



Calcul intensif pour l'évaluation de la vulnérabilité en utilisant une approche d'Ingénierie Dirigée par les Modèles : application à la vulnérabilité des prairies au changement climatique sous contraintes de plans d'expériences

Romain Lardy

► To cite this version:

Romain Lardy. Calcul intensif pour l'évaluation de la vulnérabilité en utilisant une approche d'Ingénierie Dirigée par les Modèles : application à la vulnérabilité des prairies au changement climatique sous contraintes de plans d'expériences. Ingénierie assistée par ordinateur. Université Blaise Pascal - Clermont-Ferrand II, 2013. Français. NNT : 2013CLF22352 . tel-01913555

HAL Id: tel-01913555

<https://theses.hal.science/tel-01913555>

Submitted on 6 Nov 2018

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.



Université Blaise Pascal

**École Doctorale des Sciences pour l'Ingénieur
Thèse**

pour l'obtention du grade de

DOCTEUR D'UNIVERSITÉ

Discipline : Informatique

Présentée par

Romain LARDY

le 13/05/2013

**Calcul intensif pour l'évaluation de la vulnérabilité en utilisant
une approche d'Ingénierie Dirigée par les Modèles.**

**Application à la vulnérabilité des prairies au changement climatique sous
contraintes de plans d'expériences**

Composition du jury :

Rapporteur :	Éric RAMAT, Professeur, Université du Littoral Côte d'Opale Jean-Christophe SOULIÉ, Chargé de recherche, CIRAD
Examineur :	Michel DURU, Directeur de recherche, INRA Bruno BACHELET, Maître de conférences, Université Blaise Pascal
Directeur de thèse :	David HILL, Professeur, Université Blaise Pascal
Codirecteur de thèse :	Gianni BELLOCCHI, Directeur de recherche, INRA
Invité :	Claude MAZEL, Maître de conférences, Université Blaise Pascal

THE PLANS:



THE PLAN YOU
TELL YOUR
ADVISOR

• "I'M GOING TO BE A
PROFESSOR AT A MAJOR
RESEARCH UNIVERSITY
AFTER I GRADUATE."



THE REAL
PLAN

• LOOK FOR CAREER
ALTERNATIVES.



THE SECRET
PLAN

• BECOME A
BAKER/ROCKSTAR/WRITER.

JORGE CHAM © 2012

WWW.PHDCOMICS.COM

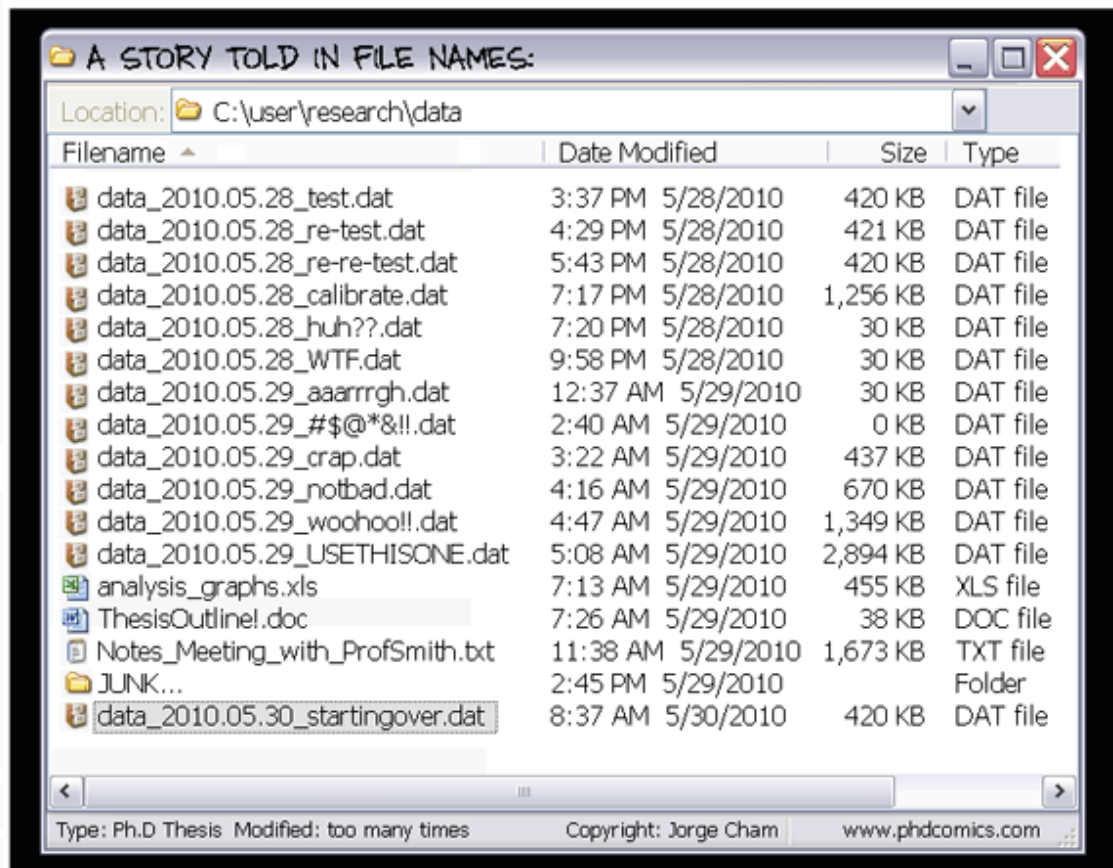


Table des matières

Remerciements	12
Résumé	14
Abstract	15
Glossaire et Abréviations	16
Introduction générale	20
Contexte	20
Objectifs	21
Plan.....	22
Chapitre 1 : La vulnérabilité des prairies au changement climatique	24
1.1. Introduction	24
1.2. Le changement climatique	24
1.3. Les prairies	28
1.4. Comment proposer un outil logiciel générique pour l'analyse de vulnérabilité ?.....	32
1.5. Description de PaSim	37
1.6. Conclusion	42
Chapitre 2 : Etat de l'art : notion de vulnérabilité et éléments de calcul.....	43
2.1. Introduction	43
2.2. La vulnérabilité (au changement climatique).....	43
2.2.1. Définitions des concepts de la vulnérabilité	45
2.2.2. Méthode de Luers pour le calcul de la vulnérabilité.....	53
2.2.3. Autres indices	60
2.2.4. Calcul régionalisé de vulnérabilité	62
2.3. Les plans d'expériences.....	66
2.3.1. Plans d'expériences classiques.....	67
2.3.2. Choix d'un plan d'expériences	69
2.3.3. La conception avancée de plans d'expériences	72
2.3.4. Surfaces de réponse ou métamodèle	77
2.4. Le calcul intensif	82
2.4.1. Les supercalculateurs	83
2.4.2. Cluster & grille.....	83
2.4.3. Le cloud computing	86
2.4.4. Le calcul hybride.....	87

2.4.5.	Outils de distribution.....	88
2.4.6.	OpenMole.....	93
2.5.	L'ingénierie dirigée par les modèles.....	97
2.5.1.	Concepts.....	98
2.5.2.	Un exemple de plateforme intégrant l'IDM.....	100
2.6.	Conclusion.....	104
Chapitre 3 : Modélisation des plans d'expériences par une approche d'ingénierie des modèles		106
3.1.	Introduction.....	106
3.2.	Proposition d'une approche en deux étapes pour l'analyse de vulnérabilité	107
3.3.	Rétro-ingénierie sur des plans d'expériences	109
3.3.1.	Métamodèle des entrées de PaSim	109
3.3.2.	Utilisation de plans d'expériences.....	110
3.3.3.	Modèle et métamodèle des plans d'expériences	114
3.4.	Plans d'expériences pour l'analyse de vulnérabilité	120
3.4.1.	Utilisation du modèle exemple ModVégé.....	120
3.4.2.	Evaluation "simple"	122
3.4.3.	Evaluation tenant compte des incertitudes climatiques.....	123
3.4.4.	Evaluation avec incertitudes sur la fréquence des événements climatiques extrêmes	124
3.4.5.	Evaluation tenant compte des incertitudes sur la gestion.....	126
3.4.6.	Evaluation combinant les incertitudes de gestion et de climat (dont les événements extrêmes).....	127
3.4.7.	Evaluation du minimum de vulnérabilité	128
3.4.8.	Métamodèle des modèles d'impact avec plans d'expériences associés, pour l'analyse de vulnérabilité.....	128
3.5.	Analyse de vulnérabilité sous EMF.....	131
3.5.1.	Modèle agro-écologique	131
3.5.2.	Le plan expérimental	132
3.5.3.	L'analyse de vulnérabilité	133
3.5.4.	Métamodèle complet.....	134
3.5.5.	Implémentation sous EMF	137
3.6.	Conclusion	141
Chapitre 4 : Applications d'analyses de vulnérabilité au changement climatique.....		143
4.1.	Introduction.....	143

4.2.	Cas exemple à l'échelle parcellaire	144
4.2.1.	Le modèle ModVégé	144
4.2.2.	Matériels et méthodes.....	145
4.2.3.	Résultats.....	148
4.2.4.	Discussion.....	149
4.3.	Cas exemple sur carte européenne.....	150
4.3.1.	Simulation à l'échelle européenne.....	150
4.3.2.	Résultats.....	154
4.3.3.	Discussion.....	158
4.4.	Compromis entre précision et temps de calcul de la matière organique du sol à l'équilibre	160
4.4.1.	Plan d'expériences	161
4.4.2.	Résultats.....	164
4.4.3.	Discussion.....	168
4.5.	Surfaces de réponse de la matière organique de PaSim à l'équilibre.....	168
4.5.1.	Choix des facteurs et du domaine de simulation.....	169
4.5.2.	Choix du plan expérimental	171
4.5.3.	Méthodes de régression	171
4.5.4.	Résultats.....	172
4.5.5.	Discussion.....	177
4.5.6.	Extension aux périodes futures.....	177
4.6.	Analyse de vulnérabilité au changement climatique de la matière organique à l'équilibre	181
4.6.1.	Analyse de vulnérabilité au changement climatique de la matière organique à l'équilibre sans prise en compte de l'adaptation.....	181
4.6.2.	Recherche d'un minimum de vulnérabilité.....	185
4.6.3.	Evaluation du minimum de vulnérabilité.....	193
4.6.4.	Discussion.....	196
4.7.	Conclusion	198
	Conclusion générale.....	200
	Retour sur la démarche.....	200
	Conclusions, observations et propositions	201
	Perspectives	204
	Bibliographie	207
	Annexes.....	228

I.	Description des projets Européens	228
i.	CARBO-Extreme.....	228
ii.	GHG Europe.....	231
iii.	ANIMALCHANGE.....	231
II.	NetCDF.....	240
III.	Description rapide de l'utilisation d'OCL dans EMF	241
IV.	Tutoriel d'utilisation de l'outil pour l'analyse de vulnérabilité	241

Table des figures

Figure 1. Les quatre catégories de scénarios SRES sont représentées ici très simplement selon la finalité du développement mondial et le degré de mondialisation. En réalité, ces canevas sont inclus dans un espace défini par davantage de dimensions (<i>i.e.</i> population, équité, technologie) [IPCC, 2007a]	26
Figure 2 Moyennes mondiales multimodèles (courbes en trait plein) du réchauffement en surface pour les scénarios A2, A1B et B1 du SRES (par rapport à la période 1980-1999). A droite du graphique sont indiquées les valeurs les plus probables du réchauffement moyen (lignes horizontales) et les incertitudes associées (zones grises) selon les six scénarios SRES de référence pour la fin du siècle [IPCC 2007a]. Ces valeurs moyennes et ces fourchettes tiennent compte des projections établies à l'aide des modèles de circulation générale, couplés atmosphère-océan, ainsi que des résultats d'une hiérarchie de modèles indépendants et des contraintes liées à l'observation.	27
Figure 3 Diagramme du cycle du carbone en prairie. Les flèches bleues représentent les entrées du système et les flèches rouges les sorties.	30
Figure 4 Diagramme du cycle de l'azote pour le système prairial "sol-plante-animal" [d'après Graux et al., 2012c]. Les flèches bleues représentent les entrées du système et les flèches rouges les sorties.	31
Figure 5 [Graux 2011] Schéma illustratif des entrées/sorties du modèle PaSim.	39
Figure 6 [Lardy et al., 2011b] Diagramme du cycle du carbone entre les compartiments du modèle de décomposition de matière organique du sol de PaSim.	40
Figure 7 [Lardy et al., 2012b] Modèle des concepts liés à la vulnérabilité par rapport au changement climatique, basé sur [Füssel et Klein, 2006]. Les flèches à doubles traits représentent les rétroactions des stratégies d'atténuations et d'adaptation.	47
Figure 8 [Scheffer et al., 2001] Métaphore du potentiel physique (axe z) pour représenter les états d'équilibre par rapport aux conditions et à l'état du système. F1 et F2 représentent deux points de bifurcation pour lesquels un faible stress fait automatiquement changer le système d'état.	49
Figure 9 Exemple de surface de vulnérabilité des prairies pour divers systèmes.	59
Figure 10 Ensemble d'indicateurs utilisés pour développer le modèle de capacité d'adaptation proposé par [Metzger et al., 2006].	64
Figure 11 Création d'un indice de vulnérabilité par le croisement de l'exposition et de la capacité d'adaptation [d'après Nelson et al., 2010].	65
Figure 12 Répartition des communautés en quatre classes de vulnérabilité selon l'intensité des impacts et la capacité d'adaptation [Downing et Patwardhan, 2005]	65
Figure 13 Exemple de plan expérimental à trois facteurs et à deux niveaux : à gauche plan "un seul facteur à la fois", au milieu plan factoriel complet et à droite plan fractionnaire dit à 2^{3-1} (pour 2^{3-1} points couverts sur un espace à 2^3 points).	69
Figure 14 Exemple de plan expérimental à trois facteurs et à deux niveaux : à gauche plan "un seul facteur à la fois", au milieu plan factoriel complet et à droite plan fractionnaire dit à 2^{3-1} (pour 2^{3-1} points couverts sur un espace à 2^3 points).	69
Figure 15 Exemples de LHD pour 2 facteurs (X1 et X2) à 6 niveaux par facteurs	73
Figure 16 Exemple de LHD symétrique, pour un cas à deux facteurs et six niveaux. On notera graphiquement, la symétrie par rapport au centre.	73

Figure 17 Exemple de LHD orthogonal pour un cas à quatre facteurs (X1, X2, X3 et X4) et huit niveaux (1 à 8). On notera que dans la matrice des plans d'expérience, centrée par rapport au centre de l'espace ((1,2,3,4,5,6,7,8) → (-4,-3,-2,-1,+1,+2,+3,+4), ce qui permet à la somme de chaque colonne d'être nulle), les colonnes sont bien orthogonales (produit scalaire nul).....	74
Figure 18 Exemple d'Orthogonal Design $L_9(3^4)$ correspondant à un OA de force 2 avec 9 simulations de 4 facteurs à 3 niveaux.....	75
Figure 19. Diagramme de fonctionnement des éléments constituant de la grille [Mahul 2011] ..	85
Figure 20 Exemple de script OpenMOLE [Reuillon 2012].	95
Figure 21 Exemple de description de l'enchaînement de tâches et du flux de communication dans l'interface d'OpenMOLE [Reuillon 2012].	95
Figure 22 Exemple de description d'une tâche d'échantillonnage (i.e. construction du plan d'expériences) dans l'interface d'OpenMOLE [Reuillon 2012].	96
Figure 23. Exemple de diagramme Ecore, représentant le métamodèle d'une librairie. Toute instance de "Library" est conforme à ce métamodèle.	102
Figure 24. Exemple d'utilisation des contraintes OCL sur le métamodèle des stratégies de gestion (extrait du métamodèle des entrées du modèle ModVege)	103
Figure 25. Proposition d'approche en deux étapes pour l'analyse de vulnérabilité [Lardy et al., 2011a]	108
Figure 26. [Lardy et al., 2011c] Métamodèle UML des entrées du modèle PaSim.....	110
Figure 27. [Lardy et al., 2011c] Niveaux statistiques définis à partir de deux fonctions de distribution cumulée : Uniforme (gris) et Gaussienne (noire).....	113
Figure 28. Modèle du plan expérimental pour les simulations des impacts du changement climatique.	115
Figure 29. Modèle du plan expérimental pour l'analyse de sensibilité.	116
Figure 30. Modèle du plan expérimental pour les simulations de gestion automatique.	117
Figure 31. Patron ou métamodèle d'un plan expérimental (la classe A) avec deux entrées élémentaires (les classes B et C) et la combinaison résultante (la classe D , qui est une forme de classe « BC »).....	118
Figure 32. Métamodèle UML des entrées du modèle ModVégé.....	121
Figure 33. Modèle de plan expérimental pour l'évaluation "simple" de la vulnérabilité sous changement climatique.	123
Figure 34. Modèle de DOE pour le calcul de la vulnérabilité sous changement climatique en tenant compte en partie des incertitudes climatiques.	124
Figure 35. Modèle de DOE pour le calcul de la vulnérabilité sous changement climatique en tenant compte de l'incertitude sur la fréquence des événements climatiques extrêmes.	125
Figure 36. Modèle de DOE pour le calcul de la vulnérabilité sous changement climatique en tenant compte de l'incertitude sur la gestion.	126
Figure 37. Modèle de DOE pour le calcul de la vulnérabilité sous changement climatique en tenant compte des incertitudes sur la gestion et le climat, y compris la fréquence des événements climatiques extrêmes.....	127
Figure 38. Métamodèle d'un modèle agro-écologique.....	128
Figure 39. Diagramme d'instance correspondant au modèle UML d'un DOE associé à un modèle agro-écologique dont le métamodèle est décrit dans la Figure 38. Tout Plan	

d'expériences est la composition d'instances d' Entrée , formant elles-mêmes un ensemble (instance de la classe Entrée 1 x 2 x ... x N).	130
Figure 40. Diagramme UML du modèle des entrées/sorties du modèle agro-écologique.....	132
Figure 41. Diagramme UML d'un métamodèle de plan d'expériences.	133
Figure 42. Diagramme UML d'un métamodèle d'analyse de vulnérabilité.	134
Figure 43. Métamodèle et modèle d'analyse de vulnérabilité.	135
Figure 44. Métamodèle d'analyse de vulnérabilité sous EMF.	136
Figure 45. Exemple d'insertion de contrainte OCL dans l'implémentation du métamodèle sous EMF.	138
Figure 46. Exemple de spécification du code contenu dans une méthode sous EMF, à travers le principe des annotations du Java.	139
Figure 47. Diagramme d'activité de l'analyse de vulnérabilité par le plugin EMF.....	141
Figure 48 [Lardy et al., 2011a] Résultats des simulations de A) Biomasse sur pied et B) digestibilité de l'herbe obtenus par ModVégé sur une prairie permanente intensive (année 1994) à Marcenat (45° 18' Nord, 02° 49' Est), en Auvergne.....	145
Figure 49 Calcul de vulnérabilité en tenant compte A) juste du pixel, B) du premier niveau de pixels voisins, C) des pixels voisins jusqu'à deux niveaux.....	154
Figure 50 Calcul de la vulnérabilité des prairies en Europe en utilisant l'indice de vulnérabilité proportionnelle pour la GPP annuelle (<i>Gross Primary Production</i>), pour la gestion Moyenne, pour la période climatique de "Futur Proche" (2020-2049), pour un seuil A) de 5%, B) de 10%, C) de 15%, D) de 20%. Les indices ont été calculés en utilisant jusqu'à deux niveaux de pixels voisins, puis normés par les valeurs de la période de référence (1980-2009).	155
Figure 51 Seuils de vulnérabilité pour la GPP annuelle (<i>Gross Primary Production</i>) pour la gestion Moyenne, calculés en utilisant jusqu'à deux niveaux de pixels voisins, pour un seuil statistique de 5% à gauche et de 20% à droite.....	156
Figure 52 Calcul de la vulnérabilité des prairies en Europe en utilisant l'indice de Luers pour la GPP annuelle (<i>Gross Primary Production</i>), pour la gestion Moyenne, pour la période climatique de "Futur Proche" (2020-2049), en tenant compte A) juste du pixel, B) du premier niveau de pixels voisins, C) des pixels voisins jusqu'à deux niveaux. Les indices ont été normés par les valeurs de la période de référence (1980-2009).....	157
Figure 53 Vulnérabilité en utilisant l'indice de Luers pour le rendement pour la période climatique de "futur lointain" (2070-2099), pour A) la gestion extensive, B) la gestion moyenne, C) la gestion intensive. Les indices ont été calculés en utilisant jusqu'à deux niveaux de pixels voisins, puis normés par les valeurs de la période de référence (1980-2009).....	158
Figure 54 Comparaison de matière organique du sol à l'équilibre en fonction du nombre d'années utilisées dans le cycle climatique, pour une prairie temporaire en monoculture de <i>Festuca arundinacea</i> . Schreb. (SF) à Bordeaux.....	164
Figure 55 Comparaison de matière organique du sol à l'équilibre en fonction du nombre d'années utilisées dans le cycle climatique, pour une prairie permanente extensive (PE) à Bordeaux.....	165
Figure 56 [Lardy et al., 2012a] Influence du changement d'une seule année climatique sur le carbone organique du sol à l'équilibre, pour A) une prairie extensive en climat aride sur sol peu profond, et pour B) une prairie intensive, en climat humide sur sol profond. .	176

Figure 57. Exemple de convergence des algorithmes d'optimisation ("force brut" en vert, l'algorithme de la première pente en bleu et l'algorithme du recuit simulé en rouge) sur la période climatique de référence (1970-2006) et l'indice de "Severity Vulnerability".	189
Figure 58. Comparaison des valeurs des indices de vulnérabilité avec prise en compte de l'adaptation, obtenues à l'aide des surfaces de réponse (avec entrée climatique P-ET ₀ ou Saison), en ordonnée, et avec PaSim en abscisse, pour les différentes périodes climatiques ("Période de Référence", PR : 1970-2006 ; "Futur Proche", FP : 2018-2051 ; "Futur Lointain", FL : 2067-2100).	194
Figure 59. Création du projet d'analyse de vulnérabilité.	243
Figure 60. Choisir le nom du projet d'analyse de vulnérabilité.	244
Figure 61. Insérer le modèle agronomique.	245
Figure 62. Ajouter une entrée au modèle.	245
Figure 63. Associer une distribution et un échantillonnage à une entrée.	246
Figure 64. Valider une distribution associée à une entrée.	247
Figure 65. Définition de la nature du plan expérimental.	248
Figure 66. Graphe de l'analyse de vulnérabilité après ajout de la surface de réponse de type régression linéaire.	249
Figure 67. Lancement de la fonctionnalité "Analyse de vulnérabilité".	250
Figure 68. Diagramme d'activité de l'analyse de vulnérabilité par le plugin EMF.	251
Figure 69. Exemple de code R généré par le plugin EMF, pour la régression linéaire.	252
Figure 70. Sélection des fichiers pour les calculs d'indices.	253
Figure 71. Exemple de résultats de calculs d'indices.	253

Table des tableaux

Table 1 Concepts liés à la vulnérabilité	51
Table 2 Exemples associés aux concepts liés à la vulnérabilité	52
Table 3 Calcul de la vulnérabilité dans le cas d'une prairie permanente extensive et sous réchauffement climatique.	54
Table 4 Comparaison de la vulnérabilité selon deux modes de calcul, sans prise en compte de l'exposition (V) ou avec (V').	55
Table 5 Comparaison de la vulnérabilité dans le système d'origine et dans le système optimisé. .	57
Table 6 Potentiel d'adaptation basé sur une gestion optimale et vulnérabilité calculée en tenant compte de l'exposition (A') ou non (A).	57
Table 7 Calcul de la vulnérabilité dans le cas d'une prairie permanente extensive et sous réchauffement climatique.	61
Table 8. Résumé des plans expérimentaux pour l'analyse de vulnérabilité sous changement climatique.	122
Table 9 Résumé des indices de vulnérabilité.	146
Table 10. Indices de vulnérabilité, calculés pour différents plans d'expériences (DOE) pour le "futur proche" (FP : 2020-2049) et le "futur lointain" (FL : 2070-2099). Les valeurs absolues des indices ont été normées par la valeur de la "période de référence" (PR : 1975-2004).	149
Table 11. Description des quatre systèmes de prairie simulés : prairie temporaire semée en monoculture (SL et SF) et prairies permanentes semi-naturelles extensive et intensive (PE et PI).	162
Table 12. Géo-localisation et type de climat pour les sites utilisés dans cette étude. Pour les détails sur la classification des types de climats, voir [Graux et al., 2012a]	163
Table 13. Erreurs relatives (moyenne \pm erreur standard) pour la quantité totale de carbone organique du sol prédit avec des cycles météorologiques de longueurs différentes pour différents systèmes prairiaux sur 12 sites.	166
Table 14. Erreurs relatives (moyenne \pm erreur standard) pour la quantité totale de carbone organique du sol prédit avec des cycles météorologiques de longueurs différentes pour 12 sites.	167
Table 15. Erreurs relatives (moyenne \pm erreur standard) pour la quantité totale de carbone organique du sol prédit avec différentes combinaisons d'ordre d'années au sein d'un cycle climatique de trois ans, pour différents systèmes prairiaux pour deux sites choisis.	167
Table 16. Description des facteurs du plan d'expériences et de leurs niveaux.	171
Table 17. Évaluation par l'indice RRMSE (" <i>Relative Root Mean-Squared Error</i> ") des régressions linéaires et quadratiques pour les cinq compartiments de carbone (C) et leur total pour les trois niveaux d'entrées climatiques (P-ET ₀ , Saison et Mois), sur les points du plan d'expériences.	173
Table 18. Évaluation par l'indice RMSRE (" <i>Root Mean-Squared Relative Error</i> ") des régressions linéaires et quadratiques pour les cinq compartiments de carbone (C) et leur total pour les trois niveaux d'entrées climatiques (P-ET ₀ , Saison et Mois), sur les points du plan d'expériences.	173

Table 19. Évaluation par l'indice d' efficacité de modélisation (EF) des régressions linéaires et quadratiques pour les cinq compartiments de carbone (C) et leur total pour les trois niveaux d'entrées climatiques (P-ET ₀ , Saison et Mois), sur les points du plan d'expériences.....	174
Table 20. Évaluation par l'indice RRMSE (" <i>Relative Root Mean-Squared Error</i> ") des régressions linéaires et quadratiques du carbone total du sol pour les trois périodes climatiques ("Référence" : 1970-2006 ; "Futur Proche" : 2018-2051; "Futur Lointain" : 2067-2100), pour les trois niveaux d'entrées climatiques (P-ET ₀ , Saison et Mois), sur les points du plan d'expériences.....	179
Table 21. Évaluation par l'indice RMSRE (" <i>Root Mean-Squared Relative Error</i> ") des régressions linéaires et quadratiques du carbone total du sol pour les trois périodes climatiques ("Référence" : 1970-2006 ; "Futur Proche" : 2018-2051; "Futur Lointain" : 2067-2100), pour les trois niveaux d'entrées climatiques (P-ET ₀ , Saison et Mois), sur les points du plan d'expériences.....	179
Table 22. Évaluation par l'indice d' efficacité de modélisation des régressions linéaires et quadratiques du carbone total du sol pour les trois périodes climatiques ("Référence" : 1970-2006 ; "Futur Proche" : 2018-2051; "Futur Lointain" : 2067-2100), pour les trois niveaux d'entrées climatiques (P-ET ₀ , Saison et Mois), sur les points du plan d'expériences.....	179
Table 23. Indices de vulnérabilité , calculées sur le plan d'expériences utilisé pour régresser les surfaces de réponse, pour le "Futur Proche" (FP : 2018-2051) et le "Futur Lointain" (FL : 2067-2100). Les valeurs des indices ont été normées par la valeur de la "Période de Référence" (PR : 1970-2006).....	182
Table 24. Indices de vulnérabilité , calculés à l'aide des régressions quadratiques pour les trois niveaux d'entrée climatique (P-ET ₀ , Saison et Mois) (sur le plan d'expériences ayant lui-même servi à les régresser) et pour les résultats de simulation à l'aide de PaSim, pour le "Futur Proche" (FP : 2018-2051) et le "Futur Lointain" (FL : 2067-2100). Les valeurs des indices ont été normées par la valeur de la "Période de Référence" (PR : 1970-2006).	183
Table 25. Indices de vulnérabilité , calculés à l'aide des régressions quadratiques pour les trois niveaux d'entrée climatique (P-ET ₀ , Saison et Mois), pour le "Futur Proche" (FP : 2018-2051) et le "Futur Lointain" (FL : 2067-2100). La vulnérabilité est évaluée en chacun des points du LHD ayant servis à régresser les surfaces, par une plan de type OAT . Les valeurs des indices calculés de manière uniforme à l'aide de PaSim sont également rappelées. Les valeurs des indices ont été normées par la valeur de la "Période de Référence" (PR : 1970-2006).	184
Table 26. Quantité de fertilisant organique annuel (kg N/ha/an) dans la gestion minimisant la vulnérabilité, en fonction de l'indice de vulnérabilité, de la période climatique ("Période de Référence", PR : 1970-2006 ; "Futur Proche", FP : 2018-2051 ; "Futur Lointain", FL : 2067-2100) et de la surface de réponse utilisée (avec entrée climatique P-ET ₀ ou Saison). Résultats obtenus avec la méthode de la "force brute".....	190
Table 27. Pourcentage de légumineuses dans la gestion minimisant la vulnérabilité, en fonction de l'indice de vulnérabilité, de la période climatique ("Période de Référence", PR : 1970-2006 ; "Futur Proche", FP : 2018-2051 ; "Futur Lointain", FL : 2067-2100) et de la surface	

de réponse utilisée (avec entrée climatique P-ET ₀ ou Saison). Résultats obtenus avec la méthode de la première pente.....	190
Table 28. Gain d'adaptation pour les périodes climatiques "Période de Référence" (PR : 1970-2006), "Futur Proche" (FP : 2018-2051) et "Futur Lointain" (FL : 2067-2100), en fonction de la surface de réponse utilisée (avec entrée climatique P-ET ₀ ou Saison). Les valeurs correspondent à la différence de l'indice de vulnérabilité sans et avec adaptation , le tout normé par la valeur de l'indice sans adaptation pour la période en question. Plus la valeur est haute plus le gain par adaptation est important.	191
Table 29. Vulnérabilité avec prise en compte de l'adaptation , pour les différentes périodes climatiques ("Période de Référence", PR : 1970-2006 ; "Futur Proche", FP : 2018-2051 ; "Futur Lointain", FL : 2067-2100) et en fonction de la surface de réponse utilisée (avec entrée climatique P-ET ₀ ou Saison). Les valeurs ont été normées par la valeur de l'indice sans adaptation à la période de référence	191
Table 30. Vulnérabilité avec prise en compte de l'adaptation , pour les périodes climatiques "Futur Proche" (FP : 2018-2051) et "Futur Lointain" (FL : 2067-2100), en fonction de la surface de réponse utilisée (avec entrée climatique P-ET ₀ ou Saison). Les valeurs ont été normées par la valeur de l'indice avec prise en compte de l'adaptation à la "Période de Référence" (PR : 1970-2006).....	192
Table 31. Vulnérabilité avec prise en compte de l'adaptation, relativement à la vulnérabilité sans adaptation à la période de référence . Les indices ont été évalués par PaSim aux « minima » trouvés par les différentes surfaces de réponse (avec entrée climatique P-ET ₀ ou Saison), pour les différentes périodes climatiques ("Période de Référence", PR : 1970-2006 ; "Futur Proche", FP : 2018-2051 ; "Futur Lointain", FL : 2067-2100). Les valeurs ont été normées par la valeur de l'indice sans adaptation à la période de référence.	195
Table 32. Vulnérabilité avec prise en compte de l'adaptation, relativement à la vulnérabilité avec adaptation à la période de référence . Les indices ont été évalués par PaSim aux « minima » trouvés par les différentes surfaces de réponse (avec entrée climatique P-ET ₀ ou Saison), pour les périodes climatiques "Futur Proche" (FP : 2018-2051) et "Futur Lointain" (FL : 2067-2100). Les valeurs ont été normées par la valeur de l'indice avec prise en compte de l'adaptation à la "Période de Référence" (PR : 1970-2006), calculée à l'aide des résultats de PaSim.	195
Table 33. Gain d'adaptation pour les périodes climatiques "Période de Référence" (PR : 1970-2006), "Futur Proche" (FP : 2018-2051) et "Futur Lointain" (FL : 2067-2100). Les indices ont été évalués par PaSim aux « minima » trouvés par les différentes surfaces de réponse (avec entrée climatique P-ET ₀ ou Saison). Les valeurs correspondent à la différence de l'indice de vulnérabilité sans et avec adaptation , le tout normé par la valeur de l'indice sans adaptation pour la période en question, calculée à l'aide des résultats de PaSim..	196

Remerciements

Je voudrais tout d'abord remercier David Hill, mon directeur de thèse, qui a toujours su se rendre disponible, et qui a un talent inégalable pour motiver ses doctorants (et pas uniquement grâce aux repas qu'il organise avec eux). Si, comme tout doctorant, il m'est arrivé de douter, je suis toujours ressorti motivé de son bureau. Je voudrais ensuite remercier Gianni Bellocchi, mon co-directeur de thèse, pour son aide importante sur la partie agronomie/écologie de mon travail, mais également pour les discussions intéressantes allant de l'Histoire à la religion. Je voudrais ensuite remercier Bruno Bachelet, mon co-encadrant informatique, qui a toujours été disponible et a su me donner un angle de vue différent sur mon travail. Merci encore à vous trois, car c'est grâce à vous que j'ai pu réaliser cette thèse dans les temps.

Je tiens également à remercier Pascal Carrère et l'unité pour m'avoir permis de faire cette thèse, et de la faire dans de bonnes conditions. Parmi mes collègues de travail à l'INRA, je voudrais tout d'abord remercier Anne-Isabelle Graux, avec qui j'ai travaillé depuis mon arrivée à l'INRA. La confrontation de nos points de vue a toujours été enrichissante. Je voudrais ensuite remercier Katja Klumpp et Raphaël Martin avec qui j'ai eu le plaisir d'interagir avant et pendant la thèse. Je voudrais également remercier et souhaiter bonne chance à Haythem Ben Touhami pour la suite de sa thèse. J'ai eu plaisir à te transmettre ma connaissance du modèle PaSim, mais surtout j'ai apprécié nos discussions ouvertes, auxquelles se sont parfois également joints Franck Ganzalco, Shaoxiou Ma et Sylvain Laperche, également membres de notre bureau, que je remercie fortement. J'aimerais aussi remercier mes collègues doctorants pour leurs échanges et leurs soutiens : Antoine, Rie, Marine, Nazia, Nian-Xun, Damien, Benoît et Pauline. Je voudrais également remercier tous ceux avec qui j'ai eu le plaisir de travailler ou simplement de discuter : Julien, Sébastien, Gaël, Olivier, Jean-François, Giorgio, Juliette, Frédérique, Annick, Laurence pour ne citer qu'eux. Je remercie aussi mes anciens collègues, notamment Vincent Maire, Mathilde Piquet, Amélie Cantarel, Tiphaine Tallec, Tanvir Shahzad, Lionel Thiery et Michael Gaurut (avec qui j'ai apprécié de partager le développement de PaSim).

Cette thèse a également été réalisée au sein d'un autre labo, le LIMOS, que je remercie pour m'avoir accueilli. Je remercie tout particulièrement Jonathan P. (oui ! je n'ai pas oublié ma dette de repas), Luc, Jonathan C., Nathalie, Pierre, Sébastien, Guillaume, Faouzi, Wajdi et Corinne pour ne citer qu'eux. Je voudrais également remercier Pascale Gouinaud pour sa disponibilité par rapport à l'administration du cluster du LIMOS. Je tiens également à remercier Romain Reuillon pour son aide et ses réponses rapides sur l'utilisation d'OpenMOLE.

Je remercie aussi Claude Mazel pour avoir accepté de relire mon manuscrit et pour les discussions très intéressantes qui s'ensuivirent.

Je suis également reconnaissant envers les projets directement ou indirectement (via l'accès à des données, notamment) rattachés à ma thèse : les ANR CLIMATOR, NITROEUROPE et VALIDATE, et les projets européens FP6 NITROEUROPE, FP7 CARBOExtreme, FP7 GHG-Europe et FP7 ANIMALCHANGE.

Je remercie toute ma famille et mes amis qui ont su me soutenir et m'inciter à me changer les idées quand c'était nécessaire.

Je tiens aussi à remercier l'ensemble des membres de mon jury qui m'ont fait l'honneur de bien vouloir y participer.

Enfin, je remercie tous ceux que j'aurais pu oublier.

Résumé

La vulnérabilité est le degré de stress ou de perturbation qu'un système humain et environnemental est capable de supporter avant d'être endommagé. Au cours des dernières années, c'est devenu un sujet central de l'étude du changement global (incluant le changement climatique), et l'évaluation de la vulnérabilité des agro-écosystèmes aux changements climatiques s'inscrit en effet dans les axes prioritaires de l'Institut National de la Recherche Agronomique (INRA). La littérature sur le changement climatique contient de nombreuses explications de la vulnérabilité, basée sur la notion de sensibilité et faisant éventuellement appel à des idées plus complexes, depuis la prise en compte de l'exposition jusqu'aux impacts résiduels du changement climatique après mise en place de mesures d'adaptation. Dans le cadre des activités de l'Unité de Recherche sur les Ecosystèmes Prairiaux (UREP), l'intérêt porte sur la vulnérabilité des systèmes prairiaux et d'élevage face aux risques de réduction de la production laitière et herbagère, ainsi que la problématique de l'augmentation des émissions de gaz à effet de serre qui accompagne la production des services d'un écosystème prairial.

Compte-tenu du fait que l'analyse de vulnérabilité s'apparente à une forme d'analyse de sensibilité et de la cascade d'incertitudes lors des évaluations des impacts du changement climatique, un grand nombre de simulations sont nécessaires. Dans ce cadre, afin de réduire le temps d'attente utilisateur, la conception d'un plan expérimental approprié est nécessaire, ainsi que l'utilisation du calcul à haute performance. De plus, l'analyse de vulnérabilité peut se composer de nombreuses étapes parmi lesquelles on trouve la conception (choix du modèle agro-écologique, des variables d'intérêt, des scénarii et des seuils de référence, la distribution des paramètres, ...), la génération de plans expérimentaux, la création de surfaces de réponse, le calcul de métriques (par exemple des indices de vulnérabilité) et l'optimisation (à travers la conception et l'évaluation de mesures d'adaptation) de la vulnérabilité. À notre connaissance, aucun outil spécifique n'a été conçu ou validé dans le but de faciliter l'exécution de la majeure partie des tâches énoncées précédemment. Ainsi, la problématique de cette thèse a été de proposer une méthode générique pour la réalisation d'une analyse de vulnérabilité sous changement climatique. Le travail réalisé dans cette thèse a donc commencé par une révision du concept de vulnérabilité et la proposition d'une démarche générique, en s'appuyant sur une synthèse critique des notions de l'état de l'art. Ensuite, avec une approche d'Ingénierie Dirigée par les Modèles, nous avons conçu un outil informatique pour l'analyse de vulnérabilité. Cet outil, implémenté avec le cadriciel *Eclipse Modeling Framework* (EMF), est générique, modulaire et permet la distribution et l'interprétation des simulations. Enfin, des exemples applicatifs d'analyse de vulnérabilité au changement climatique ont été réalisés à l'aide des solutions proposées précédemment. Cette démarche s'est notamment appuyée sur l'utilisation du modèle biogéochimique d'écosystème prairial PaSim ([Riedo et al., 1998], [Vuichard 2007a], [Graux 2011]).

Abstract

Vulnerability is the degree to which human or environmental systems are likely to experience harm due to a perturbation or a stress. In the last years, it has become a central focus of the global change (including climate change). Assessing the vulnerability of agro-ecosystems to climate change is one of the priority areas of the French National Institute of Agronomic Research (INRA). The climate change literature contains many explanations of vulnerability, stemming from the notion of sensitivity to more complex ideas, yet taking into account the exposure history of the system up to residual impacts of climate change after adaptation. In the framework of the activities of the INRA's Grassland Ecosystem Research Unit (UREP) of Clermont-Ferrand, interest is on vulnerability of grassland and livestock systems against the risk of reduced milk and forage production, and against the problem of increased greenhouse gas emissions that comes with the production of grassland ecosystem services.

Vulnerability assessment has similarities with sensitivity analysis and is based on simulations of the target system, forced to respond to the changes of stress factors. Due to the cascade of uncertainties in climate change impacts assessment, a large number of simulations are necessary. In this context, the need to reduce user waiting time calls for the conception of an appropriate experimental plan, as well as the use of high performance computing. Moreover, vulnerability assessment may consist of many steps, such as designing the experiment (choice of agro-ecological model, variables of interest, scenarios, reference thresholds, parameters distribution ...), designing of the experimental plans, regressing response surfaces, computing metrics (e.g. vulnerability indices) and optimizing vulnerability (through designing and evaluating adaptation measures). To our knowledge, no specific tool has been built or validated, in order to facilitate the implementation of most of these tasks. Thus, the goal of this thesis was to propose a generic method to perform a comprehensive vulnerability analysis to climate change. The work in this thesis has begun with a review of the concept of vulnerability and the proposal of a generic approach, based on a critical synthesis of the state of the art. Then, with a Model Driven Engineering approach, we have developed a computer tool for vulnerability analysis. This tool, implemented with the modelling framework *Eclipse Modeling Framework* (EMF) is generic, modular and allows the distribution and interpretation of simulation results. Finally, application examples of climate change vulnerability assessment were achieved with the previously proposed solutions. This approach relied, in particular, on the use of the grassland ecosystem biogeochemical model PaSim ([Riedo et al., 1998], [Vuichard 2007a], [Graux 2011]).

Glossaire et Abréviations

Abscission : processus physiologique naturel d'une plante entraînant la chute des feuilles, des fleurs ou des fruits.

API : interface de programmation, fournie par un système informatique ou une bibliothèque afin de permettre l'interaction des composants les uns avec les autres. Abréviation issue de l'anglais pour **Application Programming Interface**.

Cuticule : couche protectrice des organes aériens des végétaux.

Dicotylédone : espèce végétale contenant deux cotylédons, c'est-à-dire deux feuilles primordiales issues de la graine.

DOE : plans d'expériences. Abréviation issue de l'anglais pour **Design Of Experiment**.

Effet de serre : processus naturel par lequel les rayonnements thermiques à la surface d'une planète sont absorbés par les Gaz à Effet de Serre (GES) de l'atmosphère et réémis dans toutes les directions. Ce phénomène induit une augmentation de la température de surface par rapport à la température en l'absence de ces GES.

EMF : **Eclipse Modeling Framework**, est une plateforme de modélisation et de génération de code, s'appuyant sur l'outil Eclipse et intégrant les principes de l'IDM.

Fermentation entérique : transformation / décomposition de la matière organique sous l'action de microorganismes dans le milieu intestinal et digestif des ruminants, en anaérobie (*i.e.* sans dioxygène (O₂)).

GES : **Gaz à Effet de Serre**, composants gazeux de l'atmosphère ayant pour caractéristique commune d'absorber une partie des infrarouges émis par la surface d'une planète et contribuant ainsi à l'effet de serre. Les principaux gaz à effet de serre existant naturellement dans l'atmosphère sont : la vapeur d'eau (H₂O), le dioxyde de carbone (CO₂), le méthane (CH₄), le protoxyde d'azote (N₂O) et l'ozone (O₃).

GHG : **GreenHouse Gas**, équivalent anglais de GES.

GIEC : **Groupe d'Experts Intergouvernemental sur l'évolution du Climat** (équivalent français de l'IPCC). Créé en 1988 par l'Organisation Météorologique Mondiale (OMM) et le Programme des Nations Unies pour l'Environnement (PNUE), il a pour mission d'évaluer, de façon neutre, méthodique, claire et objective, les informations d'ordre scientifique, technique et socio-économique qui nous sont nécessaires pour mieux comprendre les fondements scientifiques des risques liés au changement climatique d'origine humaine, cerner plus précisément les conséquences possibles de ce changement et envisager d'éventuelles stratégies d'adaptation et d'atténuation. L'une des principales activités du GIEC consiste à procéder, à intervalles réguliers, à une évaluation de l'état des connaissances relatives au changement climatique. (http://www.ipcc.ch/home_languages_main_french.shtml)

GPP : **Gross Primary Production**, acronyme anglais pour production primaire brute. On appelle production primaire l'énergie accumulée par les producteurs primaires. Dans les écosystèmes terrestres, les producteurs primaires sont principalement les plantes. La

GPP est représentée par le flux en masse de carbone assimilé par unité de temps. Cette production de matière organique s'effectue principalement grâce à la photosynthèse.

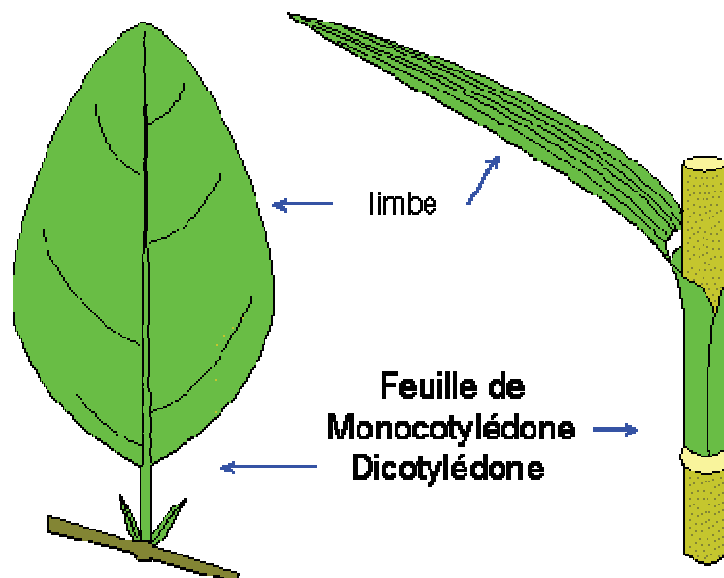
Graminées : Nom courant d'une famille de plantes monocotylédones (*i.e.* contenant une seule feuille primordiale issue de la graine), plus récemment nommées "*Poaceae*", très importantes au niveau agronomique puisqu'elles contiennent des espèces essentielles dans l'économie mondiale (par exemple les céréales). On trouve dans cette famille de nombreuses plantes fourragères comme le dactyle, la fétuque et le ray-grass.

IDM : Ingénierie Dirigée par les Modèles, branche du Génie Logiciel s'intéressant au développement, à la maintenance et à l'évolution des modèles au sein des systèmes informatiques, avec une approche de modélisation unificatrice.

IPCC : Intergovernmental Panel on Climate Change. Nom anglais du GIEC.

Légumineuses : Nom courant d'une famille de plantes dicotylédones (*i.e.* contenant deux feuilles primordiales issues de la graine), plus récemment nommées "*Fabaceae*", très importantes au niveau agronomique puisqu'elles permettent d'enrichir le sol en azote en le captant dans l'atmosphère grâce à la présence de nodules (petites boursouflures) fixateurs sur les racines (phénomène dit de fixation symbiotique car réalisé en symbiose avec des bactéries). On trouve dans cette famille la luzerne et le trèfle.

Limbe : Partie large de la feuille d'une plante. Il peut prendre des formes très variables. La diversité des limbes est due principalement à la diversité des modes de nervation (*i.e.* de constitution des nervures de la plante). Voir exemple ci-dessous.



Exemple de limbe [<http://malherbologie.cirad.fr/Adventrop/defs/limbe.html>]

LSU : LiveStock Unit, équivalent anglais de l'UGB.

Métamodèle (au sens de l'IDM) : modèle d'un langage de modélisation. Cette définition s'appuie sur la relation suivante typique de la notion de métamodèle : « un modèle **est conforme à** un métamodèle ».

Métamodèle (au sens du domaine des plans d'expériences) : modèle de modèle ou surfaces de réponse, à savoir un modèle plus simple capable de reproduire les sorties

issues du simulateur en fonction de ses entrées. Ce terme n'est utilisé que dans la partie Chapitre 2 : de ce manuscrit.

Monocotylédone : espèce végétale ne contenant qu'un cotylédon, c'est-à-dire une seule feuille primordiale issue de la graine.

MOS : Matière Organique du Sol¹. Le terme matière organique regroupe une somme importante et hétérogène de substances et composés carbonés d'origine végétale et animale : des débris en cours de décomposition issus de la végétation (feuilles, racines, herbe) qui constituent la litière du sol, jusqu'à l'humus stable solidement fixé aux particules d'argile.

OCL : Object Constraint Language, est un langage formel permettant, sur un graphe UML (Unified Modeling Language), de spécifier des contraintes (au sens de restrictions entre une ou plusieurs valeurs d'un élément du modèle).

OMM : L'Organisation Météorologique Mondiale est l'institution spécialisée des Nations Unies qui, depuis 1950, fait autorité pour tout ce qui concerne l'état et le comportement de l'atmosphère terrestre, son interaction avec les océans, le climat qui en est issu et la répartition des ressources en eau qui en résulte [http://www.wmo.int/pages/index_fr.html]

Scénarios SRES : scénarios de projections futures d'émissions de gaz à effet de serre, issues du rapport spécial sur les scénarios d'émissions (SRES) du GIEC [Nakićenović 2000].

SRES : acronyme anglais pour **Special Report on Emission Scenarios**. Le SRES est le rapport spécial du GIEC publié en l'an 2000 sur les scénarios d'émission des gaz à effet de serre. Ces scénarios décrivent un monde futur à partir d'hypothèses sur l'évolution démographique, le contexte économique, l'évolution technologique et les échanges entre pays, ainsi que les émissions de GES qui en résultent.

Stomate : orifice de petite taille présent dans l'épiderme des organes aériens des végétaux (sur la face inférieure des feuilles le plus souvent). Il permet les échanges gazeux (dioxygène (O₂), dioxyde de carbone (CO₂), vapeur d'eau (H₂O)...) entre la plante et l'air ambiant.

Système fourrager : ensemble organisé de moyens destinés à produire des fourrages. Peut également être défini comme un « ensemble d'informations et de décisions visant à équilibrer les besoins et les ressources en fourrages en vue de répondre à un objectif de production dans un cadre de contraintes donnés » [Moreau et al., 2010]

Tige et Gaine foliaire : organes aériens de la plante. La tige est l'axe aérien qui prolonge les racines et porte les bourgeons et les feuilles. La gaine foliaire est la base élargie des feuilles, enveloppant la tige dans une partie de sa longueur. Voir exemple ci-dessous.

¹ <http://www.vignevin-sudouest.com/publications/fiches-pratiques/matiere.organique.php>



Exemple de la partie formée par la tige, entourée par la gaine foliaire, et par les feuilles, pour du vulpin (graminée).

UGB : Unité de Gros Bétail² (équivalent français du LSU) est une unité de référence permettant d'agréger le bétail de différentes espèces et de différents âges en utilisant des coefficients spécifiques établis initialement sur la base des besoins nutritionnels ou alimentaires de chaque type d'animal. L'unité standard utilisée pour le calcul du nombre d'unités de gros bétail (= 1 UGB) est l'équivalent pâturage d'une vache laitière produisant 3 000 kg de lait par an, sans complément alimentaire concentré.

² http://ec.europa.eu/agriculture/envir/report/fr/lex_fr/report.htm

Introduction générale

Contexte

On pourrait définir de manière générique la vulnérabilité comme le degré de stress ou de perturbation qu'un système humain ou environnemental est capable de supporter avant d'être endommagé [Turner et al., 2003a]. Au cours des dernières années, la vulnérabilité est devenue un sujet central de l'étude du changement global (incluant le changement climatique). Le changement climatique, vu comme une modification durable du climat, est devenu une préoccupation majeure depuis que les impacts de ce dernier sont observables [IPCC 2007b]. Par exemple, l'évaluation de la vulnérabilité des agro-écosystèmes aux changements climatiques s'inscrit en effet dans les axes prioritaires de l'Institut National de la Recherche Agronomique (INRA). La proposition et la mise en place de politiques d'adaptation est devenue une priorité [Hinkel 2011]. Par conséquent, les porteurs d'enjeux demandent souvent quels pays, régions ou secteurs, sont les plus vulnérables, dans le but de prioriser les efforts qui doivent être fournis pour minimiser les risques et atténuer les conséquences (e.g., [Füssel et Klein, 2006]). Dans le cadre des activités de l'Unité de Recherche sur les Ecosystèmes Prairiaux (UREP) de l'INRA, un intérêt particulier est porté sur la vulnérabilité des systèmes prairiaux et d'élevage face aux risques de réduction de la production laitière et herbagère, ou de perte de la diversité biologique, ainsi que sur la problématique de l'augmentation des émissions de gaz à effet de serre qui accompagne la production des services d'un écosystème prairial.

Dans le cadre des études sur le changement climatique, le terme de vulnérabilité est utilisé avec différentes significations au sein de la communauté scientifique. De plus il est souvent mal défini et parfois même utilisé sans définition (e.g., [Ionescu et al., 2009]). Il existe donc toute une gamme de méthodes d'évaluation de la vulnérabilité ([Eakin et Luers, 2006] [Füssel et Klein, 2006]). Au fil du temps, les méthodes sont montées en complexité, incluant de plus en plus de sous-systèmes, de processus, de facteurs, de rétroactions et de types d'impacts à prendre en compte. La littérature sur le changement climatique contient de nombreuses explications de la vulnérabilité, basée sur la notion de sensibilité et faisant éventuellement appel à des idées plus complexes, depuis la prise en compte de l'exposition jusqu'aux impacts résiduels du changement climatique, une fois les mesures d'adaptation mises en place.

Même si la littérature contient beaucoup de méthodes, réaliser une analyse de vulnérabilité reste un problème difficile, compte-tenu du manque de consensus sur la signification du terme lui-même, mais surtout parce que la vulnérabilité n'est pas un phénomène directement observable [Downing et al., 2001]. Si certains aspects de la

vulnérabilité, tels que la résilience, peuvent être observés expérimentalement, certains autres aspects doivent nécessairement faire appel à la modélisation. Compte-tenu de la cascade d'incertitudes lors des évaluations des impacts du changement climatique et du fait que l'analyse de vulnérabilité s'apparente à une forme d'analyse de sensibilité, un grand nombre de simulations sont nécessaires. Dans ce cadre, afin de réduire le temps d'attente utilisateur, la conception d'un plan expérimental approprié est nécessaire, ainsi que l'utilisation du calcul à haute performance. En effet, même avec l'utilisation de plans d'expériences, cette thèse a nécessité environ 24 ans de temps de calcul mono-CPU, qui ont été distribués sur cluster et sur grille de calcul. De plus, l'analyse de vulnérabilité peut se composer de nombreuses étapes, parmi lesquelles on trouve la conception (choix du modèle agro-écologique, des variables d'intérêt, des scénarios et des seuils de référence, distribution des paramètres, ...), la génération de plans expérimentaux, la création de surfaces de réponse, le calcul de métriques (par exemple des indices de vulnérabilité) et l'optimisation de la vulnérabilité (à travers la conception et l'évaluation de mesures d'adaptation), comme illustré dans [Lexer et Seidl, 2009] [Metzger et al., 2005] et [Luers et al., 2003]. À notre connaissance, aucun outil spécifique n'a été conçu ou validé dans le but de faciliter l'exécution de la majeure partie des tâches énoncées précédemment. C'est pourquoi la conception d'une approche générique et la création d'un outil l'implémentant et facilitant la réalisation des différentes étapes d'une analyse de vulnérabilité serait un gain pour les chercheurs qui s'intéressent aux impacts du changement climatique (même au-delà du domaine scientifique traité dans ce travail de thèse).

Objectifs

Notre problématique est de proposer une méthode générique pour la réalisation d'une analyse de vulnérabilité sous changement climatique. Notre travail se décline en plusieurs objectifs.

Notre premier objectif a consisté en une révision du concept même de vulnérabilité. Pour cela, une étude a été réalisée, en s'appuyant sur une synthèse critique des notions présentes dans l'état de l'art, tout en cherchant à proposer une démarche générique.

En lien avec ce premier objectif, nous cherchons à élaborer et à créer un outil informatique pour l'analyse de vulnérabilité. Ce dernier doit :

- être générique par rapport au modèle agro-écologique utilisé, mais également par rapport aux hypothèses sous-jacentes à l'analyse de vulnérabilité ;
- permettre la conception d'une analyse de vulnérabilité selon la démarche proposée dans le cadre de cette thèse ;

- permettre de gérer la distribution et l'interprétation des simulations ;
- être modulaire en termes d'utilisation (*i.e.* utilisation de seulement certains aspects de l'outil) et être extensible (*i.e.* facilité d'intégration de nouveaux éléments et de nouvelles fonctionnalités).

Pour le développement des applications informatiques pour l'analyse de vulnérabilité, nous souhaitons faire appel aux concepts de l'Ingénierie Dirigée par les Modèles, notamment en proposant des modèles et métamodèles pour servir de base à la conception d'un outil d'aide à l'analyse de la vulnérabilité.

Un dernier objectif était de faire la preuve de l'utilité d'une telle démarche et d'illustrer l'intérêt d'un tel outil sur des exemples concrets d'analyse de vulnérabilité des écosystèmes prairiaux au changement climatique : l'un avec un modèle d'impact simple basé sur le logiciel ModVégé [Jouven et al., 2006ab] et un autre utilisant le modèle biogéochimique d'écosystème prairial PaSim [Riedo et al., 1998] et ses améliorations les plus récentes [Vuichard 2007a] [Graux 2011].

Plan

Ce manuscrit aborde dans le premier chapitre la problématique de l'analyse de vulnérabilité des prairies au changement climatique. Il dépeint le contexte scientifique du changement climatique avec un focus particulier sur l'écosystème prairial, explique le besoin d'analyse de vulnérabilité et les problèmes associés et enfin présente le modèle de prairie PaSim (*Pasture Simulation model*). Il s'agit d'un modèle biogéochimique largement utilisé dans les études d'impact, surtout en France et en Europe, et qui est au cœur des nombreuses applications décrites dans ce manuscrit.

Le deuxième chapitre présente la littérature internationale sur la vulnérabilité et les méthodes et les outils permettant de répondre à la problématique, à savoir les plans d'expériences, le calcul intensif et l'Ingénierie Dirigée par les Modèles.

Le troisième chapitre propose des éléments de solution dans le contexte de la problématique d'analyse de vulnérabilité. Tout d'abord, en se basant sur l'étude de la littérature que nous avons effectuée, nous présentons un modèle des concepts liés à la vulnérabilité et une méthode originale d'analyse de vulnérabilité. Puis, en appui à cette approche, nous employons une approche d'Ingénierie Dirigée par les Modèles pour exposer les différentes étapes de la conception d'une solution pour l'analyse de vulnérabilité au changement climatique sous contrainte de plans d'expériences. Enfin, nous abordons l'implémentation de notre solution avec le cadre *Eclipse Modeling Framework* (EMF) présenté au chapitre précédent.

Le quatrième chapitre présente trois exemples d'analyses de vulnérabilité au changement climatique, réalisés à l'aide des solutions proposées dans le chapitre précédent. Après avoir présenté un exemple d'analyse de vulnérabilité sur une prairie permanente auvergnate, nous étudions ensuite le cas de l'analyse de vulnérabilité des prairies permanentes européennes à l'aide de simulations spatialisées. Enfin, nous documentons la démarche complète d'analyse de vulnérabilité avec et sans mise en place de mesures d'adaptation. Ce chapitre correspond également à la mise en œuvre des propositions du chapitre précédent.

Enfin, un bilan des résultats est présenté et des perspectives de recherches issues de cette thèse sont proposées dans une conclusion générale.

Chapitre 1 : La vulnérabilité des prairies au changement climatique

1.1. Introduction

La notion de vulnérabilité peut prendre de multiples définitions, les variantes apparaissent notamment lorsque le domaine concerné est précis. Cependant, une définition relativement générique pourrait être que la vulnérabilité est le degré de stress ou de perturbation qu'un système humain ou environnemental est capable de supporter avant d'être endommagé [Turner et al., 2003a]. Au cours des dernières années, la vulnérabilité est devenue un sujet central de l'étude des changements globaux (incluant le changement climatique). En effet, l'analyse de vulnérabilité fait partie du processus permettant de déterminer les menaces dues à l'exposition d'un système aux aléas climatiques, et elle participe donc à la proposition et à l'évaluation de contre-mesures face au changement climatique. Dans cette partie, nous présenterons l'intérêt d'effectuer une analyse de vulnérabilité des écosystèmes au changement climatique et les problèmes associés. Nous commencerons par rappeler ce qu'est le changement climatique et nous détaillerons quelques aspects du futur climatique. Puis, nous nous focaliserons sur un écosystème particulier, à savoir celui des prairies. Ensuite nous détaillerons le besoin d'analyse de vulnérabilité et les problèmes associés. Enfin, nous présenterons le modèle de prairie PaSim (*Pasture Simulation model*). Il s'agit d'un modèle biogéochimique qui est au cœur des nombreuses applications décrites dans ce manuscrit.

1.2. Le changement climatique

Le changement climatique se définit comme une modification durable du climat. On considère généralement que ces changements sont dus aux variations naturelles du climat, auxquels s'ajoute le changement climatique dû au réchauffement planétaire induit par l'augmentation d'origine anthropique des émissions de gaz à effet de serre (GES), modifiant la composition de l'atmosphère de la planète [IPCC 2007a]. Au cours du siècle dernier, la Terre s'est réchauffée de 0,76 °C en moyenne, et le réchauffement s'est encore accéléré. D'après l'Organisation Météorologique Mondiale (OMM³), les 13 années les plus chaudes ont toutes été enregistrées au cours des 15 dernières années. La seconde moitié du 20^{ème} siècle a été la période la plus chaude dans l'hémisphère nord depuis probablement un millier d'année. La température en Europe a augmenté d'environ 1 °C en une centaine d'années, et ce phénomène a été plus rapide que le réchauffement moyen de la planète. Même à travers la variabilité climatique, les impacts du changement climatique sont d'ores et déjà visibles : fonte de la banquise arctique, montée du niveau de la mer dans certaines régions, modification de la végétation et des pratiques agricoles (par exemple avancée des

³ http://www.wmo.int/pages/index_fr.html

dates de vendanges depuis les années 1950 de trois semaines à un mois selon les régions, en France).

Le changement climatique est associé aux concentrations de Gaz à Effet de Serre (GES), qui à travers leur pouvoir radiatif, influent sur le bilan énergétique terrestre. Parmi les principaux GES participant à l'effet de serre anthropique, on trouve une contribution à hauteur de 76,7% pour le dioxyde de carbone (CO₂), de 14,3% pour le méthane (CH₄), de 7,9% pour le protoxyde d'azote (N₂O) et enfin de 1,1% pour les halocarbures [IPCC 2007a].

Devant la complexité des interactions des différents facteurs du changement climatique, les approches de modélisation sont la seule possibilité d'amélioration de nos connaissances et de la compréhension que nous avons des phénomènes en jeu. Des scénarios sont fournis aux différents modèles climatiques existants par le Groupe d'experts Intergouvernemental sur l'Evolution du Climat (GIEC ou IPCC : Intergovernmental Panel on Climate Change). Ils intègrent des projections avec différentes évolutions possibles de la société : allant d'une forte croissance de la population, de la consommation et donc des GES émis, à des évolutions plus raisonnées. Dans son 4^{ème} rapport d'évaluation, le GIEC a proposé quatre catégories de scénarios (Figure 1, Figure 2) [IPCC 2007a] :

- Les scénarios de la famille A1 se traduisent par une croissance économique très rapide, une augmentation de la population jusqu'à 9 milliards en 2050 puis un déclin après, et par le développement de nouvelles technologies plus efficaces. Au sein de cette famille on distingue les scénarios A1FI caractérisés par une utilisation forte des combustibles fossiles, A1T caractérisés par l'utilisation de sources d'énergie autres que fossiles et A1B traduisant un équilibre des sources énergétiques ;
- Les scénarios A2 correspondent à un monde très hétérogène, avec une économie régionalisée, à accroissement continu de la population mondiale ;
- Les scénarios B1 représentent un changement rapide de l'économie vers une économie de services et d'information, l'introduction de technologies propres et une utilisation plus efficiente des ressources, couplés à une augmentation de la population mondiale jusqu'aux années 2050 puis un déclin ;
- Les scénarios B2 traduisent une économie régionalisée dans le sens d'une viabilité économique, sociale et environnementale, à accroissement continu de population mondiale.

Ces scénarios, détaillés dans le rapport SRES (Special Report on Emission Scenario), ne correspondent qu'à des représentations possibles du futur [Nakićenovič 2000]. Ils sont majoritairement utilisés par la communauté scientifique, cependant d'autres types de scénarios pourraient être envisagés tels que les scénarios d'atténuation des émissions, basés

sur une hypothèse de stabilisation de la concentration en CO₂ à la fin du 21^{ème} siècle [Arnel et al., 2002] [Tubiello et Fischer, 2007].

Finalité Mondialisation	Economie	Environnement
Globalisation (monde homogène)	A1 Croissance économique rapide (groupes : A1T, A1B, A1FI) 1,4-6,4 °C	B1 Durabilité environnementale globale 1,2-2,9 °C
Régionalisation (monde hétérogène)	A2 Développement économique régionalisé 2,0-5,4 °C	B2 Durabilité environnementale locale 1,4-3,8 °C

Figure 1. Les quatre catégories de scénarios SRES sont représentées ici très simplement selon la finalité du développement mondial et le degré de mondialisation. En réalité, ces canevas sont inclus dans un espace défini par davantage de dimensions (*i.e.* population, équité, technologie) [IPCC, 2007a]

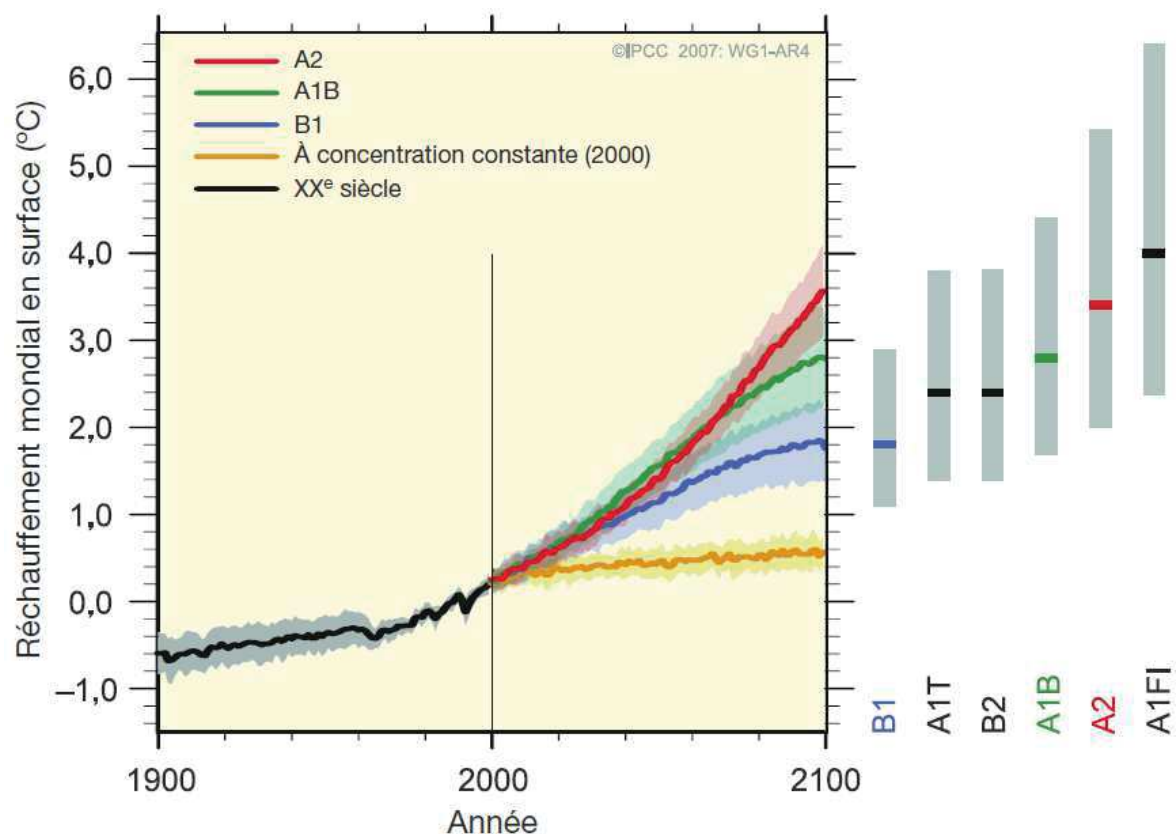


Figure 2 Moyennes multimodèles (courbes en trait plein) du réchauffement en surface pour les scénarios A2, A1B et B1 du SRES (par rapport à la période 1980-1999). A droite du graphique sont indiquées les valeurs les plus probables du réchauffement moyen (lignes horizontales) et les incertitudes associées (zones grises) selon les six scénarios SRES de référence pour la fin du siècle [IPCC 2007a]. Ces valeurs moyennes et ces fourchettes tiennent compte des projections établies à l'aide des modèles de circulation générale, couplés atmosphère-océan, ainsi que des résultats d'une hiérarchie de modèles indépendants et des contraintes liées à l'observation.

Cependant, tous les scénarios (du GIEC) se traduisent par une augmentation du CO₂ atmosphérique de 540 à 950 ppm en 2100 [IPCC 2007a], pour une concentration actuelle (2012) de 392 ppm. De même, une augmentation de la température allant de 1 à 4 °C entre 1990 et 2100 est suggérée par les modèles.

En accord avec la définition du potentiel de réchauffement global (PRG, en anglais Global Warming Potential (GWP)), qui est un moyen simple de comparer les gaz à effet de serre en tenant compte de leur pouvoir de réchauffement et de leur durée de vie [IPCC 2007a], les scénarios climatiques sont prévus pour une centaine d'année. En effet, on considère que le CO₂, premier gaz à effet de serre, a une durée de vie dans l'atmosphère d'environ 100 ans. Beaucoup d'acteurs et de décideurs ne sont en réalité intéressés que par les changements à court et à moyen terme, cependant la variabilité climatique interannuelle couvre largement le changement climatique sur de petits horizons temporels (décennies)

[Solomon et al., 2011]. L'Organisation Météorologique Mondiale a également montré qu'il fallait *a minima* une trentaine d'année pour évaluer une période climatique. Bien que le changement climatique puisse se considérer à des échelles temporelles plus grandes (plusieurs siècles), peu de scénarios les considèrent, notamment à cause de l'augmentation très forte des incertitudes au cours du temps.

Toutes ces projections climatiques contiennent une cascade d'incertitudes [Graux 2011]. Les scénarios d'émissions couvrent toute une gamme de possibilités de concentration de GES en fonction des évolutions démographiques, sociétales, technologiques et d'utilisation des terres. A cela, s'ajoute l'incertitude des modèles climatiques globaux et régionaux, qui comme tous les modèles, contiennent des erreurs structurelles et de paramétrage. C'est pourquoi le GIEC a utilisé toute une série de modèles pour les projections climatiques dans ses rapports [IPCC 2007a]. Afin d'obtenir le climat en un lieu précis, différentes méthodes de régionalisation peuvent être appliquées, chacune contenant un biais propre et ses incertitudes. Le lecteur souhaitant plus de détails sur les scénarios d'émissions, les modèles climatiques et les méthodes de régionalisation pourra se référer à [Graux 2011] et [IPCC 2007a]. C'est à partir de ces climats régionalisés que des évaluations d'impacts sont effectuées à l'aide de modèles d'impact.

Un autre aspect important du climat, associé à un haut niveau d'incertitudes, sont les événements extrêmes. En effet, s'il est clair que la canicule européenne de 2003 et celle de Russie en 2010, sont bien des événements extrêmes tant à travers l'aspect climatique que pour leurs substantielles conséquences sur la santé humaine et sur les écosystèmes (e.g., [Hegerl et al., 2011]), définir les événements extrêmes n'est pas une tâche aisée. Il n'est en effet pas facile de savoir si c'est la fréquence, l'intensité, la durée ou une combinaison des trois qui les caractérise. De même, en fonction du positionnement temporel de l'évènement climatique extrême, celui-ci peut ne pas induire d'impact extrême. On distingue deux grandes familles de définition d'un événement extrême dans la littérature : celles qui le définissent comme étant un événement climatique sortant de la distribution classique (e.g., 10^{ème} percentile) et celles qui le définissent par son impact sur le système. Le rapport [IPCC 2012] souligne la grande probabilité pour que le réchauffement global augmente la fréquence et l'intensité des événements extrêmes climatiques, notamment les canicules et la sécheresse. Ces changements vont vraisemblablement avoir de lourds impacts sur les systèmes de production agricole, dont les prairies.

1.3. Les prairies

En écologie, un écosystème se définit comme l'ensemble formé par une communauté d'êtres vivants en interaction avec son environnement biologique, géologique, hydrologique, climatique, ... Les écosystèmes qui sont modifiés/gérés par l'homme afin d'en exploiter la matière organique produite sont appelés agro-écosystèmes. On y trouve

notamment les prairies et les grandes cultures. Plus précisément, les (agro-)écosystèmes prairiaux sont l'ensemble formé par le couvert herbacé, le sol et les herbivores (si présents). C'est au sein de cet ensemble que l'on s'intéresse au cycle du carbone (C), de l'azote (N), de l'eau, et de l'énergie. Si le fonctionnement interne de l'écosystème présente un fort intérêt scientifique, ce sont les flux d'entrée et de sortie du système qui intéressent les porteurs d'enjeux (les agriculteurs par exemple), notamment à travers les exportations de fourrage, les émissions/absorptions de GES ou encore les retours à la nappe de l'eau.

En agriculture, la prairie est une culture de plantes fourragères, principalement composée de graminées et de dicotylédones (dont des légumineuses⁴), destinée à être pâturée ou fauchée. Au sein des prairies, on distingue classiquement les prairies permanentes des prairies temporaires et artificielles. Les prairies permanentes sont des surfaces herbacées composées de plusieurs espèces. Pour être considérées comme permanentes, elles doivent être des surfaces herbacées depuis un certain nombre d'années. Il n'y a pas de consensus sur cette durée qui dépend donc de la définition utilisée, comme par exemple cinq ans selon la notion adoptée par la Politique Agricole Commune (PAC). A l'inverse les prairies temporaires sont des prairies semées et périodiquement retournées (labour). Les prairies artificielles sont des prairies ensemencées exclusivement en légumineuses fourragères (pures ou en mélanges). Les prairies représentent une part importante du territoire français (11 millions d'hectares en 2008 soit 24% de la superficie du territoire national) et du secteur agricole (45% de la Surface Agricole Utile (SAU)) [Agreste 2008].

Les différents modes de conduite de l'herbe sont fonction des objectifs de production de l'exploitation (à laquelle la parcelle de prairie est rattachée), du climat, du sol et du type de prairies (permanente ou temporaire). Dans les facteurs de gestion de la prairie on trouve l'exploitation de l'herbe (fauche et pâturage), la fertilisation et parfois l'irrigation. L'herbe fauchée sert à nourrir les animaux à d'autres moments de l'année (à l'étable ou par complémentation de la ration au pâturage). La fertilisation sert à compenser les exports et à combler les carences de la prairie en azote (N), phosphore (P), potassium (K) et soufre (S).

L'éleveur agit sur les flux de matière entre la prairie et son environnement au sein du système de production. A travers la production, l'éleveur modifie la qualité des eaux superficielles et les émissions de GES émis par la prairie (CO₂, N₂O, CH₄). Pour mieux comprendre comment la gestion de la prairie peut impacter l'environnement, nous allons maintenant décrire le cycle du carbone puis le cycle de l'azote.

Schématiquement, le cycle du carbone peut se résumer ainsi (Figure 3) : le CO₂ atmosphérique est absorbé par la plante par photosynthèse, et alloué aux différents organes de la plante, que ce soit dans la partie souterraine (racines) ou aérienne (tiges, gaines, épis,

⁴ Voir glossaire pour une définition des graminées, des légumineuses et des dicotylédones

limbes, ...). Une partie du carbone est réémise par respiration, tandis qu'une autre va au sol sous forme de litière à cause de la sénescence. Le carbone de la litière rentre alors dans la matière organique du sol et est stabilisé ou réémis dans l'atmosphère par respiration en fonction de l'activité microbienne. Une fraction du carbone de la matière organique du sol peut également être perdue par lessivage. La fertilisation organique se traduit par un apport de carbone au niveau de la matière organique du sol. Dans le cas d'une prairie fauchée, un export de carbone a lieu, 30 à 40% de la biomasse fauchée étant du carbone. Lors du pâturage, le carbone contenu dans l'herbe ingérée par les animaux sera réémis dans l'atmosphère, soit par la respiration, soit sous forme de méthane entérique. Seule une petite fraction du carbone est assimilée par l'animal ou exportée dans le lait et la différence est restituée aux sols par les déjections.

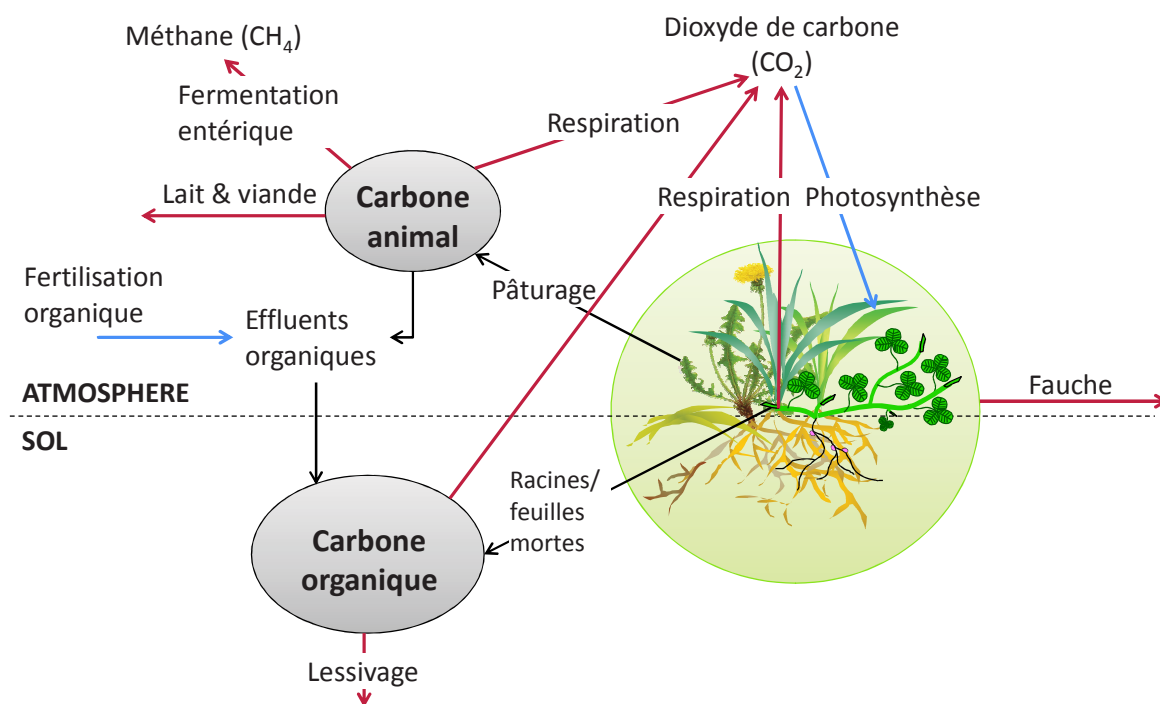


Figure 3 Diagramme du cycle du carbone en prairie. Les flèches bleues représentent les entrées du système et les flèches rouges les sorties.

Dans la prairie, le cycle de l'azote (Figure 4) et celui du carbone sont liés, cependant l'entrée d'azote dans le système se fait principalement par la fertilisation minérale (sous forme d'ammonium ou de nitrate) ou organique (fumier, lisier,...), la déposition atmosphérique d'ammoniac (vers la plante ou le sol), ou bien par fixation symbiotique des légumineuses à partir du diazote (N₂) atmosphérique. L'azote sous forme de nitrate (NO₃⁻) et d'ammonium (NH₄⁺) est assimilé par les plantes au niveau des racines dans le sol. De même que pour le carbone, une partie de l'azote de la plante peut être perdu par sénescence et entrer dans la matière organique du sol. En fonction de la dynamique de la matière organique du sol, de l'azote peut être libéré sous forme d'ammonium ou de nitrate dans le sol (minéralisation), ou inversement être immobilisé dans la matière organique. Au sein de

l'azote inorganique (minéral) du sol, des réactions de nitrification⁵ et de dénitrification⁶ ont lieu, entraînant la formation de protoxyde d'azote (N_2O) et de diazote (N_2). Une faible fraction de l'azote de la plante peut également être réémis vers le sol sous forme d'ammoniac (NH_3) à travers le phénomène de dissolution de l'azote des cuticules des feuilles par l'eau des pluies. Cet ammoniac peut lui-même être réémis vers l'atmosphère via volatilisation. Lors de la fauche une partie de l'azote est bien évidemment exportée au sein de la biomasse. De même lors du pâturage, une fraction de l'azote est assimilée par l'animal ou exportée dans le lait, tandis que le reste est rendu au système par les déjections.

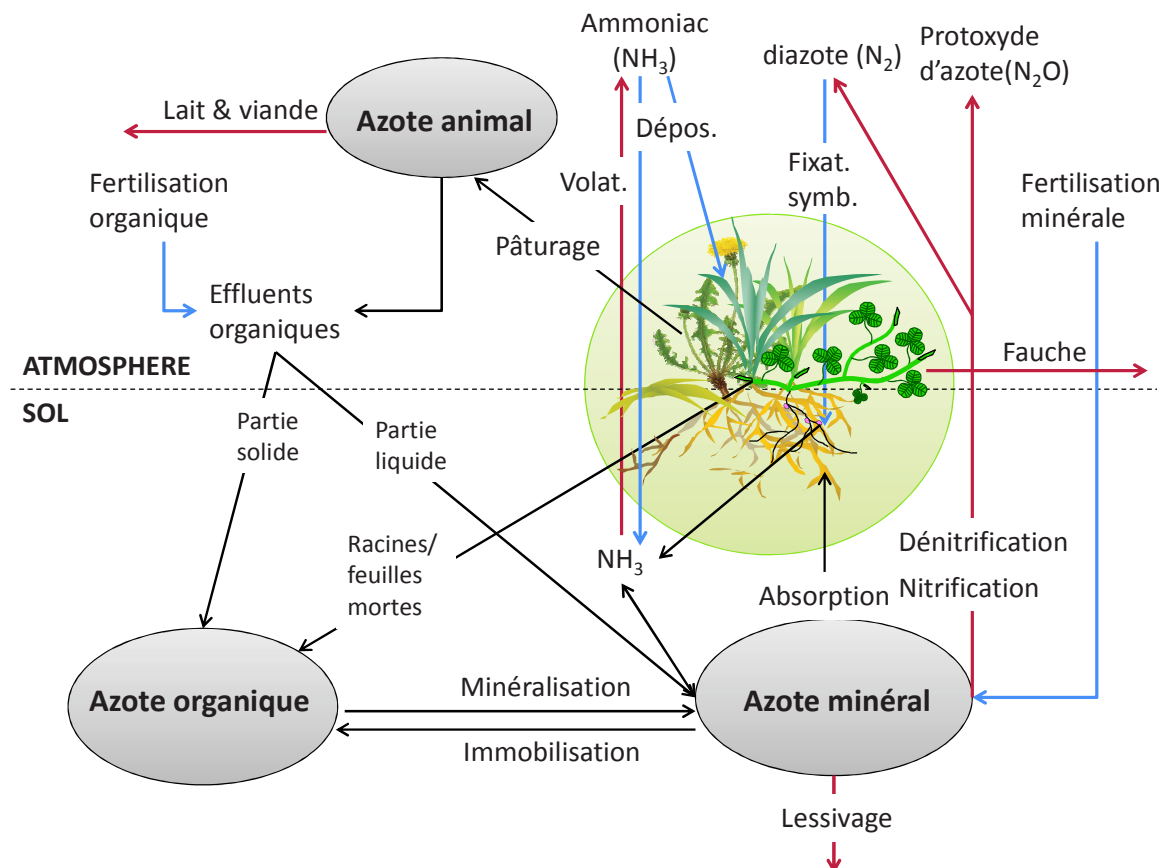


Figure 4 Diagramme du cycle de l'azote pour le système prairial "sol-plante-animal" [d'après Graux et al., 2012c]. Les flèches bleues représentent les entrées du système et les flèches rouges les sorties.

Le système fourrager, qui est l'ensemble organisé des moyens de production des fourrages, constitue la base de fonctionnement des systèmes d'élevage bovins, ovins et caprins [Graux 2011]. Il peut également être défini comme « un ensemble d'informations et de décisions visant à équilibrer les besoins et les ressources en fourrages en vue de répondre à un objectif de production » [Moreau et al., 2010]. Le système d'élevage est défini par le triptyque « *éleveur* (son projet, ses décisions), *troupeau* (ensemble des

⁵ Dénitrification : réaction permettant de passer des nitrates au diazote (N_2)

⁶ Nitrification : réaction permettant de passer de l'ammoniac au nitrate en passant par l'ammonium

animaux) qui produit et se reproduit, *ressources* (fourrages, main d'œuvre, matériel, terres) » [Landais, 1987] [Gibon et al., 1996]. Il s'agit d'un système complexe, piloté par des décisions humaines (avec un objectif de production) et dépendant des fonctionnements biologiques des animaux et des végétaux. Pour gérer ce système l'éleveur tient compte des contraintes et des aléas internes et externes. La France, à travers une grande diversité de systèmes d'élevage, est le premier éleveur bovin de l'Union Européenne (UE), avec 19 millions de têtes en 2010 soit un cinquième du cheptel européen [Agreste 2010]. La quasi-totalité des systèmes d'élevage bovins français comprennent des surfaces en herbe et sont donc par conséquent vulnérables aux impacts du changement climatique sur les prairies. Parmi les impacts attendus du changement climatique, on trouve [Graux et al., 2012a] pour la France, une augmentation de la production annuelle (à travers des opportunités de production de fourrage au printemps et à l'automne) et des baisses du drainage et de la qualité du fourrage. Ces changements sont dus principalement à l'augmentation de la température, aux changements (de quantité et de distribution temporelle) des précipitations et à l'augmentation de la concentration en CO₂ atmosphérique. L'un des principaux facteurs influençant le bilan en carbone (et en azote) du système est la gestion. Dans certains cas à gestion constante, on observe une compensation entre l'augmentation de la respiration du système et l'augmentation de la photosynthèse [Graux et al., 2012a] [Bachman et al., 2010]. De plus, l'augmentation de la séquestration de carbone due à l'augmentation de la concentration en CO₂ atmosphérique pourrait être contrebalancée par une probable augmentation de la fréquence des événements extrêmes, telles que les sécheresses. Ainsi, l'évolution des stocks de carbone dépendra fortement de l'utilisation des terres et des adaptations de la gestion des prairies et des systèmes d'élevage [IPCC 2007b]. Cependant pour définir des stratégies d'adaptation au changement climatique, il faut au préalable évaluer la vulnérabilité de ces systèmes.

1.4. Comment proposer un outil logiciel générique pour l'analyse de vulnérabilité ?

Le but de l'évaluation de la vulnérabilité n'est pas seulement d'identifier quels systèmes sont les plus vulnérables et de proposer des adaptations mais également de comprendre pourquoi ils le sont [Luers et al., 2003]. Cette information est cruciale pour les acteurs/décideurs qui doivent prioriser leurs ressources (qui sont limitées) dans le cadre d'un plan de réduction de la vulnérabilité. Si certains aspects de la vulnérabilité telle que la résilience (*i.e.* la capacité du système à revenir à son état initial après un stress) peuvent être observés expérimentalement, certains autres aspects ne peuvent être analysés qu'à l'aide d'une approche de modélisation. Compte-tenu de la cascade d'incertitudes lors des évaluations des impacts du changement climatique, une analyse de vulnérabilité au changement climatique implique un grand nombre de simulations. Une démarche d'analyse de vulnérabilité peut se résumer de la manière suivante :

1. Conception de l'analyse de vulnérabilité :
 - spécification du système étudié (par exemple la prairie) ;
 - choix des variables d'intérêt (par exemple les flux de GES) ;
 - définition des facteurs d'exposition (par exemple une projection climatique selon le scénario A1B [IPCC 2007a]) ;
 - spécification de l'échelle spatiale (par exemple la parcelle ou la région) ;
 - choix des références : période temporelle de référence, seuil de vulnérabilité ...
 - définition des horizons temporels (par exemple, "futur lointain" : 2040-2069) ;
2. Construction d'un plan d'expériences (de type "Analyse de sensibilité") ;
3. Réalisation des simulations à l'aide du ou des modèles précédemment choisi(s) ;
4. Calcul de statistiques sur les résultats précédents (par exemple indice quantitatif ou qualitatif de vulnérabilité) ;
5. Interprétation des résultats ; dans certains cas, comme la prise en compte de stratégie d'adaptation, des simulations supplémentaires peuvent être réalisées.

Cependant, si la majorité des approches contiennent ces étapes, il n'y a actuellement aucun consensus sur la définition de la vulnérabilité et encore moins sur la méthode à appliquer. De plus, les approches quantitatives ou génériques sont rares (par exemple [Luers et al., 2003], [Vías et al., 2010]). C'est pourquoi une des premières questions de cette thèse a consisté en une étude du concept de vulnérabilité. Cette étude s'est effectuée en s'appuyant sur une revue de l'état de l'art permettant de proposer une démarche générique. De plus, l'analyse de vulnérabilité se compose d'un certain nombre d'étapes incluant *a minima* la conception du plan d'expériences, la réalisation des simulations (ayant souvent besoin d'une importante puissance de calcul) et le calcul de statistiques (par exemple indices de vulnérabilité), comme illustré dans [Lexer et Seidl, 2009] [Metzger et al., 2005] [Seidl et al., 2010] [UNFCCC 2004] [Warren et al., 2008]. D'autres étapes peuvent également avoir lieu, telles que la construction de surfaces de réponse [Luers 2005], l'optimisation (adaptation/ atténuation) de vulnérabilité [Luers et al., 2003]. Tout cela implique une certaine diversité de tâches.

A notre connaissance, aucun outil ne permet de faciliter la majeure partie des tâches énoncées précédemment. Dans ce contexte, notre principale problématique a été de proposer un outil permettant de faciliter l'analyse de vulnérabilité au changement climatique. Notre second souhait était que cet outil soit générique par rapport au modèle agro-écologique utilisé, mais également par rapport aux hypothèses sous-jacentes. De plus, cet outil devrait permettre la conception d'une analyse de vulnérabilité suivant la démarche proposée dans le cadre de cette thèse, mais également permettre de gérer la distribution de simulations, et leur interprétation.

D'autre part, il nous semble que la conception d'un tel outil devrait s'appuyer sur l'ingénierie dirigée par les modèles et ce, afin de s'appuyer sur les meilleures avancées en termes de Génie Logiciel. L'Ingénierie Dirigée par les Modèles (IDM) est une branche du Génie Logiciel qui s'intéresse, depuis plus d'une dizaine d'années, au développement, à la maintenance et à l'évolution des logiciels, avec une approche de modélisation unificatrice [Favre et al., 2006]. L'IDM repose sur trois concepts fondamentaux : le "modèle", le "métamodèle" et la "transformation de modèle". Le "modèle" est une représentation simplifiée d'un système étudié et se définit par la relation « **est une représentation de** » entre lui-même et le système étudié [Hill 1996] [Atkinson et Kuhne, 2003] [Seidewitz 2003] [Bézivin 2004]. Dans le cadre de l'IDM, le modèle peut également être défini comme « une description d'un (ou d'une partie d'un) système écrit dans un langage bien-défini » [Kleppe et al., 2003]. La notion de langage bien-défini fait appel indirectement au deuxième principe de l'IDM, à savoir le "métamodèle", qui peut être vu comme « un modèle qui définit un langage pour exprimer un modèle » [OMG 2002]. Cela se traduit par la relation le modèle « **est conforme au** » métamodèle [Bézivin 2005]. A la différence des deux autres principes, il n'y a pas de consensus sur la définition du troisième concept fondamental de l'IDM, à savoir la transformation de modèle [Rahim et Mansoor, 2008] [Lano et Clark, 2008] [Iacob et al., 2008]. D'après [Favre 2004a], elle devrait pouvoir se définir par la relation « **est transformé en** » entre le modèle source et le modèle cible. On pourrait également la voir comme « la génération d'un ou de plusieurs modèles cibles à partir d'un ou plusieurs modèles sources » [Bézivin 2004]. Parmi les langages de méta-modélisation, on trouve notamment Eclipse Modeling Framework (EMF, [Eclipse-EMF]), un outil dans lequel la transformation de modèles revient à l'exécution d'un métaprogramme. Il s'agit, en effet, d'une plateforme de modélisation et de génération de code qui fut utilisée dans le cadre de cette thèse pour implémenter l'outil répondant à notre problématique.

L'analyse de vulnérabilité nécessite un grand nombre de simulations, réalisées à l'aide de modèles d'impacts dont le temps de calcul est généralement non négligeable. Par exemple, dans le cadre de cette thèse, environ 200 000 heures (environ 23 ans) de temps de calcul mono-CPU ont été nécessaires. Pour cela, l'utilisation du calcul haute performance était nécessaire. A la jonction des mathématiques et de l'informatique, le calcul intensif et la simulation illustrent bien le triptyque scientifique fondateur « théorie-expérience-simulation » [ANR 2010]. Dans les environnements pour le calcul distribué, on distingue notamment les fermes de calcul et les grilles. Une ferme de calcul ou "cluster" est un ensemble d'unités de calculs regroupées physiquement en un même lieu afin d'additionner les puissances élémentaires et de fournir une plus grande puissance de calcul et/ou de stockage distribuée. Si les clusters permettent "simplement" d'accéder à une grande puissance de calcul, ils ne sont cependant pas toujours suffisants pour faire face à des demandes massives. De plus l'utilisation de ces ressources n'est pas continue (besoins non uniformes dans le temps). Il a donc été décidé par un certain nombre de centres de

recherche, de mettre en commun leurs ressources [Shiers 2009] [Foster et al., 2001]. On passe alors d'une puissance de calcul de l'ordre de grandeur de quelques centaines de processeurs à des dizaines et centaines de milliers de processeurs, dans le cadre d'une grille de calcul [Foster et Kesselman, 1999]. Depuis la naissance du concept dans les années 1990 jusqu'à maintenant, la grille (vue comme une infrastructure de calcul distribué pour la science et l'ingénierie de pointe) a pris une place centrale en tant qu'outil dans les domaines de simulation. Certains outils, tels que OpenMOLE⁷ [Reuillon et al., 2010], permettent de gérer la distribution de calculs tout en masquant l'hétérogénéité des environnements. Ce dernier a été choisi comme outil de distribution de plan d'expériences dans le cadre de cette thèse. Il a ainsi permis de distribuer l'équivalent de 15 années de calcul mono-CPU sur la grille d'organisation virtuelle Biomed.

De plus, il nous paraissait nécessaire de faire appel au domaine des plans d'expériences pour réduire le nombre de simulations et donc les temps de calcul. La conception des plans d'expériences (DOE, "Design of Experiment") a une histoire riche, incluant de nombreuses avancées théoriques et des applications pratiques dans des domaines variés. Depuis le début des simulations par ordinateur jusqu'à nos jours, le domaine de la conception des plans d'expériences est un champ de recherche très actif [Kempthorne 1952] [Amblard et al., 2003]. En effet, les DOE sont, dans le domaine de la modélisation, un outil nécessaire pour tester et analyser efficacement le comportement d'un modèle [Kleijnen 1987]. Les DOE font partie intégrante du processus de développement d'un modèle [Kleijnen et al., 2005]. En effet, que ce soit pour la vérification, la validation ou les différentes utilisations possibles du modèle, un grand nombre de simulations du modèle sont nécessaires. En particulier, dans la modélisation des environnements dynamiques, les modèles sont devenues incroyablement plus complexes au fur et à mesure que la puissance de calcul des ordinateurs augmentait. Du fait du grand nombre de paramètres requis par le modèle et du temps de calcul nécessaire pour une seule simulation, le temps nécessaire pour un plan d'expériences factoriel, uniforme et complet est généralement trop coûteux pour les meilleures machines séquentielles. C'est pourquoi l'utilisation d'autres types de plans d'expériences, ainsi que la distribution des simulations sont nécessaires. L'utilisation d'un plan d'expériences approprié permettra d'avoir toutes les informations recherchées, tout en ayant le nombre minimum de simulations nécessaires pour un résultat de bonne qualité et pour un temps de calcul aussi petit que possible. Par exemple, dans le cas de l'analyse de sensibilité (qui est une des composantes de l'analyse de vulnérabilité), le plan d'expériences est important pour avoir l'information de la sensibilité de tous les paramètres, sans pour autant négliger leurs interactions. La réalisation de ces plans d'expériences permet également de calculer des surfaces de réponse, qui sont des approximations de la fonction Entrée/Sortie du simulateur. Dans le cadre de cette thèse un certain nombre de plans d'expériences (par exemple des plans "hypercubes latins" [McKay

⁷ <http://www.openmole.org/>

et al., 1979]) ont été réalisés, tandis que des surfaces de réponses ont été établies par régression (par exemple des régressions quadratiques).

Cette thèse se situe dans le contexte applicatif d'un "work package" (WP) du projet européen EU-FP7 ANIMALCHANGE (voir annexe II.iii). Si la première partie de la thèse (révision du concept de vulnérabilité pour proposer une démarche générique) n'est pas directement influencée par les projets, en revanche le cadre applicatif, lui, est fortement dépendant des projets. En effet, une application est attendue afin de valider la faisabilité d'une telle approche et valider l'aspect opérationnel de l'outil développé. Pour cela nous présenterons un exemple applicatif complet sur l'analyse de vulnérabilité des écosystèmes prairiaux. Bien que l'exemple applicatif se focalise sur les prairies, une forte généricité par rapport à l'agro-écosystème est cependant attendue. L'objectif global du projet ANIMALCHANGE est de fournir des orientations scientifiques permettant l'intégration d'options d'adaptation et d'atténuation dans des politiques de développement durable de la production animale en Europe, en Afrique du nord et sub-saharienne, et en Amérique latine. Le projet ANIMALCHANGE va contribuer à l'élaboration des politiques publiques en Europe, et portera des actions de coopération concernant de petites exploitations d'élevage dans certains pays en voie de développement. Le "work package 15" (WP15, « *Data infrastructure and model support* »), notamment, doit servir à fournir un support méthodologique et technique au projet. Cela inclut la base de données (tâche 15.1), des outils pour l'évaluation et le paramétrage des modèles (tâche 15.2) et enfin des méthodes d'analyse de vulnérabilité (tâche 15.3). Ainsi, les résultats de la tâche 15.3 (*i.e.* les méthodes d'analyse de vulnérabilité), alimentés par les résultats de cette thèse, seront appliqués dans d'autres parties du projet, notamment dans la tâche 5.2 du WP5 (« *Climate change vulnerability of livestock systems* ») portant sur l'analyse de vulnérabilité des systèmes d'élevage au changement climatique et dans le WP8 (« *Integrating adaptation and mitigation options* ») concernant les stratégies d'adaptation et d'atténuation sur des systèmes d'exploitation d'élevage, où la vulnérabilité avec application de stratégies sera évaluée.

Ce projet est lui-même lié à deux autres projets européens : EU-FP7 GHG-Europe⁸ (voir annexe I.ii) et EU-FP7 CARBO-Extreme⁹ (voir annexe I.i). D'une part, le projet GHG-Europe a pour but d'améliorer notre capacité de prédiction et notre compréhension des bilans sur l'Europe des émissions de GES (CO₂, CH₄, N₂O) et du cycle du carbone terrestre. D'autre part, le but du projet CARBO-Extreme est de parvenir à une meilleure connaissance du cycle du carbone terrestre en réponse à la variabilité du climat et aux phénomènes extrêmes, de représenter et d'appliquer ces connaissances sur l'Europe à travers une modélisation prédictive du cycle du carbone terrestre, et enfin d'interpréter les prédictions des modèles en termes de vulnérabilité des compartiments de carbone terrestre

⁸ <http://www.ghg-europe.eu/>

⁹ <http://www.carbo-extreme.eu/>

(en particulier ceux du sol) et d'en déduire des conseils à donner à l'Union Européenne pour des politiques de protection des sols. Dans ce cadre, des données et des méthodes sont partagées (par exemple des données climatiques). De plus, d'objectifs parfois identiques (analyse de vulnérabilité) découle une certaine richesse de méthodes (par exemple analyse de la sensibilité et de l'exposition de manière probabiliste pour CARBO-Extreme, et analyse de vulnérabilité par des indices dans le cadre d'ANIMALCHANGE). Ainsi cette thèse, qui se situe en amont par rapport au projet ANIMALCHANGE, doit servir également à "alimenter" le projet CARBO-Extreme (en remplissant également les tâches demandées dans le WP7, « Carbon vulnerability synthesis »).

1.5. Description de PaSim

Notre objectif étant d'évaluer la vulnérabilité des prairies au changement climatique, il était naturel de choisir un modèle mécaniste détaillant suffisamment finement les mécanismes du cycle du carbone, de l'azote, de l'eau et de l'énergie de façon à prendre en compte autant que possible les impacts d'un climat futur.

Plusieurs modèles pourraient répondre à notre problématique. Parmi les modèles classiques de prairie à l'échelle de la parcelle et sous conditions climatique tempérées, on trouve notamment DayCent ([Parton et al., 1998], [Del Grosso et al., 2001]), DNDC ([Li 2000] [Li et al., 1992]), GRASIM ([Mohtar et al., 1997] [Mohtar et al., 2000]), HPM ([Thornley 2001], [Thornley et Cannell, 1997]), PaSim ([Riedo et al., 1998]) et STICS ([Brisson et al., 1998], [Brisson et al., 2002], [Ruget et al., 2006]). Tous ces modèles sont mécanistes, déterministes et dynamiques. Les points forts de PaSim par rapport à nos objectifs et comparativement aux autres modèles sont :

- Il a déjà été utilisé pour une gamme assez large de variables d'intérêt : GES, production fourragère, production animale, lessivage, stocks de carbone, ... (e.g., [Vuichard et al., 2007b] [Graux et al., 2012abc]). PaSim est le plus complet des modèles cités précédemment en termes de disponibilité de variables de sortie. Par exemple, STICS, DNDC et DayCent ne gèrent pas réellement le pâturage (sinon en le "simulant" par des fauches) et ne peuvent par conséquent donner aucune information de production animale.
- Il possède à la fois un cycle du carbone et de l'azote finement couplés et suffisamment détaillés, ainsi qu'une représentation détaillée des processus végétaux (à la différence de DayCent et DNDC) pour permettre des simulations sous changement climatique. Il est par conséquent opérationnel et a déjà été utilisé pour des simulations d'impact. Une connaissance *a priori* du comportement du modèle est toujours une aide pour l'évaluation des résultats. De plus, l'existence de campagnes de simulation pouvant faire partie des étapes d'une analyse de vulnérabilité est une aide à la conception d'un outil pour une analyse de

vulnérabilité sous changement climatique. Il est ainsi, par exemple, possible de les utiliser dans une approche de rétro-ingénierie des plans d'expériences.

- Il inclut de manière fine l'interaction herbe-animal et la prédiction des productions animales, à la différence de la majorité des autres modèles de prairies.
- La prise en compte de la dynamique horaire du climat par PaSim et HPM est également un point fort pour la modélisation des événements extrêmes et du changement climatique car il y a changement de distribution et d'intensité horaire du climat (e.g., [Hanel et Buishand, 2010]).

Toutes ces raisons confortent le choix du modèle, qui était déjà en partie fortement suggéré par le cadre applicatif de cette thèse (*i.e.* le projet ANIMALCHANGE). En effet, même si une bonne marge de liberté était accordée dans le choix des méthodes, les données d'entrée sont fournies par le projet et certaines simulations sont à réaliser avec PaSim dans le cadre de ces mêmes projets. D'autre part, ma connaissance du modèle (3 ans d'utilisation et de développement), sa disponibilité pour utilisation, et le fait qu'il soit maintenu au sein de notre unité ont également influés sur le choix du modèle PaSim, décrit ci-après, pour répondre à nos objectifs.

PaSim (<https://www1.clermont.inra.fr/urep/modeles/pasim.htm>) est un modèle de prairies permanentes ou semées, initialement développé par [Riedo et al., 1998] et inspiré du *Hurley Pasture Model* de [Thornley et al., 1998] pour sa simulation du couvert en réponse à la lumière, à la température et aux ressources carbone, azote et eau. Ce modèle simule les flux de matière et d'énergie au sein du système sol-végétation-animal en réponse aux conditions pédoclimatiques et aux pratiques, à un pas de temps proche de la demi-heure sur une à plusieurs années enchaînées. Il prend en compte les pratiques de fertilisation minérale ou organique, d'irrigation, de retournement du sol et d'exploitation de l'herbe par la fauche et/ou le pâturage (Figure 5). La gestion peut être prescrite ou déterminée/optimisée par le modèle [Vuichard et al., 2007b] [Graux 2011]. Les simulations se limitent à l'échelle de la parcelle, et les animaux ne sont modélisés qu'au pâturage (c'est-à-dire que les périodes d'étable ne sont pas simulées). La végétation y est représentée à l'échelle de la communauté et a été initialement paramétrée pour une prairie permanente, composée majoritairement de graminées et de légumineuses.

PaSim est composé de plusieurs sous-modèles représentant les processus de la végétation, du microclimat, de la biologie et de la physique du sol, des herbivores. Il se compose actuellement de 60 000 lignes de code majoritairement en Fortran 90 et 95, compilé à l'aide du compilateur Intel. Récemment, le Fortran 2003, qui est une version du langage supportant la programmation objet, a commencé à être introduit dans le modèle. Cependant il n'utilise encore que faiblement ces propriétés. En plus d'un guide utilisateur détaillé [INRA-UREP 2011], une interface utilisateur a été développée. La version 5.3 du modèle a été déposée à l'Agence pour la Protection des Programmes (APP) sous le numéro IDDN.FR.001.220024.000.R.P.2012.000.10000.

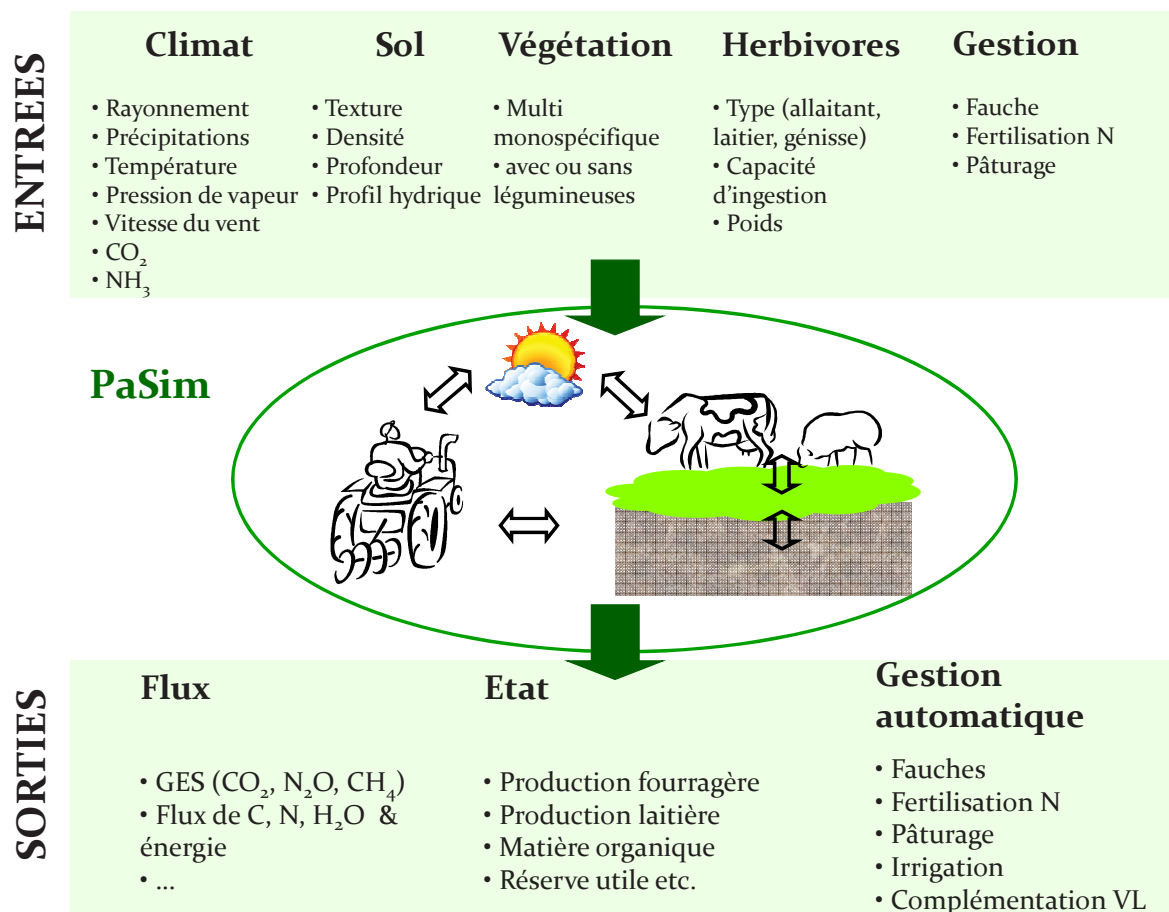


Figure 5 [Graux 2011] Schéma illustratif des entrées/sorties du modèle PaSim.

Le carbone assimilé par photosynthèse est alloué dynamiquement à un compartiment racinaire et à trois compartiments aériens (limbes, tiges et gaines, épis). De façon à modéliser la sénescence, chacun de ces compartiments est décliné en quatre sous-classes d'âges. Le carbone est réémis par respiration, par production de méthane entérique par les herbivores ou via les exportations liées à la fauche et à la production de lait. L'azote est apporté au système via la déposition d'azote atmosphérique, l'apport de fertilisants azotés, la fixation symbiotique par les légumineuses. L'azote minéral du sol, disponible pour la croissance de la végétation, peut être réduit par immobilisation, par lessivage, par volatilisation de NH₃ et via les processus de nitrification/dénitrification qui conduisent à la production de N₂O. Les processus de sol sont basés sur le modèle CENTURY de [Parton et al., 1988]. Le carbone organique du sol est représenté par cinq compartiments de vitesses de décompositions différentes (de 0,5 à 1500 ans d'âge moyen) (Figure 6). Il est basé sur des équations du premier ordre, c'est-à-dire que les flux de sortie des compartiments sont fonction de leur taille respective. Le même principe est utilisé pour les compartiments d'azote, cependant les flux d'azote sont induits par les flux de carbone via un ratio C/N. Les vitesses de décomposition sont fonction de la température, de l'humidité et de la texture du sol, ainsi que d'autres variables. Comme la majorité des modèles

biogéochimiques, l'initialisation du modèle se fait par mise à l'équilibre (des variables d'état) du modèle avec le climat et la gestion, faute de données et de méthodes pour paramétrer finement les différents horizons de sol. La méthode native de mise à l'équilibre consiste à réaliser une simulation sous des conditions de climat et de gestion répétées jusqu'à stabilisation du système (en fait de la matière organique, car c'est la composante qui a le plus grand temps de résidence, et est par conséquent la plus lente à se stabiliser). De façon à accélérer cette mise à l'équilibre une méthode algébrique a été développée [Lardy et al., 2011b]. Le principe de cette méthode est dans un premier temps d'écrire les équations différentielles de la matière organique sous forme matricielle. Puis par récurrence on construit une suite de matrices liant les valeurs des compartiments à l'instant du premier pas de temps aux valeurs des compartiments à la fin du cycle climatique. L'état d'équilibre se traduisant par des valeurs identiques à ces deux instants, il suffit alors de résoudre le système d'équations linéaires associées. On peut alors forcer le modèle à l'aide de ces valeurs d'équilibre. Cependant, les flux entrants dans le sol sont dépendants de la plante qui est elle-même sensible à la quantité de matière organique du sol. Ainsi quelques itérations de cycle météo sont nécessaires pour équilibrer le sol et la plante.

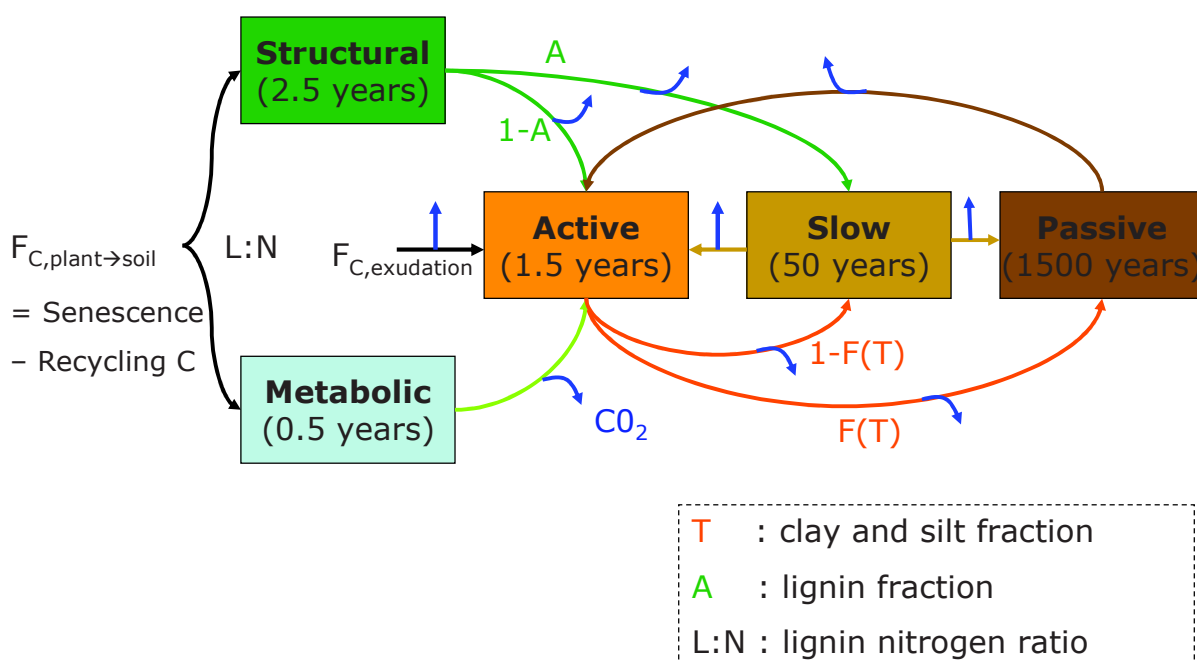


Figure 6 [Lardy et al., 2011b] Diagramme du cycle du carbone entre les compartiments du modèle de décomposition de matière organique du sol de PaSim.

Dans PaSim, les formalismes des processus biophysiques du sol, basés sur ceux du modèle CENTURY [Parton et al., 1988], ont été améliorés par [Schmid et al., 2001] pour la simulation des émissions de N_2O , puis par [Riedo et al., 2002] pour la simulation des émissions de NH_3 . Puis [Vuichard et al., 2007a] améliorèrent les processus de sénescence, les effets de la qualité de l'herbe sur les émissions de méthane, le stress hydrique ainsi que l'optimisation des pratiques de la prairie (*i.e.* détermination de la fauche, de la fertilisation

(date et quantité) et du pâturage (date, durée et nombre d'animaux) par le modèle). Depuis, le paramétrage de la végétation et du stress hydrique ont été améliorés [Graux 2011] et les réserves de la végétation ont été ajoutées (voir les annexes de [Graux 2011]). La possibilité de simuler des prairies temporaires a également été ajoutée par [Graux 2011]. En parallèle, le calcul de la température du sol a été amélioré : ajout d'une couche litière jouant le rôle d'isolant et ajustement automatique de la température de la couche limite basse (voir les annexes de [Graux 2011]). Un nouveau module animal simulant l'ingestion et les émissions de méthane des bovins au pâturage (vaches allaitantes avec leurs veaux, vaches laitières, génisses) en réponse au climat et à la gestion, a été développé par [Graux et al., 2011]. La gestion optimisée par le modèle a été fortement améliorée par [Graux 2011] et [Graux et al., 2013]. Dans le cadre d'une collaboration avec l'Institut de Radioprotection et de Sécurité Nucléaire (IRSN), les cycles du carbone 14 (^{14}C) et du carbone 13 (^{13}C) ont été ajoutés à celui du carbone existant (*i.e.* celui du ^{12}C) [Aulagnier et al., 2012] [Duclos 2010]. Ce travail a permis d'évaluer la fiabilité du modèle quant au temps de résidence du carbone dans les différents compartiments. Le lecteur pourra se référer à [Graux 2011] pour une description détaillée des équations et dans le guide utilisateur ([INRA-UREP 2011]¹⁰) pour une description de l'utilisation du modèle.

Il est tout de même à noter que PaSim possède certaines limitations par rapport à nos objectifs :

- Un temps de calcul non négligeable (environ 10 minutes pour une mise à l'équilibre du modèle par rapport à un climat et à une gestion, et une heure pour simuler 150 ans) ;
- Un besoin de données en entrée important : climat horaire, gestion détaillée (dates de fauches, dates de fertilisation et nature de cette dernière, dates et intensité du pâturage, propriétés des animaux au pâturage), propriétés du sol, paramètres de la végétation, etc. ...
- Capacité non complète et non suffisamment évaluée à reproduire certains événements extrêmes (notamment à cause du manque actuel de connaissances et de données expérimentales) ;
- Faible représentation du rôle fonctionnel de la diversité végétale pour les cycles biogéochimiques et non prise en compte de l'hétérogénéité au sein de la parcelle.

C'est dans ce contexte et en connaissant bien les atouts et faiblesses que le modèle PaSim a été utilisé comme modèle support pour la réalisation d'analyses de vulnérabilité des prairies au changement climatique.

¹⁰ https://www1.clermont.inra.fr/urep/modeles/Pasim_User_Guide-pasim_v5-3_201212.pdf

1.6. Conclusion

Nous avons vu dans ce chapitre que le changement climatique va entraîner des modifications significatives par rapport au climat actuel. De fortes incertitudes existent sur l'intensité de ces changements et leurs impacts. Il est toutefois très probable que les systèmes agro-écologiques seront lourdement impactés, créant ainsi des opportunités (par exemple l'augmentation de la production hivernale des prairies) et des risques (par exemple la baisse de la qualité du fourrage). La prairie représente une part importante de ces systèmes (environ 45% de la surface agricole utile française) et est quasiment incontournable dans les systèmes d'élevage bovins. Elle pourrait subir des changements qui nécessitent une forte adaptation des pratiques de gestion. Afin de concevoir des stratégies d'adaptations, il convient d'abord d'évaluer les faiblesses potentielles (vulnérabilités). L'analyse de vulnérabilité nécessite un certain nombre d'étapes incluant la génération et la réalisation de plan d'expériences puis l'analyse des résultats ainsi générés. C'est pourquoi la création d'une approche générique et d'un outil l'implémentant et facilitant la réalisation des différentes étapes de cette analyse serait un gain pour le chercheur. Ce travail se place dans le contexte applicatif de l'analyse de vulnérabilité des prairies au changement climatique. En effet, tout au long de la thèse, c'est-à-dire des phases préliminaires de rétro-ingénierie jusqu'à l'analyse de vulnérabilité de la matière organique des sols prairiaux au changement climatique, des simulations de prairies sont utilisées. La réalisation de ces objectifs nécessite de faire appel aux concepts modernes de conception logicielle (Ingénierie Dirigée par les Modèles), au calcul à hautes performances, ainsi qu'à la conception et à l'analyse des plans d'expériences. Ces domaines et leurs apports par rapport à nos objectifs sont décrits dans le chapitre suivant, après une discussion sur l'analyse de vulnérabilité au changement climatique.

Chapitre 2 : Etat de l'art : notion de vulnérabilité et éléments de calcul

2.1. Introduction

Ce travail de thèse se place dans un contexte fortement interdisciplinaire centré sur la modélisation et la simulation, alliant des domaines informatiques et mathématiques (calcul intensif, ingénierie dirigée par les modèles, plans d'expériences) aux domaines de l'agronomie et de l'écologie (prairie, changement climatique et analyse de vulnérabilité sous ce contexte).

L'analyse de vulnérabilité peut être vue comme s'apparentant à une forme d'analyse de sensibilité. En tant que telle, elle nécessite de réaliser un très grand nombre de simulations. C'est pourquoi afin de réduire ce nombre et de maximiser l'efficacité des simulations réalisées, nous ferons appel au domaine de la conception des plans d'expériences ou « *Design of Experiments* » (DOE). Même si ces techniques permettent de réduire sensiblement le nombre de simulations nécessaires, il n'en reste pas moins important de faire appel au calcul intensif pour réaliser ces dernières dans un temps d'attente utilisateur réduit. De plus, l'analyse de vulnérabilité n'est pas forcément facile à mettre en œuvre. En effet, elle se compose d'un assez grand nombre d'étapes (définitions des simulations, distribution de ces dernières, analyse des résultats, création de surfaces de réponse, calcul d'indices de vulnérabilité, ...). Pour la mise en œuvre informatique nous nous sommes penchés sur les dernières avancées du génie logiciel inspirées de la modélisation, en particulier l'Ingénierie Dirigée par les Modèles (IDM).

Ce chapitre décrit les quatre domaines-clé sur lesquels notre thèse s'appuie. Tout d'abord, nous verrons la problématique de la vulnérabilité au changement climatique (2.2) à travers sa définition, ses concepts, ses indices. Ensuite, nous détaillerons la problématique de la conception et de l'analyse des plans d'expériences (2.3). La partie suivante (2.4) expliquera les principes du calcul intensif (cluster et grille), les outils de distribution de simulations et plus particulièrement l'outil OpenMOLE qui a retenu notre attention pour distribuer nos simulations. Enfin, la dernière partie de ce chapitre (2.5) décrira les principes de l'Ingénierie Dirigée par les Modèles et son application dans l'outil Eclipse Modeling Framework (EMF).

2.2. La vulnérabilité (au changement climatique)

Les sciences humaines furent le premier domaine dans lequel on étudia la vulnérabilité, notamment dans les études sur les risques de danger et en géographie [Kasperson et al., 2005]. La vulnérabilité y est généralement définie comme un potentiel de

perte [Mitchell et al., 1989]. On distingue souvent deux aspects : un côté externe à travers l'exposition (c'est-à-dire les chocs et les perturbations auxquels le système est soumis), et un aspect interne qui représente la capacité (ou l'absence de capacité) à répondre et à récupérer d'un stress externe [Chambers 1989]. Dans l'écologie, les études se sont plus centrées sur le concept opposé de stabilité. La définition précise de la stabilité dépend de l'écosystème en question, de la variable ou des variables d'intérêt, et du contexte global. L'écologie emprunte notamment au domaine des systèmes dynamiques certaines de ses idées. En raison de l'utilisation du concept de stabilité dans des domaines différents, beaucoup ont proposé d'employer des termes plus spécifiques.

Les concepts écologiques associés à la stabilité prennent une connotation dans un continuum de mesures telles que constance, résistance et inertie (ou persistance), **résilience**, **élasticité** et **amplitude** [Grimm et Wissel, 1997], qui apparaissent implicitement ou explicitement dans la littérature de la vulnérabilité. Si la constance est utilisée pour décrire des systèmes qui peuvent demeurer sans changements, les autres termes traitent des réactions du système à une perturbation. La persistance (ou inertie) implique que le système vivant est capable de résister à des fluctuations externes, tandis que la résistance (ou **sensibilité**) est une mesure de la façon dont une variable d'intérêt peut changer en réponse aux pressions externes. La **résilience** est la capacité, pour un système, à revenir dans un état précédent après une perturbation. Plus précisément, en écologie, la **résilience** peut être définie à la fois comme le maximum de perturbation qu'un système peut supporter avant de changer d'état et comme la capacité d'un groupement de structure et de processus à se renforcer mutuellement afin de résister à des perturbations et des stress [Holling 1973] [Brand et Jax, 2007]. Les concepts complémentaires d'**élasticité** et d'**amplitude** sont des mesures de résilience. L'élasticité est la vitesse avec laquelle un système retourne à son état initial. L'**amplitude** (ou **résilience écologique**) est une mesure de la distance maximale de laquelle un système peut être déplacé à partir d'un état précédent en conservant la possibilité de pouvoir encore y retourner.

Le but de l'évaluation de la vulnérabilité n'est pas seulement d'identifier quels systèmes sont les plus vulnérables mais également de comprendre pourquoi ils le sont [Luers et al., 2003]. Cette information est cruciale pour les décideurs qui doivent prioriser leurs ressources (qui sont limitées) dans le cadre d'un plan de réduction de la vulnérabilité. La vulnérabilité peut donc être le but de l'analyse ou une étape intermédiaire dans le processus d'apprentissage et d'adaptation [Folke et al., 2005] [UNFCCC 2004].

Pearson et al. [2011] ont estimé, à travers leur revue des méthodes d'analyse de vulnérabilité et des modèles de recherche appliqués à l'agriculture australienne, que du fait que la nature de la vulnérabilité est spécifique à chaque situation, il n'est pas possible de trouver d'approche unificatrice pour conceptualiser la vulnérabilité. Ils soulignent aussi, comme l'a montré [Adger 2006] dans sa revue, que depuis 2000 il y a eu une évolution des méthodes, depuis une estimation linéaire des impacts jusqu'à des méthodes plus

intégratives comprenant une approche systémique. Parmi ces nouvelles méthodes on trouve notamment celle de [Warren et al., 2008] et [Nelson et al., 2010].

Cette section sur la vulnérabilité se décompose en quatre sous-parties. Tout d'abord, nous essaierons de définir ce que sont la vulnérabilité et tous les autres concepts qui lui sont liés. Ensuite nous verrons en détail la méthode proposée par [Luers et al., 2003], [Luers 2005], qui a l'avantage d'être complète. Puis, nous passerons en revue les autres méthodes existantes. Enfin, nous nous intéresserons au cas particulier des cartes de vulnérabilité dans le cadre du changement climatique.

2.2.1. Définitions des concepts de la vulnérabilité

Une difficulté pour définir la vulnérabilité peut être qu'elle n'est pas souvent un phénomène directement observable [Downing et al., 2001]. Il s'agit, en fait, d'une notion relative. En effet, peut-on dire qu'un système est vulnérable, sans comparaison à un cas de référence ? Que ce soit explicite ou implicite la vulnérabilité est toujours comparée à un système de référence. Une autre difficulté est qu'elle peut changer selon l'échelle de temps considérée dans notre étude. Comme Fussel [2007] l'a souligné pour décrire une situation de vulnérabilité, il faut préalablement bien définir le système analysé (par exemple, l'écosystème prairial), les aspects considérés (par ex. le rendement), les stress externes (par ex. le changement climatique) et la référence temporelle (par ex. une période de trente ans). En effet, considérons, par exemple, deux prairies (A et B) soumises à une sécheresse au milieu de l'été. On observe que juste après la sécheresse la prairie A est beaucoup plus touchée (70% de biomasse en moins) que la prairie B qui n'a subi qu'une baisse de 30%. On en déduira alors que la prairie A est plus vulnérable que la prairie B. En même temps, si l'on regarde ces mêmes prairies au milieu de l'automne, alors la prairie A a intégralement récupéré de la sécheresse tandis que la prairie B est encore affectée d'une baisse de production. La prairie B est-elle donc la plus vulnérable parce qu'elle met plus de temps à récupérer, même si l'impact est moins fort ? Cette question fait, entre autres, référence au problème de résilience.

[Füssel et Klein, 2006] ont identifié trois catégories de définitions pour évaluer la vulnérabilité :

- La définition des "risques-aléas" ou "*risk-hazard framework*". La vulnérabilité est alors définie comme la relation de réponse entre un stress exogène au système et ses effets négatifs [Downing et Patwardhan, 2003]. Cette définition correspond assez bien à la définition de la **sensibilité** dans [IPCC 2001]. Cette définition est très présente dans la littérature technique sur la gestion des risques et des catastrophes naturelles.
- Le modèle social constructiviste [Adger et Kelly, 1999]. Dans ce cas la vulnérabilité est vue comme l'ensemble des causes socio-économiques expliquant

des différences de sensibilité et d'exposition. Cette définition correspond à l'ensemble des facteurs non-climatiques intervenant sur la vulnérabilité qui seront définis par la suite. Cette définition est dominante dans le domaine de l'économie politique et dans la géographie humaine.

- La définition de l'IPCC (3ème rapport [IPCC 2001]) : *"Mesure dans laquelle un système est sensible – ou incapable de faire face – aux effets défavorables des changements climatiques, y compris la variabilité du climat et les phénomènes extrêmes. La vulnérabilité est fonction de la nature, de l'ampleur et du rythme de la variation du climat à laquelle le système considéré est exposé, de la sensibilité de ce système et de sa capacité d'adaptation"*. Cette définition est celle principalement utilisée dans le domaine du changement climatique. La définition de la vulnérabilité selon l'IPCC est centrée sur le changement climatique. Mais, elle peut facilement être étendue : "Degré de stress ou de perturbation qu'un système humain et environnemental est capable de supporter avant d'être endommagé" [Kasperson et al., 2005] et [Turner et al., 2003a].

Nous nous baserons sur cette troisième définition. [Füssel et Klein, 2006] ont retracé en partie l'histoire de la conceptualisation de la vulnérabilité. On retiendra que le concept de vulnérabilité s'est affiné jusqu'au modèle de la Figure 7.

Les émissions de gaz à effet de serre (GES) influent directement sur les concentrations en GES atmosphériques. Ces concentrations sont le principal facteur influençant le climat et donc le changement climatique. Les autres facteurs (par exemple, activité solaire et activité volcanique) ne sont pas considérés dans notre étude. Le climat est composé du climat global (cycle moyen annuel de température, de rayonnement...) et de son aspect variabilité climatique, qui intègre les événements extrêmes. Le changement climatique se traduit par une modification de ses deux aspects.

L'**exposition** est l'ensemble des chocs et des perturbations auxquels le système est soumis, ou plus précisément dans notre cas, le degré et la nature du changement environnemental auquel l'écosystème est soumis. L'exposition est en fait influencée à la fois par le changement global et la variabilité climatique, les concentrations de GES (par exemple, l'ozone) et les facteurs non-climatiques (ensemble de facteurs environnementaux (non climatiques), politiques, économiques, sociaux, démographiques et techniques). Les **facteurs non-climatiques** sont eux-mêmes définis par les **scénarii non-climatiques** (par exemple, scénario du cours du blé, ou scénario du cours du pétrole).

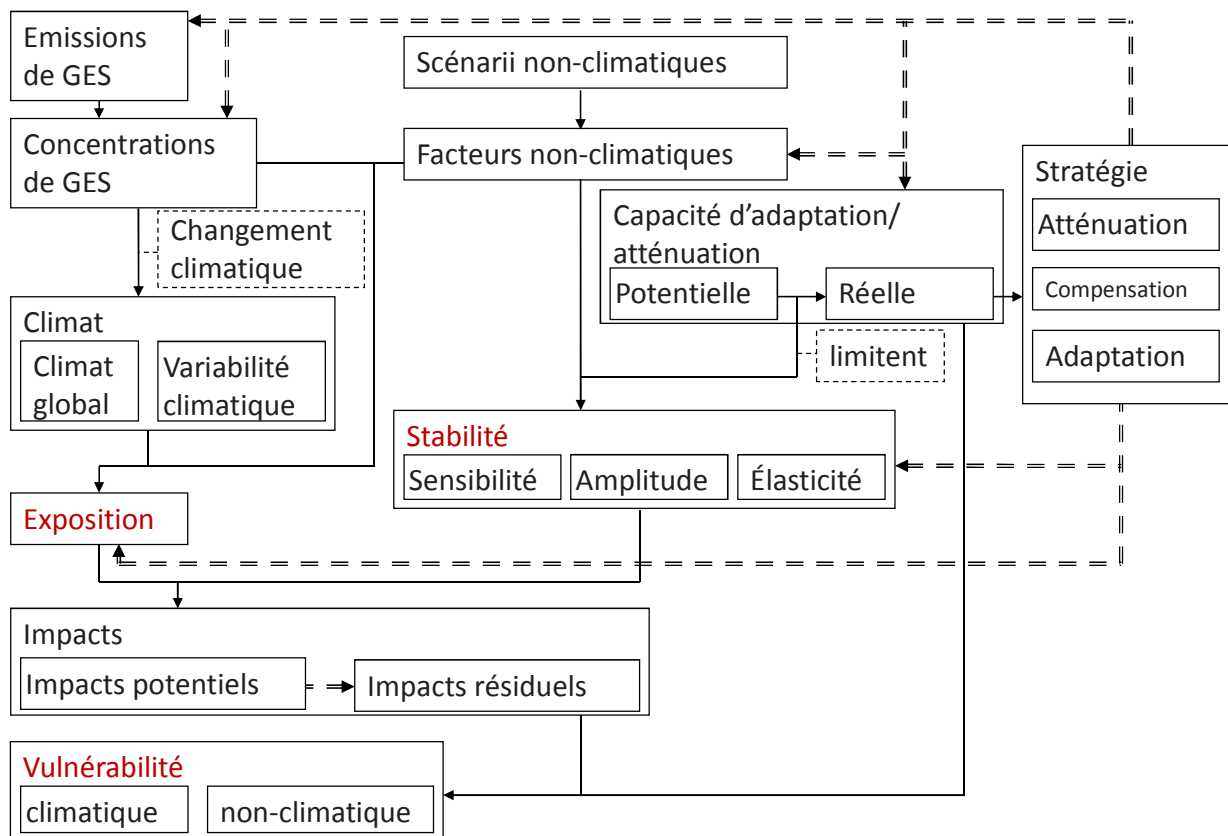


Figure 7 [Lardy et al., 2012b] Modèle des concepts liés à la vulnérabilité par rapport au changement climatique, basé sur [Füssel et Klein, 2006]. Les flèches à doubles traits représentent les rétroactions des stratégies d'atténuations et d'adaptation.

La **sensibilité** est la proportion dans laquelle un système est influencé, favorablement ou défavorablement, par des stimuli, en l'occurrence liés au climat. Elle est surtout significative pour de faibles variations, tandis que pour de forts stimuli (du type événements extrêmes), on s'intéressera plus aux propriétés de récupération du système, à savoir la résilience écologique ou l'élasticité. La résilience écologique (ou amplitude, c'est-à-dire le maximum de perturbation qu'un système peut supporter avant de changer d'état) n'apparaît pas explicitement dans le modèle conceptuel de [Füssel et Klein, 2006]. Certains auteurs définissent la capacité d'adaptation comme le potentiel de récupération d'un écosystème [De Lange et al., 2010], ce qui correspond, en fait, à la capacité d'adaptation interne par opposition à celle qui fait appel aux éléments externes au système. Cette capacité d'adaptation interne pourrait être définie comme la résilience écologique. Nous ne considérons dans la suite que la capacité d'adaptation externe. La vitesse de récupération face à des petites perturbations ("*engineering resilience*" ou vitesse de récupération, [Holling 1996]), définie comme la vitesse de retour à l'état (ou à la dynamique) de référence après une perturbation temporaire [Grimm et Wissel, 1997] est également appelé **élasticité**. Celle-ci est un assez bon indicateur de la résilience écologique [Van Nes et Scheffer, 2007]. La résilience écologique et l'élasticité sont deux aspects du

concept de stabilité. La résilience écologique ne peut malheureusement pas être mesurée en pratique. [Van Nes et Scheffer, 2007] ont montré que la vitesse de récupération tend vers zéro quand on est proche du seuil de changement de système, mais que cela peut également se produire dans d'autres situations. Dans la plupart des modèles, la vitesse de récupération change en même temps que la résilience écologique. La sensibilité, la résilience écologique et l'élasticité que l'on pourrait considérer comme représentatifs de la **stabilité** sont principalement influencées par les facteurs non climatiques [Füssel et Klein, 2006]. Une approche conceptuelle pour considérer les changements de système, dans la continuité de [Holling 1973] consiste à voir le système selon la métaphore d'une balle et d'une colline (Figure 8). Dans cette métaphore, on s'intéresse aux états d'équilibre stable vers lequel le système tend à revenir malgré les faibles perturbations continues. Deux équilibres différents se caractérisent généralement par des valeurs différentes des fonctions écologiques. Les changements progressifs de conditions peuvent altérer la résilience sans pour autant changer l'état d'équilibre [Scheffer et al., 2001]. [Scheffer et al., 2009] explique que, à nature du système constant, le fait, pour le celui-ci, d'approcher d'un point de bifurcation (sommet de la colline) pourrait être détecté à travers l'observation d'une baisse de récupération entre deux perturbations (*i.e.*, moins bon retour vers l'état de référence) et d'une augmentation de l'auto-corrélation (entre le système à l'instant t et l'instant $t+1$), le tout traduisant une perte d'élasticité.

L'exposition et les propriétés de sensibilité et de résilience sont les principaux paramètres dont dépendent les **impacts**. Au sein des impacts, on peut distinguer les **impacts potentiels** et les **impacts résiduels** qui sont tous les impacts pouvant survenir suite au changement climatique, respectivement avant ou après prise en compte d'adaptation planifiée. La vulnérabilité est parfois définie comme les impacts résiduels. C'est par exemple le cas dans le rapport de la FAO [FAO, 1996] où la vulnérabilité est vue comme le potentiel d'impacts négatifs non évitables par adaptation.

La notion d'impacts résiduels fait appel à la notion de **capacité d'adaptation**, à savoir la capacité que possède un système à se modifier de façon à se rendre moins vulnérable. Dans le cadre du changement climatique, on peut la définir comme la capacité d'un système à s'adapter aux changements climatiques (notamment à la variabilité du climat et aux phénomènes extrêmes), de façon à atténuer les dommages potentiels, à tirer parti des possibilités offertes et à faire face aux conséquences. La capacité d'adaptation est directement fonction des facteurs non-climatiques (voir des exemples dans la table 2). La **vulnérabilité** est donc une fonction des impacts et de la capacité d'adaptation. Au sein de la capacité d'adaptation (et d'atténuation) on distingue la **capacité d'adaptation potentielle** de la **capacité d'adaptation réelle** qui, elle, est limitée par des facteurs non-climatiques (cf. table 2).

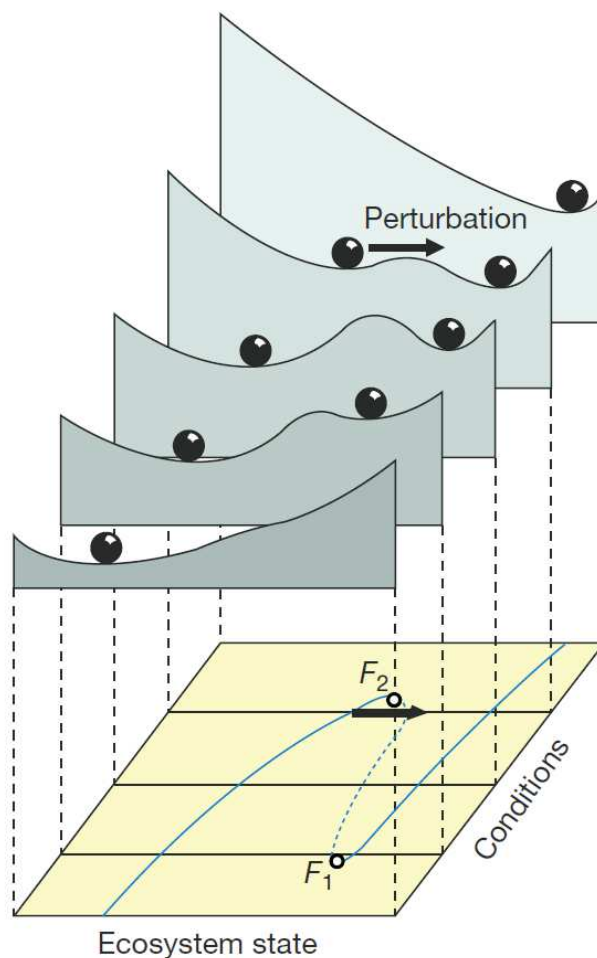


Figure 8 [Scheffer et al., 2001] Métaphore du potentiel physique (axe z) pour représenter les états d'équilibre par rapport aux conditions et à l'état du système. F1 et F2 représentent deux points de bifurcation pour lesquels un faible stress fait automatiquement changer le système d'état.

La politique d'**atténuation** ("*mitigation*") consiste à limiter le changement climatique global en réduisant les sources et en renforçant les puits de GES [Füssel et Klein, 2006]. Tandis que la politique d'**adaptation planifiée** consiste à réduire les effets négatifs et inévitables du changement climatique. Le pré-requis majeur à ces **stratégies** est la suffisance des ressources nécessaires pour mettre celles-ci en œuvre. Historiquement, la politique d'atténuation a retenu plus d'attention car, si d'un côté une politique d'atténuation permet de réduire les impacts sur tous les systèmes sensibles au changement climatique, d'un autre côté, le potentiel des politiques d'adaptation est très limité pour certains systèmes. Pour une comparaison plus précise entre adaptation et atténuation, le lecteur pourra se référer à [Füssel et Klein, 2006]. Ces deux stratégies à objectif premier différents ne sont pas contradictoires. Elles peuvent influencer un certain nombre de facteurs. L'adaptation cherche principalement à influencer les facteurs non-climatiques, la stabilité et l'exposition du système et donc par voie de conséquence les impacts du changement climatique sur celui-ci. Tandis que la politique d'atténuation aura

principalement des impacts sur les émissions et les concentrations de GES. Une troisième stratégie, celle de **compensation**, consiste à mettre en place des paiements et/ou autres assistances des pays qui contribuent de manière disproportionnelle aux émissions de GES vers ceux qui souffrent disproportionnellement des conséquences (par exemple [Paavola et Adger, 2002]). Cette troisième politique a été pour le moment nettement moins étudiée.

Les scénarios d'adaptations devront chercher non pas à être simplement les plus optimaux possibles mais surtout les plus robustes par rapport aux incertitudes [Kleijnen et al., 2005]. En effet, compte-tenu des incertitudes fortes de l'environnement futur, une politique de gestion optimale pourrait échouer complètement. Ces politiques de gestion robustes sont parfois appelées "*satisficing*" (combinaison de "*satisfy*" et "*suffice*") [Simon 1981]. Il est intéressant de noter que les politiques robustes devraient en fait être des politiques optimales par rapport à la vulnérabilité. Par exemple, une politique robuste pourrait allier un fort rendement tout en ayant une faible sensibilité (et donc une forte stabilité) face aux aléas environnementaux. Elle serait alors peu vulnérable. Un exemple de recherche de solution robuste peut être trouvé dans [Sanchez et al., 1996] ou [Kleijnen et Gaury, 2003]. La philosophie des conceptions robustes a été inspirée par [Taguchi 1987]. Dans le cas particulier des systèmes environnementaux, une méthode pour obtenir une adaptation robuste pourrait consister à se focaliser sur la résilience. En effet, comme le souligne [Scheffer et al., 2001], construire et maintenir une forte résilience dans un écosystème est certainement une manière efficace et pragmatique de le gérer face à une augmentation du changement climatique. Cette idée est confortée par le fait qu'un système à forte résilience devrait avoir une faible vulnérabilité. Les tables 1 et 2 récapitulent les définitions des concepts-clés de la vulnérabilité.

Table 1 Concepts liés à la vulnérabilité

Terme	Définition	Références
Adaptation planifiée	Résultat d'une politique de gestion délibérée, basée sur la prise de conscience que les conditions sont sur le point de changer, ou ont déjà changé, dans le but de retourner à, maintenir ou atteindre un certain état.	[IPCC 2001] [Metzger et al., 2006]
Atténuation	Politique consistant à limiter le changement climatique global en réduisant les sources et en renforçant les puits de GES.	[Füssel et Klein, 2006]
Capacité d'adaptation	Capacité que possède un système à se modifier de façon à se rendre moins vulnérable. Capacité d'un système à s'adapter aux changements climatiques (notamment à la variabilité du climat et aux phénomènes extrêmes), de façon à atténuer les dommages potentiels, à tirer parti des possibilités offertes et à faire face aux conséquences.	[Luers et al., 2003] [IPCC 2001]
Capacité d'adaptation potentielle	Capacité d'adaptation sans prise en compte de la limitation par les facteurs non-climatiques.	
Capacité d'adaptation réelle	Capacité d'adaptation avec prise en compte de la limitation par les facteurs non-climatiques.	
Elasticité ou "Engineering resilience"	Vitesse de récupération après un stress. Vitesse de retour à l'état (ou à la dynamique) de référence après une perturbation temporaire.	[Grimm et Wissel, 1997] [Holling 1996]
Exposition	Chocs et perturbations auxquels le système est soumis. Degré et nature du changement environnemental auquel l'écosystème est soumis.	[Luers 2005] [IPCC 2001]
Impacts potentiels	Tous les impacts pouvant survenir suite au changement climatique, sans prise en compte d'adaptation planifiée.	[IPCC 2001] [Metzger et al., 2006]
Impacts résiduels	Tous les impacts pouvant survenir suite au changement climatique, après prise en compte d'adaptation planifiée.	[IPCC 2001] [Metzger et al., 2006]
Résilience (écologique) ou Amplitude	Capacité d'un groupement de structure et de processus à se renforcer mutuellement afin de résister à des perturbations et des stress. Maximum de perturbation qu'un système peut supporter avant de changer d'état.	[Holling 1973] [Brand et Jax, 2007]
Sensibilité	Proportion dans laquelle un système est influencé, favorablement ou défavorablement, par des stimuli (en l'occurrence liés au climat).	[IPCC 2001]
Stabilité	Concept à multiple facettes incluant notamment l'élasticité, la résilience écologique et la sensibilité.	[Grimm et Wissel, 1997]
Vulnérabilité	Degré de stress ou de perturbation qu'un système humain et environnemental est capable de supporter avant d'être endommagé. Mesure dans laquelle un système est sensible – ou incapable de faire face – aux effets défavorables des changements climatiques, y compris la variabilité du climat et les phénomènes extrêmes.	[Kasperson et al., 2005] [Turner et al., 2003a] [IPCC 2001]

L'approche par indicateur peut conduire à un manque de correspondance entre la définition conceptuelle de la vulnérabilité et sa mesure. En effet, en se concentrant sur des

mesures qui représentent la vulnérabilité d'un lieu (ou d'un système), l'approche par indicateur (qui utilise tout un ensemble de variables) se verra difficilement interprétable dans d'autres configurations. Par exemple un indicateur du volume de pluie annuel aura un sens dans une région mais pourrait ne pas être représentatif de vulnérabilité dans une autre (où la pluie serait un facteur non limitant). C'est-à-dire qu'un indicateur construit pour une zone n'est pas forcément utilisable dans une autre zone, même après calibration. Il peut manquer une composante ou, à l'inverse, être basé sur une notion qui n'a pas de sens à ce nouvel endroit. Cependant, comme le souligne [Adger 2006], les méthodes proposées par [Luers et al., 2003] et [Luers 2005] ont l'avantage d'être des méthodes générales de mesure de la vulnérabilité, qui prennent en compte l'intensité de la vulnérabilité et ses dynamiques (ce qui est vulnérable à une période ne l'est pas forcément à la période suivante). Il souligne également le problème récurrent du choix du seuil de vulnérabilité (ou de l'état de référence).

Table 2 Exemples associés aux concepts liés à la vulnérabilité

Terme	Exemple
Adaptation planifiée	Mise en place d'irrigation
Atténuation	Extensification ou arrêt de fertilisation d'une prairie permanente pâturée
Capacité d'adaptation	Capacité de mise en place de l'irrigation
Capacité d'adaptation potentielle	Capacité de mise en place de l'irrigation
Capacité d'adaptation réelle	Capacité de mise en place de l'irrigation seulement sur une partie de l'année
Elasticité ou "Engineering resilience"	Vitesse de récupération d'une prairie après une sécheresse
Exposition	Augmentation de la fréquence et de l'intensité des sécheresses
Impacts potentiels	Baisse de la production agricole
Impacts résiduels	Légère baisse de production agricole, même après la mise en place d'irrigation
Résilience (écologique) ou Amplitude	Récupération de la prairie après une sécheresse
Sensibilité	Ecosystème sensible à une augmentation de la sécheresse
Stabilité	Un système de production à rendement constant
Vulnérabilité	Augmentation de la probabilité de pertes liées aux sécheresses

L'un de nos objectifs est d'utiliser plusieurs méthodes pour quantifier la vulnérabilité des écosystèmes. Pour cela nous avons d'abord étudié la méthode générique de mesure de la vulnérabilité de Luers (2.2.2). Puis, nous nous sommes intéressés à d'autres indices possibles de vulnérabilité (2.2.3) et enfin nous nous sommes penchés sur la question de l'analyse de vulnérabilité régionalisée (2.2.4).

2.2.2. Méthode de Luers pour le calcul de la vulnérabilité

Dans la "méthode de Luers" (décrite dans [Luers et al., 2003] [Luers 2005]), l'évaluation de la vulnérabilité se fait en utilisant quatre concepts :

- l'état du système relativement à un seuil de dégâts ;
- la sensibilité ;
- l'exposition ;
- la capacité d'adaptation.

Les sous-parties suivantes décrirons le mode de calcul de la vulnérabilité sans prise en compte de l'exposition, puis avec. Ensuite, nous nous intéresserons à la capacité d'adaptation et enfin à la notion de surface de vulnérabilité. Afin de faciliter la compréhension, nous accompagnerons ces descriptions d'un exemple simple.

Calcul de la vulnérabilité

Les pré-requis à la définition de la vulnérabilité d'un système sont la sensibilité de celui-ci aux différents stress et l'identification du seuil au-delà duquel le système est endommagé. [Luers et al., 2003] ont défini la vulnérabilité (V) à de faibles stress de la manière suivante :

$$V = f(\text{Sensibilité/valeur relative de la variable d'état par rapport à la valeur-seuil})$$

$$= f\left(\frac{|\partial W / \partial X|}{W/W_0}\right), \text{ où } W^{11} \text{ est la variable d'état, } W_0 \text{ est la valeur-seuil, } X \text{ est le stress.}$$

Un faible stress est un stress qui n'entraîne pas de changement du système. D'autres calculs pour la sensibilité peuvent être utilisés, comme le coefficient de variation. Il est intéressant de noter que l'unité de la vulnérabilité est exprimée en unité de la variable d'état par unité de stress. Cependant, en normalisant par la valeur à un état de référence on obtient une vulnérabilité adimensionnelle et indépendante de la valeur-seuil (mais dépendante du choix de l'état de référence). Comme le souligne [Luers 2005] ce seuil peut évoluer au cours du temps, à la fois parce que la capacité de résilience a changé (par exemple, un éleveur ayant subi une sécheresse l'année N-1, ne pourra pas supporter une production minimale identique pour l'année N), mais également car le seuil peut « socialement » ou « économiquement » varier au cours du temps (par exemple, le rendement minimum nécessaire pour la période passée est probablement inférieur à celui nécessaire pour la période future).

¹¹ Noter que, comme W est une variable représentative du bien-être du système, il ne faut pas qu'elle oscille autour de 0. Par exemple, dans le cas de la NBP (« Net Biome Productivity »), où 0 n'est pas l'optimum, il faut appliquer une transformation à la variable (un « offset » devrait suffire).

Prenons un exemple « simple » : considérons le rendement annuel une prairie permanente clermontoise gérée de manière extensive sur sol peu profond. On simule un scénario climatique futur (scénario A2 de l'IPCC SRES [Nakićenovič 2000]), avec le modèle climatique ARPEGE régionalisé par la méthode dynamique quantile-quantile [Déqué et al., 2007], sur lequel on force la concentration de CO₂ à rester constante. Un indicateur du stress climatique (exposition) est l'indice d'aridité de De Martonne-Gottmann [De Martonne 1942]. On considère les résultats de la simulation avec PaSim pour le passé proche (1970-1999), le futur proche (2020-2049) et le futur lointain (2070-2099). On fixe (à titre d'exemple) le seuil critique de rendement à 5 tMS ha⁻¹. Le calcul de la vulnérabilité selon la première méthode (V) est fait en utilisant le coefficient de variation.

Table 3 Calcul de la vulnérabilité dans le cas d'une prairie permanente extensive et sous réchauffement climatique.

	W Rendement moyen (tMS ha ⁻¹)	Écart-type du rendement (tMS ha ⁻¹)	Vulnérabilité V	Indice moyen d'aridité
Période 1970-1999	4,25	0,57	0,157	29,1
Période 2020-2049	4,62	0,75	0,174	27,5
Période 2070-2099	4,45	0,81	0,204	16,5

Cet exemple (Table 3) est intéressant car il est basé sur un gradient progressif d'augmentation de la sécheresse (diminution de l'indice d'aridité). Cependant la production moyenne augmente, ce qui devrait réduire la vulnérabilité. Mais dans le même temps l'écart-type et donc le coefficient de variation augmente significativement, ce qui induit l'effet inverse.

Vulnérabilité en tenant compte de l'exposition

Les différents écosystèmes sont exposés à des perturbations d'amplitude et de fréquences différentes ([IPCC 2001], [Turner et al., 2003a, b]). Cette information (l'exposition) est capturée et incluse dans la vulnérabilité en introduisant l'information de distribution du stress [Luers et al., 2003] :

$$V' = f(\text{Sensibilité/valeur relative de la variable d'état par rapport à la valeur-seuil})$$

$$= \int \left(\frac{|\partial W / \partial X|}{W/W_0} \right) P_X dX, \text{ où } P_X \text{ est la densité de probabilité du stress } X.$$

Reprenons l'exemple précédent (rendement annuel sur une période climatique de 30 ans) et calculons ce nouvel indice de variabilité. Comme variable X, nous allons utiliser l'indice d'aridité. Comme estimateur de $|\partial W / \partial X|$ nous utiliserons un estimateur simple de la dérivée :

$$(|\partial W / \partial X|)_i = \frac{1}{2} \left(\frac{W_{i+1} - W_i}{X_{i+1} - X_i} + \frac{W_i - W_{i-1}}{X_i - X_{i-1}} \right), \text{ où } i \text{ est l'indice de l'année}$$

climatique (ces dernières ayant été classées par X croissant). Derrière le classement de ces années, nous faisons l'hypothèse que les années sont indépendantes, ou que l'autocorrélation temporelle est négligeable. De plus, le fait d'indicer W suivant les années classées par X croissant permet bien sûr d'exprimer W comme une fonction de X, mais surtout, de prendre en compte la distribution de X dans la somme (qu'elle soit intégrale ou discrétisée). Dans notre cas, nous considérons chaque année comme équiprobable, ainsi la densité de probabilité de chacune est de 1/30. Dans notre cas, la formule de V' est donc discrétisée :

$$V' = \frac{1}{30} \sum_{i=1}^{30} \left(\frac{(|\partial W / \partial X|)_i}{W_i / W_0} \right)$$

Si on l'applique à notre exemple, on retrouve (Table 4) les mêmes tendances d'augmentation de vulnérabilité. Cependant, si l'on regarde la vulnérabilité relativement à la période passée, l'augmentation de vulnérabilité est beaucoup plus sensible avec l'indice V' (tenant compte de l'exposition) qu'avec l'indice de base :

- entre le passé proche et le futur proche : +11% avec V contre +63% avec V' ;
- entre le passé proche et le futur lointain : +30% avec V contre +121% avec V'.

La prise en compte de l'exposition semble améliorer la qualité des résultats. Cette mesure de la vulnérabilité permet notamment de distinguer des systèmes avec des états moyens et des sensibilités identiques mais des expositions (*i.e.* des distributions du stress) différentes.

Table 4 Comparaison de la vulnérabilité selon deux modes de calcul, sans prise en compte de l'exposition (V) ou avec (V').

	W Rendement moyen (tMS ha ⁻¹)	Écart-type du rendement (tMS ha ⁻¹)	Indice moyen d'aridité	Vulnérabilité V	Vulnérabilité V'
Période 1970-1999	4,25	0,57	29,1	0,157	0,223
Période 2020-2049	4,62	0,75	27,5	0,174	0,363
Période 2070-2099	4,45	0,81	16,5	0,204	0,493

Capacité d'adaptation

La capacité d'adaptation est définie par [Luers et al., 2003] comme la capacité que possède un système à se modifier de façon à se rendre moins vulnérable. Un exemple classique est celui de l'éleveur qui repense son système d'exploitation de façon à mieux

supporter les sécheresses. Ainsi définie, la capacité d'adaptation (A) peut se calculer de la manière suivante :

$$A = V'(\text{système actuel}) - V'(\text{système modifié}).$$

$V'(\text{système modifié})$ est également appelé le potentiel minimum de vulnérabilité. On peut réduire la vulnérabilité d'un système, à travers des mesures d'adaptation, en jouant sur trois axes :

- en modifiant sa fonction d'état (*i.e.*, fonction caractéristique du bien-être du système) de façon à la rendre moins sensible au stress critique ;
- en changeant la valeur-seuil (en-deçà de laquelle le système est considéré comme vulnérable);
- en modifiant l'exposition du système aux stress concernés.

De nombreux facteurs peuvent influencer la capacité d'un système à s'adapter : la vitesse de changement des stress, le contexte social et les capitaux du système (par exemple les réserves financières, les bâtiments d'élevage, les machines et leurs capacités nécessaires pour certains travaux) ...

Pour poursuivre l'exemple, on pourrait considérer que l'éleveur décide d'optimiser la gestion de sa prairie en fertilisant autant que nécessaire et en fauchant autant que possible [Graux 2011, chapitre 6]. De façon à pouvoir comparer ce nouveau système avec le précédent, nous considérons que l'éleveur nourrit ses animaux à l'aide des fourrages obtenus pendant l'équivalent de la durée de la période de pâturage du système non optimisé. On retirera donc ce besoin (0.36 tMS ha^{-1}) du rendement optimal. Cet exemple est biaisé car une telle gestion serait difficile à faire et ne tient compte ni de l'impact économique (coût des fertilisants, coût d'entretien des bâtiments, ...) ni de l'aspect environnemental. Cependant les résultats sont intéressants (Table 5, Table 6) car ils permettent de mettre en évidence une différence significative entre les deux modes de calculs de la vulnérabilité. En effet si dans le cas de la vulnérabilité sans prise en compte de l'exposition (V), la gestion optimisée représente une bonne adaptation (A) ; dans le cas de la prise en compte de l'exposition la prairie optimisée n'est pas toujours moins vulnérable (période passée proche) et entraîne une valeur d'adaptation (A') négative. La première conclusion est que l'optimisation ne permet pas forcément de rendre le système plus stable et moins vulnérable. Dans le cadre de PaSim, il faudra réfléchir à une manière différente de celle utilisée dans cet exemple pour obtenir le potentiel minimum de vulnérabilité (parcours d'une gamme de gestions prédéterminées et fixes dans le temps ou utilisation de scénarios adaptatifs). Compte-tenu du nombre de combinaisons possibles, il faudra choisir un plan d'expérience avec soin. La seconde conclusion est qu'il paraît donc important de tenir compte de l'exposition et donc le calcul V' devrait être préféré à V.

Table 5 Comparaison de la vulnérabilité dans le système d'origine et dans le système optimisé.

	W Rendement moyen (tMS ha ⁻¹)	Écart-type du rendement (tMS ha ⁻¹)	Indice moyen d'aridité	Vulnérabilité V	Vulnérabilité V'
Gestion normale					
Période 1970-1999	4,25	0,57	29,1	0,157	0,223
Période 2020-2049	4,62	0,75	27,5	0,174	0,363
Période 2070-2099	4,45	0,81	16,5	0,204	0,493
Gestion optimisée					
Période 1970-1999	14,88	2,70	29,1	0,0061	0,388
Période 2020-2049	13,97	1,55	27,5	0,0040	0,139
Période 2070-2099	10,16	2,76	16,5	0,0134	0,403

Table 6 Potentiel d'adaptation basé sur une gestion optimale et vulnérabilité calculée en tenant compte de l'exposition (A') ou non (A).

	Adaptation A	Adaptation A'
Période 1970-1999	0,151	-0,165
Période 2020-2049	0,170	0,224
Période 2070-2099	0,191	0,090

Il existe un biais récurrent à ce genre d'analyse, c'est la corrélation interannuelle entre les valeurs de la variable X , ici l'indice d'aridité. En effet, l'effet d'une année ne sera probablement pas le même sur une prairie selon que l'année précédente ait été une année pluvieuse ou déjà une année de sécheresse. Pour des raisons de simplicité, les exemples ont été présentés pour un cas unidimensionnel, mais les principes restent valables pour plusieurs dimensions de stress. À noter également que certains auteurs restreignent la capacité d'adaptation au potentiel de mise en place de mesures planifiées d'adaptation et non comme l'effet de ces mesures sur la vulnérabilité [Metzger et Schröter, 2006].

La méthode de Luers est appropriée pour quantifier l'impact d'une mesure d'adaptation. Il reste cependant le problème de définir ces systèmes adaptés. Une autre approche est celle de [Metzger et al., 2006] qui définit la capacité d'adaptation comme une fonction directe de l'environnement socio-économique. On pourrait alors utiliser cette information comme contrainte dans la minimisation de la vulnérabilité. Par exemple, la contrainte de capacité d'adaptation peut alors être l'accès au prêt pour la mise en place d'une irrigation. Les différents niveaux de contrainte de capacité d'adaptation entraînent alors une minimisation plus ou moins forte de la vulnérabilité.

Surface de vulnérabilité

La définition « mathématique » de la vulnérabilité peut être étendue en considérant la vulnérabilité comme une fonction à la fois de la sensibilité (1^{ère} dimension) et de l'exposition (2^{nde} dimension). Ainsi définie, il s'agit d'une surface de réponse [Luers 2005]:

Surface de vulnérabilité = $f\left[\frac{(Sensibilité, Exposition)}{W/W_0}\right]$, où W est la variable d'état, W₀ est la valeur-seuil.

Dans ce cas-là, la vulnérabilité est calculée selon l'approche la plus simple (mode de calcul V, voir ci-dessus, cependant un autre mode de calcul aurait pu être utilisé). Cette surface peut être un outil pour trier les systèmes. Reprenons notre exemple de réchauffement climatique (toujours à concentration en CO₂ atmosphérique constante) et considérons six systèmes prairiaux différents [Graux et al., 2012a] :

- prairie temporaire de ray-grass anglais irriguée (Ray-grass GI) ;
- prairie temporaire de fétuque irriguée (Fétuque GI) ;
- prairie temporaire de ray-grass anglais non irriguée (Ray-grass G) ;
- prairie temporaire de fétuque non irriguée (Fétuque G) ;
- prairie permanente pâturée et fauchée, gérée de manière intensive (PP5) ;
- prairie permanente pâturée et fauchée, gérée de manière extensive (PP20).

Si l'on trace le rendement par rapport à sa sensibilité (coefficient de variation) pour tous les systèmes sur les trois périodes (Figure 9), on peut alors observer des classes de systèmes en terme de vulnérabilité : les systèmes irrigués sont les moins vulnérables (on n'a pas tenu compte d'une probable diminution d'eau disponible pour le futur) mais possèdent peu de marge d'adaptation. À l'inverse les systèmes de prairie temporaire non irriguée sont plus vulnérables, mais possèdent une grande marge de capacité d'adaptation. Dans le cadre d'un scénario climatique réel (avec augmentation du CO₂), il pourrait être intéressant de quantifier le déplacement des systèmes par rapport aux lignes d'iso-vulnérabilité. On peut trouver des exemples d'application des idées de Luers dans [Lexer et Seidl, 2009] [Seidl et al., 2010], où différents indicateurs sont agrégés de façon à former une seule valeur de sensibilité et de variable d'état. Comme le souligne [Lexer et Seidl, 2009] et [Drechsler 2004], l'agrégation de paramètre peut parfois induire une faible qualité (compte-tenu de la compensation d'un facteur par les autres). Certains des facteurs utilisés par [Seidl et al., 2010] sont en fait des seuils de changement de rendement moyen sur la période future par rapport à la période de référence : par exemple ±10% (seuil de détection) ; ±15% (seuil de tolérance)). Ils sont ensuite rendu unitaires {-1,...,+1}, par exemple, une augmentation de rendement de 15% vaudra 1.

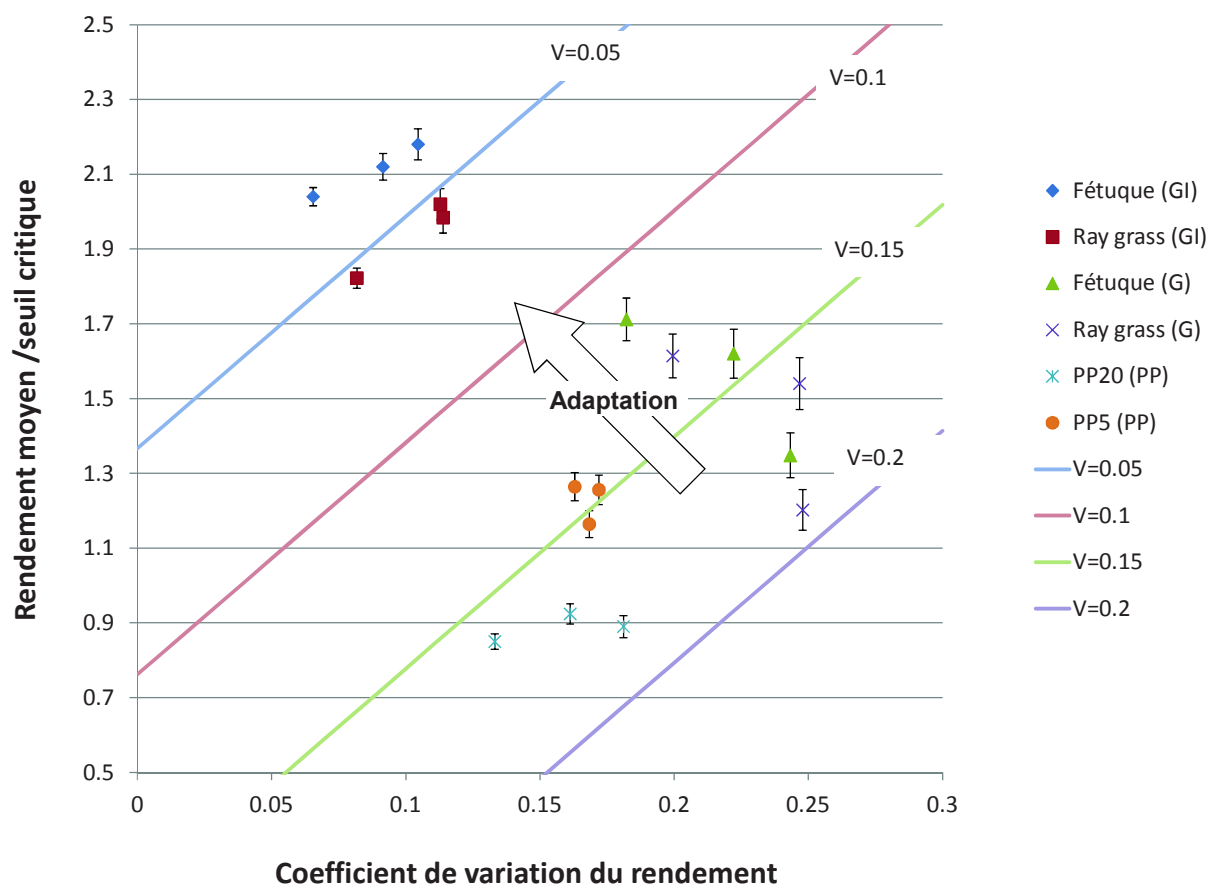


Figure 9 Exemple de surface de vulnérabilité des prairies pour divers systèmes.

Les termes de "stress" ou "perturbations" font généralement référence à des forces exogènes ayant le potentiel de créer des impacts négatifs sur le système [Turner et al., 2003a] [Kelly et Adger, 2000]. Dans notre cas, il s'agira principalement du changement climatique, mais il pourrait être étendu aux variations du marché. Cependant les forces endogènes ne sont pas à négliger, mais comme [Luers 2005] le précise, les stress endogènes sont considérés via la sensibilité.

Les écologues ont mis en évidence, à travers diverses études (par exemple [Folke et al., 2002]) que des changements progressifs dans les variables clés pour la sensibilité et la résilience peuvent entraîner une augmentation de la vulnérabilité aux événements extrêmes (comme la sécheresse de 2003). Cela implique le fait qu'il faut tenir compte de la vulnérabilité aux événements extrêmes et aux stress de plus petite amplitude. Ceci étant vrai à la fois pour l'évaluation de la vulnérabilité, comme pour les politiques décisionnelles qui sont généralement focalisées sur les politiques à court terme. Cependant, les méthodes de Luers n'ont pas été conçues pour évaluer la vulnérabilité aux forts stress (comme les extrêmes climatiques). De plus, [Schneider et Lane, 2006] insistent sur le fait qu'une seule mesure peut ne pas être suffisante et qu'un ensemble d'indices est nécessaire pour capturer toute l'importance du changement climatique. Il conviendra donc d'utiliser d'autres indicateurs en parallèle pour pouvoir tenir compte des faibles et forts stress.

2.2.3. Autres indices

La vulnérabilité peut également être calculée comme la somme pondérée d'un certain nombre d'indices, comme dans [Penghua et al., 2007] :

$V = \sum_i a_i I_i$, où a_i est un coefficient de pondération et I_i un indice utilisé pour le calcul de la vulnérabilité.

Comme [Adger 2006] l'a expliqué, il est possible d'utiliser la mesure généralisée de la pauvreté [Foster et al., 1984] pour calculer la vulnérabilité :

$$V_\alpha = \frac{1}{n} \left[\sum_{i=1}^q \left(\frac{W_0 - W_i}{W_0} \right)^\alpha \right]$$
, où n est le nombre d'individus, W_i la valeur d'état pour l'individu i , W_0 la valeur seuil, q le nombre d'individus au dessous du seuil et enfin α est l'indice de sensibilité.

Ainsi pour $\alpha = 0$, cet indice est appelé vulnérabilité proportionnelle : $V_0 = q/n$ et correspond simplement à la proportion du nombre d'individus vulnérables. Cet indice ne tient pas compte du degré de vulnérabilité des individus.

Pour $\alpha = 1$, $V_1 = \frac{1}{n} \left[\sum_{i=1}^q \left(\frac{W_0 - W_i}{W_0} \right) \right]$ est appelé "*vulnerability gap*".

Cet indice intègre en partie le degré de vulnérabilité des individus. Améliorer cet indice revient à réduire le nombre d'individus vulnérables (sous le seuil) ou améliorer globalement leur état (réduire l'écart moyen par rapport à la valeur seuil). Noter que l'on suppose W positif.

Pour $\alpha = 2$, $V_2 = \frac{1}{n} \left[\sum_{i=1}^q \left(\frac{W_0 - W_i}{W_0} \right)^2 \right]$ est appelé l'indice de "*Vulnerability severity*".

Cet indice intègre une prise en compte quadratique des distances par rapport au seuil, des individus vulnérables. Plus la variance est forte, plus grand est l'indice de sévérité. C'est en améliorant l'état des plus vulnérables que l'on améliore le plus fortement cet indice.

On peut également définir V'_α par :

$$V'_\alpha = \frac{1}{n W_0} \left[\sum_{i=1}^q (W_0 - W_i)^\alpha \right]^{\frac{1}{\alpha}}$$

On pourrait également utiliser l'indice :

$$V'_{\infty} = \frac{W_0 - \min_i(W_i)}{nW_0}$$

Cet indice informe sur la vulnérabilité de l'individu le plus vulnérable. Réduire cet indice revient à réduire la vulnérabilité des individus les plus vulnérables. Cependant afin d'avoir un indice compris entre 0 et 1, on peut le multiplier par n , comme cela est fait dans [Lardy et al., 2012b].

Table 7 Calcul de la vulnérabilité dans le cas d'une prairie permanente extensive et sous réchauffement climatique.

	W Rendement moyen (tMS ha ⁻¹)	Écart-type du rendement (tMS ha ⁻¹)	Vulnérabilité proportionnelle V ₀	Vulnerability gap V ₁	Vulnerability severity V ₂
Période 1970-1999	4,25	0,57	0,866	0,157	0,0344
Période 2020-2049	4,62	0,75	0,733	0,101	0,0236
Période 2070-2099	4,45	0,81	0,733	0,130	0,0362

Si l'on applique ces trois indices sur les mêmes exemples que précédemment (Table 7), on trouve des résultats différents. Pour la vulnérabilité proportionnelle, on trouve que la période passée est la plus vulnérable car dans le futur, la probabilité de descendre en dessous du seuil est plus faible. L'indice V₁ (Vulnerability gap) montre une plus grande vulnérabilité pour le passé que pour les périodes suivantes car la moyenne des écarts à la valeur de référence, pour les cas vulnérables, est plus forte dans la période passée que dans le futur lointain. À l'inverse, quand la distribution en-deçà du seuil de référence devient étalée, cette information est perçue par l'indice V₂ ; ce qui se traduit par une plus grande vulnérabilité dans le futur lointain que dans le passé. L'influence du seuil de référence semble beaucoup plus importante dans cet indice que dans ceux de [Luers et al., 2003] [Luers 2005] où la sensibilité est transférée sur le choix, non plus de la valeur-seuil, mais sur celui de la situation de référence.

A noter qu'il serait possible d'estimer la vulnérabilité par la logique floue en utilisant les indicateurs (capacité d'adaptation, impacts potentiels...) précédemment calculés [Cheng et Tao, 2010]. On peut également l'utiliser pour assembler des indicateurs de vulnérabilité calculés selon un seul aspect (par exemple, indicateur de vulnérabilité du rendement ou de stockage de carbone).

[Stephen et Downing, 2001] testèrent différentes approches pour estimer la vulnérabilité à la famine. Ils utilisèrent notamment les réseaux de neurones pour classer et donner une valeur numérique de vulnérabilité. Cet outil peut permettre de passer d'un indicateur de vulnérabilité purement qualitatif à une valeur numérique, mais également, sans évaluation

préalable de la vulnérabilité, de classer les éléments d'entrée (zones géographiques, individus,...).

Dans [Skondras et al., 2011], la vulnérabilité est évaluée via un indice agrégé d'indicateurs préalablement redimensionnés dans une échelle d'entiers de 1 à 7. Des classes de vulnérabilité peuvent ainsi être définies. Ce même principe peut être réutilisé pour agréger la vulnérabilité sur différents aspects (par exemple celle liée au stock de C et celles des émissions de N₂O).

2.2.4. Calcul régionalisé de vulnérabilité

Calcul de l'indice stratifié d'Impact potentiel

Pour comparer la vulnérabilité sur différentes zones géographiques, il est nécessaire de replacer les valeurs quantitatives des services rendus par l'écosystème dans leur contexte régional [Metzger et Schröter, 2006]. Une approche peut être de créer des strates d'écosystèmes, notamment grâce à l'outil Environmental Stratification [Metzger et al., 2005] [Jongman et al., 2006]. Pour découper un territoire en strates, différentes approches ont déjà été utilisées, basées sur l'expertise [Noirfalize 1987] [Bohn et al., 2000] [Olson et al., 2001] [Roekaerts 2002] ou sur les statistiques [Metzger et al., 2005] [Jongman et al., 2006] [Mücher et al., 2009]. La stratification calculée dans [Metzger et al., 2005] s'est faite en quatre étapes :

1. sélection des variables environnementales significatives ;
2. extraction des principaux gradients environnementaux, par analyse en composante principales (ACP) et classification statistique de données ("*statistical clustering*") ;
3. "post-processing" pour minimiser l'existence de groupes de cellules isolées ;
4. validation de la stratification obtenue par comparaison, et évaluation de corrélation avec d'autres jeux de données.

Au sein de ces strates, la meilleure valeur de service rendu par l'écosystème est considérée comme valeur de référence (*i.e.* valeur optimale). On peut alors définir l'indice adimensionnel suivant : (ESstr : "*stratified ecosystem service provision*", Efficience de l'écosystème par rapport à sa strate)

$ESstr = \text{Valeur de service environnemental} / \text{Valeur de service environnemental de référence.}$

Noter que la strate et la valeur de référence sont à redéfinir à chaque horizon temporel. L'impact du changement climatique sur l'écosystème peut se mesurer via l'indice suivant :

Indice stratifié d'impact potentiel ("*Stratified Potential Impact Index*"), $PIstr = \Delta ESstr$,

ΔES_{str} étant la différence d'efficience du service entre la période considérée et la période de référence. Une valeur positive traduit un gain de service.

Estimation de la capacité d'adaptation

La capacité d'adaptation (vue comme potentiel de mise en place de mesures planifiées d'adaptation) est, dans [Metzger et al., 2006], estimée par un modèle d'agrégation d'indicateurs via la logique floue. Dans ce cadre des indicateurs socio-économiques sont choisis (nombre de docteurs, taux d'emploi des femmes...) puis agrégés via la logique floue jusqu'à un indicateur (compris entre 0 et 1) représentant la capacité d'adaptation (Figure 10). Ce calcul est bien évidemment réalisé de manière spatialisée, généralement à l'aide des statistiques régionales et nationales. Ce modèle de la capacité d'adaptation est ensuite validé par une analyse de sensibilité et d'incertitude. Les valeurs d'indicateurs pour les périodes futures sont obtenues par extrapolations linéaires. Pour notre cas, compte-tenu des incertitudes sur le choix et les valeurs des indicateurs, cette méthode pourrait être difficile à mettre en œuvre.

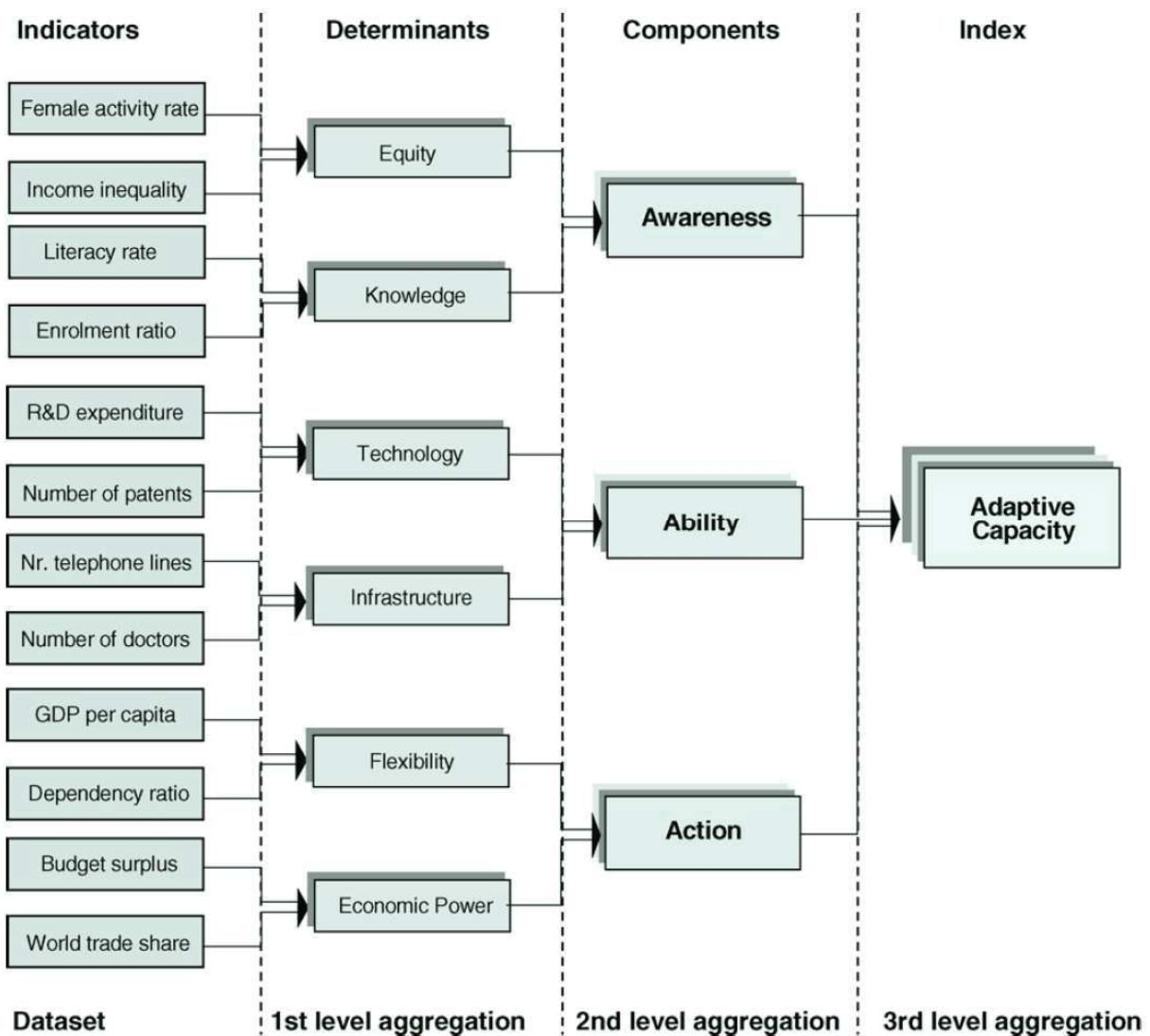


Figure 10 Ensemble d'indicateurs utilisés pour développer le modèle de capacité d'adaptation proposé par [Metzger et al., 2006].

Une autre manière d'estimer la capacité d'adaptation [Nelson et al., 2005, 2007, 2010] est d'utiliser une mesure pondérée des capitaux sociaux, humains, naturels, physiques et financiers. Cette pondération pouvant être uniforme ou basée sur une analyse en composante principale.

Création de cartes de vulnérabilité

L'un des objectifs d'une analyse de vulnérabilité régionalisée est de réaliser des cartes de vulnérabilité, afin de pouvoir en analyser les patrons régionaux. Partant du principe que la vulnérabilité est une fonction de la capacité d'adaptation et des impacts potentiels, [Metzger et Schröter, 2006] créèrent des cartes de vulnérabilité dont l'échelle de couleur illustre les impacts, tandis que le niveau de transparence de l'image représente la capacité d'adaptation. Cependant compte-tenu de la difficulté d'interprétation de ces cartes à

échelle bidimensionnelle, une approche d'analyse séparée de la capacité d'adaptation et des impacts potentiels semble préférable dans un premier temps [Metzger et al., 2006]. Dans un deuxième temps, la combinaison de ces résultats permet de déterminer qualitativement la vulnérabilité. Une autre possibilité est celle utilisée par [Nelson et al., 2010], qui consiste à définir trois classes pour la capacité d'adaptation et pour l'exposition. Le croisement de ces classes permettant de générer un indice de vulnérabilité (Figure 11). On peut également utiliser le croisement entre l'information de la capacité d'adaptation et l'intensité des impacts [Downing et Patwardhan, 2005] (Figure 12).

		Exposition		
		Forte	Moyenne	Faible
Capacité d'adaptation	Faible	Forte	Forte	Moyenne
	Moyenne	Forte	Moyenne	Faible
	Forte	Moyenne	Faible	Faible

Figure 11 Création d'un indice de vulnérabilité par le croisement de l'exposition et de la capacité d'adaptation [d'après Nelson et al., 2010].

		Capacité d'adaptation	
		Faible	Forte
Intensité des impacts	Forte	Communautés vulnérables	Opportunités de développement
	Faible	Risques résiduels	Communautés viables

Figure 12 Répartition des communautés en quatre classes de vulnérabilité selon l'intensité des impacts et la capacité d'adaptation [Downing et Patwardhan, 2005]

Pour comparer les indices de vulnérabilité à l'échelle spatiale (carte de vulnérabilité), on peut utiliser une analyse statistique [Marín et al., 2012] (analyse de sensibilité, coefficient de variation...), mais également un paramètre global de vulnérabilité permettant d'avoir une vue globale de la vulnérabilité sur une zone. L'indice global de vulnérabilité (G_v) [Vías et al., 2010] se définit comme la moyenne pondérée des classes de vulnérabilité :

$$G_v = \frac{\sum cl \times w}{100},$$
 où cl est la valeur de la classe de vulnérabilité et w le pourcentage de surface occupé par cette classe.

L'analyse de vulnérabilité, contient nécessairement une phase d'évaluation de la sensibilité au climat. De plus les modèles biogéochimiques utilisés pour les projections d'impacts du changement climatique (par exemple PaSim [Graux et al., 2012a, b]) sont généralement consommateurs en temps de calcul et requièrent beaucoup d'entrées. C'est pourquoi il faut faire appel aux plans d'expériences, afin de réduire le nombre de simulations, et au calcul intensif pour distribuer les calculs.

2.3. Les plans d'expériences

La conception de plans d'expériences (DOE, "*Design of Experiment*") a une histoire riche, incluant de nombreuses avancées théoriques et des applications pratiques dans des domaines variés. Depuis le début des simulations par ordinateur jusqu'à nos jours, le domaine de la conception des plans d'expériences est un champ de recherche très actif [Kempthorne 1952] [Amblard et al., 2003]. En effet, les DOE sont, dans le domaine de la modélisation, un outil nécessaire pour tester et analyser efficacement le comportement d'un modèle [Kleijnen 1987]. La plupart des simulations ont pour but d'explorer et de tester le comportement d'un modèle. En DOE, un paramètre, ou une entrée, est appelé un facteur. Il peut être qualitatif ou quantitatif [Kleijnen et al., 2005]. Chaque facteur peut prendre différentes valeurs que l'on appelle niveaux. Si, d'un côté les expérimentations sur des systèmes réels ne peuvent que très difficilement contrôler plus de dix facteurs, la réalité, quant à elle, intègre un beaucoup plus grand nombre de facteurs. De même, les modèles complexes possèdent un grand nombre de facteurs qui sont, soit fixés par l'utilisateur, soit définis par un plan d'expériences. Les DOE font partie intégrante du processus de développement d'un modèle [Kleijnen et al., 2005]. En effet, que ce soit pour la vérification, la validation ou les différentes utilisations possibles du modèle, un grand nombre de simulations du modèle sont nécessaires. En particulier, dans la modélisation des environnements dynamiques, les modèles sont devenus incroyablement plus complexes au fur et à mesure que la puissance de calcul des ordinateurs augmentait. A cause du grand nombre de paramètres requis par le modèle et du temps de calcul nécessaire pour une seule simulation, le temps nécessaire pour un plan d'expériences factoriel, uniforme et complet est généralement trop coûteux pour les meilleures machines séquentielles. C'est pourquoi l'utilisation d'autres types de plans d'expériences, ainsi que la distribution des simulations sont nécessaires. L'utilisation d'un plan d'expériences approprié permettra d'avoir toutes les informations recherchées, tout en ayant le nombre minimum de simulations nécessaires pour un résultat de bonne qualité et avec un temps d'attente pour l'obtention des résultats aussi petit que possible. Par exemple, dans le cas de l'analyse de sensibilité, le plan d'expériences est important pour avoir l'information de la sensibilité de tous les paramètres, sans pour autant négliger leurs interactions.

Le domaine des plans d'expériences est également parfois appelé "Conception et analyse des expériences de simulations" ("*Design and Analysis of Simulation Experiment*", DASE) [Kleijnen 2008a]. Dans les simulations expérimentales (*i.e.* réalisations d'expériences à l'aide de simulations), on change les valeurs des paramètres des variables d'entrée, et parfois la structure du modèle, dans le but d'analyser les effets d'un tel changement [Kleijnen et van Groenendaal, 1992]. Une conception expérimentale (*experimental design*) est un ensemble de combinaisons de niveaux de facteurs. Dans la majorité des cas l'évaluation de toutes les combinaisons des niveaux de chaque facteur, appelée plan complet, serait trop coûteuse en temps de calcul. Le but de la théorie des

plans d'expériences est de gagner en connaissance sur le comportement du modèle tout en observant relativement peu de combinaisons de facteurs. Il s'agit de trouver un échantillonnage pertinent de l'espace des paramètres. On parle également de problème de remplissage de cet espace (problèmes dits de « *space filling* ») afin d'observer au mieux le comportement des modèles.

Les plans d'expériences s'utilisent pour trois objectifs :

- augmenter notre compréhension d'un système ou d'un modèle de simulation ;
- trouver des politiques ou des décisions robustes ;
- comparer les mérites de différentes politiques et décisions.

Si l'on cherche à comparer un petit échantillon (de politiques ou de scénarios), différentes méthodes peuvent être utilisées, telles que les méthodes de classements et de sélections, simples ou multiples. On distingue deux approches possibles [Kleijnen et al., 2005] :

1. sélectionner avec une forte probabilité le système ou la politique qui (pour notre objectif) est le meilleur parmi les choix potentiels ;
2. filtrer les systèmes ou les politiques potentielles de façon à obtenir un échantillon de "bons" systèmes ou de "bonnes" politiques.

Le choix des critères qui mesurent la performance d'un système est crucial car il influence fortement le résultat. Par exemple, la gestion d'une prairie qui est optimale par rapport au rendement n'est pas la même que celle qui est optimale par rapport au stockage de carbone. Or, il s'avère que très souvent le problème à résoudre possède en fait plusieurs mesures de performance (par exemple, pour la prairie : le rendement, sa qualité et le stockage de carbone). La plupart des méthodes citées plus loin n'utilisent qu'une seule mesure de performance qu'elle maximise ou minimise. L'approche par conception (*design*) robuste de [Taguchi 1987] offre une alternative intéressante dans le cas de multiples mesures de performance. En effet, si les réponses sont converties en pertes dans une échelle appropriée, on peut reconstruire un modèle de perte totale attendue et ainsi ne minimiser qu'un seul critère. Par exemple, si l'on convertit le rendement fourrager, la production de lait et le stockage de carbone (ici, nos mesures de performance) en euros (ici, notre échelle commune), nous avons alors un modèle n'ayant qu'un seul critère à maximiser. Une autre approche consiste à minimiser une sortie tandis que les autres sorties doivent respecter certaines contraintes fixées par l'utilisateur [Angün et al., 2002].

2.3.1. Plans d'expériences classiques

[Kleijnen 1987] a distingué trois types de plan d'expériences classiques (Figure 13, Figure 14) :

1. Un seul facteur à la fois : comme son nom l'indique, on ne bouge qu'un seul facteur à la fois par rapport à la position initiale. Il est intéressant de noter que la définition est différente de celle donnée par [Morris 1991] en ce sens que, selon Morris, un plan OAT ("*One factor At a Time*") se caractérise par le fait que l'on ne change qu'un facteur (autrement dit on ne se déplace que dans une seule direction) par rapport à la position précédente.
2. Plan factoriel (ou maillé) complet : on étudie toutes les combinaisons possibles, c'est-à-dire que si nous avons f facteurs et que chaque facteur a n niveaux alors le nombre de simulations est n^f . Bien qu'ils soient coûteux et du coup généralement utilisés pour un faible nombre de facteurs (le nombre d'expériences augmente exponentiellement avec f), les plans factoriels complets ont l'avantage d'être faciles à créer, à tracer et à analyser, et les puissances de calcul autorisent maintenant de plus en plus cette approche sur un nombre de facteurs et de niveaux raisonnables (*i.e.* en adéquation avec la puissance disponible).
3. Plan factoriel incomplet, que l'on appelle en général plan fractionnaire. Un inconvénient du plan factoriel complet (outre son nombre de simulation, 2^f pour deux niveaux de facteur) est que s'il n'y a pas d'interaction¹² entre facteurs, une fraction seulement de ces observations étaient nécessaires (égale à $1/2^p$). Le principe du plan fractionnaire est d'utiliser une fraction du plan factoriel complet, choisie de telle manière qu'elle contienne distinctement les effets principaux, et éventuellement, en fonction de la taille, les interactions. En pratique, si on a $k+1$ effets, il suffit de $k+1$ observations. Ce type de plan est souvent préféré par rapport au plan factoriel complet pour son faible nombre de simulations nécessaires et pour son efficacité. Plus il y a de paramètres, plus le plan fractionnaire devient efficace, à la différence du plan à un seul facteur à la fois : la variance de l'estimateur de l'effet principal est de σ/f pour le plan factoriel tandis que pour le plan à un seul facteur à la fois, la variance est de $\sigma/2$. Le principal inconvénient de cette approche est que ce plan fractionnaire n'est pas adapté dans le cas où il y a des interactions entre les facteurs. Il peut alors être à la source de fausses conclusions.

La représentation spatiale et sous forme de tableau (ci-dessous, Figure 13 et Figure 14) pour 3 facteurs et 2 niveaux nous donne une vision de la couverture spatiale obtenue par les différents types de plans d'expériences.

¹² On distingue les effets principaux des effets d'interactions. Les premiers étant les impacts sur la fonction d'intérêt d'une modification élémentaire d'un seul facteur à la fois, tandis que les effets d'interaction représentent l'impact sur la fonction d'intérêt de la modification simultanée de plusieurs facteurs

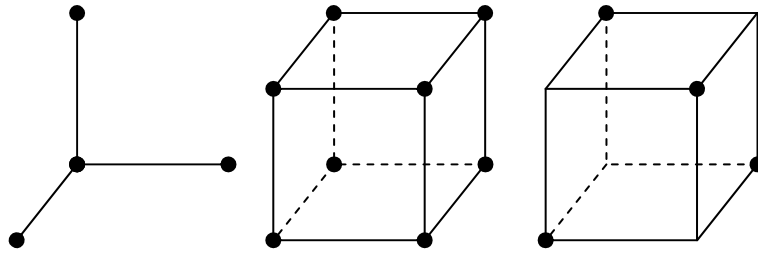


Figure 13 Exemple de plan expérimental à trois facteurs et à deux niveaux : à gauche plan "un seul facteur à la fois", au milieu plan factoriel complet et à droite plan fractionnaire dit à 2^{3-1} (pour 2^{3-1} points couverts sur un espace à 2^3 points).

		Facteurs		
		X1	X2	X3
N° expérience	1°	0	0	0
	2°	1	0	0
	3°	0	1	0
	4°	0	0	1

		Facteurs		
		X1	X2	X3
N° expérience	1°	0	0	0
	2°	1	0	0
	3°	0	1	0
	4°	0	0	1
	5°	0	1	1
	6°	1	1	0
	7°	1	0	1
	8°	1	1	1

		Facteurs		
		X1	X2	X3
N° expérience	1°	1	0	0
	2°	0	1	0
	3°	0	0	1
	4°	1	1	1

Figure 14 Exemple de plan expérimental à trois facteurs et à deux niveaux : à gauche plan "un seul facteur à la fois", au milieu plan factoriel complet et à droite plan fractionnaire dit à 2^{3-1} (pour 2^{3-1} points couverts sur un espace à 2^3 points).

2.3.2. Choix d'un plan d'expériences

Trois stratégies de conception et d'analyse des plans d'expériences

D'après [Gauchi 2007], trois stratégies de conception et d'analyse des plans d'expériences sont possibles :

1. On ne possède pas de base d'exemples, ni de métamodèle au sens statistique (métamodèle étant dans ce contexte un modèle de modèle [Kleijnen 1987]). Notre but sera donc de répartir au mieux les simulations dans l'espace (des entrées), ce qui permettra d'avoir une première base d'exemples.

2. On a une base d'exemples et la démarche adoptée est non-séquentielle. Grâce à ces exemples on estime un métamodèle. Un métamodèle étant (dans le domaine des plans d'expériences) un modèle de modèle, à savoir un modèle plus simple capable de reproduire les sorties issues du simulateur en fonction de ses entrées, grâce par exemple à une technique basée sur l'apprentissage, comme un modèle supervisé à base de réseaux de neurones (par exemple, [Aussem et Hill, 1999]). On peut alors compléter notre base d'exemples d'un autre ensemble d'expériences, de façon à raffiner notre métamodèle. Ce nouveau plan d'expériences pourra être choisi pour répondre à certains critères définis ci-après (comme l'optimalité). Cette approche présente l'avantage d'être distribuable, puisque chaque expérience est indépendante des autres expériences du plan d'expériences.
3. On dispose d'une base d'exemples et la démarche adoptée est séquentielle. On construit d'abord un métamodèle à l'aide des exemples. Puis, grâce au métamodèle, on choisit une expérience où l'incertitude est maximale (si on sait la calculer). Une fois réalisée cette nouvelle expérience, on met à jour le métamodèle, puis on recommence à choisir une expérience là où l'incertitude est maximale, jusqu'à ce que notre objectif soit atteint (par exemple, optimisation d'une variable d'intérêt). Cette démarche est bien séquentielle dans la mesure où pour déterminer une expérience, il faut préalablement réaliser les précédentes.

Critères d'évaluation des plans d'expériences

Le choix d'un plan d'expériences se fera en fonction des objectifs (stratégie, niveau d'interaction recherché...) en s'appuyant sur certains critères [Gauchi 2007] [Kleijnen et al., 2005] :

- le critère d'**indépendance** des facteurs deux à deux. Il est calculé par le déterminant de la matrice des corrélations. Pour de meilleures performances, ce critère est à maximiser, le maximum étant 1 et représentant l'absence de toute corrélation.
- le critère du **remplissage de l'espace** ("*space-filling*") et celui de la **protection contre les biais**. Un bon remplissage de l'espace permet de ne pas faire d'hypothèse sur la nature de la surface de réponse. Tandis que la protection contre les biais permet d'éviter de favoriser certains résultats par rapport à d'autre. Une mesure de ce remplissage de l'espace est la **discrépance**¹³, qui représente la

¹³ La discrépance en un point x de l'espace unitaire cubique de dimension s (C^s), pour un plan d'expériences D_n de n éléments, est définie comme $\left| \frac{N(D_n, [\mathbf{0}, x])}{n} - Vol([\mathbf{0}, x]) \right|$, où $N(D_n, [\mathbf{0}, x])$ est le nombre de points du plan d'expériences appartenant au rectangle à s dimension $[\mathbf{0}, x]$ et $Vol([\mathbf{0}, x])$ le volume de ce même rectangle. La L_p -discrépance étoilée ($D_p(D_n)$) est la moyenne par la norme L_p , sur l'espace, de la discrépance :

distance à une répartition uniforme, cette valeur étant à minimiser. Une autre mesure est le critère **maximin**, qui est un indicateur de remplissage. Il est calculé comme la distance entre les deux points les plus proches. Pour de meilleures performances, ce critère est à maximiser. Dans cette catégorie, on trouve notamment les échantillonnages par hypercube latin ("*Latin Hypercube Sampling*") [McKay et al., 1979]. Les plans d'expériences de remplissage de l'espace apportent de la flexibilité dans l'estimation d'un grand nombre d'effets, qu'ils soient linéaires, non-linéaires ou d'interactions. Par conséquent, ils apportent une protection générale contre les biais lors de la conception de métamodèles statistiques par des techniques de régression.

- le critère du **nombre de scénarios**, et donc le temps du calcul associé. Un plan d'expériences est dit "saturé" si le nombre de combinaisons de facteurs est égal au nombre de paramètres du métamodèle.
- le critère d'**orthogonalité**. Une conception est dite **orthogonale** si les colonnes de la matrice du plan d'expériences sont orthogonales (le produit scalaire de deux colonnes vaut zéro). L'intérêt réside dans le fait que les colonnes de paramètres d'entrée sont ainsi **décorrélées**, ce qui facilite l'interprétation des résultats. Cependant l'utilisation de ces plans n'est pas toujours possible, comme par exemple dans le cas du problème de définition de l'espace des entrées ou dans les cas d'instabilité (c'est-à-dire dans les cas où une légère variation dans l'espace des entrées entraîne un résultat très différent). Un plan peut être orthogonal grâce à son codage des facteurs (par exemple si les valeurs possibles sont uniquement -1, +1), mais n'est pas forcément orthogonal pour les valeurs initiales de ceux-ci, et ce changement d'échelle peut entraîner des problèmes.
- Le critère d'**efficience** ou d'**optimalité**. Le plan d'expériences détermine les erreurs standards pour les estimations du métamodèle [Kleijnen et al., 2005]. Différents critères existent dans la littérature. Les plus populaires sont : la **A-optimalité** signifie que la somme des erreurs standards est minimale ; la **D-optimalité** considère l'intégralité de la matrice de covariance des paramètres estimés et cela signifie que le déterminant de la matrice est minimal ; la **G-optimalité** minimise l'erreur quadratique moyenne des sorties prédites par le métamodèle [Koehler et Owen, 1996]. Malheureusement ces critères nécessitent de fortes hypothèses sur le métamodèle, qui sont généralement invalidées par les simulations. En résumé,

$D_p(D_n) = \left[\int \left| \frac{N(D_n, [\mathbf{0}, x])}{n} - Vol([\mathbf{0}, x]) \right|^p \right]^{\frac{1}{p}}$. On appelle également discrédance étoilée ($D(D_n)$) la L_p -discrédance étoilée pour $p \rightarrow \infty$, $D(D_n) = \max_{x \in \mathcal{C}^s} \left| \frac{N(D_n, [\mathbf{0}, x])}{n} - Vol([\mathbf{0}, x]) \right|$. Pour plus de détails le lecteur pourra se référer par exemple à [Fang et Sudjianto, 2006].

l'efficacité est le critère le plus critique quand les simulations sont très gourmandes en temps de calcul [Kleijnen et al., 2005].

- le critère de capacité à **gérer les contraintes** sur les combinaisons de niveaux de facteur. Certaines combinaisons de facteurs sont impossibles et doivent donc être enlevées du plan d'expériences ou celui-ci doit être construit de façon à l'éviter, comme par exemple à travers des espaces d'entrée de formes particulières (sphères, ellipsoïde...).
- le critère de **facilité de construction**. En effet si un plan d'expériences doit être utilisé en pratique, il doit être facile à construire ou déjà disponible sur un grand nombre de plateforme pour ceux qui sont incorporés dans des logiciels.

[Kleijnen et al., 2005] suggère de commencer généralement par un plan d'expériences contenant beaucoup de paramètres, quitte à faire des simplifications par la suite. De plus, si la première expérimentation n'a pas permis de répondre entièrement à l'objectif, les premiers résultats devraient permettre de construire une nouvelle expérimentation pour se focaliser sur les paramètres ou les régions les plus intéressants.

2.3.3. La conception avancée de plans d'expériences

Les plans par "hypercube latins"

Pour la première stratégie de parcours de l'espace (voir 2.3.2; cas où l'on cherche à parcourir au mieux l'espace des entrées), comme le plan factoriel complet se révèle souvent impraticable (temps de calcul prohibitif), d'autres familles de plans d'expériences de remplissage ("*space filling designs*") existent. L'échantillonnage par hypercube latin (LHS, "*Latin Hypercube Sampling*") [McKay et al., 1979] est souvent utilisé. Sa propriété majeure est que la variance de la moyenne de la fonction d'intérêt sur les points du LHS est inférieure à la variance de la moyenne obtenue à partir de l'échantillonnage aléatoire uniforme simple si la fonction considérée est monotone en chacun de ses arguments (entrées). Un LHS se construit par la division de chaque domaine des facteurs en n strates de probabilité $1/n$, puis une seule valeur est tirée aléatoirement pour chaque strate. Ainsi pour chaque facteur, les strates ne sont présentes qu'une fois. Le LHD ("*Latin Hypercube Design*") est une forme dérivée où le point n'est pas tiré aléatoirement dans la strate mais placé au milieu.

Le LHD se définit par les propriétés suivantes (Figure 15) :

- chaque facteur a le même nombre de niveaux
- pour chacun des facteurs, les niveaux ne sont pris qu'une seule et unique fois

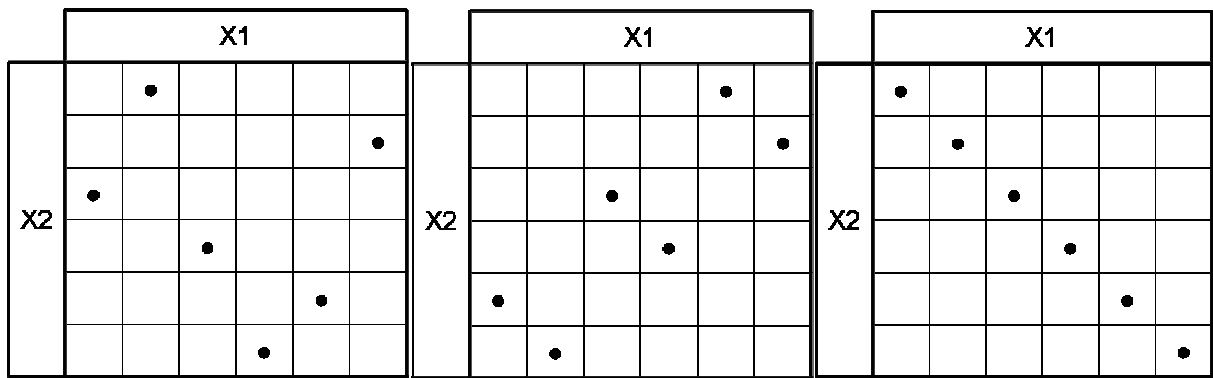


Figure 15 Exemples de LHD pour 2 facteurs (X1 et X2) à 6 niveaux par facteurs

Il existe des variantes du LHD, notamment les LHD symétriques (Figure 16) et les LHD orthogonaux (Figure 17). Parmi l'ensemble des LHD possibles, le choix du candidat peut se faire à l'aide des critères d'optimalité. On parle alors de LHD optimal.

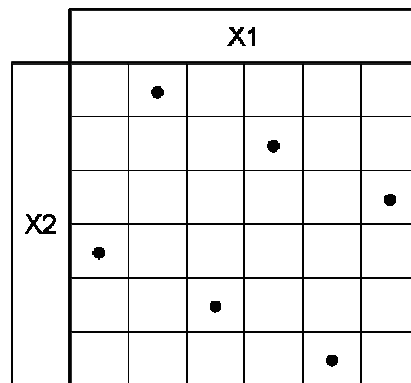


Figure 16 Exemple de LHD symétrique, pour un cas à deux facteurs et six niveaux. On notera graphiquement, la symétrie par rapport au centre.

		Facteurs			
		X1	X2	X3	X4
N° expérience	1°	5	3	1	7
	2°	6	5	2	1
	3°	7	1	6	4
	4°	8	7	5	6
	5°	1	2	4	3
	6°	2	8	3	5
	7°	3	4	7	8
	8°	4	6	8	2

→

		Facteurs			
		X1	X2	X3	X4
N° expérience	1°	+1	-2	-4	+3
	2°	+2	+1	-3	-4
	3°	+3	-4	+2	-1
	4°	+4	+3	+1	+2
	5°	-4	-3	-1	-2
	6°	-3	+4	-2	+1
	7°	-2	-1	+3	+4
	8°	-1	+2	+4	-3

Figure 17 Exemple de LHD orthogonal pour un cas à quatre facteurs (X1, X2, X3 et X4) et huit niveaux (1 à 8). On notera que dans la matrice des plans d'expérience, centrée par rapport au centre de l'espace ((1,2,3,4,5,6,7,8) → (-4,-3,-2,-1,+1,+2,+3,+4)), ce qui permet à la somme de chaque colonne d'être nulle), les colonnes sont bien orthogonales (produit scalaire nul).

Les plans "orthogonal array" et "orthogonal design tables"

Les plans "orthogonal array" (OA) [Hedayat et al., 1999] [Owen 1992] [Owen 1994] [Tang 1993] sont des plans fractionnaires orthogonaux qui servent à étudier les effets principaux de facteurs à plus de deux niveaux par facteur, mais pas leurs interactions. Un OA de force r avec n simulations de s facteurs à q niveaux, nommé $OA(n,s,q,r)$ est un plan factoriel fractionnaire dont chaque sous-plans de n simulations et m facteurs ($m < r$) est un plan factoriel complet [Fang et al., 2006]. Les OA sont souvent exprimés sous forme d'"orthogonal design tables", notées $L_n(q_1 \times \dots \times q_s)$, qui sont des matrices de taille $n \times s$ avec q_i niveaux de facteurs par colonnes tel que :

- tous les niveaux ont le même nombre d'occurrences dans chaque colonne ;
- toutes les combinaisons de niveaux ont le même nombre d'occurrences dans n'importe quel couple de colonne.

Dans le cas où tous les facteurs ont le même niveau, la table est notée $L_n(q^s)$ et correspond à un $OA(n,s,q,2)$. Dans l'exemple de la Figure 18, on notera que tout couple de colonnes forme bien un plan factoriel complet.

		Facteurs			
		X1	X2	X3	X4
N° expérience	1°	1	1	1	1
	2°	1	2	2	2
	3°	1	3	3	3
	4°	2	1	2	3
	5°	2	2	3	1
	6°	2	3	1	2
	7°	3	1	3	2
	8°	3	2	1	3
	9°	3	3	2	1

Figure 18 Exemple d'Orthogonal Design $L_9(3^4)$ correspondant à un OA de force 2 avec 9 simulations de 4 facteurs à 3 niveaux.

Les OA peuvent également servir de base aux "*OA-based LHD*". Pour plus de détails, le lecteur pourra se référer à [Owen 1992, 1994] et [Tang 1993].

Les plans Minimax et Maximin

On distingue également les plans Minimax et Maximin [Johnson et al., 1990] :

- Minimax: on minimise le maximum, sur l'ensemble des points du plan d'expérience, de la distance de ceux-ci à leur plus proche voisin ; c'est-à-dire que l'on minimise $\max_i(\min_{j \neq i} d_{ij})$.
- Maximin: on maximise la distance minimale entre deux points du plan d'expériences ; c'est-à-dire que l'on maximise $\min_{i,j \neq i} d_{ij}$. Les plans maximin sont plus faciles à construire que les plans minimax.

Les plans uniformes de Fang

Les plans uniformes de Fang [Fang 1980] [Wang et Fang, 1981] ont été construits afin de minimiser l'écart, en tout point de l'espace, entre la valeur d'une fonction objectif et celle prédite par le modèle. La fonction objectif désigne la fonction qui sert de critère pour déterminer la meilleure solution à un problème d'optimisation, ce problème d'optimisation pouvant être ici d'estimer l'espérance sur le domaine de la fonction liant la sortie aux entrées. Si l'on considère comme fonction objectif l'espérance (E) et comme modèle

l'estimateur du métamodèle (\hat{y}), cela revient à minimiser la différence absolue entre E et \hat{y} : $\Delta = |E - \hat{y}|$. L'inégalité de Koksma-Hlawka fournit une borne supérieure à Δ :

$$\Delta \leq V(y) D(D_n),$$

où D est la discrétion du plan D_n , composé de n points et $V(y)$ est une mesure de la variation de y (la définition de $V(y)$, au sens de Hardy et Krause, a été donnée par Niederreiter [1992, p. 19]).

L'idée de base des plans uniformes de Fang est de minimiser la discrétion. On notera que les plans uniformes sont des *Latin Hypercube Designs*.

Les plans hypercubes latins quasi-orthogonaux

En se basant sur les idées de [Ye 1998], [Cioppa et Lucas, 2007] ont proposé un algorithme permettant de créer des plans LHD quasi-orthogonaux. La quasi-orthogonalité permet d'avoir une très faible corrélation entre les colonnes, ce qui permet d'avoir une non-corrélation entre les effets linéaires et quadratiques. Ces LHD permettent d'avoir un très bon remplissage de l'espace. Bien que la construction de ces plans soit coûteuse en temps de calcul, un grand nombre de ces tables déjà construites sont disponibles [Sanchez 2005]¹⁴.

Les plans quasi-aléatoires

Les plans quasi-aléatoires (ou "*Quasi-Random*") sont basés sur des générateurs de nombres quasi-aléatoires. Ces derniers parcourent l'espace de manière très uniforme, c'est-à-dire qu'ils minimisent la discrétion¹⁵ (en quelque sorte l'écart à une uniformité parfaite). Ces générateurs visent à remplir progressivement et systématiquement les « trous » laissés dans l'espace parcouru jusque-là. Certains générateurs quasi-aléatoires sont de type fractal (séquence de Van der Corput, ou de Halton pour plusieurs dimensions...). À la différence des générateurs pseudo-aléatoires classiques, les générateurs quasi-aléatoires ne satisfont pas les tests statistiques prévus pour une distribution aléatoire. En effet, le but de ces générateurs n'est pas d'être le plus proche possible du véritable aléatoire, mais de parcourir l'espace le plus uniformément possible, jusqu'à respecter une certaine densité. On peut recommander les générateurs quasi-aléatoires de Sobol ([Sobol 1967]) jusqu'à 40 dimensions. Une implémentation de ce

¹⁴ <http://diana.cs.nps.navy.mil/SeedLab/>

¹⁵ Pour plus de détails, le lecteur pourra se référer par exemple à http://en.wikipedia.org/wiki/Low-discrepancy_sequence

générateur est disponible dans la GSL (Gnu Scientific Library¹⁶) en 32 bits. Pour des implémentations en 64 bits se référer à [Reuillon 2008].

2.3.4. Surfaces de réponse ou métamodèle

Les métamodèles (au sens littéral de modèle de modèle), également appelés surfaces de réponse, sont des approximations de la fonction Entrée/sortie du simulateur.

Création du métamodèle

Les hypothèses standards dans le domaine des plans d'expériences sont que la composante déterministe peut être reproduite par un modèle polynomial des niveaux de facteurs (peut-être après transformation des facteurs ou des réponses), tandis que la composante stochastique peut être caractérisée comme un bruit "blanc" additif.

On appelle « analyse » l'étape consistant à créer le métamodèle. Un certain nombre de méthodes sont possibles. Elles s'apparentent aux techniques de classification. On dissocie les méthodes d'apprentissage supervisée et celles qui ne le sont pas (*i.e.* qui n'ont pas besoin d'un jeu de données d'entraînement). Dans la deuxième catégorie de méthode on trouve de nombreux outils de régression, mais on utilise également la méthode de Krigage, les "fonctions radiales de bases" (RBF, "*radial basis functions*"), l'analyse de Fourier, les splines et les ondelettes. Tandis que parmi les méthodes d'apprentissage supervisé on utilisera les réseaux de neurones et les machines à support de vecteurs (SVM *Support Vector Machines*). D'après [Gauchi 2007], deux critères d'estimation du métamodèle sont principalement utilisés :

- la minimisation du maximum ou de la moyenne de l'erreur quadratique de prédiction sur l'espace considéré ;
- la maximisation de l'entropie ; dans le cas de la méthode de Krigage (voir ci-après), cela revient à maximiser le déterminant de la matrice de corrélations.

Dans un premier temps, on peut se contenter des simulations existantes pour construire le métamodèle, calculer les estimations, et les comparer. Sinon, on peut également postuler un métamodèle, et chercher un plan optimal. L'optimalité du plan peut être évaluée à partir des critères de 'G-' et 'D-optimalité' [Issanchou et Gauchi, 2008]. Il faudra cependant que le métamodèle soit une bonne approximation du modèle.

Il existe une autre stratégie s'appuyant également sur un jeu de données existant et servant à construire le métamodèle. Cependant, cette fois, la démarche est séquentielle. L'idée est de faire les simulations là où l'incertitude est la plus forte. Malgré le problème majeur d'estimation de l'incertitude, cette stratégie est celle qui est la plus riche en termes

¹⁶ <http://www.gnu.org/software/gsl>

de qualité d'information et de données statistiques apportées par les simulations. Une méthode possible est la LDR ("*Leaner Disagreement from experiment Resampling*") [Gazut et al., 2006] [Gazut et al., 2008]. Cette méthode est composée de trois étapes : d'abord une phase d'apprentissage actif permettant la construction de plusieurs métamodèles (réseaux de neurones) ; cette étape s'appuie sur des répliques de données obtenues par la méthode du "*bootstrap*"¹⁷ [Efron et Tibshirani, 1993] ; la deuxième phase consiste à créer une variance de prédiction par une approche de type "*Bagging*" (ou "*Bootstrap Aggregating*") ; et enfin, on cherche le maximum de variance de prédiction afin de le réduire.

Validation des hypothèses

Après avoir choisi un plan d'expériences et une catégorie de métamodèles, il convient de vérifier les hypothèses faites (notamment sur la nature du métamodèle). Ces vérifications peuvent prendre plusieurs formes :

- Vérification du signe : par exemple, est-ce qu'une augmentation (seule) de la concentration en CO₂ atmosphérique induit bien une augmentation du rendement ? Bien que ce genre de validation soit basée uniquement sur l'expertise, elle peut s'avérer très importante, à la fois pour détecter un problème potentiel ou encore pour prendre conscience d'un mécanisme contre-intuitif.
- Comparaison entre les prédictions du métamodèle et un échantillon de nouvelles simulations (celles-ci peuvent être issues d'un petit LHS).
- On peut également vérifier que les erreurs sont bien normales, indépendantes et identiquement distribuées.

¹⁷ Les méthodes de "*bootstrap*" sont des méthodes de ré-échantillonnage permettant d'estimer la distribution d'échantillonnage de presque n'importe quelle statistique en utilisant uniquement des méthodes très simples. Dans les cas les plus simples, de nouveaux échantillonnages sont créés à partir de tirages aléatoires avec remise depuis l'échantillon initial. Pour plus détails, le lecteur pourra se référer à http://fr.wikipedia.org/wiki/Bootstrap_%28statistiques%29

Méthode de Krigage

Il s'agit d'une méthode de prédiction basée sur la combinaison linéaire pondérée de toutes les valeurs de sorties déjà observées [Sacks et al., 1989] [Meckesheimer et al., 2001] [Koehler et Owen, 1996] [Jones et al., 1998]. Ces poids dépendent de la distance entre l'entrée (\mathbf{d}_0) pour laquelle on cherche à prédire la valeur de sortie et des entrées ($\mathbf{d}_1, \dots, \mathbf{d}_n$) déjà simulées : plus \mathbf{d}_0 est proche d'une entrée déjà simulée, plus le poids est important. Ces poids sont déterminés de façon à minimiser l'erreur quadratique moyenne entre les sorties $\mathbf{w}_1, \dots, \mathbf{w}_n$ (respectivement associées à $\mathbf{d}_1, \dots, \mathbf{d}_n$) et les estimations $y(\mathbf{d}_1), \dots, y(\mathbf{d}_n)$ prédites par le métamodèle de Krigage.

La méthode de Krigage ("*Kriging*" en anglais) possède un avantage fort par rapport aux autres méthodes de régression linéaire. En effet, dans le cas de simulations déterministes, les modèles de Krigage donnent, pour les valeurs d'entrée des observations, des valeurs prédites qui sont strictement égales à celles obtenues par simulation. De plus, à la différence des modèles de régression de faible ordre, le Krigage est une méthode à l'échelle globale plutôt que locale.

La forme ordinaire de l'estimateur du métamodèle de Krigage (y) est :

$$y = \lambda'(\mathbf{d}) \mathbf{w},$$

où $\lambda(\mathbf{d})$ est le vecteur des poids, qui dépend de l'entrée \mathbf{d} , et \mathbf{w} les résultats des simulations déjà réalisées.

On peut démontrer que les poids optimaux ont pour valeurs :

$$\lambda_0 = \Gamma^{-1} \left[\gamma + \mathbf{1} \frac{\mathbf{1}' \Gamma^{-1} \gamma}{\mathbf{1}' \Gamma^{-1} \mathbf{1}} \right]$$

où λ_0 est le vecteur des poids en \mathbf{w}_0 ; $\mathbf{1}$ est le vecteur ne comportant que des 1 ; Γ est la matrice de variance-covariance entre les n résultats des précédentes simulations : $\Gamma = \text{cov}(w_i, w_{i'}), i, i' = 1, \dots, n$, tandis que γ représente le vecteur de dimension n contenant les covariances entre les sorties des anciennes simulations w_i et celles de la nouvelle, w_0 :

$$\gamma = \text{cov}(w_i, w_0), i = 1, \dots, n$$

Cependant, comme le vecteur des covariances des w_i par rapport à w_0 (γ) est fonction de w_0 , les poids optimaux sont par conséquent fonction de w_0 . Un problème majeur avec les poids optimaux du Krigage est donc que les valeurs λ_i sont dépendantes de la covariance entre les sorties déjà simulées et celle que l'on cherche à prédire. Or, par définition, celle-ci est inconnue [Kleijnen 2009].

Une solution à ce problème consiste à considérer, moyennant une hypothèse de stationnarité, que les corrélations entre \mathbf{w}_i et \mathbf{w}_g sont uniquement fonctions de la distance $\mathbf{h}(i, g)$ entre les entrées \mathbf{d}_i et \mathbf{d}_g (ce qui permettra en particulier d'estimer $\mathbf{h}(0, i)$ et donc la corrélation entre \mathbf{w}_0 et \mathbf{w}_i). La corrélation est donnée sous la forme générale :

$$\rho(\mathbf{h}) = \prod_{j=1}^k \rho(h_j),$$

où h_j désigne la composante de \mathbf{h} associée au facteur j ($h_j(i, g) = |\mathbf{d}_{ij} - \mathbf{d}_{gj}|$).

Différentes fonctions peuvent être utilisées pour $\rho(h_j)$:

i) $\rho(h_j) = \max(1 - \theta_j h_j, 0),$

ii) $\rho(h_j) = \exp(-\theta_j h_j),$

iii) $\rho(h_j) = \exp(-\theta_j h_j^2),$

où θ_j représente l'importance du facteur d'entrée j (plus θ_j augmente, moins le facteur d'entrée j a d'importance). Les valeurs des θ_j sont obtenues par régression des moindres carrés.

Les plans de type hypercube latin semblent les plus appropriés pour une surface de réponse de type krigeage, compte-tenu de son bon remplissage de l'espace. Sinon (ou en complément) les plans séquentiels sont appropriés ; en effet on possède une erreur d'estimation qui est connue. Le lecteur trouvera plus de détails pour l'utilisation du Krigeage pour les simulations stochastiques dans [Van Beers et Kleijnen, 2003].

Conception de plan d'expérience en fonction du métamodèle

Conception dite de résolution R

La résolution d'un plan d'expérience détermine la complexité du métamodèle que l'on peut lui associer, des plans avec des résolutions plus grandes autorisant des métamodèles plus complexes. Une conception dite de résolution R est un plan d'expériences dans lequel aucun effet à p-facteurs n'est confondu avec un autre effet contenant moins de R-p facteurs [Box et al., 1978, p. 385]. Deux effets sont confondus quand on ne peut pas les estimer séparément. Pour un métamodèle ne contenant que les effets principaux ($p = 1$), il peut être prouvé que les conceptions de type R3 sont les plus efficaces. Pour un plan d'expériences à k facteurs, si k+1 est une puissance de 2, les expérimentations R3 sont des plans fractionnaires (2^{k-p}). Si k+1 n'est pas une puissance de deux mais est multiple de quatre, alors les plans d'expérience R3 sont ceux de Plackett-Burman [Box et al., 1978]. Dans le cas où les interactions entre facteurs sont présentes mais que seuls les effets de premier ordre nous intéressent alors les expériences de type R4 sont celles à utiliser. Si l'utilisateur

est intéressé par les interactions de second ordre alors les conceptions de type R5 sont nécessaires. Cependant leurs temps de calculs sont plus importants : ils nécessitent $O(k^2)$ combinaisons de facteurs. Il est également possible de faire une première expérimentation de type R4 pour déterminer quels sont les facteurs importants pour ensuite réduire la conception R5 à ces seuls facteurs. L'avantage des plans fractionnaires est d'offrir une bonne D-optimalité.

Les plans composites centraux ("Central Composite Design")

Pour un modèle du second ordre (qui donc prend en compte les effets quadratiques et les effets d'interaction entre deux facteurs), le plan expérimental le plus utilisé est celui des plans composites, avec cinq valeurs par facteur ($0, \pm 1, \pm c$ avec $c \neq 0,1$) [Box et al., 1978] [Myers et Montgomery, 2002]. Il est également possible d'utiliser les « saturated designs », ces derniers étant des plans factoriels fractionnaires dans lesquels le nombre d'effets (principal ou d'interaction) estimés est égal au nombre de simulations [Fang et al., 2006].

Frequency based design

Dans ce type de conception, chaque facteur (forcément quantitatif) oscille sinusoidalement entre sa valeur minimale et sa valeur maximale à une fréquence qui lui est spécifique. Cette approche qui initialement faisait bouger les paramètres pendant la simulation, peut également être appliquée sur les paramètres en entrée (ceux-ci changeant d'une simulation sur l'autre) [Lucas et al., 2002] [Sanchez et Wu, 2003]. Ces plans ont un bon remplissage de l'espace. Ils peuvent permettre de créer un métamodèle du second et troisième ordre. De plus, ils sont relativement faciles à construire et à implémenter [Sanchez et Wu, 2003].

"Sequential bifurcation"

Ce type de plan est très utile dans le cas où seul un petit nombre de facteurs sont importants. Cette méthode est basée sur la création de groupes qui seront subdivisées jusqu'à l'obtention de groupes de facteurs dont les effets agrégés sont soit insignifiants, soit négatifs, ou bien jusqu'à l'obtention de facteurs individuellement significatifs. Cette méthode est décrite dans [Bettonvil et Kleijnen, 1997] et [Cheng 1997]. Cependant de nombreuses autres méthodes de screening existent. La méthode SB ("*Sequential Bifurcation*") utilisent les trois hypothèses suivantes sur le métamodèle [Kleijnen 2008b] :

- Un polynôme de premier ordre avec des interactions de premier ordre est un bon métamodèle ;
- Tous les effets de premier ordre sont positifs ou nuls ;

- Tout facteur n'ayant pas d'effet principal important, n'aura pas d'effet d'interaction avec les autres facteurs [Wu et Hamada, 2000].

Plans d'expériences croisés et combinés

Dans le cas où l'on s'intéresse principalement à l'effet d'une décision, seule une partie des facteurs est importante. L'utilisateur peut alors décider d'utiliser un plan d'expériences croisé, c'est-à-dire un plan croisant une forte résolution pour les facteurs de décision et une plus faible résolution pour les autres.

Par opposition à l'approche classique qui consiste à créer un plan d'expériences pour les facteurs de contrôle et un pour les facteurs dits environnementaux (qui génèrent un bruit), puis de faire le produit de ces plans d'expériences, les plans d'expériences combinés [Shoemaker et al., 1991] [Myers et al., 1992] sont des plans dans lesquels les facteurs de contrôle et environnementaux sont combinés dans la même table. Comme cela est le cas pour un certain nombre de conceptions expérimentales, il faut utiliser la connaissance *a priori* sur les interactions recherchées entre facteurs. Mais [Myers et al., 1992] suggèrent que les seules interactions qu'il est nécessaire d'estimer sont celles entre deux facteurs de décision, ce qui réduit la dimension du plan d'expérience.

De façon à clarifier le manuscrit le terme métamodèle, ne sera à partir d'ici plus utilisé dans le sens de « surface de réponse » qui est lié au domaine des plans d'expériences et aux techniques statistiques. En effet, dans ce manuscrit qui présente principalement des travaux d'informatique, il ne prendra plus que le sens donné par le domaine de l'ingénierie dirigée par les modèles. Même si les plans d'expériences permettent de réduire le nombre de simulations nécessaires, le temps de calcul mono-CPU reste généralement problématique pour une analyse de vulnérabilité au changement climatique et nécessite l'usage d'outils adaptés au calcul intensif.

2.4. Le calcul intensif

À la jonction des mathématiques et de l'informatique, le calcul intensif et la simulation illustrent bien le triptyque scientifique fondateur « théorie-expérience-simulation » [ANR 2010]. La société se pose des questions cruciales sur le réchauffement climatique. Les réponses sont très largement dépendantes de l'accès à des moyens de calcul intensif. En effet, réduire les incertitudes, quantifier la probabilité d'événements extrêmes, dénombrer les puits de carbone et leurs possibles évolutions, étudier les impacts sur les écosystèmes, nécessitent une très grande puissance de calcul.

Si la recherche représente 15% de l'utilisation des ressources informatiques, on peut être surpris de noter que les secteurs tel que la finance, les services de l'information, la géophysique, les services de logistique ou les semi-conducteurs représentent en cumul

32%. Dans les autres secteurs minoritaires on trouve notamment la santé et la biologie dont les besoins sont en augmentation de manière significative.

2.4.1. Les supercalculateurs

La première réponse au besoin croissant en puissance de calcul fut les supercalculateurs, qui sont des ordinateurs uniquement dédiés à l'exécution de tâches de calcul intensif. Pour améliorer les performances de ces derniers, une première solution consiste à augmenter les performances des processeurs séquentiels. Mais cela n'est pas suffisant pour répondre au besoin toujours croissant des utilisateurs, aussi des solutions de parallélisation ont été envisagées. Parmi ces solutions on trouve, à partir de 1970, les instructions vectorielles qui permettent aux ordinateurs de traiter une même instruction simultanément sur plusieurs données. D'autres moyens pour paralléliser des calculs plus généralistes ont donc été mis en place et les supercalculateurs se sont rapidement dotés de plusieurs processeurs travaillant en parallèle [Hoffman et Traub, 1989]. Les supercalculateurs sont passés d'une dizaine de processeurs dans les années 80 (avec une puissance de l'ordre de grandeur du GFlops¹⁸) jusqu'à atteindre (en 2012) environ 500 000 processeurs (pour une puissance de 17,6 PFlops, pour la machine Titan [Top500 -2012¹⁹]). L'avantage de cette forte augmentation du nombre de processeurs est que, sur ces derniers, différentes instructions (appartenant éventuellement à différents exécutables) peuvent être exécutées en parallèle. Cependant, cette méthode de parallélisation n'est pas parfaite. D'une part, tous les problèmes ne permettent pas d'utiliser la parallélisation. D'autre part, il faut repenser l'application pour l'adapter à ce genre d'architecture, notamment en la découpant en parties aussi indépendantes que possible, mais également en gérant les communications entre processus.

2.4.2. Cluster & grille

Dans les environnements pour le calcul distribué, on distingue notamment les fermes de calcul et les grilles. Une ferme de calcul ou "Cluster" est un ensemble d'unités de calculs regroupées physiquement en un même lieu afin d'additionner les puissances élémentaires et de fournir une plus grande puissance de calcul et/ou de stockage distribué. Si les clusters permettent "simplement" d'accéder à une grande puissance de calcul, ils ne sont cependant pas toujours suffisants pour faire face à des demandes massives. De plus l'utilisation de ressources de type ferme de calcul n'est pas continue (les besoins sont répartis dans le temps de façon non uniforme, dans un laboratoire par exemple).

¹⁸ Les Flops sont unité de mesure de vitesse d'un système informatique. Ils correspondent au nombre d'opérations à virgule flottante par seconde (en anglais, **F**loating point **O**perations **P**er **S**econd).

¹⁹ <http://www.top500.org/list/2012/11/>

Sans avoir à sa disposition de supercalculateur, il reste possible d'absorber des demandes massives de calcul et de stockage, telles que, par exemple, celles liées à la gestion et à l'analyse des données du *Large Hadron Collider* (LHC) ou du télescope *Large Scale Spacio-Telescope* (LSST) en faisant appel à une approche distribuée. Celle-ci consiste à mettre en commun et en réseau les ressources de nombreux centres de recherche. On passe alors d'une puissance de calcul de l'ordre de grandeur de quelques centaines de processeurs pour des fermes à des dizaines et des centaines de milliers de processeurs dans le cadre d'une grille de calcul [Foster et Kesselman, 1999]. Depuis la naissance du concept, dans les années 1990, jusqu'à maintenant, la grille (vue comme une infrastructure de calcul distribué pour la science et l'ingénierie de pointe) a pris une place centrale en tant qu'outil dans le domaine de la simulation. En effet, suite au développement de l'intergiciel Globus [Foster et Kesselman, 1997] [Foster 2001], le CERN adopte, en 1999, le principe de la grille pour traiter les données du futur LHC (*Large Hadron Collider* : Grand collisionneur de hadrons, opérationnel à partir de 2010) et crée le LCG (*LHC Computing Grid* : grille de calcul du LHC). Puis entre 2001 et 2003, le projet européen DATAGRID permet la création d'une grille de test européenne. De 2004 à 2009 trois projets européens EGEE (*Enabling Grid for E-sciencE*²⁰) permettent la mise en place d'une infrastructure de grille de production. Cette dernière est véritablement opérationnelle à partir de 2007. À partir de 2010 est créée l'organisation EGI (*European Grid Infrastructure*²¹) dont le but est de soutenir les services de la grille. Afin d'uniformiser les composantes de la grille, des standards ont été générés par l'OGF (*Open Grid Forum*²²), l'organisme de standardisation des grilles. A l'échelle de la France, des initiatives ont également lieu, notamment l'Institut des Grilles²³ créé en 2007 (devenu en 2011 l'Institut Des Grilles et du Cloud) qui est la structure du CNRS en charge de la coordination et de l'animation de l'ensemble des travaux du CNRS dans le domaine des grilles et du Cloud Computing. Il pilote également le Groupement d'Intérêt Scientifique (GIS) France Grilles²⁴, créé en 2010 par huit établissements scientifiques dont le but est de coordonner le déploiement et la gestion de grilles au niveau national pour l'ensemble de la communauté scientifique.

De façon à répondre aux besoins de partage de ressources définissant clairement les pourvoyeurs et les utilisateurs de ressources et alliant un haut niveau de sécurité, des **Organisations Virtuelles** (*Virtual Organisations, VO*), c'est-à-dire des communautés d'utilisateurs de la grille ont été créées [Foster et al., 2001]. Ces dernières se regroupent par domaine, par exemple « Biomed » pour les recherches dans le domaine médical et en biologie. Au sein de chaque VO, on définit les restrictions (autorisations) d'accès aux

²⁰ <http://www.eu-egee.org/>

²¹ <http://www.egi.eu/>

²² www.ogf.org

²³ <http://www.idgrilles.fr/>

²⁴ <http://www.france-grilles.fr/>

ressources (clusters, bases de données, logiciels ...). L'affectation de ces mêmes ressources à la VO est faite localement par l'administrateur.

L'un des aspects-clé de la grille est son interopérabilité, c'est-à-dire sa capacité à fonctionner avec divers produits ou systèmes informatiques, existants ou futurs, sans restriction d'accès ou de mise en œuvre. L'interopérabilité requiert des protocoles (*i.e.* des ensembles de règles permettant de définir l'échange d'information dans un système de télécommunication) et des services (*i.e.* des surcouches logicielles permettant d'accéder à des fonctionnalités telles que le transfert de fichiers).

La Figure 19 récapitule les éléments constituant de la grille et explicite leurs interactions. L'accès à la grille par l'utilisateur, se fait via une machine nommée **UI** (*User Interface*). Il s'agit d'un serveur de connexion à la grille. Le pré-requis à toute soumission sur la grille est la création d'un proxy **VOMS** (*VO Membership System*), via l'utilisation d'un certificat (clé publique, clé privée). Les fichiers utilisés par la grille sont stockés sur un ou plusieurs **SE** (*Storage Element* : serveur de stockage) et référencés par le **LFC** (*LCG File Catalog* : un catalogue des fichiers). Les jobs sont soumis au **WMS** (*Workload Management Server* : un système de gestion des jobs) qui avec l'aide du **BDII** (*Berkeley Database Information Index* : Système d'information de la grille) ordonnance les tâches. Ces dernières seront transférées aux **CE** (*Computing Element* : Ressources de calcul, e.g., clusters) puis exécutées sur les **WN** (*Worker Node* : Nœud de calcul).

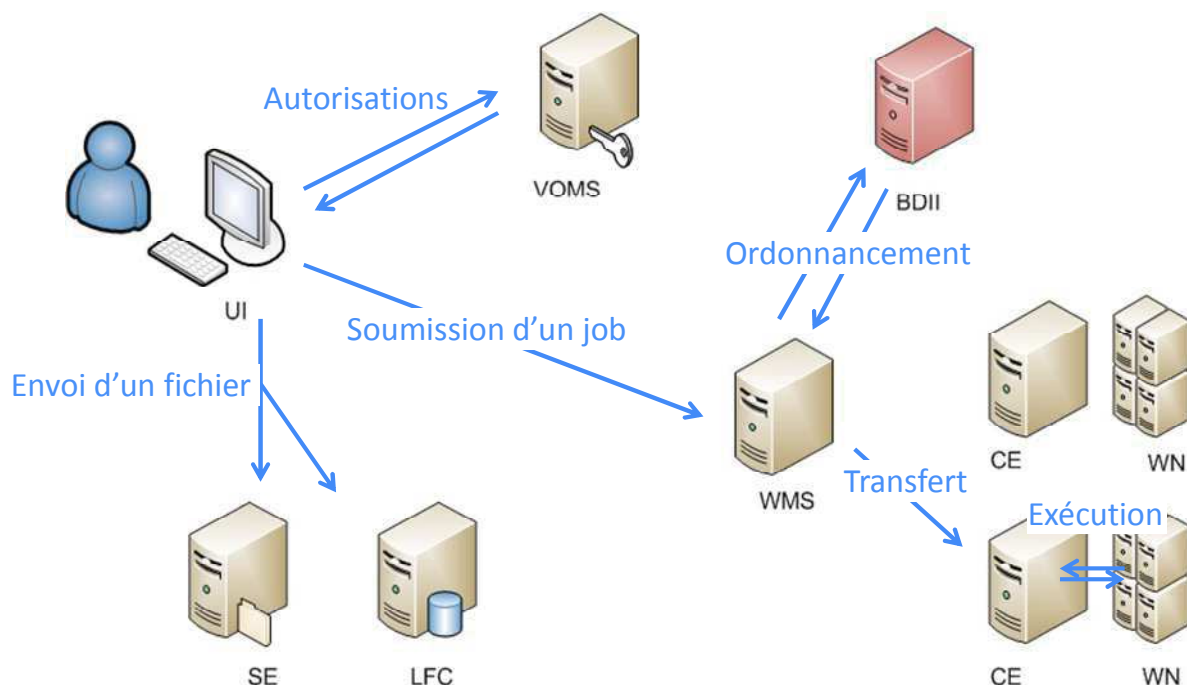


Figure 19. Diagramme de fonctionnement des éléments constituant de la grille [Mahul 2011]

Les systèmes d'authentification pour les environnements de VO ont les caractéristiques suivantes [Butler et al., 2000] :

- une seule authentification nécessaire pour accéder à toutes les ressources de la grille ;
- la possibilité de déléguer à un programme un droit d'accès aux ressources similaire à celui de l'utilisateur ;
- une complète compatibilité entre toutes les solutions locales de sécurité ;
- des relations de confiance basée sur les utilisateurs, dans le sens où un utilisateur ayant accès aux sites A et B, doit pouvoir utiliser ces deux sites sans intervention des administrateurs de sécurité.

Ces aspects de sécurité sont obtenus à travers l'utilisation de certificats et de systèmes de gestion des utilisateurs et de leurs autorisations (**VOMS** : *VO Membership System*).

2.4.3. Le cloud computing

L'intérêt des grilles et des clouds est que les ressources sont utilisables par les utilisateurs locaux, mais également par d'autres collaborateurs ou utilisateurs distants [Shiers 2009]. Cela répond en partie au problème de lissage des besoins en ressources de calcul qui se réalise plus sous forme de « pics » au sein d'un laboratoire. Cette approche favorise ainsi une utilisation plus optimale des ressources. Les clouds se comportent comme des grilles, mais avec un niveau d'interface supérieur et ils sont généralement accessibles de manière payante (coût à la ressource), tandis que l'accès aux grilles est généralement simplement lié à la participation et à la maintenance en locale de cette dernière. Le principe du cloud computing peut être résumé comme l'utilisation de ressources informatiques (matérielles et logicielles) délivrées sous formes de services à travers un réseau (généralement internet). Il se base sur un principe de virtualisation des ressources. La grande flexibilité du cloud (notamment par rapport à la grille) est due à trois relations de service :

1. *Software as a Service* (SaaS) : l'application est découpée en services ;

Le principe est de fournir des applications accessibles vers le web, en fonction des besoins du client. Dans cette catégorie de services on trouve par exemple Google Apps²⁵ et Salesforce²⁶. L'intérêt de ces services réside dans son ergonomie (aucun déploiement du logiciel, interface web) et son accessibilité (par exemple par téléphone portable). Cependant, d'après [Blanchet et Loomis, 2010], ce type d'approche pose des problèmes en termes de sécurité (accès et propriété des informations), mais aussi sur la pérennité des

²⁵ <http://www.google.fr/apps/>

²⁶ <http://www.salesforce.com/fr/>

services (voir par exemple [CNRS-COCIN 2012, p36]). De plus, il est parfois difficile d'utiliser plusieurs services ensemble.

2. *Platform as a Service (PaaS)* : la plate-forme est granulaire ;

Le client crée ses applications à partir des outils et des bibliothèques fournis par la plateforme. Le client peut contrôler le déploiement de son application, tandis que le fournisseur fournit le réseau, les serveurs, le stockage et les autres services [Mell et Grance, 2011]. Parmi les plateformes on trouve Google App Engine²⁷, Windows Azure²⁸, ... L'avantage de ce genre de service réside dans le fait que les fonctionnalités, comme l'équilibrage des charges, les redondances des services, ... sont fournies par le système, et dans le fait que les programmeurs peuvent se concentrer sur l'application en n'ayant pas à gérer les problèmes et le code de plus bas niveau. Cependant, la plateforme requiert un langage de programmation qui lui est spécifique et les applications créées ne sont pas forcément portables sur une autre plateforme.

3. *Infrastructure as a Service (IaaS)* : l'infrastructure est virtualisée.

Le principe est de fournir par le moyen de machine virtuelle (VM : *Virtual Machine*) les ressources demandées par l'utilisateur. Ainsi, ce dernier aura l'impression d'accéder à du véritable matériel physique (CPU, disque mémoire, réseau, ...). Parmi les applications commerciales de cloud computing on trouve par exemple Amazon Web Services²⁹, GoGrid³⁰, FlexiScale³¹, ElasticHosts³². D'après [Blanchet et Loomis, 2010], l'avantage d'une telle approche, c'est qu'elle permet l'exécution d'environnements personnalisés par l'utilisateur, que cet environnement est accessible à tout moment par une interface de programmation (API, *Application Programming Interface*) simple et que l'utilisateur a un contrôle complet sur sa ressource virtualisée. Cependant, les interfaces sont non standardisées et la création de machine virtuelle est difficile. Comme souligné par [CNRS-COCIN 2012], « l'émergence des clouds basés sur des techniques de virtualisation permet d'envisager de créer des infrastructures informatiques beaucoup plus souples et accessibles à l'ensemble des communautés scientifiques ».

2.4.4. Le calcul hybride

La baisse de l'augmentation des performances des microprocesseurs a mis en avant une autre forme de processeurs : les processeurs graphiques (GPU : *Graphical Processing Unit*). Le traitement des données graphiques est généralement répétitif (*i.e.* qu'une même

²⁷ <https://cloud.google.com/products/>

²⁸ www.windowsazure.com/

²⁹ <http://aws.amazon.com/fr/>

³⁰ <http://www.gogrid.com/>

³¹ <http://www.flexiscale.com/>

³² <http://www.elastichosts.com/>

opération s'applique à tout un ensemble de données) et nécessite une précision moindre par rapport aux calculs généralement effectués sur CPU (*Central Processor Unit*). Pour cela des architectures massivement parallèles ont été créées. L'apparition, au début des années 2000, d'API (OpenGL [Woo et al., 1999], DirectX [Bargen et Donnelly, 1998]) permettant de programmer des instructions sur GPU, ont permis de commencer à utiliser la puissance des GPU (par exemple [Rumpf et Strzodka, 2001], [Harris et al., 2002] et [Govindaraju et al., 2004]). Ces GPU utilisés à des fins de calculs généralistes sont couramment appelés GPGPU (*General-Purpose computation on Graphics Processing Units*). Puis, en 2007, le lancement d'une nouvelle API (CUDA³³) pour permettre le développement d'applications généralistes et, en 2008, l'apparition d'un standard ouvert OpenCL [Khronos OpenCL Working Group 2008], ont marqué un véritable essor du codage d'applications généralistes sur carte graphique.

Bien que les GPGPU puissent être utilisés de manière efficace pour améliorer les performances de certains calculs, certains algorithmes, à l'inverse, ne se prêtent pas à ce genre d'approche. C'est la raison pour laquelle, à partir de la fin des années 2000, l'utilisation de machines hybrides possédant à la fois des processeurs très puissants et des cartes GPGPU s'est généralisée. Ces machines laissent ainsi la possibilité de développer des applications qui exploitent le meilleur des CPU et des GPU. On trouve maintenant ces machines hybrides parmi les supercalculateurs, mais il est également possible d'accéder à la puissance des GPU à bas coûts, à travers des stations de travail pour un surcoût réduit. Pour plus de détails, le lecteur pourra se référer à [Caux 2012].

2.4.5. Outils de distribution

La parallélisation de simulations permet de réduire le temps d'attente utilisateur en distribuant les calculs sur différentes machines et en gérant les communications nécessaires entre les calculs. Dans certains cas, les calculs sont strictement indépendants et par conséquent la parallélisation s'apparente alors uniquement à une distribution de simulations. Dans notre cas, nous cherchons à distribuer des plans d'expériences qui sont par nature des calculs indépendants, aussi nous nous limiterons à considérer les outils permettant la distribution de calculs et nous ne nous intéresserons pas à la problématique des algorithmes parallèles.

Parmi les outils classiques pour la parallélisation de code, on trouve (de manière non exhaustive) des outils permettant la parallélisation au niveau du code (les plus connus étant MPI, OpenMP, PVM, JSAGA et Gridscale), des intergiciels (tel que CORBA, RMI, Maisie et Proactive), et d'autres outils qui sont à un niveau supérieur d'abstraction (et donc qui masquent une partie de la complexité de la distribution des simulations) tel que DIRAC, Taverna, Ganga, Distme, SimExplorer ou OpenMOLE.

³³ http://www.nvidia.com/object/cuda_home_new.html

Des bibliothèques et API pour code parallèle

OpenMP³⁴ (*Open Multi-Processing*, [Dagum et Menon, 1998]) est une interface de programmation (API, *Application Programming Interface*) pour le calcul parallèle sur architecture à mémoire partagée. Elle est supportée sur de nombreuses plateformes, incluant Unix et Windows, pour les langages de programmation C/C++ et Fortran. Elle se présente sous la forme d'un ensemble de directives, d'une bibliothèque logicielle et de variables d'environnement.

MPI³⁵ (*Message Passing Interface*, [Walker 1994]) est une norme pour la programmation parallèle, utilisable avec les langages C, C++ et Fortran. Elle permet d'exploiter des ordinateurs multiprocesseurs (et/ou éventuellement distants, mais avec une performance réduite, dépendante du réseau d'interconnexion) par passage de messages. C'est un standard de communication pour des nœuds exécutant des programmes parallèles sur des systèmes à mémoire distribuée. MPI a été écrit pour obtenir de bonnes performances aussi bien sur des machines massivement parallèles à mémoire partagée que sur des clusters d'ordinateurs hétérogènes à mémoire distribuée. Les constructeurs en fournissent des versions optimisées pour leur architecture ou leur réseau d'interconnexion. Elle est disponible sur de très nombreux matériels et systèmes d'exploitation. MPI est très portable (car implantée sur un grand nombre d'architectures de mémoires) et performant car, comme nous l'avons précisé, il est possible de trouver des versions optimisées de cette API. Le principal problème concernant les performances réside dans le fait que les meilleures performances sont obtenues pour de petits nombres de nœuds car les coûts de communication inter-nœuds sont rapidement significatifs (loi d'Amdahl [Amdahl 1967]).

PVM^{36,37} (*Parallel Virtual Machine*, [Geist et al., 1994]) est une bibliothèque logicielle permettant de gérer les communications entre ordinateurs (éventuellement très hétérogènes en performance). Il est opérationnel pour les langages C, C++, Fortran, Python et Perl. Le principe est de faire apparaître un réseau d'ordinateur comme un seul ordinateur, appelé machine virtuelle. Le but général de PVM est de proposer une solution simple et hautement productive pour développer des applications fortement partagées. PVM gère tout le routage de messages, la conversion de données et la répartition des tâches au sein d'un réseau d'architectures hétérogènes. PVM a rapidement cédé le pas à MPI.

L'interface de programmation est très simple :

- 1) l'utilisateur écrit une application comme une collection de tâches coopératives ;

³⁴ <http://www.openmp.org>

³⁵ http://fr.wikipedia.org/wiki/Message_Passing_Interface

³⁶ http://www.csm.ornl.gov/pvm/pvm_home.html

³⁷ http://fr.wikipedia.org/wiki/Parallel_Virtual_Machine

- 2) les tâches accèdent aux ressources PVM à travers une bibliothèque de routines standards ;
- 3) ces routines permettent l'initialisation et l'exécution des tâches à travers le réseau ainsi que les communications et les synchronisations entre les différentes tâches.

JSAGA³⁸ (*Java Simple API for Grid Application*) est une implémentation Java, sous licence libre, de la norme SAGA (*Simple API for Grid Application*, [OGF-SAGA 2011]) de l'OGF (*Open Grid Forum*) [Bachar et Benyezza, 2011]. Cette implémentation permet une utilisation uniforme des mécanismes et composants de sécurité, de gestion des données et de gestion de l'exécution, provenant de différents intergiciels de grille (Globus Toolkit³⁹, Unicore⁴⁰, gLite⁴¹, etc.) et de technologies plus classiques (certificats X509, PKI, HTTPS, SFTP, SSH, ...).

Gridscale⁴² est une librairie Scala permettant masquer une partie de la complexité des différents outils de soumission de jobs sur plateformes distribuées. Il est ainsi possible, à l'aide d'une syntaxe très simple, d'accéder notamment à des grilles, des clusters PBS et SLURM ou des serveurs SSH.

Des middlewares/intergiciels

CORBA⁴³ (*Common Object Request Broker Architecture*) est un intergiciel permettant à des morceaux de logiciel, codés dans différents langages et tournant sur différents ordinateurs, de travailler ensemble comme une seule application. Il utilise un langage de définition d'interface (IDL : *Interface Definition Language*) pour définir la communication entre un objet et les autres objets de l'application distribuée. C'est un standard maintenu par l'*Object Management Group* qui n'a pas connu un grand succès. La version 3.2 a été proposée en 2011 [OMG-CORBA 2011].

RMI⁴⁴ (*Remote Method Invocation*) est une interface de programmation (API) pour le langage Java, qui permet d'appeler des méthodes distantes. Cette bibliothèque, qui se trouve en standard dans Java J2SE, est une technologie qui permet la communication via le protocole HTTP entre des objets Java éloignés physiquement les uns des autres, autrement dit s'exécutant sur des machines virtuelles Java distinctes. RMI facilite le développement des applications distribuées en masquant au développeur la communication client / serveur. Cette bibliothèque rentre en concurrence avec la norme CORBA décrite précédemment. Une intégration de RMI a été faite par [Page et al., 1997] pour permettre la distribution de

³⁸ <http://grid.in2p3.fr/jsaga/>

³⁹ <http://www.globus.org/toolkit/>

⁴⁰ <http://www.unicore.eu/>

⁴¹ <http://glite.cern.ch/>

⁴² <https://github.com/romainreuillon/gridscale>

⁴³ <http://www.corba.org/>

⁴⁴ <http://www.oracle.com/technetwork/java/javase/tech/index-jsp-136424.html>

simulations implémentées à l'aide SimJava. SimJava [Howell et McNab, 1998] [McNab et Howell, 1996] est un paquet java permettant l'implémentation de simulation à évènement discrets.

Maisie⁴⁵ [Bagrodia et Liao, 1994] est un langage de simulation basé sur le C, dont le but est de séparer proprement le modèle de simulation, de l'algorithme (séquentiel ou parallèle) utilisé pour l'exécution du modèle. L'intérêt de ce langage est que le programme est indépendant du mode d'exécution (séquentiel ou parallèle). Les communications interprocessus sont en fait basées sur MPI.

Proactive⁴⁶ [Baduel et al., 2006] est un intergiciel java permettant l'exécution de calculs parallèles, distribués et multi-threads. Il fournit un framework et un modèle de programmation pour simplifier la programmation et l'exécution en parallèle d'applications sur des processeurs multi-cœurs, distribués dans un réseau local, sur des clusters et des grilles de calculs. Il s'agit d'un programme open-source.

Des outils à plus haut niveau d'abstraction

DIRAC⁴⁷ (*Distributed Infrastructure with Remote Agent Control*) [Tsaregorodtsev et al., 2003] [Bargiotti et Smith, 2008] [Paterson et Tsaregorodtsev, 2008] est un outil pour gérer des calculs distribués. Il a initialement été développé pour répondre aux besoins de l'expérience LHC (*Large Hadron Collider*) du CERN, pour le traitement de ses données sur la grille de calcul. Il a pour but d'intégrer les ressources de calculs hétérogènes tels que les PC individuels, les clusters et les grilles, tout en simplifiant au maximum leur utilisation. Il est léger, robuste et extensible, adaptable aux évolutions rapides de la grille. Il est d'ailleurs disponible sur un certain nombre de grilles autres que la grille Européenne EGI (*European Grid Initiative*). DIRAC est programmé principalement en Python et les APIs sont fournies dans ce langage.

Taverna⁴⁸ [Missier et al., 2010] [Hull et al., 2006] [Oinn et al., 2006] est une application open source qui permet de construire et d'exécuter des workflows (enchaînements de tâches élémentaires). Chaque étape de ces chaînes de traitements correspond à l'exécution d'un outil disponible localement ou sous forme de service web. L'outil est fortement typé bioinformatique et propose dans sa version de base des outils classiques de bioinformatique. Des API Java peuvent également être intégrées dans l'outil.

⁴⁵ <http://may.cs.ucla.edu/projects/maisie/>

⁴⁶ <http://proactive.activeeon.com/>

⁴⁷ <http://diracgrid.org/>

⁴⁸ <http://www.taverna.org.uk/>

Ganga⁴⁹ [Egede et al., 2005] [Mościckif et al., 2009] est un outil, codé en Python, permettant la gestion de tâches de calcul. Il permet la spécification, la soumission, la gestion de la compatibilité avec les ressources et le post-processing des tâches de calcul sur un ensemble de ressources distribuées. Il masque l'hétérogénéité des ressources. Ganga possède une API, mais il peut également être utilisé à travers son interface graphique. Il s'agit d'un logiciel libre.

GEL (*Grid Execution Language*) [Lian et al., 2005] est un langage de scripts permettant la description et l'exécution de calculs sur grille. Il permet de masquer à la fois l'hétérogénéité de la grille et les temps de communications avec des temps de latence importants. Bien que conçu initialement pour le calcul sur grille, il est opérationnel sur clusters et machines à mémoire partagée, notamment à travers sa capacité à soumettre des jobs en utilisant des gestionnaires de tâches tel que PBS (*Portable Batch System*) ou Globus Toolkit.

CoG Kits⁵⁰ (*Commodity Grid Kits*) [von Laszewski et al., 2001] [von Laszewski et al., 2006] permet aux utilisateurs de la grille, aux développeurs d'applications sur grille et aux administrateurs de grille d'utiliser, programmer et administrer la grille à l'aide d'un outil de haut niveau. Cet outil permet un développement rapide et facile d'applications sur grille, tout en encourageant la réutilisation collaborative de code.

G-Eclipse⁵¹ [Wolniewicz et al., 2007] [Gjermundrod et al., 2008] [Kornmayer et al., 2008] est un environnement de travail permettant d'accéder à la puissance des infrastructures de grille et du cloud. Cet environnement s'appuie sur le fonctionnement de la communauté de l'outil Eclipse. Tout comme Eclipse, il s'agit d'un projet open-source. Il fournit des outils permettant aux utilisateurs de la grille de personnaliser leurs applications, de gérer les ressources sur la grille et de participer au cycle de développement de nouvelles applications sur grille (compilation et débogage distants). G-Eclipse fournit un modèle de grille de haut niveau permettant de manipuler les objets de grille indépendamment de la grille (ou du cloud) cible.

SimExplorer⁵² ([Amblard et al., 2003] [Faure et Deffuant, 2006] [Chuffart et al., 2010]) est un outil open-source, codé en Java, incluant à la fois une API et une interface. Il a été conçu pour permettre la conception et l'analyse d'expérimentations par ordinateur. Sa conception a commencé en 2002 [Amblard et al., 2003], puis son développement a été couplé à celui de l'outil OpenMOLE [Reuillon et al., 2010] (décrit ci-après) pour offrir une plateforme numérique de flux de tâches. SimExplorer combine la gestion de la distribution

⁴⁹ <http://ganga.web.cern.ch/ganga/>

⁵⁰ http://wiki.cogkit.org/wiki/Main_Page

⁵¹ <http://www.geclipse.org/>

⁵² <http://www.simexplorer.org/>

des calculs, des outils d'analyses de sensibilité et une génération automatique de rapport de synthèse. On peut considérer que l'outil contient cinq sous-parties [Chuffart et al., 2008] :

- un outil d'échantillonnage qui permet de définir/sélectionner les valeurs des paramètres en fonction du plan d'expériences sélectionné ;
- un lanceur de modèle qui permet la distribution des simulations ;
- des bibliothèques statistiques qui gèrent l'agrégation et l'affichage graphique pour faciliter l'interprétation du grand volume de données générées ; il est également possible de faire appel à des programmes externes tels que l'outil statistique R ;
- un outil de synthèse qui permet de générer des rapports sur l'expérimentation, synthétisant des données et métadonnées issues des simulations ;
- un système d'information qui garde les informations relatives aux différentes étapes et les résultats de l'exploration, il peut également être synchronisé avec un serveur dans le but de partager ces informations avec d'autres utilisateurs.

A l'état actuel SimExplorer n'est pas opérationnel pour la distribution de simulations sur grille. Bien que SimExplorer permette de répondre à notre problématique d'analyse de vulnérabilité, le logiciel OpenMOLE (dérivé de SimExplorer et décrit ci-après) lui a été préféré car plus "mûr".

Distme⁵³ [Reuillon et al., 2008] est un logiciel open-source codé en Java qui permet la parallélisation de simulations stochastiques. Il s'adapte à la fois aux outils de soumission de jobs et aux différents types d'environnement d'exécutions distribuées. Il permet la distribution des simulations stochastiques selon une approche boîte noire [Reuillon 2008]. C'est-à-dire que l'expérimentateur décrit la simulation stochastique qu'il souhaite distribuer et le logiciel la prépare pour une exécution sur cluster ou grille de calcul. Il a été intégré dans SimExplorer.

OpenMOLE ⁵⁴ (*Open MODeL Experiment*) [Reuillon et al., 2010] est un logiciel gérant de manière générique les flux de tâches de simulations de modèle en utilisant le calcul distribué. Il est décrit plus finement dans la partie suivante.

2.4.6. OpenMole

OpenMole [Reuillon et al., 2010] est issu d'une distribution spécifique de SimExplorer pour gérer les plans d'expériences, dont le but est d'étendre la qualité du processus de modélisation. OpenMole masque l'hétérogénéité des environnements (physique et logiciel)

⁵³ <http://sourceforge.net/projects/distme/>

⁵⁴ <http://www.openmole.org/>

de plateforme de calculs. Il intègre la gestion du flux de simulations (gestion des plans d'expériences : "*workflow engine*") des premières versions de SimExplorer [Faure et al., 2006] [Amblard et al., 2003], permettant de lancer et de distribuer des simulations en local, sur des clusters, sur une grille de calcul ou une grille de PC ("*desktop grid*").

Ses principales caractéristiques et intérêts sont :

- de permettre facilement des analyses de modèle, à travers un logiciel permettant de définir facilement des expérimentations numériques ;
- de faciliter l'intégration de modèles (Java, exécutable compilé ...) ;
- de faciliter la distribution des calculs sur différents environnements (machine multi-cœurs, cluster, grille ou desktop-grid) tout en masquant les spécificités et l'hétérogénéité de ces environnements ;
- de supporter la charge jusqu'à des millions de tâches et des To de données.

Il contient deux modes de fonctionnement. Le premier est le mode "console" où la description du flux de simulation est faite via un script (Figure 20). Ce mode de fonctionnement s'apparente à l'écriture d'un code Scala⁵⁵, à l'aide d'objets et de procédures propres à OpenMOLE. Mais il comporte également une interface graphique (Figure 21, Figure 22) le rendant très facile d'accès. Cette interface permet notamment d'avoir une meilleure vision d'ensemble du flux de tâches. Les informations de l'interface sont stockées dans un fichier XML.

```
val i1 = Prototype[Int]("i1")
val i2 = Prototype[Int]("i2")
val j = Prototype[Int]("j")

val hello = GroovyTask("hello", "j = Model.compute(i1, i2)")

hello addInput i1
hello addInput i2
hello addOutput j
hello addLib "/path/to/model.jar"

val exploration = ExplorationTask(
  "exploration",
  Factor(i1, 0 to 100 by 2 toDomain) x
  Factor(i2, new UniformIntDistribution take 10)
)

val ex = exploration -< (hello by 10 on biomed) toExecution
ex.start
```

⁵⁵ <http://www.scala-lang.org/>

Figure 20 Exemple de script OpenMOLE [Reuillon 2012]. Cet exemple traduit la création d'une tâche "hello" exécutée à l'aide de la classe Model, en fonction de deux paramètres $i1$ et $i2$. Cette tâche, distribuée sur la grille de VO Biomed, est exécutée pour un plan d'expériences composé de l'ensemble des valeurs comprises entre 0 et 100 par pas de 2 pour le facteur $i1$, et de 10 valeurs tirées uniformément entre 0 et 1 pour le facteur $i2$.

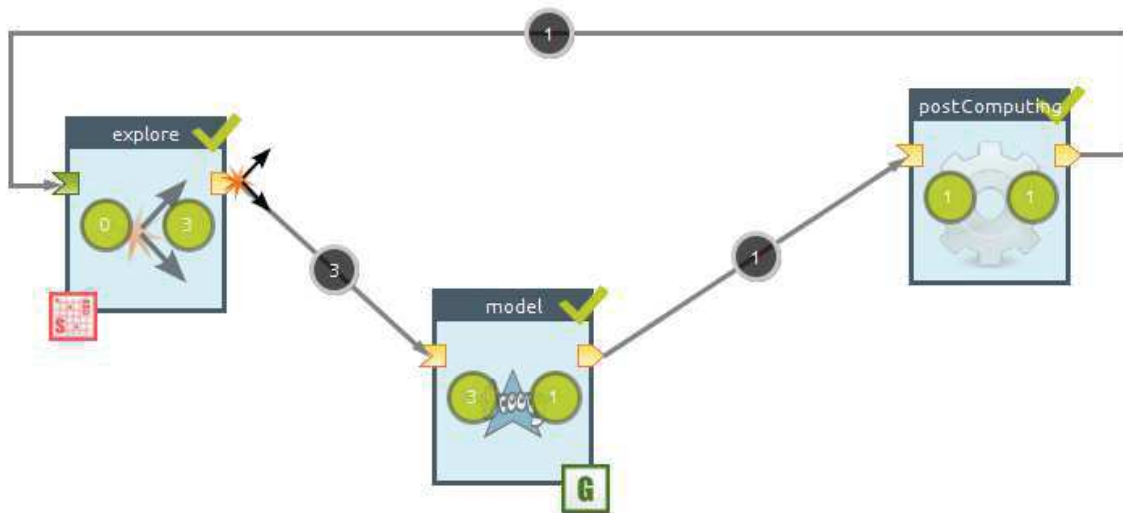


Figure 21 Exemple de description de l'enchainement de tâches et du flux de communication dans l'interface d'OpenMOLE [Reuillon 2012].

L'outil s'utilise de la manière suivante : on commence par définir les facteurs d'un plan d'expériences ainsi que leur nature (entier, double, booléen, programme externe...). Puis, en fonction du type de plan d'expériences souhaité on définit l'ensemble des combinaisons de facteurs à simuler (Figure 22). L'échantillonnage d'une valeur peut également être défini dans un fichier texte de type CSV (*Comma Separated Variables*). OpenMOLE permet d'effectuer des calculs entre les facteurs d'entrée ou de sortie des tâches. De même, on peut définir un certain nombre de conditions entre les tâches (par exemple, si tel facteur est supérieur à un certain seuil, alors on effectue telle nouvelle tâche) (Figure 21). Le tout permettant de définir une « mole », c'est-à-dire un ensemble de tâches. Enfin l'environnement de simulation (*i.e.* grille, cluster, desktop-grid ou PC personnel) est défini.

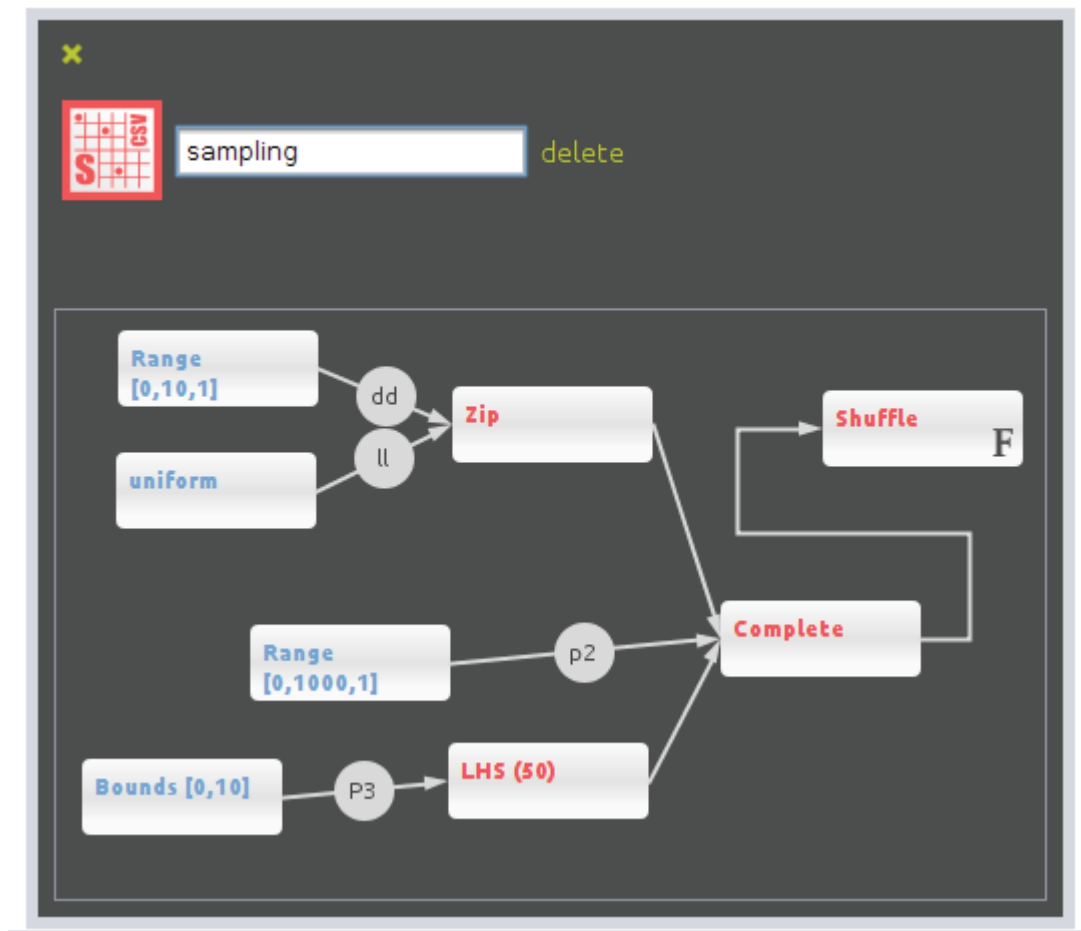


Figure 22 Exemple de description d'une tâche d'échantillonnage (i.e. construction du plan d'expériences) dans l'interface d'OpenMOLE [Reuillon 2012].

Pour résumer, les raisons pour lesquelles nous avons choisi d'utiliser OpenMOLE par rapport à notre problématique sont les suivants :

- OpenMOLE répond à nos besoins de distribution des simulations sur grille et sur cluster ;
- il supporte de gros volumes de données et de simulations ;
- OpenMOLE est un outil ergonomique car son utilisation se fait à travers une interface ou à l'aide de scripts (mode "console") ;
- il est accompagné d'une documentation détaillée et mise à jour rapidement à chaque version ; de plus, l'outil possède une communauté active d'utilisateurs, ainsi qu'un support utilisateur performant (à travers un retour rapide des développeurs aux questions posées) ;
- il peut aisément être couplé avec d'autres outils ; cela est facilement réalisable en utilisant des scripts de syntaxe relativement simple (pour le mode "console" de l'outil) ou un fichier XML (pour le mode interface) ; noter cependant que la DTD (Document Type Definition) du fichier d'interface n'est pas encore stable ;
- il est sous licence libre AGPLv3.

Il est tout de même à noter que son installation nécessite de faire appel à Java 7 qui n'est pas toujours installé sur les machines. D'autre part, les messages d'erreur ne sont pas toujours explicites.

L'analyse de vulnérabilité nécessite de réaliser un grand nombre de simulations. Aussi, un outil générique pour l'analyse de vulnérabilité doit faire appel au calcul intensif. Nous venons de voir qu'il existe une certaine diversité de ressources de calculs (supercalculateurs, clusters, grille, calcul hybride, clouds). Afin d'utiliser ces ressources, il existe un certain nombre d'outils de distribution de plus ou moins haut niveau d'abstraction telle que des bibliothèques et API pour code parallèle (par exemple, OpenMP, MPI, GridScale), des middlewares/intergiciels (par exemple, CORBA, RMI, Maise) et des outils à plus haut niveau d'abstraction (par exemple, Ganga, SimExplorer, OpenMOLE). Parmi ceux-ci nous avons retenu OpenMOLE pour nous aider dans la réalisation d'un outil générique pour l'analyse de vulnérabilité. En effet celui-ci nous permet de réaliser certaines étapes de l'analyse de vulnérabilité (réalisation de plans d'expériences et distribution des calculs). Cependant, la conception de l'outil devant être la plus générique possible (par rapport aux modèles, aux hypothèses et aux outils externes utilisés), l'utilisation de l'Ingénierie Dirigée par les Modèles nous a paru naturelle. Cette dernière apportant généricité et bonne pratique.

2.5. L'ingénierie dirigée par les modèles

L'Ingénierie Dirigée par les Modèles (IDM) est une branche du génie logiciel s'intéressant au développement, à la maintenance et à l'évolution des modèles au sein des systèmes informatiques.

La première architecture logicielle suivant ces principes est nommée MDA (architecture dirigée par les modèles, *Model Driven Architecture* [OMG-MDA 2001]). Elle a été proposée par l'*Object Management Group* (OMG), pour succéder à l'*Object Management Architecture* (OMA). Cette architecture, proposée en 2000, est la déclinaison de l'Ingénierie des Modèles au sein de l'OMG.

On pourrait résumer les approches de la MDA et de l'IDM de la manière suivante [Favre et al., 2006] :

La MDA est une collection de standards industriels promus par l'OMG. Elle se caractérise par une séparation de la description de la partie des systèmes indépendants des plateformes (*PIM Platform Independent Model*) de celle spécifique à une plateforme (*PSM Platform Specific Model*). Tandis que l'IDM offre un cadre méthodologique et technologique qui permet d'unifier différentes façons de faire dans un processus homogène.

2.5.1. Concepts

L'Ingénierie Dirigée par les Modèles s'appuie sur trois éléments, concepts ou notions fondamentales :

1. le concept de modèle ;
2. le concept de métamodèle ;
3. le concept de transformation de modèles.

Le "modèle" est une représentation simplifiée d'un système étudié. Le système étant l'entité que l'on modélise afin de l'étudier, de le comprendre et de l'utiliser à des fins prédictives dans un contexte maîtrisé autre que le contexte réel. [Minsky 1965] a proposé la définition suivante : « Pour un observateur B, un objet A* est un modèle d'un objet A, dans la mesure où B peut utiliser A* pour répondre aux questions qui l'intéressent à propos de A ». [Popper 1973] résume ainsi trois caractéristiques communes à tous les modèles :

1. un modèle doit avoir un caractère de ressemblance avec le système réel ;
2. un modèle doit constituer une simplification du système réel ;
3. un modèle est une idéalisation du système réel.

Notamment dans le cadre de l'IDM, ces trois caractéristiques peuvent se traduire par la relation « **est une représentation de** » entre le modèle et le système étudié ([Hill 1996] [Atkinson et Kuhne, 2003] [Seidewitz 2003] [Bézivin 2004]). Un modèle n'est pas supposé être parfait ni contenir toutes les informations à propos du système étudié. Il est simplement considéré comme « suffisamment bon » pour un certain objectif [Favre 2004b].

Cependant, cette définition n'est pas suffisante dans le cadre de l'IDM car elle ne permet pas de rendre un modèle « productif » (c'est-à-dire interprétable et manipulable par une machine). C'est pourquoi certains auteurs utilisent la définition suivante [Kleppe et al., 2003] :

« Un modèle est une description d'un (ou d'une partie d'un) système écrit dans un langage bien-défini »

La notion de langage bien-défini fait appel indirectement au deuxième principe de l'IDM, à savoir la notion de métamodèle.

Tout comme pour le concept de modèle, plusieurs définitions existent dans la littérature :

- « Un métamodèle est un modèle qui définit un langage pour exprimer un modèle » [OMG 2002] ;

- « Un métamodèle est un modèle de spécification pour une classe de systèmes étudiés, où chaque système étudié dans cette classe est lui-même un modèle valide exprimé dans un certain langage de modélisation » [Kleppe et al., 2003] ;

Contrairement à une idée courante, le métamodèle n'est pas dans ce contexte un modèle de modèles, comme cela pouvait être le cas quand nous parlions des méthodes statistiques. On pourrait plutôt définir le métamodèle comme le modèle d'un langage de modélisation. Cette définition s'appuie sur la relation suivante typique, de la notion de métamodèle : « un modèle **est conforme** à un métamodèle ».

Par exemple, dans le cadre de la programmation par objets, si l'on considère l'objet comme un modèle de la réalité, alors la classe est le métamodèle et la relation entre les deux est bien du type : l'objet **est conforme** à la classe.

Mais les métamodèles ont des formes spécifiques selon les espaces technologiques étudiés, par exemple :

- pour les bases de données : une base est conforme au schéma de la base de données ;
- pour la technologie *Extensible Markup Language* XML : un fichier XML est conforme à un schéma ou à une *Document Type Definition* (DTD) ;
- pour la théorie du langage et de la compilation : un code source est conforme à une grammaire ;
- dans le cadre de la *Standard Template Library* (STL) le vector de type int (vector<int>) est conforme au modèle Vector<T> ;
- dans le cadre de la cartographie, si notre système est la France, le modèle pourrait être une carte IGN et le métamodèle la légende : la carte est conforme à la légende.

On notera qu'à la différence du MDA, les principes de l'IDM sont applicables à tout type de modèle, objet ou non. L'IDM n'est pas restreint à un espace technologique.

L'apport majeur de l'IDM par rapport aux autres méthodes de génie logiciel est sa capacité à rendre les modèles/métamodèles productifs plutôt que « contemplatifs ». Or, pour rendre productif un modèle, il faut décrire comment transformer celui-ci. Cet aspect correspond au troisième principe de l'IDM : « la transformation de modèles ». À la différence des deux autres principes, il n'y a pas de consensus sur sa définition ([Rahim et Mansoor, 2008], [Lano et Clark, 2008], [Iacob et al., 2008]). Elle devrait pouvoir se définir par la relation « **est transformé en** » d'après Favre [2004a]. On pourrait également la définir comme « la génération d'un ou de plusieurs modèles cibles à partir d'un ou plusieurs modèles sources » [Bézivin 2004]. Pour [Diaw et al., 2008], une transformation de modèle comporte deux étapes : tout d'abord une identification des correspondances entre concepts des modèles cibles et sources, au niveau des métamodèles (création des fonctions de transformation) ; puis une étape d'application des fonctions de transformation

par un programme appelé moteur de transformation. On peut différencier les transformations, selon qu'elles soient endogènes ou exogènes, c'est-à-dire selon que les métamodèles cibles et sources soient identiques ou non. Les métamodèles sur lesquels va porter la transformation pouvant appartenir à des domaines techniques différents, on fait alors appel aux "*Domain-Specific Modeling languages*" (DSML). Ces derniers, également appelés langages de modélisation dédiés offrent ainsi aux utilisateurs des concepts propres à leur métier et dont ils ont la maîtrise [Combemale 2008]. Les DSML sont généralement implicites car ils ne sont pas représentés explicitement mais "codés" d'une part dans le code source via des patrons et d'autres part sous forme d'artefacts de bas niveaux ([Favre et al., 2001], [Favre et al., 2006]).

De façon identique à ce que nous avons vu pour la notion de métamodèle, la transformation prend des formes différentes selon les espaces techniques ou niveaux d'abstraction étudiés [Favre et al., 2006] :

- pour les bases de données, nous trouvons les scripts de migrations ;
- pour le langage XML : les transformations sont exprimées en *eXtensible Stylesheet Language* (XSLT) ;
- pour la théorie des langages : les transformations sont prise en charge par la compilation et plus particulièrement par la génération de code ;
- Pour les modèles (M2M, "*Model to Model*") : il existe des langages tels que l'*ATLAS Transformation Language* (ATL) ou QVT (*Query View Transform*) [OMG-QVT 2011].

Le développement dirigé par les modèles (MDD : *Model Driven Development*) est une approche de développement logicielle basé sur l'IDM. On y trouve une utilisation successive de transformations de modèles [Cetinkaya et al., 2011].

Parmi les langages de méta-modélisation les plus populaires on trouve la spécification MOF (*Meta Object Facility*, [OMG-QVT 2011]), UML, *Eclipse Modeling Framework* (EMF, [Eclipse-EMF]). Dans le cas d'EMF par exemple, la transformation de modèles revient à l'exécution d'un métaprogramme. Il s'agit, en effet, d'une plateforme de modélisation et de génération de code.

2.5.2. Un exemple de plateforme intégrant l'IDM

Eclipse Modeling framework (EMF)

Initié par IBM (*International Business Machines*) en 2001, puis développé par la fondation Eclipse depuis 2004, Eclipse⁵⁶ est un environnement de développement (IDE,

⁵⁶ <http://www.eclipse.org/>

Integrated Development Environment) très fortement diffusé. Il a pour but de couvrir toutes les phases de développement (modélisation, conception, implémentation, tests, ...). L'environnement Eclipse couvre un grand nombre de langages de programmation tel que Java, C, C++, C#, Construit de manière modulaire, il intègre de nombreux outils dont le fameux Eclipse Modeling Framework permettant d'intégrer les concepts de l'Ingénierie Dirigée par les Modèles.

L'outil EMF⁵⁷ (*Eclipse Modeling Framework*) est une plateforme de modélisation et de génération de code, s'appuyant sur Eclipse. Il intègre les principes de l'IDM. Commencé par IBM en 2002, le projet EMF avait pour but d'unifier les outils de développement. Il permet de créer et d'intégrer facilement de nouveaux plugins Eclipse contenant les modèles conformes au métamodèle initialement défini. EMF est basé sur le métamodèle Ecore, qui est approximativement une implémentation de "Essential MOF" [OMG-MOF 2006], un sous-échantillon de la spécification MOF ne contenant que les éléments fondamentaux. Les projets EMF permettent de créer le métamodèle (ici un métamodèle de type "Ecore", c'est-à-dire qu'il est conforme au méta-métamodèle Ecore), puis dans un second temps l'instanciation de ce dernier permet de créer les modèles. Les modèles sont générés sous forme de fichiers XMI, ce qui les rend facilement manipulables et réutilisables (éditeurs spécifiques, transformation). XMI [OMG-XMI 2011], signifiant "XML Metadata Interchange", représente un standard, géré par l'OMG, basé sur XML et permettant l'échange de métadonnées UML.

On peut résumer les différentes étapes d'utilisation de la manière suivante⁵⁸ :

1. création du diagramme Ecore (métamodèle) ;
2. spécification au sein de ce diagramme des différentes classes, attributs, références,... (Figure 23) ;
3. à l'aide du générateur de modèles EMF, générer le code du modèle, des éditeurs de modèles et le code des tests unitaires ;
4. création du plugin Eclipse ;
5. à l'aide du plugin, créer un modèle, l'interface du modèle ayant été générée précédemment.

Le code généré est facile d'accès et aisément adaptable en fonction de nos besoins. Un système d'annotation dans Ecore permet à la génération du code de rajouter des commentaires et de spécifier du code. Tandis que des annotations dans le code permettent de désactiver ou d'activer la génération du code de certaines classes ou méthodes. En effet, en java, les annotations⁵⁹ sont des méta-données insérées au sein du code et associées aux classes, méthodes, attributs, paramètres, variables locales et paquets. Ainsi, la JVM (*Java*

⁵⁷ <http://www.eclipse.org/modeling/emf/>

⁵⁸ <http://www.vogella.com/articles/EclipseEMF/article.html>

⁵⁹ http://en.wikipedia.org/wiki/Java_annotation

virtual machine, machine virtuelle Java) et d'autres programmes peuvent accéder à ces méta-données pour déterminer comment interagir avec les composants du programme ou changer leur comportement. Les annotations Java peuvent (si besoin) être accessibles au programmeur lors de l'exécution du programme grâce à la réflexion. Au sein d'EMF il est également possible de faire des héritages (au sens classique de la programmation orientée objet) entre métamodèles.

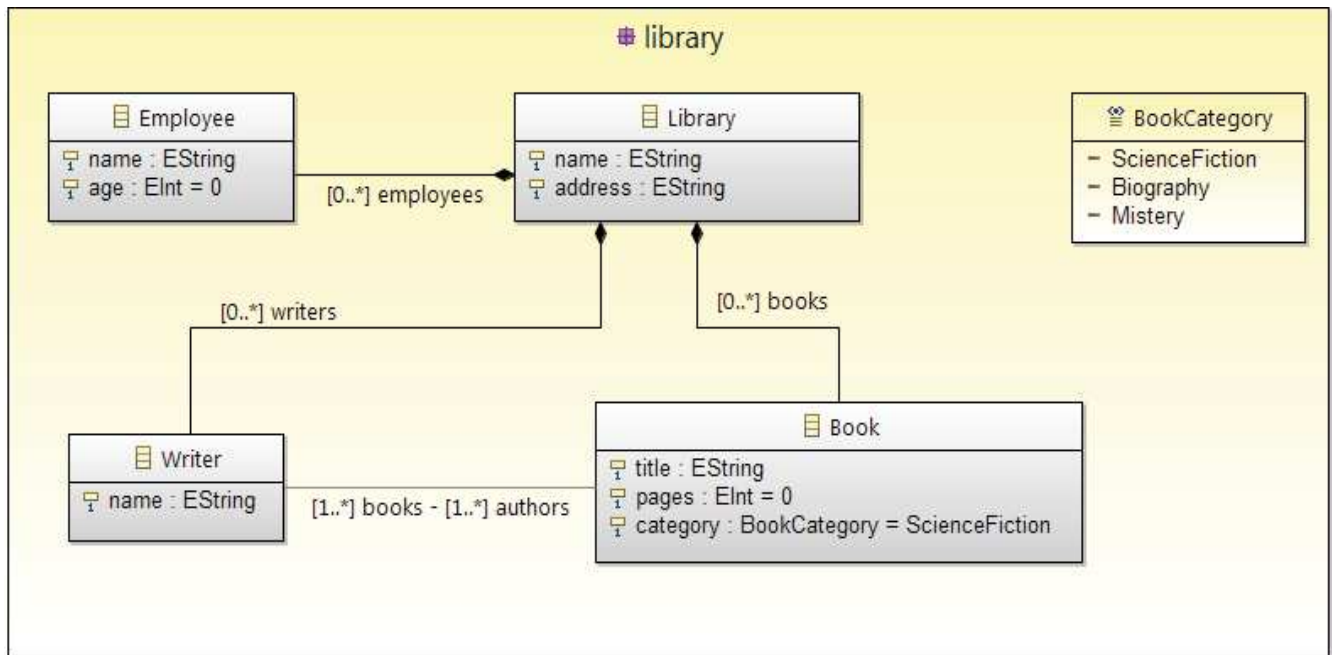


Figure 23⁶⁰. Exemple de diagramme Ecore, représentant le métamodèle d'une librairie. Toute instance de "Library" est conforme à ce métamodèle.

Bien que cet outil nécessite de connaître et de comprendre Eclipse et en particulier les plugins Eclipse, sa portabilité (indépendance du système d'exploitation), sa relative ergonomie (basée sur celle d'Eclipse) et son intégration relativement complète des principes de l'IDM (modèle, métamodèles, transformations de modèles) en font un outil puissant et populaire. De plus, EMF bénéficie de beaucoup des avantages d'Eclipse. Il est ainsi possible de faire appel à d'autres outils, tel que l'*Object Constraint Language* que nous allons détailler dans la partie suivante.

OCL et programmation par contrat

La programmation par contrat a été mise en relief par Bertrand Meyer avec l'approche de conception suivante : "*Designing by contracting*" ([Meyer 1988], [Meyer 1992]). Cette approche peut être vue comme un style de programmation (dans le sens "bonnes pratiques"), dont l'objectif est de produire un code le plus fiable possible au regard de spécifications incluses dans le logiciel. Pour cela on s'appuie sur l'idée de contrat

⁶⁰ http://fr.wikipedia.org/wiki/Eclipse_Modeling_Framework

appliquée aux routines du code. En effet, comme dans un contrat, le client (celui qui appelle la routine et qui passe les arguments) et le fournisseur (la routine qui exécute la tâche demandée par le client) ont des obligations (par exemple, le client doit fournir une valeur positive, et le fournisseur renvoyer une valeur négative). Ce style de programmation est une suite logique du "*defensive programming*", dont le principe est d'inclure autant de vérifications que possible et de faire des fonctions les plus génériques possibles. Le principal problème de cette approche est qu'elle entraîne généralement une forte redondance de code. La programmation par contrat s'appuie sur les assertions, c'est-à-dire des morceaux de code servant à décrire les cas attendus (par exemple argument1 strictement positif) et non les exceptions.

OCL (*Object Constraint Language*) est inspiré des concepts de la programmation par contrats. L'OCL est un langage formel permettant, sur un graphe UML, de spécifier des contraintes (au sens de restrictions entre une ou plusieurs valeurs d'un élément du modèle). Sa description complète peut être trouvée dans [OMG 2010]. L'OCL fait maintenant partie du standard UML.

On peut distinguer trois types de contraintes :

- **invariant** de classe (contrainte devant être satisfaite par toutes les instances de la classe) : par exemple, une date de fauche doit être strictement comprise entre 0 et 366. Cette condition se traduit en OCL par « context Fauche inv : $0 < \text{date} < 366$ » ;
- **pré-condition** d'une opération (contrainte devant être valide avant l'exécution d'une opération) : par exemple, au moment de la fauche il faut qu'il y ait au minimum 1 tMS/ha ; cette condition se traduit par « context Typename::Fauche(Biomasse : float, ...): float pre parameterOk: Biomasse > 1 » ;
- **post-condition** d'une opération (contrainte devant être valide après l'exécution d'une opération) : par exemple, au moment de la fauche il faut qu'il reste au minimum 1 tMS/ha ; cette condition se traduit par « context Typename::Fauche(Biomasse : float, ...): float post resultOk : Biomasse > 1 ».

Ce formalisme peut également être ajouté aux graphiques UML, comme dans l'exemple de l'invariant de classe de la Figure 24.

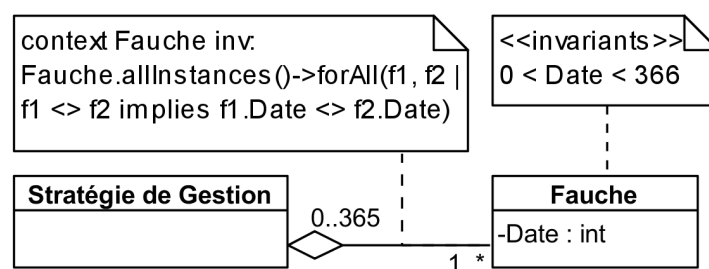


Figure 24. Exemple d'utilisation des contraintes OCL sur le métamodèle des stratégies de gestion (extrait du métamodèle des entrées du modèle ModVege)

Ce formalisme peut être utilisé dans EMF⁶¹, permettant ainsi facilement de vérifier les contraintes sur les métamodèles, à travers la phase de validation naturellement présente dans l'outil.

2.6.Conclusion

Nous avons vu dans ce chapitre quatre domaines-clés sur lesquels la thèse s'appuie. Tout d'abord, nous nous sommes intéressés à la question de la définition de la vulnérabilité et de ses concepts et indices associés. En effet, la littérature sur le changement climatique contient de nombreuses explications de la vulnérabilité, basée sur la notion de sensibilité et faisant éventuellement appel à des idées plus complexes, depuis la prise en compte de l'exposition jusqu'aux impacts résiduels du changement climatique après mise en place de mesures d'adaptation. Et même s'il existe un certain nombre de méthodes pour quantifier la vulnérabilité, réaliser une analyse de vulnérabilité au changement climatique reste un problème difficile, compte-tenu du manque de consensus sur la signification exacte du terme, mais surtout parce que la vulnérabilité n'est pas un phénomène directement observable. Par nature, et compte-tenu de la cascade d'incertitudes (scénarios climatiques, modèles climatiques, méthodes de régionalisation, modèles d'impacts, ...), une analyse de vulnérabilité au changement climatique va comporter beaucoup de simulations.

De façon à réduire notre nombre de simulations et à augmenter leur efficacité, nous nous sommes ensuite intéressés au domaine de la conception et de l'analyse des plans d'expériences. En effet, il existe tout un gradient de complexité entre les DOE les plus simples tels que les plans factoriels, jusqu'aux plans combinés en passant par les très fréquents LHD. Certains sont disponibles directement (tables construites), tandis que pour d'autres, des algorithmes pour les générer sont implémentés dans de nombreux logiciels (par exemple Matlab ou R).

Si l'utilisation des plans d'expériences permet de réduire le nombre de simulations requises, il est cependant nécessaire de faire appel au calcul intensif de façon à réduire le temps d'attente utilisateur. Pour cela nous pouvons faire appel à une certaine diversité de ressources de calculs (clusters, grille, ...), aidés par un certain nombre d'outils de distribution. À notre connaissance aucun outil ne permet de faciliter la majeure partie des étapes de l'analyse de vulnérabilité. C'est pourquoi, afin de faciliter la réalisation d'analyses de vulnérabilité, il a été décidé de développer une méthode d'analyse de vulnérabilité et de créer un outil permettant sa mise en œuvre. La méthode et l'outil se veulent les plus génériques possible (par rapport aux modèles utilisés et aux hypothèses sous-jacentes à l'analyse de vulnérabilité). Pour son ergonomie et sa facilité de couplage, le logiciel OpenMOLE a été choisi pour nous aider dans la réalisation de cet outil.

⁶¹ <http://wiki.eclipse.org/MDT/OCLinEcore>

L'ingénierie dirigée par les modèles est une branche du génie logiciel s'intéressant au développement, à la maintenance et à l'évolution des modèles au sein des systèmes informatiques. Elle est basée sur les trois concepts de "modèle", "métamodèle" et de "transformation de modèle". Son utilisation permet de répondre au besoin de genericité recherchée dans l'outil d'analyse de vulnérabilité. Pour faciliter la conception et la réalisation de l'outil, nous nous sommes appuyés sur Eclipse Modeling framework (EMF), une plateforme de modélisation et de génération de code intégrant les principes de l'IDM.

Chapitre 3 : Modélisation des plans d'expériences par une approche d'ingénierie des modèles

3.1. Introduction

Dans ce chapitre nous proposons des éléments de solution dans le contexte de la problématique d'analyse de vulnérabilité que nous avons exposée dans les chapitres précédents. Tout d'abord, en se basant sur l'étude de la littérature que nous avons effectuée, nous présentons une méthode originale d'analyse de vulnérabilité. Puis, en appui à cette approche, nous emploierons une approche d'ingénierie guidée par les modèles pour exposer les différentes étapes de la conception d'une solution pour l'analyse de vulnérabilité au changement climatique sous contrainte de plans d'expériences. Enfin, nous aborderons l'implémentation de notre solution avec le cadriciel *Eclipse Modeling Framework* (EMF) présenté au chapitre 2.

La première partie de ce chapitre est une proposition d'approche d'analyse de vulnérabilité en deux étapes. Après avoir brièvement rappelé l'état de l'art, et explicité les origines de cette approche, nous détaillerons la méthode. Cette méthode a été initialement proposée lors de la conférence European Simulation and Modelling Conference 2011 (ESM 2011) [Lardy et al., 2011a].

La deuxième partie de ce chapitre a été publiée et présentée lors de la conférence SummerSim 2011 [Lardy et al., 2011c]. Il s'agit d'un travail de rétro-ingénierie sur des plans d'expériences utilisés lors de campagnes de simulations précédentes. Ce travail, pouvant être considéré comme préliminaire, avait pour but de permettre la proposition de modèles pour la conception de plans d'expériences dits « DOE » (*Design Of Experiments*). Nous avons également conçu le métamodèle qui est associé à ces modèles.

La troisième partie de ce chapitre présente un jeu de plans d'expériences pour l'analyse de vulnérabilité au changement climatique. Ces DOEs ont à la fois été conçus dans le but de tenir compte de différents niveaux d'incertitudes, mais également de proposer un métamodèle de modèle agro-écologique et des plans d'expériences associés lors de l'analyse de vulnérabilité. Ce travail a été présenté lors de la conférence European Simulation and Modelling Conference 2011 (ESM 2011) [Lardy et al., 2011a].

La quatrième partie présente le métamodèle complet d'analyse de vulnérabilité au changement climatique et son intégration dans EMF. L'outil ainsi généré et son utilisation seront également décrits.

3.2. Proposition d'une approche en deux étapes pour l'analyse de vulnérabilité

Comme nous l'avons vu dans le chapitre précédent, la majorité des analyses de vulnérabilité ne tiennent pas compte de la mise en place de stratégies d'adaptation ou interprètent les impacts indépendamment des actions d'adaptation mises en œuvre. Or, nous souhaitons développer une approche qui permette d'analyser la vulnérabilité et d'interpréter les résultats qui en découlent avec et sans prise en compte de l'adaptation. Dans ce premier cas, l'adaptation influence la vulnérabilité. De plus, nous voulons une analyse quantitative (et non uniquement qualitative) de la vulnérabilité. Nous recherchons donc une approche qui soit générique par rapport aux modèles agro-écologiques utilisés, aux indices quantitatifs de vulnérabilité employés et aux variables d'intérêt (par exemple : rendements de la prairie ou émissions de gaz à effets de serre de l'écosystème prairial), mais qui puisse également tenir compte des effets de différentes hypothèses sur lesquelles se fonde l'analyse (par exemple, l'ensemble des sources d'incertitudes). Comme il paraît difficile d'évaluer la vulnérabilité avec et sans adaptation en même temps, nous avons donc proposé une approche en deux étapes. De plus, comme l'analyse de vulnérabilité fait appel à un nombre important d'évaluations, l'utilisation de surfaces de réponse nous permet de rendre réalisable (en terme de temps de calcul) certaines phases (comme la recherche d'un minimum de vulnérabilité) de l'analyse de vulnérabilité. Une surface de réponse est un modèle, ou une approximation de la fonction implicite qui caractérise la relation entre variables d'entrée et variables de sortie, et qui s'exprime en des termes plus simples que dans la simulation complète (le plus généralement à travers une équation polynomiale). On peut résumer les deux étapes de notre approche de la manière suivante (Figure 25) [Lardy et al., 2011a] :

1. premièrement, une étape d'analyse de sensibilité, dont le but sera de permettre d'estimer la vulnérabilité sans adaptation et de calculer une surface de réponse ;
2. dans un second temps, on cherche à minimiser la vulnérabilité sous contrainte de la capacité d'adaptation réelle.

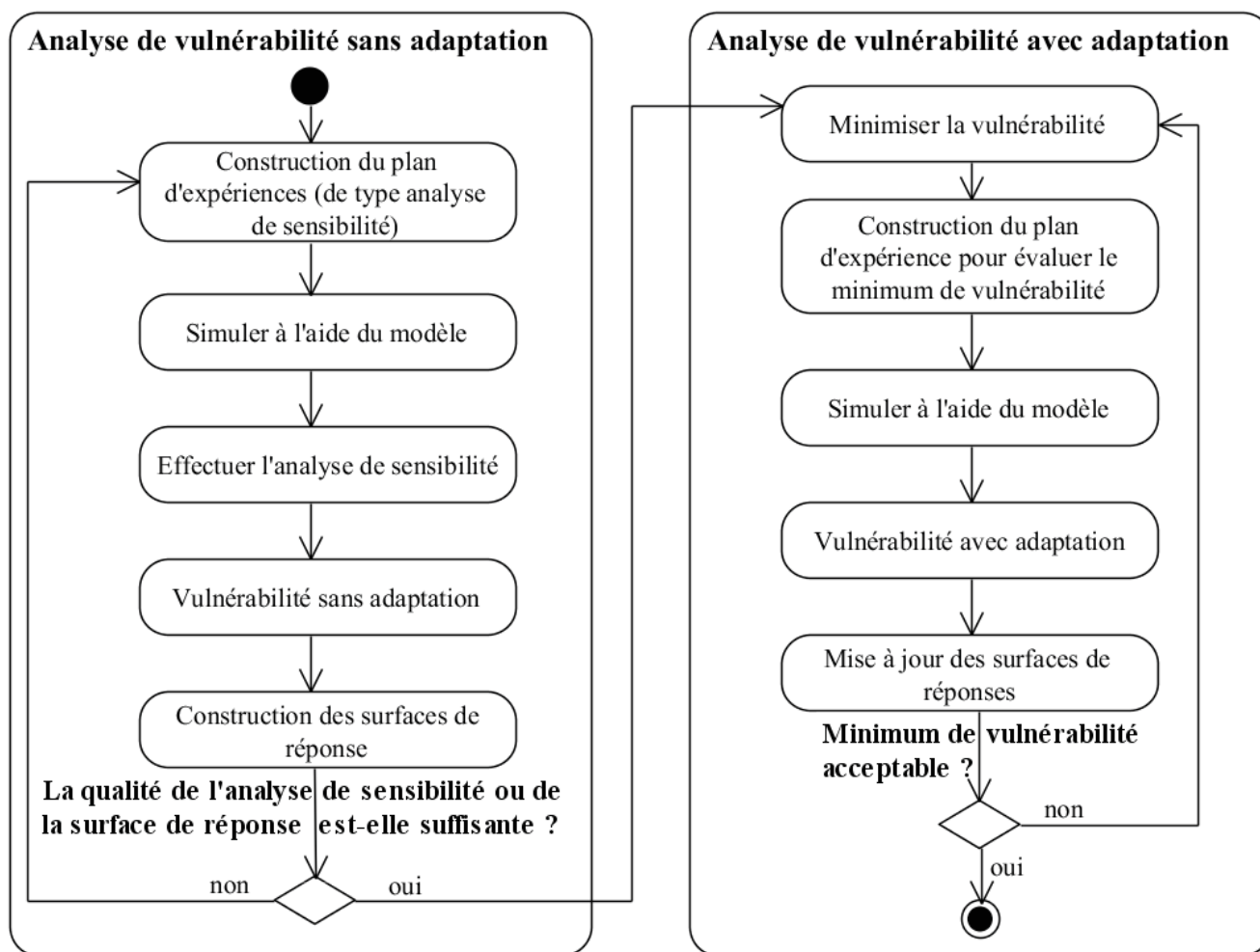


Figure 25. Proposition d'approche en deux étapes pour l'analyse de vulnérabilité [Lardy et al., 2011a]

Au sein de la première phase, la démarche consistera donc à créer un plan d'expériences approprié à l'analyse de sensibilité, puis à réaliser les simulations à l'aide du modèle agro-écologique choisi. L'analyse de ces résultats nous donnera une estimation de la vulnérabilité sans prise en compte de l'adaptation, et servira également à calculer des surfaces de réponse (en fait une surface par variable de sortie considérée). Si l'on estime que la qualité de l'analyse de sensibilité (par exemple, robustesse de l'analyse) et de la surface de réponse est suffisante (par exemple, une "bonne" erreur relative moyenne), alors on peut passer à la phase suivante, sinon on recrée ou on complète un plan d'expériences pour l'analyse de sensibilité.

La deuxième phase commence par une étape de minimisation de la vulnérabilité à l'aide d'algorithmes d'optimisation itératifs (métaheuristiques) et des surfaces de réponse précédemment calculées, en tenant compte de la capacité d'adaptation réelle. Puis, on construit un plan d'expériences permettant d'évaluer le minimum de vulnérabilité et sa robustesse, la robustesse étant définie, ici, comme l'inverse de la sensibilité aux incertitudes (e.g., climat, gestion). Une étude de sensibilité locale au minimum de

vulnérabilité pourrait convenir. Une fois les simulations réalisées à l'aide d'un modèle agro-écologique, on évalue ces résultats et on met à jour les surfaces de réponse. Si l'on considère le minimum de vulnérabilité trouvé comme acceptable (par exemple, si le minimum est suffisamment robuste), alors on s'arrête et on a donc une estimation de la vulnérabilité tenant compte de la capacité d'adaptation réelle. Sinon, on recommence la boucle en commençant par chercher de nouveau le minimum de vulnérabilité.

L'approche proposée est générique pour tout type de modèle d'impact. Dans le but de faciliter la mise en œuvre de cette méthode, nous avons décidé de créer un outil générique pour la conception et la réalisation d'analyses de vulnérabilité au changement climatique. En complément de cette approche, nous proposons également l'utilisation d'un ensemble complémentaire d'indices quantitatifs. En effet, même pondéré par la probabilité d'exposition, un indice unique ne peut pas donner une image complète de la vulnérabilité du système. Le chapitre suivant illustre la pertinence de l'utilisation de ces indices (voir 4.2, 4.3 et Table 9).

3.3. Rétro-ingénierie sur des plans d'expériences

Une première étape dans la conception d'un outil générique pour la conception et la réalisation d'analyse de vulnérabilité a été l'analyse de jeux de simulations déjà réalisés. Une approche de rétro-ingénierie basée sur l'IDM a été appliquée dans le but de proposer des modèles UML des plans d'expériences et d'identifier un schéma commun aux simulations réalisées (au sens des patrons de conception [Gamma et al., 1995]). Ce schéma nous donnera un métamodèle auquel nos modèles de plans d'expériences seront conformes.

3.3.1. Métamodèle des entrées de PaSim

PaSim est un logiciel de simulation implémenté en Fortran 95 (environ 60 000 lignes de code). Le code est divisé en modules, chacun traitant de la modélisation d'une partie spécifique de l'écosystème prairial:

- module microclimat pour l'interception de la lumière et les bilans d'énergie ;
- module physique du sol pour les profils d'humidité et de température du sol ;
- module biologie des sols pour les dynamiques des teneurs en matière organique, nitrate, ammonium et oxyde nitreux ;
- module animal pour l'ingestion, les performances animales et les émissions de méthane au pâturage.

Tous ces modules contiennent de nombreuses entrées, certaines d'entre elles étant des variables, tandis que d'autres sont généralement considérées comme constantes (paramètres) dans le domaine de validité du modèle. Le diagramme UML ci-dessous (Figure 26) présente l'organisation retenue pour gérer les entrées PaSim. Ces paramètres peuvent être classés en trois catégories fonctionnelles ayant trait au **Site**, au **Sol** et à la **Végétation**. En outre, le modèle utilise des variables d'entrée météorologiques au pas de temps horaire : la température de l'air, la vitesse du vent, le rayonnement solaire global, les précipitations, la pression de vapeur d'eau, les concentrations atmosphériques en CO₂ et ammoniac. Toutes ces entrées sont dans la classe **Climat**. La gestion de la parcelle est contenue dans la classe **Politique de gestion**, qui a deux sous-classes de **gestion** : la **Gestion proposée par le modèle** et la **Gestion prescrite**. Le modèle peut exécuter ces deux sous-classes de gestion indépendamment ou simultanément. La **Gestion proposée par le modèle** donne un ensemble de règles de gestion automatique qui affectent le comportement du modèle. La **Gestion** précise les dates de fertilisation azotée, la nature et la quantité de l'engrais apporté (classe **Fertilisation**), les dates de fauche de l'herbe (classe **Fauche**), et les informations relatives au pâturage (classe **Pâturage**), incluant les dates de mise à l'herbe et de sortie de parcelle des animaux, les chargements instantanés (*i.e.* le nombre d'animaux par unité de surface pendant la période de pâturage), la note d'état, le poids vif initial des animaux et la complémentation alimentaire au pâturage si nécessaire.

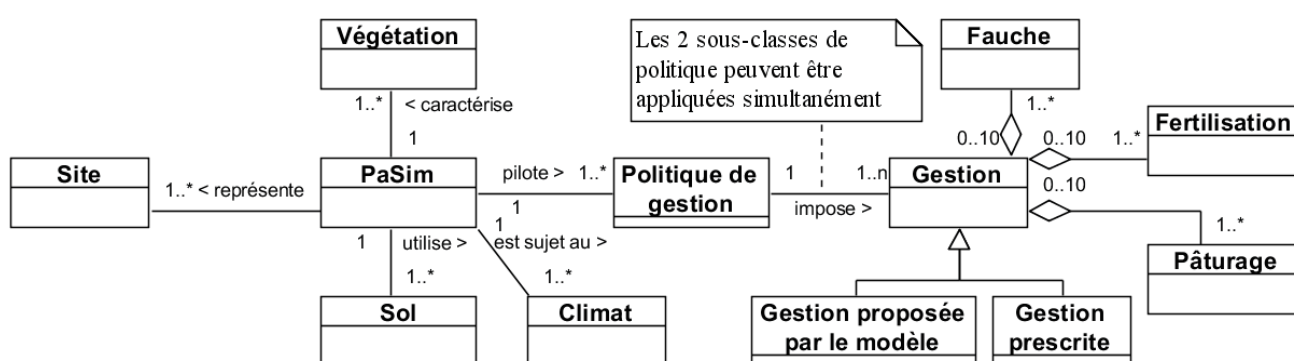


Figure 26. [Lardy et al., 2011c] Métamodèle UML des entrées du modèle PaSim.

Construire un métamodèle des entrées du modèle PaSim était une étape préliminaire à la recherche d'un métamodèle des plans d'expériences dans le cadre d'une analyse de vulnérabilité.

3.3.2. Utilisation de plans d'expériences

Pour réaliser toutes les simulations nécessaires, dans le cadre des études récemment menées avec PaSim, nous avons dû faire appel au Calcul à Hautes Performances. Ces études visaient à évaluer les impacts du changement climatique sur les prairies, à effectuer une analyse de sensibilité (par rapport à différentes sorties) de la plupart des paramètres de PaSim, et à optimiser la gestion des prairies sous changement climatique. Les plans

d'expériences (DOE, "*Design Of Experiments*") associés à ces trois études sont décrits ci-après.

Projections d'impact du changement climatique

Un travail important, basé sur un grand nombre de simulations, a été réalisé à la fois pour évaluer les impacts du changement climatique sur les prairies et pour caractériser les différents niveaux d'incertitude sur ces impacts. Pour cela, un plan factoriel incomplet a été utilisé. Des projections d'impacts climatiques ont été réalisées pour 12 sites français. Pour chacun de ces sites, PaSim a été forcé avec un éventail de 12 conditions climatiques futures possibles, combinant une gamme de scénarios d'émission de gaz à effet de serre reflétant différents scénarios socio-économiques du SRES (*Special Report on Emission Scenarios*) [Nakićenović 2000], de modèles climatiques globaux et de méthodes de régionalisation. Il n'était pas possible de générer toutes les combinaisons de scénarios, de modèles climatiques et de méthodes de régionalisation, par conséquent un plan factoriel complet n'était pas réalisable. Toutefois, si l'on considère ces trois éléments : SRES, modèles climatiques globaux et méthodes de régionalisation, comme une entité unique (*i.e.* le climat), alors le DOE est de type factoriel complet (*i.e.* la combinaison complète de 12 sites par 12 conditions climatiques futures possibles).

Cinq profils de sols ont également été choisis depuis une base de données incluant les principaux types de sols (*i.e.* texture, caractéristiques hydrologiques et profondeur du sol) et l'utilisation des terres en France (base DONESOL⁶²). On peut considérer que chaque type de sol peut également être divisé en deux classes, selon que la nappe phréatique soit à proximité ou non, considérant donc, ou non, le fait qu'il puisse y avoir des remontées capillaires depuis la couche limite du sol.

Parmi quatre politiques de gestion de parcelle de prairie et quatre types de végétation, six couples de « gestion x type de végétation » ont été retenus. Le DOE est par conséquent le produit complet de :

- un site ;
- un sol ;
- une gestion et le type de végétation associée ;
- un climat (à noter que les climats disponibles ne couvrent pas l'ensemble du produit scénario, modèle climatique global et méthode de régionalisation).

Toutes ces simulations prendraient environ un an de temps de calcul sur un seul processeur moderne (2012). Puisque toutes ces simulations sont indépendantes, elles furent distribuées sur des fermes de calcul (*computing clusters*). Cela a été fait en divisant la quantité de simulations par le nombre de processeurs disponibles. Un script supplémentaire

⁶² <http://www.gissol.fr/outil/donesol/donesol.php>

a été créé, qui liste toutes les simulations incomplètes (e.g., non encore réalisées ou ayant été arrêtées en cours de simulation), puis génère les scripts nécessaires pour lancer les simulations restantes jusqu'à ce que toutes les simulations aient été lancées. L'interprétation des résultats obtenus dans cette campagne de simulation peut être trouvée dans [Graux 2011] et [Graux et al., 2012a, b].

Analyse de sensibilité

Afin de réduire le nombre de paramètres de PaSim à estimer, une analyse de sensibilité a été effectuée. Une implémentation de la méthode de "*screening*" (criblage) de Morris [Morris 1991], basée sur [Campolongo et al., 2007], a été utilisée pour évaluer la sensibilité du modèle à 133 paramètres d'entrée sur trois sites français. Pour chaque site, trois années de données climatiques ont été utilisées, chacune d'entre elles représentant un niveau de conditions d'aridité (année d'aridité minimale, médiane et maximale sur la période 1950-1999), telles qu'elles sont définies par l'indice de De Martonne-Gottmann [De Martonne 1942]. Deux politiques de gestion ont été définies sur un sol générique. Les paramètres utilisés dans l'analyse de sensibilité appartiennent aux classes de **Végétation**, **Site**, **Sol** et **Gestion prescrite** (Figure 26). L'idée principale de la méthode de Morris [Morris 1991] est de déterminer pour chaque paramètre si son effet peut être considéré comme négligeable, linéaire et additif, non linéaire ou en interaction avec d'autres facteurs.

La méthode de Morris utilise un plan expérimental de type OAT ("*One factor At a Time*"), c'est-à-dire que chaque facteur est bougé aléatoirement, à son tour, par rapport à la combinaison précédente. Grâce à cela, une trajectoire est décrite dans l'espace formé de tous les facteurs. La sensibilité de chaque paramètre est estimée à travers l'effet sur la variable d'intérêt de la variation élémentaire du paramètre, et cela à partir de plusieurs points de l'espace. Cette méthode a l'avantage de pallier aux inconvénients des approches locales où la sensibilité est généralement mesurée en analysant les variations de la sortie du modèle lorsqu'un seul paramètre d'entrée est modifié, les autres étant fixés à leurs moyennes. Peu de trajectoires étant nécessaires pour obtenir un résultat robuste, six trajectoires de 134 points (133 paramètres + 1 point initial) ont été utilisées dans notre cas. Pour chaque paramètre, six niveaux ont été définis, chaque niveau correspondant aux valeurs suivantes de la fonction de répartition : 0 ; 0,2 ; 0,4 ; 0,6 ; 0,8 ; 1 (Figure 27). Cette conception a été répétée en se basant sur deux hypothèses de distribution des paramètres du modèle, une distribution uniforme et une distribution gaussienne. Pour des raisons de calcul, la fonction gaussienne a été tronquée aux deux extrémités de la distribution.

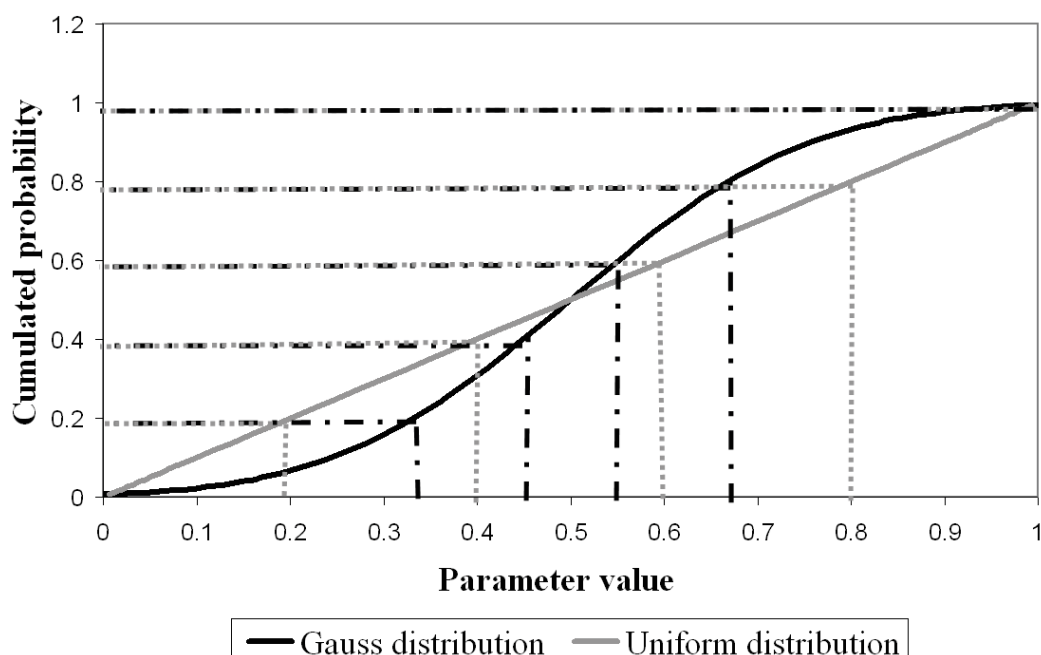


Figure 27. [Lardy et al., 2011c] Niveaux statistiques définis à partir de deux fonctions de distribution cumulée : Uniforme (gris) et Gaussienne (noire).

Le même design a été utilisé pour réaliser une analyse de la sensibilité de tous les types de sorties du modèle (émissions de gaz à effet de serre, croissance des plantes, teneurs en carbone ou en azote dans la matière sèche de la végétation...), de sorte que le même jeu de simulation était suffisant pour effectuer une analyse de sensibilité pour chacune des différentes sorties. Bien que le temps de calcul soit plus faible que celui requis pour les simulations des impacts du changement climatique, il était cependant nécessaire de distribuer les simulations sur les fermes de calcul locales. Pour résumer, le DOE était le produit des éléments suivants :

- un site ;
- un climat ;
- une gestion ;
- un jeu de valeurs pour les paramètres de sol, de site, de végétation et de gestion, pour une distribution uniforme ou gaussienne.

Simulations avec gestion automatique

Pour tester les adaptations possibles des systèmes d'élevage herbagers français à l'augmentation des aléas climatiques, une procédure en deux étapes a été élaborée par [Vuichard et al., 2007b], puis améliorée par [Graux 2011] et [Graux et al., 2013]. On s'intéresse ici à des systèmes exclusivement herbagers, dont les surfaces pâturées d'une part, et fauchées d'autre part, sont supposées homogènes en termes de végétations et de sols. Ces deux types de surfaces sont représentés très grossièrement par deux parcelles,

l'une fauchée, l'autre pâturée. Tout d'abord, le modèle détermine la gestion optimale de la fertilisation et de la fauche pour une prairie uniquement fauchée. Puis, en fonction des ressources fourragères potentielles, le modèle optimise itérativement le chargement animal à l'échelle du système fourrager, afin d'atteindre l'équilibre entre les ressources fourragères potentielles et les besoins pour alimenter les animaux à l'étable, en tenant compte de la période couverte par le pâturage. Cette procédure suppose que toutes les ressources fourragères (pas de fourrage acheté) sont entièrement consommées par des animaux à l'étable. Chacune des deux étapes de cette procédure utilise la même année climatique. Pour réaliser l'optimisation de chaque année, les informations relatives à l'état du système à la fin de l'année précédente sont nécessaires.

La gestion automatique a été déterminée pour deux scénarios climatiques, deux sols et deux sites associés à six conditions spécifiques à chaque site (combinaisons de périodes de vèlage, de qualités de fourrage / concentré et de types d'animaux), ce qui fait donc uniquement six couples (site, conditions). Pour l'ensemble des 150 années de la série climatique, deux algorithmes ont été utilisés, dépendant du fait que l'on considère ou non la matière organique du sol à l'équilibre avec la gestion et le climat. Pour résumer, dans ce cas le DOE est la combinaison de :

- un site ;
- un sol ;
- un scénario climatique ;
- une gestion (dépendante du site) associée à une méthode d'optimisation.

À la différence des simulations de projections climatiques et de l'analyse de sensibilité, toutes les simulations ne peuvent pas être parallélisées. En effet, chaque simulation pour une année climatique donnée doit attendre les résultats de l'année précédente. Dans notre cas, on ne peut donc distribuer les calculs que par bloc de 150 années. L'interprétation des résultats peut être trouvée dans [Graux 2011].

3.3.3. Modèle et métamodèle des plans d'expériences

Le but de cette partie de notre travail est de proposer des abstractions pour les plans d'expériences. Une bonne manière de concevoir un métamodèle des plans expérimentaux pour l'analyse de vulnérabilité dans notre contexte est de commencer par construire des modèles de plans expérimentaux sur des campagnes de simulations déjà réalisées, même si ces dernières ne représentent qu'une étape de l'analyse de vulnérabilité. Commençons par le plan expérimental sur les projections d'impacts du changement climatique.

Comme décrit précédemment, ce DOE est presque un plan factoriel complet. Il n'est pas complet dans le fait que le climat, qui est une composition du scénario d'émission SRES, du modèle climatique global et de la méthode de régionalisation, n'est pas complet.

Cependant, si l'on considère le climat comme une entité unique, alors le plan est factoriel complet comme nous l'avons déjà dit. Le modèle résultant du plan expérimental est donné par la Figure 28.

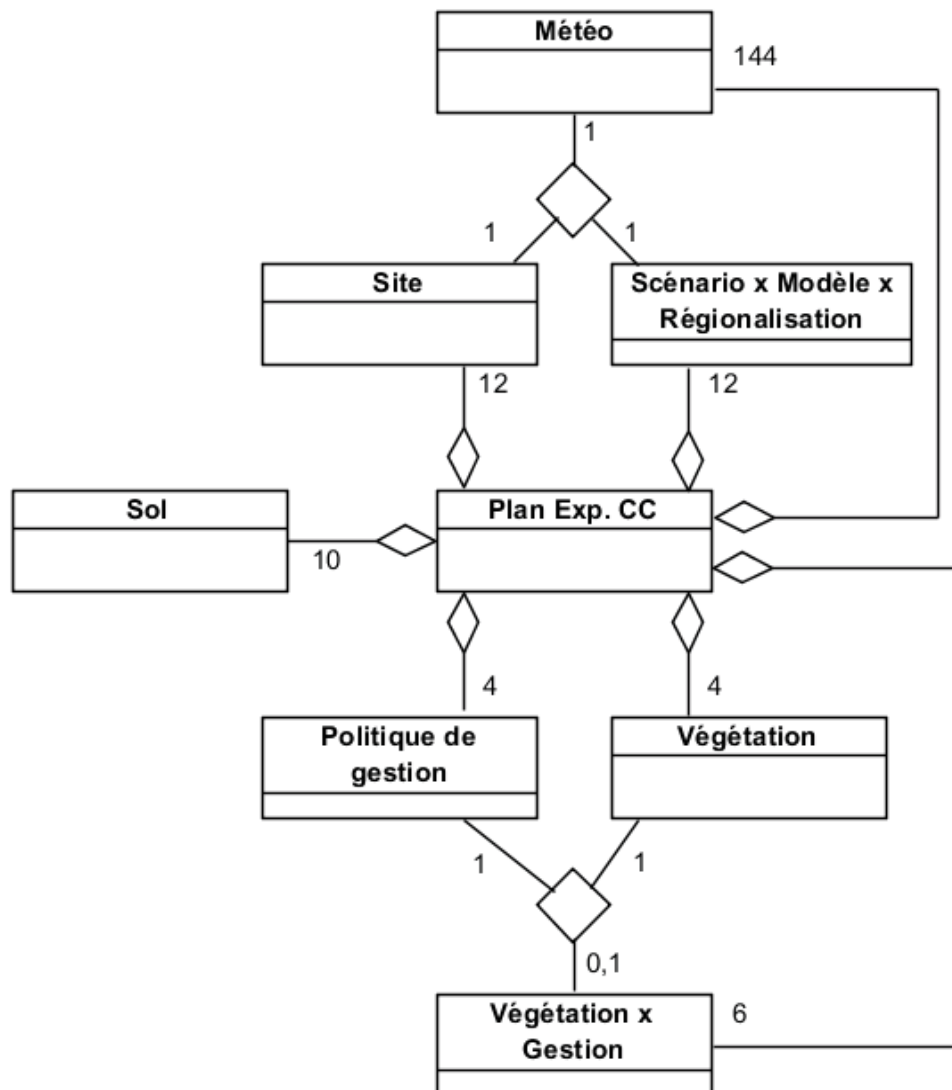


Figure 28. Modèle du plan expérimental pour les simulations des impacts du changement climatique.

Il est intéressant de noter que pour un **Site** et pour un **Climat**, il existe un unique ensemble de données **Météorologiques**. Cette unicité est importante dans la mesure où la cardinalité résultante est le produit de la cardinalité de **Site** et de **Climat**. De la même manière, la **politique de Gestion** et la **Végétation** peuvent être assemblée de manière cohérente dans une classe d'association "**Végétation et Gestion**". Cette classe d'association nous permet d'éviter des combinaisons incompatibles de politique de gestion et de végétation. Par exemple, une prairie permanente (depuis suffisamment longtemps) ne peut pas être mono-spécifique. Dans notre expérience, nous avons quatre politiques de gestion et quatre types de végétation et seulement six combinaisons valides. Le dernier

élément du plan expérimental est donné par la classe du **Sol**, qui n'a pas de relations directes avec les autres classes du modèle.

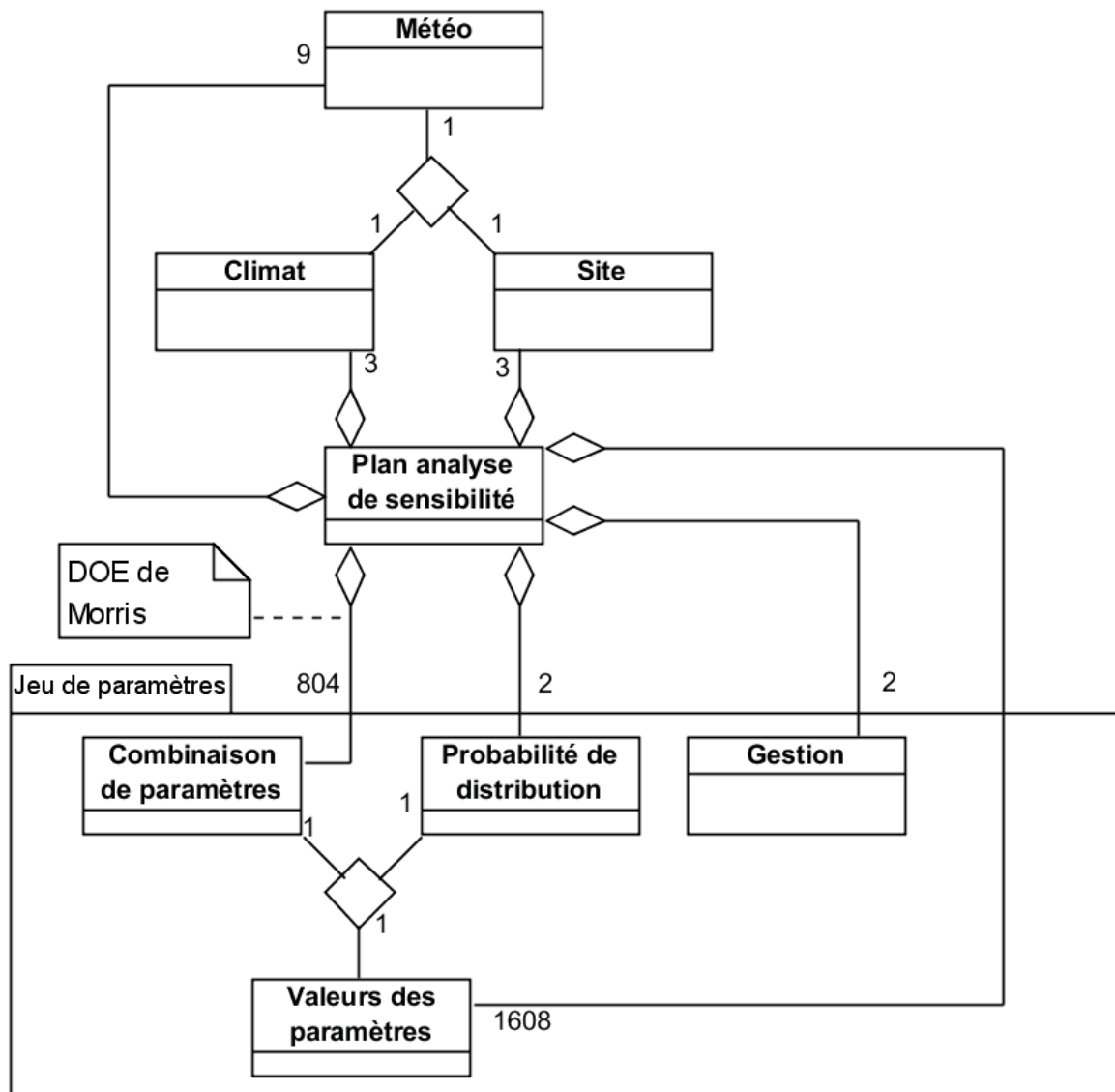


Figure 29. Modèle du plan expérimental pour l'analyse de sensibilité.

Si nous considérons maintenant le plan expérimental pour l'analyse de sensibilité, nous pouvons également proposer un modèle (Figure 29). Comme pour le plan expérimental des simulations de changement climatique, il y a une association spécifique du **Climat** et du **Site** donnant les données **Météorologiques**. La classe **Valeurs des paramètres** résulte de l'association entre les **Combinaisons de paramètres** et la **Probabilité de distribution**. La cardinalité de la classe **Combinaison de paramètres** est de 804 pour six trajectoires de 134 points (dans un design de Morris). Dans ce cas, la **Gestion** est plus simple et ne présente pas d'interaction particulière avec d'autres classes.

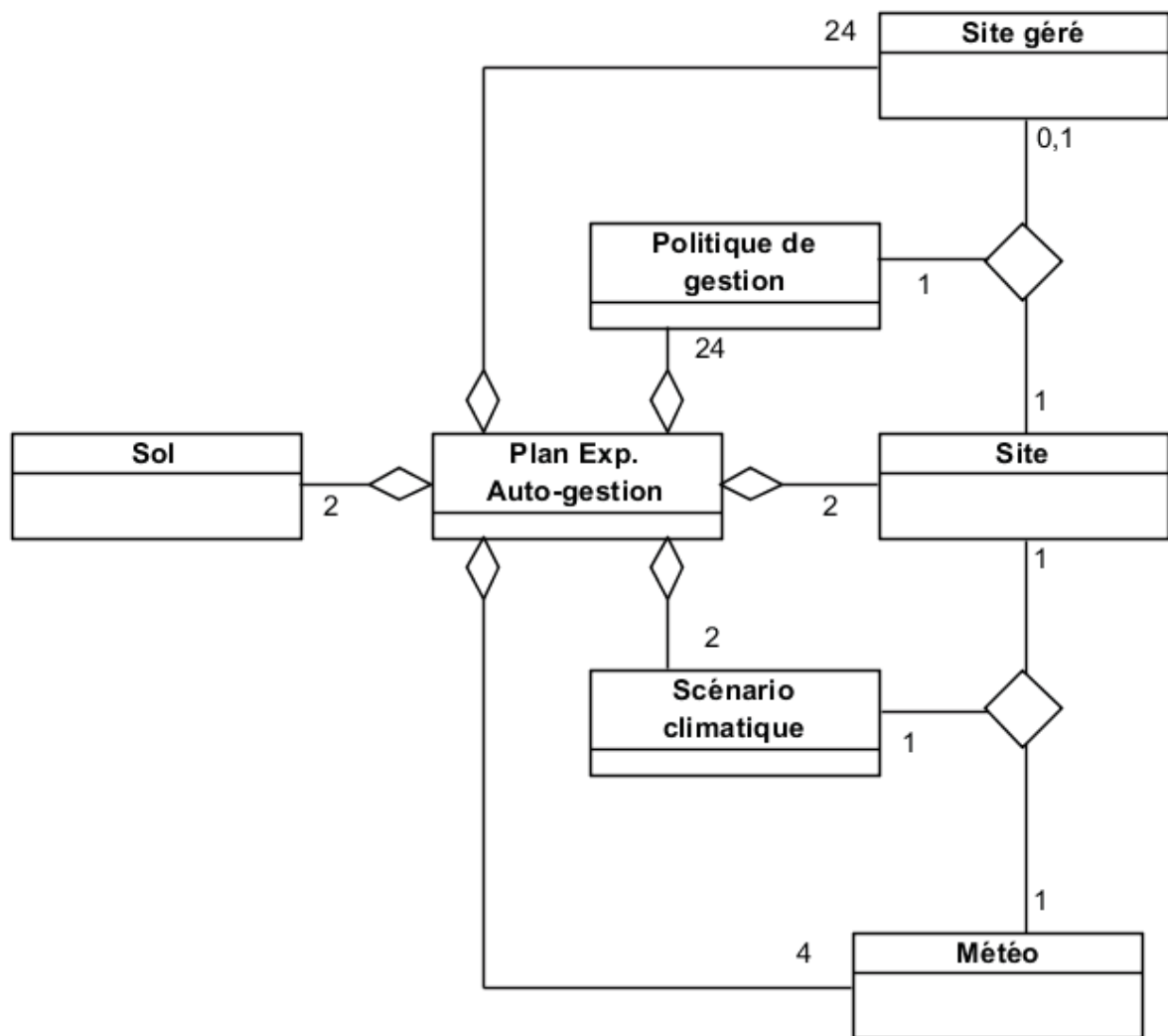


Figure 30. Modèle du plan expérimental pour les simulations de gestion automatique.

Dans le cas de la simulation de gestion automatique, le plan expérimental (Figure 30) contient deux classes d'association. La première, comme dans les deux modèles précédents, gère les données **Météorologiques**. La seconde est le croisement de l'information du **Site** avec sa **Politique de gestion**. Dans ce cas, nous pouvons noter que le plan factoriel n'est pas complet (c'est-à-dire que nous ne simulons pas toutes les politiques de gestion pour tous les sites). Le dernier élément du plan expérimental est la classe de **Sol**, qui n'a aucune interaction particulière avec les autres classes.

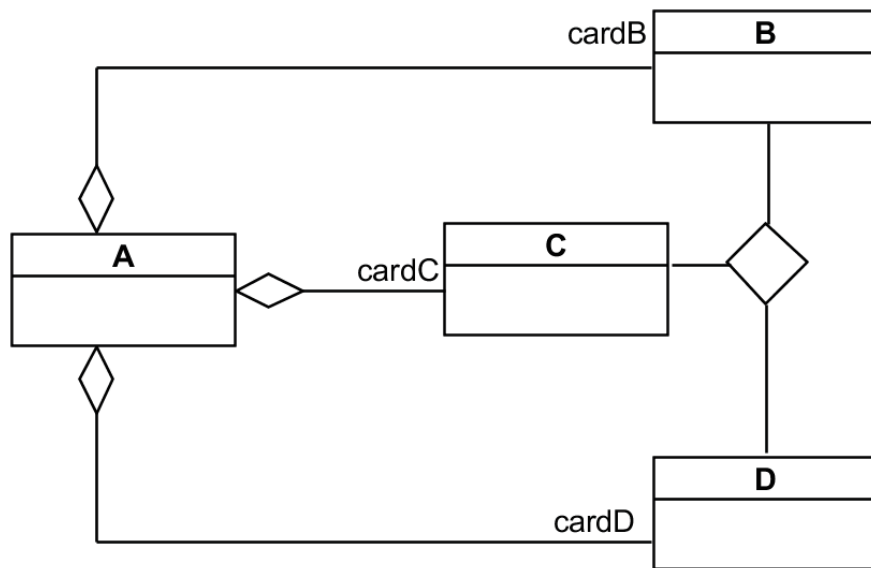


Figure 31. Patron ou métamodèle d'un plan expérimental (la classe **A**) avec deux entrées élémentaires (les classes **B** et **C**) et la combinaison résultante (la classe **D**, qui est une forme de classe « **BC** »).

À travers ces trois modèles de plans expérimentaux, nous pouvons identifier un patron commun [Gamma et al., 1995] (Figure 31). Le patron peut être vu comme un métamodèle auquel certaines parties des modèles sont conformes. La généralité est la suivante : notons deux entrées élémentaires de PaSim, de classe **B** et **C**, faisant partie du plan expérimental (classe **A**) ; dans le même temps, les classes **B** et **C** sont combinées pour former la classe **D**, qui peut aussi être considérée comme faisant partie du plan.

Si dans nos plans expérimentaux, nous considérons la partie formée par la classe **Site**, la classe **Scénario Climatique**, la classe **Météorologique** et la classe du **Plan expérimental**, nous pouvons voir l'instanciation de notre patron/métamodèle dans chacun de nos trois exemples. Ces trois sous-parties de modèle sont conformes au métamodèle [Bézivin 2005]. Par exemple dans le modèle des simulations de gestion automatiques (Figure 30), nous pouvons considérer les classes **Site** et **Scénario Climatique** comme les classes **B** et **C**, **Météo** comme la classe **D** et la classe **Plan expérimental** comme la classe **A**. Un autre exemple est le sous-modèle issu du modèle du plan expérimental pour les simulations des impacts du changement climatique (Figure 28), dans lequel les classes **Végétation** et **Politique de gestion** peuvent être vues comme les classes **B** et **C**, la classe "**Végétation x Gestion**" comme la classe **D**, la classe **A** étant représentée par la classe "**Plan Exp. CC**". Deux autres exemples peuvent être trouvés. L'un dans le modèle du plan sur l'analyse de sensibilité (**A** : **Plan d'analyse de sensibilité** ; **B** et **C** : **Combinaison de paramètres et Probabilité de distribution** ; **D** : **Valeur des paramètres**) et l'autre dans le modèle sur la gestion automatique (**A** : **Plan Exp. Auto-gestion** ; **B** et **C** : **Site et Politique de Gestion** ; **D** : **Site géré**).

La nature du DOE déterminera la valeur de la cardinalité (**cardD**) entre **A** et **D** : si le plan est factoriel complet alors **cardD** est le produit de **cardB** par **cardC**, comme, par exemple, dans le sous-modèle "Site – Scénario Climatique" du modèle de plan expérimental pour les simulations des impacts du changement climatique, Figure 28 ; si le design n'est pas complet, alors **cardD** sera une fonction de **cardB** et **cardC** (et sera inférieur à **cardB** × **cardC**), comme c'est le cas pour le sous-modèle "Végétation – Politique de gestion" du modèle présenté sur la Figure 28.

Ce métamodèle, qui est écrit pour un exemple à trois éléments, pourrait être généralisé. En effet, ce métamodèle peut être appliqué à n'importe lequel des trois exemples de DOE précédemment décrits. La conformité au métamodèle s'applique de manière récursive. Pour qu'un modèle soit conforme au métamodèle de la Figure 31, il faut que les classes **B** et **C** soient elles-mêmes conformes à ce même métamodèle. Ainsi de manière récursive, on peut montrer que tout modèle est bien conforme à ce métamodèle. On pourrait également considérer, cette conformité de manière itérative. En effet, si la classe d'association **D** (et ses deux classes associées **B** et **C**) est elle-même considérée comme une nouvelle classe élémentaire elle peut, en tout conformité avec le métamodèle, être associée à d'autres classes pour former des classes d'association de niveau supérieur. Il est ainsi également possible de reconstruire itérativement un modèle qui est complètement conforme à ce métamodèle.

L'analyse de vulnérabilité utilise souvent des analyses de sensibilité et des options d'adaptations (dont l'optimisation fait partie). Ainsi, en considérant l'exemple de ces trois campagnes de simulations, représentant chacune une composante de l'analyse de vulnérabilité, nous pouvons conclure qu'un outil dont le but sera de permettre la réalisation d'analyses de vulnérabilité nécessitera la réalisation d'un grand volume de simulations. Le nombre de simulations est un problème important dans les projections de changement climatique à cause de la multiplicité des scénarios et de la cascade d'incertitudes. Il est clair que, d'un côté, le choix d'un plan expérimental approprié est nécessaire pour réduire le nombre de simulations autant que possible, et que, d'un autre côté, les calculs distribués apparaissent comme une solution absolument indispensable. À titre de référence, les simulations présentées précédemment représentent environ 1,5 année de temps de calcul mono-CPU. Pour la distribution des calculs, différentes plateformes sont à considérer : clusters, grille ou cloud computing. La conception d'un logiciel qui permet la distribution des calculs d'un DOE de manière indépendante de la plateforme [Amblard et al., 2003], [Reuillon et al., 2008], [Reuillon et al., 2010] a déjà été réalisée. Toutefois, dans leur état actuel, ces outils ne tiennent pas compte de l'évaluation de la vulnérabilité. Pour les adapter, nous avons besoin d'un modèle des entrées de PaSim et des modèles des plans expérimentaux (ceux présentés précédemment étant relativement similaires à ceux requis pour une analyse de vulnérabilité) et du métamodèle des plans d'expériences.

Dans cette partie nous avons présenté notre expérience d'une approche de rétro-ingénierie de modèles. La rétro-ingénierie, comme définie par [Chikofsky et Croos, 1990], est une tâche préliminaire dans la conception d'un cadre logiciel à partir des expériences passées. Ce travail préliminaire a permis la proposition de modèles et d'un métamodèle des plans d'expériences et constitue la première étape vers la conception d'un outil générique d'évaluation de la vulnérabilité.

3.4. Plans d'expériences pour l'analyse de vulnérabilité

Cette partie présente une approche pour l'analyse de vulnérabilité sous changement climatique et sous contrainte de plans d'expériences en utilisant une approche d'ingénierie dirigée par les modèles. Afin de présenter notre approche, nous utilisons ModVégé [Jouven et al., 2006a], un modèle mécaniste pour la dynamique de la production, de la composition et de la digestibilité de l'herbe des prairies permanentes gérées. Ce modèle a été retenu pour étudier la faisabilité de notre approche parce que, d'une part, il est suffisamment complexe (environ 20 équations) pour reproduire les impacts de la variabilité climatique sur un pâturage [Jouven et al., 2006b] et, d'autre part, il nécessite relativement peu d'entrées tout en ayant un temps de calcul limité (quelques secondes). Notre approche sera ensuite appliquée aux gros modèles biogéochimiques classiques généralement utilisés pour ce genre d'études (et notamment le logiciel PaSim que nous avons déjà présenté, [Graux et al., 2012a]).

Après une brève description du modèle ModVégé, nous détaillerons différents plans d'expériences pour une analyse de vulnérabilité sous changement climatique en fonction du degré d'incertitude évalué. Nous concluons par la proposition d'un métamodèle des modèles d'impacts et des plans d'expériences associés pour l'analyse de vulnérabilité.

3.4.1. Utilisation du modèle exemple ModVégé

Une description plus agronomique/écologique du modèle peut être trouvée dans le chapitre suivant à la section 4.2.1.

Le modèle ModVégé, initialement développé en Python, a été récemment ré-implémenté en Java dans le but de l'intégrer dans une plate-forme de modélisation. Comme la plupart des décisions de gestion de la parcelle sont prises sur une base quotidienne, le modèle fonctionne sur un pas de temps journalier.

Pour résumer, les entrées du modèle peuvent être regroupées en quatre catégories (Figure 32):

- les données sur la **Végétation** : traits fonctionnels moyens de la communauté végétale (caractéristiques de fonctionnement de la plante) et de l'état initial du couvert (biomasse et âge des quatre compartiments) ;
- les propriétés du **Site**, à savoir l'indice de nutrition du sol et la capacité de rétention en eau du sol ;
- les données journalières d'**Environnement** : rayonnement photosynthétique actif, température de l'air, précipitations et évapotranspiration de référence ;
- la **Stratégie de gestion**, qui est composée d'un certain nombre d'événements de **Fauche**.

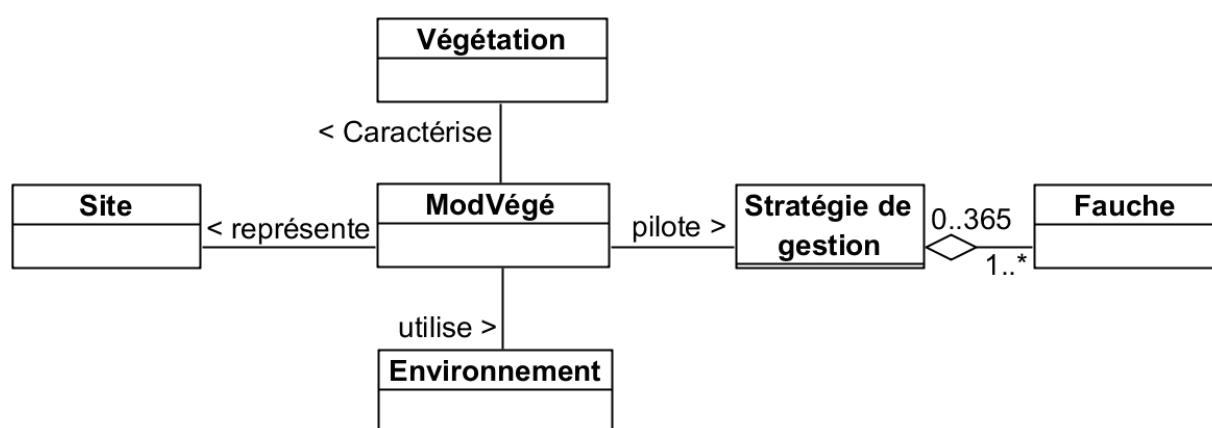


Figure 32. Métamodèle UML des entrées du modèle ModVégé.

Différentes approches peuvent être choisies en fonction de ce que l'on souhaite prendre en compte dans l'analyse de vulnérabilité. Par exemple, on pourrait supposer que tous les paramètres (comme les traits de la végétation) et les variables (comme les données climatiques) de forçage du modèle sont connues avec suffisamment de précision. Dans le cas contraire, nous pourrions tenir compte des incertitudes d'une ou de plusieurs composantes du forçage du modèle. En fait, on peut définir un ensemble d'approches en fonction des degrés de liberté relâchés. Toutes ces approches sont résumées dans la Table 8. Pour illustrer ces différentes approches, nous allons considérer un cas simple : un pâturage permanent avec un climat observé sur 30 années de données météorologiques et une gestion connue. Dans les différents plans expérimentaux présentés ci-après, chaque simulation est donc effectuée sur une série climatique constituée par 30 couples (climat, gestion) (sauf dans le cas de l'évaluation tenant compte des incertitudes sur la gestion). Ces couples sont prélevés parmi les instances de la classe **Année Gestion × Environnement** (voir, par exemple, Figure 34). Selon les plans d'expériences, ces prélèvements sont faits sans remise ou bien avec remise. Dans ce dernier cas, des instances identiques **Année Gestion × Environnement** peuvent être répétées dans la série (d'où une cardinalité de la classe **Année Gestion × Environnement** éventuellement inférieure à 30). Nous nous restreindrons aux incertitudes possibles sur le climat et la gestion. Une illustration des

résultats obtenus à l'aide de ces conceptions de plans d'expériences est présentée dans le chapitre suivant.

Table 8. Résumé des plans expérimentaux pour l'analyse de vulnérabilité sous changement climatique.

Évaluation	Description	Plan expérimental
"Simple"	Le coefficient de variation (CV) nous donne la sensibilité dans le calcul de la vulnérabilité. Pas de prise en compte des incertitudes.	1 site × 1 végétation × 1 Environnement × 1 Stratégie de gestion
"Incertitudes climatique"	Le CV nous donne la sensibilité dans le calcul de la vulnérabilité. Incertitudes climatiques : l'autocorrélation entre années consécutives est négligeable devant leur appartenance à la période considérée.	1 site × 1 végétation × n2 Environnements × Gestions (Bootstrapping sans remise sur le couple Année Gestion × Environnement)
"Incertitudes sur la fréquence des événements climatiques extrêmes"	La sensibilité, dans le calcul de la vulnérabilité avec prise en compte de l'exposition, est pondérée par les probabilités d'occurrence des événements extrêmes. Prise en compte des incertitudes sur la fréquence des événements extrêmes.	1 site × 1 végétation × n3 séries climatiques Environnements × Gestions,
"Incertitudes sur la gestion"	La sensibilité, dans le calcul de la vulnérabilité avec prise en compte de l'exposition, est pondérée par les probabilités d'occurrence des écarts par rapport à la gestion de référence. Prise en compte de l'incertitude sur la gestion.	1 site × 1 végétation × 1 Environnement × n4 Stratégies de gestion (composées de 30 années de gestion identique aux années de référence ou avec un "delta")
"Incertitudes sur la gestion, le climat et les événements extrêmes"	La sensibilité, dans le calcul de la vulnérabilité avec prise en compte de l'exposition, est pondérée par les probabilités d'occurrence des écarts par rapport à la gestion de référence et par les probabilités d'occurrence des événements extrêmes. Prise en compte des incertitudes sur la gestion, sur le climat et l'occurrence des événements extrêmes.	1 site × 1 végétation × n5 séries climatiques Environnements × Gestions, (composées de n5' Stratégies de gestion et n5" Environnements)

3.4.2. Evaluation "simple"

Le premier cas consiste simplement à simuler le système sur 30 ans. Le coefficient de variation (CV) sera utilisé comme mesure de sensibilité dans le calcul de la vulnérabilité. Toutes les années étant considérées comme équiprobables, l'exposition est donc uniforme, ainsi le CV nous permet de calculer la vulnérabilité en tenant compte de l'exposition à la variabilité climatique. Cette première approche ne tient compte d'aucune source

d'incertitudes. Le modèle du plan d'expériences est simple (Figure 33), puisque chaque facteur (c'est-à-dire les éléments du métamodèle des entrées de ModVégé) n'apparaît qu'une seule fois. Cette approche permet d'avoir une évaluation "simple" des impacts du changement climatique et de la vulnérabilité associée.

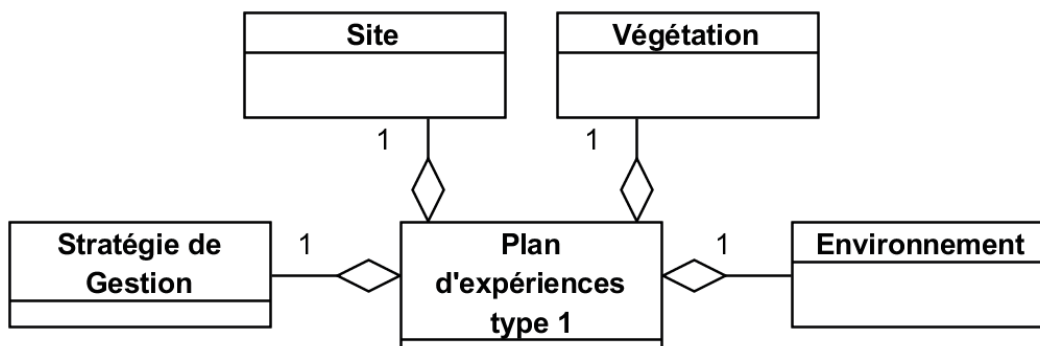


Figure 33. Modèle de plan expérimental pour l'évaluation "simple" de la vulnérabilité sous changement climatique.

3.4.3. Evaluation tenant compte des incertitudes climatiques

On considère maintenant que les années climatiques sont représentatives de périodes définies par des caractéristiques climatiques propres. On observe alors que la corrélation entre les années N et $N+1$ est négligeable devant la corrélation entre les années appartenant à une même période. Cela revient à considérer que l'ordre d'occurrence des années est aléatoire (on fait ce raisonnement sans tenir compte de l'augmentation de CO_2 d'une année sur l'autre). On peut alors, par bootstrapping (tirage aléatoire sans remise [Efron et Tibshirani, 1993]), générer des séries climatiques et les simuler. De même que précédemment, le coefficient de variation pourrait être utilisé pour évaluer la sensibilité pour le calcul de la vulnérabilité. Il y a deux manières de gérer notre plan d'expériences dans ce cas-là :

- soit on détermine *a priori* le nombre de séries climatiques (e.g., le nombre de tirages dans le cas du bootstrapping) à simuler de façon à parcourir suffisamment l'espace formé par les facteurs et leurs niveaux. Dans ce cas, on pourrait substituer au bootstrapping un LHD qui par nature parcourrait de manière plus uniforme l'espace ;
- soit on adapte notre plan d'expériences au fur et à mesure des simulations (dans le cas d'un bootstrapping, on se contentera d'en augmenter la taille) jusqu'à convergence de la vulnérabilité.

Le plan d'expériences résultant (Figure 34) est, dans ce cas-là, formé de facteurs fixes (ici, **Site** et **Végétation**) et d'un certain nombre de **Stratégies de Gestion** et d'**Environnements**, ces derniers étant respectivement composés de 30 années de gestion

(**Gestion Annuelle**) et de 30 années environnementales (**Environnement Annuel**) différentes (sans répétition). Cependant, les années de gestion et d'environnement sont contraintes par une condition d'unicité, ce qui fait que le couple nommé **Année Gestion x Environnement** n'a que 30 occurrences. L'intérêt de ce plan d'expériences réside dans le fait que la vulnérabilité calculée tient compte d'une partie des incertitudes climatiques.

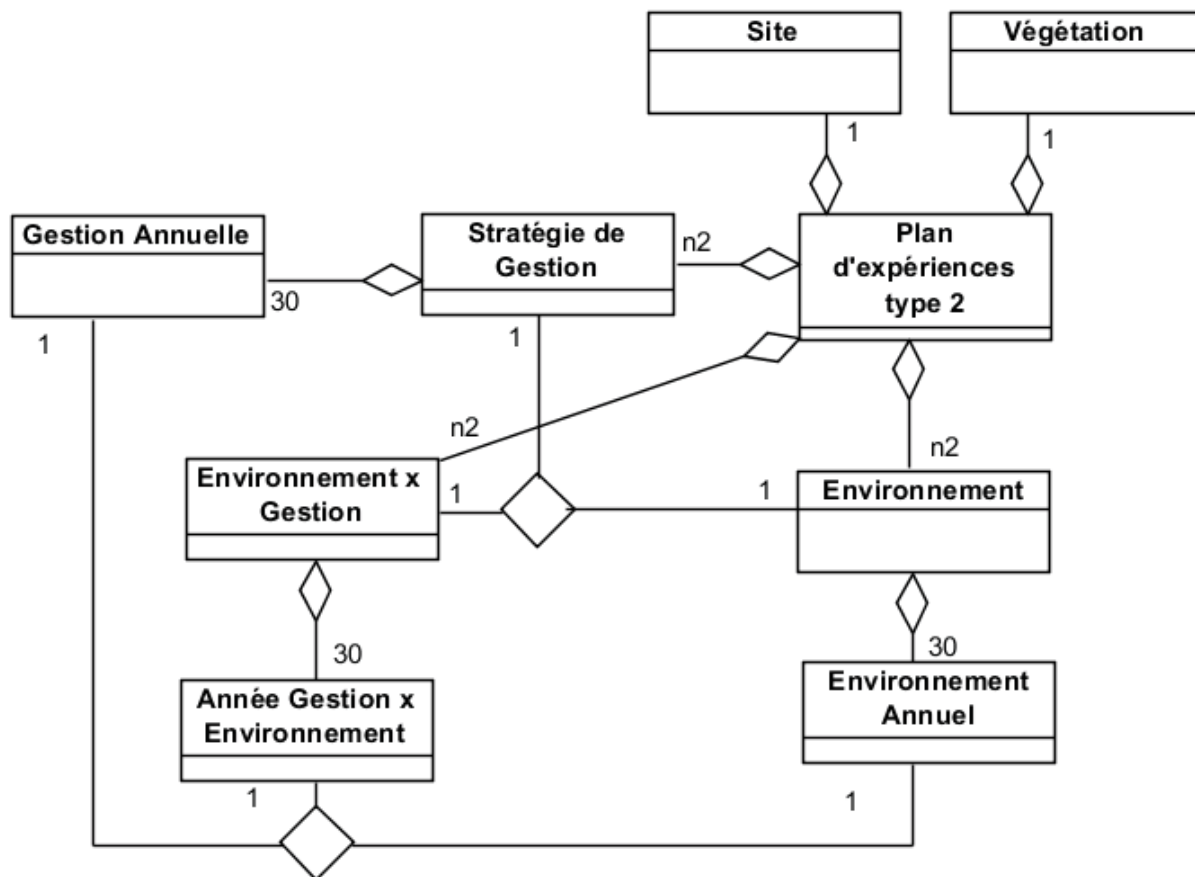


Figure 34. Modèle de DOE pour le calcul de la vulnérabilité sous changement climatique en tenant compte en partie des incertitudes climatiques.

3.4.4. Evaluation avec incertitudes sur la fréquence des évènements climatiques extrêmes

On suppose cette fois que dans la série climatique initiale, la fréquence des évènements extrêmes aurait pu être supérieure à celle observée. On simule donc des séries de 30 ans dans lesquelles on remplace une ou plusieurs années climatiques par une année avec extrêmes climatiques. Pour le calcul de la vulnérabilité, il faudra alors pondérer les séries climatiques par leur probabilité d'occurrence. Par exemple, supposons que la série climatique observée ne contient qu'une année extrême. En supposant que la probabilité d'occurrence des évènements extrêmes suit une loi de Poisson d'espérance 1, on pourrait alors considérer que la probabilité de la série observée est de 37%, que la probabilité

d'avoir un événement extrême de plus (*i.e.* deux événements extrêmes sur 30 ans) est de 18%, et de 6% et 2% pour les séries contenant deux ou trois événements extrêmes de plus (c'est-à-dire trois ou quatre années extrêmes sur une série de 30 ans). La série ne contenant aucun événement extrême aurait alors une probabilité de 37%.

Le plan d'expériences résultant (Figure 35) se compose toujours des deux éléments fixes : **Site** et **Végétation**. La principale différence avec le plan d'expériences précédent (Figure 34) se situe principalement dans les cardinalités. En effet, les **Stratégies de Gestion** et les **Environnements** sont toujours composées de 30 années, cependant, certaines années (les années extrêmes) peuvent y être répétées. Ce plan d'expériences force à distinguer les années climatiques dites "normales" des années dites "extrêmes". Ces dernières pouvant se substituer aux premières. L'intérêt de ce plan d'expériences réside dans le fait que la vulnérabilité calculée tient compte des incertitudes sur la fréquence des événements climatiques extrêmes.

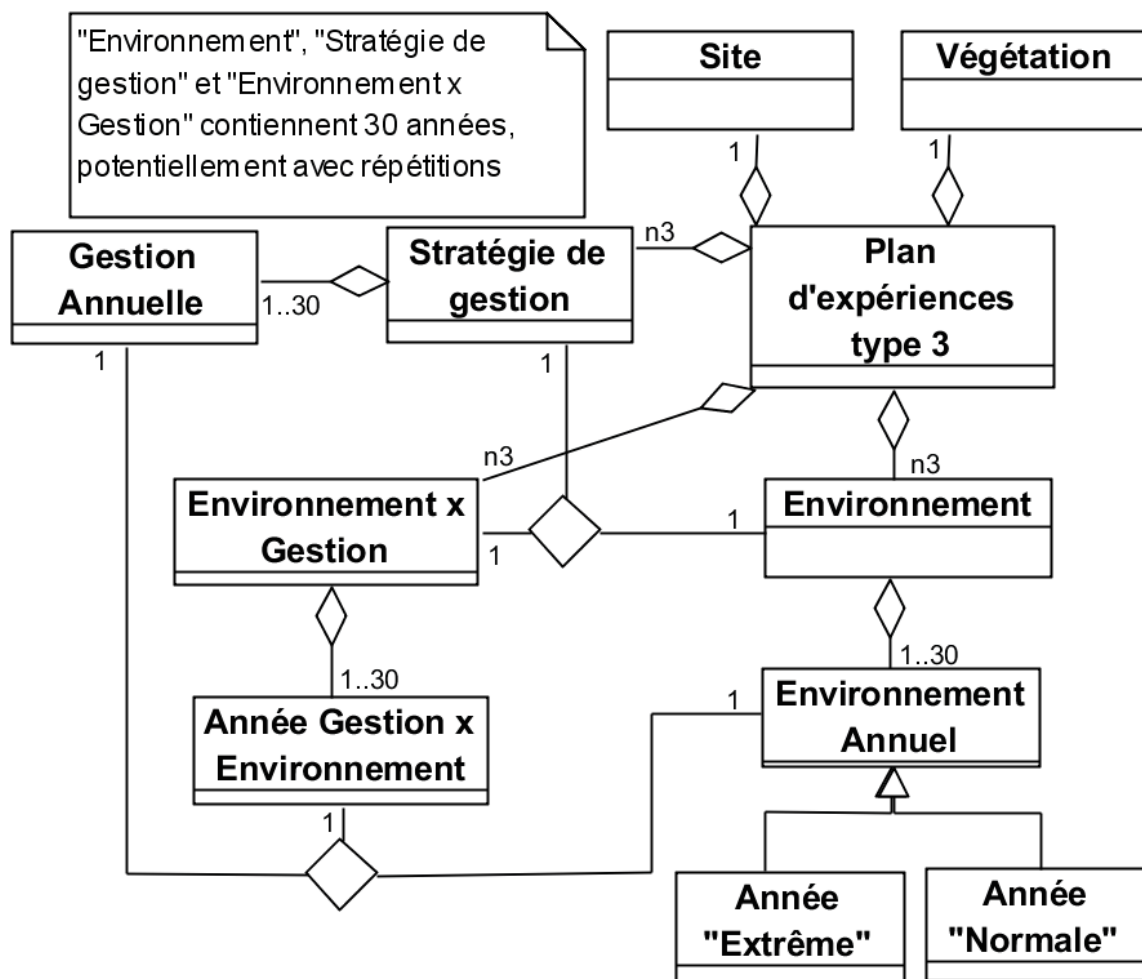


Figure 35. Modèle de DOE pour le calcul de la vulnérabilité sous changement climatique en tenant compte de l'incertitude sur la fréquence des événements climatiques extrêmes.

3.4.5. Evaluation tenant compte des incertitudes sur la gestion

Une autre possibilité serait de considérer que la gestion aurait pu être légèrement différente (par exemple, une date de fauche légèrement décalée). On utilisera donc, dans ce cas, un plan d'expériences testant un écart des dates par rapport à leur positionnement initial. Ainsi, la date de fauche de la première année peut être décalée une semaine plus tôt, puis celle de la deuxième année trois jours plus tard, ainsi de suite... On pourra alors calculer la vulnérabilité, en tenant compte de l'exposition, en estimant que l'écart du nombre de jours suit une loi de distribution connue (gaussienne par exemple).

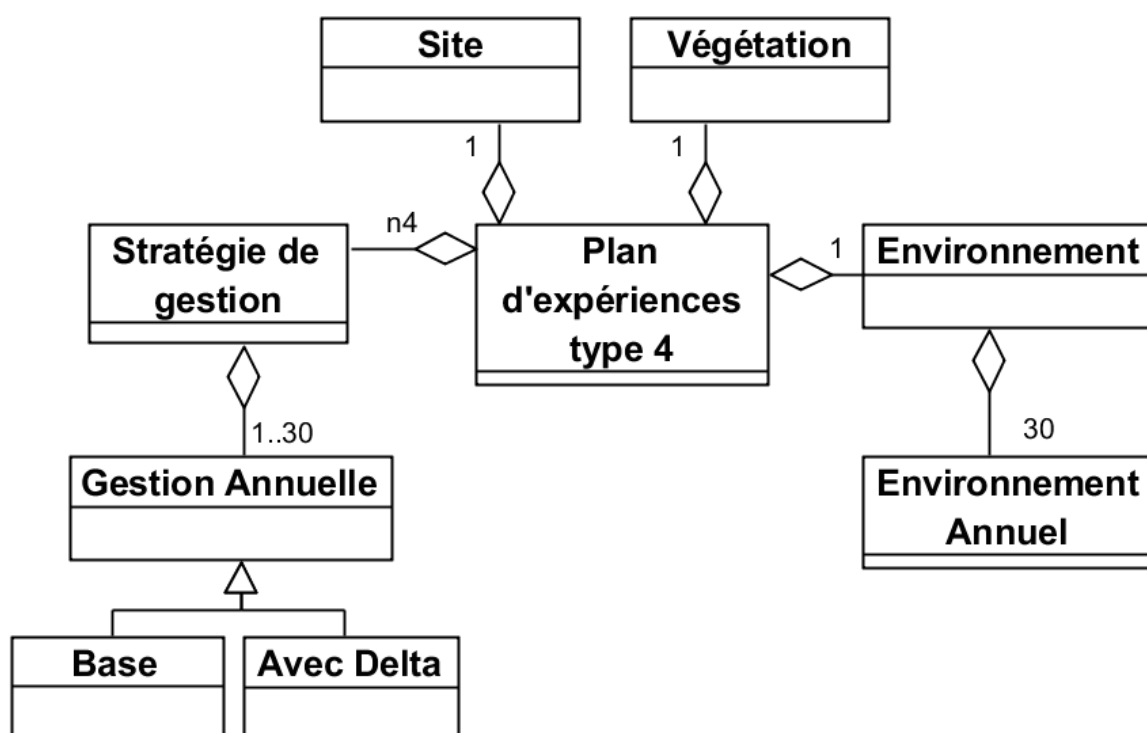


Figure 36. Modèle de DOE pour le calcul de la vulnérabilité sous changement climatique en tenant compte de l'incertitude sur la gestion.

Le plan d'expériences résultant (Figure 36) se compose cette fois de trois composantes fixes : **Site**, **Végétation**, **Environnement** et d'un certain nombre de **Stratégies de Gestion**. Ces dernières étant la composition de **Gestion Annuelle** conforme à l'observation (**Base**) ou sur laquelle on a appliqué une modification (**Avec Delta**). L'intérêt de ce plan d'expériences réside dans le fait que la vulnérabilité calculée tient compte de l'incertitude sur la gestion.

3.4.6. Evaluation combinant les incertitudes de gestion et de climat (dont les évènements extrêmes)

Pour tenir compte de toutes les incertitudes et avoir une bonne évaluation de la sensibilité, on peut combiner les approches précédentes (3.4.3, 3.4.4 et 3.4.5). L'inconvénient majeur cette évaluation est la taille du plan d'expériences résultant. L'utilisation d'un échantillonnage par hypercube latin (LHS : *Latin Hypercube Sampling*) pourrait être une bonne manière de réduire le nombre de simulations nécessaires [McKay et al., 1979].

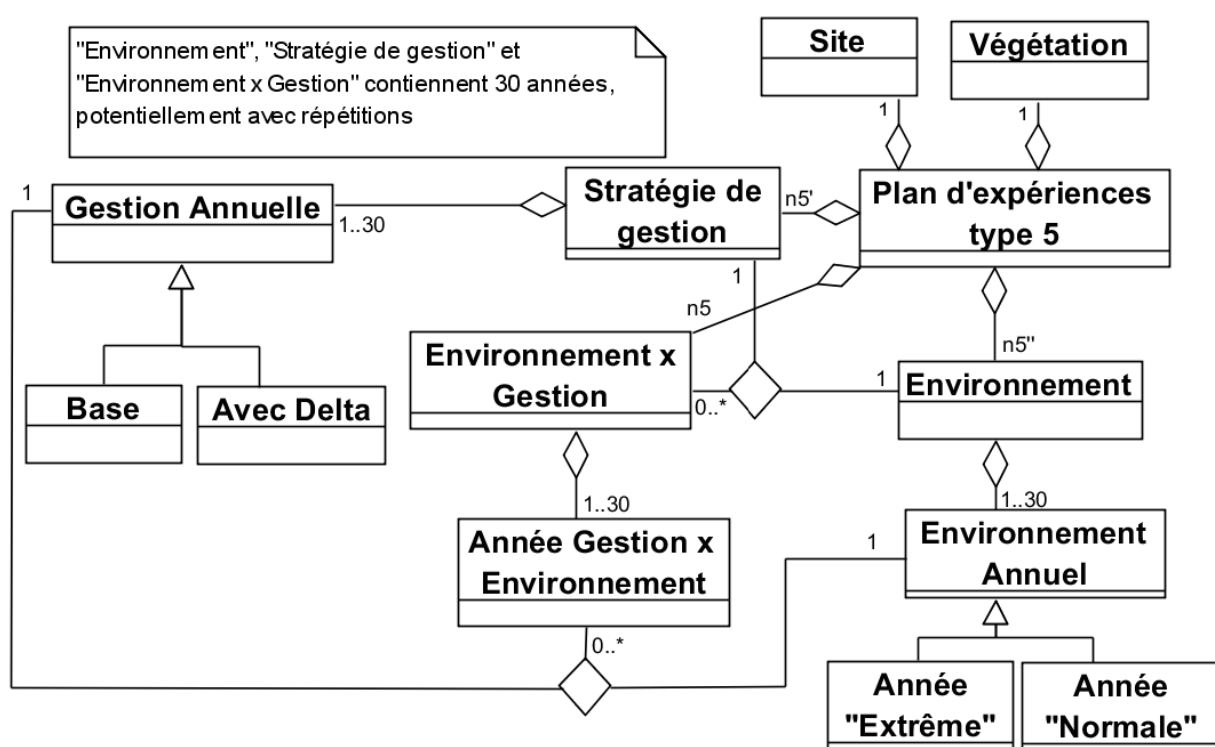


Figure 37. Modèle de DOE pour le calcul de la vulnérabilité sous changement climatique en tenant compte des incertitudes sur la gestion et le climat, y compris la fréquence des évènements climatiques extrêmes.

Ce plan d'expériences (Figure 37) a l'avantage de tenir compte des différentes sources d'incertitudes (les plus probables uniquement car on n'a pas inclus le **Site** ou la **Végétation**). De plus, il apporte une bonne connaissance *a priori* pour le calcul des surfaces de réponse et pour le calcul de la vulnérabilité. Dans ce plan d'expériences on retrouve une combinaison des **Stratégies de Gestion** et des **Environnements**. Ces derniers étant composés de 30 éléments éventuellement avec répétitions. Il est à noter cependant que la multiplication des incertitudes sur les probabilités utilisées pour le calcul de la vulnérabilité peut fausser les résultats. Une analyse de sensibilité de la vulnérabilité calculée par rapport aux probabilités utilisées pourrait être envisagée.

3.4.7. Evaluation du minimum de vulnérabilité

Une fois les surfaces de réponse construites par régression, un minimum de vulnérabilité sera recherché. Cependant, ce que nous recherchons réellement est plus une solution robuste pour réduire la vulnérabilité qu'une solution optimale (au sens mathématique). En effet, en raison du niveau élevé des incertitudes environnementales une solution optimale peut complètement échouer [Kleijnen et al., 2005]. Les solutions robustes pourraient être satisfaisantes par rapport à la vulnérabilité. Par exemple, une solution robuste devrait fournir un rendement prairial élevé avec une faible sensibilité à l'environnement (c'est-à-dire une grande stabilité). Il devrait par nature avoir une faible vulnérabilité. Un exemple de conception robuste, inspiré de [Taguchi 1987] peut être trouvé dans [Sanchez et al., 1996] et [Kleijnen et Gaury, 2003]. L'évaluation de la vulnérabilité trouvée (à l'aide des surfaces de réponse) aura deux objectifs : d'une part, de vérifier que nous avons vraiment trouvé un minimum de vulnérabilité et, d'autre part, d'effectuer une analyse de sensibilité afin d'évaluer la robustesse de la solution. En conclusion, le DOE nécessaire pour cette étape sera similaire à celui décrit précédemment (Figure 37).

3.4.8. Métamodèle des modèles d'impact avec plans d'expériences associés, pour l'analyse de vulnérabilité

Considérant les modèles de plans d'expériences précédemment construits, on peut observer un motif commun (Figure 38 et Figure 39). En effet, tout **Modèle agro-écologique** utilise pour se paramétrer un certain nombre d'**entrées**. Au sein de ces dernières, on peut distinguer les entrées de type **Gestion**, c'est-à-dire celles qui appliquent une stratégie de gestion pouvant fortement modifier les résultats (et qui doivent donc être optimisées en fonction des critères de l'étude), des entrées de type **Environnement**, dont on subit les incertitudes.

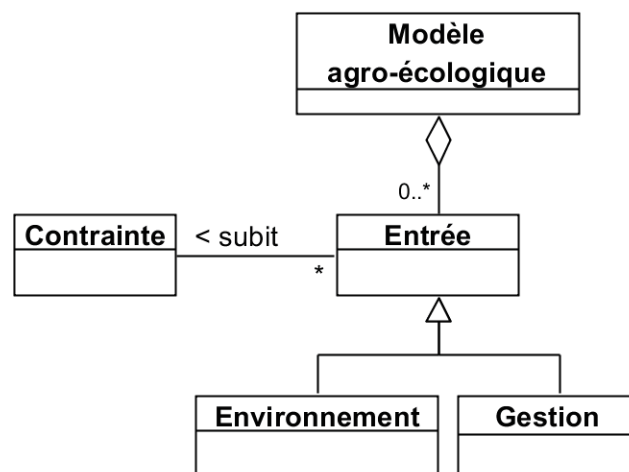


Figure 38. Métamodèle d'un modèle agro-écologique.

Cette distinction entre entrées de type gestion et celles de type environnement correspond à l'approche de Taguchi [Taguchi 1987]. Cette classification aide à proposer un métamodèle auquel tout modèle des entrées d'un modèle agro-écologique serait conforme. On notera également qu'il existe une relation liant les **Entrées** entre elles, cette information étant contenue dans une classe que l'on appellera **Contrainte**. Par exemple, une entrée peut subir la contrainte d'être strictement positive. Comme l'illustre la Figure 39, tout plan d'expériences est bien construit en composant des instances d'**Entrée** (**élément 1, élément 2 ... élément N**). Ces dernières peuvent être combinées (selon le type de plan d'expériences T) pour former un jeu d'entrées cohérent par rapport au modèle agro-écologique (**Entrée 1 x 2 x... x N**). Toutes ces entrées sont contraintes par les instances des classes **Contrainte**. Par exemple, deux valeurs de **Entrée 1** et de **Entrée 2** pourraient s'exclure mutuellement. Ces contraintes peuvent être spécifiées de manière plus précise grâce à OCL (*Object Constraint Language*, [OMG 2010]). OCL est un langage formel pour spécifier les contraintes entre les éléments d'un schéma UML. La cardinalité résultante ($cardT$) est une fonction directe du type de plan d'expériences. Par exemple, si le plan est un plan factoriel complet alors :

$$cardT = \prod_{i=1}^N (card_i),$$

où $card_i$ est la cardinalité de l'espace de définition de l'instance de l'Entrée i . Si le plan est de type "One factor At the Time" [Kleijnen 1987], alors :

$$cardT = 1 + \sum_{i=1}^N (card_i - 1).$$

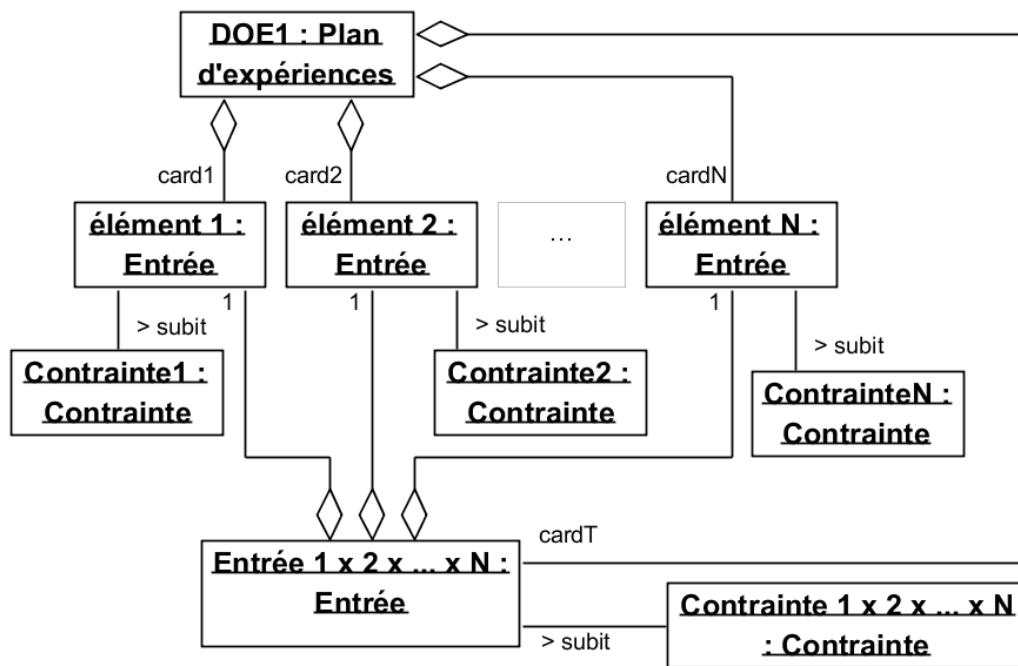


Figure 39. Diagramme d'instance correspondant au modèle UML d'un DOE associé à un modèle agro-écologique dont le métamodèle est décrit dans la Figure 38. Tout **Plan d'expériences** est la composition d'instances d'**Entrée**, formant elles-mêmes un ensemble (instance de la classe **Entrée 1 x 2 x ... x N**).

Tous les exemples de DOE que nous avons présentés sont adaptés aux modèles déterministes. Dans le cas des modèles stochastiques, l'approche reste presque la même. Une légère différence existe à travers l'ajout de répliquations de simulations dans le DOE. Il sera également nécessaire de tenir compte du bruit dans la surface de réponse. Toujours dans le cas des plans d'expériences stochastiques, il sera intéressant de préciser le générateur de nombres pseudo-aléatoires utilisé et, pour être rigoureux, nous devons varier le choix du générateur parmi un ensemble de générateurs fiables : Mersenne Twister [Matsumoto et Nishimura, 1998], les générateurs WELL [Panneton et al., 2006] ou un générateur du type *Multiple Recursive Generator* de L'Ecuyer, comme MRG32k3a [L'Ecuyer 1999]. Les techniques d'initialisation et de parallélisation des générateurs de nombres aléatoires devront également être spécifiées au sein d'un DOE de grande qualité [Hill et al., 2012].

Nous avons défini différentes expérimentations en fonction du niveau d'incertitude dont nous souhaitions tenir compte. Pour chacun de ces designs, nous avons proposé un modèle. L'étude des modèles obtenus nous a permis de franchir une étape supplémentaire grâce à la proposition d'un métamodèle de modèles agro-écologiques et des DOE associés. Chacun des modèles proposés précédemment est conforme à ce métamodèle. Ce travail, faisant suite à un travail de rétro-ingénierie, peut être considéré comme une étape préliminaire dans la conception d'un outil dédié à l'évaluation de la vulnérabilité sous changement climatique. La conception de ce dernier est décrite dans la section suivante.

3.5. Analyse de vulnérabilité sous EMF

Notre objectif est de concevoir un outil qui permet de proposer une méthode pour l'analyse de vulnérabilité, et de l'assister à travers une interface relativement simple d'utilisation. C'est-à-dire qu'il doit permettre l'automatisation de tâches constitutives de l'analyse de vulnérabilité via notamment la génération de codes et de scripts. De façon à apporter genericité et bonne pratique, l'Ingénierie Dirigée par les Modèles (IDM) a été utilisée. C'est pourquoi, il a été fait appel à EMF (*Eclipse Modeling Framework*) car ce cadre intègre au sein du logiciel Eclipse les concepts d'IDM. Cette partie va décrire les principes de l'outil pour l'analyse de vulnérabilité (*i.e.* modèles et métamodèles utilisés) et son implémentation sous EMF.

L'utilisation d'EMF pour la conception de l'outil a abouti à la création d'un plugin Eclipse. Le principe est le suivant : on renseigne dans EMF le métamodèle de l'analyse de vulnérabilité, on spécifie certaines fonctionnalités, puis on génère le code associé. Pour une description plus précise d'EMF, on pourra se référer à la partie 2.5.2. Nous proposons donc un plugin Eclipse, auquel est associée une interface dans laquelle l'utilisateur saisit les renseignements relatifs à son analyse de vulnérabilité, de manière conforme au métamodèle renseigné dans EMF. Le métamodèle utilisé s'inspire du travail de rétro-ingénierie précédemment décrit et des différentes propositions faites de métamodèles et modèles pour les modèles agro-écologiques et les DOE associés pour l'analyse de vulnérabilité. Il est composé de trois sous-parties :

- le modèle agro-écologique et ses propriétés (*i.e.* entrées, sorties) ;
- le plan expérimental ;
- l'analyse de vulnérabilité.

3.5.1. Modèle agro-écologique

Considérant les modèles précédemment proposés, et compte-tenu des besoins pour l'analyse de vulnérabilité, nous avons proposé le modèle suivant (Figure 40) pour les entrées/sorties d'un modèle agro-écologique. Tout **Modèle Agro-écologique** produit une ou plusieurs **Sorties** par simulation. Il utilise un ensemble d'**Entrées** au sein desquelles on distingue : les entrées de type **Gestion**, c'est-à-dire celles qui servent à appliquer une gestion ayant une action sur les résultats et qui, par conséquent, peuvent être optimisées (lors de la phase d'analyse de vulnérabilité avec adaptation) ; les entrées de type **Environnement** servant uniquement à prendre en compte les incertitudes du milieu. Cette distinction entre les deux natures d'entrée, correspondant à l'approche de Taguchi [Taguchi 1987], servira pour la phase d'analyse de vulnérabilité avec adaptation. Chaque entrée peut subir des contraintes (**Contrainte Entrée**), comme par exemple une valeur minimale ou maximale. Dans le cadre d'une analyse de vulnérabilité, une entrée prend ses

valeurs dans différents niveaux définis par un **Echantillonnage**, et l'**Exposition** au stress est représentée par la distribution de probabilité de ces entrées (qui peut être **Gaussienne**, **Uniforme** ou **spécifiée par l'utilisateur**). Suivant cette **Exposition** et cet **Echantillonnage**, chaque entrée va prendre des **Valeurs**, formant ainsi des **Combinaisons** de valeurs (*i.e.* ensemble formé par une valeur pour chacune des entrées). Les combinaisons peuvent éventuellement subir des contraintes spécifiques (**Contrainte Combinaison**), comme par exemple l'exclusion mutuelle de deux valeurs. Ce modèle étant en fait un métamodèle, tous les modèles agro-écologiques sont conformes à celui-ci.

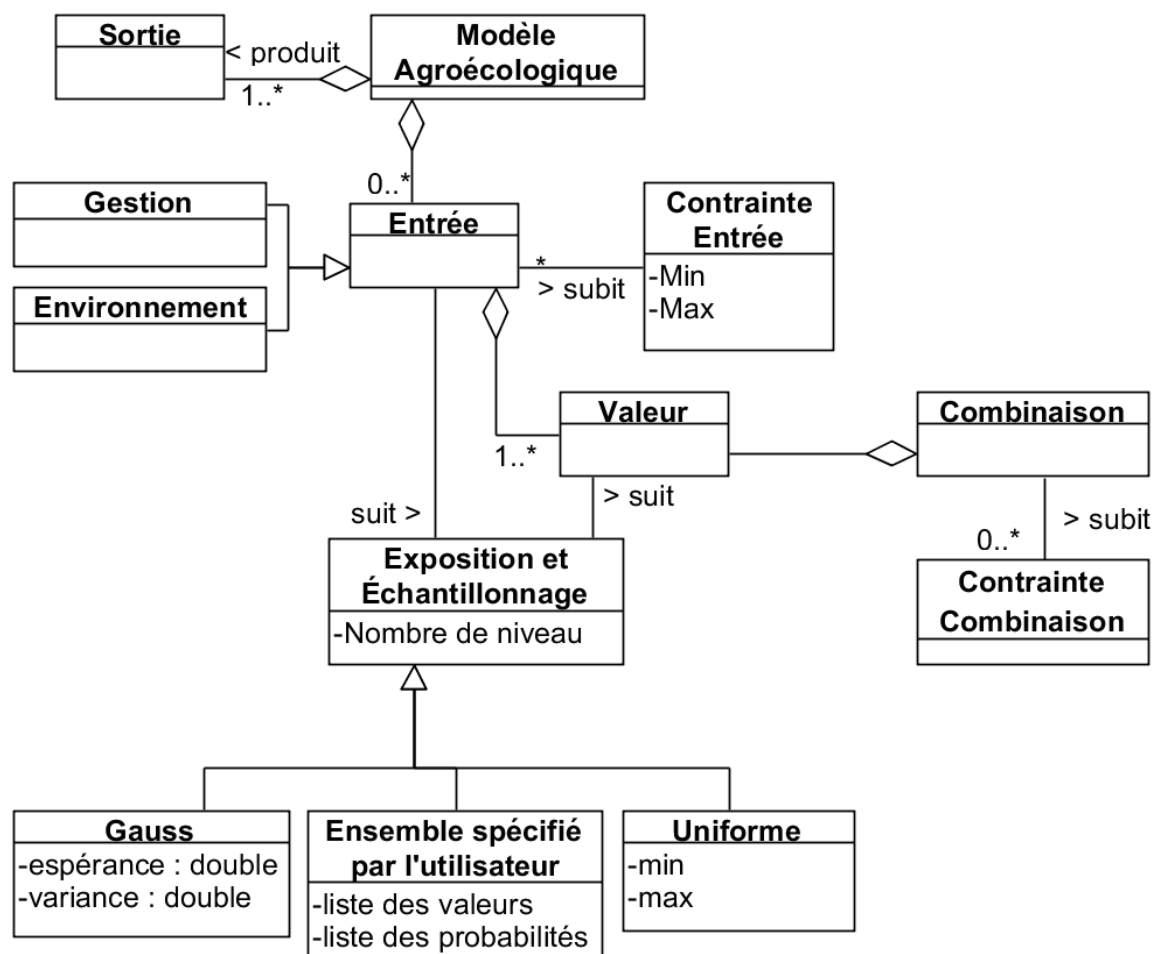


Figure 40. Diagramme UML du modèle des entrées/sorties du modèle agro-écologique, dans le cadre d'une analyse de vulnérabilité.

3.5.2. Le plan expérimental

De même que pour le modèle agro-écologique, en se basant sur le travail précédent, nous avons proposé un métamodèle des plans d'expériences pour l'analyse de vulnérabilité (Figure 41). Tout **Plan d'expériences** est généré à partir d'une **Méthode**, comme par exemple un plan **Factoriel Complet** ou un **LHD** (*Latin Hypercube Design*). Le plan se

compose d'un ensemble de **Combinaisons**, ces dernières correspondant finalement à une simulation.

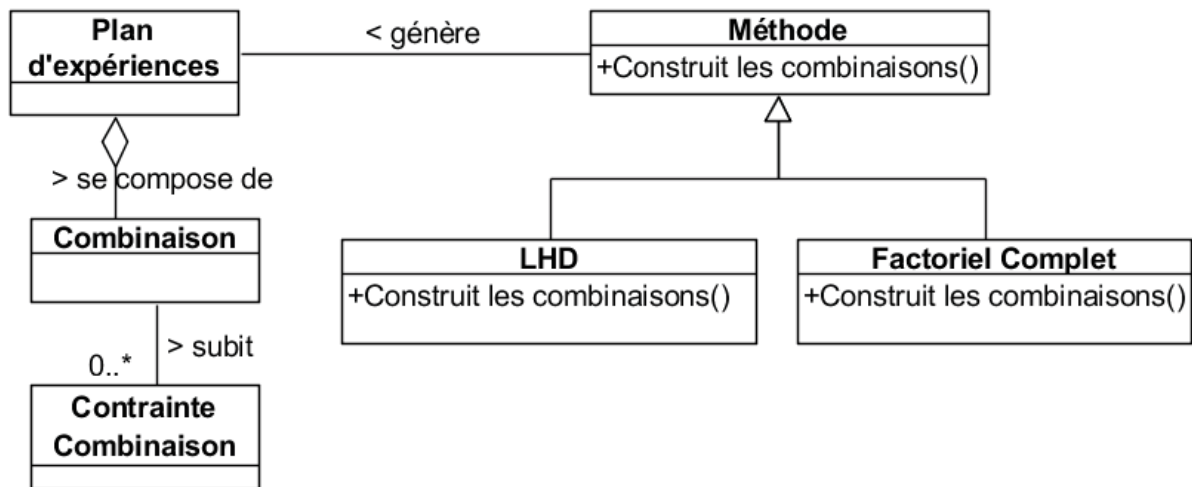


Figure 41. Diagramme UML d'un métamodèle de plan d'expériences, dans le cadre d'une analyse de vulnérabilité.

3.5.3. L'analyse de vulnérabilité

Nous proposons également un métamodèle de l'analyse de vulnérabilité (Figure 42), auquel toute analyse est conforme. En effet, l'**Analyse de Vulnérabilité** utilise un **Indice de Vulnérabilité**, qui peut prendre plusieurs formes tel que l'indice de **Vulnérabilité Proportionnelle** [Foster et al., 1984], celui de **Luers** [Luers et al., 2003], ... Les surfaces de réponses pour l'analyse de vulnérabilité avec adaptation seront construites à l'aide d'une **Méthode de Surface de Réponse** telle que, par exemple, les **Régressions Linéaire** ou **Quadratique** ou le **Krigeage**. Cette surface de réponse sera optimisée avec une **Méthode d'Optimisation** afin de trouver un minimum de vulnérabilité sous contrainte de la capacité d'adaptation. Enfin, les simulations sur lesquelles l'analyse s'appuie seront lancées via un **Lanceur de simulation**. Ce dernier peut être par exemple le logiciel **OpenMOLE**, ou bien les simulations peuvent être simplement lancées par l'utilisateur à l'aide de simples fichiers **Ascii** comme fichiers d'entrée.

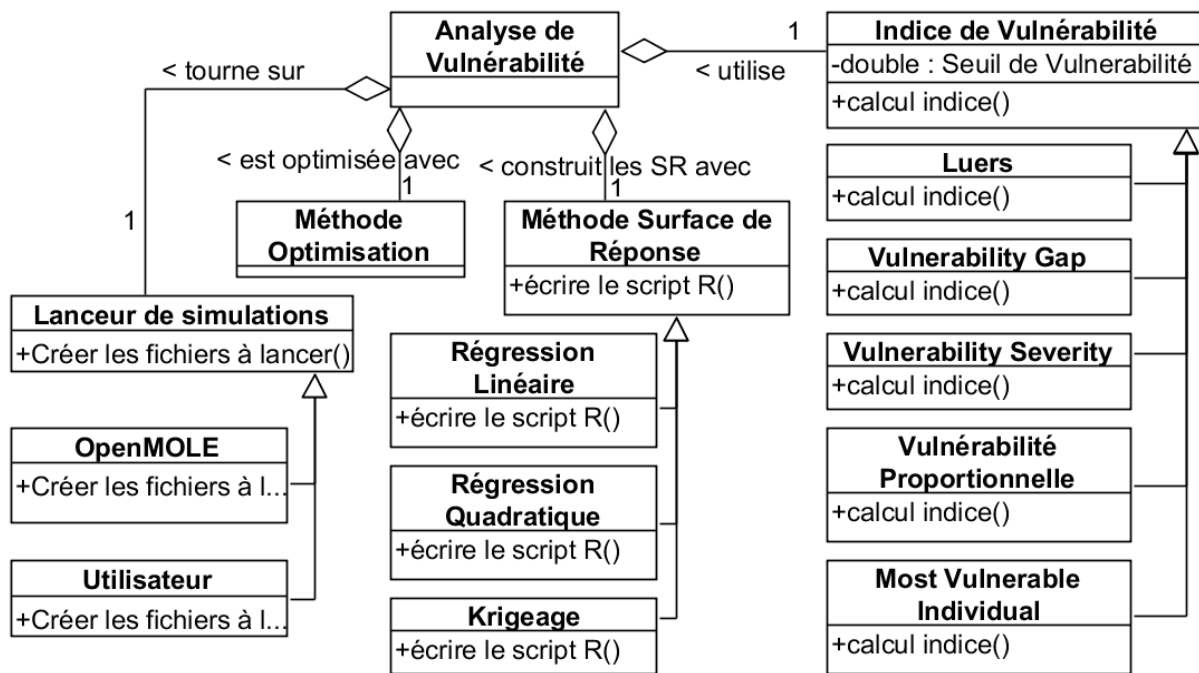
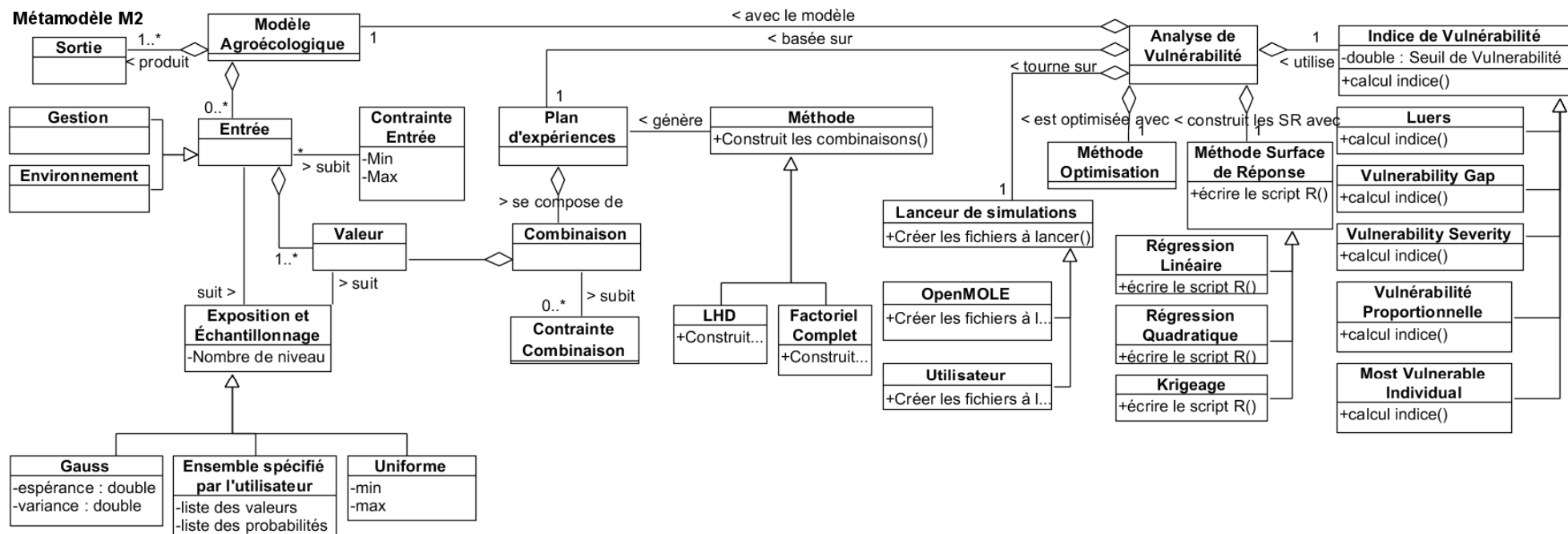


Figure 42. Diagramme UML d'un métamodèle d'analyse de vulnérabilité dans le cadre d'une analyse de vulnérabilité.

3.5.4. Métamodèle complet

L'application assemble les trois modèles précédemment décrits (Figure 40, Figure 41, Figure 42) en un seul diagramme (Figure 43). On notera le lien entre l'analyse de vulnérabilité et le modèle agro-écologique, l'analyse est faite « avec un modèle ». De même l'analyse de vulnérabilité « se base sur » un plan d'expériences. Partant de ce modèle de classe (Figure 43), on peut écrire le diagramme d'instanciation dans lequel on notera l'existence d'un "chemin d'instanciation", spécifiant que certaines classes doivent être instanciées avant d'autres. En l'occurrence les classes **Valeur**, **Combinaison** et **Contrainte Combinaison** ne peuvent être instanciées qu'après l'instanciation des classes **Exposition** et **Echantillonnage** et **Méthode**. Si l'on considère le diagramme des instances (M1) comme étant une représentation (un modèle) de la réalité (M0), alors le diagramme de classe M2 est le métamodèle et le diagramme d'instanciation M1 est bien conforme au métamodèle M2. C'est le métamodèle suivant que nous avons implémenté dans EMF.



Les classes "Valeur", "Combinaison" et "Contrainte Combinaison" sont instanciées en dernier. Leur instanciation nécessite l'instanciation préalable des classes "Exposition et Échantillonnage" et "Méthode".

Figure 43. Métamodèle et modèle d'analyse de vulnérabilité.

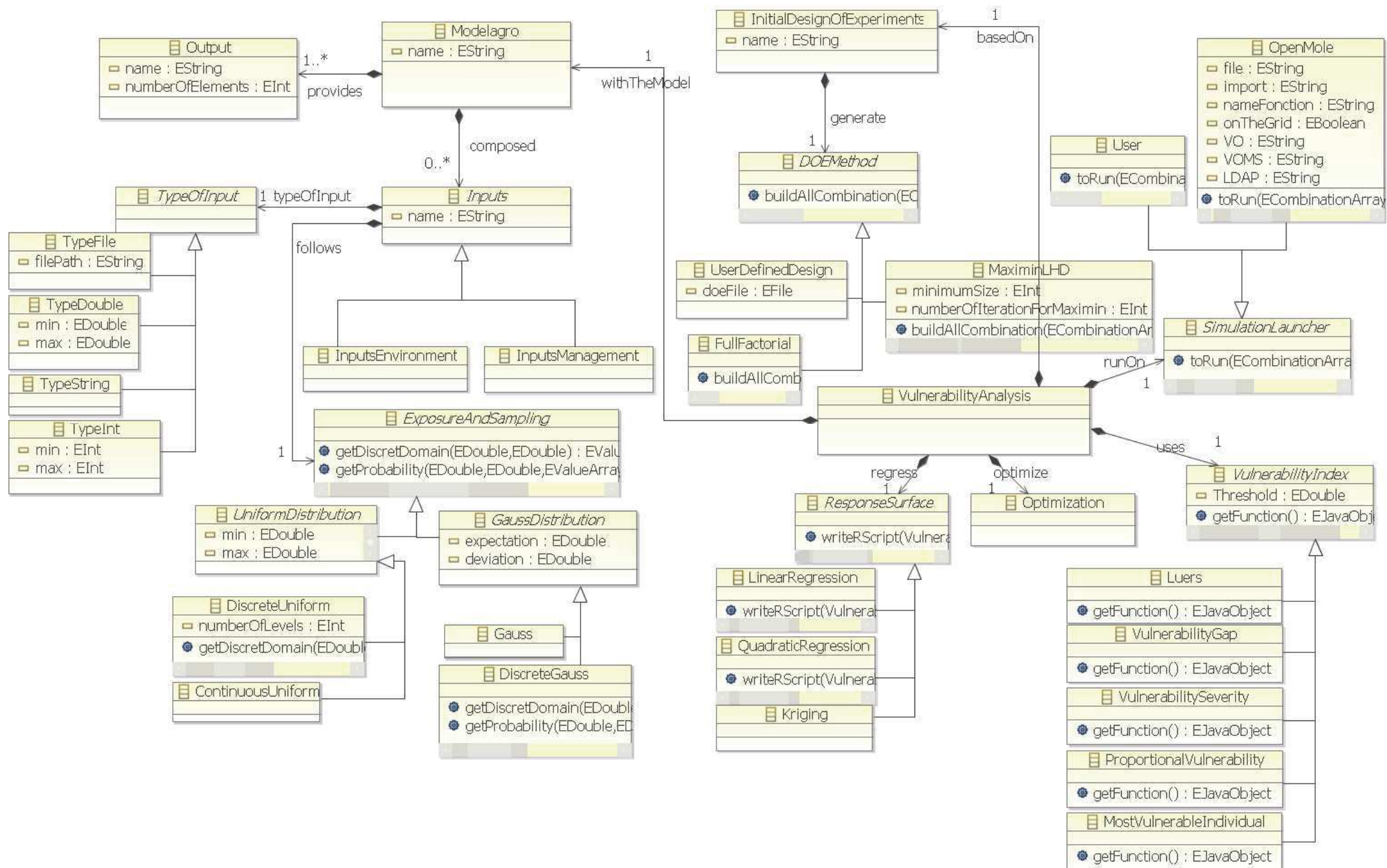


Figure 44. Métamodèle d'analyse de vulnérabilité sous EMF.

3.5.5. Implémentation sous EMF

L'implémentation sous EMF du métamodèle (Figure 44) implique des modifications par rapport au métamodèle théorique (Figure 43) :

- de façon à gérer le chemin d'instanciation, les classes de **Valeur**, **Combinaison** et **Contrainte Combinaison** n'ont pas été saisies dans le métamodèle sous EMF. En revanche, elles sont bien implémentées après la saisie des données dans l'interface. Pour des raisons de simplicité et devant l'absence de nécessité immédiate (c'est-à-dire pour nos applications), la classe **Contrainte Combinaison** n'a pas encore été implémentée ;
- pour des raisons de codage en Java, il est nécessaire de connaître la nature de l'entrée (par exemple double ou int). Pour résoudre ce problème, la classe **TypeOfInput** a été ajoutée. Elle peut être de type **TypeDouble**, **TypeString**, **TypeInt** ou **TypeFile**, selon que l'entrée soit un double, une chaîne de caractères, un entier ou un fichier ;
- les classes **Gauss** et **Uniforme** sont devenues des classes virtuelles et deux sous-classes en héritent de façon à distinguer les cas où l'échantillonnage est discret (c'est-à-dire qu'un nombre défini de valeurs sont prises par les entrées), de celui où il est continu (tirage aléatoire, en suivant une certaine loi, des valeurs prises par les entrées) ;
- la possibilité pour l'utilisateur de saisir, pour une entrée spécifique, la liste de ses valeurs et des probabilités associées (classe **Ensemble spécifié par l'utilisateur**) n'a pas encore été implémentée dans EMF. En revanche, cette fonctionnalité est partiellement couverte par la classe **UserDefinedDesign** qui permet de renseigner son propre plan expérimental en spécifiant les combinaisons et probabilités associées.

L'application s'appuie sur le principe de transformation de modèle. En effet, après avoir instancié le modèle, différentes méthodes vont transformer celui-ci en d'autres modèles, comme par exemple en fichier texte contenant le plan d'expériences, à savoir l'ensemble des combinaisons. Ce fichier texte est un "modèle" des simulations qui seront exécutées pour ce plan d'expériences. De même, le modèle peut être transformé en script R pour construire les surfaces de réponse, ce script étant spécifique aux entrées et aux sorties du modèle agro-écologique de cette analyse de vulnérabilité.

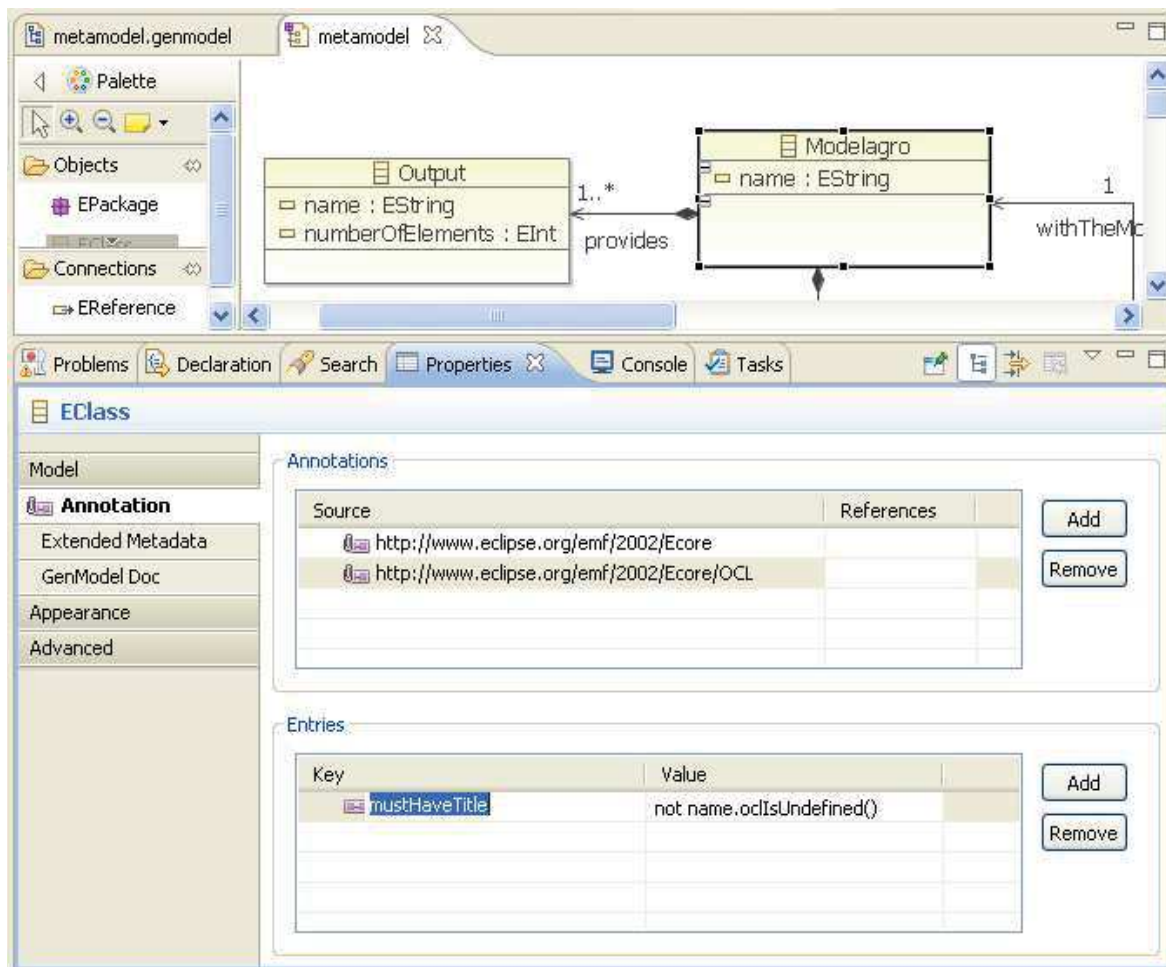


Figure 45. Exemple d'insertion de contrainte OCL dans l'implémentation du métamodèle sous EMF.

Notre instantiation sous EMF peut se résumer ainsi :

1. création du métamodèle (diagramme Ecore) ;
2. spécification au sein de ce dernier (Figure 44) :
 - des classes (par exemple "Modelagro") ;
 - des attributs au sein des classes (par exemple, l'attribut "name" de type EString dans la classe "Modelagro") ;
 - des contraintes, exprimées via OCL (par exemple, Figure 45, l'obligation pour l'attribut "name" de la classe "Modelagro" d'être défini) ;
 - des méthodes (par exemple, Figure 46, le code de la méthode "getProbability" au sein de la classe "DiscreteGauss").

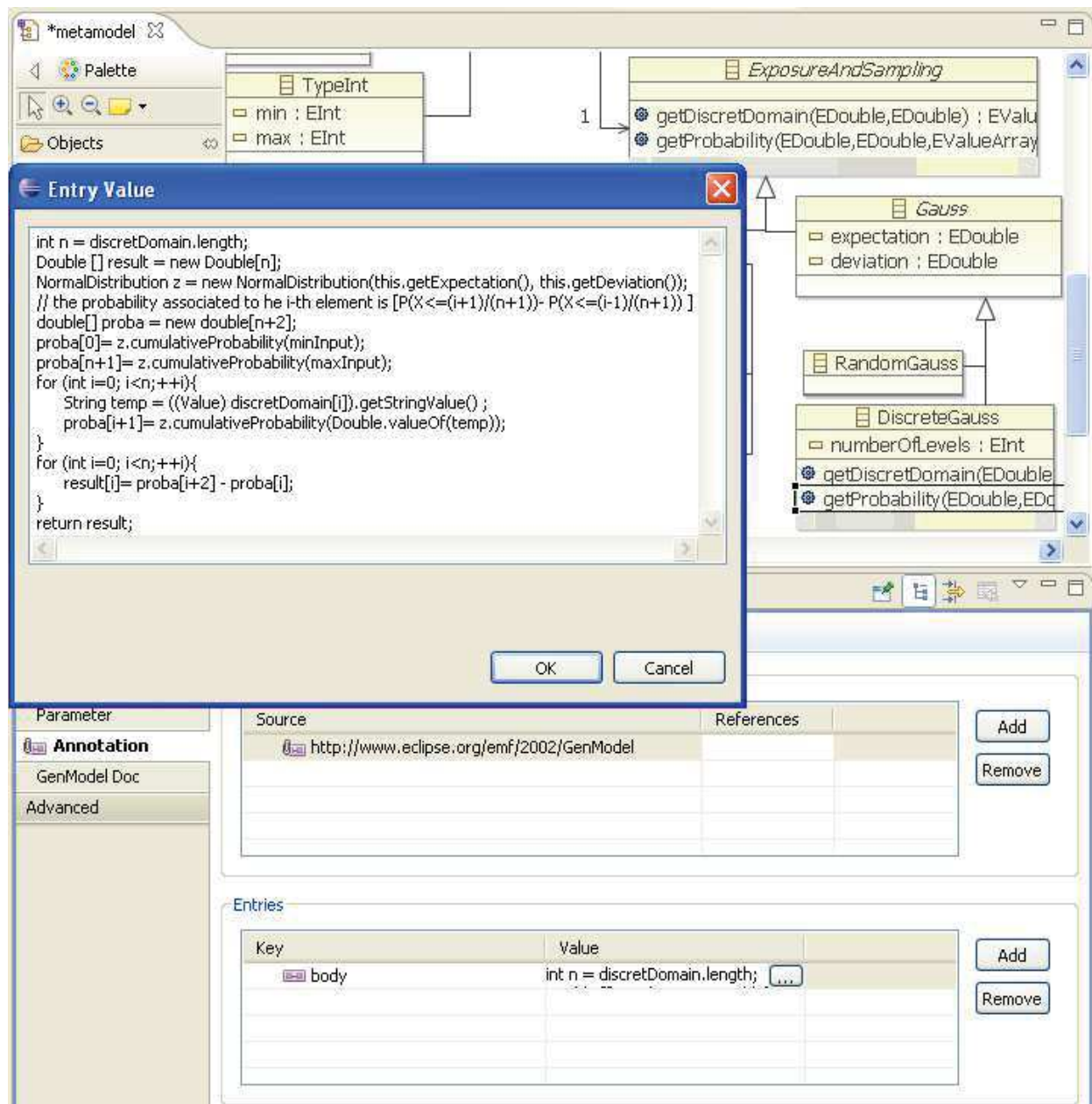


Figure 46. Exemple de spécification du code contenu dans une méthode sous EMF, à travers le principe des annotations du Java.

3. génération du code du modèle, des éditeurs de modèles et des tests unitaires ;
4. création puis export des plugins Eclipse liés au modèle ;
5. création du plugin spécifique pour l'analyse de vulnérabilité (dont le fonctionnement est synthétisé dans la Figure 47) :
 - création dans l'interface d'un menu spécifique pour lancer l'analyse de vulnérabilité ;

- gestion de la lecture des données du modèle, stockées sous forme de fichier XML et appel des méthodes permettant de vérifier la validité des données ;
- instanciation du modèle à partir des données lues dans le fichier XML, y compris les classes **Valeur** et **Combinaisons** qui ne sont pas explicitement décrites par l'utilisateur ;
- implémentation des fonctionnalités avancées de l'analyse de vulnérabilité telle que la création du plan d'expériences, l'estimation de l'exposition, la génération des scripts (pour lancer les simulations ou pour le calcul des surfaces des réponses), l'appel à la bibliothèque⁶³ de calcul des indices d'analyse de vulnérabilité.

Un tutoriel d'utilisation de l'outil, ainsi que des exemples de résultats peuvent être trouvés en annexe III.

Grâce à l'approche retenue en utilisant le cadrage EMF, la genericité et la modularité de l'outil est très grande, que ce soit en termes d'utilisation (par exemple, utilisation simple pour générer un DOE, ou juste pour créer le script R permettant de construire une régression linéaire), ou de potentialité d'évolutions. Il permet en effet très simplement d'intégrer une classe traduisant une nouvelle fonctionnalité (par exemple, un nouvel indice de vulnérabilité, des scripts Matlab...). Il est à noter, cependant, que la phase d'analyse de vulnérabilité avec adaptation n'a été que partiellement implémentée dans le plugin. Cette phase se traduira à l'avenir par la recherche d'un minimum de vulnérabilité sur les surfaces de réponses, puis par un nouveau plan d'expériences pour évaluer ce minimum. Cependant, cette étape n'ayant pas encore été suffisamment testée et mise en œuvre préalablement, son implémentation s'est limitée à la création de la classe associée (*i.e.* la classe **Méthode Optimisation**).

⁶³ Un ensemble de classes et de méthodes spécifiquement dédiées aux calculs des indices de vulnérabilité a été implémenté. Une classe *vulnerability* contient l'ensemble des données (valeurs, exposition, seuil,...) sur lesquelles on souhaite calculer la vulnérabilité, tandis qu'un ensemble de classes implémentant une interface permettent de spécifier le calcul de chaque indice de vulnérabilité.

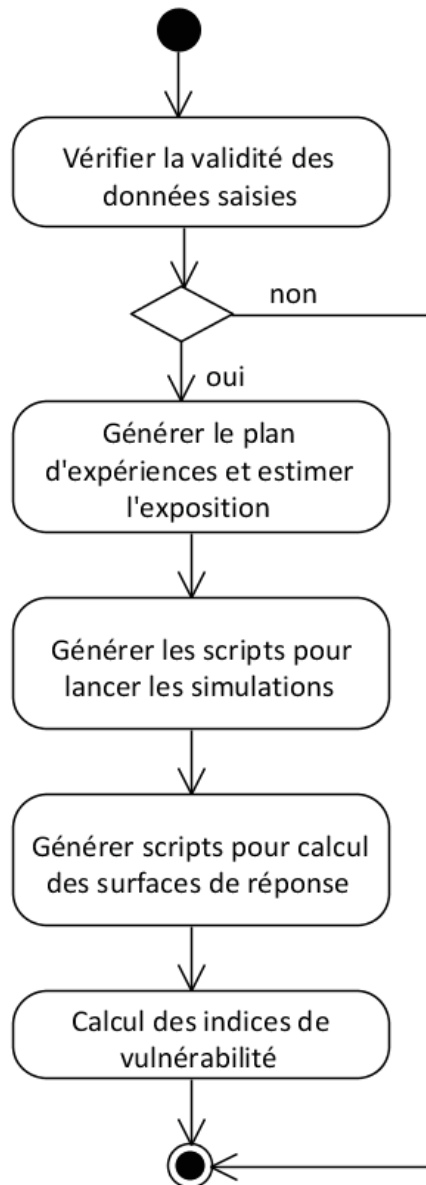


Figure 47. Diagramme d'activité de l'analyse de vulnérabilité par le plugin EMF.

3.6. Conclusion

L'analyse de vulnérabilité se compose de nombreuses étapes. Ainsi, en se référant à l'expérience acquise sur trois campagnes de simulations réalisées avant cette thèse, à l'aide du modèle de simulation de prairie PaSim, et représentant chacune une étape de l'analyse de vulnérabilité, nous savons qu'un outil dont le but sera de permettre la réalisation d'analyses de vulnérabilité nécessitera la réalisation d'un grand volume de simulations. Certains outils permettent déjà la distribution de plans d'expériences tout en masquant l'hétérogénéité de la plateforme. Cependant, pour les adapter à l'analyse de vulnérabilité, nous avons commencé, à travers une approche de rétro-ingénierie, par chercher quels sont les modèles des entrées de logiciels tels que PaSim et les modèles des plans expérimentaux déjà réalisés (ceux présentés précédemment étant relativement similaires à ceux requis

pour une analyse de vulnérabilité). Cette étude nous a conduits à proposer un métamodèle des plans d'expériences dans notre contexte. Puis, pour raffiner nos modèles et métamodèles, nous avons proposé des expérimentations en fonction du niveau d'incertitude que l'on souhaite prendre en compte dans l'analyse de vulnérabilité au changement climatique. Ensuite, pour chacune de ces expérimentations nous avons proposé un modèle. Cela nous a permis d'affiner notre proposition de métamodèles pour des modèles agro-écologiques, avec des plans expérimentaux associés à une analyse de vulnérabilité. Dans une quatrième étape, nous avons combiné ces métamodèles (modèle agro-écologique, plan expérimental et analyse de vulnérabilité) pour former un métamodèle complet, de telle sorte que toute analyse de vulnérabilité à l'aide d'un modèle agro-écologique et sous contrainte de plans d'expériences soit conforme à ce métamodèle. Ce dernier a été implémenté dans le cadre EMF (*Eclipse Modeling Framework*) sous forme de plugin Eclipse de façon à créer un outil permettant la réalisation et l'automatisation d'une analyse de vulnérabilité. Cet outil peut éventuellement être utilisé pour réaliser seulement certains aspects de l'analyse de vulnérabilité (par exemple, juste l'analyse de sensibilité). L'outil est l'implémentation logicielle de la méthode d'analyse de vulnérabilité proposée au début de ce chapitre. La particularité de cette méthode est de permettre, à travers deux phases, d'analyser la vulnérabilité avec et sans mesures d'adaptation. L'aspect opérationnel de l'outil et de la méthode est conforté par les résultats présentés dans le chapitre suivant. Cependant, l'aspect "analyse de vulnérabilité avec prise en compte d'adaptations" ayant été plus faiblement étudié, l'outil proposé prévoit l'existence de cette phase (*i.e.* des classes ont été créées) mais il ne l'implémente pas encore.

Chapitre 4 : Applications d'analyses de vulnérabilité au changement climatique

4.1. Introduction

Le chapitre précédent proposait une démarche d'analyse de vulnérabilité pour répondre à notre problématique, ainsi que l'implémentation logicielle de cette solution afin de faciliter la réalisation et l'interprétation des résultats qui en découlent. À travers ce chapitre, nous verrons certaines étapes constitutives de notre démarche d'analyse de vulnérabilité.

Dans la première partie (4.2), nous allons détailler un exemple illustrant la démarche d'analyse de vulnérabilité, sans adaptation, de la productivité prairiale à l'aide d'un modèle simplifié de prairies, ModVégé [Jouven et al., 2006a], en tenant compte de différentes sources d'incertitudes. Cette partie permet également de mettre en avant l'utilisation d'un jeu d'indices de vulnérabilité. Cette partie a été présentée lors de la conférence de « *l'International Environmental Modelling and Software Society* » en 2012 à Leipzig – Allemagne (iEMSs 2012) [Lardy et al., 2012b].

La deuxième partie (4.3) illustre le cas particulier d'une analyse de vulnérabilité spatialisée. Nous expliciterons ce contexte particulier à l'aide de simulations de prairies permanentes européennes, réalisées à l'aide du modèle mécaniste de prairie, PaSim [Riedo et al., 1998].

La troisième partie (4.4) discute la question du nombre d'années à utiliser dans le cycle climatique servant, à l'aide du modèle PaSim, à mettre en équilibre la matière organique du sol en fonction de la gestion prescrite. Cette partie sert de pré-requis à la partie suivante.

La quatrième partie (4.5) a été publiée et présentée à la conférence iEMSs 2012 [Lardy et al., 2012a]. Elle décrit la construction de surfaces de réponse de la matière organique du sol à l'équilibre pour une gamme de conditions représentatives de différents climats, sols et gestions.

À l'aide des surfaces de réponses réalisées, nous avons proposé dans la section suivante (4.6) une évaluation de la vulnérabilité au changement climatique de la matière organique, à l'équilibre, des prairies permanentes. Pour cela, nous décrirons les phases de calculs de la vulnérabilité sans prise en compte de l'adaptation. D'abord nous rechercherons un minimum de vulnérabilité et enfin une phase d'évaluation de ces minima nous permettra d'évaluer la vulnérabilité avec adaptation.

4.2. Cas exemple à l'échelle parcellaire

Dans le but d'illustrer la démarche d'analyse de vulnérabilité décrite dans les chapitres précédents, le modèle ModVégé [Jouven et al., 2006a], décrit ci-dessous, a été utilisé. Une campagne de simulations basée sur différents plans d'expériences tenant compte de différentes sources d'incertitudes a été réalisée, puis les résultats ont été analysés à l'aide d'un jeu d'indices de vulnérabilité. Ce travail a permis dans le chapitre précédent de faire la proposition d'un certain nombre de plans d'expériences à partir desquels nous avons fait des propositions de modèles et de métamodèle de plans d'expériences.

4.2.1. Le modèle ModVégé

ModVégé [Jouven et al., 2006a] est un modèle mécaniste multi-années pour la simulation de la dynamique de la production, de la structure et la digestibilité de l'herbe des prairies permanentes gérées. Il est conçu pour répondre aux différents régimes de défoliation. Il repose sur cinq hypothèses. Tout d'abord, la valeur moyenne des propriétés biologiques de la végétation (traits fonctionnels) peuvent expliquer le fonctionnement d'une prairie permanente [Louault et al., 2005]. Deuxièmement, l'hétérogénéité du couvert peut être modélisé par l'abondance relative des composants structurels de la plante [Carrère et al., 2002] (c'est-à-dire gaines et feuilles vertes, gaines et feuilles mortes, fleurs et tiges vertes, fleurs et tiges mortes). Troisièmement, comme dans beaucoup de modèles dynamiques de prairie (par exemple, HPM, [Thornley 1998]), la sénescence, la croissance et l'abscission peuvent être modélisées par des flux continus, calculés sur un pas de temps journalier. Quatrièmement, en raison du stockage des réserves et de la mobilisation dans les organes végétaux, la croissance du couvert, basée sur une approche d'efficacité d'utilisation de la lumière, est modulée par une tendance saisonnière (voir par exemple [Volenc et al., 1996]). Cette dernière est considérée comme un trait fonctionnel. La dernière hypothèse est que la qualité du compartiment vert, l'abscission et la sénescence sont influencées par le vieillissement des compartiments. Le modèle a été évalué pour les prairies de montagne dans la région Auvergne, en France [Jouven et al., 2006b]. Le choix de ce modèle est lié à sa complexité (suffisante pour reproduire les effets de la variabilité climatique sur la prairie [Jouven et al., 2006b]) et son faible nombre d'entrées nécessaires, couplé à un temps de calcul limité (quelques secondes pour 30 années climatiques). Juste quelques sorties sont utilisées de manière directe ou couplées avec d'autres modèles (par exemple un modèle animal). Les plus utilisées sont la biomasse sur pied et la digestibilité de l'herbe. Un exemple de résultats de simulation pour une prairie de montagne auvergnate est donné par la Figure 48.

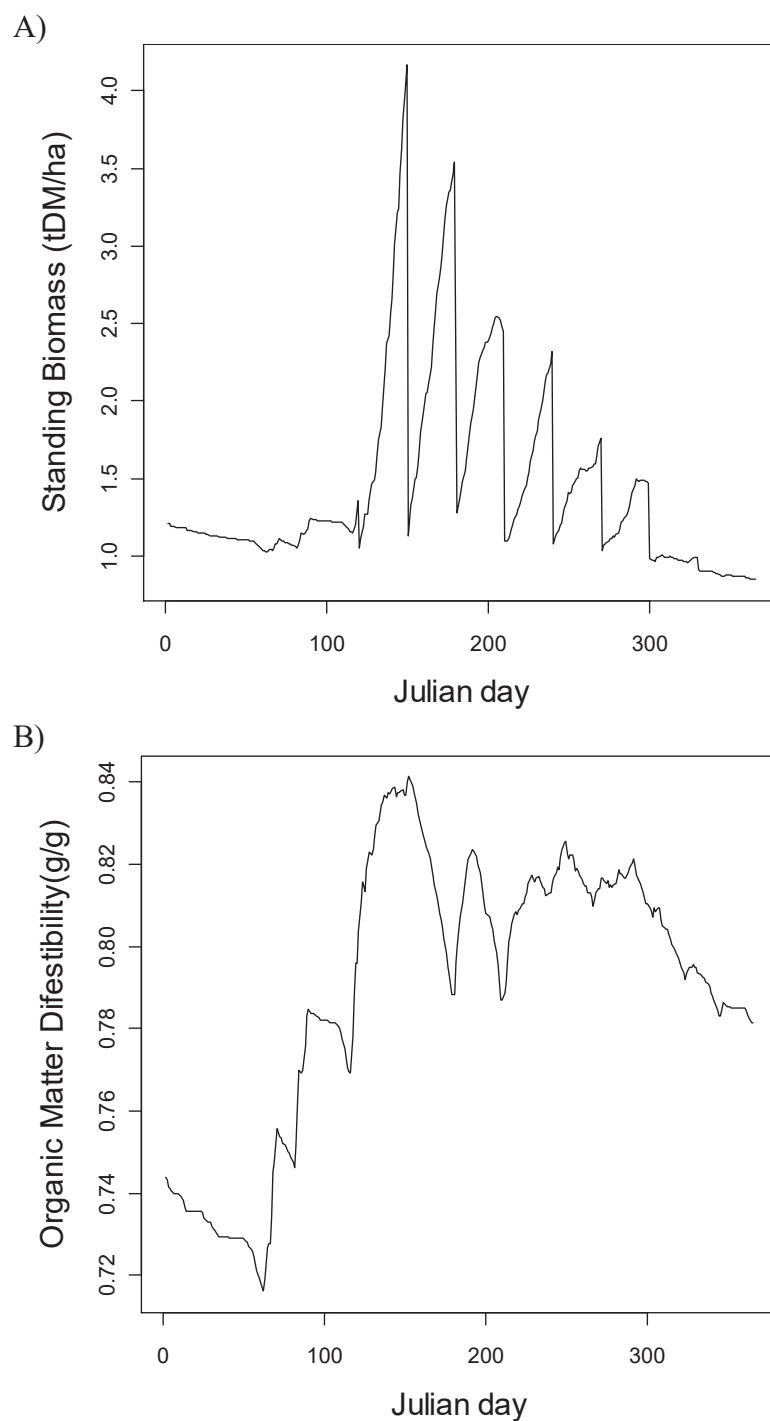


Figure 48 [Lardy et al., 2011a] Résultats des simulations de A) Biomasse sur pied et B) digestibilité de l'herbe obtenus par ModVégé sur une prairie permanente intensive (année 1994) à Marcenat (45° 18' Nord, 02° 49' Est), en Auvergne.

4.2.2. Matériels et méthodes

Dans la littérature internationale, l'analyse de vulnérabilité est souvent une analyse qualitative et seulement dans quelques cas basés sur des indices quantitatifs. Notre approche consiste à évaluer de manière quantitative la vulnérabilité à l'aide de l'indice de

Luers [Luers et al., 2003], des mesures généralisées de pauvreté [Foster et al., 1984] et une extension de ces dernières (Table 9). La vulnérabilité est une notion relative et les valeurs absolues d'un indice de vulnérabilité ne sont pas très pertinentes [Downing et al., 2001]. Généralement, définir la vulnérabilité d'un système requiert préalablement d'identifier un seuil en dessous ou au dessus duquel le système est considéré comme endommagé.

Table 9 Résumé des indices de vulnérabilité. W est la variable d'état (*i.e.* productivité), W_0 est le seuil (*i.e.* 750 kg MS ha⁻¹), n est le nombre d'éléments (*i.e.* nombre d'années), q le nombre d'éléments en dessous de la valeur du seuil. Lors du calcul, les indices sont pondérés de l'exposition.

Indice	Formule	Interprétation
Vulnérabilité proportionnelle, [Foster et al., 1984]	$V_0 = \frac{q}{n}$	Il correspond au nombre d'individus vulnérables dans la population (les années dans notre cas).
Vulnerability gap, [Foster et al., 1984]	$V_1 = \frac{1}{n} \left[\sum_{i=1}^q \frac{(W_0 - W_i)}{W_0} \right]$	Il représente le déficit moyen des individus vulnérables.
Vulnerability severity, [Foster et al., 1984]	$V_2 = \frac{1}{n} \left[\sum_{i=1}^q \left(\frac{(W_0 - W_i)}{W_0} \right)^2 \right]$	La distance moyenne au seuil est utilisée comme pondération. Plus de poids est donné aux cas les plus vulnérables.
Most vulnerable individual	$V_\infty = 1 - \frac{\min W_i}{W_0}$	C'est la distance relative au seuil du cas le plus vulnérable.
Luers et al. [2003]	$V_L = f \left(\frac{ \partial W / \partial X }{W / W_0} \right)$	Le coefficient de variation est utilisé pour quantifier la sensibilité du système.

Pour illustrer les résultats qui peuvent être obtenus par les différents indices, une prairie permanente de montagne à faible productivité à Theix (45° 43' Nord, 03° 01' Est, à 850 m d'altitude) a été simulée pour trois périodes climatiques de 30 ans chacune : période de "référence" (1975-2004), période de "futur proche" (2020-2049), période de "futur lointain" (2070-2099). Les projections climatiques futures sont basées sur le scénario d'émissions A1B [Nakićenović et al., 2000], généré à l'aide du modèle climatique ARPEGE⁶⁴ [Déqué et al., 1994], puis régionalisé selon la méthode dite "des anomalies" [Déqué 2007]. Le sol est caractérisé par une capacité de rétention d'eau de 200 mm. La gestion de la prairie a été simplifiée à une seule coupe, le 15 juin de chaque année.

Notre variable d'intérêt est la valeur de production annuelle estimable en fauche en matière sèche (MS). Cette dernière est calculée comme 1,5 fois la valeur au pic de biomasse, ce dernier correspondant au potentiel de croissance de la prairie. La valeur au pic de biomasse est prise à 1100 °C jour (*i.e.* la somme jour par jour de la température moyenne de l'air au dessus de 0 °C, depuis le 1^{er} janvier) ou à la première fauche si elle a lieu avant.

⁶⁴ <http://www.cnrm.meteo.fr/gmgec/arpege/arpege.html>

Cette sortie a été simulée par le modèle ModVégé sous conditions climatiques actuelles et futures. Le seuil arbitraire (W_0) de 750 kg MS ha⁻¹ a été utilisé comme valeur en deçà de laquelle le système est considéré comme vulnérable. Nous avons utilisé la période climatique actuelle comme période de référence pour calculer des valeurs relatives pour tous les indices de vulnérabilité au changement climatique. L'indice relatif est par conséquent défini comme la valeur pour le climat futur divisée par la valeur de la période de référence. Dans certains cas, comme pour l'indice de Luers, la valeur seuil (W_0) disparaît dans l'indice relatif, qui est alors juste influencé par le climat de référence.

De plus, afin d'évaluer comparativement les différents indices de vulnérabilité, mais aussi pour tenir compte des différentes sources d'incertitudes, nous avons conçu une série d'expérimentations illustrant une gamme de cas de figures. Les plans d'expériences (DOE) utilisés sont les mêmes que ceux proposés dans le chapitre précédent (3.4.1, Table 8). Ces derniers peuvent être résumés de la manière suivante.

Le premier design ("simple") consiste à simplement simuler le système sur 30 ans pour une gestion et des conditions environnementales données.

Dans la deuxième expérimentation ("incertitudes climatiques"), les années climatiques sont représentatives d'une période, et la corrélation entre l'année N et $N-1$ est négligeable devant leur appartenance à cette même période. Autrement dit l'ordre d'occurrence des années est aléatoire. Une simplification a cependant été introduite dans la conception car on ne tient pas compte de l'augmentation de la concentration de CO₂ atmosphérique au cours des années (tel que prévu par le scénario d'émissions utilisé). Par tirage aléatoire sans remplacement (bootstrapping [Efron et Tibshirani, 1993]), nous avons généré 10 000 climats.

Dans le troisième DOE, nous avons considéré que la fréquence des événements extrêmes aurait pu être supérieure à celle observée. Ainsi, la série de 30 ans a été simulée en remplaçant une à trois années par des années avec événement climatique extrême. L'année climatique extrême a été définie comme étant l'année la plus aride de la période, d'après l'indice d'aridité de De Martonne-Gottmann [De Martonne 1942]. Pour l'évaluation de la vulnérabilité, il a été nécessaire de pondérer les séries climatiques par la probabilité d'occurrence des événements extrêmes. Ainsi la probabilité d'occurrence⁶⁵ d'1, 2, 3 ou 4 événements extrêmes dans la série climatique était respectivement de 50%, 40%, 9% et 1%.

La quatrième expérimentation considère la possibilité que la gestion aurait pu être légèrement différente (par exemple, modification des dates de fauche). Pour tenir compte

⁶⁵ Bien que l'on puisse considérer que la fréquence des événements extrêmes suit une loi de Poisson, nous avons préféré, par simplicité, ne pas considérer dans notre exemple les séries climatiques sans événements extrêmes.

de ces incertitudes, le DOE permet jusqu'à ± 9 jours de différence par rapport aux dates initialement prévues, avec une distribution connue (gaussienne dans ce cas).

4.2.3. Résultats

Les simulations ont été réalisées grâce à l'outil de distribution de simulations OpenMOLE⁶⁶ (voir 2.4.6). Puis les indices de vulnérabilité ont été calculés pour les différentes expérimentations (Table 10). En négligeant les effets de l'augmentation du CO₂, le changement climatique augmente la vulnérabilité du système. Cela s'explique par une diminution de la production des prairies, couplée à une faible réduction de la variabilité interannuelle. Les simulations tenant compte des incertitudes sur la fréquence des événements extrêmes n'ont pas montré beaucoup de différences par rapport au "simple" design, probablement à cause d'une faible sensibilité du modèle ModVégé aux événements climatiques extrêmes. La prise en compte de l'incertitude sur l'ordre des années (deuxième design) augmente globalement les valeurs des indices de vulnérabilité. Cela montre que les incertitudes sur les scénarios d'émissions de GES, les modèles climatiques et les méthodes de régionalisation doivent être prises en compte dans les études d'analyses de vulnérabilité au changement climatique. Quelque soit l'indice considéré, la prise en compte de l'incertitude sur la gestion tend à réduire la vulnérabilité. Cela signifie que l'on doit tenir compte de la sensibilité et des incertitudes de la gestion (dans notre cas les dates de fauche) pour l'évaluation de la vulnérabilité, mais également dans la recherche d'options d'adaptation. En effet, l'adaptation cherche à réduire la vulnérabilité à travers une solution robuste (*i.e.* qui traduit une faible sensibilité du modèle aux perturbations climatiques). Toutefois quelque soient les sources d'incertitudes prises en compte, une tendance claire est observée, c'est-à-dire une augmentation de la vulnérabilité du système de prairie permanente étudiée. Dans le cas où les indices de vulnérabilité ne produiraient pas des résultats similaires (ou en cas d'absence de tendance claire), ne pas considérer certaines incertitudes associée à la gestion, au climat, ou plus globalement au contexte environnemental, pourrait fausser les résultats.

Utiliser une gamme d'indices dans l'évaluation de la vulnérabilité est important parce que chacun d'entre eux contient des informations complémentaires. Par exemple, l'indice "Most vulnerable individual" nous informe que la productivité de l'année la plus vulnérable devrait être jusqu'à trois fois plus faible dans le futur lointain que dans la période actuelle (DOE climatique). Tandis que l'indice "Vulnerability severity" nous informe dans quelle mesure le système pourrait être sévèrement endommagé. Dans le même temps, grâce à l'indice "Vulnerability gap", nous savons que la biomasse manquante moyenne pour les années vulnérables pourrait augmenter jusqu'à huit fois, tandis que le nombre d'années vulnérables augmente de 2 à 2,5 pour le "futur proche" et la période du "futur lointain",

⁶⁶ <http://www.openmole.org/>

d'après l'indice de vulnérabilité proportionnelle. L'indice de Luers (calculé ici en utilisant le coefficient de variation comme mesure de sensibilité du système au climat) est une sorte d'indice moyen, qui combine des informations sur la productivité globale et sur la variabilité du système. Il a également l'avantage d'être indépendant du seuil fixé, et par conséquent de ne pas nécessiter de choix arbitraire concernant un seuil. Pour les autres indices, dans le cadre d'une analyse de vulnérabilité complète, une analyse de sensibilité de la valeur de seuil doit être réalisée pour vérifier la robustesse des résultats obtenus.

Table 10. Indices de vulnérabilité, calculés pour différents plans d'expériences (DOE) pour le "futur proche" (FP : 2020-2049) et le "futur lointain" (FL : 2070-2099). Les valeurs absolues des indices ont été normées par la valeur de la "période de référence" (PR : 1975-2004). Plus la valeur de l'indice est haute, plus le système est vulnérable.

DOE \ Indice	Most vulnerable individual		Indice de Luers		Vulnérabilité proportionnelle		Vulnerability gap		Vulnerability severity	
	FP	FL	FP	FL	FP	FL	FP	FL	FP	FL
Simple	1,62	2,76	1,66	2,23	2,13	2,63	3,14	7,64	4,49	20,37
Climatique	1,57	3,00	1,42	2,33	1,92	2,58	2,98	8,35	4,38	23,58
Evènements extrêmes	1,62	2,77	1,67	2,22	2,16	2,64	3,16	7,64	4,48	20,22
Gestion	1,45	2,45	1,67	2,19	2,06	2,55	3,08	7,42	4,12	18,65

4.2.4. Discussion

Ces résultats servent de preuve de concept de notre démarche d'évaluation de la vulnérabilité. Cette démarche sera étendue à d'autres études plus complètes à l'aide de modèles biogéochimiques tel que le modèle PaSim [Graux et al., 2012b]. Dans notre étude illustrative, le modèle ModVégé a été appliqué à une prairie permanente de haute montagne en France de façon à montrer la pertinence d'utiliser un ensemble complémentaire d'indices quantitatifs pour évaluer la vulnérabilité. En effet, même pondéré par la probabilité d'exposition, un indice unique ne peut pas donner une image complète de la vulnérabilité du système. L'étude souligne également l'avantage d'utiliser un plan d'expériences bien pensé pour tenir compte des différents niveaux d'incertitude associée au système étudié. Il est à noter que nous aurions également pu combiner toutes les approches précédentes, afin de tenir compte de toutes les incertitudes. Le principal problème serait alors la taille du plan d'expériences. L'utilisation d'un plan de type hypercube latin pourrait être un bon moyen de réduire le nombre de simulations nécessaires ([McKay et al., 1979]). Pour résumer, l'originalité de cette approche réside dans l'utilisation combinée de plusieurs indices de vulnérabilité et dans l'utilisation de plusieurs plans d'expériences tenant respectivement compte de différents niveaux d'incertitudes. Un autre intérêt de ce travail réside dans le fait qu'à partir des différents

plans d'expériences, nous avons pu proposer, dans le chapitre précédent, des modèles de DOE et ainsi affiner notre métamodèle. Cela nous a également permis de prendre en main et de tester l'outil OpenMOLE par rapport à notre problématique.

4.3. Cas exemple sur carte européenne

De façon à avoir un exemple d'application des indices de vulnérabilité à l'aide d'un modèle de prairie plus complexe, mais également à une échelle différente, nous avons décidé de réaliser une simulation spatialisée des prairies permanentes européennes et de leur appliquer un jeu d'indices de vulnérabilité. Ce travail va nous permettre de montrer que notre proposition de méthode d'analyse de vulnérabilité (démarche, plans d'expériences, jeux d'indices ...) est également opérationnelle dans le cas d'analyse régionalisée.

4.3.1. Simulation à l'échelle européenne

Noter que bien que spatialisées les données européennes n'intègrent pas d'interactions entre pixels. De fait, elles sont plus "pixélisées" que spatialisés au sens classique du terme. Les simulations ont eu lieu sur une grille de 170 x 278 de résolution 0,25 x 0,25 degrés de latitude et longitude, correspondant à [29,125 N ; 71,375 N] et [-23,875 E ; 45,375 E].

Le climat

Les données climatiques utilisées ont été fournies par Christian Beer (Max-Planck Institute for Biogeochemistry, Jena, Allemagne) dans le cadre du projet EU-FP7 CARBO-Extreme⁶⁷ (voir annexe I.i pour une description du projet), et se compose de deux parties. La période de 1901-2010 est composée du climat reconstruit à partir des observations, tandis que la période 2011-2100 suit le scénario d'émission A1B [IPCC 2007a]. Ce scénario traduit une croissance économique forte, une population mondiale atteignant les neuf milliards en 2050 et déclinant ensuite, associée à un développement rapide de nouvelles technologies plus efficaces et à une utilisation équilibrée de toutes les ressources.

Le projet européen EU-FP6 WATCH⁶⁸ a fournit les données journalières pour une échelle de 0,5 degrés [Weedon et al., 2011]. Une phase de régionalisation pour atteindre les 0,25 degrés a été effectuée en se basant sur le jeu de données CRU CL 2.0⁶⁹ [New et al., 2002]. Ce jeu de données contient les grilles mensuelles des données climatiques moyennes observées depuis 1961-1990 et couvrent le globe avec une résolution spatiale de 10'. En plus de l'altitude, huit variables climatiques (mensuelles) sont disponibles :

⁶⁷ <http://www.carbo-extreme.eu/>

⁶⁸ <http://www.eu-watch.org/>

⁶⁹ http://www.cru.uea.ac.uk/~timm/grid/CRU_CL_2_0.html

l'amplitude diurne de la température de l'air, les précipitations, la température moyenne de l'air, la fréquence des jours pluvieux, la fréquence des jours de gel, l'humidité relative, l'ensoleillement et la vitesse moyenne du vent. Les températures minimales et maximales peuvent être déduites de la température moyenne et de l'amplitude thermique diurne.

La période 1979-2010 du climat a été générée en utilisant l'outil ERA-Interim [Dee et al., 2011] pour ajuster les moyennes et écart-types des données de forçage WATCH pour la période 1979-2010. De même que pour la période précédente, une phase de régionalisation a eu lieu à l'aide du CRU CL 2.0 [New et al., 2002].

Pour la période 2011-2100, les données ont été générées à l'aide des résultats obtenus par l'ensemble de modèles climatiques régionaux utilisés dans le projet européen ENSEMBLES⁷⁰. Ces dernières ont été ajustées de façon à faire coïncider les moyennes et écart-types des données de la période 1970-2010 avec les données climatiques WATCH-ERA-Interim.

Les variables climatiques que nous avons utilisées pour les simulations couvrent le rayonnement journalier, la température minimale et maximale de l'air, l'humidité de l'air, la vitesse du vent et les précipitations. Le modèle PaSim nécessite normalement des données horaires, c'est pourquoi une fonctionnalité du modèle permet de prendre en entrée des données journalières et en interne de les désagréger en horaire [INRA-UREP 2011].

Les données de forçage de la concentration atmosphérique en CO₂ sont sur une base annuelle. On ne tient pas compte de la dynamique saisonnière et horaire. Il s'agit en fait d'une valeur identique pour toute l'Europe. Les valeurs de 1901 à 2010 sont basées sur les mesures des carottes de glace et des observations atmosphériques de l'agence américaine "*National Oceanic and Atmospheric Administration*" (NOAA⁷¹). A partir de 2011, les valeurs de concentration en CO₂ atmosphérique sont celles du SRES A1B [IPPC 2007a].

Les propriétés du sol, c'est-à-dire la texture (sable/limon/argile), le pH, la densité apparente, les propriétés hydriques, la profondeur, ... sont issues de la base de données harmonisée mondiale de sol de la FAO (HWSD⁷², "*Harmonized World Soil Database*" [Nachtergaele et al., 2008]). Les données de la base ont été agrégées de façon à avoir les mêmes coordonnées spatiales que la météo.

La gestion

Les données relatives à la gestion des prairies pour l'Europe proviennent du protocole défini dans le projet européens EU-FP7 GHG-Europe⁷³. Il a été conçu à partir des résultats

⁷⁰ <http://www.ensembles-eu.org/>

⁷¹ <http://www.noaa.gov/>

⁷² <http://webarchive.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML/>

⁷³ <http://www.ghg-europe.eu/>

obtenus lors du projet européen EU-FP6 NitroEurope⁷⁴ [communication personnelle d'Adrian Leip, Centre commun de recherche de la Commission européenne]. À partir des données statistiques par région et par pays, trois niveaux d'intensité de gestion (extensive, moyenne et intensive), correspondant aux trois terciles, ont été générés. De façon à rendre la gestion dépendante du climat, les dates de fauches ont été calculées en utilisant des sommes de températures en base 5 °C (c'est-à-dire la somme jour par jour de la température moyenne de l'air au dessus de 5 °C, depuis le 1^{er} janvier). Ainsi la fauche a lieu à 900 °C jour pour la gestion extensive, à 500 °C jour pour la gestion moyenne et à 500 et 1200 °C jour pour la gestion intensive. Aucune fauche n'est réalisée si la somme de température n'est pas atteinte. La fertilisation a lieu trois jours après la fauche. Le pâturage commence 30 jours après la dernière fauche et se finit à la fin de l'année ou à la première période de cinq jours de gel consécutifs. Comme pour la coupe et la fertilisation, si les sommes de température ne sont pas atteintes, alors aucun pâturage n'a lieu.

Les statistiques ne donnant les informations que pour l'année 2010, des règles de transformation ont été appliquées pour recréer la gestion passée à partir des données actuelles :

- De 1901 à 1950 : pour chaque niveau d'intensification, la gestion de type extensive est utilisée. Cependant aucune fertilisation minérale n'est appliquée. Ainsi, seule la fraction de légumineuse diffère entre les niveaux d'intensité ;
- De 1951 à 2010 : cette période se traduit par une intensification progressive de la gestion dans le but d'atteindre en 2010 les valeurs prévues par les statistiques pour tous les niveaux de gestion. Dès le début de la période, la fertilisation minérale apparaît dans la gestion, avec une quantité égale à celle prévue pour la gestion extensive. Puis, il y a une augmentation linéaire des quantités de fertilisant apporté, une intensification linéaire du chargement animal et une avancée progressive de la date de fauche (de 900 °C jour vers 500 °C jour). A partir de 1981, la deuxième date de fauche apparaît pour la gestion intensive ;
- De 2010 à 2100 : la gestion est gardée constante par rapport aux statistiques observées. Devant la gamme possible de scénarios économiques et surtout la difficulté d'en reporter l'impact sur la gestion, nous avons préféré conserver une gestion constante.

Il est à noter que devant l'absence ou la trop mauvaise qualité des données statistiques initiales, certains pays de l'Union Européenne ont été exclus de la simulation.

⁷⁴ <http://www.nitroeurope.eu>

Réalisation et analyse des simulations

Le format de stockage des données d'entrée est le NetCDF⁷⁵ (voir annexe II). Le choix du format de données a été imposé par le projet européen EU-FP7 CARBO-Extreme⁷⁶. Le modèle PaSim n'étant pas actuellement encore entièrement opérationnel pour gérer directement ce format d'entrée, une solution temporaire a été utilisée. Les données en NetCDF (~130 Go) ont été désagrégées en texte (sous la forme classique des fichiers d'entrée de PaSim [INRA-UREP 2011]) pour les pixels de prairie, puis transférées sur le cluster du LIMOS, où les simulations ont été réalisées (~1 an de calcul mono-cœur sur CPU « moderne » pour un niveau de gestion). Enfin les résultats, préalablement compressés pour réduire le temps de transfert, ont été rapatriés en local afin d'être retransformés en fichier NetCDF.

Le modèle biogéochimique de prairie PaSim [Riedo et al., 1998] (décrit en 1.5) a donc été lancé sur cette grille de climat-sol-gestion européenne. Après une mise à l'équilibre sur les cinq premières années climatiques (voir 4.4 pour le choix du nombre d'années), PaSim 5.3 a simulé la prairie de 1901 à 2100. Pour des raisons de simplification, le manque de données et les incertitudes qu'aurait entraîné l'extrapolation des données manquantes, l'ancienne version du module animal (la version 4.5 de PaSim) a été utilisée. Le paramétrage de base pour les prairies permanentes ([Riedo et al, 1998]) a été utilisé, ainsi qu'une fraction de légumineuse constante dans le temps, fonction de l'intensité de la gestion en 2010.

Pour analyser les résultats obtenus, nous avons définis trois périodes climatiques :

- la période de référence, de 1980 à 2009, qui correspond au climat actuel ;
- la période "futur proche", de 2020-2049, qui correspond à la prochaine période climatique ;
- la période "futur lointain", de 2070-2099, qui correspond à la période climatique de la fin du siècle.

⁷⁵ <http://www.unidata.ucar.edu/software/netcdf/>

⁷⁶ <http://www.carbo-extreme.eu/>

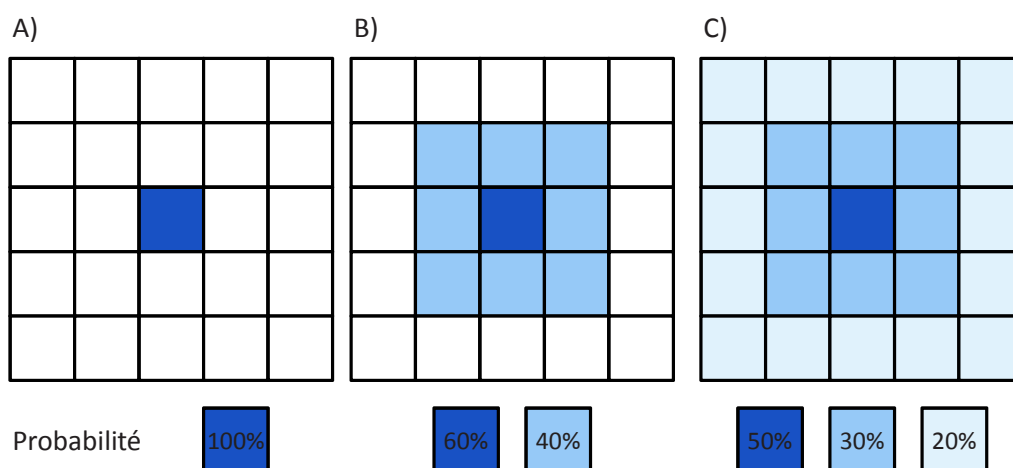


Figure 49 Calcul de vulnérabilité en tenant compte A) juste du pixel, B) du premier niveau de pixels voisins, C) des pixels voisins jusqu'à deux niveaux.

À partir de la période de référence, nous avons construit un seuil statistique. Ainsi, pour les 30 ans de la période de référence, nous avons considéré les valeurs du pixel, des huit pixels directement autour et des seize pixels à deux pixels de distance (Figure 49-C). Considérant que la série du pixel a une probabilité d'occurrence de 50%, que le premier carré en a une de 30% et que le deuxième en a une de 20%, nous avons fixé un seuil statistique à 10%, considérant ainsi que toute valeur en dessous représente une situation vulnérable. L'idée principale derrière ce calcul est que le climat en un pixel n'est qu'un tirage aléatoire de ce qui aurait pu se produire. Or les pixels voisins sont globalement proches⁷⁷ en termes de climat, sol et gestion, on peut donc les considérer comme des tirages aléatoires (bien que moins probable) de ce qui aurait pu se passer à ce pixel donné. Ce mode de calcul permet d'être plus représentatif d'une zone mais en contrepartie il "lisse" la variabilité spatiale. À l'aide de ce principe de calcul, les indices de vulnérabilité présentés précédemment (Table 10) ont été calculés pour les variables de rendement et de photosynthèse ("*Gross Primary Production*", GPP) pour les trois périodes climatiques, puis les indices ont été normés par la valeur de la période de référence.

4.3.2. Résultats

Différentes valeurs de seuil statistique ont été testées (5%, 10%, 15% et 20%), de façon à vérifier la robustesse des résultats ainsi obtenus (Figure 50). Même si les valeurs diffèrent, le patron spatial est conservé. De plus, comme attendu la valeur du seuil change que l'on considère un seuil de 5% à 20%, en revanche le patron spatial est également largement conservé (Figure 51). Noter que l'indice de Luers est indépendant du seuil considéré. En effet l'action de normer par la valeur de la période de référence a pour effet de faire disparaître le seuil du calcul.

⁷⁷ Noter qu'un pixel représente approximativement une zone de 25km x 25km

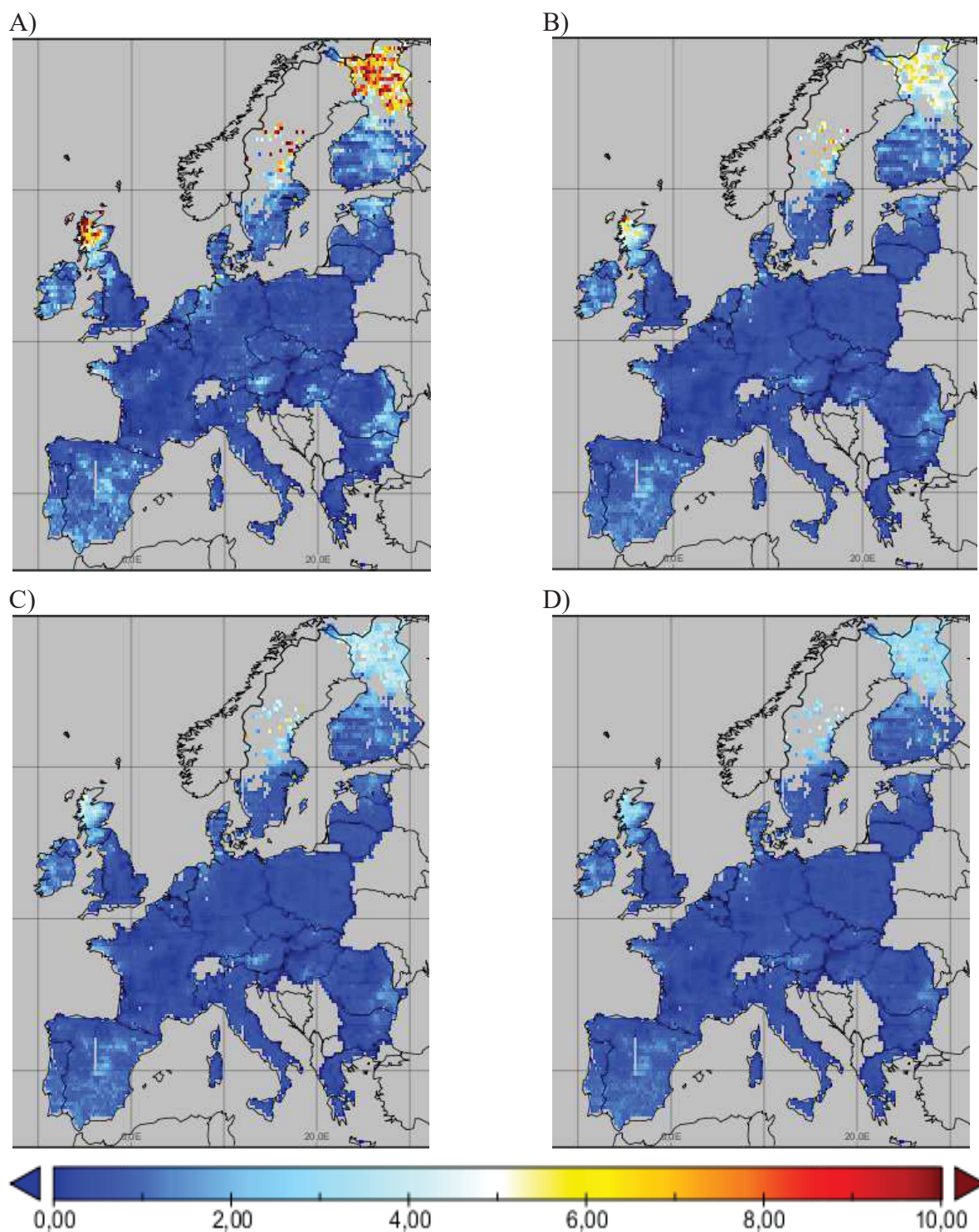


Figure 50 Calcul de la vulnérabilité des prairies en Europe en utilisant l'indice de vulnérabilité proportionnelle pour la GPP annuelle (*Gross Primary Production*), pour la gestion Moyenne, pour la période climatique de "Futur Proche" (2020-2049), pour un seuil A) de 5%, B) de 10%, C) de 15%, D) de 20%. Les indices ont été calculés en utilisant jusqu'à deux niveaux de pixels voisins, puis normés par les valeurs de la période de référence (1980-2009). Plus la valeur de l'indice est haute plus le système est vulnérable.

Seuil de GPP à 5% de probabilité

Seuil de GPP à 20% de probabilité

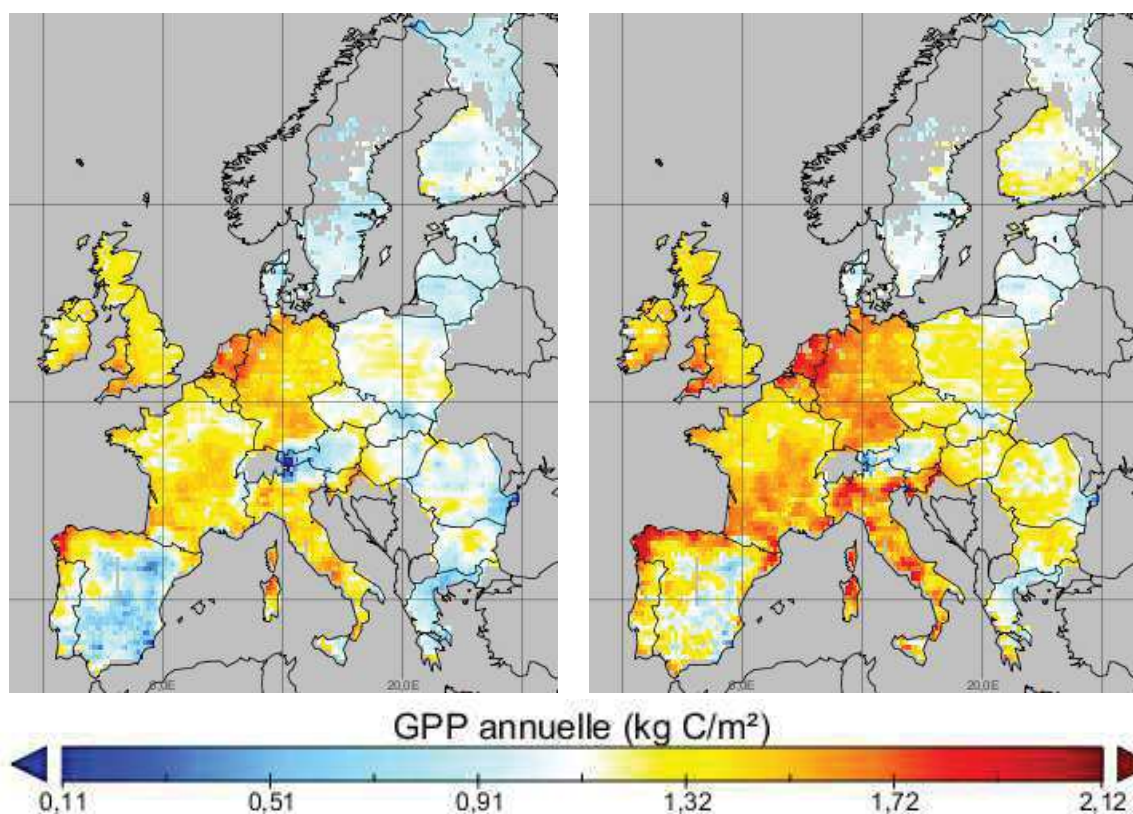


Figure 51 Seuils de vulnérabilité pour la GPP annuelle (*Gross Primary Production*) pour la gestion Moyenne, calculés en utilisant jusqu'à deux niveaux de pixels voisins, pour un seuil statistique de 5% à gauche et de 20% à droite.

Comme attendu le fait de tenir compte des voisins du pixel (Figure 49) a pour effet de "lisser" la carte (Figure 52), mais il permet également de la rendre moins sensible au seuil de vulnérabilité [données non présentées].

L'intensification de la gestion semble avoir globalement pour effet d'augmenter la vulnérabilité de la photosynthèse au changement climatique. On observe en effet une augmentation des indices moyennés sur l'Europe en passant d'une gestion extensive à intensive [tableau non présenté]. La période de "futur proche" (FP) se caractérise pour la photosynthèse (GPP) par une vulnérabilité plus élevée que la période de "futur lointain" (FL) et ce pour tous les indices (par exemple l'indice "vulnerability severity" pour la période FP est de $2,33 \pm 0,089$ et de $0,56 \pm 0,042$ pour la période FL). A l'exception de l'indice de Luers et "vulnerability severity" pour la gestion intensive, la période de futur lointain se caractérise dans tous les autres cas par une baisse en moyenne (sur les pixels) de la vulnérabilité de la GPP par rapport à la période de référence.

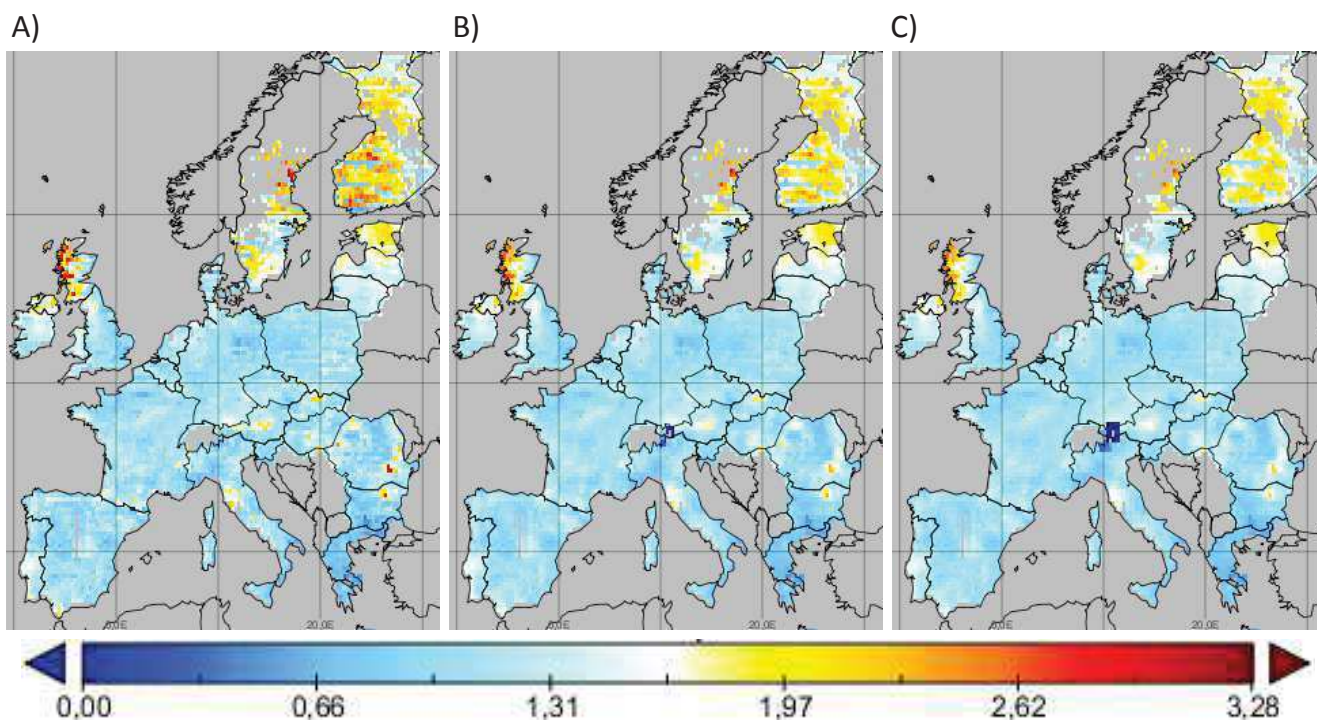


Figure 52 Calcul de la vulnérabilité des prairies en Europe en utilisant l'indice de Luers pour la GPP annuelle (*Gross Primary Production*), pour la gestion Moyenne, pour la période climatique de "Futur Proche" (2020-2049), en tenant compte A) juste du pixel, B) du premier niveau de pixels voisins, C) des pixels voisins jusqu'à deux niveaux. Les indices ont été normés par les valeurs de la période de référence (1980-2009). Plus la valeur de l'indice est haute plus le système est vulnérable.

La vulnérabilité, normée par la valeur de la période de référence, du rendement exprimé en biomasse (*i.e.* kg MS/m²) ou en carbone (*i.e.* kg C/m²) est similaire [données non présentées]. Nous allons donc simplement considérer les résultats pour le rendement exprimé en biomasse. Noter que pour certaines zones, la variable rendement n'a pas été calculée (Figure 53). En effet dans certaines zones (comme une partie de l'Espagne), le nombre de fauches réalisées par rapport à la gestion prescrite est trop faible pour conserver la robustesse des indices. L'absence de valeur caractérise en fait des zones très vulnérables.

La gestion a une grande influence sur la vulnérabilité du rendement au changement climatique. Ainsi la prairie est, en moyenne sur l'espace, moins vulnérable dans les périodes climatique futures que dans la période actuelle pour la gestion la plus intensive. À l'inverse la gestion moyenne se traduit, en moyenne sur les pixels, par une augmentation de la vulnérabilité. De plus les patrons spatiaux sont très influencés par le niveau de gestion, comme l'illustre la Figure 53, où par exemple la partie Ouest de la France devient très vulnérable pour la gestion moyenne et pour la période "futur lointain" tandis que sous gestion intensive, cette même zone est moins vulnérable que la période de référence.

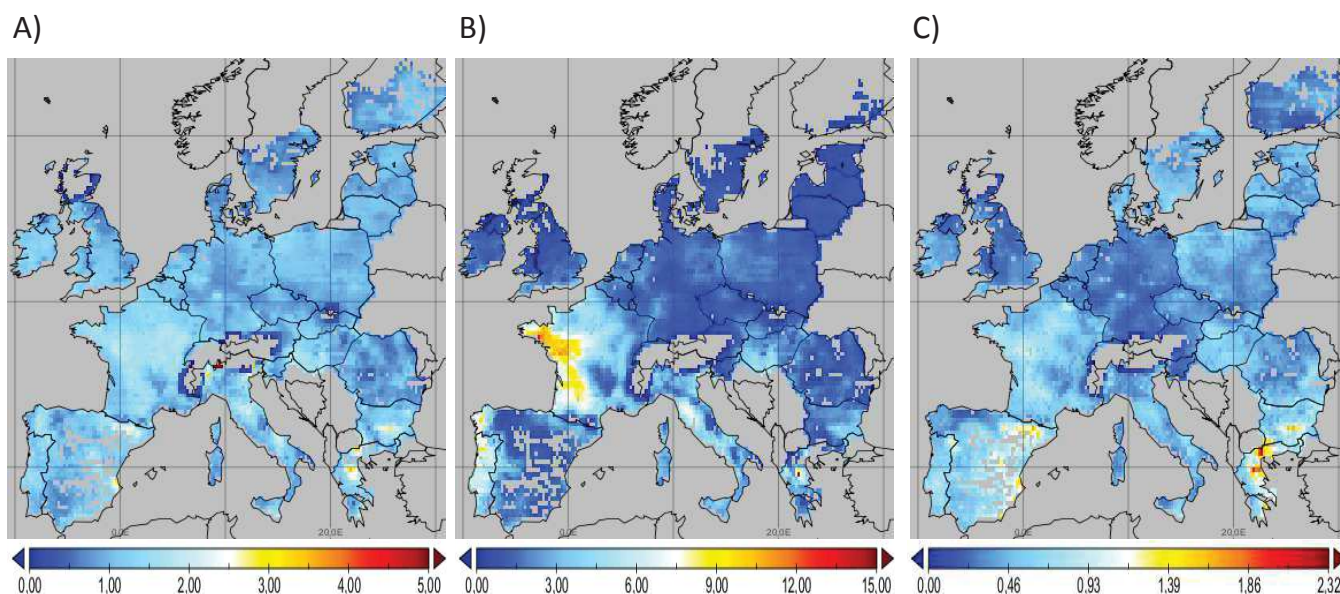


Figure 53 Vulnérabilité en utilisant l'indice de Luers pour le rendement pour la période climatique de "futur lointain" (2070-2099), pour A) la gestion extensive, B) la gestion moyenne, C) la gestion intensive. Les indices ont été calculés en utilisant jusqu'à deux niveaux de pixels voisins, puis normés par les valeurs de la période de référence (1980-2009). Plus la valeur de l'indice est haute plus le système est vulnérable.

4.3.3. Discussion

Compte-tenu du fait que les trois gestions utilisées sont représentatives des gestions possibles à un endroit donné, nous aurions pu envisager de combiner cette information lors du calcul de la vulnérabilité de façon à avoir un seul indice agrégé. La méthode la plus simple pour cela aurait été de faire la moyenne des indices pour les trois gestions. Les résultats pourraient également être analysés à l'aide d'une étude statistique plus poussée qui permettrait de mieux quantifier les patterns régionaux (par exemple [Marín et al., 2012]). De plus, il aurait pu être intéressant de vérifier la robustesse des indices ainsi calculés par rapport aux probabilités utilisées dans le calcul (*i.e.* 50% pour le pixel central, 30% pour le premier niveau et 20% pour les pixels à une distance de 2).

Ces résultats ne tiennent compte que d'une faible fraction des incertitudes. Seules les incertitudes sur la gestion, à travers trois séries, ont été prises en compte. De plus, un seul climat a été utilisé, c'est-à-dire un seul modèle climatique, un seul scénario SRES et une seule méthode d'initialisation du modèle climatique. De même, une seule initialisation du modèle d'impact (PaSim) a été utilisée (*i.e.* une mise à l'équilibre sur cinq ans), alors qu'il s'agit d'un facteur très influent sur les résultats d'évolution des stocks de carbone. De façon à également tenir compte des erreurs structurelles du modèle d'impact, il aurait fallu tester plusieurs modèles. Pour tenir compte de l'erreur due au paramétrage de PaSim, nous aurions pu utiliser plusieurs jeux de paramètres (potentiellement issus d'une calibration

bayésienne préalablement effectuée avec le modèle [Ben Touhami et al., 2012a] [Ben Touhami et al., 2012b]). De même, cette simulation ne tient pas compte de la dérive de la végétation (*i.e.* du changement d'espèces végétales) au cours du temps, à cause du changement de climat et de gestion.

Cependant, cette étude n'a pas vocation à tenir compte de toutes les sources importantes d'incertitudes, elle sert principalement à valider et à adapter la proposition de méthode d'analyse de vulnérabilité dans un contexte régionalisé, notamment en mettant en avant la pertinence de l'utilisation d'un ensemble d'indices et en montrant les différences avec l'évaluation de la vulnérabilité en un site donné. En effet, à la différence d'une modélisation à l'échelle d'un site, le nombre de simulations pour couvrir les incertitudes, même avec du calcul haute performance (*i.e.* calculs distribués sur grille ou cluster), reste prohibitif. Une solution alternative pourrait être :

1. Sélectionner un échantillon de points sur la carte, de manière relativement uniforme, un plan par hypercube latin (LHD) ou un plan minimax/maximin pourrait le permettre. Dans le choix des pixels, la distance entre pixels pourrait être considérée, mais également la gestion et l'environnement (climat, altitude, végétation) ;
2. Effectuer pour ses points une analyse de vulnérabilité, tenant compte d'une grande part des incertitudes. Pour cela un plan d'expériences combinant les différentes sources d'incertitude pourrait être utilisé pour obtenir la vulnérabilité sans adaptation ;
3. Ensuite, des surfaces de réponses de la vulnérabilité en fonction de la gestion pourraient être régressées ;
4. À partir de ces dernières, un minimum de vulnérabilité pourrait être cherché ;
5. La robustesse du minimum ainsi trouvé devra être évaluée sur un nouveau design, nous permettant ainsi d'évaluer la vulnérabilité avec adaptation. Si nécessaire (*i.e.* minimum non robuste), on recommence à partir de l'étape 3 ;
6. Enfin une dernière étape de Krigeage [Sacks et al., 1989] pourrait permettre d'extrapoler de manière plus précise la vulnérabilité avec et sans adaptation sur la carte.

Cette méthode permettrait de réaliser notre démarche d'analyse de vulnérabilité dans un contexte spatialisé, c'est-à-dire une méthode spatialisée d'analyse de vulnérabilité avec et sans adaptation et tenant compte de multiples sources d'incertitudes.

Même si dans le cadre de cette thèse, nous n'avons pas eu le temps de réaliser cette méthode, certaines étapes de cette méthode ont cependant été réalisées. Ainsi la partie 4.5,

après la validation de certaines hypothèses par la partie 4.4, détaillera la construction de surface de réponse (étape 3 de la méthode décrite précédemment). Enfin la partie 4.6 décrit également la recherche et l'évaluation du minimum de vulnérabilité (étape 4 et 5 de la méthode décrite précédemment).

4.4. Compromis entre précision et temps de calcul de la matière organique du sol à l'équilibre

La matière organique du sol (MOS) est le principal réservoir de carbone (C) et d'azote (N) organique dans les prairies et joue un rôle-clé dans l'atténuation des émissions des gaz à effet de serre (GES). Les modèles biogéochimiques classiques intègrent une vision mécaniste de la dynamique de la matière organique du sol et fournissent des informations sur les compartiments de matière organique [Guo et Giffort, 2002]. Dans la plupart des modèles biogéochimiques, la MOS est divisée en compartiments conceptuels/fonctionnels avec des temps de résidences ("*turnover*") différents variant de quelques mois pour les produits labiles de la décomposition microbiennes à plusieurs milliers d'années pour les substances organiques résistantes (par exemple [Yadav et Malanson, 2007]). Les temps de résidence varient en fonction de la température et de l'humidité du sol et d'autres variables. Il est probable que la matière organique du sol soit vulnérable au réchauffement climatique futur [IPCC 2001], ce qui souligne la nécessité d'améliorer notre compréhension de la dynamique de la matière organique à l'aide de simulations réalisées à partir de modèles biogéochimiques du cycle du carbone qui sont en mesure de faire des simulations à long terme.

Dans les modèles biogéochimiques, les compartiments de carbone et d'azote sont habituellement initialisés à l'équilibre à l'aide de simulations reproduisant l'histoire du climat et des pratiques agricoles. Cependant, comme les pratiques agricoles et climatiques changent, les sols ne sont généralement pas à l'équilibre et par conséquent les simulations de mise à l'équilibre ne sont pas toujours valables ([Wutzler et Reichstein, 2007], [Foereid et al, 2012]). La question de l'initialisation hors équilibre est difficile. En effet, par exemple, pour une quantité totale de MOS connue, il faut spécifier des valeurs pour chacun des compartiments, et une infinité de solutions existe. Peu d'études se concentrent sur cette question, on peut cependant citer [Yeluripati et al., 2009] et [Basso et al., 2011]. Les simulations de mise à l'équilibre sont très chronophages. La méthode la plus courante pour réaliser une mise à l'équilibre consiste à répéter le climat et la gestion jusqu'à stabilisation du système. Afin de réduire les temps de calcul, quelques méthodes existent telles que dans [Thornton et Rosenbloom, 2005], [Martin et al., 2007], [Lardy et al., 2011] et [Xia et al., 2012]. Avec ces approches, des questions demeurent cependant, comme le nombre d'années à prendre en compte dans le cycle de données météorologiques répété pour obtenir la matière organique à l'équilibre. D'un côté, plus on utilise d'années climatiques,

plus les résultats seront précis, mais d'un autre côté, plus nous utilisons d'années plus le temps de calcul pour la recherche de l'équilibre est important. En outre, si l'on suppose qu'un ensemble restreint d'années est suffisant pour obtenir la plupart des informations climatiques pour la matière organique du sol à l'équilibre, on peut se demander si des changements dans l'ordre des années se traduiront par des changements visibles dans le résultat. En effet, nous savons que pour la croissance de la végétation l'ordre des années a une importance. Si, par exemple, nous avons deux années de sécheresse sur cinq ans, le fait qu'elles soient consécutives va augmenter les effets négatifs sur la végétation. Toutefois, considérant de longues périodes, la végétation des prairies est résiliente, et les compartiments de matière organique les plus grands et les plus lents ont des temps de résidence élevés (siècles à millénaires). C'est pourquoi la matière organique du sol pourrait ne pas être sensible à l'ordre des années.

La section suivante décrit le plan d'expériences utilisé pour étudier le comportement de la matière organique du sol à l'équilibre par rapport au nombre et à l'ordre des années climatiques utilisées dans la mise à l'équilibre. Puis nous analyserons les erreurs relatives induites par des cycles météorologiques plus courts ou des ordres d'années différents, et enfin nous discuterons les résultats.

4.4.1. Plan d'expériences

Le modèle biogéochimique de prairie, PaSim, a été utilisé pour réaliser les simulations de mise à l'équilibre du sol. Ce modèle a précédemment été décrit dans la section 1.5. La version 5.3 de PaSim a été utilisée dans cette étude. Toutefois, pour des raisons de simplification de données d'entrée, l'ancienne version du module animal (la version 4.5 de PaSim) a été utilisée. La méthode algébrique, développée pour réduire le temps de calcul pour la mise à l'équilibre [Lardy et al., 2011b], a été utilisée dans cette étude. Cette méthode permet, grâce à une suite de matrices, de proposer des estimations de la matière organique du sol à l'équilibre. Lorsque la différence entre les deux solutions candidates est inférieure à 0,1%, l'équilibre est considéré comme atteint.

Dans le but d'obtenir des informations sur le comportement de la MOS à l'équilibre par rapport au nombre et à l'ordre des années climatiques utilisées dans le cycle météorologique, nous avons réalisé un simple plan factoriel complet. Les mises à l'équilibre ont été effectuées sur les conditions climatiques françaises, les sols et la gestion telle que définis dans [Graux et al., 2012a]. Quatre prairies ont été simulées (Table 11) pour des prairies de monocultures semées (tous les cinq ans) de *Lolium perenne* L. (SL) ou *Festuca arundinacea* Schreb. (SF) et des prairies permanentes intensives (PI) et extensives (PE). Les prairies temporaires SL et SF ont été fauchées quatre fois par an, du printemps à l'automne. Les prairies permanentes ont été fauchées deux fois au printemps, suivie par deux périodes de pâturage de 15 jours par des vaches laitières. Une fertilisation minérale

par nitrate d'ammonium a été appliquée en plusieurs fois (un jour après chaque fauche) pour une quantité totale de 200 kg N ha⁻¹ an⁻¹, à l'exception des prairies permanentes extensives (PE) qui n'ont pas été fertilisées. On suppose que les prairies PE et PI contiennent respectivement 20% et 5% de légumineuses. Pour la simulation des prairies semées, le labour du sol, ayant lieu tous les cinq ans, augmente le taux de décomposition de la matière organique du sol. Les simulations ont été réalisées, soit sur un sol brun modérément drainant, spécifique des grandes cultures (de 1,40 m de profondeur), soit sur un sol calcaire brun spécifique des prairies (0,60 m de profondeur).

Table 11. Description des quatre systèmes de prairie simulés : prairie temporaire semée en monoculture (SL et SF) et prairies permanentes semi-naturelles extensive et intensive (PE et PI).

Systèmes de prairie	Temporaire / Semée		Permanente	
	Ray-grass anglais (SL)	Fétuque élevée (SF)	Extensive (PE)	Intensive (PI)
Végétation	monoculture		plurispécifique	
	<i>Lolium perenne</i> L.	<i>Festuca arundinacea</i> Schreb.	Avec 20% <i>Trifolium repens</i> L.	Avec 5% <i>Trifolium repens</i> L.
Fertilisation (kg N ha ⁻¹)	200		0	200
Dates de fauches	15 avril		15 avril	
	30 juin		1 juin	
	15 août		-	
	15 octobre		-	
Périodes de pâturage	-		20 juillet au 5 août 15 octobre au 1 Novembre	
Chargement animal (UGB ha ⁻¹)	-		0,8	1,5

Douze sites ont été utilisés (Table 12), pour leur représentativité de la diversité des climats français actuels, notamment des climats du littoral, semi-continental, montagnard et subméditerranéen [Graux et al., 2012a]. Pour chaque site, un ensemble de 30 années observées à partir de 1970 (ou plus tard si non disponible) à 1999 a été utilisé. Afin d'éviter l'effet du CO₂ sur la MOS, la concentration atmosphérique en CO₂ a été maintenue à 365 ppm (valeur observée en 1999).

Table 12. Géo-localisation et type de climat pour les sites utilisés dans cette étude. Pour les détails sur la classification des types de climats, voir [Graux et al., 2012a]. Les valeurs des moyennes de température de l'air et de somme des précipitations concernent la période des années observables.

Site	Latitude	Longitude	Altitude (m)	Type de Climat	Précipitations (mm)	Température de l'air (°C)
Avignon	43° 54' N	04° 54' E	37	Subméditerranéen, semi-aride à aride	702	14,0
Bordeaux	44° 49' N	00° 41' W	20	Océanique, sub-humide à semi-aride	886	13,4
Colmar	48° 04' N	07° 20' E	202	Semi-continentale, semi-aride à aride	541	10,3
Dijon	47° 16' N	05° 05' E	211	Semi-continentale, sub-humide	663	10,2
Lusignan	46° 25' N	00° 07' E	150	Océanique, sub-humide à semi-aride	820	11,5
Mirecourt	48° 18' N	06° 08' E	265	Semi-continentale, humide à sub-humide	877	9,2
Mons en Chaussées	49° 87' N	03° 01' E	85	Océanique, semi-aride	617	10,5
Rennes	48° 06' N	01° 42' W	35	Océanique, sub-humide à semi-aride	727	11,4
Saint-Etienne	45° 32' N	04° 17' E	400	Semi-continentale, sub-humide à semi-aride	707	11,1
Clermont-Theix	45° 43' N	03° 01' E	890	Mountagnard, humide à sub-humide	774	7,9
Toulouse	43° 33' N	01° 29' E	150	Semi-continentale, semi-aride à aride	710	13,1
Versailles	48° 48' N	02° 08' E	125	Continental, sub-humide à semi-aride	651	10,6

Pour chaque combinaison de site, de sol et de gestion, une mise à l'équilibre a été réalisée avec un cycle météorologique de une à 30 années, grâce à la méthode algébrique décrite dans [Lardy et al., 2011b]⁷⁸. La valeur totale de carbone organique du sol pour un cycle de 30 ans a été considérée comme la valeur de référence, c'est-à-dire comme représentative du carbone organique du sol à l'équilibre pour la période climatique. Les valeurs de carbone organique du sol pour des périodes plus courtes ont été comparées à cette valeur de référence. Une autre expérimentation à l'aide d'un cycle météorologique de trois ans a été effectuée afin d'évaluer la sensibilité à l'ordre des années de la MOS à l'équilibre. Pour chaque site, sol et gestion, des simulations de mise à l'équilibre ont été effectuées avec les six combinaisons d'ordre d'années climatiques. Les erreurs induites par différents ordres d'années ont été estimées par rapport à la valeur de matière organique du sol obtenue avec l'ordre climatique observé.

⁷⁸ Ce travail a été publié pendant la thèse mais réalisé antérieurement

4.4.2. Résultats

Les simulations furent réalisées sur le cluster de 256 cœurs de calcul au LIMOS⁷⁹, Clermont-Ferrand (France). La question du nombre d'années à utiliser dans le cycle climatique pour la recherche de l'équilibre a été analysée par rapport aux pratiques de gestion (Table 13) et par rapport aux conditions climatiques locales (c'est-à-dire les sites) (Table 14).

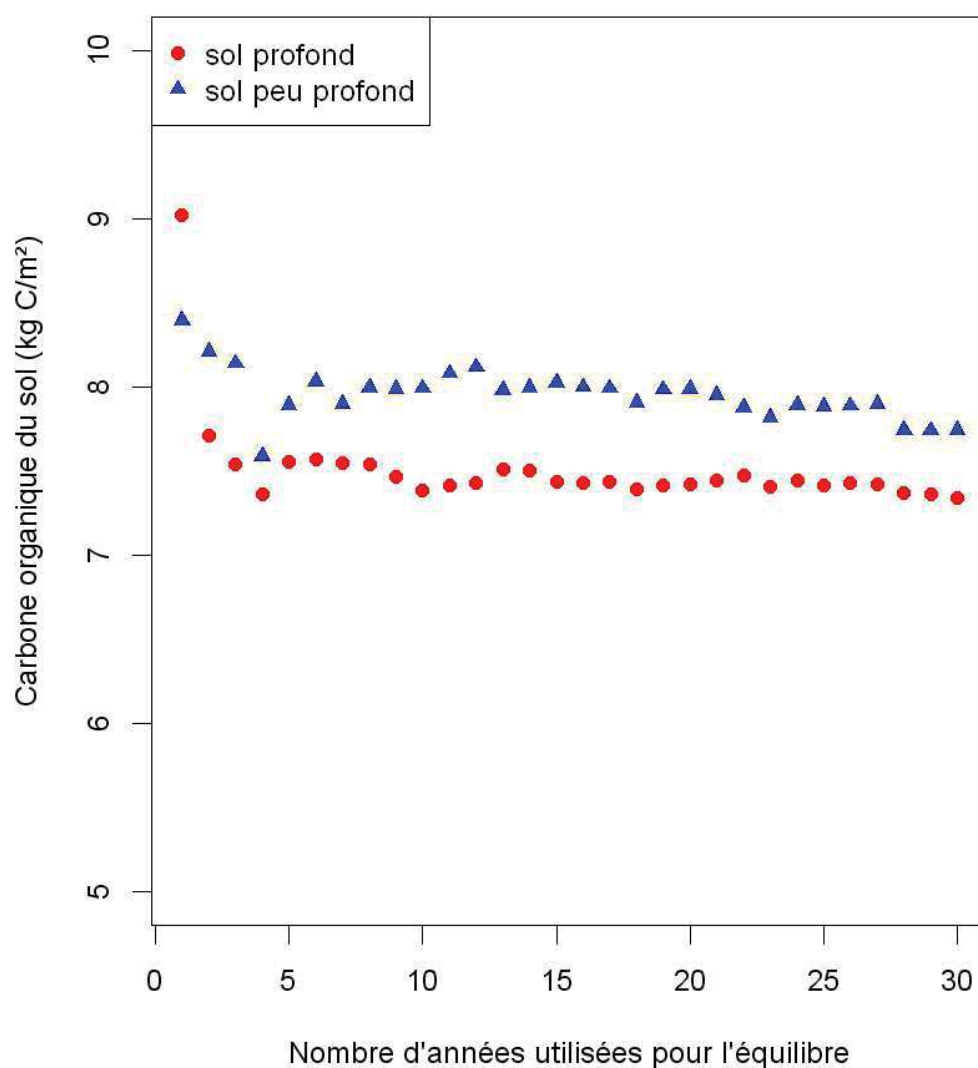


Figure 54 Comparaison de matière organique du sol à l'équilibre en fonction du nombre d'années utilisées dans le cycle climatique, pour une prairie temporaire en monoculture de *Festuca arundinacea*. Schreb. (SF) à Bordeaux.

Les résultats montrent une convergence logique des quantités de carbone organique du sol pour une augmentation du nombre d'années dans le cycle climatique, comme l'illustre la Figure 54 et la Figure 55. Les propriétés du sol ont une influence sur les quantités de

⁷⁹ <http://limos.isima.fr>

MOS, cependant l'influence climatique semble suivre la même tendance indépendamment du type de sol, comme le montre l'exemple de la Figure 54 et de la Figure 55.

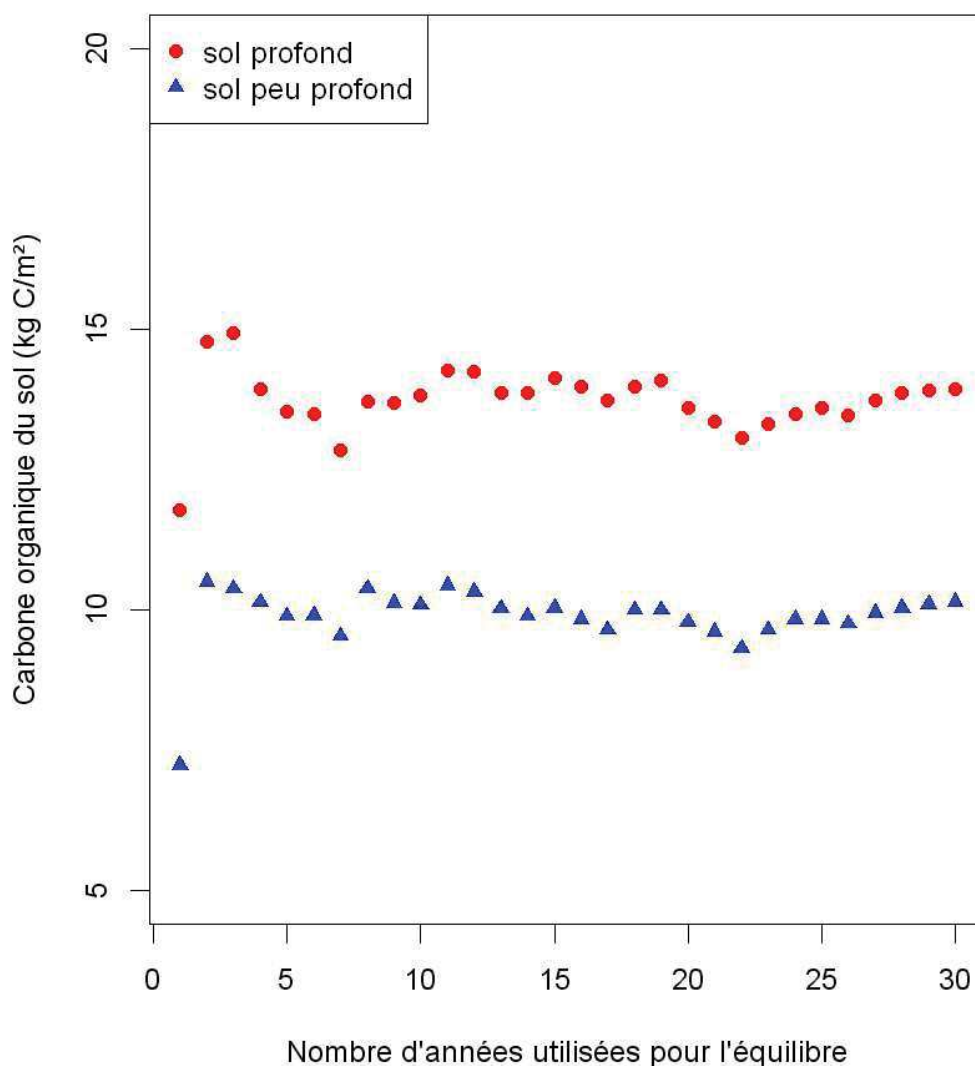


Figure 55 Comparaison de matière organique du sol à l'équilibre en fonction du nombre d'années utilisées dans le cycle climatique, pour une prairie permanente extensive (PE) à Bordeaux.

La Table 13 montre que le type de prairies et les pratiques de gestion sont des facteurs-clé de la sensibilité du carbone organique du sol au nombre d'années climatiques. En effet, les erreurs pour les prairies fauchées uniquement (SL et SF) sont plus faibles (environ la moitié) que pour les prairies qui sont fauchées et pâturées (PE et PI). Ceci suggère que la principale source de différences sur la sensibilité ne viendrait pas des paramètres de la végétation (qui sont maintenus constants durant la simulation), mais des effets du pâturage. L'activité de pâturage augmente probablement la difficulté du modèle à converger, en raison du prélèvement de l'herbe et des déchets animaux, qui jouent un rôle important sur les cycles du carbone et de l'azote [Vuichard et al., 2007a]. Il est intéressant de noter que pour les prairies semées et uniquement fauchées, on peut utiliser seulement trois années

climatiques pour obtenir une approximation correcte de la matière organique du sol à l'équilibre ($3,53\% \pm 0,71\%$).

Table 13. Erreurs relatives (moyenne \pm erreur standard) pour la quantité totale de carbone organique du sol prédit avec des cycles météorologiques de longueurs différentes pour différents systèmes prairiaux sur 12 sites.

Nombre d'années dans le cycle climatique	Semée en Fétuque (SF)	Semée en Ray-grass (SL)	Permanente Extensive (PE)	Permanente Intensive (PI)	Tous les systèmes
1	$5,68\% \pm 2,24\%$	$5,8\% \pm 2,78\%$	$18,47\% \pm 6,79\%$	$15,76\% \pm 5,75\%$	$11,43\% \pm 2,43\%$
2	$4,77\% \pm 1,35\%$	$4,88\% \pm 1,55\%$	$15,79\% \pm 3,99\%$	$13,59\% \pm 3,54\%$	$9,76\% \pm 1,50\%$
3	$3,3\% \pm 0,96\%$	$3,77\% \pm 1,07\%$	$13,59\% \pm 2,01\%$	$11,76\% \pm 1,9\%$	$8,1\% \pm 0,90\%$
4	$3,6\% \pm 0,94\%$	$4,55\% \pm 1,05\%$	$8,3\% \pm 1,63\%$	$7,28\% \pm 1,68\%$	$5,93\% \pm 0,70\%$
5	$3,83\% \pm 0,83\%$	$4,1\% \pm 0,92\%$	$8,46\% \pm 1,67\%$	$7,42\% \pm 1,54\%$	$5,95\% \pm 0,67\%$
6	$3,62\% \pm 0,75\%$	$3,97\% \pm 0,78\%$	$7,91\% \pm 1,75\%$	$7,02\% \pm 1,46\%$	$5,63\% \pm 0,65\%$
7	$3,66\% \pm 0,64\%$	$4,16\% \pm 0,83\%$	$8,39\% \pm 1,66\%$	$7,73\% \pm 1,50\%$	$5,98\% \pm 0,65\%$
8	$3,24\% \pm 0,71\%$	$3,77\% \pm 0,85\%$	$10,86\% \pm 1,55\%$	$10,1\% \pm 1,32\%$	$6,99\% \pm 0,67\%$
9	$3,05\% \pm 0,60\%$	$3,37\% \pm 0,70\%$	$10,73\% \pm 1,37\%$	$10,07\% \pm 1,14\%$	$6,8\% \pm 0,62\%$
10	$2,60\% \pm 0,59\%$	$2,68\% \pm 0,60\%$	$10,28\% \pm 1,27\%$	$9,74\% \pm 1,04\%$	$6,32\% \pm 0,59\%$
15	$2,38\% \pm 0,48\%$	$2,14\% \pm 0,56\%$	$8,76\% \pm 1,45\%$	$8,47\% \pm 1,12\%$	$5,44\% \pm 0,59\%$
20	$1,57\% \pm 0,43\%$	$1,28\% \pm 0,41\%$	$3,75\% \pm 1,83\%$	$3,83\% \pm 1,38\%$	$2,61\% \pm 0,60\%$
29	$-0,06\% \pm 0,12\%$	$-0,09\% \pm 0,11\%$	$0,44\% \pm 0,20\%$	$0,39\% \pm 0,19\%$	$0,17\% \pm 0,08\%$

En comparant le nombre d'années de simulation requis pour la mise à l'équilibre (*i.e.* nombre de répétitions du cycle climatique x longueur du cycle climatique) [données non présentées] et les erreurs associées, nous pouvons conclure qu'un cycle de cinq ans offre un bon compromis entre les erreurs ($5,95\% \pm 0,67\%$) et le temps de calcul (~ 24 minutes avec $59,3 \pm 1$ années de la simulation), pour tout type de prairie. La sensibilité de la matière organique du sol par rapport au nombre d'années nous donne des informations sur le comportement de la matière organique du sol par rapport au climat. En effet, si cinq années suffisent pour obtenir la majeure partie de l'information climatique pour la quantité de carbone organique du sol, alors nous savons que les stocks potentiels de matière organique du sol (c'est-à-dire la matière organique à l'équilibre) changeront lentement au cours du temps. Cela signifie que quelle que soit la variabilité présente au sein de la période climatique (généralement considérée de 30 ans), nous avons besoin de passer à la période climatique suivante pour avoir un changement visible dans l'état d'équilibre de la matière organique du sol.

La variabilité climatique est très dépendante du site et peut légèrement biaiser la représentativité d'une petite série d'années pour la mise à l'équilibre de la matière organique du sol (Table 14). Tester toutes les combinaisons d'ordre pour un ensemble de trois années climatiques sur deux sites d'exemple, a montré que l'influence de l'ordre des

années est négligeable (erreur < 0,5%) sur la matière organique à l'équilibre (Table 15). Cela signifie que le carbone organique du sol est résilient face à l'hétérogénéité climatique au sein de la même période climatique.

Table 14. Erreurs relatives (moyenne \pm erreur standard) pour la quantité totale de carbone organique du sol prédit avec des cycles météorologiques de longueurs différentes pour 12 sites.

Longueur du cycle climatique (nombre d'années)	1	2	3	4	5	10	15	20
Avignon	-27,53% \pm 4,47%	-9,19% \pm 2,64%	2,95% \pm 2,16%	4,79% \pm 1,48%	7,90% \pm 1,35%	5,58% \pm 0,61%	1,71% \pm 0,37%	-0,28% \pm 1,09%
Bordeaux	-3,6% \pm 20,32%	5,82% \pm 2,96%	4,78% \pm 2,25%	-0,09% \pm 1,68%	0,21% \pm 3,13%	0,77% \pm 1,67%	1,44% \pm 1,70%	-0,02% \pm 2,61%
Clermont	23,12% \pm 18,7%	21,43% \pm 17,22%	11,84% \pm 16,81%	4,81% \pm 8,63%	5,59% \pm 9,12%	3,57% \pm 6,2%	3,11% \pm 5,39%	-0,03% \pm 2,10%
Colmar	36,84% \pm 30,33%	17,61% \pm 13,32%	12,64% \pm 5,49%	14,81% \pm 7,88%	10,11% \pm 3,87%	10,37% \pm 9,43%	6,05% \pm 4,34%	0,72% \pm 1,53%
Dijon	28,99% \pm 28,04%	5,78% \pm 6,57%	5,88% \pm 8,22%	1,61% \pm 4,25%	4,69% \pm 5,61%	11,44% \pm 7,75%	13,6% \pm 11,69%	14,63% \pm 13,1%
Lusignan	10,36% \pm 16,82%	6,42% \pm 5,1%	4,12% \pm 3,97%	3,43% \pm 2,00%	5,33% \pm 2,01%	4,22% \pm 1,27%	3,63% \pm 3,11%	0,92% \pm 1,63%
Mirecourt	7,10% \pm 2,21%	0,74% \pm 1,96%	4,52% \pm 3,05%	0,43% \pm 5,92%	4,95% \pm 2,50%	8,67% \pm 4,51%	10,42% \pm 6,01%	4,16% \pm 2,22%
Mons	22,2% \pm 8,75%	26,7% \pm 14,14%	19,21% \pm 10,48%	10,79% \pm 2,53%	1,46% \pm 4,86%	5,41% \pm 2,13%	6,21% \pm 4,24%	0,73% \pm 1,50%
Rennes	23,55% \pm 7,70%	9,61% \pm 8,94%	10,19% \pm 3,56%	9,04% \pm 3,04%	8,00% \pm 1,28%	6,97% \pm 4,46%	6,16% \pm 2,71%	4,82% \pm 3,21%
Saint-Etienne	-8,05% \pm 16,24%	15,01% \pm 12,88%	10,92% \pm 7,33%	11,18% \pm 4,77%	14,65% \pm 9,30%	6,31% \pm 5,50%	5,24% \pm 2,19%	5,02% \pm 1,52%
Toulouse	9,71% \pm 18,19%	14,76% \pm 24,72%	2,42% \pm 7,79%	5,62% \pm 9,89%	5,06% \pm 10,24%	6,01% \pm 8,63%	3,28% \pm 4,65%	-2,25% \pm 1,08%
Versailles	14,44% \pm 7,01%	2,41% \pm 4,57%	7,79% \pm 7,80%	4,78% \pm 4,51%	3,47% \pm 3,89%	6,56% \pm 2,49%	4,42% \pm 2,13%	2,86% \pm 0,59%

Table 15. Erreurs relatives (moyenne \pm erreur standard) pour la quantité totale de carbone organique du sol prédit avec différentes combinaisons d'ordre d'années au sein d'un cycle climatique de trois ans, pour différents systèmes prairiaux pour deux sites choisis.

	Semé en Fétuque (SF)	Semé en Ray-grass (SL)	Permanente Extensive (PE)	Permanente Intensive (PI)	Tous les systèmes
Bordeaux	-0,31% \pm 0,15%	0,05% \pm 0,10%	0,56% \pm 0,14%	0,38% \pm 0,13%	0,17% \pm 0,08%
Mirecourt	0,55% \pm 0,10%	1,01% \pm 0,08%	-0,22% \pm 0,12%	0,17% \pm 0,13%	0,38% \pm 0,14%

4.4.3. Discussion

Cette étude visait à améliorer notre connaissance du comportement de la matière organique du sol, en se concentrant sur la question de l'influence du nombre et de l'ordre des années dans le cycle climatique utilisé lors de la recherche de l'équilibre, tout en distinguant les différentes gestions de la parcelle. Nous avons montré qu'un cycle de trois ans peut être suffisant pour capturer les informations relatives à la période climatique pour les prairies temporaires uniquement fauchées. Tandis que de manière plus générale, un cycle de cinq ans peut être considéré comme un bon compromis entre précision et coût de calcul. Nous avons également mis en évidence que l'influence de l'ordre dans lequel les années sont introduites dans le cycle climatique est négligeable. Cette étude peut aider à résoudre le problème de l'initialisation du modèle à l'équilibre, et en particulier aider à répondre à la question de combien d'années climatiques devraient être utilisées pour définir la matière organique du sol à l'équilibre. Le modèle biogéochimique utilisé dans cette étude représente l'état de l'art des processus et interactions sol-plante-animal-atmosphère pour les prairies. En conséquence, nous pouvons raisonnablement supposer que d'autres modèles biogéochimiques fonctionnant à la même échelle et avec le même niveau de détails montreront des comportements semblables.

Cette partie sert également de pré-requis à la partie suivante dans laquelle on calcule des surfaces de réponse de la matière organique à l'équilibre dans le but d'effectuer la seconde phase de l'analyse de vulnérabilité (*i.e.* analyse de vulnérabilité avec adaptation). Cependant pour réaliser ces surfaces certains postulats nécessitaient d'être vérifiés tels que la représentativité de trois années climatiques par rapport à la période climatique et la non influence de l'ordre des années climatiques permettant ainsi de les trier pour le calcul des surfaces de réponses.

4.5. Surfaces de réponse de la matière organique de PaSim à l'équilibre

Comme rappelé précédemment, la matière organique du sol (MOS) est le principal réservoir de carbone (C) et d'azote (N) organique dans les prairies et joue un rôle-clé dans l'atténuation des émissions des gaz à effet de serre (GES). Les modèles biogéochimiques classiques intègrent une vision mécaniste de la dynamique de la matière organique du sol.

Il existe un grand nombre de paramètres qui influent sur la vulnérabilité au changement climatique. De plus, il y a un niveau élevé d'incertitudes associées aux études d'impacts du changement climatique (notamment les scénarios d'émissions, la modélisation du climat, les méthodes de régionalisation et d'initialisation et la modélisation des impacts sur le système cible). Par conséquent l'évaluation de la vulnérabilité des stocks de matière organique du sol nécessite beaucoup de simulations pour obtenir une idée claire et précise de l'influence des variables environnementales. Cela suggère qu'un plan d'expériences

(DOE) bien conçu devrait être employé pour réduire le temps nécessaire à la réalisation des simulations ([Lardy et al. 2011a], voir 3.4). Une telle quantité de simulations nécessite également de faire appel au calcul haute performance pour distribuer les calculs.

En raison du manque fréquent de données expérimentales suffisantes et / ou la difficulté de relier les mesures avec les valeurs d'entrée de MOS dans le modèle, la méthode la plus courante pour initialiser les modèles consiste à le mettre en équilibre avec le climat et la gestion. Pour cette initialisation, des simulations itératives sont généralement effectuées pour amener les compartiments de carbone et d'azote du sol à un état stationnaire. Pour éviter les temps de calcul prohibitifs exigés par cette approche classique, nous pouvons construire des surfaces de réponse de matière organique du sol, c'est-à-dire des approximations de la relation entre les entrées et les sorties, en termes beaucoup plus simples que le modèle de simulation complet [Kleijnen et al., 2005]. Dans cette partie nous allons documenter la création de ces surfaces de réponse pour la matière organique du sol sous une gamme de conditions climatiques, de sols et de gestions représentatives de la France, simulées à l'aide du modèle PaSim [Riedo et al., 1998]. Cette étude est aussi une première étape dans l'évaluation de la vulnérabilité de la matière organique du sol, en tenant compte des différentes sources d'incertitudes, comme le climat, la gestion, le sol ou les espèces végétales [Lardy et al., 2011a]. Cette partie avait également pour but de créer les scripts R pour les surfaces de réponses, le script OpenMOLE pour la distribution des simulations et le code nécessaire à la conception du plan d'expériences décrit plus loin. En effet, avant d'implémenter les transformations de modèles vers les scripts, dans le plugin EMF, il nous a paru nécessaire de posséder d'abord une base d'exemples de scripts.

Les sections suivantes présentent les étapes de conception et de réalisation du plan d'expériences (choix des facteurs et du domaine de simulation, choix du plan d'expériences, méthodes de régression), réalisé à l'aide de PaSim. La version 5.3 de PaSim a été utilisée dans cette étude. Toutefois, pour des raisons de simplification de données d'entrée, l'ancienne version du module animal (la version 4.5 de PaSim) a été utilisée. La méthode algébrique de mise à l'équilibre, développée pour réduire le temps de calcul [Lardy et al., 2011b], a été utilisée dans cette étude. Après avoir discuté les résultats, une dernière section étend les surfaces de réponses aux périodes climatiques futures.

4.5.1. Choix des facteurs et du domaine de simulation

La taille du domaine de simulation a été réduite en tenant compte de climats et de conditions de gestion représentatifs de la France. De plus, le nombre de pratiques agricoles potentielles a été réduit aux simples prairies fauchées. Les paramètres par défaut de la végétation ont été utilisés (voir [Riedo et al., 1998]). L'expérimentation a été construite sur la combinaison de trois types de facteurs:

- Le sol. Nous avons utilisé 102 sols prairiaux dominants de France, fournis par le projet ANR VALIDATE⁸⁰. Les sols furent caractérisés par leur texture (fraction d'argile, sable et limon), leur profondeur, leur densité apparente et leur pH. Les autres caractéristiques du sol n'ont pas été prises en compte dans la conception car elles sont très liées aux quatre propriétés précédentes. Comme les données ne contenaient pas l'ensemble des combinaisons de profondeur × texture × densité apparente × pH, une valeur qualitative a été affectée à chaque sol (*i.e.* un numéro par sol) pour la conception du DOE. Cependant, les propriétés du sol ont été utilisées lors de l'estimation des surfaces de réponse ;
- Le climat. Pour chaque simulation de mise à l'équilibre, nous avons répété un cycle climatique de trois ans jusqu'à ce que l'équilibre soit atteint. Nous avons montré que l'utilisation de trois années est un bon compromis entre vitesse et performance (il s'agit ici de prairies uniquement fauchées). Les données climatiques sont celles observées de 1970 (ou ultérieurement) à 2006 sur 12 sites représentatifs du climat français [Graux et al., 2012a]. Les 408 années de données météorologiques au pas de temps horaire, ont été fournies par le projet ANR CLIMATOR⁸¹. En raison de la complexité du choix d'un ensemble de variables représentatives du climat, et afin d'avoir une couverture cohérente du domaine, le climat a été considéré comme un facteur qualitatif dans la conception du DOE. Nous avons également ajouté, comme facteur indépendant, la concentration atmosphérique en CO₂ et l'altitude du site au design ;
- La gestion. Elle est considérée comme constante au cours des trois années du cycle répété lors de la mise à l'équilibre. Différentes pratiques agricoles furent appliquées, en fonction du nombre de coupes et des niveaux de fertilisation azotée apportée. De un à quatre événements de fauche par an peuvent avoir lieu, avec environ un mois d'intervalle entre chacune d'elles. Les quantités de fertilisant apportée (sous la forme d'ammonitrate) varient de 0 à 120 kg N ha⁻¹ par an, appliquées 30 jours avant la première fauche et deux jours après les autres fauches (exceptée la dernière). La même quantité de fertilisant a été apportée à chaque date, en respectant un minimum de 40 kg N ha⁻¹ (si nécessaire pour cela, les dates de fertilisation les plus tardives sont ignorées). La présence de légumineuses dans la prairie étant également liée aux pratiques, elles ont été ajoutées au design (sous forme de fraction du couvert). Cette dernière restera fixe durant toute la simulation avec PaSim.

⁸⁰ <http://www1.clermont.inra.fr/validate>

⁸¹ http://w3.avignon.inra.fr/projet_climator/

4.5.2. Choix du plan expérimental

En raison du nombre de facteurs et de leurs niveaux respectifs dans cette étude (Table 16), la plupart des plans d'expériences classiques (comme par exemple le plan d'expériences factoriel) ne sont pas directement applicables. Nous avons donc décidé d'utiliser un LHD ("*Latin Hypercube Design*", [McKay et al., 1979]) pour ses bonnes propriétés de remplissage d'espace pour un nombre relativement faible de points d'échantillonnage. La taille du LHD est le plus petit commun multiple des niveaux (c'est-à-dire 157 080). Une simple transformation a été appliquée sur chaque facteur pour obtenir le niveau correct (c'est-à-dire, par exemple, pour repasser de 157 080 niveaux à 408 pour les années climatiques). Nous l'avons affiné en optimisant le critère "Maximin" [Johnson et al., 1990] sur une centaine de designs. C'est-à-dire que nous avons maximisé la distance minimale entre deux points du plan d'expériences.

Table 16. Description des facteurs du plan d'expériences et de leurs niveaux.

Nom du facteur	Nombre de niveaux	Information
Année climatique 1	408	Trois niveaux de grain temporel sont utilisés dans les surfaces de réponse.
Année climatique 2	408	
Année climatique 3	408	
Concentration atmosphérique en CO₂	8	320–390 ppm
Sol	102	Profondeur, densité apparente, pH et texture sont utilisés dans les surfaces de réponse.
Nombre de fauche	4	1; 2; 3; 4
1^{ère} date de fauche	7	1 ^{er} avril au 1 ^{er} mai
2^{ème} date de fauche	7	16 mai au 15 juin
3^{ème} date de fauche	7	1 ^{er} juillet au 31 juillet
4^{ème} date de fauche	7	16 août au 15 septembre
Fraction de légumineuse	5	0; 10; 20; 30; 40%
Taux de fertilisation azotée	6	0; 40; 60; 80; 100; 120 kg N ha ⁻¹ année ⁻¹
Altitude	20	50–1000 m

4.5.3. Méthodes de régression

Le plan d'expériences décrit précédemment, nous permet de construire des surfaces de réponse de la matière organique du sol au climat, au sol et à la gestion. Dans notre cas, nous avons limité notre étude aux approches les plus simples, c'est-à-dire des régressions linéaires et quadratiques, avec différentes entrées climatiques. En effet, nous avons utilisé trois niveaux de granularité temporelle pour représenter le climat :

- Une seule variable par an : $P-ET_0$, où P est la somme des précipitations annuelles, ET_0 la somme annuelle de l'évapotranspiration potentielle de référence définie par [Walter et al., 2000] et ajustée de l'effet de la concentration

en CO₂ atmosphérique [Olioso et al., 2010]. Il s'agit d'un indice de bilan hydrique climatique (Climatic Water Balance indicator) ;

- Cinq variables par saison⁸² : la température moyenne de l'air, le rayonnement global moyen, l'humidité moyenne de l'air, la vitesse moyenne du vent et la somme des précipitations ;
- Cinq variables par mois : la température moyenne de l'air, le rayonnement global moyen, l'humidité moyenne de l'air, la vitesse moyenne du vent et la somme des précipitations.

Comme l'influence de l'ordre des années est négligeable sur la quantité totale de matière organique du sol (erreur < 0,5 %, Table 15, voir 4.4.2), les années ont été triées par conditions d'aridité pour le calcul de la régression (*i.e.* par P-ET₀ dans le premier cas, et par l'indice d'aridité de De Martonne-Gottmann ([De Martonne 1942]) pour les deux autres cas). Toutes les régressions et les tests statistiques ont été effectués en utilisant le package "biglm" [Lumley 2011] dans le logiciel R [R Development Core Team 2011].

4.5.4. Résultats

Nous avons réalisé les simulations en les distribuant sur la grille de calcul européenne en utilisant l'organisation virtuelle « Biomed », grâce au logiciel OpenMOLE ([Reuillon et al., 2010]) dans le but de réduire le temps d'attente nécessaire. Il a suffi d'un weekend, tandis qu'il aurait en effet fallu environ quatre ans à un seul cœur de processeur « moderne » pour réaliser toutes les simulations.

La qualité des surfaces de régression de la matière organique du sol a été analysée à travers trois indices de performance : la racine carrée relative de l'erreur quadratique moyenne (RRMSE, "*Relative Root Mean Squared Error*", Table 17), la racine carrée de l'erreur relative quadratique moyenne (RMSRE, "*Root Mean Squared Relative Error*", Table 18) et l'efficacité de modélisation (EF, Table 19). Tous ces indices utilisent la comparaison entre les valeurs prédites par régression (P_i) et les valeurs données par PaSim (O_i) pour chaque $i^{ème}$ simulation du plan d'expériences. Pour la RRMSE et la RMSRE, plus la valeur est proche de zéro, meilleure est la prédiction, tandis que pour l'efficacité de modélisation ($[-\infty ; +1]$) les meilleures prédictions sont obtenues pour des valeurs proches de 1. Tous les indices calculés pour l'azote organique du sol ont montré les mêmes tendances (données non présentées) que ceux calculés pour le carbone organique du sol. Les résultats des régressions pour le carbone organique du sol (Table 17, Table 18, Table 19) montrent clairement que les informations sur le climat contenues dans la variable annuelle P-ET₀ sont suffisantes pour une régression linéaire des compartiments de carbone organique du sol à l'équilibre. En effet, la RRMSE décroît lors du passage de la variable

⁸² Au sens calendaire, *i.e.* janvier-mars ; avril-juin ; juillet- septembre; octobre-décembre

annuelle (P-ET₀) à des variables saisonnières et mensuelles (*i.e.* précipitations, rayonnement ...) car l'ajout de variables explicatives permet de réduire le carré des erreurs (et d'améliorer l'efficacité) des régressions de moindres carrés. Cependant, la RMSRE, qui donne plus de poids à l'erreur relative et moins à l'erreur totale est seulement améliorée dans cinq des 22 cas.

Table 17. Évaluation par l'indice **RRMSE** ("*Relative Root Mean-Squared Error*") des régressions linéaires et quadratiques pour les cinq compartiments de carbone (C) et leur total pour les trois niveaux d'entrées climatiques (P-ET₀, Saison et Mois), sur les points du

plan d'expériences. $RRMSE = \sqrt{\frac{\sum_{i=1}^n (P_i - O_i)^2}{n}} \times 100$; où P_i est la i^{ème} valeur prédite par la régression et O_i la i^{ème} valeur donnée par PaSim.

	Linéaire			Quadratique		
	P-ET ₀	Saison	Mois	P-ET ₀	Saison	Mois
C métabolique	38,87	36,21	36,09	34,17	29,14	27,78
C structurel	30,25	28,81	28,75	21,56	16,60	15,69
C actif	27,54	26,59	26,49	18,34	14,99	14,08
C lent	21,03	19,85	19,75	14,99	11,18	10,38
C passif	20,89	19,71	19,62	14,85	11,04	10,25
C total	21,39	20,12	20,03	15,26	11,20	10,42

Table 18. Évaluation par l'indice **RMSRE** ("*Root Mean-Squared Relative Error*") des régressions linéaires et quadratiques pour les cinq compartiments de carbone (C) et leur total pour les trois niveaux d'entrées climatiques (P-ET₀, Saison et Mois), sur les points du

plan d'expériences. $RMSRE = \sqrt{\frac{\sum_{i=1}^n \left(\frac{P_i - O_i}{O_i}\right)^2}{n}} \times 100$; où P_i est la i^{ème} valeur prédite par la régression et O_i la i^{ème} valeur donnée par PaSim.

	Linéaire			Quadratique		
	P-ET ₀	Saison	Mois	P-ET ₀	Saison	Mois
C métabolique	83,05	81,61	81,44	60,95	56,70	54,59
C structurel	38,99	39,95	39,94	21,32	19,98	19,55
C actif	35,77	36,06	36,05	21,94	20,67	19,71
C lent	28,47	28,55	28,49	17,33	15,39	14,42
C passif	28,38	28,44	28,37	17,10	15,11	14,15
C total	28,62	28,75	28,69	17,23	15,20	14,27

Le passage d'une régression linéaire à une quadratique peut considérablement améliorer la qualité de l'approximation, et ce quel que soit le critère statistique pris en compte (par exemple l'efficacité de modélisation augmente de ~ 0,7 à ~ 0,9 pour le carbone total du sol). Cette amélioration n'est pas uniquement due à l'augmentation du nombre de variables explicatives. En effet, la régression quadratique P-ET₀ contient 172 variables contre 196

variables dans la régression linéaire avec entrées climatiques mensuelles. À la différence de la régression linéaire, passer d'une variable climatique annuelle à cinq variables climatiques par saisons entraîne une amélioration de la régression quadratique (par exemple, la RRMSE a diminuée, passant de 15% à 11% pour le carbone total du sol).

Table 19. Évaluation par l'indice d'efficacité de modélisation (EF) des régressions linéaires et quadratiques pour les cinq compartiments de carbone (C) et leur total pour les trois niveaux d'entrées climatiques (P-ET₀, Saison et Mois), sur les points du plan d'expériences. $Efficiency = 1 - \frac{\sum_{i=1}^n (P_i - O_i)^2}{\sum_{i=1}^n (\bar{O} - O_i)^2}$; où P_i est la i^{ème} valeur prédite par la régression et O_i la i^{ème} valeur donnée par PaSim.

	Linéaire			Quadratique		
	P-ET ₀	Saison	Mois	P-ET ₀	Saison	Mois
C métabolique	0,183	0,360	0,367	0,468	0,662	0,702
C structurel	0,662	0,703	0,704	0,853	0,918	0,927
C actif	0,609	0,645	0,649	0,858	0,909	0,921
C lent	0,688	0,731	0,735	0,863	0,928	0,939
C Passif	0,687	0,731	0,734	0,863	0,929	0,939
C total	0,695	0,739	0,742	0,865	0,931	0,941

La faible qualité de la régression du compartiment métabolique est en partie due à son faible temps moyen de résidence (0,5 ans [Parton et al., 1988]) et sa forte relation avec l'état de la biomasse végétale. Compte tenu de son taux de décomposition élevé, ce compartiment est plus influencé par le climat de la dernière année que par le cycle climatique complet. En effet, les régressions ont été améliorées par l'utilisation d'années non triées (données non présentées), avec une efficacité de modélisation de 0,629 contre 0,367 pour une régression linéaire en utilisant les données climatiques mensuelles. Toutefois, le gain de prédiction n'est apparent que pour le compartiment métabolique, qui représente en moyenne que 0,89% de la biomasse totale du sol. L'absence de tri conduit à une légère baisse de la précision de la prédiction selon certains critères (par exemple 0,71 vs 0,69 pour l'efficacité de modélisation de la matière organique totale avec P-ET₀).

PaSim a montré un taux d'échec de 4,0 %, uniquement sur des sols alcalins (8 ≤ pH ≤ 8,5), où l'instabilité numérique dans le calcul de la concentration en ammoniac a été exacerbée par le climat sec et la gestion intensive. Il est également intéressant de noter que 0,95 % des équilibres n'existent pas (même après 100 cycles de recherche d'équilibre avec la méthode matricielle [Lardy et al., 2011b]). Ces situations sont caractérisées par une fréquence de fauche plus élevées que la moyenne (environ 2,8 contre 2,5), une fraction de légumineuses plus faible (en moyenne, 10% contre 20% sur tout le plan d'expériences) et des taux de fertilisation plus faibles (46 kg N ha⁻¹ an⁻¹ par rapport à une moyenne de 66 kg N ha⁻¹ an⁻¹). Ces conditions ne sont pas viables pour les prairies, l'exploitation est excessive (fauche) par rapport aux ressources (ressources limitées d'azote) et provoque une

absence d'équilibre si les conditions climatiques et de gestion étaient maintenues indéfiniment. En effet, en absence d'équilibre le modèle se déplace lentement vers une biomasse herbacée nulle.

Bien que les taux de décomposition de la matière organique du sol soient dépendants de la température dans le modèle, l'altitude ne change pas significativement les valeurs de matière organique dans cette étude. En effet, les effets de l'altitude et du climat ont été testés de manière indépendante (c'est-à-dire que l'effet de l'altitude a été testé sans changement des données climatiques d'entrée et inversement). Comme attendu, la concentration atmosphérique en CO₂, la fraction de légumineuses et la quantité d'engrais appliquée ont un effet positif sur la matière organique, car ils favorisent la croissance des plantes et donc augmentent les flux entrant dans la matière organique du sol. De même, la fauche a un effet négatif en raison de l'exportation de matière. Comme le modèle simule la matière organique du sol sur tout le profil du sol, il est logique de trouver un effet positif de la profondeur du sol, de même pour la densité apparente qui a globalement un effet positif. L'augmentation dans chacune des fractions de la texture du sol a un effet négatif, cependant, qui est plus faible pour la teneur en limon. Il est intéressant de noter que l'effet de l'indice agro-météorologique P-ET₀ a un effet positif sur la matière organique. En fait, un climat aride induit une baisse de productivité de la prairie associée à une réduction des flux de matières. De plus, l'effet des deux années les plus arides est presque deux fois celui de l'année la plus humide (coefficients de régression standardisés, pour la totalité du carbone, avec une régression linéaire avec P-ET₀ : 6,0 ; 5,2 ; 3,0 kg C m⁻² mm⁻¹). Ces résultats montrent que le comportement du modèle PaSim est cohérent avec l'état de l'art des interactions climat-sol-plante.

On peut observer (Figure 56) que la réponse au climat de la matière organique du sol à l'équilibre est bruitée. En effet, on peut facilement couvrir une gamme très importante de valeurs de quantité de carbone organique du sol pour chaque valeur de P-ET₀ (d'une seule année climatique). Nous pouvons évaluer sur ces deux exemples un écart-type de 5-7%. Ceci peut s'expliquer par le fait qu'*in situ* la dynamique intra-annuelle de la végétation contient des effets de seuil (par exemple la phénologie) et par la présence de non linéarités dans les équations (par exemple le bilan énergétique) et dans les entrées (par exemple, les précipitations). D'une certaine manière, cela prouve l'intérêt de l'utilisation des modèles biogéochimiques (par rapport aux modèles d'approximation) pour simuler la dynamique de la matière organique du sol. Cela montre qu'un modèle comme PaSim est capable de produire de l'information qui ne peut pas être complètement capturée par des régressions. En outre, avec les modèles de régression, il est plus difficile de détecter l'absence d'équilibre qu'avec les simulations usuelles de mise à l'équilibre. En effet, les régressions peuvent fournir des valeurs d'équilibre de matière organique du sol même si il n'y a presque plus de biomasse herbacée (situations irréalistes).

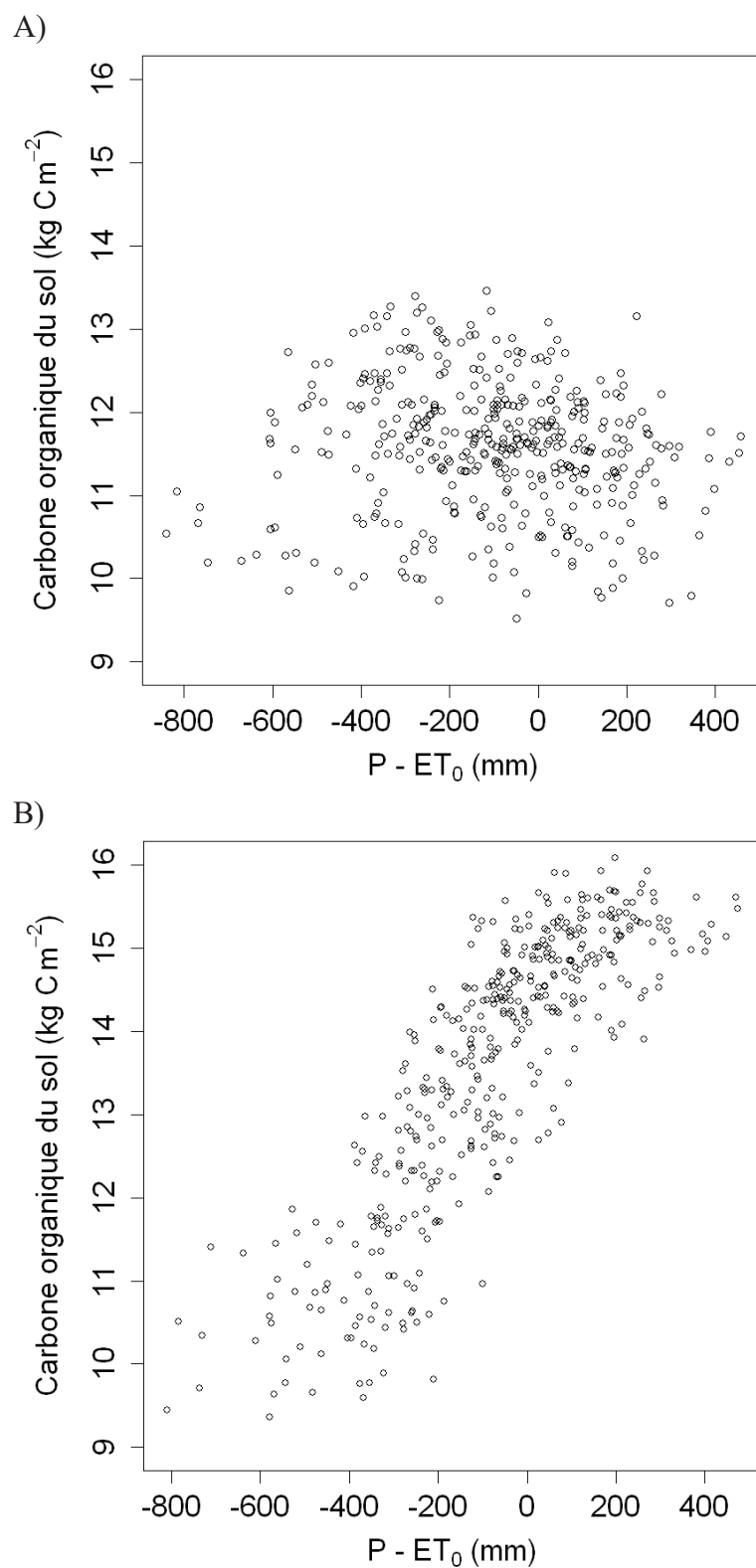


Figure 56 [Lardy et al., 2012a] Influence du changement d'une seule année climatique sur le carbone organique du sol à l'équilibre, pour A) une prairie extensive en climat aride sur sol peu profond , et pour B) une prairie intensive, en climat humide sur sol profond.

4.5.5. Discussion

Cette étude démontre la faisabilité d'approximer la matière organique du sol à l'équilibre par une régression quadratique dans le but d'initialiser plus rapidement les modèles mécanistes biogéochimiques. Utiliser P-ET₀ peut être suffisant pour capturer les informations des entrées climatiques. Par exemple, l'initialisation par régression quadratique avec P-ET₀ semble être environ 500 fois plus rapide qu'avec une simulation de mise à l'équilibre avec PaSim, avec un écart de 10-15% en raison de la variabilité de la réponse de la matière organique du sol au climat (par exemple ~7% d'écart-type pour une des trois années climatiques). L'analyse de vulnérabilité est un processus en plusieurs étapes, qui comprend des analyses de sensibilité et d'incertitudes à l'aide de modèles [Lardy et al., 2011a], et nécessite donc une très grande quantité de simulations. Les régressions quadratiques présentées ici pourraient également être utilisées dans les analyses de sensibilité nécessitant un nombre prohibitif de simulations de modèles complexes.

Cette partie nous a permis de créer les scripts R pour les surfaces de réponses, le script OpenMOLE pour la distribution des simulations et le code nécessaire à la conception du plan d'expériences. Nous avons ainsi pu implémenter, dans le plugin EMF, les transformations de modèles vers les scripts R et OpenMOLE. Cette implémentation du plugin EMF a été notamment testée dans la partie suivante.

En outre, ce travail peut être étendu aux projections de conditions climatiques futures et ainsi permettre le calcul d'indices de vulnérabilité sur le stock potentiel de la matière organique (c'est à dire à l'état d'équilibre) en utilisant le climat actuel comme référence. De plus, les évolutions dans le domaine de la non-existence de l'équilibre peuvent aussi présenter un intérêt. Cependant, des examens complémentaires, pourraient être nécessaires pour prouver que ce modèle contient de manière suffisante, pour notre objectif (l'analyse de vulnérabilité), les tendances spécifiques et les maxima et minima globaux. Il pourrait également être envisagé de tester d'autres modèles d'approximation, tel que la méthode du Krigeage. En réalité dans le cas du Krigeage, la taille du plan d'expériences (et donc du volume des données associées) le rend impossible à calculer. Il faudrait alors découper le domaine parcouru par le plan d'expériences en sous-domaines de taille plus raisonnable de façon à pouvoir effectuer les régressions de Krigeage.

4.5.6. Extension aux périodes futures

Ayant prouvé la faisabilité d'approximer la matière organique du sol à l'équilibre pour la période actuelle, il a été décidé d'effectuer le même travail pour deux périodes futures. Ces surfaces de réponses devraient nous servir dans l'évaluation de la vulnérabilité de la matière organique à l'équilibre. Les surfaces de réponse décrites précédemment concernaient la période 1970-2006. Les données climatiques de cette période que l'on

considérera comme la période de référence étaient des données observées sur 12 sites représentatifs du climat français (Table 12). Pour les périodes futures, nous avons utilisé le climat (fourni par le projet ANR CLIMATOR⁸³) sur ces sites pour le scénario A1B [Nakićenović 2000] [IPCC 2007a]. Ce scénario traduit une croissance économique forte, une population mondiale atteignant les neuf milliards en 2050 et déclinant ensuite, un développement rapide de nouvelles technologies plus efficaces et une utilisation équilibrée de toutes les ressources. Ce scénario a été régionalisé à l'aide de la méthode dite des anomalies [Déqué et al., 2007]. Nous avons considéré la période 2018-2051 comme période de "Futur Proche" (FP), et la période 2067-2100 comme période de "Futur Lointain" (FL). Nous avons conservé le même plan d'expériences que pour la période de référence (Table 16), en changeant juste les années climatiques et la concentration en CO₂ atmosphérique. Ainsi, tout en conservant le même nombre de niveau, les concentrations en CO₂ atmosphérique varient (par pas de 15 ppm) de 415 à 520 ppm pour la période de "futur proche" et de 590 à 695 ppm pour la période de "futur lointain".

Les scripts pour la distribution des simulations via l'outil OpenMOLE ont été générés par l'outil décrit dans le chapitre précédent. PaSim a été encapsulé dans un petit exécutable java. Il prend en entrée les informations du plan d'expériences et génère les fichiers nécessaires à la simulation de mise à l'équilibre avec PaSim. Puis, il lance PaSim et récupère les résultats de la simulation (à savoir les valeurs des variables à l'équilibre). Les simulations pour une période climatique représentant environ 4 ans de calcul mono-cœur sur des CPU modernes, elles furent distribuées sur la grille à travers la VO Biomed. Les scripts R pour le calcul des régressions furent générés à l'aide de l'outil proposé dans le chapitre précédent. La génération de ces surfaces de réponses avait comme premier objectif de permettre une analyse de vulnérabilité de la matière organique du sol, et comme objectif secondaire de vérifier l'opérationnalité du plugin EMF que nous avons proposé pour l'analyse de vulnérabilité.

Les régressions pour les périodes climatiques "Futur Proche" et "Futur Lointain" ont également été évaluées à l'aide des mêmes critères que précédemment (RRMSE : Table 20 ; RMSRE : Table 21 et Efficience de modélisation : Table 22). Comme le résume ces tables, la qualité des régressions n'est pas altérée par le changement de période climatique.

⁸³ http://w3.avignon.inra.fr/projet_climator/

Table 20. Évaluation par l'indice **RRMSE** ("*Relative Root Mean-Squared Error*") des régressions linéaires et quadratiques du carbone total du sol pour les trois périodes climatiques ("Référence" : 1970-2006 ; "Futur Proche" : 2018-2051; "Futur Lointain" : 2067-2100), pour les trois niveaux d'entrées climatiques (P-ET₀, Saison et Mois), sur les

points du plan d'expériences.
$$\text{RRMSE} = \sqrt{\frac{\sum_{i=1}^n (P_i - O_i)^2}{n}} \times 100$$
 ; où P_i est la i^{ème} valeur prédite par la régression et O_i la i^{ème} valeur donnée par PaSim.

	Linéaire			Quadratique		
	P-ET ₀	Saison	Mois	P-ET ₀	Saison	Mois
C total période "Référence"	21,39	20,12	20,03	15,26	11,20	10,42
C total période "Futur proche"	20,64	19,39	19,31	14,68	10,73	9,99
C total période "futur lointain"	20,02	18,81	18,73	14,17	10,28	9,58

Table 21. Évaluation par l'indice **RMSRE** ("*Root Mean-Squared Relative Error*") des régressions linéaires et quadratiques du carbone total du sol pour les trois périodes climatiques ("Référence" : 1970-2006 ; "Futur Proche" : 2018-2051; "Futur Lointain" : 2067-2100), pour les trois niveaux d'entrées climatiques (P-ET₀, Saison et Mois), sur les

points du plan d'expériences.
$$\text{RMSRE} = \sqrt{\frac{\sum_{i=1}^n \left(\frac{P_i - O_i}{O_i} \right)^2}{n}} \times 100$$
 ; où P_i est la i^{ème} valeur prédite par la régression et O_i la i^{ème} valeur donnée par PaSim.

	Linéaire			Quadratique		
	P-ET ₀	Saison	Mois	P-ET ₀	Saison	Mois
C total période "Référence"	28,62	28,75	28,69	17,23	15,20	14,27
C total période "Futur proche"	28,58	28,81	28,76	17,02	15,20	14,23
C total période "futur lointain"	29,29	29,71	29,67	16,99	15,38	14,43

Table 22. Évaluation par l'indice d'**efficience de modélisation** des régressions linéaires et quadratiques du carbone total du sol pour les trois périodes climatiques ("Référence" : 1970-2006 ; "Futur Proche" : 2018-2051; "Futur Lointain" : 2067-2100), pour les trois niveaux d'entrées climatiques (P-ET₀, Saison et Mois), sur les points du plan d'expériences.
$$\text{Efficience} = 1 - \frac{\sum_{i=1}^n (P_i - O_i)^2}{\sum_{i=1}^n (\bar{O} - O_i)^2}$$
 ; où P_i est la i^{ème} valeur prédite par la régression et O_i la i^{ème} valeur donnée par PaSim.

	Linéaire			Quadratique		
	P-ET ₀	Saison	Mois	P-ET ₀	Saison	Mois
C total période "Référence"	0,695	0,739	0,742	0,865	0,931	0,941
C total période "Futur proche"	0,725	0,765	0,768	0,878	0,938	0,947
C total période "futur lointain"	0,745	0,781	0,783	0,887	0,943	0,951

À l'image de la période climatique de référence, nous avons observé un taux d'échecs (respectivement de 3,0 % et de 6,2 %, pour la période climatique du "Futur Proche" et du "Futur Lointain"), dus à certaines instabilités numériques dans les calculs du cycle de l'azote au sein du logiciel PaSim. On peut noter une diminution pour les périodes futures du nombre de simulations se concluant par une absence d'équilibre: 0,90 % pour la période "Future Proche" et 0,88 % pour la période "Futur lointain", comparativement à un taux de 0,95 % pour la période de référence. Cette zone du domaine du plan d'expériences où les conditions ne sont pas supportables pour la prairie conserve globalement les mêmes propriétés que pour la période de référence, c'est-à-dire une fréquence de fauche plus élevée que la moyenne (~2,9 fauches contre 2,5 sur l'ensemble du plan d'expériences), une faible fraction de légumineuses (9 à 10% contre 20%) et enfin un plus faible apport en fertilisant (46 à 49 kg N ha⁻¹ an⁻¹ par rapport à une moyenne de 66 kg N ha⁻¹ an⁻¹). On aurait pu penser que le type de sol soit un facteur important. En effet, environ la moitié (39) des sols concernés par l'absence d'équilibre sont en communs à travers les périodes climatiques (entre 69 et 78 sols concernés pour une seule période). Cependant les propriétés des sols ne diffèrent que faiblement par rapport à la moyenne du plan conçu. En réalité, on peut supposer que le nombre de simulations se traduisant par une absence d'équilibre (139-150) est trop faible pour qualifier une possible évolution des bornes du domaine d'existence de l'équilibre. De plus, cela limite l'interprétation de l'influence du climat sur les conditions inacceptables pour une mise à l'équilibre. Ainsi, par exemple, on note juste une tendance à un climat plus aride pour les « non-équilibres » que pour le reste du design (par exemple pour la variable climatique P-ET₀, pour la période de référence, la valeur est de -57,7±9,9 pour les simulations hors équilibres contre -47±0,32 pour le reste du design). Bien que les interprétations sur l'effet du changement climatique soient limitées, on peut cependant noter que la légère diminution d'absence d'équilibre traduit un effet supérieur de l'effet de l'augmentation du CO₂ atmosphérique sur les effets négatifs des autres composantes du climat. En effet, la concentration en CO₂ atmosphérique tend à augmenter l'efficacité en utilisation de l'eau à travers la fermeture des stomates [Leuning 1995] et augmente la productivité de la prairie (photosynthèse plus efficace [Farquhar et Caemmerer, 1982]). Il convient cependant de pondérer l'interprétation de ces résultats de la validité du modèle, qui bien que représentatif de l'état de l'art, peut contenir des erreurs structurelles. Ainsi, une confirmation de ces tendances par d'autres modèles (tel que DNDC [Li et al., 1992] [Li 2000], par exemple) serait nécessaire. De même, une campagne de simulations avec un plan d'expériences "autour" des points de "non-équilibre" pourrait nous permettre de mieux appréhender la frontière entre équilibre et non-équilibre, notamment pour les facteurs de gestion. Ce sont en effet uniquement sur ces derniers que l'on agit pour éviter ces situations de non-équilibre.

La réalisation des surfaces de réponses fait partie des étapes de l'analyse de vulnérabilité. C'est l'étape-clé qui permet de passer de l'analyse de vulnérabilité sans adaptation à l'analyse de vulnérabilité avec adaptation. Ces surfaces sont en effet

nécessaires pour la recherche de minimisation de vulnérabilité. En effet cette dernière nécessite un trop grand nombre de simulations pour être réalisée à l'aide du modèle d'origine. Cette phase est décrite dans la partie suivante.

4.6. Analyse de vulnérabilité au changement climatique de la matière organique à l'équilibre

À l'aide des simulations réalisées dans la partie précédente, nous allons pouvoir évaluer la vulnérabilité de la matière organique du sol des prairies au changement climatique, sans prise en compte de l'adaptation. Puis, à l'aide des surfaces de réponse de la matière organique au climat, à l'environnement et à la gestion, pour les trois périodes climatiques, nous allons pouvoir proposer une évaluation de la vulnérabilité avec prise en compte de l'adaptation. Cette partie, nous permet d'illustrer la méthode d'analyse de vulnérabilité, proposée dans le chapitre précédent, notamment l'étape de recherche du minimum de vulnérabilité. Ce travail nous servira également à proposer des améliorations dans le plugin EMF, décrit dans le chapitre précédent.

4.6.1. Analyse de vulnérabilité au changement climatique de la matière organique à l'équilibre sans prise en compte de l'adaptation

Dans une première approche et pour des raisons de simplicité, nous allons considérer l'exposition comme uniforme. C'est-à-dire que chaque élément du plan d'expériences réalisé précédemment sera considéré comme équiprobable. Il est alors relativement simple d'appliquer le calcul des indices de vulnérabilité sur tout le design. Considérant la valeur totale de matière organique du sol à l'équilibre, un seuil à 75 tC/ha a été défini (cela correspond à 5,45% des valeurs de la période de référence et peut être considéré comme représentatif de sols pauvres [Brunelle et Savoie, 2000]⁸⁴). Les indices ont été normés par la valeur de la période de référence (Table 23).

⁸⁴ 75 tC/ha correspond à une teneur en carbone organique de 2% pour un sol de 50 cm de profondeur, de densité apparente 1,4 t/m³ et un taux de conversion du carbone organique en teneur de matière organique du sol de 1,9 [Pribyl 2010].

Table 23. **Indices de vulnérabilité**, calculées sur le plan d'expériences utilisé pour régresser les surfaces de réponse, pour le "Futur Proche" (FP : 2018-2051) et le "Futur Lointain" (FL : 2067-2100). Les valeurs des indices ont été **normées par la valeur de la "Période de Référence"** (PR : 1970-2006). Plus la valeur de l'indice est haute, plus le système est vulnérable.

Indice Période	Most vulnerable individual	Indice de Luers	Vulnérabilité proportionnelle	Vulnerability gap	Vulnerability severity
"Futur Proche"	0,98	0,87	0,73	0,78	0,81
"Futur Lointain"	0,90	0,76	0,57	0,64	0,67

Le modèle traduit globalement une augmentation du stock de carbone pour les périodes futures. En effet, sachant que le plan d'expériences est uniforme, on peut dans un premier temps observer une augmentation du stock moyen à l'équilibre sur les différentes périodes climatiques (*i.e.* moyenne \pm erreur standard (kg C/m²) : période de référence 16,75 \pm 0,019 ; période de "futur proche" 19,38 \pm 0,022 ; période de "futur lointain" 22,26 \pm 0,025), ce qui cohérent avec la littérature (par exemple [Jones et Donnelly, 2004] [Rees et al., 2005]). Bien que l'état actuel des sols soient hors équilibre, et que par conséquent certains compartiments puissent être au dessus ou en dessous de la valeur d'équilibre, on peut cependant attendre un stockage de carbone dans les périodes futures de façon à atteindre ce potentiel de stock. Il est intéressant de noter que tous les indices convergent pour dire qu'il y aura une diminution de la vulnérabilité de la matière organique à l'équilibre dans le "futur lointain", due à une augmentation plus forte de la teneur moyenne que de la variabilité. De plus, l'indice "Most vulnerable individual" traduit le fait que parmi les situations viables, la quantité de matière organique du sol à la plus faible teneur sera quasiment identique entre la période de référence et la période de "futur proche" (FP), tandis que dans la période de "futur lointain" (FL), il sera sensiblement plus élevé. De même, on peut en déduire que le nombre de sols en dessous du seuil critique de 75 tC/ha (sur toute la profondeur du sol) diminue de 27% puis de 43% pour la période FP et FL tandis que la distance euclidienne (incluse dans l'indice de "Vulnerability severity") ne diminue que de 19% et 33%.

Table 24. **Indices de vulnérabilité**, calculés à l'aide des régressions quadratiques pour les trois niveaux d'entrée climatique (P-ET₀, Saison et Mois) (sur le plan d'expériences ayant lui-même servi à les régresser) et pour les résultats de simulation à l'aide de PaSim, pour le "Futur Proche" (FP : 2018-2051) et le "Futur Lointain" (FL : 2067-2100). Les valeurs des indices ont été **normées par la valeur de la "Période de Référence"** (PR : 1970-2006). Plus la valeur de l'indice obtenue avec la surface de réponse (P-ET₀, Saison et Mois) est proche de celle obtenue avec PaSim, plus la régression est fiable pour l'évaluation de vulnérabilité.

	Période "Futur Proche"				Période "Futur Lointain"			
	Régressions quadratiques			PaSim	Régressions quadratiques			PaSim
	P-ET ₀	Saison	Mois		P-ET ₀	Saison	Mois	
Most vulnerable individual	1,002	1,001	1,001	0,984	1,001	1,003	0,999	0,899
Indice de Luers	0,873	0,871	0,871	0,869	0,761	0,758	0,758	0,755
Vulnérabilité proportionnelle	0,713	0,698	0,695	0,725	0,548	0,527	0,528	0,575
Vulnerability gap	0,792	0,780	0,778	0,783	0,621	0,621	0,619	0,641
Vulnerability severity	0,851	0,834	0,844	0,810	0,689	0,698	0,706	0,671

Avant de pouvoir utiliser les surfaces de réponses de la matière organique à l'équilibre, une question demeure : ces modèles d'approximation contiennent-ils toute l'information nécessaire pour donner une évaluation cohérente de la vulnérabilité ? Ainsi, l'étape suivante de notre démarche consiste à calculer les indices de vulnérabilité sur le même plan d'expériences, mais à l'aide des surfaces de réponses. Les résultats de la Table 24 illustrent les valeurs des indices ainsi calculés et comparés aux valeurs obtenues à partir des simulations de PaSim. Ainsi plus la valeur est proche de celle obtenue avec PaSim, plus l'on peut considérer que la régression est fiable pour approximer la matière organique à l'équilibre lors de l'évaluation de la vulnérabilité. Lors de l'évaluation des indices une centaine de points (soit moins de 1‰) du plan d'expériences ont été ignorés car hors du domaine de validité des régressions (*i.e.* valeurs négatives). Cela explique que les valeurs de l'indice "Most vulnerable individual" ne soient pas exploitables par les surfaces de réponse. En effet, cet indice utilise la plus basse valeur de matière organique du sol, or nous l'avons borné à 0, d'où une valeur de l'indice très proche de 1. En revanche, les valeurs et les tendances des autres indices de vulnérabilité sont conservées par les régressions.

Ayant confirmé la validité de l'utilisation des surfaces de réponses pour notre objectif, nous pourrions raffiner les calculs des indices en ne se concentrant plus simplement sur l'intégralité du domaine, mais en évaluant la vulnérabilité de manière plus locale, puis en agrégeant les informations. L'idée est que la sensibilité à un facteur étant elle-même liée aux autres facteurs, elle ne peut donc réellement être évaluée que localement.

Ainsi, nous avons réévalué la vulnérabilité, en évaluant localement la sensibilité pour chacune des combinaisons du LHD ayant servi à régresser les surfaces. Ainsi en chacun des points de l'espace des paramètres définis par le LHD, nous avons localement appliqué un plan "One factor At a Time" [Kleijnen 1987] sur la gestion. Ce plan OAT se caractérise par ± 5 jours pour les dates de fauches, $\pm 5\%$ de la proportion de légumineuses dans la prairie et $\pm 10 \text{ kg N ha}^{-1} \text{ an}^{-1}$ pour la fertilisation minérale, tout en respectant les bornes définies dans le design (Table 16) utilisé pour la régression des surfaces de réponse (dans le but d'éviter de sortir du domaine de validité des surfaces). Ce plan contient donc entre 7 et 13 éléments (en fonction du nombre de fauches). La vulnérabilité a alors été évaluée en chacun des 157 800 points de l'espace formé par le LHD.

Table 25. Indices de vulnérabilité, calculés à l'aide des régressions quadratiques pour les trois niveaux d'entrée climatique (P-ET₀, Saison et Mois), pour le "Futur Proche" (FP : 2018-2051) et le "Futur Lointain" (FL : 2067-2100). La vulnérabilité est évaluée en chacun des points du LHD ayant servis à régresser les surfaces, par une plan de type OAT. Les valeurs des indices calculés de manière uniforme à l'aide de PaSim sont également rappelées. Les valeurs des indices ont été normées par la valeur de la "Période de Référence" (PR : 1970-2006).

	Période "Futur Proche"				Période "Futur Lointain"			
	Régressions quadratiques			PaSim	Régressions quadratiques			PaSim
	P-ET ₀	Saison	Mois		P-ET ₀	Saison	Mois	
Most vulnerable individual	0,778	0,764	0,772	0,984	0,597	0,582	0,591	0,899
Indice de Luers	0,934	0,906	0,949	0,869	0,840	0,830	0,871	0,755
Vulnérabilité proportionnelle	0,703	0,696	0,696	0,725	0,527	0,517	0,518	0,575
Vulnerability gap	0,772	0,761	0,765	0,783	0,601	0,592	0,597	0,641
Vulnerability severity	0,824	0,806	0,819	0,810	0,663	0,652	0,665	0,671

Il est intéressant de noter que le fait d'évaluer plus localement la vulnérabilité par rapport à une approche plus globale permet en théorie d'affiner les valeurs des indices. Cela nous a permis de mettre en évidence (Table 25) que la non prise en compte du contexte local (traduit par notre plan d'expériences OAT) entraîne une surestimation de la diminution de la vulnérabilité à travers l'indice de Luers et une sous-estimation de cette diminution avec l'indice de vulnérabilité "Most vulnerable individual". Tandis que les autres indices ne semblent que faiblement impactés par la prise en compte plus ou moins fine de la sensibilité. Cette approche de prise en compte locale de la sensibilité aurait pu être améliorée en ne tenant pas compte uniquement de la sensibilité locale de la gestion, mais en ajoutant par exemple la sensibilité au sol et au climat. On aurait également pu substituer un plan factoriel complet au plan OAT lors de l'évaluation "locale" de la vulnérabilité.

4.6.2. Recherche d'un minimum de vulnérabilité

Le principe de cette étape est de chercher à minimiser un indice de vulnérabilité, à climat et environnement fixés, en jouant sur les modalités de gestion de la prairie. Cependant comme l'on souhaite une solution robuste, il faudra que le calcul de notre indice de vulnérabilité intègre la robustesse (en fait, on utilisera la sensibilité qui est l'inverse de la robustesse). Trois approches ont été testées pour rechercher le minimum de vulnérabilité : une première approche « en force brute », une métaheuristique de recherche locale (algorithme de la première pente, "*first descent*" [Hansen et Mladenović, 2005]) et une méthode plus globale qu'est le recuit simulé [Kirkpatrick et al., 1983].

Recherche par "force brute"

Comme le temps nécessaire pour l'évaluation d'une valeur de vulnérabilité en un point donné de l'espace est quasiment non limitant, une première approche simple a été retenue. Elle s'apparente à une recherche en "force brute" et consiste à parcourir aléatoirement l'espace des possibilités de gestion, et en chacun de ces points d'évaluer la sensibilité au climat x environnement x gestion. C'est-à-dire qu'en chacun de ces points un plan croisé composé d'un LHD de taille 408 contenant le climat, le sol et la [CO₂], et multiplié par 12 petits deltas de gestions, permet d'évaluer la vulnérabilité en ce point.

Ainsi, un LHD de taille 408 pour cinq variables (trois années climatiques, la concentration en CO₂ atmosphérique et le type de sol) a été généré et optimisé parmi 2 000 000 par le critère maximin [Johnson et al., 1990] (*i.e.* maximisation du minimum de distance entre deux points du plan d'expériences). Ensuite, un plan croisé a été construit en combinant ce LHD avec plan "One factor At a Time" (OAT) [Kleijnen 1987] pour la gestion. Ce plan OAT se caractérise par ± 5 jours pour les dates de fauches, $\pm 5\%$ de légumineuses et $\pm 10 \text{ kg N ha}^{-1} \text{ an}^{-1}$, tout en respectant les bornes définies dans le design (Table 16) utilisé pour la régression des surfaces de réponse (dans le but d'éviter de sortir du domaine de validité des surfaces). Ce plan contient donc entre 7 et 13 éléments (en fonction du nombre de fauches). Le croisement de ces deux plans d'expériences crée alors un plan d'une taille de 2856 à 5304 points, sur lequel les indices pourront être évalués. La sensibilité à l'altitude étant considérée comme négligeable devant celles associées aux autres facteurs, elle sera ignorée dans notre design. Les évaluations ont été distribuées en 100 paquets de 10 000 combinaisons de gestion. Une stratégie plus efficace qu'un simple tirage aléatoire aurait pu consister à parcourir l'espace des paramètres de manière uniforme à l'aide d'un plan factoriel complet (en conservant, par exemple, les mêmes niveaux que ceux du design ayant servi à régresser les surfaces de réponses). Puis, une fois le point de plus faible vulnérabilité évaluée, nous pourrions alors éventuellement raffiner notre recherche en recréant un design autour de ce point. Cette approche en "force brute" nous a

permis d'évaluer les performances (exactitude et temps de calcul) des deux autres approches d'optimisation.

Il s'avère que par construction les indices de vulnérabilité "Vulnérabilité Proportionnelle", "Vulnerability Severity" et "Vulnerability Gap" sont extrêmement sensibles au seuil de vulnérabilité. En effet, l'indice se base sur les éléments en dessous du seuil de vulnérabilité. Or, il existe des zones de l'espace des paramètres dans lesquelles, toutes les valeurs totales de matière organique du sol sont supérieures au seuil de 75 tC/ha, et donc les indices valent 0. Nous pourrions éventuellement chercher à définir la frontière caractérisant la zone de matière organique supérieure au seuil. Bien que l'iso-quantité de matière organique au niveau du seuil soit intéressante (et notamment son évolution à travers le changement climatique), le but de cette étude est de montrer la faisabilité de notre approche d'analyse de vulnérabilité avec adaptation. Aussi, le seuil de vulnérabilité a été monté à la valeur arbitraire de 300 tC/ha de façon à ce que l'indice puisse être calculé en tout point de l'espace.

Recherche par la méthode de la première pente

La deuxième méthode utilisée pour rechercher le minimum de vulnérabilité est une méthode de recherche locale. Nous avons choisi l'algorithme de la première pente, "*first descent heuristic*". Le principe de cet algorithme est le suivant : on part d'une solution acceptable (dans notre cas particulier, n'importe quel point). Puis on évalue dans le voisinage de ce point des directions d'amélioration (minimisation) de l'indice de vulnérabilité. Une fois qu'une pente d'amélioration est déterminée, on se déplace dans cette direction. On itère l'opération jusqu'à ce qu'aucune direction ne donne d'amélioration. Dans notre cas le voisinage d'un point a été discrétisée par l'espace couvert par un plan factoriel complet à 3 niveaux des facteurs de gestion, privé du point central. Ces niveaux sont $\{-1 ; 0 ; +1\}$ pour le nombre de fauches et les dates de fauches, $\{-0,5\% ; 0 ; +0,5\%\}$ pour la fraction de légumineuses et $\{-1 ; 0 ; +1\}$ kg N/ha/an pour le niveau de fertilisation. Ainsi l'indice de vulnérabilité est évalué à l'aide du même design que dans la méthode précédente, jusqu'à ce qu'aucun de ces 2186 points (3^7-1 pour la position actuelle) ne donne d'amélioration. Le déplacement s'il a lieu se fait sur le premier point du plan factoriel à plus faible valeur d'indice de vulnérabilité.

Pour une description plus détaillée de cette métaheuristique ou des variantes (telle que l'algorithme du gradient), le lecteur pourra se référer à [Hansen et Mladenović, 2003] [Hansen et Mladenović, 2005]. L'avantage de ces heuristiques de recherche locale est leur rapidité (dans notre cas 3 400 évaluations en moyenne) pour une efficacité très dépendante des situations. Pour vérifier la convergence ou non de ces heuristiques, une vingtaine de chaînes par indice a été lancée. Le défaut majeur de ce genre de méthode est que l'on peut trouver un minimum local qui ne correspond pas au minimum globale recherché. En effet

une fois engagée dans la vallée d'un minimum local, il n'est plus possible d'en ressortir. Il n'est en effet pas possible de se diriger vers une solution moins bonne. C'est pourquoi une troisième méthode plus globale a été testée : le recuit simulé [Kirkpatrick et al., 1983].

Recherche par la méthode du recuit simulé

Le principe de cet algorithme est le suivant. On part d'un point aléatoire de l'espace, et à chaque itération on regarde un point du voisinage de notre position actuelle. Si ce nouveau point a une meilleure vulnérabilité (*i.e.* plus basse), on s'y déplace. Sinon, un tirage aléatoire selon une certaine probabilité permet de déterminer si l'on accepte ou non ce nouveau point. Cette probabilité est influencée à la fois par la différence entre la valeur de l'indice de vulnérabilité à la position candidate et celle à la position actuelle, et par la "température". Dans ce contexte, la "température" est une variable qui, à l'image de la température dans la physique moléculaire, va en refroidissant forcer le système dans une direction minimisant son énergie (pour nous l'indice de vulnérabilité). Il est prouvé que si l'algorithme respecte certaines conditions (vitesse de décroissance de la température, réversibilité et accessibilité), le système converge en probabilité [Hajek 1988]. Dans notre cas particulier, nous sommes partis d'un point aléatoire de l'espace, puis un candidat est choisi aléatoirement (loi uniforme) dans le voisinage formé de $\{-1 ; 0 ; +1\}$ pour le nombre de fauches et les dates de fauches, de $[-0,5\% ; +0,5\%]$ pour la fraction de légumineuses et de $[-1 ; +1]$ kg N/ha/an pour le niveau de fertilisation. La définition de ce voisinage, nous permet de respecter les conditions de réversibilité et d'accessibilité nécessaire au théorème de [Hajek 1988]. En revanche, le critère de vitesse de décroissance de la température ($T_n \geq \frac{k}{\ln(n+1)}$; où T_n est la température au pas de temps n , et k un scalaire positif) est difficile à respecter car entraînant une décroissance très lente et donc un trop grand nombre d'itérations. C'est pourquoi, une loi plus rapide de décroissance a été adoptée ($T_n = 0,999 T_{n-1}$) en partant d'une température initiale de 1. Une telle configuration se traduit par une acceptation très forte dans les premières itérations (par exemple, 99 % de chance d'accepter une augmentation de la valeur de l'indice de vulnérabilité de 0,01 pendant les dix premières itérations) et qui décroît fortement au fil des itérations (par exemple, 23 % et seulement 2×10^{-8} % de probabilité d'accepter une augmentation de la valeur de l'indice de vulnérabilité de 0,01 au bout de 5 000 et 10 000 itérations). La recherche continue jusqu'à stabilité (dans notre cas 5000 itérations sans amélioration). Pour vérifier la méthode, 10 réplifications de la méthode ont été lancées. Pour une description plus détaillée de cette méthode le lecteur pourra se référer à [Salamon et al, 2002] [Aarts et Korst, 1989] [Aarts et al., 2005]. Il est à noter également que cet algorithme revient à générer une chaîne de Markov [Aarts et Van Laarhoven, 1985] [Lundy et Mees, 1986].

Résultats

En se servant de la méthode de recherche par force brute comme référence, on peut comparer l'efficacité de la méthode de recherche par algorithme de la première pente et celle du recuit simulé. Nous avons pu noter que la méthode de recherche par l'algorithme de la première pente aboutit au même résultat (*i.e.* même minimum de vulnérabilité) avec une erreur absolue relative de $0,16 \% \pm 0,05\%$. De même les conditions de gestion pour ces minimums sont similaires [comparaisons non montrées]. Cependant, dans le cas particulier de l'indice de vulnérabilité proportionnelle, certaines chaînes ne convergent pas vers le minimum de vulnérabilité. En effet, compte-tenu de la valeur de seuil élevée, il existe des conditions pour lesquelles, toutes les valeurs de matières organiques à l'équilibre sont en dessous du seuil, dans le voisinage du point. Cela se traduit par une valeur de l'indice valant 1. De plus, il s'avère que dans ces conditions, un faible changement de gestion tel que défini par notre algorithme (*i.e.* dû à la définition de notre voisinage) ne permet pas de passer à une zone où au moins une valeur de matière organique est au dessus du seuil. On se trouve en réalité dans un plateau, ce qui a le même effet qu'un minimum local pour notre algorithme de recherche. Ce problème représente en moyenne 20% des cas.

De même, la méthode du recuit simulé permet de trouver le minimum de vulnérabilité avec une précision légèrement meilleure que la méthode dite de "force brute", pour toutes les périodes et les indices (gain relatif de $-0,22 \% \pm 0,06\%$), pour un coût moyen de 22 000 évaluations pour les indices de Luers, de "Vulnerability Gap" et "Vulnerability Severity". Bien que la méthode du recuit simulé assure une grande précision dans tous les cas (à la différence de la méthode de la première pente) pour coût assez peu élevé (~22 000 vs 1 000 000 pour la "force brute"), elle est quand même impactée par la particularité de l'indice de Vulnérabilité proportionnelle. Ce dernier est un ratio du nombre d'éléments au dessus du seuil et donc forme localement dans l'espace des paliers. Cela induit un problème de convergence, ainsi 80 % des recherches avec cet algorithme, et pour cet indice, se sont arrêtées au nombre maximum d'évaluations autorisées (*i.e.* 100 000), cependant la solution trouvée reste optimale. La vitesse et l'acuité de ces méthodes d'optimisation sont résumées dans la Figure 57.

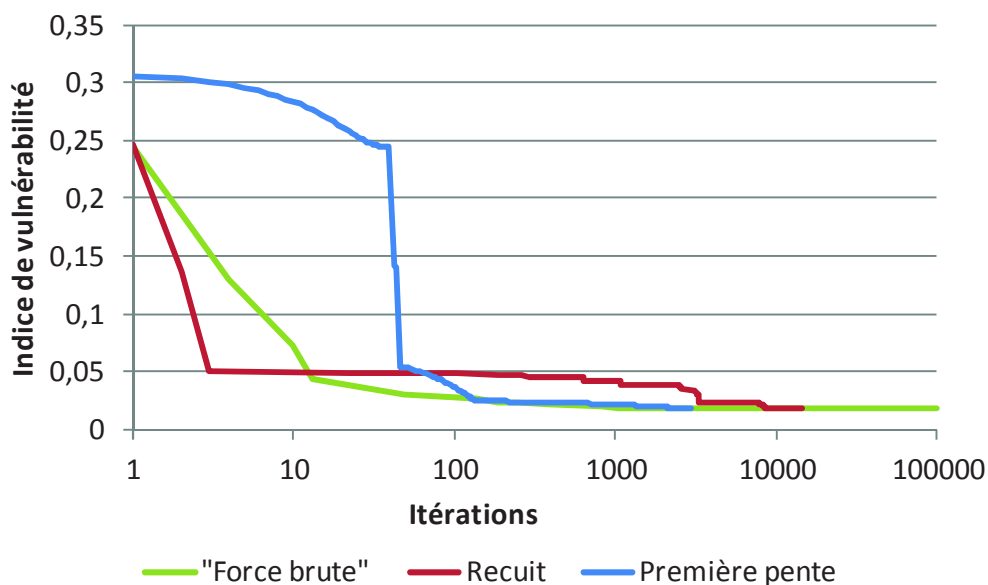


Figure 57. Exemple de convergence des algorithmes d'optimisation ("force brut" en vert, l'algorithme de la première pente en bleu et l'algorithme du recuit simulé en rouge) sur la période climatique de référence (1970-2006) et l'indice de "Severity Vulnerability".

La recherche d'un minimum de vulnérabilité a permis de mettre en évidence certaines limites de ces indices de vulnérabilité. Il est ainsi incorrect d'utiliser, pour évaluer la vulnérabilité, un indice basé sur un seuil potentiellement toujours dépassé dans certaines zones de l'espace. L'analyse des résultats ainsi obtenus (Table 26, Table 27) met en évidence qu'une gestion extensive induit une plus faible vulnérabilité de la matière organique du sol au changement climatique. C'est pourquoi, pour tous les indices et toutes les périodes, on trouve une gestion "optimale" se caractérisant par une seule fauche avec la date la plus avancée possible (donc moins de biomasse prélevée) et un niveau de fertilisation maximal (c'est-à-dire les 120 kg N ha⁻¹ an⁻¹ autorisés par le design) (Table 26). En revanche la fraction de légumineuse (Table 27) n'est pas "collée" à une des bornes du domaine du plan d'expériences (0-40%), elle est comprise entre 0,3% et 16% en fonction de l'indice de vulnérabilité et de la période climatique considérée. Cette gestion optimale est cohérente avec l'état de l'art (e.g., [Conant et al., 2001]). Il est intéressant de noter que le niveau d'entrée des surfaces de réponse a peu d'impacts sur la gestion obtenue (par exemple, on obtient une erreur absolue moyenne de 0,92% ± 0,15% pour la fraction de légumineuses). En revanche, le changement climatique induit un glissement visible de la fraction optimale pour l'indice de vulnérabilité proportionnelle (de 16 à 11%) et pour l'indice de Luers (de 0,3 à 7,8%). Cela traduit vraisemblablement un changement dans la sensibilité de la matière organique. C'est-à-dire, pour la vulnérabilité proportionnelle, moins de risques d'être en-dessous du seuil pour une fraction de 11-12% et pour l'indice de Luers un coefficient de variation meilleur à 8% de légumineuses. Ces valeurs de gestion traduisent probablement le fait que, pour la matière organique du sol à l'équilibre, les zones extensives sont également des zones à faible sensibilité, ou tout au moins des zones

où la forte quantité de matière organique compense la sensibilité au climat, à la gestion et ce pour toute une gamme de sols.

Table 26. **Quantité de fertilisant minéral annuel** (kg N/ha/an) dans la gestion minimisant la vulnérabilité, en fonction de l'indice de vulnérabilité, de la période climatique ("Période de Référence", PR : 1970-2006 ; "Futur Proche", FP : 2018-2051 ; "Futur Lointain", FL : 2067-2100) et de la surface de réponse utilisée (avec entrée climatique P-ET₀ ou Saison). Résultats obtenus avec la méthode de la "force brute".

	PR		FP		FL	
	P-ET ₀	Saison	P-ET ₀	Saison	P-ET ₀	Saison
Vulnerability Severity	119,7	119,9	120,0	119,9	119,9	119,8
Vulnerability Gap	120,0	119,7	120,0	120,0	119,4	119,8
Vulnérabilité Proportionnelle	119,8	119,9	119,7	119,6	120,0	119,9
Indice de Luers	119,3	120,0	119,9	119,9	119,8	119,8

Table 27. **Pourcentage de légumineuses** dans la gestion minimisant la vulnérabilité, en fonction de l'indice de vulnérabilité, de la période climatique ("Période de Référence", PR : 1970-2006 ; "Futur Proche", FP : 2018-2051 ; "Futur Lointain", FL : 2067-2100) et de la surface de réponse utilisée (avec entrée climatique P-ET₀ ou Saison). Résultats obtenus avec la méthode de la première pente.

	PR		FP		FL	
	P-ET ₀	Saison	P-ET ₀	Saison	P-ET ₀	Saison
Vulnerability Severity	8,2%	9,0%	8,2%	8,8%	9,0%	9,4%
Vulnerability Gap	10,9%	10,8%	9,9%	10,5%	9,5%	10,5%
Vulnérabilité Proportionnelle	16,0%	14,2%	12,9%	12,2%	11,0%	11,9%
Indice de Luers	0,3%	1,2%	2,2%	3,8%	6,2%	7,8%

Il est intéressant de noter que d'après les surfaces de réponses, le gain par adaptation augmente avec le changement climatique (Table 28). Bien que le gain soit très variable d'un indice à l'autre, il souligne un potentiel d'adaptation grandissant dans le futur. Ainsi, par exemple, d'après la vulnérabilité proportionnelle, une adaptation de la gestion ne permet actuellement de réduire que de 40% le nombre de cas vulnérables, tandis que dans le "futur lointain", ce nombre peut être réduit de 90%. Il est à noter cependant que l'indice de Luers ne permet pas de conclure à une véritable augmentation du potentiel d'adaptation. Toujours dans l'hypothèse où les surfaces de réponse sont suffisamment fiables, nous pouvons observer une réduction de la vulnérabilité avec prise en compte de l'adaptation (Table 29, Table 30). Ce qui est cohérent avec une augmentation de l'efficacité de l'adaptation (Table 28) et une baisse de la vulnérabilité sans adaptation avec le changement climatique. Bien que les indices bougent dans le même sens, leurs amplitudes diffèrent à

cause de leurs représentations différentes des aspects vulnérables. On peut ainsi supposer que la sensibilité au seuil explique en partie les forts changements de ces indices par rapport à l'indice de Luers. On peut également supposer que l'augmentation du niveau général de teneur en matière organique du sol est plus impactant sur le calcul de ces indices que sur celui de Luers (bien que ce dernier en tienne également compte), à cause d'un effet de compensation de la sensibilité.

Table 28. **Gain d'adaptation** pour les périodes climatiques "Période de Référence" (PR : 1970-2006), "Futur Proche" (FP : 2018-2051) et "Futur Lointain" (FL : 2067-2100), en fonction de la surface de réponse utilisée (avec entrée climatique P-ET₀ ou Saison). Les valeurs correspondent à la **différence de l'indice de vulnérabilité sans et avec adaptation**, le tout normé par la valeur de l'indice sans adaptation pour la période en question. Plus la valeur est haute plus le gain par adaptation est important.

	PR		FP		FL	
	P-ET ₀	Saison	P-ET ₀	Saison	P-ET ₀	Saison
Vulnerability Severity	92,9%	91,2%	98,1%	97,0%	99,7%	99,2%
Vulnerability Gap	81,5%	79,3%	93,2%	91,3%	98,4%	97,1%
Vulnérabilité Proportionnelle	37,9%	41,0%	69,5%	65,3%	89,5%	86,4%
Indice de Luers	76,1%	73,1%	77,1%	74,3%	76,9%	74,3%

Table 29. **Vulnérabilité avec prise en compte de l'adaptation**, pour les différentes périodes climatiques ("Période de Référence", PR : 1970-2006 ; "Futur Proche", FP : 2018-2051 ; "Futur Lointain", FL : 2067-2100) et en fonction de la surface de réponse utilisée (avec entrée climatique P-ET₀ ou Saison). Les valeurs ont été **normées par la valeur de l'indice sans adaptation à la période de référence**. Plus la valeur est haute plus le système est vulnérable.

	PR		FP		FL	
	P-ET ₀	Saison	P-ET ₀	Saison	P-ET ₀	Saison
Vulnerability Severity	7,09%	8,80%	1,52%	2,32%	0,18%	0,48%
Vulnerability Gap	18,47%	20,68%	5,68%	7,25%	1,06%	2,00%
Vulnérabilité Proportionnelle	62,08%	59,02%	28,35%	32,28%	8,82%	11,51%
Indice de Luers	23,86%	26,89%	20,02%	22,41%	17,59%	19,51%

Table 30. **Vulnérabilité avec prise en compte de l'adaptation**, pour les périodes climatiques "Futur Proche" (FP : 2018-2051) et "Futur Lointain" (FL : 2067-2100), en fonction de la surface de réponse utilisée (avec entrée climatique P-ET₀ ou Saison). Les valeurs ont été **normées par la valeur de l'indice avec prise en compte de l'adaptation à la "Période de Référence" (PR : 1970-2006)**. Plus la valeur est haute plus le système est vulnérable.

	FP		FL	
	P-ET ₀	Saison	P-ET ₀	Saison
Vulnerability Severity	21,5%	26,3%	2,5%	5,5%
Vulnerability Gap	30,8%	35,1%	5,8%	9,7%
Vulnérabilité Proportionnelle	45,7%	54,7%	14,2%	19,5%
Indice de Luers	83,9%	83,3%	73,7%	72,6%

Discussion

Pour être plus efficace que la première méthode (force brute), tout un ensemble de méthodes existent. Dans notre cas, nous n'avons testé que deux métaheuristiques [Glover 1986], une locale (algorithme de la première pente) et une globale (recuit simulé). Dans le cas où l'espace serait uniquement discret, il aurait pu être intéressant d'utiliser des algorithmes génétiques [Holland 1975]. L'avantage des métaheuristiques⁸⁵ est de pouvoir relativement facilement incorporer les connaissances des experts sur le système et donc d'être possiblement les plus efficaces en termes de vitesse de convergence. D'autre part, la recherche de minimisation de vulnérabilité devrait également être limitée par la capacité d'adaptation réelle (par exemple, impossibilité d'augmenter le nombre de fauche). Il sera sans doute plus facile d'inclure ce genre d'information dans une métaheuristique que dans les méthodes analytiques/exactes, où la capacité d'adaptation se traduira par une restriction du domaine de validité et par une distance maximale d'éloignement par rapport à la position initiale. La contrepartie des métaheuristiques est généralement une efficacité plus variable selon le contexte (trouver une solution optimale n'est pas garanti contrairement aux méthodes exactes) et un paramétrage (notamment l'initialisation) à adapter à chaque cas. Il sera intéressant d'implémenter sous EMF deux ou trois types de méthodes de recherche de minimum de vulnérabilité (classe Optimisation) et ainsi de générer les scripts pour R et pour OpenMOLE permettant la distribution des calculs (dans la mesure où la méthode le permet).

Les résultats obtenus sont à prendre avec précaution, car ils ont été obtenus à partir de surfaces de réponses. Afin d'affirmer la validité des résultats, il conviendrait d'abord de vérifier que l'on obtient les mêmes résultats avec de véritables simulations à l'aide de

⁸⁵ Le lecteur intéressé pourra se référer à [Burke et Kendall, 2005]

PaSim. D'une part, PaSim, comme tout modèle, contient des erreurs structurelles. D'autre part, les simulations ont été réalisées à fraction de légumineuses constante, et ne tiennent donc pas compte de la compétition entre graminées et légumineuses. Or, avec une seule fauche annuelle comme gestion, la végétation doit normalement atteindre une hauteur à laquelle les plantes fixatrices (légumineuses) sont pénalisées par rapport aux plantes non fixatrices (par exemple les graminées) [Woledge et al., 1992] [Soussana et al., 1995] [Faurié et al., 1996] entraînant une diminution de cette fraction. Aussi une fraction de 16% n'est peut-être pas réalisable avec une telle gestion. De façon à confirmer les résultats précédemment décrits une dernière phase d'évaluation du minimum de vulnérabilité a lieu à l'aide du modèle agro-écologique (ici PaSim) (voir Figure 25). Cette phase est décrite dans la partie suivante.

4.6.3. Evaluation du minimum de vulnérabilité

À partir des gestions précédemment trouvées, nous pouvons évaluer la vulnérabilité avec adaptation (avec capacité d'adaptation non limitante dans notre cas). Les résultats de la Table 24 laissent à penser que l'évaluation obtenue à partir des surfaces de réponses sera proche de celles obtenues à l'aide de PaSim. Ainsi, comme prévu dans notre démarche d'analyse de vulnérabilité, nous avons évalué ce minimum à l'aide des simulations avec PaSim. C'est-à-dire que pour chacune des gestions trouvées par minimisation de la vulnérabilité à l'aide des surfaces de vulnérabilité, nous avons lancé des simulations permettant de couvrir le design d'évaluation de la vulnérabilité localement (*i.e.* le plan croisé formé d'un LHD pour le climat et le sol et d'un plan OAT pour la gestion). Les simulations de mise à l'équilibre ont été distribuées sur le cluster du LIMOS permettant ainsi de réaliser l'équivalent de 2 ans de calculs monoprocesseur en quelques jours. La Figure 58 illustre la comparaison entre les valeurs obtenues par PaSim et par les surfaces de réponse. Les surfaces de réponse sous-estiment les valeurs des indices de vulnérabilité, cependant les tendances sont conservées. La corrélation entre indices calculés par les surfaces et par PaSim peut se traduire par une équation de type puissance entre ces derniers. Le coefficient positif dans la puissance illustre le fait que les plus petites valeurs de vulnérabilité sont proportionnellement plus sous-estimées que les plus grandes valeurs (puisque dans notre cas les indices sont inférieurs à 1)⁸⁶. Cependant, même si la représentation de l'espace par les surfaces de réponse est déformée, on peut quand même supposer que les surfaces de réponse contiennent les informations nécessaires à notre minimisation de la vulnérabilité. Il est très probable que l'erreur des surfaces de réponse soit liée à la nature de ces dernières (dans notre cas quadratique).

⁸⁶ Par construction seul l'indice de Luers aurait pu être supérieur à 1. Dans notre cas cela aurait uniquement pu se produire si l'écart-type avait été supérieur au carré de la moyenne divisé par la valeur du seuil.

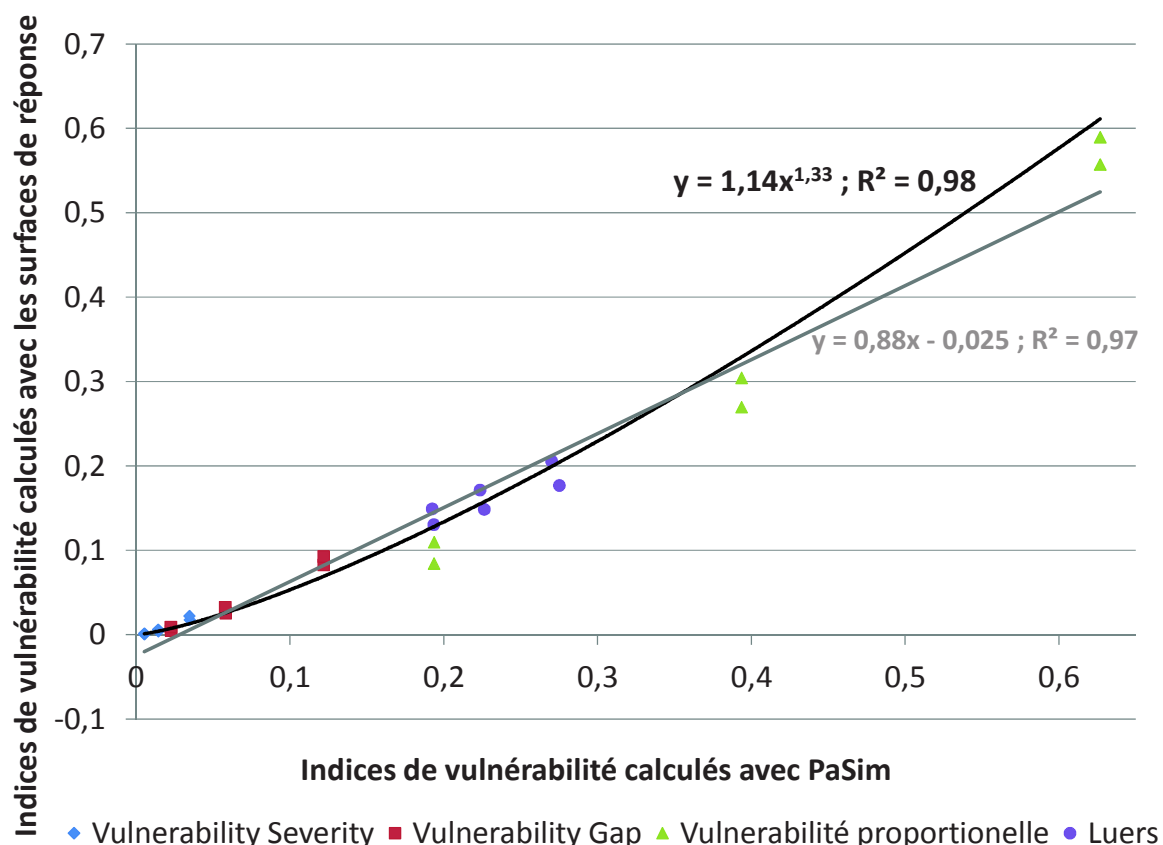


Figure 58. Comparaison des valeurs des indices de vulnérabilité avec prise en compte de l'adaptation, obtenues à l'aide des surfaces de réponse (avec entrée climatique P-ET₀ ou Saison), en ordonnée, et avec PaSim en abscisse, pour les différentes périodes climatiques ("Période de Référence", PR : 1970-2006 ; "Futur Proche", FP : 2018-2051 ; "Futur Lointain", FL : 2067-2100).

Bien que les valeurs absolues des indices obtenus par les simulations avec PaSim diffèrent de ceux obtenus sur les surfaces de réponse, il est tout de même intéressant de noter que les tendances en réduction de vulnérabilité sont conservées (Table 31, Table 32, Table 33). Les simulations de PaSim à défaut de pouvoir confirmer qu'il s'agit réellement de minima de vulnérabilité, permettent quand même d'affirmer que la gestion "optimisée" permet une réduction importante de la vulnérabilité (par exemple, Table 33, entre 68% et 96% de réduction pour la période de futur lointain). On retrouve un effet fort du changement climatique sur la vulnérabilité de la matière organique (Table 31, Table 32). Les simulations à l'aide de PaSim confirment (Table 33) également la faible amélioration du gain par adaptation selon la période climatique pour l'indice de Luers (de 65% à 68%).

Table 31. **Vulnérabilité avec prise en compte de l'adaptation, relativement à la vulnérabilité sans adaptation à la période de référence.** Les indices ont été évalués par PaSim aux « minima » trouvés par les différentes surfaces de réponse (avec entrée climatique P-ET₀ ou Saison), pour les différentes périodes climatiques ("Période de Référence", PR : 1970-2006 ; "Futur Proche", FP : 2018-2051 ; "Futur Lointain", FL : 2067-2100). Les valeurs ont été normées par la valeur de l'indice sans adaptation à la période de référence. Plus la valeur est haute plus le système est vulnérable.

	PR		FP		FL	
	P-ET ₀	Saison	P-ET ₀	Saison	P-ET ₀	Saison
Vulnerability Severity	13,7%	13,7%	5,7%	5,6%	2,1%	2,1%
Vulnerability Gap	27,0%	27,0%	12,9%	12,8%	5,0%	5,0%
Vulnérabilité Proportionnelle	66,7%	66,7%	41,9%	41,9%	20,6%	20,7%
Indice de Luers	34,7%	34,1%	28,6%	28,2%	24,4%	24,3%

Table 32. **Vulnérabilité avec prise en compte de l'adaptation, relativement à la vulnérabilité avec adaptation à la période de référence.** Les indices ont été évalués par PaSim aux « minima » trouvés par les différentes surfaces de réponse (avec entrée climatique P-ET₀ ou Saison), pour les périodes climatiques "Futur Proche" (FP : 2018-2051) et "Futur Lointain" (FL : 2067-2100). Les valeurs ont été normées par la valeur de l'indice avec prise en compte de l'adaptation à la "Période de Référence" (PR : 1970-2006), calculée à l'aide des résultats de PaSim. Plus la valeur est haute plus le système est vulnérable.

	FP		FL	
	P-ET ₀	Saison	P-ET ₀	Saison
Vulnerability Severity	41,3%	41,3%	15,1%	15,2%
Vulnerability Gap	47,7%	47,5%	18,6%	18,5%
Vulnérabilité Proportionnelle	62,8%	62,8%	30,9%	31,0%
Indice de Luers	82,2%	82,8%	70,3%	71,3%

Table 33. **Gain d'adaptation** pour les périodes climatiques "Période de Référence" (PR : 1970-2006), "Futur Proche" (FP : 2018-2051) et "Futur Lointain" (FL : 2067-2100). Les indices ont été **évalués par PaSim aux « minima »** trouvés par les différentes surfaces de réponse (avec entrée climatique P-ET₀ ou Saison). Les valeurs correspondent à la **différence de l'indice de vulnérabilité sans et avec adaptation**, le tout normé par la valeur de l'indice sans adaptation pour la période en question, calculée à l'aide des résultats de PaSim. Plus la valeur est haute plus le gain par adaptation est important.

	PR		FP		FL	
	P-ET ₀	Saison	P-ET ₀	Saison	P-ET ₀	Saison
Vulnerability Severity	86,3%	86,3%	92,8%	92,8%	96,5%	96,5%
Vulnerability Gap	73,0%	73,0%	84,6%	84,7%	92,7%	92,7%
Vulnérabilité Proportionnelle	33,3%	33,3%	55,1%	55,1%	75,7%	75,7%
Indice de Luers	65,3%	65,9%	67,1%	67,5%	67,7%	67,8%

Dans le but d'évaluer si les minima trouvés par optimisation correspondent réellement à des minima, nous avons calculé tous les indices par chacune des gestions. On peut noter que l'écart relatif à la meilleure valeur d'indice est en moyenne de 1,51% \pm 0,36%. Il est aussi intéressant de noter [données non montrées] que la gestion trouvée pour l'indice de vulnérabilité proportionnelle s'avère dans 95% des cas la plus optimale des gestions. Bien que les écarts de valeurs de vulnérabilité soient faibles, la fraction de légumineuse proposée pour la gestion "optimale" selon l'indice de Luers, s'avère en réalité dans tous les cas la "moins optimale" des gestions. Il pourrait également être envisagé d'effectuer d'autres simulations avec PaSim de façon à mieux évaluer la robustesse des minima.

Si le minimum de vulnérabilité ainsi trouvé ne nous convient pas, alors nous pourrions inclure les résultats des simulations que nous venons d'effectuer dans le calcul des régressions pour les surfaces de réponse. Il peut également être envisagé de faire un plan d'expériences couvrant des zones supposées comme *a priori* contenant les minima de vulnérabilité. Une fois les nouvelles surfaces de vulnérabilité réalisées, nous pouvons recommencer à l'étape de minimisation de la vulnérabilité.

4.6.4. Discussion

La validité de ces résultats est réduite par la non prise en compte de la corrélation « sol x climat ». En effet la répartition des sols et des climats n'est pas homogène sur la France. Ainsi certaines combinaisons « sol x climat » sont plus probables que d'autres. S'il est cohérent pour régresser les surfaces de réponses de considérer toutes les occurrences de climat représentatives de la France et ainsi, de les décorréler spatialement lors de l'évaluation de la vulnérabilité, il conviendrait d'inclure ce genre d'information (*i.e.* exposition).

Il est à noter que même si cette analyse utilise la matière organique à l'équilibre, elle nous donne des informations sur la vulnérabilité de la matière organique réelle. En effet, la matière organique à l'équilibre est un potentiel vers lequel le système va tendre (qu'il soit au-dessus ou au-dessous). Aussi une augmentation ou une diminution de la vulnérabilité de la matière organique à l'équilibre, traduit la même tendance pour la teneur réelle en matière organique du sol.

Cette étude se voulant volontairement simple, une seule variable a été considérée (la matière organique du sol) et donc un seul objectif de minimisation. De plus une capacité d'adaptation non limitante a été utilisée. Il aurait cependant été plus cohérent de considérer qu'il existe quatre systèmes distincts avec chacun un nombre différent de fauches et de considérer qu'au sein de ces sous-systèmes il n'est pas possible d'augmenter ou de diminuer le nombre de fauches. Avec cette même approche, nous aurions pu également partir d'une situation réelle (par exemple, une gestion bien définie) et chercher à minimiser la vulnérabilité de la matière organique, tout en respectant une certaine cohérence (par exemple en introduisant une relation entre la teneur maximale de légumineuses et la quantité de fertilisation). Dans une approche plus complexe, où plus d'une variable aurait été considérée, nous aurions pu chercher à minimiser la vulnérabilité de la matière organique tout en respectant certains niveaux de production. De même, il aurait été possible de combiner/pondérer les indices de vulnérabilité.

Malgré certaines limitations, ce travail nous permet quand même de conclure sur la viabilité de notre approche d'analyse de vulnérabilité avec adaptation. Cette étude pourrait servir comme exemple de la démarche en vue d'effectuer d'autres analyses de vulnérabilité. Cette démarche est dans une certaine mesure générique, c'est-à-dire qu'elle est applicable à tout type de modèle agroécologique et pas uniquement aux modèles prairiaux. Cette méthode nécessite cependant certaines adaptations propres aux variables considérées et aux modèles utilisés. Parmi ces adaptations, on trouve notamment la définition du seuil de vulnérabilité (à adapter à chaque variable d'intérêt), mais également certains paramètres liés aux méthodes d'optimisation (par exemple, définition du voisinage, nombre maximum d'itérations).

Cette partie, en plus d'illustrer l'étape de recherche et d'analyse du minimum de vulnérabilité, nous a permis de réaliser des exemples de scripts R. Ces scripts pour la recherche du minimum de vulnérabilité, nous permettrons plus tard de proposer des améliorations dans le plugin EMF, par exemple, l'implémentation des méthodes de recherches de minimum de vulnérabilité, à l'aide de scripts R.

4.7. Conclusion

Nous avons vu dans ce chapitre trois exemples d'analyses de vulnérabilité. Tout d'abord, nous nous sommes intéressés, en guise d'exemple préliminaire, au cas de l'analyse de vulnérabilité au changement climatique d'une prairie permanente auvergnate, à l'aide du modèle ModVégé. Cette première analyse avait pour but de prouver la validité conceptuelle de notre démarche d'évaluation de la vulnérabilité. Elle justifiait également l'intérêt de l'utilisation d'un ensemble complémentaire d'indices quantitatifs pour évaluer la vulnérabilité et d'un plan d'expériences conçu pour tenir compte des différents niveaux d'incertitudes associés au système étudié. Enfin, elle permettait de concevoir différents plans d'expériences, à partir desquels nous avons pu, dans le chapitre précédent, proposer des modèles de DOE (*Design Of Experiment*) et ainsi affiner notre métamodèle.

Dans un deuxième temps, nous avons étudié l'application d'un jeu d'indices quantitatifs sur des simulations spatialisées de prairies permanentes à l'échelle de l'Europe, à l'aide du modèle PaSim. Cet exemple d'analyse de vulnérabilité sans adaptation a permis à la fois de conforter l'utilisation d'un jeu d'indices et de mettre en avant les spécificités de ce genre d'analyse (par exemple, prise en compte des pixels voisins ou problématique du temps de calcul).

Ayant réalisé deux applications d'analyses de vulnérabilité sans adaptation, nous avons alors décidé de documenter la démarche complète d'analyse de vulnérabilité avec et sans adaptation. Le cas particulier de la matière organique à l'équilibre pour les prairies permanentes françaises a été retenu pour sa relative facilité de mise en œuvre (disponibilité des données, temps de calcul raisonnables à l'aide du modèle PaSim). Après avoir préalablement vérifié que trois années suffisaient dans le cycle climatique servant, à l'aide du modèle PaSim, à mettre en équilibre la matière organique du sol en fonction de la gestion prescrite et que l'ordre des années était sans importance, nous avons construit des surfaces de réponse de la matière organique du sol à l'équilibre pour la période climatique actuelle et deux périodes futures. Après avoir évalué ces surfaces de réponse et vérifié qu'elles contenaient les informations nécessaires à notre problématique, nous avons réalisé le calcul de la vulnérabilité sans, puis avec prise en compte de l'adaptation. La prise en compte de l'adaptation a permis de comparer l'efficacité de trois algorithmes d'adaptation (*i.e.* de minimisation de la vulnérabilité : "Force brute", "Recuit simulé" et "Première pente"). En plus d'illustrer cet aspect particulier de notre méthode d'analyse de vulnérabilité, le travail réalisé dans ce chapitre nous a permis de concevoir des exemples de scripts qui nous serviront dans l'implémentation future des méthodes de recherche de minimum de vulnérabilité dans le plugin EMF (*Eclipse Modeling Framework*).

La réalisation de l'ensemble des résultats de ce chapitre s'est accompagnée d'un important temps de calcul (à peu près 24 ans de calcul mono-cœur sur CPU « moderne »). Il a ainsi été nécessaire de faire appel au calcul intensif à travers l'utilisation du cluster du

LIMOS et de la grille de calcul européenne (en utilisant l'organisation virtuelle « Biomed »). La distribution des calculs a en partie été réalisée à l'aide de l'outil OpenMOLE.

En conclusion, ce chapitre a permis à la fois d'illustrer, à travers trois exemples, notre démarche d'analyse de vulnérabilité, les outils utilisés (clusters, grille, conception et analyse de plans d'expériences), mais également de vérifier le caractère opérationnel des scripts générés par le plugin EMF présenté dans le chapitre précédent.

Conclusion générale

Après un rappel des buts poursuivis et de la démarche appliquée dans le cadre de la thèse, nous reviendrons sur les principales conclusions, observations et propositions qui émergent de cette étude sur l'analyse de vulnérabilité des prairies au changement climatique. Enfin, nous terminerons par la présentation de quelques pistes de réflexion pour l'avenir, concernant l'évolution des outils d'analyse de vulnérabilité et leurs applications dans des contextes différents.

Retour sur la démarche

À l'origine, le but premier de cette étude doctorale était d'ordre très pratique : fournir des outils conceptuels et techniques aux scientifiques œuvrant dans le domaine des impacts des changements climatiques et de l'adaptation à ceux-ci, afin d'aboutir à un instrument d'aide à l'analyse de la vulnérabilité au changement climatique à l'échelle du système de production agricole. Dans la continuité des études scientifiques récentes (e.g., [Graux et al., 2012ab]), menées en France et en Europe, qui se sont intéressées à l'évolution sous changement climatique des services fournis par les prairies et à la contribution des systèmes prairiaux au réchauffement global, le système prairial a été choisi comme objet d'application de notre travail de recherche. En effet un certain nombre d'études visent à concevoir et à évaluer des mesures d'atténuation des émissions de gaz à effet de serre et/ou des mesures d'adaptation aux changements climatiques. Le cadre applicatif de cette thèse est donc celui des changements climatiques qui affectent les systèmes prairiaux gérés par l'homme.

Dans ce cadre, l'originalité de notre travail de thèse a été d'élaborer une méthode intégrative et de montrer ses avantages potentiels dans l'évolution d'applications existantes en agro-écologie, en combinant notamment différentes techniques informatiques (par exemple, la rétro-ingénierie utilisant l'Ingénierie Dirigée par les Modèles et le calcul intensif) et des méthodes du domaine de la modélisation et de la simulation (plans d'expériences, métriques synthétiques, distributions de probabilités) pour l'analyse de vulnérabilité.

Conclusions, observations et propositions

Conceptualisation et méthode

La vulnérabilité est le degré de stress ou de perturbation qu'un système donné (humain ou environnemental) est capable de supporter avant d'être endommagé. La littérature sur le changement climatique contient de nombreuses explications de la vulnérabilité, basée sur la notion de sensibilité et faisant éventuellement appel à des idées plus complexes, depuis la prise en compte de l'exposition jusqu'aux impacts résiduels du changement climatique après mise en place de mesures d'adaptation. Une connaissance accrue des éléments et des enjeux de chaque notion dans un contexte de changement climatique nous paraissait la meilleure façon de parvenir à une conceptualisation plus approfondie de la vulnérabilité, ne négligeant pas les incertitudes et posant les bases indispensables à toute méthode ultérieure. Cette conceptualisation est essentielle car, même si un certain nombre de méthodes sont disponibles pour caractériser la vulnérabilité, réaliser une analyse de vulnérabilité complète demeure un problème difficile, compte-tenu du manque de consensus sur la signification même du terme, mais surtout parce que la vulnérabilité n'est pas un phénomène directement observable.

Or, à notre connaissance, aucun outil n'a été conçu pour faciliter la réalisation des différentes étapes qui constituent une analyse de vulnérabilité complète. C'est pourquoi, afin de faciliter la réalisation d'analyses, nous avons décidé de développer une méthode d'analyse de vulnérabilité et de créer un outil permettant sa mise en œuvre. La méthode et l'outil se veulent le plus génériques possible par rapport aux modèles utilisés et aux hypothèses sous-jacentes à l'analyse de vulnérabilité. De plus, cet outil permet la conception d'analyses de vulnérabilité qui suivent la démarche proposée dans cette thèse, et il permet également de gérer la distribution et l'interprétation de simulations basées sur les modèles agro-écologiques nécessaires pour ce type d'études.

Après avoir établi un modèle des concepts de la vulnérabilité par rapport au changement climatique, une méthode d'analyse de vulnérabilité en deux étapes a été conçue (3.2). Premièrement, une étape d'analyse de sensibilité, dont le but est de permettre d'estimer la vulnérabilité sans adaptation et de calculer une surface de réponse. Deuxièmement, la méthode cherche à minimiser la vulnérabilité sous contrainte de la capacité d'adaptation réelle, à partir de la surface de réponse créée précédemment.

Conception guidée par l'Ingénierie Dirigée par les Modèles

Par nature et compte-tenu de la cascade d'incertitudes (scénarios climatiques, modèles climatiques, méthodes de régionalisation, modèles d'impacts, ...) une analyse de

vulnérabilité va comporter beaucoup de simulations. Sans l'utilisation d'un plan d'expériences (DOE) approprié et du calcul intensif, une analyse de vulnérabilité qui suit la méthode décrite précédemment serait impossible. Certains outils permettent déjà la distribution de plans d'expériences tout en masquant l'hétérogénéité de la plateforme. Cependant, pour les adapter à l'analyse de vulnérabilité, nous avons commencé, grâce à une approche de rétro-ingénierie (3.3), par concevoir les modèles des entrées de logiciels tels que PaSim et les modèles des plans expérimentaux déjà réalisés (ceux présentés et réalisés préalablement à cette thèse étant relativement similaires à ceux requis pour une analyse de vulnérabilité). Cette étude nous a conduits à proposer un métamodèle des plans d'expériences dans notre contexte. Puis, pour raffiner nos modèles et métamodèles, nous avons proposé des expérimentations en fonction du niveau d'incertitudes que l'on souhaite prendre en compte (3.4). Ensuite, pour chacune de ces expérimentations nous avons proposé un modèle. Ceci nous a permis d'affiner notre proposition de métamodèles pour des modèles agro-écologiques, avec plans expérimentaux associés, dans le cadre d'une analyse de vulnérabilité. Enfin, nous avons combiné ces métamodèles (modèle agro-écologique, plan expérimental et analyse de vulnérabilité) pour former un métamodèle complet (3.5), de telle sorte que toute analyse de vulnérabilité à l'aide d'un modèle agro-écologique et sous contrainte de plans d'expériences soit conforme à ce métamodèle. Ce métamodèle a ensuite été implémenté dans EMF sous forme de plugin Eclipse de façon à créer un outil permettant la réalisation et l'automatisation d'une analyse de vulnérabilité. Cet outil est l'implémentation logicielle de la méthode d'analyse de vulnérabilité proposée dans le cadre de ce travail de thèse. Il est modulaire, car il y a la possibilité de ne réaliser seulement que certains aspects de l'analyse de vulnérabilité (par exemple juste l'analyse de sensibilité). Mais il est également générique dans la mesure où il peut être utilisé pour toute analyse de vulnérabilité à l'aide d'un modèle agro-écologique et ce quel que soit le plan d'expériences utilisé. En effet, du fait de la conformité de l'analyse de vulnérabilité avec le métamodèle, l'outil est opérationnel quel que soit le modèle agro-écologique et les hypothèses sous-jacentes (qui viennent notamment influencer le plan d'expériences).

Indices quantitatifs de vulnérabilité et prise en compte des incertitudes et des adaptations

Un ensemble de simulations ont été exécutées de façon à tester et à illustrer notre proposition de démarche d'analyse de vulnérabilité et son implémentation logicielle permettant d'en faciliter la réalisation et l'interprétation. Les résultats des applications (chapitre 4) servent avant tout à mettre en avant les forces et les faiblesses de la méthode d'analyse de vulnérabilité proposée.

Il s'agit, en effet, d'une approche quantitative basée sur un ensemble d'indices, chacun reflétant un aspect particulier de la vulnérabilité. De plus une approche par indices

quantitatifs (à la différence par exemple d'une approche par densité de probabilité) est relativement appropriée à la diffusion vers un public plus large que la communauté scientifique. La principale faiblesse de cette approche réside dans la sensibilité des indices à la période de référence et au seuil de vulnérabilité (à l'exception de celui de Luers [Luers et al., 2003] qui n'est sensible qu'à la période de référence). Cette faiblesse est en partie inhérente au concept même de vulnérabilité. En effet la vulnérabilité est toujours une notion relative [Downing et al., 2001]. De plus, pour être appliqués à certaines variables, certains indices nécessitent que ces dernières soient préalablement transformées. En effet, les variables d'état utilisées doivent représenter le bien-être du système. En tant que tel, plus la valeur de la variable est haute, meilleur est l'état du système, tandis que la valeur 0 représente la plus basse valeur possible. Ainsi, si on s'intéresse, par exemple, à une variable oscillant entre des valeurs positives et négatives, telle que le bilan carbone du système (*Net Biome Productivity*, NBP), il faudra préalablement la transformer (un décalage pourrait suffire pour éviter des valeurs négatives).

La méthode proposée a également l'avantage de pouvoir prendre en compte différentes sources d'incertitudes (par exemple celles liées au climat ou à la gestion). De plus, à travers ces deux phases, elle permet de réaliser, à la différence de la majorité des autres méthodes (e.g., [Metzger et Schröter, 2006] et [Downing et Patwardhan, 2005]), des analyses de vulnérabilité avec et sans prise en compte de l'adaptation.

Un autre avantage de cette méthode d'analyse de vulnérabilité est d'utiliser une approche de modélisation pour chercher à réduire la vulnérabilité. En effet, la phase exploratoire de minimisation de la vulnérabilité ne saurait être réalisée expérimentalement compte-tenu du nombre d'expérimentations à réaliser, qui est encore augmenté si l'on tient compte de différentes sources d'incertitude.

Comme cela est présenté dans le manuscrit, la phase de minimisation de la vulnérabilité peut être réalisée à l'aide de métaheuristiques [Glover 1986] de façon à éviter la méthode dite de "force brute" qui est coûteuse en calcul. L'utilisation de métaheuristiques permettrait de d'incorporer facilement la connaissance que les experts ont sur le système et donc d'être potentiellement plus efficaces en termes de vitesse de convergence. Notre approche conceptuelle de la vulnérabilité distingue la capacité d'adaptation potentielle de la capacité d'adaptation réelle, cette dernière devant limiter la minimisation de vulnérabilité. Il sera sans doute plus facile d'inclure ce genre d'information dans une métaheuristique que dans des algorithmes plus classiques (méthodes analytiques, par exemple). Il est également possible d'inclure des connaissances (approche experte) dans la conception du plan d'expériences (à travers par exemple la prise en compte de combinaisons impossibles, ou la connaissance *a priori* de la forme des surfaces de réponse). La contrepartie des métaheuristiques est généralement une efficacité (qualité de solution et vitesse de convergence) plus variable selon le contexte (par rapport aux autres méthodes) et un paramétrage (notamment l'initialisation) à adapter à chaque cas.

Un exemple d'application de la méthode sur des données spatialisées a également été réalisé. Il a permis de mettre en avant la pertinence de l'utilisation d'un ensemble d'indices et de montrer les différences avec l'évaluation de la vulnérabilité en un site donné. En effet, lors de l'évaluation d'un pixel il est également possible de tenir compte des pixels voisins car on peut les considérer comme partiellement représentatifs du pixel. De plus, à la différence d'une modélisation à l'échelle d'un site, le nombre de simulations pour couvrir les incertitudes devient trop important pour pouvoir être traité en un temps acceptable, même avec du calcul à haute performance. Une méthode alternative a également été proposée (voir 4.3.3).

Nous avons proposé une approche générique, dans le sens où elle peut permettre de tenir compte de multiples hypothèses (par exemple différentes sources d'incertitudes) et de différents types de modèles agro-écologiques (par exemple, dans notre cas, ModVégé et PaSim). De plus, cette méthode, conçue initialement pour l'analyse de vulnérabilité au changement climatique, est également opérationnelle dans le cas d'analyse de vulnérabilité à d'autres changements. En effet, une analyse de vulnérabilité à d'autres facteurs que le changement climatique va à la fois contenir des étapes s'apparentant à des analyses de sensibilité, et être confrontée au problème de l'analyse de vulnérabilité avec et sans adaptation. Or, notre approche basée sur la vision de [Taguchi 1987], qui consiste à distinguer les facteurs environnementaux des facteurs de gestion, permet une grande genericité par rapport au domaine d'application de notre méthode.

Perspectives

L'optimisation (minimisation) et l'analyse de vulnérabilité avec adaptation n'ont pas encore été implémentées dans le plugin EMF. Il sera intéressant d'implémenter sous EMF deux ou trois types de méthodes de recherche de minimum de vulnérabilité (classe Optimisation) et, ainsi, de générer les scripts pour R et pour OpenMOLE permettant la distribution des calculs (dans la mesure où la méthode le permet).

Pour mûrir, l'outil généré, ainsi que la méthode d'analyse de vulnérabilité, auront également besoin de plus d'exemples d'utilisation et donc de plus d'utilisateurs. Même si la genericité et la modularité de l'outil devrait garantir sa viabilité à long terme, de nombreuses améliorations peuvent encore être apportées (par exemple, différentes surfaces de réponse ou d'autres indices de vulnérabilité) en fonction des besoins des utilisateurs.

De façon à rendre plus accessibles les résultats de cette thèse, il a également été décidé de concevoir une interface graphique pour la bibliothèque d'indices de vulnérabilité. Ce travail est en partie réalisé par deux étudiants de l'école d'ingénieur ISIMA. L'idée est de pouvoir lancer et visualiser facilement les résultats des calculs d'indices de vulnérabilité, à

l'aide d'une interface ergonomique, ces calculs pouvant être effectués pour une analyse à l'échelle d'un site (fichiers ASCII) ou à une échelle spatialisée (fichiers NetCDF).

Cette thèse a mis en évidence le côté opérationnel de la méthode proposée et de l'outil qui a été réalisé pour l'analyse de vulnérabilité. Ils seront utilisés en partie dans le cadre du projet européen ANIMALCHANGE. En 2012, un projet avait également été soumis en réponse à un appel d'offre⁸⁷ du Ministère de l'Écologie, du Développement Durable, des Transports et du Logement (aujourd'hui Ministère de l'Écologie, du Développement durable et de l'Énergie). Cet appel d'offre, centré sur les impacts et l'analyse de vulnérabilité, avait notamment pour but de faciliter la constitution d'une communauté française « Climat - Impacts, Vulnérabilité et Adaptation ». Le projet soumis (nommé "FRANCE HERBE-IVA"), coordonné par INRA-UREP⁸⁸ à l'aide de deux partenaires (LIMOS⁸⁹ et LSCE⁹⁰), proposait la construction d'une plateforme d'interaction entre climatologues, agro-écologues et informaticiens autour de l'analyse de vulnérabilité de prairies. Il proposait notamment la réalisation de simulations spatialisées de climat et d'impacts, pour les nouvelles trajectoires de concentration de gaz à effet de serre (*representative concentration pathways* [Moss et al., 2008]) du 5^e rapport du GIEC (en préparation), dans le but d'analyser la vulnérabilité des prairies en France, tout en se basant sur la méthode développée dans ce manuscrit. Bien que non retenu, ce projet met en avant la possibilité de prolonger ce travail de thèse dans un cadre applicatif différent et plus complet que celui dans lequel il a été conçu.

Parmi les travaux futurs qui sont possibles, on pourrait également envisager la réalisation de simulations réellement spatialisées (et non simplement "pixélisées" comme dans les résultats présentés dans ce manuscrit). En effet, il devient important, compte-tenu de l'état d'avancement des études sur la vulnérabilité au changement climatique, de raisonner en terme de dimension spatiale (idéalement continue), ou *a minima* d'approximer les interactions entre pixels. Les principales limites à ce genre d'approche étaient les temps de calculs, la disponibilité des données d'entrée et la capacité des modèles à reproduire les phénomènes que l'on cherche à observer. Ces limites ont été partiellement levées à travers le développement du Calcul à Haute Performance, l'amélioration des modèles et de leurs paramétrages (bien qu'ils puissent rester de fortes incertitudes) et la création de certaines bases de données. En effet, les Systèmes d'Information Géographique (SIG) actuels pourraient aisément nous permettre d'évaluer la vulnérabilité à l'échelle régionale en tenant compte de certains facteurs locaux telle que la pente, l'exposition au soleil et la qualité des sols. Une telle approche serait cependant toujours limitée par la disponibilité de certaines données (par exemple, les pratiques agricoles, qui ne sont pas forcément connues de manière suffisamment précise). Une autre limite majeure serait sans doute que, en plus

⁸⁷ <http://www.developpement-durable.gouv.fr/GICC-2012-Gestion-et-Impacts-du.html>

⁸⁸ <https://www1.clermont.inra.fr/urep/>

⁸⁹ <http://limos.isima.fr/>

⁹⁰ <http://www.lsce.ipsl.fr/>

de leurs incertitudes, les modèles utilisés n'intègrent pas forcément certains mécanismes spatiaux tels que ceux relatifs à l'hétérogénéité spatiale de la biomasse et des espèces au sein d'une parcelle de prairie.

Tout cela ouvre un vaste champ d'actions scientifiques qui ne font pas partie directement des objectifs de cette thèse. Notamment, la démarche de recherche et les résultats obtenus peuvent aider à progresser vers une approche méthodologique commune permettant de répondre aux questions soulevées dans les études sur les impacts du changement climatique et l'adaptation à celui-ci.

Bibliographie

- Aarts, E., H., L., Van Laarhoven, P., J., M., 1985. Statistical cooling: A general approach to combinatorial optimization problems. *Philips Journal of Research*, 40, 193-226.
- Aarts, E., Korst, J., 1989. *Simulated annealing and Boltzmann machines: a stochastic approach to combinatorial optimization and neural computing*. Wiley, Chichester, 272p.
- Aarts, E., Korst, J., Michiels, W., 2005. *Search methodologies. Introductory tutorials in optimization and decision support techniques. Chapter 7 Simmulated annealing*. Springer, Burke, E., K., Kendall, G., eds., 187-210.
- Adger, W., N., 2006. Vulnerability. *Glob. Environ. Change*, 16, 268-281.
- Adger, W., N., Kelly, M., 1999. Social vulnerabililty to climate change and the architecture of entitlements. *Mitig. Adapt. Strat. Global Change* 4, 253–266.
- Agreste (La statistique, l'évaluation et la prospective agricole), 2008. Statistique agricole annuelle. <http://www.agreste.agriculture.gouv.fr>.
- Agreste (La statistique, l'évaluation et la prospective agricole), 2010. Agreste Infos rapides – Animaux de boucherie- décembre 2010. <http://www.agreste.agriculture.gouv.fr>.
- Amblard, F., Hill, D., R., C., Bernard, S., Truffot, J., Deffuant, G., 2003. MDA Compliant Design of SimExplorer A Software Tool to Handle Simulation Experimental Frameworks. Summer Computer Simulation Conference, Montreal, Canada, July 20-24, 279-284.
- Amdahl, G., 1967. Validity of the Single Processor Approach to Achieving Large-Scale Computing Capabilities. *AFIPS Conference Proceedings*, 30, 483-485.
- Angün, E., den Hertog, D., Gürkan, G., Kleijnen, J., P., C., 2002. Response surface methodology revisited. E. Yücesan, C. H. Chen, J. L. Snowdon, J. M. Charnes, eds. *Proc. 2002 Winter Simulation Conf.* Institute of Electrical and Electronics Engineers, Piscataway, NJ, 377–383.
- ANR (Agence National pour la recherche), 2010. Le calcul intensif : technologie clé pour le futur. *Les Cahiers de l'ANR*, 3, Badillo P., -Y, Tarrier F. eds., 172p., disponible en ligne via http://www.agence-nationale-recherche.fr/fileadmin/user_upload/documents/uploaded/2010/ Cahier-ANR-3-Calcul-intensif.pdf.
- Arnell, N., W., Cannell, M., G., R., Hulme, M., Martens, P., Mitchell, J., F., B., McMichael, A., J., Nicholls, R., J., Parry, M., L., White, A., 2002 .The consequences of CO2 stabilisation for the impacts of climate change? *Climatic Change*, 53, 413–446
- Atkinson, C., Kuhne, T., 2003. Model-driven development: a metamodeling foundation. *IEEE Software*, 20, 36-41.
- Aulagnier, C., Le Dizès, S., Maro, D., Hébert, D., Lardy, R., Martin, R., Gonze, M., A., 2012. Modelling the transfer of ¹⁴C from the atmosphere to grass: A case study in a

- grass field near AREVA-NC La Hague. *Journal of Environmental Radioactivity*, 112, pp. 52-59.
- Aussem, A., Hill, D., 1999. Wedding connectionist and algorithmic modelling, *Ecological modelling*, 120, 225-236.
- Bachar Y., Benyessa T., 2011. *Implémentation d'un greffon en vue de son intégration dans la distribution principale de JSAGA*. Rapport de projet de 3^{ème} année, Institut Supérieur d'Informatique de Modélisation et leurs Application (ISIMA), 34p.
- Bachman, S., Heisler-White, J., L., Pendall, E., Williams, D., G., Morgan, J., A., Newcomb, J., 2010. Elevated carbon dioxide alters impacts of precipitation pulses on ecosystem photosynthesis and respiration in a semi-arid grassland. *Oecologia* 162, 791-802.
- Baduel, L., Baude, F., Caromel, D., Contes, A., Huet, F., Morel, M., Quilici, R., 2006. Chapter 9. Programming, composing, deploying for the grid. *Grid Computing. Software Environments and Tools*. Cunha, J., C., Rana, O., F. Eds., 205-229.
- Bagrodia, R., L., Liao, W.-T., 1994. Maisie: A Language for the Design of Efficient Discrete-Event Simulations, *IEEE Transactions on Software Engineering*, 20, pp. 225-238.
- Bargen, B., Donnelly, P., 1998. Inside DirectX: in-depth techniques for developing high-performance multimedia applications. Microsoft Press, 545p.
- Bargiotti, M., Smith, A., C., 2008. DIRAC Data Management: consistency, integrity and coherence of data. *Journal of Physics: Conference Series*, 119, 062013.
- Basso, B., Gargiulo, O., Paustian, K., Porter, C., H., Robertson, G., P., Grace, P., R., Jones, J., W., 2011. Procedures for initializing soil organic carbon pools in DSSAT-Century model for agricultural systems. *Soil Sci. Soc. Am. J.*, 75-1-69-78.
- Ben Touhami, H., Lardy, R., Bahn, M., Bellocchi G., Barra V., 2012a. Bayesian calibration of the Pasture Simulation Model (PaSim) to simulate European grasslands under climate extremes: case study at Stubai, *Proceedings of Probabilistic Safety Assessment and Management Conference (PSAM'12)*, Helsinki, Finland, 25–29 June 2012, 567-577.
- Ben Touhami, H., Lardy, R., Gilgen, A., Buchmann, N., Barra, V., Bellocchi, G., 2012b. Bayesian calibration of the Pasture Simulation Model (PaSim) to simulate Swiss grasslands under climate extremes, *Proceedings of ESA12 - 12th Congress of the European Society for Agronomy*, Helsinki, Finland, 20-24 August, 2012, 2p.
- Bettonvil, B., Kleijnen, J., P., C., 1997. Searching for important factors in simulation models with many factors: Sequential bifurcation. *Eur. J. Oper. Res.* 96 180–194.
- Bezivin, J., 2004. In Search of Basic Principle for Model Driven Engineering. *Novatica Journal*, Special Issue, March-April, 2004, 2, 21-24.
- Bezivin, J., 2005. On the unification power of models. *Software and Systems Modeling*, 4, 171-188.

- Blanchet et Loomis, 2010. La technologie « cloud ». Ecole GRISBI (Roscoff). 1 octobre 2010.
- Bohn, U., Gollub, G., Hettwer, C., 2000. *Karte der natürlichen vegetation Europas: Maßstab 1 : 2.500.000*. Bundesamt für Naturschutz, Bonn-Bad Godesberg.
- Box, G. E., Hunter W. G., Hunter, J.S., 1978. *Statistics for Experimenters: An introduction to Design, Data Analysis and Model Building*, Wiley, New York, 653p.
- Brand, F., S., Jax, K., 2007. Focusing the meaning(s) of resilience: resilience as a descriptive concept and a boundary object. *Ecology and Society* 12, 23.
- Brisson, N., Mary, B., Ripoche, D., Jeuffroy, M., H., Ruget, F., Nicoullaud, B., Gate, P., Devienne-Barret, F., Antonioletti, R., Durr, C., Richard, G., Beaudoin, N., Recous, S., Tayot, X., Plenet, D., Cellier, P., Machet, J., M., Meynard, J., M., Delecolle, R., 1998. STICS: a generic model for the simulation of crops and their water and nitrogen balances. I. Theory and parameterization applied to wheat and corn. *Agronomie*, 18, 311–346.
- Brisson, N., Ruget, F., Gate, P., Lorgeau, J., Nicoullaud, B., Tayot, X., Plenet, D., Jeuffroy, M., H., Bouthier, A., Ripoche, D., Mary, B., Justes, E., 2002. STICS: a generic model for simulating crops and their water and nitrogen balances. II. Model validation for wheat and maize. *Agronomie*, 22, 69–92.
- Brunelle, A., Savoie, V., 2000. *Guide des pratiques de conservation en grandes cultures*. Module 3 : Gestion de la matière organique, 12p.
- Burke, E., K., Kendall, G., 2005. *Search methodologies. Introductory tutorials in optimization and decision support techniques*. Springer, 620p.
- Butler, R., Welch, V., Engert, D., Foster, I., Tuecke, S., Volmer, J., Kesselman, C., 2000. Design and deployment of a national-scale authentication infrastructure. *IEEE Computer*, 33, 60-66.
- Campolongo, F., Caribioni, J., Saltelli, A., 2007. An effective screening design for sensitivity, *Environmental Modelling & Software*, 22, 1509-1518.
- Carrère, P., Force, C., Soussana, J.-F., Louault, F., Dumont, B., Baumont, R., 2002. Design of a spatial model of perennial grassland grazed by a herd of ruminants: the vegetation sub-model. *Grassland Science in Europe* 7, 282–283.
- Caux, J., 2012. *Parallélisation et optimisation d'un simulateur de morphogénèse d'organes Application aux éléments du rein*. Université Blaise Pascal, Thèse de doctorat, 267p.
- Cetinkaya, D., Verbraeck, A., Seck, M., D., 2011. MDD4MS: A Model Driven Development Framework for Modeling and Simulation. 2011, SummerSim'11, The Hague, Netherlands, 203-211.
- Chambers, R., 1989. Vulnerability, coping and policy. *Institute of Developmental Studies Bulletin* 20, 1–7.
- Cheng, J., Tao, J.-P., 2010. Fuzzy Comprehensive Evaluation of Drought Vulnerability Based on the Analytic Hierarchy Process: --An Empirical Study from Xiaogan City in

- Hubei Province. Agriculture and Agricultural Science Procedia, Volume 1, International Conference on Agricultural Risk and Food Security 2010, 126-135.
- Cheng, R., C., H., 1997. Searching for important factors: Sequential bifurcation under uncertainty. S. Andradottir, K. J. Healy, D. H. Withers, B. L. Nelson, eds. Proc. 1997 Winter Simulation Conf. Institute of Electrical and Electronics Engineers, Piscataway, NJ, 275–280.
- Chikofsky, E., Cross, J., 1990. Reverse Engineering and Design Recovery: A taxonomy, IEE Software 13-17.
- Chuffart, F., Dumoulin, N., Faure, T., Deffuant, G., 2010. SimExplorer: Programming Experimental Designs on Models and Managing Quality of Modelling Process. International Journal of Agricultural and Environmental Information Systems 1, 55-68.
- Chuffart, F., Faure, T., Dumoulin, N., Reuillon, R., Deffuant, G., 2008. Design and Analysis of Computer Experiments with SimExplorer, ESM'2008, October 27-29, 2008, Le Havre, France, 575-577
- Cioppa, T., M., Lucas, T., W., 2007. Efficient Nearly Orthogonal and Space-Filling Latin Hypercubes. Technometrics. February 1, 2007, 49, 45-55.
- CNRS, Comité d'Orientation pour le Calcul INTensif (Cocin), 2012. Livre blanc sur le calcul intensif au CNRS. 44p., accessible en ligne via http://www.cnrs.fr/inp/IMG/pdf/livre_blanc-calcul_intensif-version_finale_092512.pdf
- Combemale, B., 2008. *Ingénierie Dirigée par les Modèles (IDM) État de l'art*. 19p. disponible en ligne via <http://hal.archives-ouvertes.fr/docs/00/37/15/65/PDF/mde-stateoftheart.pdf>
- Conant, R., T., Paustian, K., Elliott, E., T., 2001. Grassland management and conversion into grassland: effects on soil carbon. Ecological Applications, 11, 343-355
- Dagum, L., Menon, R., 1998. OpenMP: an industry standard API for shared-memory programming. Computational Science and Engineering, IEEE 5, 46-55.
- De Lange, H., J., Sala, S., Vighi, M., Faber J., H., 2010. Ecological vulnerability in risk assessment – A review and perspectives. Science of The Total Environment, Volume 408, Issue 18, 3871-3879.
- De Martonne, E., 1942. Nouvelle carte mondiale de l'indice d'aridité. Annales de Géographie, 51, 242-250.
- Dee, D., P., Uppala, S., M., Simmons, A., J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M., A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A., C., M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A., J., Haimberger, L., Healy, S., B., Hersbach, H., Hólm, E., V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A., P., Monge-Sanz, B., M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., Vitart, F., 2011. The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. Quarterly Journal of the Royal Meteorological Society, 137, 553-597.

- Del Grosso, S., J., Parton, W., J., Mosier, A., R., Hartman, M., D., Brenner, L., Ojima, D., S., Schimel, D., S., 2001. Simulated interaction of carbon dynamics and nitrogen- trace gas fluxes using the DAYCENT model. Schaffer, M., et al. (Eds.), Modeling Carbon and Nitrogen Dynamics for Soil Management. CRC Press, Boca Raton, FL, 303–332.
- Déqué, M., 2007. Frequency of precipitation and temperature extremes over France in an anthropogenic scenario: Model results and statistical correction according to observed values. *Global Planet. Change*, 57, 16–26.
- Déqué, M., Dreveton, C., Braun, A., Cariolle, D., 1994. The ARPEGE / IFS atmospheremodél: a contribution to the French community climate modelling. *Climate Dyn.* 10, 249–266.
- Déqué, M., Rowell, D., P., Lüthi, D., Giorgi, F., Christensen, J., H., Rockel, B., Jacob, D., Kjellström, E., de Castro, M., van den Hurk, B., 2007. An intercomparison of regional climate simulations for Europe: assessing uncertainties in model projections. *Climatic Change*, 81, 53-70.
- Diaw, S., Lbath, R., Coulette, B., 2008. Etat de l’art sur le développement logiciel dirigé par les modèles, TSI-X, 2008, Toulouse.
- Downing, T., E., Butterfield, R., Cohen, S., Huq, S., Moss, R., Rahman, A., Sokona, Y., Stephen, L., 2001. *Climate Change Vulnerability: Linking Impacts and Adaptation. UNEP Policy Series*, UNEP, Nairobi., 91 pp.
- Downing, T., E., et Patwardhan, A., 2003. *Vulnerability Assessment for Climate Adaptation*, APF Technical Paper 3, United Nations Development Programme, New York City, NY. Final draft.
- Downing, T., E., Patwardhan, A., 2005. *Assessing Vulnerability for Climate Adaptation. Technical Paper 3*, in Bo Lim, Erika Spanger-Siegfried, Ian Burton, Elizabeth Malone and Saleemul Huq, eds. *Adaptation Policy Frameworks for Climate Change*. Cambridge: Cambridge University Press, pp. 67-89.
- Drechsler, M., 2004. Model-based conservation decision aiding in the presence of goal conflicts and uncertainty, *Biodiversity and Conservation* 13, pp. 141–164.
- Duclos, E., 2010. *Prise en compte du carbone 14 dans le modèle PASIM*, Rapport de stage de troisième année d’école ingénieur ISIMA, 52 p.
- Eakin, H., Luers, A., L., 2006. Assessing the vulnerability of social-environmental systems, *Annual Review of Environment and Resources*, 31, 365–394.
- Eclipse-EMF, <http://www.eclipsesetotale.com/index.html?keywords=emf>;
<http://www.eclipse.org/modeling/emf/>
- Efron, B., Tibshirani, R., J., 1993. *An introduction to the Bootstrap*, CRC Press, 456 p.
- Egede, U., Harrison, K., Jones, R., W., L., Liko, D., Maier, A., Mościckif, J., T., Patrick, G., N., Soroko, A., Tan, C., L., 2005. Ganga user interface for job definition and management. Proc. Fourth International Workshop on Frontier Science: New Frontiers in Subnuclear Physics. - Milan, Italie, 367-374.

- Fang, K., T., 1980. The Uniform Design: Application of Number-Theoretic Methods in Experimental Design. *Acta Mathematicae Applagatae Sinica*, 3, 363-372.
- Fang, K., T., Li, R., Sudjianto, A., 2006. *Design and modelling for computer experiments*. London: Chapman and Hall / CRC, 290p.
- FAO, 1996. *Global climate change and agricultural production. Direct and indirect effects of changing hydrological, pedological and plant physiological processes*, John Wiley & Sons Ltd, 345p.
- Farquhar, G., D., von Caemmerer, S., 1982. Modelling of photosynthetic response to environmental conditions. *Physiological Plant Ecology II. Encyclopedia of Plant Physiology, New Series*, Vol. 12B (eds Lange, OL Nobel, PS Osmond, CB Ziegler, H), pp. 549-587. Springer, Berlin.
- Faure, T., Deffuant, G., 2006. SimExplorer: A software tool for programming and executing experiemental designs on complex models. Summer Simulation Multiconf., Calgary, Canada, 218-225
- Faurié, O., Soussana, J., -F., Sinoquet, H., 1996. Radiation interception, partitioning and use in grass-clover mixtures. *Annals of Botany*, 77, 35–45.
- Favre, J.,-M., Duclos, F., Estublier, J., Sanlaville, R., Auffret, J., J, 2001. Fifth European Conference on Software Maintenance and Reengineering. 95-104.
- Favre, J.-M, Estublier, J., Blay-Fornarino, M., 2006. *L'ingénierie dirigée par les modèles*. Lavoisier, 302p.
- Favre, J.-M., 2004a. Towards a basic Theory to Model Driven Engineering. 3rd Workshop in Software Model Engineering, WiSME 2004, 8p., <http://www-adele.imag.fr/~jmfavre>.
- Favre, J.-M., 2004b. Foundations of Meta-Pyramids: Languages vs. Metamodels, Episode II: Story of Thotus the Baboon, Dagstuhl Seminar 04101 on Language Engineering for Model-Driven Software Development (Dagsthul, Germany), 70-80.
- Folke, C., Carpenter, S., Elmqvist, T., Gunderson, L., Holling, C., Walker, B., 2002. Resilience and sustainable development: building adaptive capacity in a world of transformations. *Ambio* 31, 437-440.
- Folke, C., Hahn, T., Olsson, P., Norberg, J., 2005. Adaptive governance of social ecological systems. *Annu. Rev. Environ. Resour.* 30, 441-473.
- Foereid, B., Bellamy, P., H., Holden, A., Kirk, G., J., D., 2012. On the initialization of soil carbon models and its effects on model predictions for England and Wales. *European Journal of Soil Science*, 63: 32–41.
- Foster, I., 2001. The Globus Toolkit for Grid Computing. *Cluster Computing and the Grid*, IEEE International Symposium, 1, pp. 2.
- Foster, I., Kesselman, C., 1997. Globus: A Metacomputing Infrastructure Toolkit, *International Journal on Supercomputer Applications* 11, 115-128.

- Foster, I., Kesselman, C., 1999. *The Grid: Blueprint for a new computing infrastructure*, Morgan Kaufmann, 675p.
- Foster, I., Kesselman, C., Tuecke, S., 2001. The Anatomy of the Grid - Enabling Scalable Virtual Organizations, *International Journal of High Performance Computing Applications* 15, 200-222.
- Foster, J., Greer, J., Thorbecke, E., 1984. A class of decomposable poverty measures. *Econometrica* 52, 761–766.
- Füssel, H.-M., 2007. Vulnerability: A generally applicable conceptual framework for climate change research, *Global Environmental Change*, 17, 155-167.
- Füssel, H.-M., Klein R., 2006. Climate change vulnerability assessments: an evolution of conceptual thinking. *Climate Change*, 75, 301–329.
- Gamma, E., Helm, R., Johnson, R., and Vlissides, J., 1995. *Design Patterns: Elements of Reusable Object-Oriented Software*. Addison-Wesley: Reading, MA, 295 p.
- Gauchi, J.-P., 2007. Plans d'expériences numériques. Journée de la Modélisation au Cemagref, 26-27 novembre 2007, Clermont-Ferrand.
- Gazut, S., Martinez, J., Dreyfus, G., 2008. Towards the optimal design of numerical experiments. *IEEE Transactions on Neural Networks* 19, 874–882.
- Gazut, S., Martinez, J.-M., Issanchou, S., 2006. Plans d'expériences itératifs pour la construction de modèles non linéaires, 38^{èmes} Journées de Statistique, Clamart, France.
- Geist, A., Beguelin, A., Dongarra, J., Jiang, W., Manchek, R., Sunderam, V., 1994. *PVMParallel Virtual Machine : a Users' Guide and Tutorial for Networked Parallel Computing*. MIT press, 285 p.
- Gibon, A., Rubinor, R., Sibbald, A., R., Sorensen, J., T., Flamant, J., C., Lhoste, P., Revilla, R., 1996. A review of current approaches to livestock farming systems in Europe: towards a common understanding. *Livestock farming systems: research, development, socioeconomics and the land manager*, Dent, McGregor, Sibbald Eds., EAAP, 79, 7–19.
- Gjermundrod, H., Dikaiakos, M., D., Stumpert, M., Wolniewicz, P., Kornmayer, H., 2008. g-Eclipse-an integrated framework to access and maintain Grid resources. 9th IEEE/ACM International Conference on Grid Computing, 57-64.
- Glover, F., 1986. Future paths for integer programming and links to artificial intelligence. *Computers and Operations Research*, 13, 533-549.
- Govindaraju, N., K., Lloyd, B., Wang, W., Lin, M., C., Manocha, D., 2004. Fast Computation of Database Operations using Graphics Processors. *International Conference on Management of Data* 2004, 215-226.
- Graux A.-I., 2011. Modélisation des impacts du changement climatique sur les écosystèmes prairiaux. Voies d'adaptation des systèmes fourragers, Université Blaise Pascal, Thèse de doctorat, 535 p.

- Graux, A.-I., Bellocchi, G., Lardy, R., Soussana, J.-F., 2012a. Ensemble modelling of climate change risks and opportunities for managed grasslands in France. *Agricultural and Forest Meteorology*, 170, 114-131.
- Graux, A.-I., Gaurut, M., Agabriel, J., Baumont, R., Delagarde, R., Delaby, L., Soussana, J.-F., 2011. Development of the Pasture Simulation Model for assessing livestock production under climate change. *Agriculture, Ecosystems & Environment*, 144, 69–91
- Graux, A.-I., Lardy R., Bellocchi G., Soussana J.-F., 2012b. Global warming potential of French grassland-based dairy livestock systems under climate change. *Regional Environmental Change*, 12, 751-763.
- Graux, A.-I., Lardy, R., Gaurut, M., Soussana, J.-F., Hill, D., R., C., 2013. Adaptations of French grassland based livestock systems to increasing climatic hazards. (en préparation).
- Graux, A.-I., Lardy, R., Manneville, V., Carrère, P., 2012c. Peut-on limiter les pollutions azotées en prairie permanente, en jouant sur la période d'épandage des effluents d'élevage ?. Accepté dans *Fourrage*.
- Grimm, V., Wissel, C., 1997. Babel, or the ecological stability discussions: an inventory and analysis of terminology and a guide for avoiding confusion. *Oecologia (Berlin)* 109, 323–334.
- Guo, L., B., Gifford, R., M., 2002. Soil carbon stocks and land use change: a meta analysis. *Global Change Biology*, 8, 345-360.
- Hajek, B., 1988. Cooling schedules for optimal annealing. *Mathematics of operations research*, 13, 311-329
- Hanel, M., Buishand, T., A., 2010. On the value of hourly precipitation extremes in regional climate model simulations. *Journal of Hydrology* 393, 265–273.
- Hansen, P., Mladenović, N., 2003. *Handbook of metaheuristics. Variable neighborhood search*. Springer, Glover, F. et Kochenberger, G. eds, Kluwer, Dordrecht, 145-184
- Hansen, P., Mladenović, N., 2005. *Search methodologies. Introductory tutorials in optimization and decision support techniques. Chapter 8 Variable Neighborhood Search*. Springer, Burke, E., K., Kendall, G., eds. 211-238.
- Harris, M., J., Coombe, G., Scheuermann, T., Lastra, A., 2002. Physically-based visual simulation on graphics hardware. *SIGGRAPH/EUROGRAPHICS Workshop On Graphics Hardware 2002*, 109-118.
- Hedayat, A.,S., Sloane, N.,J.,A., Stufken, J., 1999. *Orthogonal arrays: theory and applications*. Springer New York, 416p.
- Hegerl, G., C., Hanlon, H., Beierkuhnlein, C., 2011. Elusive extremes, *Nature Geoscience*, 4, 142-143.
- Hill, D., Passerat-Palmbach, J., Mazel, C., Traore, M., K., 2012. Distribution of Random Streams for Simulation Practitioners. *Concurrency and Computation: Practice and Experience*, 2012. First Published Online Nov. 2012, doi:10.1002/cpe.2942. To be published.

- Hill, D., R., C., 1996. *Object-oriented Analysis and Simulation*. Addison-Wesley, 291 p.
- Hinkel, J., 2011. Indicators of vulnerability and adaptive capacity: Towards a clarification of the science-policy interface. *Global Environmental Change*, 21, 198–208.
- Hoffman, A., R., Traub, J., F., 1989. *Supercomputers: directions in technology and applications*. National Academies Press, 102p.
- Holland, J., H., 1975. *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. University of Michigan Press, 183p.
- Holling, C., S., 1996. Engineering resilience vs. ecological resilience. Pages 31–43 in P. C. Schulze, ed. *Engineering within ecological constraints*. National Academy, Washington, DC.
- Holling, C., S., 1973. Resilience and stability of ecological systems. *Annual Review of Ecology and Systematics* 4, 1–23.
- Howell, F., McNab, R., 1998. SimJava: A discrete event simulation library for java. *Simulation Series*, 30, 51-56.
- Hull, D., Wolstencroft, K., Stevens, R., Goble, C., Pocock, M., Li, P., Oinn, T., 2006. Taverna: a tool for building and running workflows of services. *Nucleic Acids Research*, 34, Web Server issue, pp. 729-732.
- Iacob, M., Stehen, M., Heerink, L., 2008. Reusable Model Transformation Patterns. Workshop on Models and Model-driven Method for Enterprise Computing (3M4EC), page 1.
- INRA-UREP 2011. Graux, A.-I., Lardy, R., Gaurut, M., Duclos, E., Klump, K., *PaSim V5.3 User's Guide*, p. 80. Rapport interne, disponible via https://www1.clermont.inra.fr/urep/modeles/Pasim_User_Guide-pasim_v5-3_201212.pdf
- Ionescu, C., Klein, R., Kumar, K., K., Hinkel, J., Kavi Kumar, K., S., Klein, R., 2009. Towards a Formal Framework of Vulnerability to Climate Change. *Environmental Modeling & Assessment*, 14, 1-16.
- IPCC (Intergovernmental Panel on Climate Change), 2001. *Impacts, adaptation, and vulnerability climate change 2001. Third Assessment Report of the IPCC*. University Press, Cambridge, UK, 1042p.
- IPCC (Intergovernmental Panel on Climate Change), 2007a. Climate change 2007: the physical science basis. *Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom, 976 p.
- IPCC (Intergovernmental Panel on Climate Change), 2007b. Climate change: impacts, adaptation and vulnerability. *Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom. 982p.

- IPCC, 2012. Field, C.B., V. Barros, T.F. Stocker, D. Qin, D.J. Dokken, K.L. Ebi, M.D. Mastrandrea, K.J. Mach, G.-K. Plattner, S.K. Allen, M. Tignor, and P.M. Midgley (Eds.). Cambridge University Press, Cambridge, 582 pp.
- Issanchou, S., Gauchi, J.,-P., 2008. Computer-aided optimal designs for improving neural network generalization, *Neural Networks*, 21, 945-950.
- Holland, J., H., 1975. *Adaption in natural and artificial systems*. The University of Michigan Press, Ann Harbor, MI, 237p.
- Johnson, M., E., Moore, L., M., Ylvisaker, D., 1990. Minimax and maximin distance designs. *J. Stat. Plann. Infer.* 26 131–148.
- Jones, D., R., Schonlau, M., Welch, W., J., 1998. Efficient global optimization of expensive black-box functions. *J Global Optim* 13: 455–492.
- Jones, M., B., Donnelly, A., 2004. Carbon sequestration in temperate grassland ecosystems and the influence of management, climate and elevated CO₂. *New Phytol.* 164, 423–439.
- Jongman, R., H., G., Bunce, R., G., H., Metzger, M., J., Mùcher, C., A., Howard, D., C., Mateus, V., L., 2006. Objectives and applications of a statistical environmental stratification of Europe. *Landscape Ecology* 21, 409-419.
- Jouven, M., Carrere, P., Baumont, R., 2006a. Model predicting dynamics of biomass, structure and digestibility of herbage in managed permanent pastures. 1. Model description. *Grass and Forage Science*, 61, 112-124.
- Jouven, M., Carrere, P., Baumont, R., 2006b. Model predicting dynamics of biomass, structure and digestibility of herbage in managed permanent pastures. 2. Model evaluation. *Grass and Forage Science*, 61, 125-133.
- Kasperson, J., X., Kasperson, R., E., Turner II, B., L., Schiller, A., Hsieh, W.-H., 2005. Vulnerability to global environmental change. In: Kasperson J.X., Kasperson R.E.. (Eds.), *Social contours of Risk. Vol. II: Risk Analysis Corporations and Globalizations of Risk*. Earthscan, London, pp. 245-285 (Chapter 14).
- Kelly, P., M., Adger, W., N., 2000. Theory and practice in assessing vulnerability to climate change and facilitating adaptation. *Climatic Change* 47, 325–352.
- Kempthorne, O., 1952. *The Design and Analysis of Experiments*, Wiley, Publications in Statistics, 631p.
- Khronos OpenCL Working Group, 2008. *The OpenCL specification. Version 1.0*. A. Munshi, Ed., 302p., disponible en ligne via <http://www.khronos.org/registry/cl/specs/opencl-1.0.29.pdf>
- Kirkpatrick, S., Gelatt, C., D., Vecchi, M., P., 1983. Optimization by Simulated Annealing. *Science*, 220, 671–680
- Kleijnen, J., P., C., 1987. *Statistical tools for simulation practitioners*. New York: Marcel Dekker. 448p.
- Kleijnen, J., P., C., 2008a. *Design and Analysis of simulation experiments*. Springer, 216 p.

- Kleijnen, J., P., C., 2008b. Design of experiments: Overview. WSC '08 Proceedings of the 40th Conference on Winter Simulation, December 2008.
- Kleijnen, J., P., C., 2009. Kriging metamodeling in simulation: A review. *European Journal of Operational Research* 192, 707-716.
- Kleijnen, J., P., C., Gaury E., 2003. Short-term robustness of production management systems: A case study. *Eur. J. Oper. Res.* 148 452–465.
- Kleijnen, J., P., C., Sanchez, S., M., Lucas, T., W., Cioppa, T., M., 2005. A User's Guide to the Brave New World of Designing Simulation Experiments. *INFORMS Journal on Computing*, 17, 263-289.
- Kleijnen, J., van Groenendaal, W., 1992. *Simulation a statistical perspective*, Wiley, Chichester, UK, 241p.
- Kleppe, A., Warmer, S., Bast W., 2003. *MDA Explained The Model Driven Architecture: Practice and Promise*. Addison-Wesley, 192 p.
- Koehler, J., R., Owen, A., B., 1996. Computer experiments. S. Ghosh, C. R. Rao, eds. *Handbook of Statistics*, Volume 13. Elsevier, Amsterdam, The Netherlands, 261–308.
- Kornmayer, H., Stümpert, M., Gjermundrød, H., Wolniewicz, P., 2008. g-Eclipse - a contextualised framework for grid users, grid resource providers and grid application developers, *Computational Science-ICCS 2008*, 399-408.
- L'Ecuyer, P., 1999. Good parameters and implementations for combined multiple recursive random number generators. *Operations Research* 47, 159–164.
- Landais, E., 1987. *Recherches sur les systèmes d'élevage: questions et perspectives*. Document de travail, INRA Département de recherches sur les Systèmes Agraires et le Développement, INRA publications, Toulouse, France, 75 pp.
- Lano, K., Clark, D., 2008. Model Transformation Specification and Verification. *Quality Software. 2008. QSIC'08. The Eighth International Conference* , 45-54.
- Lardy, R., Bachelet, B., Bellocchi, G., Hill, D., R., C., 2011a. Climate Change vulnerability assessment with constrained design of experiments, using a model driven approach. *European Simulation and Modelling conference (ESM) 2011*, 24-26 Oct, Guimaraes, Portugal, pp. 354-362.
- Lardy, R., Bellocchi, G., Soussana, J.-F., 2011b. A new method to determine soil organic carbon equilibrium, *Environmental Modelling & Software*, 26, 1759-1763.
- Lardy, R., Graux, A., -I., Bachelet, B., Hill, D., R., C., Bellocchi, G., 2012a. Steady-state soil organic matter approximation model: application to the Pasture Simulation Model. *iEMSs 2012, 6th Biennial Meeting of International Congress on Environmental Modelling and Software: Managing Resources of a Limited Planet: Pathways and Visions under Uncertainty*. Seppelt, R., Voinov, A., A., Lange, S., Bankamp, D., (Eds.), 1-5 July, 2012, Leipzig, Germany, 769-776. Disponible en ligne via <http://www.iemss.org/sites/iemss2012/proceedings.html>
- Lardy, R., Graux, A.-I., Gaurut, M., Bellocchi, G., Hill, D., R., C., 2011c. Model Driven Reverse Engineering For A Grassland Model With Design Of Experiment In The

Context Of Climate Change, The SummerSim'11 The Hague, 27-30 June, 199-206, Book 3.

- Lardy, R., Martin, R., Bachelet, B., Hill, D., R., C., Bellocchi, B., 2012b. Ecosystem Climate Change Vulnerability Assessment Framework. iEMSs 2012, 6th Biennial Meeting of International Congress on Environmental Modelling and Software: Managing Resources of a Limited Planet: Pathways and Visions under Uncertainty. Seppelt, R., Voinov, A., A., Lange, S., Bankamp, D., (Eds.), 1-5 July, 2012, Leipzig, Germany, 777-784. Disponible en ligne via <http://www.iemss.org/sites/iemss2012/proceedings.html>
- Leuning, R., 1995. A critical appraisal of a combined stomatal-photosynthesis model for C₃ plants. *Plant Cell Environment*, 18, 339-355.
- Lexer, M., J., Seidl, R., 2009. Addressing biodiversity in a stakeholder-driven climate change vulnerability assessment of forest management. *Forest Ecology Management* 258:S158–S167.
- Li, C., 2000. Modelling trace gas emissions from agricultural ecosystems. *Nutrient Cycling in Agroecosystems*, 58, 259–276.
- Li, C., Frohling, S., Frohling, T., A., 1992. A model of nitrous oxide evolution from soil driven by rainfall events.1. Model structure and sensitivity. *Journal of Geophysical Research-Atmospheres*, 97, 9759–9776.
- Lian, C., C., Francis Tang, F., Issac, P., Krishnan, A., 2005. GEL: Grid execution language. *Journal of Parallel and Distributed Computing*, 7, 857-869.
- Louault, F., Pillar, V., D., Aufrère, J., Garnier, E., Soussana, J., -F., 2005. Plant traits and functional types in response to reduced disturbance in semi-natural grassland. *Journal of Vegetation Science*, 16, 151–160.
- Lucas, T., W., Sanchez, S., M., Brown, L., Vinyard, W., 2002. Better designs for high-dimensional explorations of distillations. G. Horne, S. Johnson, eds. *Maneuver Warfare Science 2002*. USMC Project Albert, Quantico, VA, 17–46.
- Luers A., L., 2005. The surface of vulnerability an analytical framework for examining environmental change, *Global Environmental Change* 15, 214-223.
- Luers A., L., Lobell, D., Sklar, L., S., Addams, C., L., Matson, P., M., 2003. A method for quantifying vulnerability, applied to the agricultural system of the Yaqui Valley, Mexico, *Global Environmental Change* 13, 255-267.
- Lumley, T., 2011. biglm: bounded memory linear and generalized linear models. R package version 0.8., <http://CRAN.R-project.org/package=biglm>.
- Lundy, M., Mees, A., 1986. Convergence of an annealing algorithm. *Mathematical programming*, 34, 111-124
- Mahul, A., 2011. Infrastructure de la grille européenne EGI – Cours de 3ème année ISIMA 2011. <http://frog.isima.fr/antoine/grid/3-egi.pdf>.

- Marín, A., I., Dörfliger, N., Andreo, B., 2012. Comparative application of two methods (COP and PaPRIKa) for groundwater vulnerability mapping in Mediterranean karst aquifers (France and Spain). *Environmental Earth Sciences*, 65, 2407-2421.
- Martin, P., M., Cordier, S., Balesdent, J., Arrouays, D., 2007. Periodic solutions for soil carbon dynamics equilibriums with time-varying forcing variables. *Ecological Modelling* 204, 523-530.
- Matsumoto, M., Nishimura, T., 1998. Mersenne twister: A 623-dimensionnally equidistributed uniform pseudorandom number generator. *ACM Transactions on Modeling and Computer Simulations: Special Issue on Uniform Random Number Generation*, 8, 3–30.
- McKay, M., D., Beckman, R., J., Conover, W., J., 1979. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* 21, 239-245.
- McNab, R., F., Howell, W., 1996. Using Java for Discrete Event Simulation. *Proceedings of the Twelfth UK Computer and Telecommunications Performance Engineering Workshop*, University of Edinburgh, UK, 219-228.
- Meckesheimer, M., Barton, R., R., Simpson T., W., Booker A., J., 2001. Computationally inexpensive metamodel assessment strategies. *Am Inst Aeronaut Astronaut J* 40: 2053-2060.
- Mell, P., Grance, T., 2011. The NIST Definition of Cloud Computing. 7p. Accessible en ligne via <http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf>
- Metzger, M., J., Bunce, R., G., H., Jongman, R., H., G., Mu"cher, C., A., Watkins, J., W., 2005. A climatic stratification of the environment of Europe. *Global Ecol. Biogeogr.* 14, 549-563.
- Metzger, M., J., Rounsevell, M., D., A., Acosta-Michlik, L., Leemans, R., Schröter, D., 2006. The vulnerability of ecosystem services to land use change, *Agric. Ecosyst. Environ.*, 114, 69-85.
- Metzger, M., J., Schröter D., 2006. Towards a spatially explicit and quantitative vulnerability assessment of environmental change in Europe. *Regional Environmental Change* 6, 201–16.
- Meyer, B., 1988. *Object-Oriented Software Construction*. Prentice Hall, Englewood Cliffs N.J., 480p.
- Meyer, B., 1992. Applying "Design by contract". *Computer (IEEE)*, 25, 40-51.
- Minsky, M., L., 1965. Matter, Minds and Models. *Proceedings of Inter, Federation of Information Processing Congress*, vol. 1, 45-49.
- Missier, P., Soiland-Reyes, S., Owen, S., Tan, W., Nenadic, A., Dunlop, I., Williams, A., Oinn, T., Goble, C., 2010. Taverna, reloaded. *SSDBM 2010*, Heidelberg, Germany, 18p.
- Mitchell, J., Devine, N., Jagger, K., 1989. A contextual model of natural hazards. *Geographical Review* 79, 391–409.

- Mohtar R., H, Zhai, T, Chen, X., W., 2000. A world wide web-based grazing simulation model (GRASIM). *Comput. Electron. Agr.*, 29, 243–250.
- Mohtar, R., H., Buckmaster, D., R., Fales, S., L., 1997. A grazing simulation model: GRASIM .A. Model development. *Transactions of the ASAE*, 40, 1483–1493.
- Moreau, J., C., Delaby, L., Duru, M., Guérin, G., 2010. Advices given about the forage system : evolutions and conceptions concerning the steps to be taken and the tools to be used. *Fourrages*, 200, 565–586.
- Morris, M., D., 1991. Factorial sampling plans for preliminary computational experiments, *Technometrics*, 33, 161-174.
- Mościckif, J., T., Brochu, F., Ebke, J., Egede, U., Elmsheuser, J., Harrison, K., Jones, R., W., L., Lee, H., C., Liko, D., Maier, A., Muraru, A., Patrick, G., N., Pajchel, K., Reece, W., Samset, B., H., Slater, M., W., Soroko, A., Tan, C., L., van der Ster, D., C., Williams, M., 2009. Ganga: a tool for computational-task management and easy access to Grid resources, *Computer Physics Communications*, 180, 2303-2316
- Moss, R., Babiker, M., Brinkman, S., Calvo, E., Carter, T., Edmonds, J., Elgizouli, I., Emori, S., Erda, L., Hibbard, K., Jones, R., Kainuma, M., Kelleher, J., Lamarque, J., F., Manning, M., Matthews, B., Meehl, J., Meyer, L., Mitchell, J., Nakićenovič, N., O'Neill, B., Pichs R., Riahi, K., Rose, S., Runci, P., Stouffer, R., van Vuuren, D., Weyant, J., Wilbanks, T., van Ypersele, J., P., Zurek, M., 2008. *Towards new scenarios for analysis of emissions, climate change, impacts, and response strategies*. Intergovernmental Panel on Climate Change, Geneva, p. 132.
- Mücher, C., A., Klijn, J., A., Wascher, D., M., Schaminée, J., H., J., 2009. A new European Landscape Classification (LANMAP): a transparent, flexible and user-oriented methodology to distinguish landscapes. *Ecological Indicators* 10, 87-103.
- Myers, R., H., Khuri, A., I., Vining, G., 1992. Response surface alternatives to the Taguchi robust design parameter approach. *Amer. Statist.* 46 131–139.
- Myers, R., H., Montgomery, D., C., 2002. *Response Surface Methodology: Process and Product Optimization using Designed Experiments*, 2nd ed. Wiley, New York.
- Nachtergaele, F., Van Velthuisen, H., Verelst, L., Batjes, N., Dijkshoorn, K., Van Engelen, V., Fischer, G., Jones, A., Montanarella, L., Petri, M., Prieler, S., Teixeira, E., Wiberg, D., Shi, X., 2008. *Harmonized World Soil Database (Version 1.0)*. Food and Agriculture Organization of the United Nations, Rome, Italy and IIASA, Laxenburg, Austria.
- Nakićenovič, N., 2000. Greenhouse gas emissions scenarios. *Technological Forecasting and Social Change* 65, 149–166.
- Nelson, R., Brown, P., R., Darbas, T., Kokic, P., Cody, K., 2007. *The potential to map the adaptive capacity of Australian land managers for NRM policy using ABS data*. CSIRO, Australian Bureau of Agricultural and Resource Economics, prepared for the National Land & Water Resources Audit. 55p.
- Nelson, R., Kokic, P., Crimp, S., Martin, P., Meinke, H., Howden, S., M., 2010. The vulnerability of Australian rural communities to climate variability and change: part II-integrating impacts with adaptive capacity. *Environ. Sci. Pol.* 13, 18-27.

- Nelson, R., Kokic, P., Elliston, L., King, J., 2005. Structural adjustment: a vulnerability index for Australian broadacre agriculture. *Australian Commodities* 12, 171–179.
- New, M., Lister, D., Hulme M., Makin, I., 2002. A high-resolution data set of surface climate over global land areas. *Climate Research*, 21, 1-25.
- Niederreiter, H., 1992. *Random Number Generation and Quasi-Monte Carlo Methods*. CBMS-NSF Regional Conference Series in Applied Mathematics. Society of Industrial and Applied Mathematics, Philadelphia, 243p.
- Noirfalise, A., 1987. *Map of the natural vegetation of the member countries of the European Community and the Council of Europe*. Office for Official Publications of the European Communities, Luxembourg, 80p.
- OGF-SAGA, 2011. *A Simple API for Grid Applications (SAGA) Version: 1.1*, 341p., disponible en ligne via <http://www.ogf.org/documents/GFD.90.pdf>
- Oinn, T., Greenwood, M., Addis, M., Alpdemir, N., Ferris, J., Glover, K., Goble, C., Goderis, A., Hull, D., Marvin, D., Li, P., Lord, P., Pocock, M., Senger, M., Stevens, R., Wipat, A., Wroe, C., 2006. Taverna: lessons in creating a workflow environment for the life sciences, *Concurrency and Computation: Practice and Experience*, 18, 1067-1100.
- Olioso, A., Huard, F., Guilioni, L., 2010. Prise en compte du CO₂ dans le calcul de l'évapotranspiration de référence. Présentation des méthodes et des résultats du projet CLIMATOR, Versailles, France, Juin 17-18, 66–67.
- Olson, D., M., Dinerstein, E., Wikramanayake, E., D., Burgess, N., D., Powell, G., V., N., Underwood, E., C., d'Amico, J., A., Itoua, I., Strand, H., E., Morrison, J., C., Loucks, C., J., Allnutt, T., F., Ricketts, T., H., Kura, Y., Lamoreux, J., F., Wettengel, W., W., Hedao, P., Kassem, K., R., 2001. Terrestrial ecoregions of the world: a new map of life on Earth. *BioScience*, 51, 933–938.
- OMG (Object Management Group), 2002. *Meta Object Facility (MOF) Specification. Version 1.4, April 2002*. 358p. Disponible en ligne via <http://www.omg.org/spec/MOF/1.4/>
- OMG (Object Management Group), 2010. *Object Constraint Language, Version 2.2, February 2010*, 238p. <http://www.omg.org/spec/OCL/2.2>.
- OMG-CORBA (Object Management Group), 2011. Common Object Request Broker Architecture (CORBA) Specification, Version 3.2, <http://www.omg.org/spec/CORBA/3.2/>.
- OMG-MDA (Object Management Group), MDA, 2001. <http://www.omg.org/mda/specs.htm>.
- OMG-MOF, (Object Management Group), 2006. *Meta Object Facility (MOF) Core Specification OMG Available Specification Version 2.0*. 88p. Disponible en ligne via <http://www.omg.org/spec/MOF/2.0/>
- OMG-QVT (Object Management Group), 2011. *Meta Object Facility (MOF) 2.0 Query/View/Transformation Specification. Version 1.1, Janvier 2011*. 246p. Disponible en ligne via <http://www.omg.org/spec/QVT/>

- OMG-XMI, (Object Management Group), 2011. *OMG MOF 2 XMI Mapping Specification. Version 2.4.1.* 108p. Disponible en ligne sur <http://www.omg.org/spec/XMI/2.4.1>
- Owen, A., B., 1992. Randomly orthogonal arrays for computer experiments, integration and visualization, *Statistica Sinica*, 2, 439-452.
- Owen, A., B., 1994. Lattice sampling revisited: Monte Carlo variance of means over randomized orthogonal array. *Ann. Statist.*, 22, 930-945.
- Paavola, J., Adger, W.N., 2002. Justice and adaptation to climate change. Working Paper 23, Tyndall Centre for Climate Change Research, Norwich, UK.
- Page, E., H., Moose, Jr, R., L., Griffin, S., P., 1997. Web-based simulation in Simjava using remote method invocation. *Proceedings of the 29th conference on Winter simulation*, 468-474.
- Panneton, F., L'Ecuyer, P., Matsumoto, M., 2006. Improved long-period generators based on linear recurrences modulo 2. *ACM Transactions on Mathematical Software* 32, 1-16.
- Parton, W., J., Hartman, M., D., Ojima, D., S., Schimel, D., S., 1998. DAYCENT and its land surface sub-model: description and testing. *Global Planet. Change*, 19, 35-48.
- Parton, W., J., Stewart, J., B., W., Cole, C., V., 1988. Dynamics of C, N, P and S in grassland soils: a model. *Biogeochemistry*, 5, 109-131.
- Paterson, S., K., Tsaregorodtsev, A., 2008. DIRAC Optimized Workload Management. *Journal of Physics: Conference Series*, 119, 062040.
- Pearson, L., J., Nelson, R., Crimp, S., Langridge, J., 2011. Interpretive review of conceptual frameworks and research models that inform Australia's agricultural vulnerability to climate change. *Environmental Modelling & Software* 26, 113-123.
- Penghua Q., Songjun X., Genzong X., Benan T., Hua B., Longshi Y., 2007. Analysis of the ecological vulnerability of the western Hainan Island based on its landscape pattern and ecosystem sensitivity. *Acta Ecol Sin* 2007, 27, 1257-64.
- Popper, J., 1973. *La Dynamique des systèmes, principes et applications*. Paris, Éditions d'organisation. 273p.
- Pribyl, D., W., 2010. A critical review of the conventional SOC to SOM conversion factor, *Geoderma*, 156, 75-83
- R Development Core Team, 2011. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria, 2011, 409p.
- Rahim, L., et Mansoor, S., 2008. Proposed Design Notation for Model Transformation, *Software Engineering*, 2008. ASWEC 2008. 19th Australian Conference, 589-598.
- Rees, R., M., Bingham, I., J., Baddeley, J., A., Watson, C., A., 2005. The role of plants and land management in sequestering soil carbon in temperate arable and grassland ecosystems. *Geoderma* 128, 130-154

- Reuillon, R., 2008. *Simulations stochastiques en environnements distribués / Applications aux grilles de calcul*. Blaise Pascal University, Ph.D. thesis, 255p.
- Reuillon, R., 2012. Experimenting on complex-system models in the cloud. LIMOS, Clermont-ferrand, 14 décembre 2012, disponible via <http://public.iscpif.fr/~reuillon/limos2012/>
- Reuillon, R., Chuffart, F., Leclaire, M., Faure, T., Dumoulin, N., Hill, D., 2010. Declarative task delegation in OpenMOLE. HPCS 2010, Caen, June 28 – July 2, 55-62.
- Reuillon, R., Hill, D., El Bitar, Z., Breton, V., 2008. Rigorous distribution of stochastic simulations using the DistMe toolkit, IEEE Transactions On Nuclear Science, 55, pp. 595-603.
- Riedo, M., Grub, A., Fuhrer, J., 1998. A pasture simulation model for dry matter production, and fluxes of carbon, nitrogen, water and energy. Ecological Modelling 105, 141–183.
- Riedo, M., Milford, C., Schmid, M., Sutton, M., A., 2002. Coupling soil-plant-atmosphere exchange of ammonia with ecosystem functioning in grasslands. Ecological Modelling, 158, 83–110.
- Roekaerts, R., 2002. *The biogeographical regions map of Europe. Basic principles of its creation and overview of its development*. European Environment Agency, Copenhagen, 17p.
- Ruget, F., Novak, S., Granger, S., 2006. Use of the ISOP system, based on STICS model for the assessment of forage production. Adaptation to grassland and spatialized application. Fourrages, 186, 241–256.
- Rumpf, M., Strzodka, R., 2001. Level set segmentation in graphics hardware. International Conference on Image Processing 2001, 1103-1006.
- Sacks, J., Welch, W.,J., Mitchell, T.,J.,Wynn, H.,P., 1989. Design and analysis of computer experiments. Stati Sci 4: 409–435.
- Salamon, P., Sibani, P., Frost, R., 2002. *Facts, Conjectures, and Improvements for Simulated Annealing*. Society for Industrial and Applied Mathematics, Philadelphia, PA. 155p.
- Sanchez, S., M., 2005. NOLH designs spreadsheet. Disponible en ligne via <http://diana.cs.nps.navy.mil/SeedLab/>, consulté le 27/11/2012.
- Sanchez, S., M., Sanchez, P., J., Ramberg, J., S., Moeeni, F., 1996. Effective engineering design through simulation. Internat. Trans. Oper. Res. 3 169–185.
- Sanchez, S., M., Wu, H., -F., 2003. Frequency-based designs for terminating simulations: A peace-enforcement application. S. Chick, P. J. Sánchez, D. Ferrin, D. J. Morrice, eds. Proc. 2003 Winter Simulation Conf. Institute of Electrical and Electronics Engineers, Piscataway, NJ, 952–959.
- Scheffer, M., Bascompte, J., Brock, W., A., Brovkin, V., Carpenter, S., R., Dakos, V., Held, H., van Nes, E., H., Rietkerk, M., Sugihara, G., 2009. Early-warning signals for critical transitions. Nature 461, 53-59.

- Scheffer, M., Carpenter, S., Foley, J., A., Folke, C., Walker, B., 2001. Catastrophic shifts in ecosystems. *Nature* 413, 591-596.
- Schmid, M., Neftel, A., Riedo, M., Fuhrer, J., 2001. Process-based modelling of nitrous oxide emissions from different nitrogen sources in mown grassland. *Nutr. Cycl. Agroecosys.*, 60, 177–187.
- Schneider, S., H., Lane, J., 2006. Dangers and thresholds in climate change and the implications for justice. Dans: Adger W.N., Paavola J., Huq S., Mace MJ, eds. *Fairness in adaptation to climate change*. Cambridge, MA: MIT Press, 23-51.
- Seidewitz, E., 2003. What models mean. *Software, IEEE*, 20, 26-32.
- Seidl, R., Rammer, W., Lexer, M., 2010. Climate change vulnerability of sustainable forest management in the Eastern Alps. *Climatic Change*: 1-30.
- Shiers, J., 2009. Grid today, clouds on the horizon. *Computer Physics Communication* 180, 559-563.
- Shoemaker, A., C., Tsui, K.-L., Wu., C., F., J., 1991. Economical experimentation methods for robust design. *Technometrics* 33 415–427.
- Simon, H., A., 1981. *The Sciences of the Artificial*, 2nd ed. MIT Press, Cambridge, MA, 247p.
- Skondras, N., A., Karavitis, C., A., Gkotsis, I., I., Scott, P., J., B., Kaly, U., L., Alexandris, S., G., 2011. Application and assessment of the Environmental Vulnerability Index in Greece, *Ecological Indicators*, Volume 11, Issue 6, 1699-1706
- Sobol', I., M., 1967. On the distribution of points in a cube and the approximate evaluation of integrals, *U.S.S.R Computational Mathematics and Mathematical Physics*, 7, 86–112.
- Solomon, S., Battisti, D., Doney, S., Hayhoe, K., Held, I., M., Lettenmaier, D., P., Lobell, D., Matthews, D., Pierrhumbert, R., Raphael, M., Richels, R., Root, T., L., Steffen, K., Tebaldi, C., Yohe, G., W., 2011. *Climate stabilization targets: emissions, concentrations, and impacts over decades to millennia*. The National Academies Press, Washington, D.C., 298 pp.
- Soussana, J., -F., Vertès, F., Arregui, M., C., 1995. The regulation of clover shoot growing points density and morphology during short-term clover decline in mixed swards. *European Journal of Agronomy*, 4, 205–215.
- Stephen, L., Downing, T., E., 2001. Getting the scale right: a comparison of analytical methods for vulnerability assessment and household level targeting. *Disasters*, 25, 113-135.
- Taguchi, G., 1987. *System of Experimental Designs*, Vol. 1, 2. UNIPUB/Krauss International, White Plains, NY, 1176p.
- Tang, B., 1993. Orthogonal array-based Latin hypercubes. *J. Amer. Statist. Assoc.*, 88, 1392-1397.

- Thornley J., H., M., 1998. Grassland dynamics: an ecosystem simulation model. CAB International, Wallingford, United Kingdom, p 241.
- Thornley, J., H., M., 2001. Modelling grassland ecosystems. Proceedings of the XIX International Grassland Congress, in: Mattos W.R.S., da Silva S.C. (Eds.), Sao Paulo, Brazil, 1029–1035.
- Thornley, J., H., M., Cannell, M., G., R., 1997. Temperate grassland responses to climate change: An analysis using the Hurley pasture model. *Ann. Bot.*, 80, 205–221.
- Thornton, P., E., Rosenbloom, N., A., 2005. Ecosystem model spin-up: estimating steady state conditions in a coupled terrestrial carbon and nitrogen cycle model. *Ecological Modelling* 189, 25–48.
- Tsaregorodtsev, A., Garonne, V., Closier, J., Frank, M., Gaspar, C., van Herwijnen, E., Loverre, F., Ponce, S., Graciani Diaz, R., Galli, D., Marconi, U., Vagnoni, V., Brook, N., Buckley, A., Harrison, K., Schmelling, M., Egede, U., Bogdanchikov, A., Korolko, I., Washbrook, A., Palacios, J.P., Klous, S., Saborido, J., J., Khan, A., Pickford, A., Soroko, A., Romanovski, V., Patrick, G., N., Kuznetsov, G., Gandelman, M., 2003. DIRAC-distributed infrastructure with remote agent control, Computing in High Energy and Nuclear Physics, 24–28 March 2003, La Jolla, California, TUAT006.
- Tubiello, F., N., Fischer, G., 2007. Reducing climate change impacts on agriculture: global and regional effects of mitigation, 2000–2080. *Technological Forecasting and Social Change*, 74, 1030–1056.
- Turner, B., L., II, Kasperson, R., E., Matson, P., A., McCarthy, J., J., Corell, R., W., Christensen L., Eckley, N., Kasperson, J., X., Luers, A., Martello, M., L., Polsky, C., Pulsipher, A., Schiller A., 2003a. A framework for vulnerability analysis in sustainability science. *Proceedings of the National Academy of Sciences of the United States of America*, 100, 8074–8079.
- Turner, II, B., L., Matson, P., A., McCarthy, J., J., Corell, R., W., Christensen, L., Eckley, N., Hovelsrud-Broda, G., Kasperson, J., X., Kasperson, R., E., Luers, A., Martello, S., Mathiesen, M., L., Naylor, R., Polsky, C., Pulsipher, A., Schiller, A., Selin, H., Tyler, N., 2003b. Illustrating the coupled human-environment system for vulnerability analysis: three case studies. *Proceedings of the National Academy of Science* 100, 8080–8085.
- UNFCCC, 2004. Application of methods and tools for assessing impacts and vulnerability, and developing adaptation responses: background paper. FCCC/SBSTA/2004/INF.13. Bonn, Germany. <http://unfccc.int/resource/docs/2004/sbsta/inf13.pdf>.
- van Beers, W., C., M., Kleijnen, J., P., C., 2003. Kriging for interpolation in random simulation. *J Oper Res Soc* 54, 255–262.
- Van Nes, E., H., Scheffer, M., 2007. Slow recovery from perturbations as a generic indicator of a nearby catastrophic shift. *Am Nat* ;169:738–47.
- Vías, J., Andreo, B., Ravbar, N., Hötzl, H., 2010. Mapping the vulnerability of groundwater to the contamination of four carbonate aquifers in Europe. *J Environ Manag* 91, 1500–1510.

- Volenc, J., J., Ourry, A., Joern, B., C., 1996. A role for nitrogen reserves in forage regrowth and stress tolerance. *Physiologia Plantarum*, 97, 185–193.
- von Laszewski, G., Foster, I., Gawor, J., Lane, P., 2001. A Java Commodity Grid Kit.. *Concurrency and Computation: Practice and Experience*, 13, 643-662, 2001.
- von Laszewski, G., Hategan, M., Kodeboyina, D., 2006. Work Coordination for Grid Computing. *Grid Technologies: Emerging from Distributed Architectures to Virtual Organizations*, 5, 309-329.
- Vuichard, N., Ciais, P., Viovy, N., Calanca, P., Soussana, J.-F., 2007b. Estimating the greenhouse gas fluxes of European grasslands with a process-based model: 2. Simulations at the continental level. *Global biogeochemical cycles* 21, GB1005.
- Vuichard, N., Soussana, J.-F., Ciais, P., Viovy, N., Ammann, C., Calanca, P., Clifton-Brown, J., Fuhrer, J., Jones, M., Martin, C., 2007a. Estimating the greenhouse gas fluxes of European grasslands with a process-based model: 1. Model evaluation from in situ measurements, *Global Biogeochem. Cycles*, 21, GB1004.
- Walker, D., W., 1994. The design of a standard message passing interface for distributed memory concurrent computers. *Parallel Computing*, 20, 657-673.
- Walter, I., Allen, R., Elliott, R., Jensen, M., Itenfisu, D., Mecham, B., Howell, T., Snyder, R., Brown, P., Echings, S., Spofford, T., Hattendorf, M., Cuenca, R., Wright, J., and Martin, D., 2000. ASCE's Standardized Reference Evapotranspiration Equation. *Watershed Management and Operations Management* 2000, 1-11.
- Wang, Y., Fang, K., T., 1981. A Note on Uniform Distribution and Experimental Design. *KeXue TongBao*, 26, 485-489.
- Warren, R., de la Nava Santos, S., Arnell, N., Bane, M., Barker, T., Barton, C., Ford, R., Füssel, H.-M., Hankin, R., K., Klein, R., Linstead, C., Kohler, J., Mitchell, T., Osborn, T., Pan, H., Raper, S., Riley, G., Schellnhuber, H., Winne, S., Anderson, D., 2008. Development and illustrative outputs of the Community Integrated Assessment System (CIAS), a multi-institutional modular integrated assessment approach for modelling climate change. *Environ. Model. Software* 23, 592-610.
- Weedon, G., P., Gomes, S., Viterbo, P., Shuttleworth, W., J., Blyth, E., Österle, H., Adam, J., C., Bellouin, N., Boucher, O., Best, M., 2011. Creation of the WATCH Forcing Data and Its Use to Assess Global and Regional Reference Crop Evaporation over Land during the Twentieth Century. *Journal of Hydrometeorology*, 12, 823-848.
- Williams, D., N., Ananthakrishnan, R., Bernholdt, D., E., Barathi, S., Brown, D., Chen, M., Chervenak, A., L., Cinquini, L., Drach, R., Foster, I., T., Fox, P., Fraser, D., Garcia, J., Hankin, S., Jones, P., Middleton, D., E., Schwidder, J., Schweitzer, R., Schuler, R., Shoshani, A., Siebenlist, F., Sim, A., Strand, W., G., Su, M., Wilhelmi, N., C., 2009. The Earth System Grid: enabling access to multimodel climate simulation data. *Bulletin of the American Meteorological Society* , 90, 195-205.
- Woledge, J., Davidson, K., Dennis, W., D., 1992. Growth and photosynthesis of tall and short cultivars of white clover with tall and short grasses. *Grass and Forage Science*, 47, 230-238.

- Wolniewicz, P., Meyer, N., Stroinski, M., Stuempert, M., Kornmayer, H., Polak, M., Gjermundrod, H., 2007. Accessing Grid computing resources with g-Eclipse platform. *Computational Methods in Science and Technologie*, 13, 131—141.
- Woo, M., Neider, J., Davis, T., Shreiner, D., 1999. *OpenGL Programming Guide: the official guide to learning OpenGL, version 1.2*. Addison-Wesley Longman Publishing Co., Inc., 784p.
- Wu, C., F., J., et Hamada, M., 2000. *Experiments: planning, analysis, and parameter design optimization*. New York: Wiley, 638p.
- Wutzler, T., Reichstein, M., 2007. Soils apart from equilibrium – consequences for soil carbon balance modelling. *Biogeosciences* 4, 125-136.
- Xia, J., Luo, Y., Wang, Y., -P., Weng, E., Hararuk, O., 2012. A semi-analytical solution to accelerate spin-up of a coupled carbon and nitrogen land model to steady state. *Geoscientific Model Development*, 5, 1259-1271.
- Yadav, V., Malanson, G., 2007. Progress in soil organic matter research: litter decomposition, modelling, monitoring and sequestration. *Progress in Physical Geography* 31, 131–154.
- Ye, K., Q., 1998. Orthogonal Column Latin Hypercubes and Their Application in Computer Experiments. *Journal of the American Statistical Association*, 93, 1430-1439.
- Yeluripati, J., B., van Oijen, M., Wattenbach, M., Neftel, A., Ammann, A., Parton, W., J., Smith, P., 2009. Bayesian calibration as a tool for initialising the carbon pools of dynamic soil models. *Soil Biology & Biochemistry* 41, 2579-2583.

Annexes

I. Description des projets Européens

i. CARBO-Extreme

Abstract of the European project FP7 n°226701, called CARBO-Extreme (Greenhouse gas management in European land use systems):

The aim of this project is to achieve an improved knowledge of the terrestrial carbon cycle in response to climate variability and extremes, to represent and apply this knowledge over Europe with predictive terrestrial carbon cycle modelling, to interpret the model predictions in terms of vulnerability of the terrestrial – in particular soil – carbon pools and give according advice to EU climate and soil protection policies. This objective will be achieved by integrating three major types of recent and new solid scientific carbon cycle data, from: (i) soil process studies, (ii) a network of established ecosystem manipulation experiments, and (iii) longterm observations spanning several times-scales (e.g. eddy covariance data, tree rings and growth, crop yields, long-term remote sensing data on soil moisture and vegetation activity and soil carbon inventories). The integration will be reached by establishing a consistent and harmonized data base and by confronting the terrestrial carbon cycle models with the multiple data sets within a Bayesian model identification and improvement procedure. Specific model development concerning processes affected by extreme events (e.g. soil carbon destabilization, tree growth response incl. lag effects and mortality) will be included and followed by model testing and improvement against the data made available in the project. The improved models will simulate terrestrial processes relevant to carbon balance and soil erosion at pan- European scale using regionalized climate scenarios with explicit inclusion of extreme climatic events. Since we are using several climate scenarios and an ensemble of models we will be able to characterize the uncertainties in prediction coming from models and climate scenarios. We will interpret the empirical evidence from the observational work and the model simulations in a framework of vulnerability assessment and disseminate and discuss results with stakeholders at EU level.

Description of WP7 (“Carbon vulnerability synthesis”):

In this work package we aim at providing an integrated assessment of the vulnerability of the European carbon cycle to climate change and extreme events based on the scientific and quantitative basis created in the observational and modelling work packages.

We define here vulnerability as the probability of impacts on the carbon cycle to exceed a given threshold. The threshold itself is defined in terms of impacts on ecosystem services (thereby incorporating the human dimension of vulnerability) and the evaluation is carried out for model results derived in WP6 using a range of likely scenarios of climate and land use change. Impacts will be assessed both in terms of C pools (e.g. soil, biomass) and of C fluxes (e.g. annual GPP, Reco, NPP and NBP). For each emission scenario and for an ensemble of ecosystem models, and for a range of land use change scenarios, the conditional probability of reducing by 20, 50, 80 %

the annual continental scale ecosystem C sink of Europe will be assessed in **Task 7.1**. Moreover, the mean return period (in years) of an annual continental scale C source from European ecosystems will be assessed for the contrasted emission scenarios and for time slices ranging from 2020 to 2100. The consistency of these projections will be assessed by comparisons with observations (WP5) establishing the sensitivity of the C cycle to climatic variability.

Changes in the continental scale C balance will be translated into a key policy relevant indicator in **Task 7.2**. This indicator shows the additional emission reduction commitments (for the continent, for EU27 and for given countries) which will be needed to reach targets accounting for the decline in European C sinks. This will be calculated as the conditional probability that the quantified emission reduction commitments of Europe should be increased by 5, 10 or 20 % to counter reductions in ecosystems C sinks by 2020 and 2050 for contrasted SRES scenarios (cf. **Task 7.1**). A major product of the CARBO-Extreme project for European post-Kyoto policies will therefore be maps showing the risks, in a full C accounting approach, of not attaining targets in emission reduction, as a result of a weakening of the European C sink. While this Europe-centred approach neglects the vulnerability of major C pools worldwide (e.g. tropical forests, savannas, permafrost), it is of high policy relevance given the need to stabilize atmospheric CO₂ concentration, by reducing in each region of the world the net amounts of CO₂ that are added each year to the atmosphere by all sectors and by the ecosystem feedbacks.

The overall continental scale assessment by **Task 7.3** will be strengthened by comparing the conditional vulnerability of the main ecosystem types (e.g. forests, crops, pastures) under these scenarios, taking into account both adaptations of agricultural and forest management and of land use in response to risks from climate extremes. This approach will allow estimating the vulnerability of each ecosystem type and the extent of adaptation showing to which extent adaptation allows buying time and to delay impacts (Howden et al., 2007). The outputs from tasks 5.3, 6.2, 6.3, and 6.4 will be used to define those aspects of variability / extremes that provoke specific vulnerabilities in specific ecosystem-dependent economic sectors (e.g., forest fire risk, drought-triggered crop failure, soil carbon loss and erosion), in addition to other C fluxes. Maps of thresholds, risks and vulnerability will be produced, and where possible, the probability range of a given consequence will be quantified and mapped.

Objectives:

- 1) To provide an assessment of the vulnerability of the carbon cycle to climate change and extreme events by sector and ecosystem type
- 2) To calculate the probability of different reductions of the European terrestrial carbon sink conditional on emission scenarios, as a map and for the whole continent
- 3) To develop an indicator that shows the additional emission reduction commitment necessary when a full carbon accounting approach is adopted for long-term emission goals.
- 4) To develop recommendations to reach the '2°C-guardrail' from a European perspective

Description of work (possibly broken down into tasks) and role of partners:

Task 7.1 Conditional probabilities of % change in continental scale annual C sink – lead: UNIABDN (Martin Wattenbach)

- The outputs from WP6 will be used to calculate for each emission scenario the probability that certain reductions of carbon sinks will occur or that carbon sinks turn into sources
- This probability will be given spatially explicit and aggregated by sectors and ecosystem types and appropriate administrative boundaries
- Mean return period of an annual continental scale C source from European ecosystems will be assessed for the contrasted emission scenarios and for time slices ranging from 2020 to 2100.

Task 7.2 Continental scale carbon vulnerability indicator – lead: INRA (Jean-Francois Soussana)

- Changes in the continental scale ecosystem source/sink C balance will be translated into a key policy relevant indicator.
- This indicator shows the additional emission reduction commitments (for the continent, for EU27 and for given countries) which will be needed to reach targets accounting for the decline in European C sinks in a full carbon accounting approach.
- This will be calculated as the conditional probability that the quantified emission reduction commitments of Europe should be increased by 5, 10 or 20 % to counter reductions in ecosystems C sinks by 2020 and 2050 for contrasted SRES scenarios

Task 7.3 Recommendation of adaptation and mitigation strategies – lead: PIK (Wolfgang Cramer)

- The overall continental scale assessment will be strengthened by comparing the conditional vulnerability of the main ecosystem types (e.g. forests, crops, pastures) under these scenarios, taking into account both adaptations of agricultural and forestry management and of land use in response to risks from climate extremes.
- This approach will allow estimating the vulnerability of each ecosystem type and the extent of adaptation showing to which extent adaptation allows to buy time and to delay impacts

ii. GHG Europe

Abstract of the European project FP7 n°244122, called GHG Europe (Greenhouse gas management in European land use systems):

The GHG-Europe project aims to improve our understanding and capacity for predicting the European terrestrial carbon and greenhouse gas (GHG) budget by applying a systematic, comprehensive and integrative approach. GHG-Europe quantifies the annual to decadal variability of the carbon and GHG budgets of terrestrial ecosystems in EU27 plus Switzerland and in six data-rich European regions via data-model integration, diagnostic and predictive modelling. Models are calibrated by multi-site observations. Research includes CO₂, CH₄ and N₂O in forests, croplands, grasslands, shrublands, peatlands and soils. Via an integrated approach, GHG Europe scales up consistently from local to regional and continental scale via scale dependent error propagation and systematic quantification of uncertainties, model validation at different scales and top-down verification by atmospheric inversion models. At regional and European scale lateral C transport by land use, trade and rivers are included. Variability in C and GHG budgets is attributed to natural (climate) and anthropogenic drivers (N deposition, land use, past and present management) by synthesis of past and emerging experiments, targeted observations in hot spots and hot moments and model sensitivity analyses. For this purpose, observations are extended to under-sampled regions and ecosystems with likely high importance for the European C budget: forests and land use change in Eastern Europe and Mediterranean shrublands. The future vulnerability of carbon pools and risks of positive feedbacks in the climate-carbon system are assessed by scenario analyses with biophysical models and by integrating feedbacks with socio-economic changes and EU climate and land use policies. GHG-Europe uses a bidirectional interaction with stakeholders to provide regular and timely scientific advice targeted to the emerging needs of the UNFCCC process and for implementing post-2012 climate commitments in Europe.

iii. ANIMALCHANGE

Abstract of the European project FP7-KBBE-2010-4, called ANIMALCHANGE (AN Integration of Mitigation and Adaptation options for sustainable Livestock production under climate CHANGE):

ANIMALCHANGE will provide scientific guidance on the integration of adaptation and mitigation objectives and design sustainable development pathways for livestock production in Europe, in Northern and Sub-Saharan Africa and Latin America. ANIMALCHANGE will inform public policy development in EU27 and propose cooperation programs addressing smallholder livestock farming in selected developing countries. The core analytical spine of the project is a series of coupled biophysical and socio-economic models combined with experimentation. This allows exploring future scenarios for the livestock sector under baseline and atmospheric CO₂ stabilization scenarios. These scenarios are first constructed in Component (CP) 1. They are

elaborated and enriched by breakthrough mitigation and adaptation options from CP 2 at field and animal scales, integrated and evaluated at farm scale in CP 3 and used to assess policy options and their socio-economic consequences in CP 4.

ANIMALCHANGE will:

- Quantify and reduce uncertainties in greenhouse gas (GHG) emissions and assess climate change impacts on livestock systems (including grasslands)
- Revise estimates of the GHG balance of livestock systems and integrate soil carbon sequestration
- Integrate climate variability and extremes into the assessment of impacts, adaptation and vulnerability of livestock systems to climate change
- Develop breakthrough technologies for adaptation and mitigation to climate change for both ruminants and monogastrics
- Study and quantify trade-offs and synergies between adaptation and mitigation options
- Assess the potential societal and sectoral costs and benefits of these options for the livestock sector in Europe and in study regions of Africa and Latin America
- Assess climate change vulnerability of animal production and of associated GHG emissions
- Provide direct support through the design of an integrated and consistent mitigation and adaptation policy framework for the livestock sector

Description of Task 5.2 (“Vulnerability analysis of changes in production”) of WP5 (“Climate change vulnerability of livestock systems”):

Leader: Gianni Bellocchi (INRA)

Involved partners: INRA, CEA, IIASA, ILRI, DLO, EMBRAPA, UP, FAO

Methods:

A range of statistical analyses (see Task 5.3 for the vulnerability analysis methods) will be used to provide metrics of vulnerability crop, pasture and animal production to the increased climatic risks taking into account: i) biotic interactions through increased incidence of pests, weeds, diseases and through likely changes in the botanical composition of grasslands (see WP4) and ii) autonomous adaptation of agronomic practices (e.g. for sowing dates, mowing and grazing dates).

Task 5.2 will map the probability of production losses of more than 20, 40 and 60 % compared to means for the reference period centered around 1990. Such maps will be generated for Europe and other project regions in Africa and Latin America and for contrasted time slices: e.g. 2030;

2050; 2100. Moreover, the comparison of the models and of the different climate scenarios for 2050 will provide an estimate of the uncertainty associated to these production losses.

Outcomes:

An assessment of risk of production failures and/or migration and transformation of crop types, pasture types and of their associated livestock systems without planned adaptation and mitigation measures.

Description of WP8 (“Integrating adaptation and mitigation options”):

Objectives:

In WP6 and WP7 breakthrough mitigation and adaptation options have been explored. The aim of WP8 is to test the implications of mitigation on the potential to adapt to climate change, and to test the implications of adaptation on the potential to mitigate greenhouse gas emissions.

Description of work and role of partners:

WP Leader: Agnes van den Pol-van Dasselaar (DLO)

Background:

Potential mitigation and adaptation options for animal agriculture have been summarized in various reviews focussing on livestock production, manure management and crop production. Mitigation options include fertilization, animal feeding, manure management, and soil, water and harvest management. In recent years, studies on adaptation options have been added. Adaptation options include among others plant breeding, animal genetics, change in forages and grassland management, and in grazing management. Some options can be foreseen: increase in grassland species diversity, increased use of legumes and other available forages better adapted to drought, winter grazing, increased housing during heat, and so on. However, there is a lack of data on the mitigating effect of these options, whereas consequences on GHG emissions will be important. Similarly, there is a lack of data on the effect of mitigation options on the potential to adapt to climate change. Trade-offs between mitigation and adaptation have hardly been explored by the literature, although there is circumstantial evidence of side effects of mitigation on adaptation (and conversely of adaptation on mitigation). For instance:

- Mitigating enteric methane per unit product in dairy systems requires animal breeds with a high production potential, leading to large heat dissipation needs and, hence, suffering more from heat events than breeds with low production potential;
- Coping with the after-effects of a severe drought may require ploughing up degraded grasslands leading to soil organic carbon loss.

Evaluating mitigation / adaptation:

Gathering evidence of the mitigation options and the effects of adaptation should not be limited to empirical approaches (work package 6 and 7; component 3). Interactions occur between various aspects of farm management and between various emission sources. Gathering information on a more theoretical basis is essential to provide an understanding of the mechanisms involved and a realistic view of how various emission sources are related to some key farm parameters (e.g. level of fertilization, crop yield, animal feed intake, animal productivity), and how they are interrelated. An example are the interactions between climate change, roughage quality, roughage digestibility, enteric methane emission, animal production level achieved, animal excreta and manure quality, manure application, soil fertility and soil emissions. Furthermore, stakeholders will not utilise or evaluate one mitigation and/or adaptation tool but rather a combination and how these interact is really important to understand, either by modelling or measuring.

Detailed models aid in the interpretation of experimental results and identify the consequences of the experimental set-up chosen and the key factors underlying the observations. Also these models aid in pinpointing the difficulties and uncertainties (see Component 1) with inventories of greenhouse gas emissions on a regional or even higher scale. Process based models are to be preferred (Ellis et al., 2008; Kebreab et al., 2006) at the animal, field and manure scales. The same suite of process based models used in Component 1 to reduce uncertainties in GHG emissions (WP3) and in climate change impacts on crop, pasture and animal productivity (WP4) will be used in WP8 for investigating possible side-effects of mitigation options on adaptive capacity to climate change (including climatic variability) and – conversely – of adaptation options on GHG emissions (including carbon sequestration).

Approach:

WP 2.3 will:

- Develop a framework to link details gathered on various farming aspects addressed in CP1 and in WP6 and WP7 in a systematic manner. Aspects addressed are 1) ruminant nutrition (feed intake, enteric fermentation, digestion, excretion, excreta composition and characteristics, animal productivity), 2) manure storage and application (storage conditions, manure characteristics, application to soil), and 3) soil conditions and productivity (crop yield, carbon sequestration, manure application, leaching and emissions). Potential synergies and trade-offs between adaptation and mitigation options will be context dependent (e.g. type of region or climate, type of production system, type of farm management). The framework will assist in the development of working and communication schedules that enable a linkage between information gathered in CP1, WP6, WP7 and CP3. It will allow the identification of options which are likely beneficial both for adaptation and for mitigation.
- Interpret results for options tested or evaluated in work package 6 (mitigation) and work package 7 (adaptation) by using detailed process based models also used in Component 1. Translate applicability and consequences (as far as possible) of these options for the wider range of conditions to be evaluated in CP3.

- Assess major options in terms of both mitigation and adaptation potential derived from modelling approaches for a range of systems and regions, or from consultation of stakeholders, literature and results gathered in CP1 and CP3 when the quality of data is insufficient for appropriate use of the detailed process based models applied in this work package.
- Evaluate effects of farm adaptation on the potential to mitigate emissions, and evaluate effects of mitigation measures on the potential to adapt to climate change. Furthermore, it will evaluate how to integrate both.

For detailed description of tasks, see full document:

Task 8.1: Gathering of site-specific information and assessment of mitigation and adaptation options for a range of systems and regions.

Task 8.2: Integrating mitigation and adaptation options in intensive ruminant production systems by process based modelling.

Task 8.3: Integrating mitigation and adaptation options in extensive pasture based systems by process based modeling.

Task 8.4 Integrating adaptation and mitigation scenarios

Description of WP15 (“Data Infrastructure and Model support”):

Objectives:

The first objective of this component establishes the data infrastructure of the project through a series of interconnected data bases, that will link the activities in the four research components (CP1 to CP4), receive data inputs from partner research projects (e.g. meteorological data as simulated by regional climate models of IPCC) and provide data outputs to the scientific community and to stakeholders. Therefore, CP5 is the backbone of the project in terms of workflow.

The second objective of CP5 is to facilitate the use of models, their parametrisation and their interpretation by providing a series of methods and of tools that will be used throughout the project for comparing models with data from experiments and from long-term observations, by developing Bayesian methods to reduce uncertainties in key parameter values and, when relevant, by using data-model fusion methods to improve model accuracy. Moreover, given the stochastic nature of climate risks, the interpretation of climate change impacts is not trivial and requires advanced statistical methods to disentangle the climate change signal from the ‘noise’ linked to uncertainties and to other causes of variability in pasture, crop and animal production as well as in GHG emissions. Therefore, CP5 will supply a set of tools and methods for developing climate change vulnerability analyses.

Description of work and role of partners:

Component Leader: Raphaël Martin (INRA)

Background:

ANIMALCHANGE requires a systematic, comprehensive and integrative approach, where a multitude of experimental data, of socio-economic data, of GHG measurements, of farm data and of long-term observations along with meteorological observations is brought together with state-of-the-art process modelling approaches.

This objective will be achieved by integrating three major types of scientific data, along with meteorological data and scenarios, from: (i) long-term observations spanning several times-scales (e.g. eddy covariance data, crop yields, animal production, long-term remote sensing data on soil moisture and vegetation activity, soil data bases), (ii) a network of established ecosystem and animal manipulation experiments in CP1 and CP2, (iii) a network of highly documented case-study farms in CP3. Moreover, detailed geo-referenced agricultural statistics data will be used in CP1 and in CP4 for the regional modelling of livestock systems.

The inter-operability of databases in which these data are stored needs to be achieved through an efficient framework, building on the wealth of efforts that has already been made for the consolidation of data in Europe and elsewhere. ANIMALCHANGE has all the key partners that already have these data bases and are willing to share them within the project, including major international institutions (FAO, ILRI, IIASA), the Joint Research Centre of the European Commission, and other partner institutions which have already an integrated data management system in place for these issues, including CEA, DLO, TEAGASC, AU, EMBRAPA and UP. The integration will be reached by establishing a consistent and harmonized data base and by confronting the models with the multiple data sets within a Bayesian model identification and improvement procedure.

Approach:

WP15 will create the interface between the data collections and the modelling and data mining activities. Data collection, quality control, technical standardization and harmonization together with detailed metadata preparation and uncertainty quantification will be the main activities. In addition, a range of statistical methods and tools will be developed and provided to other components for uncertainty, sensitivity and vulnerability analyses in models outputs. All the information will be available inside the project through a database interface.

Work breakdown:

Task 15.1 Data management, system of the project: database architecture and consistency checks (M1 – M48)

Leader: Raphaël Martin (INRA)

Involved partners: IIASA, CEA, ILRI, JRC, FAO, DLO, TEAGASC, AU, UP, EMBRAPA

Objectives:

Large amounts of data will be produced and collected in WP1 to WP4 concerning different variables, acquired with different time and spatial scales and different methodologies. The use of this important information by the modelling activities and for data analysis purposes, needs a pre-processing of the datasets with the aim to control their quality, to standardize and harmonize them and to provide all the necessary meta information essential for a correct data use and interpretation in a “standard and ready-to-use” format. Care will be taken to associate relevant uncertainties to each dataset, in close interaction with the “data-experts”. These uncertainties are indeed crucial for Bayesian model parameters estimation approaches that will be used to reduce uncertainties in key model parameter values.

Methods:

The data infrastructure will be based on past positive experiences (e.g. in the Carboeurope-IP and NitroEurope IP projects), realizing that datasets in ANIMALCHANGE are even more heterogeneous ranging from ecosystem manipulation experiments and long term observations at site level (eddy fluxes, yield history, etc.), to regional and continental level (soil maps, farming systems, remote sensing...) in CP1 and CP2, through farm scale in CP3, but also the model output (this way observed data can be used by modelling WPs, the main flow of information in the project, but also modelled data can be used by the experimentally working partners). New quality control, standardization and harmonization methods will be developed with the aim to give a consistent data structure to the database users particularly concerning the concurrent use of different datasets coming from different sources. Different levels of data processing will be considered (Level 0: collected raw data, units and time steps not harmonised, Level 1: data harmonized; Level 2: data where observational operator have been applied).

The task is divided into three steps, one responsible for the database system design, building and maintenance and the other two for data collection, QA/QC, standardization, harmonization and ingestion. The large differences in the datasets characteristics suggested to split these last activities in two groups based mainly on the time-series length because of consequent differences in the data pre-processing methods and also in the metadata collection. Data quality control will be performed using standard and objective methods to detect and correct possible inconsistencies or measurement problems. All the quality control activities will be recorded and also the original data will be maintained. The quality controlled data will be then harmonized to ensure the consistencies in format and units also between different data type and loaded in the database system.

Task 15.1 will define, build-up and maintain the database system. The database system will function as link between the experimental and data collection parts of the project and the modelling and data analysis activities. In order to avoid problems with the data availability, it is important to have a clear picture of the possible datasets that will be part of the database to design the system. We will design the structure of the system; therefore examples of all the datasets that will be part of the database will be collected. The interaction with both, the “data experts” and “data-users” of the project, is pivotal to understand what is needed and what is available in order to define the data processing needed. The database structure will be created based on an open access web interface where it will be possible to have an overview of all the information available and a protected section where the data will be accessible to the project participants. Moreover, quality control tools will be developed and applied metadata information prepared according to international standards.

Outcomes:

Task 15.1 will deliver a consistent quality controlled data infrastructure for ANIMALCHANGE with full WEB access for partners and with publicly accessible metadata information.

Task 15.2 Tools for models evaluation and parameterization (M1 – M48)

Leader: Gianni Bellocchi (INRA)

Involved partners: INRA

Objectives:

When using empirical information in modelling, it is important to not only account for uncertainties about model structure, but also to recognize measurement uncertainties associated with the data. Data uncertainty needs to be accounted for in model parameterisation as well. The iterative nature of Bayes’ theorem incorporates existing knowledge and updates the joint probability distribution of model parameters as new information becomes available. Bayesian parameter estimation is especially useful when scattered, multivariate information on the modelled system is available to improve the understanding of model parameters (Arhonditsis et al., 2007). This kind of inference changes the perspective from seeking a single optimal value for each model parameter, to seeking a joint distribution of parameter sets, which then provide the basis for estimating model prediction error. We aim to employ a Bayesian approach to model parameterisation (Van Oijen et al., 2005), which will provide a translation of uncertainty in data to uncertainty in model parameters, and this is the primary aim of Task 15.2. The Bayesian approach is central to the integration of data and modelling in the project, and in effect provides a bookkeeping of uncertainties, from data to model parameterisation to model application. Therefore an additional responsibility for Task 15.2 is to ensure that information in ANIMALCHANGE is properly qualified with a measure of uncertainty.

Methods:

Practical guidance will be provided to all project partners in the data-model application chain about uncertainty quantification. Provided that quite a lot of this work will be done in other projects (such as FP6 NitroEurope, FP7 Carbo-Extreme, and FP7 GHG-Europe), ANIMALCHANGE will

do extra Bayesian analysis as required, based on similar methods and data/modelling protocols. Bayesian calibration will then be used to refine our knowledge of model input parameters, obtain insight into the degree of information the data contain about model inputs, and obtain estimates and uncertainty bounds for modelled output variables. In the first step, Bayesian calibration will be carried out using data already available at the start of the project, and starting from prior probability distributions for each model's parameters. Initial information on prior probabilities w. r. t. soil processes will be gained from WP1's data. Bayesian calibration will be applied at intensively studied sites by using local meteorological and soil data, and site history as drivers for site-level runs. In addition, it will be applied using spatially explicit data from WP3.

Outcomes:

With this task, the modelling work integrated with Bayesian analysis will provide improved model estimates and management actions over time and space. The result of this exercise will be updated parameter probability distributions. After model development is concluded in CP1 and CP3, and more data have become available in these components, a second and final Bayesian calibration will produce the posterior probability distribution for the parameters that is then available for all further model application in CP1 to CP4.

Task 15.3 Vulnerability analyses methods (M1 –M48)

Leader: Gianni Bellocchi (INRA)

Involved partners: INRA

The second part of the integration of data-driven and theory-based information will be combined analyses of observations and modelled patterns at different scales by Task 15.3.

Methods:

A range of statistical analyses will be developed to provide metrics of vulnerability to the increased climatic risks. A range of methods are available and already developed for assessing vulnerabilities and hazards in the field of risk management (e.g., Locatelli et al., 2008). Various vulnerability analysis methods will therefore be assessed for their relevance to the purposes of the project, i.e.:

- Variance decomposition analysis in order to weigh the climate change and climate variability signals compared to uncertainties
- Bayesian analyses to calculate the probability of different percentage reduction levels of crop, pasture and animal production conditional on baseline and stabilisation scenarios
- Logistic (yet multivariate) regression of vulnerability measures (e.g. probability to exceed critical thresholds) versus independent variables (either categorical and continuous)
- Fuzzy-logic based systems to relate to events that have no well-defined meaning, e.g. vulnerability thresholds. Fuzzy theory enables to deal with polymorphous and

ambiguous concepts for which a straightforward quantification is difficult, while mathematically handling the reasoning for these concepts and producing concrete unambiguous answers (also incorporating expert decisions, e.g. from stakeholders)

- Indicators for describing the different components of vulnerability while providing a clue to a matter of larger significance or makes perceptible a trend or phenomenon that is not immediately detectable

Outcomes:

A set of tested and consistent methods for vulnerability analyses in the project with their associated guidelines.

Contingency plan:

INRA has large experience in data management and quality insurance. The main risk would be that some partners would not provide their data. In such case, the Executive Committee will take action to ensure that data are delivered. With statistical methods, external advice will be sought if problems are encountered during method development.

II. NetCDF

Le format de fichier NetCDF (Network Common Data Form, <http://www.unidata.ucar.edu/software/netcdf/>) a été conçu pour gérer de grand volume de données scientifiques. Il s'agit d'un format ouvert de données « auto-documentées », indépendant de l'architecture matérielle qui permet la création, l'accès et le partage de données scientifiques stockées sous la forme de tableaux. Il est également accompagné d'un ensemble de bibliothèques avec une API pour le C, C++ et le FORTRAN, et un ensemble d'utilitaires (par exemple, Panoply⁹¹, ncBrowse⁹², ncview⁹³). Depuis la version 4, le format possède une extension permettant la gestion des entrées/sorties en parallèle⁹⁴.

Ce format possède cependant certaines faiblesses. Par exemple, le format alloue la mémoire même en absence de données (entraînant gaspillage de mémoire et réduction d'efficacité d'accès), même si pour compenser ce défaut la dernière version du format permet une certaine compression des données. Le format a été conçu pour être le plus efficace possible pour de gros accès de données (par exemple toute une ligne du tableau). À l'inverse, comme les fichiers sont généralement très volumineux (~10-0,1 Go), l'accès à une seule donnée est très coûteux en temps. Ce format représente un compromis entre taille et nombre de fichiers.

⁹¹ <http://www.giss.nasa.gov/tools/panoply/>

⁹² <http://www.epic.noaa.gov/java/ncBrowse/>

⁹³ http://meteora.ucsd.edu/~pierce/ncview_home_page.html

⁹⁴ <http://trac.mcs.anl.gov/projects/parallel-netcdf/>

Dans le cadre de cette thèse, le choix du format de données a été imposé par le projet européen EU-FP7 Carbo-Extreme⁹⁵. Il s'agit en effet d'un format classique de sortie des modèles climatiques, comme par exemple pour le cinquième rapport du GIEC [Williams et al., 2009]. Il existe des outils ergonomiques pour manipuler ce format et notamment créer des cartes (par exemple l'outil Panoply).

III. Description rapide de l'utilisation d'OCL dans EMF

Le formalisme OCL (*Object Constraint Language*) peut être utilisé dans EMF⁹⁶. Pour cela il faut :

1. dans les annotations du modèle, définir la source suivante :
<http://www.eclipse.org/emf/2002/Ecore> ;
2. spécifier les clés et les valeurs suivantes : invocationDelegates, settingDelegates, validationDelegates et pour valeur
<http://www.eclipse.org/emf/2002/Ecore/OCL> ;
3. sur l'élément sur lequel on souhaite faire porter les contraintes, définir deux sources :
 - <http://www.eclipse.org/emf/2002/Ecore> ; définir en entrée « constraints » et en valeurs le nom des contraintes à appliquer (par exemple mustHaveTitle, nonNegativeValue) ;
 - <http://www.eclipse.org/emf/2002/Ecore/OCL> ; définir en entrée le nom de la clé (par exemple mustHaveTitle, nonNegativeValue) et la contrainte en OCL (par exemple not name.oclIsUndefined(), Value > 0).

Ce formalisme permet facilement de vérifier les contraintes sur les métamodèles, via la phase de validation naturellement présente dans EMF.

IV. Tutoriel d'utilisation de l'outil pour l'analyse de vulnérabilité

De manière simplifiée, la saisie dans l'outil des données nécessaires à l'analyse de vulnérabilité se fait en six grandes étapes :

1. On commence par la racine : "Analyse de vulnérabilité"

On choisit un **modèle**

⁹⁵ <http://www.carbo-extreme.eu/>

⁹⁶ <http://wiki.eclipse.org/MDT/OCLinEcore>

- a. On définit les entrées qui peuvent être de type "Environnement" ou "Gestion". Seuls les paramètres de type gestion sont considérés pour l'adaptation, tandis que ceux de type environnement sont subis. Pour chacun d'eux on spécifie :
 - i. Leur nature : entier, double ou chaîne de caractère.
 - ii. Leur distribution et leur échantillonnage, par exemple un échantillonnage uniforme discret de 10 valeurs.
 - b. On définit également les sorties.
2. On définit ensuite le **plan expérimental initial** pour l'analyse de vulnérabilité sans adaptation, dont on spécifie sa nature, par exemple plan factoriel complet.
 3. On choisit la **nature des scripts de simulation** (par exemple OpenMOLE).
 4. On choisit le type de **surface de réponse** à construire (actuellement uniquement par des scripts R), par exemple régression linéaire.
 5. On choisit l'**indice de vulnérabilité**, par exemple Luers.
 6. Enfin on choisit une méthode d'optimisation

De façon à mieux expliciter les étapes, nous allons faire l'analyse de vulnérabilité d'un modèle de paire appliquée à des simulations de 30 ans, avec une fauche par an. Nous nous intéresserons à l'analyse de vulnérabilité sans adaptation.

Initialisation de l'analyse de vulnérabilité.

Lancer Eclipse. Créer un projet vide (File → New → Project ; Choisir General → Project ; on l'appellera « tuto_Evaluation_Vulnerability »).

Créer un nouveau projet pour l'analyse de vulnérabilité :

File→New → Other. Choisir dans "Example EMF Model Creation Wizards" la catégorie "Vulnerability Model" (Figure 59).



Figure 59. Création du projet d'analyse de vulnérabilité.

Puis choisir le nom de l'analyse. On choisira comme nom "MyVulnerabilityAssessmentTutorial" (Figure 60). Puis cliquer sur "Finish".

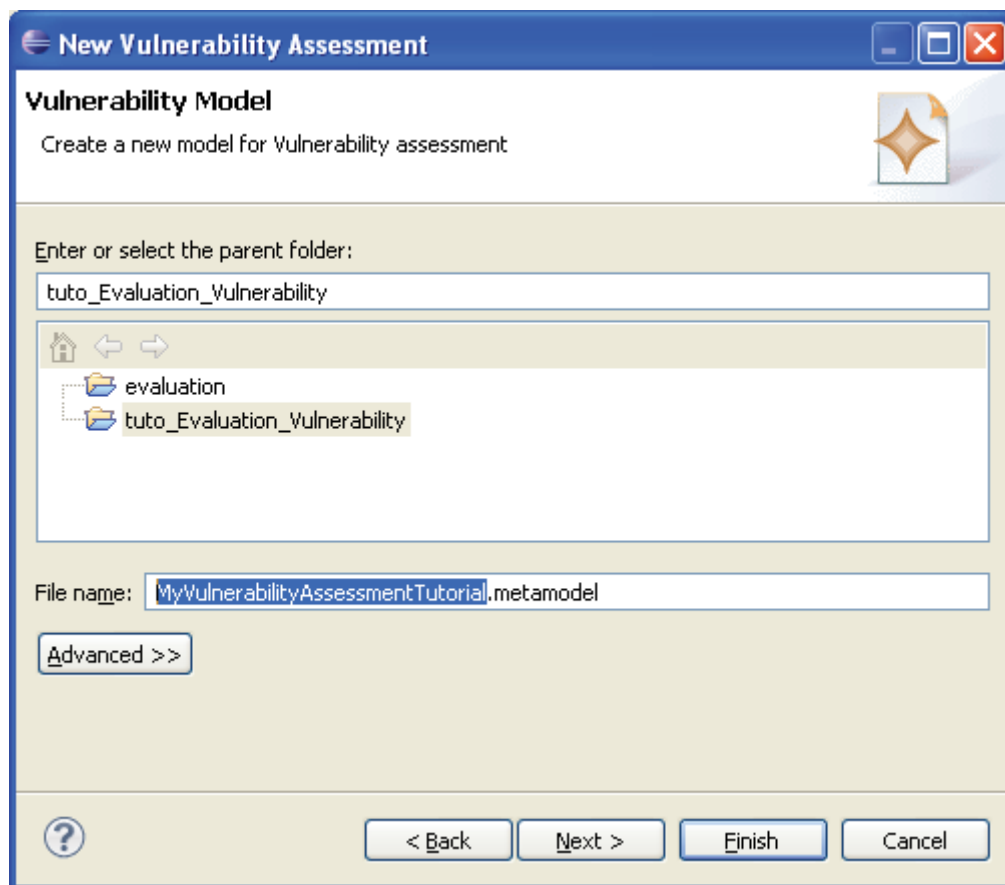


Figure 60. Choisir le nom du projet d'analyse de vulnérabilité.

On choisit un modèle

Sélectionner le nœud "Vulnerability Analysis". Clic droit → New Child → Modelagro (Figure 61). Cliquer sur ModelAgro. Clic droit → "Show Properties View". Dans cette fenêtre spécifier le nom du modèle (par exemple "Mon Modèle").

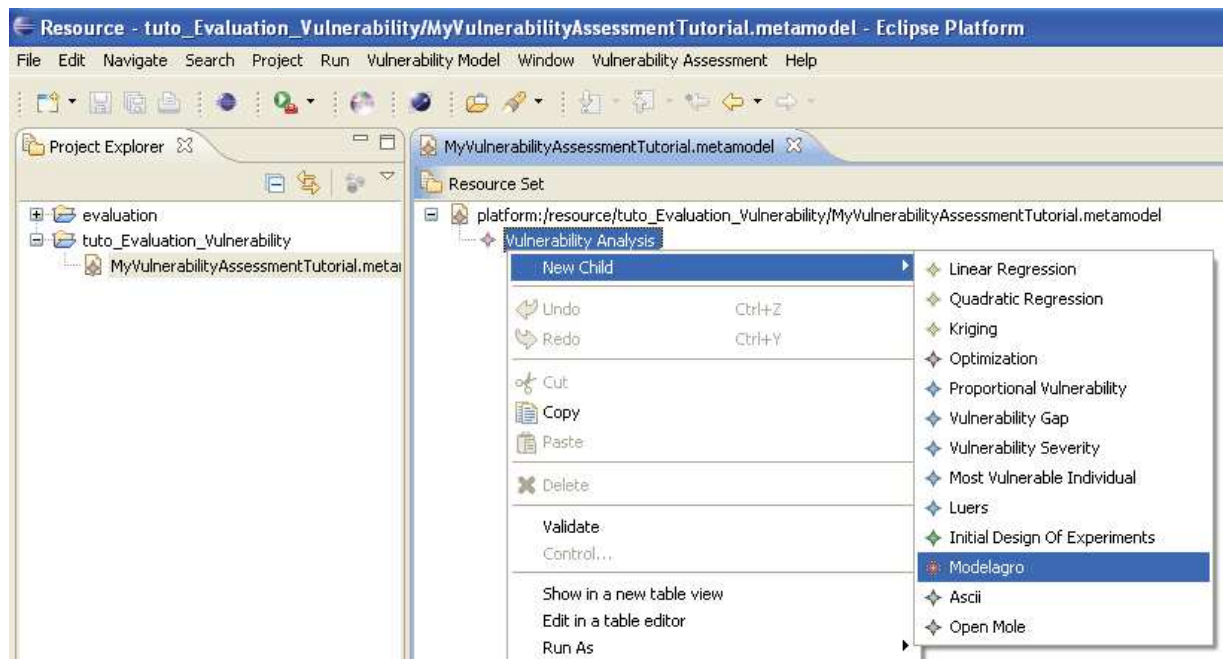


Figure 61. Insérer le modèle agronomique.

Nous allons maintenant définir les entrées et les sorties du modèle, à savoir la date de fauche en entrée et les valeurs de rendement sur 30 ans en sortie.

Se placer le modèle. Clic droit → New Child→Inputs Management (Figure 62). Nous rappelons que les entrées de type management sont des entrées sur lesquelles l'adaptation peut porter. Dans l'onglet de propriété, on lui affectera un nom (par exemple. "Fauche1"). Puis on lui choisira une nature (au sens type informatique : int, double, string, file), ici un entier : Clic droit → New Child→ Type Int. On pourra également lui affecter une valeur minimale et maximale (par exemple 150 et 200).

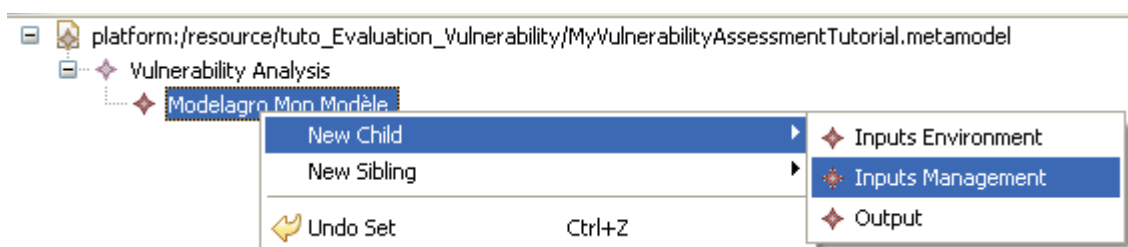


Figure 62. Ajouter une entrée au modèle.

Ensuite, il faut définir la distribution que ce suit cette variable (par exemple Uniforme ou Gaussienne) et quel échantillonnage on souhaite pour l'analyse de vulnérabilité sans adaptation. Autrement dit, on cherche à définir quel est l'ensemble des valeurs que peut prendre cette entrée ainsi que la probabilité associée. Dans notre exemple, on choisira une distribution Uniforme discrète (*i.e.* juste un certain nombre défini de valeurs et chaque valeur est équiprobable) : en partant du type Int, clic droit → new Sibling → Discret

Uniform (Figure 63). Noter que la nature de l'entrée ayant déjà été définie, la partie associée est grisée.

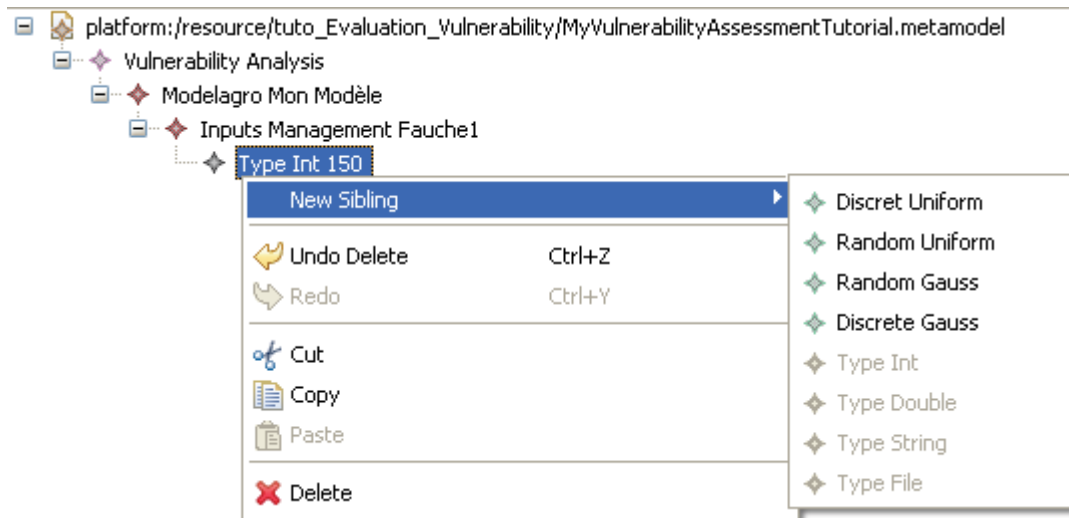


Figure 63. Associer une distribution et un échantillonnage à une entrée.

De même que pour tous les éléments de l'analyse de vulnérabilité, on renseigne les propriétés demandées dans l'onglet "Properties", à savoir le nombre de niveau (c'est-à-dire le nombre de valeur que peut prendre "Fauche1", ainsi que le minimum et le maximum pour l'échantillonnage. Noter que le domaine parcouru par l'échantillonnage doit être inclus dans le domaine possible des valeurs défini dans le Type Int. Le domaine possible des valeurs servira pour l'analyse de vulnérabilité avec adaptation, tandis que le domaine de l'échantillonnage sert pour l'analyse de vulnérabilité sans adaptation. Pour notre exemple on choisira $\max = 160$, $\min = 169$ et 10 niveaux. Faites un clic droit → valider et observer le report d'un problème sur la définition du minimum et du maximum (Figure 64). Il faut en effet nécessairement que \min soit inférieur ou égal à \max . De manière générale, au cours de la saisie ou à la fin, on peut vérifier la validité de la saisie via le menu contextuel "Validate" qui va vérifier qu'il ne manque aucun élément (par exemple pas d'oubli du choix de la méthode d'optimisation), mais également la cohérence des données saisies via OCL (Object Constraint Language).

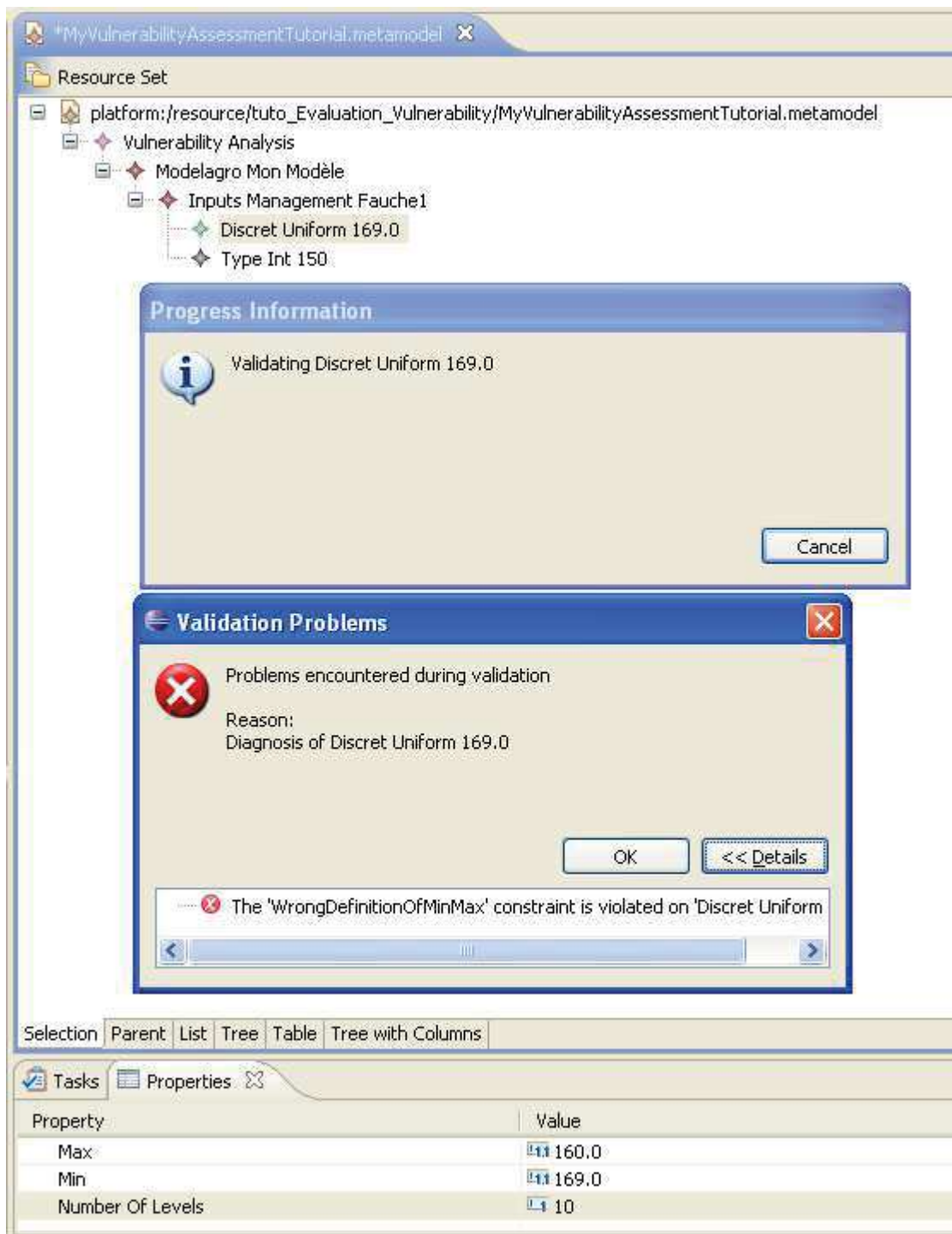


Figure 64. Valider une distribution associée à une entrée.

Inverser les valeurs du minimum et du maximum, puis sélectionner le modèle agro. Ajouter une sortie : clic droit → New Child → Output. Nous l'appellerons "Rendement" et lui affecterons une taille de 30 éléments (1 valeur par année).

Définition du plan expérimental initial pour l'analyse de vulnérabilité sans adaptation

Remplacez-vous sur le nœud "Vulnerability Analysis". Clic droit → New Child → Initial Design Of Experiment. Nous le nommerons "Doe1". Nous allons maintenant définir la nature du design (Figure 65). Choisissons un plan factoriel complet : New Child → Full Factorial.

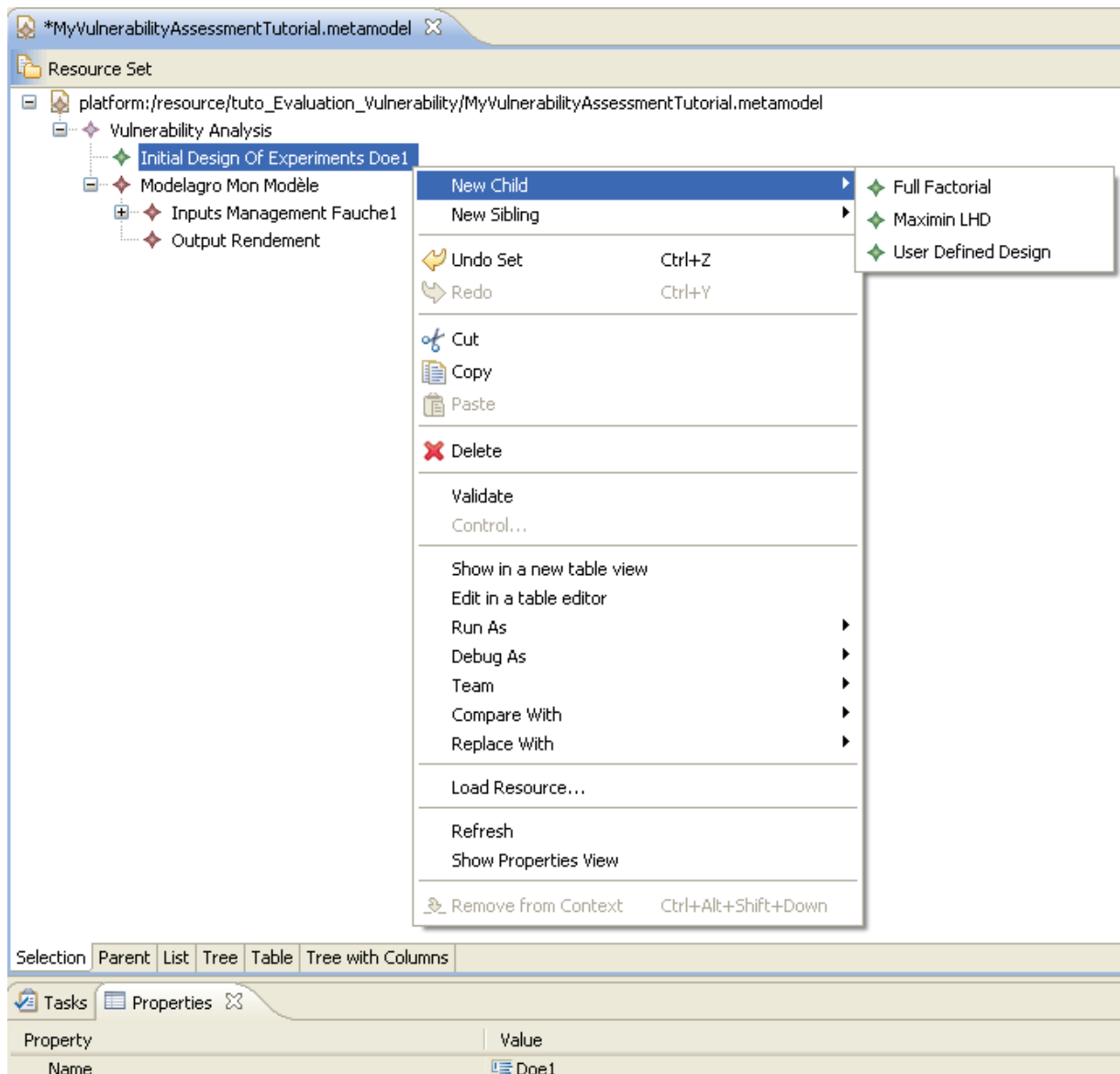


Figure 65. Définition de la nature du plan expérimental.

On choisit la nature des scripts de simulation

Actuellement, seules deux natures de scripts de simulation sont disponibles. Les scripts pour OpenMOLE⁹⁷ ou la solution "ascii", qui génèrent simplement le design et calcule la probabilité de chaque combinaison. Nous allons simplement appliquer cette solution : repartir du nœud d'analyse de vulnérabilité, Vulnerability Model → New Child → Ascii.

On choisit le type de surface de réponse à construire

Actuellement l'outil permet uniquement la génération des scripts pour le logiciel R. Nous allons choisir une surface de réponse de type régression linéaire. En repartant du nœud d'analyse de vulnérabilité, Vulnerability Model → New Child → Linear Regression. On obtient une arborescence équivalente à celle de la Figure 66.

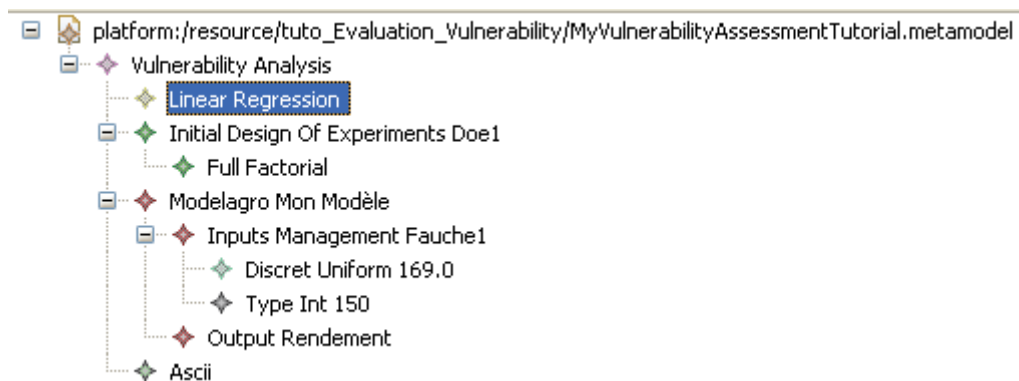


Figure 66. Graphe de l'analyse de vulnérabilité après ajout de la surface de réponse de type régression linéaire.

On choisit l'indice de vulnérabilité

Actuellement cinq différents indices sont disponibles : "Proportional Vulnerability", "Vulnerability Gap", "Vulnerability Severity", "Most Vulnerable individual", et l'indice de Luers basé sur un coefficient de variation pour l'estimation de la sensibilité. On va choisir l'indice de Luers et lui affecter un seuil de 0.75 : en repartant du nœud d'analyse de vulnérabilité, Vulnerability Model → New Child → Luers.

Enfin on choisit une méthode d'optimisation

À l'avenir, nous souhaitons implémenter une méthode d'optimisation qui se base sur les surfaces de réponses pour chercher un minimum de vulnérabilité et proposer un plan d'expériences autour de ce dernier afin d'évaluer la pertinence de celui-ci et ainsi d'estimer

⁹⁷ Comme la (Document Type Definition) du fichier d'interface n'est pas encore stable, nous avons préféré générer les scripts pour le mode "console" plutôt que le fichier xml utilisé par l'interface.

la vulnérabilité avec adaptation. Au niveau interface cela correspond simplement au choix de la méthode de la manière suivante : en repartant du nœud d'analyse de vulnérabilité, Vulnerability Model → New Child → Optimization. L'arborescence finale obtenue est représentée dans la Figure 67.

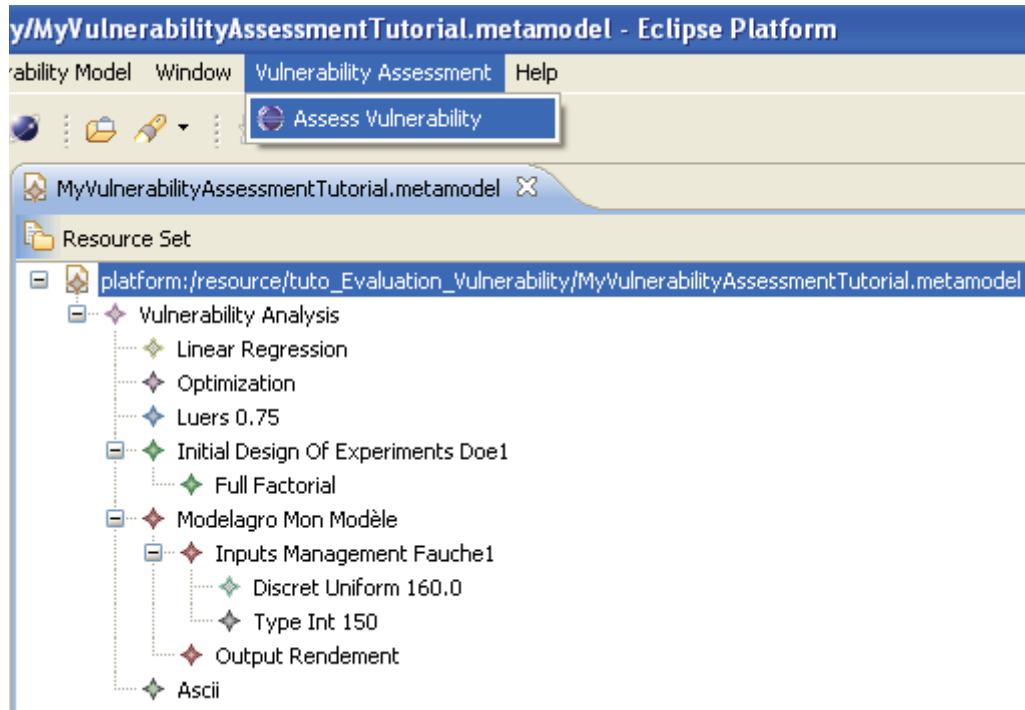


Figure 67. Lancement de la fonctionnalité "Analyse de vulnérabilité".

Lancement de l'analyse de vulnérabilité

Une fois la saisie terminée, on lance l'analyse de vulnérabilité : menu : "Vulnerability Assessment → Assess Vulnerability" (Figure 67). L'outil va ainsi (Figure 68) générer le plan expérimental, le script pour lancer les simulations ainsi que le script R⁹⁸ (dont un exemple est donné dans la Figure 69) pour calculer la surface de réponse. Il a été décidé de ne pas enchaîner de manière automatique les étapes. En effet, les temps de calculs pouvant être importants (par exemple une journée de calcul sur la grille Européenne de calcul), il est préférable d'être libre de choisir la machine et le moment du lancement des calculs. De plus l'automatisation aurait entraîné l'obligation de configurer les logiciels (R, OpenMOLE) sur la machine de l'utilisateur. D'autre part, le découplage des étapes permet de n'utiliser qu'une ou plusieurs étapes (par exemple juste le calcul des indices de vulnérabilité). Cet aspect modulaire est important car il augmente le nombre d'utilisateurs potentiels.

⁹⁸ <http://www.r-project.org/>

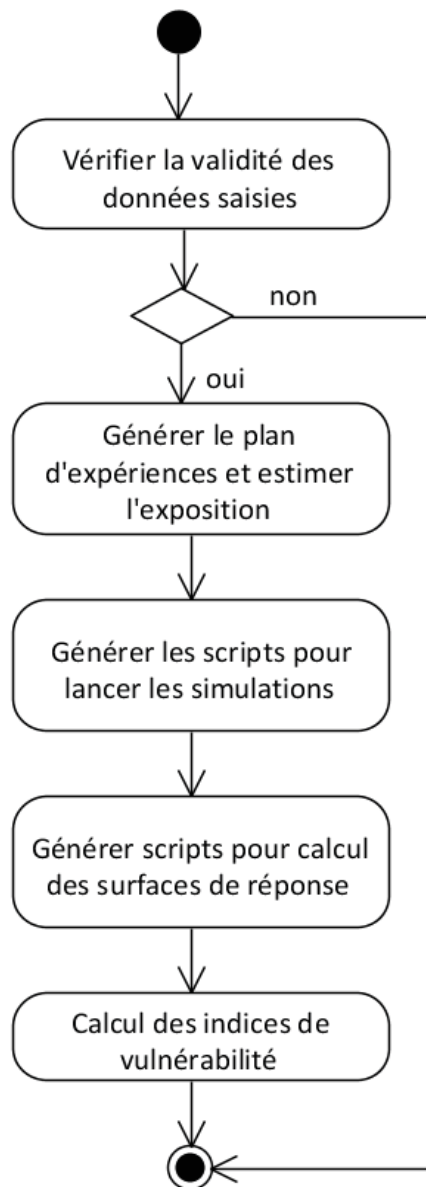


Figure 68. Diagramme d'activité de l'analyse de vulnérabilité par le plugin EMF.

```

1  # Script for response surface calculation.
2  # We will do one linear regression per
3  # output with respect to all input except
4  # String ones.
5
6  resultAndDOEFile = read.csv("Mon Modèle_Doe1.csv",header=FALSE)
7  names(resultAndDOEFile)[1]="Fauche1"
8  names(resultAndDOEFile)[2]="proba"
9  names(resultAndDOEFile)[3]="Rendement_1"
10 ...
11 names(resultAndDOEFile)[30]="Rendement_28"
12 names(resultAndDOEFile)[31]="Rendement_29"
13 names(resultAndDOEFile)[32]="Rendement_30"
14
15 inputToConsider=c(1)
16 dim1 = 2 + length(inputToConsider)
17 DOE = resultAndDOEFile[,inputToConsider]
18
19 # Might be necessary for windows system
20 # memory.size(max = TRUE)
21 # memory.size(4095)
22 # memory.limit(4095)
23
24 #to follow the progression of calculation and results
25 sink("regLinear_Mon Modèle_Doe1.txt")
26
27 tab=as.data.frame(t(1:dim1))
28 for (i in 1:30){
29     result= resultAndDOEFile[,i + 2]
30     modell= "result~ "
31     for (variable in names(DOE)){
32         modell = paste(modell,variable,sep=" + " )
33     }
34     library(bigm)
35     reg = biglm(as.formula(modell),data=cbind(DOE,result))
36     print(names(resultAndDOEFile[i + 2]))
37     print(summary(reg))
38     tab[i,1]=names(resultAndDOEFile)[i + 2]
39     tab[i,2:dim1]=coef(reg)
40     names(tab)[2:dim1]=names(coef(reg))
41 }
42 sink()
43 write.csv(tab, file="coefLinear_Mon Modèle_Doe1.csv", row.names=F)

```

Figure 69. Exemple de code R généré par le plugin EMF, pour la régression linéaire.

Enfin, une fenêtre permet de choisir les fichiers de résultats sur lesquels on souhaite calculer les indices (Figure 70, Figure 71). Les valeurs des indices seront affichées à l'écran, mais seront également écrites dans un fichier "log.txt". Noter également que certaines alertes, comme par exemple l'impossibilité de discrétiser la variable fauche1 en

20 valeurs entières entre 160 et 169, peuvent être affichés à l'écran et sont également écrites dans un fichier "log_erreur.txt".

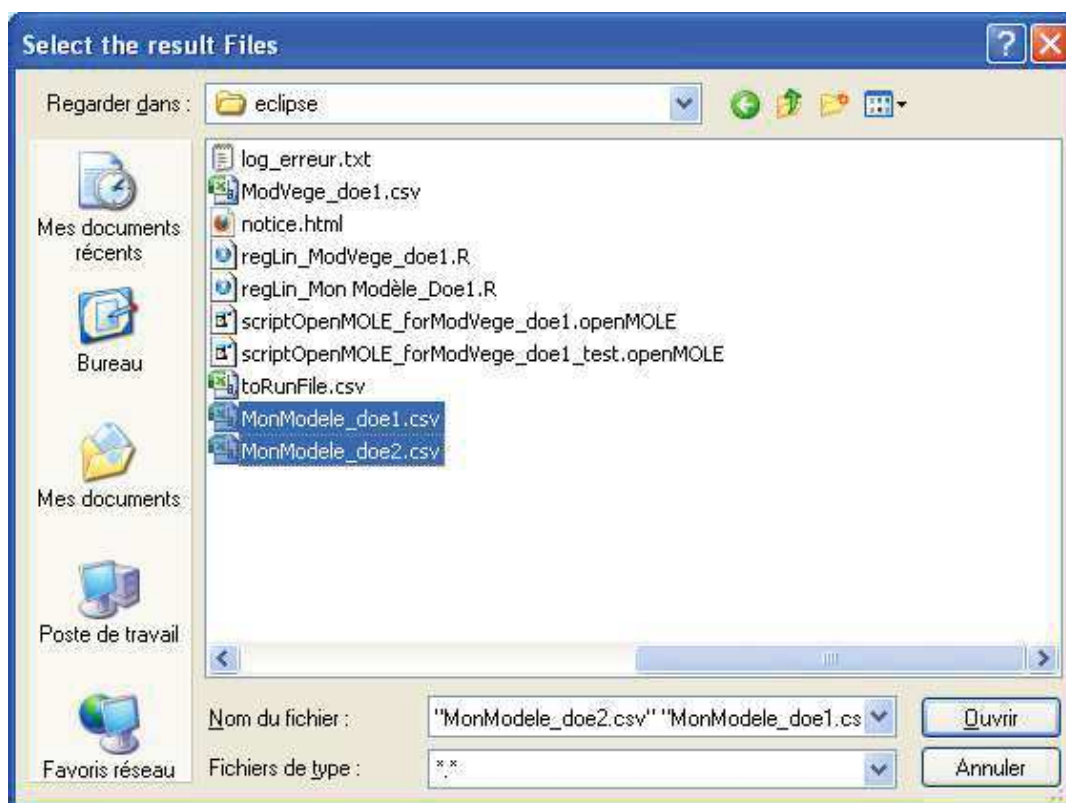


Figure 70. Sélection des fichiers pour les calculs d'indices.

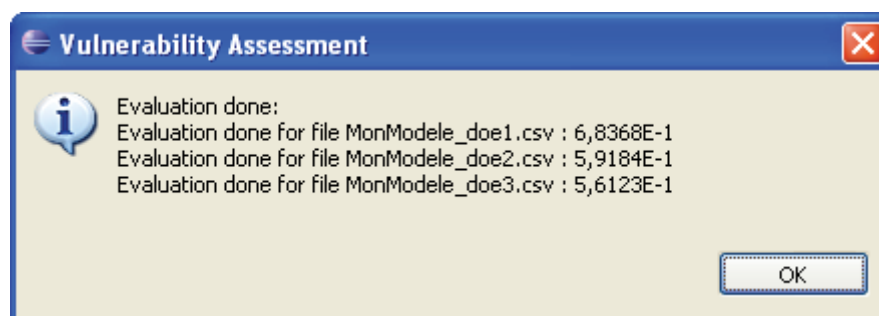


Figure 71. Exemple de résultats de calculs d'indices.

Contenu de l'approche retenue en utilisant le cadriciel EMF, la généricité et la modularité de l'outil est très grande, à la fois à travers son utilisation (par exemple utilisation simple pour générer un DOE, ou pour juste créer le script R pour construire une régression linéaire), mais également à travers son évolution. Il permet en effet très simplement d'intégrer une classe traduisant une nouvelle fonctionnalité (par exemple nouvel indice de vulnérabilité, scripts Matlab...). A noter cependant que la phase d'analyse de vulnérabilité avec adaptation, n'a été que partiellement implémenté dans le plugin. Cette phase se traduira à l'avenir par la recherche d'un minimum de vulnérabilité sur les surfaces de réponses puis par un nouveau plan d'expériences pour évaluer ce minimum. Cependant,

cette étape n'ayant pas encore été suffisamment testée et mise en œuvre préalablement, son implémentation s'est limitée à la création des classes associées (*i.e.* classe **Méthode Optimisation**).