and the large intestine
- γ_{cp} , the link's strength between the crop and the stomach

In particular, we observe that male's organism seem to stay highly efficient longer than female's (a fact that we observe in ground truth), characterized by a slightly better β_2 on the male. At the same time, we observe that on female's, the amount of material sent from the small intestine to the big intestine (which is going to be lost) is higher than male's, where a bigger part is sent to the storage compartment.

Those observations would need to be confirmed on other strains on to be confronted to a zootechnician expertise, but is by itself a first step of knowledge feedback.

3.2. Accuracy

Figure 5 shows that the built model fits the data with a goodness of fit close to our target. We chose to compare our model with two types of models : a Gompertz growth (17) law usually used to simulate growth phenomena, and a two order Polynomial Model (16) took as a basis with the same number of parameters. The latter one doesn't have a biological likelyhood, so the first goal is to perform better.

$$W(t) = p_0 + p_1t + p_2t^2 \tag{16}$$

$$W(t) = g_0 \exp\left(-g_1 \exp(-g_2t)\right) \tag{17}$$

We fitted the parameters of the Polynomial Model (p_0, p_1, p_2) and the ones of the Gompertz's Model (g_0, g_1, g_2) on the data concerning the female and the male chickens.

We calculated for each model the mean of the R^2 and the $RMSE$ values associated the Female and the Male Models. Figures 5, 6 and 7 and the results contained in Table 2 show that the accuracy of those different models is globally equivalent. Therefore we built a model satisfying in terms of accuracy. But the real advantage of the Biomimetic Model is its capability to integrate biological knowledge and assimilate input data.

Model	$RMSE$	R^2
Gompertz's Model	94.2	0.997
Order 2 Polynomial Model	147.4	0.992
Biomimetic Model	102.1	0.996

TABLE 2. Comparison of the models in terms of R^2 and $RMSE$ values.

¹its ability to metabolize a high ratio of ingered food with respect of time

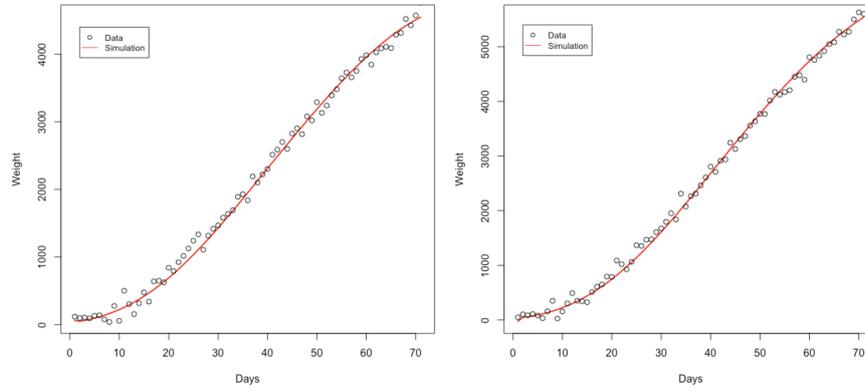


FIGURE 5. Comparison between the initial curve (Data) and the curve simulated via the Biomimetic Model (Simulation) for the Female (left) and the Male Model (right).

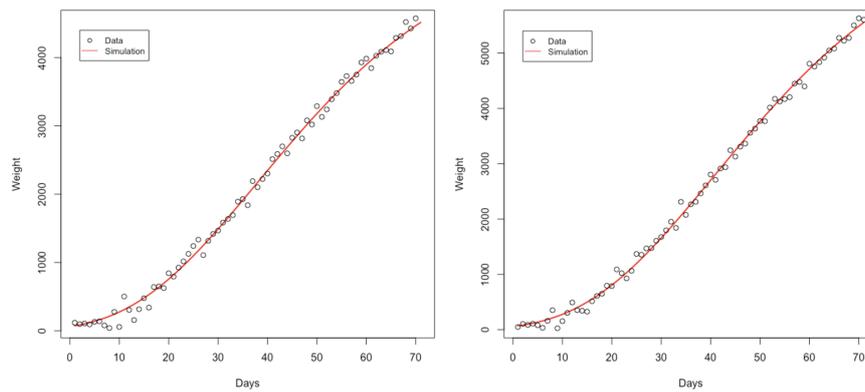


FIGURE 6. Comparison between the initial curve (Data) and the curve simulated via the Gompertz's Model (Simulation) for the Female (left) and the Male Model (right).

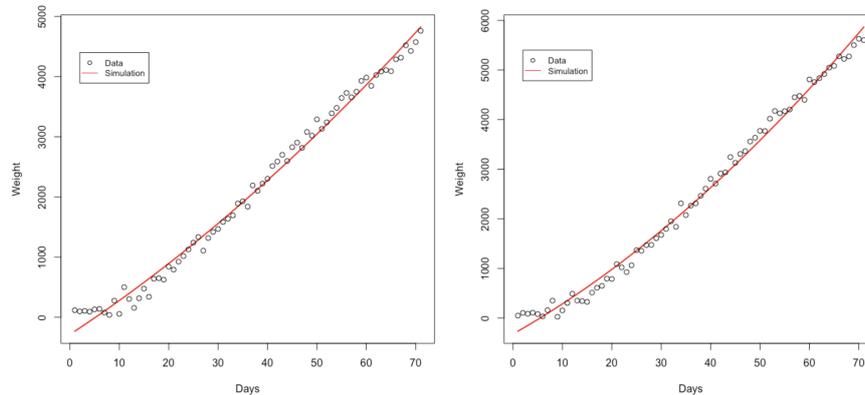


FIGURE 7. Comparison between the initial curve (Data) and the curve simulated via the Polynomial Model (Simulation) for the Female (left) and the Male Model (right).

4. PERSPECTIVES AND CONCLUSION

4.1. Perspectives

The developed approach allows us to stay very modular with regards to the nature of the simulated model, so it could adapt to a higher level of information for a few adaptations.

4.1.1. Fluxes multiplication

In the scope of this work, the model was only trained to deal with simple weightings data, and so not in a context where it could out-perform classic growth laws. The actual added value to those models would appear as soon as more fine data would be available.

Indeed, the objective at mid-term is to exploit an eventual higher level of knowledge (theoretical or measured) about modeled phenomena, helping to define in a more precise way what happens inside the organism, typically with data centered on an organ's input and output. Unlike classic growth laws based on temporal axis (Gompertz, Verhulst for example), this kind of model could handle data made of several dimensions (time, ingested proteins, drugs, etc.), and among everything, could pretend to follow a biological likelihood.

4.1.2. Model reinterpretation

At a longer term, and in a context where the previous hypothesis about new data would be validated, the goal would be to optimize automatically the weight parameters between biological-like functions.

Knowing the organism's structure to a sufficient level of details, and having a sufficient number of measures, it would be possible to refine the knowledge concerning the nature of the connections between these different virtual organs, by set of parameters optimization (genetical optimization for example).

As those parameters do have a relationship with actual phenomenon due to their biological likelihood, it would be interesting to confront the final set of optimized parameters to a business expertise.

4.2. Conclusion

We explored a Data-Model Coupling approach and built a biomimetic dynamic tool containing Ordinary Differential Equations. Those Differential Equations contain parameters conferring a learning capability to the developed tool. Therefore, our tool is a Statistical Learning tool able to learn the value of parameters from field data.

An optimization process permits to learn the parameters and then parametrize the mathematical model by integrating input data concerning the feed consumption of the chickens. After the learning step we obtained a model able to simulate the growth of chickens with an accuracy similar to the one associated to the classic growth models. But the real advantage of our model in comparison with those classical logistical models, is its capability to integrate input variables. Furthermore, the structure of the model is create, manage and flexible. Hence, the structure of the model can be adapted if we are able to collect more detailed data.

The number of parameters to determine is less important than in the realistic models, but it is still high and only one structure of the model was tested. Therefore model selection methods could be used to determine the optimal structure of the model in terms of accuracy and the number of parameters to learn.

Considering the number of parameters to determine and the number of data, a problem of overfitting can be expected. We suppose that the integrated knowledge and the imposed profile of the growth curve could permit to avoid overfitting, but this aspect has to be verified on a Test Dataset.

Furthermore, it could be a major improvement to have a higher quantity of data to fit the model. In particular, more detailed data concerning the processes occurring in the different organs of the digestive system of the chickens could permit to fit the parameters compartment by compartment.

REFERENCES

- Aviagen. Aviagen, ross 308. URL <http://en.aviagen.com/brands/ross/products/ross-308>.
- Damian Barrett, M Hill, Lindsay Hutley, Jason Beringer, Johnny H. Xu, Garry Cook, John Carter, and Richard J. Williams. Prospects for improving savanna biophysical models by using multiple-constraints model-data assimilation methods. *Australian Journal of Botany*, 53(7), 01 2005. .
- Damian J. Barrett. Steady state turnover time of carbon in the australian terrestrial biosphere. *Global Biogeochemical Cycles*, 16(4):55–1–55–21, 2002. . URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2002GB001860>.
- D. Bastianelli and D. Sauvant. Modelling the mechanisms of pig growth. *Livestock Production Science*, 1997.
- S. Büchel and A. Sundrum. Short communication: Decrease in rumination time as an indicator of the onset of calving. *Journal of Dairy Science*, 97(5):3120 – 3127, 2014. ISSN 0022-0302. . URL <http://www.sciencedirect.com/science/article/pii/S0022030214001684>.
- Colin PD Birch. A new generalized logistic sigmoid growth equation compared with the richards growth equation. *Annals of Botany*, 83(6):713–723, 1999.
- Cobb-Vantress. Cobb-vantress, cobb 500. URL <https://www.cobb-vantress.com/products/cobb500>.
- G. C. Emmans. Problems in applying models in practice. *Publication-European Association For Animal Production*, 78:223–223, 1995.
- Patricia Filipe, Carlos Braumann, and Carlos Roquete. Multiphasic individual growth models in random environments. *Methodology and Computing in Applied Probability*, 14:49–56, 03 2012. .
- Michael T. Gorczyca, Hugo Fernando Maia Milan, Alex Sandro Campos Maia, and Kifle G. Gebremedhin. Machine learning algorithms to predict core, skin, and hair-coat temperatures of piglets. *Computers and Electronics in Agriculture*, 151:286 – 294, 2018. ISSN 0168-1699. . URL <http://www.sciencedirect.com/science/article/pii/S0168169918303612>.
- V. Guardabasso, D. Rodbard, and P. J. Munson. A model-free approach to estimation of relative potency in dose-response curve analysis. *American Journal of Physiology-Endocrinology and Metabolism*, 252(3):E357–E364, 1987. . URL <https://doi.org/10.1152/ajpendo.1987.252.3.E357>. PMID: 3826362.
- A. Holman, J. Thompson, J. E. Routly, J. Cameron, D. N. Jones, D. Grove-White, R. F. Smith, and H. Dobson. Comparison of oestrus detection methods in dairy cattle. *Veterinary Record*, 169(2):47–47, 2011. ISSN 0042-4900. . URL <https://veterinaryrecord.bmj.com/content/169/2/47>.
- Steven G. Johnson. The nlopt nonlinear-optimization package. 2008.
- Gwangseob Kim and Ana P Barros. Space–time characterization of soil moisture from passive microwave remotely sensed imagery and ancillary data. *Remote Sensing of Environment*, 81(2):393 – 403, 2002. ISSN 0034-4257. . URL <http://www.sciencedirect.com/science/article/pii/S0034425702000147>.
- William L. Crosson, Charles A Laymon, Ramarao Inguva, and Marius P. Schamschula. Assimilating remote sensing data in a surface flux–soil moisture model. *Hydrological Processes*, 16:1645–1662, 06 2002. .
- James C. W. Locke, Andrew J. Millar, and Matthew S. Turner. Modelling genetic networks with noisy and varied experimental data: the circadian clock in arabidopsis thaliana. *Journal of theoretical biology*, 234 3: 383–93, 2005.
- T. Even M. C. Kathleen. The use of gompertz models in growth analyses, and new gompertz-model approach: An addition to the unified-richards family. *PLoS One*, 12(6):e0178691., 2017.
- D.Scott Mackay, Sudeep Samanta, Ramakrishna R. Nemani, and Lawrence E. Band. Multi-objective parameter estimation for simulating canopy transpiration in forested watersheds. *Journal of Hydrology*, 277 (3):230 – 247, 2003. ISSN 0022-1694. . URL <http://www.sciencedirect.com/science/article/pii/S0022169403001306>.
- Bettina Miekley, Imke Traulsen, and Joachim Krieter. Detection of mastitis and lameness in dairy cows using wavelet analysis. *Livestock Science*, 148(3):227 – 236, 2012. ISSN 1871-1413. . URL <http://www.sciencedirect.com/science/article/pii/S187114131200220X>.
- Yanjun Qi, Ziv Bar-Joseph, and Judith Klein-Seetharaman. Evaluation of different biological data and computational classification methods for use in protein interaction prediction. *Proteins: Structure, Function, and Bioinformatics*, 63(3):490–500, 2006. . URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/prot.20865>.

- R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2018. URL <https://www.R-project.org/>.
- P. J. Rayner, M. Scholze, W. Knorr, T. Kaminski, R. Giering, and H. Widmann. Two decades of terrestrial carbon fluxes from a carbon cycle data assimilation system (ccdas). *Global Biogeochemical Cycles*, 19(2), 2005. . URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2004GB002254>.
- Luigi J. Renzullo, Damian J. Barrett, Alan S. Marks, Michael J. Hill, Juan P. Guerschman, Qiaozhen Mu, and Steve W. Running. Multi-sensor model-data fusion for estimation of hydrologic and energy flux parameters. *Remote Sensing of Environment*, 112(4):1306 – 1319, 2008. ISSN 0034-4257. . URL <http://www.sciencedirect.com/science/article/pii/S0034425707003227>. Remote Sensing Data Assimilation Special Issue.
- WB Roush, WA Dozier 3rd, and SL Branton. Comparison of gompertz and neural network models of broiler growth. *Poultry Science*, 85(4):794–797, 2006.
- WILLIAM J. SACKS, DAVID S. SCHIMEL, RUSSELL K. MONSON, and BOBBY H. BRASWELL. Model-data synthesis of diurnal and seasonal co2 fluxes at niwot ridge, colorado. *Global Change Biology*, 12(2):240–259, 2006. . URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2486.2005.01059.x>.
- A. J. Simmons and A. Hollingsworth. Some aspects of the improvement in skill of numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society*, 128(580):647–677, 2002. . URL <https://rmets.onlinelibrary.wiley.com/doi/abs/10.1256/003590002321042135>.
- J Suganthi Jemila and S Suja Priyadharsini. A sensor-based forage monitoring of grazing cattle in dairy farming. *International Journal on Smart Sensing and Intelligent Systems*, 11:1–9, 01 2018. .
- Masoomah Taghipoor. *Modélisation du transport, de la dégradation et de l'absorption des aliments dans l'intestin grêle*. PhD thesis, Université François Rabelais-Tours, 2012.
- John Joseph Valletta, Colin Torney, Michael Kings, Alex Thornton, and Joah Madden. Applications of machine learning in animal behaviour studies. *Animal Behaviour*, 124:203 – 220, 2017. ISSN 0003-3472. . URL <http://www.sciencedirect.com/science/article/pii/S0003347216303360>.
- Rik van der Tol and Arjen van der Kamp. Time series analysis of live weight as health indicator. In *Proceedings of the first North American conference precision dairy management*, pages 230–231, 2010.
- D. Wallach, B. Goffinet, J. E. Bergez, P. Debaeke, D. Leenhardt, and J. N. Aubertot. Parameter estimation for crop models. *Agronomy journal*, 93(4):757–766, 2001.

BIBLIOGRAPHIE

- ADAMS, R. A. (1975). « Sobolev Spaces, volume 65 ». In : *Pure and applied mathematics*.
- AILLIOT, P., E. FRÉNOT et V. MONBET (2006). « Long term object drift in the ocean with tide and wind. » In : *SIAM Journal on Multiscale Modeling and Simulation : A SIAM Interdisciplinary Journal* 5 (2), p. 514 -531. URL : <https://hal.archives-ouvertes.fr/hal-00129093>.
- ALFONSI, Aurélien, Eric CANCES, Gabriel TURINICI, Barbara DI VENTURA et Wilhelm HUISINGA (2005). « Adaptive simulation of hybrid stochastic and deterministic models for biochemical systems ». In : *ESAIM : proceedings*. T. 14. EDP Sciences, p. 1-13.
- ALINHAC, S. et P. GÉRARD (2012). *Opérateurs pseudo-différentiels et théorème de Nash-Moser*. SAVOIRS ACTUELS. EDP Sciences. ISBN : 9782759802821. URL : <https://books.google.fr/books?id=oAVBSvJz6NYC>.
- ALLAIRE, G. (2005). *Analyse numérique et optimisation : une introduction à la modélisation mathématique et à la simulation numérique*. Mathématiques appliquées. Editions de l'Ecole polytechnique. ISBN : 9782730212557. URL : <https://books.google.fr/books?id=vReEuE4margC>.
- ANDERSON, Laurence A, Allan R ROBINSON et Carlos J LOZANO (2000). « Physical and biological modeling in the Gulf Stream region :: I. Data assimilation methodology ». In : *Deep Sea Research Part I : Oceanographic Research Papers* 47.10, p. 1787-1827.
- ARCHER, JA, EC RICHARDSON, RM HERD et PF ARTHUR (1999). « Potential for selection to improve efficiency of feed use in beef cattle : a review ». In : *Australian Journal of Agricultural Research* 50.2, p. 147-162.
- AREL, I., D. C. ROSE et T. P. KARNOWSKI (2010). « Deep machine learning-a new frontier in artificial intelligence research ». In : *IEEE computational intelligence magazine* 5.4, p. 13-18.
- AUROUX, Didier et Jacques BLUM (2005). « Back and forth nudging algorithm for data assimilation problems ». In : *Comptes Rendus Mathématique* 340.12, p. 873-878.

-
- BALDWIN, R. L. (1995). *Modeling ruminant digestion and metabolism*. Springer Science & Business Media.
- BARRETT, D., M. HILL, L. HUTLEY, J. BERINGER, J. H. XU, G. COOK, J. CARTER et R. J. WILLIAMS (2005). « Prospects for improving savanna biophysical models by using multiple-constraints model-data assimilation methods ». In : *Australian Journal of Botany* 53(7). DOI : 10.1071/BT04139.
- BARRETT, D. J. (2002). « Steady state turnover time of carbon in the Australian terrestrial biosphere ». In : *Global Biogeochemical Cycles* 16.4, p. 55-1-55-21. DOI : 10.1029/2002GB001860. eprint : <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2002GB001860>. URL : <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2002GB001860>.
- BARTOCCI, Ezio, Luca BORTOLUSSI et Guido SANGUINETTI (2014). « Data-Driven Statistical Learning of Temporal Logic Properties ». In : *Formal Modeling and Analysis of Timed Systems*, p. 23.
- BASARAB, J., M. PRICE, J. L. AALHUS, E. OKINE, W. M. SNELLING et K. L. LYLE (2003a). « Residual Feed intake and body composition in young growing cattle ». In : *Canadian Journal of Animal Science* 83, p. 189-204. DOI : 10.4141/A02-065.
- BASARAB, JA, MA PRICE, JL AALHUS, EK OKINE, WM SNELLING et KL LYLE (2003b). « Residual feed intake and body composition in young growing cattle ». In : *Canadian Journal of Animal Science* 83.2, p. 189-204.
- BASTIANELLI, D. et D. SAUVANT (1997). « Modelling the mechanisms of pig growth. » In : *Livestock Production Science*.
- BASTIANELLI, D., D. SAUVANT et A. RERAT (1996). « Mathematical modeling of digestion and nutrient absorption in pigs. » In : *Journal of animal science*.
- BENGIO, Y. (2012). « Deep learning of representations for unsupervised and transfer learning ». In : *Proceedings of ICML workshop on unsupervised and transfer learning*, p. 17-36.
- BOUSQUET, Olivier, Ulrike von LUXBURG et Gunnar RÄTSCH (2011). *Advanced Lectures on Machine Learning : ML Summer Schools 2003, Canberra, Australia, February 2-14, 2003, Tübingen, Germany, August 4-16, 2003, Revised Lectures*. T. 3176. Springer.
- BRAZDIL, Pavel et Peter CLARK (1990). « Learning from imperfect data ». In : *Machine Learning, Meta-Reasoning and Logics*. Springer, p. 207-232.

-
- BRÉZIS, H., P.G. CIARLET et J.L. LIONS (1999). *Analyse fonctionnelle : théorie et applications*. Collection Mathématiques appliquées pour la maîtrise. Dunod. ISBN : 9782100043149. URL : <https://books.google.fr/books?id=aykKPQAACAAJ>.
- BRUN-LAFLEUR, L., E. CUTULLIC, P. FAVERDIN, L. DELABY et C. DISENHAUS (2013). « An individual reproduction model sensitive to milk yield and body condition in Holstein dairy cows ». In : *Animal : an international journal of animal bioscience* 7, p. 1-12. DOI : 10.1017/S1751731113000335.
- BÜCHEL, S. et A. SUNDRUM (2014). « Short communication : Decrease in rumination time as an indicator of the onset of calving ». In : *Journal of Dairy Science* 97.5, p. 3120 -3127. ISSN : 0022-0302. DOI : <https://doi.org/10.3168/jds.2013-7613>. URL : <http://www.sciencedirect.com/science/article/pii/S0022030214001684>.
- BUYSE, J., B. GEYPENS, R. D. MALHEIROS, V. M. MORAES, Q. SWENNEN et E. DECUYPERE (2004). « Assessment of age-related glucose oxidation rates of broiler chickens by using stable isotopes ». In : *Life sciences* 75.18, p. 2245-2255.
- CARTER, RG, JM GABLONSKY, A PATRICK, Carl Tim KELLEY et OJ ESLINGER (2001). « Algorithms for noisy problems in gas transmission pipeline optimization ». In : *Optimization and engineering* 2.2, p. 139-157.
- CASTILLO, A., E. KEBREAB, D. BEEVER et J. FRANCE (2000). « A review of efficiency of nitrogen utilisation in lactating dairy cows and its relationship with environmental pollution ». In : *Journal of Animal and Feed Sciences* 9.1, p. 1-32. ISSN : 1230-1388. DOI : 10.22358/jafs/68025/2000. URL : <http://dx.doi.org/10.22358/jafs/68025/2000>.
- CHASSAGNOLE, C., N. NOISOMMIT-RIZZI, J. SCHMID, K. MAUCH et M. REUSS (2002). « Dynamic modeling of the central carbon metabolism of Escherichia coli ». In : *Biotechnology and bioengineering* 79.1, p. 53-73.
- CHAVE, AD, M ARROTT, C FARCAS, E FARCAS, I KRUEGER, M MEISINGER, JA ORCUTT, FL VERNON, C PEACH, Oscar SCHOFIELD *et al.* (2009). « Cyberinfrastructure for the US Ocean Observatories Initiative : Enabling interactive observation in the ocean ». In : p. 1-10.
- CHEN, Ting, Hongyu L HE et George M CHURCH (1999). « Modeling gene expression with differential equations ». In : *Biocomputing'99*. World Scientific, p. 29-40.
- CHO, Y. et L. K. SAUL (2009). « Kernel methods for deep learning ». In : *Advances in neural information processing systems*, p. 342-350.

-
- COMTE, Eloïse (déc. 2017). « Groundwater pollution from agricultural activities : coupling hydrogeological and economical approaches ». Theses. Université de La Rochelle. URL : <https://hal.archives-ouvertes.fr/tel-01669150>.
- COURANT, R. et D. HILBERT (2008). *Methods of Mathematical Physics : Partial Differential Equations*. Wiley Classics Library. Wiley. ISBN : 9783527617241. URL : <https://books.google.fr/books?id=fcZV4ohrerwC>.
- COURANT, R., K. FRIEDRICHS et H. LEWY (1928). « Über die partiellen Differenzgleichungen der mathematischen Physik ». In : *Mathematische annalen* 100.1, p. 32-74.
- CROSSON, W. L., C. A. LAYMON, R. INGUVA et M. P. SCHAMSCHULA (2002). « Assimilating remote sensing data in a surface fluxsoil moisture model ». In : *Hydrological Processes* 16, p. 1645-1662. DOI : 10.1002/hyp.1051.
- DA ROS, S., G. COLUSSO, T. A. WESCHENFELDER, L. de MARSILLAC TERRA, F. de CASTILHOS, M. L. CORAZZA et M. SCHWAAB (2013). « A comparison among stochastic optimization algorithms for parameter estimation of biochemical kinetic models ». In : *Applied Soft Computing* 13.5, p. 2205 -2214. ISSN : 1568-4946. DOI : <https://doi.org/10.1016/j.asoc.2013.01.019>. URL : <http://www.sciencedirect.com/science/article/pii/S1568494613000380>.
- DAVID, C. et P. GOSSELET (2015). *Equations aux dérivées partielles-2e éd. : Cours et exercices corrigés*. Dunod.
- DI MENZA, L. (2009). *Analyse numérique des équations aux dérivées partielles*. Casini.
- DINIZ, Wellison JS, Kamila O ROSA, Luiz L COUTINHO, Polyana C TIZIOTO, Priscila SN de OLIVEIRA, Marcela M de SOUZA, Amália S CHAVES, Dante PD LANNA, Gerson B MOURÃO et Luciana CA REGITANO (2018). « KCNJ11 gene expression is associated to feed consumption and growth traits in Nelore beef cattle ». In : *Agri Gene* 9, p. 1-4.
- DOMINGOS, P. (2012). « A Few Useful Things to Know About Machine Learning ». In : *Commun. ACM* 55, p. 7887. DOI : 10.1145/2347736.2347755.
- DOMINGOS, Pedro et Matthew RICHARDSON (2007). « 1 markov logic : A unifying framework for statistical relational learning ». In : *Statistical Relational Learning*, p. 339.

-
- DOWD, Michael (2007). « Bayesian statistical data assimilation for ecosystem models using Markov Chain Monte Carlo ». In : *Journal of Marine Systems* 68.3-4, p. 439-456.
- DUMAS, A., J. DIJKSTRA et J. FRANCE (2008). « Mathematical modelling in animal nutrition : A centenary review ». In : *Journal of Agricultural Science* 146 (2008) 2 146. DOI : 10.1017/S0021859608007703.
- DUMPALA, S. H., R. CHAKRABORTY et S. K. KOPPARAPU (2017). *k-FFNN : A priori knowledge infused Feed-forward Neural Networks*. arXiv : 1704.07055 [cs.LG].
- DUNSON, David B. (2018). « Statistics in the big data era : Failures of the machine ». In : *Statistics Probability Letters* 136. The role of Statistics in the era of big data, p. 4-9. ISSN : 0167-7152. DOI : <https://doi.org/10.1016/j.spl.2018.02.028>. URL : <http://www.sciencedirect.com/science/article/pii/S0167715218300737>.
- ENRIGHT, W. (1989). « The Numerical Analysis of Ordinary Differential Equations : Runge Kutta and General Linear Methods ». In : *SIAM Review* 31.4, p. 693-693. DOI : 10.1137/1031147. eprint : <https://doi.org/10.1137/1031147>. URL : <https://doi.org/10.1137/1031147>.
- ESMEIJER, Jop, Tom BAKKER, Merel OOMS, Bas KOTTERINK et Sponsor Ministerie van ECONOMISCHE ZAKEN (2015). « Data-driven innovation in agriculture : case study for the OECD KBC2-programme ». In : *TNO Report TNO*, R10154.
- EVANS, L.C. (2010). *Partial Differential Equations, Second Edition*. Graduate studies in mathematics. American Mathematical Society. ISBN : 9780821849743. URL : https://books.google.fr/books?id=Xnu0o_EJrCQC.
- EVERITT, B. S., S. LANDAU, M. LEESE et D. STAHL (2011). « Cluster Analysis. —John Wiley & Sons ». In : *Ltd., New York*, p. 330.
- FAULKNER, A et K CEBUL (2014). « Agriculture Gets Smart : The Rise of Data and Robotics, Cleantech Agriculture Report ». In : *Cleantech Group*.
- FINKEL, D. E. (2003). *DIRECT Optimization Algorithm*. North Carolina State University.
- FINKEL, Daniel E et CT KELLEY (2004). « Convergence analysis of the DIRECT algorithm ». In : *Optimization Online* 14.2, p. 1-10.
- FLEGG, J. A., H. M. BYRNE, M. B. FLEGG et S. D. L. MCELWAIN (2012). « Wound healing angiogenesis : The clinical implications of a simple mathematical model ». In : *Journal of Theoretical Biology* 300, p. 309-316. ISSN : 0022-5193. DOI : <https://doi.org/10.1016/j.jtbi.2012.01.043>. URL : <http://www.sciencedirect.com/science/article/pii/S002251931200063X>.

-
- FOKOUÉ, Ernest (2014). « A Taxonomy of Big Data for Optimal Predictive Machine Learning and Data Mining ». In : *Serdica Journal of Computing* 8.2, p. 111-136.
- FOOD et Agriculture ORGANIZATION (2009). *How to Feed the World in 2050 ? Executive Summary*. Rapp. tech.
- FRÉNOT, E. (2017). « A PDE-like Toy-Model of Territory Working ». In : *Understanding Interactions in Complex Systems - Toward a Science of Interaction*. Understanding Interactions in Complex Systems - Toward a Science of Interaction. Cambridge Scholar Publishing, p. 37-47. URL : <https://hal.archives-ouvertes.fr/hal-00817522>.
- FRITSCH, Stefan, Frauke GUENTHER et following earlier work by MARC SULING (2012). *neuralnet : Training of neural networks*. R package version 1.32. URL : <http://CRAN.R-project.org/package=neuralnet>.
- GAO, Chao, Han WANG, Ensheng WENG, S LAKSHMIVARAHAN, Yanfen ZHANG et Yiqi LUO (2011). « Assimilation of multiple data sets with the ensemble Kalman filter to improve forecasts of forest carbon dynamics ». In : *Ecological Applications* 21.5, p. 1461-1473.
- GETTO, Philipp et Marcus WAURICK (2016). « A differential equation with state-dependent delay from cell population biology ». In : *Journal of Differential Equations* 260.7, p. 6176 -6200. ISSN : 0022-0396. DOI : <https://doi.org/10.1016/j.jde.2015.12.038>. URL : <http://www.sciencedirect.com/science/article/pii/S0022039615007032>.
- GHAHRAMANI, Z. (2003). « Unsupervised learning ». In : p. 72-112.
- GLOROT, X., A. BORDES et Y. BENGIO (2011). « Domain adaptation for large-scale sentiment classification : A deep learning approach ». In : *Proceedings of the 28th international conference on machine learning (ICML-11)*, p. 513-520.
- GOMPERTZ, B. (1825). « XXIV. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. In a letter to Francis Baily, Esq. FRS &c ». In : *Philosophical transactions of the Royal Society of London* 115, p. 513-583.
- GONZE, D, J HALLOY et A GOLDBETER (2003). « Modèles déterministes et stochastiques pour les rythmes circadiens Deterministic and stochastic models for circadian rhythms ». In : *Pathologie Biologie* 51.4, p. 227-230.
- GORCZYCA, M. T., H. F. M. MILAN, A. S. C. MAIA et K. G. GEBREMEDHIN (2018). « Machine learning algorithms to predict core, skin, and hair-coat temperatures of pi-

- glets ». In : *Computers and Electronics in Agriculture* 151, p. 286 -294. ISSN : 0168-1699. DOI : <https://doi.org/10.1016/j.compag.2018.06.028>. URL : <http://www.sciencedirect.com/science/article/pii/S0168169918303612>.
- GREGG, Watson W, Marjorie AM FRIEDRICHS, Allan R ROBINSON, Kenneth A ROSE, Reiner SCHLITZER, Keith R THOMPSON et Scott C DONEY (2009). « Skill assessment in ocean biological data assimilation ». In : *Journal of Marine Systems* 76.1-2, p. 16-33.
- GUZMÁN-CRUZ, R., R. CASTANEDA-MIRANDA, J. GARCIA-ESCALANTE, L. SOLIS-SÁNCHEZ, D. ALANIZ-LUMBRERAS, J. MENDOZA-JASSO, A. LARA-HERRERA, G. ORNELAS-VARGAS, E. GONZALEZ-RAMIREZ et R. MONTOYA-ZAMORA (2011). « Evolutionary Algorithms in Modelling of Biosystems ». In : ISBN : 978-953-307-171-8. DOI : 10.5772/16231.
- HASSOUN, M. (1995). *Fundamentals of artificial neural networks*. MIT press.
- HASTIE, T., R. TIBSHIRANI et J. FRIEDMAN (2009). « The elements of statistical learning : data mining, inference, and prediction, Springer Series in Statistics ». In :
- HE, Jian, Layne T WATSON, Naren RAMAKRISHNAN, Clifford A SHAFFER, Alex VERS-TAK, Jing JIANG, Kyung BAE et William H TRANTER (2002). « Dynamic data structures for a direct search algorithm ». In : *Computational Optimization and Applications* 23.1, p. 5-25.
- HERD, RM, VH ODDY et EC RICHARDSON (2004). « Biological basis for variation in residual feed intake in beef cattle. 1. Review of potential mechanisms ». In : *Australian Journal of Experimental Agriculture* 44.5, p. 423-430.
- HOLMAN, A., J. THOMPSON, J. E. ROUTLY, J. CAMERON, D. N. JONES, D. GROVE-WHITE, R. F. SMITH et H. DOBSON (2011). « Comparison of oestrus detection methods in dairy cattle ». In : *Veterinary Record* 169.2, p. 47-47. ISSN : 0042-4900. DOI : 10.1136/vr.d2344. eprint : <https://veterinaryrecord.bmj.com/content/169/2/47.full.pdf>. URL : <https://veterinaryrecord.bmj.com/content/169/2/47>.
- HOUTEKAMER, Peter L et Herschel L MITCHELL (1998). « Data assimilation using an ensemble Kalman filter technique ». In : *Monthly Weather Review* 126.3, p. 796-811.
- HUBBARD, T. et A. REINHARDT (1998). « Using neural networks for prediction of the subcellular location of proteins ». In : *Nucleic Acids Research* 26.9, p. 2230-2236. ISSN : 0305-1048. DOI : 10.1093/nar/26.9.2230. eprint : <http://oup.prod>.

sis.lan/nar/article-pdf/26/9/2230/9471729/26-9-2230.pdf. URL : <https://dx.doi.org/10.1093/nar/26.9.2230>.

INGALLS, B. (2013). « Mathematical Modelling in Systems Biology : An Introduction ». In :

INIESTA, R, D STAHL et P MCGUFFIN (2016). « Machine learning, statistical learning and the future of biological research in psychiatry ». In : *Psychological medicine* 46.12, p. 2455-2465.

IP, R. H. L., L. M. ANG, K. P. SENG, J. C. B. et J. E. PRATLEY (2018). « Big data and machine learning for crop protection ». In : *Computers and Electronics in Agriculture* 151, p. 376 -383. ISSN : 0168-1699. DOI : <https://doi.org/10.1016/j.compag.2018.06.008>. URL : <http://www.sciencedirect.com/science/article/pii/S0168169917314588>.

ISHII, K *et al.* (2014). « Big data analysis in medicine, agriculture and environmental sciences. » In : *Seibutsu-kogaku Kaishi* 92.2, p. 92-93.

JAMES, Gareth, Daniela WITTEN, Trevor HASTIE et Robert TIBSHIRANI (2013). *An introduction to statistical learning*. T. 112. Springer.

JAYARAMAN, Prem, Ali YAVARI, Dimitrios GEORGAKOPOULOS, Ahsan MORSHED et Arkady ZASLAVSKY (2016). « Internet of things platform for smart farming : Experiences and lessons learnt ». In : *Sensors* 16.11, p. 1884.

JEMILA, J. S. et S. S. PRIYADHARSINI (2018). « A Sensor-Based Forage Monitoring of Grazing Cattle in Dairy Farming ». In : *International Journal on Smart Sensing and Intelligent Systems* 11, p. 1-9. DOI : 10.21307/ijssis-2018-014.

JOHNSON, S. G. (2008). « The NLOpt nonlinear-optimization package ».

JONES, Donald R, Cary D PERTTUNEN et Bruce E STUCKMAN (1993). « Lipschitzian optimization without the Lipschitz constant ». In : *Journal of optimization Theory and Applications* 79.1, p. 157-181.

KALNAY, Eugenia (2003). *Atmospheric modeling, data assimilation and predictability*. Cambridge university press.

KIM, G. et A. P. BARROS (2002). « Spacetime characterization of soil moisture from passive microwave remotely sensed imagery and ancillary data ». In : *Remote Sensing of Environment* 81.2, p. 393 -403. ISSN : 0034-4257. DOI : [https://doi.org/10.1016/S0034-4257\(02\)00014-7](https://doi.org/10.1016/S0034-4257(02)00014-7). URL : <http://www.sciencedirect.com/science/article/pii/S0034425702000147>.

-
- KITANO, H. (avr. 2002). « Systems Biology : A Brief Overview ». In : *Science (New York, N.Y.)* 295, p. 1662-4. DOI : 10.1126/science.1069492.
- KOTSIANTIS, Sotiris B, I ZAHARAKIS et P PINTELAS (2007). « Supervised machine learning : A review of classification techniques ». In : *Emerging artificial intelligence applications in computer engineering* 160, p. 3-24.
- KUTTER, Thomas, Silja TIEMANN, Rosemarie SIEBERT et Spyros FOUNTAS (2011). « The role of communication and co-operation in the adoption of precision farming ». In : *Precision Agriculture* 12.1, p. 2-17.
- LADÉAU, Shannon L, Gregory E GLASS, N Thompson HOBBS, Andrew LATIMER et Richard S OSTFELD (2011). « Data–model fusion to better understand emerging pathogens and improve infectious disease forecasting ». In : *Ecological Applications* 21.5, p. 1443-1460.
- LADYENSKAJA, O. A., V. A. SOLONNIKOV et N. N. URALCEVA (1968). *Linear and quasi-linear equations of parabolic type*. Translations of mathematical monographs. Providence, RI : American Mathematical Society. URL : <https://cds.cern.ch/record/2623252>.
- LECUN, Y., Y. BENGIO et G. HINTON (2015). « Deep learning ». In : *nature* 521.7553, p. 436.
- LEE, C. H. (2004). « From knowledge-ignorant to knowledge-rich modeling : a new speech research paradigm for next generation automatic speech recognition ». In :
- LENSINK, BL et B GUERMONPREZ (2003). « L'impact de la génétique des animaux et leur alimentation sur la qualité de la viande ». In : *L'Information diététique* 1, p. 41-47.
- LEWIS, John M, Sivaramakrishnan LAKSHMIVARAHAN et Sudarshan DHALL (2006). *Dynamic data assimilation : a least squares approach*. T. 13. Cambridge University Press.
- LGUENSAT, Redouane, Pierre TANDEO, Pierre AILLIOT, Manuel PULIDO et Ronan FABLET (2017). « The analog data assimilation ». In : *Monthly Weather Review* 145.10, p. 4093-4107.
- LGUENSAT, Redouane, Phi Huynh VIET, Miao SUN, Ge CHEN, Tian FENGLIN, Bertrand CHAPRON et Ronan FABLET (2019). « Data-driven Interpolation of Sea Level Anomalies using Analog Data Assimilation ». In : *Remote Sensing* 11.7, p. 858.

-
- LI, XiuFeng, ShouHe CHEN, LeiFeng GUO *et al.* (2014). « Technological innovation of agricultural information service in the age of big data. » In : *Journal of Agricultural Science and Technology (Beijing)* 16.4, p. 10-15.
- LOCKE, J. C. W., A. J. MILLAR et M. S. TURNER (2005). « Modelling genetic networks with noisy and varied experimental data : the circadian clock in *Arabidopsis thaliana*. » In : *Journal of theoretical biology* 234 3, p. 383-93.
- LU, H., R. SETIONO et H. LIU (1996). « Effective data mining using neural networks ». In : *IEEE transactions on knowledge and data engineering* 8.6, p. 957-961.
- LU, H., K. N. PLATANIOTIS et A. VENETSANOPOULOS (2013). *Multilinear subspace learning : dimensionality reduction of multidimensional data*. Chapman et Hall/CRC.
- LUO, Yiqi, Kiona OGLE, Colin TUCKER, Shenfeng FEI, Chao GAO, Shannon LADEAU, James S CLARK et David S SCHIMEL (2011). « Ecological forecasting and data assimilation in a data-rich era ». In : *Ecological Applications* 21.5, p. 1429-1442.
- MA, C., H. H. ZHANG et X. WANG (2014). « Machine learning for Big Data analytics in plants ». In : *Trends in Plant Science* 19.12, p. 798 -808. ISSN : 1360-1385. DOI : <https://doi.org/10.1016/j.tplants.2014.08.004>. URL : <http://www.sciencedirect.com/science/article/pii/S1360138514002192>.
- MACH, J. et Z. KRISTKOVA (2010). « Modelling The Cattle Breeding Production in the Czech Republic ». In : *AGRIS on-line Papers in Economics and Informatics* 2.
- MACKAY, D. S., S. SAMANTA, R. R. NEMANI et L. E. BAND (2003). « Multi-objective parameter estimation for simulating canopy transpiration in forested watersheds ». In : *Journal of Hydrology* 277.3, p. 230 -247. ISSN : 0022-1694. DOI : [https://doi.org/10.1016/S0022-1694\(03\)00130-6](https://doi.org/10.1016/S0022-1694(03)00130-6). URL : <http://www.sciencedirect.com/science/article/pii/S0022169403001306>.
- MARIA, Anu (1997). « Introduction to modeling and simulation ». In : 29.1, p. 7-13.
- MARTIN, O. et D. SAUVANT (2007). « Dynamic model of the lactating dairy cow metabolism ». In : *animal* 1.8, p. 1143-1166. DOI : [10.1017/S1751731107000377](https://doi.org/10.1017/S1751731107000377).
- (2010a). « A teleonomic model describing performance (body, milk and intake) during growth and over repeated reproductive cycles throughout the lifespan of dairy cattle. 1. Trajectories of life function priorities and genetic scaling. » In : *Animal*.
- (2010b). « A teleonomic model describing performance (body, milk and intake) during growth and over repeated reproductive cycles throughout the lifespan of dairy cattle. 2. Voluntary intake and energy partitioning ». In : *Animal* 4.12, 2048ñ2056. DOI : [10.1017/S1751731110001369](https://doi.org/10.1017/S1751731110001369).

-
- McPHEE, M. (2009a). « Mathematical modelling in agricultural systems : A case study of modelling fat deposition in beef cattle for research and industry ». In :
- McPHEE, MJ (2009b). « Mathematical modelling in agricultural systems : A case study of modelling fat deposition in beef cattle for research and industry ». In : *Proceeding in 18th World IMACS.-MODSIM Congress, Cairns, Australia, 13-17 July*, p. 59.
- MERTENS, Christopher, W Kent TOBISKA, David BOUWER, Brian KRESS, Michael WILTBERGER, Stanley SOLOMON et John MURRAY (2009). « Development of nowcast of atmospheric ionizing radiation for aviation safety (nairas) model ». In : p. 3633.
- MIEKLEY, Bettina, Imke TRAUlsen et Joachim KRIETER (2012). « Detection of mastitis and lameness in dairy cows using wavelet analysis ». In : *Livestock Science* 148.3, p. 227 -236. ISSN : 1871-1413. DOI : <https://doi.org/10.1016/j.livsci.2012.06.010>. URL : <http://www.sciencedirect.com/science/article/pii/S187114131200220X>.
- MORIK, K., P. BROCKHAUSEN et T. JOACHIMS (1999). *Combining statistical learning with a knowledge-based approach : a case study in intensive care monitoring*. Rapp. tech. Technical Report, SFB 475 : Komplexitätsreduktion in Multivariaten
- MOURET, J. B. (2016). « Micro-data learning : The other end of the spectrum ». In : *arXiv preprint arXiv :1610.00946*.
- NADARAYA, E. A. (1964). « On estimating regression ». In : *Theory of Probability & Its Applications* 9.1, p. 141-142.
- NAHM, KH (2002). « Efficient feed nutrient utilization to reduce pollutants in poultry and swine manure ». In : *Critical Reviews in Environmental Science and Technology* 32.1, p. 1-16.
- NESETRILOVA, H. (2005). « Multiphasic growth models for cattle ». In : *Czech Journal of Animal Science* 50, p. 347-354. DOI : 10.17221/4176-CJAS.
- NICHOLAS, Paul J et David L ALDERSON JR (2014). *Method for optimal transmitter placement in wireless mesh networks*. US Patent 8,654,672.
- NICHOLAS, PE (2014). « A dividing rectangles algorithm for stochastic simulation optimization ». In : *Proc. INFORMS Comput. Soc. Conf. T. 14*, p. 47-61.
- NIU, Shuli, Yiqi LUO, Michael C DIETZE, Trevor F KEENAN, Zheng SHI, Jianwei LI et F Stuart CHAPIN III (2014). « The role of data assimilation in predictive ecology ». In : *Ecosphere* 5.5, p. 1-16.
- NKRUMAH, J. D., J. BASARAB, Z. WANG, C. LI, M. PRICE, E. OKINE, D. H. CREWS et S. S. MOORE (2007). « Genetic and phenotypic relationships of feed intake and

-
- measures of efficiency with growth and carcass merit of beef cattle ». In : *Journal of animal science* 85, p. 2711-20. DOI : 10.2527/jas.2006-767.
- O'GRADY, Michael J et Gregory MP O'HARE (2017). « Modelling the smart farm ». In : *Information processing in agriculture* 4.3, p. 179-187.
- O'REGAN, D. (2013). *Existence Theory for Nonlinear Ordinary Differential Equations. Mathematics and Its Applications*. Springer Netherlands. ISBN : 9789401715171. URL : <https://books.google.fr/books?id=ijjpCAAAQBAJ>.
- PERASSO, Antoine (2009). « Identifiabilité de paramètres pour des systèmes décrits par des équations aux dérivées partielles. Application à la dynamique des populations ». Thèse de doct. Université Paris Sud-Paris XI.
- PHUA, Y., T. RIBEIRO, S. TOURET et K. INOUE (2017). « Learning Logic Program Representation for Delayed Systems With Limited Training Data ». In :
- PUILLET, L., O. MARTIN, D. SAUVANT et M. TICHIT (2011). « Introducing efficiency into the analysis of individual lifetime performance variability : a key to assess herd management ». In : *animal* 5.1, p. 123-133. DOI : 10.1017/S175173111000162X. URL : <https://hal.archives-ouvertes.fr/hal-01137029>.
- QI, Y., Z. BAR-JOSEPH et J. KLEIN-SEETHARAMAN (2006). « Evaluation of different biological data and computational classification methods for use in protein interaction prediction ». In : *Proteins : Structure, Function, and Bioinformatics* 63.3, p. 490-500. DOI : 10.1002/prot.20865. eprint : <https://onlinelibrary.wiley.com/doi/pdf/10.1002/prot.20865>. URL : <https://onlinelibrary.wiley.com/doi/abs/10.1002/prot.20865>.
- R CORE TEAM (2014). *R : A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL : <http://www.R-project.org/>.
- RAYNER, P. J., M. SCHOLZE, W. KNORR, T. KAMINSKI, R. GIERING et H. WIDMANN (2005). « Two decades of terrestrial carbon fluxes from a carbon cycle data assimilation system (CCDAS) ». In : *Global Biogeochemical Cycles* 19.2. DOI : 10.1029/2004GB002254. eprint : <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2004GB002254>. URL : <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2004GB002254>.
- RENZULLO, L. J., D. J. BARRETT, A. S. MARKS, M. J. HILL, J. P. GUERSCHMAN, Q. MU et S. W. RUNNING (2008). « Multi-sensor model-data fusion for estimation of hydrologic and energy flux parameters ». In : *Remote Sensing of Environment* 112.4.

-
- Remote Sensing Data Assimilation Special Issue, p. 1306 -1319. ISSN : 0034-4257. DOI : <https://doi.org/10.1016/j.rse.2007.06.022>. URL : <http://www.sciencedirect.com/science/article/pii/S0034425707003227>.
- RIZZI, M., M. BALTES, U. THEOBALD et M. REUSS (1997). « In vivo analysis of metabolic dynamics in *Saccharomyces cerevisiae* : II. Mathematical model ». In : *Biotechnology and bioengineering* 55.4, p. 592-608.
- ROBERTSON, T. B. (1916). « Experimental studies on growth II. The normal growth of the white mouse ». In : *Journal of Biological Chemistry* 24.3, p. 363-383.
- (1923). *The chemical basis of growth and senescence*. JB Lippincott Company.
- ROMÁN-ROMÁN, P. et F. TORRES-RUIZ (2012). « Modelling logistic growth by a new diffusion process : Application to biological systems ». In : *Biosystems* 110.1, p. 9-21.
- ROUSSEAU, A. et M. NODET (2013). « Modélisation mathématique et assimilation de données pour les sciences de l'environnement ». In : *Bulletin de l'APMED* 505, p. 467-472.
- RYU, Minwoo, Jaeseok YUN, Ting MIAO, Il-Yeup AHN, Sung-Chan CHOI et Jaeho KIM (2015). « Design and implementation of a connected farm for smart farming system ». In : *2015 IEEE SENSORS*. IEEE, p. 1-4.
- SABER, Nima et John M SHAW (2008). « Rapid and robust phase behaviour stability analysis using global optimization ». In : *Fluid Phase Equilibria* 264.1-2, p. 137-146.
- SACKS, W. J., D. S. SCHIMEL, R. K. MONSON et B. H. BRASWELL (2006). « Model-data synthesis of diurnal and seasonal CO₂ fluxes at Niwot Ridge, Colorado ». In : *Global Change Biology* 12.2, p. 240-259. DOI : 10.1111/j.1365-2486.2005.01059.x. eprint : <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-2486.2005.01059.x>. URL : <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2486.2005.01059.x>.
- SACKS, W. J., D. S. SCHIMEL et R. K. MONSON (2007). « Coupling between carbon cycling and climate in a high-elevation, subalpine forest : a model-data fusion analysis ». en. In : *Oecologia* 151.1, p. 54-68. ISSN : 0029-8549, 1432-1939. DOI : 10.1007/s00442-006-0565-2. URL : <http://link.springer.com/10.1007/s00442-006-0565-2> (visité le 22/11/2018).
- SAKOMURA, N. K., F. A. LONGO, E. O. OVIEDO-RONDON, C. BOA-VIAGEM et A. FERRAUDO (2005). « Modeling energy utilization and growth parameter description for broiler chickens ». In : *Poultry Science* 84.9, p. 1363-1369.

-
- SCHULZE, ED, S LUYSSAERT, P CIAIS, A FREIBAUER, IA JANSSENS, JF SOUSSANA, P SMITH, J GRACE, I LEVIN, B THIRUCHITTAMPALAM *et al.* (2009). « Importance of methane and nitrous oxide for Europe's terrestrial greenhouse-gas balance ». In : *Nature Geoscience* 2.12, p. 842.
- SHAVLIK, J., L. HUNTER et D. SEARLS (1995). « Introduction ». In : *Machine Learning* 21.1, p. 5-9. ISSN : 1573-0565. DOI : 10.1007/BF00993376. URL : <https://doi.org/10.1007/BF00993376>.
- SIMMONS, A. J. et A. HOLLINGSWORTH (2002). « Some aspects of the improvement in skill of numerical weather prediction ». In : *Quarterly Journal of the Royal Meteorological Society* 128.580, p. 647-677. DOI : 10.1256/003590002321042135. URL : <https://rmets.onlinelibrary.wiley.com/doi/abs/10.1256/003590002321042135>.
- SOETAERT, K., T. PETZOLDT et R. Woodrow SETZER (2010). « Solving Differential Equations in R : Package deSolve ». In : *Journal of Statistical Software* 33.9, p. 1-25. ISSN : 1548-7660. DOI : 10.18637/jss.v033.i09. URL : <http://www.jstatsoft.org/v33/i09>.
- SOLTANI, A. (2012). *Modeling physiology of crop development, growth and yield*. CABI.
- TAN, A. C. et D. GILBERT (2003). « An empirical comparison of supervised machine learning techniques in bioinformatics ». In : *Proceedings of the First Asia-Pacific bioinformatics conference on Bioinformatics 2003-Volume 19*. Australian Computer Society, Inc., p. 219-222.
- TAYLOR, M.E. (1996). *Partial Differential Equations I : Basic Theory*. Applied mathematical sciences. Springer. ISBN : 9780387946535. URL : <https://books.google.fr/books?id=0xtSyLjsphkC>.
- TESCHL, G. (2012). *Ordinary Differential Equations and Dynamical Systems*. Graduate studies in mathematics. American Mathematical Society. ISBN : 9780821883280. URL : <https://books.google.fr/books?id=FZ0CAQAAQBAJ>.
- TOL, R. van der et A. van der KAMP (2010). « Time series analysis of live weight as health indicator ». In : *Proceedings of the first North American conference precision dairy management*, p. 230-231.
- TOMLIN, Claire J et Jeffrey D AXELROD (2007). « Biology by numbers : mathematical modelling in developmental biology ». In : *Nature reviews genetics* 8.5, p. 331.
- TYSON, J., K. CHEN et B. NOVAK (2003). « Sniffers, buzzers, toggles and blinkers : dynamics of regulatory and signaling pathways in the cell ». In : *Current opinion in cell biology* 15.2, p. 221-231.

-
- VALLETTA, J. J., C. TORNEY, M. KINGS, A. THORNTON et J. MADDEN (2017). « Applications of machine learning in animal behaviour studies ». In : *Animal Behaviour* 124, p. 203 -220. ISSN : 0003-3472. DOI : <https://doi.org/10.1016/j.anbehav.2016.12.005>. URL : <http://www.sciencedirect.com/science/article/pii/S0003347216303360>.
- VAPNIK, Vladimir N (1995). *The nature of statistical learning theory*. Rapp. tech.
- VÁZQUEZ-CRUZ, M. A., A. ESPINOSA-CALDERÓN, A. R. JIMÉNEZ-SÁNCHEZ et R. GUZMÁN-CRUZ (2014). « Mathematical Modeling of Biosystems ». In : *Biosystems Engineering : Biofactories for Food Production in the Century XXI*. Cham : Springer International Publishing, p. 51-76. DOI : 10.1007/978-3-319-03880-3_2. URL : https://doi.org/10.1007/978-3-319-03880-3_2.
- VERHULST, P. F. (1838). « Notice sur la loi que la population suit dans son accroissement ». In : *Corresp. Math. Phys.* 10, p. 113-126. URL : <https://ci.nii.ac.jp/naid/10015246307/en/>.
- WALTER, Achim, Robert FINGER, Robert HUBER et Nina BUCHMANN (2017). « Smart farming is key to developing sustainable agriculture ». In : *Proceedings of the National Academy of Sciences* 114.24, p. 6148-6150. ISSN : 0027-8424. DOI : 10.1073/pnas.1707462114. eprint : <https://www.pnas.org/content/114/24/6148.full.pdf>. URL : <https://www.pnas.org/content/114/24/6148>.
- WALTER, Eric (2014). *Identifiability of parametric models*. Elsevier.
- WANG, Bin, Xiaolei ZOU et Jiang ZHU (2000). « Data assimilation and its applications ». In : *Proceedings of the National Academy of Sciences* 97.21, p. 11143-11144.
- WANG, L., H. ZHANG, K. C. L. WONG, H. LIU et P. SHI (2010). « Physiological-model-constrained noninvasive reconstruction of volumetric myocardial transmembrane potentials ». In : *IEEE Transactions on Biomedical Engineering* 57.2, p. 296-315.
- WANG, Ye-Yi, Alex ACERO, Milind MAHAJAN et John LEE (2006). « Combining statistical and knowledge-based spoken language understanding in conditional models ». In : p. 882-889.
- WATSON, G. S. (1964). « Smooth regression analysis ». In : *Sankhyā : The Indian Journal of Statistics, Series A*, p. 359-372.
- WAYNE, M. et W. S. DAVID (2007). « Computational systems biology in drug discovery and development : methods and applications ». In : *Drug Discovery Today* 12.7, p. 295 -303. ISSN : 1359-6446. DOI : <https://doi.org/10.1016/j.drudis>.

2007.02.013. URL : <http://www.sciencedirect.com/science/article/pii/S1359644607000943>.

- WEISSTEIN, E. W. (2014). « Courant-friedrichs-lewy condition ». In : *Wolfram MathWorld—A Wolfram Web Resource*.
- WILLIAMS, Mathew, Paul A SCHWARZ, Beverly E LAW, James IRVINE et Meredith R KURPIUS (2005). « An improved analysis of forest carbon dynamics using data assimilation ». In : *Global change biology* 11.1, p. 89-105.
- WILSON, P. et J. DAINTY (1999). « Modelling in nutrition : An introduction ». In : *Proceedings of the Nutrition Society* 58.1, p. 133-138. DOI : 10.1079/PNS19990018.
- WINSOR, C. P. (1932). « The Gompertz Curve as a Growth Curve ». In : *Proceedings of the National Academy of Sciences* 18.1, p. 1-8. ISSN : 0027-8424. DOI : 10.1073/pnas.18.1.1. eprint : <https://www.pnas.org/content/18/1/1.full.pdf>. URL : <https://www.pnas.org/content/18/1/1>.
- WOLFERT, S., D. GOENSE et C. A. G. SØRENSEN (2014). « A Future Internet Collaboration Platform for Safe and Healthy Food from Farm to Fork ». In : *2014 Annual SRII Global Conference*, p. 266-273. DOI : 10.1109/SRII.2014.47.
- WOLFERT, Sjaak, Lan GE, Cor VERDOUW et Marc-Jeroen BOGAARDT (2017). « Big data in smart farming—a review ». In : *Agricultural Systems* 153, p. 69-80.
- ZHU, Hong et David B BOGY (2002). « DIRECT algorithm and its application to slider air-bearing surface optimization ». In : *IEEE Transactions on Magnetics* 38.5, p. 2168-2170.
- ZÚÑIGA, E. C. T., I. L. L. CRUZ et A. R. GARCÍA (2014). « Parameter estimation for crop growth model using evolutionary and bio-inspired algorithms ». In : *Applied Soft Computing* 23, p. 474 -482. ISSN : 1568-4946. DOI : <https://doi.org/10.1016/j.asoc.2014.06.023>. URL : <http://www.sciencedirect.com/science/article/pii/S156849461400297X>.

Titre : Construction et analyse d'outils de Statistical Learning biomimétiques basés sur des systèmes d'Equations aux Dérivées Partielles pour l'Assimilation de Données d'élevage.

Mots clés : Statistical Learning, EDP, Prédiction, Assimilation de Données, Couplage Modèle-Données, Modélisation mathématique, Intelligence Artificielle

Résumé : Ces travaux de thèse ont montré que les Equations aux Dérivées Partielles (EDP) sont des outils très intéressants pour construire d'efficaces Intelligences Artificielles, et tout particulièrement pour traiter des problématiques liées à l'élevage de précision. L'Assimilation de Données semble être l'outil qui permettra de piloter les élevages de demain. Or pour réaliser de l'Assimilation de Données d'élevage il est nécessaire de disposer d'outils ayant une très forte capacité d'apprentissage et qui soient capables de prédire l'évolution de variables biologiques. Pour atteindre ces objectifs nous avons construit des outils de Statistical Learning biomimétiques basés sur des systèmes d'EDP

embarquant l'expression mathématique de processus biologiques clés. Ces équations permettent d'embarquer une modélisation synthétique de la dynamique interne de l'animal. Elles contiennent des paramètres, associés à des facteurs biologiques, qui peuvent être appris sur des données.

Au cours de ces travaux de thèse nous avons montré que ces outils se distinguent des outils existants par leur capacité à s'ajuster sur très peu de données, sans surajuster les données d'apprentissage. Les outils construits se distinguent également par leur capacité d'extrapolation et leur capacité à intégrer des informations au cours du temps.

Title : Construction and analysis of Statistical Learning tools, based on systems of Partial Differential Equations, for livestock Data Assimilation

Keywords : Statistical Learning, PDE, Forecasting, Data Assimilation, Data- Model Coupling, Biological Mathematical Modeling, Artificial Intelligence

Abstract: This thesis demonstrates that Partial Differential Equations can be powerful to build efficient Artificial Intelligence, especially in the Smart Farming sector. Data Assimilation seems to be the tool that will permit to oversee the breeding farms of tomorrow. Yet, to perform Data Assimilation for livestock data, it is necessary to have tools having a very strong learning capability from data et able to predict the evolution of biological variables. To achieve these objectives, we built biomimetic Statistical Learning tools, based on PDE systems embarking the mathematical expression of key

biological processes. These equations permit to involve synthetic modeling of the internal dynamics of the animal. They contain parameters, corresponding to biological-like factors which can be learnt from collected data. Those thesis work put demonstrated that those tools distinguished themselves from the existing ones by their capability to be fitted on few data without overfitted the training data. They also distinguished themselves by their capability to perform extrapolation and to integrate information over time.